

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

The 15-Minute City for Tourists

Mario Tapia Montero

MSc. in Data Science

Supervisor: Prof. Francesco Silvestri

September 19th, 2024



Main objective

This thesis investigates how the availability of **attractions, restaurants, and public transportation within walking distance** of each **accommodation** influences its **attractiveness and price**





Agenda

- Context
- Methodology
- Results
- Conclusions and future work

Context



About walking



Most inclusive
form of mobility



Improves
health

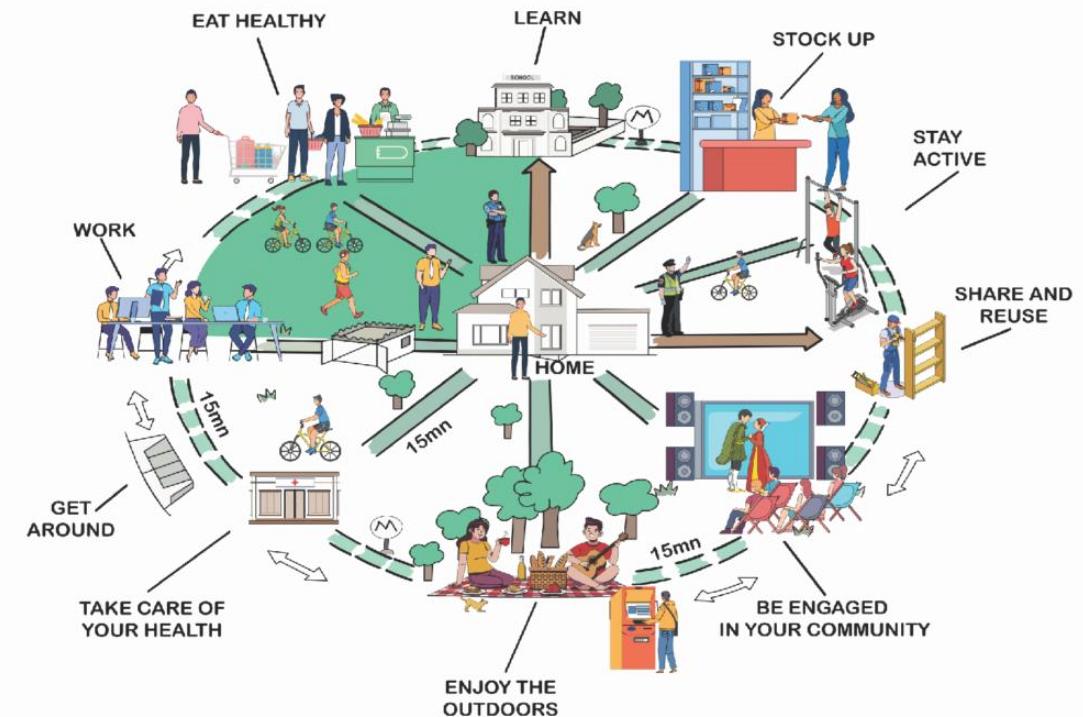


Improves
sustainability



15-minute cities concept

- Ensures essential services are **within a 15-minute walk from home**.
- Key services include commerce, healthcare, and education.





About tourism



Number of attractions within walking distance*



A large portion of the tourist's time is spent near accommodation**

* N. Shoval, B. McKercher, E. Ng, and A. Birenboim, "Hotel location and tourist activity in cities," *Annals of Tourism Research*, vol. 38, no. 4, pp. 1594–1612, 2011.

** S. Aksoy and M. Ozbuluk, "Multiple criteria decision making in hotel location: does it relate to postpurchase consumer evaluations?" *Tourism Management Perspectives*, vol. 22, pp. 73–81, 2017.



Walkable Tourism: A 15-Minute Perspective

- By replacing:
 - **Residents with Tourists**
 - Healthcare facilities, grocery stores, and banking services with **Attractions, Restaurants, and public transport stops**
- Also considering **quantity** and **quality** of **POIs near** lodging



Assumptions



Tourist Type

Prefers walking
and public
transportation

Walking Speed

Considered
constant
(1 m/s)

Shortest-Path

When moving
from
accommodation to
POI



Cities under analysis



Naples

Hotels:	~ 200
BnBs:	~ 7k
Attractions:	~ 600
Restaurants:	~ 2k
Metro stations:	23





Cities under analysis



Venice

Hotels:	~ 300
BnBs:	~ 5k
Attractions:	~ 600
Restaurants:	~ 1k
Vaporetto stops:	20



Methodology



Steps

- 1) Mapping each POI to the nearest node of the city network
- 2) Create attractions/restaurants ranking and filter top- k
- 3) Calculate shortest-path distance from accommodations to each POIs in the city

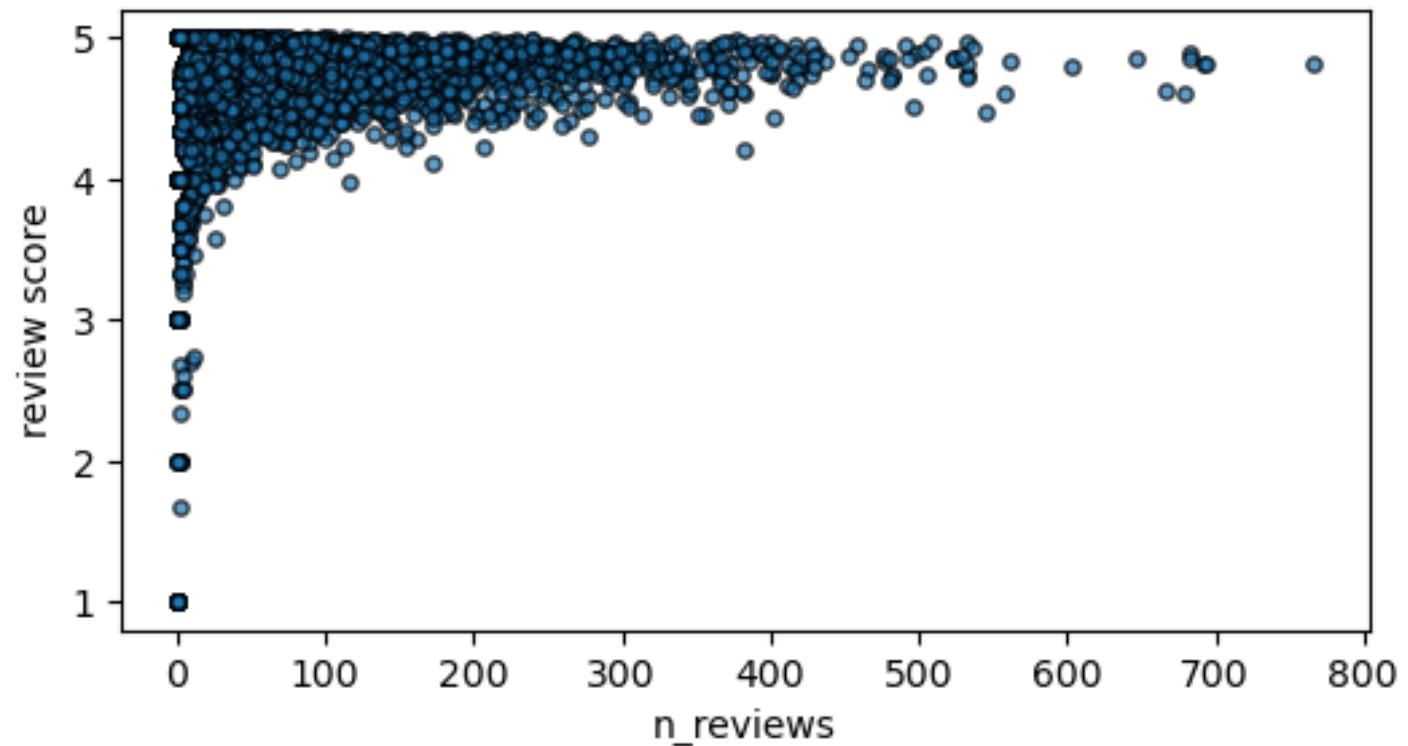


Steps

- 1) Mapping each POI to the nearest node of the city network
- 2) Create attractions/restaurants ranking and filter top- k
- 3) Calculate shortest-path distance from accommodations to each POIs in the city

Attractions and Restaurants ranking

- Variability in **review score** as the **number of reviews** changes
- Indicator of reliability: POIs with higher number of reviews are more trustworthy





Attractions and Restaurants ranking

The Bayesian Score allows to adjust the product score in the following way:

- For a POI with a **substantial number of ratings**, the score should be **the POI's review score** as reported by TripAdvisor/AirBnB
- For a POI with a **fewer than average** number of ratings, the score should be **around the overall arithmetic mean**

Ranking	Name	Overall Score	Review count	Bayesian score
1	Trattoria Al Gazzettino	4.5	12,295	4.477414
5	Trattoria Bar Pontini	4.5	3,397	4.425122





Steps

- 1) Mapping each POI to the nearest node of the city network
- 2) Create attractions/restaurants ranking and filter top- k
- 3) Calculate shortest-path distance from accommodations to each POIs in the city



Shortest-path calculation





Shortest-path calculation





Shortest-path calculation





Shortest-path calculation





Steps

- 4) Tourism Walkability Index determination for each accommodation
- 5) Evaluate the city's attractiveness using the TWI histogram
- 6) Conduct price analysis



Tourism Walkability Index (TWI)

- Variables **Count of Attractions** and **Count of Restaurants** within 15-minutes are considered

$$TWI = 2Z_{Count_Attractions} + Z_{Count_Restaurants}$$

- Scale the index (0-100) and classify index

TWI range	Walkability category
$0 \leq TWI \leq 20$	Poor
$20 < TWI \leq 40$	Basic
$40 < TWI \leq 60$	Average
$60 < TWI \leq 80$	High
$80 < TWI \leq 100$	Excellent



Steps

- 4) Tourism Walkability Index determination for each accommodation
- 5) Evaluate the city's attractiveness using the TWI histogram
- 6) Conduct price analysis



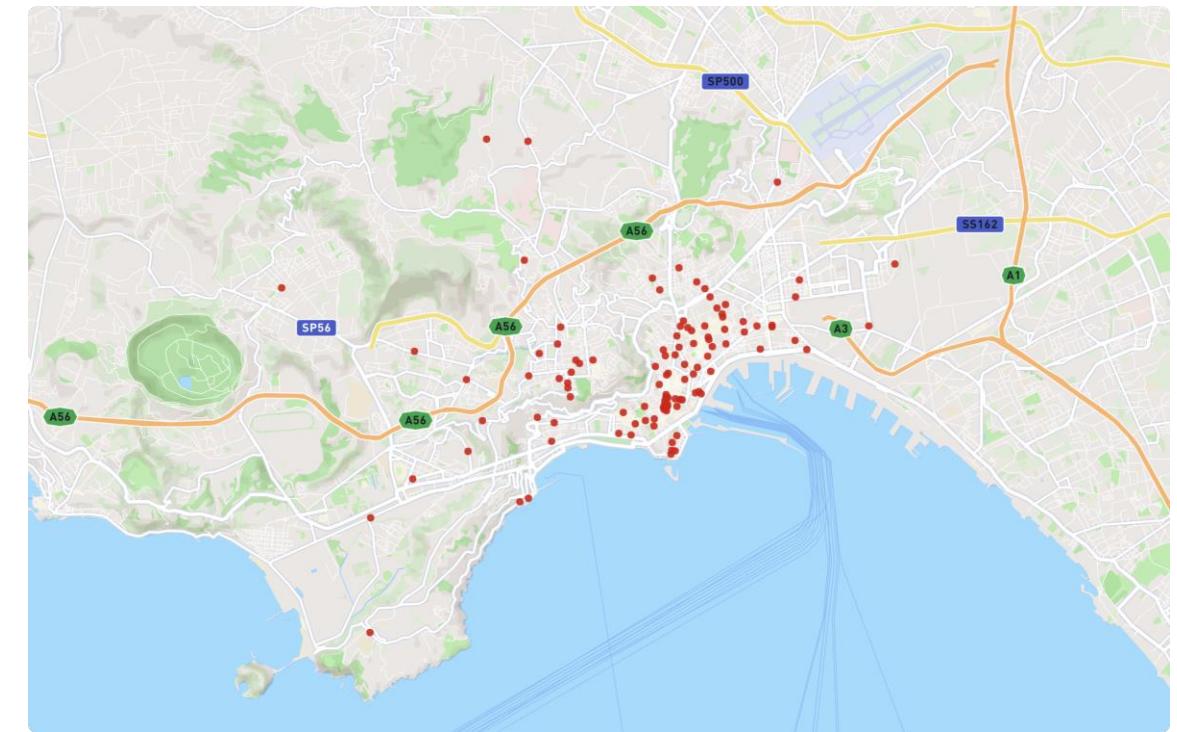
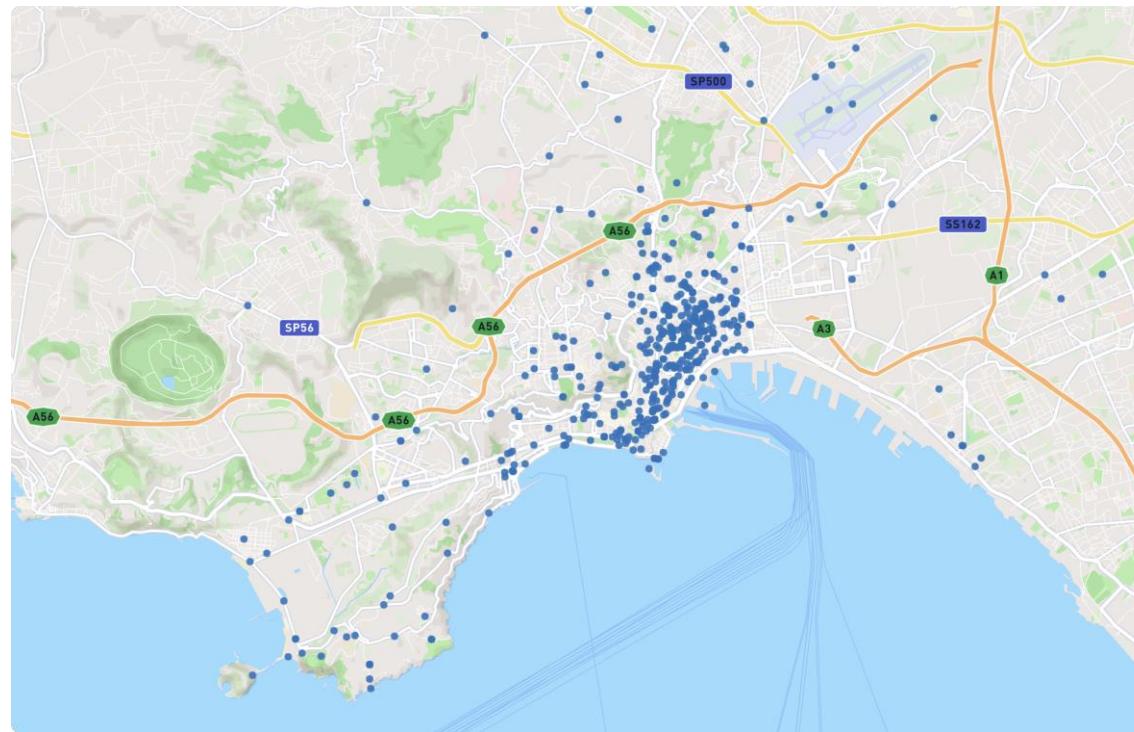
Steps

- 4) Tourism Walkability Index determination for each accommodation
- 5) Evaluate the city's attractiveness using the TWI histogram
- 6) Conduct price analysis

Results

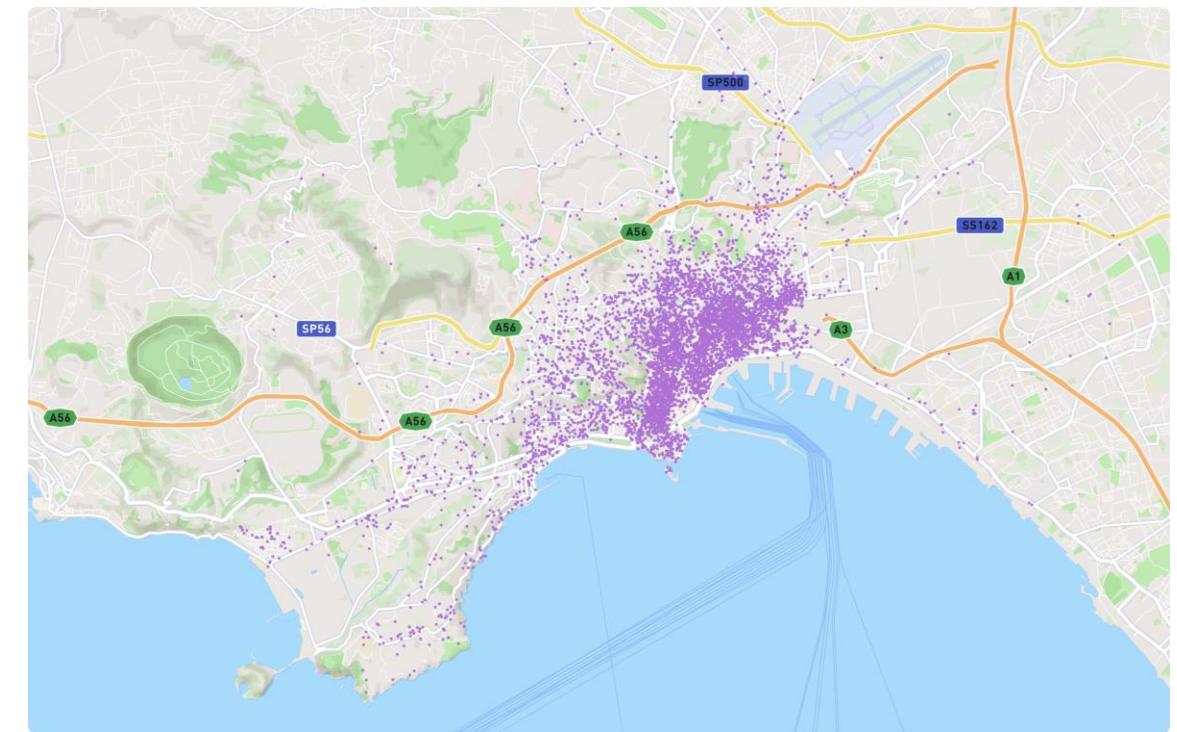
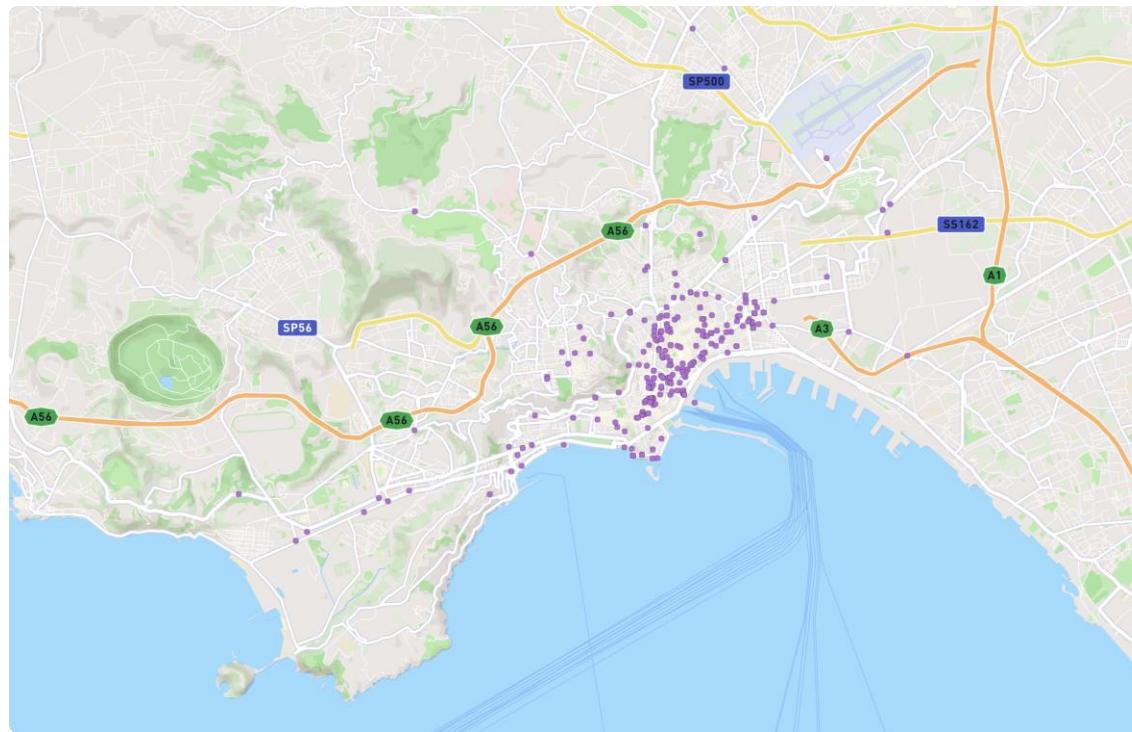


Attractions and Restaurants distribution: Naples



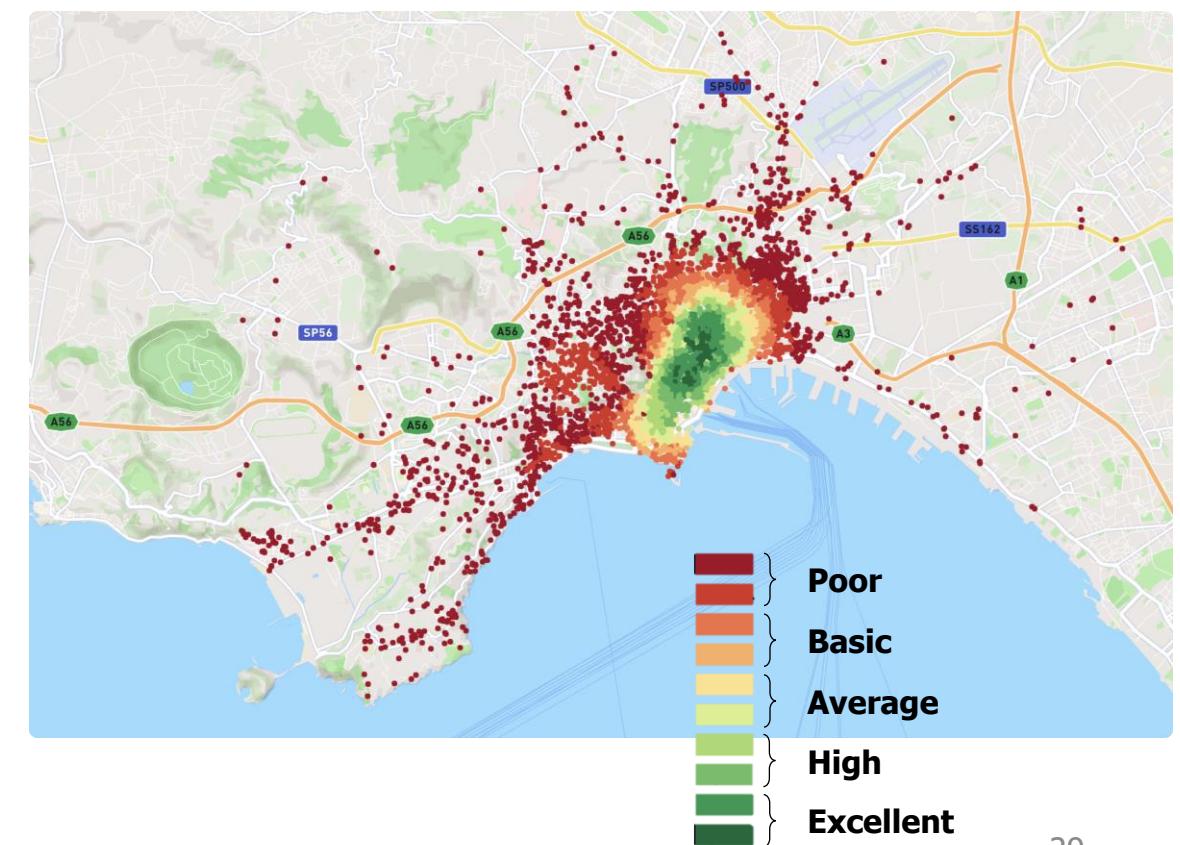
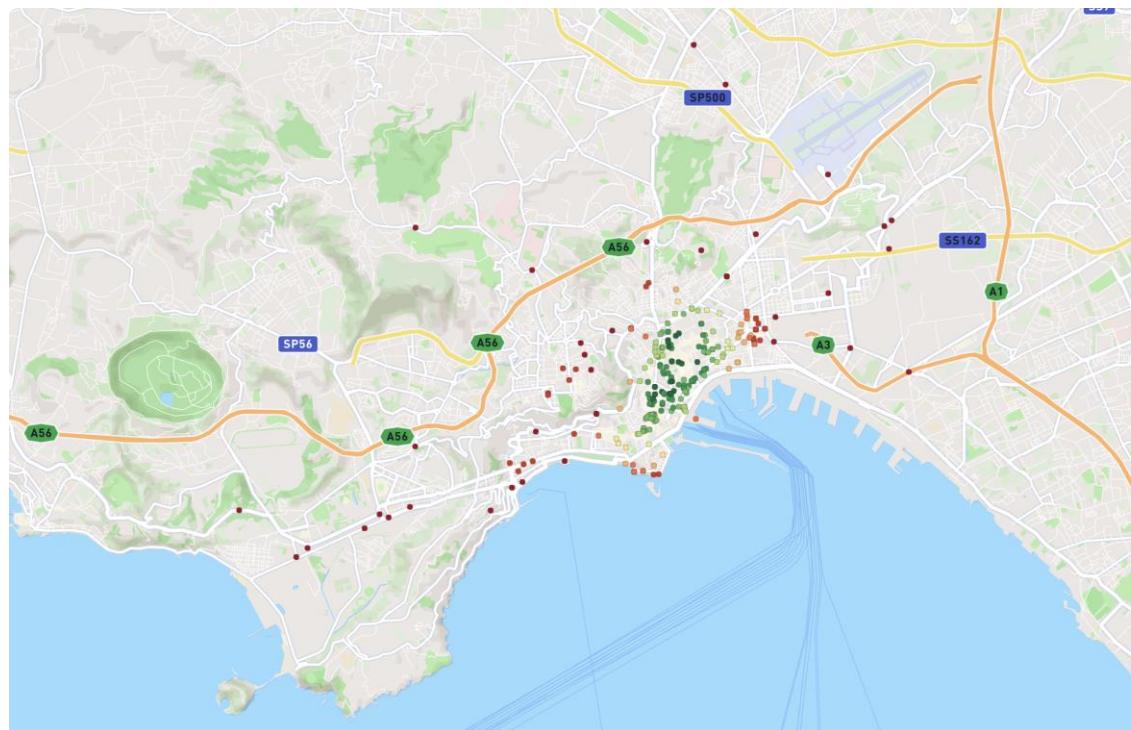


Hotels and BnBs distribution: Naples



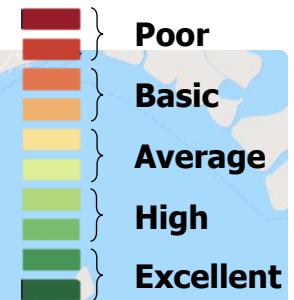
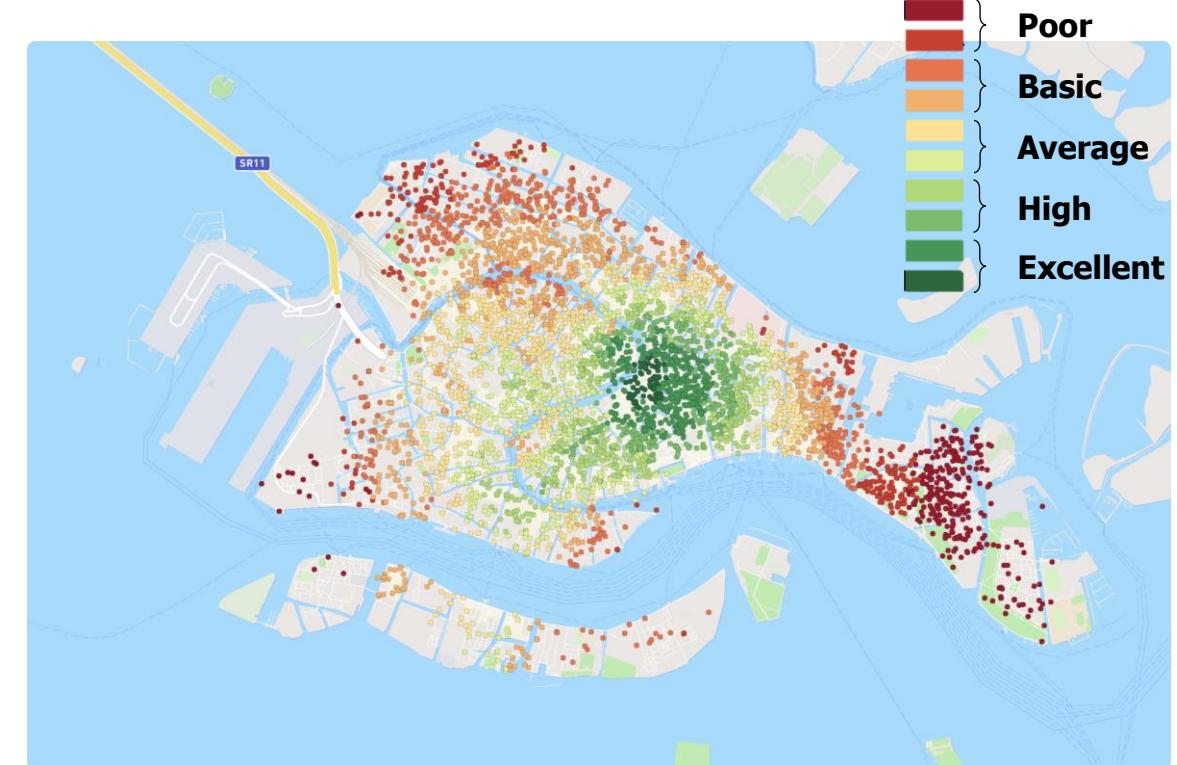
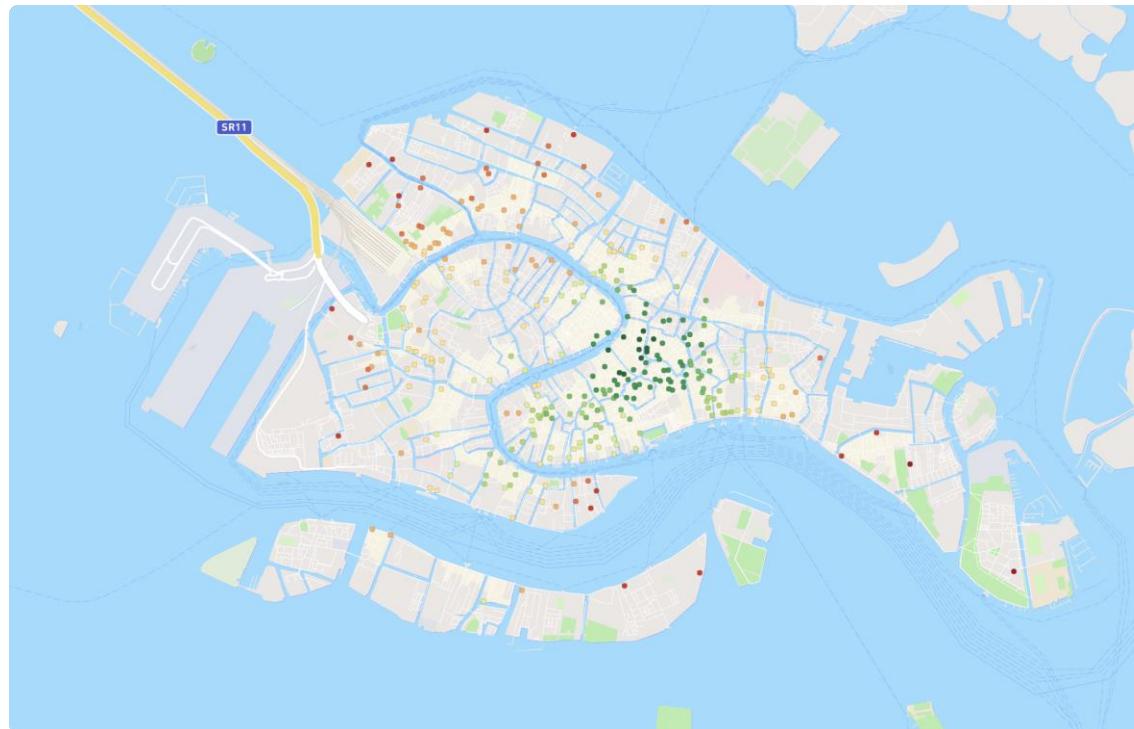


TWI distribution: Hotels and BnBs in Naples



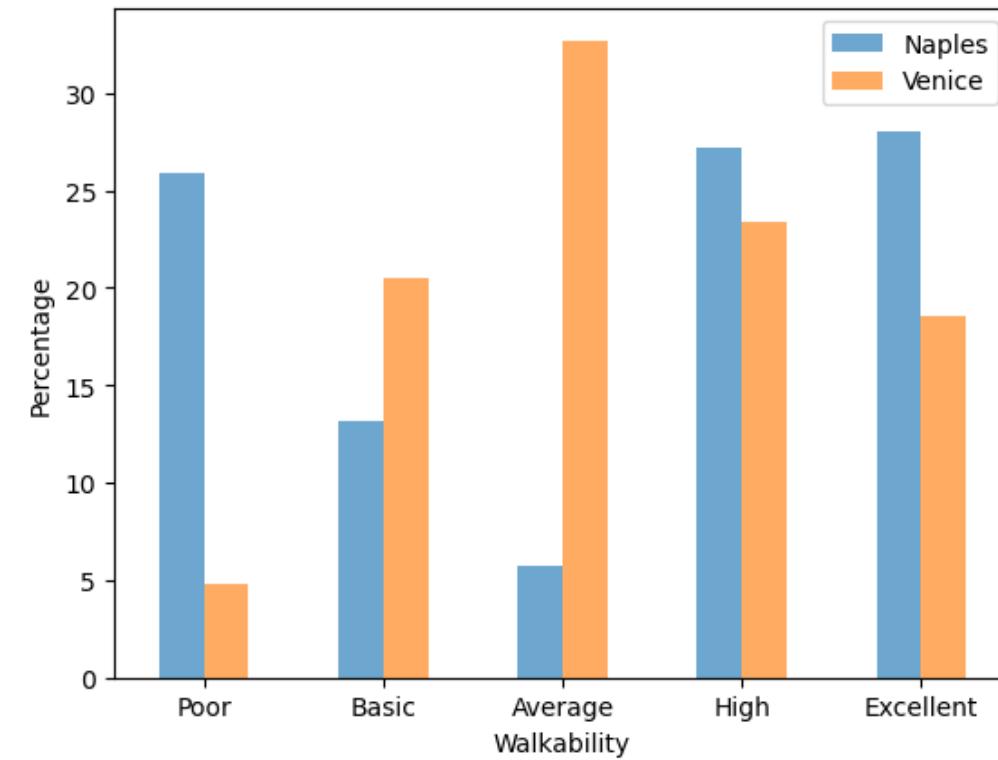


TWI distribution: Hotels and BnBs in Venice





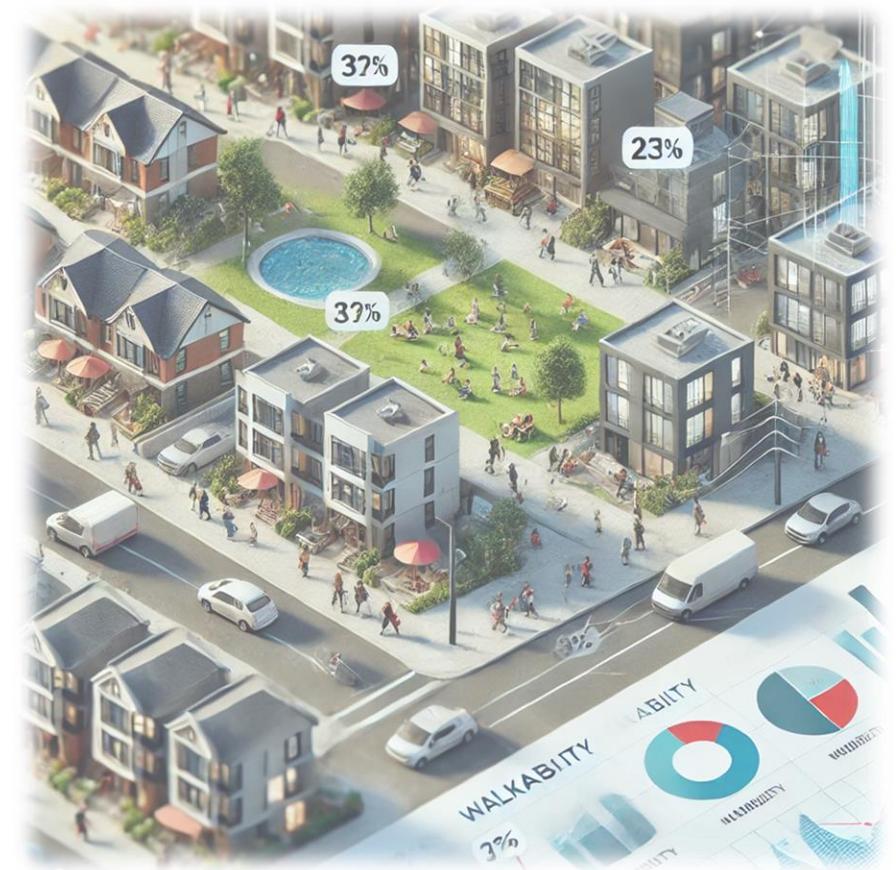
TWI distribution comparison: Naples and Venice at hotel level





Price analysis

- Objective: to determine whether the **price** of an accommodation is influenced by its **walkability**
- Multiple Linear Regression and Regression Trees are used





Variables considered



Location and Accessibility

Availability of Public Transport
TWI (Numerical and categorical)
Location Score
Neighborhood (BnBs)



Accommodation Characteristics

Hotel Class (hotels)
Number of rooms (hotels)
Amenity Type (BnBs)
Room and Property Types (BnBs)
Bathroom, Bedroom, and Bed Number (BnBs)



Review and Rating Metrics

Number of Reviews
Score
Bayesian Score
Mean Bayesian Score for Attractions
Mean Bayesian Score for Restaurants



Multiple Linear Regression results

- All levels of the TWI are significantly different from the Poor level, for hotels and BnBs

- For Venice, the TWI of hotels is dropped in the process of backward elimination

Venice BnB model

Model 1	
Intercept	4.8539*** (0.018)
Walkability_num_2	0.0759*** (0.021)
Walkability_num_3	0.1354*** (0.016)
Walkability_num_4	0.1405*** (0.018)
Walkability_num_5	0.1505*** (0.021)
Bayesian_avg_H	0.0870*** (0.005)
TVs	0.0532*** (0.010)
Bathroom_Amenities	0.0246* (0.011)
neighbourhood_cleansed_4	0.1128*** (0.015)
property_type_1	-0.1771*** (0.023)
property_type_4	0.1090*** (0.031)
bathrooms_2	0.1653*** (0.012)
bathrooms_3	0.3535*** (0.028)
bedrooms_2	0.0611*** (0.008)
beds_2	-0.0236* (0.011)
beds_3	-0.0263* (0.013)
amenities_count	0.0434*** (0.008)
Review_count_H	-0.0972*** (0.005)
location_score	0.0163*** (0.003)
accommodates	0.1628*** (0.007)
maximum_nights	0.0595*** (0.010)
availability_365	0.0708*** (0.008)
AIC	2,276
R ²	0.458
Adj.R ²	0.456
N	4,535



Regression Tree

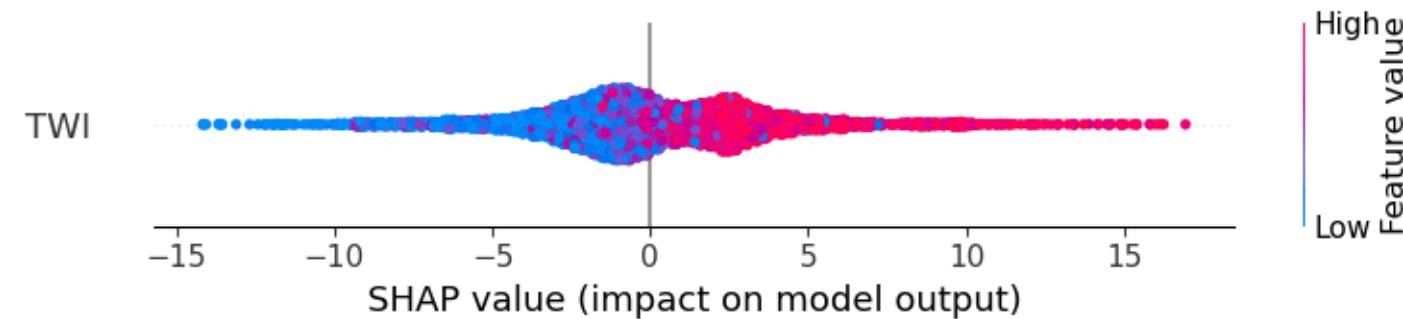
- Objective: Confirm findings regarding the significance and impact of TWI on price



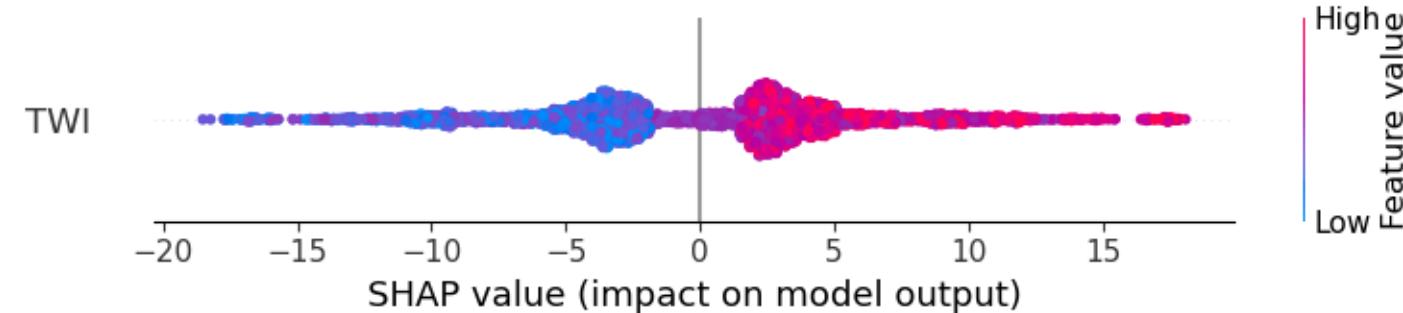


SHAP analysis

➤ Naples

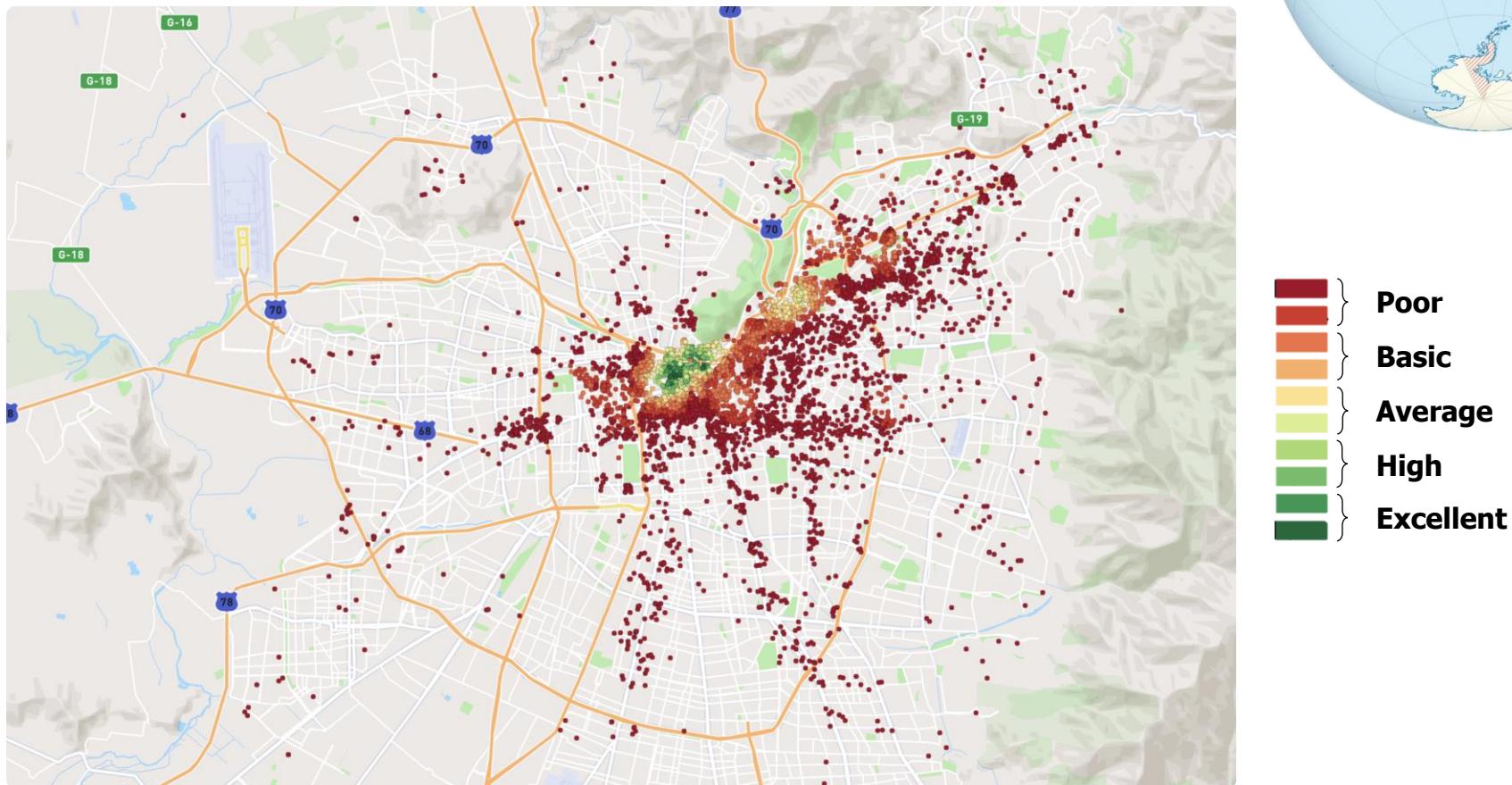


➤ Venice





Insight from larger cities: Santiago, Chile



Conclusions

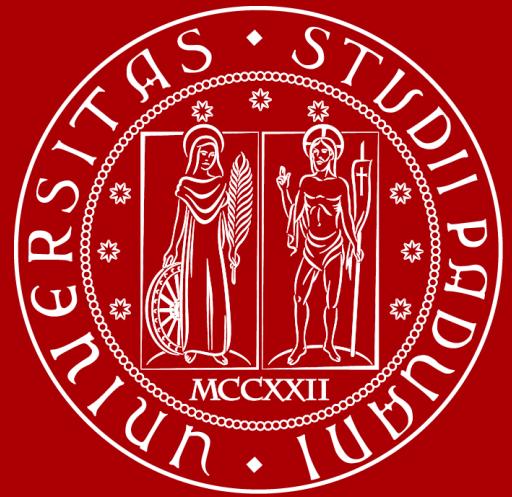


- **Hotels vs. BnBs:** Hotels have **higher percentages of good walkability scores** compared **BnBs**
- **Venice vs. Naples:** Venice presents **consistent walkability experience**; Naples shows **more variability**, with more percentage of **High** and **Excellent** walkability accommodations
- **Higher walkability** is linked to **higher accommodation prices**
- **Larger cities** may have **multiple clusters** of **high walkability accommodations**

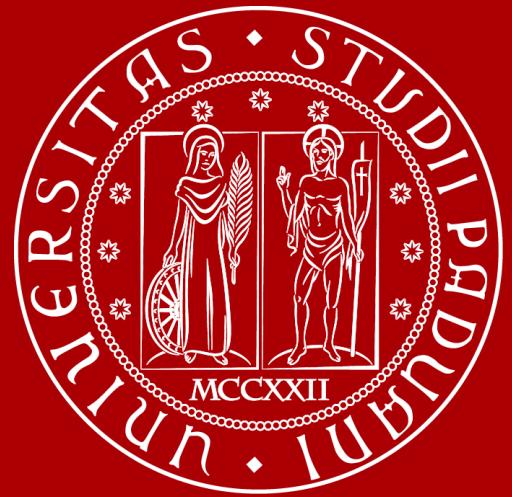
Future Work



- **Refine TWI** with **other weights** and **additional factors** like safety and slope
- Study **30-minute walkability** in **larger cities** and its effect on prices
- **Adjust POI quality threshold** in the rankings
- Consider metro **connectivity** in the city network
- Apply **TWI** for **comparisons at country level**



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



Frank's Walkability Index

- **Intersection density** measures street network connectivity, calculated as the **ratio of true intersections** (three or more legs) **to the land area of the block group**. Higher intersection density provides more direct paths between destinations.

"The street connectivity z score was **weighted by a factor of two** within the walkability index. **This was based on prior evidence** regarding reported utilitarian walking distances and the resulting strong influence of street connectivity on non motorized travel choice."

The development of a walkability index:
application to the Neighborhood
Quality of Life Study

L D Frank,¹ J F Sallis,² B E Saelens,³ L Leary,⁴ K Cain,² T L Conway,² P M Hess⁵

From 2009, published in British Journal of Sports Medicine

} **Validated with expert's opinion and against data on pedestrian travel**



Introduction

- *t*-minute studies measure how attractive a place is by assessing accessibility to points of interest (POI) from a residential perspective
- Tourism plays an important role in cities
- How can we combine these two concepts?



Main objective

The overall objective of this thesis is to determine which areas of a city are **attractive to tourists from a walkability perspective**. To achieve this, the thesis will investigate the relationship between the **location of tourist accommodations** and their **proximity to points of interest** using a developed index and assess how it relates to accommodation pricing.



Particular objectives

- To determine whether well-connected, central accommodations with high accessibility (t-minutes) to well-reviewed points of interest like restaurants, attractions, and metro stations, have higher accommodation prices compared to less connected ones. To examine if there is an association between higher accessibility to these points of interest and high accommodation review score and number of reviews.
- To develop a metric or index that captures these variables and assigns scores to different accommodations.
- To identify the attractiveness of different city areas for tourists based on the previous concepts.



About walking mode

Walking is the **most inclusive form of mobility**

Improvements in **walkability** have been identified as **improving** public and private **health**

In Munich, **60 percent of visitors** engaged in walking as a means of active transport

In London, international **tourists** were more likely to use **walking** as a means of transport than domestic tourists

Despite growth in research on walking in a tourism context, **relatively little has been written on the contribution that walkability might make to tourism**



Key factors for tourism

The most important criterion of **tourists' hotel location choice** is the **number of tourist attractions within walking distance of the hotel**

Airbnb listings **benefit** from a greater **proximity** to the **most visited places** in the city

The **location of tourist accommodation** is a major determinant of the spatial patterns of movements of tourists in urban destinations

Major **visitor attractions** stand out as the **strongest pull factors of destinations** and are key tourism resources

Walkability alone is not attractive to all tourists. But given the interest in encouraging visitors to engage in more **sustainable** forms of mobility it is something that needs to be investigated further



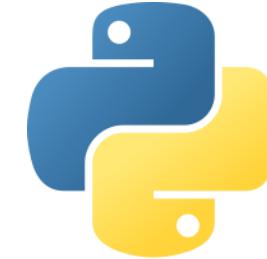
Main variables involved

- Hotels
- Attractions
- Restaurants
- Metro availability



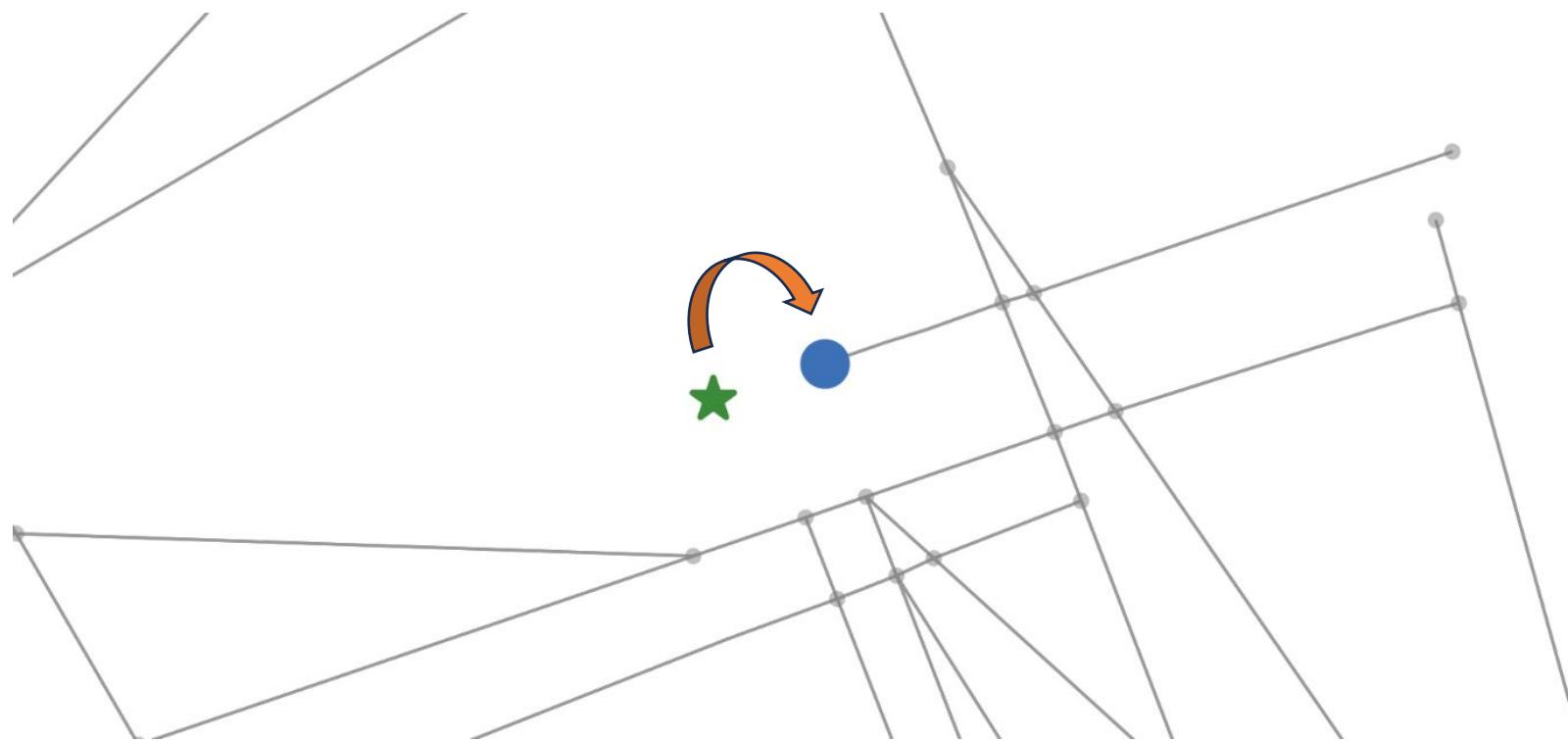
Tools and Data

- Python
- Selenium
- TripAdvisor (hotels, POI)
- InsideAirbnb (BnB)
- OpenStreetMap (Pandana/OSMnx)





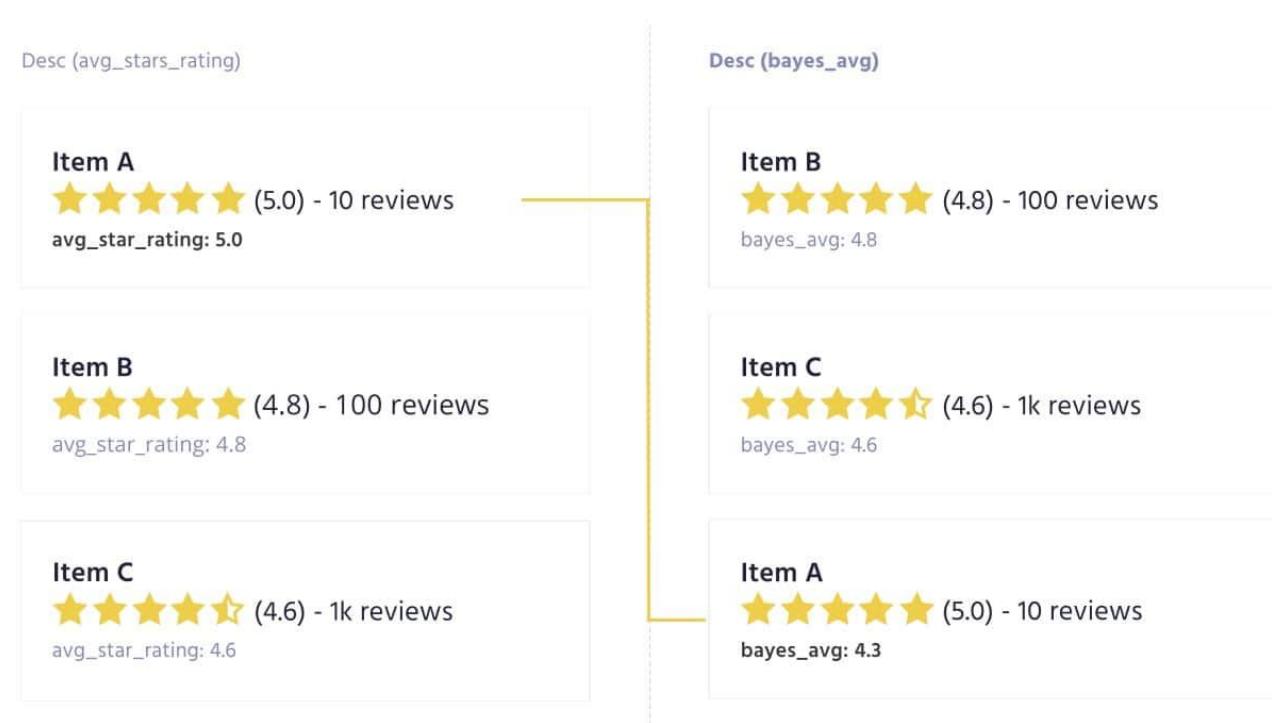
Mapping of nodes





Attractions and Restaurants ranking

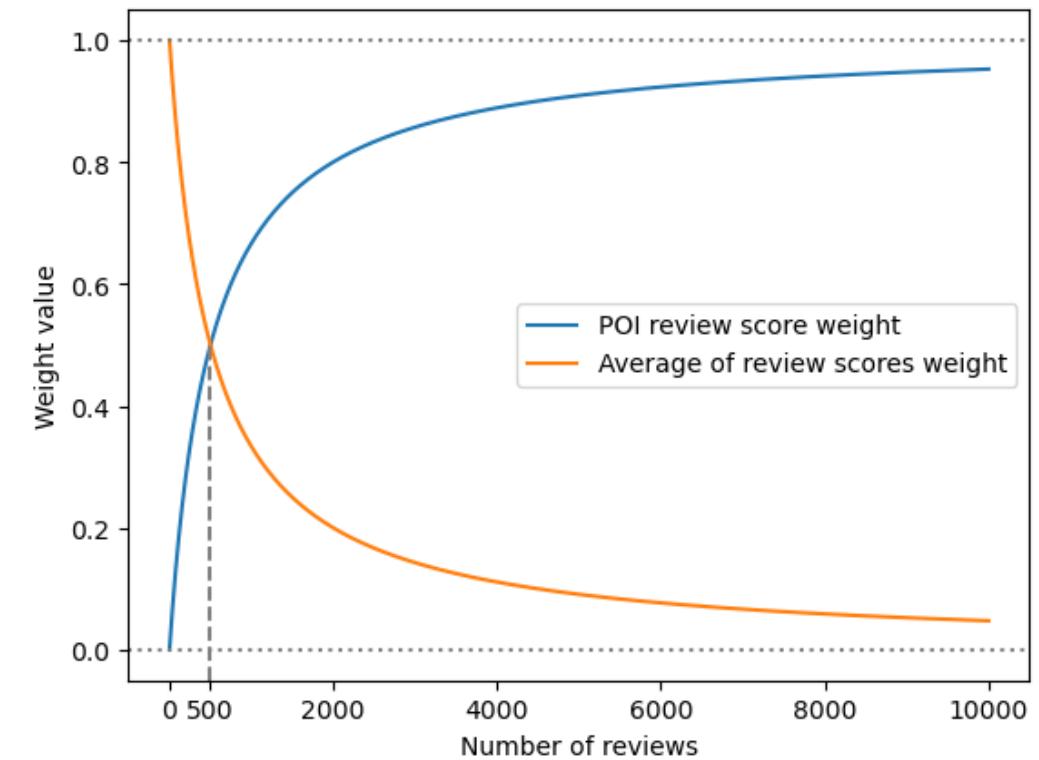
- A metric that could be used to build a ranking based on **review** and **number of reviews** is needed
- **Bayesian score** adjusts a product's average rating by how much it varies from the catalog average. This favors products with a higher quantity of ratings.



Attractions and Restaurants ranking

- As the number of ratings that a POI receives increases, the score should gradually move from the overall mean to the POI's mean

$$B.Score = \frac{n_{reviews}}{n_{reviews} + n_{reviews\ avg}} S + \left(1 - \frac{n_{reviews}}{n_{reviews} + n_{reviews\ avg}}\right) S_{avg}$$





Attractions and Restaurants ranking



Ranking	Name	Score	Review count	Bayesian score
1	Centro Storico di Venezia	5.0	2,218	4.859084
2	Gioielleria Eredi Jovon	5.0	476	4.573338
3	Canal Grande	4.5	41,167	4.495788
4	Piazza San Marco	4.5	36,935	4.495311
5	Basilica di San Marco	4.5	28,570	4.493956

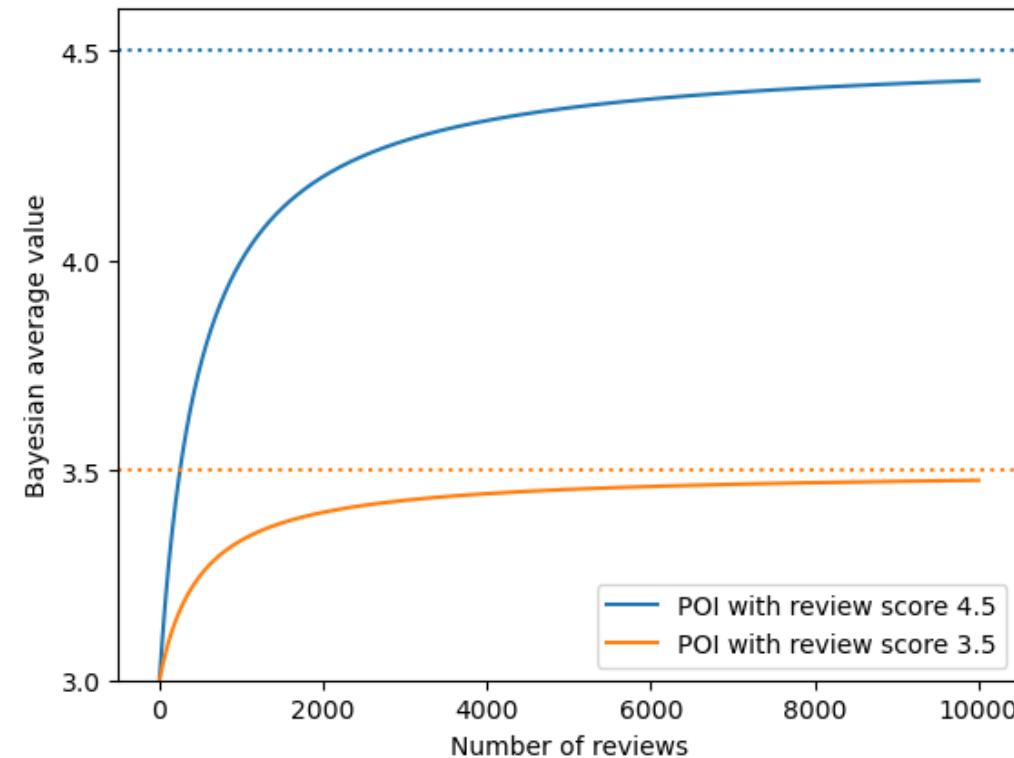
Ranking	Name	Score	Review count	Bayesian score
1	Trattoria Al Gazzettino	4.5	12,295	4.477414
2	Da Mamo	4.5	3,696	4.430527
3	Ristorante Florida	4.5	3,640	4.429575
4	Bistrot de Venise	4.5	3,600	4.428879
5	Trattoria Bar Pontini	4.5	3,397	4.425122



Attractions and Restaurants ranking

Example:

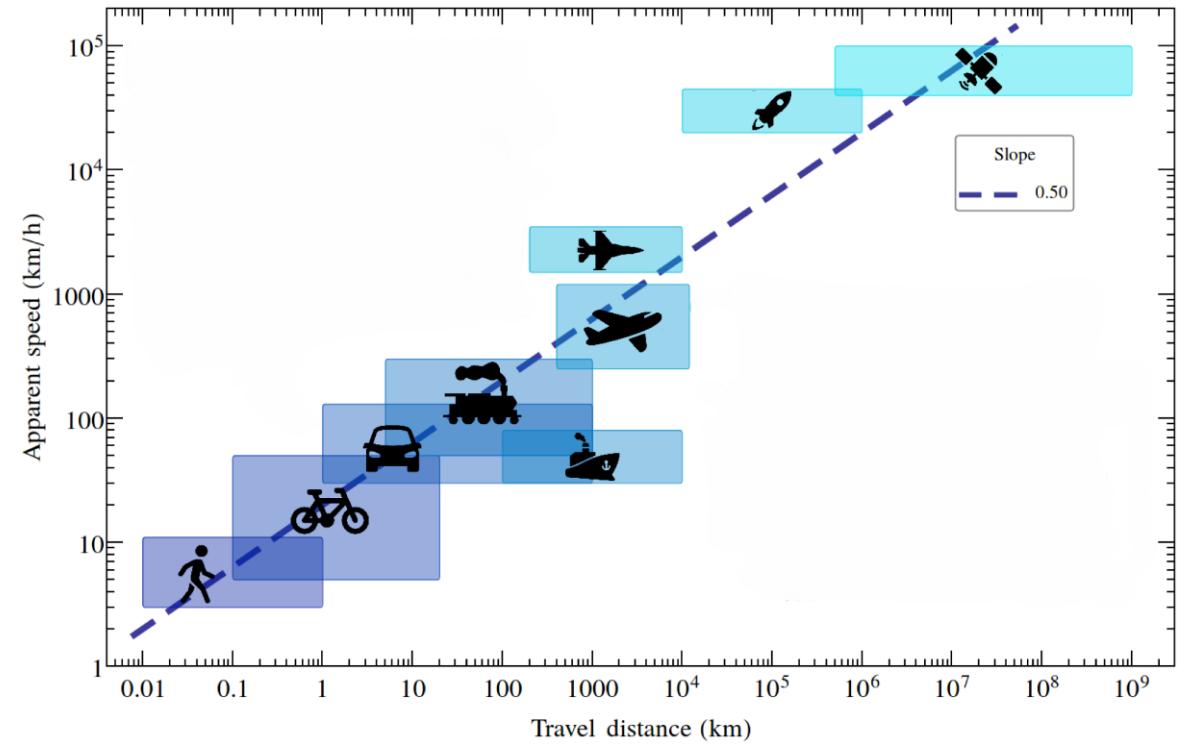
- POI 1: **4.5**
- POI 2: **3.5**
- Mean of review scores: **3.0**
- Mean of number of reviews: **500**





Defining t-minute and walking speed

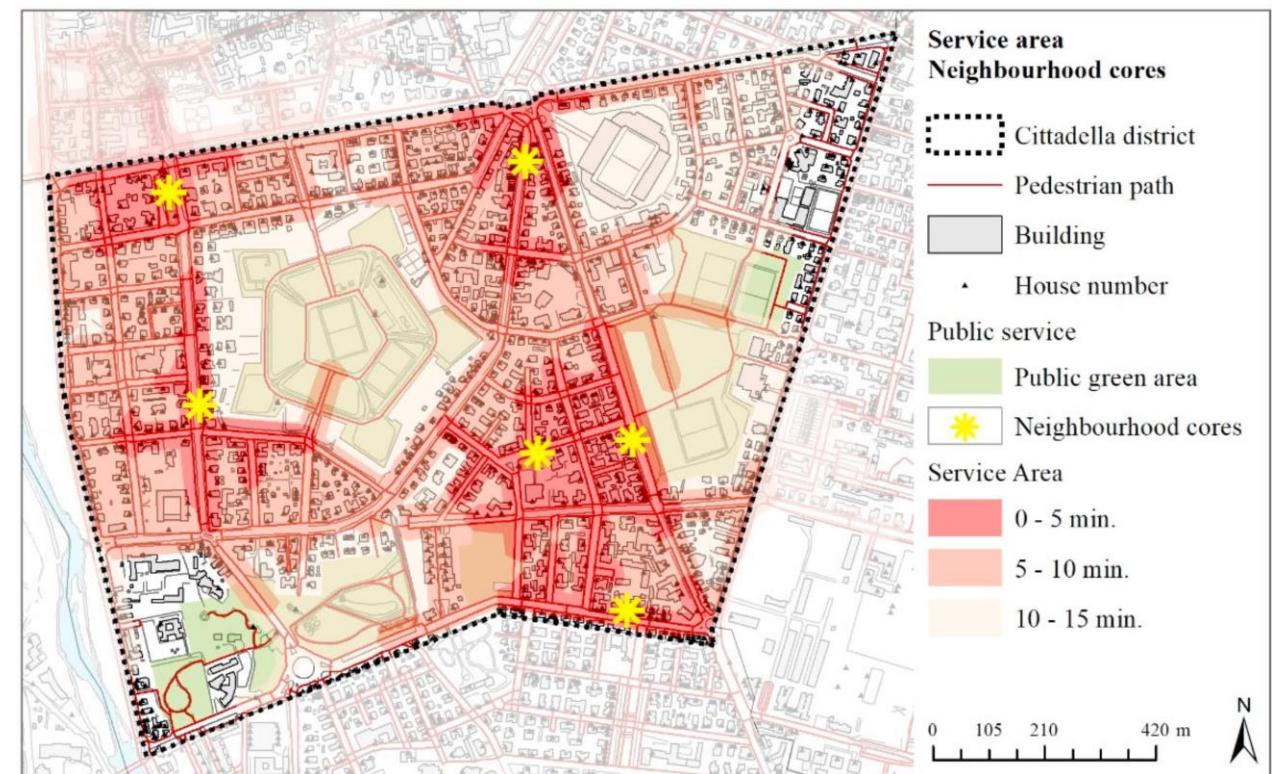
- Walking speed: Between **1** and 1.51 m/s
- Inclusion criterion





Isochrone maps

- Places reachable from the accommodations in t -minutes
- Visual inspection





Feature engineering

- For each hotel, calculate **Closeness centrality** as

$$C(u) = \frac{n - 1}{\sum_{v=1}^{n-1} d(u, v)}$$

where $n - 1$ is the number of reachable nodes (POIs) from node u , v is the destination node, and d is the shortest-path distance from u to v

- The **Count** feature considers the number of attractions and number of restaurants (separately) **within** walking distance



Variables involved for Hotels



Location and Accessibility

Availability of Public Transport
Closeness Attractions and Restaurants
Centrality TWI (Numerical and categorical)
and Location score



Hotel Characteristics

Hotel Class
Number of rooms
Log of mean price
(response var.)



Review and Rating Metrics

Hotel number of reviews
Hotel score
Hotel Bayesian score
Mean Bayesian score for Attractions
Mean Bayesian score for Restaurants



Variables involved for BnBs



Location and Accessibility

Availability of Public Transport
Closeness Centrality to Attractions and Restaurants
TWI (Numerical and categorical)
Location score
Neighborhood



BnB Characteristics

Amenities count
Amenity type
Room and Property types
Nights: Min, Max, and Availability
Superhost
Bathroom, Bedroom, and Bed number
Log of Price (response var.)

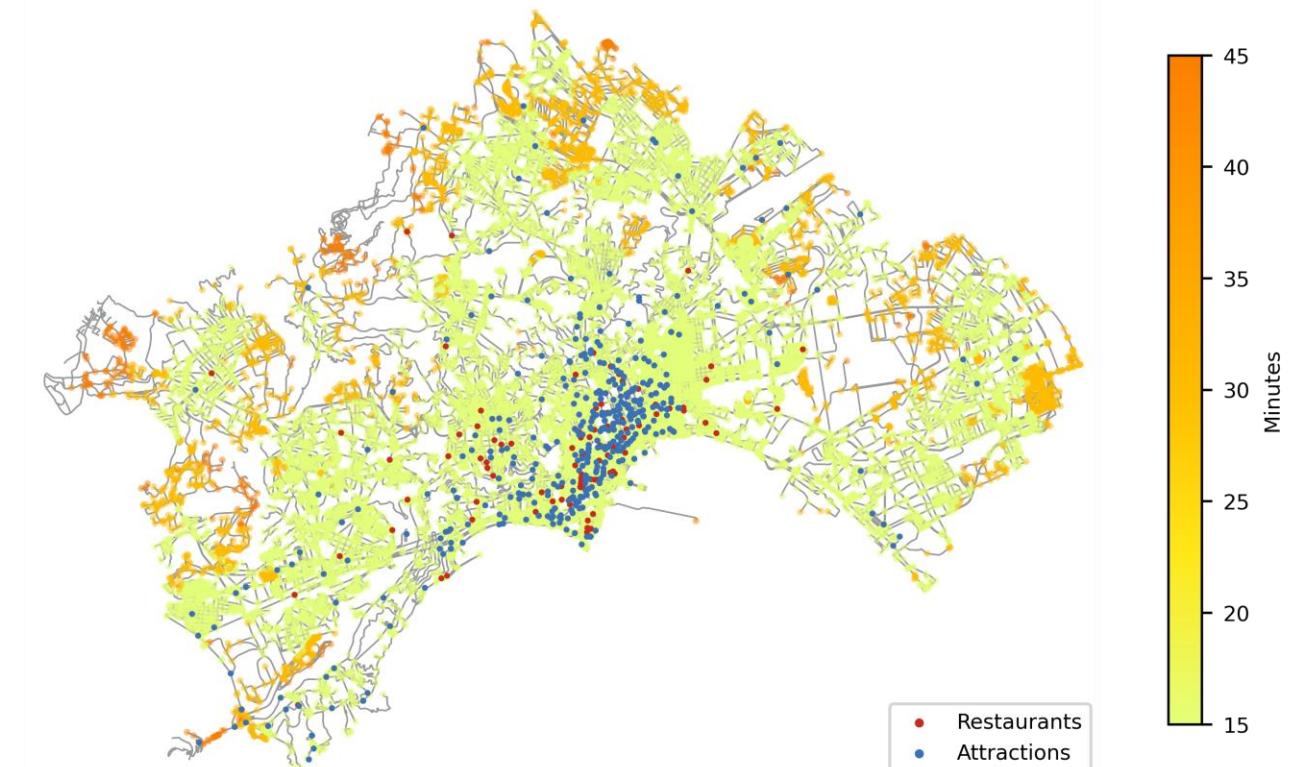
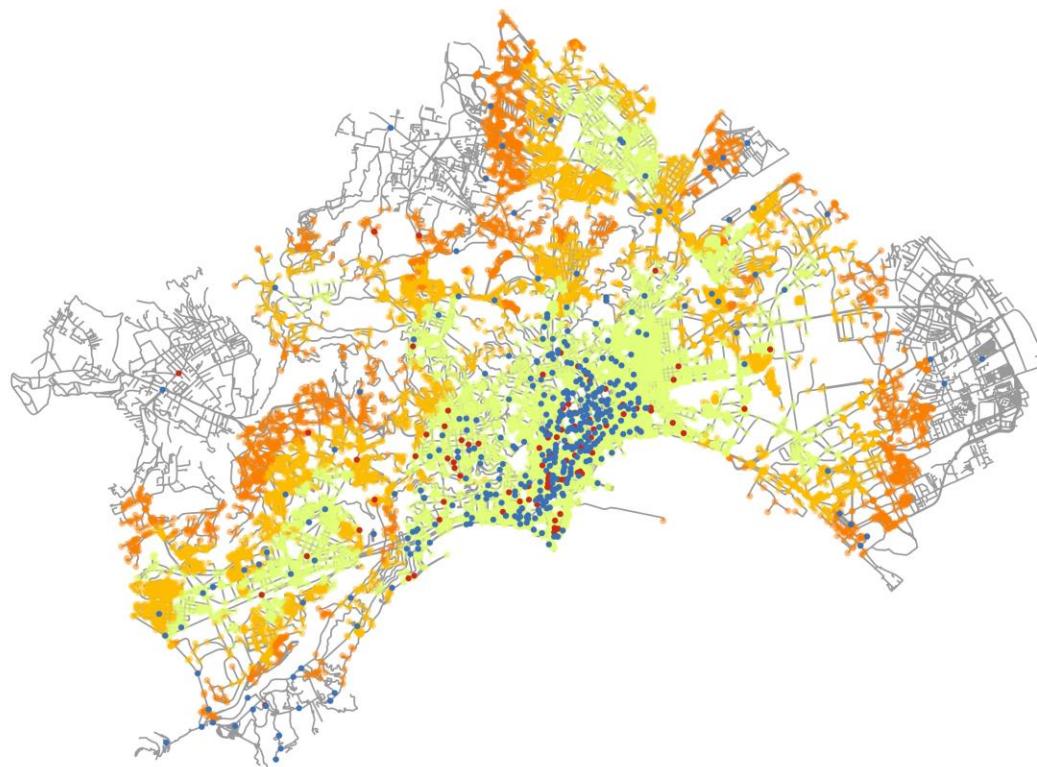


Review and Rating Metrics

BnB number of reviews
BnB score (accuracy, cleanliness, checkin, communication, value)
BnB Bayesian score
Mean Bayesian score for Attractions
Mean Bayesian score for Restaurants



Isochrone maps: Hotels and BnBs



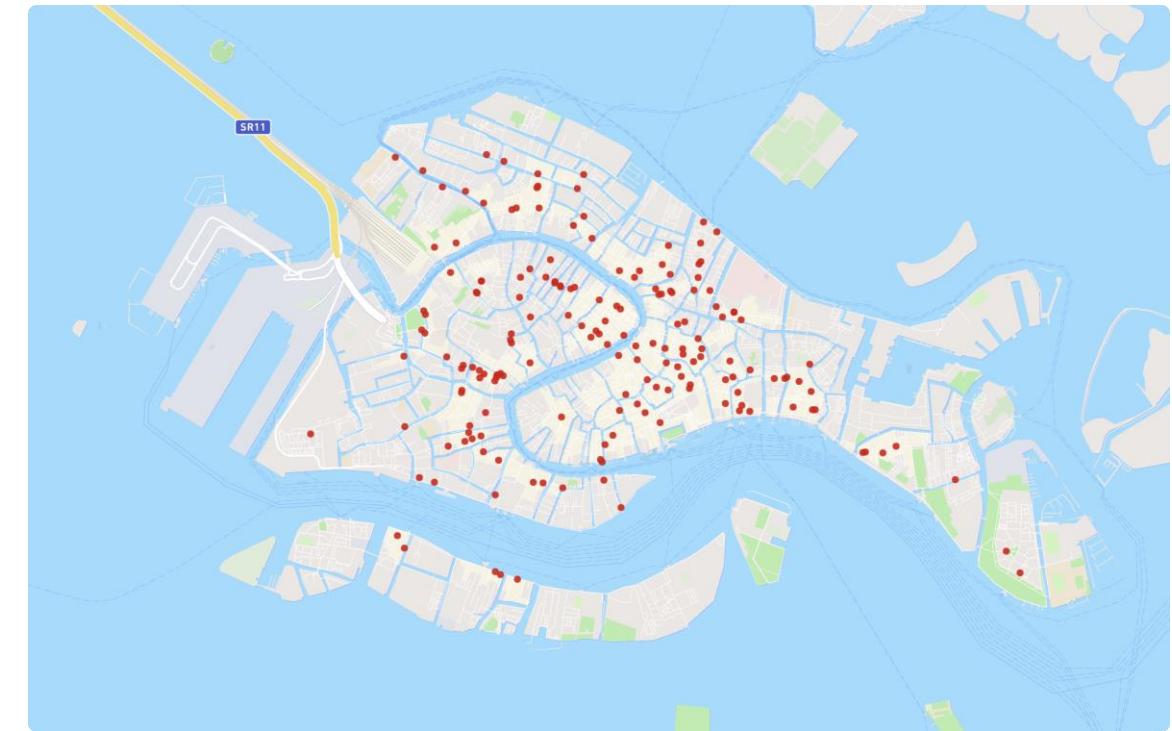
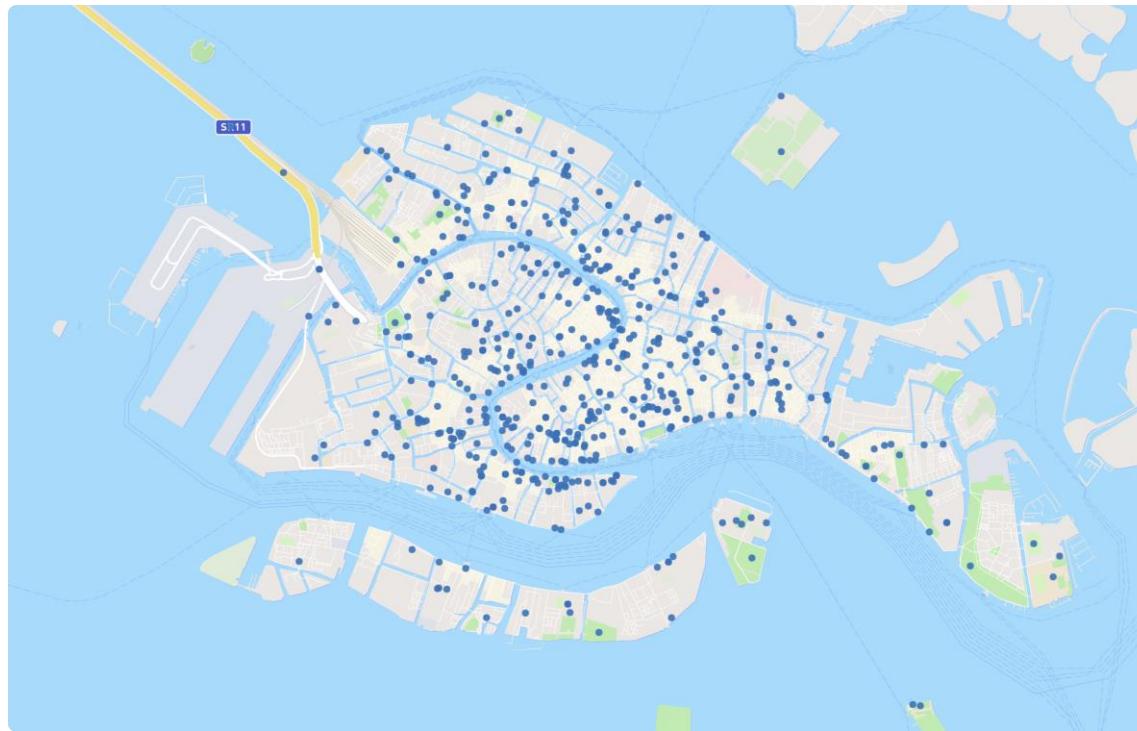


Isochrone maps: Hotels and BnBs



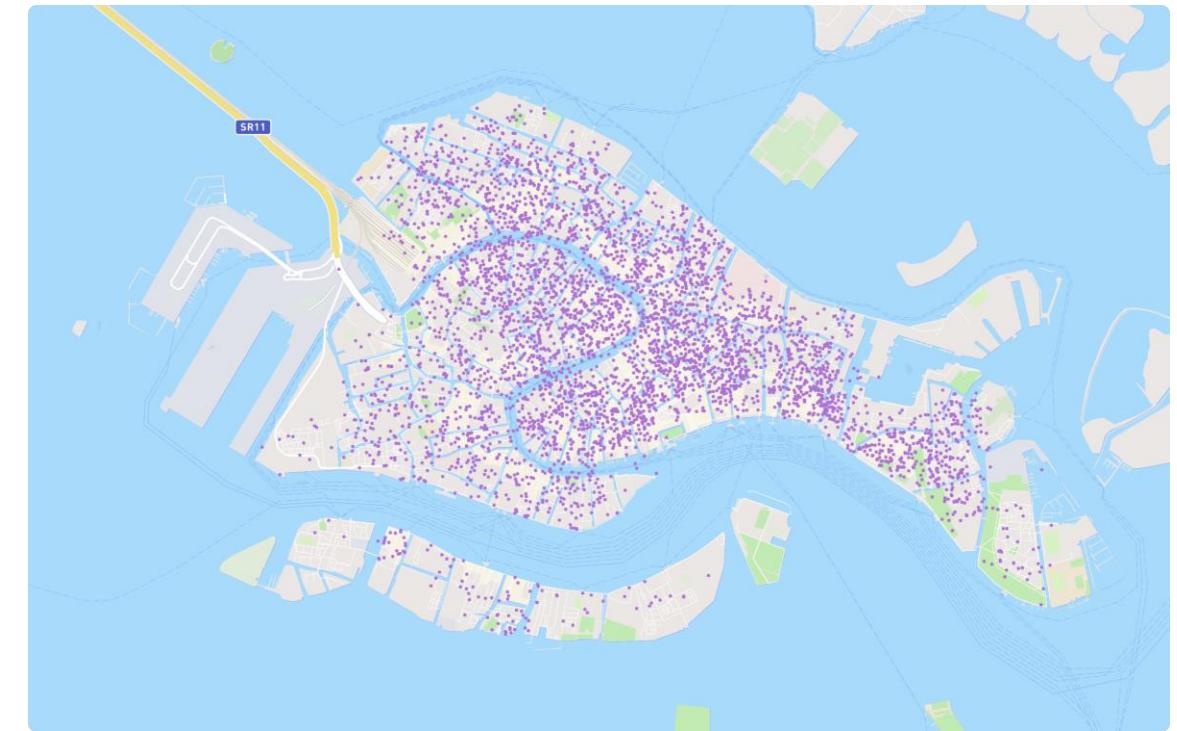
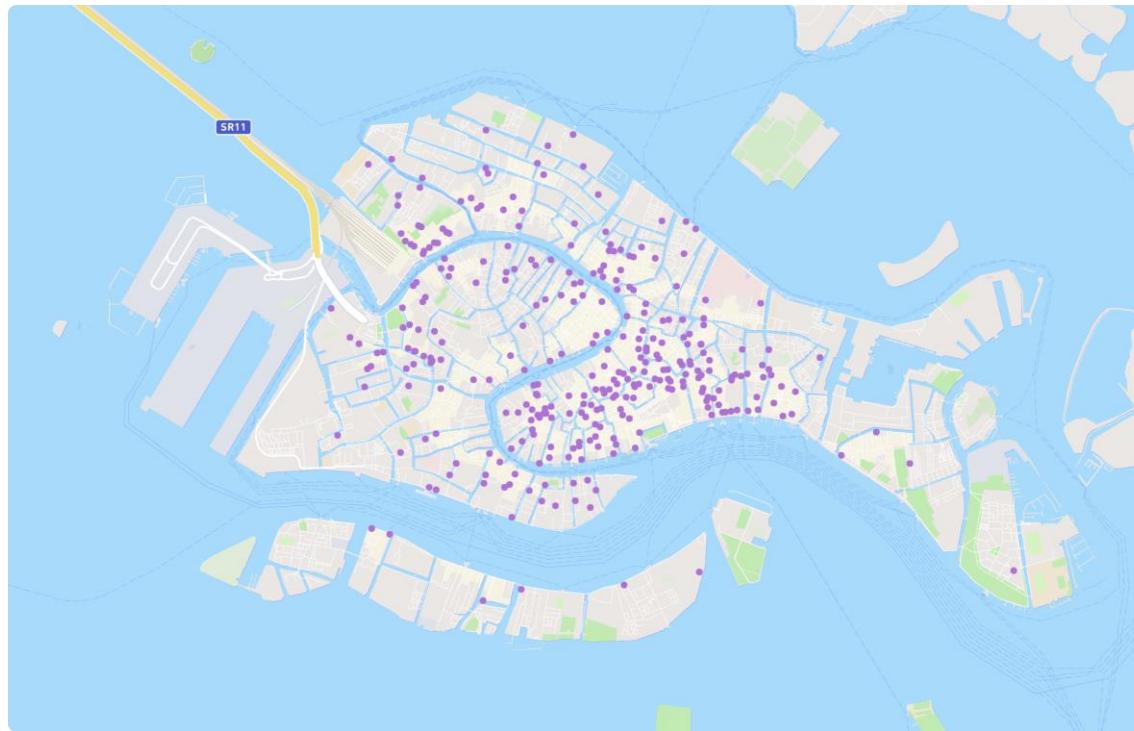


Attractions and Restaurants distribution: Venice



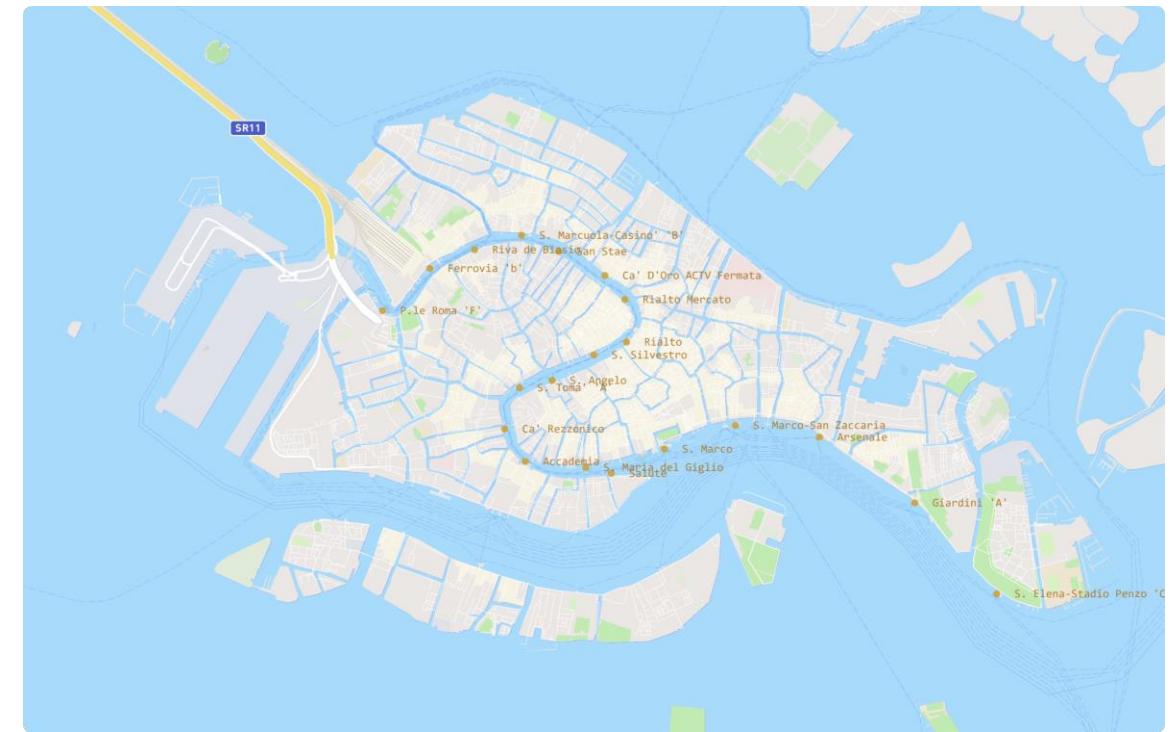
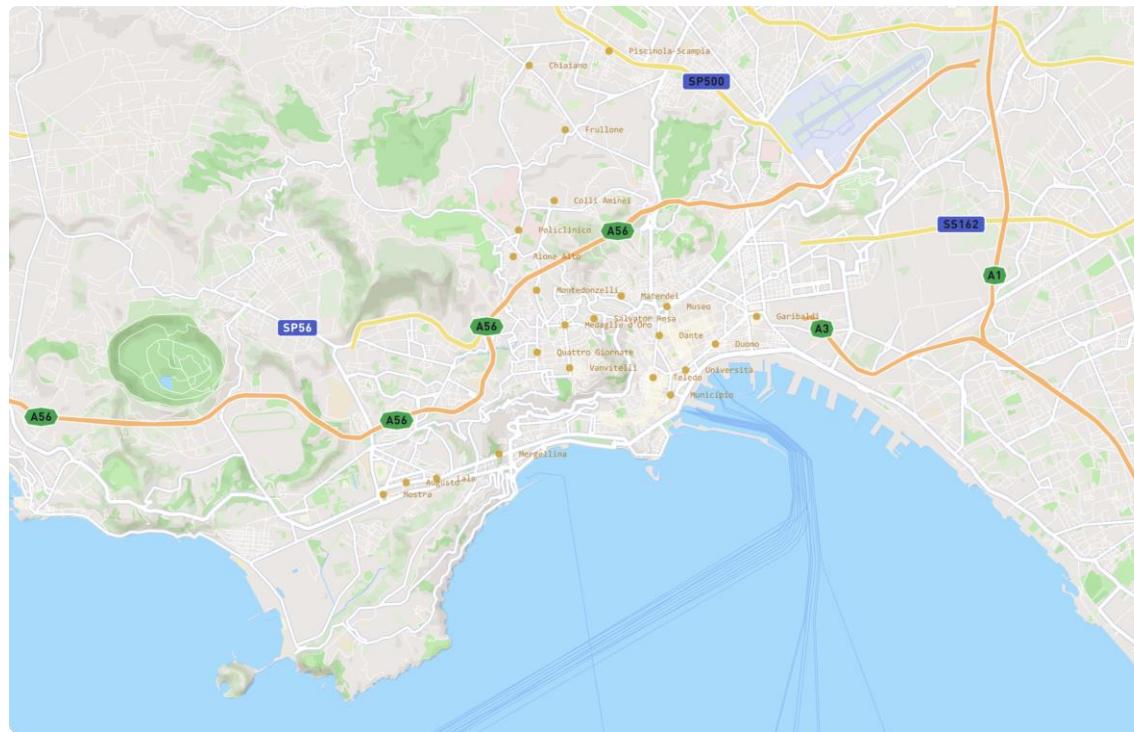


Hotels and BnBs distribution: Venice



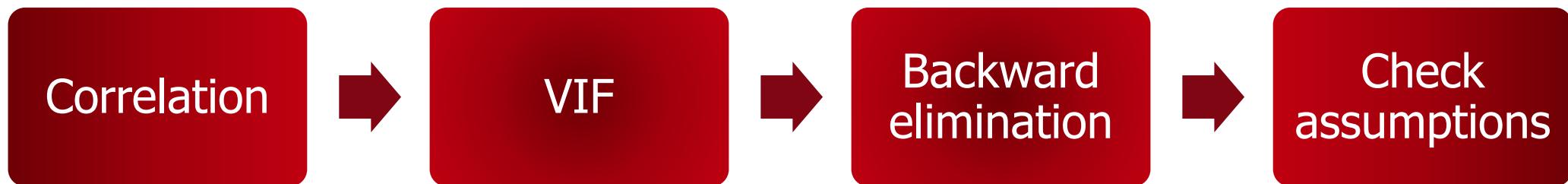


Public transport distribution



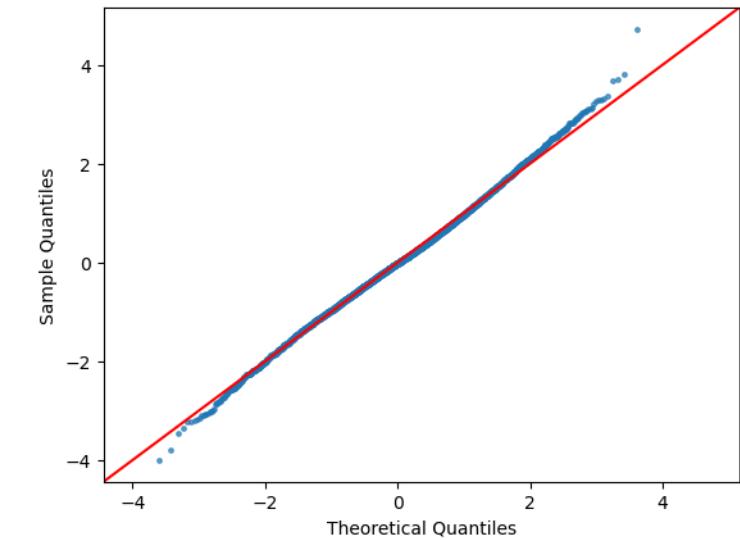
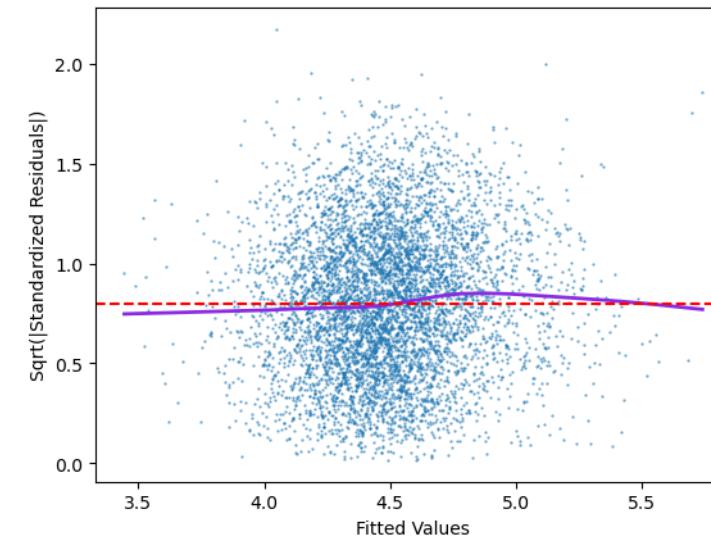
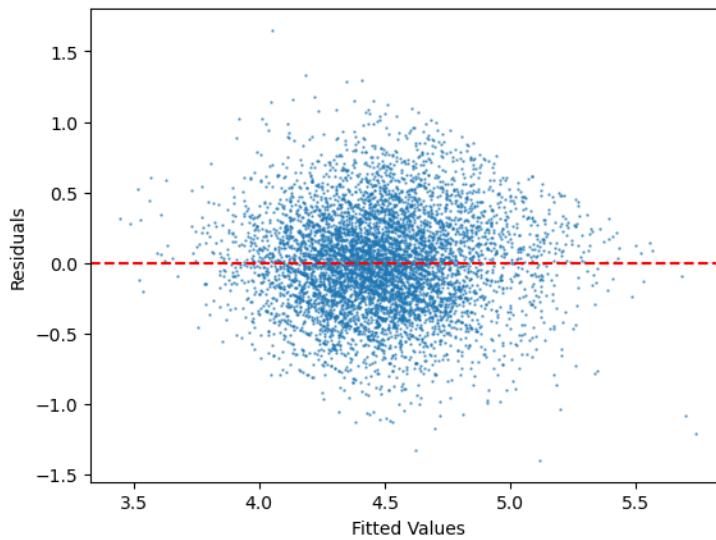


Multiple Linear Regression



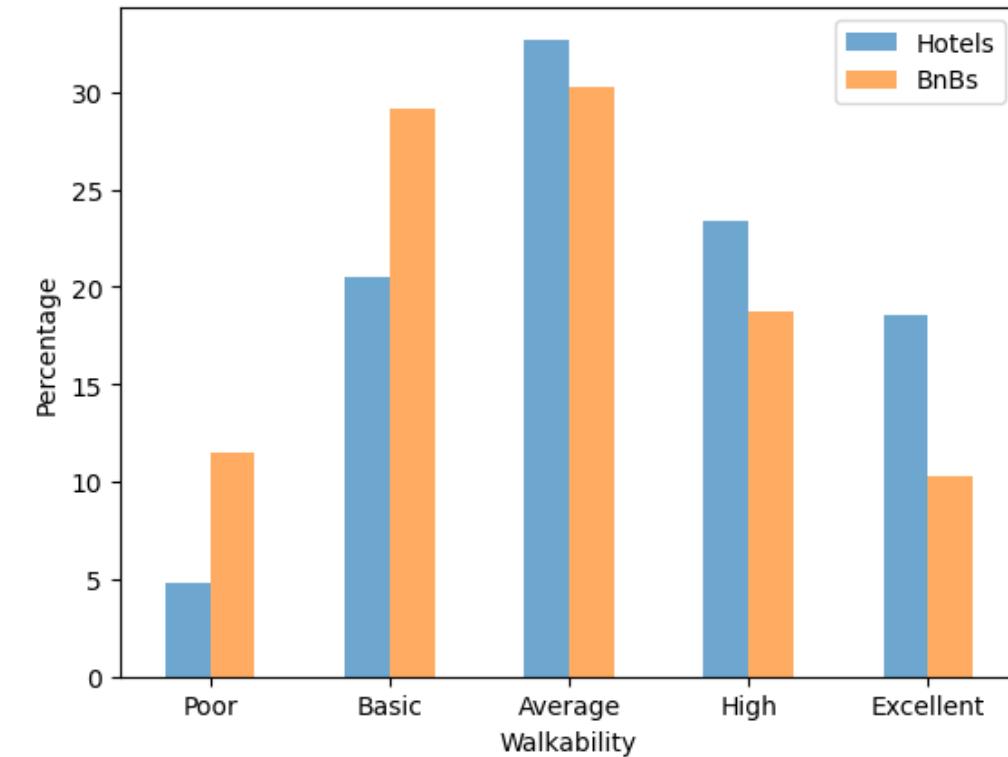
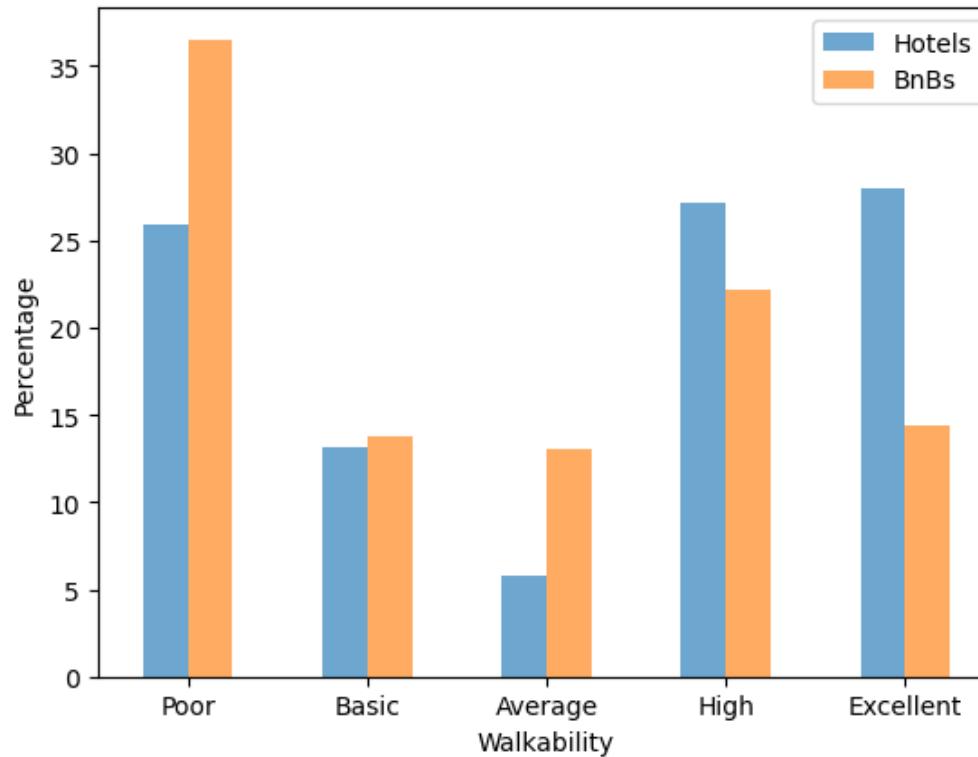


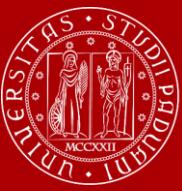
Linearity, homoscedasitcity, and normality: BnBs in Naples



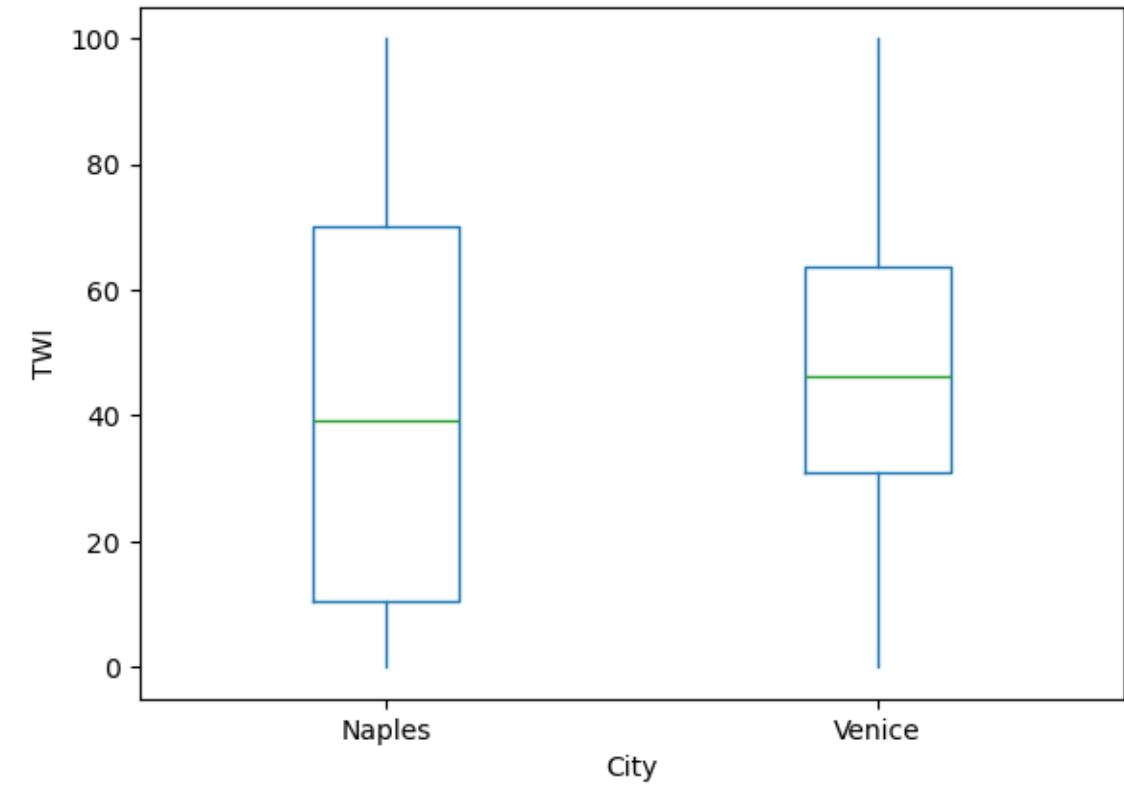
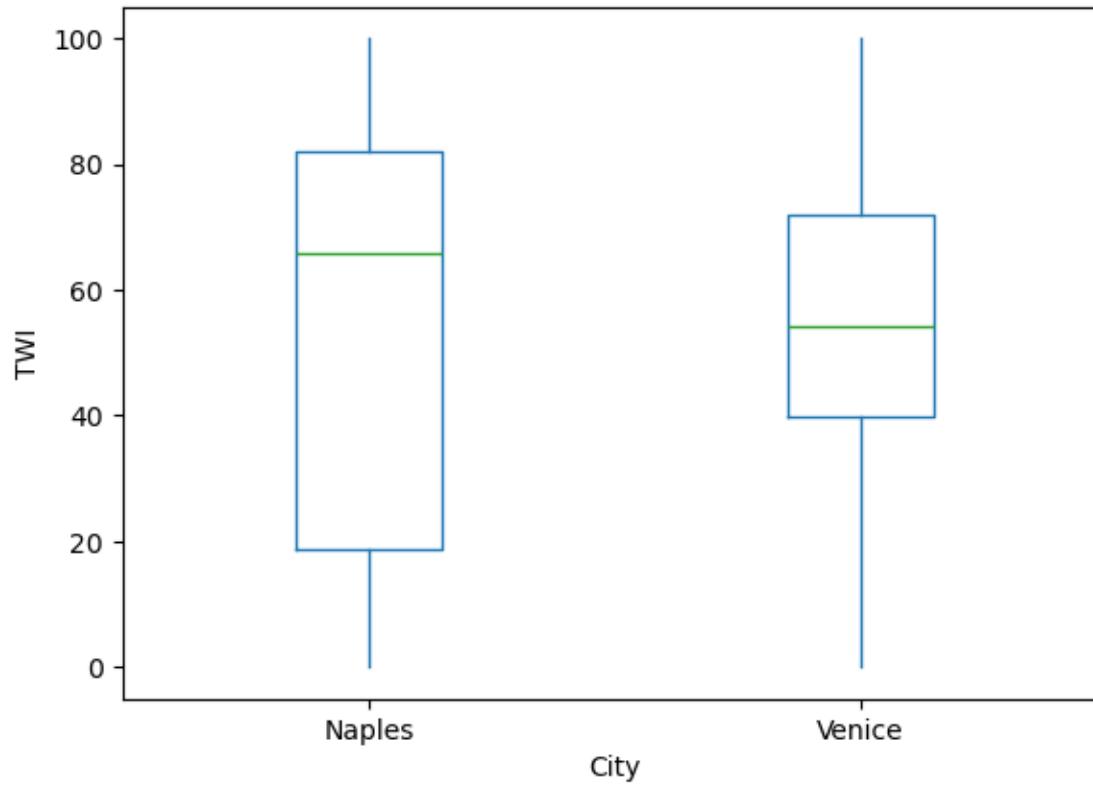


TWI distribution comparison: Naples and Venice





Comparison of Cities: Hotels and BnBs





Summary statistics: Hotels in Naples

- Higher walkability is associated with higher prices
- Average walkability level does not significantly differ from the base level
- Hotel quality is positively related to the price
- Model without TWI presents higher AIC

	Model 1	Model 2
Constant	4.7450*** (0.049)	4.8434*** (0.043)
n_rooms_quant_0	-0.1441*** (0.040)	-0.1361*** (0.041)
Hotel_class_3	0.1210** (0.044)	0.1186** (0.046)
Hotel_class_45	0.2854*** (0.054)	0.2738*** (0.055)
Walkability_num_2	0.1247* (0.044)	-
Walkability_num_4	0.1319** (0.054)	-
Walkability_num_5	0.1503*** (0.043)	-
Bayesian_avg_H	0.0655*** (0.018)	0.0745*** (0.018)
Review_count_H	0.0561*** (0.012)	0.0565*** (0.012)
AIC	18.23	27.95
R ²	0.353	0.308
Adj.R ²	0.330	0.293
N	233	233



Summary statistics: Hotels in Venice

- TWI eliminated in the process of backward elimination
- Hotel quality seems to impact less
- The average restaurant quality appears significant

	Model
Constant	5.3465*** (0.030)
Hotel_class_3	0.1357*** (0.036)
Hotel_class_45	0.4711*** (0.041)
Bayesian_avg_H	0.0936*** (0.016)
Bayes_mean_R	0.0361* (0.017)
AIC	18.40
R ²	0.441
Adj.R ²	0.433
N	301



Summary statistics: BnBs in Naples

- Higher walkability is associated with higher prices
- BnB quality is positively related to the price
- Model without TWI presents higher AIC

	Model 1	Model 2
Intercept	4.3190*** (0.014)	4.4202*** (0.012)
Walkability_num_2	0.0891*** (0.015)	-
Walkability_num_3	0.1239*** (0.015)	-
Walkability_num_4	0.1459*** (0.014)	-
Walkability_num_5	0.1808*** (0.018)	-
Bayesian_avg_H	0.0543*** (0.004)	0.0542*** (0.004)
neighbourhood_cleansed_o	-0.0787*** (0.023)	-0.1622*** (0.022)
neighbourhood_cleansed_i	-0.0925*** (0.022)	-0.1484*** (0.021)
neighbourhood_cleansed_7	0.1423*** (0.016)	0.1792*** (0.016)
neighbourhood_cleansed_9	0.0647** (0.021)	0.1327*** (0.020)
neighbourhood_cleansed_10	0.2692*** (0.035)	0.1803*** (0.034)
neighbourhood_cleansed_11	-0.0900*** (0.021)	-0.1140*** (0.021)
neighbourhood_cleansed_12	0.2165*** (0.018)	0.1586*** (0.017)
neighbourhood_cleansed_13	0.1875*** (0.023)	0.0966*** (0.022)
neighbourhood_cleansed_16	0.1102*** (0.022)	0.1890*** (0.019)
property_type_1	-0.0664** (0.022)	-0.0536* (0.022)
property_type_2	-0.1144*** (0.017)	-0.1186*** (0.018)
property_type_3	0.1097*** (0.016)	0.1036*** (0.016)
property_type_4	0.0864*** (0.019)	0.0866*** (0.020)
property_type_5	-0.0774*** (0.017)	-0.0757*** (0.017)
property_type_6	-0.1311*** (0.026)	-0.1378*** (0.026)
property_type_8	0.0271* (0.013)	0.0310* (0.013)
bathrooms_2	0.1171*** (0.013)	0.1218*** (0.014)
bathrooms_3	0.1913*** (0.031)	0.1999*** (0.032)
bedrooms_2	0.0945*** (0.013)	0.0851*** (0.013)
bedrooms_3	0.1883*** (0.024)	0.1740*** (0.024)
Parking	0.0326*** (0.010)	0.0174 (0.010)
Child_Amenities	0.0672*** (0.010)	0.0709*** (0.010)
accommodates	0.1282*** (0.007)	0.1307*** (0.007)
availability_365	0.0816*** (0.007)	0.0785*** (0.007)
minimum_nights	-0.0111*** (0.003)	-0.0101** (0.003)
Review_count_H	-0.0605*** (0.003)	-0.0573*** (0.003)
review_scores_cleanliness	0.0264*** (0.004)	0.0255*** (0.004)
AIC	4,881	5,023
R ²	0.376	0.362
Adj.R ²	0.373	0.359
N	6,528	6,528



Summary statistics: BnBs in Venice

- Higher walkability is associated with higher prices
- BnB quality is positively related to the price
- Neighborhood San Polo appears to be significantly different from all the others
- Model without TWI presents higher AIC

	Model 1	Model 2
Intercept	4.8539*** (0.018)	4.9482*** (0.013)
Walkability_num_2	0.0759*** (0.021)	-
Walkability_num_3	0.1354*** (0.016)	-
Walkability_num_4	0.1405*** (0.018)	-
Walkability_num_5	0.1505*** (0.021)	-
Bayesian_avg_H	0.0870*** (0.005)	0.0855*** (0.005)
TVs	0.0532*** (0.010)	0.0573*** (0.010)
Bathroom_Amenities	0.0246* (0.011)	0.0249* (0.011)
neighbourhood_cleansed_4	0.1128*** (0.015)	0.1530*** (0.013)
property_type_1	-0.1771*** (0.023)	-0.1764*** (0.023)
property_type_4	0.1090*** (0.031)	0.1148*** (0.032)
bathrooms_2	0.1653*** (0.012)	0.1733*** (0.012)
bathrooms_3	0.3535*** (0.028)	0.3654*** (0.028)
bedrooms_2	0.0611*** (0.008)	0.0631*** (0.011)
beds_2	-0.0236* (0.011)	-0.0262* (0.011)
beds_3	-0.0263* (0.013)	-0.0279* (0.013)
amenities_count	0.0434*** (0.008)	0.0401*** (0.008)
Review_count_H	-0.0972*** (0.005)	-0.0931*** (0.005)
location_score	0.0163*** (0.003)	0.0210*** (0.003)
accommodates	0.1628*** (0.007)	0.1619*** (0.007)
maximum_nights	0.0595*** (0.010)	0.0609*** (0.011)
availability_365	0.0708*** (0.008)	0.0747*** (0.008)
AIC	2,276	2,358
R ²	0.458	0.447
Adj.R ²	0.456	0.445
N	4,535	4,535



- Both Naples and Venice offer good accessibility, with key amenities within a 15-minute walk of at least one accommodation.