

Case Study Project - Office Supplies

Kevin Chu



An overview of the project and business goals

Background:

- Pens & Printers is a national office supplies chain. At the moment, they send office supplies out of warehouses in four regions: East, West, South, and Central, and all four warehouses stock the same products. The Head of Sales thinks this leads to large amounts of unsold products in some locations.

Data Integrity and Inconsistence:

- Na values existed for “Profit” Column.
- Number of unique values between Variable “Product.ID” and “Product.Name” not aligned. Duplicated Product names are used for different ID

Business Goal:

- Are there products that do not sell as well in some locations?
- Are there any other patterns over time in each region that you can find in the data?



Are there products that do not sell as well in some locations?

- Sold product quantity for each region by Category/Sub-category

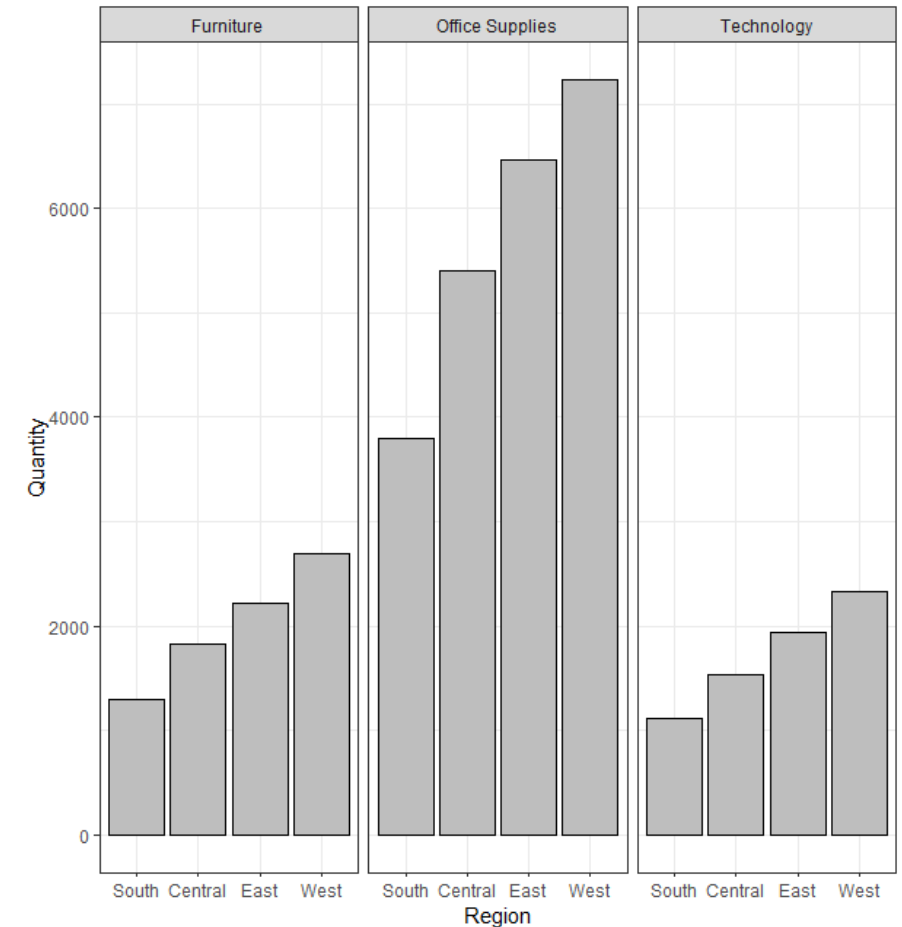
- Descriptive Summary and the bar plots shows the comparison between region.
- The result shows that there is indeed unbalance situation existed.

Descriptive Summary for Sales and Quantity sold

By region and Category

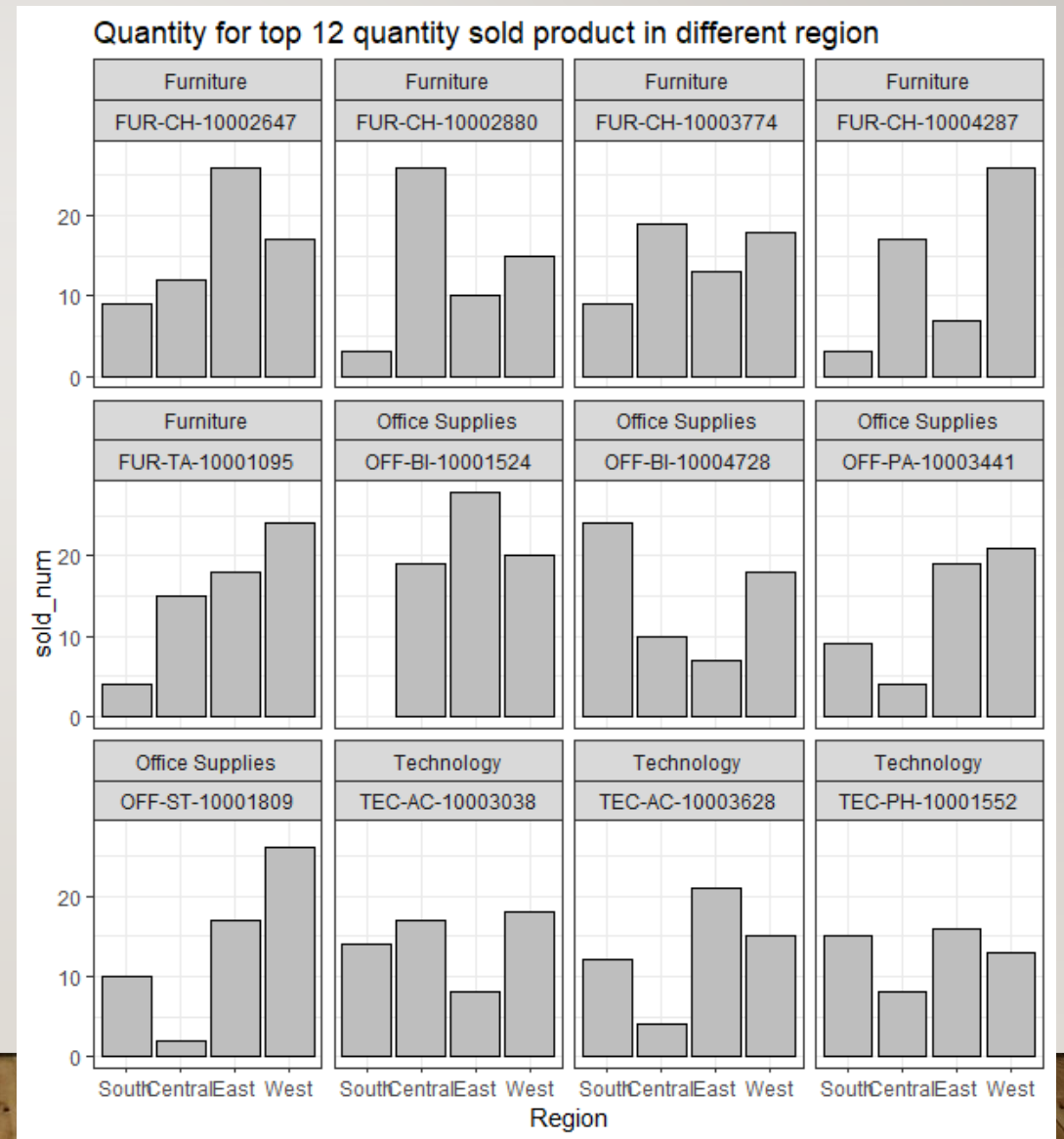
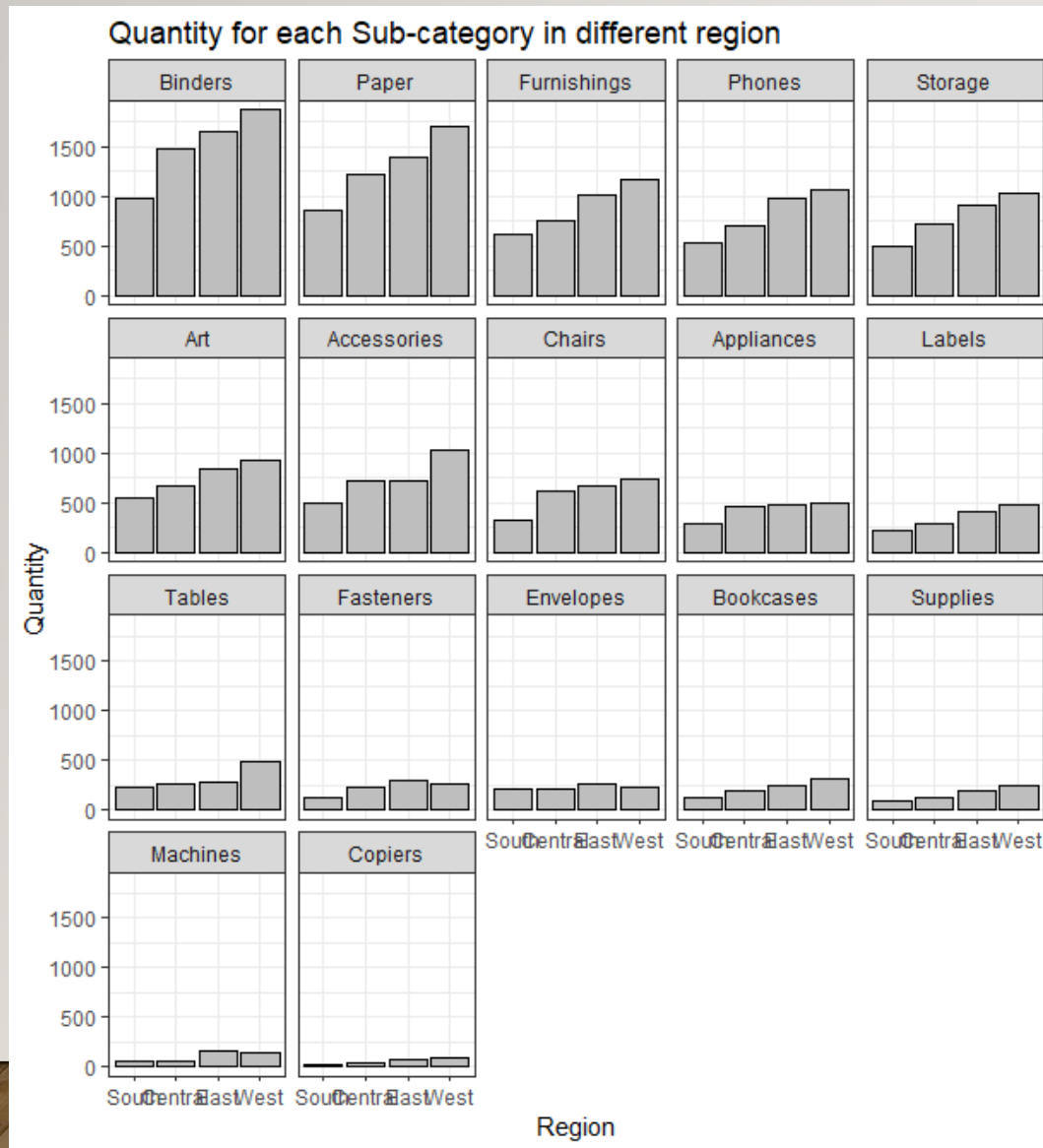
| Region | Total_sold | Total_sales |
|-----------------|------------|-------------|
| Furniture | | |
| West | 2696 | 252612.7 |
| East | 2214 | 208291.2 |
| Central | 1827 | 163797.2 |
| South | 1291 | 117298.7 |
| Office Supplies | | |
| West | 7235 | 220853.2 |
| East | 6462 | 205516.1 |
| Central | 5409 | 167026.4 |
| South | 3800 | 125651.3 |
| Technology | | |
| West | 2335 | 251991.8 |
| East | 1942 | 264974.0 |
| Central | 1540 | 170152.4 |
| South | 1118 | 148771.9 |

Quantity for each category in different region



Are there products that do not sell as well in some locations?

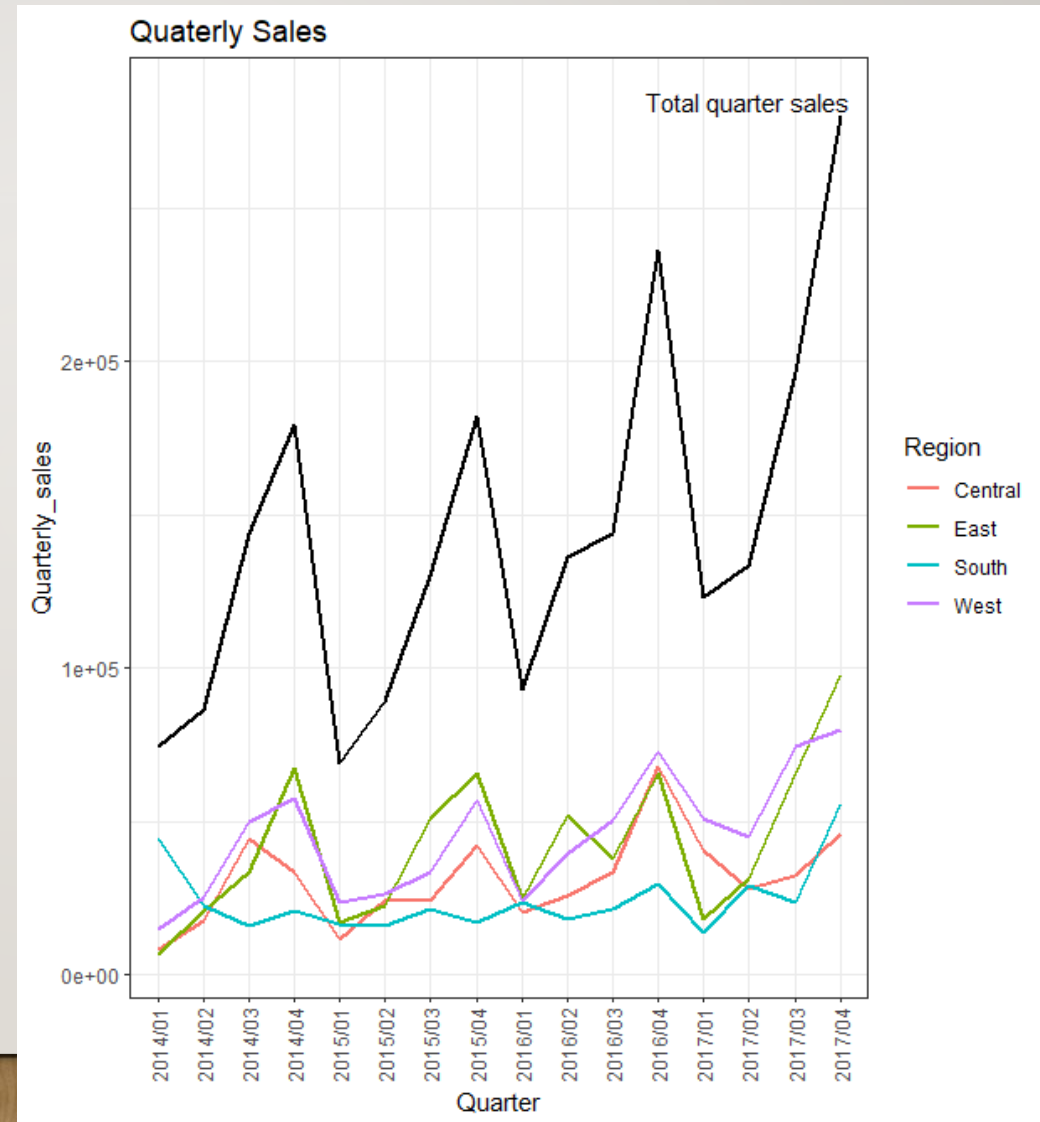
- Sold product quantity for “Different Sub-category” and “top 12 total quantity sold product”



Are there any other patterns over time in each region that you can find in the data?

- Sales performance

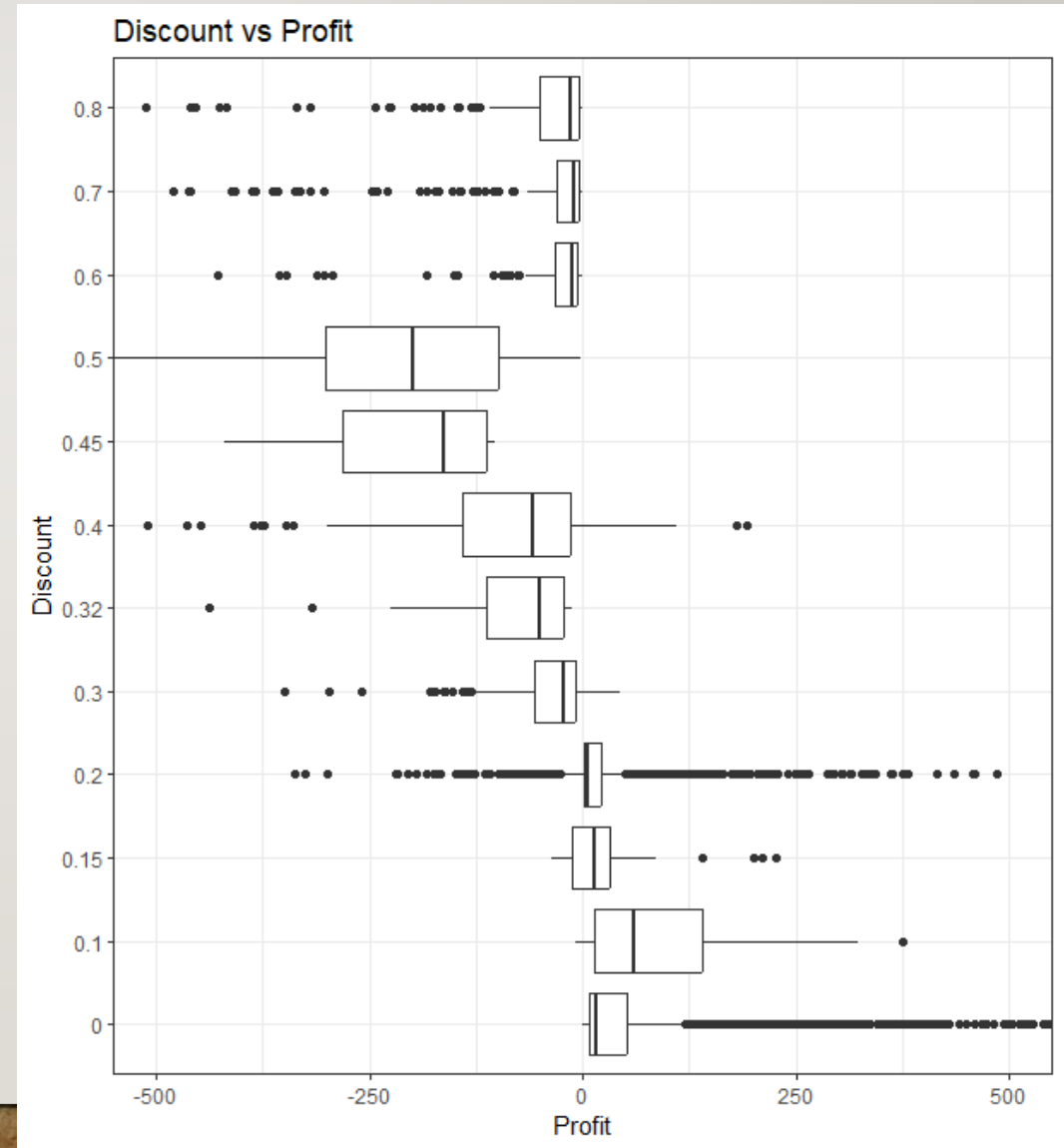
- Trend chart shows the quarterly sales situation from 2014-Q1 to 2017-Q4
- The peak sales usually appears in Q4
- Region sales lines show the same result as we observed in previous session
 - Sales of South and West are better than other two



Are there any other patterns over time in each region that you can find in the data?

- Discount vs Profit

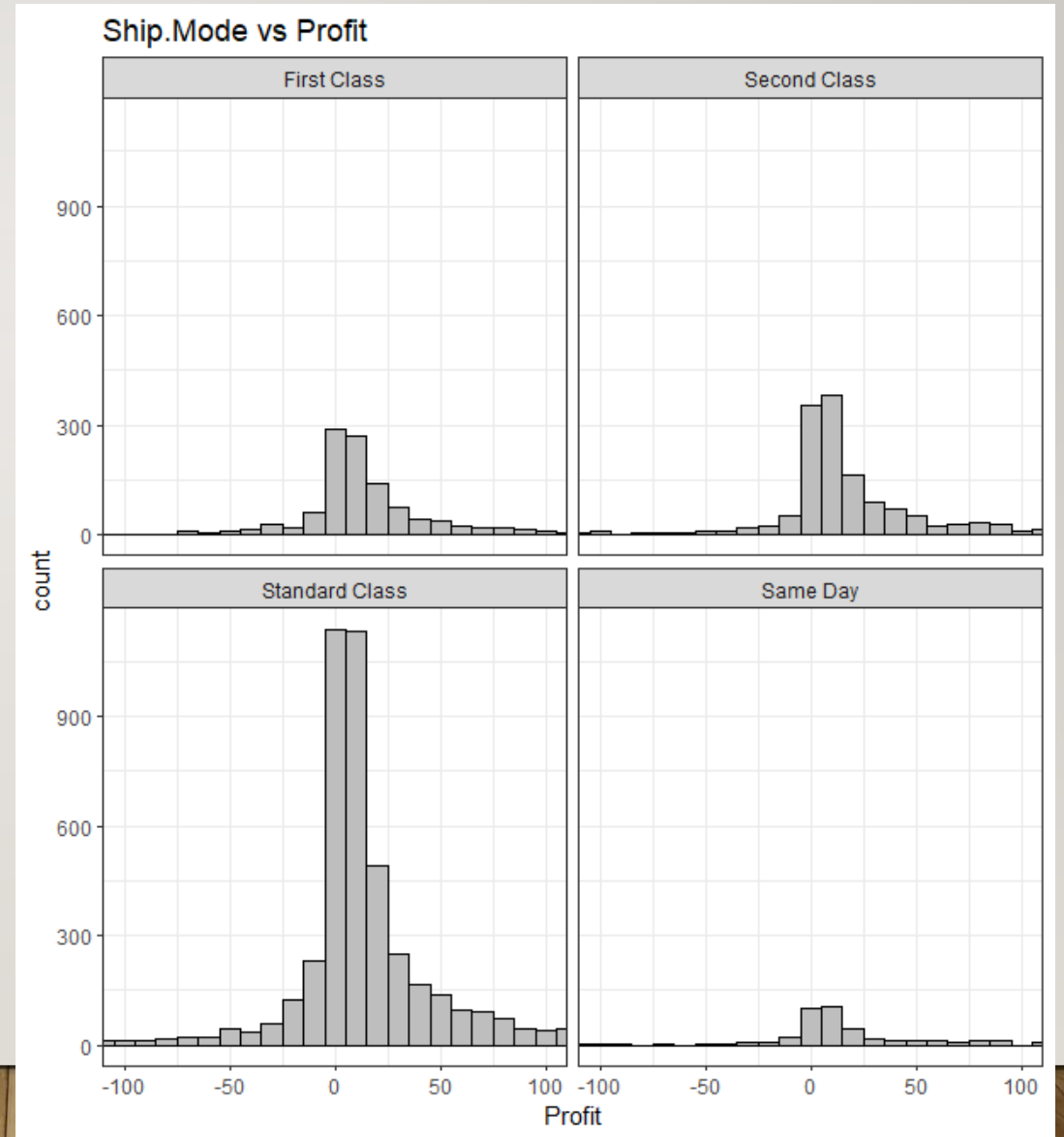
- Remove Na data (1993 rows) in Profit. Will be more accurate if the data is complete.
- Visualization shows that the profit starts turning to negative when discount is larger than 20% off.
- Should control the discount value



Are there any other patterns over time in each region that you can find in the data?

- Shipment vs Profit

- Visualization shows that there isn't obvious difference between profit distribution
- The number of the shipment mode varies though
- Could adjust the shipment fee accordingly.



Summary and recommendation for future work

Overview:

- Product do not sell as well in some locations
- Sales peak usually appear in Q4, probably related to the holiday season
- High discount usually results in negative profit
- Consumers prefer to use standard shipment

Recommendation:

- Adjust product number according to the analysis, store more product on those with higher demands (Especially for West and East area)
- Prepare more product in Q4 in case any of them sold out
- Control the discount value for each product, limit them within 20% off
- Consider to adjust shipment fee for different method. Also, can negotiate with shipment companies to reduce the cost of standard shipment.
- Data completeness of profit values would increase the accuracy of the results

