



UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE E TECNOLOGIE

Corso di Laurea Triennale in Informatica

**CONFRONTO TRA AUTOMOBILE E
MEZZI DI TRASPORTO ALTERNATIVI
NEL COMUNE DI MILANO BASATO SU
STIME DI TEMPI DI PERCORRENZA**

Tesi di Laurea di
MAURO MASTRAPASQUA
Matricola 892629

Relatore
Andrea Trentini

Correlatore
Dario Malchiodi

Anno Accademico 2019/2020

Indice

1	Introduzione	3
1.1	Automobile in città	3
1.2	Problemi	4
1.3	Vantaggi	5
1.4	Confronto con mezzi alternativi	6
2	Metodi	7
2.1	Lavori precedenti	8
2.2	Primo prototipo del progetto	9
2.2.1	Assenza di Open Data	10
2.3	L'alternativa: simulazione	12
2.3.1	Crowdsourced data: il caso Waze	12
3	Progettazione	13
3.1	Ricerca delle API	14
3.1.1	Google	14
3.1.2	Waze	14
3.1.3	Moovit	15
3.1.4	Here	15
3.1.5	Enjoy	15
3.1.6	OpenStreetMap	15
3.2	Scelta delle API	16
3.2.1	Lo scraping di Waze	17
3.2.2	Lo scraping di Moovit (scartato)	18
3.2.3	Estratti del codice sorgente degli scraper	19
3.2.4	Stima del percorso per il car sharing	21
3.3	Scelta tra percorsi random e prefissati	21
3.4	Generatore random dei percorsi	22
3.5	Timing delle richieste	23
3.6	Codice del programma	23

4	Risultati	25
4.1	Raccolta dati	25
4.1.1	Le problematiche	25
4.1.2	Filtro dati errati	26
4.1.3	Salvataggio	26
4.1.4	Distribuzione lunghezza tratte generate	27
4.2	Performance dei singoli mezzi	28
4.2.1	Automobile	30
4.2.2	Enjoy	35
4.2.3	Mezzi pubblici ATM, bici e a piedi	42
4.3	Confronto tra mezzi	45
4.3.1	Vittoria dell'automobile	45
4.3.2	Parziale sconfitta del car sharing	45
4.4	Confronto pre e post lockdown	48
4.4.1	Automobile	48
4.4.2	Enjoy	48
5	Conclusioni	51
5.1	Riflessioni sul risultato	51
5.2	Possibili estensioni	53

Capitolo 1

Introduzione

La mobilità a corto raggio è una percentuale dominante in Italia. È stato stimato che il 75% degli spostamenti sul territorio italiano è lungo meno di 10 km, e di questi circa la metà è inferiore ai 2 km [13]. Il mezzo principale usato per spostarsi all'interno delle aree urbane è l'automobile privata, che conta circa il 53% del totale degli spostamenti. Non stupisce che da diversi anni l'Italia è il secondo Paese in Europa col maggior numero di automobili per abitanti, con circa 640 auto ogni 1000 persone, dietro solamente al Lussemburgo [9].

1.1 Automobile in città

Nello studio annuale di TomTom del 2019 Milano rientra tra le prime 110 città su scala globale per livello di congestione stradale, con un traffico tale da far aumentare di quasi il doppio il tempo impiegato a percorrere una tratta durante le ore di picco¹. Molti sono stati i cambiamenti apportati a questa città per fronteggiare la domanda crescente di spostamenti e al contempo diminuire il traffico stradale, tra queste si citano la realizzazione della Linea M5, l'introduzione dell'Area C e delle numerose zone a traffico limitato (ZTL), l'aumento di piste ciclabili, fino alla concessione ad aziende nel campo della mobilità condivisa.

A fronte di questi dati sul traffico e delle restrizioni introdotte viene spontaneo chiedersi quanto sia influenzato il tempo di percorrenza in automobile su strade cittadine e, soprattutto, se questa influenza sia tale da favorire altri mezzi di trasporto. Sebbene sia un dato di fatto che, su strada, un'automobile sia più veloce di un tram o di una bicicletta, dall'altra bisogna tener conto dei vantaggi di cui godono gli altri mezzi. Basti pensare alle numerose strade

¹https://www.tomtom.com/en_gb/traffic-index/ranking/. Accessed: 20/01/2021

con corsia preferenziale o carreggiata a parte per bus e tram, e a mezzi che non soffrono minimamente dei problemi del traffico, come le metropolitane ATM e i passanti ferroviari Trenord.

L'obiettivo dello studio è quello di analizzare a livello di tempistiche l'impiego dell'automobile privata a motore termico, che rappresenta il 99.5% del parco auto², per tratte brevi all'interno del Comune di Milano e confrontarlo con mezzi di trasporto alternativi, al fine di capire se l'influenza delle congestioni che caratterizzano la città è tale da favorire uno di questi. Le principali motivazioni risiedono nel costo di questo mezzo e nei problemi derivati dal suo largo impiego.

1.2 Problemi

L'automobile a motore termico ha diversi problemi. Primo tra tutti è sicuramente l'inquinamento. I dati del 2018 dell'Automobile Club d'Italia (ACI) riportano come il 63% delle auto in circolazione sia di categoria Euro IV o minore, uno standard sulle emissioni introdotto nel 2006 e superato nel 2008 dall'Euro V, ben 12 anni fa². Se visto in prospettiva, l'inquinamento prodotto dalle auto non è una percentuale rilevante dell'inquinamento dell'aria, difatti è emerso in numerosissimi studi effettuati in diverse città del mondo, tra cui Milano e Brescia [4][2], che, durante il periodo locale di lockdown per l'emergenza da COVID-19, quindi a fronte di un traffico pressoché nullo, nell'aria si sono registrati gli stessi livelli di sostanze chimiche nocive dei periodi precedenti in condizioni di traffico abituale, con lievi abbassamenti riguardanti le sostanze riconducibili direttamente alla combustione nei motori termici. Questi studi suggeriscono che la maggior parte dell'inquinamento dell'aria è legata a fattori ambientali. Una conclusione simile è stata ottenuta analizzando i dati dell'inquinamento all'interno dell'Area C del Comune di Milano (ex Ecopass), un'area in cui sono state introdotte delle restrizioni sulla circolazione di alcuni veicoli, che concretamente ha dato scarsi risultati [15]. Nonostante questo non bisogna dimenticare che, per esempio, i gas di scarico derivanti dai motori diesel risultano sicuri cancerogeni per l'uomo, al pari del fumo di tabacco, con la sottile differenza che quest'ultimo è qualcosa che si può evitare³. L'inquinamento dunque rimane un problema in Italia, visto il parco auto ancora poco aggiornato e dal momento che la maggior parte del traffico è presente nelle città, dove le automobili, a livello di consumo di combustibile, risultano meno efficienti per via delle basse velocità, del traffico e delle numerose accelerazioni.

²<https://www.anfia.it/it/automobile-in-cifre/>. Accessed: 20/01/2021

³<https://monographs.iarc.who.int/>. Accessed: 20/01/2021

Secondo problema dell'automobile, se comparato con gli altri mezzi di trasporto per uso cittadino, è il costo. Una stima svolta da SosTariffe⁴ nel 2019 mostra come mantenere un'auto in Italia costi circa €1800 all'anno, compresi di assicurazione, benzina e bollo auto, senza contare il prezzo iniziale del mezzo, che si aggira intorno a €11000 per un'utilitaria a chilometro zero, il cambio delle gomme ed eventuali imprevisti. Seppur questo problema risulti molto relativo, dato che è il reddito e patrimonio di una persona a incidere sulla percezione di questa spesa, il solo costo di mantenimento di un'auto risulta molto più dispendioso rispetto alle sue alternative in città, come un abbonamento annuale per i mezzi pubblici o come l'acquisto e il mantenimento di una bicicletta o di mezzi alternativi. Va tenuto conto però che il costo assoluto del mezzo raramente è tra i fattori principali dietro l'investimento, guidato spesso da esigenze personali e lavorative.

Ultimo ma non ultimo dei problemi è l'efficienza dell'automobile in termini di tempo, che viene ridotta per via delle congestioni stradali ricorrenti che caratterizzano le città. Tra i contro dello spostamento in auto, a livello collettivo, vi è sicuramente l'ingombro stradale rapportato al numero di persone in viaggio, ovvero l'utilizzo dell'automobile per lo spostamento del solo conducente, nonostante la capacità maggiore dell'auto, tradotto nella congestione delle strade per un flusso di persone molto inferiore rispetto a quello di un bus. A documentarlo è l'Associazione Nazionale Comuni Italiani (ANCI) stimando che, degli 1.8 milioni di veicoli che si spostano quotidianamente nelle città italiane per tragitti casa-lavoro o casa-studio, 1.2 milioni viaggiano con solo il conducente a bordo⁵.

1.3 Vantaggi

A giustificare la grande diffusione di questo mezzo di trasporto tra la popolazione sono sicuramente i grandi vantaggi che si acquisiscono dal possederne una. L'auto può essere considerata sinonimo di libertà di muoversi: non si hanno limiti di orario di partenza; la decisione dell'orario di inizio viaggio è esclusivamente del conducente, al contrario dei mezzi pubblici dove l'orario di partenza è offerto in numero limitato e prestabilito, con scarse percorrenze negli orari notturni; si possono trasportare merci molto pesanti, difatti, tra gli impieghi più comuni dell'auto vi sono quello di fare la spesa al supermercato e di trasportare merce ingombrante acquistata, difficili da attuare in assenza di un veicolo. Inoltre, in caso di forte pioggia o maltempo, viaggiare

⁴<https://www.sostariffe.it/info/supporito-giornalisti>. Accessed: 20/01/2021

⁵<http://www.anci.it/mobilita-sostenibile-ricerca-anci-nelle-citta-serpenti-di-auto-vuote-ma-cala-inquinamento/>. Accessed: 20/01/2021

in auto risulta l'opzione più comoda e sicura, talvolta l'unica funzionante. Al momento dunque la comodità e praticità di un'automobile sono senza pari.

1.4 Confronto con mezzi alternativi

È sul terzo problema che si concentra questo studio, che ha come obiettivo capire quanto sia vantaggioso o svantaggioso in termini di tempi di percorrenza, a fronte di problemi quali traffico e congestioni stradali, l'uso dell'automobile rispetto a mezzi di trasporto alternativi all'interno del Comune di Milano, dal punto di vista di un'utenza con sole esigenze di spostamento, senza vincoli legati a trasporto merci, intemperie o raggiungimento di luoghi oltre l'area considerata. Per capire quale sia il mezzo più vantaggioso la soluzione ideale sarebbe quella di organizzare diverse gare in diversi punti della città in cui delle persone, ognuna con un mezzo di trasporto diverso, vengono cronometrate su un percorso cittadino comune, ovvero partendo da uno stesso punto di partenza e arrivando a uno stesso punto di arrivo, più volte al giorno e per diversi percorsi, al fine di raccogliere un campione abbastanza significativo di dati sul quale effettuare i confronti. Purtroppo, non avendo a disposizione le risorse e il tempo per effettuare un confronto del genere, si è optato per una soluzione informatica.

Capitolo 2

Metodi

L'idea alla base di questo studio è, dunque, quella di confrontare diversi mezzi di trasporto su percorsi comuni, ovvero con punto di partenza, destinazione, data e orario di partenza tutti uguali, salvando il tempo impiegato da ciascun mezzo per compiere il tragitto e successivamente usare questi valori per analizzarne le performance singolarmente e compararle. Non ci si aspetta un risultato diverso da quello che vede l'automobile prima in classifica, ma si vuole quantificare la percentuale di vittorie e studiare eventuali variazioni lungo l'arco della giornata di questo vantaggio, visti i problemi relativi al traffico e alle congestioni stradali di cui le grandi città soffrono. Un'idea sul possibile posizionamento in classifica la danno i dati ISFORT del 2019 basati su interviste a campione [13], riportati nella Tabella 2.1, che mostrano le velocità medie percepite dagli utenti basate sui loro personali viaggi.

	auto	bicicletta	mezzi pub.
V.media	22 km/h	15 km/h	14 km/h

Tabella 2.1: Dati ISFORT 2019, osservatorio *Audimob*, velocità media.

Avendo a disposizione i dati relativi al percorso di confronto, tra cui il tempo impiegato da ciascun mezzo, sarebbe possibile ricreare la tabella per ogni fascia oraria, avendo l'opportunità di effettuare ulteriori analisi, come la variazione della velocità media di ora in ora per ogni mezzo o per esempio l'influenza che ha la lunghezza della tratta sulla velocità media. Purtroppo però, un dataset contenente i tempi di percorrenza di tragitti comuni effettuati da diversi utenti su diversi mezzi di trasporto non esiste. Gli unici che possiedono dataset simili di tratte realmente effettuate ma con un singolo mezzo di trasporto, salvo rare eccezioni, sono le aziende private che offrono il servizio di navigazione, come per esempio Google Maps, Waze,

Moovit e TomTom. Tali aziende infatti offrono applicazioni per poter guidare l'utente da un punto a ad un punto b con un determinato mezzo, per esempio mezzi pubblici, in tempo reale e col tracciamento GPS, e dato che il servizio nella maggior parte delle volte è offerto gratis, il pagamento per tale servizio è effettuato coi dati del viaggio che l'utente produce passivamente.

2.1 Lavori precedenti

Non sono molti gli articoli accademici riguardanti un lavoro simile, molto probabilmente per la difficoltà materiale di effettuare un numero significativo di confronti reali su strada o nel reperire dei dati storici di viaggi realmente effettuati su diversi mezzi di trasporto.

Uno di questi studi è stato svolto nel 2011 a Sidney, Australia, per confrontare la competitività dei tempi di percorrenza in bicicletta rispetto a quelli in automobile [7]. Si legge dall'articolo che sono stati impegnati 178 partecipanti volontari residenti in 6 sobborghi diversi della città per monitorare tramite GPS i loro spostamenti in automobile. Inoltre sono stati incaricati di fornire, tramite un sito web a loro dedicato, dati aggiuntivi riguardo il loro viaggio, tra cui il luogo di destinazione. Per i tempi di percorrenza in bicicletta invece è stato usato un programma per richiedere a Google Maps di calcolare un itinerario a piedi in tempo reale, ovvero ogni volta che tramite GPS veniva segnalato l'inizio di un viaggio in auto da parte di un partecipante, dal punto di partenza al punto di arrivo indicati, a cui è stata applicata la velocità media di un ciclista di 16 km/h per derivarne il tempo di percorrenza. I risultati ottenuti hanno evidenziato come più del 90% dei tragitti in auto fino a 5 km di lunghezza potesse essere coperto in bicicletta con una penalità di tempo entro i 10 minuti, un dato rilevante visto che a Sidney, sempre secondo l'articolo, il 43% degli spostamenti in auto è impiegato per tragitti inferiori a 5 km.

Un altro studio simile è stato fatto nel 2015 a Cape Town, Sudafrica, per confrontare i tempi di percorrenza tra mezzi pubblici e automobile sulle medio-lunghe distanze [12]. A differenza dello studio di Sidney in questo caso è stato chiesto ai partecipanti di svolgere dei tragitti prestabiliti, studiati per coprire geograficamente tutta l'area della città, di cui i luoghi di partenza sono stati scelti strategicamente per essere vicini entro 2 km dalla più vicina fermata del bus o del treno e i cui orari di partenza coincidono con gli orari di picco e non del traffico della città. Ai partecipanti è stato chiesto quindi di effettuare tali tragitti ognuno con un mezzo di trasporto diverso, dall'automobile di proprietà alle alternative proposte dal trasporto pubblico, quali bus, treno e bus a lunga percorrenza, registrando i dati riguardanti il loro

viaggi, tra cui il tempo impiegato a raggiungere le fermate e i tempi di attesa per il gruppo dedicato ai mezzi pubblici. I risultati hanno evidenziato come il trasporto pubblico sia nettamente in svantaggio rispetto all'automobile in ogni singolo fronte, dalla lunghezza del percorso più lunga alla velocità media inferiore, persino scartando i tempi di percorrenza a piedi e i tempi di attesa. Facendo leva sui dati a disposizione relativi a ogni singola tappa del tragitto, gli autori hanno analizzato e proposto diverse aree di intervento per migliorare la velocità e l'efficienza del trasporto pubblico della città.

In uno studio del 2017 svolto a New York City, Stati Uniti, sono stati messi a confronto i tempi di percorrenza tra una corsa in taxi e l'utilizzo di un servizio di bike sharing presente nella città [10]. Gli autori si sono provvisti dei dataset del 2014 liberamente messi a disposizione dal sito web dei corrispettivi servizi associati, un caso eccezionale di adempimento alla filosofia Open Data. Una volta acquisiti anche i dati relativi alle stazioni di servizio del bike sharing presenti nella città hanno disegnato intorno a ciascuna di esse un cerchio del raggio di 250 metri per selezionare tutte le tratte percorse in taxi che sono iniziate e terminate entro questi cerchi, agglomerando insieme tratte in bike sharing e in taxi che hanno percorso lo stesso tragitto allo stesso orario. Dai risultati è emerso che, durante le ore di congestione stradale del mattino e del tardo pomeriggio, nelle tratte brevi inferiori ai 3 km di lunghezza, il bike sharing ha impiegato un tempo minore o uguale al taxi il 50% delle volte. Tali risultati sono stati ulteriormente analizzati, scoprendo che la maggior parte delle vittorie del bike sharing è associata a tratte con un punto di partenza e/o di destinazione con maggior densità di luoghi di lavoro. Gli autori concludono suggerendo dei punti di intervento per favorire i percorsi in bicicletta anche sulle medie distanze, dove in media i taxi hanno avuto la meglio.

2.2 Primo prototipo del progetto

Avendo l'accesso allo storico dei viaggi effettuati dagli utenti di più servizi, ognuno legato a un mezzo di trasporto diverso, come Moovit per i mezzi pubblici e Waze per l'automobile, si potrebbero cercare delle tratte simili nei dataset di ognuno di essi, riguardanti percorsi di viaggiatori che inconsapevolmente hanno condiviso punto di partenza e punto di arrivo, per confrontarne i tempi di percorrenza. Tenendo conto che sarebbe poco efficace usare un operatore di uguaglianza su delle coordinate geografiche prodotte da viaggiatori, dato che risulterebbero poche quelle che effettivamente coincidono, il che vorrebbe dire persone diverse che hanno fatto lo stesso identico viaggio, il modo migliore per catturare dei tragitti comuni sarebbe quello di considerare tali

coordinate di partenza e di arrivo a gruppi, ovvero circoscritti entro un certo raggio oppure vicini tra loro entro una certa soglia, e quindi considerabili come lo stesso viaggio. Per catturare questo concetto di gruppi si potrebbe utilizzare uno degli algoritmi di clusterizzazione che sono stati pensati apposta per questo tipo di problemi. Un possibile approccio infatti sarebbe quello di fondere i dataset dei diversi servizi e clusterizzare i punti di partenza e destinazione più frequenti, usando un algoritmo come il DBSCAN [8] o una delle sue varianti ottimizzate appositamente per dati geospaziali [17][1] che fanno della distanza tra i punti il fulcro dell'algoritmo, per mettere a confronto i mezzi di trasporto usando i viaggi appartenenti allo stesso cluster sia di partenza che di destinazione. Un approccio del genere però risulterebbe problematico sotto alcuni aspetti, primo tra tutti l'esecuzione dell'algoritmo stesso, infatti il DBSCAN ha una complessità spaziale di $O(n^2)$ se si usa una matrice triangolare superiore di grandezza $(n^2 - n)/2$ per evitare calcoli già effettuati e velocizzare l'algoritmo, il che risulterebbe problematico da eseguire su un comune PC in caso di un numero abbondante di tratte, senza contare i doverosi tentativi di stima del parametro ϵ riguardante la distanza tra punti, fondamentale per un buon risultato.

Un altro problema legato a questo modo di procedere, indipendentemente dall'algoritmo usato per raggruppare le tratte, è quello di trovare dei viaggi percorsi nella stessa data e allo stesso orario, il che restringerebbe il campione dei viaggi drasticamente. Difatti avrebbe poco senso confrontare tratte percorse in diversi giorni o a orari diversi per via del traffico che caratterizza una città, nello specifico quella di Milano. Secondo diversi servizi di monitoraggio del traffico infatti un viaggio con l'automobile in una settimana lavorativa alle 8:00 è molto sfavorito per via del picco di congestione stradale che caratterizza quel giorno e quell'orario, di conseguenza, confrontarlo con un qualsiasi altro viaggio che non abbia sofferto di questo rallentamento sarebbe poco significativo. Quello che si potrebbe fare avendo a disposizione un dataset di viaggi effettuati realmente sarebbe l'analisi delle performance di ogni mezzo rapportate alle condizioni storiche del traffico per misurarne l'impatto, oppure, come nel caso dello studio di New York City [10], si potrebbero scegliere dei punti geografici prestabiliti e cercare nei dati delle tratte che sono iniziate e terminate entro questi punti.

2.2.1 Assenza di Open Data

La pubblicazione totale o parziale di questi dati è una pratica ancora poco diffusa che rientra nella filosofia dell'Open Data. Secondo questa filosofia, analoga a quella dell'Open-Source, i dati dovrebbero essere liberamente accessibili da chiunque, senza nessun tipo di restrizione derivante da copyright,

patentini o altri meccanismi di controllo che ostacolano la redistribuzione. Uno dei campi che godrebbe maggiormente di questa filosofia sarebbe proprio la ricerca scientifica, che grazie all'accesso facilitato a questi dati riuscirebbe ad accelerare i tempi delle scoperte, e di conseguenza perfino le aziende stesse che possiedono questi dati ne gioverebbero. Non a caso sono già molti gli studi effettuati basati su dati acquisiti tramite scraping di siti web che sono riusciti a dare un contributo sia in ambito di ricerca e sia all'azienda stessa. Di questi vi è una sottocategoria dedicata ai trasporti. In questa categoria, un tema di spicco tra tutti è l'analisi dei dati prodotti dai servizi di mobilità condivisa. Per esempio, in uno studio effettuato a Milano nel 2014 sul bike sharing dell'azienda ATM chiamato BikeMi è stato osservato un maggior utilizzo del servizio nelle vicinanze di stazioni ferroviarie, università e zone a traffico limitato (ZTL) rispetto ad altri luoghi [5]. Un altro studio del 2017 effettuato sempre a Milano propone un sistema di nome Knowledge Discovery System basato su scraping dei dati di tre servizi di car sharing, Car2Go, Enjoy e Twist, in grado di offrire stime in tempo reale dei tempi di percorrenza e di rilevare le zone in cui la richiesta del servizio è alta e le auto a disposizione scarseggiano [14]. In questo studio inoltre viene evidenziato che durante il giorno le auto sono maggiormente concentrate nella zona centrale di Milano mentre durante la sera si riversano nelle zone suburbane. In entrambi i due studi presi come esempio sono emerse delle informazioni che risulterebbero molto utili per le aziende fornitrici dei servizi e che, di conseguenza, se venissero prese in considerazione aumenterebbero lo sfruttamento del servizio da parte degli utenti, una soluzione win-win.

Nonostante ciò, nessuna delle aziende che possiede dati generati dagli utenti ha mai reso disponibile apertamente quello che ha raccolto, dato che non ha alcun incentivo nel pubblicarlo nemmeno in parte, nonostante siano i loro utenti i veri proprietari dei dati generati. Alcune di esse si limitano a pubblicare regolarmente delle infografiche e dei dati statistici basati su di essi. Altre ancora si limitano a usare tali dati per promuoversi online e quindi farsi pubblicità. Queste pratiche però risultano problematiche sotto diversi aspetti, primo tra tutti è l'impossibilità di verificarne la veridicità, dato che la sorgente è nascosta, risultando quindi poco credibili se si considera l'interesse primario di un'azienda. Un secondo problema, dal punto di vista dell'azienda stessa, è quello di dover schierare, e quindi pagare, analisti di dati per produrre tali statistiche che però non hanno alcuna valenza scientifica, quando potrebbero solamente incaricarsi di rendere anonimi dei dati e pubblicarli per ricevere gratuitamente studi fatti su tali dati da studenti e ricercatori di tutte le parti del mondo.

2.3 L'alternativa: simulazione

Non avendo nemmeno quest'ultima possibilità di avere accesso a dataset di viaggi compiuti con diversi mezzi e visti i problemi che avrebbe introdotto tale approccio, si è scelto di creare un dataset di gare tra mezzi attraverso una simulazione. Quando si effettua una richiesta a un servizio come Google Maps per un tragitto, quello che fa il servizio è calcolare sulla base di dati in tempo reale di traffico, incidenti e dati GPS una stima del tempo di percorrenza per diversi mezzi di trasporto per compiere tale tragitto, come automobile, mezzi pubblici, tragitto a piedi e bicicletta. Sebbene siano solo stime e non cronometraggi realmente effettuati, esse risultano estremamente precise o quanto di più simile alla realtà. Lo stato attuale della tecnologia e la loro vasta e affermata presenza nella vita quotidiana delle persone permette a tali servizi di avere informazioni sempre più precise, utilizzando tecniche come quella del crowdsourcing di dati dagli smartphone degli utilizzatori e sfruttando lo storico dei viaggi effettuati. Dunque, si potrebbero usare questi dati non troppo lontani dalla realtà e costantemente aggiornati per confrontare i vari mezzi di trasporto su percorsi comuni, generati a random. Non avendo altra scelta, quest'ultimo è stato scelto definitivamente come metodo per questo studio.

2.3.1 Crowdsourced data: il caso Waze

Sul blog ufficiale di Waze è stata pubblicata la descrizione ad alto livello dell'algoritmo usato per fornire una stima del tempo di arrivo (in inglese noto come ETA: Estimated Time of Arrival), tradotta graficamente in stima di tempo di percorrenza¹. A fronte di una richiesta di viaggio, per prima cosa vengono prelevati i dati storici della velocità media di ogni segmento di strada compreso nell'itinerario calcolato. Questo particolare dato viene aggiornato ogni 30 minuti usando i dati prodotti dallo smartphone dei loro utenti, e nel caso non sia presente nessun dato la ricerca procede a ritroso fino alle ultime 8 settimane. In secondo luogo usa i dati in tempo reale che ha a disposizione dagli utenti connessi e dei dati del traffico. Infine, tramite filtri per eliminare dati fuori scala, ovvero gli utenti estremamente veloci o lenti, ed equazioni matematiche elaborano una stima del tempo di percorrenza. Inoltre, tale stima viene aggiornata in tempo reale man mano che si procede nel viaggio, sempre tramite i dati collezionati da altri utenti. Quello che ci tiene a sottolineare Waze infatti, e che la rende competitiva, è che la stima dei tempi di percorrenza viene basata interamente sulla loro comunità di utenti.

¹<https://medium.com/waze>. Accessed: 20/01/2021

Capitolo 3

Progettazione

Per attuare questo confronto tra mezzi di trasporto si è scelto di appoggiarsi a servizi di navigazione in grado di fornire stime di percorrenza in tempo reale per usarle come valori effettivi su cui basare le analisi. Le gare sono state tradotte in richieste per un tragitto comune, effettuate più volte al giorno e per diversi percorsi al fine di rendere possibile uno studio su degli stessi viaggi iniziati nello stesso istante. Sono stati presi in considerazione l'automobile e i principali mezzi di trasporto alternativi con cui è possibile spostarsi all'interno del Comune di Milano:

- automobile;
- servizio di car sharing Enjoy;
- mezzi pubblici ATM e passanti ferroviari Trenord;
- bicicletta;
- a piedi.

Sebbene il servizio di car sharing riguardi comunque l'utilizzo di un'automobile, è stato preso in considerazione come mezzo alternativo per spostamenti occasionali. Questo servizio infatti permette di sfruttare molti dei vantaggi dell'auto senza possederne una e solo quando strettamente necessario, vantaggi validi fintanto che il costo dell'uso occasionale risulta più economico dell'acquisto e del mantenimento di un'auto privata. L'utilizzo di questo servizio in determinate condizioni riuscirebbe ad abbattere, oltre il costo dell'auto, il problema dell'inquinamento dovuto dal parco auto non aggiornato. La maggior parte delle aziende nel settore infatti offre una flotta molto aggiornata, a partire dalle Fiat 500 con standard Euro V per Enjoy¹ fino alle

¹<https://enjoy.eni.com/it/milano/flotta>. Accessed: 20/01/2021

Smart Fortwo con motori Euro VI di ShareNow² (ex Car2Go). Il percorso a piedi, seppur non diretto competitore, è stato scelto per essere usato come riferimento.

3.1 Ricerca delle API

Il primo passo per la realizzazione di questo studio è stato quello di cercare delle Application Programming Interface (API) disponibili per ogni mezzo di trasporto in grado di fornire stime dinamiche sui tempi di percorrenza basate su dati in tempo reale.

3.1.1 Google

Inevitabilmente Google è stata la prima opzione che si è cercata, in grado di offrire stime di percorrenza per auto, bici, mezzi pubblici e percorsi a piedi. Google infatti, grazie a Maps e grazie all'acquisizione di Waze, è una dei leader nel settore delle informazioni di navigazione che vende la propria conoscenza e potenza di calcolo tramite API. Tra le varie opzioni in vendita al momento della ricerca era presente la Directions API³, in grado di fornire stime di percorrenza molto precise per l'auto, per il percorso a piedi e per la bicicletta, quest'ultima solo in presenza di piste ciclabili lungo il percorso. Questa opzione è stata subito scartata per via del costo eccessivo del servizio di 10\$ per 1000 richieste al mese⁴, l'equivalente di circa 1 richiesta all'ora, troppo poco per l'obiettivo dello studio di catturare la variazione oraria delle performance dei mezzi e troppo costoso per essere portato sulla scala voluta di almeno 10 richieste l'ora, una ogni dieci minuti.

3.1.2 Waze

Nonostante sia stata acquisita da Google, Waze è rimasta indipendente. È specializzata nelle stime di percorrenza in automobile e in moto. Quello che offre Waze però non sono esattamente delle API a cui fare richiesta per un determinato viaggio, ma semplicemente degli URL da inserire nel proprio sito web o nella propria applicazione Android/iOS per aprire il client Waze coi dati della richiesta parametrizzati nell'URL⁵.

²<https://www.share-now.com/it/it/fleet/>. Accessed: 20/01/2021

³<https://developers.google.com/maps/documentation>. Accessed: 20/01/2021

⁴<https://cloud.google.com/maps-platform/pricing/sheet>. Accessed: 20/01/2021

⁵<https://www.waze.com/sdk>. Accessed: 20/01/2021

3.1.3 Moovit

Anche Moovit, specializzata nei mezzi pubblici e nei treni, non offre API per interrogare direttamente i loro servizi e ricevere delle stime, ma solamente l'opzione di incorporare un widget all'interno della propria pagina web o applicazione che rimanda direttamente ai loro servizi⁶. Al momento della ricerca, Moovit offriva un servizio gratuito di API chiamato DeepLink la cui pagina risultava inaccessibile, restituendo il codice di stato HTTP 500 "Internal Server Error".

3.1.4 Here

Il servizio Here, specializzato nei percorsi in auto, scooter e mezzi pubblici, offre una REST API per le informazioni riguardo il routing e la stima del tempo di percorrenza⁷. Queste stime però risultano accurate, ovvero basate su dati in tempo reale, solo per determinate città comunicate sul loro sito, e tra queste non era inclusa Milano. Per le città non in lista la stima è fatta staticamente sulla base di dati geografici o dei dati tabellari degli orari dei passaggi dei mezzi nel caso dei trasporti pubblici.

3.1.5 Enjoy

Non sono stati trovati servizi in grado di fornire stime di tempi di percorrenza utilizzando un mezzo dei servizi car sharing. Il servizio Enjoy non offre API per tale calcolo ma dispone di una mappa costantemente aggiornata sul loro sito web con l'elenco delle auto libere, ognuna con allegato dati riguardo il tipo di veicolo, tra auto e moto, e il carburante rimasto. Un software in grado di acquisire le informazioni delle auto libere è stato gentilmente offerto da Losacco Federico, che lo ha scritto e utilizzato per uno studio sul comportamento del traffico basato sui viaggi compiuti dagli utenti del servizio [16].

3.1.6 OpenStreetMap

Il progetto collaborativo di OpenStreetMap offre, tra le altre cose, il calcolo di un tragitto usando come backend servizi offerti da terze parti, tra cui Open Source Routing Machine (OSRM), per stimare percorsi in auto, in bici e a piedi. Purtroppo però, tali servizi non usano dati in tempo reale⁸.

⁶<https://moovit.com/developers/>. Accessed: 20/01/2021

⁷<https://developer.here.com/documentation>. Accessed: 20/01/2021

⁸<https://wiki.openstreetmap.org/wiki/Routing>. Accessed: 20/01/2021

3.2 Scelta delle API

Di seguito sono elencate le scelte fatte per i servizi e i loro rispettivi ruoli di copertura dei mezzi di trasporto:

- Here API
 - mezzi pubblici ATM e passanti ferroviari Trenord;
- OpenStreetMap con OSRM
 - bicicletta e a piedi;
- scraping di Enjoy Map usando il software di Losacco Federico [16]
 - servizio di car sharing Enjoy;
- scraping di Waze personalmente programmato
 - automobile di proprietà.

Nonostante OSRM non offra soluzioni in tempo reale è stato scelto ugualmente per via della scarsa influenza del traffico sui tempi di percorrenza in bicicletta e a piedi, che dipendono per la maggior parte dall'itinerario calcolato e dalle caratteristiche delle strade, come segnali stradali di divieti, contromano e di aree pedonali. Le API di Here sono state utilizzate per mancanza di valide alternative. In un primo momento si è tentato lo scraping del sito web di Moovit per interagire col sito e comunicare le richieste senza intervento umano, ma la difficoltà nel codificare i parametri delle richieste e a decodificarne le poche risposte immerse in svariati messaggi di errore ne hanno scoraggiato l'utilizzo. Nella tabella 4.1 sono stati riportati in percentuale gli errori dovuti all'impiego di Moovit. Per il servizio di car sharing si è scelto di usare la lista delle auto libere come base, ottenuta tramite scraping, per ottenere una stima del tempo di percorrenza basata su due percorsi, per primo il tragitto a piedi per raggiungere l'auto libera più vicina al punto di partenza richiesto e come secondo il tragitto in auto per completare il percorso dal punto in cui si trova l'auto selezionata alla destinazione richiesta. Non sono state trovate valide alternative per il servizio di routing in automobile. Google risulta essere eccessivamente costosa se si considera un minimo di 10 richieste all'ora per avere un campione rappresentativo di un giorno da diversi punti della città. Il prezzo per una prestazione del genere è di circa 72\$ al mese. Waze invece non offre nessuna API per essere interrogata direttamente. Per ovviare a questo problema, si è scelto

di programmare un browser headless in grado di interagire direttamente col sito web di Waze per interrogare il servizio e prelevare i dati della stima di percorrenza, simulando per intero l'interazione umana con la pagina. Le API di Here e quelle di OpenStreetMap invece chiedono in input le coordinate di partenza e di destinazione e restituiscono in output una stima di percorrenza del tragitto in formato JSON di quel determinato momento, oltre ad altri dati relativi alla tratta consigliata.

3.2.1 Lo scraping di Waze

Dato che i servizi per sviluppatori offerti da Waze sono gratuiti, si è scelto di sfruttare tali dati ma prendendoli per vie laterali, direttamente dall'applicazione web <https://www.waze.com/live-map>. La maggior parte dei browser disponibili sul web come Mozilla Firefox offrono, oltre al browser in sè, gli strumenti di debugging usati principalmente dagli sviluppatori di siti web. Tra questi vi è la console da sviluppatore, apribile tramite la scorciatoia da tastiera Ctrl+Shift+K. Con questa console è possibile vedere le richieste di rete che il sito effettua in background per il suo funzionamento. L'applicazione web di Waze è un'applicazione a singola pagina e per il suo funzionamento effettua delle richieste tramite il linguaggio di programmazione JavaScript in modo asincrono direttamente ai loro server. Questa pratica prende il nome di Asynchronous JavaScript and XML (AJAX)⁹. Guardando i comportamenti a livello di rete del sito web mentre si svolge una richiesta per un viaggio a Waze si presenta subito un URL di riferimento usato come punto di ricevimento delle richieste di routing, come mostrato nella Figura 3.1. In particolare, la richiesta <https://www.waze.com/live-map/routing/?> contiene esattamente il punto di partenza e di arrivo richiesti parametrizzati nell'URL e il risultato restituito è in formato JSON (JavaScript Object Notation)¹⁰. Per proteggere i loro server da sovraccarichi, Waze risponde a questo URL solamente se si possiede il permesso ricevuto preventivamente tramite un token, ovvero una stringa di caratteri che funge da chiave segreta. Per ottenere questa chiave e raggiungere l'obiettivo finale di avere delle API per interrogare il servizio, è stato scritto un programma in Go¹¹ per emulare l'interazione col sito web e ricevere il token necessario per effettuare le richieste successive.

⁹[https://en.wikipedia.org/wiki/Ajax_\(programming\)](https://en.wikipedia.org/wiki/Ajax_(programming)). Accessed: 20/01/2021

¹⁰<https://en.wikipedia.org/wiki/JSON>. Accessed: 20/01/2021

¹¹<https://golang.org/>. Accessed: 20/01/2021

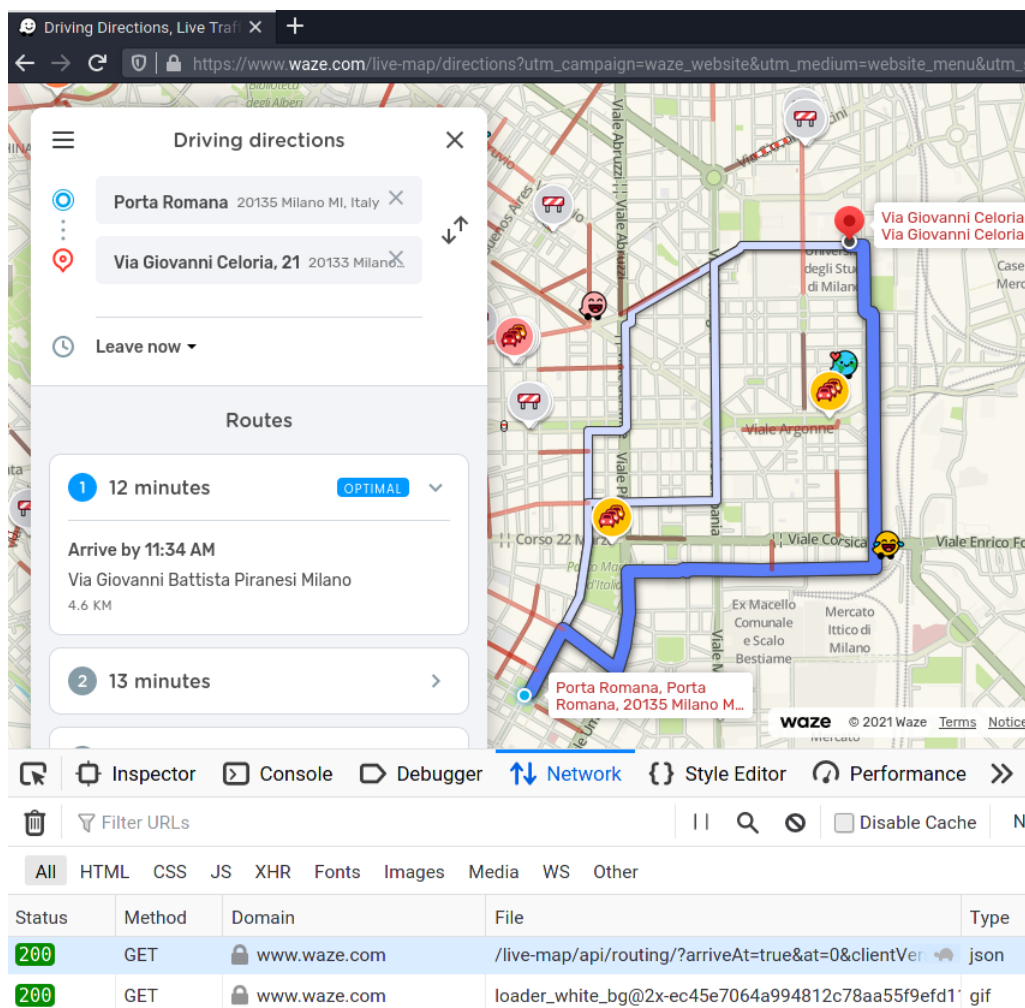


Figura 3.1: Console da sviluppatore su Mozilla Firefox.

Questo processo ha richiesto tempo sia per la codifica dei parametri e sia per la decodifica del risultato stesso e dei risultati intermedi per la richiesta del token. In altre parole, è stato fatto il reverse engineering dell'applicazione web.

3.2.2 Lo scraping di Moovit (scartato)

Inizialmente, prima di usare le API di Here, che offre stime statiche dei tragitti coi mezzi pubblici, si è tanto lo scraping dell'applicazione web di Moovit <https://moovitapp.com/> per via della qualità migliore delle stime di percorrenza, basate su dati in tempo reale. Come per lo scraping di Waze, è stato eseguito lo stesso procedimento, guardando tramite la console da sviluppa-

tore del browser le richieste HTTP svolte in background dal sito, studiando tutte le interazioni e decodificando le risposte. Sfortunatamente l'applicazione web è risultata molto più complessa del previsto. Dopo diversi giorni di studio è stato scritto un software per emulare le interazioni della pagina web coi server di Moovit ed è stato affiancato agli altri moduli per risolvere le richieste di tragitto. Le stime di percorrenza relative a Moovit raccolte nella prima settimana di esecuzione del programma sono risultate nulle nell'80% delle risposte. Visti gli scarsi risultati e vista l'enorme complessità del programma di scraping, si è scelto di optare per le API di Here al costo di perdere precisione nelle stime per i mezzi pubblici.

3.2.3 Estratti del codice sorgente degli scraper

Nei listati seguenti sono riportate le funzioni principali, scritte in Go, addette a interrogare i server dietro ai relativi siti web dei servizi per la tratta richiesta, espressa sotto forma di coordinate geografiche di partenza e arrivo, il cui risultato viene successivamente elaborato e salvato da un altro modulo del programma. Per questioni di leggibilità sono stati omessi i controlli di errore presenti dopo ogni richiesta ed è stata semplificata la sintassi del linguaggio. Ogni chiamata di funzione corrisponde una richiesta HTTP appositamente preparata e compilata coi parametri indicati per ottenere in risposta i dati necessari da inserire nella richiesta dopo, al fine di avere accesso al servizio di routing nascosto dietro il sito web.

Codice dello scraper di Waze

```
1 func
2 GetRoutes(startLat, startLon, endLat, endLon string) Result {
3     // Richiesta dei cookies necessari per effettuare
4     // le richieste di routing
5     cookies := Cookies()
6     _      := setCookieConsent()
7
8     // Richiesta di tragitti e stime di percorrenza per la
9     // tratta desiderata
10    return Routes(startLat, startLon, endLat, endLon, cookies)
11 }
```

Codice dello scraper di Moovit

```
1 func
2 GetRoutes(startLat, startLon, endLat, endLon string) Result {
3     // Richiesta del nome canonico della via corrispondente
4     // perche' i server non accettano coordinate
5     from := LocationName(startLat, startLon)
6     to   := LocationName(endLat, endLon)
7
8     // Richiesta di parametri per fare richieste ai server.
9     // Senza di essi i server non accettano altre interazioni
10    HTTPHeaders := HeadersNeededForHTTP(from, to)
11    cookie      := Cookie(HTTPHeaders)
12    _           := SetMagicKey(HTTPHeaders, cookie)
13
14    // Richiesta di dati che descrivono la tratta in formato
15    // binario (illeggibile)
16    fromData := LocationInfo(from, HTTPHeaders, cookie)
17    toData   := LocationInfo(to, HTTPHeaders, cookie)
18
19    // Richiesta del token, ovvero la chiave per effettuare
20    // le domande ai server. Ne serve una nuova per ogni
21    // richiesta
22    token := Token(fromData, toData, HTTPHeaders, cookie)
23
24    // Richiesta di tragitti e stime di percorrenza per la
25    // tratta desiderata
26    return Routes(fromData, toData, token, cookie)
27 }
```

3.2.4 Stima del percorso per il car sharing

Avendo a disposizione un servizio di stime di percorrenza per tragitti a piedi e avendo trovato un modo per risolvere quelli in automobile, per il servizio di car sharing Enjoy si è scelto di calcolare una stima sulla base di questi due. Nello specifico, per ogni percorso richiesto viene scaricata la lista aggiornata delle auto libere presenti sul territorio e viene scelta l'auto più vicina calcolando la distanza dal punto di partenza richiesto a ogni auto presente nella lista, scegliendo l'auto con la distanza minima. Una volta trovata vengono richieste e sommate la stima di percorrenza per raggiungere l'auto selezionata a piedi e la stima in automobile dalla posizione di quest'ultima alla destinazione. Sebbene non sia l'approccio migliore, dato che non tiene conto della direzione e del verso di marcia per compiere il viaggio, e quindi rendendo possibile la scelta di un'auto più vicina al punto di partenza ma più lontana dal luogo di destinazione, è risultato il più semplice da programmare e testare.

3.3 Scelta tra percorsi random e prefissati

Si è scelto di usare diverse tratte su cui basare il confronto allo scopo di ricoprire a livello geografico una buona parte del Comune di Milano e di diversificare i confronti in base alle caratteristiche delle tratte, quali lunghezza e punti di partenza e destinazione, per simulare al meglio la posizione di un possibile utente nella mappa. Per rendere automatiche le richieste da inoltrare ai servizi di navigazione è stata scritta una funzione per generare le tratte. Sono stati presi in considerazione tre principali approcci per generarle: hard-coded random; hard-coded di tratte realmente percorse; generazione a random just-in-time.

I primi due approcci sono risultati fin da subito problematici. In primo luogo, la scelta della destinazione avrebbe introdotto bias riguardo al possibile utente del percorso, portando ad analizzare la mobilità solamente dal punto di vista di una determinata categoria, per esempio: selezionare tratte con delle università come punto di arrivo porterebbe ad analizzare la mobilità solamente dal punto di vista degli studenti e dipendenti presso quelle determinate strutture. Altri problemi simili sarebbero sorti nello scegliere la partenza, la lunghezza, la distanza dal centro città e numerosità delle tratte. Seconda problematica molto più rilevante è che, involontariamente, si sarebbero introdotti dei percorsi che avrebbero favorito un mezzo piuttosto che un altro. Nella città di Milano infatti si contano numerosi tratti stradali di questo genere: strade con corsia preferenziale per mezzi pubblici e taxi;

strade a singola corsia e con numerosi semafori, quindi più soggetta a incollamenti; tangenziali; tratte coperte da passanti ferroviari, che sono più veloci delle metropolitane e con meno fermate. Tale scelta avrebbe portato ad analizzare dati non rappresentativi della città, con conseguenti risultati sbilanciati. Per questi stessi motivi si è deciso di non usare collezioni di dati disponibili da lavori precedenti riguardi servizi di mobilità condivisa, tra cui quelli del car sharing raccolti nel lavoro di Losacco [16], ovvero per evitare di analizzare la mobilità dal punto di vista di una sola utenza e per evitare l'inclusione di percorsi strutturalmente più favorevoli a un mezzo rispetto che a un altro.

Usando l'approccio della generazione a random non solo si sarebbero evitati questi problemi, ma i problemi stessi si sarebbero trasformati in analisi da poter eseguire a posteriori, per esempio selezionando da tutte le tratte generate a random quelle che hanno portato nei pressi di un'università, o tratte che in linea d'aria hanno coperto particolari strade favorevoli a determinati mezzi di trasporto. Visti i vantaggi e la flessibilità dell'approccio si è optato per quest'ultimo.

3.4 Generatore random dei percorsi

Prima ancora di scrivere una funzione per generare tratte a random è stata scelta l'area geografica all'interno della quale generarle. Siccome il servizio di car sharing Enjoy non permette di usare la propria flotta al di fuori del Comune di Milano, tale area è stata selezionata come terreno per i confronti, rappresentata sotto forma di rettangolo per questioni di semplicità nella programmazione. Nello specifico sono state scelte le coordinate (45.450562° , 9.158959°) e (45.482032° , 9.206763°) rispettivamente come vertice in basso a sinistra e in alto a destra del rettangolo rappresentativo dell'area selezionata. In questo modo la generazione di punti a random è stata ridotta alla generazione di coordinate maggiori o uguali della prima e minori o uguali della seconda. Una volta scritta la funzione per la generazione a random delle tratte è stato introdotto un vincolo per simulare dei percorsi scomodi da fare a piedi, ovvero per generare tratte che supererebbero i 20 minuti di camminata per essere coperte, e al tempo stesso che giustificerebbero l'utilizzo dell'automobile. Si è implementato tale vincolo scegliendo di usare 2 km in linea d'aria dal punto di partenza a quello di destinazione come misura minima della lunghezza di una tratta. Un secondo vincolo è stato introdotto per avere una maggiore eterogeneità delle tratte generate a livello geografico. Molte delle linee principali dei mezzi pubblici infatti attraversano il centro storico di Milano e tale area rappresenta più della metà dell'area

selezionata dal vincolo precedente. Siccome è stato scelto di ricoprire a livello geografico tutta l'area di Milano, si è programmato il generatore in modo da creare in maniera equidistribuita le seguenti tipologie di tratte: da dentro il centro storico a fuori; da dentro a dentro; da fuori a fuori; da fuori a dentro. Il rettangolo rappresentativo del centro storico è stato disegnato usando le coordinate geografiche (45.450562°, 9.158959°) e (45.482032°, 9.206763°) rispettivamente come vertice in basso a sinistra e in alto a destra.

3.5 Timing delle richieste

Le richieste sono state programmate per essere effettuate dalle 7:00 alle 23:59 di ogni giorno. La scelta di questo intervallo è stata vincolata dall'orario di servizio dei mezzi pubblici ATM in cui viene garantito il pieno regime. Per rispettare i limiti giornalieri delle varie API si è scelto di effettuare 1 richiesta al minuto per un totale di circa 1000 richieste al giorno, dove ogni richiesta rappresenta un tragitto generato a random mandato simultaneamente a ogni servizio di navigazione per ottenere una stima da ognuno di essi.

3.6 Codice del programma

```
1 func main() {
2     var a, b, c coordinate
3
4     for {
5         a, b = creaTragittoRandom()
6         c = enjoy.trovaAutoPiuVicinaOra(a)
7
8         risultato := []stime{
9             here.stimaTragitto(a, b),
10            waze.stimaTragitto(a, b),
11            osm.stimaTragitto(a, b, "foot"),
12            osm.stimaTragitto(a, b, "bike"),
13            osm.stimaTragitto(a, c, "foot") +
14                waze.stimaTragitto(c, b)}
15
16        save(risultato)
17        sleep("1min")
18    }
19 }
```


Capitolo 4

Risultati

4.1 Raccolta dati

I dati riguardanti la simulazione sono stati raccolti a partire dall'1 marzo 2020 fino al 27 giugno 2020 per un arco di tempo di circa 4 mesi, al ritmo di 1 richiesta al minuto dalle 7:00 alle 23:59 di ogni giorno.

4.1.1 Le problematiche

Sebbene questo studio fosse stato pensato per un confronto in condizioni di traffico e congestione stradale abituali che caratterizzano Milano, a influenzare questo studio è stata da subito l'emergenza epidemiologica da COVID-19. Il 22 gennaio 2020 il Governo Italiano ha dichiarato lo stato di emergenza, mettendo in atto le prime misure di contenimento in alcuni Comuni dove si sono verificati i primi casi del contagio. A partire dal 23 febbraio sono stati emanati DPCM (decreto del Presidente del Consiglio dei ministri) sempre più stringenti riguardo la circolazione e lo svolgimento delle attività commerciali, fino a ordinare un lockdown nazionale l'11 marzo 2020 [11]. Nonostante lo studio sia stato fortemente influenzato da questo fattore, per via della circolazione ridotta se non totalmente assente durante il lockdown e per la riduzione del numero dei mezzi pubblici ATM, si è deciso comunque di procedere, tenendo conto dell'avvenuto sul giudizio finale dei risultati. A ostacolare lo studio, inoltre, nei primi giorni di utilizzo del programma per eseguire le richieste, è stato il servizio precedentemente utilizzato per i mezzi pubblici, ovvero Moovit, che ha risposto alle richieste con valori nulli o totalmente fuori scala. Per questo motivo è stato sostituito il servizio con quello offerto da Here e sono stati scartati i risultati precedentemente raccolti.

	1 marzo - 12 marzo	4 maggio - 27 giugno
Coerenti	1988	49560
Errati	7962	18290
Totale	9680	67850
% Errati	79.5	27.0

Tabella 4.1: Dati errati che contengono uno o più valori nulli nelle stime.

4.1.2 Filtro dati errati

I dati raccolti dalla prima esecuzione del programma, avvenuta l'1 marzo 2020, fino all'entrata in vigore dello stato di lockdown nazionale, il 12 marzo 2020, sono risultati per la maggior parte errati, ovvero con stime di percorrenza nulle, causati dai problemi con lo scraper di Moovit. Successivamente a questo evento si è cercato e trovato un servizio alternativo che sostituisse la stima del percorso coi mezzi pubblici, ovvero Here, che sfortunatamente non offre stime in tempo reale ma solo statiche. I dati raccolti nel periodo di lockdown dal 13 marzo 2020 al 3 maggio 2020 compresi sono stati scartati per via della circolazione dei mezzi pubblici e privati quasi totalmente assente, e quindi poco utili nell'obiettivo finale dello studio di confrontare i mezzi di trasporto in situazioni di traffico caratteristico della città. Per questo motivo si è scelto di usare solamente i dati a partire dall'allentamento delle misure di restrizioni da emergenza sanitaria, ovvero quelli a partire dal 4 maggio 2020. Anche i dati presi in considerazione da quest'ultima data contengono delle stime nulle. Questo fenomeno è stato probabilmente dovuto alla generazione a random delle coordinate di partenza e arrivo, che ha permesso richieste in qualsiasi punto della mappa, compresi punti non situati direttamente sulla strada come in giardini pubblici, condomini e altre zone private, o semplicemente dovuto a disservizi. Per un confronto alla pari sono state considerate solo le righe prive di zeri nel file formato CSV. Come indicato dalla Tabella 4.1 su circa 68000 richieste, 50000 sono risultate prive di errori, circa il 73% del totale, e quindi adatte per un confronto 1 a 1 tra mezzi di trasporto.

4.1.3 Salvataggio

Nella Tabella 4.2 è riportato l'header del file CSV in cui sono stati salvati i dati. Nei campi di partenza e arrivo sono state salvate le coordinate espresse in gradi del tragitto generato, nei restanti campi le stime di percorrenza di tale tragitto espresse in minuti, insieme alla data e all'orario in cui è stata effettuata la richiesta di tale tratta. Oltre ai dati principali sono state salvate informazioni secondarie come il numero di auto libere Enjoy al momento della

richiesta, la lunghezza in via aerea della tratta e quella totale del percorso di ogni mezzo, tutte espresse in chilometri.

Data	Orario	Partenza	Arrivo	Auto	ATM	Enjoy	Bici	Piedi
------	--------	----------	--------	------	-----	-------	------	-------

Tabella 4.2: Header del file CSV con i dati salvati.

4.1.4 Distribuzione lunghezza tratte generate

Nella Figura 4.1 e nella Tabella 4.3 sono riportate le statistiche riguardo la lunghezza in via aerea delle tratte generate. Si può notare come più della metà siano lunghe meno di 4 km in linea aerea, risultato voluto e ottenuto dai vincoli imposti nella generazione riguardanti l'area del Comune di Milano e per l'ingresso e l'uscita dal centro storico al semicentro. La Figura 4.2 mostra come la lunghezza delle tratte è stata equidistribuita nelle diverse fasce di orario.

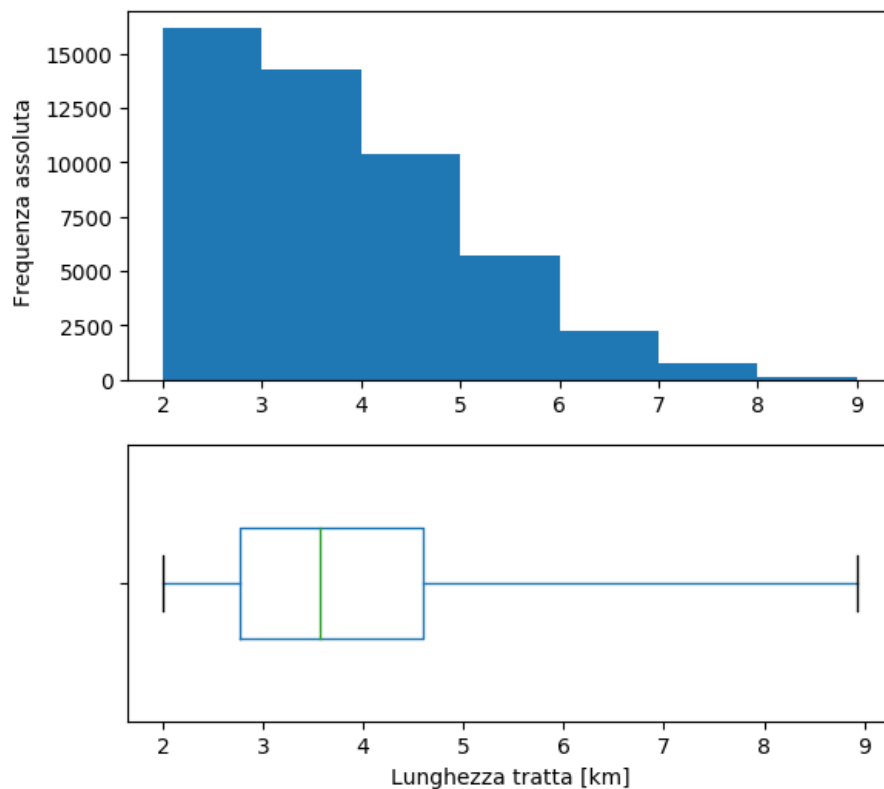


Figura 4.1: Frequenza assoluta e distribuzione lunghezza tratte [km].

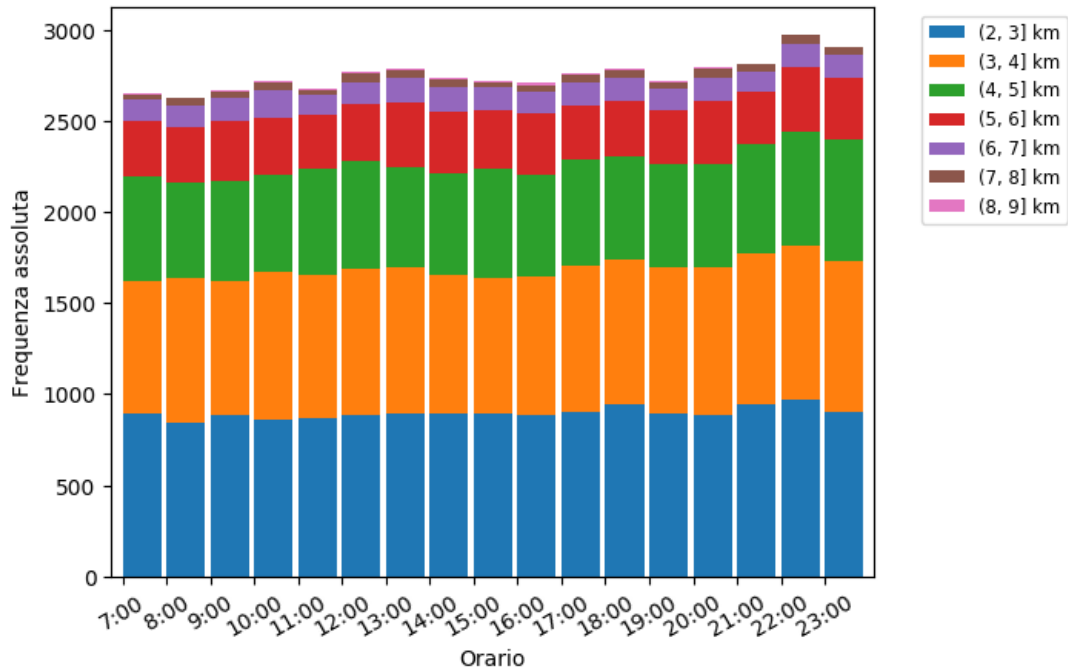


Figura 4.2: Frequenza assoluta lunghezza tratte [km] in base all'orario.

Abs. freq.	49560
Media	3.80 km
Mediana	3.57 km
Std	1.27 km
Min	2.00 km
Max	9.52 km

Tabella 4.3: Statistiche lunghezza tratta [km].

4.2 Performance dei singoli mezzi

Dalla Figura 4.3 si può subito notare come i tragitti a piedi e in bicicletta abbiano dei baffi molto corti se non totalmente assenti rispetto agli altri mezzi, con conseguente varianza molto bassa, il che è un risultato aspettato dato che non c'è nessun fattore oltre la distanza a influenzare drasticamente il tempo di percorrenza per questi metodi di spostamento, motivo per cui i provider si limitano a calcolarne una stima statica.

	Auto	Enjoy	Bici	ATM	Piedi
Media	23.7	16.4	11.9	10.3	4.5
Mediana	23.6	16.2	11.9	9.8	4.5
Std	3.9	4.2	1.3	3.0	0.0
Min	8.9	3.5	6.5	3.7	4.3
Max	59.0	46.4	16.9	41.0	4.6

Tabella 4.4: Statistiche velocità media [km/h].

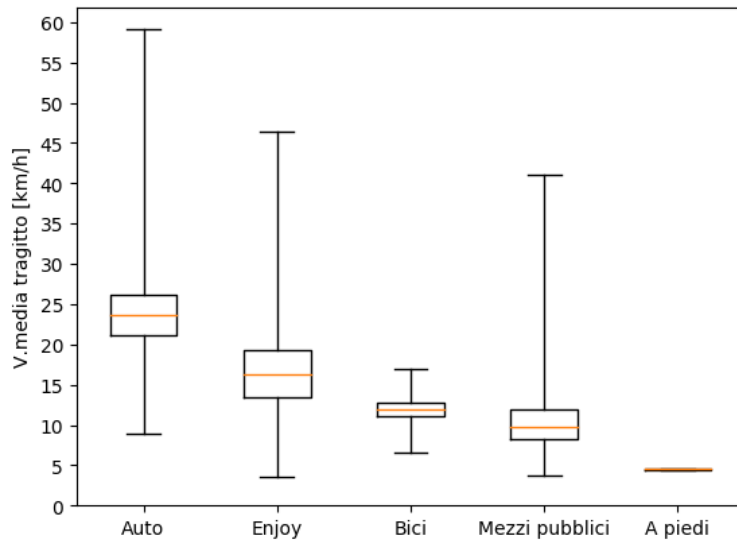


Figura 4.3: Distribuzione velocità media dei singoli mezzi.

Risultato opposto invece è stato ottenuto per i tragitti in auto, car sharing e mezzi pubblici, che mostrano una varianza molto più alta, come riportato nella Tabella 4.4, dovuta alla presenza di diverse variabili nel calcolo del tragitto. Nonostante le stime di percorrenza dei mezzi pubblici siano state calcolate dal provider usando solamente le tabelle degli orari dei mezzi, la varianza è risultata comunque alta per via delle numerosi variabili che influenzano sul tempo di percorrenza, come la distanza tra il punto di partenza e la fermata del mezzo, la scelta stessa del tipo di mezzo, come bus, tram o metro, e la corrispettiva velocità nominale. Questi primi risultati sono in linea con quelli dell'ISFORT riportati nella Tabella 2.1, che vede l'automobile in testa alla classifica intorno a 22 km/h di media, con la bici al secondo posto ma con 15 km/h e i mezzi pubblici al terzo con 14 km/h, stime poco più alte di quelle sopra riportate.

4.2.1 Automobile

Velocità media

Una delle prime analisi effettuate per ogni mezzo è stata quella di calcolare la velocità media per ogni tragitto effettuato riguardante l'intero periodo di raccolta dati, con l'obiettivo di osservare eventuali variazioni di ora in ora. Come riportato nella Figura 4.4, si evidenziano due situazioni, caratterizzate da una differenza nella velocità media: più bassa nelle fasce di orario delle 8:00-11:00 e delle 17:00-19:00, in cui è stata registrata la velocità media minima di 23 km/h, e più alta nelle ore restanti. La massima velocità media invece è stata rilevata nei tragitti dopo le 22:00, di circa 26 km/h. Tali risultati sono in linea, tra i tanti studi a riguardo del traffico in auto a Milano, con quelli dello studio di TomTom¹ effettuato sulla città di Milano nel 2019, che evidenzia le 9:00 e le 18:00 come orari di picco del traffico stradale dal lunedì al venerdì, con un livello di congestione rispettivamente del 70% e del 60%.

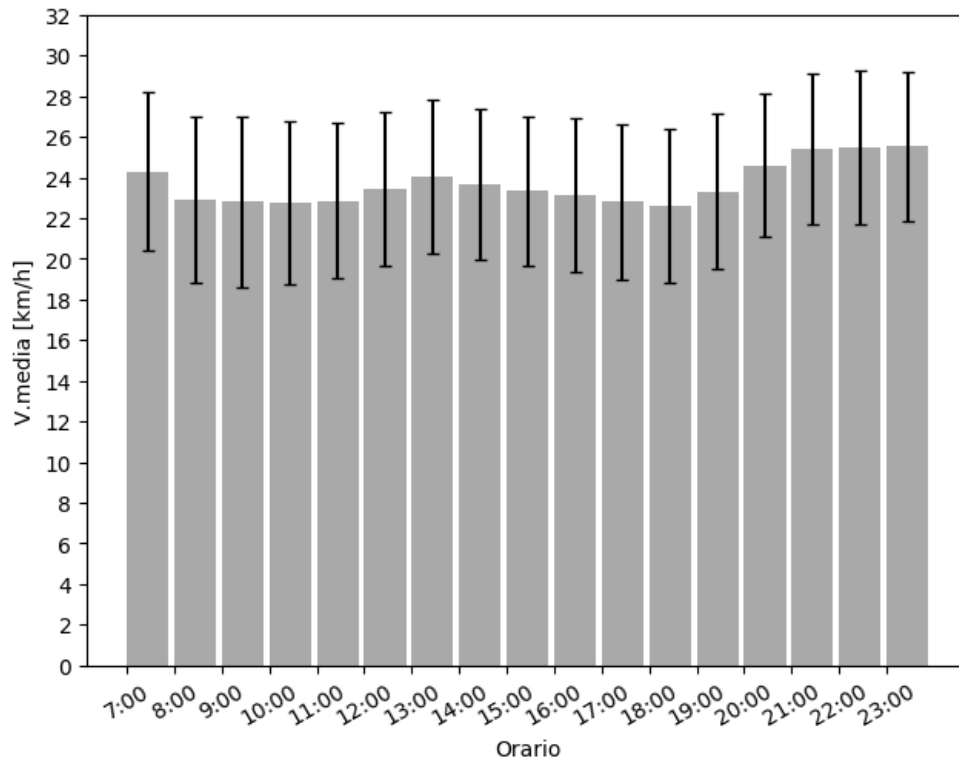


Figura 4.4: Velocità media in auto [km/h] di ora in ora. I baffi neri rappresentano la deviazione standard campionaria σ rispettiva di ogni barra.

¹https://www.tomtom.com/en_gb/traffic-index/ranking/. Accessed: 20/01/2021

Risultano in linea con lo studio anche le ore precedenti e successive agli orari individuati come picchi. Si nota inoltre come le deviazioni standard campionarie di ogni ora, rappresentate da baffi neri incastonati nelle barre, rimangano invariate. Si è scelto di rappresentare la deviazione standard invece che l'errore standard per via dell'elevata numerosità del campione che renderebbe i baffi non visibili e quindi non distinguibili. Difatti, basandosi sui dati della Tabella 4.4 che riporta una deviazione standard di 3.9 e della Figura 4.2 che mostra una media di circa 2600 tratte per ogni fascia di orario, l'errore standard per ogni barra risulterebbe di: $\frac{\sigma_x}{\sqrt{n}} = \frac{3.9}{\sqrt{2600}} = 0.07$.

Velocità media lunedì-venerdì e sabato-domenica

La Figura 4.5 mostra il risultato di una ripartizione dei dati effettuata sulla base del giorno della settimana, in particolare sono stati divisi i tragitti effettuati dal lunedì al venerdì da quelli del sabato e domenica. Per rendere leggibile la differenza è stata spostata l'origine del grafico al valore 16 sull'asse delle ordinate.

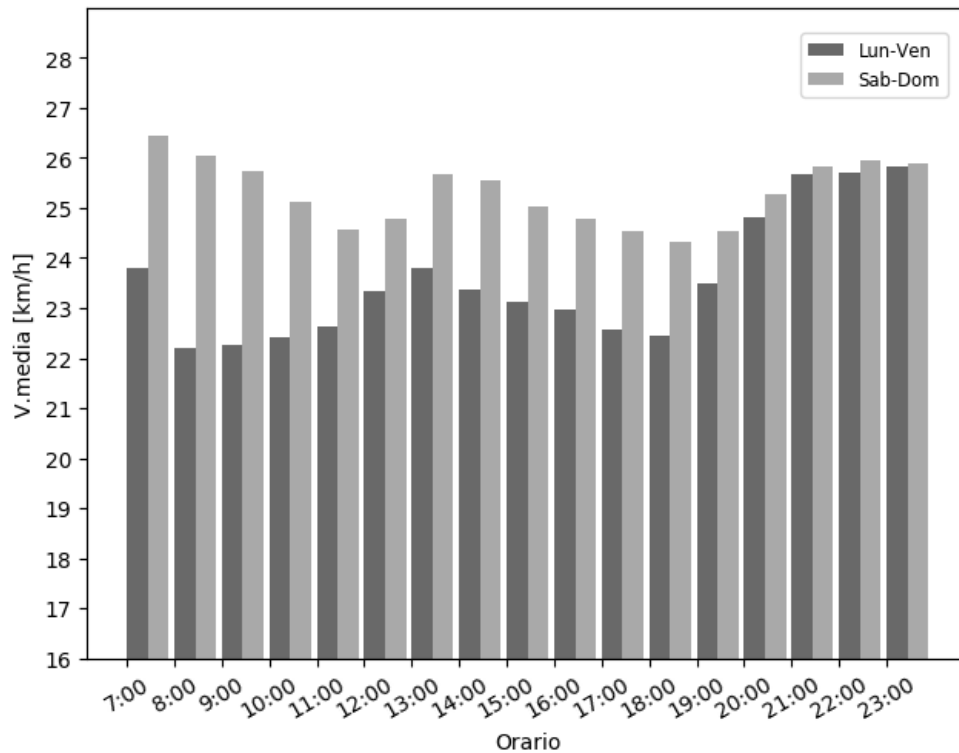


Figura 4.5: Velocità media in auto [km/h] di ora in ora. I valori sull'asse delle ordinate variano tra 16 e 28.

La differenza è nel complesso bassa, di circa 2 km/h di media, che tende a diminuire nel primo pomeriggio e negli orari notturni e ad aumentare nella fascia oraria delle ore 8:00 e delle 16:00, con una differenza di quasi 4 km/h, equivalente a un incremento della velocità media nel fine settimana del 14.6% in più rispetto al lunedì-venerdì. La variazione della velocità media di ora in ora nel sabato-domenica risulta meno evidente di quella del lunedì-venerdì, con i picchi di rallentamento che si spostano nelle fasce orarie 10:00-12:00 e 16:00-19:00. Anche questi dati risultano in linea con quelli dello studio di TomTom, che vede un minor livello di congestione nel weekend con dei picchi nelle ore 10:00 e 18:00. La deviazione standard delle fasce di orario del lunedì-venerdì e sabato-domenica, che sono state omesse per motivi di leggibilità del grafico, rimangono sempre costanti intorno al valore 3.7 km/h, con un errore standard di $\frac{3.7}{\sqrt{1622}} = 0.09$ per i giorni dal lunedì al venerdì e $\frac{3.7}{\sqrt{622}} = 0.14$ per il fine settimana, dove 1622 e 622 rappresentano rispettivamente la numerosità del campione di una sola fascia di orario lunedì-venerdì e di sabato-domenica.

Velocità media settimana dopo settimana dal fine lockdown

La Figura 4.6 mostra il risultato di una ripartizione dei dati effettuata in base alla settimana, in particolare sono stati partizionati a gruppi di 2 settimane consecutive i tragitti effettuati a partire dal 4 maggio 2020, primo giorno dell'allentamento delle restrizioni imposte dal governo italiano per l'emergenza COVID-19 [11]. Nel grafico risulta evidente una degradazione della velocità media generale e lineare rispetto al passare delle settimane. Si nota inoltre che le curve corrispondenti alle settimane successive al 17 maggio, ovvero dopo le prime 2, presentano dei flessi sempre più accentuati in prossimità degli orari di picco del traffico evidenziati dalla Figura 4.4. Anche in questo caso, non sono state riportate differenze nella deviazione standard. Nella Figura 4.7 sono stati usati i box plot per descrivere la distribuzione delle velocità medie a gruppi di 2 settimane dalla fine del lockdown. Si può notare come il quartile zero, primo, secondo e terzo siano rimasti uguali, con un lieve andamento a ribasso col passare delle settimane. La differenza risulta più accentuata nel quarto quartile, che vede la diminuzione dei valori di massima velocità media per tragitto, probabilmente dovuta a un aumento del traffico, come emerso dal grafico della Figura 4.6 con valori decrescenti della velocità media per ogni fascia di orario al passare delle settimane dalla fine del lockdown.

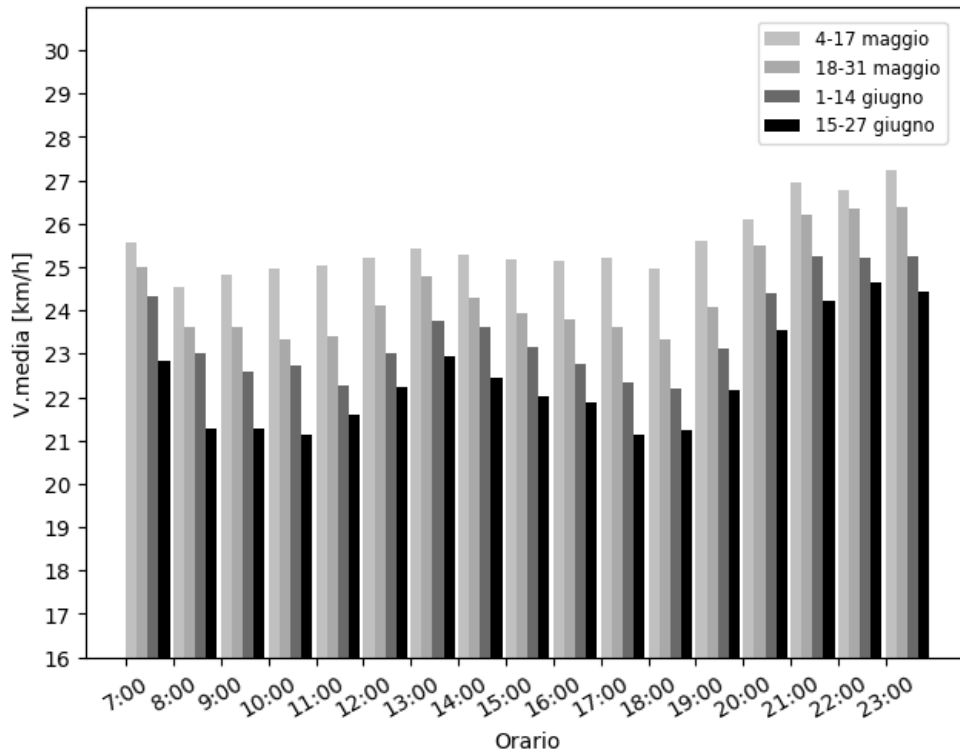


Figura 4.6: Velocità media in auto [km/h] di ora in ora. I valori sull'asse delle ordinate variano tra 16 e 30.

Velocità media in base alla lunghezza del tragitto

Nella Figura 4.8 sono riportati, tramite una heatmap, i dati della velocità media divisi per fasce d'orario e per lunghezza della tratta. Si evidenziano due gobbe in corrispondenza degli orari di picco del traffico intorno alle ore 9:00 e 18:00, e le tratte brevi, lunghe dai 2 ai 3 chilometri, risultano le più colpite. Si evidenzia inoltre un'area di forma rettangolare di colori chiari a partire dalle ore 20:00, dove la velocità media si stabilizza intorno ai 27 km/h per le tratte superiori ai 3 km.

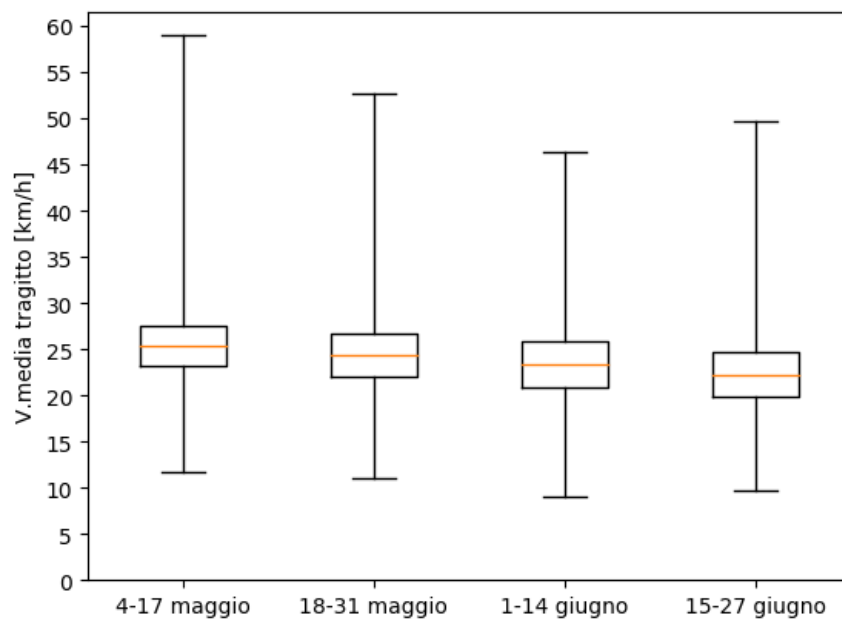


Figura 4.7: Distribuzione della velocità media in auto [km/h] calcolata a gruppi di 2 settimane dalla fine del lockdown.

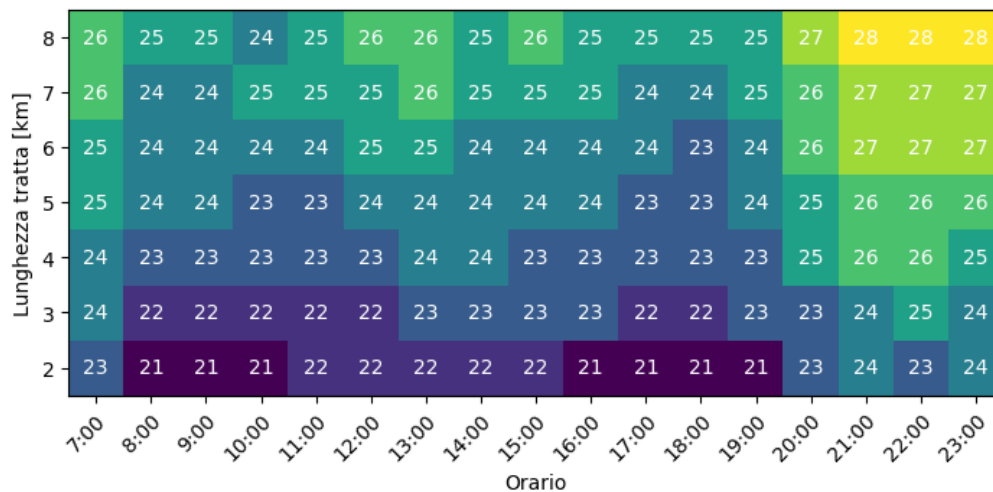


Figura 4.8: Velocità media in auto [km/h] di ora in ora in base alla lunghezza della tratta.

4.2.2 Enjoy

Velocità media

I tragitti col servizio di car sharing Enjoy hanno subito una lieve variazione di velocità media nell'arco 8:00-11:00 e in quello delle 17:00-19:00 in cui si è registrata la velocità media minima di 16 km/h e di 17 km/h di massima. La variazione è all'atto pratico poco significativa. La deviazione standard rimane costante per tutte le fasce di orario. Rispetto al grafico della Figura 4.4 della velocità media in auto, la curva risulta molto simile, traslata verticalmente di 6 km/h. Anche in questo caso è stata usata la deviazione standard invece dell'errore standard per motivi di leggibilità, data l'elevata numerosità del campione.

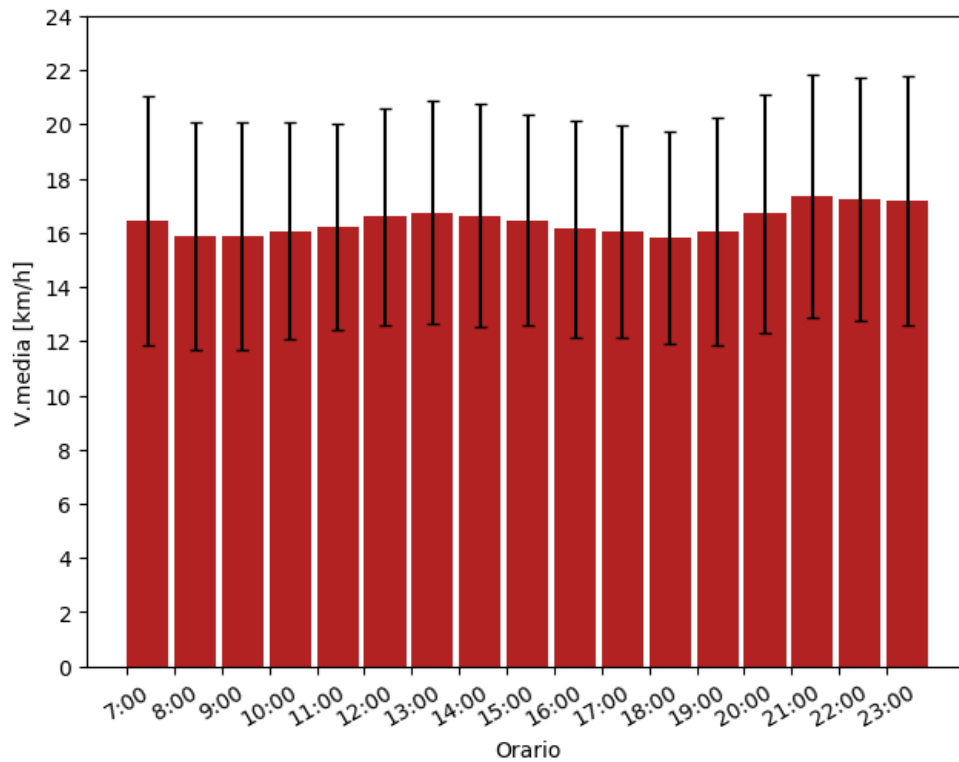


Figura 4.9: Velocità media usando il car sharing Enjoy [km/h] di ora in ora. I baffi neri rappresentano la deviazione standard campionaria σ rispettiva di ogni barra.

Velocità media lunedì-venerdì e sabato-domenica

Non sono state registrate particolari differenze effettuando la ripartizione dei dati in base ai giorni dal lunedì al venerdì e del fine settimana, come mostrato nella Figura 4.10. La velocità media rimane costante, con una lieve differenza nella fascia oraria delle 8:00. La deviazione standard rimane costante e identica a quella della Figura 4.9 per tutte le fasce di orario.

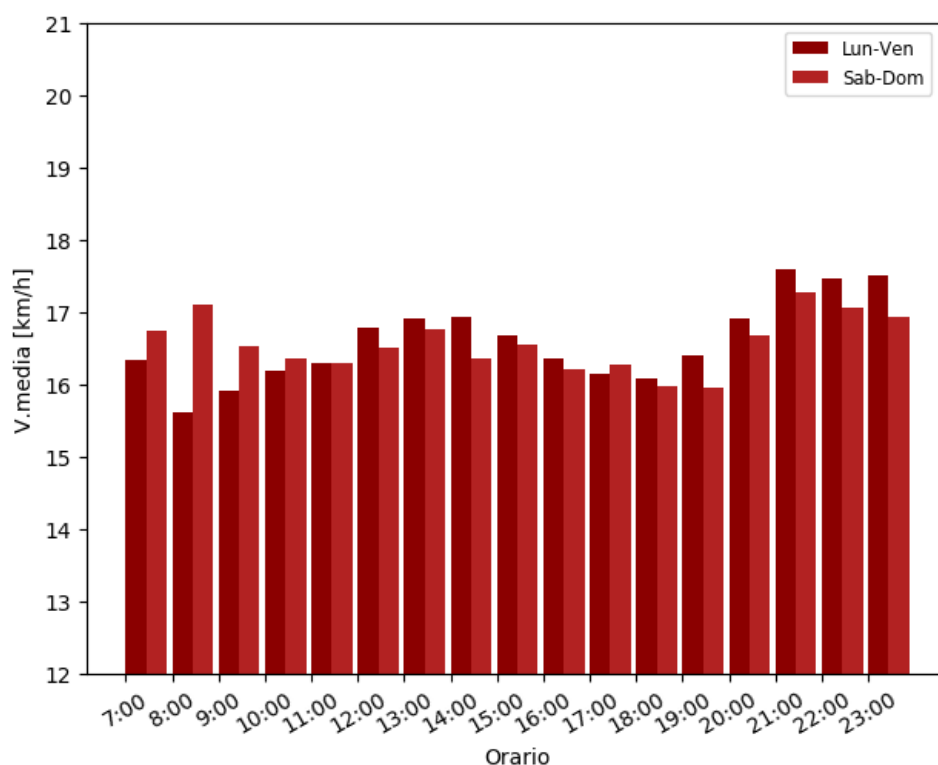


Figura 4.10: Velocità media usando il car sharing Enjoy [km/h] di ora in ora. I valori sull'asse delle ordinate variano tra 12 e 21.

Velocità media settimana dopo settimana dal fine lockdown

Anche per il car sharing risulta evidente una degradazione della velocità media generale e lineare rispetto al passare delle settimane, come riportato nel grafico della Figura 4.11. Non risultano differenze nelle fasce orarie meno trafficate dalle 12:00 alle 15:00 e dalle 21:00 alle 23:00. Le prime due settimane dall'allentamento delle restrizioni risultano più veloci di circa 1 km/h di media dalle altre settimane. Nel grafico della Figura 4.12, come per l'automobile di proprietà, il quartile zero, primo, secondo e terzo risultano uguali,

con un andamento decrescente del baffo del quarto quartile entro cui ricadono i valori massimi di velocità media per tragitto, sempre dovuto con molta probabilità al traffico.

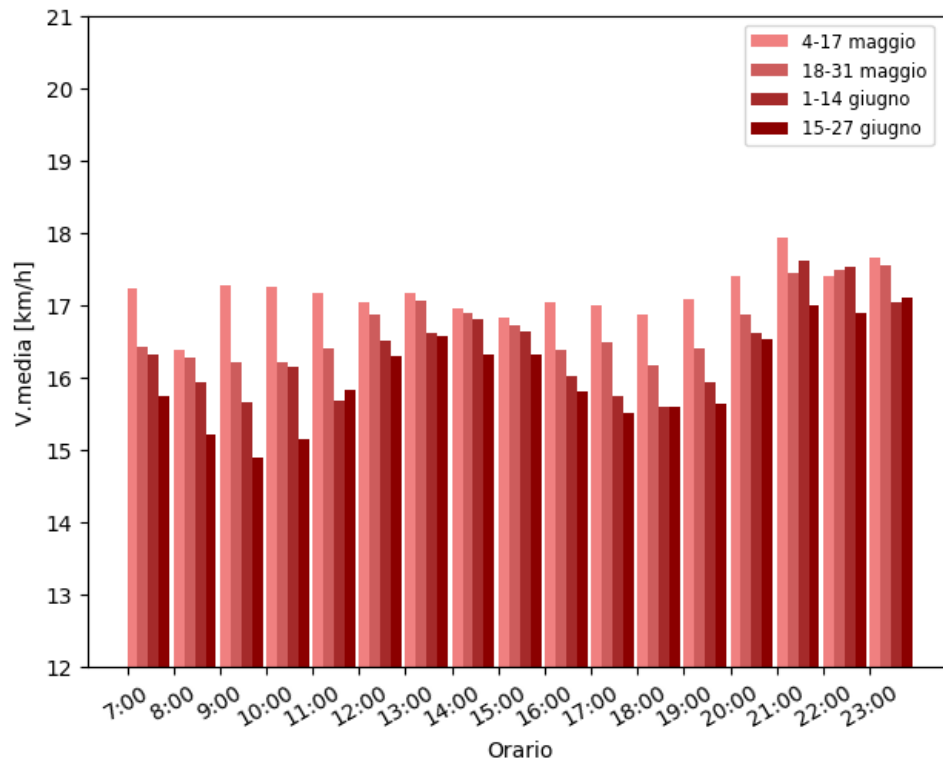


Figura 4.11: Velocità media in car sharing Enjoy [km/h] di ora in ora. I valori sull'asse delle ordinate variano tra 12 e 21.

Velocità media in base alla lunghezza del tragitto

Nel grafico della Figura 4.13 sono riportati, tramite una heatmap, i dati della velocità media divisi per fasce d'orario e per lunghezza della tratta. In questo caso, a differenza della heatmap dell'auto di proprietà, corrispondente alla Figura 4.8, le velocità medie risultano mediamente costanti in base alla lunghezza della tratta, indipendentemente dall'orario, con lievi diminuzioni nelle ore intorno alle 9:00 e 18:00. A differenza dell'auto, la velocità media sembra non aumentare in corrispondenza degli orari notturni, questo probabilmente è dovuto al grande impatto che ha il tratto a piedi per raggiungere un mezzo libero della flotta.

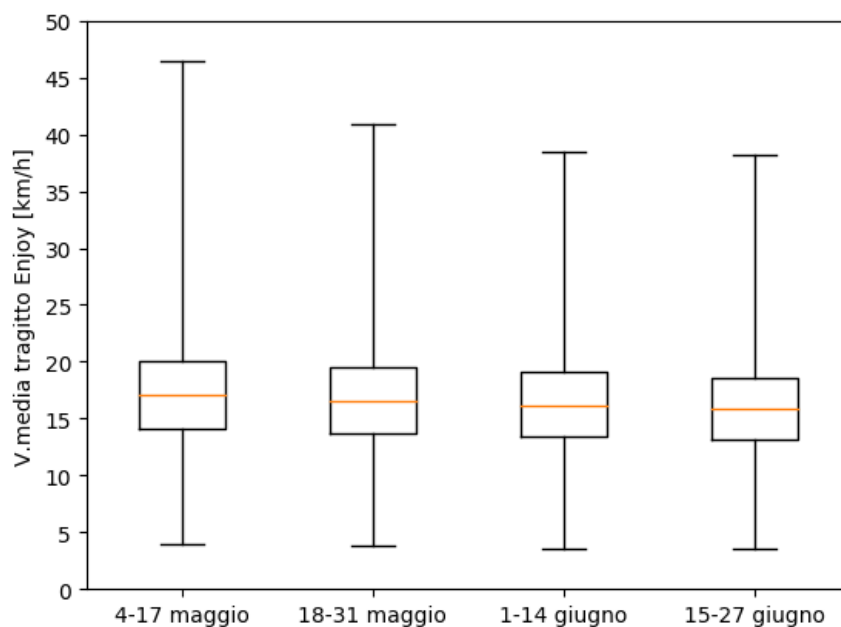


Figura 4.12: Distribuzione della velocità media in car sharing Enjoy [km/h] calcolata a gruppi di 2 settimane dalla fine del lockdown.

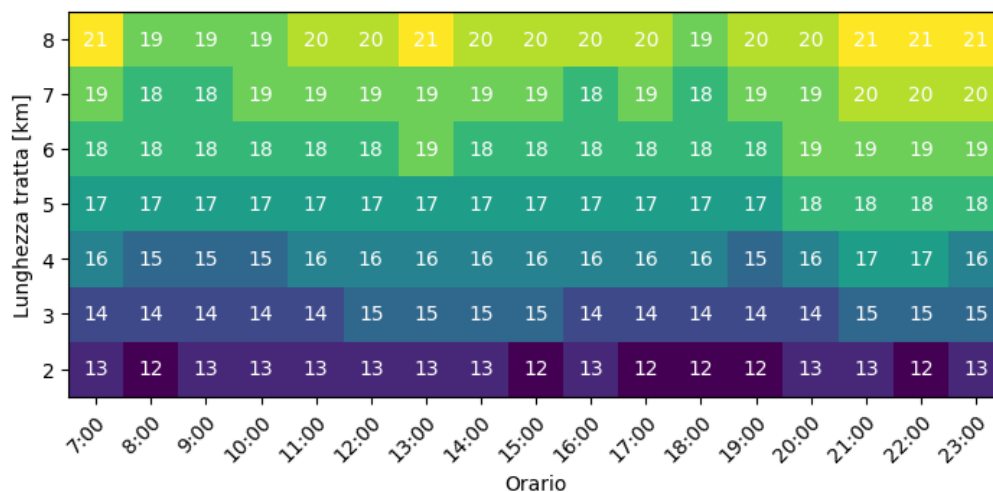


Figura 4.13: Velocità media in car sharing Enjoy [km/h] di ora in ora in base alla lunghezza della tratta.

Auto libere e tempo medio per raggiungerle

Sottraendo il tempo impiegato in auto da quello impiegato col car sharing Enjoy è stata ottenuta una stima indicativa del tempo medio per raggiungere un'auto libera a piedi. Dai calcoli è risultato che il tempo medio oscilla intorno ai 6 minuti, come riportato dalla Tabella 4.5. Il box plot della Figura 4.14 evidenzia il range interquartile nell'intervallo dai 3 ai 9 minuti, con una mediana di 6 minuti. Anche per questa analisi i dati sono stati partizionati per i giorni da lunedì a venerdì e per sabato e domenica. Il grafico della Figura 4.15 mostra una lieve differenza nel tempo medio per raggiungere un'auto libera Enjoy di circa 1 minuto a sfavore dei giorni del fine settimana. Dunque, in settimana, le auto sono risultate più facilmente raggiungibili.

Per contestualizzare il tempo medio per raggiungere un'auto Enjoy è stato usato il dato sul numero delle auto libere salvato per ogni richiesta ai servizi. Il massimo numero di auto libere in circolazione è stato di 870. Nella Figura 4.16 viene mostrata in media la variazione di questo conteggio lungo l'arco della giornata, ripartito per lunedì-venerdì e sabato-domenica. Si può notare come il picco di utilizzo del servizio, denotato da un calo delle auto libere a disposizione, inizia verso le 15:00 e finisce verso le 21:00 indipendentemente dal giorno. L'unica differenza tra i due gruppi si può notare nelle ore del mattino, che vede meno auto libere durante la settimana. Il grafico della Figura 4.17 mostra chiaramente come il numero delle auto libere a disposizione sia aumentato con la fine del lockdown. Si può notare infatti come siano state immesse circa 150 nuove auto nell'arco di un mese e altre 50 nel mese successivo. Questo dato è stato probabilmente dovuto al ritiro di una parte del parco auto da parte delle aziende per motivi di manutenzione, vista l'occasione della scarsa domanda di servizio.

T.medio ragg.auto	
Media	6.6
Mediana	6.0
Std	4.6
Min	0.0
Max	43.0

Tabella 4.5: Statistiche tempo medio per raggiungere auto libera [min].

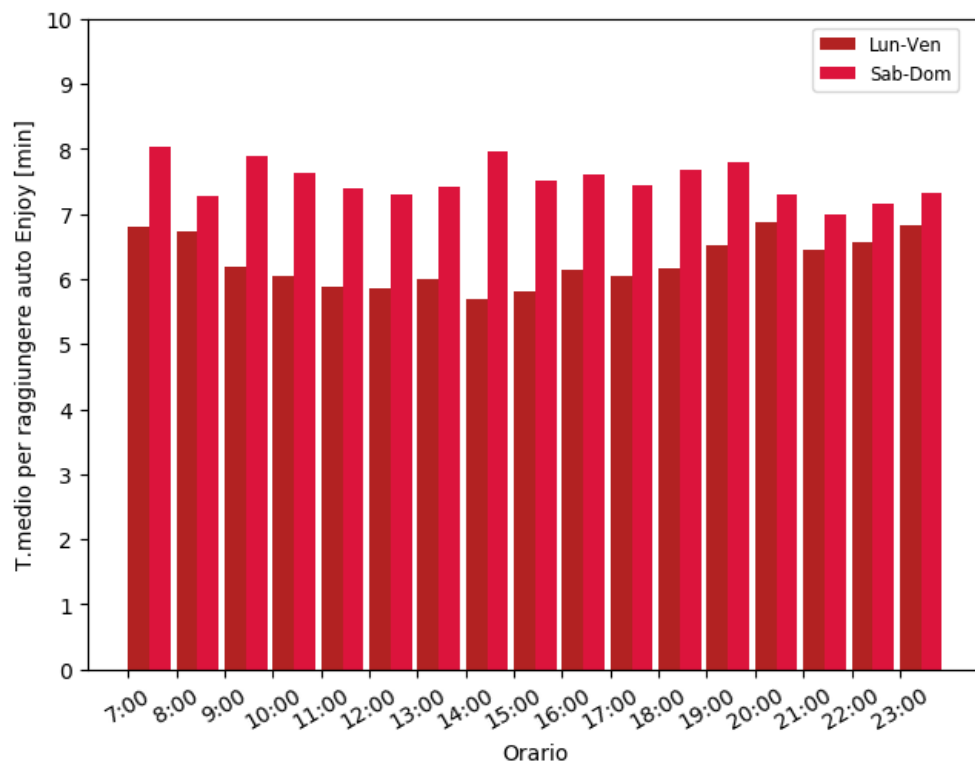


Figura 4.15: Tempo medio per raggiungere auto libera [min].

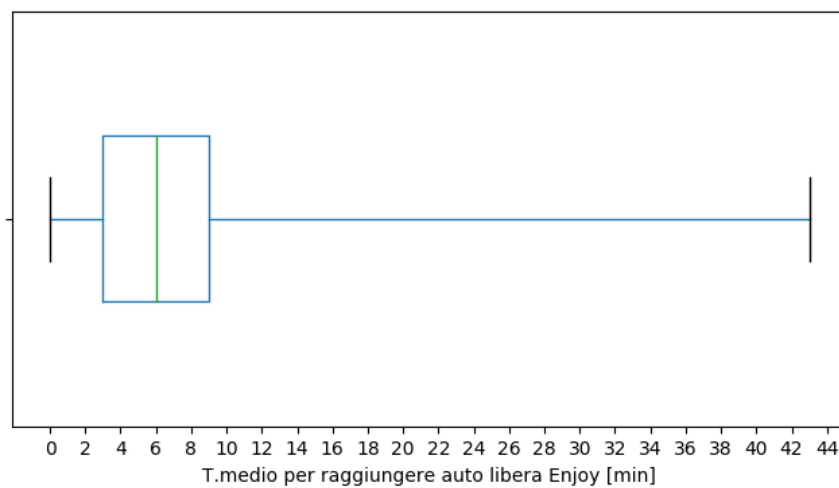


Figura 4.14: Distribuzione del tempo impiegato a raggiungere auto libera [min].

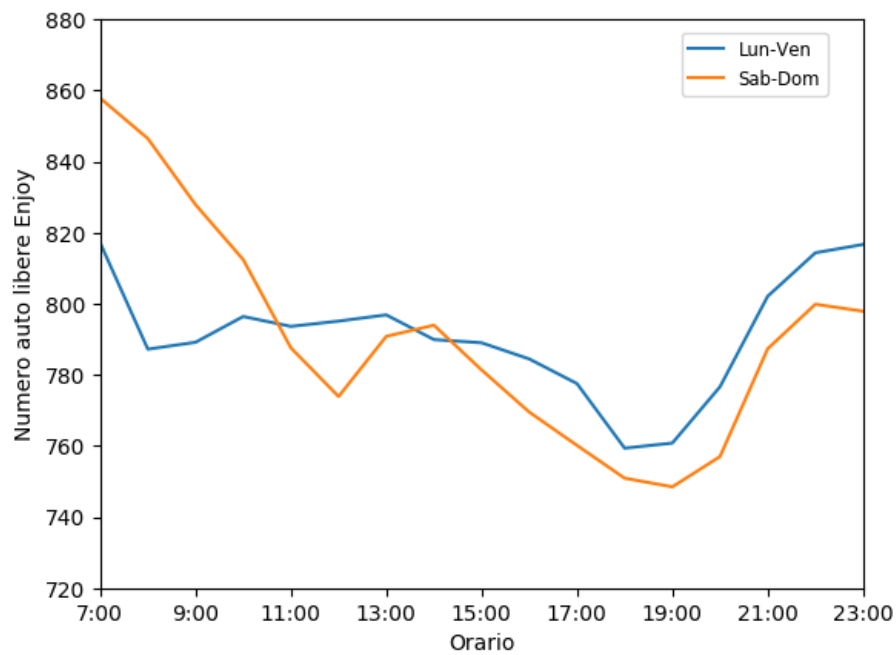


Figura 4.16: Numero di auto libere Enjoy lungo l'arco della giornata.

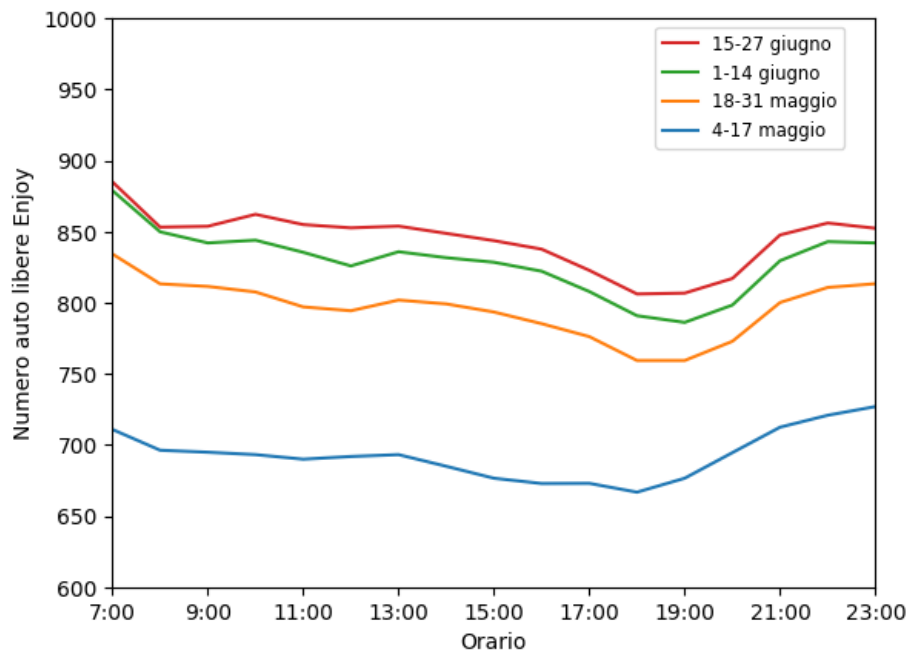


Figura 4.17: Numero di auto libere Enjoy settimana dopo settimana.

4.2.3 Mezzi pubblici ATM, bici e a piedi

Analizzando le stime di percorrenza riguardo i mezzi pubblici, la bicicletta e i percorsi a piedi, non si è verificata alcuna variazione significativa lungo l'arco della giornata, tradotto graficamente come nelle Figure 4.18, 4.20 e 4.19 in una retta parallela all'asse x per ognuno dei mezzi considerati. Il risultato è aspettato, dato che tali servizi non hanno fornito risposte dinamiche in base alle condizioni attuali del traffico e dei dati storici, e, nel caso della bicicletta e a piedi, perché l'influenza è praticamente nulla per questi mezzi di trasporto.

Si è deciso di analizzare ulteriormente le performance dei mezzi pubblici per via della loro grande varianza, al pari di quella dell'auto e del car sharing. Nel grafico della Figura 4.21 è riportata un'analisi tramite box plot per visualizzare la distribuzione dei valori di velocità media di settimana in settimana dalla fine del lockdown. Dal risultato emerge un netto cambiamento del quarto quartile a partire dal 1 giugno 2020, che vede un accorciamento del baffo corrispondente ai valori massimi di velocità media registrati. Nel tentativo di trovare una risposta a questo cambiamento, dato che può dipendere solamente dal cambio delle tabelle degli orari dei mezzi, visto che le API Here usate per questo mezzo si basano solo su di esse, si è tentato di accedere agli archivi pubblici delle notizie riguardanti cambiamenti di servizio dal sito dell'azienda fornitrice ATM, ma al momento della scrittura questa pagina risulta inaccessibile per via di un errore interno del server².

²<https://atm.it/it/AtmNews/AtmInforma/>. Accessed: 20/01/2021

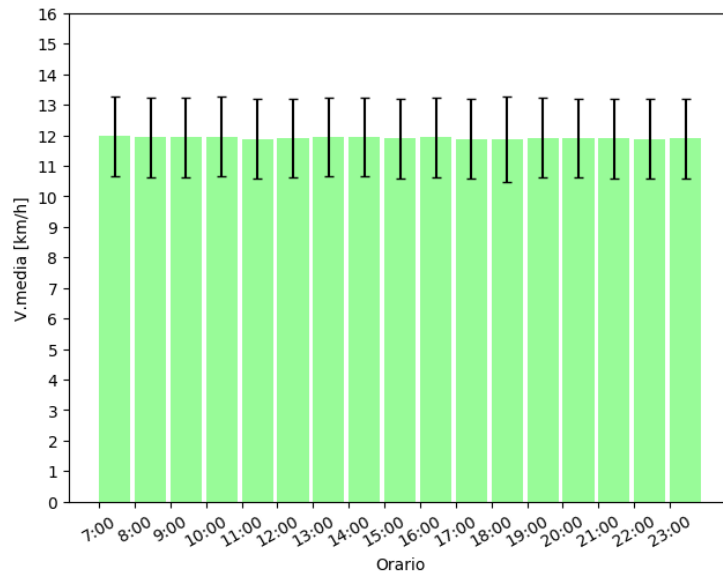


Figura 4.18: Velocità media in bici [km/h] di ora in ora. I baffi neri rappresentano la deviazione standard campionaria σ rispettiva di ogni barra.

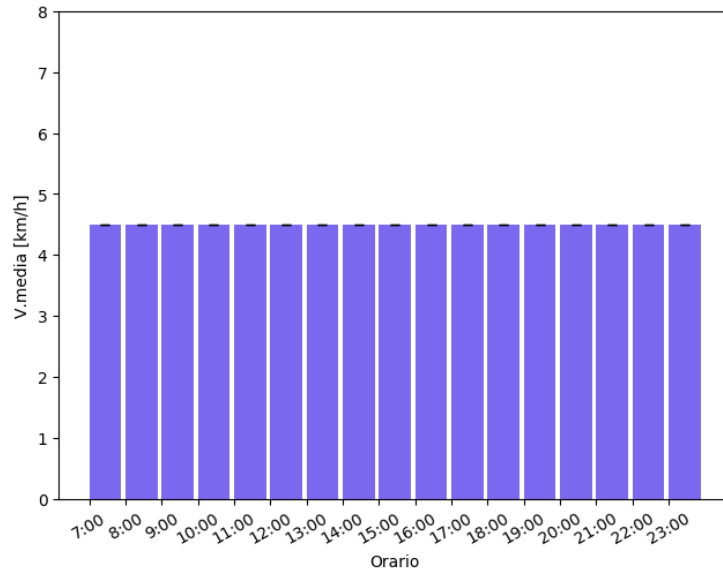


Figura 4.19: Velocità media a piedi [km/h] di ora in ora. I baffi neri rappresentano la deviazione standard campionaria σ rispettiva di ogni barra.

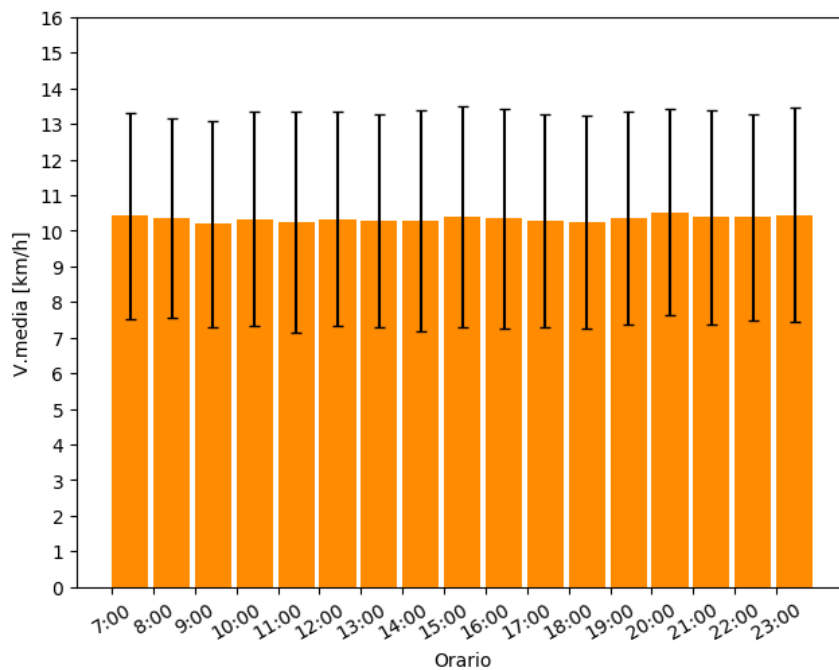


Figura 4.20: Velocità media coi mezzi pubblici [km/h] di ora in ora. I baffi neri rappresentano la deviazione standard campionaria σ rispettiva di ogni barra.

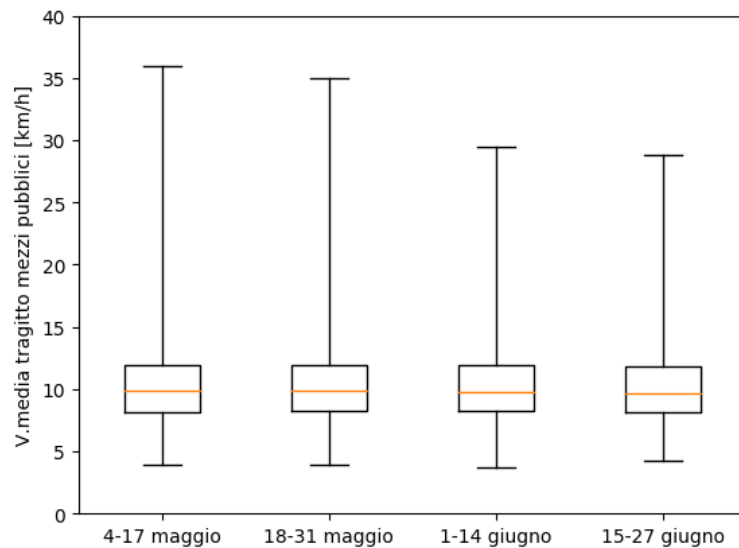


Figura 4.21: Distribuzione della velocità media coi mezzi pubblici [km/h] calcolata a gruppi di 2 settimane dalla fine del lockdown.

4.3 Confronto tra mezzi

In questo paragrafo sono stati raccolti i risultati più interessanti emersi dal confronto tra tutti i mezzi di trasporto a disposizione per spostarsi nel Comune di Milano. Il confronto più atteso che ha fatto da guida verso questo studio è stato quello di verificare se, per un uso interno al Comune di Milano a solo scopo di spostamento e senza vincoli particolari come trasporto merci o passeggeri, l'automobile privata, insieme ai suoi svantaggi, potesse essere surclassata da un mezzo più economico e pulito come i mezzi pubblici in una città dove il traffico influisce in buona parte sui tempi di percorrenza, come mostrato nella Figura 4.24.

4.3.1 Vittoria dell'automobile

Nonostante la grande influenza del traffico sui tempi di percorrenza, i tragitti in auto sono risultati sempre i più veloci di ogni sua controparte, a qualsiasi ora del giorno e su ogni distanza, con una vittoria sopra il 99% delle volte.

4.3.2 Parziale sconfitta del car sharing

Mezzi pubblici ATM vs. Enjoy

I risultati del confronto tra i mezzi pubblici ATM e il servizio di car sharing Enjoy illustrati nella Tabella 4.6 mostrano una percentuale di vittorie di circa il 9%, dove per vittoria si intende che il tempo impiegato a percorrere una tratta coi mezzi pubblici è stato minore o uguale al tempo impiegato utilizzando il car sharing. Nel grafico della Figura 4.22 viene visualizzata la distribuzione di queste vittorie di ora in ora. Si può notare come vicino alle ore di picco del traffico individuate nell'analisi delle performance dell'auto, ovvero le ore 8:00-10:00 e 17:00-20:00, si concentri la percentuale maggiore di sconfitte da parte del car sharing. Per esempio, su tutte le tratte richieste nelle ore 8:00, nell'11% delle volte i mezzi pubblici hanno pareggiato o superato la performance del car sharing. Al contrario, al di fuori degli orari di picco del traffico, la percentuale di vittorie è vicina allo 0. Dalla Tabella 4.6 si evince anche che dell'8.7% di queste vittorie, circa metà di esse sono tratte brevi dai 2 ai 5 km.

Bicicletta vs. Enjoy

Risultato ancora più interessante è quello del confronto tra bicicletta di proprietà e car sharing. La tabella 4.7 mostra una percentuale delle vittorie del 36%, poco più di un terzo del totale delle tratte. La maggior parte di queste

vittorie è concentrata nelle tratte brevi dai 2 ai 5 km. Anche in questo confronto è emerso che negli orari di picco si accentua la percentuale di vittorie che arriva a toccare quasi il 45% del totale nelle ore 8:00 e 18:00 come riportato dal grafico della Figura 4.23. Al contrario dei mezzi pubblici, questa percentuale resta alta anche al di fuori degli orari di punta, restando sempre intorno al 30%.

	Abs. freq.	% win
(2, 5] km	2072	4.2
(5, 7] km	1512	3.1
(7, 10] km	670	1.4
totale	4294	8.7

Tabella 4.6: Vittoria dei mezzi pubblici su Enjoy per lunghezza tratta.

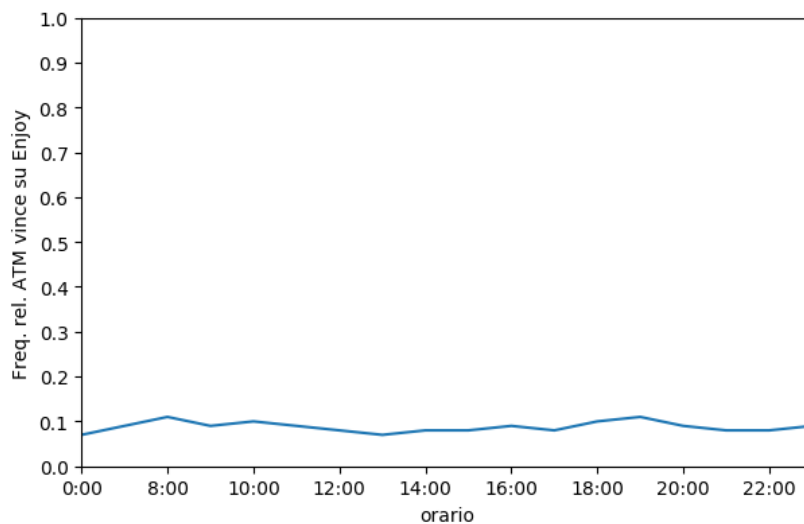


Figura 4.22: Frequenza relativa vittorie dei mezzi pubblici su Enjoy di ora in ora.

	Abs. freq.	% win
(2, 5] km	15215	30.7
(5, 7] km	2423	4.9
(7, 10] km	278	0.6
totale	17917	36.2

Tabella 4.7: Vittoria della bicicletta su Enjoy per lunghezza tratta.

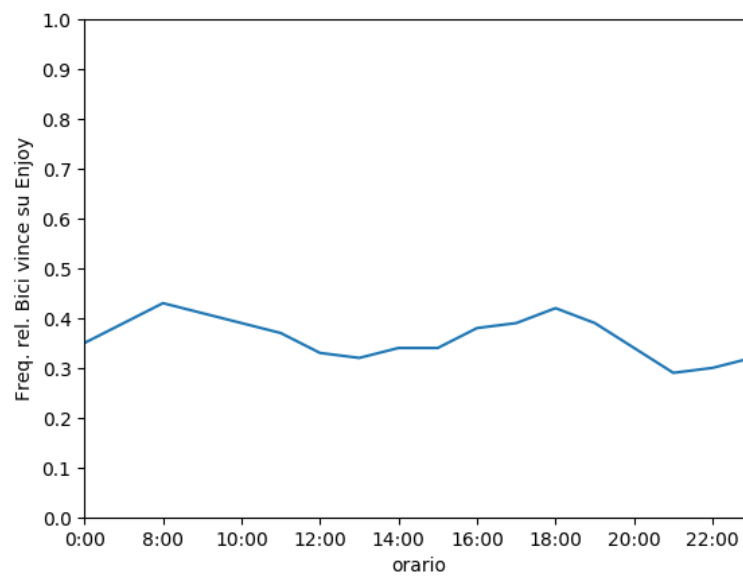


Figura 4.23: Frequenza relativa vittorie della bicicletta su Enjoy di ora in ora.

4.4 Confronto pre e post lockdown

Nonostante le prime versioni del programma di raccolta dati abbiano portato a diversi errori nelle stime, quelle relative al percorso in automobile e car sharing sono sempre state corrette. Si è deciso quindi di analizzare l'unica settimana a disposizione prima del lockdown, dal lunedì 2 a domenica 8 marzo 2020, con una delle settimane più trafficate dopo il lockdown, che dalle analisi riportate dalle Figure 4.6 e 4.11 sono risultate corrispondenti al periodo dal 15 al 27 giugno 2020.

4.4.1 Automobile

Nel grafico della Figura 4.24 si può notare una lieve differenza di km/h di media tra una settimana pre e post lockdown dal lunedì al venerdì. Questa differenza risulta più evidente negli orari di picco del traffico vicino alle 8:00 e alle 18:00, con una differenza di 2 km/h di media. Risulta evidente inoltre la differenza di velocità media nel pre lockdown tra le ore 8:00, dove si tocca il minimo di 17.5 km/h, e le ore 23:00, dove si trova il massimo di 24.5 km/h, una differenza di velocità del 28%.

4.4.2 Enjoy

I tempi di percorrenza in car sharing Enjoy sono risultati molto simili, come mostrato dal grafico della Figura 4.25. La maggior differenza è presente anche questa volta vicino le ore di picco delle 8:00 e delle 18:00, sebbene la differenza sia minima, di media 1 km/h.

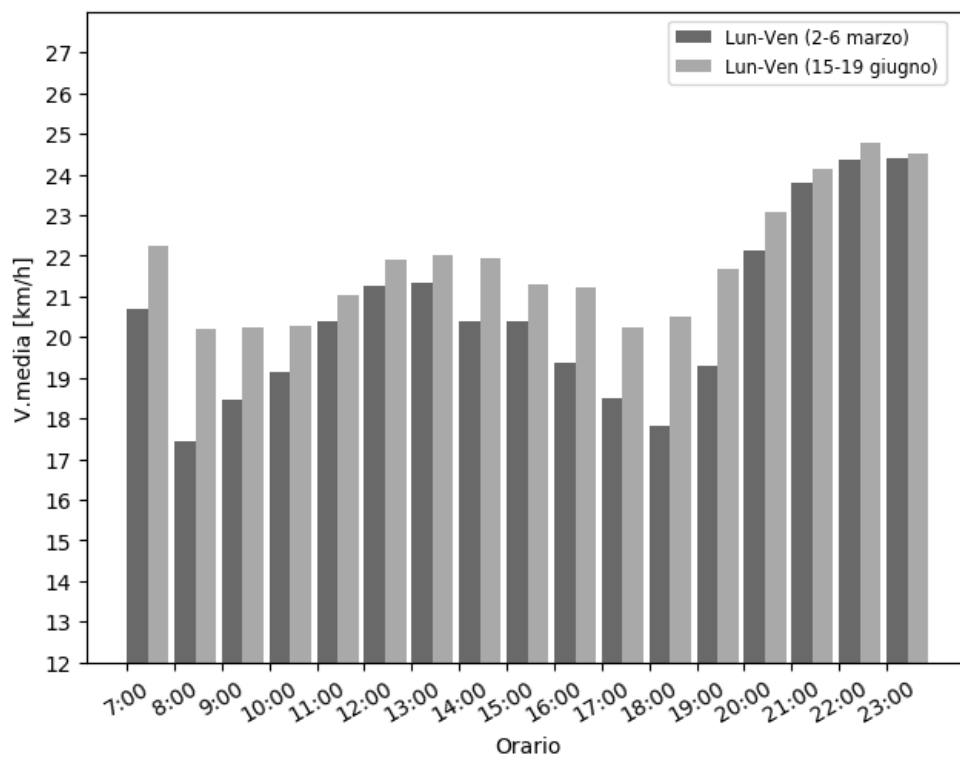


Figura 4.24: Velocità media in auto [km/h] di ora in ora pre e post lockdown. I valori sull'asse delle ordinate variano tra 12 e 27.

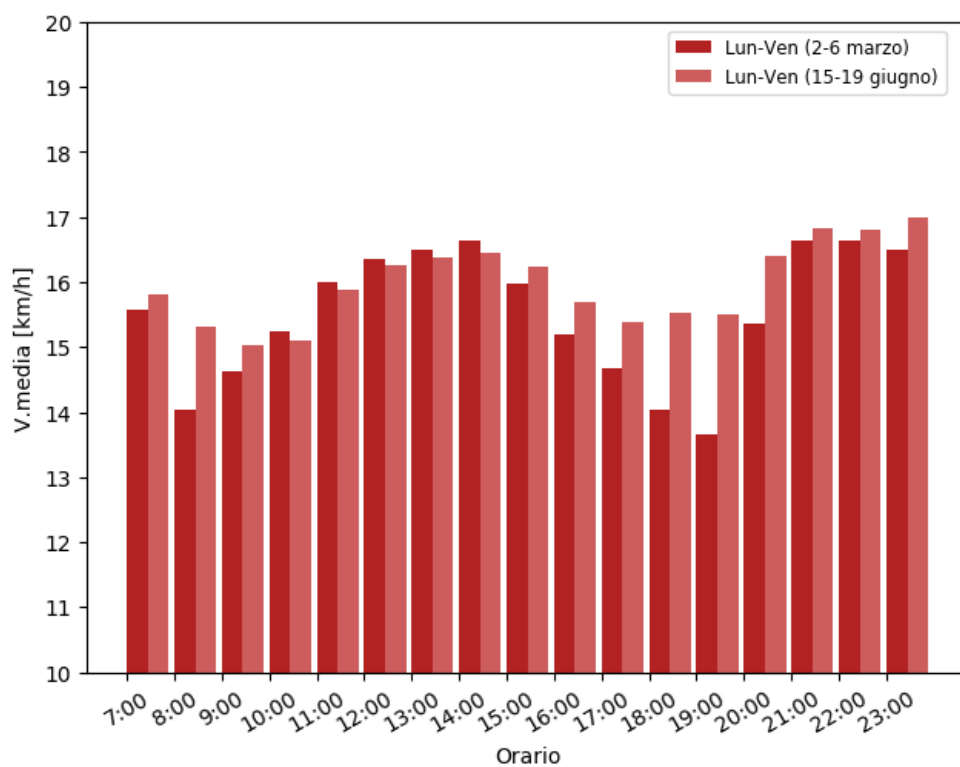


Figura 4.25: Velocità media in Enjoy [km/h] di ora in ora pre e post lockdown. I valori sull'asse delle ordinate variano tra 10 e 20.

Capitolo 5

Conclusioni

5.1 Riflessioni sul risultato

Il risultato ottenuto dal confronto tra bicicletta e servizio di car sharing Enjoy sulla base dell'orario risulta molto interessante sotto diversi aspetti. Il primo è sicuramente la relazione inversamente proporzionale del grafico stesso, 4.23, con quello delle prestazioni del car sharing nel grafico 4.9, che vede la frequenza relativa delle vittorie della bicicletta aumentare al diminuire della velocità media in car sharing dovuta al traffico. Questa relazione cattura la grande influenza della congestione stradale sull'efficienza in termini di tempo dell'uso dell'auto, sia essa di proprietà o condivisa. L'influenza del traffico sul servizio preso in considerazione ha fatto sì che negli orari intorno alle 8:00 e alle 18:00 il medesimo percorso venisse coperto in bicicletta nello stesso tempo se non più velocemente il 45% delle volte, quasi 1 volta su 2, risultati ottenuti nella maggior parte sulle tratte brevi dai 2 ai 5 km di lunghezza. Il car sharing dunque sembra soffrire maggiormente del problema dell'efficienza, che viene amplificato dal tempo impiegato a raggiungere un'auto libera a piedi, stimato di circa 6 minuti e mezzo di media, a tal punto che una bicicletta è in grado di pareggiare le sue performance negli orari di picco del traffico.

All'atto pratico però non risulta fattibile sostituire il servizio di car sharing o dei mezzi pubblici per fare così tanti chilometri pedalando in bicicletta, dato che si tratta di un'attività fisica medio-intensa e che può risultare stancante se fatta per più di 20 minuti, oltre che a essere poco pratica in caso di maltempo e poco accessibile alla fascia meno giovane della popolazione. Tuttavia, per effettuare questo studio non si è preso in considerazione uno dei mezzi alternativi in gran diffusione in svariate città del mondo, diretto concorrente della bicicletta e delle sue varianti a pedalata assistita: il mono-

pattino elettrico. Questo mezzo infatti ha una velocità massima di 25 km/h contro i 15 km/h di una pedalata normale in bicicletta, e può essere comprato per un costo che va dai €500 in sù oppure noleggiato, in mobilità condivisa, da uno dei tanti provider come Helbiz per €0.25 al minuto previo €1.00 di sblocco¹. Al momento questo mezzo risulta il più pratico soprattutto in termini di custodia e anche di parcheggio nel caso dell'uso in condivisione, dato che, al momento della scrittura, il *Codice della Strada* prevede la sosta di questi mezzi negli appositi stalli a bordo strada, e non viene dunque limitato a stazioni apposite come nel caso del servizio di bike sharing BikeMi di ATM². Visti i suoi pregi e la sua velocità più alta rispetto alla bici, non sarebbe improbabile che tale mezzo riesca a battere il car sharing negli orari di punta più del 50% delle volte e aumentando la percentuale anche nelle ore con meno traffico.

Al momento queste sono solo conclusioni preliminari che andrebbero ulteriormente analizzate e verificate, sebbene presentino alcune somiglianze con lavori precedenti. Per esempio, il risultato ottenuto riguardo la vittoria della bicicletta sul car sharing negli orari di picco e sulle tratte brevi non è molto dissimile da quello ottenuto nello studio svolto a Sidney citato nella sezione 2.1, dove nelle tratte inferiori ai 5 km la bici ha eguagliato il tempo di percorrenza in auto nel 90% delle volte entro una penalità di 10 minuti. Si ricorda infatti che il tempo di percorrenza in car sharing è stato basato sul percorso in auto con la sola aggiunta del tempo impiegato a raggiungere un'auto libera, risultato in media di 6 minuti e mezzo, col 75esimo percentile a 9 minuti, che potrebbero essere letti come i minuti di penalità dichiarati dallo studio citato per raggiungere una discreta percentuale di vittorie. La somiglianza è ancora più accentuata coi risultati dello studio svolto a New York City sempre citato nella sezione 2.1, dove il bike sharing ha battuto una corsa in taxi il 50% delle volte durante gli orari di picco del traffico. In conclusione dunque si potrebbe confermare la presenza evidenziata dai lavori precedenti di una zona grigia composta da tratte brevi e orari di picco del traffico in cui l'automobile sembra risentirne al punto tale da favorire un mezzo alternativo più economico e pulito al costo di una piccola penalità di tempo. Si ricorda però che tali conclusioni hanno senso solamente se considerate in un contesto in cui l'automobile viene usata solamente per esigenze di spostamenti, senza vincoli legati al trasporto di merci pesanti, al trasporto di passeggeri di minore età o parzialmente invalidi e senza contare i disagi causati dal maltempo.

¹<https://helbiz.com/>. Accessed: 20/01/2021

²<https://www.bikemi.com/>. Accessed: 20/01/2021

5.2 Possibili estensioni

Tra le possibili estensioni di questo lavoro vi è sicuramente quella di analizzare più a fondo i confronti tra mezzi di trasporto andando a quantificare il distacco in termini di minuti tra la vittoria dell'automobile e i restanti mezzi per diversi periodi e fasce orarie. Inoltre, vista l'enorme quantità di dati raccolti in formato JSON che sono stati scartati al fine di prelevare solamente il tempo di percorrenza e la lunghezza del tragitto, si potrebbero decodificare dati aggiuntivi riguardanti, per esempio, il tipo di mezzo pubblico selezionato da Here per percorrere la tratta, il traffico registrato in tempo reale al momento del calcolo del percorso in automobile da Waze, o ancora il numero di segmenti che compongono l'itinerario suggerito per ogni mezzo. Risulterebbe interessante inoltre rifare tale studio usando dei servizi che forniscono dati in tempo reale per tutti i mezzi di trasporto usati in questo studio, aggiungendo anche quelli non considerati, come i taxi, i monopattini elettrici in sharing e non, le biciclette in sharing e soprattutto le moto. Si potrebbero anche mettere a confronto i costi singoli di ogni viaggio in base al mezzo di trasporto selezionato per determinare la spesa complessiva da parte dell'utente.

Avendo a disposizione un dataset di viaggi realmente compiuti, o stimati come nel caso di questo studio, risulterebbe molto utile socialmente la scrittura di un software in grado di sfruttare tale conoscenza per consigliare il miglior mezzo di trasporto per affrontare tratte quotidiane o regolarmente percorse. Per esempio, sulla base di un input contenente una tratta regolarmente percorsa, indicando il punto di partenza, il punto di arrivo insieme al giorno e l'orario in cui viene affrontata abitualmente, un possibile software potrebbe effettuare le stesse statistiche di questo studio per fornire infografiche contenenti dati quali il tempo medio per raggiungere i mezzi in car sharing dal punto di partenza indicato, il mezzo più veloce, il più economico o il migliore secondo una media pesata, lasciando all'utente la libertà di scegliere le priorità. Un software del genere infatti potrebbe evidenziare che il mezzo usato abitualmente dall'utente per effettuare determinati tragitti risulta molto inefficiente in assoluto o nei giorni e nelle fasce orarie in cui le percorre, spingendolo a valutare altre opzioni per spostarsi. Ovviamente tale software richiederebbe un affinamento dei calcoli usati per stimare diverse proprietà oltre che a un database ottimizzato per l'indicizzazione dei dati storici. Progetti simili sono già stati modellati in diversi studi, dove la stima del tempo di percorrenza viene calcolata su dati storici oppure interpolando dati storici con quelli in tempo reale forniti da utenti, e comparata con viaggi realmente effettuati, portando a ottimi risultati riguardo il livello di accuratezza della predizione ed evidenziando l'efficacia dei dati storici per le stime durante le ore di picco del traffico [6][3].

Bibliografia

- [1] Bhogeswar Borah and Dhruba K Bhattacharyya. An improved sampling-based dbscan for large spatial databases. In *International conference on intelligent sensing and information processing, 2004. proceedings of*, pages 92–96. IEEE, 2004.
- [2] Michela Cameletti. The effect of corona virus lockdown on air pollution: Evidence from the city of brescia in lombardia region (italy). *Atmospheric Environment*, 239:117794, 2020.
- [3] Steven I-Jy Chien and Chandra Mouly Kuchipudi. Dynamic travel time prediction with real-time and historic data. *Journal of transportation engineering*, 129(6):608–616, 2003.
- [4] Maria Cristina Collivignarelli, Claudio De Rose, Alessandro Abbà, Marco Baldi, Giorgio Bertanza, Roberta Pedrazzani, Sabrina Sorlini, and Marco Carnevale Miino. Analysis of lockdown for covid-19 impact on no2 in london, milan and paris: What lesson can be learnt? *Process Safety and Environmental Protection*, 146:952–960, 2021.
- [5] Edoardo Croci and Davide Rossi. Optimizing the position of bike sharing stations. the milan case. 2014.
- [6] M Deeshma and Ashish Verma. Travel time modeling for bus transport system in bangalore city. *Transportation Letters*, 7(1):47–56, 2015.
- [7] Richard B Ellison and Stephen Greaves. Travel time competitiveness of cycling in sydney, australia. *Transportation research record*, 2247(1):99–108, 2011.
- [8] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.
- [9] Eurostat. Energy, transport and environment statistics. <https://ec.europa.eu/eurostat>. Accessed: 20/01/2021.

- [10] Ahmadreza Faghieh-Imani, Sabreena Anowar, Eric J Miller, and Naveen Eluru. Hail a cab or ride a bike? a travel time comparison of taxi and bicycle-sharing systems in new york city. *Transportation Research Part A: Policy and Practice*, 101:11–21, 2017.
- [11] Presidenza del Consiglio dei Ministri Governo Italiano. Coronavirus, le misure adottate dal Governo. <http://www.governo.it/it/coronavirus-misure-del-governo>. Accessed: 20/01/2021.
- [12] Gerhard Hitge and Marianne Vanderschuren. Comparison of travel time between private car and public transport in cape town. *Journal of the South African Institution of Civil Engineering*, 57(3):35–43, 2015.
- [13] ISFORT. Osservatorio "Audimob". 16° Rapporto sulla mobilità degli italiani. <https://www.isfort.it/>. Accessed: 20/01/2021.
- [14] Alessio Pagani, Francesco Bruschi, and Vincenzo Rana. Knowledge discovery from car sharing data for traffic flows estimation. In *2017 Smart City Symposium Prague (SCSP)*, pages 1–6. IEEE, 2017.
- [15] Andrea Trentini. Lombardy epa obtorto collo data and anti-pollution policies fallacies. *Journal of e-Learning and Knowledge Society*, 10(2), 2014.
- [16] Andrea Trentini and Federico Losacco. Sampling car-sharing data to evaluate urban traffic behaviour. In *International Conference Applied Computing. IADIS*, pages 295–299, 2017.
- [17] Shuigeng Zhou, Aoying Zhou, Jing Cao, Jin Wen, Ye Fan, and Yunfa Hu. Combining sampling technique with dbscan algorithm for clustering large spatial databases. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 169–172. Springer, 2000.

Ringraziamenti

Vorrei dedicare questo spazio del mio elaborato alle persone che mi hanno accompagnato lungo questo percorso di crescita personale e professionale e che hanno contribuito, con il loro supporto, alla realizzazione dello stesso.

Ringrazio il mio relatore Andrea Trentini per avermi pazientemente seguito, in questi 12 mesi, nella ricerca, nello studio e nella stesura di questo elaborato, e per gli enormi contributi nei suoi corsi di Laboratorio di Programmazione e Sistemi Embedded, con spunti di riflessione e stimoli a livello culturale atti a diffondere il patrimonio di conoscenze della scienza informatica. Ringrazio il mio correlatore Dario Malchiodi per avermi pazientemente seguito nell'elaborazione dei dati raccolti e per il grande contributo nel suo corso di Statistica e Analisi dei Dati, dove dà l'opportunità agli studenti di testare e vedere in funzione formule matematiche su una vera collezione di dati tramite i notebook Jupyter, rendendo più digeribili e stimolanti argomenti che altrimenti risulterebbero astratti e complicati. Vorrei ringraziare il Prof. Andrea Visconti, responsabile del corso di Crittografia, per avermi insegnato con estrema semplicità una materia che ho tentato di studiare svariate volte da autodidatta prima dell'università, ma a cui ho sempre rinunciato per la sua enorme complessità mostrata sui libri di testo.

Ringrazio di cuore i miei genitori per avermi permesso di portare a termine gli studi universitari coi loro enormi sforzi e mia sorella per avermi dato preziosi consigli sulla base della sua esperienza universitaria pregressa. Ringrazio la mia ragazza, Sara, al quale dedico questo lavoro, per aver creduto in me e per avermi sostenuto nei momenti più difficili di questo percorso. Ultimo, ma non per importanza, ringrazio il mio amico di lunga data, Giuseppe, il cui motivo del ringraziamento non mi è ancora chiaro.

