SUPPLY CHAIN EDA ANALYSIS

Decision Making Via Data analysis Project
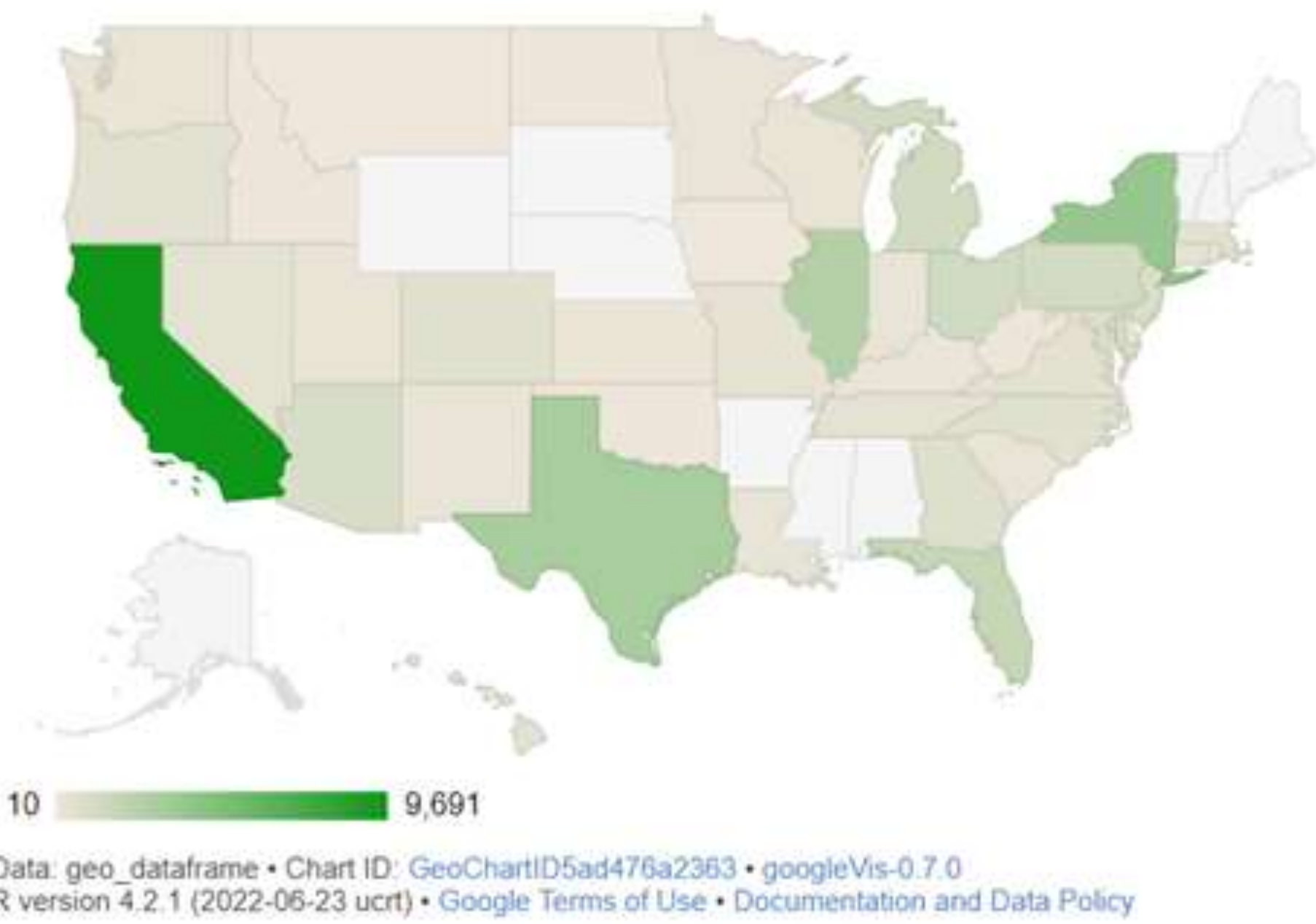
Creator: Harsh Shah

Introduction

Supply chain is a key measure to consider for any business. A business with good supply chain network has the capability to be pioneer in the industry. Supply Chains are the Totality of processes spanning operations from supplier to end-customer, focused on material, work and information flow. Supply chain management is a crucial process because an optimized supply chain results in lower costs and a more efficient production cycle.

Methods

This data is taken from Kaggle. This dataset is used by the company DataCo Global for analysis of supply chain and making decisions. EDA which is exploratory data analysis is presented here with the help of R software and various libraries like maps, mapproj, Ggplot2, data.table etc. Live and interactive graph are also presented with the help of GoogleVis library

Primary question answered

- Whether shipping affects profits and supply chain?
- Which products affect the profits?
- Which state is vulnerable in shipping?

Data: geo_dataframe • Chart ID: GeoChartID5ad476a2363 • googleVis-0.7.0
R version 4.2.1 (2022-06-23 ucrt) • Google Terms of Use • Documentation and Data Policy

Data Analysis

Our dataset contains all order information shipped from USA to various regions of world. Our data analysis is carried out on R studio, and we have installed various libraries like maps, mapproj, treemap, tidyverse etc, for our analysis. There is no null value in dataset, and we have taken orders from only USA, and which are already complete.

A. Figure is plotted for highest risk by region to understand shipping risks.

B. For profit analysis department name, benefit per order, delivery status and quantity of order to understand which department have loss and to focus on that.

C. Also plotted heat map for various regions, sales, profit, delivery scheduled and profit analysis by states and shipping mode for better analysis.

A

Highest shipping risk by region

B

D

Heatmap for profit and delivery variables according to order regions

E

Profit Analysis by states and shipping mode

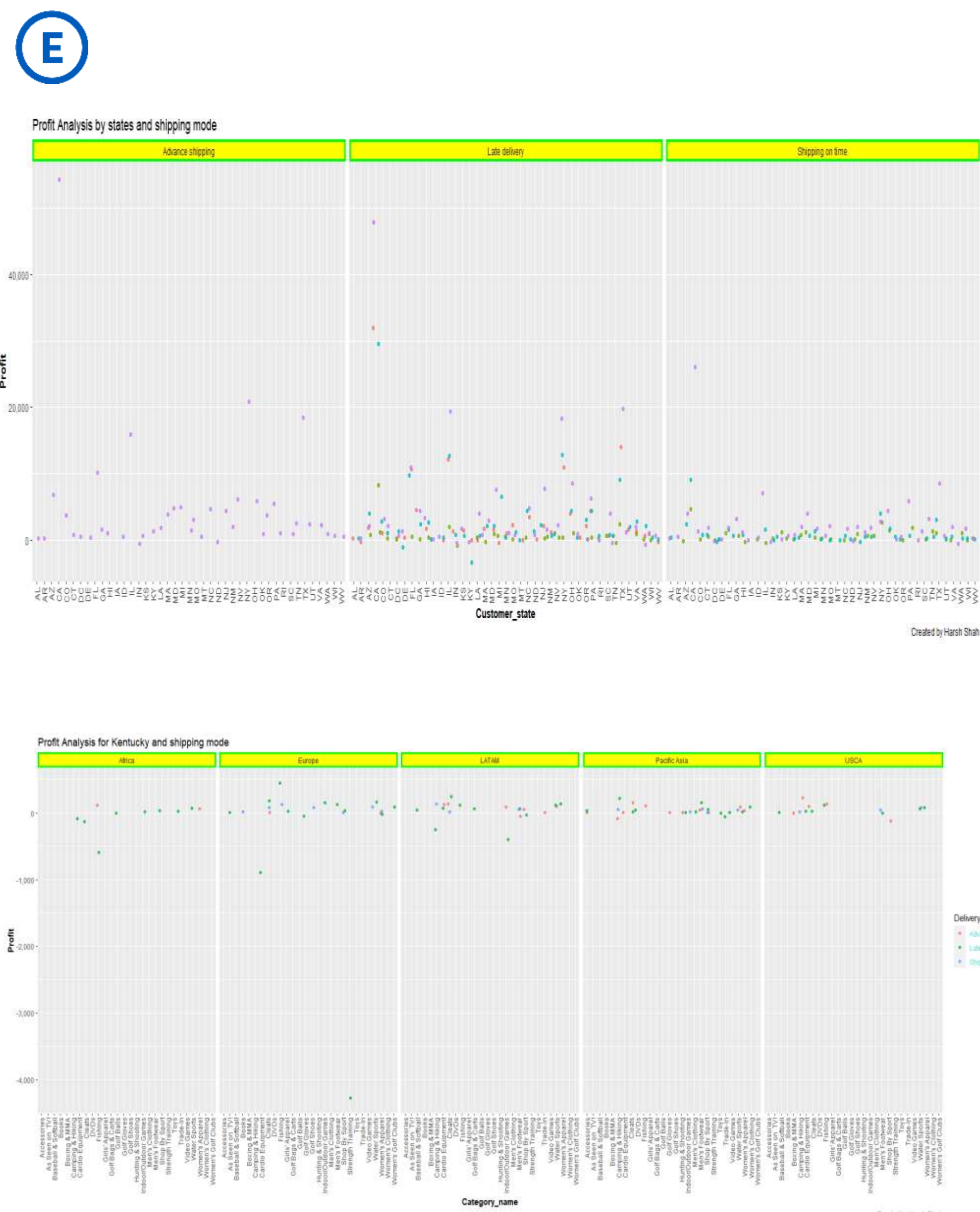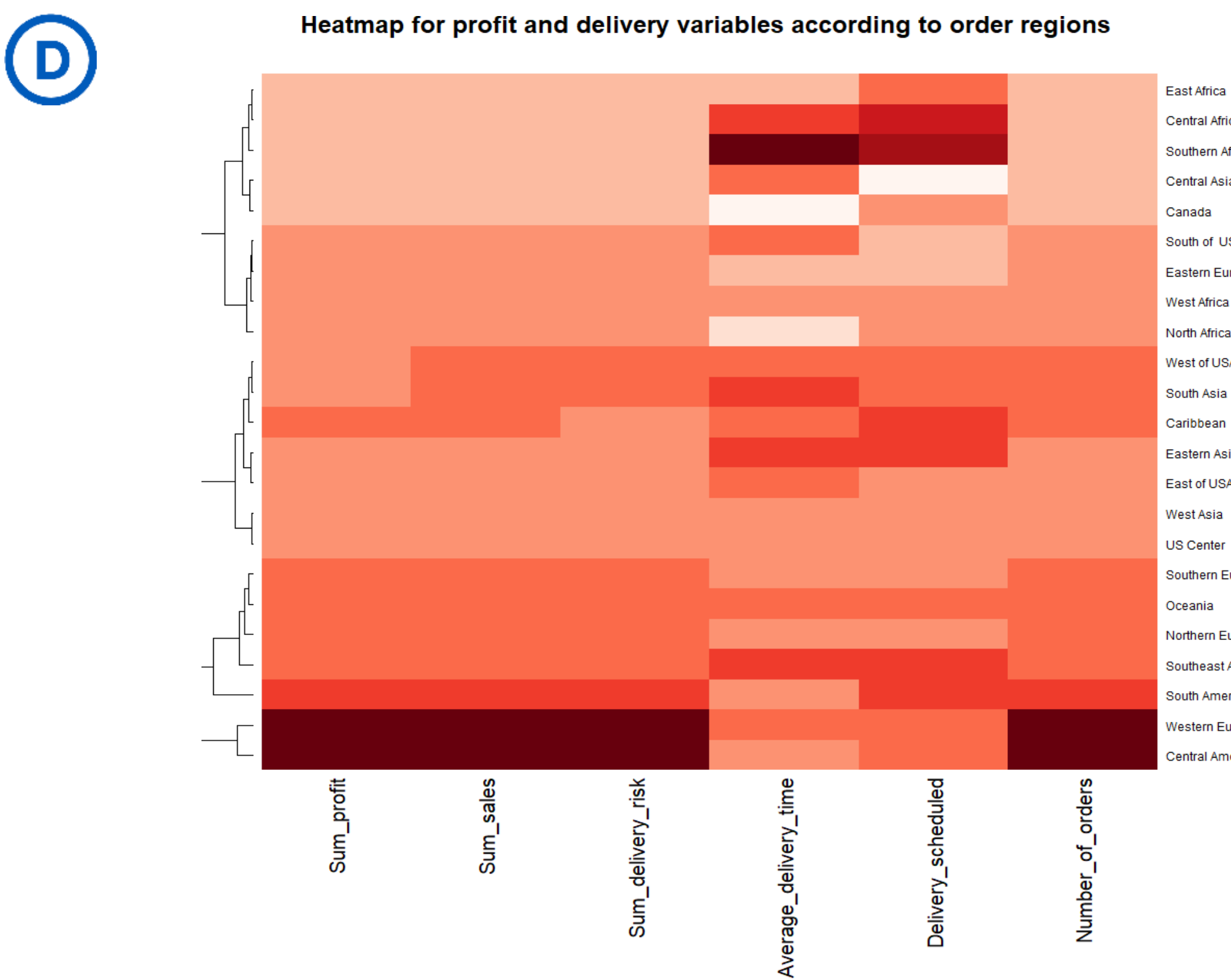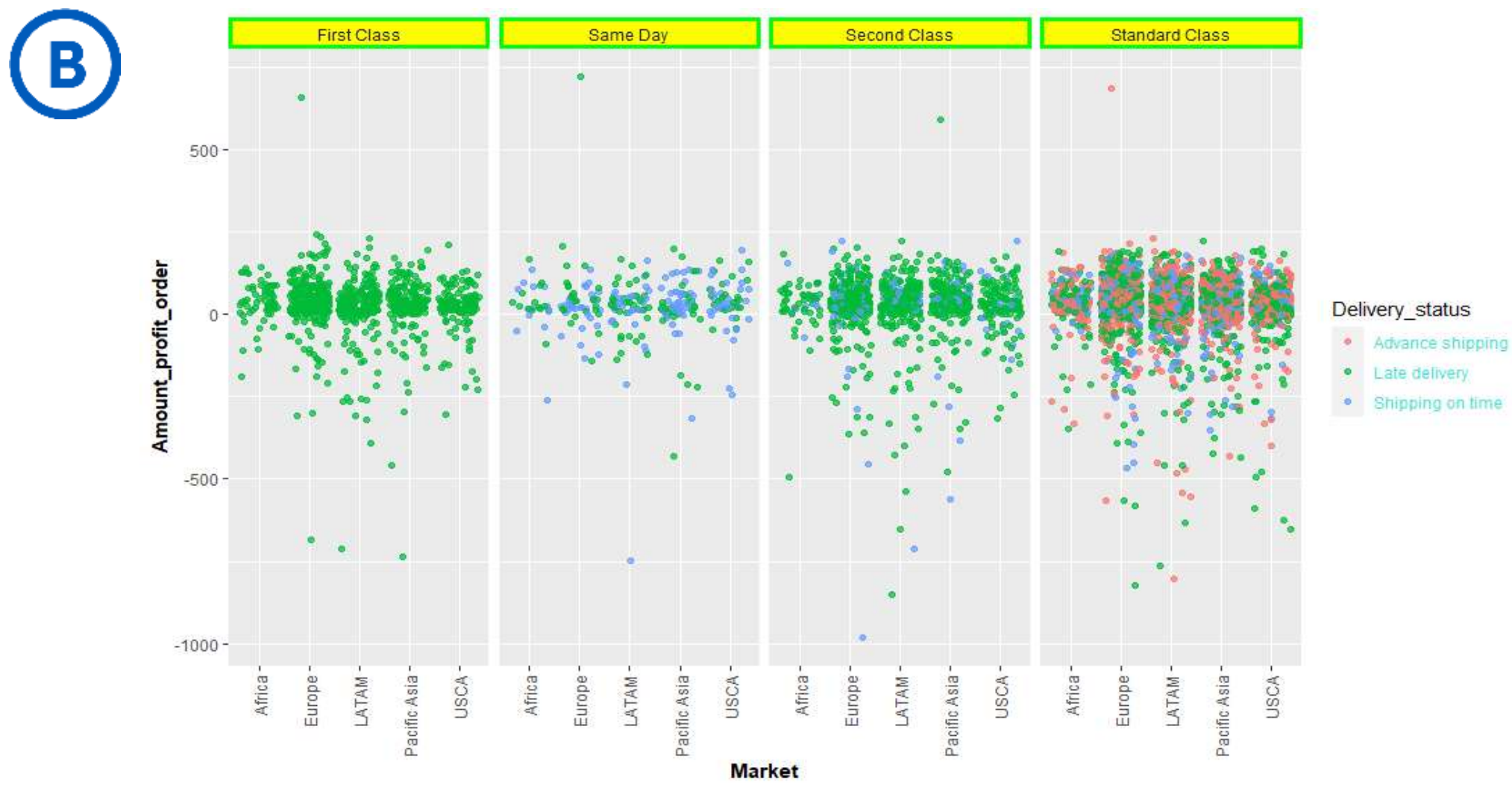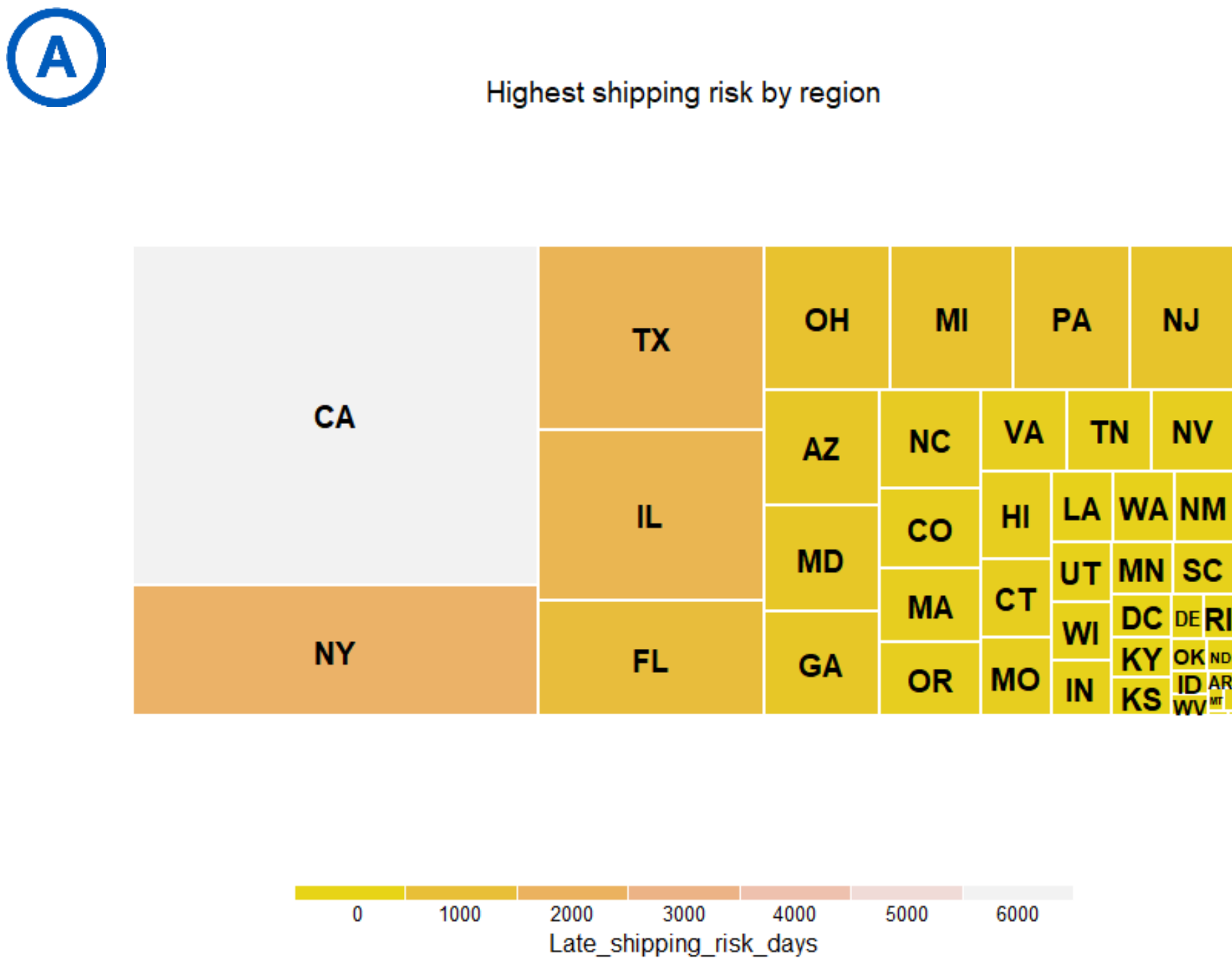Profit Analysis for Kentucky and shipping mode

Figure A: Highest Shipping risk by region

Figure B: Profit analysis by region

Figure C: Heat map for variables according to region

Figure D: Profit analysis by states and shipping mode

Conclusion

- From the initial graph California has the highest number of orders followed by New York, Texas, Illinois etc.
- Coming further to our analysis highest sales were Western Europe followed by Central America followed by Southern Europe and South America.
- For our analysis on department sales, it was observed that Fan shop has the highest sales followed by Apparel and Golf.
- California, New York, Illinois, Florida has highest risk in shipping. Higher the number of orders higher is the risk of shipping it seems.
- From profit analysis it seems there is loss with many products which mainly were delivered late and higher quantities. Some of departments with highest loss were apparel, Fitness, Golf and Fan shop.
- Coming to our shipping analysis it clearly seems shipping done by First class were all late deliveries and majority of shipping done by second class were also late. Whereas loss is highest on products shipped with standard shipping.
- From the figure it seems that for advance shipping only standard shipping service and there was no loss in advance shipping which seems that shipping might affect the profits. And there is loss when delivery is late delivery. And when there is loss, it was shipped with second class. And we took example of Kentucky for further analysis.
- From analysis it seems that with terms of sum profit, sum sales, and delivery risk, Western Europe and Central America have the highest profit, sales and delivery risk. But with respect to average delivery time southern Africa has the highest time of delivery. And lowest have for Canada. With comparison to average delivery time and average delivery scheduled delivery time is bigger than delivery scheduled for regions like Southern Africa and Central Asia are among the highest. Whereas for Canada average delivery time is less than delivery scheduled. Whereas for East Africa also the average delivery time is less than delivery scheduled.
- And majority of loss have been shipped late. Hence it can be reasonable to conclude that shipping might affect the profit for the orders. For Africa region losses are in department of Fishing, Cardio equipment and Camping and Hiking that too with late shipped. Hence this problem should be addressed. For Europe region also cardio equipment is a loss product and also golf products. For Latin America, Camping and Hiking equipment faces loss and even indoor and outdoor games. There is not much loss in ASIA region and USCA region.

References

https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis

Department: Engineering Management
Email: hshah64@stevens.edu