

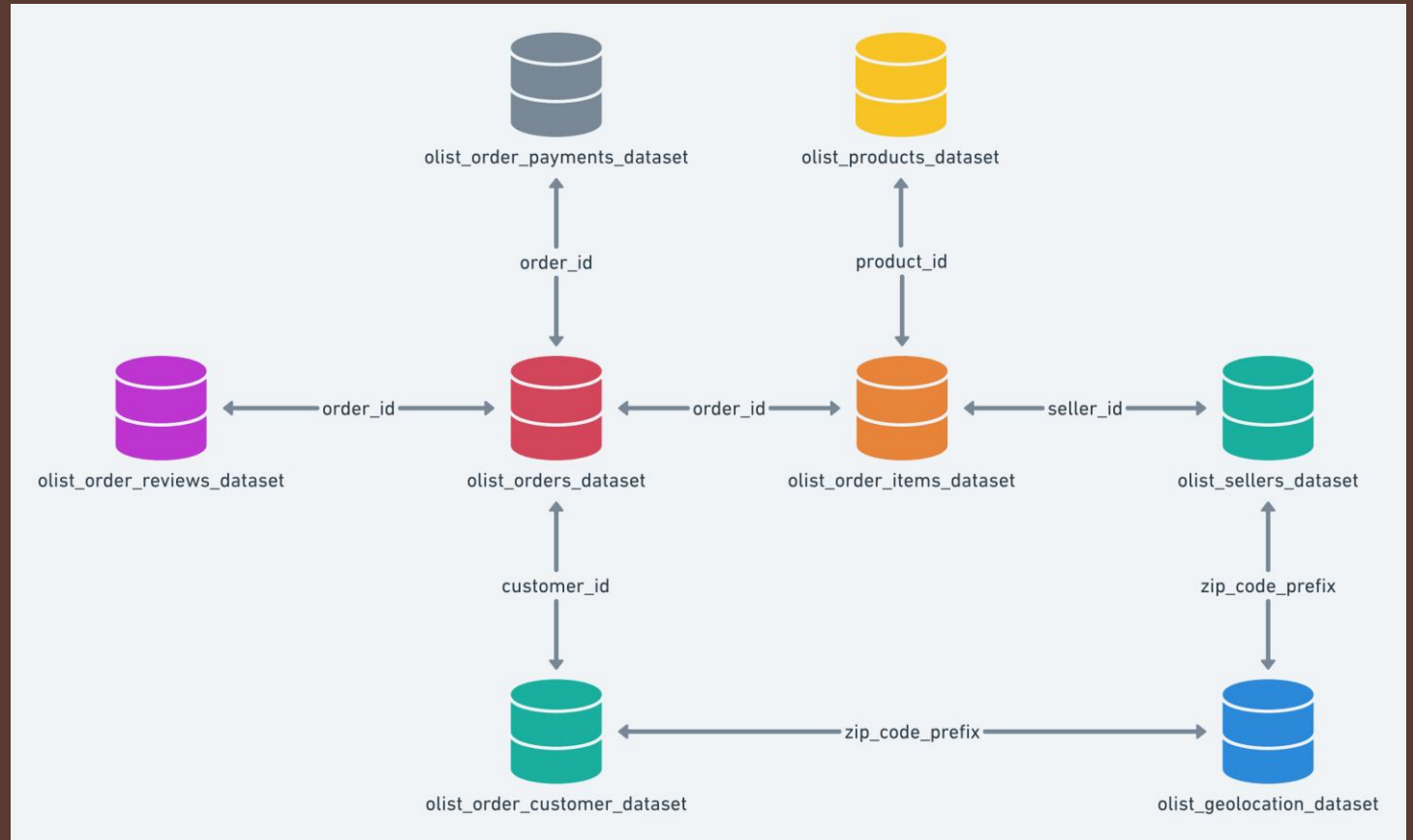
# Olist Dataset BigQuery and Data Visualization

Chiao-I Lin



# Introduction

- Brazilian E-Commerce Public Dataset by Olist Welcome!  
This is a Brazilian ecommerce public dataset of orders made at Olist Store.



- Data source: <https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce>

# Outline

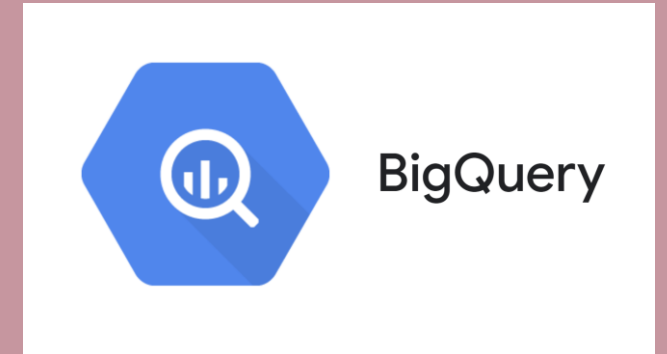
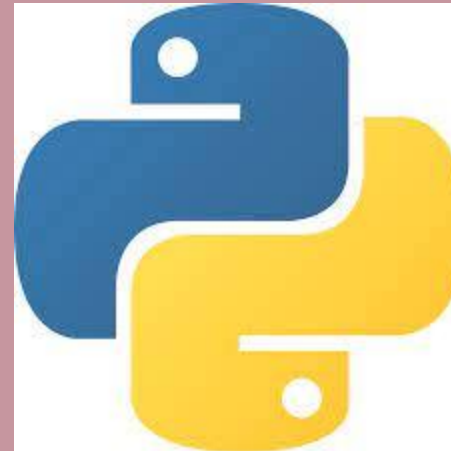
- Tools
- List of Visualization
- Visualization

**olist**



# Tools

- Jupyter
- Python
- Pandas
- Numpy
- Matplotlib.pyplot
- Seaborn
- GoogleBigCuery



*NumPy*



seaborn

# About this project

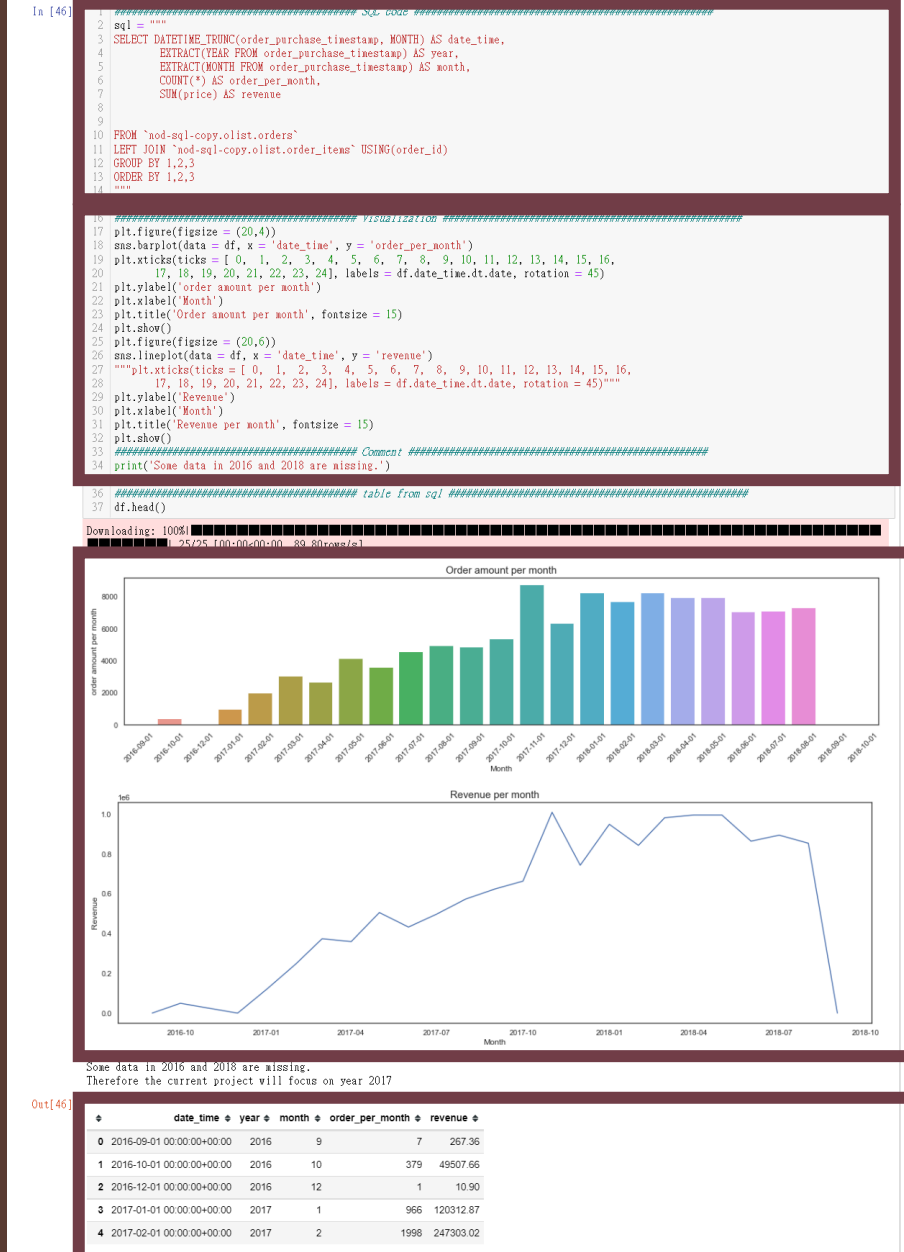
- Using google BigQuery to organize data
- And applying Seaborn and Matplotlib to visualize the data
- The structure of code and output in Jupyter notnook look like the chart in right

SQL code

Code for data visualization

data visualization

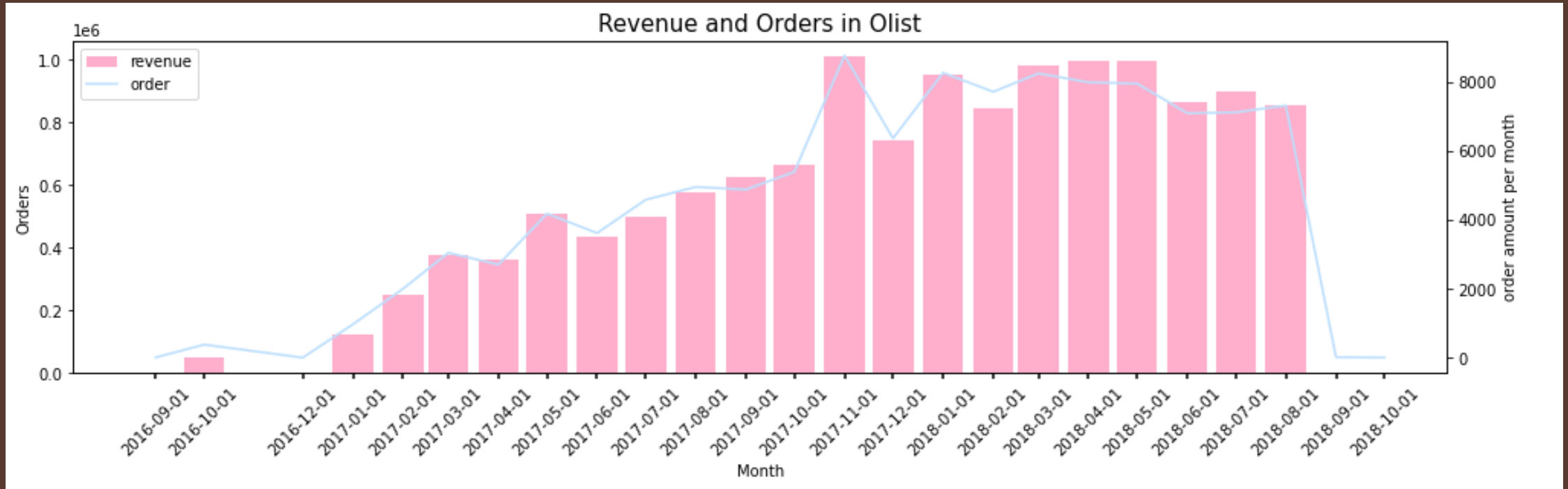
Table obtained by SQL



# List of visualization

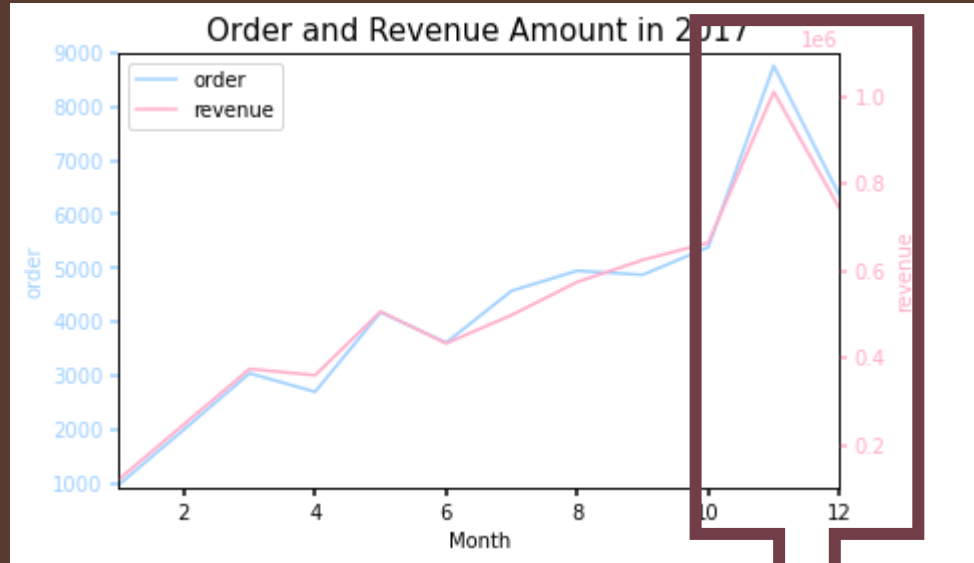
- Data: data per year
- Percentage change of order number/revenue per month in 2017
- Monthly Revenue in 2017: KDE plot
- Cancellation per month in 2017
- 2017 best monthly seller
- 2017 best selling state
- Monthly order per state
- Majority payment method in 2017
- Most popular product category

# Data: data per year



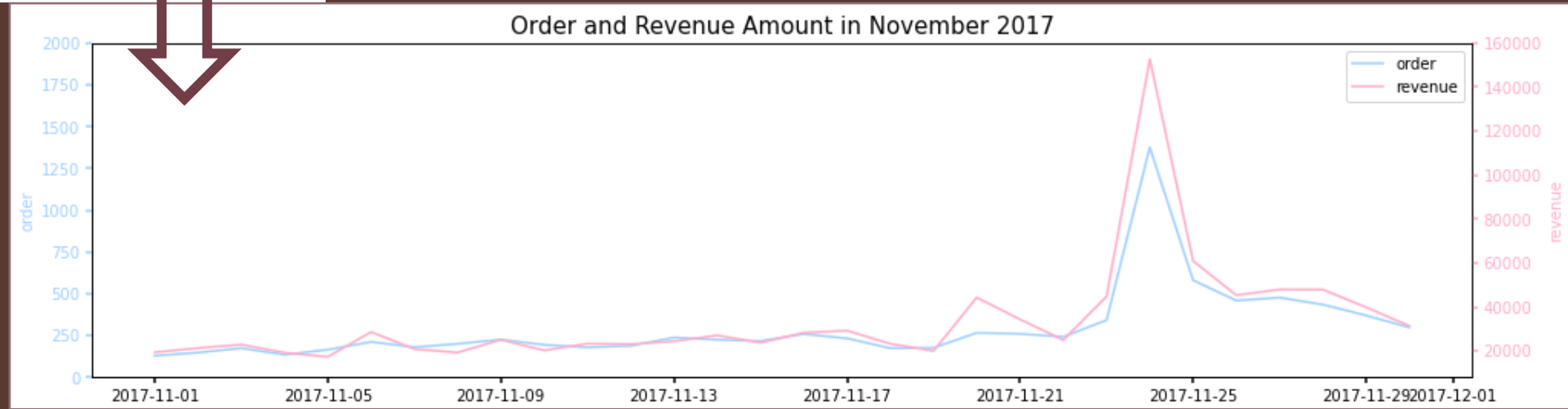
Some data in 2016 and 2018 are missing.  
Therefore, the current project will focus on year 2017

# Percentage change of order number/revenue per month in 2017



Order and revenue in 2017  
The peak shows in November

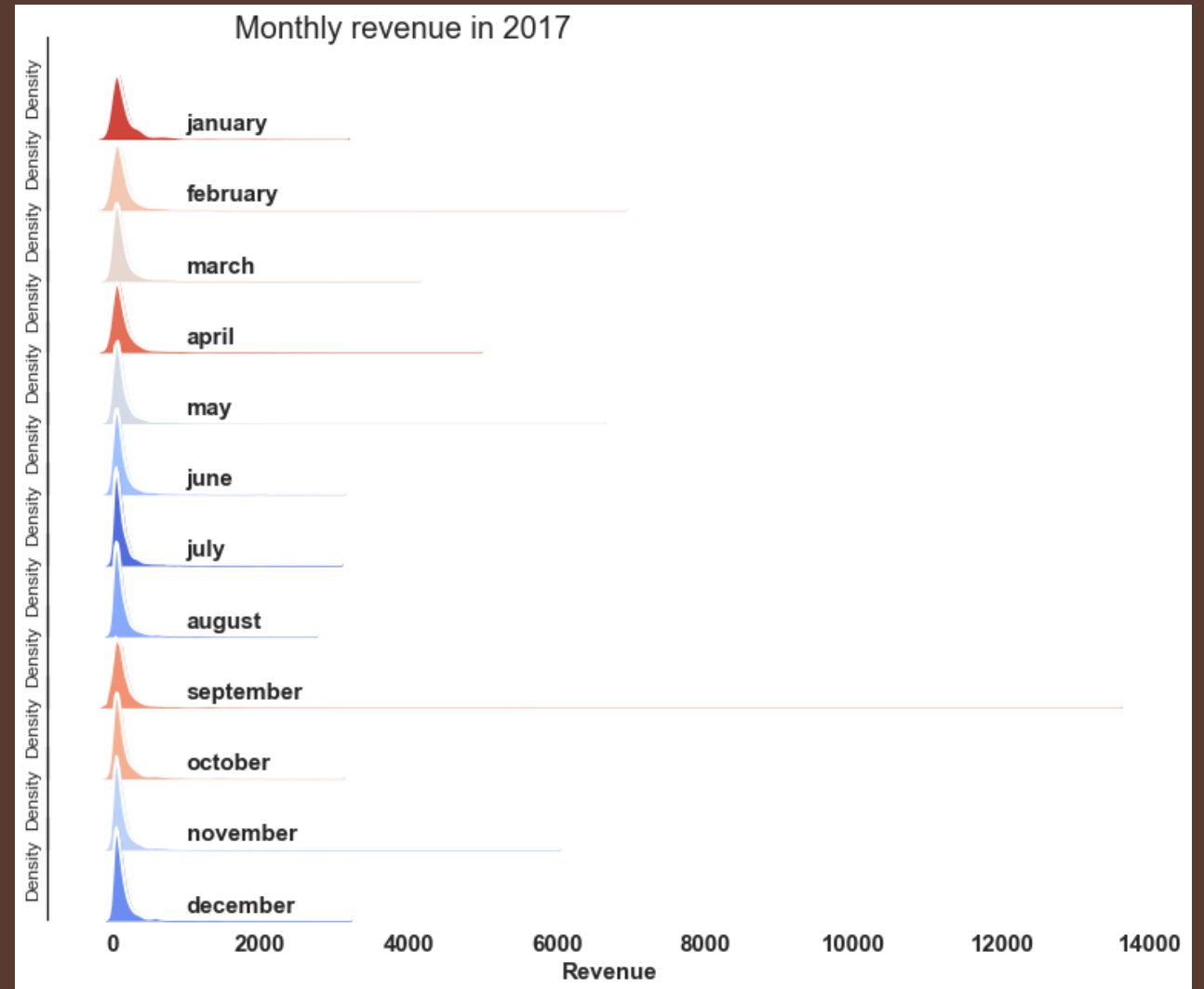
The peak was caused by Black Friday (November 24<sup>th</sup>)



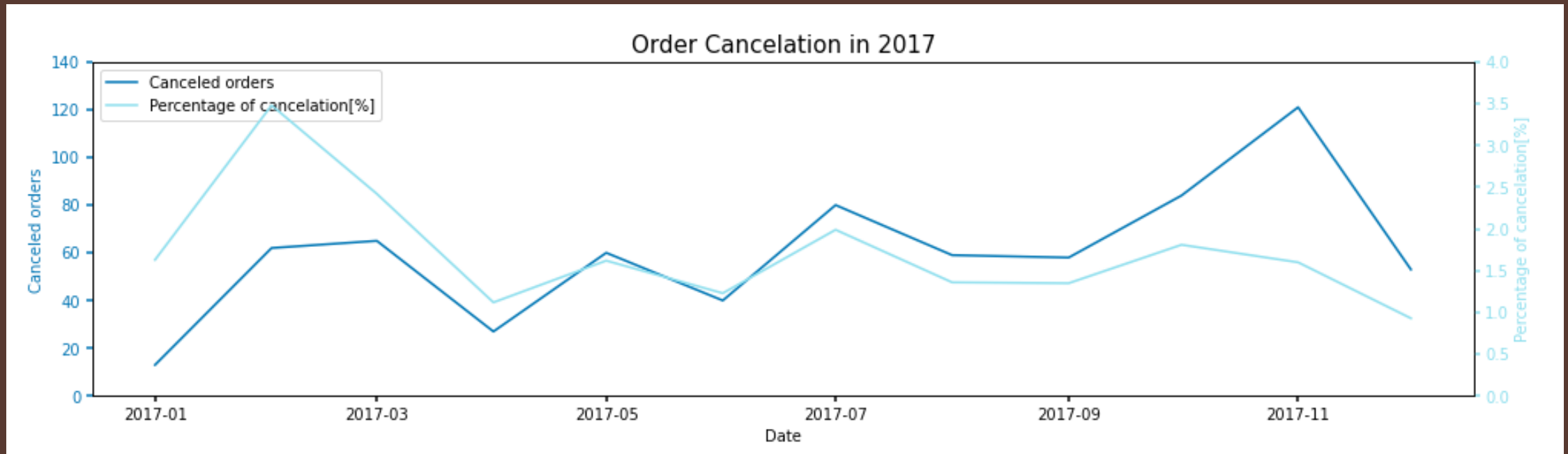


# Monthly Revenue in 2017: KDE plot

It seems like the mean of revenue per order are similar in each month.  
The increasement of revenue may caused by the order amount.



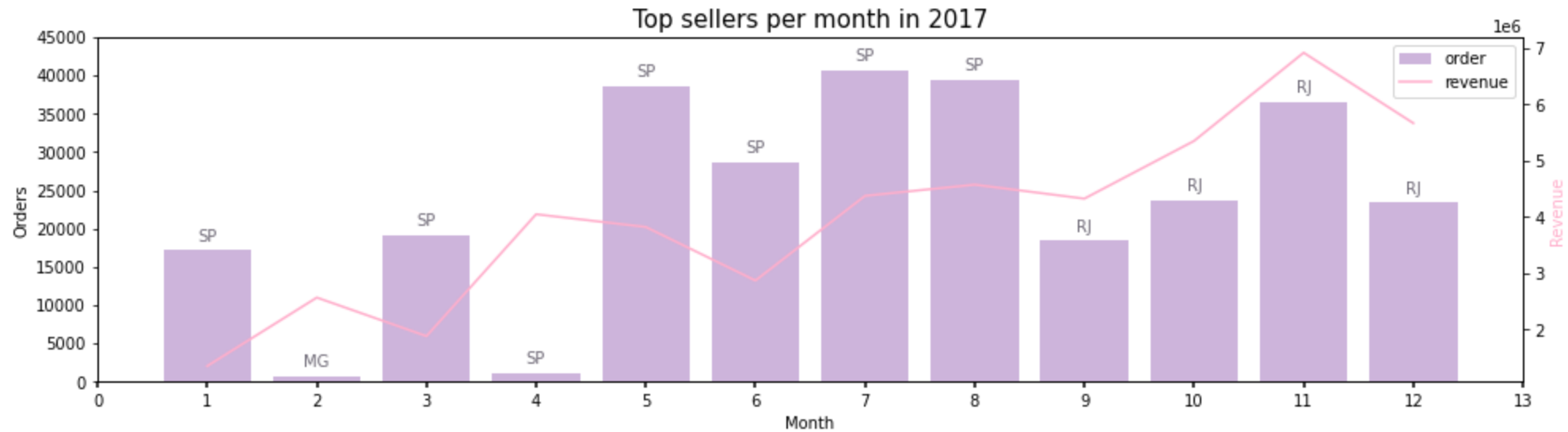
# Cancellation per month in 2017



Finding:

- In February percentage of cancellation is relatively high.
- In November the amount of canceled order is high but the total order is also high.  
Therefore, the percentage of cancellation is not that high

# Popular product category in 2017



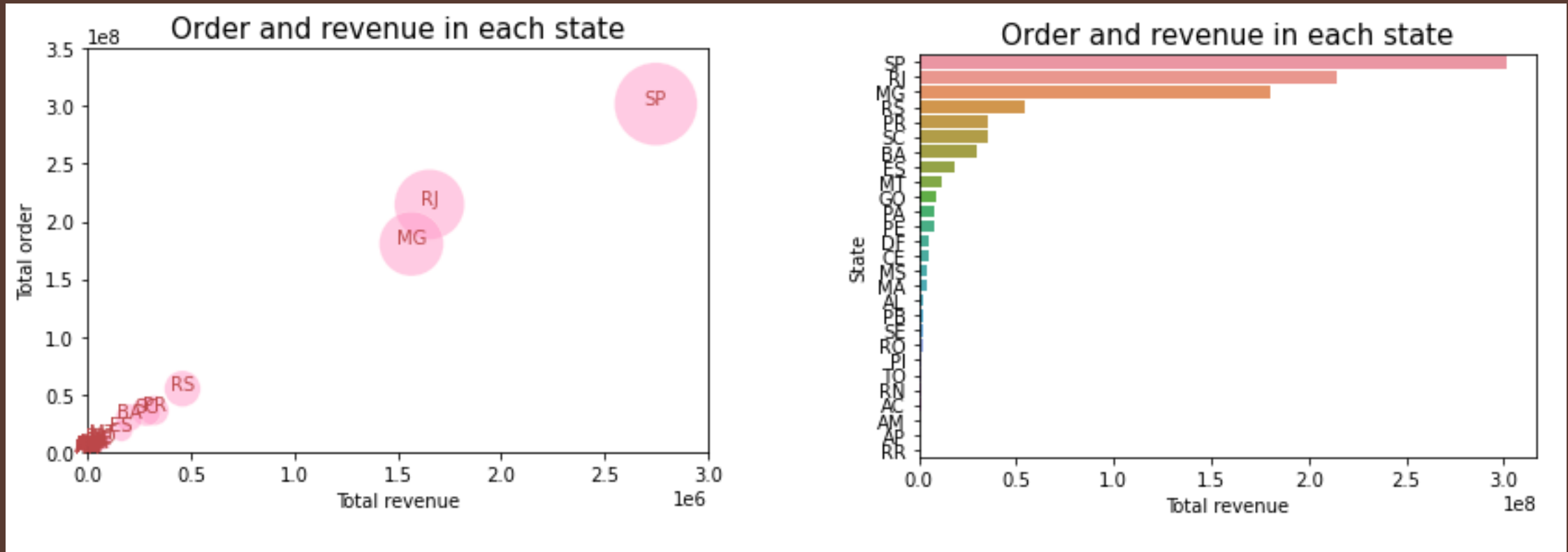
The chart show the top seller per month in 2017.

The ids of top sellers show in the table

The chart shows the total orders of the top seller in each month (left axis)  
and the total revenue of the top seller in each month (right axis)

Texts above bars are the state where sellers located

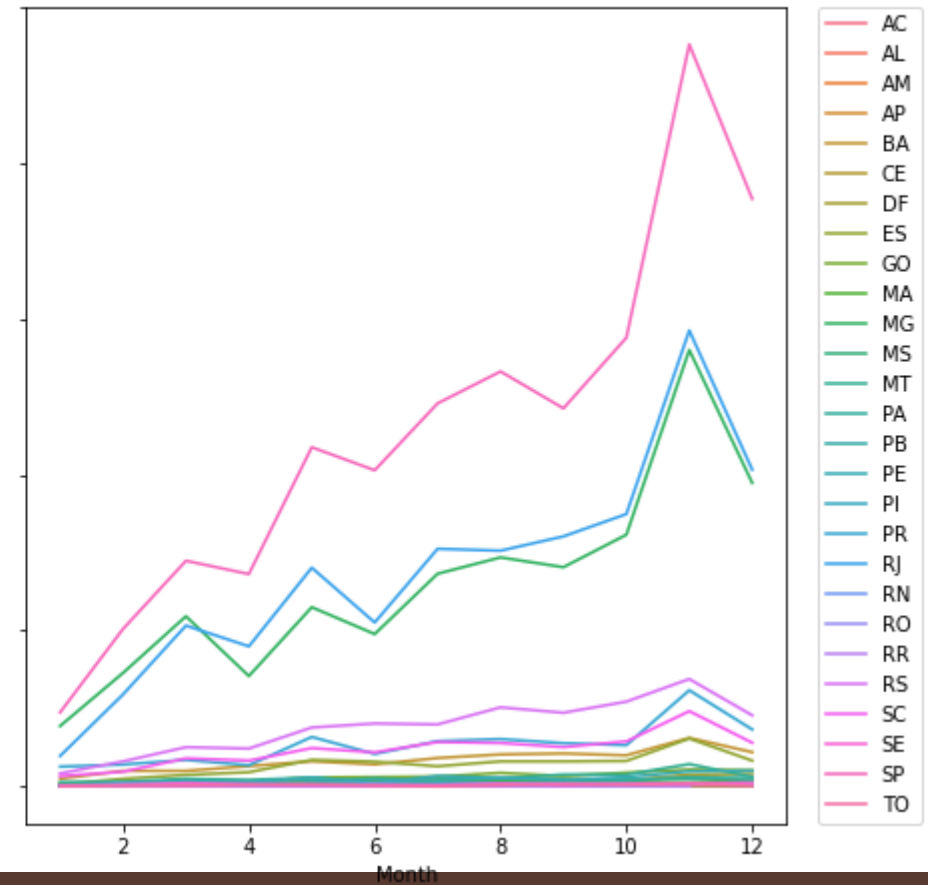
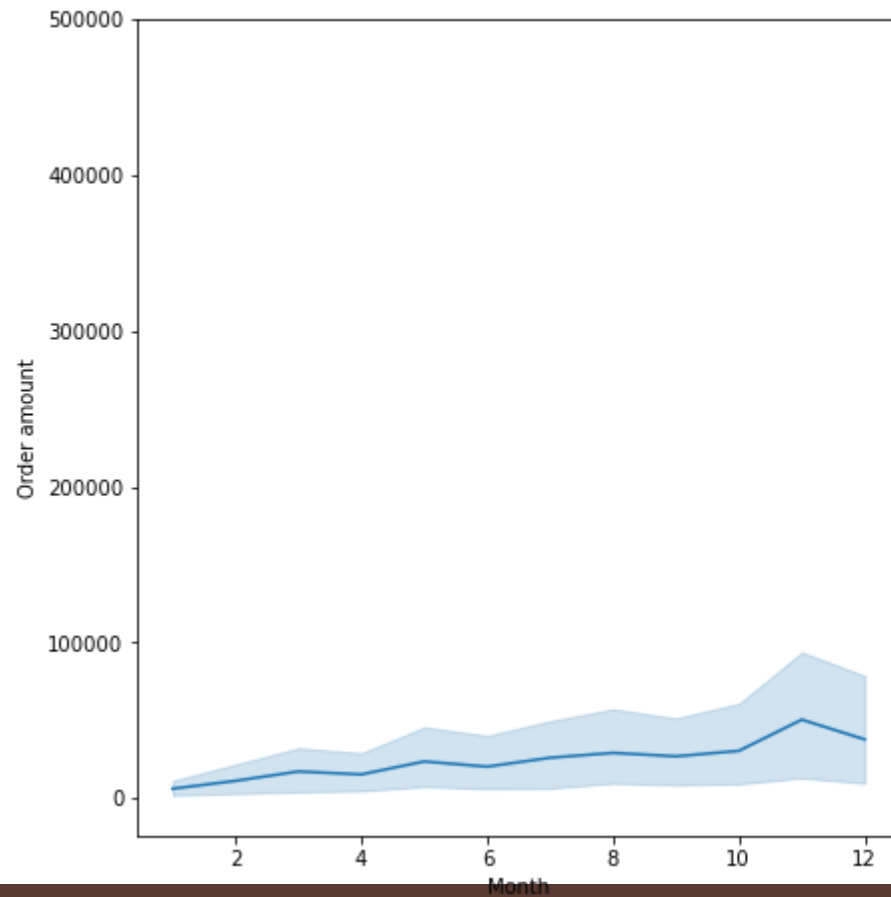
# 2017 best selling state



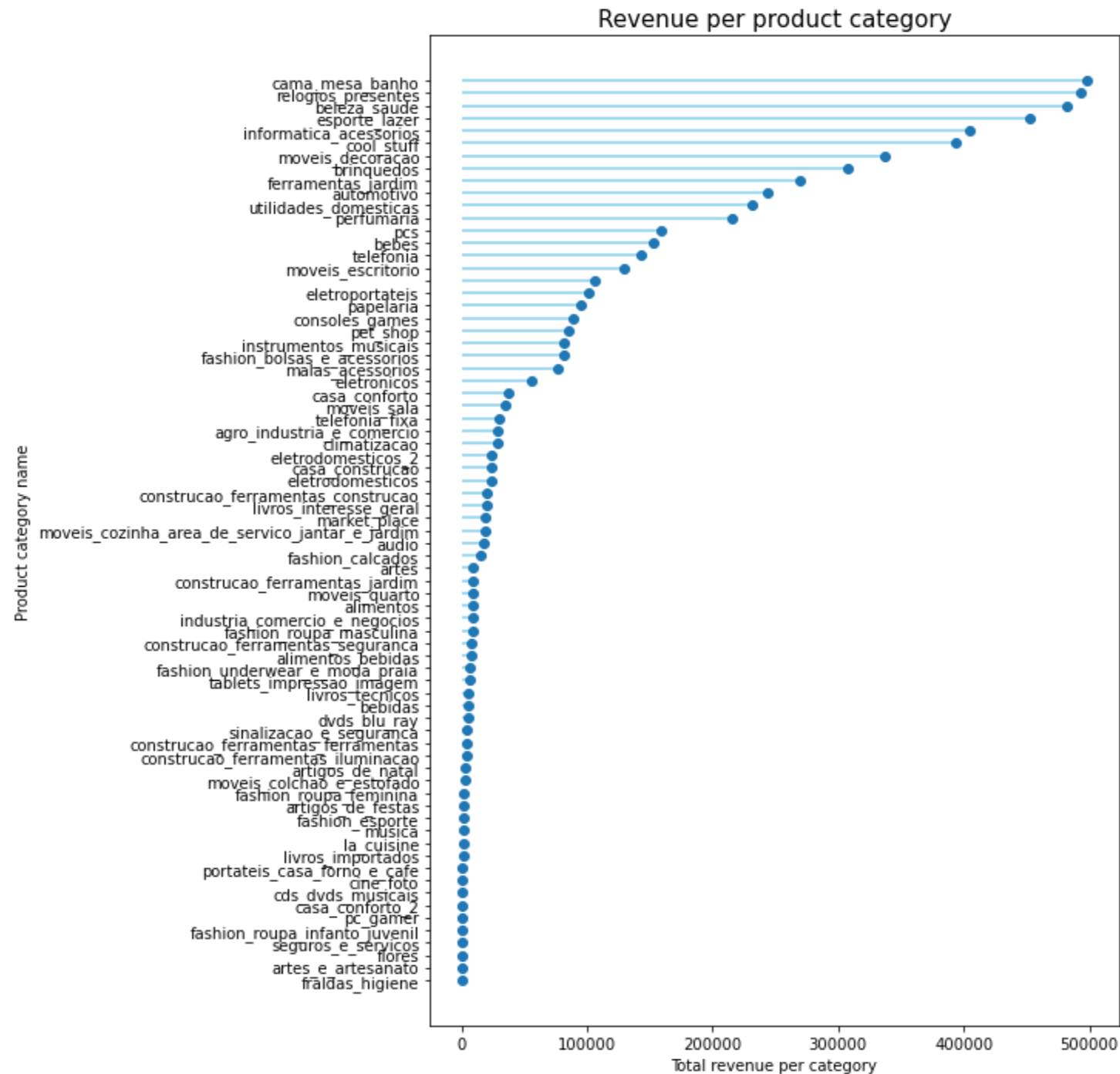
States SP, RJ and MG created the most revenue.

# Monthly order per state

Orders in 2017



# Most popular product category



# Majority payment method in 2017

