

# STATISTIEK VLAANDEREN

## Strava gegevens als bron van statistieken Kasimir Putseys (VSA)



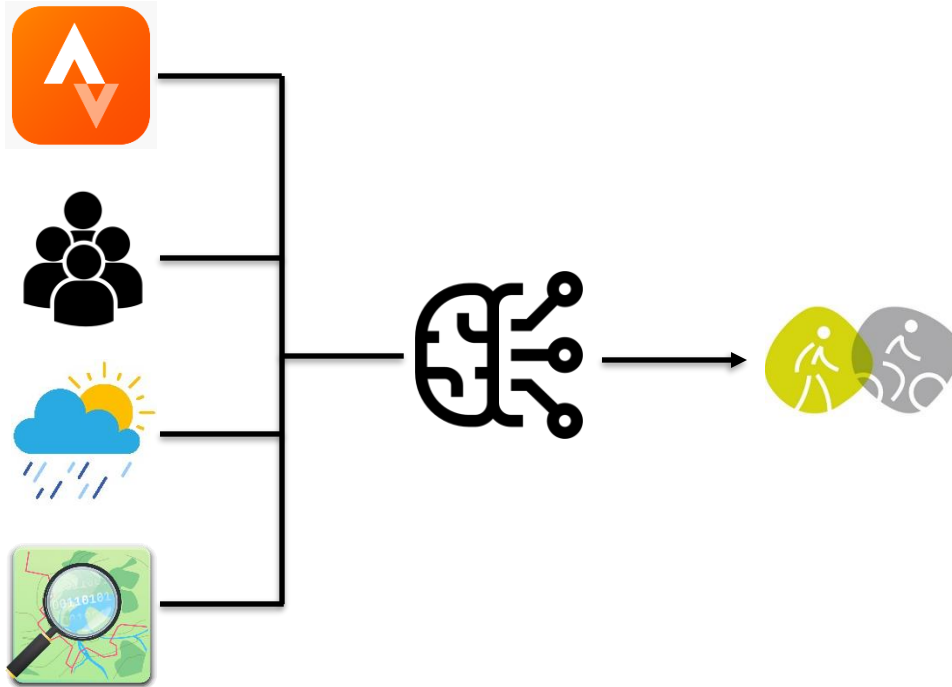
# Statistiek vraag

- ➔ Wekelijks sportgebruik in Vlaanderen
  - ⊕ Sport Vlaanderen heeft 25 fietsentellers
  - ⊕ Hoeveel wordt er gefietst op andere routes?
    - > Machine learning model om dit te voorspellen.



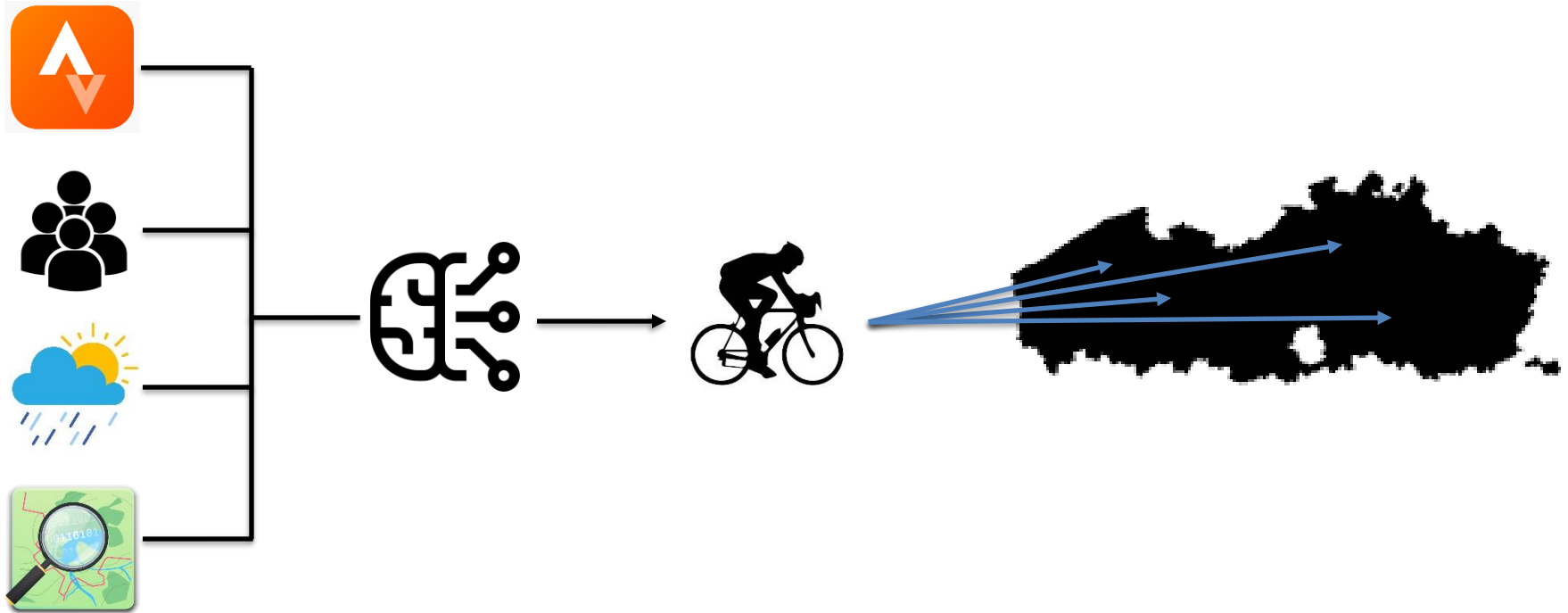
# Methode

## ➔ Training

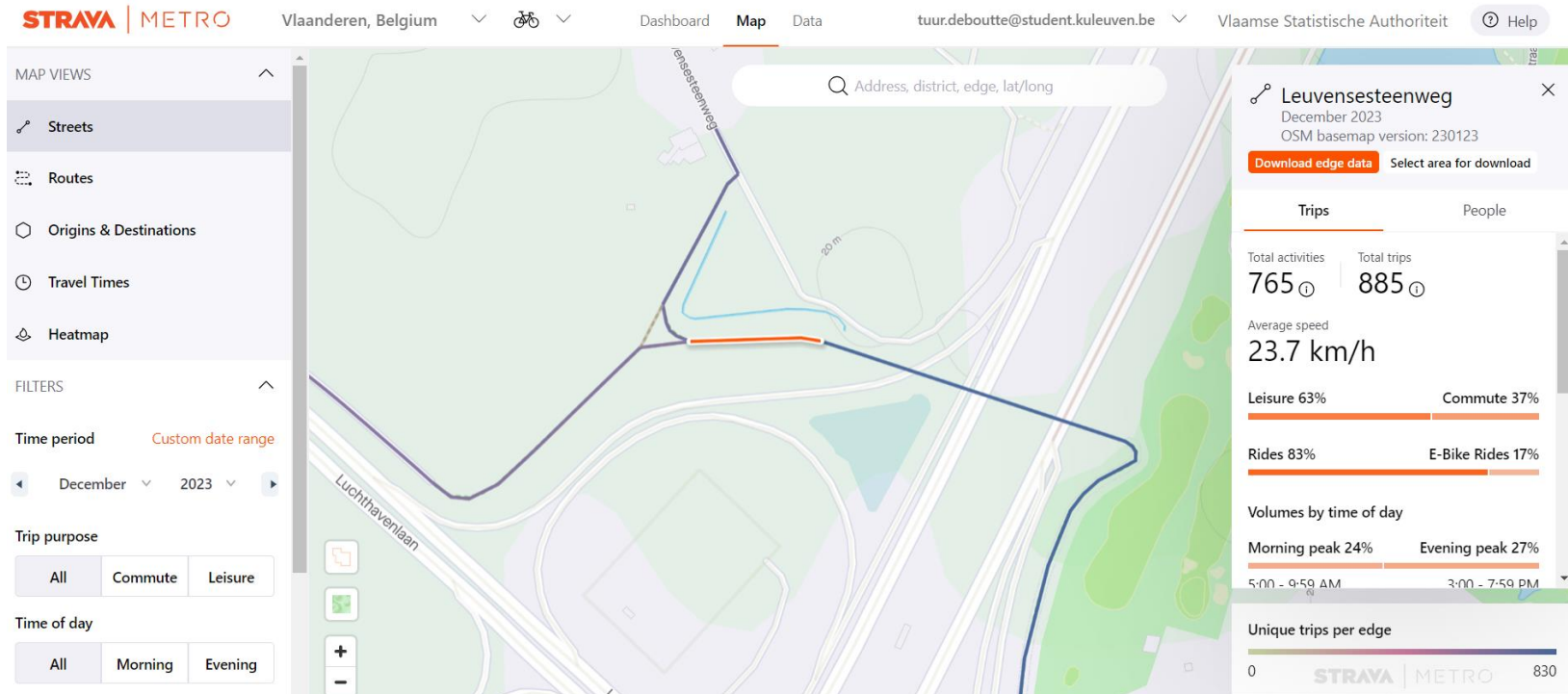


# Methode

## ➔ Predicting



# ➔ Strava Metro



Tellocaties

Domein

Dashboard

Niet-directionele site-analyse

?

+ Nieuw Dashboard

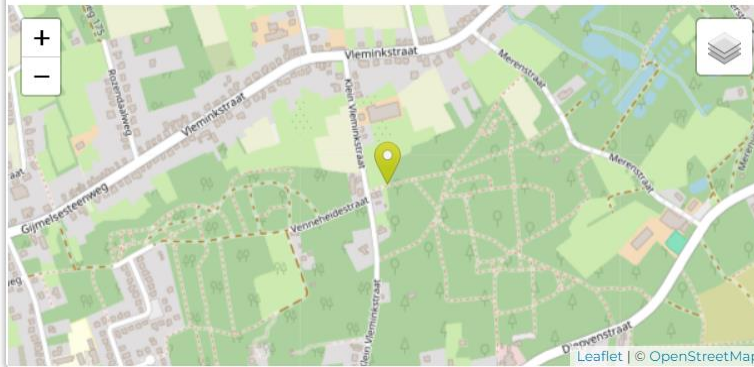
Aarschot

Huidig Jaar — 1 januari 2024 → 13 maart 2024

+ New widget

Plaats

Aarschot



Fotogallerij

Aarschot



Huidig Jaar

Aarschot

Dagelijks  
gemiddelde  
**27**

**-21,9%**

Vergeleken met Vorig jaar

Huidig Jaar

Aarschot

Drukste dag

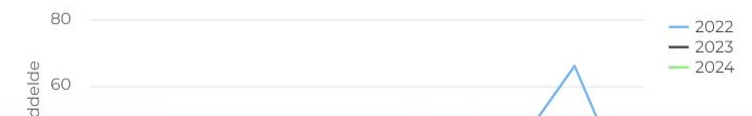
**167**

zondag  
14 jan. 2024

Jaarlijkse vergelijking per maand

@3y → @0d

Aarschot



## ➔ Pre-Processing



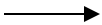
## ➔ Pre-Processing



date	location	total_trip_count
2021-03-26	lommel	15
2023-10-15	lommel	25
2022-11-06	lommel	25
2020-04-05	lommel	25
2022-06-18	lommel	10
2022-08-08	lommel	10
2020-04-21	lommel	20
2021-06-28	lommel	5
2020-04-16	lommel	20



# ➔ Pre-Processing



refnis_name	Age 0-17	Age 18-34	Age 35-54	Age 55-64	Age 65+
Aalst	18010	17927	23379	12281	17557
Aalter	5321	5599	7693	4218	6379
Aarschot	5155	5802	8033	4596	6929
Aartselaar	2693	2609	3592	2103	3494
Affligem	2893	2375	3676	1852	2679
Alken	2063	2184	2992	1891	2580
Alveringem	1037	1091	1266	694	884
Antwerpen	121863	131384	140642	57057	82220
Anzegem	3000	3038	3977	1994	3085
Ardooi	1771	1778	2270	1298	2103
Arendonk	2569	2575	3522	1941	2604
As	1518	1616	2169	1249	1618
Asse	7969	7304	9189	4304	6129
Assenede	2607	2759	3718	2236	3107
Avelgem	2082	1975	2610	1408	2077

# ➔ Pre-Processing



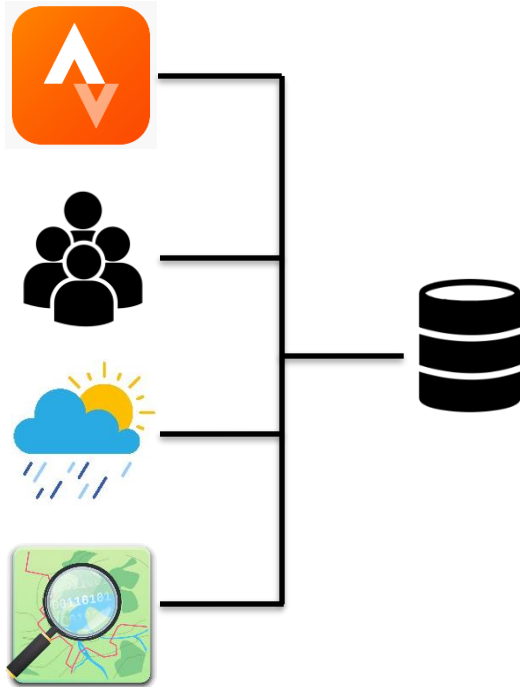
datetime	temp	prcp	wspd	tsun
2019-12-05 00:00:00	-1.4	0.0	3.6	0.0
2019-12-05 01:00:00	-1.7	0.0	3.6	0.0
2019-12-05 02:00:00	-1.7	0.0	3.6	0.0
2019-12-05 03:00:00	-1.7	0.0	7.2	0.0
2019-12-05 04:00:00	-2.1	0.0	3.6	0.0
2019-12-05 05:00:00	-2.7	0.0	3.6	0.0
2019-12-05 06:00:00	-2.9	0.0	3.6	0.0
2019-12-05 07:00:00	-3.2	0.0	3.6	0.0
2019-12-05 08:00:00	-3.1	0.0	7.2	0.0
2019-12-05 09:00:00	-0.9	0.0	3.6	0.0
2019-12-05 10:00:00	0.0	0.0	10.8	0.0
2019-12-05 11:00:00	1.5	0.0	10.8	0.0
2019-12-05 12:00:00	2.0	0.0	7.2	0.0

## ➔ Pre-Processing



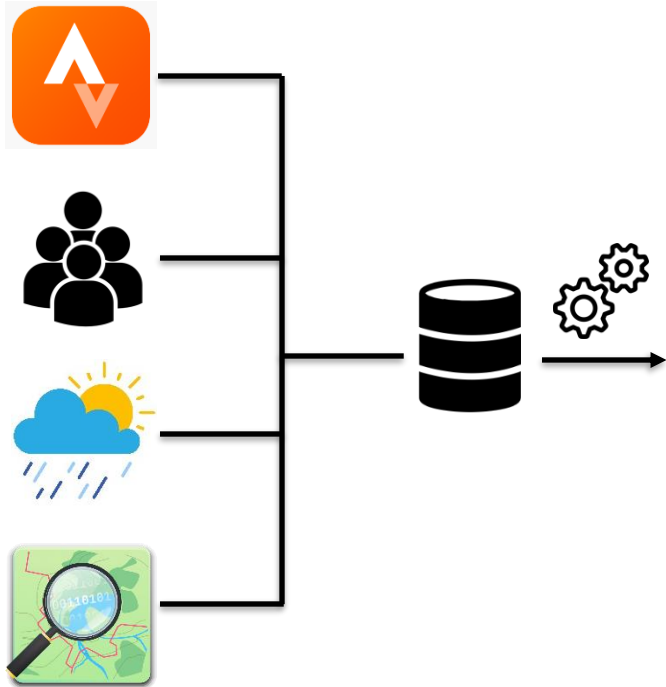
# Methode

## ➔ Pre-Processing

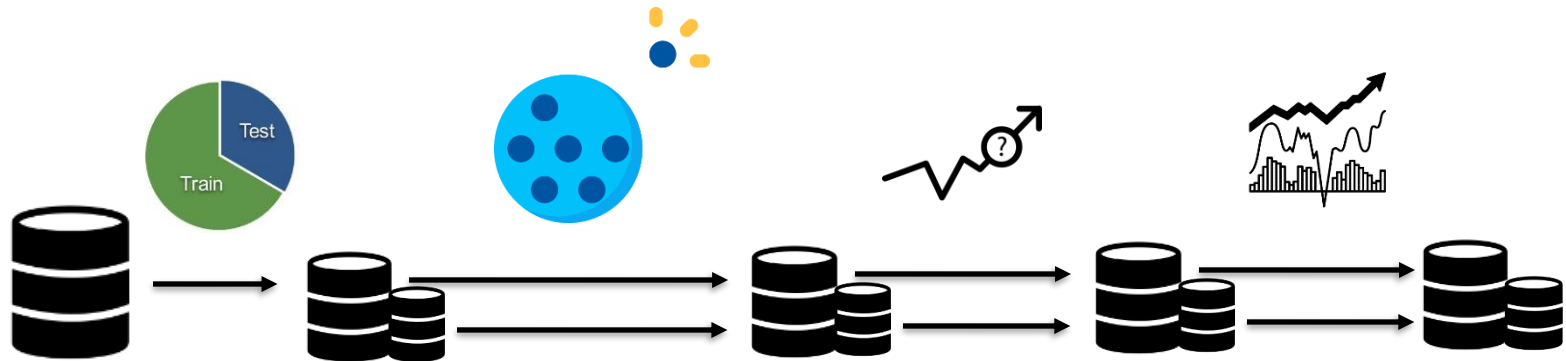


# Methode

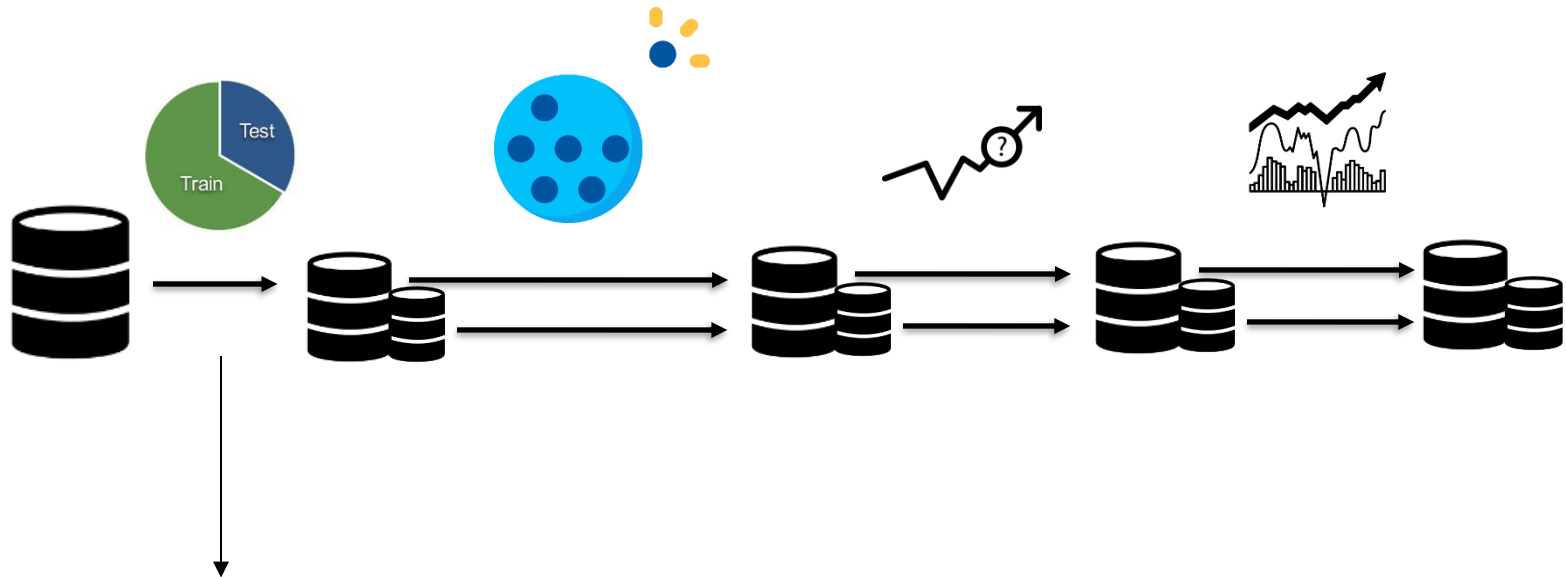
## ➔ Pre-Processing



# ➔ Pre-Processing

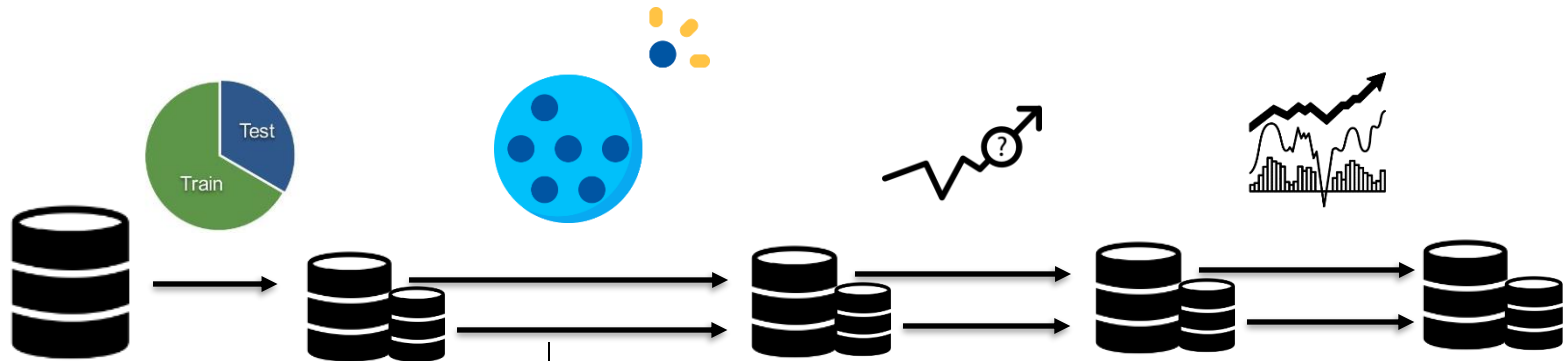


## ➔ Pre-Processing



- Train: 80% of counters, dates until 2023-01-01
- Test: 20% of counters, dates after 2023-01-01

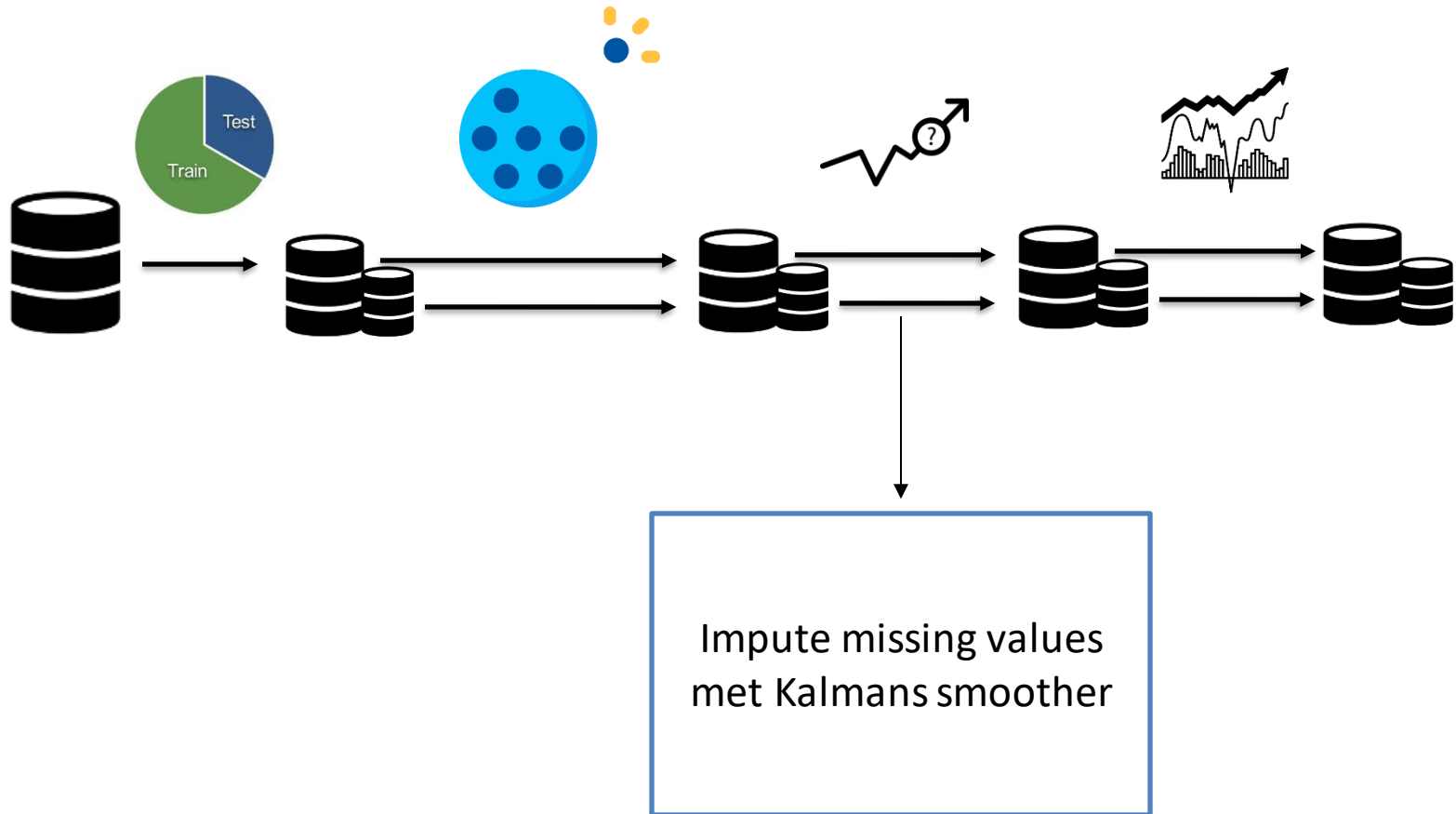
## ➔ Pre-Processing



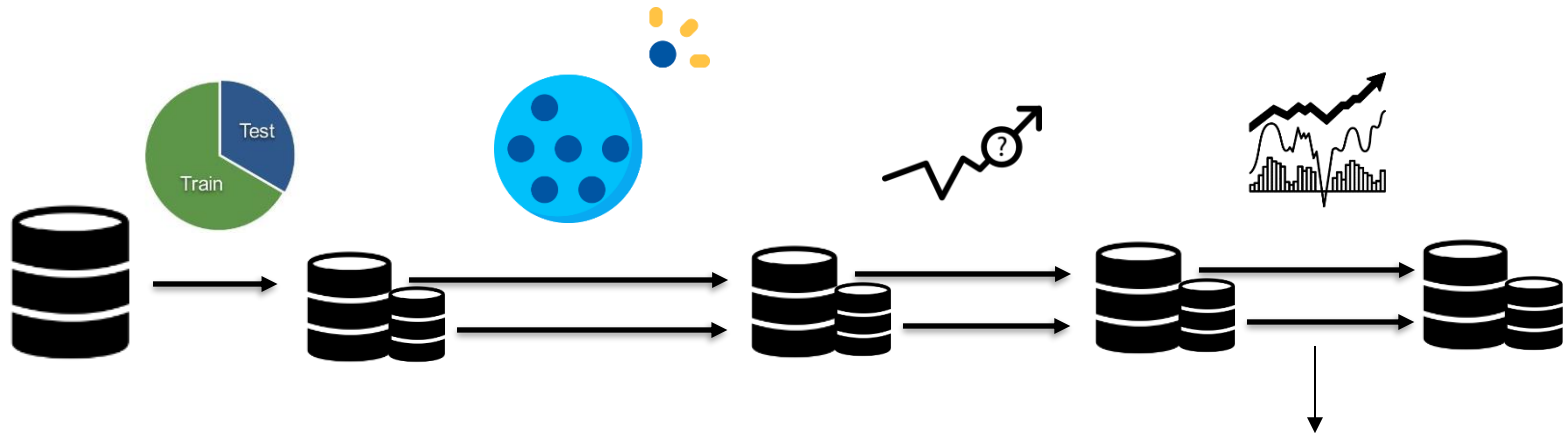
Detecteer outliers voor elke locatie en weergeef deze als missing value



## ➔ Pre-Processing



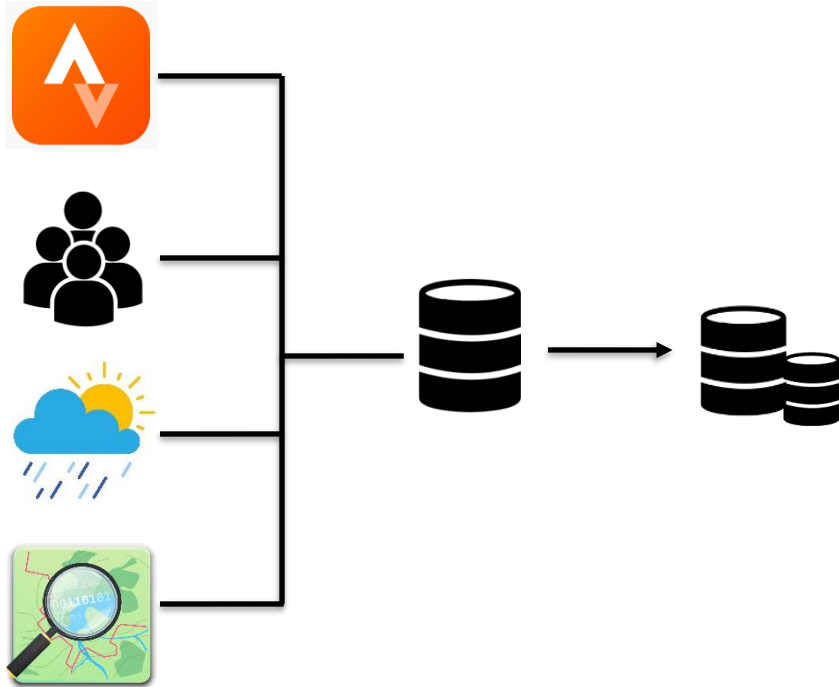
## ➔ Pre-Processing



- Creëer lags:  $t-1$ ,  $t-2$ ,  $t-3$ ,  $t-7$ ,  $t-365$
- creëer `is_event` variable om speciale events te detecteren met ARIMA model
- Aangeven of een dag een Belgische feestdag is

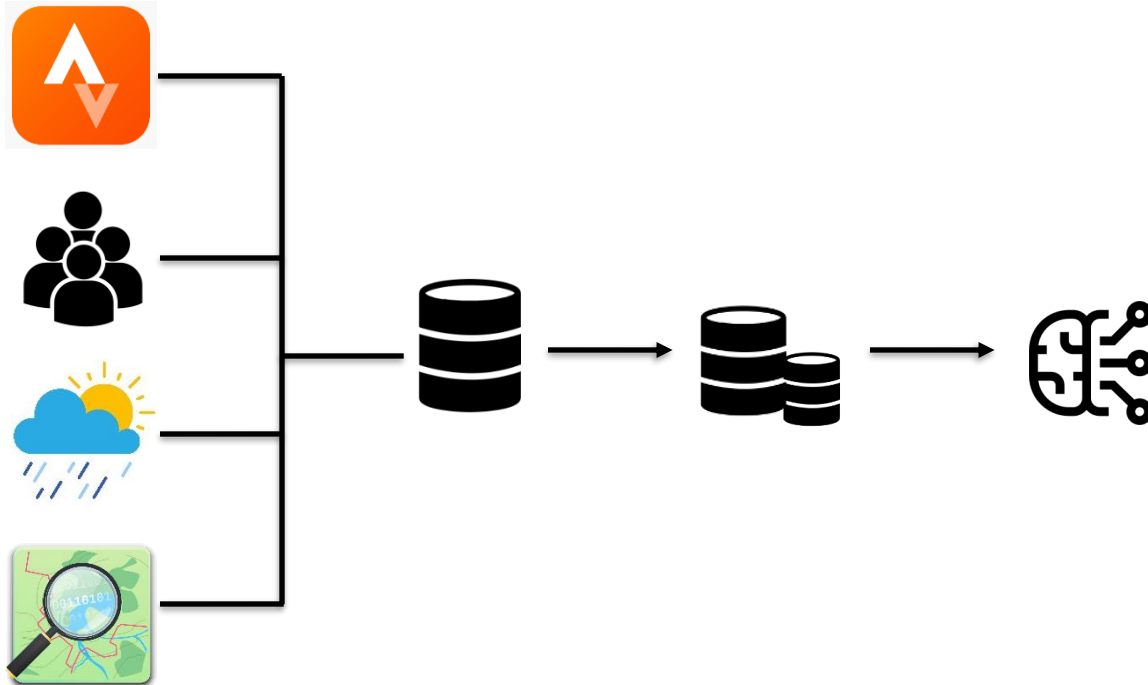
# Methode

## ➔ Pre-Processing



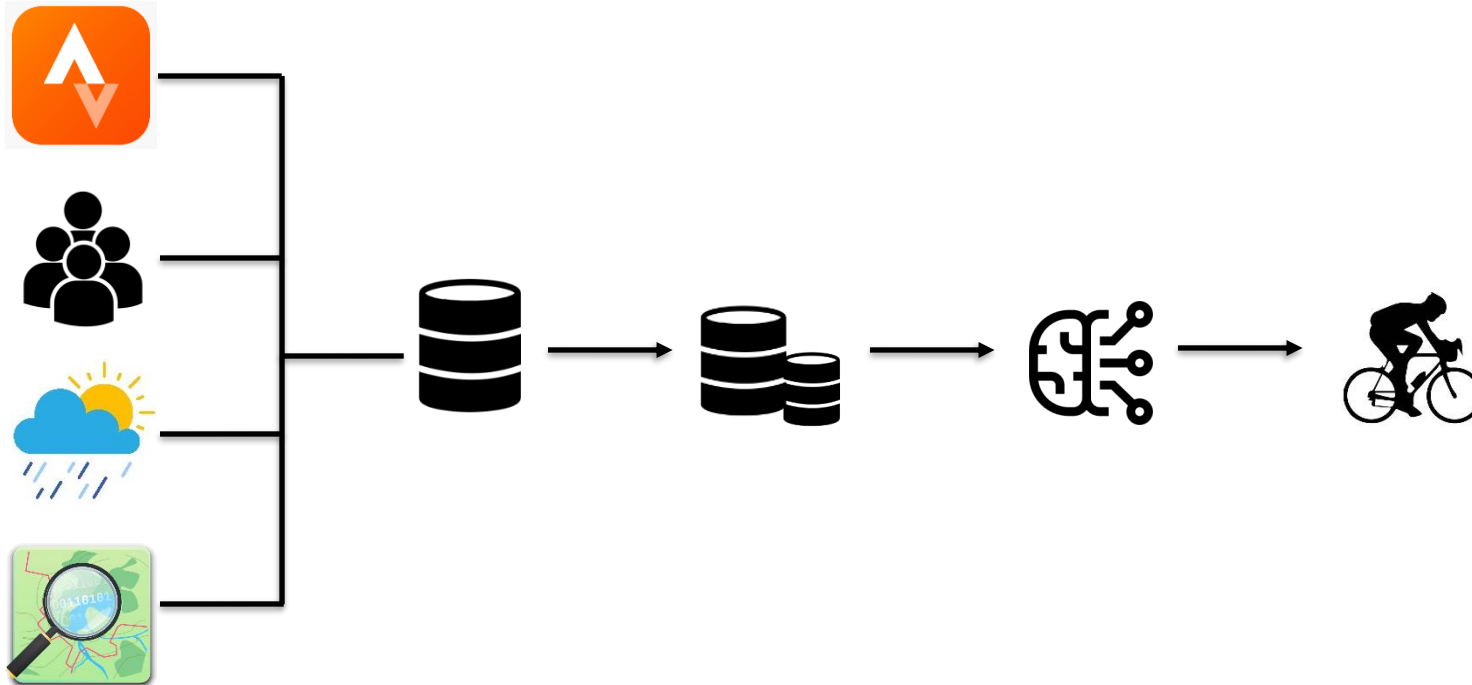
# Methode

## ➔ Pre-Processing



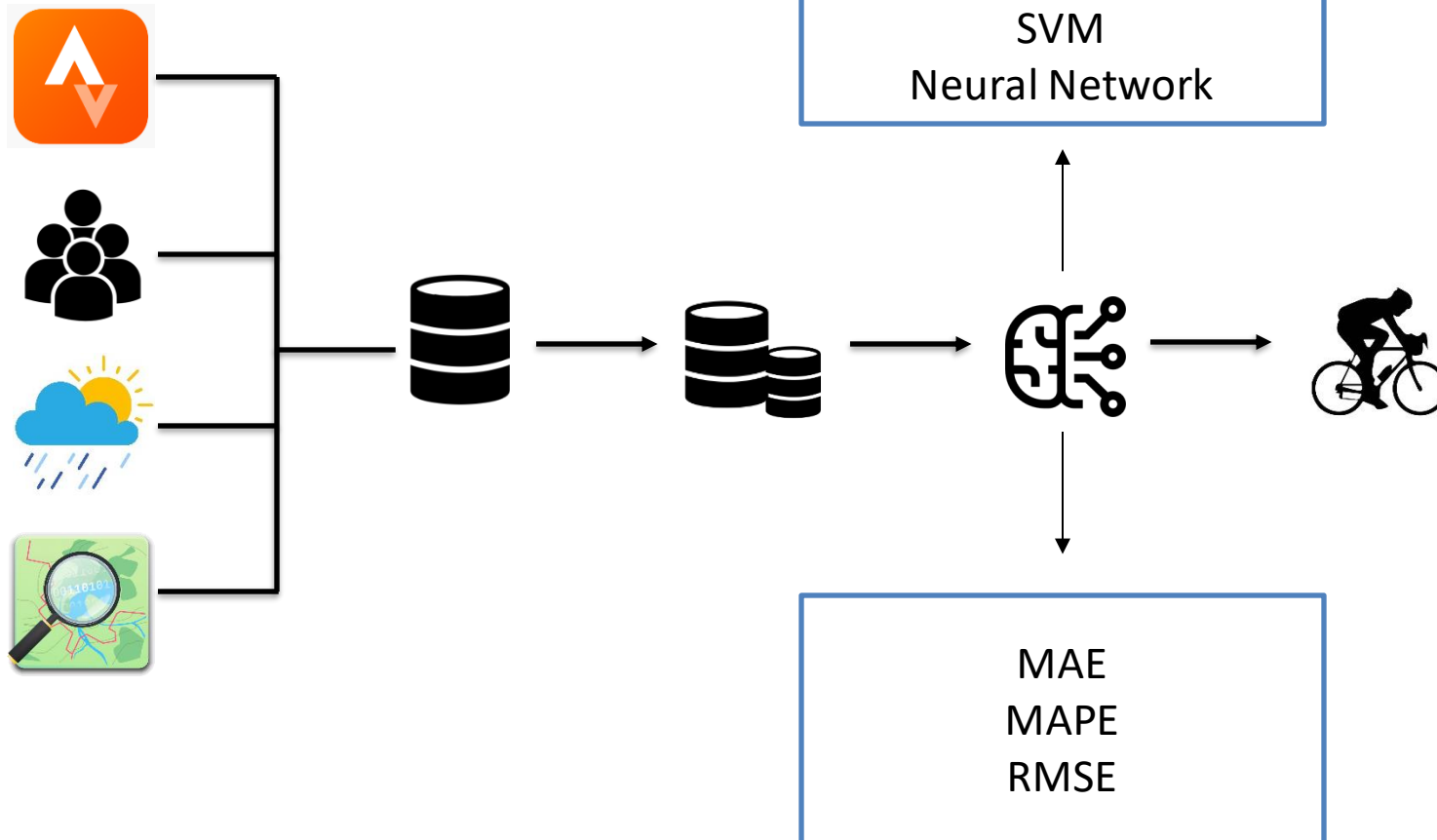
# Methode

## ➔ Pre-Processing

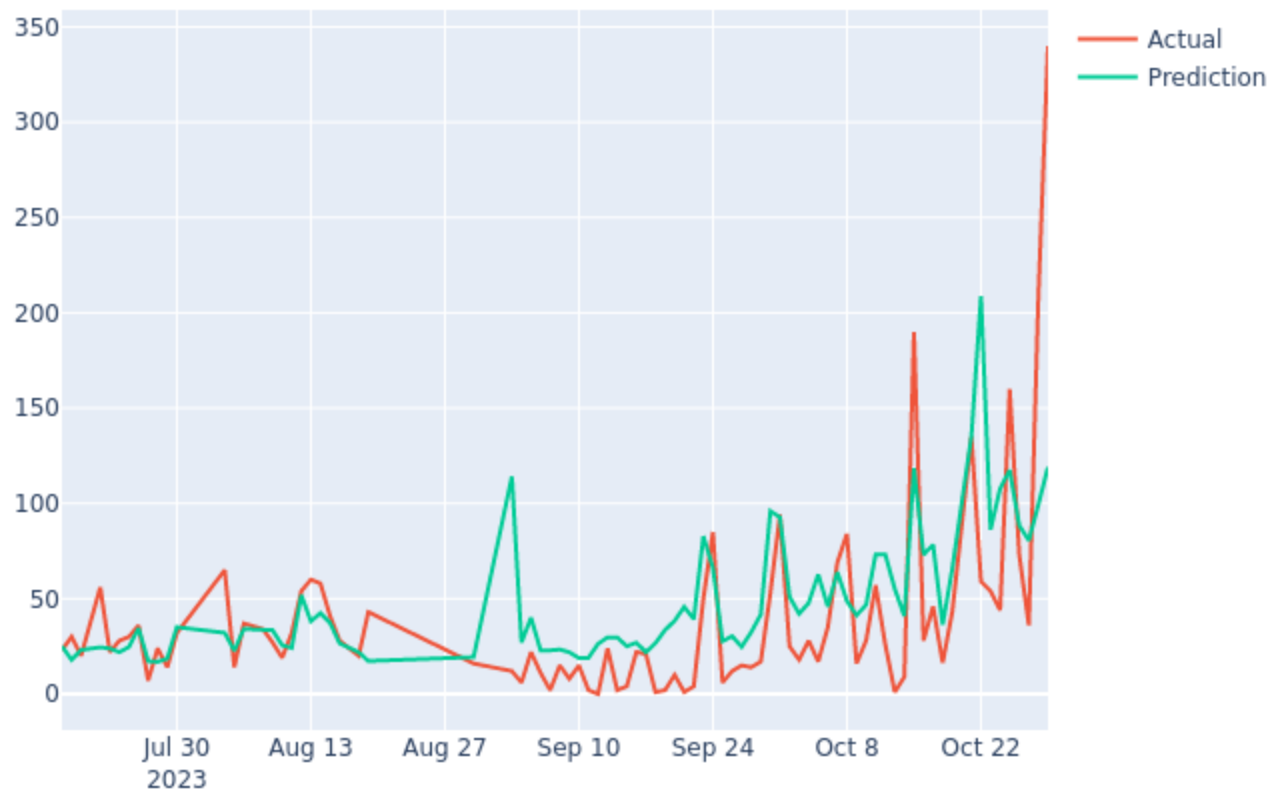


# Methode

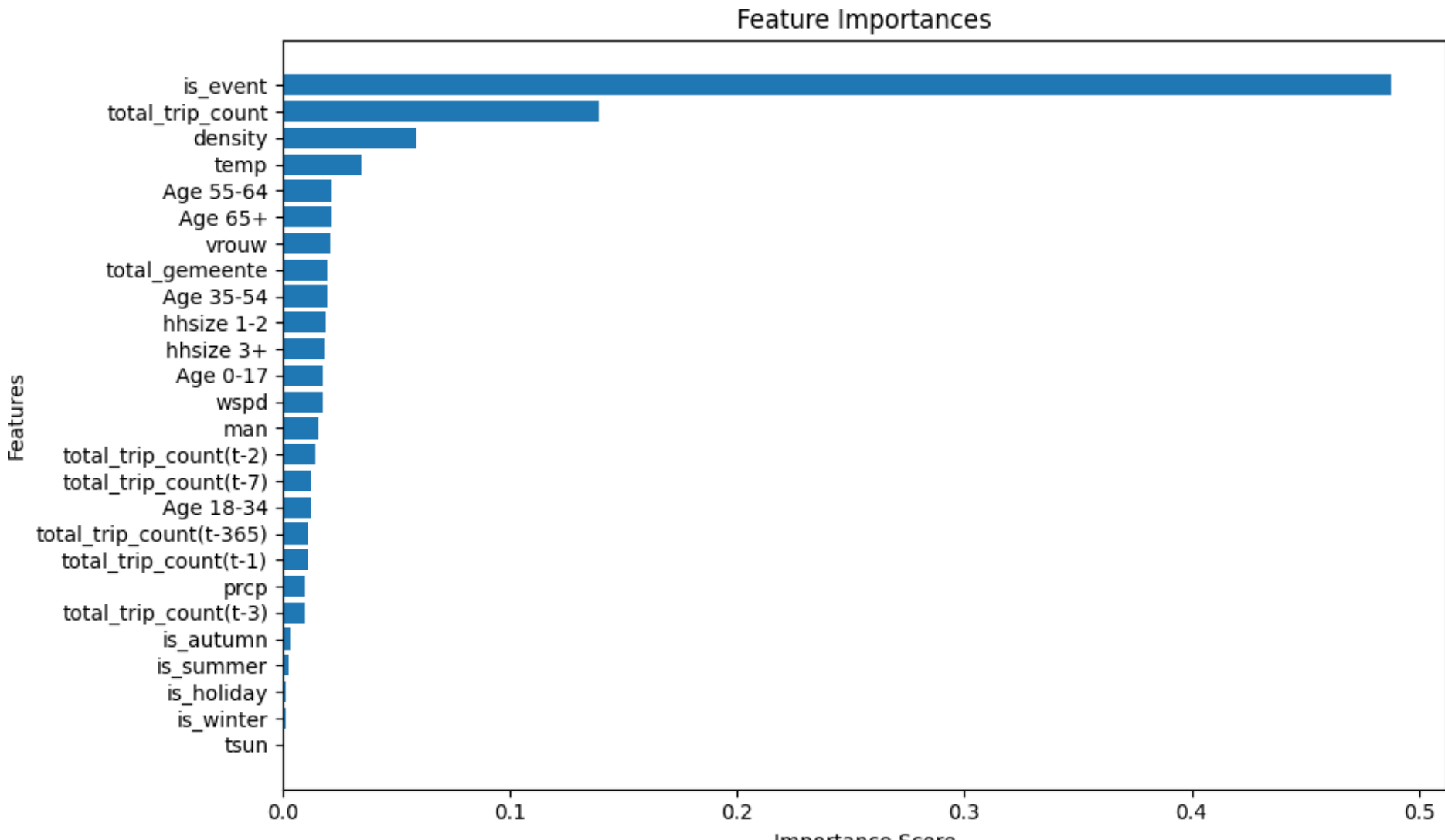
## ➔ Pre-Processing



## ➔ Random Forest

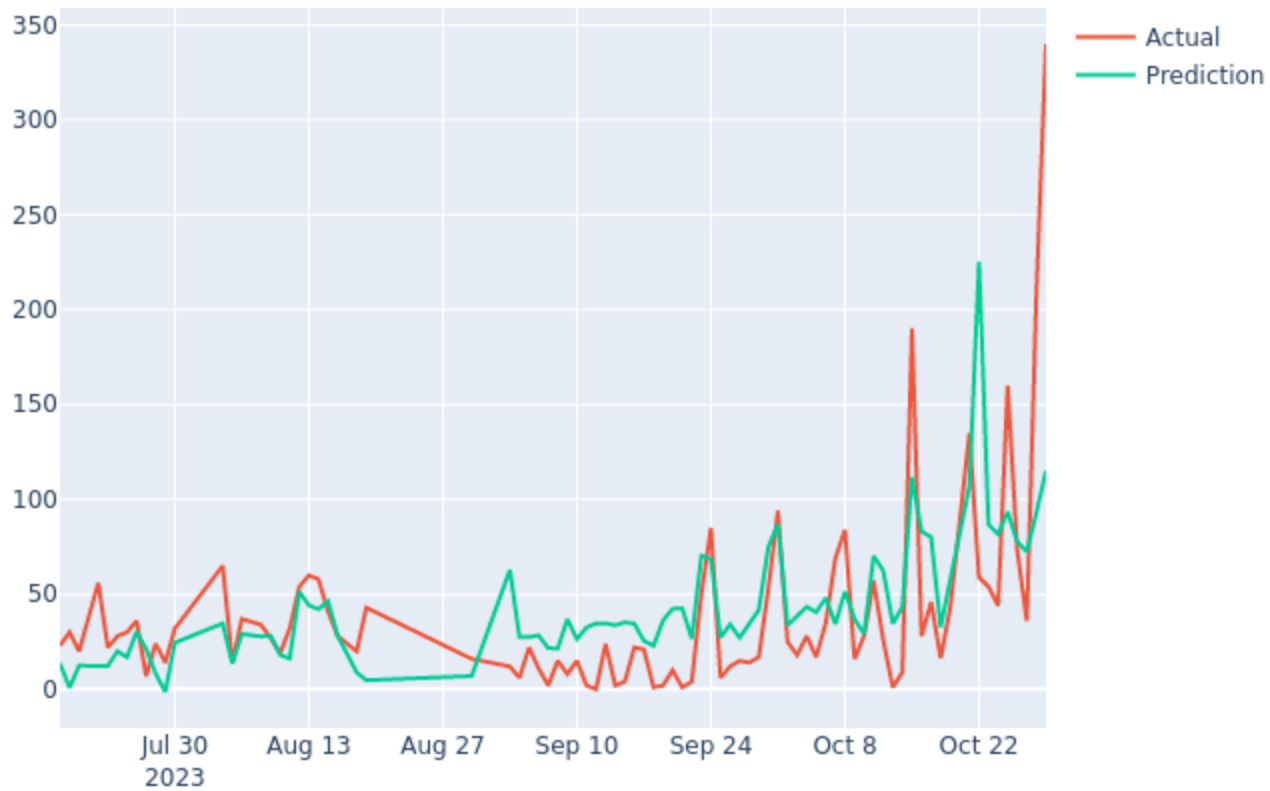


# ➔ Feature Importance

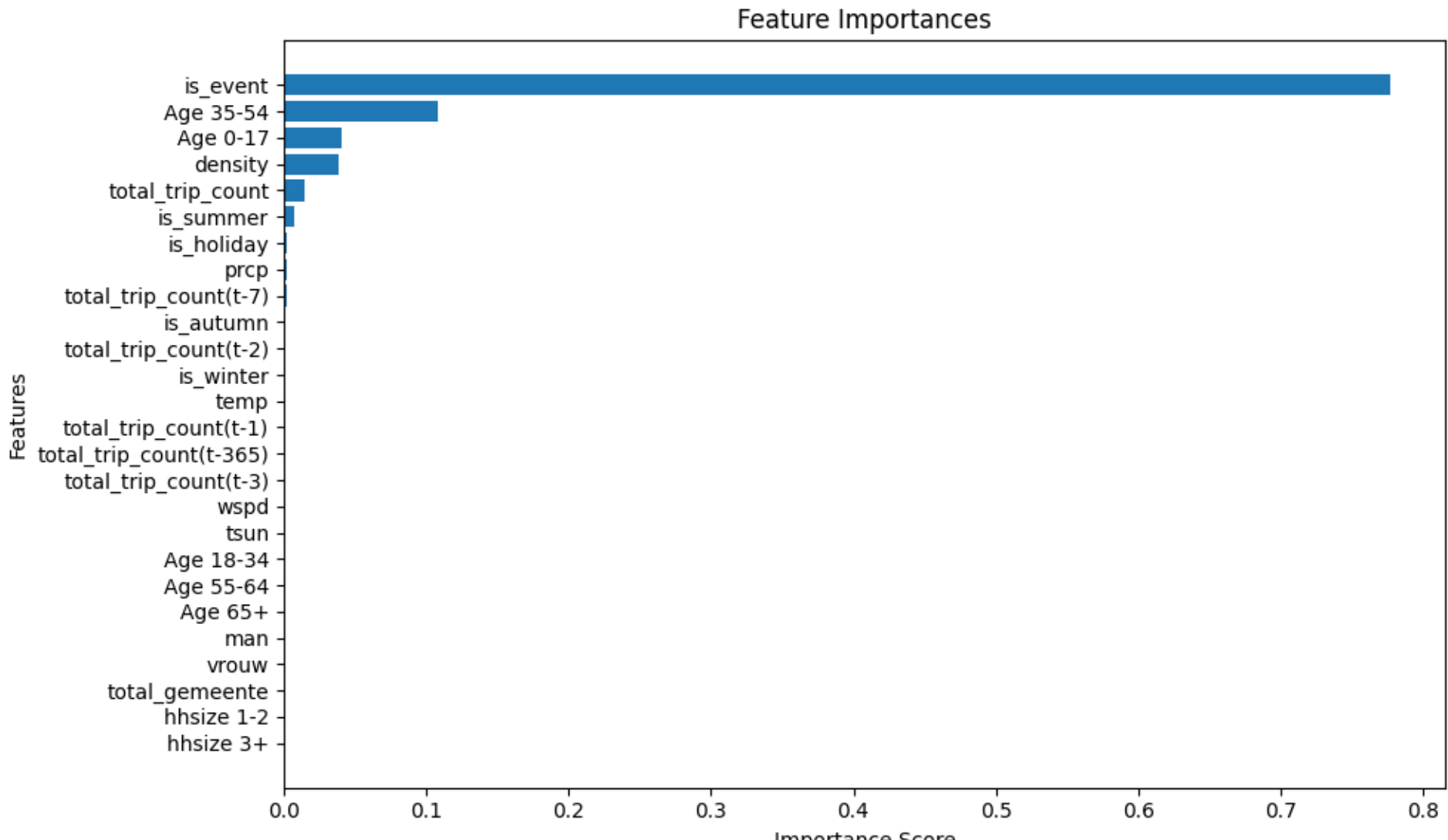




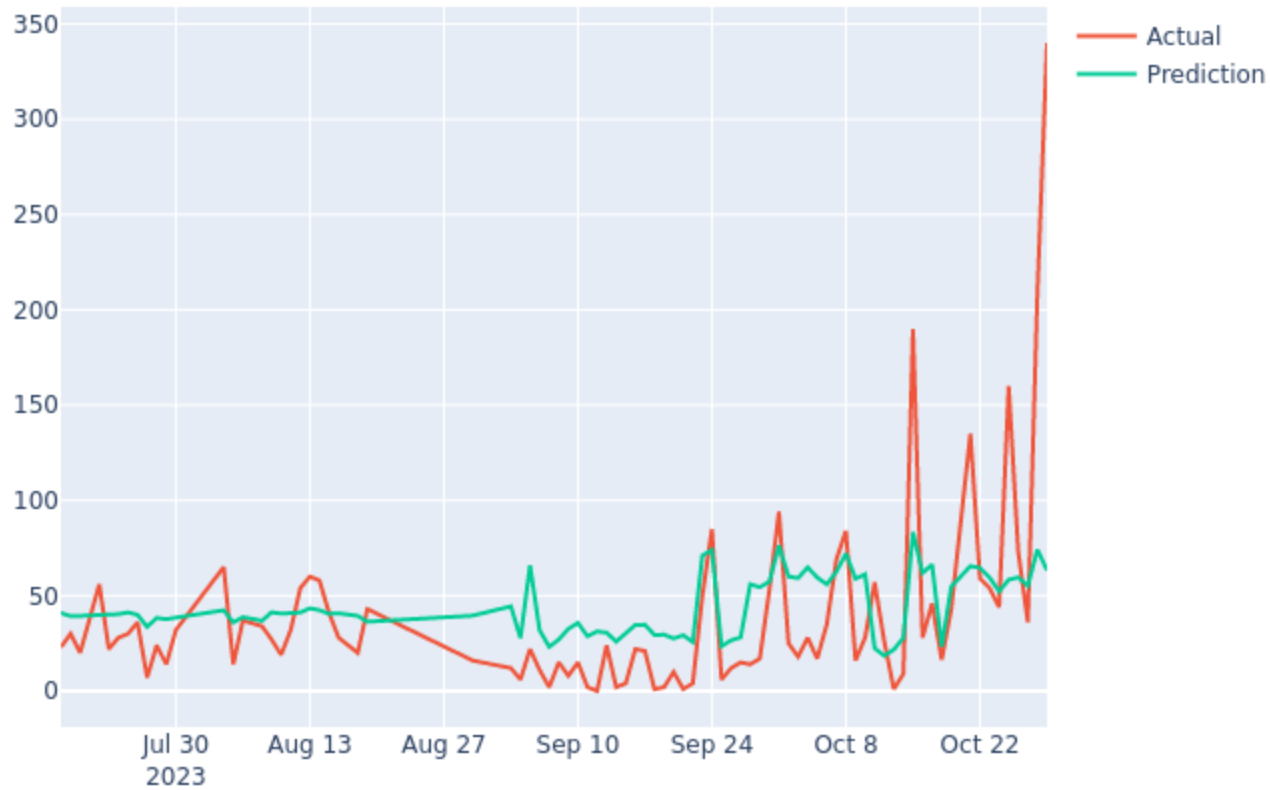
## ➔ XG Boost



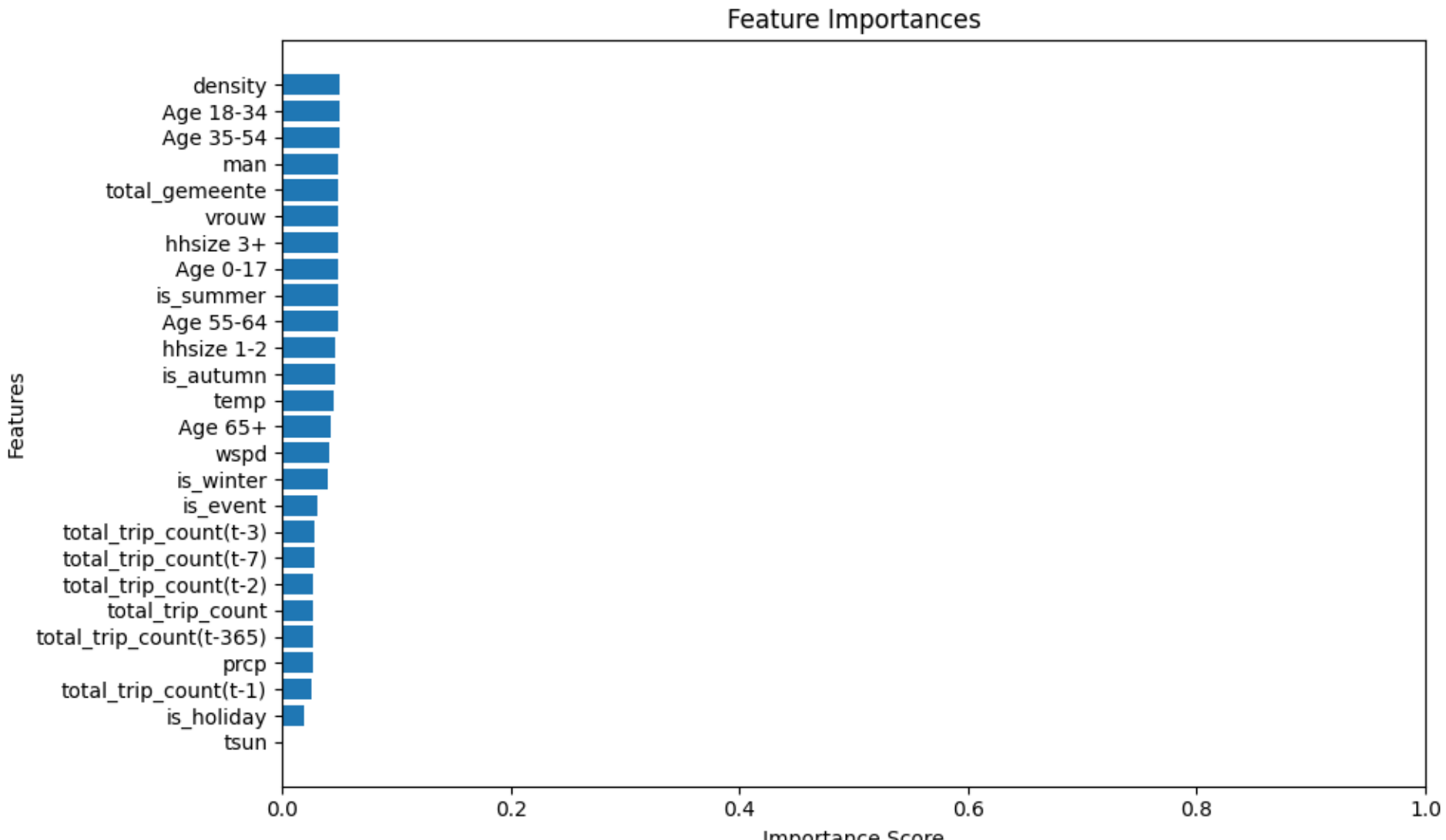
# ➔ Feature Importance



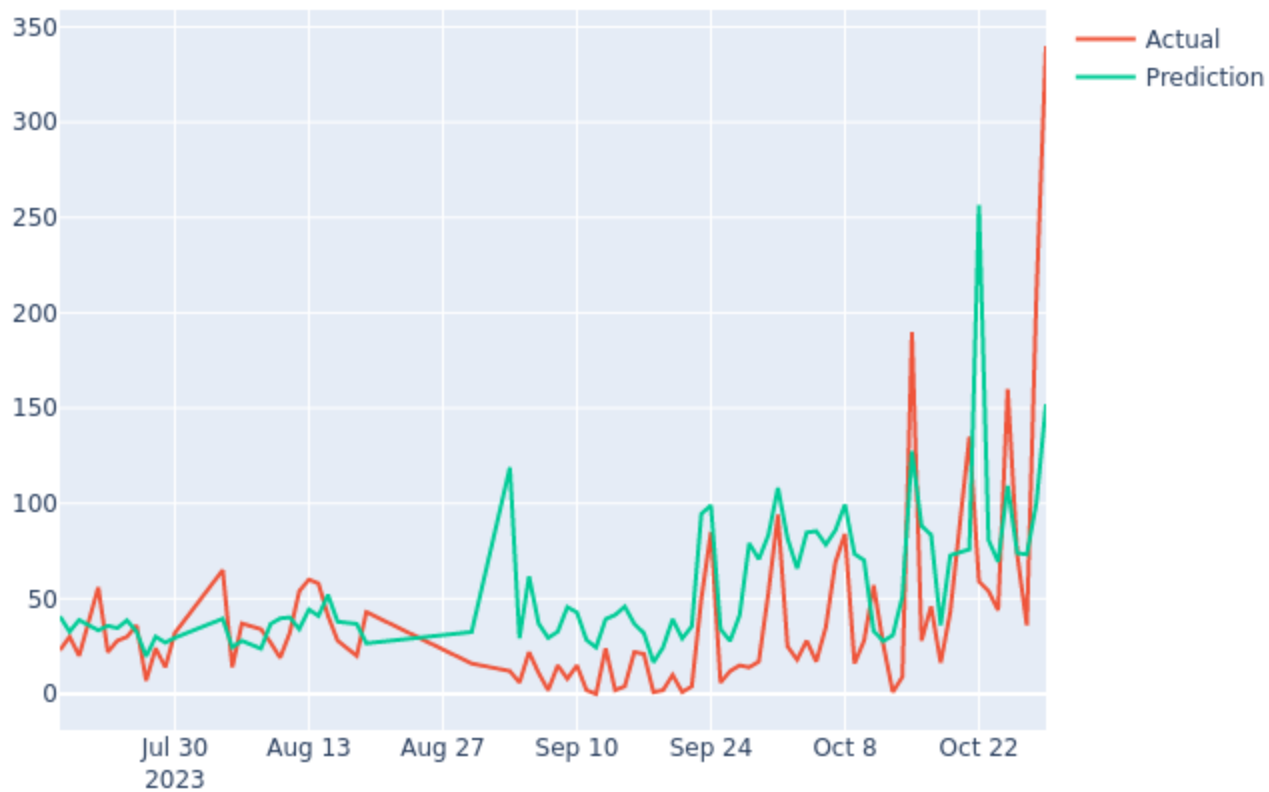
# ➔ SVM



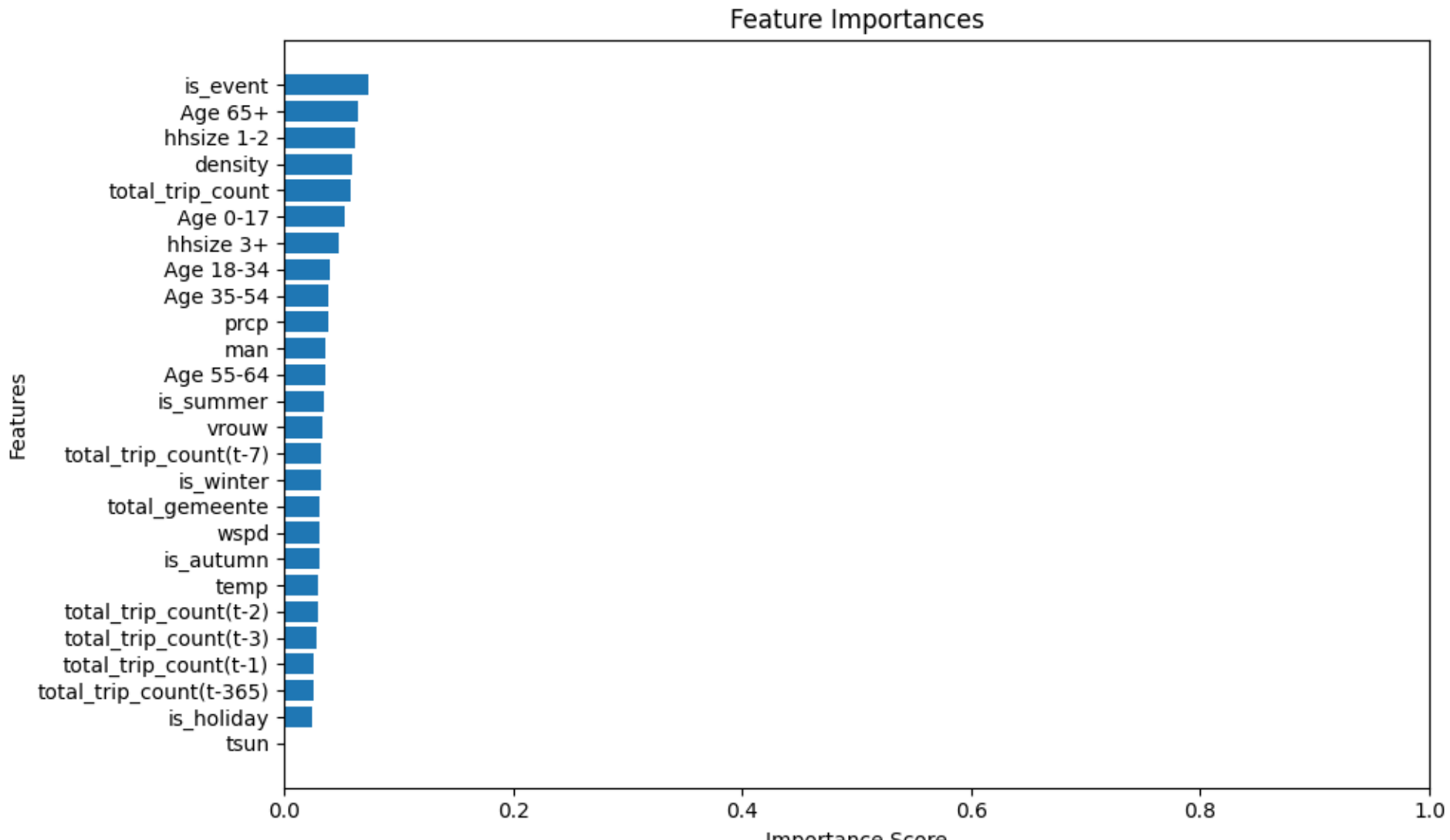
# ➔ Feature Importance



# ➔ Neural Network



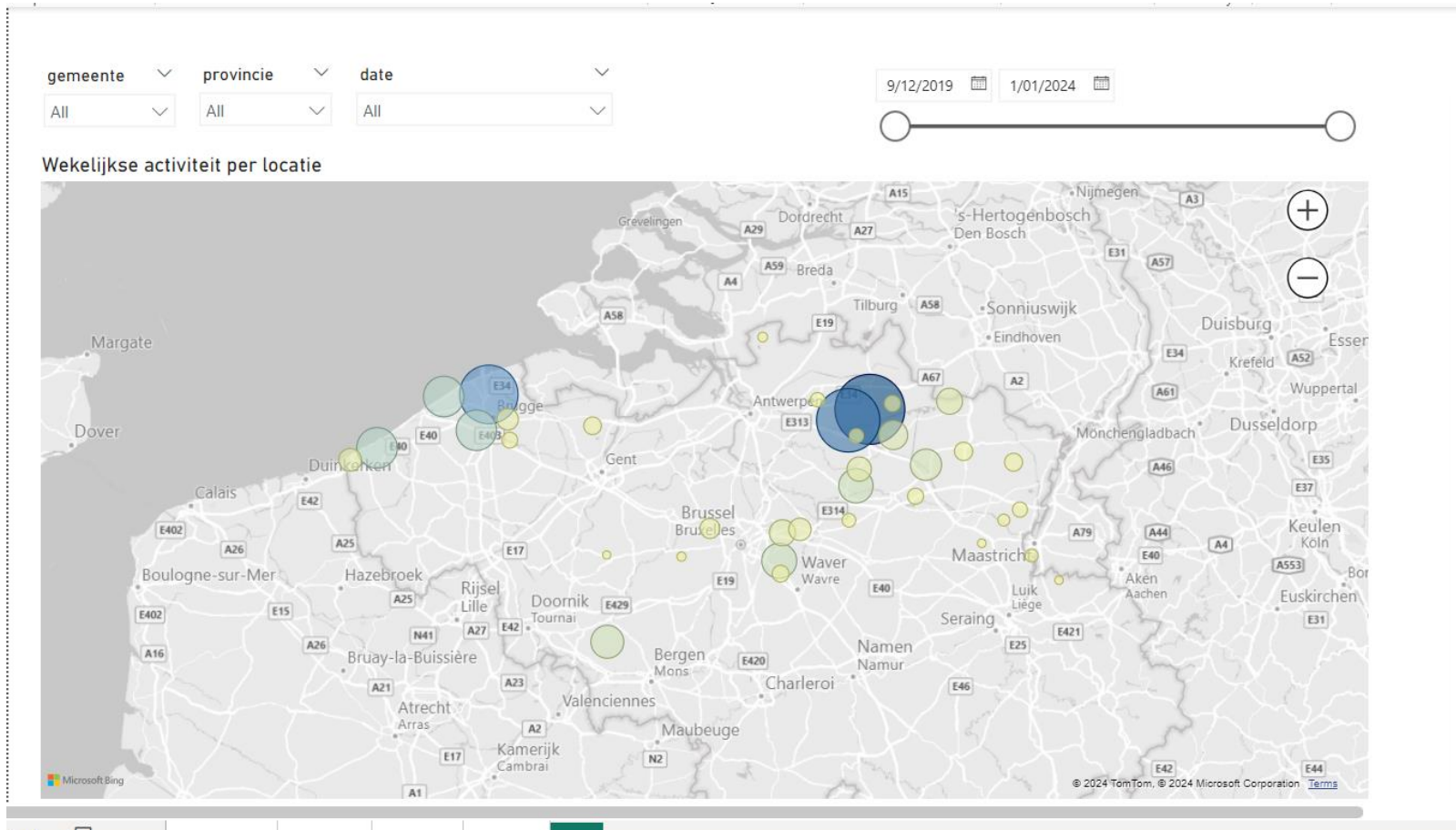
# ➔ Feature Importance



## ➔ Model Performance

Model	MAPE	MAE	RMSE
Random Forest	3472619372 641579	21.864	36.145
XGBOOST	3865953443 437424	20.465	26.520
SVM	2215952163 532702	21.006	33.869
Neural Network	2563902244 477275	25.264	35.975

# ➔ Power Bi





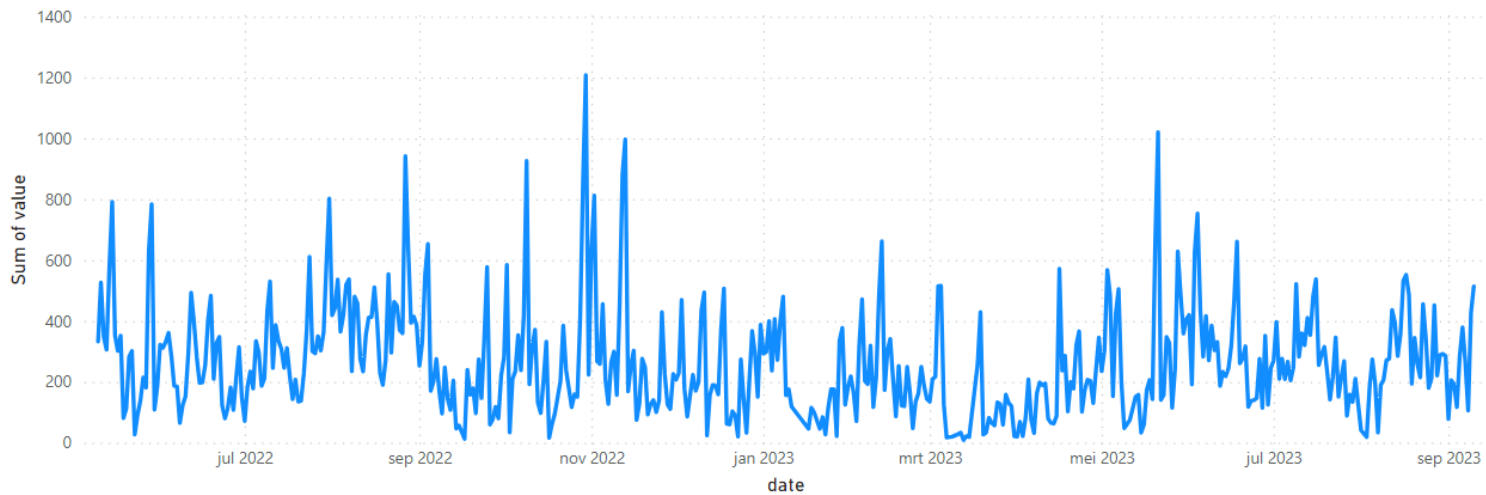
# ➔ Power Bi

## Voorspelde Wekelijkse Activiteit Per Locatie - Trend

location  provincie

10/05/2022  10/09/2023

Wekelijkse activiteit per locatie in de tijd



# Conclusie

- ② Strava activiteit en de aanwezigheid van evenementen zijn belangrijke variabelen om sport te voorspellen
- ② Voorspellingen kunnen sportgebruik extrapoleren naar nieuwe locaties zonder teller
- ② Dit kan een duidelijker beeld scheppen over het sportgebruik waar geen tellers staan

# Q&A