

Order delivery time

Based on
industry benchmarks¹

21-30
min

Customer satisfaction CSAT score

Based on
CSAT benchmarks²

75-80
%

Food delivery KPIs*

Average order value AOV

Bigger orders
Larger revenue

\$\$\$

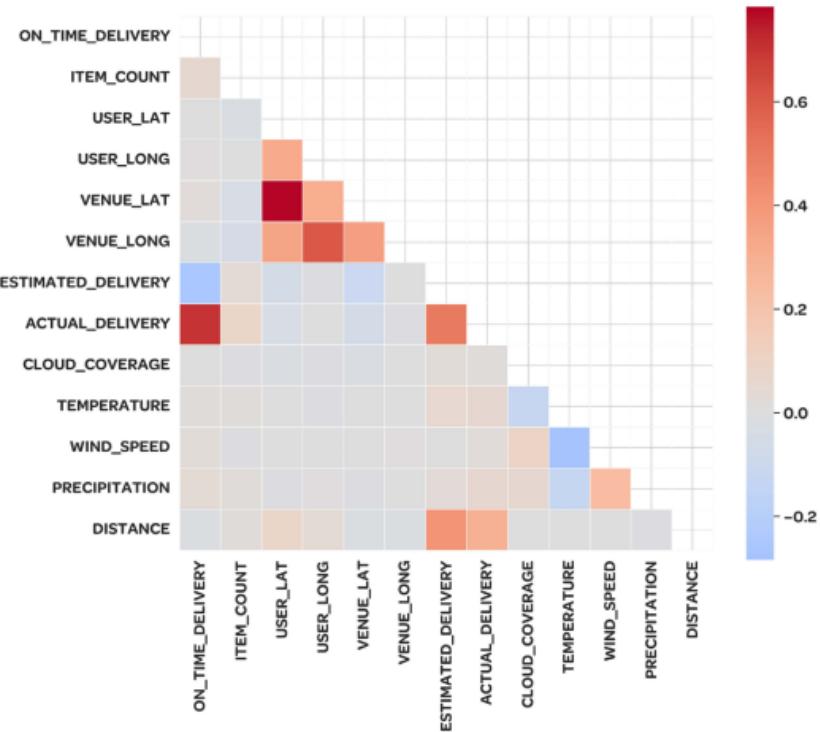
Courier availability

When and where
Couriers are needed^{3,4}



*samples based on available data for the assignment

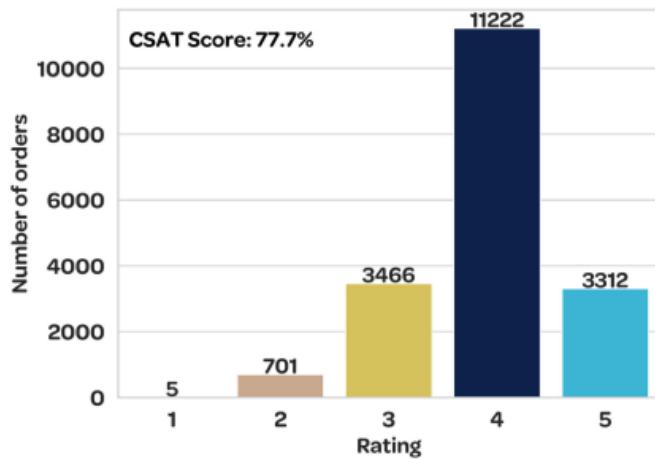
Correlations between features



No unusual strong correlations; only expected

- temperature – wind speed
- user – venue locations
- delivery times: estimated – actual – on-time

Customer satisfaction

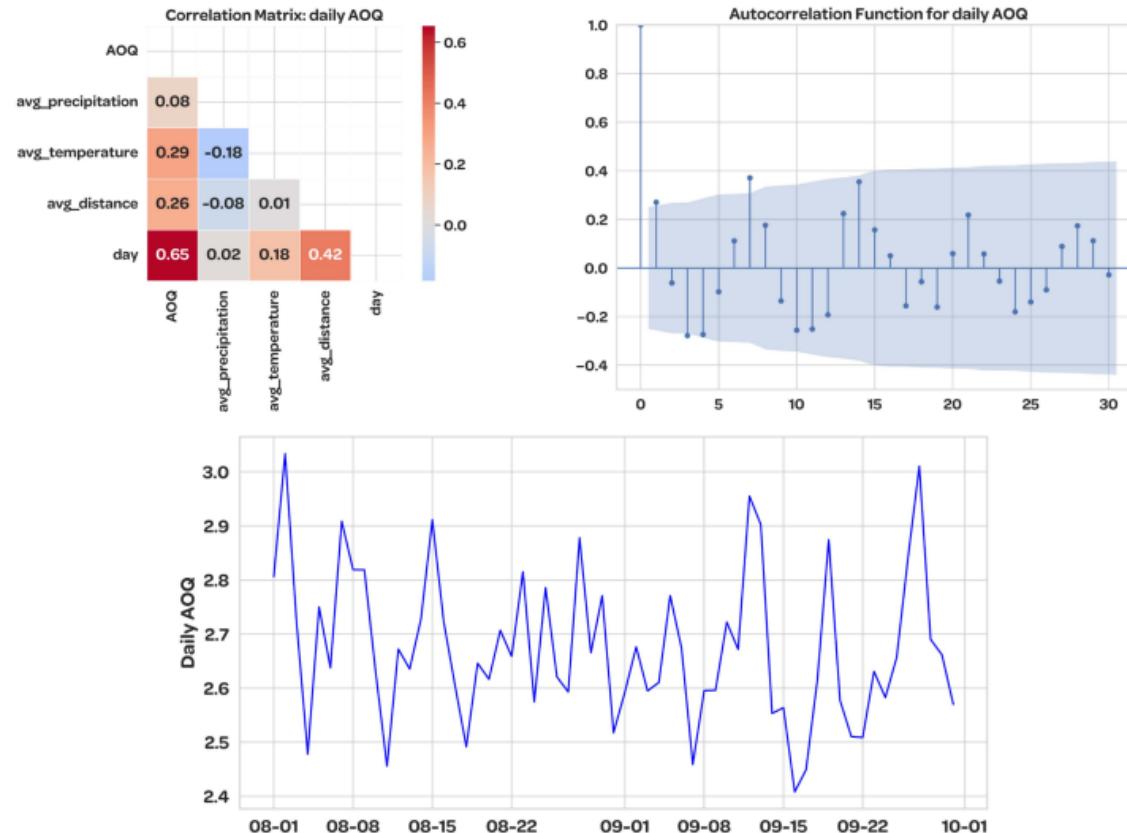


CSAT score computed by

- aggregating the time saved on orders via $\text{ACTUAL_DELIVERY_MINUTES} - \text{ESTIMATED_DELIVERY_MINUTES}$
- creating a custom **rating system**
- **CSAT** computed with
number of 4,5 ratings / number of orders

Rating	1	2	3	4	5
On-time delivery	>30 min	15-30 min	5-15 min	-10 - 5 min	<-10 min

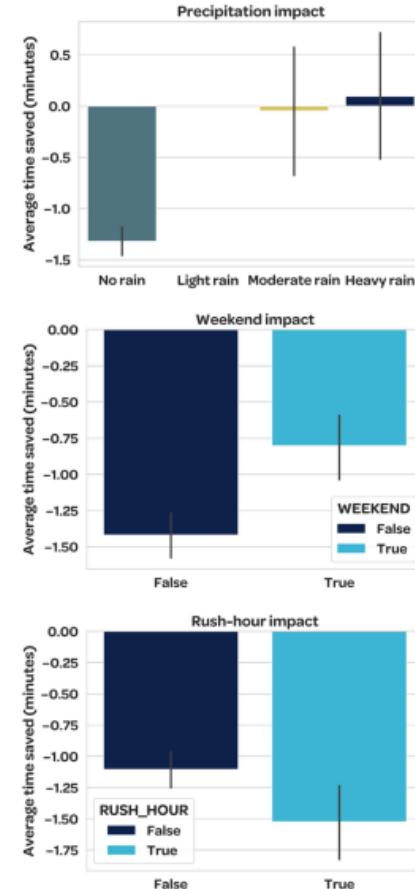
Average order quantity (AOQ)*



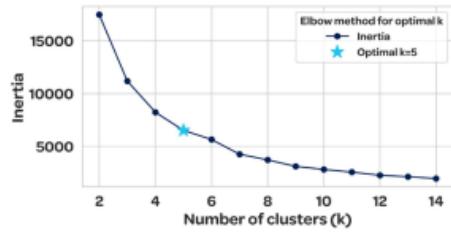
Low impact of weather, **limited** bad weather days in Aug-Sep

*Replaced AOV with **AOQ** (available data) – we correlate value with quantity

On-time delivery impact

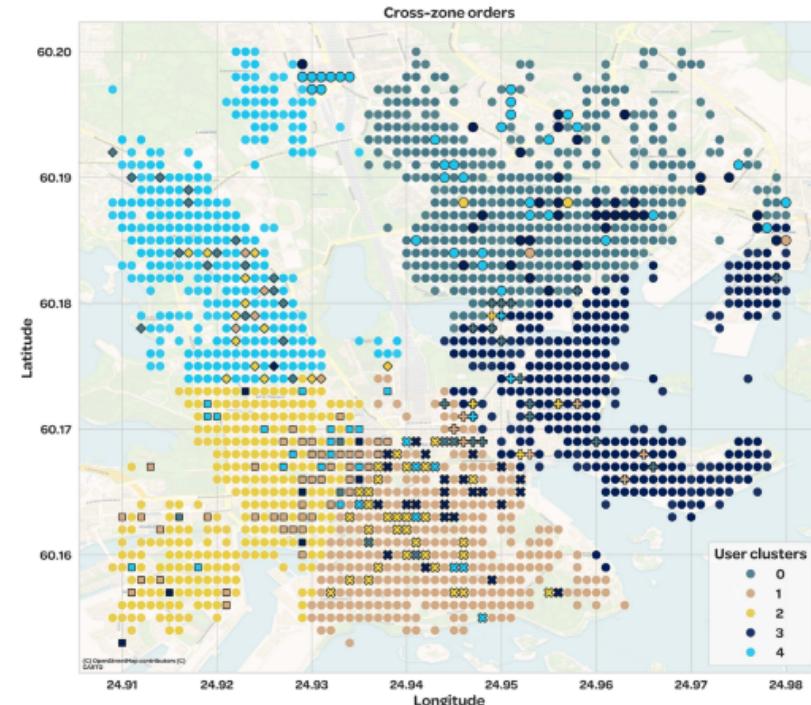
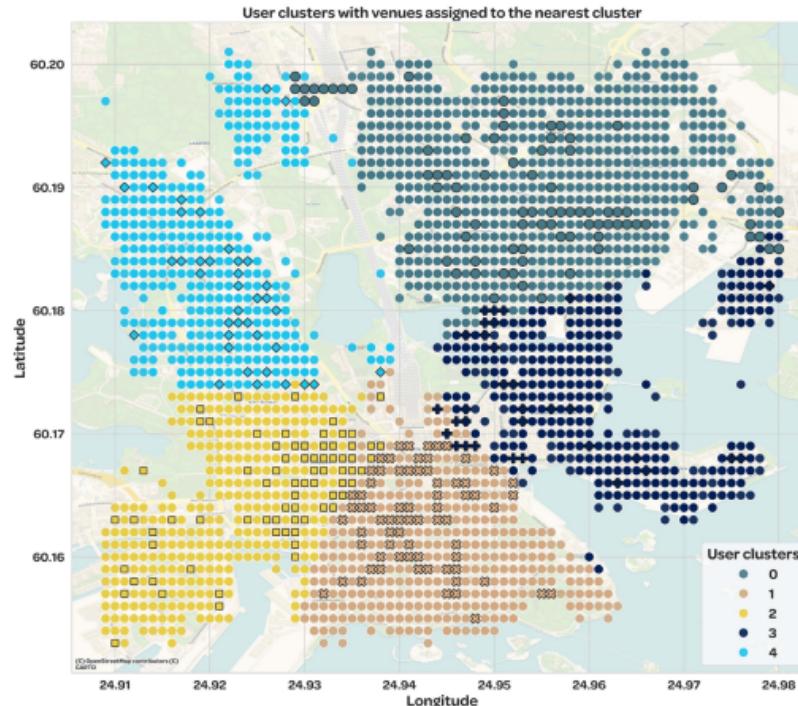


No significant impact on saved time^{5,6}

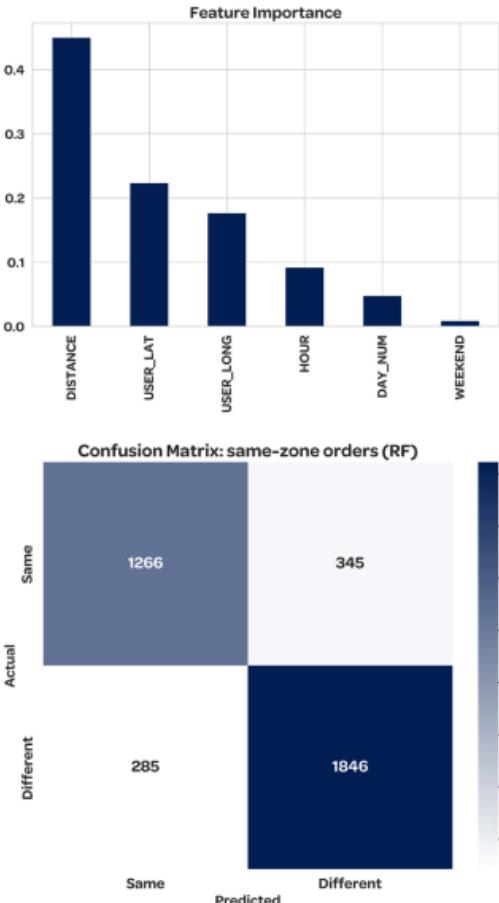


Clustering based on **K-means**

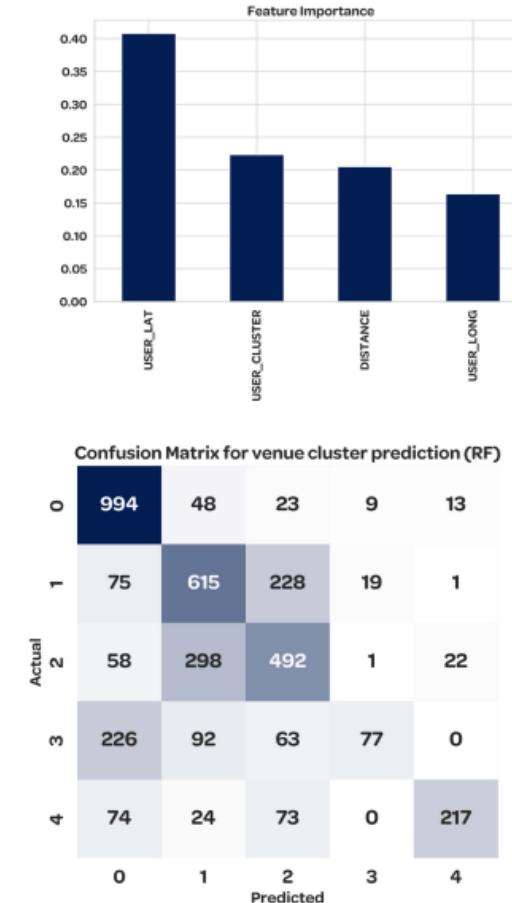
- assign users on clusters based on **location** (user latitude and longitude)
- elbow method** determines **optimal cluster number**
- differentiate between **same-zone orders** and **cross-boundary orders**
- helps towards **efficient courier distribution**



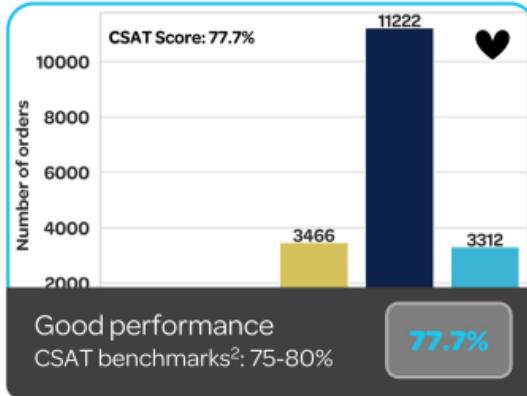
Same-zone orders prediction



Venue cluster prediction



Data exploration and KPIs



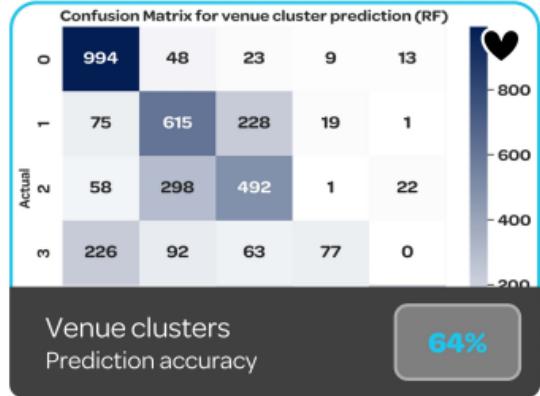
1. No major correlations
2. Impact on AOQ
 - low weather impact
 - bigger weekends orders
3. Low impact on-time delivery
 - weather
 - weekday/weekend
 - rush hour orders

Clustering by location



1. No major clustering correlations
2. Good zone separation
 - 5 clusters
 - 42% orders cross-zone
3. Zones used for classification
 - same-zone orders
 - venue clusters

Classification and prediction

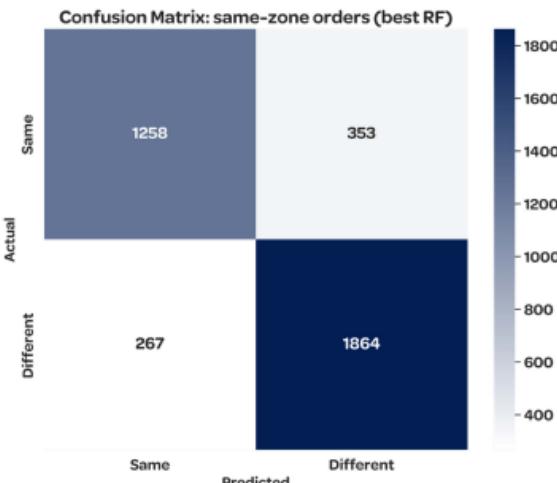


Hyper-parameter search

- optimal parameters for same-zone orders prediction from

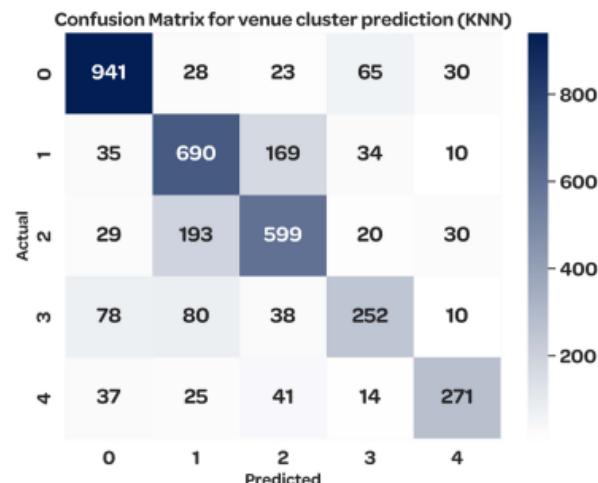
RandomForestClassifier

- max_depth=19
- n_estimators=330



Test with other classifiers

- cluster prediction with **KNN**
 - neighbors=5



Time-series analysis

- limited dataset for daily TSA
- try **SARIMA** on
 - AOQ
 - daily orders
 - weekly orders
- forecast number of orders for
 - particular day
 - user
 - clusters

*preliminary experimental ideas and results