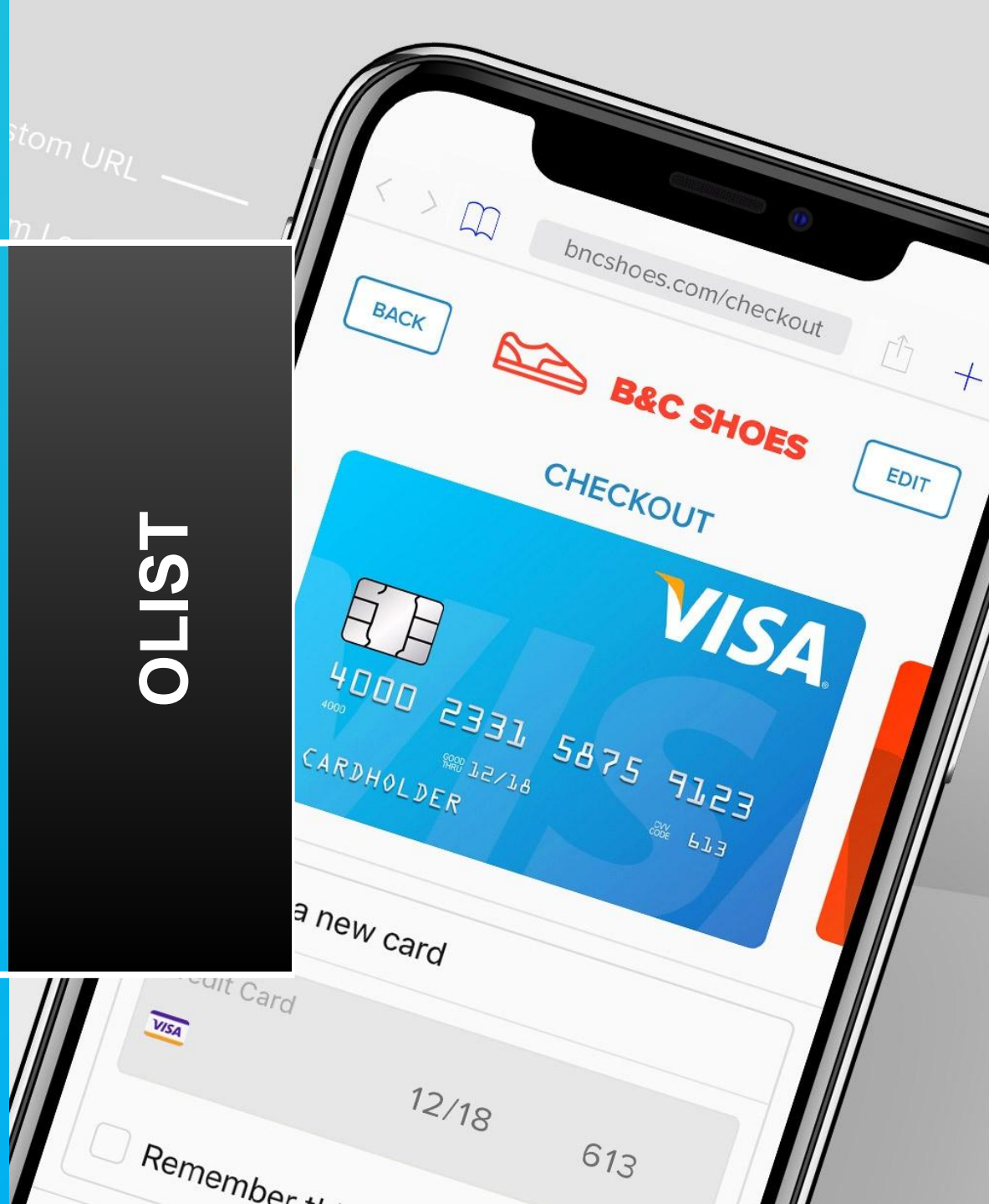
By Anna Barentz

REVENUE PREDICTION

Brazilian E-Commerce

OLIST



Goals

- Revenue forecasting
- Customer satisfaction prediction

Data

Brazilian ecommerce store

- 9 datasets
- 100,000 orders
- 66 product categories

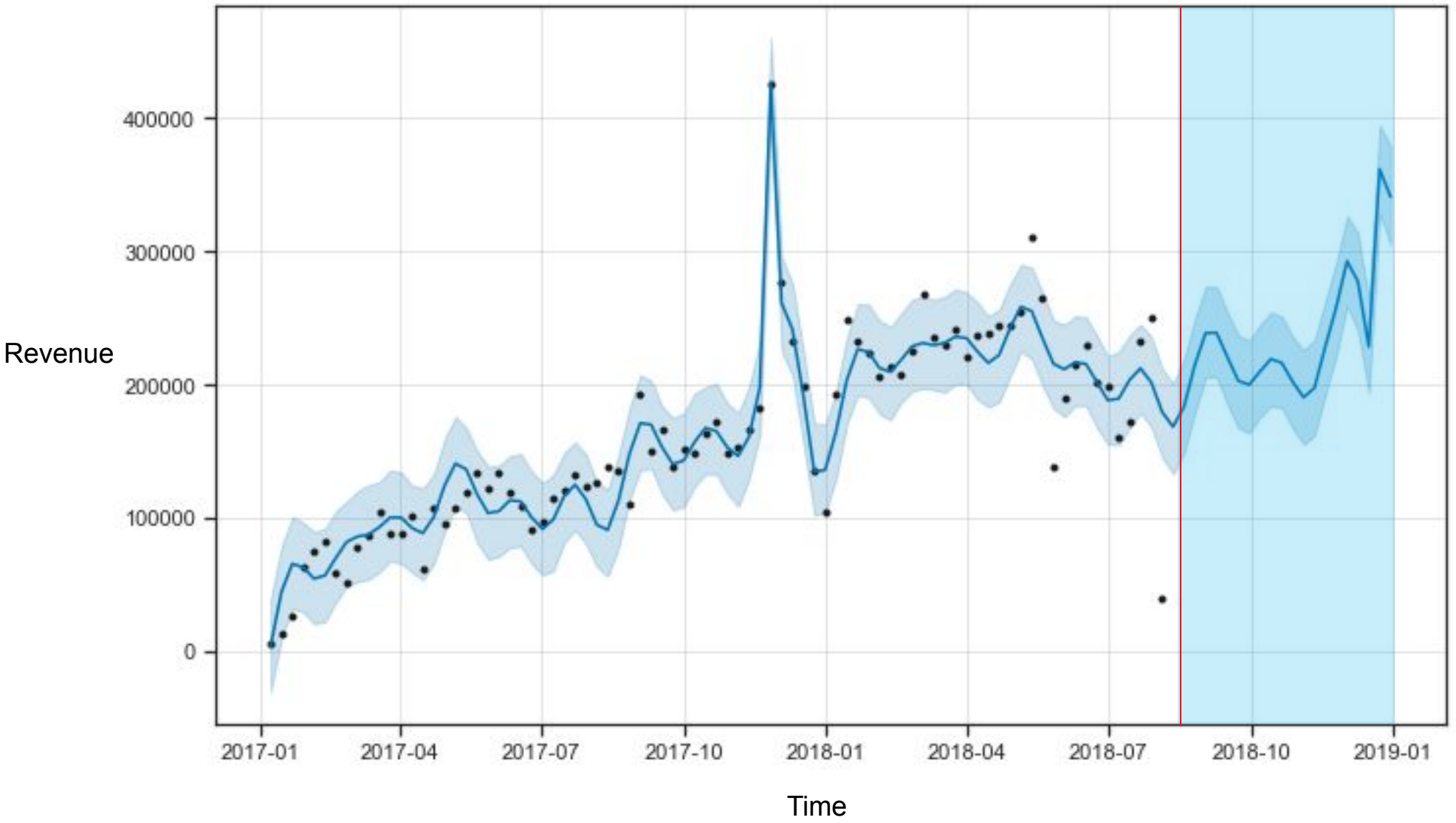
January 2017 to August 2018
weekly data

Exploratory Data Analysis

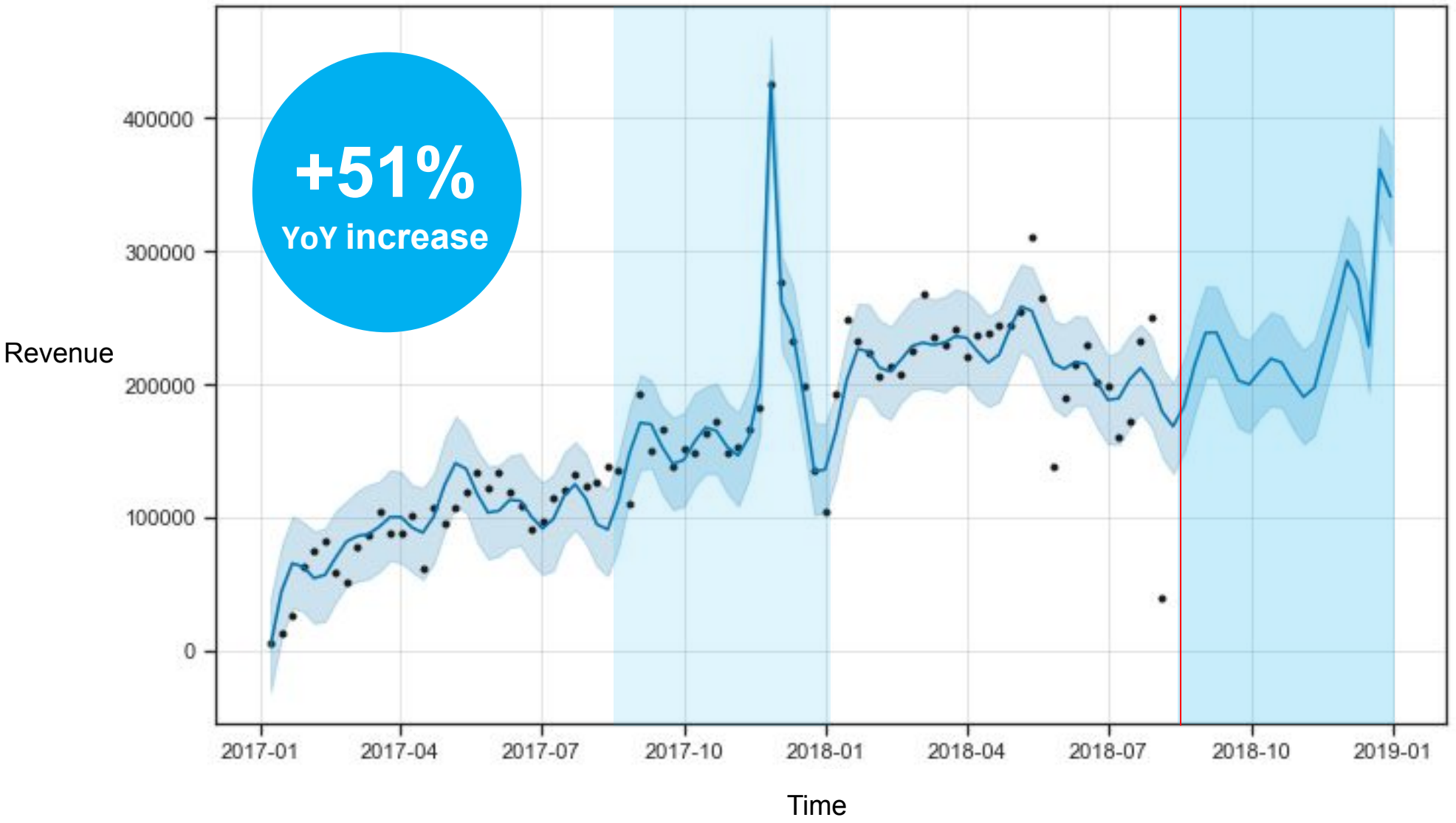
Logistic regression and Random Forest

Facebook Prophet, ARIMA

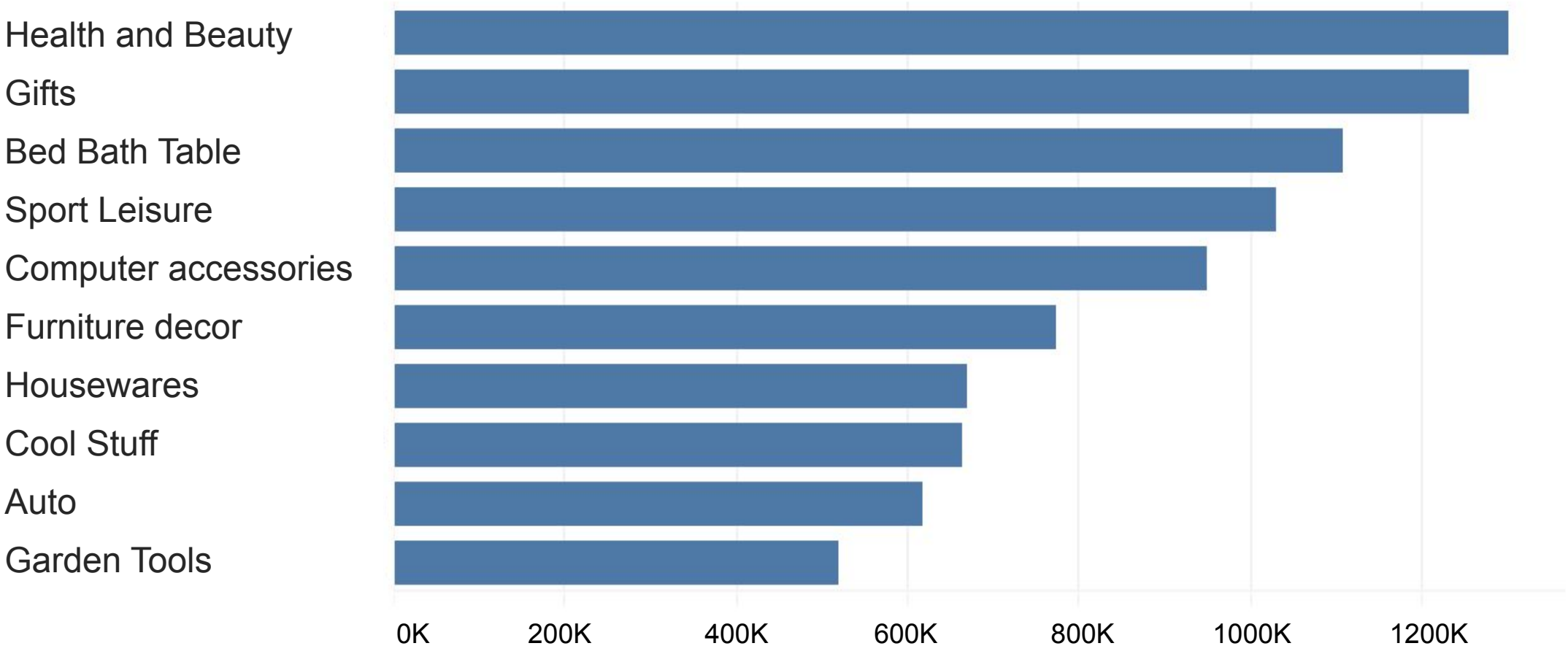
Revenue prediction



Revenue prediction: YoY



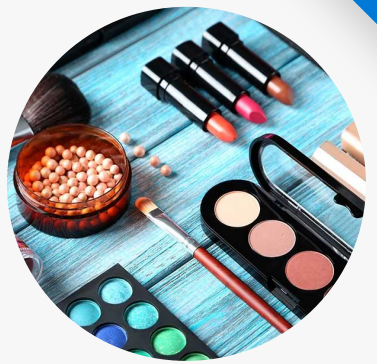
Revenue by Category



Categories

Best Review Predictors

Positive



Health/Beauty

Positive

Sport Leisure

Negative

Furniture/Decor

Negative

Computer accessor.

PRODUCT PHOTO QUALITY

the most important positive feature

LATE DELIVERY

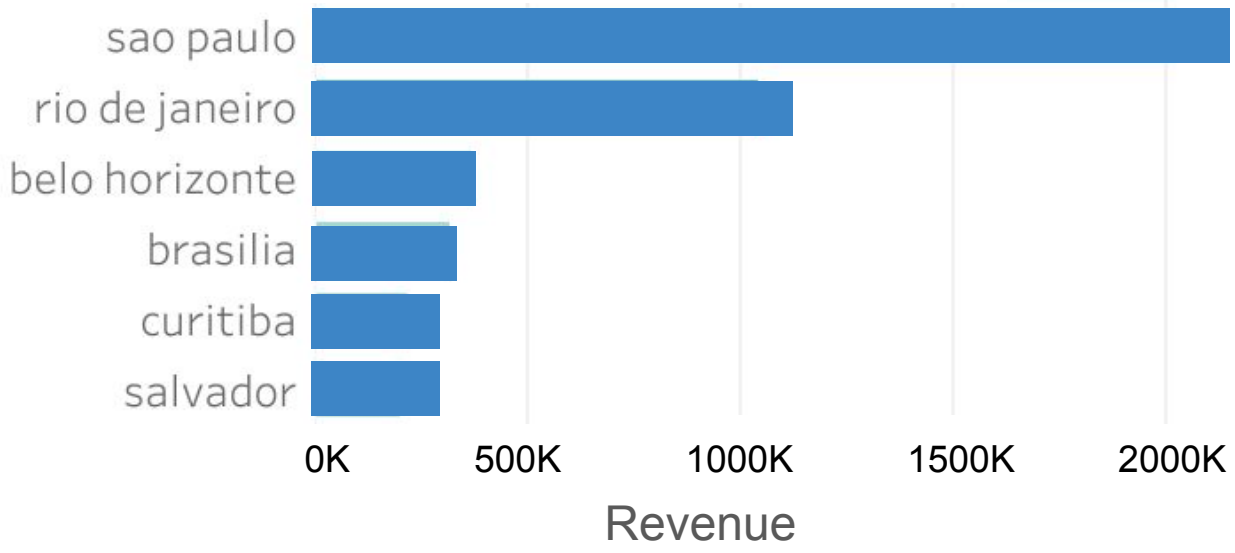
the most important negative feature

Cities

Best Review Predictors:

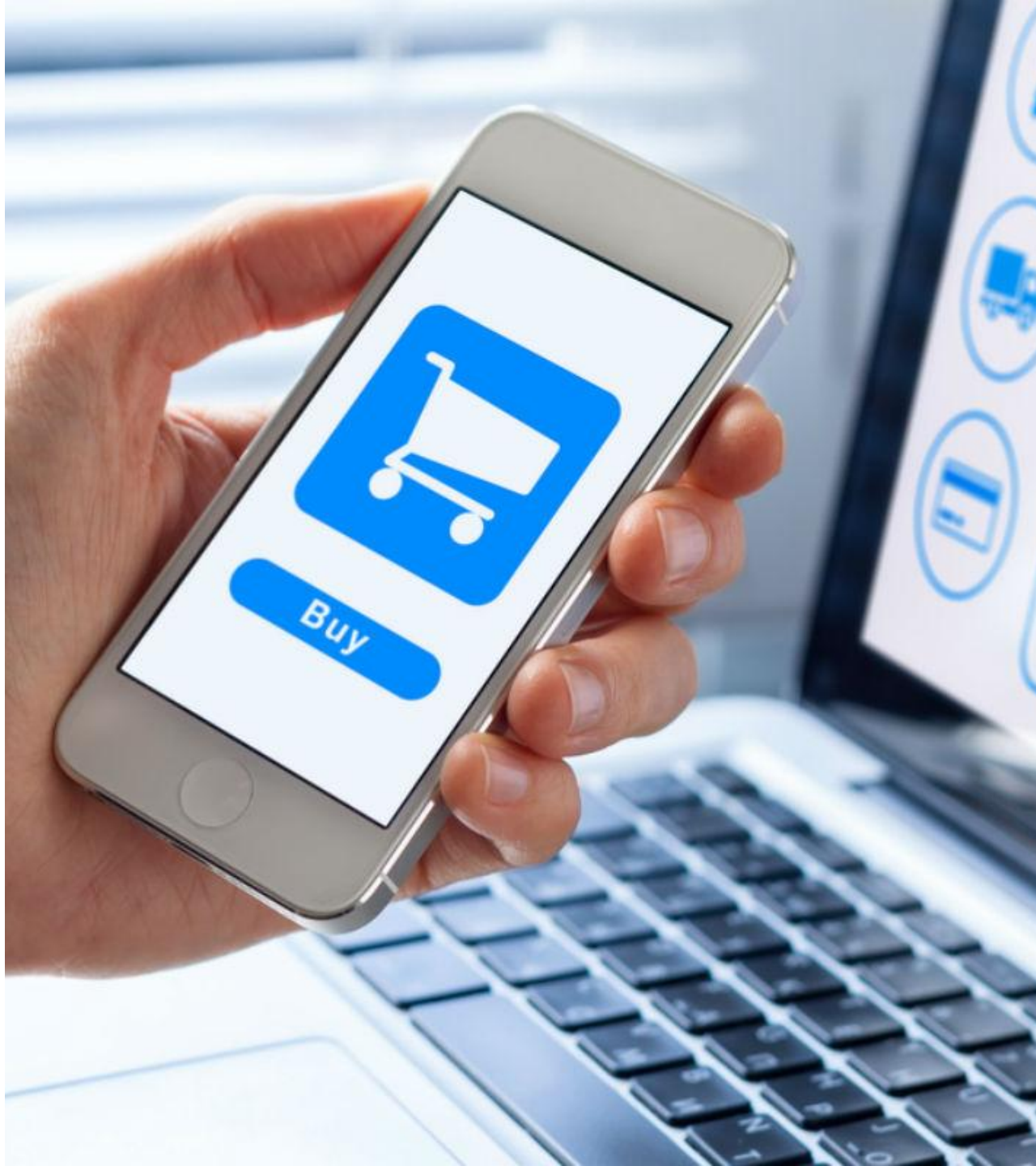
- Sao Paulo
- Curitiba
- Rio De Janeiro
- Salvador
- Brasilia

TOP 6 Cities:



RECOMMENDATIONS

- **Sao Paulo and Curitiba**
  - Add more beauty health and sport options
- **Rio De Janeiro and Salvador**
  - Eliminate late deliveries (11% and 16,7%)
- **Brasilia** - ½ population of Rio but lagging in revenue:
  - Capture products that are in demand



Anna Barentz

anna@wizarly.com

annabarentz@com

Anna-barentz

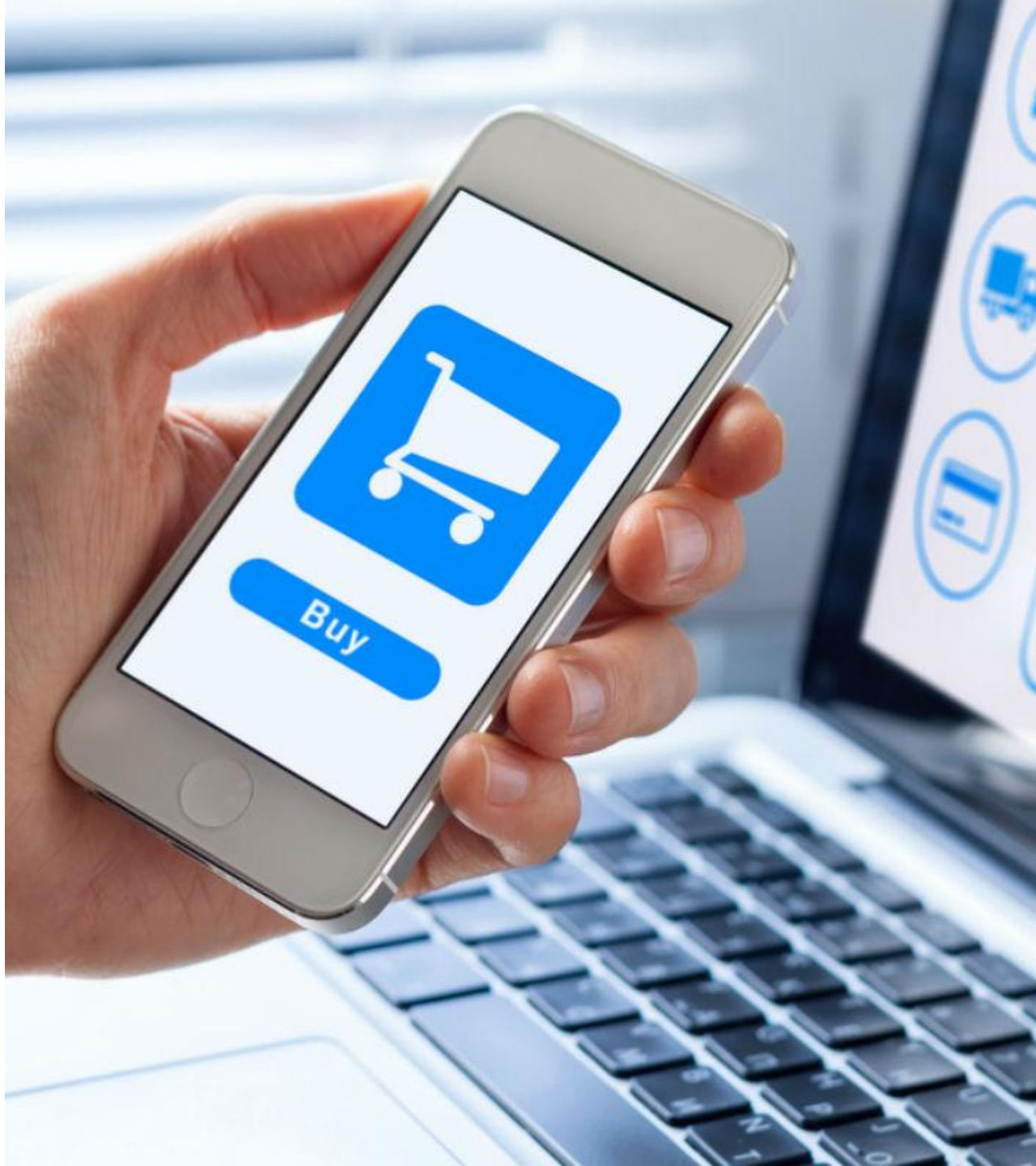
AnnaB24

THANK YOU!

Paulo and add more beauty
ons

RECOMMENDATIONS

e Janeiro
ate Late Deliveries
ore drivers
a - ½ population of Rio. But not
enue
nd capture



Additional features

Price

Good price - good quality

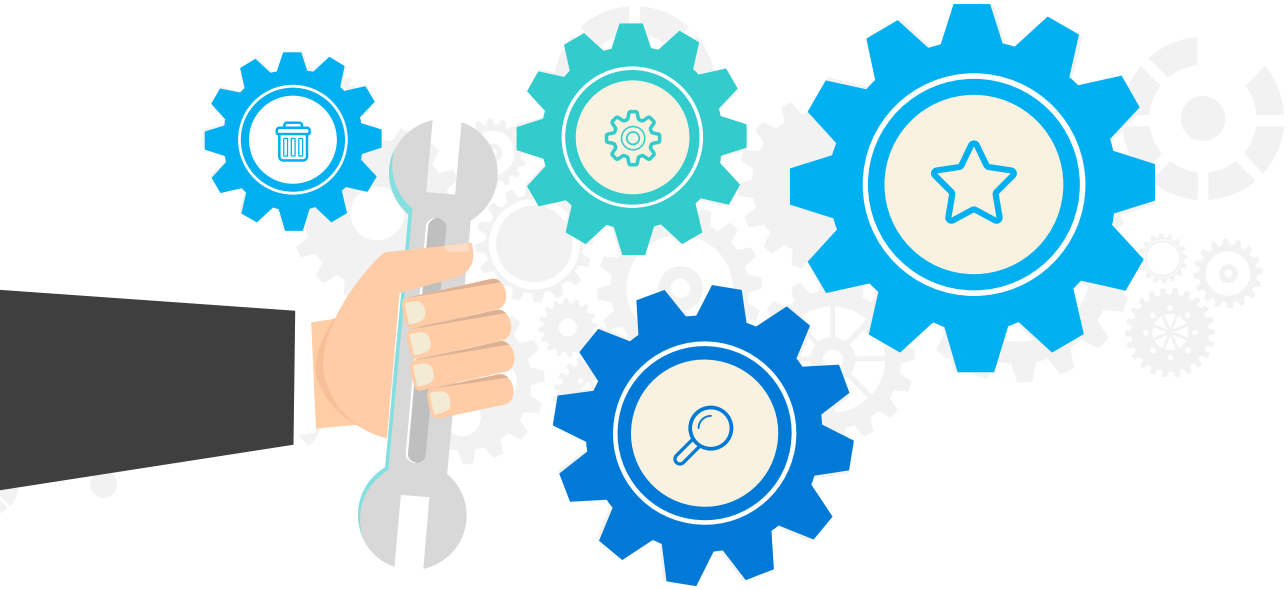
Payment installments

Nobody likes to pay for long time

Freight value

Nobody likes to pay for shipping

Methodology



Exploratory Data Analysis

- Analyzing data sets to summarize their main characteristics
- Visualization

Time Series Models

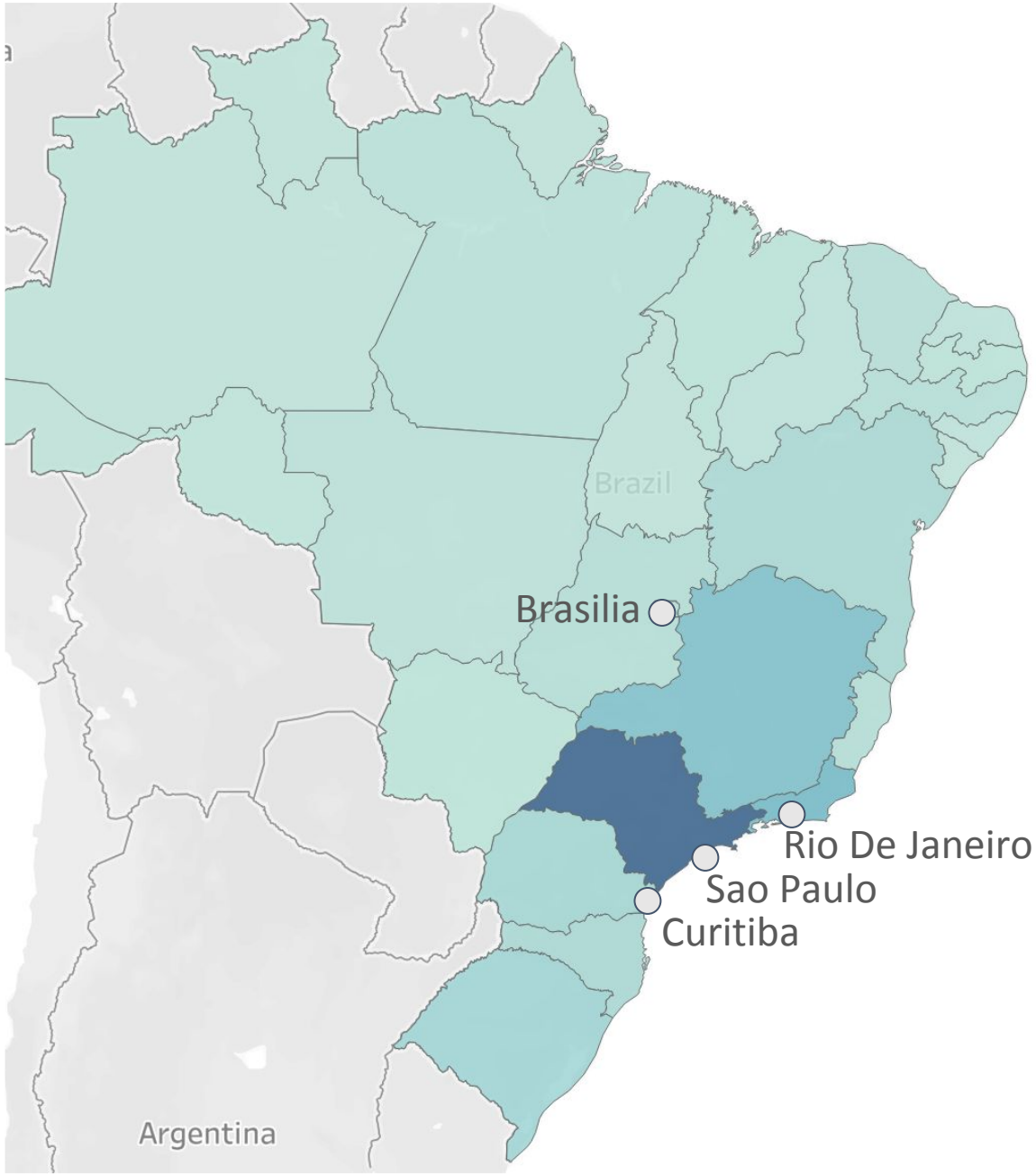
Facebook Prophet, ARIMA

- Revenue prediction
- Visualization

Classification Models

Logistic regression, Random Forest

- Predict customer satisfaction (based on reviews score)
- Find the most important features



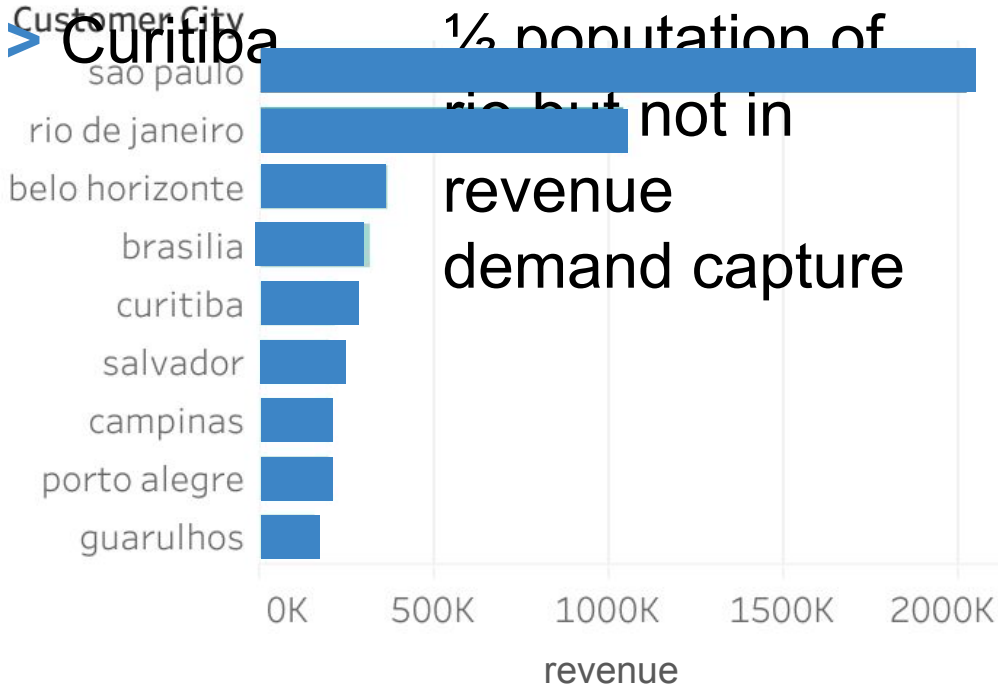
Top Cities

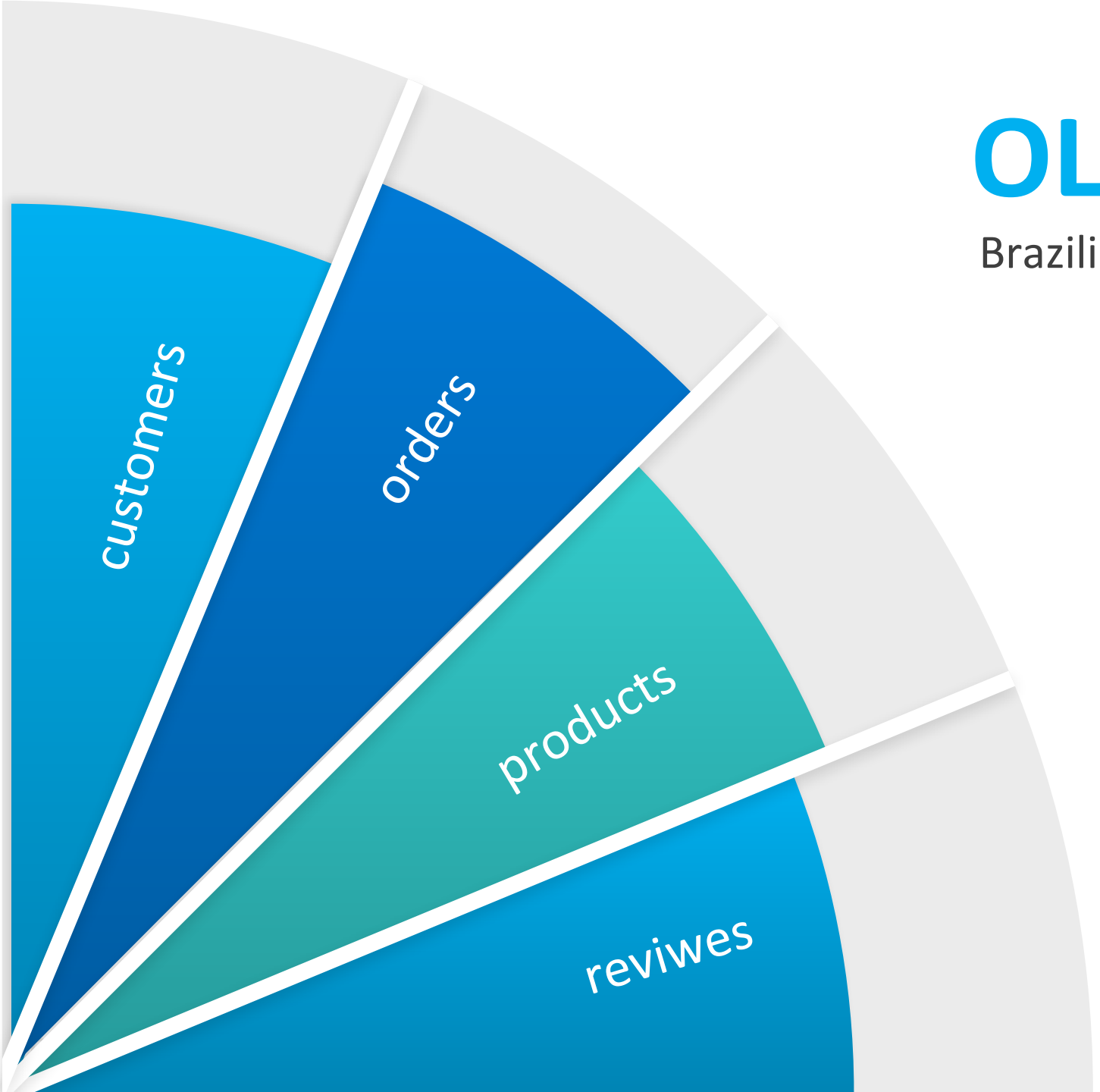
> Sao Paulo add more beauty options

> Curitiba

< Rio De Janeiro late delivery more drivers

< Brasilia ⅓ poputation of rio but not in revenue demand capture





OLIST Data

Brazilian ecommerce store

- 9 datasets
- 100,000 orders
- 66 product categories
- January 2017 to August 2018 weekly data

Exploratory Data Analysis

- Analyzing data sets to summarize their main characteristics

Classification Models

Logistic regression, Random Forest

- Predict customer satisfaction (based on reviews score)
- Find the most important features

Time Series Models

Facebook Prophet, ARIMA

Data

- Olist Brazilian ecommerce
- 9 datasets
- 100,000 orders
- 70 product categories
- January 2017 to August 2018 weekly data

2017
2018

PRODUCT PHOTO QUALITY

(x0.4) the most important positive feature