

Case Study

Data Science



Franco Troncoso

Problem statement

Dream Inc. has partnered with the city of Freiburg in Baden-Württemberg to improve vehicle parking experience by understanding the demand patterns and developing an app which tells the customer if a parking slot will be available on arrival time.

Specific tasks

- 👉 Data analysis.
- 👉 Parking garages comparison
- 👉 External data incorporation.
- 👉 Parking availability prediction model building.



Train station Garage



272 Places

City Center Garage



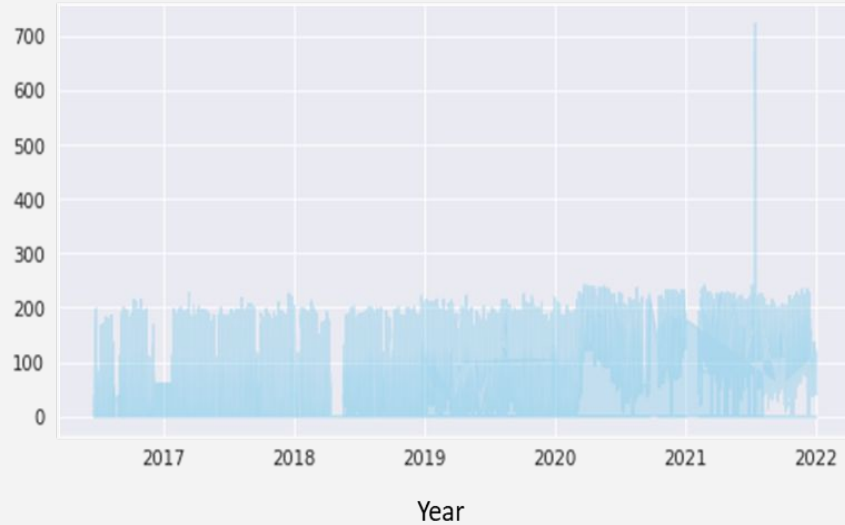
516 Places

Data

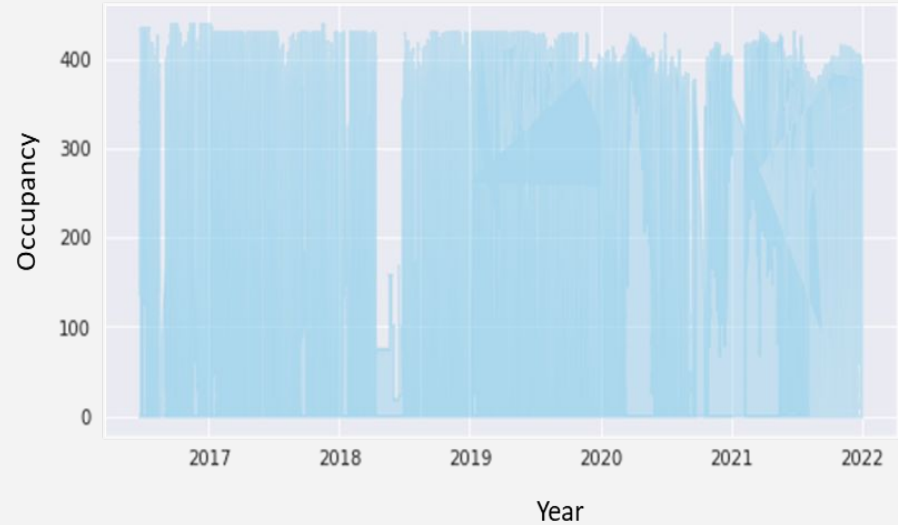
Parking Data: Occupancy 2016-2021

Extra Data: Holiday & Weather Forecast

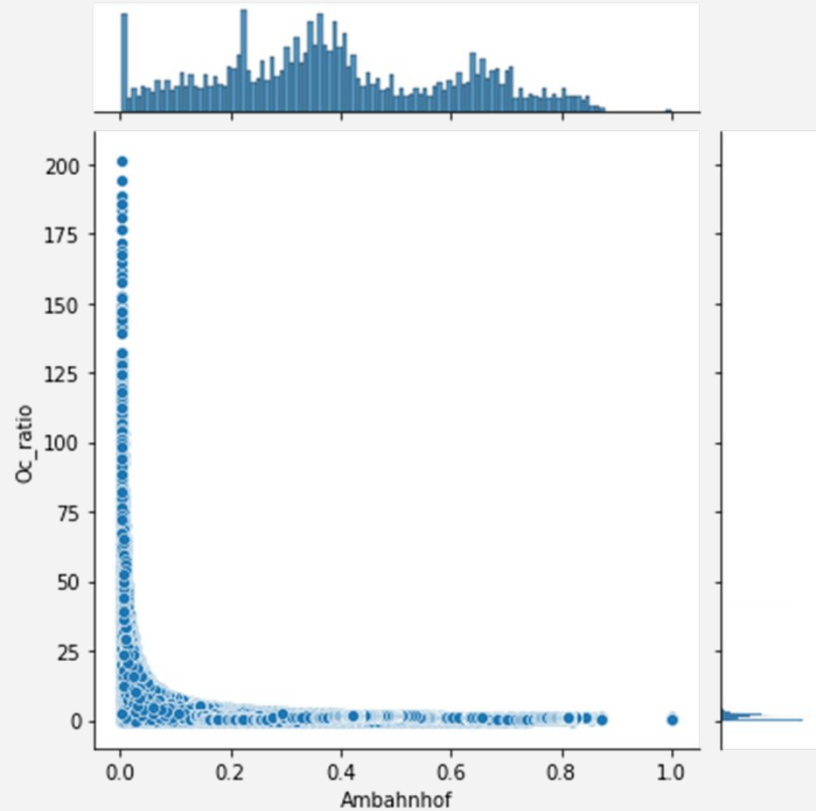
Train Station



City Center



Parking Garage Comparison

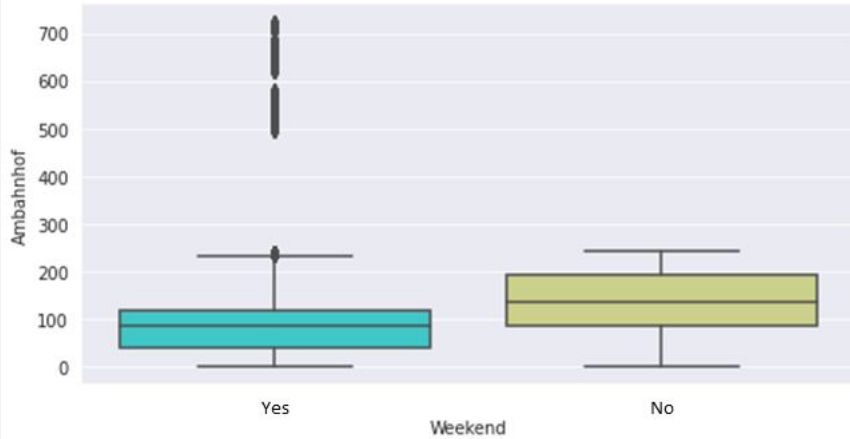


$$\text{Occupation Ratio} = \frac{\text{City Center Garage Occupancy}}{\text{Train Station Garage Occupancy}}$$

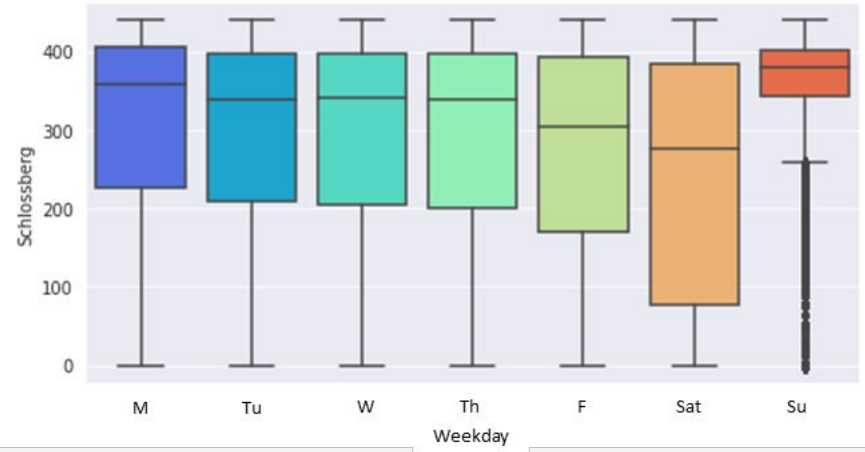
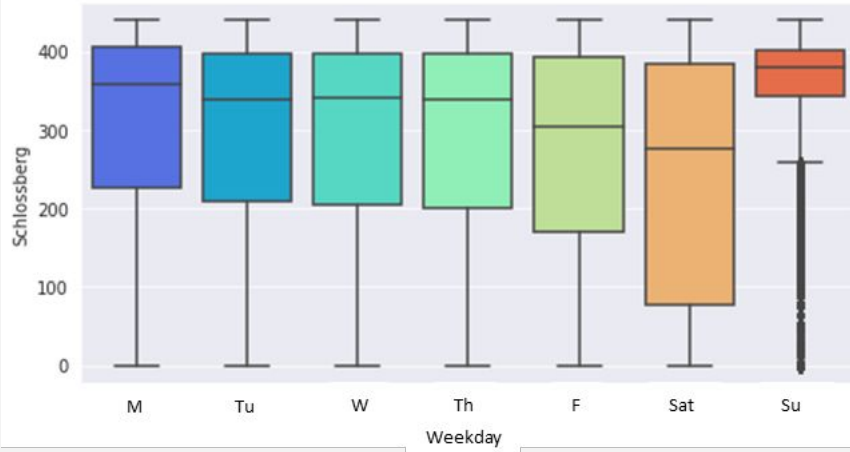
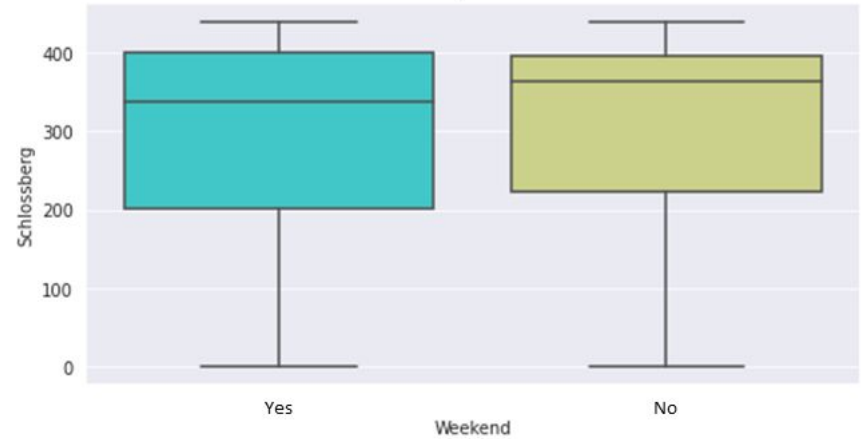
🔍 City center has higher occupation rate (99.99% data)

Boxplot

Train Station

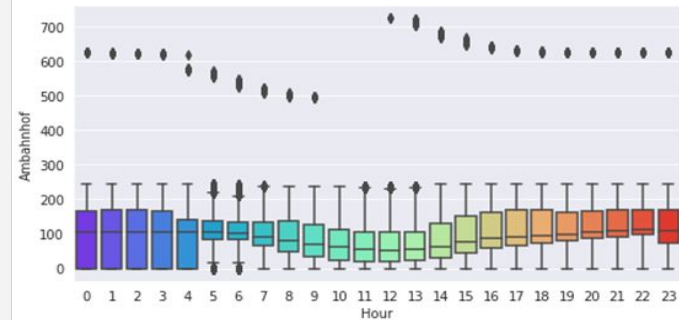
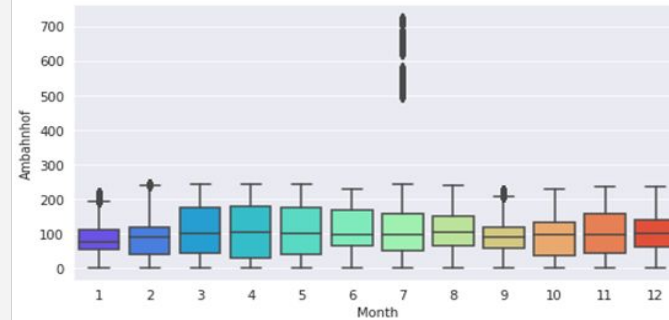
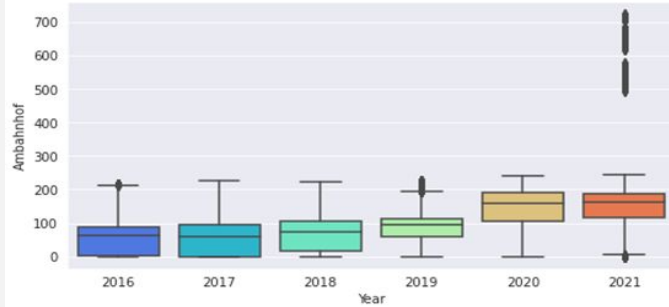


City Center

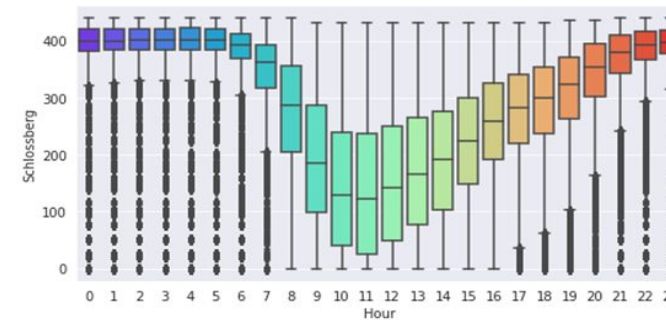
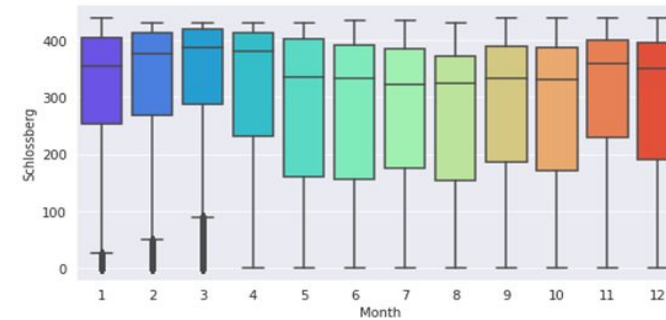


Boxplot

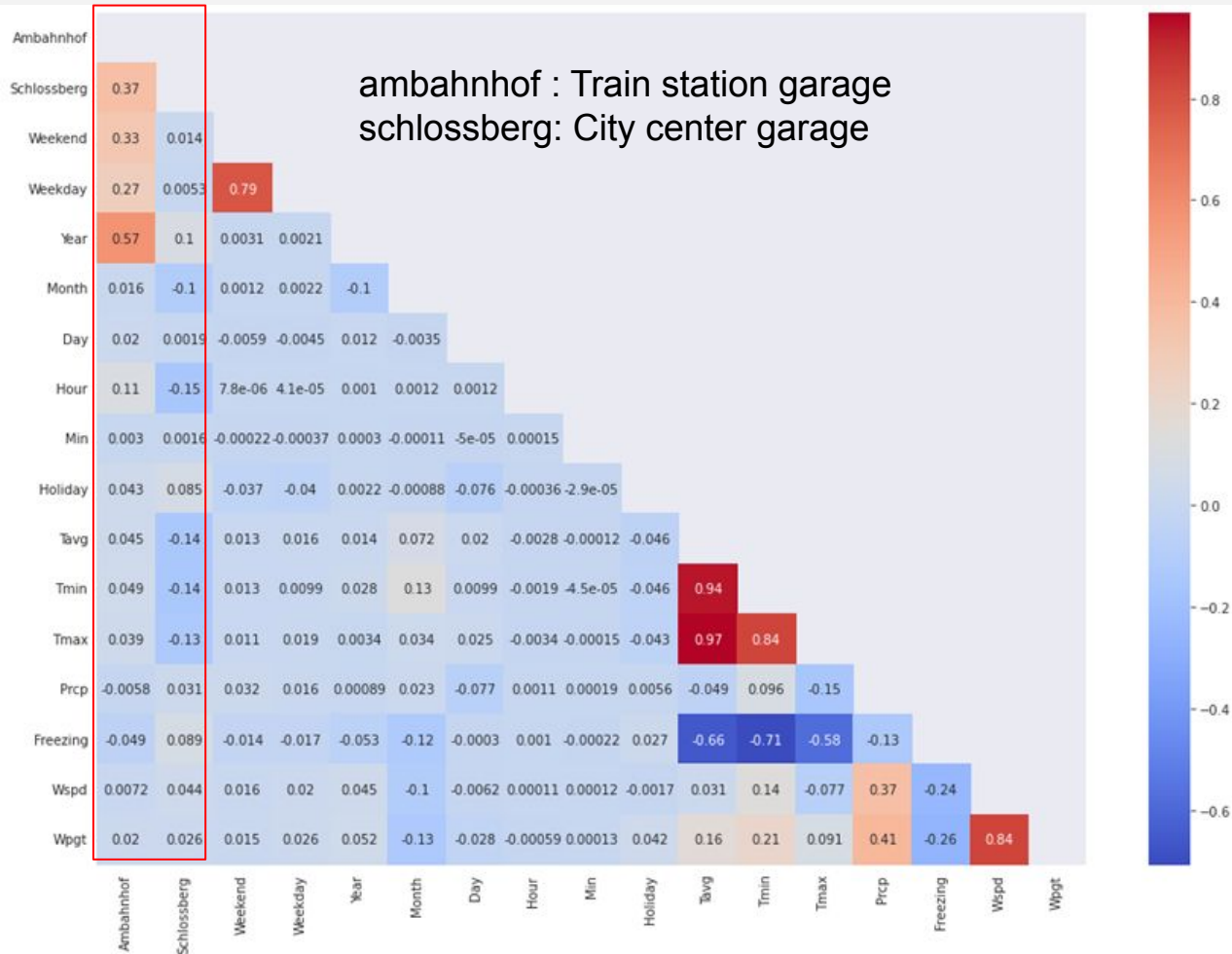
Train Station



City Center



Correlation Matrix Heatmap



Features with more predictive power

Train station garage:

1. Year (0.57)
2. Weekend (0.33)
3. Weekday (0.27)
4. Hour (0.11).

City Center garage:

1. Hour (-0.15)
2. Tmax (-0.14)
3. Tmin(-0.14)
4. Tavg(-0.13)
5. Month(-0.1)
6. Year (0.1)

Quantile Processing

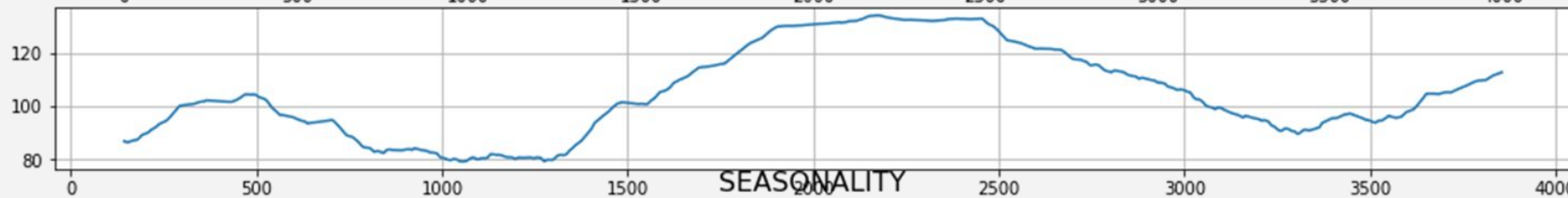
Feature	Train Station			City Center		
	Before	After	Abs. Dif [%]	Before	After	Abs. Dif [%]
Weekend	0.33	0.3	9.1	0	0	0
Weekday	0.27	0.26	3.7	0	0	0
Year	0.57	0.57	0	0.1	0.061	39.0
Month	0	0	0	0.1	0.12	20.0
Hour	0.11	0.12	9.1	0.15	0.22	46.7
Tavg	0	0	0	0.13	0.16	23.1
Tmin	0	0	0	0.14	0.17	21.4
Tmax	0	0	0	0.14	0.14	0

Seasonal Behaviour

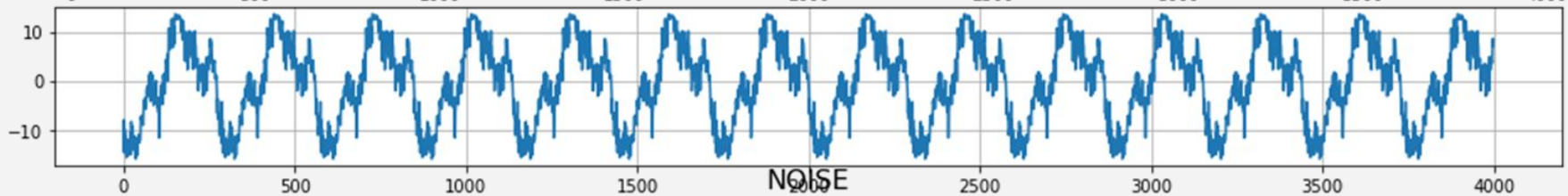
OBSERVED



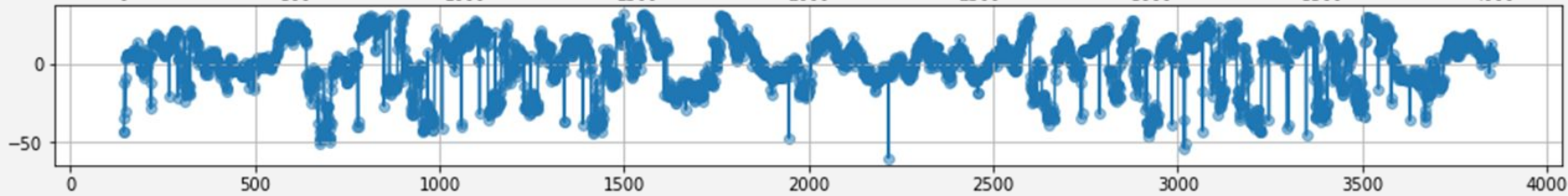
TREND



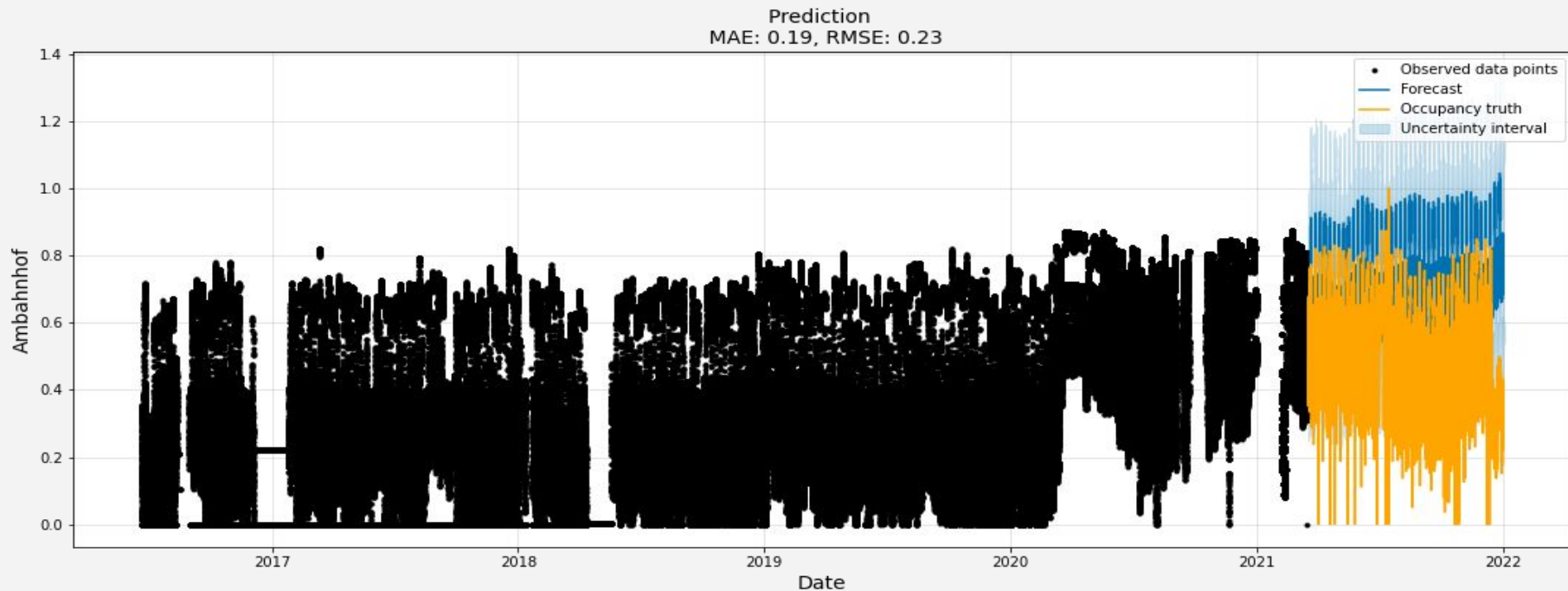
SEASONALITY



NOISE

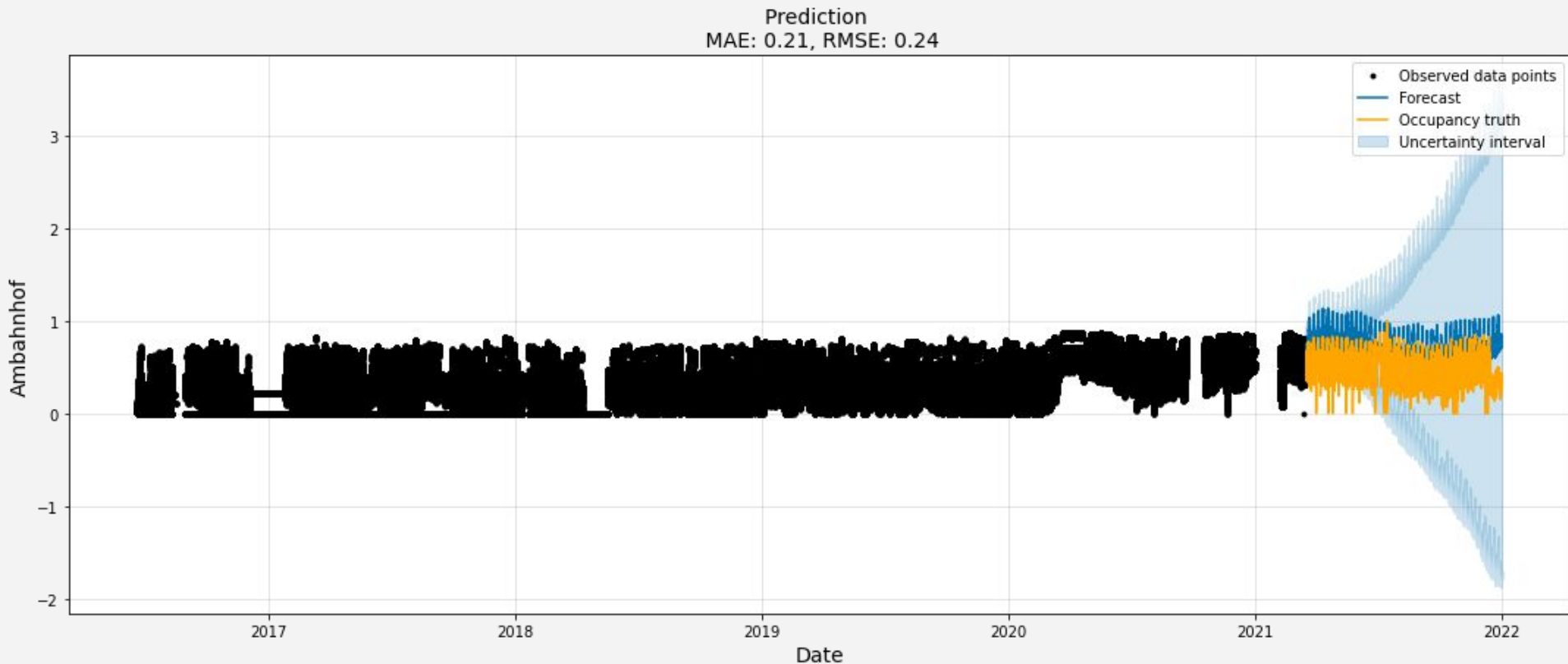


Univariate Prophet



Hyperparameters: changepoint_prior_scale=0.001, b. seasonality_prior_scale=10, yearly_seasonality=True,
weekly_seasonality=True, daily_seasonality=True, interval_width=0.90

Multivariate Prophet

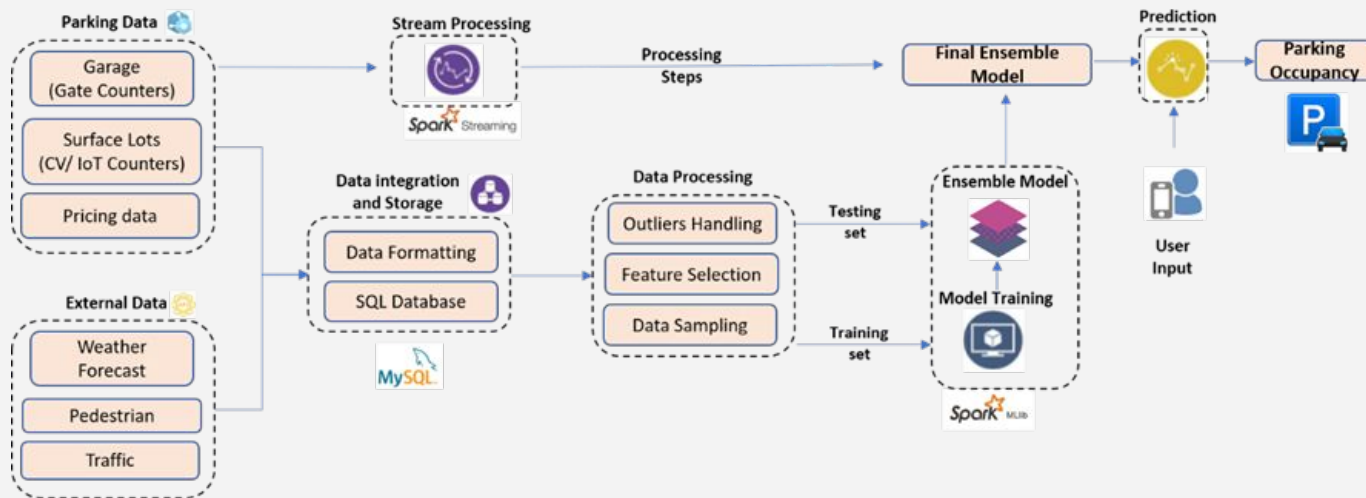


Hyperparameters: changepoint_prior_scale=0.001, b. seasonality_prior_scale=10, yearly_seasonality=True, weekly_seasonality=True, daily_seasonality=True, interval_width=0.90

Conclusions

- 🔍 Exploration of data anomalies in train station garage is required.
- 🔍 Different data range selection had a paramount effect on model metrics, surpassing even the effect of the hyperparameters tuning.
- 🔍 A different seasonal behavior was identified from the second half of 2020 to first quarter of 2021 for Train Station garage, leading to a significant increment in daily minimum occupation. This behavior was followed by a decreasing in occupation rate of train garage.
- 🔍 Train station occupation showed strong seasonal behavior. City Center garage is influenced by weather temperature.
- 🔍 The appropriate selection of training and validation data represent the most important parameter to enhance the model performance, followed by time shift and hyperparameter tuning. For this reason, ensemble method is recommended

Data Pipeline



Suggested future work

- 👉 Explore how the data was collected and outlier generation cause.
- 👉 Test the assumption performed in this work, especially the total occupancy
- 👉 Analyze the effect of pedestrian and traffic data.
- 👉 Use ensemble method to build a appropriate method for each garage.
- 👉 Set the appropriate data for each prediction (weeks/months)

THANK YOU

GRACIAS
ARIGATO
SHUKURIA
JUSPAXAR
DANKSCHEEN
TASHAKKUR ATU
YAQHANYELAY
SUKSAMA
EKHMET
BIYAN
SHUKRIA
TINGKI
MAAKE
GRAZIE
MEHRBANI
PALDIES
BOLZİN
MERCI
GOZAIMASHITA
EFCHARISTO
KOMAPSUMNIDA
LAH
FAKAAUE
AGUYJE
SAIKO
GAJEJTHO
MERASTAWHY
TAVTAPUCH
MEDAHAGSE
BAIKKA
SPASSIBO
SHACHALHUYA
NUHUN
CHALTU
WADEEJA
MAITEKA
HUI
YUSPAGADATAM
DHANYABAD
ANNA
ATTO
SPASIBO
DENKAUJA
NEHACHALHYA
UNALCHEESH
HATUR GUR
EXOUJI
SIKOMO
MAKETAI
MINMONCHAR