



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

---

ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI INGEGNERIA CIVILE, CHIMICA, AMBIENTALE E DEI MATERIALI

INGEGNERIA CIVILE

---

**Tesi di Laurea**

---

IN

Tecnica ed Economia dei Trasporti

**LA MICROMOBILITÀ CONDIVISA:  
I CASI DI FIRENZE E BOLOGNA**

*Candidato:*

Antonio José Gregorio Rega

*Relatore:*

Chiar.mo Prof. F. Rupi

*Correlatore:*

Dott. Ing. G. Bernieri

ANNO ACCADEMICO 2024/25

SESSIONE II



# Indice

Premessa . . . . .	
<b>1 Introduzione</b>	<b>1</b>
1.1 Contesto . . . . .	2
1.2 Finalità e obiettivi della ricerca . . . . .	6
1.3 Nascita ed evoluzione della mobilità urbana . . . . .	7
1.4 Caratteristiche della mobilità condivisa . . . . .	10
1.4.1 Accessibilità e flessibilità . . . . .	11
1.4.2 Integrazione tecnologica . . . . .	12
1.4.3 Sostenibilità e impatto ambientale . . . . .	13
1.5 Impatti della mobilità condivisa sulle città . . . . .	16
1.5.1 La città degli X-minuti . . . . .	16
1.6 Presentazione del servizio . . . . .	20
1.7 Confronto con altri servizi di Bike Sharing . . . . .	23
<b>2 Metodi</b>	<b>25</b>
2.1 Strumenti e software . . . . .	26
2.2 Natura dei dati . . . . .	27
2.3 Natura dell'analisi . . . . .	29
2.3.1 Analisi esplorativa dei dati (EDA) . . . . .	31
<b>3 Risultati</b>	<b>35</b>
3.1 Presentazione dei risultati . . . . .	36
3.2 La città di Bologna . . . . .	37
3.2.1 Analisi generale . . . . .	37
3.2.2 Analisi spaziale . . . . .	47
3.2.3 Analisi tempi/velocità . . . . .	51
3.2.4 Analisi applicativa . . . . .	56
3.2.5 Analisi temporale . . . . .	57
3.2.6 Analisi dei turnover . . . . .	59
3.3 La città di Firenze . . . . .	61
3.3.1 Analisi generale . . . . .	61

3.3.2	Analisi spaziale . . . . .	71
3.3.3	Analisi tempi/velocità . . . . .	75
3.3.4	Analisi applicativa . . . . .	80
3.3.5	Analisi temporale . . . . .	81
3.3.6	Analisi dei turnover . . . . .	83
<b>4</b>	<b>Discussione</b>	<b>85</b>
4.1	Introduzione e sintesi dei risultati chiave . . . . .	86
4.2	Interpretazione dei pattern d'uso del servizio . . . . .	87
4.3	Analisi comparativa: Bologna vs. Firenze . . . . .	88
4.3.1	Pattern settimanale: uso funzionale e uso misto . . . . .	88
4.3.2	Pattern orario: una bimodalità condivisa . . . . .	88
4.4	Limiti dello studio . . . . .	89
4.5	Implicazioni e direzioni future . . . . .	90
4.5.1	Implicazioni per gestori e pianificatori urbani . . . . .	90
4.5.2	Direzioni per la ricerca futura . . . . .	91
4.6	Conclusioni . . . . .	92
<b>A</b>	<b>Algoritmo d'analisi</b>	

# Elenco delle figure

1.1	Mappa della rete ciclabile. . . . .	3
1.2	I sistemi di <i>governance</i> politica e operativa <sup>[1]</sup> . . . . .	5
1.3	Una <i>bicicletta bianca</i> sollevata in aria come segno di protesta. . . . .	8
1.4	Sostituzione modale nelle città analizzate. . . . .	11
1.5	Ripartizione delle emissioni di gas serra per modalità di trasporto (2019). . . . .	13
1.6	Numero dei servizi di sharing mobility in Italia attivi al 2017. . . . .	18
1.7	Zona operativa di Bologna per il servizio di mobilità condivisa. . . . .	21
1.8	Zona operativa di Firenze per il servizio di mobilità condivisa. . . . .	21
1.9	Un ciclista pedala davanti a un enorme cumulo di biciclette condivise inutilizzate a Xiamen, in Cina. . . . .	24
2.1	Alcuni degli obiettivi di sviluppo sostenibile legati alla mobilità. . . . .	30
3.1	Istogramma della distribuzione delle distanze per il dataset di Bologna. . . . .	38
3.2	Quantile-Quantile (QQ) plot della distribuzione delle distanze. . . . .	39
3.3	Funzione di ripartizione (CDF) delle distanze. . . . .	40
3.4	Curva di densità di una distribuzione normale teorica. . . . .	41
3.5	Istogramma della distribuzione delle distanze <i>map-matched</i> per il dataset di Bologna. . . . .	43
3.6	Quantile-Quantile (QQ) plot della distribuzione delle distanze. . . . .	44
3.7	Funzione di ripartizione (CDF) delle distanze. . . . .	45
3.8	Curva di densità di una distribuzione normale teorica. . . . .	46
3.9	Distribuzione spaziale dei punti di partenza sul territorio urbano di Bologna. . . . .	47
3.10	Distribuzione spaziale dei punti di arrivo sul territorio urbano di Bologna. . . . .	48
3.11	Mappe di densità dei punti di partenza. . . . .	49
3.12	Mappe di densità dei punti di arrivo. . . . .	50
3.13	Istogramma della distribuzione dei tempi di percorrenza, suddivisi per tipologia di bicicletta: <i>e-bike</i> (arancione) e <i>muscolari</i> (azzurro). . . . .	52
3.14	Istogramma della distribuzione delle velocità medie, suddivise per tipolo- gia di bicicletta: <i>e-bike</i> (arancione) e <i>muscolari</i> (azzurro). . . . .	53

3.15	Grafico a dispersione ( <i>scatter plot</i> ) della relazione tra distanza percorsa (km) e tempo impiegato (ore). Ogni punto rappresenta un singolo viaggio, colorato in base alla tipologia di bicicletta: <i>e-bike</i> (arancione) e <i>muscolare</i> (azzurro). . . . .	54
3.16	Grafico a dispersione ( <i>scatter plot</i> ) che illustra la relazione tra la distanza percorsa (km) e la velocità media (km/h). Ogni punto rappresenta un singolo viaggio, colorato in base alla tipologia di bicicletta: <i>e-bike</i> (arancione) <i>muscolare</i> (azzurro). . . . .	55
3.17	Distribuzione oraria del numero di viaggi, suddivisi per tipologia di bicicletta: <i>e-bike</i> (arancione) e <i>muscolare</i> (azzurro). . . . .	57
3.18	Distribuzione del numero di viaggi per giorno della settimana, suddivisi per tipologia di bicicletta: <i>e-bike</i> (arancione) e <i>muscolare</i> (azzurro). . . .	58
3.19	Istogramma della distribuzione del numero totale di viaggi per bicicletta.	59
3.20	Funzione di distribuzione cumulativa (CDF) del numero totale di viaggi per bicicletta. . . . .	60
3.21	Istogramma della distribuzione delle distanze per il dataset di Firenze. .	62
3.22	Quantile-Quantile (QQ) plot della distribuzione delle distanze. . . . .	63
3.23	Funzione di ripartizione (CDF) delle distanze. . . . .	64
3.24	Curva di densità di una distribuzione normale teorica. . . . .	65
3.25	Istogramma della distribuzione delle distanze per il dataset di Firenze. .	67
3.26	Quantile-Quantile (QQ) plot della distribuzione delle distanze. . . . .	68
3.27	Funzione di ripartizione (CDF) delle distanze. . . . .	69
3.28	Curva di densità di una distribuzione normale teorica. . . . .	70
3.29	Distribuzione spaziale dei punti di partenza sul territorio urbano di Bologna.	71
3.30	Distribuzione spaziale dei punti di arrivo sul territorio urbano di Bologna.	72
3.31	Mappe di densità dei punti di partenza. . . . .	73
3.32	Mappe di densità dei punti di arrivo. . . . .	74
3.33	Istogramma della distribuzione dei tempi di percorrenza. . . . .	76
3.34	Istogramma della distribuzione delle velocità medie. . . . .	77
3.35	Grafico a dispersione ( <i>scatter plot</i> ) della relazione tra distanza percorsa (km) e tempo impiegato (ore). . . . .	78
3.36	Grafico a dispersione ( <i>scatter plot</i> ) che illustra la relazione tra la distanza percorsa (km) e la velocità media (km/h). . . . .	79
3.37	Distribuzione oraria del numero di viaggi, suddivisi per tipologia di bicicletta: <i>e-bike</i> (arancione) e <i>muscolare</i> (azzurro). . . . .	81
3.38	Distribuzione del numero di viaggi per giorno della settimana, suddivisi per tipologia di bicicletta: <i>e-bike</i> (arancione) e <i>muscolare</i> (azzurro). . . .	82
3.39	Istogramma della distribuzione del numero totale di viaggi per bicicletta.	83
3.40	Funzione di distribuzione cumulativa (CDF) del numero totale di viaggi per bicicletta. . . . .	84

## Premessa

Lo scopo della presente ricerca è fornire un'analisi attenta del servizio di **bike sharing** nelle città di Bologna e Firenze. Entrambe le città, di media grandezza, presentano un tessuto urbano che combina una matrice romana, profondamente trasformata nel Medioevo ma ancora riconoscibile, con spaziose vie periferiche sviluppate in epoca post-industriale. Questa configurazione consente di osservare l'evolversi delle esigenze di spostamento dei cittadini in contesto urbano contemporaneo, che poco spazio concede al libero movimento ma, al contrario, confina gli spostamenti alle brulicate arterie cittadine, la cui via di fuga, come sarà evidente nel seguito del testo, è per noi da ricercare nella condivisione: non più il singolo, che in autonomia compie i suoi spostamenti nella solitudine del suo abitacolo, ma la moltitudine, che persegue attraverso la condivisione un obiettivo comune.

I programmi di *bike sharing*, rispetto alle alternative a motore proposte dal trasporto pubblico, raggiungono appieno l'obiettivo di promuovere la mobilità sostenibile. In particolare, i sistemi dockless, oggetto della nostra ricerca, offrono all'individuo un'autonomia che il tradizionale trasporto di massa non garantisce, senza compromettere il carattere ecologico della modalità di trasporto. Questo modello, infatti, combina la flessibilità dell'uso personale con l'efficienza di un sistema che riduce l'impatto ambientale.

L'analisi dei servizi di sharing mobility si sviluppa attraverso diverse fasi. Inizialmente, viene esaminata la storia del fenomeno, per comprendere le sue origini e l'evoluzione nel tempo. Successivamente, si esplorano gli aspetti tecnici e operativi dei vari modelli, valutando le caratteristiche che li rendono distintivi. Infine, l'attenzione si concentra sull'impatto socio-economico e ambientale, al fine di analizzare i benefici e le sfide associate a queste soluzioni di mobilità condivisa.

Questa ricerca si inserisce all'interno di un ampio contesto di studi che ha avuto inizio il secolo scorso, e non vuole rappresentare la soluzione ultima a un fenomeno complesso in continua evoluzione, quanto piuttosto un punto di vista, forse diverso, di uno dei tanti singoli che fa parte della sazia e disperata moltitudine.

Buona lettura.





# Capitolo 1

## Introduzione

## 1.1 Contesto

Le tematiche della mobilità sostenibile e dell'efficientamento energetico sono universali e prescindono, quindi, dal contesto di riferimento, sebbene si declinino in termini diversi in base alla morfologia dei differenti territori e ai tessuti socio-economici di riferimento. Il contesto che fa da sfondo a questa tesi è quello delle città di Firenze e, in particolare, Bologna.

Bologna, da sempre città di innovazione e tradizione, si è posta l'ambizioso obiettivo di diventare un modello di riferimento nazionale e internazionale per la transizione sostenibile. Con il progetto *Città della Conoscenza*, l'amministrazione comunale sta integrando politiche attive e iniziative tecnologiche per affrontare le sfide sociali, economiche e ambientali contemporanee. Al centro di questa visione si trova il **Gemello Digitale** [1] di Bologna, un modello virtuale avanzato che utilizza dati integrati e strumenti tecnologici per affrontare le sfide urbane. Come descritto dalla *Fondazione Innovazione Urbana*, questo progetto punta a trasformare la città in un laboratorio per la sperimentazione di soluzioni che combinano mobilità sostenibile, riduzione del fabbisogno energetico e partecipazione attiva dei cittadini. I principali obiettivi includono l'ottimizzazione dei flussi di traffico, il supporto alla pianificazione urbana e la simulazione predittiva degli impatti delle politiche urbane [2].

Il *Gemello Digitale* rappresenta infrastrutture fisiche, dinamiche sociali e organizzative. Questa piattaforma integra dati provenienti da fonti eterogenee, consentendo di analizzare, visualizzare e prevedere scenari urbani. Grazie a un finanziamento iniziale di 7 milioni di euro nell'ambito del programma europeo PON METRO, Bologna sta sviluppando una soluzione che combina approcci avanzati di intelligenza artificiale con la partecipazione attiva di cittadini e stakeholder. I principali obiettivi del progetto includono:

- *Mobilità*: supportare la trasformazione del sistema di trasporto urbano, integrando progetti come il Tram, la Bicipolitana e la riduzione della velocità a 30 km/h.
- *Energia e Cambiamenti Climatici*: ottimizzare il consumo energetico e simulare l'impatto di politiche ambientali, contribuendo alla neutralità climatica entro il 2030.
- *Partecipazione Civica*: coinvolgere attivamente cittadini, aziende e centri di ricerca per la co-progettazione di soluzioni innovative.

Una delle applicazioni principali del *Gemello Digitale* è la mobilità urbana: la città di Bologna sta utilizzando i dati raccolti dai sensori cittadini, come telecamere di monitoraggio e contatori di traffico, per creare mappe dinamiche che identificano punti critici e ottimizzano i flussi di traffico. Questi strumenti supportano la progettazione di soluzioni innovative come la **Bicipolitana**.

La *Bicipolitana* è una rete ciclabile metropolitana che connette i principali centri abitati, poli produttivi e funzionali dell'area metropolitana. Con un'estensione di quasi 1.000 km, di cui 493 sulle direttrici principali, la Bicipolitana rappresenta una delle strategie centrali del Piano Urbano della Mobilità Sostenibile, o PUMS della Città Metropolitana di Bologna per il riequilibrio modale [3].

Gli obiettivi principali dichiarati sono:

- *Riduzione delle emissioni climalteranti*: mira a una diminuzione del 40% delle emissioni e all'aumento di 240.000 spostamenti giornalieri in bicicletta.
- *Rivoluzione della mobilità urbana*: punta a migliorare l'accessibilità dei territori, riducendo la dipendenza dal trasporto privato.
- *Sviluppo socio-economico*: il progetto promuove il commercio di prossimità, valorizza le periferie e rafforza Bologna come polo turistico e culturale.



**Figura 1.1:** Mappa della rete ciclabile.

La *Bicipolitana* è composta da 34 linee:

- 20 linee per spostamenti quotidiani:
  - 10 linee radiali verso Bologna;
  - 2 linee attorno a Imola;
  - 4 linee trasversali nella pianura bolognese;
  - 4 linee tangenziali intorno a Bologna e Imola;
- 14 linee per il tempo libero e il cicloturismo, alcune delle quali coincidono con le principali ciclo-vie regionali e internazionali, come la *Ciclovía del Sole* (EUROVELO7) e la *Ciclovía Emilia* (ER8).

La *Bicipolitana* rappresenta un simbolo di innovazione e sostenibilità, capace di aumentare la vivibilità del territorio metropolitano bolognese. Contribuisce a creare un'infrastruttura per la mobilità sostenibile, rafforzando le comunità locali e valorizzando il patrimonio turistico e culturale di Bologna.

Le reti ciclabili che compongono la *Bicipolitana* sono state digitalizzate e integrate nel *Gemello Digitale*, così che l'amministrazione possa simulare l'impatto di politiche di mobilità, come l'introduzione di zone a traffico limitato o il miglioramento delle infrastrutture per la mobilità condivisa.

In generale, il progetto *Città della Conoscenza* si articola in tre linee principali:

- *Missione Clima*: neutralità climatica entro il 2030 tramite interventi su mobilità, energia e gestione rifiuti.
- *Impronta Verde*: creazione di infrastrutture verdi per migliorare la salute pubblica e la qualità della vita.
- *Piano per l'Abitare*: sviluppo di 10.000 nuovi alloggi nei prossimi dieci anni.

In questo contesto, il *Gemello Digitale* non è solo uno strumento tecnologico, ma un catalizzatore per politiche partecipative: Bologna sta emergendo come laboratorio per l'implementazione di tecnologie avanzate e modelli di governance *data-driven*.<sup>1</sup>

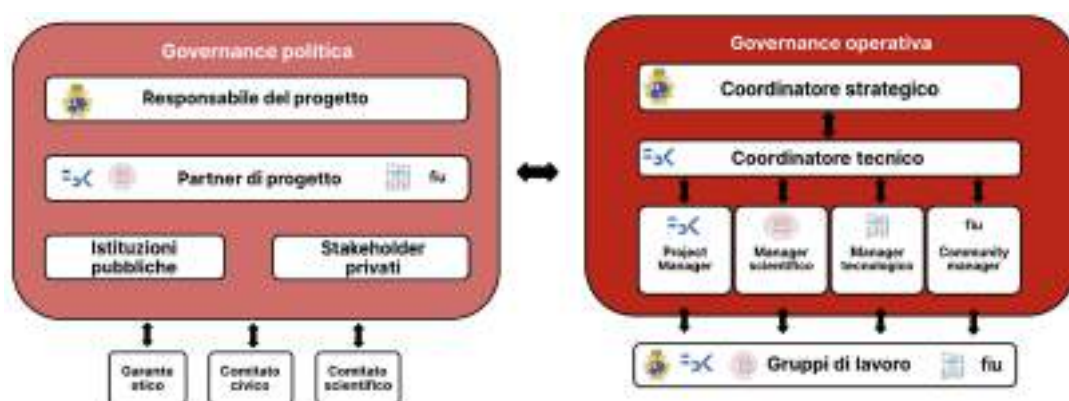
---

<sup>1</sup>*Data-driven*: di processo guidato dai numeri, ovvero di approccio basato sui dati, per prendere decisioni informate, basate su fatti oggettivi e non su sensazioni personali.

La governance del progetto si divide in due livelli: *politico* e *operativo*. La governance politica guida le scelte strategiche e comprende il responsabile del progetto, affiancato da partner istituzionali (come università e centri di ricerca), istituzioni pubbliche e stakeholder privati. Inoltre, tre comitati, etico, civico e scientifico, assicurano trasparenza, coinvolgimento cittadino e validità tecnica.

La governance operativa traduce le strategie politiche in azioni concrete. Qui operano un coordinatore strategico e un coordinatore tecnico, supportati da figure specializzate come il *project manager*, il *manager scientifico*, il *manager tecnologico* e il *community manager*. Infine, i gruppi di lavoro multidisciplinari si occupano degli aspetti esecutivi, garantendo un approccio pratico e collaborativo.

Questa struttura è pensata per garantire un continuo scambio tra visione politica e operatività concreta, in un modello che cerca di coniugare trasparenza, efficienza e innovazione.



**Figura 1.2:** I sistemi di *governance* politica e operativa<sup>[1]</sup>.

Questa tesi si inserisce quindi nel quadro delle trasformazioni promosse dal *Gemello Digitale* sotto la guida scientifica dell'ateneo, concentrandosi in particolare sulla mobilità urbana.

## 1.2 Finalità e obiettivi della ricerca

Questo lavoro indaga il contributo delle soluzioni di mobilità condivisa alla mobilità urbana sostenibile, assumendo i casi studio di Bologna e Firenze come contesto d'analisi. La ricerca valuta in che misura i servizi di *mobilità condivisa*, o *sharing mobility*, possano ridurre congestione ed emissioni, migliorare l'accessibilità e offrire indicazioni operative per la pianificazione urbana.

### Obiettivi specifici

Gli obiettivi perseguiti sono:

1. Analizzare lo stato dell'arte e le caratteristiche operative dei servizi di mobilità condivisa a Bologna e Firenze, evidenziandone differenze e punti di forza.
2. Valutare le dinamiche d'uso e l'esperienza degli utenti mediante l'elaborazione dei dati operativi disponibili.
3. Proporre raccomandazioni e linee guida replicabili per migliorare efficacia e integrazione dei servizi in città di medie dimensioni.

### Sintesi metodologica

La ricerca combina analisi quantitative (dataset operativi, analisi spaziale e temporale) e qualitative (revisione documentale). La descrizione completa dei dati, degli strumenti (p.es. NumPy, RideTool, SUMO, **hybridPY**) e del workflow analitico è riportata dettagliatamente nel Capitolo 2.

### Organizzazione dell'elaborato

L'elaborato è strutturato in quattro capitoli principali:

- **§ Cap. 1** — *Introduzione*: contesto, motivazioni e obiettivi.
- **§ Cap. 2** — *Metodi*: dati, strumenti e workflow analitico (inclusi dettagli su simulazione e tool).
- **§ Cap. 3** — *Risultati*: analisi e rappresentazione dei risultati.
- **§ Cap. 4** — *Discussione e conclusioni*: interpretazione, raccomandazioni e sviluppi futuri.

## 1.3 Nascita ed evoluzione della mobilità urbana

L'espansione delle aree urbane e l'incremento demografico nei centri cittadini, fenomeni che hanno caratterizzato la seconda metà del XX secolo, hanno esercitato una pressione crescente sulle infrastrutture di trasporto, determinando una serie di sfide complesse legate alla gestione della mobilità e alla qualità della vita nelle aree urbane. Le città, una volta caratterizzate da un rapporto relativamente equilibrato tra gli spazi dedicati alla residenza e quelli destinati al trasporto, si sono trasformate in agglomerati densamente popolati, con conseguente aumento della domanda di spostamenti quotidiani. Questo ha spinto a una riflessione profonda non solo sulle modalità di trasporto, ma anche sul concetto stesso di mobilità, che si è evoluto per rispondere alle esigenze di un mondo sempre più interconnesso e in rapido cambiamento. La mobilità urbana è divenuta, così, una delle principali sfide del nostro tempo, caratterizzata dalla necessità di ridurre la congestione, abbattere le emissioni di CO<sub>2</sub>, migliorare la qualità dell'aria e garantire spazi pubblici più vivibili e sicuri. In questo contesto, la domanda di soluzioni di trasporto sostenibili ed efficienti ha alimentato un ampio dibattito globale, che ha stimolato l'emergere di modelli di mobilità innovativi, tra cui la *mobilità condivisa*. Quest'ultima include soluzioni di noleggio condiviso di scooter, automobili e biciclette, un tema che costituisce il fulcro della ricerca di questa tesi.

A partire dal dopoguerra, molte città occidentali hanno intrapreso un processo di urbanizzazione che ha comportato una crescente dipendenza dal trasporto privato automobilistico. Questo modello, inizialmente visto come sinonimo di progresso e modernità, ha però comportato conseguenze negative ben visibili già dalle prime decadi del XXI secolo, tra cui la congestione stradale, l'inquinamento atmosferico e acustico, il deterioramento della sicurezza stradale e l'accelerazione dei processi di cambiamento climatico. A queste problematiche si è aggiunta la diminuzione dell'attività fisica dovuta alla crescente diffusione di modalità di trasporto individuale motorizzato, che hanno ridotto la mobilità pedonale e ciclistica. Tale situazione ha stimolato un processo di riflessione collettiva e politica, dando origine a un rinnovato interesse per forme di mobilità più sostenibili. In particolare, l'attenzione si è concentrata sul ciclismo urbano come alternativa ecologica e salutare, supportata dallo sviluppo di politiche pubbliche favorevoli e da un crescente impegno tecnologico. Le tecnologie moderne hanno reso possibile la realizzazione di programmi di bike sharing che, grazie all'uso di sistemi di tracciamento e pagamento digitali, si sono diffusi rapidamente, offrendo soluzioni pratiche ed economiche per le persone che desiderano spostarsi senza dover ricorrere all'auto privata. Questo processo ha dato vita a un terreno fertile per l'espansione della *mobilità condivisa*, una forma di trasporto che si inserisce perfettamente in una visione urbana sostenibile, all'interno della quale la mobilità non è solo una questione di spostamenti fisici, ma anche un'occasione per ridisegnare il rapporto tra l'individuo, lo spazio pubblico e l'ambiente [4].

Il primo esempio documentato di bike sharing risale al 1965, con l'iniziativa delle *Witte Fietsen* (Biciclette Bianche) ad Amsterdam [5].



**Figura 1.3:** Una *bicicletta bianca* sollevata in aria come segno di protesta.

Le *Witte Fietsen* furono forse l'azione di attivismo creativo più conosciuta del gruppo olandese *Provo*. Il ramo politico dei *Provo* ottenne un seggio nel consiglio comunale di Amsterdam e sviluppò i “Piani Bianchi”. Il più famoso di questi è il “Piano delle Biciclette Bianche”, che mirava a risolvere il problema dei trasporti di Amsterdam. Il rivoluzionario piano di *Luud Schimmelpenninck* prevedeva la chiusura totale del centro di Amsterdam al trasporto motorizzato, con la sola eccezione del trasporto pubblico. Le azioni anti-istituzionali dei *Provo* scatenarono aspre reazioni da parte delle istituzioni che ostacolarono la riuscita del progetto che riuscì però a cambiare profondamente il volto della mobilità urbana di Amsterdam per sempre.



La letteratura divide l'evoluzione dei sistemi di bike sharing in quattro fasi, chiamate *generazioni* [5]:

- *Prima generazione*: rappresentata dalle *Witte Fietsen*, priva di sistemi di pagamento o sicurezza.
- *Seconda generazione*: comparsa negli anni '90, introduceva un sistema a deposito di monete simile ai carrelli nei supermercati; tuttavia, la mancanza di registrazione personale esponeva il sistema a continui furti.
- *Terza generazione*: sviluppata a partire dal 2000, implementava postazioni di docking con pagamenti automatizzati tramite carta di credito e sistemi di tracciamento GPS per monitorare le biciclette. Questi miglioramenti tecnologici hanno incentivato una rapida diffusione globale dei sistemi di bike sharing.
- *Quarta generazione*: include nuove soluzioni come i sistemi *dockless*, l'assistenza elettrica e l'integrazione con smartcard per i trasporti pubblici, rendendo il servizio ancora più accessibile e flessibile.

Sebbene il responso da parte dell'utenza, soprattutto nei paesi del Nord Europa, sia stato da subito positivo, è solo negli ultimi decenni che il bike sharing ha conosciuto una crescita sensibile. Ad esempio, il numero di città con un sistema di bike sharing è passato da 13 nel 2004 a 855 nel 2014, con una flotta globale stimata di oltre 946.000 biciclette, di cui circa 750.500 solo in Cina [4].

L'introduzione delle biciclette elettriche ha rappresentato un punto di svolta significativo nella mobilità urbana. Queste biciclette consentono di affrontare percorsi più lunghi e dislivelli impegnativi senza la fatica associata alla pedalata tradizionale. Secondo un'analisi condotta in Norvegia [5], l'uso delle e-bike ha permesso di aumentare la distanza percorsa dai ciclisti, passando da una media di 4,8 km a 10,3 km.

Tuttavia, la diffusione delle e-bike ha anche evidenziato nuove sfide, in particolare in termini di sicurezza e infrastrutture adeguate. Le e-bike sono più pesanti e raggiungono velocità maggiori rispetto alle biciclette tradizionali, rendendo necessaria la creazione di piste ciclabili sicure e ben segnalate per proteggere sia i ciclisti che gli altri utenti della strada.

## 1.4 Caratteristiche della mobilità condivisa

La mobilità condivisa propone l'uso condiviso di risorse in luogo del possesso, con l'obiettivo di ridurre i veicoli privati e la congestione. L'affermazione di piattaforme digitali e la sensibilità ambientale ne hanno accelerato l'adozione nelle aree urbane; tra i servizi più diffusi figurano auto e biciclette, che svincolano l'utente dall'onere di proprietà.

Un elemento abilitante è la tecnologia: i sistemi di bike sharing impiegano GPS e piattaforme di gestione per raccogliere dati su percorsi, soste, durata e domanda. Queste informazioni supportano il *rebalancing*, la distribuzione dei mezzi e l'individuazione di *hotspot*, oltre a suggerire interventi per mitigare la congestione e migliorare l'accessibilità. Le preferenze di itinerario riflettono un compromesso tra rapidità, sicurezza e comfort: i ciclisti privilegiano percorsi scorrevoli e prevedibili, riducendo semafori e conflitti con il traffico. In questo contesto, il **costo** di un itinerario è un indicatore sintetico del “dispendio” percepito.

In termini generali, il *costo* combina componenti monetarie, temporali, psicologiche (es. stress/sicurezza percepita), ambientali e di opportunità. La scelta del percorso può essere modellata come minimizzazione di un costo totale pesato:

$$C_{\text{tot}} = \alpha \cdot C_{\text{mon}} + \beta \cdot C_{\text{temp}} + \gamma \cdot C_{\text{amb}} + \delta \cdot C_{\text{psic}} + \dots \quad (1.1)$$

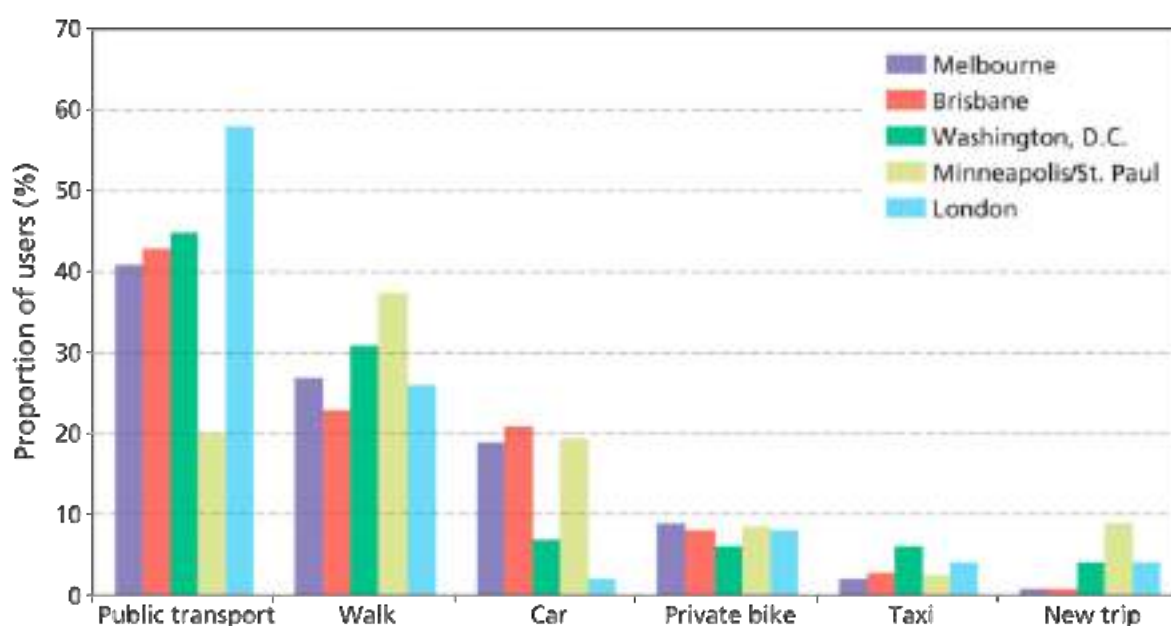
dove  $\alpha, \beta, \gamma, \delta$  rappresentano i pesi delle rispettive componenti.

Evidenze empiriche mostrano che il tempo di viaggio non è sempre determinante. Ad esempio, uno studio [6] sul comportamento dei ciclisti indica la preferenza per itinerari più sicuri e gradevoli anche se più lunghi.

L'uso dei dati è cruciale per l'efficienza e la sicurezza del sistema: il tracciamento GPS consente di localizzare aree di sosta ricorrenti per pianificare parcheggi dedicati; l'analisi dei flussi individua tratte congestionate o critiche, orientando la progettazione di infrastrutture ciclabili e interventi di traffic calming.

### 1.4.1 Accessibilità e flessibilità

La mobilità condivisa si fonda sull'accesso ai mezzi di trasporto piuttosto che sulla loro proprietà: il servizio nasce, infatti, come soluzione proposta alla battaglia al trasporto privato smodato, le cui armi sono l'*accessibilità* e la *flessibilità*. La vittoria principale che emerge dalla letteratura sul tipo di trasporto che il bike sharing sostituisce, la cosiddetta *sostituzione modale*, è che la maggior parte dei viaggi sostituisce quelli precedentemente effettuati con i mezzi pubblici e a piedi, piuttosto che sostituire viaggi effettuati con l'auto. Il tasso di sostituzione modale per una selezione di sistemi di bike sharing per i quali sono disponibili dati [4] è di seguito illustrato.<sup>2</sup>



**Figura 1.4:** Sostituzione modale nelle città analizzate.

Dai risultati di questa analisi multi-città sull'impatto del bike sharing sull'uso dell'automobile, si evince che, in tutte le città analizzate tranne una, il bike sharing riduce l'uso dell'auto. Londra è stata l'unica eccezione, a causa del basso tasso di sostituzione modale dell'auto (2%) e della grande distanza percorsa dai veicoli di redistribuzione. Poiché la riduzione dell'uso dell'auto è una motivazione chiave per l'implementazione del bike sharing, è evidente che gli attori del reparto di ricerca scientifica devono concentrare i propri sforzi nel misurare quantitativamente il grado in cui i sistemi di bike sharing incidono sull'uso delle auto e considerare modi per facilitare questo spostamento modale.

<sup>2</sup>Fonti: Fishman et al., 2014; LDA Consulting, 2012; Nice Ride Minnesota, 2010; Transport for London, 2011

## 1.4.2 Integrazione tecnologica

L'evoluzione dei *BSPs* (*Bike Sharing Programs*) è parallela alla crescita delle infrastrutture tecnologiche; le piattaforme di sharing si basano su infrastrutture digitali avanzate, che includono:

- **Geolocalizzazione e tracciamento GPS:** essenziali per individuare i mezzi disponibili in tempo reale, ottimizzando la distribuzione e riducendo i tempi di attesa.
- **Pagamenti elettronici:** semplificano le transazioni e rendono il processo di noleggio rapido e sicuro.
- **Sensori intelligenti:** monitorano lo stato del veicolo (ad esempio, livello della batteria, condizioni del mezzo) e la sua posizione, migliorando l'efficienza e la sicurezza del servizio.

La facilità con cui le persone possono iscriversi a un BSP può influire sulla probabilità di utilizzo del programma. La stragrande maggioranza dei BSP di terza generazione consente agli utenti potenziali di iscriversi immediatamente, utilizzando una carta di credito, ma quando ciò non accade, gli utenti e potenziali tali hanno segnalato che la mancanza di un'iscrizione immediata rappresenta un ostacolo all'uso. È notevole il caso dell'esperienza di *CityCycle*, uno tra i pochi BSP di terza generazione nei paesi anglofoni a non offrire l'iscrizione automatica nelle stazioni di ancoraggio [4].

Il Consiglio Comunale di Brisbane ha commissionato un sondaggio tra gli utenti registrati di *CityCycle*, inviato via e-mail a 13.495 abbonati a lungo e breve termine, con un tasso di risposta dell'11,5%. I risultati hanno mostrato che uno su tre degli utenti di *CityCycle* ha dichiarato di non voler rinnovare l'abbonamento, con il processo di iscrizione complesso segnalato come una delle ragioni principali per cui gli abbonati a breve termine non sono utenti abituali di *CityCycle*. Inoltre, il processo complesso di iscrizione è stato evidenziato dal 54% degli abbonati locali a breve termine come l'area principale che richiede miglioramenti, molti dei quali hanno dichiarato che non avrebbero più utilizzato *CityCycle*.

L'esperienza di *CityCycle* offre una lezione importante per i pianificatori di sistemi di *bike-sharing*. Gli utenti e i potenziali utenti apprezzano la spontaneità e le politiche dovrebbero cercare di minimizzare gli ostacoli associati all'iscrizione al servizio, come la mancanza di convenienza e il vantaggio competitivo rispetto ad altri mezzi di trasporto, le preoccupazioni sulla sicurezza e tutto ciò che ostacola la spontaneità.

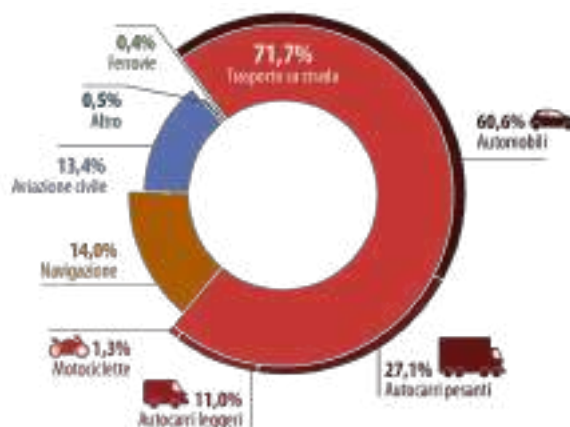
### 1.4.3 Sostenibilità e impatto ambientale

Tra le tematiche che nella ricerca scientifica hanno riscontrato più interesse negli ultimi anni vi sono sicuramente quelle legate all'*economia ambientale*, cioè l'interazione tra l'attività economica e l'ambiente naturale. L'*economia green* analizza in particolare come le risorse naturali vengano utilizzate e gestite, l'impatto delle attività umane sull'ambiente e le soluzioni economiche per affrontare i problemi ambientali. L'obiettivo principale è promuovere uno sviluppo economico sostenibile, che bilanci la crescita economica con la protezione e la conservazione delle risorse naturali.

L'obiettivo di una crescita economica che tuteli l'ambiente passa attraverso le porte della *mobilità sostenibile*, dato che secondo i dati dell'*Agenzia Europea dell'Ambiente* [7], al 2022, il settore dei trasporti è responsabile di circa un quarto delle emissioni totali di CO<sub>2</sub> in Europa, il 71,7% delle quali viene prodotto dal trasporto stradale.

Negli ultimi tre decenni l'unico settore a vedere un aumento delle emissioni di gas serra è stato quello dei trasporti, con un aumento del 33,5% tra il 1990 e il 2019. La riduzione significativa nelle emissioni dei trasporti CO<sub>2</sub> che l'*Agenda 2030* si era prefissata non è agevole, a causa di un rallentamento del tasso di riduzione delle emissioni. Le attuali proiezioni collocano la diminuzione delle emissioni nei trasporti al 22% entro il 2050, una percentuale decisamente inferiore alle ambizioni correnti.

Le emissioni di CO<sub>2</sub> nel settore del trasporto passeggeri variano significativamente a seconda della modalità di trasporto utilizzata, che include trasporti su strada, ferroviari, aerei e marittimi. In particolare, il trasporto su strada rappresenta una delle principali fonti di inquinamento atmosferico in Europa. Le autovetture sono responsabili di circa il 60,6% del totale delle emissioni di CO<sub>2</sub>, un dato che evidenzia l'impatto ambientale di questo mezzo di trasporto.



**Figura 1.5:** Ripartizione delle emissioni di gas serra per modalità di trasporto (2019).

Dall'analisi di questi dati emerge chiaramente che la sostituzione, anche parziale, dei mezzi di trasporto a combustione con biciclette o mezzi a minor consumo risulta vantaggiosa. Tuttavia, per una valutazione precisa degli effetti di tale transizione, è necessario considerare in maniera rigorosa le cosiddette **esternalità**.<sup>3</sup>

In termini generali, un'esternalità rappresenta l'insieme degli effetti, negativi o positivi, derivanti dall'attività di produzione o consumo di un soggetto, che incidono sul benessere di un altro soggetto senza che vi sia alcuna forma di compensazione. Nel caso delle esternalità *negative*, l'individuo o l'ente che subisce l'impatto non riceve alcun risarcimento proporzionale ai costi sostenuti. Viceversa, nel caso delle esternalità *positive*, il soggetto beneficiario non corrisponde alcun prezzo equivalente al vantaggio ricevuto.

Gli effetti generati da un'attività, dunque, non si limitano al soggetto che la intraprende, ma si estendono ad altri individui o entità, influenzandone la condizione. Questo implica che il calcolo complessivo dei costi legati a tale attività non si esaurisce nella somma dei soli costi diretti, ma deve includere anche i costi o benefici che ricadono su terzi. Un'esternalità, pertanto, può essere definita come l'effetto di un'attività che grava su soggetti esterni, i quali non hanno avuto alcun ruolo decisionale nella sua realizzazione. Per questo motivo, nell'ambito della ricerca scientifica, il termine *esternalità* viene frequentemente utilizzato con riferimento alle esternalità negative, in quanto la loro mitigazione e il loro controllo generano benefici tangibili per la collettività.

Le esternalità possono essere classificate in due categorie principali, distinzione che si basa sulla natura dell'impatto che l'attività di un soggetto produce su altri, senza che vi sia una corrispondente compensazione economica o un prezzo pagato. Di seguito si riportano le caratteristiche principali di ciascuna tipologia:

- Esternalità *negative*, si hanno quando il soggetto responsabile di impatti negativi non corrisponde al danneggiato un prezzo pari al danno/costo subito (ad esempio un agricoltore che utilizza prodotti chimici che si diffondono nell'ambiente).
- Esternalità *positive*, si hanno quando i soggetti beneficiari di impatti positivi prodotti da un altro soggetto non corrispondono/pagano un prezzo pari ai benefici ricevuti (ad esempio un agricoltore che, coltivando, migliora il paesaggio agricolo).

---

<sup>3</sup>*Esternalità*: gli effetti che l'attività di un'unità economica (individuo, impresa, ecc.) esercita, al di fuori delle transazioni di mercato, sulla produzione o sul benessere di altre unità<sup>[8]</sup>.

Si distinguono in particolare le esternalità di consumo, che si verificano quando il consumo di un bene influenza l'utilità altrui; di produzione, quando l'attività di un'impresa incide su quella di un'altra; e di rete, quando il valore di un bene o servizio dipende dal numero di utilizzatori o dalla disponibilità di beni complementari. Le esternalità negative possono derivare da cause fisiche, chimiche, ecologiche, di rete o sociali e politiche.

Tali effetti impattano su paesaggio, ecosistemi, biodiversità, salute e attività economiche esterne alla transazione, e la difficoltà nell'attribuire responsabilità dirette complica la loro mitigazione. Un esempio è il traffico di rete: quando una rete satura subisce congestionamento, ogni utente aggiuntivo impone un costo sociale agli altri (tempo perso o indisponibilità del servizio). Gli economisti affrontano questo fenomeno con tariffe correttive, mentre questa tesi propone di agire sulle cause strutturali piuttosto che limitarsi a mitigarne gli effetti.

In tale ottica si inserisce il concetto di trasporto condiviso, emerso come soluzione per ridurre le esternalità negative della mobilità tradizionale, soddisfare la domanda di spostamenti urbani e promuovere lo sviluppo economico. Un caso emblematico è il sistema di bike-sharing di Pechino, analizzato in uno studio che valuta le sue esternalità economiche, ambientali e sanitarie [9]. Questo modello si è sviluppato in modo esponenziale negli ultimi anni grazie all'integrazione con applicazioni mobili e all'adozione del sistema *dockless*, che consente il noleggio e la restituzione delle biciclette in qualsiasi luogo pubblico. A differenza dell'Italia, dove pochi operatori dominano il mercato, in Cina esiste una pluralità di piattaforme facilmente riconoscibili da codici cromatici distintivi, a supporto di un bacino d'utenza tra i più ampi al mondo. Il Paese ospita infatti 16 dei 20 più grandi programmi di bike-sharing esistenti.

Lo studio evidenzia tre effetti principali del *bike-sharing*: la promozione dell'intermodalità risolvendo il problema dell'ultimo chilometro; un effetto *sostituzione*, poiché le biciclette riducono l'uso di mezzi inquinanti e il relativo consumo energetico; e un effetto sulla *salute*, grazie alla riduzione dell'inquinamento e alla promozione dell'attività fisica. Queste esternalità positive, come ampiamente discusso nella letteratura, superano i benefici diretti per gli utenti, generando ricadute diffuse in termini di sostenibilità ambientale, salute pubblica ed efficienza dei trasporti urbani.

## 1.5 Impatti della mobilità condivisa sulle città

L'introduzione della mobilità condivisa ha portato numerosi cambiamenti nelle città, specialmente nelle aree urbane densamente popolate. Questo nuovo modello di mobilità urbana ha influenzato non solo il modo in cui i cittadini si spostano, ma anche la pianificazione urbana, l'ambiente e l'economia locale. Il volto delle città si trasforma in risposta alle politiche di mobilità sostenibile, rendendo possibili modelli come quello della *città dei 15 minuti*.<sup>4</sup> Questo approccio, che punta a garantire l'accesso a servizi essenziali come lavoro, istruzione, salute e svago entro un quarto d'ora a piedi o in bicicletta, si integra con politiche di mobilità sostenibile che incidono profondamente sulle abitudini degli utenti dei trasporti, rendendo gli spostamenti più efficienti e a misura d'uomo.

### 1.5.1 La città degli X-minuti

Il concetto della *città dei 15 minuti* può essere amplificato a minutaggi diversi che rispondano ad esigenze diverse, come già fatto in uno studio del 2023 nel Nord Europa [10]. Questo studio, condotto nella città di Utrecht, nei Paesi Bassi, si focalizza sull'implementazione e valutazione del concetto di "*città a X minuti*", un paradigma che mira a garantire accesso a servizi essenziali entro un determinato tempo di viaggio attraverso modalità di trasporto attive, come la bicicletta. In questo testo viene utilizzato un approccio noto come "*2-step Floating Catchment Area*" (2SFCA) per misurare l'accessibilità a determinati tipi di destinazioni, come ad esempio servizi sanitari, educativi o spazi verdi. Questo approccio tiene conto principalmente di due dimensioni dell'accesso ai servizi: l'**accessibilità** e la **disponibilità**. La "*disponibilità*" è legata all'offerta di servizi, mentre l'"*accessibilità*" riguarda la capacità delle persone di raggiungere questi servizi, considerando la distanza e il tempo di viaggio.

Secondo *Penchansky e Thomas (1981)*, l'accesso a un servizio può essere più in generale valutato attraverso cinque dimensioni: *accessibilità*, *disponibilità*, *posizione*, *accessibilità economica* e *accettabilità*. Tuttavia, l'approccio 2SFCA si concentra principalmente su accessibilità e disponibilità, come anticipato.

---

<sup>4</sup>Il concetto è stato reso popolare dal sindaco di Parigi *Anne Hidalgo* e ispirato dallo scienziato franco-colombiano *Carlos Moreno*.



I principali risultati emersi dalla ricerca condotta possono essere sintetizzati nei seguenti punti chiave:

- Il 94% della popolazione può raggiungere almeno un servizio di ogni tipo entro 10 minuti in bici, segno di una rete ciclabile molto efficace.
- Questa percentuale sale al 99,97% entro 15 minuti, confermando l'alta accessibilità complessiva del territorio urbano.
- Persistono però disuguaglianze nelle aree periferiche, con carenze di strutture sportive, sanitarie e di intrattenimento, a causa della concentrazione dei servizi nei centri principali.

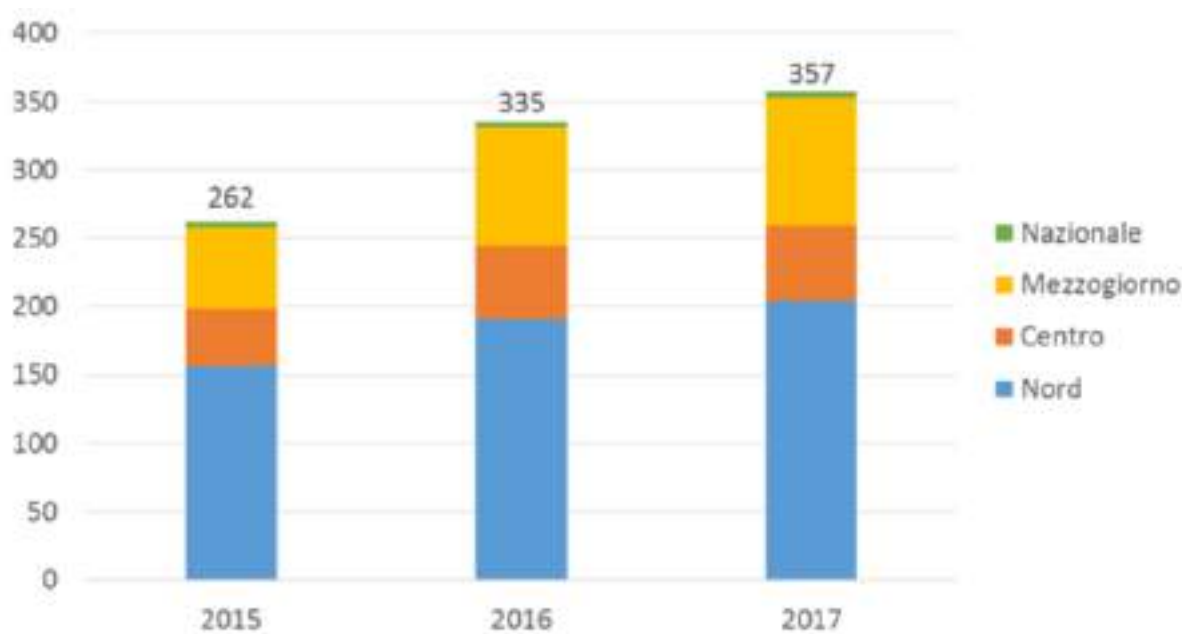
Questi risultati dimostrano l'efficacia della *metrica CS* come strumento analitico per valutare e migliorare l'accessibilità urbana. In particolare, l'applicazione al contesto di Utrecht ha evidenziato sia i punti di forza di una rete ciclabile ben sviluppata, sia le criticità legate all'inclusione delle aree periferiche. Tali evidenze supportano l'urgenza di politiche infrastrutturali mirate e di un approccio olistico nella pianificazione urbana.

Le città italiane, come Bologna e Firenze, potrebbero trarre ispirazione dal modello di Utrecht per promuovere una mobilità sostenibile e accessibile. Attraverso investimenti strategici in infrastrutture ciclabili e una pianificazione che favorisca la distribuzione equa dei servizi, queste città potrebbero allinearsi ai parametri di una "città a X minuti". In particolare, l'adozione di una rete ciclabile integrata avrebbe il potenziale di migliorare sensibilmente la qualità della vita dei residenti, riducendo i tempi di spostamento e aumentando l'accesso a servizi essenziali. Inoltre, potrebbe contribuire significativamente alla riduzione delle disuguaglianze socioeconomiche, garantendo una distribuzione più equa delle opportunità urbane tra il centro e la periferia.

Un ulteriore vantaggio derivante da tale strategia sarebbe il rafforzamento dell'immagine delle città come modelli di sostenibilità e innovazione urbana, capaci di attrarre attenzione e investimenti per lo sviluppo di politiche inclusive e resilienti. Tuttavia, un aspetto cruciale per il successo di questi interventi risiede nel monitoraggio continuo dei loro impatti. Attraverso l'adozione di sistemi di valutazione dinamici, come il DTM [1], sarebbe possibile adattare le politiche alle necessità emergenti, garantendo un progresso costante e sostenibile nel tempo.

## Stato dell'arte: la mobilità condivisa in Italia

La mobilità condivisa italiana cresce e si rafforza come settore nel suo complesso, come danno riscontro il numero di servizi in costante crescita. Nel triennio 2015-2017, infatti, il totale dei servizi di mobilità condivisa, considerando tutti i principali settori di attività (*carsharing*, *bikesharing*, ecc.) è aumentato mediamente del 17% all'anno. Dal punto di vista territoriale, le regioni del Sud sono quelle che hanno fatto registrare una crescita più forte, più 57% nel triennio considerato. Negli stessi anni, l'aumento dei servizi di mobilità condivisa è stato invece pari al 31% sia per il Centro che per il Nord Italia. A fronte di questa crescita, il totale dei servizi sparsi sul territorio italiano al 31 dicembre 2017 era 357, ripartiti con una netta maggioranza nelle regioni del nord Italia, 58% dei servizi totali, il 26% diffusi nelle regioni del Mezzogiorno, il 15% al centro e l'1% di servizi attivi su scala nazionale [11].



**Figura 1.6:** Numero dei servizi di sharing mobility in Italia attivi al 2017.

Nel *Rapporto nazionale sulla Sharing Mobility* [11] si evidenzia il ruolo crescente del *bikesharing* come componente chiave della *sharing mobility* in Italia: al 31 dicembre 2017, il numero totale di sistemi di *bikesharing* in Italia era di 286, con un incremento del 10% rispetto all'anno precedente. Tuttavia, questo tasso di crescita segna un rallentamento rispetto al periodo 2015-2016, quando i sistemi erano aumentati del 40%. Tale rallentamento è stato attribuito all'arrivo sul mercato proprio degli operatori *free-floating* interesse di questa tesi, come **Mobike**, che hanno portato a un riassetto della domanda.

Secondo il *Rapporto nazionale sulla Sharing Mobility* [11] l'introduzione del *free-floating* ha profondamente trasformato il panorama del *bike-sharing*. Il numero di biciclette disponibili è aumentato da circa 16'000 a oltre 39'000 unità nel 2017, con una crescita di quasi 2,5 volte. Questa evoluzione è stata trainata dai modelli di business innovativi degli operatori, che hanno reso le biciclette condivise più accessibili agli utenti. Anche i sistemi *station-based* hanno continuato a crescere, seppur a un ritmo più contenuto (+4% nel biennio 2016-2017), con **Bicincittà** e **Clear Channel** che detengono il 72% del mercato.

La localizzazione dei servizi di *bike-sharing* riflette importanti disparità territoriali. Il Nord Italia rimane l'area con la maggiore concentrazione di sistemi attivi, mentre il Sud Italia ha registrato un incremento modesto, passando dal 26% al 28% del totale dei Comuni con *bike-sharing*. Nonostante ciò, la maggior parte delle biciclette condivise si trova nelle grandi città del Nord e del Centro, come Milano, Torino, Firenze e Roma, che da sole ospitavano oltre due terzi del totale nazionale al momento della registrazione dei dati qui discussi.

Il *Rapporto nazionale sulla Sharing Mobility* [11] analizza poi le prestazioni dei due modelli tecnologici di bike-sharing in nove Comuni italiani, suddivisi per dimensione. I dati mostrano:

- *Percorrenze medie per noleggio*: i servizi station-based registrano un numero di chilometri percorsi per noleggio maggiore del 50-60% rispetto ai servizi free-floating.
- *Noleggi giornalieri per bicicletta*: i servizi station-based mostrano una leggera superiorità rispetto ai servizi free-floating, con un divario meno marcato.
- *Densità di biciclette per km<sup>2</sup>*: i sistemi free-floating hanno una densità maggiore nelle grandi città, con un rapporto di 3 a 1 rispetto ai sistemi station-based. Nelle città medie, questo rapporto scende a 2 a 1.

Il *bikesharing* in Italia si trova in una fase di transizione, caratterizzata dall'espansione di nuovi modelli tecnologici e dal consolidamento dei sistemi tradizionali. Le grandi città continuano a rappresentare il fulcro del mercato, evidenziando il ruolo cruciale delle infrastrutture urbane e della densità abitativa per il successo del servizio. Tuttavia, permangono significative disuguaglianze territoriali, con il Sud Italia che necessita di politiche mirate per colmare il divario con le altre regioni. Il capitolo successivo sarà dedicato alla descrizione degli operatori in attività in Italia, per proseguire poi con un'analisi il cui obiettivo è di aggiornare i dati di cui sopra e valutarne l'evoluzione ed eventuali proposte di miglioramento per il futuro.

## 1.6 Presentazione del servizio

**RideMovi** è uno dei principali operatori di mobilità condivisa in Italia, offrendo soluzioni di bike sharing, scooter sharing e monopattini elettrici in sempre più città in tutto il mondo. Il servizio è stato progettato per promuovere una mobilità urbana sostenibile e integrata, con l'obiettivo di ridurre l'impatto ambientale delle città e migliorare la qualità della vita degli utenti.

Il progetto *RideMovi* nasce con l'intento di fornire un'alternativa sostenibile e flessibile al trasporto privato nelle aree urbane. Grazie a un approccio innovativo, RideMovi ha integrato tecnologie avanzate, come sistemi di geo-localizzazione GPS e app mobili, che consentono agli utenti di localizzare facilmente e prenotare mezzi di trasporto condivisi. Il servizio è disponibile tramite un'app dedicata che permette di monitorare la disponibilità dei veicoli, localizzare la bici o il monopattino più vicino e gestire le tariffe in modo semplice e intuitivo.

*RideMovi* rappresenta una delle principali piattaforme di *bike sharing* in Italia, offrendo un servizio capillare che combina tecnologia avanzata e sostenibilità. Nato con l'obiettivo di promuovere una mobilità urbana ecologica ed efficiente, RideMovi ha sviluppato una flotta moderna composta da biciclette tradizionali ed *e-bikes* (bici elettriche), tutte dotate di tecnologie GPS integrate per il monitoraggio e la localizzazione in tempo reale.

A *Bologna*, il servizio dispone di oltre 2.000 biciclette distribuite in più di 200 stazioni virtuali e aree dedicate, consentendo agli utenti di accedere facilmente ai mezzi tramite l'applicazione mobile. La flotta include biciclette muscolari e biciclette elettriche, che costituiscono ormai gran parte della flotta.

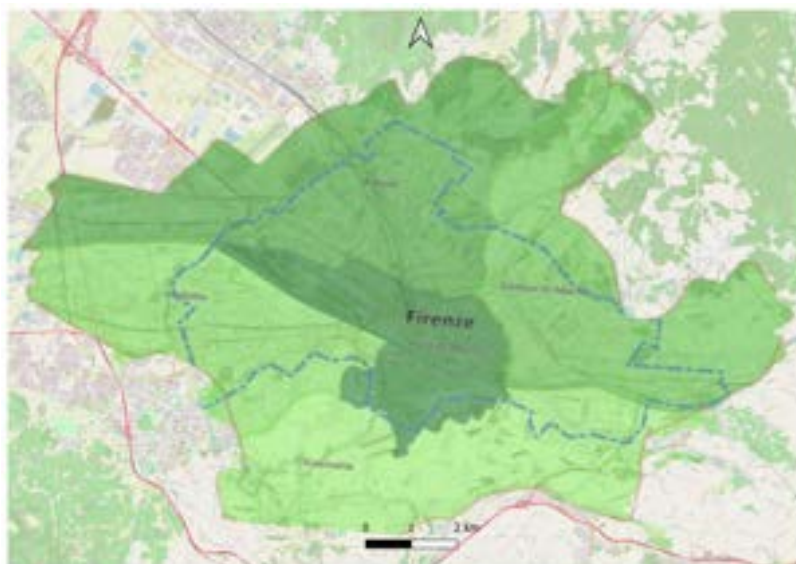
A *Firenze*, la rete RideMovi è altrettanto diffusa, con circa 2.500 biciclette operative, delle quali una parte significativa è costituita da *e-bikes*. Questa scelta è particolarmente efficace per affrontare le colline circostanti e le salite della città, rendendo il servizio accessibile a un'utenza più ampia. Inoltre, l'integrazione con il sistema di trasporto pubblico locale consente di creare sinergie utili per una mobilità multimodale. Il servizio, a differenza di quello felsineo, è arricchito dalla presenza di *monopattini elettrici*.

### **Zone operative di Bologna e Firenze**

Le seguenti immagini mostrano le zone operative per il servizio di mobilità condivisa nelle città di Bologna e Firenze. Queste aree sono state geo-referenziate per permettere una gestione più precisa e funzionale dei veicoli condivisi nelle rispettive città.



**Figura 1.7:** Zona operativa di Bologna per il servizio di mobilità condivisa.



**Figura 1.8:** Zona operativa di Firenze per il servizio di mobilità condivisa.

In entrambe le città, *RideMovi* ha sviluppato un servizio di mobilità condivisa che si è adattato alle esigenze urbane locali. Le zone operative coperte dal servizio in entrambi i Comuni includono sia i quartieri centrali che alcune aree periferiche, consentendo un'accessibilità più o meno estesa e promuovendo l'intermodalità con i trasporti pubblici. L'utilizzo di tecnologie avanzate, come il tracciamento GPS e l'integrazione con applicazioni mobili, consente ai cittadini di monitorare la disponibilità e prenotare facilmente il mezzo più vicino, rendendo l'esperienza della mobilità condivisa immediata e *user-friendly*.

I principali vantaggi di RideMovi includono:

- **Maggiore accessibilità:** con una rete di stazioni virtuali *dock-less* e aree dedicate, gli utenti possono prendere e lasciare i mezzi con maggiore flessibilità.
- **Flessibilità e convenienza:** l'utente ha la libertà di scegliere il mezzo più adatto in base alle proprie esigenze e al contesto urbano, promuovendo spostamenti più dinamici e personalizzati.
- **Tecnologia avanzata:** l'applicazione mobile offre un'interfaccia intuitiva, pagamenti digitali sicuri e un sistema di geo-localizzazione in tempo reale che facilita la gestione e la prenotazione dei mezzi.

Il servizio presenta alcune sfide operative e tecniche, come la necessità di gestire correttamente il parcheggio dei mezzi per evitare ostruzioni, garantire una manutenzione e ricarica regolare e assicurare l'accessibilità a persone con diverse capacità fisiche, adattando infrastrutture e veicoli di conseguenza.

Il funzionamento si basa su un'applicazione mobile che consente di localizzare, prenotare e sbloccare biciclette e monopattini tramite *codice QR*, dopo la registrazione su RideMovi disponibile per iOS e Android. Gli utenti possono scegliere liberamente i veicoli e parcheggiarli all'interno delle aree designate, grazie a un sistema a stazioni virtuali che offre maggiore flessibilità rispetto ai modelli tradizionali con postazioni fisse.

Le tariffe sono calcolate in base al tempo d'uso, con opzioni che spaziano da pochi minuti a intere giornate, e possono essere pagate direttamente *in-app*. Il sistema integra funzioni avanzate di geolocalizzazione, che mostrano in tempo reale la posizione e il livello di carica dei mezzi, mentre un sistema automatico di gestione e manutenzione assicura che la flotta sia costantemente funzionante e distribuita in modo efficiente sul territorio.

## 1.7 Confronto con altri servizi di Bike Sharing

Il mercato del *bike sharing* ha visto un'espansione significativa negli ultimi anni, con molteplici operatori che offrono soluzioni simili a quelle di *RideMovi*, ma con alcune differenze in termini di operatività, strategia e scala. A livello internazionale, aziende come *Lime*, *Ofo* e *Mobike*<sup>5</sup> hanno dominato la scena, portando a un rapido aumento della concorrenza e, in alcuni casi, anche alla saturazione del mercato.

L'affermazione del fenomeno a livello globale non è stata omogenea. In Nord America e in molte capitali europee, i modelli di business hanno puntato su una crescita graduale, supportata da accordi con le amministrazioni locali e da sistemi tariffari integrati con altre forme di mobilità. In Asia, e in particolare in Cina, la diffusione è stata invece estremamente rapida, favorita da bassi costi di produzione dei veicoli e da un accesso quasi illimitato al capitale di rischio. Tale approccio ha portato a una vera e propria esplosione di operatori, spesso in assenza di un adeguato quadro normativo.

Il fenomeno dell'*oversupply* si è manifestato come una delle principali criticità del modello cinese. Secondo un articolo di *The Atlantic* [13] del 2018, molte città hanno visto enormi accumuli di biciclette abbandonate e danneggiate, a causa di un'offerta che ha superato la domanda in molte aree. Come sottolinea la rivista:

*“The rapid growth vastly outpaced immediate demand and overwhelmed Chinese cities.”*<sup>6</sup>

Le conseguenze di questa dinamica non si sono limitate all'impatto estetico o ambientale delle cosiddette *bike graveyards*, le vaste aree colme di biciclette inutilizzate documentate da numerosi reportage giornalistici. Gli operatori, tra cui *Ofo* e la stessa *Mobike*, si sono trovati a dover affrontare costi insostenibili di raccolta, stoccaggio e manutenzione dei mezzi, con un conseguente ridimensionamento delle attività o, in alcuni casi, il completo ritiro dal mercato. L'eccessiva competizione ha infatti ridotto i margini di profitto, mostrando i limiti di un'espansione basata unicamente sulla quantità di veicoli distribuiti.

---

<sup>5</sup>Società che nel novembre 2019 è stata acquistata dalla S.r.l. italiana *RideMovi* [12].

<sup>6</sup>Trad.: “La rapida crescita ha superato di gran lunga la domanda immediata, travolgendo le città cinesi.”



In contrasto con la situazione cinese, il modello europeo, e in particolare quello italiano, si caratterizza per un approccio più misurato. A Bologna e Firenze, *RideMovi* calibra la distribuzione dei veicoli sulla base della domanda locale, con un numero controllato di biciclette e monopattini. Questa strategia evita il sovraffollamento e garantisce un'elevata accessibilità, migliorando la percezione del servizio da parte degli utenti. La gestione efficiente del parco veicolare e gli interventi regolari di manutenzione riducono il rischio di danneggiamenti e abbandoni, problematiche comuni in altri contesti.

Il confronto evidenzia come il successo del *bike sharing* non dipenda solo dalla quantità di mezzi disponibili, ma dall'equilibrio tra offerta, infrastruttura urbana e sostenibilità economica. RideMovi, pur operando su scala più contenuta rispetto ai colossi internazionali, sembra aver saputo trarre vantaggio dall'esperienza di altri operatori, adottando un modello orientato alla qualità del servizio e alla stabilità di lungo periodo.



**Figura 1.9:** Un ciclista pedala davanti a un enorme cumulo di biciclette condivise inutilizzate a Xiamen, in Cina.



## Capitolo 2

### Metodi

## 2.1 Strumenti e software

Per indagare l'impatto della mobilità condivisa sulla rete urbana e confrontare i dati empirici con modelli di simulazione, è fondamentale definire con chiarezza gli strumenti utilizzati, la natura dei dati analizzati e le metodologie applicate.

Questo capitolo presenta dunque gli strumenti software adottati, descrive i dataset a disposizione e introduce le tecniche di analisi esplorativa e di modellizzazione matematica impiegate nello studio. L'elaborazione dei dati e le simulazioni sono state condotte mediante strumenti open source e librerie scientifiche, selezionati per garantire efficienza computazionale e riproducibilità dei risultati.

### Python

Il linguaggio `Python` è stato scelto come principale ambiente di sviluppo per questo lavoro grazie alla sua versatilità, facilità d'uso e ampia disponibilità di librerie scientifiche. In particolare, `NumPy`, una delle librerie fondamentali di `Python`, è stata utilizzata per la gestione degli array e per l'elaborazione numerica efficiente dei dataset. La combinazione di `NumPy` e altre librerie ha permesso di manipolare grandi quantità di dati senza penalizzare le prestazioni, consentendo operazioni di aggregazione, filtraggio e calcolo statistico su milioni di record, tra cui la generazione di medie, varianze, somme cumulative e l'applicazione di funzioni matematiche vettorializzate. Grazie a queste funzionalità, è stato possibile preparare i dati in forma strutturata e pronta per le successive fasi di analisi esplorativa e modellizzazione.

### SUMO

*SUMO – Simulation of Urban MObility* [14] è un simulatore microscopico di traffico open source, adottato per replicare scenari urbani realistici e valutare l'impatto dei flussi di mobilità condivisa sulla rete viaria. SUMO è stato scelto per la sua capacità di modellare con precisione interazioni tra veicoli e infrastruttura urbana, consentendo di simulare scenari realistici e confrontare i risultati con dati empirici. In particolare, SUMO permette di:

- definire nodi (intersezioni) e archi (tratti stradali) con capacità e limiti di velocità;
- generare itinerari veicolari derivati da dati reali o sintetici;
- modellare interazioni tra differenti modalità di trasporto, includendo veicoli privati e condivisi.

## 2.2 Natura dei dati

Le analisi condotte mirano a fornire un quadro puntuale sull'uso del bike sharing nelle città di Bologna e Firenze, limitatamente all'arco temporale analizzato. Il dataset utilizzato ha origine da file CSV di grandi dimensioni fornito dall'azienda RideMovi, composto da milioni di righe, corrispondenti a singoli mesi raggruppati poi in anni. Per lo sviluppo delle successive analisi è stato selezionato, a titolo esemplificativo, il mese di maggio dell'anno 2023, per entrambe le città. Tale scelta è motivata dalle condizioni climatiche tipicamente miti che caratterizzano questo periodo, le quali favoriscono l'utilizzo dei servizi di mobilità condivisa, come il *bike sharing*. Similmente, anche i mesi di aprile, settembre e ottobre risulterebbero particolarmente indicativi per lo studio di tali fenomeni.

Il processo di analisi si basa sull'integrazione di due tipologie di dataset distinti, ognuno caratterizzato da una struttura propria e da un differente livello di dettaglio informativo. La prima tipologia corrisponde ai dataset forniti da *RideMovi*, contenenti i *raw data* (i dati grezzi) dei viaggi effettuati; la seconda, invece, è costituita da dataset ottenuti tramite *map-matching* in ambiente laboratoriale con l'ausilio di *HybridPy*, a partire dai dati originali forniti dall'azienda.

La principale differenza tra i due insiemi di dati riguarda i dati sulle *distanze percorse*: nel primo caso, tali distanze sono calcolate in maniera semplificata come rapporto spazio/tempo ( $s/t$ ), mentre nel secondo caso derivano da un processo di *map matching* avanzato, che tiene conto di variabili complesse come i costi associati agli archi della rete, le condizioni del traffico, le preferenze di percorso dell'utente e altri fattori contestuali, restituendo così una stima più realistica e aderente al comportamento reale di spostamento.

Il primo dataset, fornito direttamente da RideMovi per la città di Bologna, contiene i dati grezzi relativi ai noleggi effettuati. Per ogni noleggio sono riportati il timestamp di avvio e quello di conclusione, insieme alle coordinate geografiche di partenza e di arrivo (*latitudine* e *longitudine*). Inoltre, il dataset include la distanza percorsa in chilometri, un identificativo univoco della bicicletta (*BikeID*) e la tipologia del veicolo utilizzato, distinguendo tra biciclette muscolari e biciclette elettriche.

Il secondo dataset, anch'esso fornito da *Ridemovi*, riguarda invece la città di Firenze. Rispetto a quello di Bologna, presenta una struttura più dettagliata e un formato standardizzato a livello internazionale. Nonostante il maggior numero di campi, le informazioni fondamentali per l'analisi sono sostanzialmente analoghe a quelle del dataset bolognese. Tra queste, si evidenziano la data e l'ora di inizio e fine del noleggio (*start\_time* e *end\_time*), le coordinate di partenza e arrivo (*start\_latitude/longitude* e *start\_latitude/longitude*), la distanza percorsa (*ride\_distance*, espressa in metri) e il tipo di veicolo utilizzato (*vehicle\_type*).

Il terzo dataset deriva dal processo di map-matching applicato tramite l'algoritmo *HybridPy* ai dati di Bologna e Firenze. Questa elaborazione arricchisce i dati originali fornendo informazioni dettagliate sul percorso effettivo, ricostruito sulla rete stradale.

La struttura del dataset è decisamente più complessa e include diverse metriche derivate dal processo di matching. Tra le colonne più rilevanti ai fini dell'analisi si segnalano la durata del viaggio (*GPS duration [s]*), la distanza percorsa (*GPS distance [m]*), la velocità media e massima (*Av. speed [m/s]* e *Max. speed [m/s]*) e il tipo di veicolo (*Type*), mantenuto dai dati originali.

La presenza di queste tre diverse strutture dati ha reso necessaria una fase di *preprocessing* e armonizzazione, finalizzata a standardizzare i nomi delle colonne, le unità di misura (ad esempio convertendo i metri in chilometri) e i formati, creando così un unico dataset consolidato su cui basare le analisi successive.

## 2.3 Natura dell'analisi

L'analisi condotta si colloca a metà strada tra l'approccio tradizionale alla raccolta e al trattamento manuale dei dati e le tecniche avanzate di analisi computazionale, capaci di elaborare grandi volumi di dati eterogenei. I dati utilizzati in questo lavoro sono classificabili come *Big Data*, sia per l'estensione sia per la granularità<sup>1</sup> spaziale e temporale delle informazioni. L'approccio adottato è ibrido: strumenti tradizionali vengono affiancati a strumenti IT, in particolare Python, per automatizzare operazioni di filtraggio, pulizia e strutturazione dei dataset, sfruttando la potenza computazionale moderna pur mantenendo un controllo diretto sul processo analitico.

A questo proposito, il Technical Paper ”*Elementi metodologici per l'analisi della mobilità delle persone attraverso l'uso di big data*” [15], prodotto dal FS RESEARCH CENTRE, approfondisce l'impiego dei Big Data per colmare le lacune informative dei metodi tradizionali e migliorare la comprensione dei comportamenti di mobilità. Il documento individua indicatori chiave per descrivere dinamiche di trasporto e comportamenti individuali e sottolinea le caratteristiche dei Big Data, sintetizzate nelle “6 V”: volume, velocità, varietà, veridicità, valore e variabilità. Gestire grandi moli di dati eterogenei richiede un approccio metodologico rigoroso, capace di assicurare qualità, coerenza e affidabilità, e di estrarre informazioni significative. Pur presentando criticità legate a frammentarietà e interoperabilità, i Big Data offrono una visione più completa e aggiornata della mobilità, a condizione di adeguati investimenti in risorse, competenze e infrastrutture, nel rispetto di principi etici e normativi.

Nel caso specifico, i dati derivano dall'uso dell'applicazione *RideMovi* per il noleggio di biciclette. Le informazioni spaziali e temporali sono ottenute come risultato secondario dell'interazione degli utenti con il servizio, non tramite rilevamento progettato ad hoc per finalità analitiche; di conseguenza, i dati possono risultare passivi, parziali e soggetti a *bias* legati alle modalità di utilizzo dell'applicazione. Durante l'analisi sono emerse criticità tipiche dei Big Data, come valori anomali, dati mancanti, incoerenze temporali e difficoltà di georeferenziazione puntuale. Tali problematiche sono state gestite con procedure automatizzate di pulizia, filtraggio e normalizzazione implementate in Python, garantendo un dataset coerente, stabile e pronto per l'analisi quantitativa.

---

<sup>1</sup> *Granularità*: livello di suddivisione del sistema di dati (*grossa, fine*).

L'analisi della mobilità può essere condotta da diverse prospettive; quella oggetto della presente indagine riguarda il comportamento individuale, osservando spostamenti quotidiani, durata e distanza dei percorsi. L'uso di fonti dati innovative permette di arricchire la comprensione dei flussi urbani e delle dinamiche che li governano.

A livello internazionale, la mobilità è inclusa nei principali strumenti di valutazione dello sviluppo sostenibile. I Sustainable Development Goals (SDGs) delle Nazioni Unite attribuiscono un ruolo centrale alla mobilità nella transizione ecologica, con riferimento a salute pubblica, infrastrutture resilienti e mitigazione dei cambiamenti climatici. Anche il Rapporto BES dell'ISTAT e il PNRR sottolineano l'importanza della mobilità sostenibile, promuovendo indicatori sull'accessibilità e interventi digitali e innovativi. Il monitoraggio dei fenomeni di mobilità è supportato da raccolte di indicatori internazionali curate da Eurostat, UNECE e OCSE, che permettono confronti tra paesi e regioni europee.



**Figura 2.1:** Alcuni degli obiettivi di sviluppo sostenibile legati alla mobilità.

L'integrazione di dati tradizionali e digitali permette di superare le limitazioni dei singoli dataset e di costruire modelli interpretativi più solidi, utili per comprendere i comportamenti di mobilità e supportare politiche pubbliche efficaci. In Italia, tuttavia, manca ancora un sistema informativo unificato: i Piani Urbani di Mobilità Sostenibile e i piani casa-lavoro sono strumenti frammentari e non interoperabili.

L'uso dei Big Data offre opportunità significative, consentendo raccolta continua, su larga scala e a costi ridotti rispetto a metodi tradizionali come interviste o rilevazioni sul campo. Pur garantendo elevata copertura, i Big Data richiedono un ciclo di vita del dato specifico per la loro dimensione, velocità e valore, che include raccolta, integrazione, trasformazione, analisi e validazione. La standardizzazione dell'accesso ai dati è fondamentale, poiché molte informazioni utili sono detenute da soggetti privati; il Data Act europeo [16] promuove la condivisione obbligatoria per finalità di interesse pubblico. Per la tutela della privacy, è essenziale adottare Privacy Enhancing Technologies (PET), mentre errori non campionari, legati alla qualità o rappresentatività dei dati, devono essere controllati costantemente.

### 2.3.1 Analisi esplorativa dei dati (EDA)

Per valorizzare i dati disponibili, è essenziale disporre di strumenti metodologici come l'*Exploratory Data Analysis* (EDA), che permette di comprendere la struttura, le caratteristiche e le anomalie dei dataset [17]. L'EDA consente di familiarizzare con i dati, ispezionare le variabili (numeriche, categoriali o temporali) e gestire valori mancanti o anomalie, come ad esempio tragitti con velocità o durate non realistiche.

L'approccio è necessariamente descrittivo e adattabile alle specificità del dataset. Nel caso in esame, l'EDA comprende sotto-analisi, ciascuna focalizzata su un aspetto specifico del fenomeno, come il tempo, la distanza o altre variabili rilevanti. L'obiettivo finale è tradurre le scoperte tecniche in insight chiari e utilizzabili, supportando interpretazioni e decisioni successive, tramite l'utilizzo di strumenti matematici e statistici tra cui rientrano, in particolare, le distribuzioni gaussiane, le correlazioni lineari e altre tecniche di analisi quantitativa che costituiscono la base teorica delle elaborazioni presentate.

#### Analisi generale

L'analisi *generale* è la principale delle sei sotto-analisi, perché permette di descrivere con efficienza i parametri fondamentali che caratterizzano il dataset, mediante espressive rappresentazioni grafiche. Una volta ricevuto in input il dataset formattato, quest'analisi genera un report testuale e quattro grafici.

L'analisi dei dati ha prodotto informazioni statistiche e visualizzazioni utili per comprendere la distribuzione delle distanze dei tragitti. Le principali misure descrittive includono tendenze centrali, come media e mediana, e indicatori di dispersione e forma della distribuzione, che consentono di cogliere pattern generali, eventuali valori estremi e caratteristiche di simmetria o asimmetria.

A complemento delle statistiche, l'analisi ha prodotto quattro grafici esplorativi: una distribuzione delle distanze, che permette di osservare la forma complessiva dei dati; un QQ-plot che consente di confrontare la distribuzione osservata con una normale teorica; una distribuzione cumulativa (CDF) e una curva normale teorica, che facilitano ulteriormente l'interpretazione dei dati e il confronto tra la distribuzione reale e quella attesa. Questi strumenti, insieme alle statistiche descrittive, forniscono una base solida per le successive fasi di analisi e modellazione.

Tutte le analisi precedenti sono volte a verificare la validità dell'ipotesi di *normalità*, ovvero l'assunzione che una variabile casuale segua una distribuzione normale (o gaussiana). Questa distribuzione, di forma a campana e simmetrica rispetto alla media, è caratterizzata da una concentrazione della maggior parte dei dati intorno al valore centrale, con probabilità che decrescono all'aumentare della distanza da esso. La distribuzione normale riveste un ruolo centrale nella statistica inferenziale, poiché molte tecniche e test classici si basano proprio su tale assunzione.

Le ipotesi di normalità assumono particolare rilevanza nel calcolo di intervalli di confidenza o nell'uso di modelli probabilistici. In questi casi, la validità dei risultati è legata alla correttezza dell'assunzione che i dati provengano da una popolazione distribuita normalmente. Se tale condizione non è rispettata, l'affidabilità dei risultati può ridursi in modo significativo: i *p-value*<sup>2</sup> possono essere distorti, la potenza statistica compromessa, e le stime dei parametri poco attendibili.

«All models are wrong, but some are useful.»

---

— *George Box*

Secondo le parole del celebre statistico *George Box* in un paper del 1976, l'assunzione a priori della normalità, seppure conveniente per i test statistici, è spesso non coerente con i real-world data, cioè con il reale comportamento della popolazione statistica. Verificare la normalità, con un «*pragmatic and critical approach to statistical modeling rather than a dogmatic adherence to any single method*», è dunque una fase preliminare essenziale in molte analisi; le modalità di verifica includono sia strumenti grafici (come istogrammi, QQ-plot, boxplot) che test statistici specifici.

Valutare l'aderenza alla normalità non è solo un esercizio formale, ma una pratica necessaria per garantire la validità delle inferenze e l'adeguatezza dei modelli statistici utilizzati. Nel caso di studio in analisi, in seguito al filtraggio e l'eliminazione dei valori anomali, è stato riscontrato un comportamento statisticamente normale del dataset.

---

<sup>2</sup>*p-value*: è la probabilità, sotto l'ipotesi nulla, di osservare un risultato pari o più estremo rispetto a quello ottenuto.



## **Analisi spaziale**

L'analisi spaziale è stata condotta per studiare la distribuzione geografica dei punti di partenza e di arrivo all'interno del dataset e per quantificare la densità spaziale dei flussi di mobilità. Questo tipo di analisi permette di identificare le aree con maggiore concentrazione di traffico, note come hotspot, nonché le zone meno frequentate.

Per effettuare l'analisi sono state utilizzate metriche di densità spaziale calcolate su griglie geografiche, insieme a metodi di clustering per raggruppare i punti in aree omogenee dal punto di vista della frequenza dei tragitti. L'approccio consente di ottenere una rappresentazione quantitativa della distribuzione dei flussi di mobilità sulla rete urbana, fornendo informazioni essenziali per la modellizzazione successiva e per l'interpretazione dei pattern spaziali dei dati.

## **Analisi applicativa**

L'analisi applicativa si concentra sul comportamento effettivo degli utenti e sulle modalità di utilizzo dei veicoli, spostando il focus dalle metriche di performance generali ai modelli di impiego pratico. Lo scopo di questa analisi è segmentare i dati per far emergere le abitudini e le preferenze degli utenti, fornendo indicazioni quantitative basate sulla domanda reale.

In particolare, l'analisi prevede il calcolo della distanza media di tutti i tragitti, al fine di ottenere un indicatore sintetico del tipo di utilizzo predominante. Successivamente, ogni viaggio viene classificato in fasce di distanza predefinite, identificate come tragitti brevi, medi e lunghi, per quantificare la distribuzione della domanda sulle diverse tipologie di spostamento.

Infine, l'analisi confronta l'utilizzo dei diversi tipi di veicolo, distinguendo tra e-bike e biciclette muscolari. Questo approccio consente di valutare le preferenze degli utenti e il livello di adozione di ciascuna tecnologia, fornendo informazioni essenziali per eventuali strategie di gestione e ottimizzazione della flotta.

## Analisi tempi/velocità

L'analisi comparativa dei tempi e delle velocità mira a valutare le prestazioni dei tragitti effettuati con biciclette muscolari ed elettriche, considerando come metriche principali la durata del viaggio e la velocità media. Per garantire la robustezza dei risultati, i dati vengono filtrati per escludere spostamenti anomali, come viaggi troppo brevi o troppo lunghi e velocità fuori intervallo.

Il modulo produce statistiche descrittive per ciascun tipo di bicicletta, includendo numero di viaggi, media, mediana e intervallo dei valori per tempi e velocità. L'analisi considera anche la distribuzione dei dati, la frequenza dei tempi di percorrenza e delle velocità medie, nonché le relazioni tra distanza, durata e velocità, permettendo di descrivere i pattern d'uso e di confrontare le prestazioni dei due tipi di veicolo.

## Analisi temporale

L'analisi temporale indaga quando il servizio viene utilizzato, evidenziando ritmi e abitudini degli utenti. A differenza delle analisi precedenti, focalizzate su “dove” e “come”, questa sezione esplora il “quando” per rivelare pattern giornalieri e settimanali. Comprendere le fluttuazioni temporali è fondamentale per il ribilanciamento della flotta, la manutenzione e la definizione di strategie tariffarie. L'analisi è comparativa tra biciclette muscolari ed elettriche, permettendo di osservare sia le tendenze aggregate sia le differenze di comportamento tra i due tipi di veicolo, ad esempio l'uso delle e-bike nelle ore di punta o la distribuzione della domanda nella settimana.

Le visualizzazioni principali includono la *distribuzione oraria dei tragitti*, tramite istogramma dei viaggi avviati per ogni ora, e la *distribuzione settimanale*, tramite grafico a barre per giorno della settimana. Questi strumenti permettono di distinguere tra utilizzo sistematico (feriali) e occasionale o ricreativo (weekend).

## Analisi dei turnover

L'analisi dei turnover valuta l'efficienza operativa e la distribuzione del carico di lavoro nella flotta. Dopo la conversione delle date, la normalizzazione delle distanze e la pulizia dei dati, viene effettuata l'anonimizzazione degli ID: a ogni bicicletta viene assegnato un ID fittizio (es. “Bici 1”) ordinato per numero totale di viaggi, garantendo privacy e comparabilità delle performance. L'analisi genera un report testuale con statistiche sintetiche (numero di biciclette, media, mediana, min e max viaggi) e due visualizzazioni principali: un istogramma del numero di viaggi per veicolo, utile per evidenziare bici sottoutilizzate o sovraccariche, e la funzione di distribuzione cumulativa (CDF), che mostra la percentuale della flotta con un numero di viaggi inferiore o uguale a una soglia, fornendo una misura sintetica dell'equilibrio nell'impiego dei mezzi.

## Capitolo 3

### Risultati

## 3.1 Presentazione dei risultati

L'esecuzione dello script di analisi ha prodotto una serie di grafici, visualizzazioni e report numerici che consentono di valutare lo stato di salute e la funzionalità complessiva del servizio di bike sharing analizzato. Le analisi elencano le caratteristiche dei viaggi in base state a fascia oraria, tipologia di mezzo e distanze percorse, permettendo una lettura dettagliata delle dinamiche d'uso.

I risultati delle analisi si riferiscono a 4 dataset diversi:

- **bologna\_2023\_05.csv**: il dataset analizza i singoli noleggi del servizio di bike sharing e include, per ciascun tragitto, l'orario di inizio e fine, le coordinate geografiche di partenza e arrivo, la distanza percorsa in chilometri, l'identificativo del mezzo utilizzato e la tipologia di bici.
- **firenze\_2023\_05.csv**: il dataset raccoglie informazioni sui singoli noleggi effettuati dagli utenti, includendo un identificativo univoco dell'ordine, il paese e la città in cui è avvenuto il noleggio, il codice del veicolo e la sua tipologia, l'ID utente, l'orario di inizio e fine corsa, le coordinate GPS di partenza e arrivo, la durata del tragitto e la distanza percorsa.
- **bikesharing\_bologna\_2023\_05\_trips.csv**: i dataset contengono i risultati dell'allineamento tra tragitti GPS e percorsi di riferimento, includendo un identificativo univoco del tragitto e del GPS, un flag di selezione, il timestamp di inizio, la durata e distanza registrate via GPS, le elevazioni positive e negative, le velocità media e massima, il tipo e lo scopo dello spostamento, i dispositivi utilizzati, gli ID dei percorsi matched, più brevi e più veloci, varie misure di accuratezza del matching spaziale e temporale, indici di qualità come l'errore di distanza e il tempo di calcolo, e infine gli ID della persona e del veicolo associati al tragitto.
- **bikesharing\_firenze\_2023\_05\_trips.csv**: dello stesso tipo del dataset *mapmatched* di Bologna.

## 3.2 La città di Bologna

### 3.2.1 Analisi generale

Il *report* riassume i risultati ottenuti mediante l'analisi generale:

Misura	Valore	Misura	Valore
Media distanza	2.48 km	Range	0.15 - 15.00 km
Mediana distanza	2.12 km	IQR	1.80 km
Deviazione standard	1.63 km	Skewness	1.89
Varianza distanza	2.67 km <sup>2</sup>	Kurtosis	6.46

**Tabella 3.1:** Statistiche descrittive delle distanze dei tragitti.

L'analisi mostra come la maggior parte dei tragitti sia concentrata su distanze brevi, con la mediana inferiore alla media, suggerendo la presenza di alcuni viaggi più lunghi. La distribuzione è positivamente asimmetrica (*skewness* > 0) e leptocurtica (*kurtosis* > 3), indicando code lunghe verso valori elevati e una maggiore concentrazione dei dati attorno al centro rispetto a una distribuzione normale.

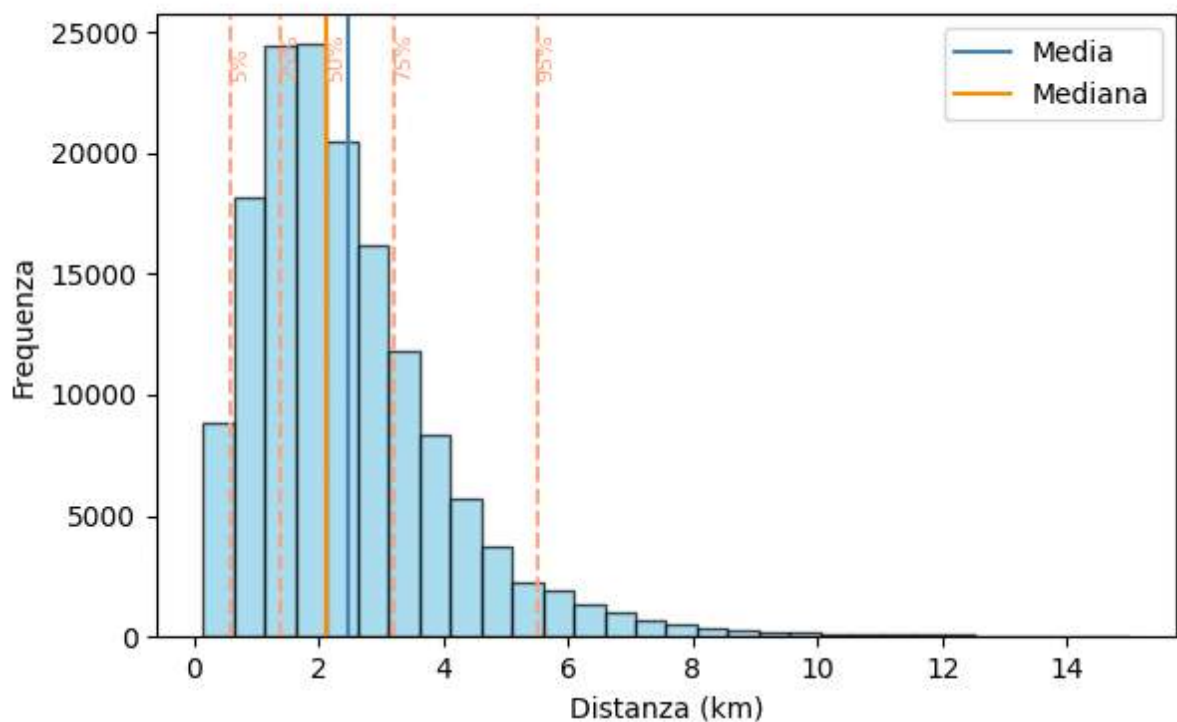
L'analisi esplorativa condotta sul dataset `bologna_2023_05.csv` ha prodotto i seguenti grafici:

- *Distribuzione* delle distanze: si osserva una distribuzione asimmetrica positiva (con coda a destra), che indica una maggiore concentrazione di osservazioni su distanze brevi. La maggior parte dei dati si colloca infatti al di sotto dei 4 km. Il valore mediano della distanza, rappresentato dalla linea verde, è di circa 2.2 km, indicando che il 50% delle osservazioni ha una distanza inferiore a questo valore. La media, pari a circa 2.5 km (linea blu), è leggermente superiore alla mediana; questa discrepanza è tipica delle distribuzioni con asimmetria positiva, dove i valori più elevati nella coda influenzano il calcolo della media.

L'analisi dei percentili (linee arancioni tratteggiate) fornisce ulteriori dettagli: il 25% delle distanze è inferiore a circa 1.5 km (primo quartile) e il 75% è inferiore a 3.5 km (terzo quartile). Di conseguenza, il 50% centrale dei dati si concentra in un intervallo di circa 2 km. Infine, si nota come solo il 5% dei valori sia inferiore a 0.5 km e il 95% sia inferiore a 5.5 km, evidenziando la rarità di distanze superiori a tale soglia.

Percentile	Distanza (km)
5°	0.60
25°	1.40
50°	2.12
75°	3.20
95°	5.50

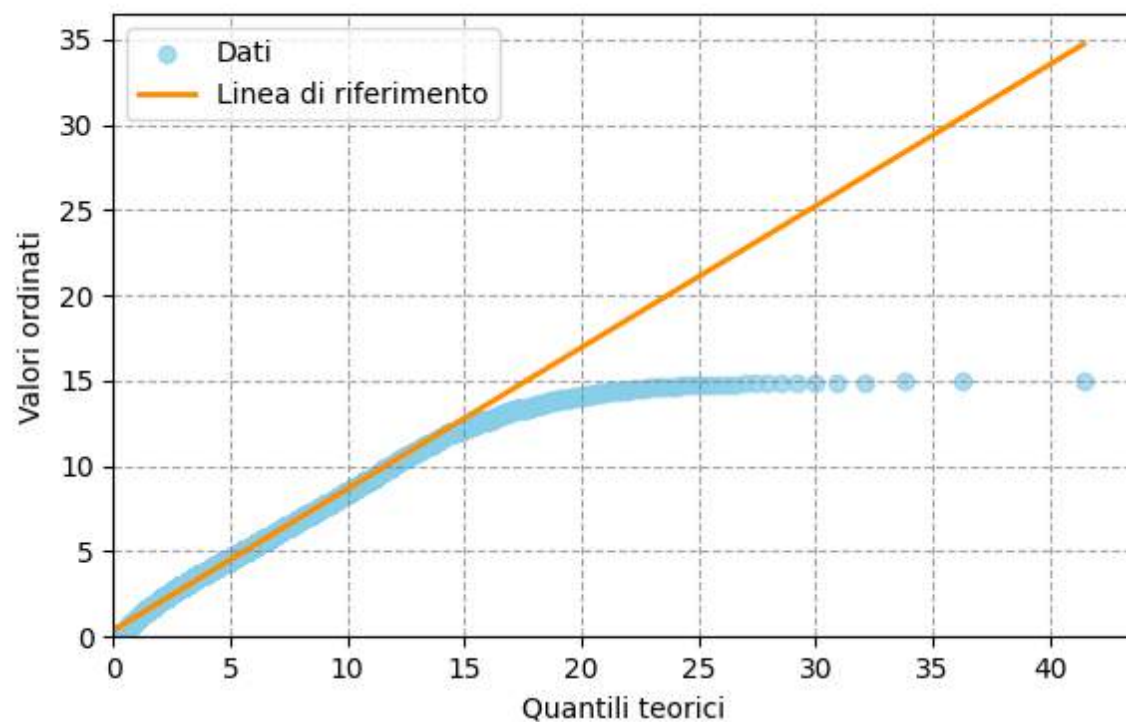
**Tabella 3.2:** Distribuzione delle distanze percorse ai diversi percentili.



**Figura 3.1:** Istogramma della distribuzione delle distanze per il dataset di Bologna.

- *QQ-plot* (Quantile-Quantile plot): il grafico mostra una chiara e sistematica deviazione dei punti (in blu) dalla retta di riferimento teorica (in arancione), confermando che la distribuzione delle distanze non è normale.

La deviazione osservata presenta una forma concava: per valori bassi, i punti si trovano vicini alla retta, ma al crescere dei quantili teorici, i valori campionari aumentano in modo meno che proporzionale. I “Valori ordinati” nella parte superiore del grafico mostrano un appiattimento rispetto alla retta teorica, segnalando una minore frequenza di valori elevati rispetto a quanto previsto da una distribuzione normale. Alcuni punti isolati all’estrema destra si discostano visibilmente dal resto della nuvola, posizionandosi in un intervallo poco popolato.

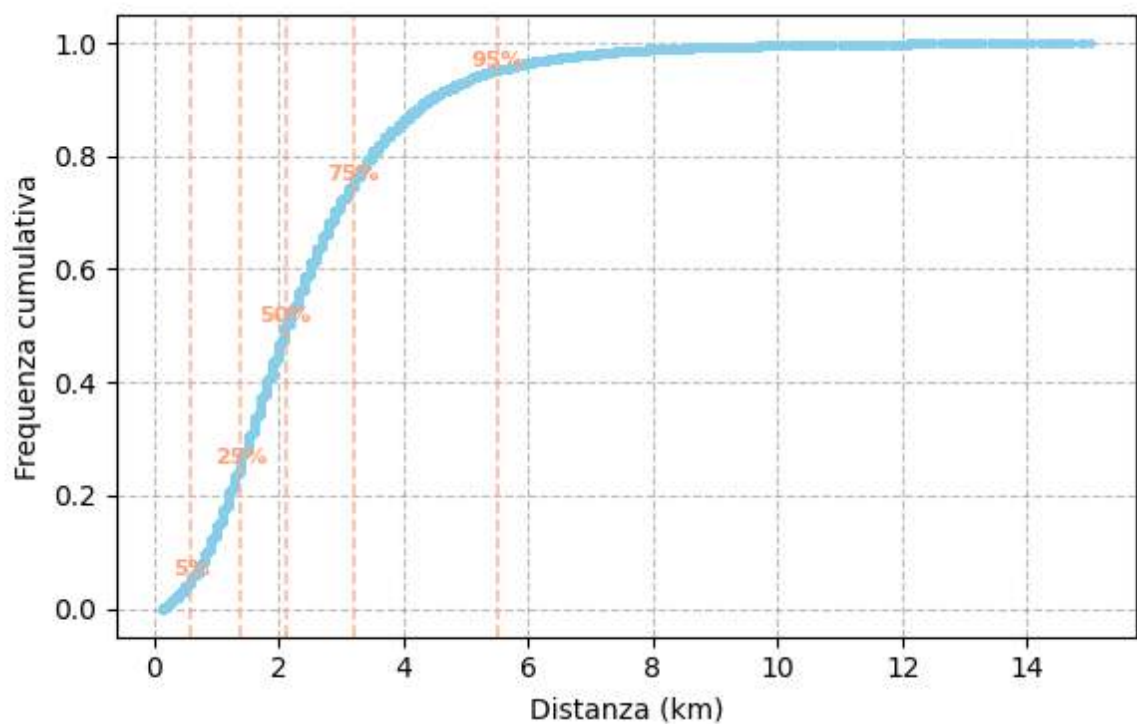


**Figura 3.2:** Quantile-Quantile (QQ) plot della distribuzione delle distanze.

- *Distribuzione cumulativa (CDF)*: la curva illustra la proporzione cumulata di osservazioni al variare della distanza. La forte pendenza iniziale, seguita da un appiattimento progressivo, conferma visivamente la concentrazione della maggior parte dei dati su distanze brevi.

Dal grafico è possibile individuare con precisione i principali percentili della distribuzione. Il 50% delle distanze (mediana) è inferiore a circa 2,2 km.

L'intervallo interquartile, che rappresenta il 50% centrale dei dati, si estende da circa 1,5 km (25° percentile) a 3,5 km (75° percentile). La curva cumulativa mostra inoltre che il 95% delle osservazioni si colloca al di sotto di 5,5 km, mentre solo una quota contenuta si trova oltre tale valore.



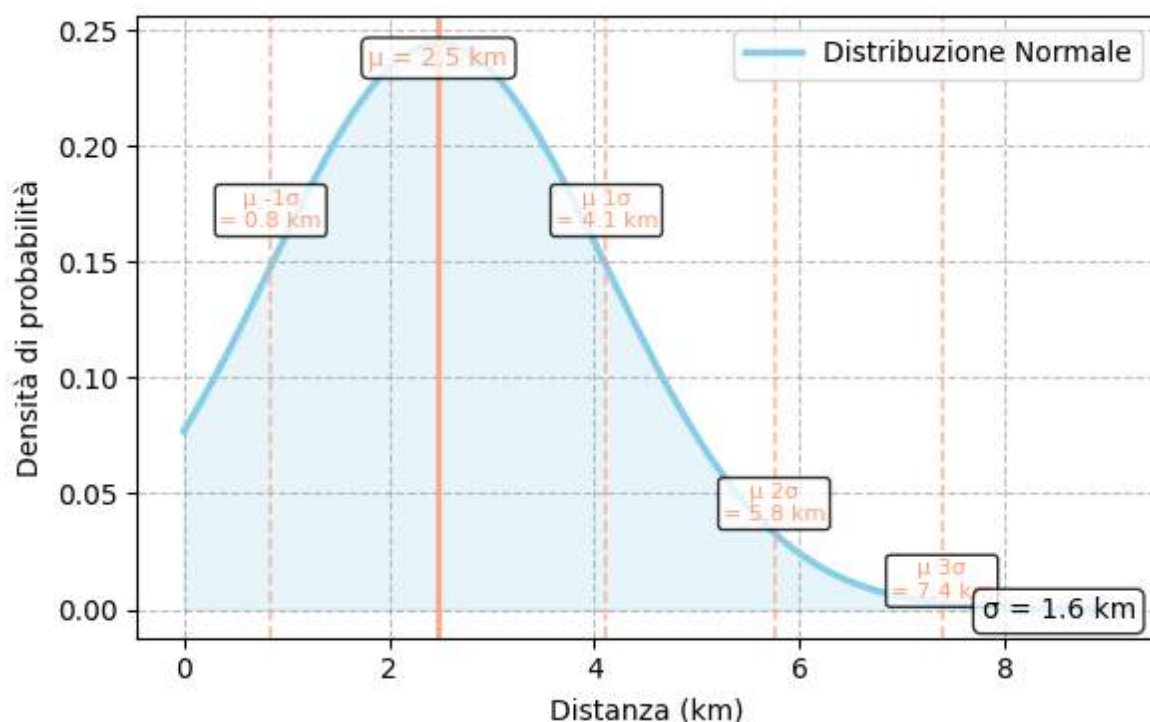
**Figura 3.3:** Funzione di ripartizione (CDF) delle distanze.



- *Distribuzione normale teorica*: la curva gaussiana rappresenta la densità di probabilità attesa assumendo una distribuzione normale delle distanze percorse. La forma simmetrica a campana è centrata attorno al valore medio  $\mu = 2,5$  km, che corrisponde al picco massimo della distribuzione.

La deviazione standard  $\sigma = 1,6$  km determina la dispersione dei dati attorno alla media. Il 68% delle osservazioni si colloca nell'intervallo  $\mu \pm \sigma$ , compreso tra 0,9 km e 4,1 km. Il 95% dei dati ricade nell'intervallo  $\mu \pm 2\sigma$ , esteso da -0,7 km a 5,7 km (considerando il troncamento a zero per le distanze negative). Il 99,7% delle osservazioni teoricamente si distribuisce tra  $\mu \pm 3\sigma$ , da -2,3 km a 7,3 km.

La rappresentazione evidenzia come la distribuzione delle distanze presenti una variabilità moderata attorno al valore medio, con la maggior parte dei percorsi concentrati nella fascia compresa tra 1 km e 4 km. Le code della distribuzione si estendono fino a circa 7 km, indicando la presenza di alcuni percorsi significativamente più lunghi della media.



**Figura 3.4:** Curva di densità di una distribuzione normale teorica.

L'analisi esplorativa condotta sul dataset `bikesharing_bologna_2023_05_trips.csv` si distingue da quella effettuata sul dataset `bologna_2023_05.csv` per la diversa natura delle distanze percorse: in questo caso, esse sono state ottenute mediante l'applicazione dell'algoritmo di map-matching fornito da HybridPy.

Il *report* riassume i risultati ottenuti mediante l'analisi generale:

Misura	Valore	Misura	Valore
Media distanza	1.89 km	Range	0.15 - 13.11 km
Mediana distanza	1.68 km	IQR	1.33 km
Deviazione standard	1.08 km	Skewness	1.30
Varianza distanza	1.17 km <sup>2</sup>	Kurtosis	3.08

**Tabella 3.3:** Statistiche descrittive delle distanze dei tragitti.

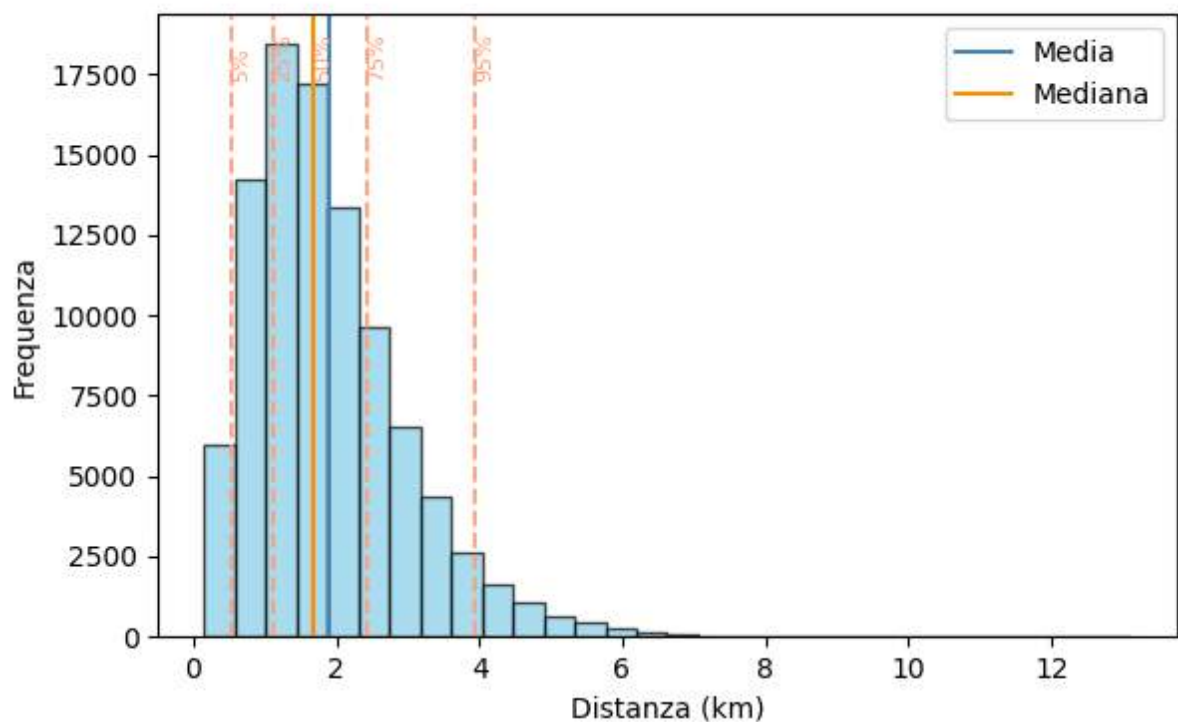
L'analisi statistica delle distanze percorse evidenzia un valore medio pari a 1.89 km e una mediana di 1.68 km, indicando che il 50% dei noleggi riguarda tragitti inferiori alla mediana. La dispersione dei dati rispetto al valore medio è rappresentata da una deviazione standard di 1.08 km, corrispondente a una varianza di 1.17 km<sup>2</sup>. Le distanze osservate variano da un minimo di 0.15 km a un massimo di 13.11 km, delineando un intervallo totale significativo. L'intervallo interquartile (IQR), che rappresenta la gamma entro cui ricade il 50% centrale dei dati, risulta pari a 1.33 km.

Dal punto di vista della forma della distribuzione, l'indice di asimmetria (*skewness*) assume un valore positivo pari a 1.30, suggerendo la presenza di una coda più estesa verso i valori più elevati. L'indice di curtosi (*kurtosis*) risulta pari a 3.08, indicando che la distribuzione presenta un picco e code di intensità comparabili a quelli di una distribuzione normale.

- *Distribuzione delle distanze*: come per il dataset originale, anche in questo caso la distribuzione è caratterizzata da una marcata asimmetria positiva (coda a destra), con la maggior parte delle osservazioni concentrate su tragitti brevi. Gli indicatori statistici mostrano valori di tendenza centrale più bassi rispetto al dataset precedente. La mediana si attesta a circa 1.8 km, mentre la media è leggermente superiore, con un valore di circa 2.0 km.

Percentile	Distanza (km)
5°	0.53
25°	1.11
50°	1.68
75°	2.44
95°	3.95

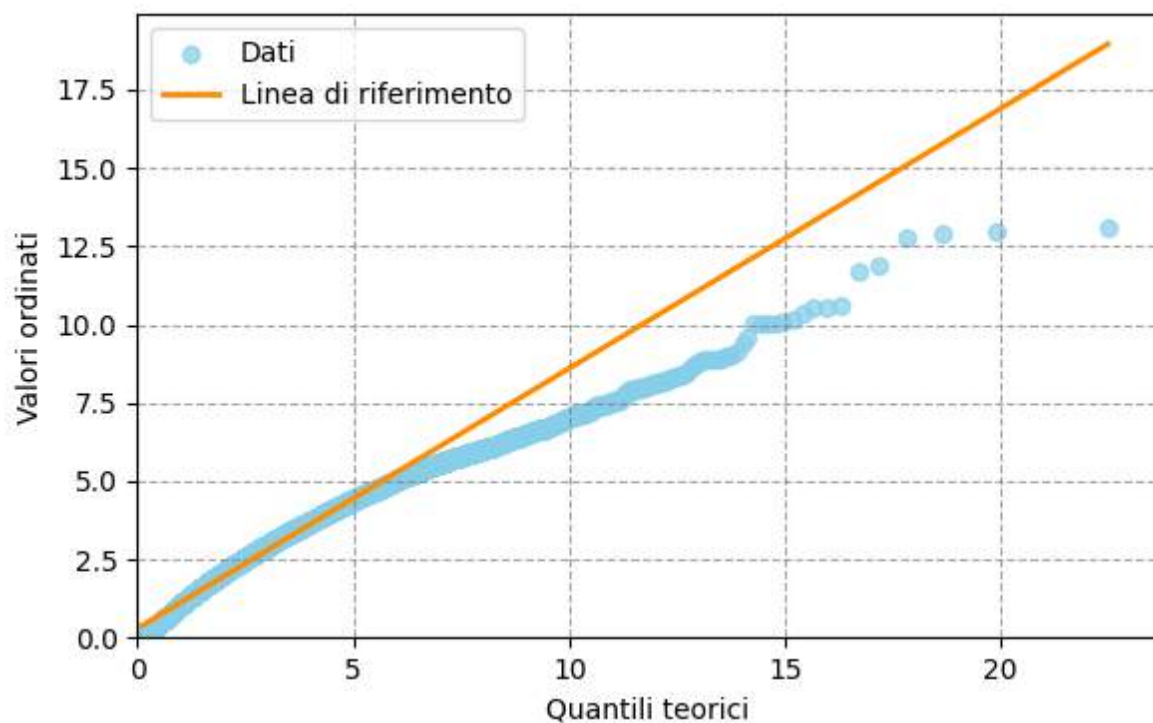
**Tabella 3.4:** Distribuzione delle distanze percorse ai diversi percentili.



**Figura 3.5:** Istogramma della distribuzione delle distanze *map-matched* per il dataset di Bologna.

- *QQ-plot* (Quantile-Quantile plot): la marcata e sistematica deviazione dei punti (in blu) dalla retta di riferimento (in arancione) permette di concludere che anche questi dati non sono distribuiti normalmente.

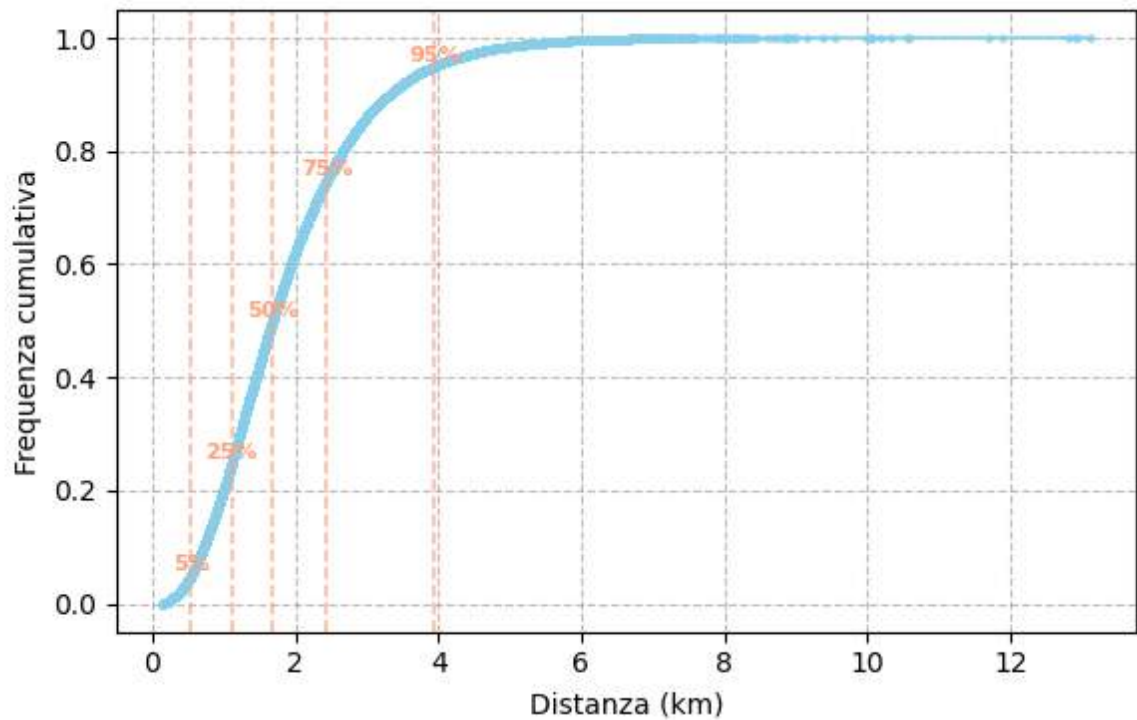
L'andamento dei punti disegna una curva concava, molto simile a quella osservata per il dataset originale. Questa forma indica che la coda destra della distribuzione è significativamente più “leggera” di quella di una distribuzione normale teorica. Anche in questo caso, la presenza di alcuni punti isolati e distanziati nella parte finale del grafico suggerisce l'esistenza di un piccolo numero di outlier.



**Figura 3.6:** Quantile-Quantile (QQ) plot della distribuzione delle distanze.

- *Distribuzione cumulativa (CDF)*: la curva permette di visualizzare la proporzione cumulata di osservazioni e di estrarre con precisione i percentili, confermando i valori discussi per l'istogramma.

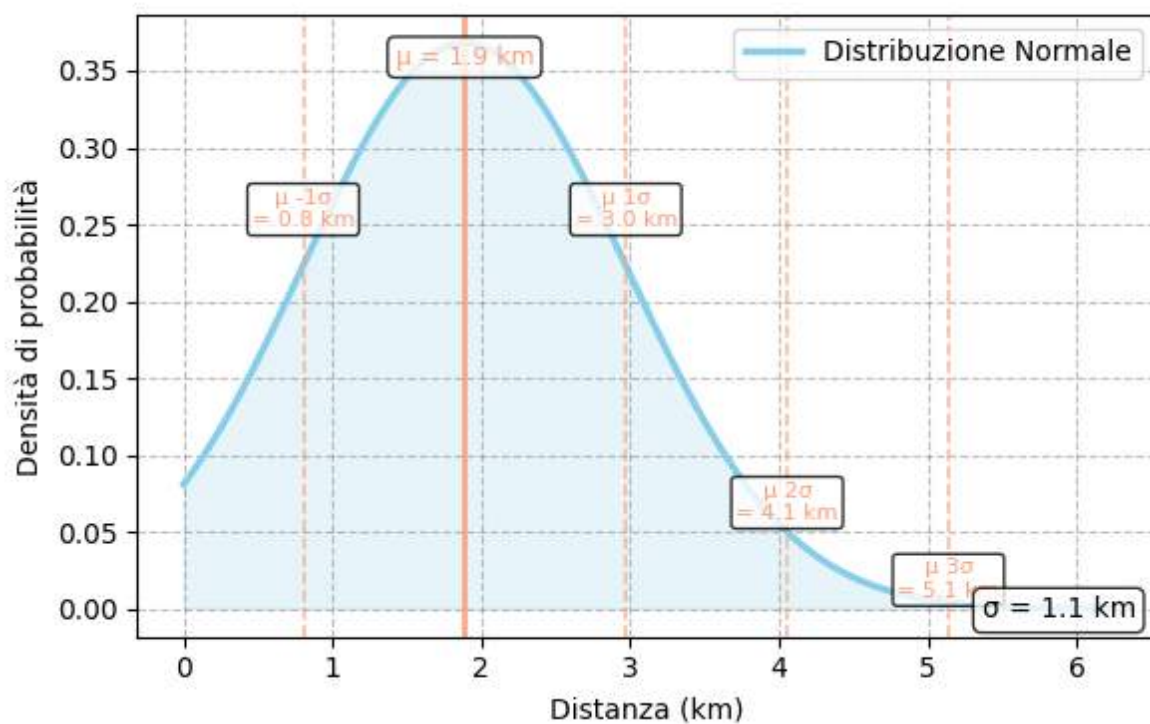
Dal grafico si evince che la mediana (50° percentile) corrisponde a una distanza di circa 1.8 km. L'intervallo che contiene il 50% centrale dei dati è compreso tra 1.0 km (primo quartile, 25%) e 2.8 km (terzo quartile, 75%). La rapida crescita della curva si esaurisce dopo i 4 km; infatti, il 95° percentile si attesta a 4.0 km, a dimostrazione che la quasi totalità delle osservazioni si concentra al di sotto di tale soglia.



**Figura 3.7:** Funzione di ripartizione (CDF) delle distanze.

- *Distribuzione normale teorica*: figura mostra il modello di una distribuzione normale teorica generata utilizzando la media ( $\mu \approx 2.0$  km) e la deviazione standard ( $\sigma \approx 1.0$  km) del dataset corretto con map-matching. Questa curva, perfettamente simmetrica attorno alla sua media, serve come termine di paragone ideale.

Il confronto visivo tra questo modello teorico e la distribuzione reale dei dati è netto. La distribuzione effettiva dei dati manifesta una chiara asimmetria positiva e una maggiore concentrazione di valori (leptocurtosi) rispetto alla forma più appiattita della curva normale. Questa differenza sostanziale conferma le conclusioni tratte dal QQ-plot e permette di rigettare l'ipotesi di normalità anche per il dataset elaborato con map-matching.

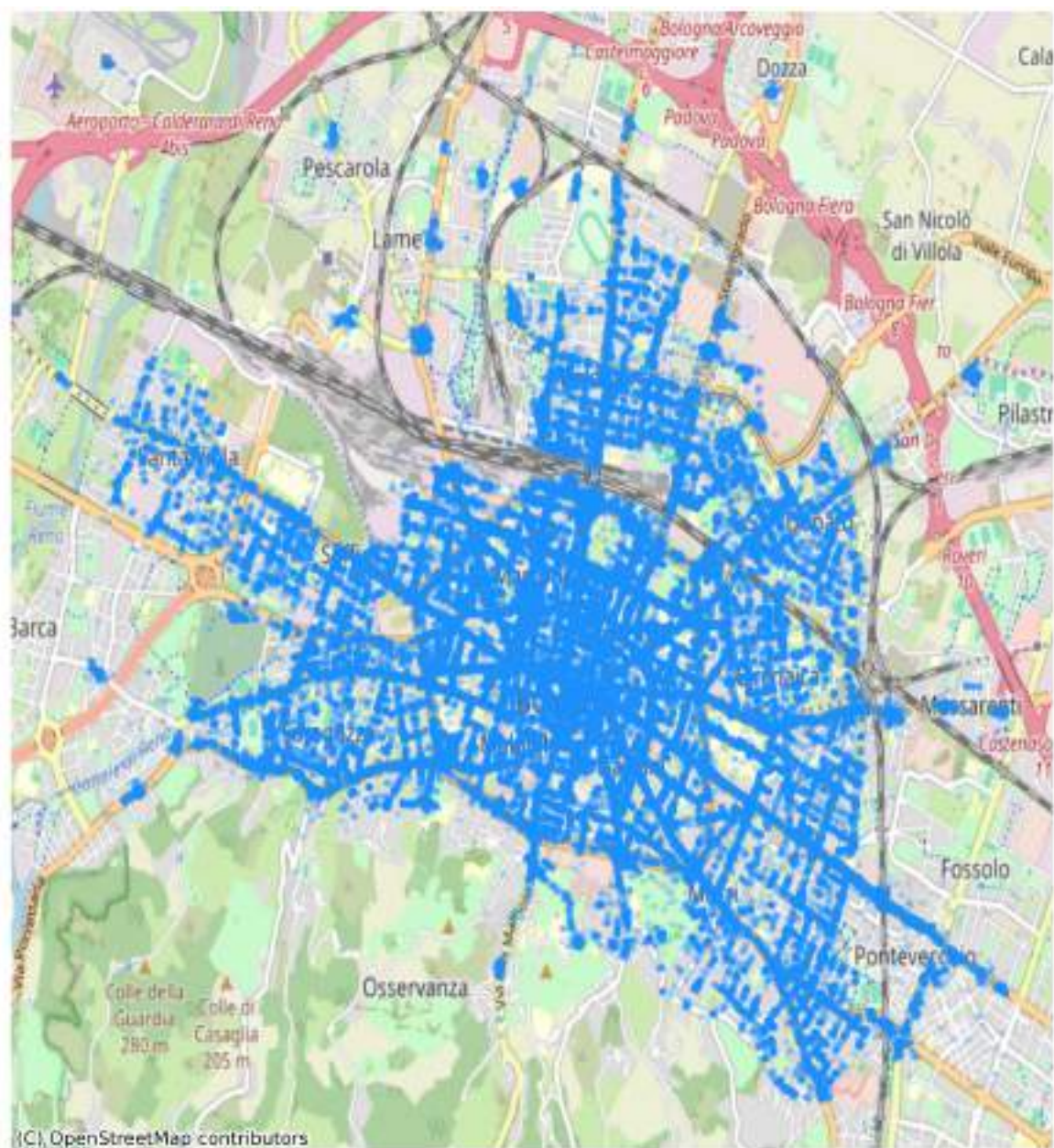


**Figura 3.8:** Curva di densità di una distribuzione normale teorica.

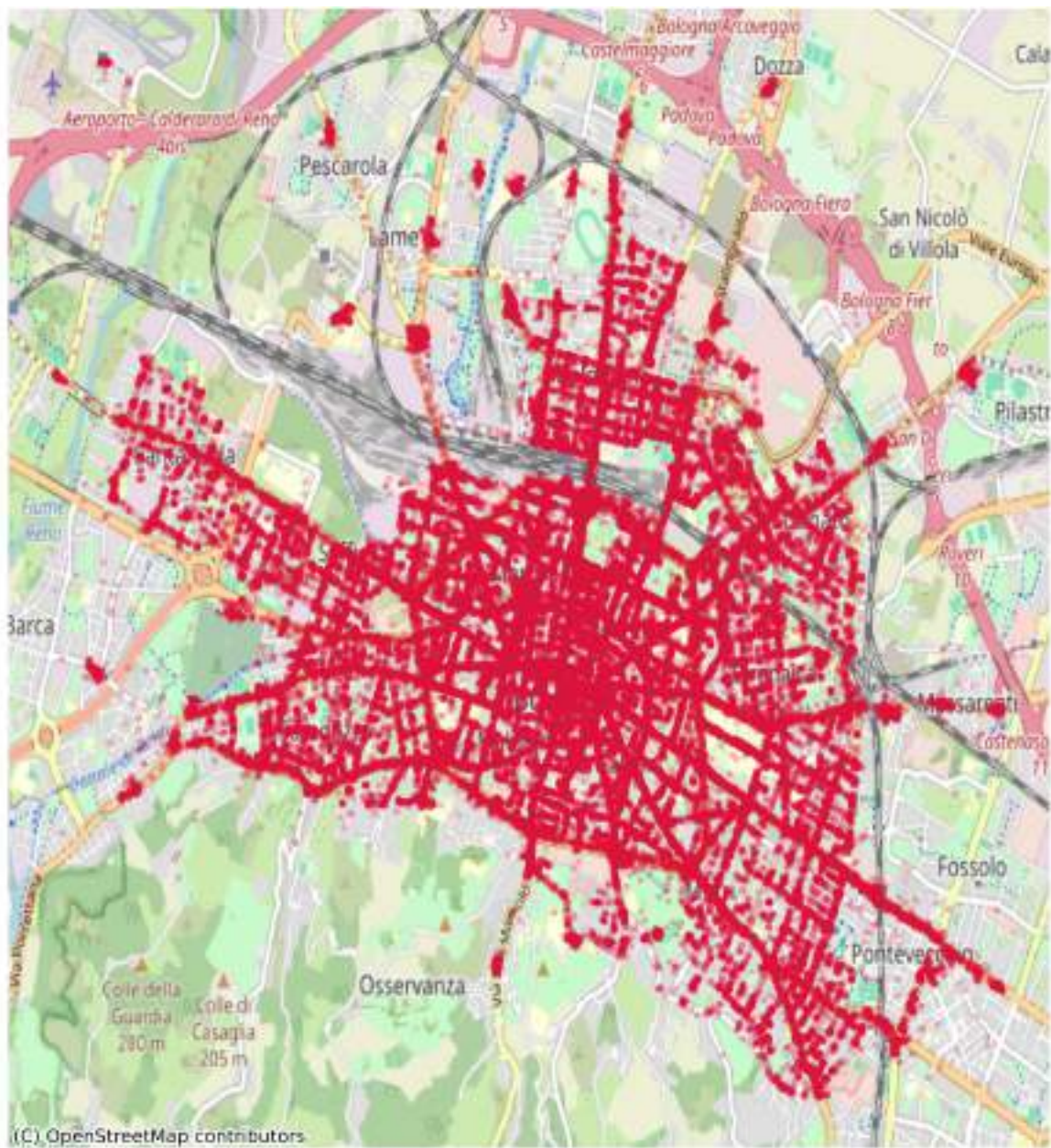


### 3.2.2 Analisi spaziale

L'analisi spaziale condotta sul dataset `bologna_2023_05.csv` ha prodotto i seguenti grafici:

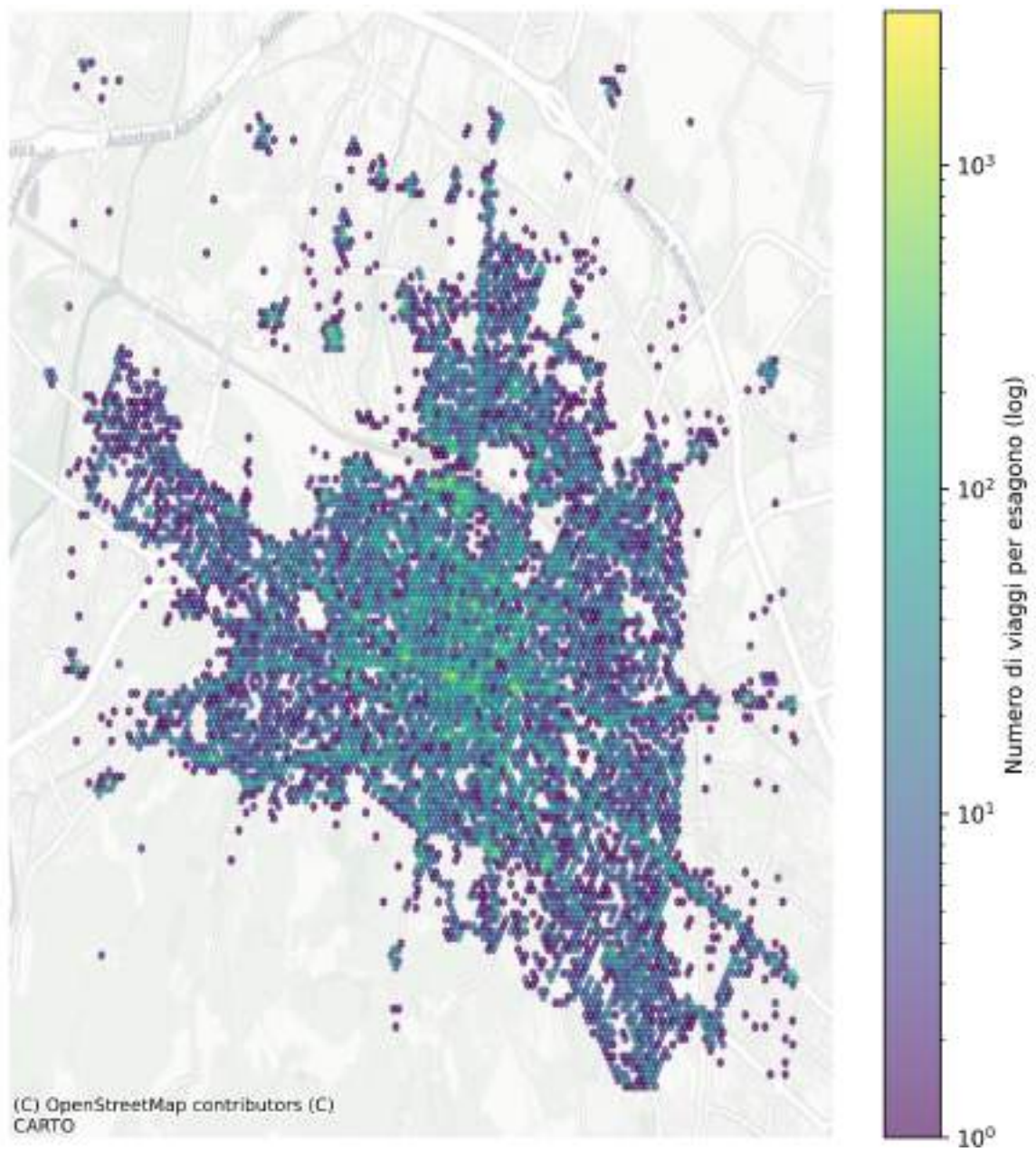


**Figura 3.9:** Distribuzione spaziale dei punti di partenza sul territorio urbano di Bologna.

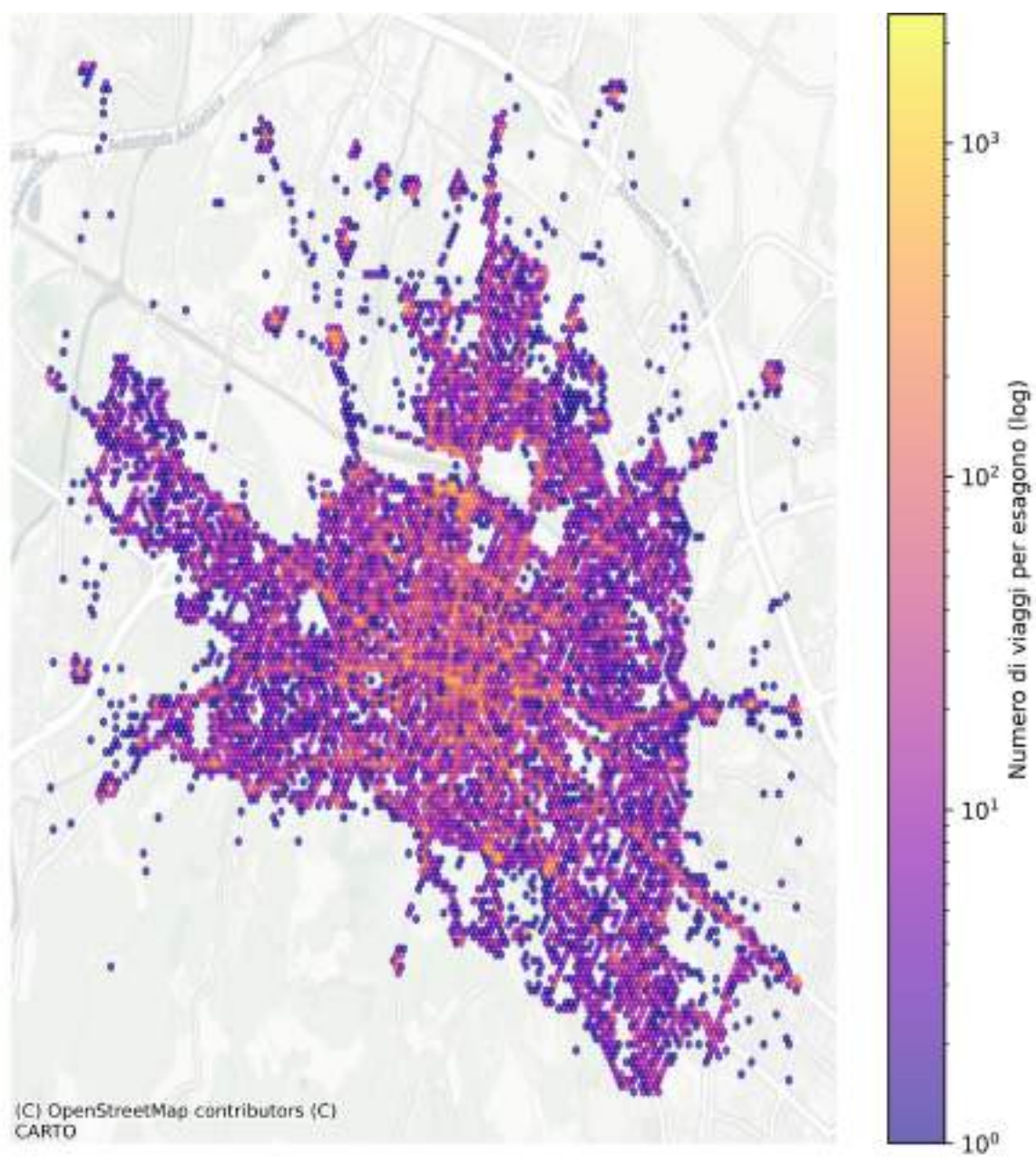


**Figura 3.10:** Distribuzione spaziale dei punti di arrivo sul territorio urbano di Bologna.





**Figura 3.11:** Mappe di densità dei punti di partenza.



**Figura 3.12:** Mappe di densità dei punti di arrivo.

### 3.2.3 Analisi tempi/velocità

L'analisi ha messo a confronto 11'360 viaggi effettuati con biciclette muscolari e 133'605 con e-bike, applicando a entrambi i gruppi i medesimi filtri per garantire l'omogeneità del confronto: sono stati considerati solo i viaggi con durata inferiore a due ore e velocità media compresa tra 5 e 30 km/h.

	Muscolari	E-bike
<b>Viaggi analizzati</b>	11'360	133'605
<b>Tempi (ore)</b>		
Media $\pm$ Dev. Std	0.196 $\pm$ 0.134	0.198 $\pm$ 0.138
Mediana	0.167	0.169
Min-Max	0.002 – 1.772	0.002 – 1.984
<b>Velocità (km/h)</b>		
Media $\pm$ Dev. Std	9.28 $\pm$ 2.53	13.21 $\pm$ 3.57
Mediana	8.99	13.24
Min-Max	5.00 – 29.38	5.00 – 29.97

**Tabella 3.5:** Report dell'analisi tempi e velocità per biciclette muscolari ed e-bike.

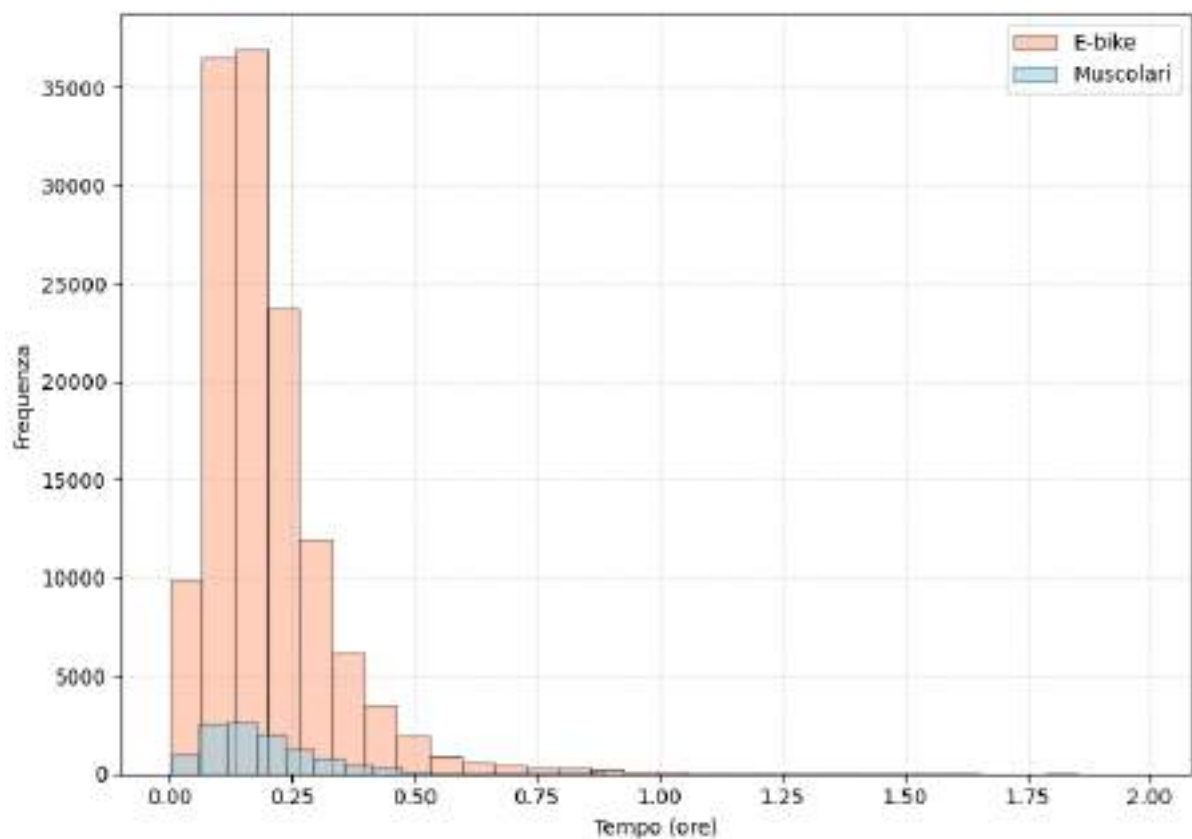
Per quanto riguarda i tempi di percorrenza, i due gruppi mostrano risultati notevolmente simili. Il tempo medio di viaggio si attesta a 11.76 minuti per le bici muscolari e a 11.88 minuti per le e-bike. Questa forte somiglianza è confermata anche dai valori mediani, quasi identici, pari a 10 minuti per le muscolari e 10.14 minuti per le e-bike.

L'analisi delle velocità mostra valori medi pari a 13,21 km/h per i viaggi effettuati con e-bike e 9,28 km/h per quelli con bici muscolari. I valori mediani si attestano rispettivamente a 13,24 km/h e 8,99 km/h. La dispersione dei dati, misurata tramite deviazione standard, è pari a  $\pm 3,57$  km/h per le e-bike e  $\pm 2,53$  km/h per le bici muscolari.

L'analisi tempi/velocità condotta sul dataset `bologna_2023_05.csv` ha prodotto i seguenti grafici:

- *Distribuzione tempi di percorrenza* (frequenza/tempo): l'istogramma confronta la distribuzione dei tempi di percorrenza tra le biciclette elettriche e quelle muscolari: il risultato più evidente è la netta prevalenza numerica dei viaggi effettuati con e-bike.

Entrambe le distribuzioni presentano un'asimmetria positiva (coda a destra), indicando che la maggior parte dei viaggi è di breve durata. Il picco di frequenza (la *moda*) per entrambe le tipologie si concentra per tempi inferiori a 0.25 h, con la stragrande maggioranza dei viaggi che si conclude entro 0.5 h. Sebbene il volume di utilizzo sia molto diverso, la forma simile delle due distribuzioni suggerisce che le due tipologie di veicoli vengono impiegate per coprire tragitti di durata comparabile.

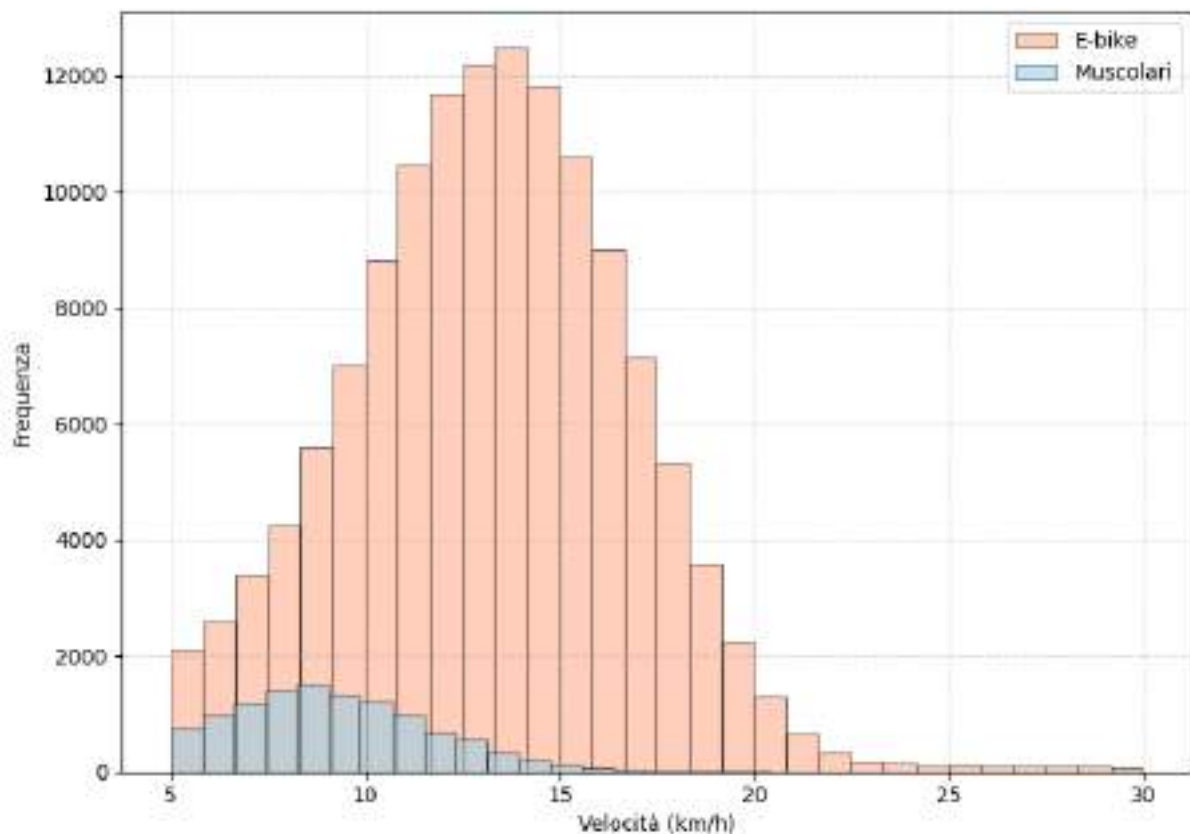


**Figura 3.13:** Istogramma della distribuzione dei tempi di percorrenza, suddivisi per tipologia di bicicletta: *e-bike* (arancione) e *muscolari* (azzurro).

- *Distribuzione velocità media* (frequenza/velocità): l'istogramma confronta le distribuzioni delle velocità medie registrate per le e-bike e per le biciclette muscolari. Il grafico evidenzia una netta differenza nelle prestazioni tra le due categorie.

La distribuzione delle velocità per le e-bike è approssimativamente simmetrica e centrata attorno a un valore modale (il picco di frequenza) di circa 13-14 km/h. Al contrario, la distribuzione per le biciclette muscolari mostra un picco a velocità inferiori, intorno ai 9-10 km/h, e una maggiore dispersione relativa.

L'intera distribuzione delle velocità delle e-bike è visibilmente traslata verso destra rispetto a quella delle biciclette muscolari, indicando che i veicoli a pedalata assistita mantengono sistematicamente una velocità media superiore.

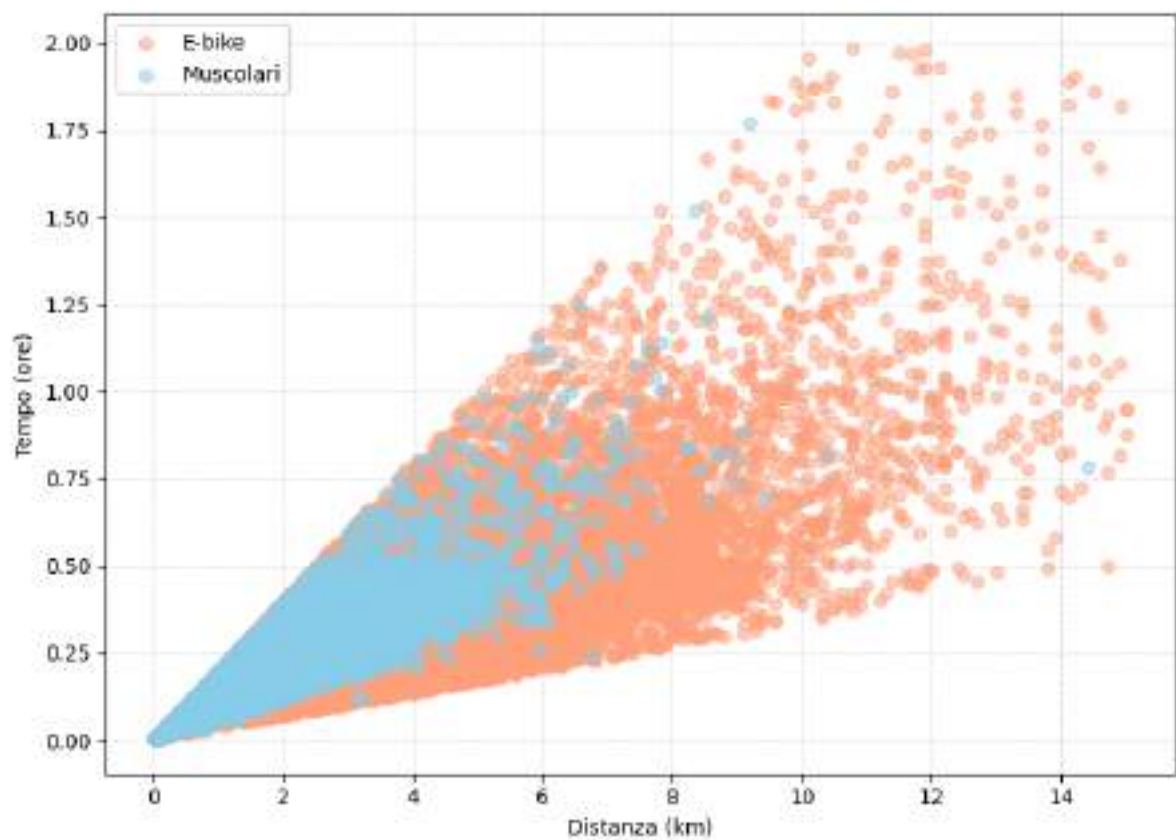


**Figura 3.14:** Istogramma della distribuzione delle velocità medie, suddivise per tipologia di bicicletta: *e-bike* (arancione) e *muscolari* (azzurro).



- *Relazione distanza-tempo*: la figura visualizza la relazione tra distanza e tempo di percorrenza per ogni singolo viaggio effettuato. Come previsto, esiste una forte correlazione positiva tra le due variabili per entrambe le tipologie di veicolo.

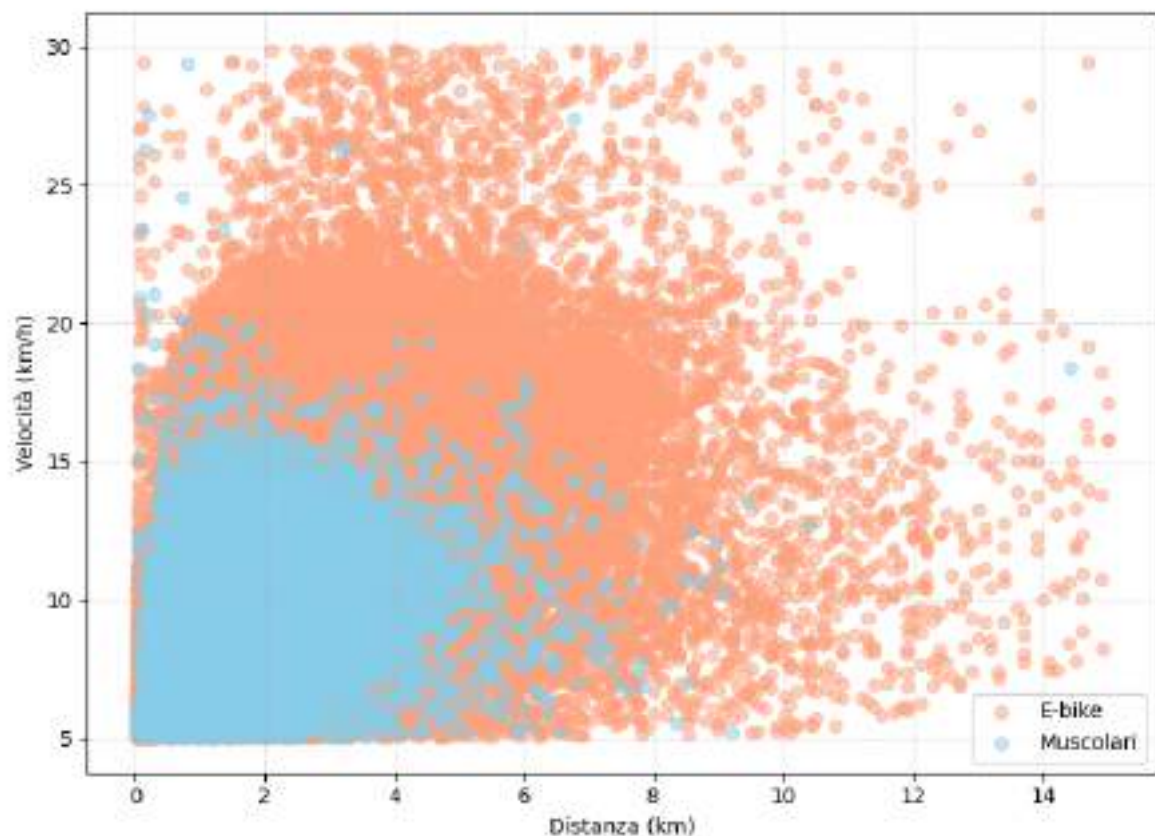
Il grafico mostra che le e-bike vengono impiegate su un intervallo di distanze più ampio, comprendente anche tragitti superiori a 10 km, meno frequenti per le biciclette muscolari. La nube di punti arancione (e-bike) presenta una pendenza mediamente inferiore rispetto a quella azzurra (muscolari), in corrispondenza di una maggiore velocità media a parità di distanza percorsa.



**Figura 3.15:** Grafico a dispersione (*scatter plot*) della relazione tra distanza percorsa (km) e tempo impiegato (ore). Ogni punto rappresenta un singolo viaggio, colorato in base alla tipologia di bicicletta: *e-bike* (arancione) e *muscolare* (azzurro).

- *Relazione distanza-velocità*: la figura analizza la relazione tra la distanza di un viaggio e la velocità media mantenuta. La relazione non è lineare; per entrambe le tipologie di veicolo, i viaggi molto brevi mostrano un'ampia dispersione delle velocità.

Tuttavia, il confronto tra le due categorie è netto. Le *e-bike* (arancione) dimostrano la capacità di raggiungere e mantenere velocità medie elevate (tra 10 e 20 km/h) in modo consistente su un vasto range di distanze, specialmente quelle superiori ai 2 km. Al contrario, le biciclette *muscolari* (azzurro), oltre ad essere impiegate su distanze più brevi, mostrano una nuvola di punti concentrata a velocità inferiori. Questo suggerisce che il vantaggio delle e-bike in termini di velocità si manifesta soprattutto sulle distanze medie e lunghe, dove la capacità di mantenere una velocità di crociera elevata diventa più rilevante.



**Figura 3.16:** Grafico a dispersione (*scatter plot*) che illustra la relazione tra la distanza percorsa (km) e la velocità media (km/h). Ogni punto rappresenta un singolo viaggio, colorato in base alla tipologia di bicicletta: *e-bike* (arancione) *muscolare* (azzurro).

### 3.2.4 Analisi applicativa

L'analisi del dataset di Bologna è stata condotta su un totale di 182'418 tragitti, che includono sia viaggi con biciclette muscolari sia con e-bike. La ripartizione per tipologia di veicolo mostra una netta prevalenza delle e-bike, utilizzate in 165'560 casi (90.8% del totale), mentre i restanti 16'858 viaggi (9.2%) sono stati effettuati con biciclette muscolari. La distanza media dei percorsi risulta pari a 2.07 km. Analizzando la distribuzione delle lunghezze, emerge che la maggior parte dei tragitti rientra nella categoria dei percorsi brevi: 98'160 spostamenti (53.8%) hanno coperto una distanza inferiore a 2 km. I tragitti medi, compresi tra 2 e 5 km, sono 74'300 (40.7%), mentre i tragitti lunghi, con percorrenza superiore a 5 km, ammontano a 9'958 (5.5%).

Il dataset `bikesharing_bologna_2023_05_trips.csv` comprende 96'919 tragitti complessivi. Anche in questo caso si osserva una prevalenza delle e-bike, con 87'179 viaggi (90.0% del totale), mentre le biciclette tradizionali (*bicycle*) sono state utilizzate in 9'740 casi (10.0%). La distanza media risulta pari a 1.88 km. Per quanto riguarda la distribuzione delle percorrenze, i tragitti brevi (inferiori a 2 km) costituiscono la maggioranza con 60'274 spostamenti (62.2%). I tragitti medi (tra 2 e 5 km) sono 35'175 (36.3%), mentre i tragitti lunghi (oltre 5 km) rappresentano solo 1'470 viaggi (1.5%).

	Dataset 1	Dataset 2
<b>Distanza media del tragitto</b>	2.07 km	1.88 km
<b>Distribuzione tragitti per classe</b>		
Brevi (<2 km)	98'160	60'274
Medi (2-5 km)	74'300	35'175
Lunghi (>5 km)	9'958	1'470
<b>Distribuzione per tipologia di bici</b>		
E-bike / ebicycle	165'560	87'179
Muscolari / bicycle	16'858	9'740

**Tabella 3.6:** Confronto dei report dell'analisi applicativa per i due dataset.

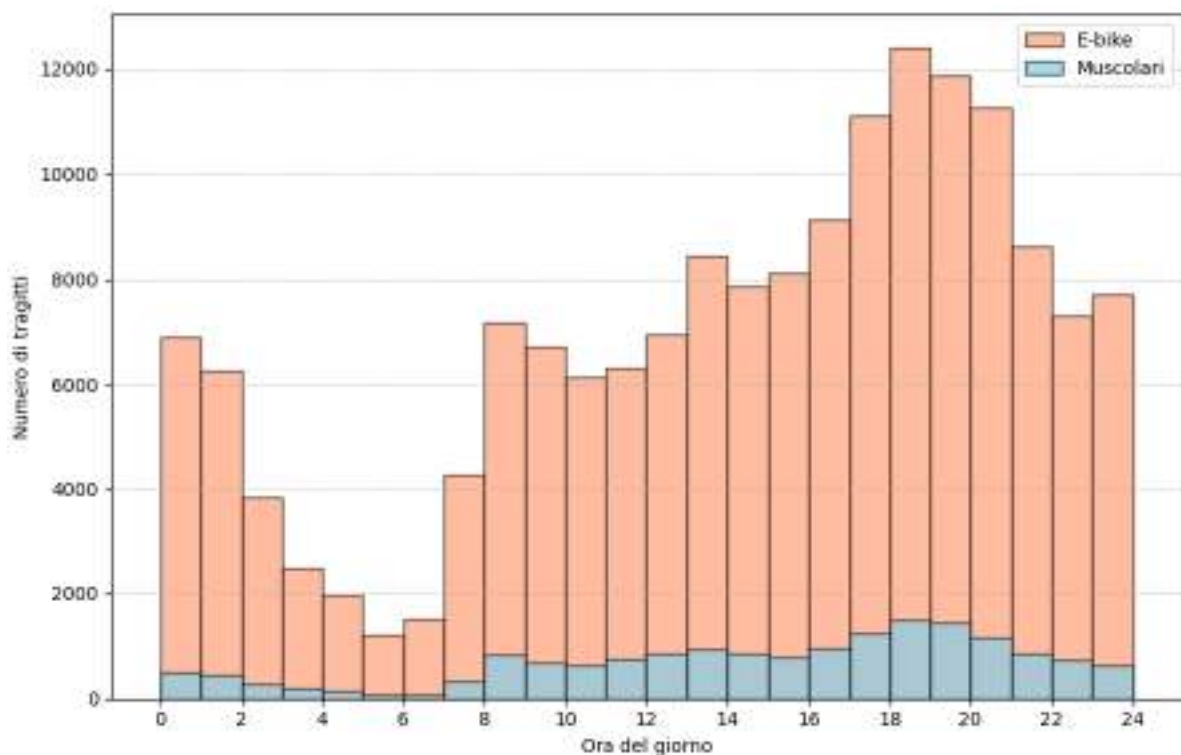


### 3.2.5 Analisi temporale

L'analisi temporale condotta sul dataset `bologna_2023_05.csv` ha prodotto i seguenti grafici:

- *Distribuzione oraria dei tragitti*: la figura illustra come l'utilizzo del servizio di bike sharing si distribuisce nell'arco della giornata. Il grafico mostra una chiara distribuzione *bimodale* con due picchi di utilizzo principali.

Il primo *picco mattutino* si verifica tra le 8:00 e le 9:00, in coincidenza con l'inizio delle attività lavorative e di studio. Dopo un utilizzo sostenuto durante la giornata, si osserva il *picco principale* nel tardo pomeriggio/sera, con la massima concentrazione di viaggi registrata tra le 18:00 e le 19:00. L'attività decresce significativamente dopo le 22:00, raggiungendo un minimo nelle ore notturne tra le 2:00 e le 6:00.

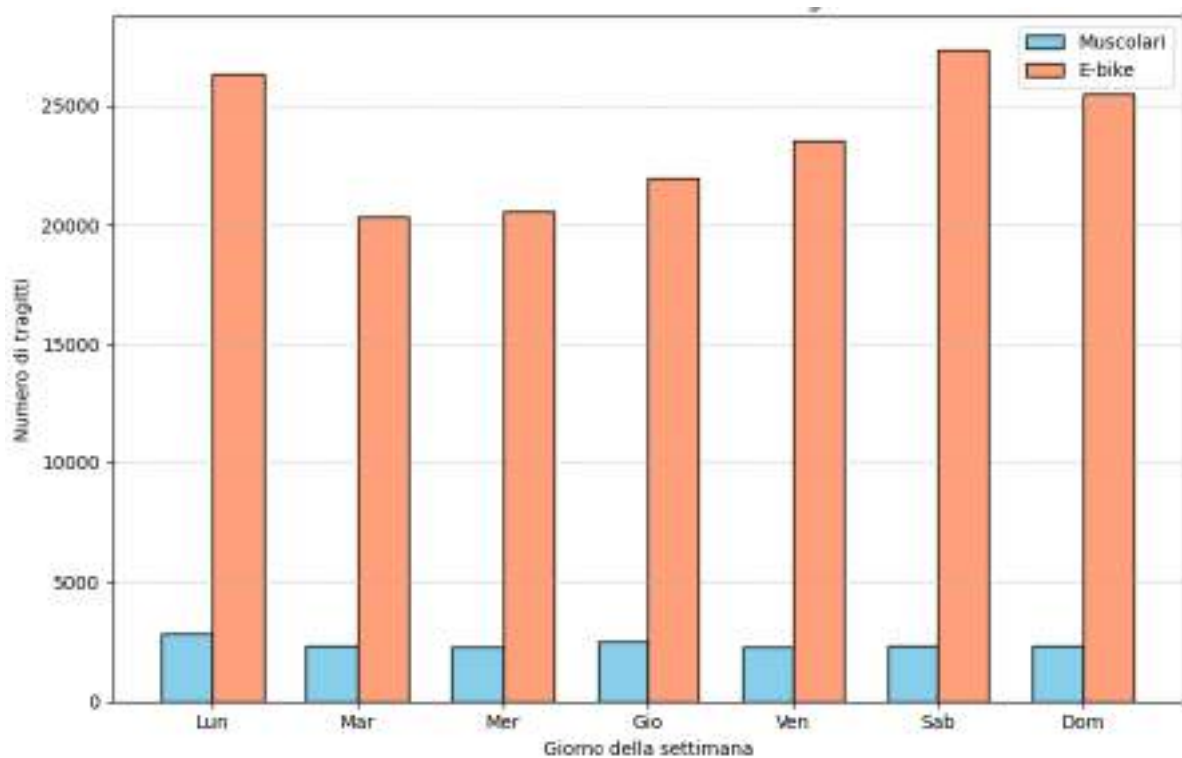


**Figura 3.17:** Distribuzione oraria del numero di viaggi, suddivisi per tipologia di bicicletta: *e-bike* (arancione) e *muscolare* (azzurro).

- *Distribuzione settimanale dei tragitti*: la figura illustra la distribuzione dei viaggi nell'arco della settimana. L'andamento non segue un tipico pattern pendolare feriale, ma mostra un modello di utilizzo più complesso.

I giorni di maggiore attività sono il sabato, che rappresenta il picco assoluto della settimana, seguito dal lunedì e dalla domenica, entrambi con un numero di viaggi molto elevato. Durante i giorni lavorativi si osserva un'intensità alta il lunedì, un leggero calo nei giorni centrali della settimana (martedì, mercoledì, giovedì) e una ripresa il venerdì.

L'elevato utilizzo durante il fine settimana suggerisce un uso significativo del servizio per scopi non legati al pendolarismo casa-lavoro, come attività ricreative e turistiche.



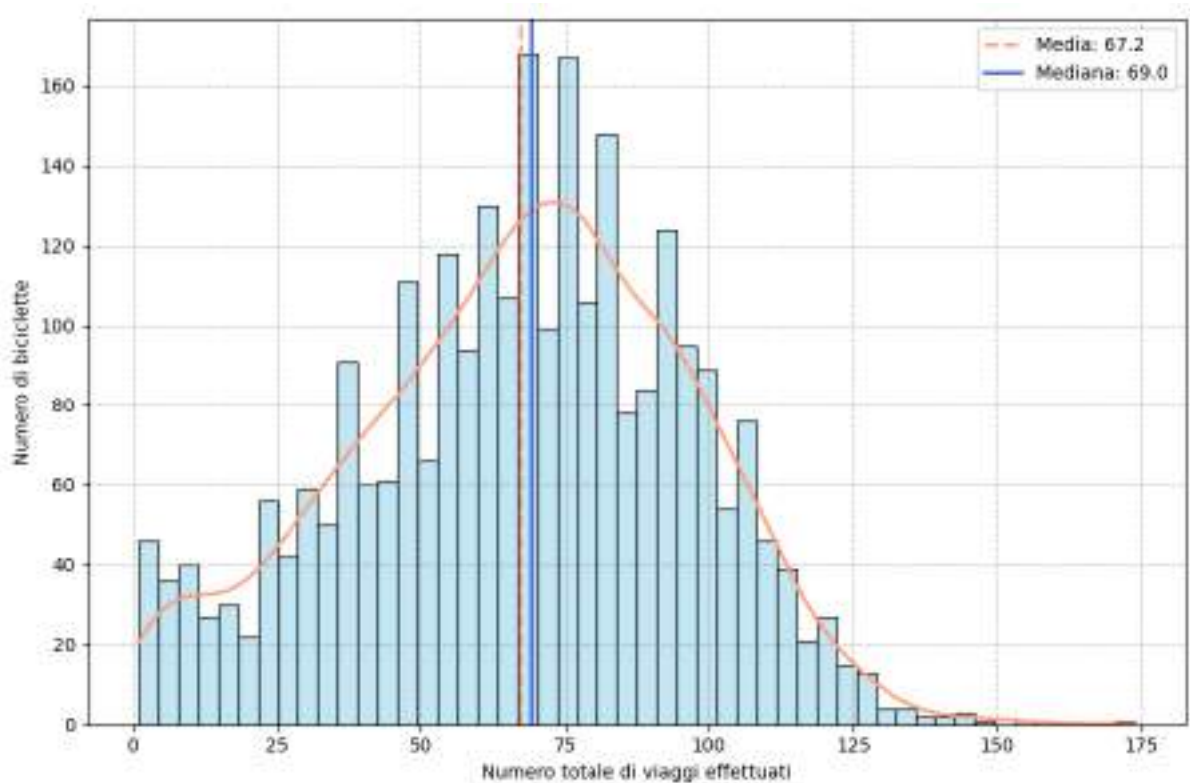
**Figura 3.18:** Distribuzione del numero di viaggi per giorno della settimana, suddivisi per tipologia di bicicletta: *e-bike* (arancione) e *muscolare* (azzurro).

### 3.2.6 Analisi dei turnover

L'analisi dei turnover condotta sul dataset `bologna_2023_05.csv` ha prodotto i seguenti grafici:

- Il primo grafico mostra come si distribuisce il carico di lavoro sull'intera flotta di biciclette. La distribuzione del numero di viaggi per veicolo è unimodale e approssimativamente simmetrica, come suggerito dalla stretta vicinanza tra i valori di media (67.2) e mediana (69.0).

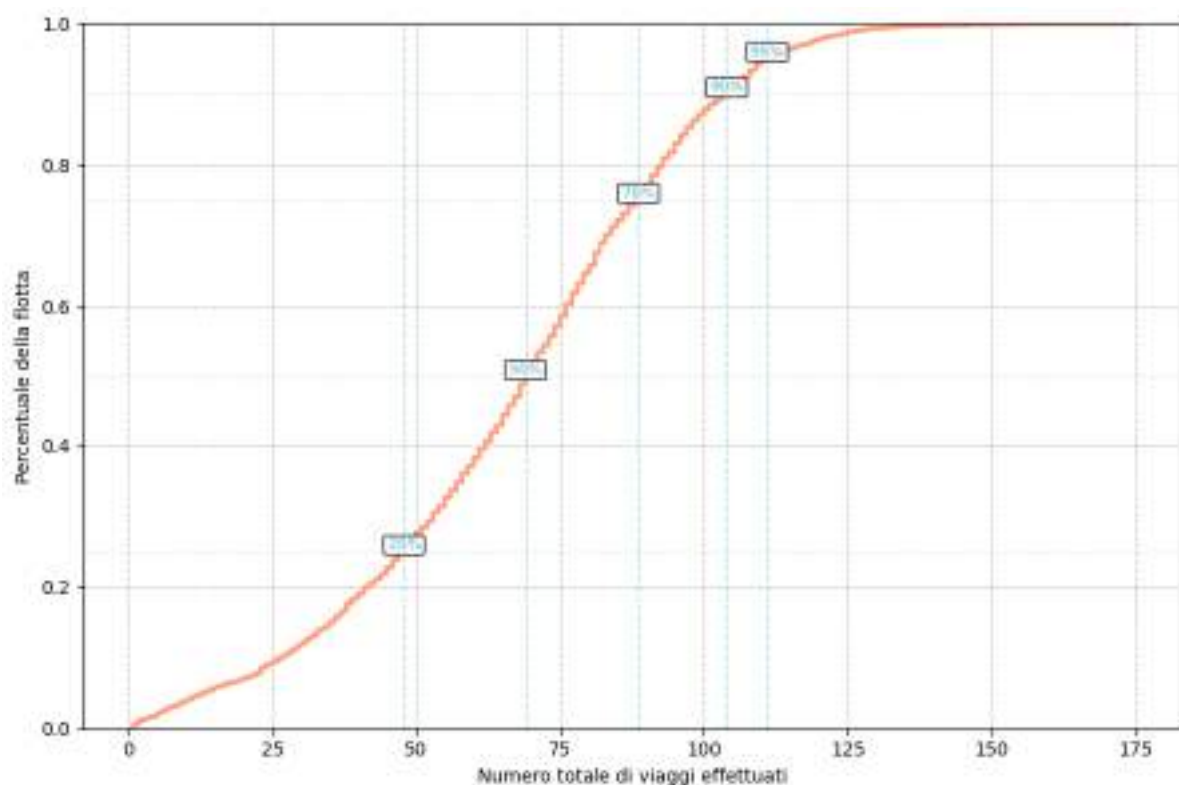
La distribuzione, con una leggera asimmetria a sinistra, mostra come la maggior parte dei veicoli abbia effettuato tra 25 e 125 viaggi, con un picco di frequenza (moda) intorno a 70-75. Questo risultato indica un utilizzo complessivamente bilanciato della flotta, in cui la maggior parte delle biciclette contribuisce in modo simile all'operatività generale, senza un numero eccessivo di mezzi inutilizzati o sovraccaricati.



**Figura 3.19:** Istogramma della distribuzione del numero totale di viaggi per bicicletta.

- Il secondo grafico illustra la funzione di ripartizione empirica, che completa l'analisi sulla distribuzione del carico di lavoro all'interno della flotta. La forma a "S" della curva conferma la forte concentrazione dei dati attorno a un valore centrale, già osservata nell'istogramma.

Dal grafico si possono leggere con precisione i percentili di utilizzo: si nota che il 50% della flotta (la *mediana*) ha effettuato meno di 69 viaggi, confermando il valore precedente. Inoltre, si evince che il 50% centrale dei veicoli (l'intervallo interquartile) ha compiuto un numero di viaggi compreso approssimativamente tra 50 (25° percentile) e 90 (75° percentile). Questo dato rafforza la conclusione di un utilizzo relativamente bilanciato della flotta, dove la maggior parte dei veicoli contribuisce in modo comparabile all'operatività del servizio.



**Figura 3.20:** Funzione di distribuzione cumulativa (CDF) del numero totale di viaggi per bicicletta.

## 3.3 La città di Firenze

### 3.3.1 Analisi generale

Il *report* riassume i risultati ottenuti mediante l'analisi generale:

Misura	Valore	Misura	Valore
Media distanza	1.93 km	Range	0.15 - 15.00 km
Mediana distanza	1.57 km	IQR	1.61 km
Deviazione standard	1.45 km	Skewness	2.01
Varianza distanza	2.09 km <sup>2</sup>	Kurtosis	7.08

**Tabella 3.7:** Statistiche descrittive delle distanze dei tragitti.

L'analisi statistica delle distanze percorse evidenzia un valore medio pari a 1.93 km, con una mediana leggermente inferiore (1.57 km), indicando una distribuzione lievemente asimmetrica verso destra. La deviazione standard di 1.45 km e la varianza di 2.09 km<sup>2</sup> mostrano una dispersione moderata rispetto al valore medio. L'intervallo di variabilità è piuttosto ampio, con valori compresi tra 0.15 km e 15.00 km, mentre l'IQR di 1.61 km conferma che la metà centrale dei dati è concentrata in un range ristretto.

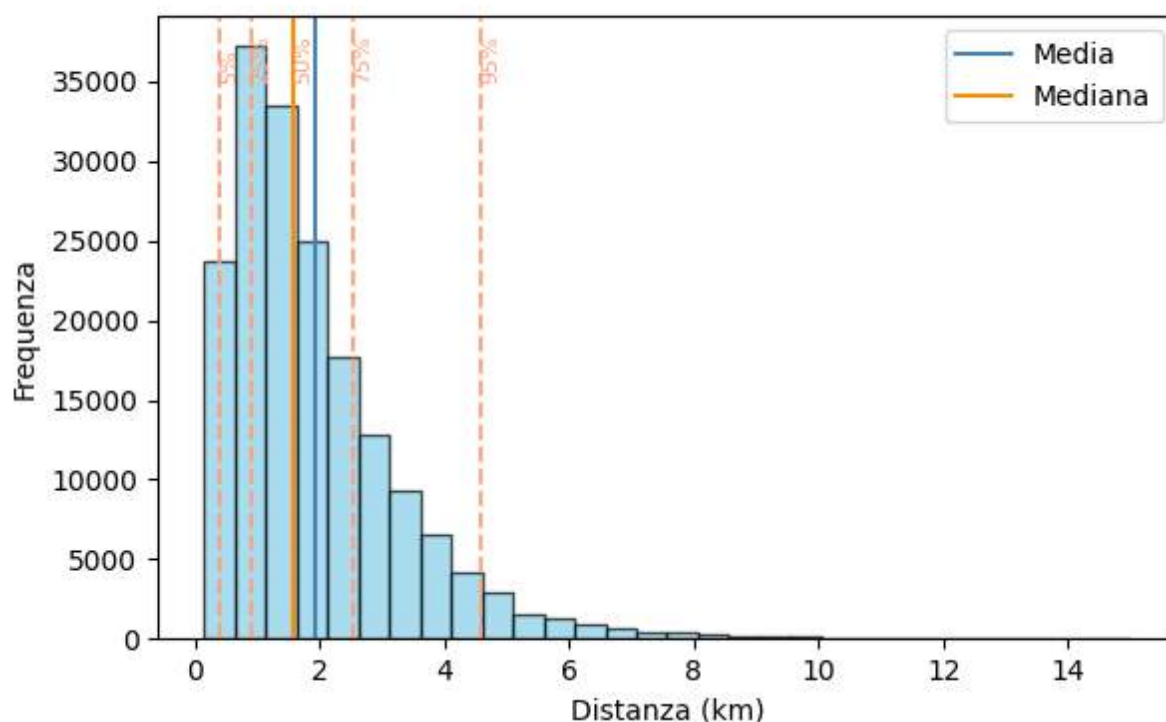
Gli indici di forma confermano questa interpretazione: la skewness positiva (2.01) descrive una distribuzione con una coda più lunga verso i valori elevati, mentre la kurtosis (7.08) indica una distribuzione più appuntita e con code più pesanti rispetto a una normale, segnalando la presenza di outlier o valori particolarmente estremi.

L'analisi esplorativa condotta sul dataset `firenze_2023_05.csv` ha prodotto i seguenti grafici:

- *Distribuzione* delle distanze: la distribuzione è marcatamente asimmetrica a destra, con una netta prevalenza di viaggi su breve distanza e una coda di tragitti più lunghi ma meno frequenti. La *mediana* si attesta a circa 1.6 km, indicando che metà dei viaggi copre una distanza inferiore a tale valore. La *media*, pari a circa 2.0 km, risulta superiore alla mediana a causa dell'influenza dei viaggi più lunghi.

Percentile	Distanza (km)
5°	0.40
25°	0.92
50°	1.57
75°	2.53
95°	4.60

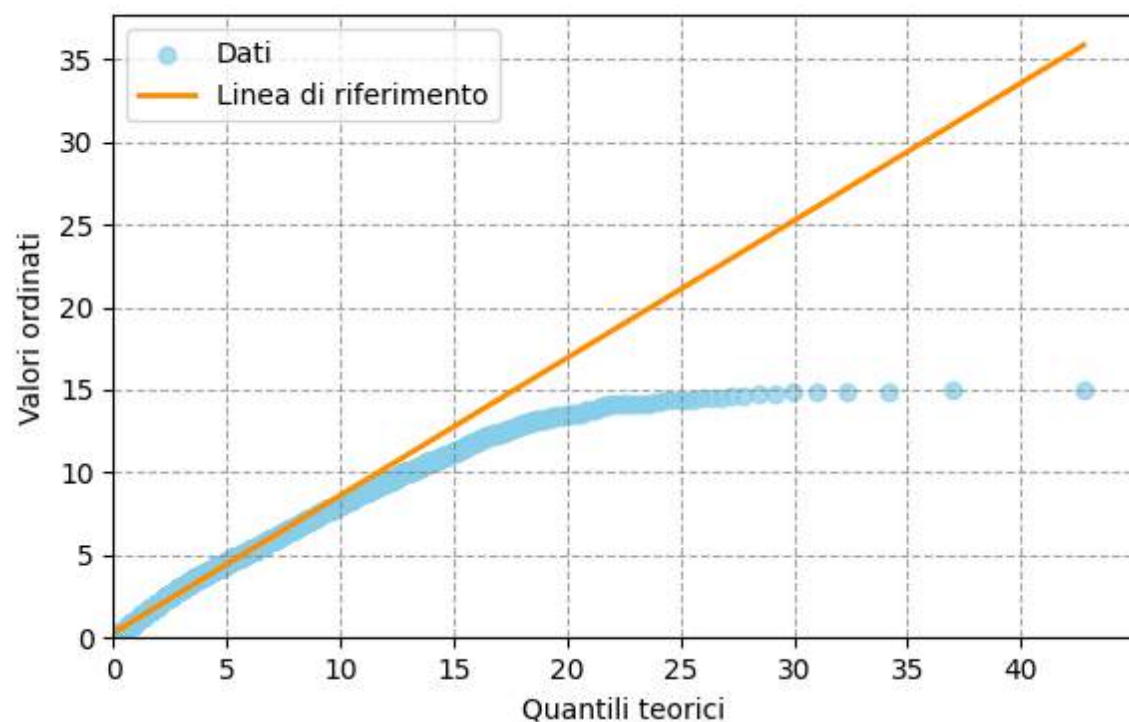
**Tabella 3.8:** Distribuzione delle distanze percorse ai diversi percentili.



**Figura 3.21:** Istogramma della distribuzione delle distanze per il dataset di Firenze.

- *QQ-plot* (Quantile-Quantile plot): il grafico confronta i quantili dei dati osservati con quelli attesi da una distribuzione normale. La sistematica e marcata deviazione dei punti dalla retta di riferimento rossa conferma in modo inequivocabile che la distribuzione delle distanze non è *normale*.

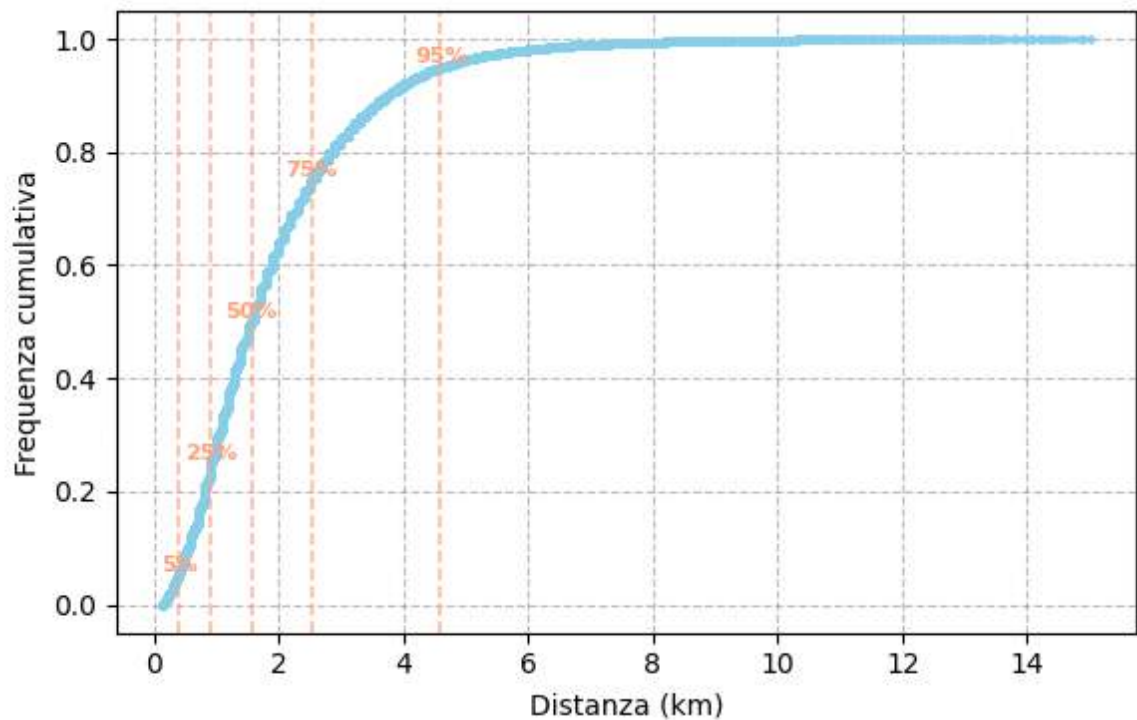
In particolare, i punti disegnano una curva *concava*: mentre i valori più bassi si allineano con una certa precisione alla retta teorica, al crescere dei quantili la deviazione si fa progressivamente più evidente. Questa tendenza è particolarmente accentuata nella coda destra, dove i valori campionari si “appiattiscono” e rimangono ben al di sotto della linea di riferimento. Ciò indica che la coda della distribuzione reale contiene valori meno estremi rispetto a quella di una distribuzione gaussiana. Infine, la presenza di alcuni punti isolati all’estrema destra suggerisce l’esistenza di un piccolo numero di outlier.



**Figura 3.22:** Quantile-Quantile (QQ) plot della distribuzione delle distanze.

- *Distribuzione cumulativa (CDF)*: la funzione di ripartizione empirica (ECDF) illustra che la pendenza della curva è massima per valori di distanza bassi, confermando la forte concentrazione dei dati in questo intervallo.

Dal grafico è possibile leggere direttamente i percentili della distribuzione. Si osserva che il 50° percentile (mediana) corrisponde a una distanza di circa 1.6 km, indicando che metà dei viaggi è più breve di tale valore. Similmente, il 25% dei dati presenta una distanza inferiore a 0.9 km (primo quartile) e il 75% una distanza inferiore a 2.6 km (terzo quartile). La curva si appiattisce notevolmente dopo i 5 km, avvicinandosi asintoticamente al valore 1.0; questo è evidenziato dal 95° percentile, che si attesta a circa 4.8 km, a dimostrazione che la quasi totalità degli spostamenti si concentra al di sotto di questa soglia.

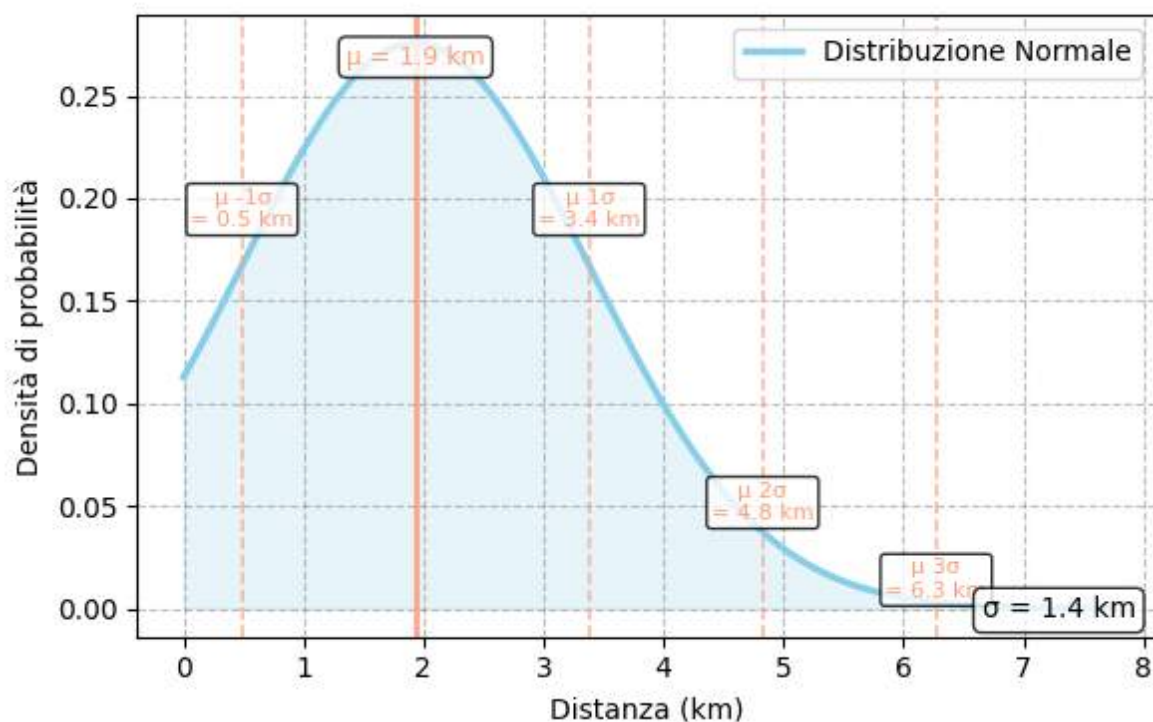


**Figura 3.23:** Funzione di ripartizione (CDF) delle distanze.



- *Distribuzione normale teorica*: la figura mostra la curva di densità di probabilità di una distribuzione perfettamente normale. Questa curva è stata generata utilizzando come parametri la media campionaria ( $\mu \approx 2.0$ ) km e la deviazione standard campionaria ( $\sigma \approx 1.4$ ) km calcolate dai dati di Firenze.

Nel grafico, la distribuzione è perfettamente simmetrica attorno alla sua media (linea blu continua), che rappresenta anche il punto di massima densità di probabilità. Le linee tratteggiate indicano gli intervalli di una, due e tre deviazioni standard dalla media. Il confronto diretto tra questa curva teorica e la distribuzione dei dati reali rende evidenti le differenze sostanziali: mentre il modello normale è simmetrico, la distribuzione campionaria presenta una marcata asimmetria positiva. Questa discrepanza giustifica il rigetto dell'ipotesi di normalità dei dati.



**Figura 3.24:** Curva di densità di una distribuzione normale teorica.

L'analisi esplorativa condotta sul dataset `bikesharing_firenze_2023_05_trips.csv`, come per la città di Bologna, si distingue da quella effettuata sul dataset `firenze_2023_05.csv` per la diversa natura delle distanze percorse.

Il *report* riassume i risultati ottenuti mediante l'analisi generale:

Misura	Valore	Misura	Valore
Media distanza	1.54 km	Range	0.15 - 14.78 km
Mediana distanza	1.29 km	IQR	1.20 km
Deviazione standard	1.03 km	Skewness	1.59
Varianza distanza	1.05 km <sup>2</sup>	Kurtosis	4.03

**Tabella 3.9:** Statistiche descrittive delle distanze dei tragitti.

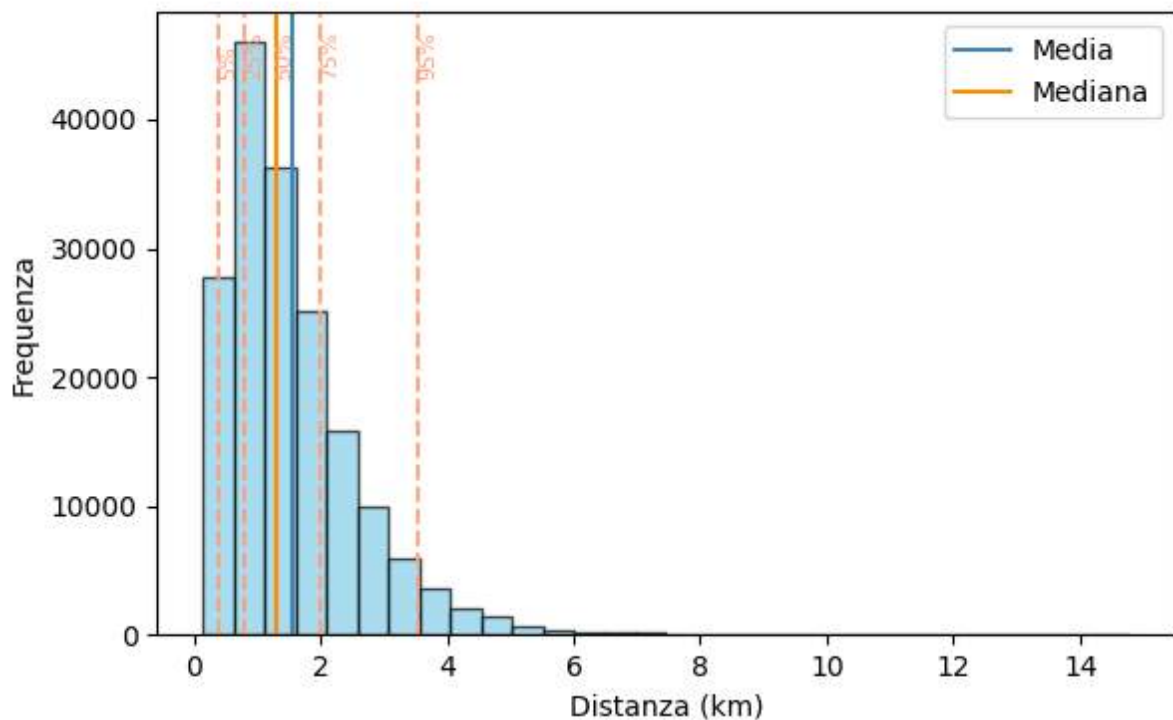
La distribuzione delle distanze percorse presenta una media di 1.54 km e una mediana pari a 1.29 km. La differenza tra i due valori indica una leggera asimmetria verso destra. La deviazione standard, pari a 1.03 km, e la varianza di 1.05 km<sup>2</sup> mostrano una variabilità moderata, con una concentrazione consistente dei dati attorno al valore medio. L'intervallo delle osservazioni è piuttosto ampio, spaziando da 0.15 km fino a 14.78 km, ma l'IQR di 1.20 km conferma che il 50% centrale dei dati resta concentrato in un range relativamente contenuto.

Gli indici di forma confermano l'asimmetria positiva (*skewness* = 1.59), che indica la presenza di viaggi più lunghi rispetto alla media, mentre la *kurtosis* (4.03) segnala una distribuzione più appuntita e con code più pesanti rispetto a quella normale, suggerendo la presenza di outlier o valori estremi, ma in misura meno marcata rispetto al dataset precedente.

- *Distribuzione delle distanze*: la distribuzione è caratterizzata da una marcata asimmetria positiva (coda a destra), con una forte concentrazione di dati su distanze brevi, a testimonianza di un utilizzo del servizio prevalentemente per tragitti corti. Gli indicatori statistici mostrano valori di tendenza centrale contenuti. La mediana si attesta a circa 1.2 km, mentre la media è leggermente superiore, con un valore di circa 1.5 km, come atteso per una distribuzione con asimmetria a destra. Si nota che il 95% di tutte le distanze è inferiore a 3.6 km.

Percentile	Distanza (km)
5°	0.38
25°	0.81
50°	1.29
75°	2.01
95°	3.54

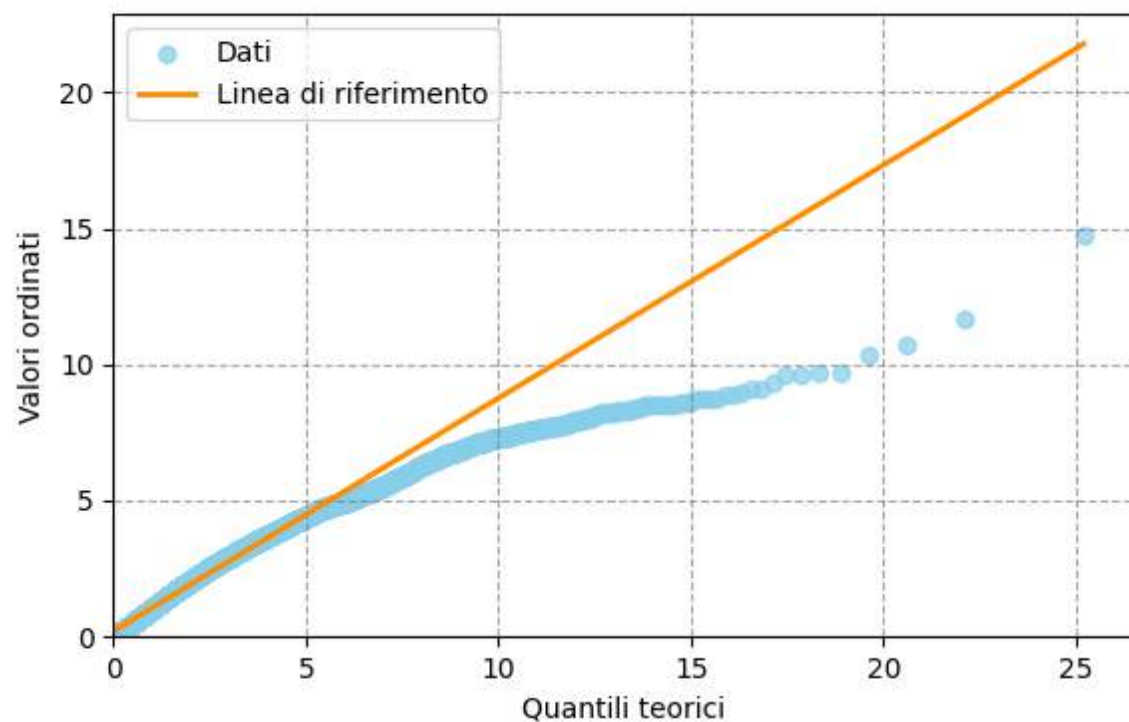
**Tabella 3.10:** Distribuzione delle distanze percorse ai diversi percentili.



**Figura 3.25:** Istogramma della distribuzione delle distanze per il dataset di Firenze.

- *QQ-plot* (Quantile-Quantile plot): la sistematica deviazione dei punti (in blu) dalla retta di riferimento (in arancione) conferma in modo evidente che la distribuzione delle distanze, anche a seguito del *map-matching*, non è normale.

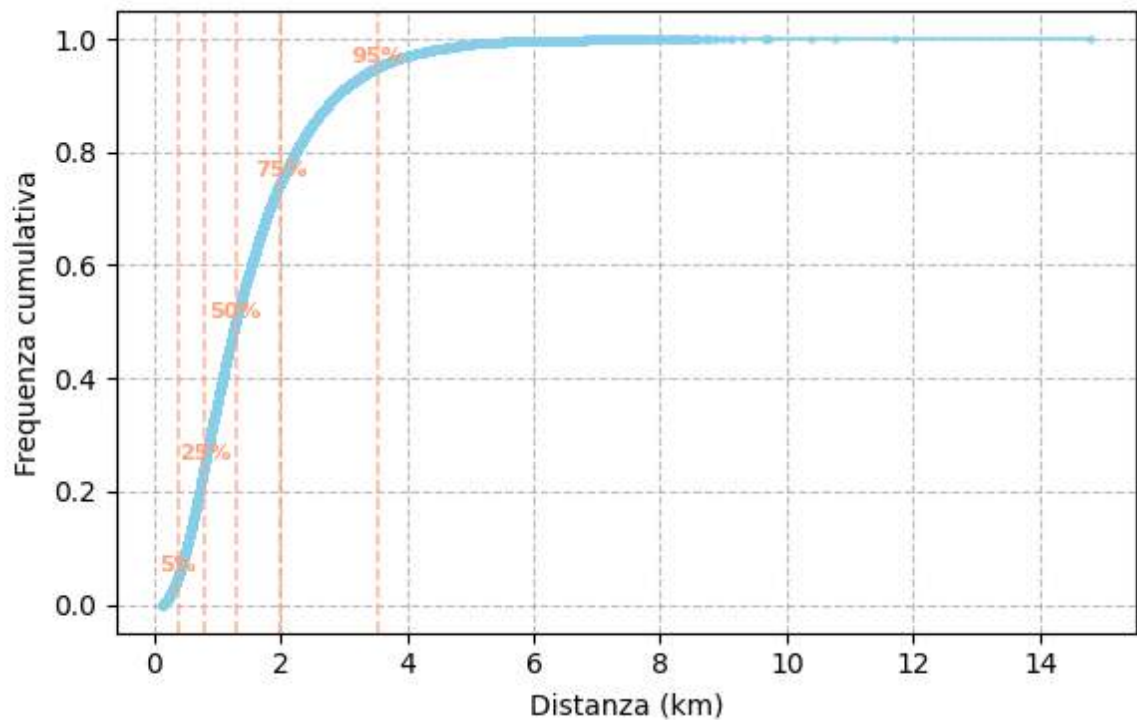
In particolare, si osserva una deviazione con una forma marcatamente concava. Mentre i valori più bassi si allineano con una certa precisione alla linea teorica, al crescere dei quantili la deviazione si accentua. Questa tendenza è particolarmente visibile nella coda destra della distribuzione: i valori ordinati si “appiattiscono” e restano ben al di sotto della retta, indicando che la coda della distribuzione campionaria è significativamente più “leggera” o “corta” rispetto a quella di una distribuzione normale. Infine, la presenza di alcuni punti isolati all’estrema destra del grafico suggerisce l’esistenza di un piccolo numero di outlier.



**Figura 3.26:** Quantile-Quantile (QQ) plot della distribuzione delle distanze.

- *Distribuzione cumulativa (CDF)*: la pendenza della curva è particolarmente accentuata per valori di distanza bassi, confermando visivamente la forte concentrazione dei dati in questo intervallo.

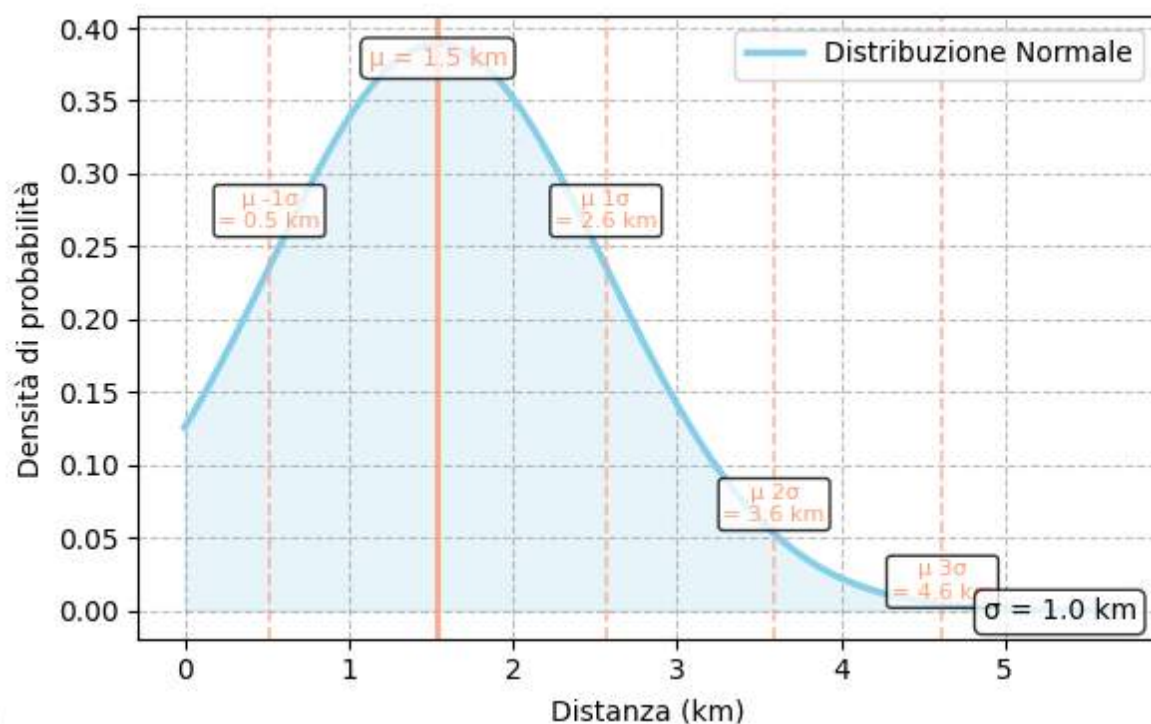
Dal grafico è possibile estrarre con precisione i percentili della distribuzione. Si osserva che il 50° percentile (la *mediana*) si attesta a circa 1.2 km, indicando che metà dei viaggi è più breve di tale valore. Similmente, si evince che il 25% dei dati presenta una distanza inferiore a 0.7 km (primo quartile) e il 75% una distanza inferiore a 2.0 km (terzo quartile). La curva si appiattisce notevolmente dopo i 4 km, avvicinandosi asintoticamente al valore 1.0; questo andamento è sottolineato dal 95° percentile, che si attesta a circa 3.6 km, a dimostrazione che la quasi totalità degli spostamenti si concentra al di sotto di questa soglia.



**Figura 3.27:** Funzione di ripartizione (CDF) delle distanze.

- *Distribuzione normale teorica*: la figura mostra la curva di densità di probabilità di una distribuzione perfettamente normale. Questa curva è stata generata utilizzando come parametri la media campionaria ( $\mu \approx 1.5$  km) e la deviazione standard campionaria ( $\sigma \approx 1.2$  km) calcolate dai dati corretti con map-matching.

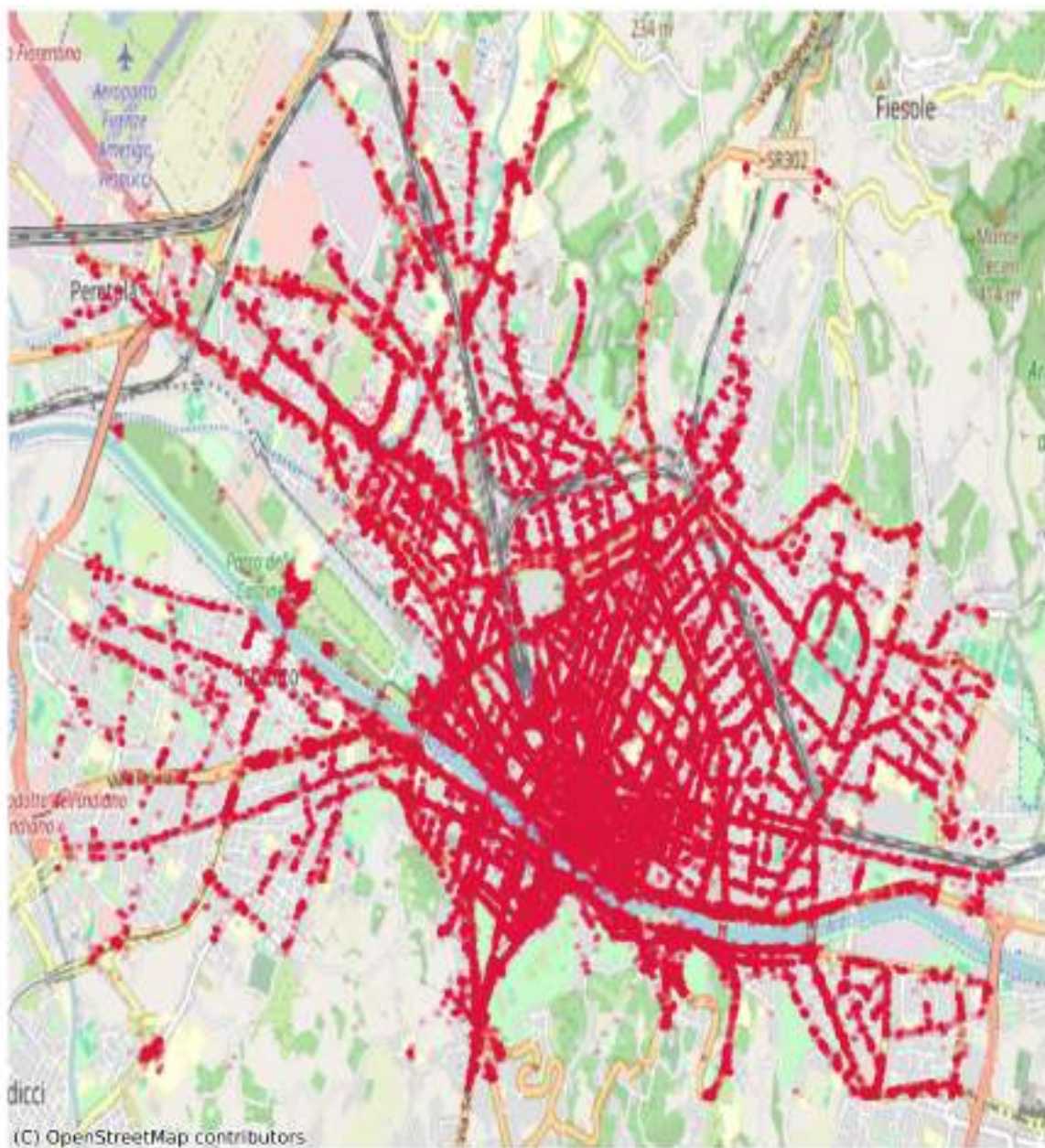
Nel grafico, la distribuzione è perfettamente simmetrica attorno alla sua media (linea blu continua), che rappresenta anche il punto di massima densità di probabilità. Le linee tratteggiate indicano gli intervalli di una, due e tre deviazioni standard dalla media. Il confronto diretto tra questo modello teorico e la distribuzione dei dati reali rende evidenti le differenze sostanziali: mentre il modello normale è simmetrico, la distribuzione campionaria presenta una marcata asimmetria positiva. Questa discrepanza giustifica il rigetto dell'ipotesi di normalità dei dati.



**Figura 3.28:** Curva di densità di una distribuzione normale teorica.

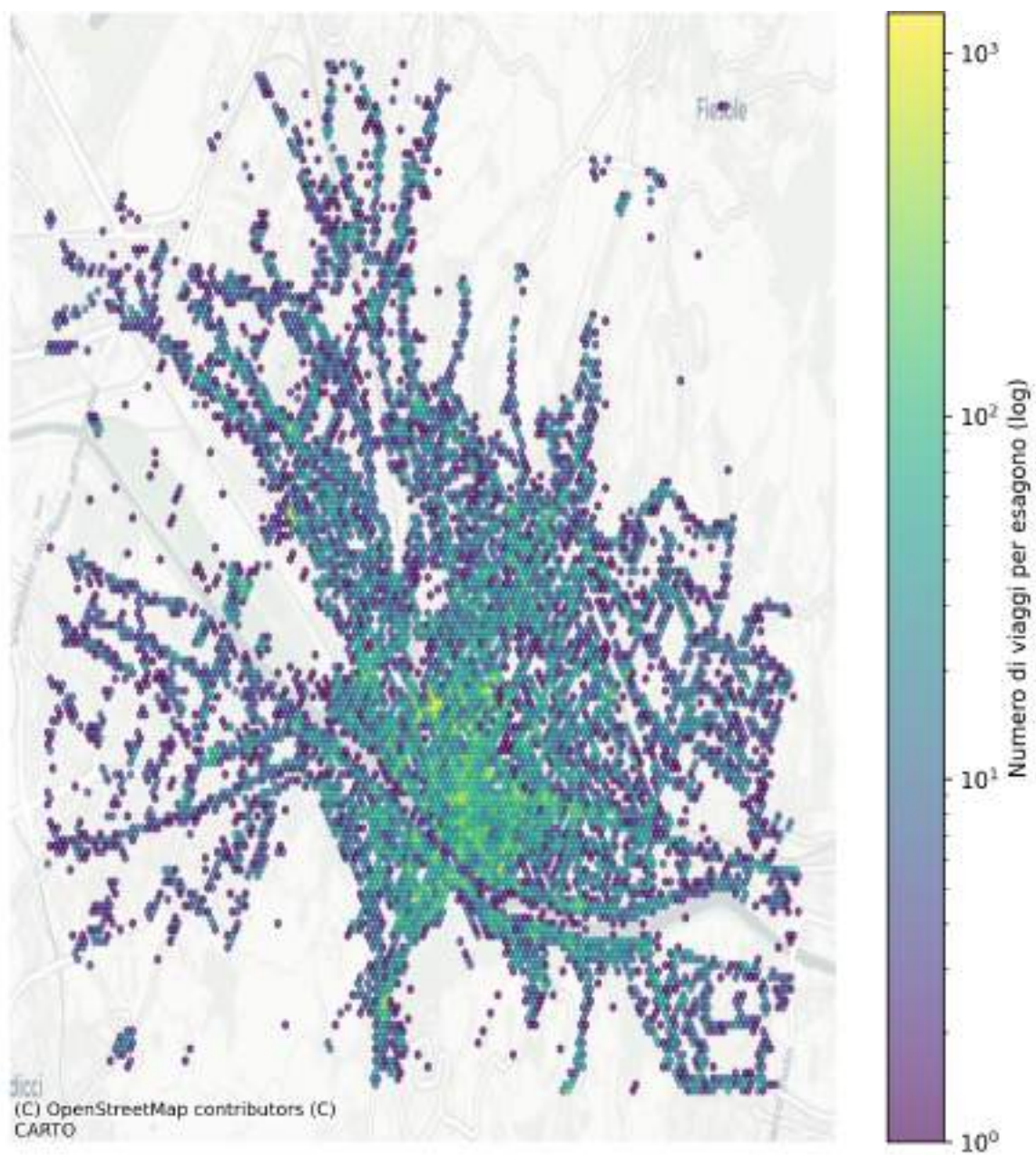




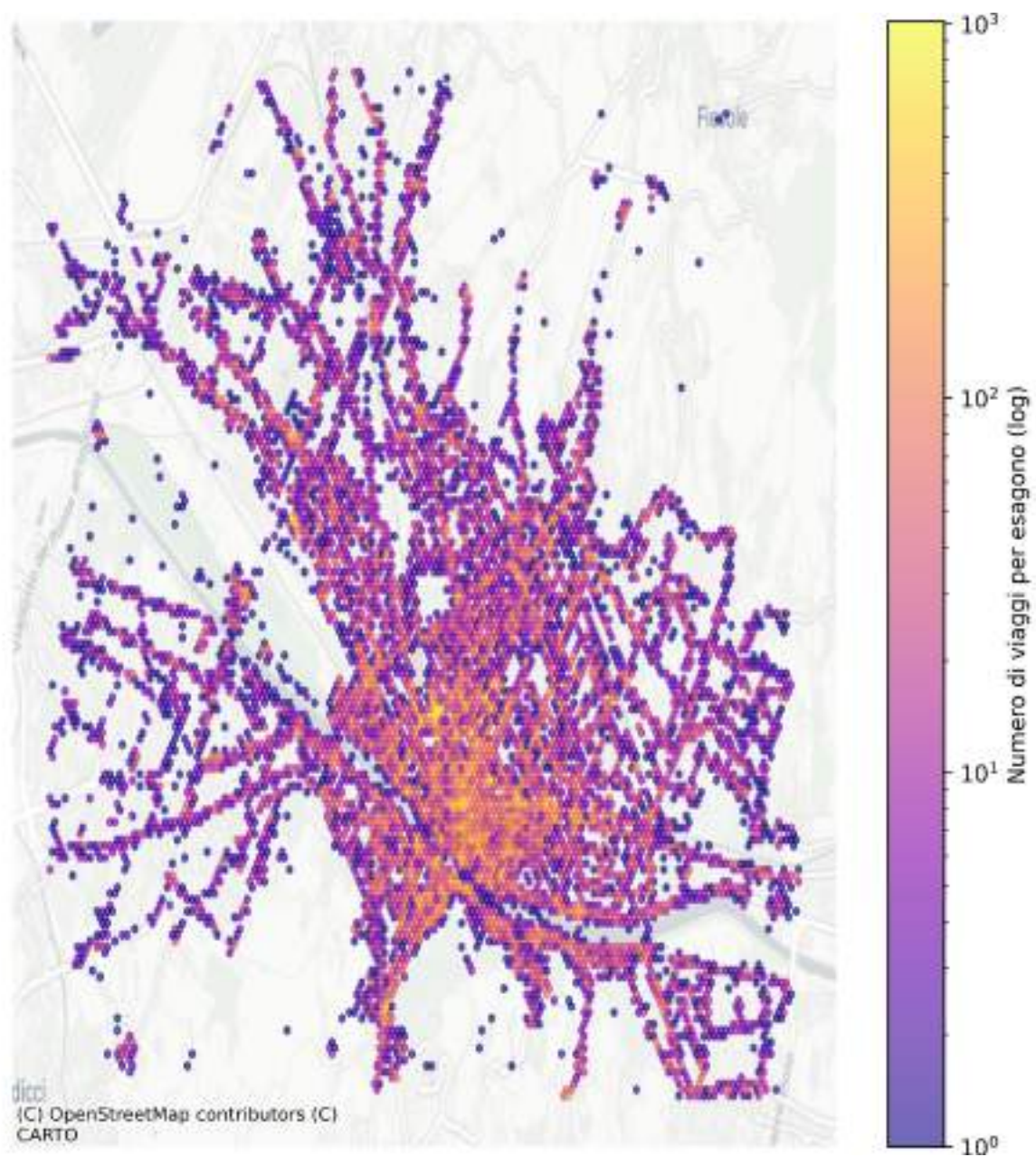


**Figura 3.30:** Distribuzione spaziale dei punti di arrivo sul territorio urbano di Bologna.





**Figura 3.31:** Mappe di densità dei punti di partenza.



**Figura 3.32:** Mappe di densità dei punti di arrivo.

### 3.3.3 Analisi tempi/velocità

L'analisi comparativa è stata condotta su 30'708 viaggi di biciclette muscolari e 122'550 viaggi di e-bike. A entrambi i gruppi sono stati applicati i medesimi filtri, considerando solo viaggi di durata compresa tra 0 e 2 ore e con una velocità media tra 5 e 30 km/h.

	Muscolari	E-bike
<b>Viaggi analizzati</b>	30'708	122'550
<b>Tempi (ore)</b>		
Media $\pm$ Dev. Std	0.198 $\pm$ 0.139	0.159 $\pm$ 0.116
Mediana	0.167	0.133
Min-Max	0.017 – 1.967	0.017 – 1.950
<b>Velocità (km/h)</b>		
Media $\pm$ Dev. Std	9.21 $\pm$ 2.63	13.41 $\pm$ 3.75
Mediana	8.89	13.46
Min-Max	5.00 – 28.20	5.00 – 29.88

**Tabella 3.11:** Report dell'analisi tempi e velocità per biciclette muscolari ed e-bike.

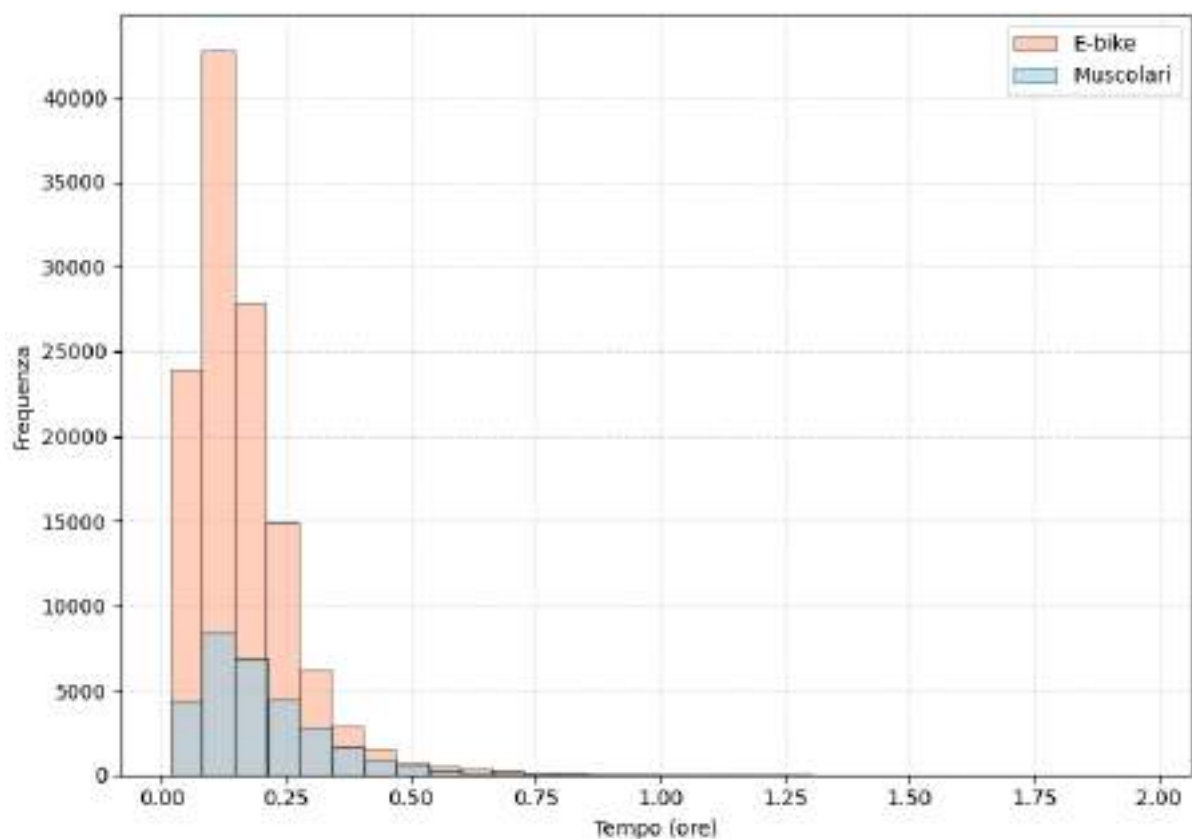
Per quanto riguarda i tempi di percorrenza, si osserva una differenza tra i due gruppi. Il tempo medio di viaggio per le e-bike si attesta a 9.54 minuti, inferiore rispetto alle bici muscolari che registrano una media di 11.88 minuti. Questa tendenza è confermata anche dai valori mediani, pari a 7.98 minuti per le e-bike e 10 minuti per le muscolari.

L'analisi delle velocità evidenzia un divario ancora più marcato. La velocità media per le e-bike è di 13.41 km/h, nettamente superiore a quella delle bici muscolari, pari a 9.21 km/h. I valori mediani confermano questa differenza, con 13.46 km/h per le e-bike contro gli 8.89 km/h delle muscolari. Si rileva inoltre una maggiore dispersione nelle velocità delle e-bike, la cui deviazione standard è di  $\pm 3.75$ , rispetto al  $\pm 2.63$  delle bici muscolari.

L'analisi tempi/velocità condotta sul dataset `bologna_2023_05.csv` ha prodotto i seguenti grafici:

- *Distribuzione tempi di percorrenza* (frequenza/tempo): il grafico confronta le distribuzioni dei tempi di percorrenza per le e-bike e le biciclette muscolari nel contesto fiorentino. L'osservazione più immediata è la netta prevalenza di viaggi effettuati con e-bike, che superano di un ordine di grandezza quelli con biciclette muscolari.

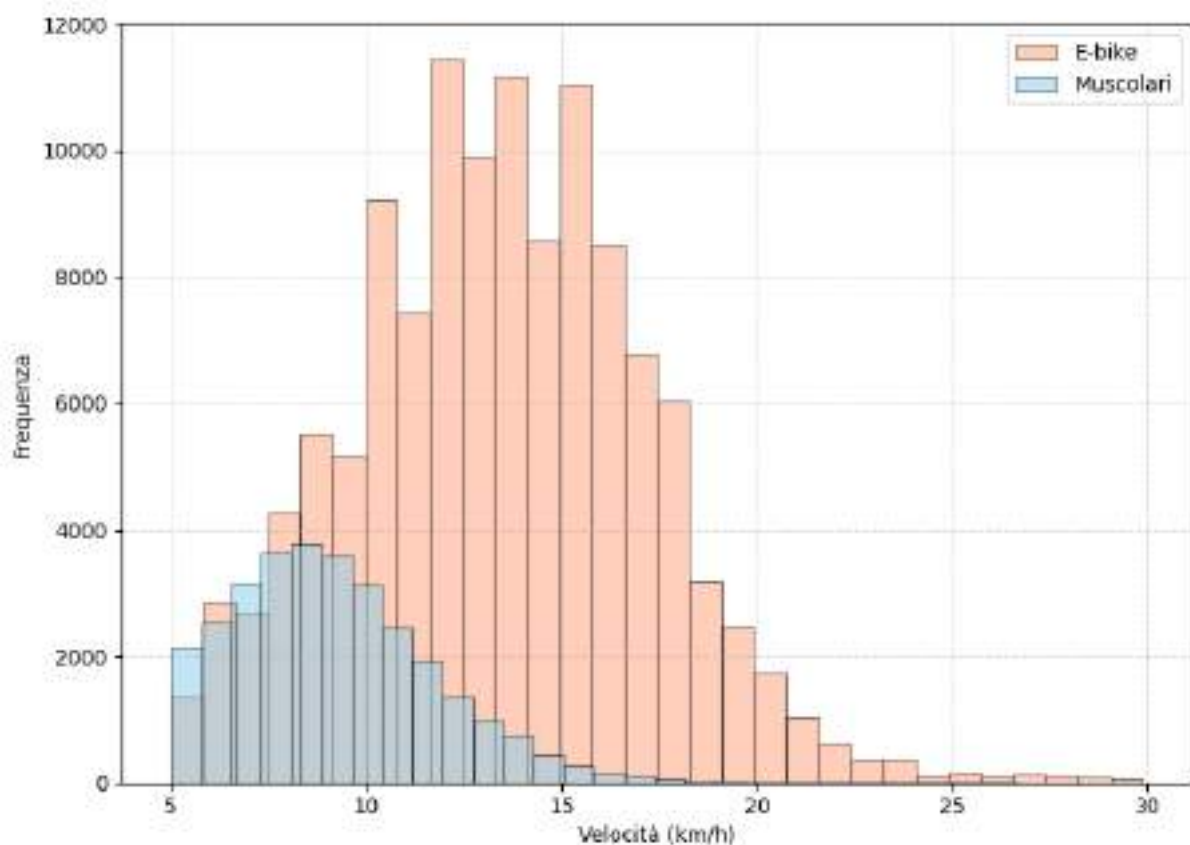
Per entrambe le categorie, le distribuzioni sono fortemente asimmetriche a destra, indicando che la maggioranza dei viaggi ha una durata molto breve. Il picco di frequenza (moda) si colloca per tempi di percorrenza inferiori a 0.25 ore (15 minuti), con la massima concentrazione di viaggi al di sotto dei 10 minuti.



**Figura 3.33:** Istogramma della distribuzione dei tempi di percorrenza.

- *Distribuzione velocità media* (frequenza/velocità): il grafico mostra una chiara e netta distinzione nelle velocità medie tenute dalle due tipologie di veicoli nel contesto fiorentino, con le e-bike che presentano velocità mediamente superiori a quelle delle biciclette muscolari.

La distribuzione delle velocità per le e-bike è approssimativamente simmetrica, con un picco di frequenza (moda) situato attorno ai 12-13 km/h. Al contrario, la distribuzione per le biciclette muscolari presenta un picco a velocità inferiori, circa 8-9 km/h. L'intera distribuzione delle e-bike è visibilmente traslata verso destra rispetto a quella delle biciclette muscolari, indicando una performance di velocità sistematicamente maggiore per i veicoli elettrici.

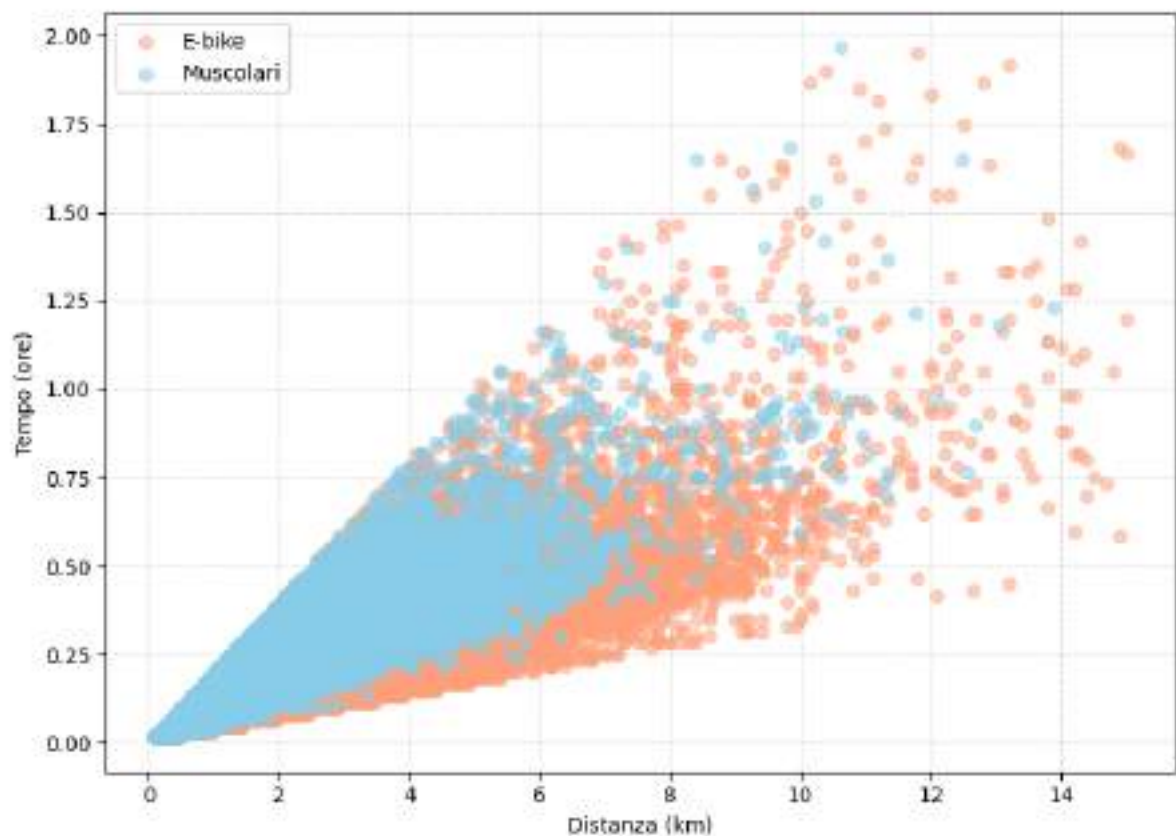


**Figura 3.34:** Istogramma della distribuzione delle velocità medie.



- *Relazione distanza-tempo*: il grafico a dispersione visualizza la relazione tra la distanza e il tempo di percorrenza per ogni singolo viaggio nel dataset di Firenze. Come atteso, si osserva una forte correlazione positiva per entrambe le tipologie di veicolo: all'aumentare della distanza percorsa, aumenta anche il tempo impiegato.

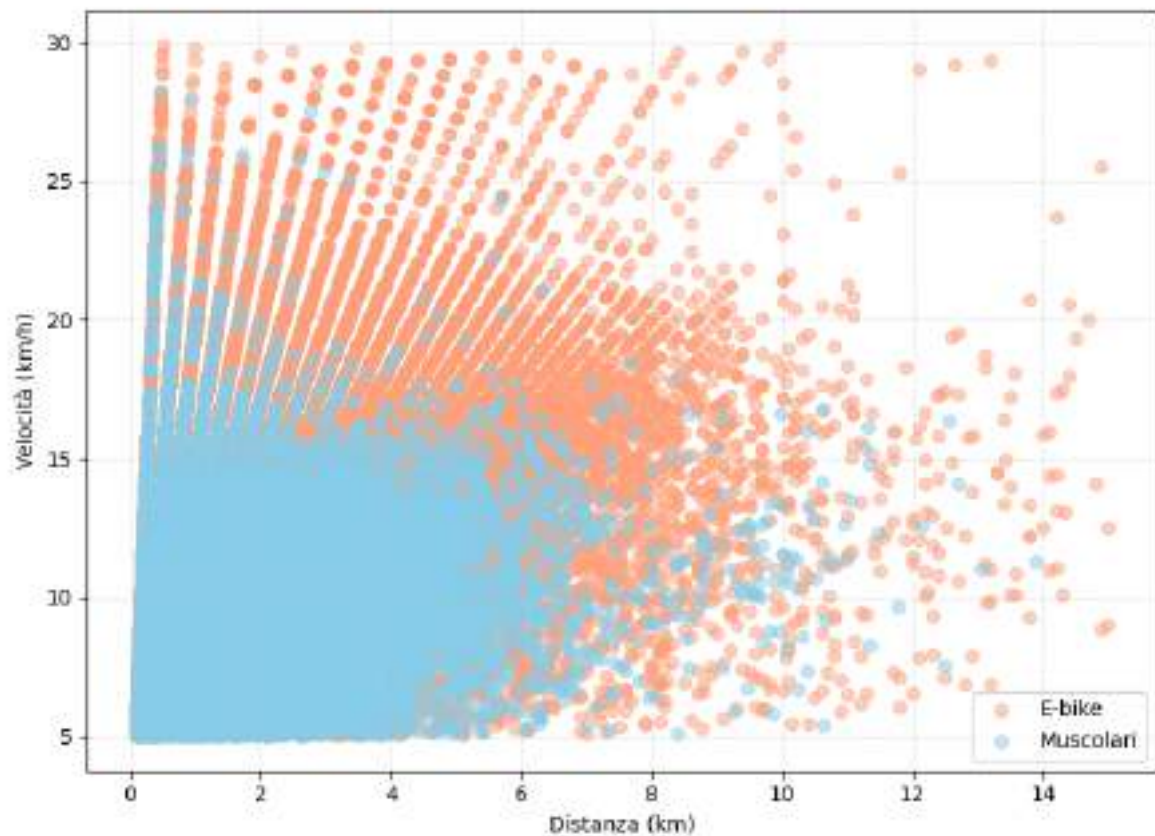
Dal grafico emerge inoltre un diverso pattern di utilizzo: le e-bike vengono impiegate per coprire un range di distanze più ampio, inclusi tragitti più lunghi che sono meno frequenti per le biciclette muscolari. Inoltre la pendenza dei punti nel grafico (tempo/distanza) è inversamente proporzionale alla velocità. Si può notare come la “nube” di punti delle e-bike presenti una pendenza mediamente inferiore rispetto a quella delle biciclette muscolari. Ciò significa che, a parità di distanza, i viaggi con e-bike richiedono generalmente meno tempo, a conferma della loro maggiore velocità media.



**Figura 3.35:** Grafico a dispersione (*scatter plot*) della relazione tra distanza percorsa (km) e tempo impiegato (ore).

- *Relazione distanza-velocità*: il grafico analizza come la velocità media di un viaggio varia in funzione della distanza percorsa. Non emerge una relazione lineare semplice; per entrambe le tipologie di veicolo, i viaggi molto brevi (indicativamente inferiori a 2 km) presentano un'ampia dispersione delle velocità medie, probabilmente a causa della maggiore incidenza delle fasi di partenza, arrivo e delle condizioni di traffico su tragitti così corti.

La differenza di performance tra le due categorie è tuttavia evidente. Le e-bike mostrano una chiara tendenza a raggiungere e mantenere velocità medie più elevate e costanti su distanze medio-lunghe. Le biciclette muscolari, oltre a essere impiegate su distanze più contenute, raggiungono il loro picco di velocità media su distanze più brevi e si concentrano in un range di velocità complessivamente inferiore.



**Figura 3.36:** Grafico a dispersione (*scatter plot*) che illustra la relazione tra la distanza percorsa (km) e la velocità media (km/h).

### 3.3.4 Analisi applicativa

L'analisi del primo dataset è stata condotta su un totale di 209·858 tragitti, comprendenti diverse tipologie di veicoli. La suddivisione mostra una chiara prevalenza delle e-bike, utilizzate in 142·252 viaggi (67.8%), seguite dalle biciclette tradizionali (**bike**) con 46·183 spostamenti (22.0%) e dagli e-scooter con 21·423 tragitti (10.2%). La distanza media complessiva è risultata pari a 1.65 km. La distribuzione delle percorrenze evidenzia una concentrazione sui tragitti brevi: 142·883 viaggi (68.1%) presentano una lunghezza inferiore a 2 km, 60·429 (28.8%) rientrano nella fascia 2–5 km, mentre solo 6·546 (3.1%) superano i 5 km.

Il secondo dataset comprende 176·149 tragitti complessivi. Anche in questo caso la maggioranza dei viaggi è stata effettuata con e-bike, pari a 121·303 spostamenti (68.9%). Seguono le biciclette tradizionali (**bicycle**) con 39·805 tragitti (22.6%) e gli e-scooter con 15·041 viaggi (8.5%). La distanza media è leggermente inferiore rispetto al dataset precedente, attestandosi a 1.54 km. La distribuzione delle lunghezze conferma una prevalenza di percorsi brevi: 131·891 viaggi (74.9%) sono inferiori a 2 km, 42·654 (24.2%) hanno una lunghezza compresa tra 2 e 5 km, mentre solo 1·604 (0.9%) superano i 5 km.

	Dataset 1	Dataset 2
<b>Distanza media del tragitto</b>	1.65 km	1.54 km
<b>Distribuzione tragitti per classe</b>		
Brevi (<2 km)	142·883	131·891
Medi (2–5 km)	60·429	42·654
Lunghi (>5 km)	6546	1604
<b>Distribuzione per tipologia di veicolo</b>		
E-bike / ebicycle	142·252	121·303
Muscolari / bicycle	46·183	39·805
Escooter / scooter	21·423	15·041

**Tabella 3.12:** Confronto dei report dell'analisi applicativa per i due dataset.

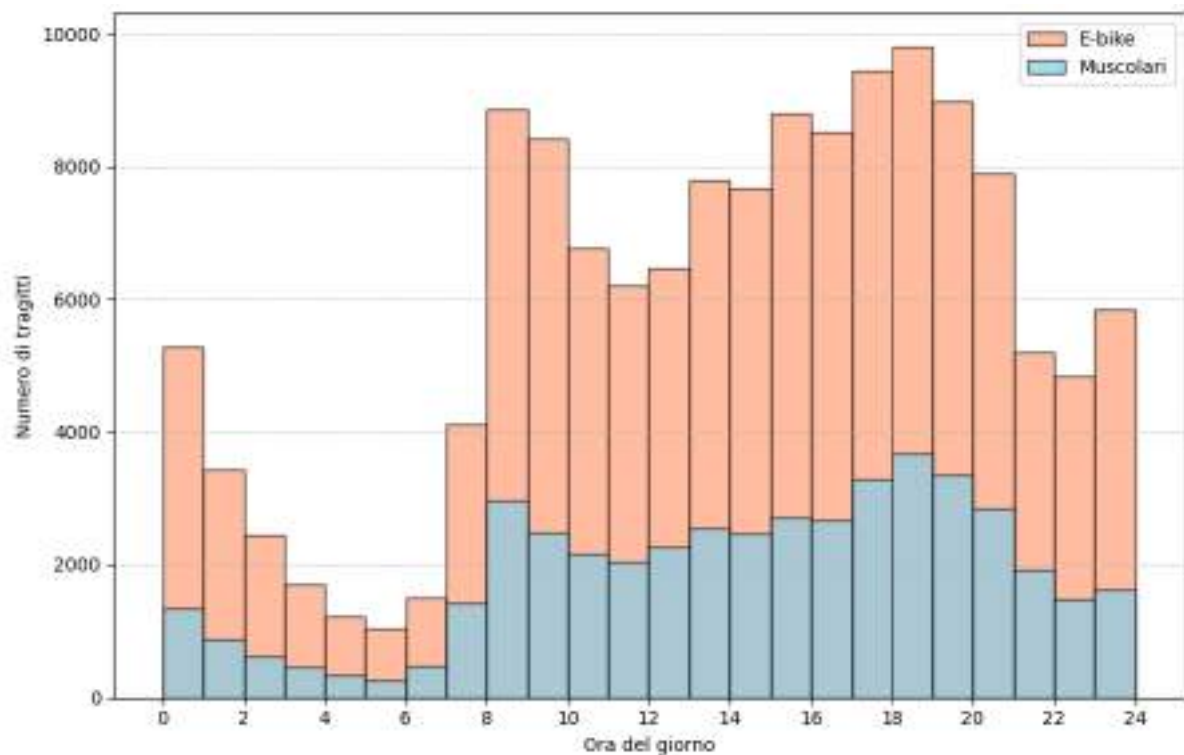


### 3.3.5 Analisi temporale

L'analisi temporale condotta sul dataset `firenze_2023_05.csv` ha prodotto i seguenti grafici:

- *Distribuzione oraria dei tragitti*: il grafico illustra come l'utilizzo del servizio di bike sharing si distribuisce nell'arco della giornata a Firenze. Emerge una chiara distribuzione bimodale, caratteristica dei flussi di mobilità urbana, con due distinti picchi di utilizzo che scandiscono la giornata.

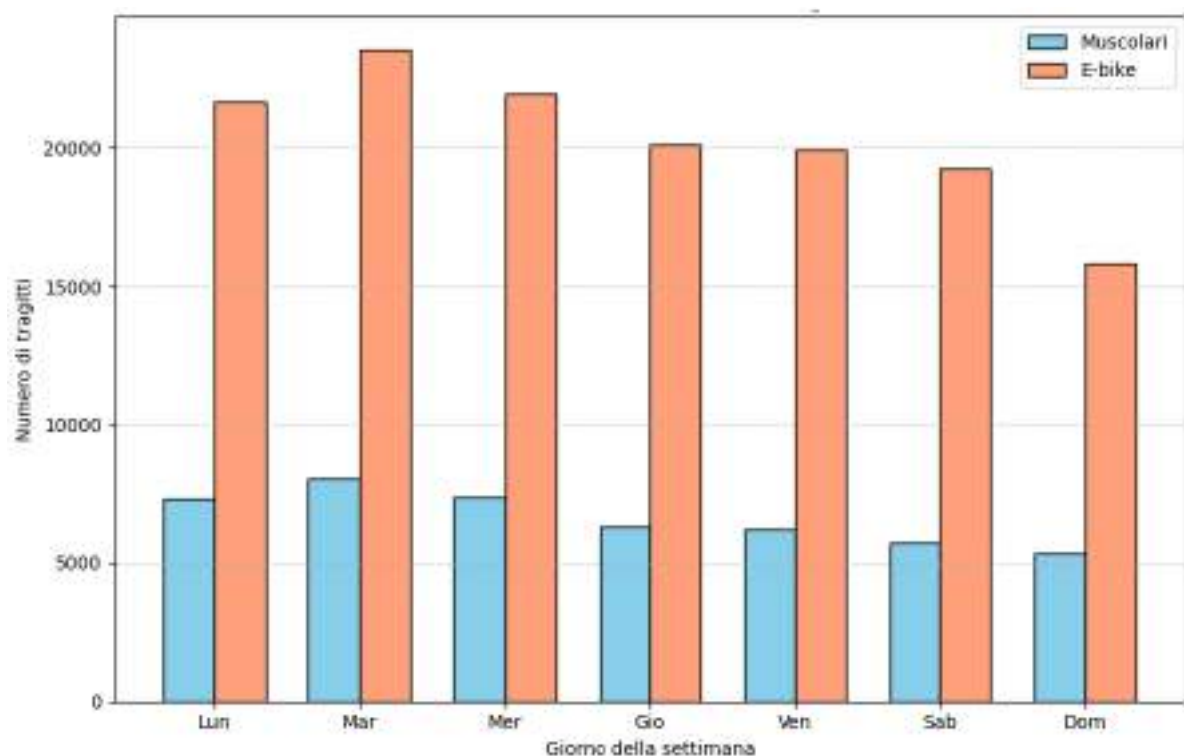
Si osserva un primo picco mattutino tra le 8:00 e le 9:00 e un secondo picco nel tardo pomeriggio, con la massima intensità di viaggi registrata tra le 18:00 e le 19:00. L'attività decresce significativamente dopo le 21:00, raggiungendo un minimo nelle ore notturne.



**Figura 3.37:** Distribuzione oraria del numero di viaggi, suddivisi per tipologia di bicicletta: e-bike (arancione) e muscolare (azzurro).

- *Distribuzione settimanale dei tragitti*: il grafico illustra la distribuzione dei viaggi nell'arco della settimana a Firenze, da cui emerge un pattern di utilizzo marcatamente feriale, indicativo di un uso prevalente del servizio per spostamenti sistematici come il pendolarismo casa-lavoro/studio.

Si osserva un'intensità di utilizzo elevata e relativamente costante durante i giorni feriali, con un picco assoluto registrato il martedì. Al contrario, si verifica un calo significativo dell'utilizzo durante il fine settimana, con la domenica che risulta essere il giorno di minore attività. Anche in questo caso, si nota la quota rilevante dei viaggi effettuati con biciclette muscolari, che seguono lo stesso andamento settimanale delle e-bike.

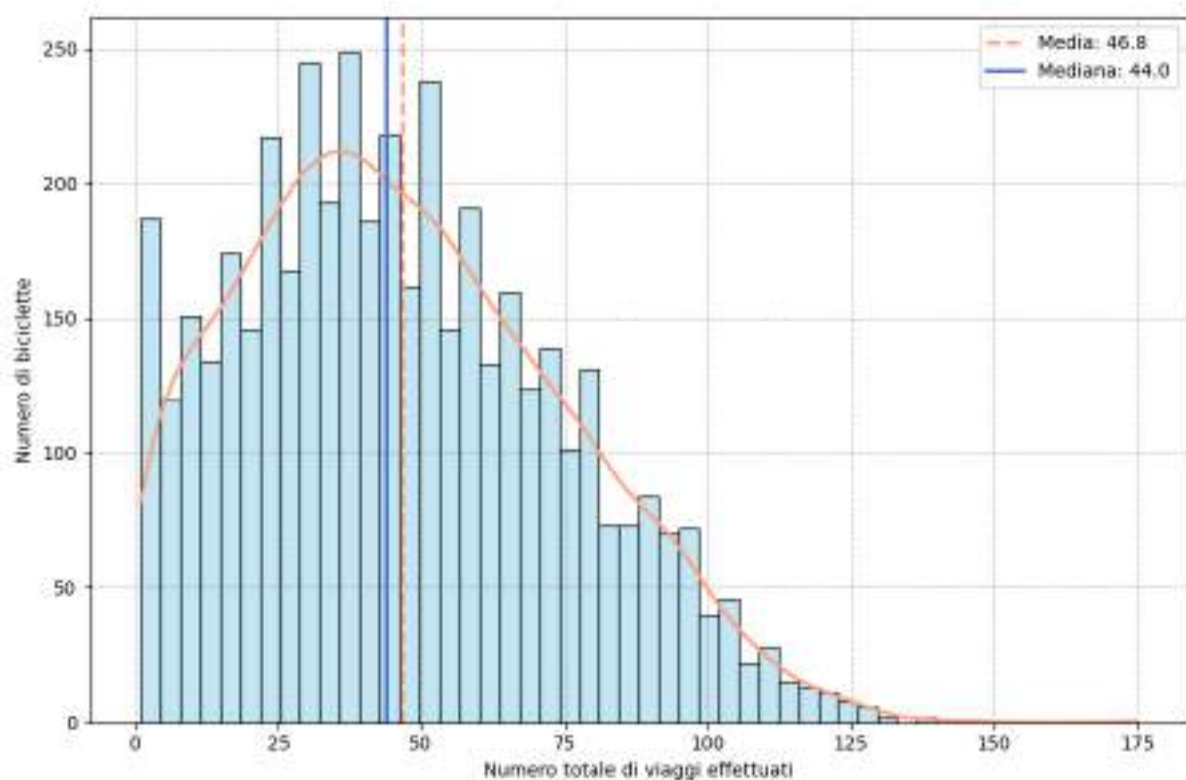


**Figura 3.38:** Distribuzione del numero di viaggi per giorno della settimana, suddivisi per tipologia di bicicletta: e-bike (arancione) e muscolare (azzurro).

### 3.3.6 Analisi dei turnover

L'analisi dei turnover condotta sul dataset `firenze_2023_05.csv` ha prodotto i seguenti grafici:

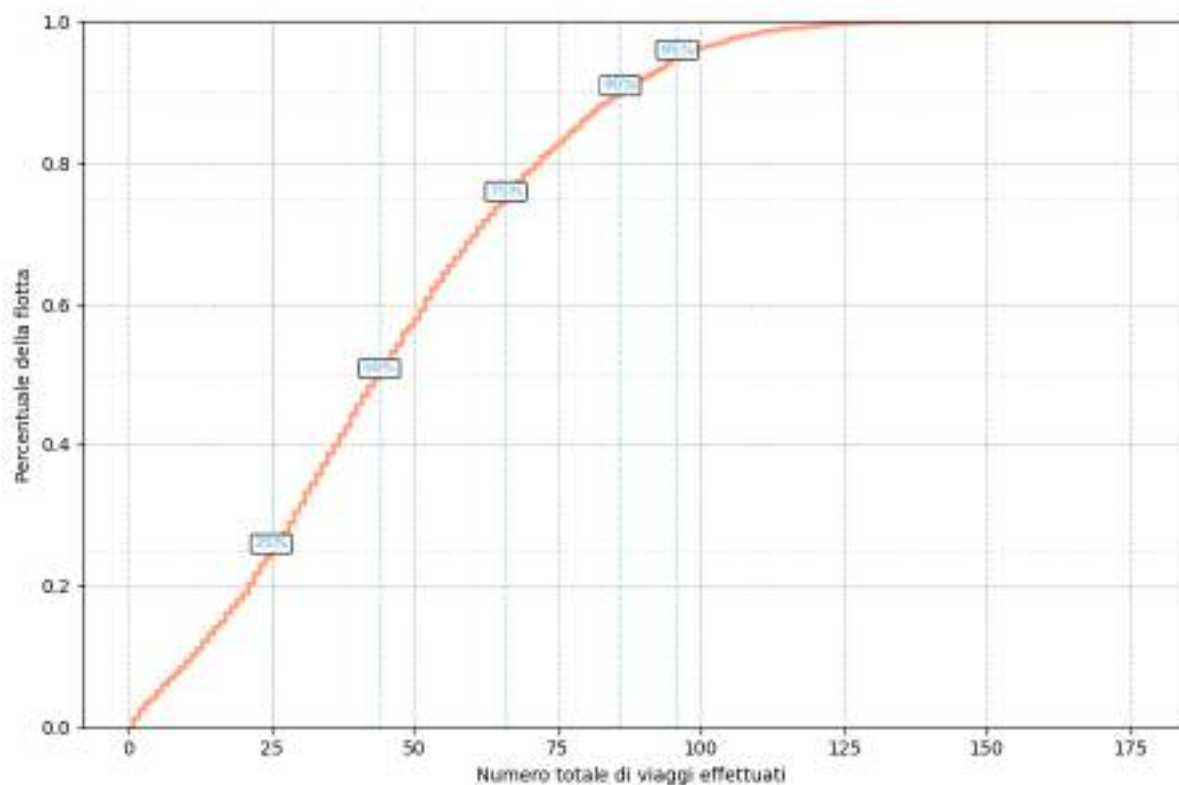
- Il primo grafico mostra come si distribuisce il carico di lavoro sull'intera flotta di biciclette a Firenze. La distribuzione del numero di viaggi per veicolo è unimodale e presenta una lieve asimmetria a destra (positiva), come si può dedurre dal valore della media (46.8), che è leggermente superiore a quello della mediana (44.0). Questo indica la presenza di un numero limitato di biciclette con un numero di viaggi molto elevato che “trascina” verso destra il valore medio. La maggior parte dei veicoli della flotta ha effettuato un numero di viaggi compreso tra 10 e 100, con un picco di frequenza (moda) intorno ai 40 viaggi. Nel complesso, questo risultato suggerisce un utilizzo della flotta relativamente bilanciato anche a Firenze, dove si osserva una contribuzione distribuita sulla maggior parte della flotta.



**Figura 3.39:** Istogramma della distribuzione del numero totale di viaggi per bicicletta.

- Il secondo grafico illustra la funzione di ripartizione empirica, che completa l'analisi sulla distribuzione del carico di lavoro all'interno della flotta di Firenze. La forma a "S" della curva conferma la concentrazione dei dati attorno a un valore centrale, come già evidenziato dall'istogramma.

Dal grafico è possibile estrarre con precisione i percentili di utilizzo: si nota che il 50% della flotta (la mediana) ha compiuto meno di 44 viaggi. L'intervallo interquartile, che rappresenta il 50% centrale dei veicoli, è compreso approssimativamente tra 29 viaggi (25° percentile) e 63 viaggi (75° percentile). Questi dati rafforzano la conclusione di un utilizzo relativamente bilanciato della flotta, dove la maggior parte dei veicoli contribuisce in modo comparabile all'operatività del servizio, senza una coda eccessivamente lunga di veicoli sottoutilizzati o sovraccaricati.



**Figura 3.40:** Funzione di distribuzione cumulativa (CDF) del numero totale di viaggi per bicicletta.

## Capitolo 4

### Discussione

## 4.1 Introduzione e sintesi dei risultati chiave

Il presente capitolo si propone di discutere e interpretare in modo critico i risultati emersi dall'analisi dei dati relativi ai servizi di bike sharing nelle città di Bologna e Firenze. L'analisi, condotta sui dataset del mese di maggio 2023, ha avuto come scopo primario la valutazione della funzionalità complessiva dei servizi nei due distinti contesti urbani. L'obiettivo di questa discussione è quello di contestualizzare i risultati, evidenziare le tendenze più significative, e istituire un confronto sistematico tra i modelli di mobilità che emergono, al fine di trarre conclusioni approfondite sullo stato dei servizi e sulle loro implicazioni per la mobilità urbana sostenibile.

Dall'indagine analitica sono scaturiti diversi risultati chiave che costituiscono la spina dorsale di questa discussione e che verranno approfonditi nelle sezioni successive.

- **La predominanza degli spostamenti brevi:** il bike sharing viene utilizzato soprattutto per coprire brevi distanze, sia a Bologna che a Firenze. In entrambe le città, le distanze percorse mostrano una forte concentrazione sotto i pochi chilometri. La distanza mediana è di 2.12 km a Bologna e 1.57 km a Firenze, confermando l'efficacia del servizio come soluzione di micromobilità.
- **La preferenza per la pedalata assistita:** i dati mostrano che le e-bike costituiscono la maggior parte dei viaggi registrati, rappresentando il 91% a Bologna e il 68% a Firenze. Sebbene non sia possibile stabilire se la maggiore quota di viaggi effettuati con e-bike sia dovuta a una preferenza degli utenti o a una maggiore disponibilità di questi mezzi, le aziende di *bike-sharing* tendono comunque a orientarsi verso l'elettrificazione dei loro parchi bici.
- **La divergenza nei modelli d'uso settimanale:** il confronto tra le due città rivela una differenza significativa negli andamenti settimanali: Firenze è orientata al pendolarismo feriale mentre Bologna mostra un uso più vario e intenso anche nei giorni festivi.

Questi quattro pilastri tematici verranno ora analizzati nel dettaglio, interpretando il loro significato e discutendone le possibili cause e implicazioni.

## 4.2 Interpretazione dei pattern d'uso del servizio

Un dato comune a entrambe le città è la caratteristica distribuzione delle distanze percorse. Sia a Bologna che a Firenze, la distribuzione è fortemente asimmetrica, con una coda a destra che indica una grande concentrazione di viaggi su tragitti brevi. A Bologna, la maggior parte dei dati si colloca al di sotto dei 4.00 km, con una distanza media di 2.48 km (2.12 km la mediana). A Firenze, le distanze sono ancora più contenute, con una mediana di 1.57 km e il 75% dei viaggi che non supera i 2.60 km. Questo pattern è coerente con l'utilizzo del bike sharing per spostamenti brevi e frequenti all'interno della città, dove la maggior parte dei viaggi riguarda percorsi di pochi chilometri, piuttosto che lunghe tratte. La micromobilità di questo tipo ha il vantaggio di ridurre i tempi di percorrenza per brevi distanze rispetto ai mezzi tradizionali come auto o trasporto pubblico. Inoltre, favorisce una maggiore accessibilità alle aree centrali, contribuendo a decongestionare il traffico e a diminuire le emissioni locali, pur garantendo agli utenti un mezzo leggero, economico e facilmente fruibile.

L'analisi ha poi mostrato in modo inequivocabile un utilizzo maggiore delle e-bike da parte degli utenti. A Bologna, i viaggi con e-bike superano di un ordine di grandezza quelli con bici muscolari. A Firenze, le e-bike rappresentano la maggioranza assoluta dei viaggi, seguite dalle bici tradizionali e dagli e-scooter. Questa preferenza può essere direttamente collegata alle performance superiori dei veicoli a pedalata assistita. In entrambe le città, le e-bike mantengono una velocità media significativamente più elevata: a Bologna, la velocità media è di 13.21 km/h per le e-bike contro i 9.28 km/h delle muscolari; a Firenze, il divario è simile, con 13.41 km/h contro 9.21 km/h. È interessante notare come questo vantaggio in termini di velocità non si traduca necessariamente in viaggi di durata maggiore. A Bologna, i tempi medi e mediani di percorrenza sono quasi identici per le due tipologie di veicolo. A Firenze, invece, i viaggi con e-bike sono mediamente più brevi (9.5 minuti) rispetto a quelli con bici muscolari (12 minuti). Ciò suggerisce che le e-bike vengono utilizzate per coprire distanze simili o maggiori in minor tempo, aumentando la velocità e, forse, la convenienza dello spostamento. Questa interpretazione è confermata visivamente dai grafici di relazione distanza-tempo, dove la pendenza mediamente inferiore della nube di punti delle e-bike indica una velocità superiore.

## **4.3 Analisi comparativa: Bologna vs. Firenze**

### **4.3.1 Pattern settimanale: uso funzionale e uso misto**

Il confronto più significativo tra le due città risulta dall'analisi della distribuzione settimanale dei viaggi. A Firenze si delinea un chiaro pattern di utilizzo feriale, indicativo di un uso prevalente per spostamenti sistematici come il pendolarismo. L'attività è elevata e costante nei giorni lavorativi, con un picco registrato il martedì, per poi calare significativamente durante il fine settimana, con la domenica come giorno di minor attività in assoluto.

Bologna, al contrario, mostra un modello d'uso più complesso e ibrido. Sebbene l'utilizzo durante la settimana sia sostenuto, i giorni di maggiore attività sono il sabato, che rappresenta il picco assoluto, il lunedì e la domenica. L'elevato utilizzo durante il weekend suggerisce fortemente che il servizio non sia legato solo al pendolarismo, ma anche ad attività ricreative, sociali e turistiche. Questa divergenza potrebbe essere legata a molteplici fattori, tra cui la diversa vocazione delle città, le caratteristiche demografiche dell'utenza o le specifiche politiche di mobilità urbana.

### **4.3.2 Pattern orario: una bimodalità condivisa**

Se l'analisi settimanale ha messo in luce forti differenze tra Bologna e Firenze, la distribuzione giornaliera dei viaggi rivela invece un punto di convergenza: entrambe le città mostrano un chiaro andamento bimodale. Si tratta di un pattern tipico della mobilità urbana, che riflette i ritmi collettivi della vita cittadina.

La giornata inizia con un primo picco di utilizzo tra le 8:00 e le 9:00 del mattino, coincidente con l'inizio delle attività lavorative e scolastiche, a testimonianza dei flussi di pendolarismo. Successivamente, il servizio continua a essere utilizzato in modo stabile nelle ore centrali, per poi registrare un secondo picco nel tardo pomeriggio, tra le 18:00 e le 19:00. Quest'ultimo momento di massima attività è alimentato sia dai rientri da lavoro e università, sia dagli spostamenti legati al tempo libero e alla socialità. Dopo le 21:00–22:00, l'utilizzo cala drasticamente, raggiungendo un minimo nelle ore notturne tra le 2:00 e le 6:00 del mattino, quando la città si ferma.

La forte somiglianza tra i due contesti conferma che il bike sharing non è un servizio marginale o occasionale, ma una componente stabile e funzionale della mobilità urbana, in grado di adattarsi ai ritmi quotidiani e rispondere in modo efficiente ai bisogni della collettività.



## 4.4 Limiti dello studio

Per garantire una valutazione onesta e trasparente dei risultati ottenuti, è fondamentale riconoscere e discutere i limiti intrinseci di questa ricerca. L'identificazione di tali confini non inficia il valore delle analisi condotte, ma ne definisce il perimetro di validità e invita a una corretta e prudente interpretazione delle conclusioni.

Una prima e importante considerazione riguarda l'orizzonte temporale dell'analisi. I dati utilizzati si riferiscono esclusivamente al mese di maggio del 2023 per entrambe le città. Sebbene questo periodo primaverile offra uno spaccato significativo dell'utilizzo del servizio in condizioni climatiche generalmente favorevoli, esso costituisce una "fotografia" puntuale che non può catturare le significative variazioni stagionali. L'uso della bicicletta, e di conseguenza dei servizi di sharing, è notoriamente influenzato da fattori come le condizioni meteorologiche, le ore di luce disponibili e i flussi turistici, che raggiungono il loro apice in altri periodi dell'anno. Di conseguenza, i pattern qui descritti potrebbero differire notevolmente in altri mesi, e una comprensione completa della domanda richiederebbe un'analisi longitudinale estesa ad almeno un intero anno solare.

Un secondo limite, intrinseco alla natura dei dati operativi a disposizione 2.3, è l'assenza di informazioni socio-demografiche sugli utenti. Le analisi si basano infatti esclusivamente sui dati di viaggio, precludendo quindi l'accesso a speculazioni basate su dettagli sul profilo di chi utilizza il servizio, come età, genere, status di residente, studente o turista. Questa mancanza impedisce di correlare in modo diretto i pattern d'uso a specifici segmenti di popolazione. Pertanto, le conclusioni sul motivo degli spostamenti come la distinzione tra uso pendolare e ricreativo rimangono delle inferenze, per quanto logicamente robuste, basate unicamente sui dati temporali e spaziali. La disponibilità di dati demografici avrebbe permesso di validare queste ipotesi con un grado di certezza superiore.

Infine, è necessario usare particolare cautela nell'interpretare i risultati relativi all'analisi del turnover della flotta. Tale analisi, infatti, è stata condotta su un'unica, specifica giornata per ciascuna città: il 1° Maggio per Bologna e il 30 Aprile per Firenze. La scelta di queste date, che corrispondono a un giorno festivo e a un prefestivo, le rende potenzialmente non rappresentative di una giornata lavorativa tipica. I comportamenti di mobilità in tali occasioni possono essere anomali, influenzando drasticamente i tassi di utilizzo. I risultati emersi, e in particolare il dato estremamente basso di utilizzo registrato a Firenze, devono quindi essere considerati come un'indicazione puntuale e non necessariamente generalizzabile all'efficienza operativa media del servizio. Una valutazione più attendibile richiederebbe il calcolo del turnover su un campione di più giorni, includendo sia giornate feriali che festive.

## 4.5 Implicazioni e direzioni future

L'analisi condotta e i risultati emersi non solo forniscono una fotografia dettagliata dei servizi di bike sharing a Bologna e Firenze, ma offrono anche una serie di implicazioni pratiche per gli stakeholder del settore e tracciano percorsi promettenti per la ricerca futura. Questa sezione finale si propone di esplorare in modo approfondito tali aspetti, collegando i dati alle decisioni strategiche e operative e delineando come le conoscenze acquisite possano essere ulteriormente consolidate.

### 4.5.1 Implicazioni per gestori e pianificatori urbani

I risultati di questo studio hanno una ricaduta diretta e pragmatica sulle strategie dei due principali attori coinvolti nella mobilità urbana: i gestori dei servizi di sharing e i pianificatori urbani responsabili dello sviluppo infrastrutturale e normativo.

**Per i gestori del servizio:** le evidenze emerse dall'analisi possono guidare decisioni cruciali sia in termini strategici che operativi; la netta preferenza degli utenti per le e-bike rappresenta un'indicazione chiara per la futura composizione della flotta: la maggiore velocità ed efficienza di questi mezzi ne confermano la superiorità rispetto alle bici muscolari, che potrebbero essere relegate a un ruolo marginale in contesti specifici come le aree universitarie pianeggianti. Inoltre, il tasso elevato di inattività dei mezzi, come osservato in particolare a Firenze, solleva interrogativi sul corretto dimensionamento della flotta, suggerendo l'opportunità di una valutazione costi-benefici per evitare investimenti eccessivi in veicoli non utilizzati.

Anche sul piano operativo, i risultati forniscono indicazioni utili per l'ottimizzazione delle attività di ribilanciamento. Le mappe di densità e i pattern temporali possono orientare le strategie di redistribuzione: a Bologna, dove l'attività è significativa anche nel weekend, sarà necessario affiancare alla copertura dei flussi pendolari dei giorni feriali una strategia mirata alle zone turistiche e ricreative nel fine settimana. A Firenze, invece, l'uso prevalentemente feriale suggerisce una focalizzazione sui nodi casa-lavoro nei momenti di picco giornaliero. In entrambi i casi, queste informazioni possono alimentare modelli predittivi in grado di anticipare la domanda e rendere le operazioni più efficienti e reattive.

**Per i pianificatori urbani e i policy maker:** l'utilizzo aggregato del servizio di bike sharing va ben oltre la sfera gestionale, diventando uno strumento fondamentale per la pianificazione urbana. Le mappe di calore che evidenziano le zone di partenza e arrivo dei tragitti permettono di individuare i corridoi preferenziali della mobilità ciclistica, utili per orientare gli investimenti infrastrutturali. Questi dati suggeriscono dove costruire nuove piste ciclabili, migliorare la sicurezza degli incroci e aumentare la capillarità della rete urbana.

La conferma che il bike sharing viene usato per tragitti brevi ne evidenzia il ruolo nell'ambito della mobilità intermodale. Le amministrazioni potrebbero migliorare l'integrazione con il trasporto pubblico, ad esempio aumentando il numero di mezzi presso snodi strategici. I diversi profili d'uso suggeriscono politiche mirate: Bologna potrebbe puntare sul cicloturismo, mentre Firenze potrebbe incentivare l'uso feriale del servizio da parte dei lavoratori, alleggerendo la pressione sul trasporto pubblico.

#### 4.5.2 Direzioni per la ricerca futura

I limiti dello studio aprono la strada a nuove prospettive di ricerca in grado di approfondire ulteriormente la comprensione del fenomeno e ampliare il potenziale applicativo dei risultati ottenuti.

Un primo ambito promettente riguarda l'estensione dell'analisi a un orizzonte temporale più ampio, superando la stagionalità del dataset attuale. Un'indagine pluriennale permetterebbe di individuare non solo pattern stagionali ricorrenti ma anche trend evolutivi nel tempo. Inoltre, l'analisi comparativa potrebbe essere estesa ad altre città italiane ed europee con caratteristiche diverse per distinguere gli elementi generalizzabili da quelli specifici del contesto locale. Un secondo filone di ricerca riguarda l'integrazione dei dati di utilizzo con fonti esterne, un tipo di arricchimento dei dati che potrebbe sostenere lo sviluppo di modelli predittivi avanzati, basati su tecniche di machine learning, in grado di anticipare con precisione la domanda a breve termine, fornendo un vantaggio strategico decisivo per i gestori.

Infine, un ulteriore passo avanti sarebbe rappresentato dall'integrazione tra analisi quantitative e approcci qualitativi: attraverso survey rivolti agli utenti, si potrebbero esplorare le motivazioni alla base delle scelte modali, le barriere all'adozione, il grado di soddisfazione e le aspettative nei confronti del servizio. Solo adottando una prospettiva realmente user-centric sarà possibile validare in modo solido le ipotesi sullo scopo degli spostamenti e progettare servizi più inclusivi, funzionali e rispondenti alle reali esigenze della cittadinanza.

## 4.6 Conclusioni

Nel contesto di un dibattito globale sempre più centrato sulla necessità di modelli di mobilità urbana sostenibile, il presente studio si è proposto di fornire una caratterizzazione dettagliata e un'analisi comparativa rigorosa dei servizi di bike sharing in due importanti città italiane: Bologna e Firenze. Partendo dalla consapevolezza che tali servizi rappresentano una componente chiave delle moderne città *smart*, l'obiettivo non era solo descrivere, ma interpretare le dinamiche d'uso, decifrare i comportamenti degli utenti e comprendere il ruolo effettivo di questi sistemi nell'ecosistema dei trasporti urbani. Attraverso un'analisi approfondita di dati reali questa tesi ha cercato di rispondere a domande fondamentali sulla funzionalità, l'efficienza e il profilo di utilizzo, offrendo una base empirica solida per valutazioni future.

Il percorso analitico ha portato a una serie di scoperte che, nel loro insieme, dipingono un quadro ricco e sfaccettato. In primo luogo, è stato confermato in modo conclusivo il ruolo del bike sharing nel dominio della micromobilità. Con le distanze mediane modeste rilevate è poi evidente come il servizio risponda primariamente a un'esigenza di spostamenti brevi, rapidi e capillari, affermandosi come soluzione ideale per l'ultimo chilometro e come alternativa flessibile al trasporto pubblico su corte distanze. In secondo luogo, lo studio ha sancito la supremazia tecnologica e la preferenza degli utenti per le biciclette a pedalata assistita. La loro schiacciante popolarità in entrambe le città è direttamente riconducibile a una performance nettamente superiore: le e-bike si sono dimostrate più veloci ed efficienti, un vantaggio che riduce lo sforzo e i tempi di percorrenza, abbassando così la barriera all'utilizzo per una fascia più ampia di popolazione. Sul piano metodologico, è inoltre emersa con chiarezza l'importanza cruciale di un'adeguata elaborazione dei dati, evidenziando come l'algoritmo di *map-matching* sia un passo imprescindibile per correggere le imprecisioni del GPS e ottenere una stima affidabile delle distanze garantendo così la robustezza di tutte le analisi successive.

Tuttavia, il risultato più rilevante di questa ricerca risiede nell'analisi comparativa, che ha messo in luce una divergenza nei modelli di utilizzo tra le due città. È emerso un duplice e distinto "carattere" dei servizi. Firenze si delinea come un modello di bike sharing **funzionale-pendolare**, il cui utilizzo è intrinsecamente legato ai ritmi dei giorni feriali, mentre Bologna, al contrario, presenta un modello ibrido **pendolare-ricreativo**, andamento che indica che il servizio non è solo un mezzo di trasporto funzionale, ma anche uno strumento per vivere la città nel tempo libero, per attività turistiche, sociali e ricreative. Questa differenza sostanziale riflette probabilmente le diverse vocazioni economiche, turistiche e culturali delle due città, nonché le diverse abitudini dei loro cittadini.

In virtù di queste scoperte, i risultati dello studio offrono spunti quantitativi e strategici di grande valore per gli operatori del settore e per i policy maker. Per i gestori, emerge una chiara indicazione a investire nelle flotte di e-bike e a utilizzare i dati spazio-temporali per ottimizzare dinamicamente le costose operazioni di ribilanciamento e manutenzione. Per i pianificatori urbani, le mappe di densità diventano uno strumento basato sull'evidenza per orientare gli investimenti in infrastrutture ciclabili sicure e per disegnare politiche di mobilità sempre più integrate. Pur nei limiti di un'analisi circoscritta nel tempo e nella granularità dei dati, questo lavoro si pone come base per eventuali future ricerche, che potranno esplorare la stagionalità, arricchire i modelli con dati esogeni e integrare analisi qualitative per dare voce agli utenti.

In definitiva, questa tesi non si limita ad una semplice visualizzazione di dati, ma fornisce un'interpretazione critica e comparativa, trasformando i dati grezzi in conoscenza strategica. Ha dimostrato come, al di là di un'apparente somiglianza, i servizi di mobilità condivisa si adattino e vengano plasmati dalle unicità del tessuto urbano in cui operano, contribuendo con spunti concreti al dialogo sulla progettazione di città più sostenibili, efficienti e a misura d'uomo.



# Bibliografia

- [1] Comune di Bologna. (2024) Il gemello digitale di bologna.  
<https://www.comune.bologna.it/notizie/gemello-digitale>. Ultimo accesso: 25 novembre 2024.
- [2] F. I. Urbana. Gemello digitale di bologna. Accesso il 25 Novembre 2024. [Online]. Available: <https://www.fondazioneinnovazioneurbana.it/progetto/gemellodigitale>
- [3] B. Bolognese. La bicipolitana bolognese - rete ciclabile metropolitana. Accesso il 25 Novembre 2024. [Online]. Available: <https://www.bicipolitanabolognese.it/bicipolitana>
- [4] E. Fishman, “Bikeshare: A review of recent literature,” *Transport Reviews*, vol. 36, no. 1, pp. 92–113, 2016. [Online]. Available: <https://doi.org/10.1080/01441647.2015.1033036>
- [5] P. DeMaio, “Bike-sharing: History, impacts, models of provision, and future,” *Journal of public transportation*, vol. 12, no. 4, pp. 41–56, 2009.
- [6] P. A. Plazier, G. Weitkamp, and A. E. van den Berg, ““cycling was never so easy!” an analysis of e-bike commuters’ motives, travel behaviour and experiences using gps-tracking and interviews,” *Journal of Transport Geography*, vol. 65, pp. 25–34, 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0966692316307566>
- [7] Parlamento Europeo. (2019) Emissioni di co2 delle auto: i numeri e i dati (infografica). Accessed: 2024-12-25. [Online]. Available: <https://www.europarl.europa.eu/topics/it/article/20190313STO31218/emissioni-di-co2-delle-auto-i-numeri-e-i-dati-infografica>
- [8] Treccani, “Esterionalità,” n.d., accessed: 2024-12-24. [Online]. Available: <https://www.treccani.it/enciclopedia/esternalita/>
- [9] L.-Y. Qiu and L.-Y. He, “Bike sharing and the economy, the environment, and health-related externalities,” *Sustainability*, vol. 10, no. 4, 2018. [Online]. Available: <https://www.mdpi.com/2071-1050/10/4/1145>

- [10] E. Knap, M. B. Ulak, K. T. Geurs, A. Mulders, and S. van der Drift, “A composite x-minute city cycling accessibility metric and its role in assessing spatial and socioeconomic inequalities – a case study in utrecht, the netherlands,” *Journal of Urban Mobility*, vol. 3, p. 100043, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2667091722000310>
- [11] O. S. Mobility. (2018) Il rapporto nazionale sulla sharing mobility: Capitolo dati. Accessed: dicembre 2024. [Online]. Available: <https://osservatoriosharingmobility.it/>
- [12] BikeItalia. (2021) Mobike diventa s.p.a., unica sopravvissuta dell’apocalisse del bike sharing a flusso libero. Accesso: novembre 2024. [Online]. Available: <https://www.bikeitalia.it/2021/05/13/mobike-diventa-spa-unica-sopravvissuta-dellapocalisse-bike-sharing-a-flusso-libero/>
- [13] T. Atlantic. (2018) Bike share oversupply in china: Huge piles of abandoned and broken bicycles. Accesso: novembre 2024. [Online]. Available: <https://www.theatlantic.com/photo/2018/03/bike-share-oversupply-in-china-huge-piles-of-abandoned-and-broken-bicycles/556268/>
- [14] E. Foundation. About sumo - simulation of urban mobility. Accesso il 25 Novembre 2024. [Online]. Available: <https://eclipse.dev/sumo/about/>
- [15] FS Research Centre, “Elementi metodologici per l’analisi della mobilità delle persone attraverso l’uso di big data,” Gruppo Ferrovie dello Stato Italiane, Technical Paper, 2024, consultato il 4 maggio 2025. [Online]. Available: [https://www.fsitaliane.it/content/dam/fsitaliane/Documents/centro-studi/FSRC\\_Technical\\_Papers\\_2024.pdf](https://www.fsitaliane.it/content/dam/fsitaliane/Documents/centro-studi/FSRC_Technical_Papers_2024.pdf)
- [16] Commissione Europea. (2023) Data act. <https://digital-strategy.ec.europa.eu/it/policies/data-act>. Disponibile online presso: <https://digital-strategy.ec.europa.eu/it/policies/data-act>. [Online]. Available: <https://digital-strategy.ec.europa.eu/it/policies/data-act>
- [17] GeeksforGeeks. (2025, May) What is exploratory data analysis? <https://www.geeksforgeeks.org/data-analysis/what-is-exploratory-data-analysis/>. Last updated: 2025-05-10, Accessed: 2025-07-07.
- [18] E. Fishman and T. Brennan, “Oil vulnerability in melbourne,” in *Australasian Transport Research Forum*. Planning and Transport Research Centre, 2010.
- [19] S. Handy, B. Van Wee, and M. Kroesen, “Promoting cycling for transport: research needs and challenges,” *Transport reviews*, vol. 34, no. 1, pp. 4–24, 2014.



- [20] L. S. Davis, "Rolling along the last mile: Bike-sharing programs blossom nationwide," *Planning*, vol. 80, no. 5, pp. 10–16, 2014.
- [21] S. D. Parkes, G. Marsden, S. A. Shaheen, and A. P. Cohen, "Understanding the diffusion of public bikesharing systems: evidence from europe and north america," *Journal of transport geography*, vol. 31, pp. 94–103, 2013.
- [22] S. Shaheen and S. Guzman, "Worldwide bikesharing," *Access Magazine*, vol. 1, no. 39, pp. 22–27, 2011.
- [23] R. Meddin and P. DeMaio, "The bike-sharing world map. retrieved january 26, 2014," 2015.
- [24] A. Fyhri and N. Fearnley, "Effects of e-bikes on bicycle use and mode share," *Transportation Research Part D: Transport and Environment*, vol. 36, pp. 45–52, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361920915000140>
- [25] Z. Ling, C. R. Cherry, H. Yang, and L. R. Jones, "From e-bike to car: A study on factors influencing motorization of e-bike users across china," *Transportation Research Part D: Transport and Environment*, vol. 41, pp. 50–63, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361920915001364>
- [26] S. Haustein and M. Møller, "E-bike safety: Individual-level factors and incident characteristics," *Journal of Transport and Health*, vol. 3, no. 3, pp. 386–394, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2214140516301979>
- [27] D. v. L. Zheyen Chen and D. Ettema, "Dockless bike-sharing systems: what are the implications?" *Transport Reviews*, vol. 40, no. 3, pp. 333–353, 2020. [Online]. Available: <https://doi.org/10.1080/01441647.2019.1710306>
- [28] E. Eren and V. E. Uz, "A review on bike-sharing: The factors affecting bike-sharing demand," *Sustainable Cities and Society*, vol. 54, p. 101882, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2210670719312387>
- [29] M. S. Yingning Xie and R. B. Noland, "Powering bikeshare in new york city: does the usage of e-bikes differ from regular bikes?" *Transportation Planning and Technology*, vol. 47, no. 6, pp. 875–902, 2024. [Online]. Available: <https://doi.org/10.1080/03081060.2024.2341301>
- [30] S. Project. (2024) Sumo - simulation of urban mobility 2024: A comprehensive overview. Accesso il 25 Novembre 2024. [Online]. Available: <https://sumo.dlr.de/pdf/2024/4-3.pdf>

- [31] T. N. Developers. About numpy. Accesso il 25 Novembre 2024. [Online]. Available: <https://numpy.org/about/>
- [32] RideMovi. Ridemovi - il servizio di mobilità condivisa. <https://www.ridemovi.com/it/>. Accesso: novembre 2024.
- [33] Wikipedia contributors. (2024) Parsing – wikipedia. <https://it.wikipedia.org/wiki/Parsing>. Ultimo accesso: 19 luglio 2025. [Online]. Available: <https://it.wikipedia.org/wiki/Parsing>
- [34] N. Gunnoo. (2024, 12) Cos'è il parsing dei dati? definizione, tecniche e sfide. Blog Parseur. Ultimo aggiornamento: 19 dicembre 2024; Ultimo accesso: 19 luglio 2025. [Online]. Available: <https://parseur.com/it/blog/cos-e-il-parsing-dei-dati>
- [35] Treccani Vocabolario. (2024) Docking Station. Ultimo accesso: 14 novembre 2024. [Online]. Available: <https://www.treccani.it/vocabolario/docking-station/>

# Appendice A

## Algoritmo d'analisi

L'algoritmo utilizzato per l'analisi dei dati rappresenta la sequenza logica delle operazioni eseguite durante il processo di trattamento del dataset, dalla fase preliminare di pulizia e preparazione dei dati, fino alla generazione finale degli output analitici e delle visualizzazioni. La struttura è del tipo:

```
1 .
2 -/data
3 -/graph
4 -/report
5 -/images
6 -- /process.py
7 -- /main.py
8 -- /gui.py
9 --- RideTool.spec
10 --- README.md
```

**Listing A.1:** Struttura del tool con cartelle e file principali.

Ogni elemento è fondamentale e svolge una funzione diversa:

- *data*, è la cartella che contiene il dataset da analizzare;
- *graph* e *report*, sono le cartella di destinazione dei grafici e dei report generati;
- *images*, è la cartella contenente le icone usate nella GUI;
- *process.py*, *main.py* e *gui.py*, sono gli algoritmi di formattazione, analisi e gestione dell'interfaccia grafica;
- *RideTool.spec*, è un file di configurazione per l'installazione dell'eseguibile;
- *README.md*, un file di documentazione scritto in Markdown che fornisce una panoramica del progetto e le istruzioni per l'uso.

## Process.py

Il primo script, *process.py*, è stato sviluppato per automatizzare e standardizzare la preparazione dei dati raccolti. La sua funzione principale è quella di elaborare file in formato CSV (Comma-Separated Values), guidando l'utente in modo interattivo attraverso la selezione del file, la normalizzazione del formato, la modifica dell'intestazione e la pulizia dei dati da valori anomali (gli *outlier*).

### Gestione dei file e creazione dell'output

L'interazione iniziale con lo script è focalizzata sulla selezione del file da elaborare. Per migliorare l'usabilità, lo script non si limita a chiedere un percorso, ma guida attivamente l'utente. La funzione `scegli_cartella` assicura che il percorso inserito sia valido, mentre `mostra_file_csv` elenca tutti i file con estensione “.csv” presenti nella directory specificata, permettendo all'utente di scegliere comodamente da una lista numerata tramite la funzione `scegli_file`.

```
1 def scegli_cartella():
2     # ... implementazione ...
3
4 def mostra_file_csv(cartella):
5     # ... implementazione ...
6
7 def scegli_file(files):
8     # ... implementazione ...
```

**Listing A.2:** Funzioni per la selezione interattiva del file CSV.

Per preservare l'integrità dei dati originali, tutti i file modificati vengono salvati in una sotto-cartella `output`, creata automaticamente dalla funzione `crea_cartella_output` all'interno della directory di lavoro. Questo approccio garantisce che i dati sorgente rimangano intatti, permettendo di ripetere l'elaborazione con parametri diversi se necessario.

## Normalizzazione del separatore e modifica dell'header

I file **CSV** possono utilizzare diversi caratteri come separatore di campo (es. virgola o punto e virgola). Lo script chiede all'utente di specificare il delimitatore del file originale e quello desiderato per il file di output, gestendo la conversione in modo trasparente.

Una delle funzioni di maggior valore è la gestione flessibile dell'intestazione (o *header*). La funzione `scegli_e_applica_header` offre un menù con opzioni predefinite per diversi set di dati (*e.g.*: “dati\_bologna”, “dati\_firenze”, “dati\_mapmatched”), oltre alla possibilità di inserire manualmente un header personalizzato. Questa funzionalità è cruciale per standardizzare le colonne prima delle analisi successive.

```
1 def scegli_e_applica_header(file_path, delimiter):
2     # ... chiede all'utente se modificare l'header ...
3
4     headers_predefiniti = {
5         '1': "start,end,lat_start,...", # header_bologna
6         '2': "order_id;country;...",   # header_firenze
7         # ... altre opzioni ...
8     }
9
10    # ... logica per la scelta e applicazione ...
11    if nuovo_header:
12        cambia_header(file_path, nuovo_header, delimiter)
13
14    return conversione_automatica
```

**Listing A.3:** Funzione per la gestione e modifica dell'header.

La modifica avviene in modo sicuro: la funzione `cambia_header` scrive le modifiche su un file temporaneo e solo al termine dell'operazione, se andata a buon fine, lo sostituisce al file di lavoro utilizzando `shutil.move`, minimizzando il rischio di corruzione dei dati.

## Elaborazione e filtraggio della distanza

La fase di pulizia dei dati prevede l'impiego di `processa_e_filtra_distanza`, funzione incaricata sia dell'eliminazione degli *outlier* sia della conversione delle unità di misura, al fine di assicurare uniformità e correttezza nei dati analizzati. Lo script è altresì in grado di modificare le distanze espresse in metri che necessitano di essere convertite in chilometri, grazie alla scelta dinamica dell'utente.

Il filtraggio viene eseguito sulla colonna `distance_km`, conservando solo i valori compresi nell'intervallo plausibile `[0, 15.0]`. La funzione utilizza `csv.DictReader`, che rende il codice più leggibile e robusto e gestisce con attenzione i valori mancanti o non numerici.

```
1 def processa_e_filtra_distanza(input_path, output_path, delimiter,
2   max_distance_km=15.0, conversione_automatica=False):
3   # ... gestione della scelta di conversione (automatica o manuale)
4   ...
5   with open(...) as f_in, open(...) as f_out:
6       reader = csv.DictReader(f_in, delimiter=delimiter)
7       writer = csv.DictWriter(f_out, fieldnames=reader.fieldnames,
8   delimiter=delimiter)
9       # ...
10      for row in reader:
11          try:
12              dist_val = row.get('distance_km')
13              # ... logica per gestire valori nulli ...
14
15              dist = float(dist_val.replace(',', ' '))
16
17              if convertire:
18                  dist = dist / 1000.0
19
20              if 0 <= dist <= max_distance_km:
21                  row['distance_km'] = f"{dist:.6f}"
22                  writer.writerow(row)
23              # ...
24          except (ValueError, TypeError):
25              # ... gestisce righe con dati non validi ...
```

**Listing A.4:** Funzione per convertire e filtrare i dati di distanza.

La gestione delle eccezioni e dei valori nulli o mal formattati assicura che il processo di pulizia sia robusto e non si interrompa inaspettatamente.

## Flusso di esecuzione principale

Il blocco `if __name__ == "__main__":` gestisce l'intero processo in una sequenza logica:

1. **Selezione del file:** l'utente sceglie la cartella e il file CSV da elaborare tramite il menù interattivo.
2. **Creazione ambiente di lavoro:** viene creata la cartella di output e una copia temporanea del file di input (`temp_processing_*`) su cui verranno eseguite tutte le operazioni.
3. **Normalizzazione del separatore:** l'utente definisce il separatore originale e quello desiderato; la modifica viene applicata direttamente al file di lavoro temporaneo.
4. **Modifica dell'header:** viene invocata la funzione per la gestione dell'header, che funzione restituisce anche un flag che indica se è necessaria una conversione automatica della distanza.
5. **Filtraggio e conversione:** la funzione `processa_e_filtra_distanza` viene eseguita sul file di lavoro, utilizzando il *flag*<sup>1</sup> per la conversione ed il risultato è salvato in un file finale con prefisso `preprocessed_`.
6. **Pulizia:** il file di lavoro temporaneo viene rimosso per mantenere pulito lo spazio di lavoro.

Questo flusso strutturato assicura che ogni operazione sia eseguita nell'ordine corretto su file temporanei, proteggendo i dati originali e garantendo un processo di pre-elaborazione sicuro ed efficiente.

---

<sup>1</sup>*Flag*: una variabile che viene usata per indicare uno stato, una condizione o un evento particolare durante l'esecuzione di un programma.

## Main.py

Lo script `main.py` è il motore computazionale dell'intero progetto. Contiene la logica per il caricamento, la preparazione e l'esecuzione di tutte le analisi sui dati del servizio di bike sharing. Questo script è stato progettato per funzionare in due modalità: come “backend” richiamato dall'interfaccia grafica (`GUI.py`) e come strumento autonomo eseguibile da riga di comando.

Per svolgere le sue funzioni, si avvale di librerie fondamentali dello stack scientifico di Python, tra cui:

- NumPy e SciPy per il calcolo numerico e statistico.
- Matplotlib e Seaborn per la creazione di visualizzazioni statiche.
- Pandas per la manipolazione avanzata e l'aggregazione dei dati.
- GeoPandas e Contextily per l'analisi e la mappatura geospaziale.

Prima di ogni analisi, i dati devono essere caricati e preparati correttamente. A questo scopo, lo script utilizza delle funzioni di supporto.

La funzione `load_csv` si occupa di leggere il file CSV e caricarlo in un array strutturato di NumPy. Questo approccio è particolarmente efficiente per dati tabellari di grandi dimensioni, come il dataset del caso studio.

```
1 def load_csv(file_path, delimiter=','):
2     # caricamento dati con tipi misti, parsing solo nominale
3     data = np.genfromtxt(file_path, delimiter=delimiter, names=True,
4                           dtype=None, encoding='utf-8-sig')
5     data.dtype.names = tuple(name.strip() for name in data.dtype.names)
6     print("[OK] File caricato. Colonne:", data.dtype.names)
7     return data
```

**Listing A.5:** Funzione per caricare dati CSV con NumPy.

Poiché le analisi temporali sono cruciali, la funzione `parse_datetime_array` è stata creata per convertire le colonne di testo contenenti date e orari in oggetti `datetime` nativi di Python, gestendo in modo robusto eventuali errori di formato o valori mancanti, essenziale per la solidità del codice.



## Analisi generale (general\_analysis)

La “Analisi generale” ha lo scopo di sintetizzare le caratteristiche principali del un dataset, individuare pattern, anomalie e ottenere una prima comprensione profonda dei dati, principalmente attraverso strumenti statistici e visuali.

Attraverso l’analisi della variabile `distance_km`, si cerca di ottenere un quadro più chiaro di quali siano, in media, le distanze percorse nei tragitti considerati.

Prima di ogni calcolo, viene applicato un filtro per escludere i viaggi con una distanza inferiore a 150 metri, operazione di pulizia necessaria per rimuovere dati “sporchi”, come viaggi avviati e interrotti immediatamente o errori di localizzazione GPS, che non rappresentano spostamenti reali e potrebbero inquinare le statistiche.

```
1 distances = data['distance_km']
2
3 # filtro per rimuovere valori non significativi
4 distances = distances[(distances > 0.150)]
```

**Listing A.6:** Filtro per rimuovere outlier da array di distanze.

Successivamente, vengono calcolati i principali indici statistici descrittivi utilizzando le librerie NumPy e SciPy. Questi indici forniscono una sintesi numerica della distribuzione:

- Misure di *tendenza centrale*: media, mediana;
- Misure di *dispersione*: deviazione standard, varianza, IQR;
- Misure di *forma*: asimmetria (*skewness*, curtosi)

I risultati vengono sia stampati a video sia salvati in un file di report per una consultazione futura.

```
1 print("\n Statistiche descrittive:")
2 print(f"Media distanza: {np.mean(distances):.2f} km")
3 print(f"Mediana distanza: {np.median(distances):.2f} km")
4 print(f"Deviazione standard: {np.std(distances):.2f} km")
5 # ...
6 print(f"Simmetria: {stats.skew(distances):.2f}")
7 print(f"Curtosi: {stats.kurtosis(distances):.2f}")
```

**Listing A.7:** Stampa statistiche descrittive sulle distanze.

La seconda parte dell'analisi è quella invece che si dedica alla produzione delle visualizzazioni, raccolte in un'unica figura composita.

```
1 # istogramma con percentili, media e mediana
2 plt.subplot(2, 2, 1)
3 plt.hist(distances, bins=30, color='skyblue', edgecolor='black',
4 alpha=0.7)
5 for p, val in zip(percentiles, perc_values):
6     plt.axvline(val, color='red', linestyle='--', alpha=0.7)
7     plt.text(val, plt.ylim()[1]*0.9, f'{p}%', rotation=90, color='
8 red', fontsize=8)
9     plt.axvline(np.mean(distances), color='blue', linestyle='-',
10 linewidth=1.5, label='Media')
11     plt.axvline(np.median(distances), color='green', linestyle='-',
12 linewidth=1.5, label='Mediana')
13 plt.legend()
14 (...)
15
16 # QQ-plot
17 plt.subplot(2, 2, 2)
18 shape, loc, scale = stats.lognorm.fit(distances, floc=0)
19 stats.probplot(distances, dist=stats.lognorm(shape, loc=loc, scale=
20 scale), plot=plt)
21 plt.title('QQ-plot')
22 (...)
23
24 # distribuzione cumulativa
25 plt.subplot(2, 2, 3)
26 sorted_distances = np.sort(distances)
27 cum_freq = np.arange(1, len(distances) + 1) / len(distances)
28 plt.plot(sorted_distances, cum_freq, marker='.')
29 for p, val in zip(percentiles, perc_values):
30     plt.axvline(val, color='green', linestyle='--', alpha=0.7)
31     plt.text(val, p/100, f'{p}%', color='green', fontsize=8)
32 plt.title('Distribuzione cumulativa')
33 (...)
34
35 # distribuzione normale teorica (Gaussiana)
36 plt.subplot(2, 2, 4)
37 mu = np.mean(distances)
38 sigma = np.std(distances)
39 x_min = max(0, mu - 4*sigma)
40 x_max = mu + 4*sigma
41 x = np.linspace(x_min, x_max, 100)
42 y = norm.pdf(x, mu, sigma)
43 plt.plot(x, y, color='purple')
44 (...)
45
```

**Listing A.8:** Quattro visualizzazioni grafiche per analisi delle distanze.

## Analisi spaziale (spatial\_analysis)

La “Analisi spaziale” è fondamentale per comprendere dove il servizio di bike sharing viene utilizzato. L’obiettivo di questa analisi è duplice: mappare la distribuzione geografica dei punti di partenza e di arrivo e identificare le “zone calde” (gli *hotspot*) della città. Per raggiungere questo scopo, sono state impiegate le librerie Python **Pandas**, per la manipolazione dei dati, **GeoPandas**, per la gestione di dati georeferenziati, e **Contextily**, per l’aggiunta di mappe di base ai grafici. Il primo passo consiste nel convertire l’array NumPy in un **DataFrame** Pandas, una struttura dati più flessibile per le operazioni di pulizia. Successivamente, per garantire l’affidabilità dell’analisi, vengono rimossi i punti geografici anomali: il filtraggio si basa su *quantili*, un approccio robusto che definisce dinamicamente un’area rettangolare (bounding box) che contiene il 99% dei dati, escludendo lo 0.5% più estremo per ogni coordinata.

```
1 # pulizia e preparazione dati
2 q_low, q_high = 0.005, 0.995
3 lat_min, lat_max = data['lat_start'].quantile([q_low, q_high])
4 lon_min, lon_max = data['long_start'].quantile([q_low, q_high])
5
6 df_filtered = data[
7     (data['lat_start'] > lat_min) & (data['lat_start'] < lat_max) &
8     # ... (condizioni analoghe per long_start, lat_end, long_end)
9 ].copy()
```

**Listing A.9:** Filtraggio dati rimuovendo outlier geografici tramite quantili.

Una volta puliti, i dati vengono trasformati in un **GeoDataFrame**, l’oggetto cardine di **GeoPandas**. A ogni riga di dati viene associata una geometria, in questo caso un punto definito dalle sue coordinate. Durante la creazione, viene specificato il *Sistema di Riferimento delle Coordinate* (CRS) di origine come EPSG:4326, lo standard globale per latitudine e longitudine (WGS 84). Tuttavia, per poter sovrapporre i punti a mappe web (come quelle di OpenStreetMap), è necessario proiettare i dati in un CRS compatibile, ovvero EPSG:3857 (Web Mercator). Questo passaggio è cruciale e viene eseguito con il metodo `.to_crs()`.

```
1 # conversione in GeoDataFrame
2 gdf_start = gpd.GeoDataFrame(
3     df_filtered,
4     geometry=gpd.points_from_xy(df_filtered['long_start'], df_filtered[
5         'lat_start']),
6     crs="EPSG:4326" # Imposta il CRS di origine
7 ).to_crs(epsg=3857) # Riproyetta nel CRS per mappe web
```

**Listing A.10:** Conversione DataFrame in GeoDataFrame con proiezione web mercator.

L'analisi produce due figure distinte per una comprensione completa.

1. La prima figura mostra la posizione esatta di ogni singolo viaggio, con partenze e arrivi su due mappe affiancate per un confronto diretto. Questo grafico offre una visione d'insieme della copertura del servizio. La mappa di base, fornita da `contextily`, aggiunge il contesto urbano (strade, parchi) indispensabile per interpretare la distribuzione.

```
1 # grafico 1: distribuzione punti
2 axes1[0].scatter(x_start, y_start, s=5, c='dodgerblue', alpha
3 =0.2)
4 axes1[0].set_title("Distribuzione partenze", fontsize=14)
5 ctx.add_basemap(axes1[0], source=ctx.providers.OpenStreetMap.
  Mapnik)
```

**Listing A.11:** Grafico distribuzione punti di partenza con basemap OpenStreetMap.

2. Per dataset con decine di migliaia di punti, un semplice grafico a dispersione diventa un'unica macchia di colore illeggibile. La seconda figura utilizza, allora, una visualizzazione `hexbin`, che raggruppa i punti in celle esagonali e colora ogni cella in base al numero di punti che contiene, creando una mappa di densità (heatmap). Per evidenziare le differenze sia nelle aree a bassa che ad alta densità, la scala di colori della mappa è *logaritmica*.

```
1 # grafico 2: heatmap
2 im1 = axes2[0].hexbin(x=x_start, y=y_start, gridsize=150, cmap
3 ='viridis',
4                      norm=LogNorm(), mincnt=1) # <- Uso della
5 scala logaritmica
6 axes2[0].set_title("Densita partenze", fontsize=14)
  ctx.add_basemap(axes2[0], source=ctx.providers.CartoDB.
    Positron)
```

**Listing A.12:** Heatmap densità punti di partenza con scala logaritmica e basemap.

I grafici generati vengono infine salvati come immagini ad alta risoluzione nella cartella `graph`, pronti per essere inclusi in report e presentazioni.

## Analisi applicativa (application\_analysis)

La “Analisi applicativa” ha il compito di capire come viene utilizzato il bike sharing, segmentando i viaggi in tre fasce di distanza (*brevi*, *medi* e *lunghi*) come indicatore del loro possibile scopo. La suddivisione avviene in modo efficiente tramite maschere booleane su array NumPy, evitando cicli espliciti e accelerando l’elaborazione.

```
1 distances = data['distance_km'].astype(float)
2
3 # Selezione vettorizzata dei segmenti
4 short = data[distances < 2.0]
5 medium = data[(distances >= 2.0) & (distances <= 5.0)]
6 long = data[distances > 5.0]
7
8 print(f"Tragitti brevi (<2 km): {len(short)}")
9 print(f"Tragitti medi (2/5 km): {len(medium)}")
10 print(f"Tragitti lunghi (>5 km): {len(long)}")
```

**Listing A.13:** Classificazione vettoriale tragitti per distanza.

Oltre alla segmentazione per distanza, la funzione esegue un’analisi secondaria, calcolando il numero totale di viaggi effettuati con ciascuna tipologia di bicicletta disponibile (es. “Muscolari” e “E-bike”). Questo fornisce un’indicazione immediata sulla popolarità relativa dei diversi tipi di veicoli offerti dal servizio.

```
1 bike_types = np.unique(data['bike_type'])
2 print()
3 for btype in bike_types:
4     count_bike = np.sum(data['bike_type'] == btype)
5     print(f"Tipo bici: {btype}")
6     print(f"Totale tragitti: {count_bike}")
7     print()
```

**Listing A.14:** Conteggio tragitti per tipo bici.

L’analisi si conclude con l’offerta di risultati chiari e incisivi, trascritti in riassunto testuale con il conteggio dei viaggi per segmento e tipo di veicolo, salvato in un report con data e ora per garantirne tracciabilità e riproducibilità.

```
1 now = datetime.now().strftime("%Y-%m-%d_%H-%M-%S")
2 report_path = os.path.join("report", f"report_applicativo_{now}.txt")
3
4 with open(report_path, 'w') as f:
5     f.write(f"Distanza media del tragitto: {average_trip:.2f} km\n\n")
6     f.write(f"Tragitti brevi (<2 km): {len(short)}\n")
7     (...)
```

**Listing A.15:** Creazione e salvataggio del report con statistiche sui tragitti.

## Analisi delle velocità (time\_speed\_analysis)

La “Analisi tempi/velocità” ha lo scopo di confrontare le caratteristiche di utilizzo delle biciclette tradizionali, le *muscolari* e quelle a pedalata assistita, le *e-bike*. L’obiettivo è quantificare le differenze nel comportamento degli utenti, in particolare riguardo ai tempi di percorrenza e alle velocità medie.

Per garantire un confronto equo e basato su dati realistici, è stato necessario un processo di elaborazione e filtraggio rigoroso. La logica è stata incapsulata nella funzione di supporto interna `_process_bike_data` per essere applicata in modo identico a entrambi i sottogruppi di dati.

I passaggi chiave di questa elaborazione sono:

1. Calcolo della *durata*: il tempo di percorrenza viene calcolato in ore.
2. Calcolo della *velocità*: la velocità media viene calcolata come rapporto tra la distanza e la durata ( $v = \frac{d}{t}$ ).
3. Filtraggio robusto: vengono applicati dei filtri severi per escludere i viaggi non plausibili, che potrebbero essere frutto di errori GPS o di utilizzi anomali, mentre vengono preservati i viaggi che soddisfano tutte le seguenti condizioni:
  - *Distanza* percorsa positiva.
  - *Durata* del viaggio compresa tra 0 e 2 ore.
  - *Velocità media* compresa tra 5 e 30 km/h.

Questi filtri sono fondamentali per escludere sia i viaggi troppo lenti (assimilabili a una camminata con la bici) sia quelli irrealisticamente veloci, garantendo che l’analisi si concentri su spostamenti effettivi e rappresentativi.

```
1 duration = (end - start).total_seconds() / 3600 # in ore
2 if 0 < duration < 2: # filtro durata
3     speed = dist / duration
4     if 5 < speed < 30:
5         durations.append(duration)
6         speeds.append(speed)
```

**Listing A.16:** Calcolo della durata e della velocità dei tragitti filtrati.

I dati filtrati vengono quindi utilizzati per creare i grafici comparativi, organizzati in una griglia 2x2 per una visione d'insieme. In ogni grafico, i dati delle e-bike e delle bici muscolari sono sovrapposti usando colori diversi, permettendo un confronto visivo immediato.

```
1 # grafico 2: distribuzione velocita
2 ax2 = axes[0, 1]
3 ax2.hist(ele_speeds, bins=30, color='#FFA07A', alpha=0.7, label='E-bike',
4          zorder=1)
5 ax2.hist(musc_speeds, bins=30, color='#87CEEB', alpha=0.7, label='Muscolari',
6          zorder=2)
7 ax2.set_title('Distribuzione velocita media')
8 (...)
```

**Listing A.17:** Istogramma della distribuzione delle velocità medie di e-bike e biciclette muscolari.

A complemento dell'analisi visiva, vengono calcolati gli indici statistici (media, mediana, deviazione standard, min/max) per tempi e velocità, separatamente per ogni tipo di bici.

Questi dati quantitativi vengono formattati e salvati in un report testuale; è di fondamentale importanza che il report specifichi chiaramente i filtri applicati durante l'elaborazione, poiché questi definiscono il contesto dei risultati e ne garantiscono la corretta interpretazione e la trasparenza scientifica.

```
1 def _get_stats_text(name, durations, speeds):
2     # ...
3     text = f"--- STATISTICHE {name.upper()} ---\n"
4     text += f"Viaggi analizzati: {len(durations)}\n"
5     text += f"VELOCITA (km/h):\n"
6     text += f"  - Media: {np.mean(speeds):.2f} +/- {np.std(speeds):.2f}\n"
7     text += f"  - Mediana: {np.median(speeds):.2f}\n"
8     # ...
9     return text
```

**Listing A.18:** Funzione per generare un riepilogo testuale delle statistiche di durata e velocità.

## Analisi temporale (temporal\_analysis)

La “Analisi temporale” si occupa di indagare quando il servizio di bike sharing viene utilizzato; lo scopo è identificare pattern ricorrenti su base oraria e settimanale, per capire le abitudini degli utenti e le dinamiche di utilizzo della flotta. Anche in questo caso, viene mantenuto un approccio comparativo, confrontando l’uso delle biciclette tradizionali con quello delle e-bike per evidenziare eventuali differenze nei pattern.

L’analisi temporale inizia con la conversione delle date in oggetti `datetime`, scartando i record non validi. Da questi si estraggono poi due variabili quantitative, l’ora del giorno e il giorno della settimana, per entrambi i dataset, generando array numerici pronti per la visualizzazione.

```
1 # ...dopo aver ottenuto 'start_times_musc' e 'start_times_ele' come
   array di datetime
2
3 # estrazione delle feature temporali
4 hours_musc = np.array([dt.hour for dt in start_times_musc])
5 weekdays_musc = np.array([dt.weekday() for dt in start_times_musc])
6
7 hours_ele = np.array([dt.hour for dt in start_times_ele])
8 weekdays_ele = np.array([dt.weekday() for dt in start_times_ele])
```

**Listing A.19:** Estrazione delle feature temporali dagli orari di inizio tragitti.

L’analisi produce un’unica figura contenente i due grafici affiancati, dedicati rispettivamente alla “distribuzione oraria dei viaggi” e alla “distribuzione settimanale dei viaggi”.

```
1 # grafico 1: distribuzione oraria
2 ax1.hist(hours_ele, bins=24, range=(0, 24), label='E-bike', ..., zorder
   =1)
3 ax1.hist(hours_musc, bins=24, range=(0, 24), label='Muscolari', ...,
   zorder=2)
4 ax1.set_title("Distribuzione oraria dei tragitti")
5 (...)
6
7 # grafico 2: distribuzione settimanale
8 x = np.arange(len(day_labels))
9 width = 0.35
10 ax2.bar(x - width/2, counts_musc, width, label='Muscolari', ...)
11 ax2.bar(x + width/2, counts_ele, width, label='E-bike', ...)
12 ax2.set_title("Distribuzione settimanale dei tragitti")
13 # ...
```

**Listing A.20:** Grafici della distribuzione oraria e settimanale dei tragitti.



## Analisi dei turnover (turnover\_analysis)

L’“Analisi dei turnover” si concentra esclusivamente sulla distribuzione del carico di lavoro all’interno dell’intera flotta di veicoli. L’obiettivo è comprendere come il numero totale di viaggi effettuati sia distribuito tra le biciclette, senza entrare nel dettaglio dei veicoli più performanti. Questa analisi si affida interamente alla libreria **Pandas**.

Il primo passo consiste in una preparazione dati mirata: conversione delle date, normalizzazione delle distanze e rimozione dei dati mancanti. Successivamente viene effettuata l’anonimizzazione degli ID dei veicoli, fondamentale per la chiarezza dei grafici e per la privacy: a ogni **bikeID** originale viene assegnato un nuovo ID ordinato (es. “Bici 1”, “Bici 2”, ...), dove “Bici 1” corrisponde alla bicicletta con il maggior numero di viaggi.

```
1 # anonimizzazione degli ID per una visualizzazione ordinata e chiara
2 classifica_viaggi = df.groupby('bikeID').size().sort_values(ascending=
    False)
3 mappa_anonima = {bike_id: f'Bici {i+1}' for i, bike_id in enumerate(
    classifica_viaggi.index)}
4 df['bikeID_anonimo'] = df['bikeID'].map(mappa_anonima)
```

**Listing A.21:** Anonimizzazione degli ID dei mezzi.

L’analisi si concentra sull’intera flotta, producendo due visualizzazioni principali: l’“istogramma dei viaggi per bicicletta” e la “funzione di ripartizione cumulativa” (CDF). Il primo mostra la distribuzione del numero di viaggi effettuati da ciascuna bicicletta, con linee che evidenziano media e mediana, mentre la CDF permette di osservare la percentuale cumulativa di biciclette in funzione dei viaggi effettuati.

```

1 # analisi intera flotta
2 print("-> Analisi della distribuzione del lavoro sull'intera flotta.")
3
4 viaggi_per_bici_anonimi = df.groupby('bikeID_anonimo').size()
5
6 fig, (ax3, ax4) = plt.subplots(1, 2, figsize=(18, 7))
7 fig.suptitle('Distribuzione del carico di lavoro sull\'intera flotta',
8             fontsize=18, weight='bold')
9
10 # grafico 3
11 sns.histplot(viaggi_per_bici_anonimi, ax=ax3, bins=50, kde=True, color=
12             'skyblue')
13 ax3.set_title('Istogramma dei viaggi per bicicletta', fontsize=14)
14 ax3.set_xlabel('Numero totale di viaggi effettuati')
15 ax3.set_ylabel('Numero di biciclette')
16 media_viaggi = viaggi_per_bici_anonimi.mean()
17 mediana_viaggi = viaggi_per_bici_anonimi.median()
18 ax3.axvline(media_viaggi, color='red', linestyle='--', label=f'Media: {
19             media_viaggi:.1f}')
20 ax3.axvline(mediana_viaggi, color='green', linestyle='--', label=f'
21             Mediana: {mediana_viaggi:.1f}')
22 ax3.legend()
23
24 # grafico 4
25 sns.ecdfplot(viaggi_per_bici_anonimi, ax=ax4, color='coral')
26 ax4.set_title('Distribuzione cumulativa (CDF)', fontsize=14)
27 ax4.set_xlabel('Numero totale di viaggi effettuati')
28 ax4.set_ylabel('Percentuale della flotta')
29 ax4.grid(True, linestyle='--')
30
31 plt.tight_layout(rect=[0, 0, 1, 0.95])
32
33 if not os.path.exists("report"):
34     os.makedirs("report")
35 if not os.path.exists("graph"):
36     os.makedirs("graph")
37
38 img_path_flotta = os.path.join("graph", f"
39     grafico_distribuzione_flotta_anonimo_{now}.png")
40 plt.savefig(img_path_flotta)
41 plt.show()

```

**Listing A.22:** Analisi della distribuzione del carico di lavoro sull'intera flotta.

L'ultima parte dell'analisi genera un report testuale, salvato all'interno della cartella **report** con un nome file contenente la data corrente, che include:

- Una sezione con statistiche sintetiche sull'intera flotta: numero di biciclette analizzate, media, mediana, massimo e minimo dei viaggi per bicicletta.
- L'indicazione del percorso del grafico generato relativo all'intera flotta, seguita dalla stampa di messaggi di conferma nel terminale.

```
1 report_path = os.path.join("report", f"report_turnover_anonimo_{now}.  
  txt")  
2 with open(report_path, 'w', encoding='utf-8') as f:  
3     f.write("=== REPORT ANALISI TURNOVER (intera flotta) ===\n")  
4     f.write(f"Report generato il: {now}\n")  
5     f.write(f"Grafico distribuzione flotta salvato in: {img_path_flotta}\n\n")  
6     f.write("--- Statistiche ---\n")  
7     f.write(f"Numero biciclette analizzate: {viaggi_per_bici_anonimi.  
  shape[0]}\n")  
8     f.write(f"Media viaggi per bici: {media_viaggi:.2f}\n")  
9     f.write(f"Mediana viaggi per bici: {mediana_viaggi:.2f}\n")  
10    f.write(f"Max viaggi su una bici: {viaggi_per_bici_anonimi.max()}\n")  
11    f.write(f"Min viaggi su una bici: {viaggi_per_bici_anonimi.min()}\n")  
12  
13 print("\n[OK] Analisi completata.")  
14 print(f"    -> Report testuale salvato in: {report_path}")  
15 print(f"    -> Grafico Flotta salvato in: {img_path_flotta}")
```

**Listing A.23:** Salvataggio del report testuale con i risultati dell'analisi turnover.

## Loop di esecuzione principale

Lo script può essere eseguito anche senza l'interfaccia grafica. Il blocco `if __name__ == "__main__":` gestisce questa modalità, trasformando lo script in uno strumento a riga di comando.

Il flusso è il seguente:

1. Viene richiesto all'utente di specificare la cartella contenente i file CSV.
2. Viene caricato il file scelto.
3. Viene avviato un ciclo `while` che presenta un menù testuale all'utente.

Questo menù è gestito da un dizionario che associa un numero a ciascuna funzione di analisi. Questa struttura è pulita, scalabile e permette all'utente di eseguire più analisi in sequenza sullo stesso dataset senza dover ricaricare i dati.

```
1 if __name__ == "__main__":
2     # ... (caricamento file) ...
3
4     analisi = {
5         1: ("Analisi generale", general_analysis),
6         2: ("Analisi spaziale", spatial_analysis),
7         3: ("Analisi tempi e velocità", time_speed_analysis),
8         4: ("Analisi applicativa", application_analysis),
9         5: ("Analisi temporale", temporal_analysis),
10        6: ("Analisi turnover veicoli", turnover_analysis)
11    }
12
13    while True:
14        scelta = main_menu(analisi)
15        if scelta == 0:
16            break
17        elif scelta in analisi:
18            nome_analisi, funzione_analisi = analisi[sceita]
19            funzione_analisi(data)
20        # ... (gestione errori)
```

**Listing A.24:** Loop principale con menu di selezione.

## GUI.py

Per rendere l'utilizzo degli script di analisi accessibile e intuitivo, è stata sviluppata infine un'interfaccia utente grafica (GUI) tramite il file 'gui.py'. Questa applicazione agisce come un front-end che permette all'utente di selezionare un file di dati e avviare le diverse analisi senza la necessità di interagire direttamente con il codice o la riga di comando.

### Architettura e tecnologie

L'interfaccia è stata realizzata in Python utilizzando la libreria `customtkinter`, una versione moderna e stilisticamente personalizzabile del toolkit standard `tkinter`. La scelta è ricaduta su questa libreria per la sua capacità di creare interfacce dall'aspetto professionale e per la sua semplicità d'uso.

L'intera applicazione è incapsulata all'interno della classe `RideMoviGUI`, che eredita da `ctk.CTk`. Questo approccio a oggetti permette di organizzare in modo ordinato tutti i componenti dell'interfaccia (i *widget*), le relative proprietà (colori, font) e le funzioni che ne gestiscono il comportamento.

```
1 import customtkinter as ctk
2 # ...
3 from main import (
4     load_csv,
5     general_analysis,
6     # ... altre funzioni di analisi
7 )
8
9 class RideMoviGUI(ctk.CTk):
10     def __init__(self):
11         super().__init__()
12         # ... impostazioni della finestra e dei widget
```

**Listing A.25:** Sezione degli import e definizione iniziale della classe GUI RideMovi.

Un aspetto cruciale dell'architettura è il disaccoppiamento tra l'interfaccia e la logica di analisi. Come si evince dalle prime righe, `gui.py` non contiene gli algoritmi di calcolo, ma li importa dal modulo di analisi `main.py`. La GUI ha quindi il solo compito di raccogliere l'input dell'utente e invocare la funzione corretta, rendendo il codice più modulare e robusto.

## Progettazione dell'interfaccia e gestione degli eventi

Il design dell'interfaccia è stato curato per garantire una buona esperienza utente, la cosiddetta *User Experience* (UX). È stato definito un tema cromatico coerente, sono stati scelti font leggibili e ogni pulsante di analisi è stato associato a un'icona rappresentativa. La disposizione degli elementi segue una griglia logica che guida l'utente dalla selezione del file all'esecuzione delle analisi.

Per evitare la duplicazione del codice, la creazione dei pulsanti di analisi è stata affidata a una funzione di supporto, `add_analysis.button`.

L'interazione centrale dell'applicazione è gestita dal metodo `run_analysis`, che collega i pulsanti della GUI alle funzioni di analisi del backend.

```
1 def run_analysis(self, analysis_func, analysis_name):
2     file_path = self.file_entry.get()
3     if not file_path or not os.path.exists(file_path):
4         messagebox.showerror("Errore", "Seleziona un file CSV valido...")
5         return
6
7     try:
8         data = load_csv(file_path)
9         analysis_func(data) # Esegue la funzione di analisi specifica
10        messagebox.showinfo("Successo", f"{analysis_name} completata con successo!")
11    except Exception as e:
12        messagebox.showerror("Errore durante l'analisi", f"Si è verificato un errore:\n{e}")
```

**Listing A.26:** Gestione e controllo del comportamento dei pulsanti.

Come mostrato nel frammento di codice, questo metodo:

1. Recupera il percorso del file inserito dall'utente e ne valida l'esistenza.
2. Utilizza un blocco `try-except` per gestire eventuali errori che potrebbero verificarsi durante il caricamento dei dati o l'esecuzione dell'analisi, notificando l'utente tramite una finestra di dialogo (`messagebox`).
3. Invoca la specifica funzione di analisi (`analysis_func`) passata come argomento.
4. Comunica all'utente il successo dell'operazione.

## Gestione delle risorse per la distribuzione

In previsione della distribuzione dell'applicazione come file eseguibile (`.exe`), è stata implementata una funzione di supporto essenziale, `resource_path`. Quando un'applicazione Python viene impacchettata, ad esempio con `PyInstaller`, i file ausiliari come le icone vengono inseriti in una cartella temporanea il cui percorso è accessibile tramite `sys._MEIPASS`. La funzione `resource_path` gestisce questa logica, garantendo che le icone vengano caricate correttamente sia durante lo sviluppo (quando lo script viene eseguito direttamente) sia in produzione (quando viene eseguito come file compilato).

```
1 def resource_path(relative_path):
2     try:
3         # PyInstaller crea una cartella temp e salva il percorso in
4         # _MEIPASS
5         base_path = sys._MEIPASS
6     except Exception:
7         base_path = os.path.abspath(".")
8
9     return os.path.join(base_path, relative_path)
```

**Listing A.27:** Gestione del percorso delle risorse nell'eseguibile creato con `PyInstaller`.

Questa accortezza tecnica è fondamentale per creare un'applicazione stand-alone e facilmente distribuibile, che non dipenda dalla struttura delle cartelle presenti sulla macchina di sviluppo.