edureka!
LEARNING CENTER

Capstone Project

Shopify Analysis

Team member

Aswini Kumar Motapothula

edureka! LEARNING CENTER

Shopify

- Shopify is an e-commerce platform that enables individuals and businesses to create their own online stores to sell products and services. It was founded in 2006 and has since become one of the most popular e-commerce platforms in the world, with over 1 million businesses in more than 175 countries using the platform.
- Shopify offers a variety of features and tools to help businesses create and manage their online stores, including customizable website templates, a secure payment gateway, inventory management, marketing and SEO tools, and 24/7 customer support. The platform also offers a range of apps and integrations to help businesses extend the functionality of their online stores.
- Shopify operates on a subscription-based model, with businesses paying a monthly fee to use the platform. Pricing plans vary depending on the features and level of support required, and there is also a free trial available for businesses to try the platform before committing to a subscription.

edureka! LEARNING CENTER

Discussion Topics

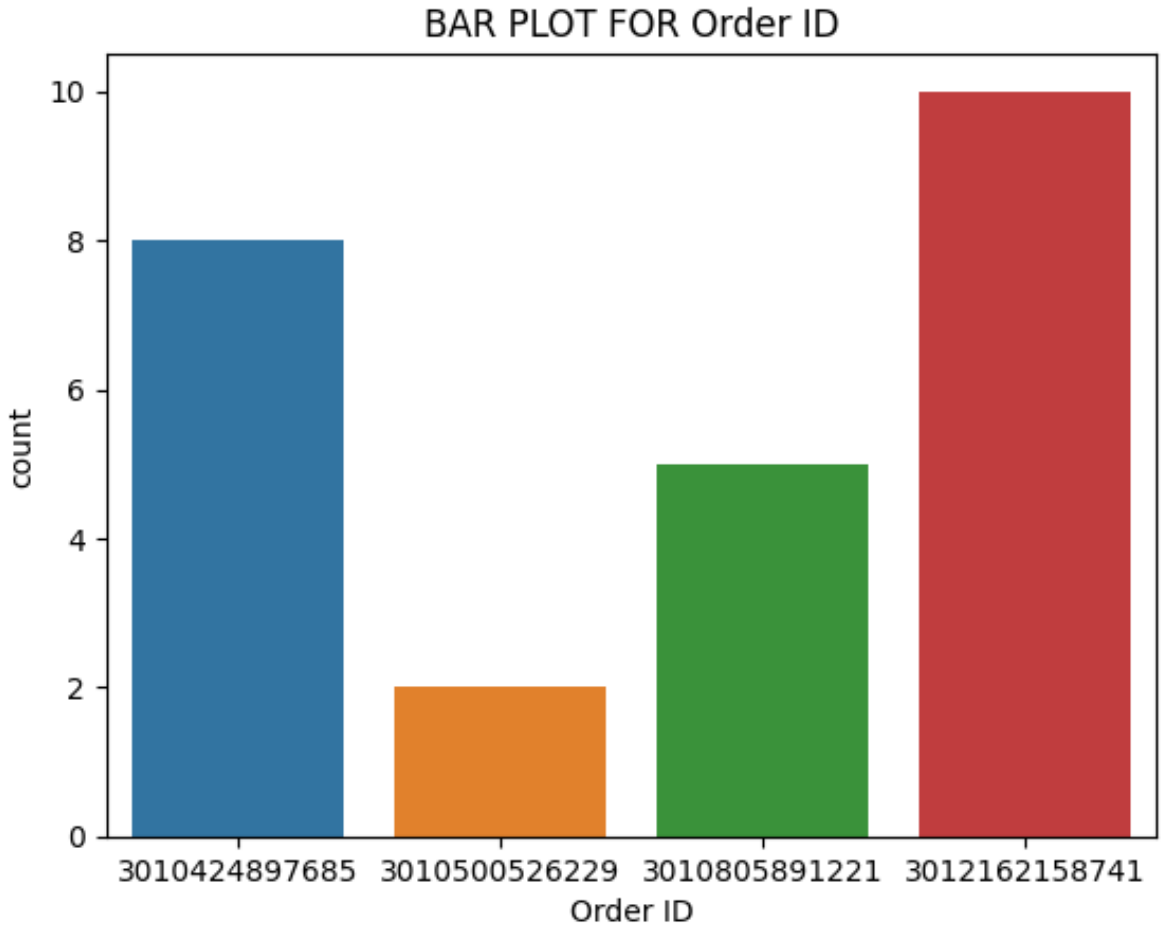
1. Which Order ID has highest Sales.
2. Which is highest Billing City
3. Product Vendor has highest Net sales
4. Network map between product and product vendor
5. Highest Sales of product vendor
6. Monthly sales
7. Co-relation between the columns
8. Plots for only text data (Network map and Word cloud)

Attributes of each variable:

1. Sale ID
2. Date
3. Order
4. Transaction type
5. Sale type
6. Sales channel
7. Billing country
8. Billing region
9. Billing city
10. Shipping country
11. Shipping region
12. Shipping city
13. Order ID
14. Product vendor
15. Product
16. variant sku
17. Net quantity
18. Gross sales
19. Discounts
20. Returns
21. Net sales
22. Shipping
23. Taxes
24. Total sales

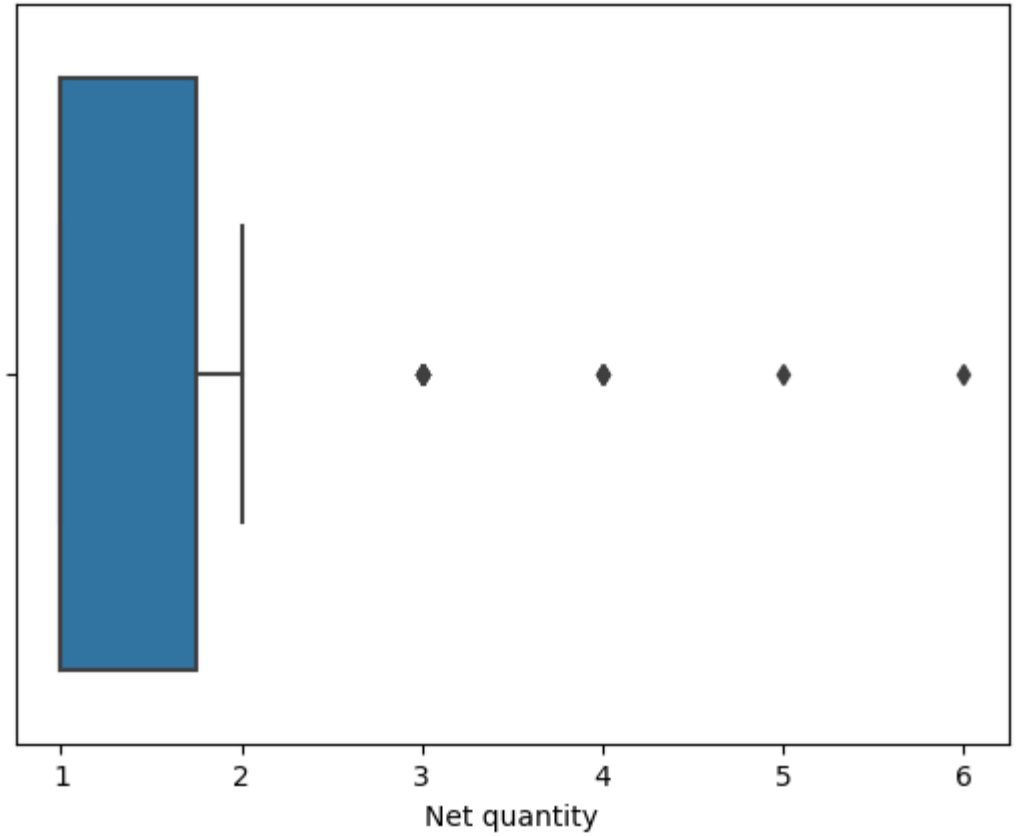
Which Order ID has Highest Sales

- The resulting plot can help in understanding the frequency distribution of the 'Order ID' column, and can be useful for identifying any patterns or outliers in the data.
- However, since it only shows the count of values in the first 25 rows, it may not provide a complete picture of the entire dataset
- Order ID = 3012162158741 has the Largest volume of sales
- Order ID = 3010500526229 has Smallest volume of sales



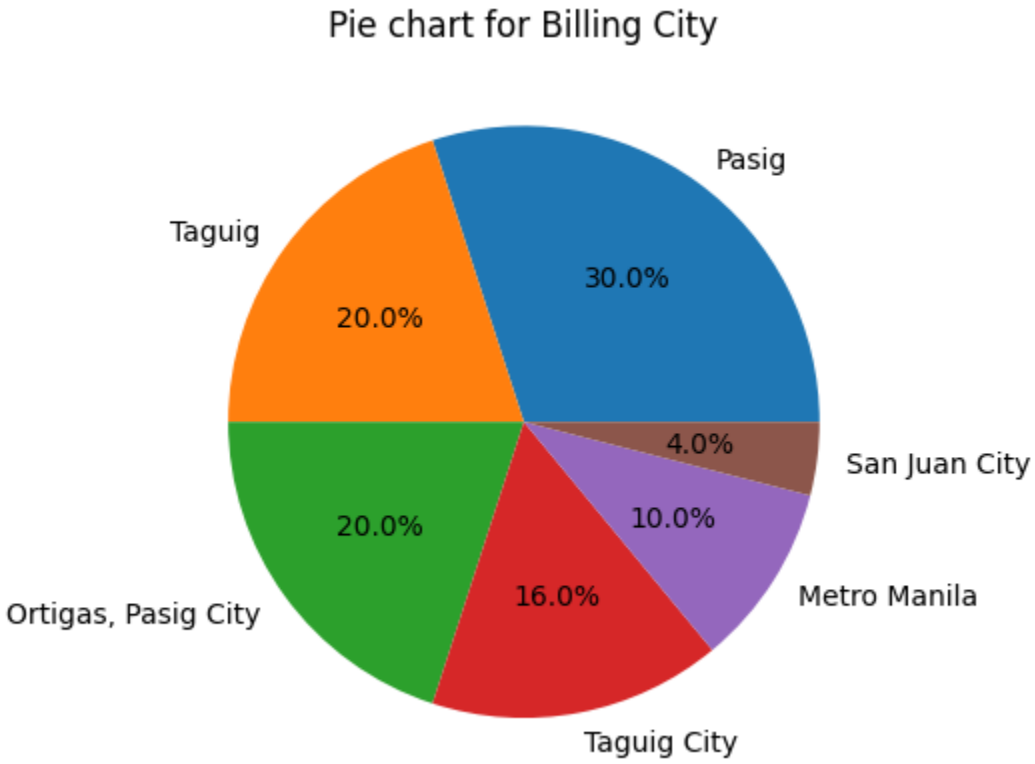
Highest Net quantity

- A boxplot provides information about the distribution of a dataset, showing the median, quartiles, and any outliers in the data.
- The resulting plot can help in understanding the range and spread of values in the 'Net quantity' column.
- However, since it only shows the first 150 values, it may not provide a complete picture of the entire dataset.
- The highest net quantity value is 1



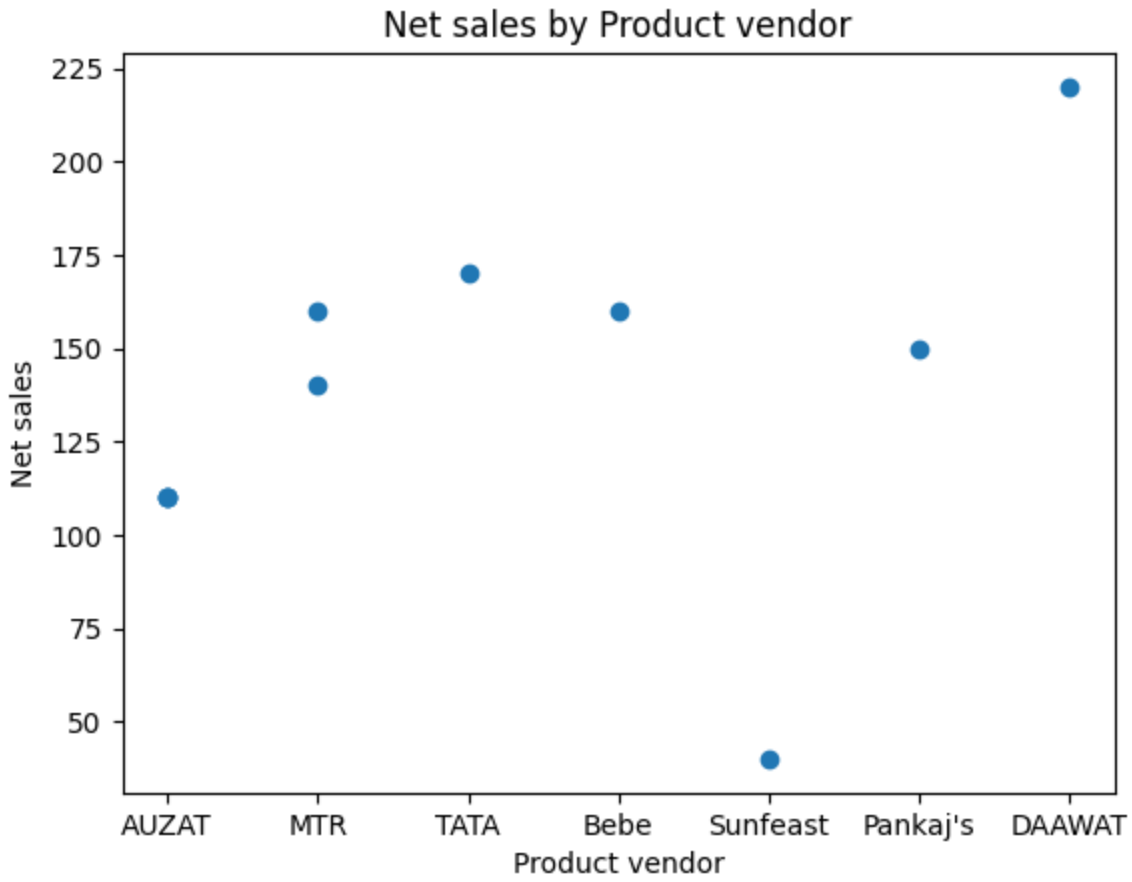
Which is Highest Billing City

- The resulting pie chart can be useful for visualizing the distribution of the 'Billing city' column values and for identifying any patterns or trends in the data. However, since it only shows the top 50 values, it may not provide a complete picture of the entire dataset.
- According to the given data, the city with the highest number of occurrences in the 'Billing city' column is Pasig.
- the city with the lowest number of occurrences in the 'Billing city' column is San Juan City



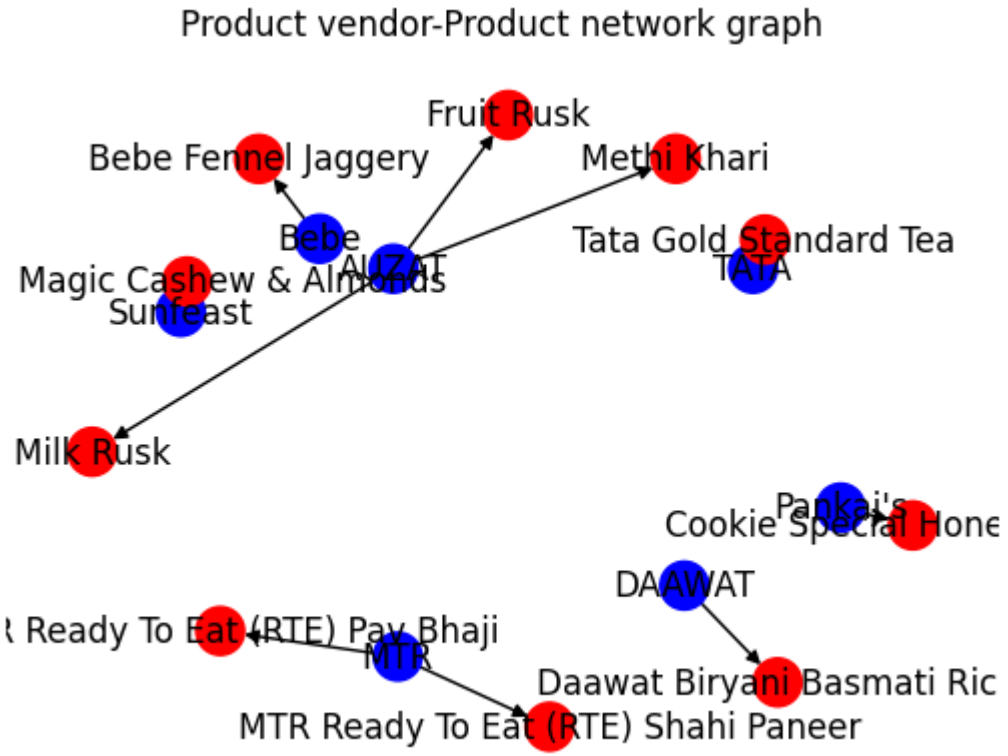
SCATTER PLOT B/W NET SALES AND PRODUCT VENDOR

- The above plot between Net sales and product vendor
- we have taken first five records from the data
- In this plot Daawat Vendor has the highest Net sales
- Sunfeast has the lowest Net sales



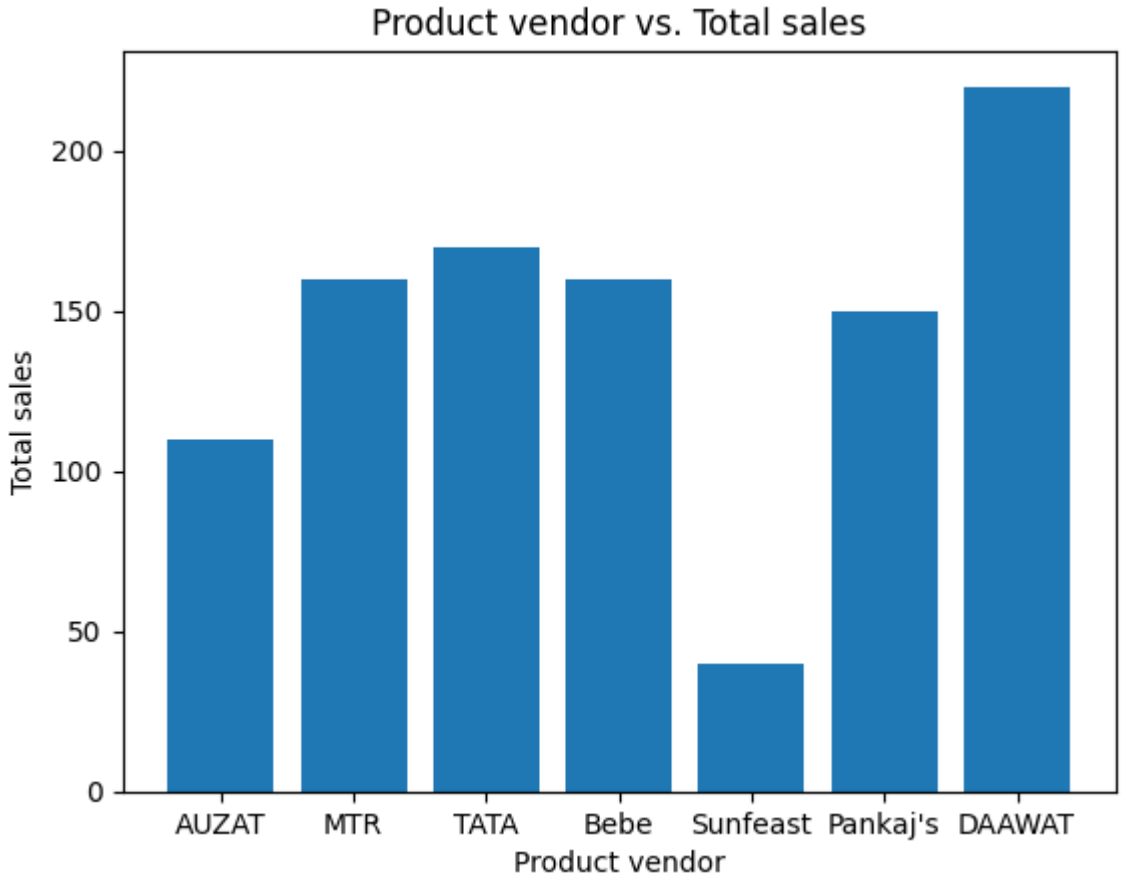
NETWORK MAP B/W PRODUCT AND PRODUCT VENDOR

- This is the Network plot, using for text data only
- in this plot we identify product of the vendor
- Red color indicates product and blue color indicates Product Vendor



BAR PLOT B/W TOTAL SALES AND PRODUCT VENDOR

- This is the Bar plot between the Product vendor and Total sales
- x axis contains Product Vendor and Y axis is Total sales
- From this plot the higest sales of product vendor is DAAWAT
- Lowest sales of the product vendor is Sunfeast

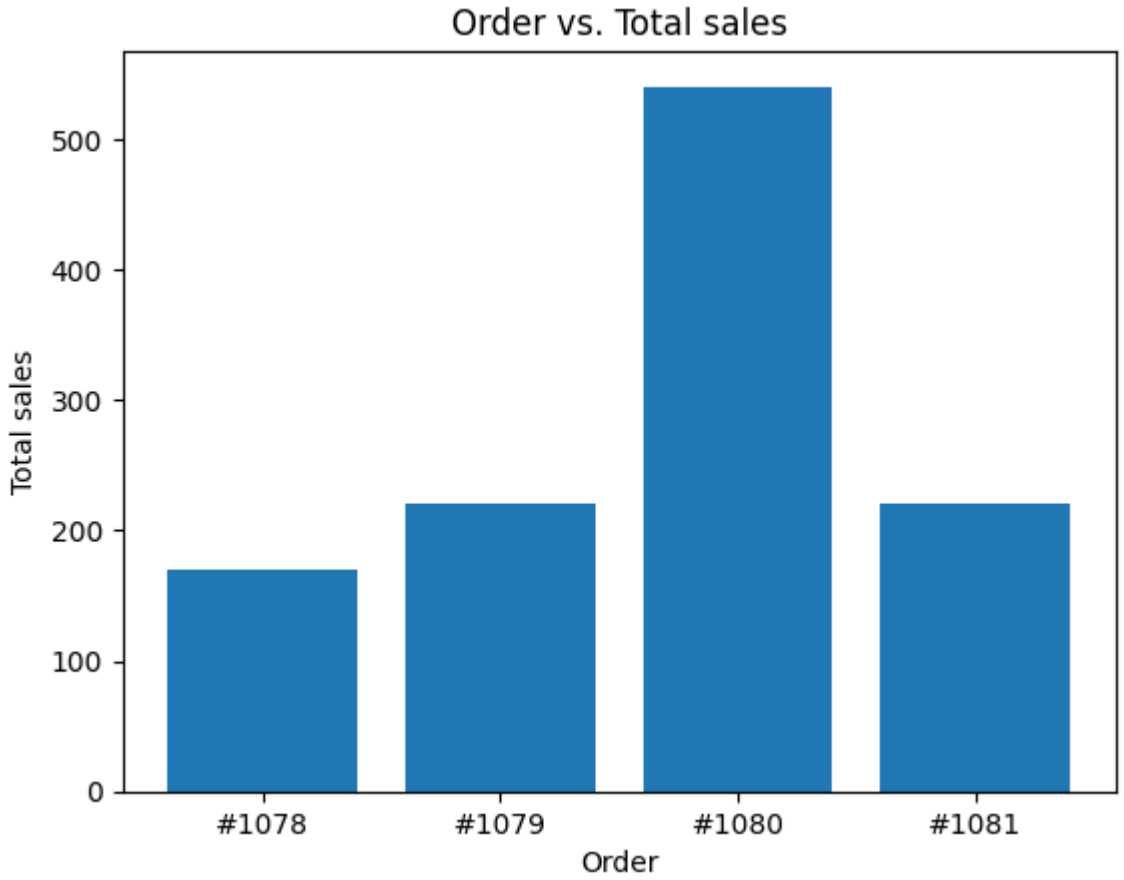


WORD CLOUD B/W SALE TYPE AND PRODUCT

- This is the Word Cloud Plot, this plot for only for text data
- we can draw plots between two text data columns
- In this plot we find relation between product and sale type
- Here all products sale type is Order

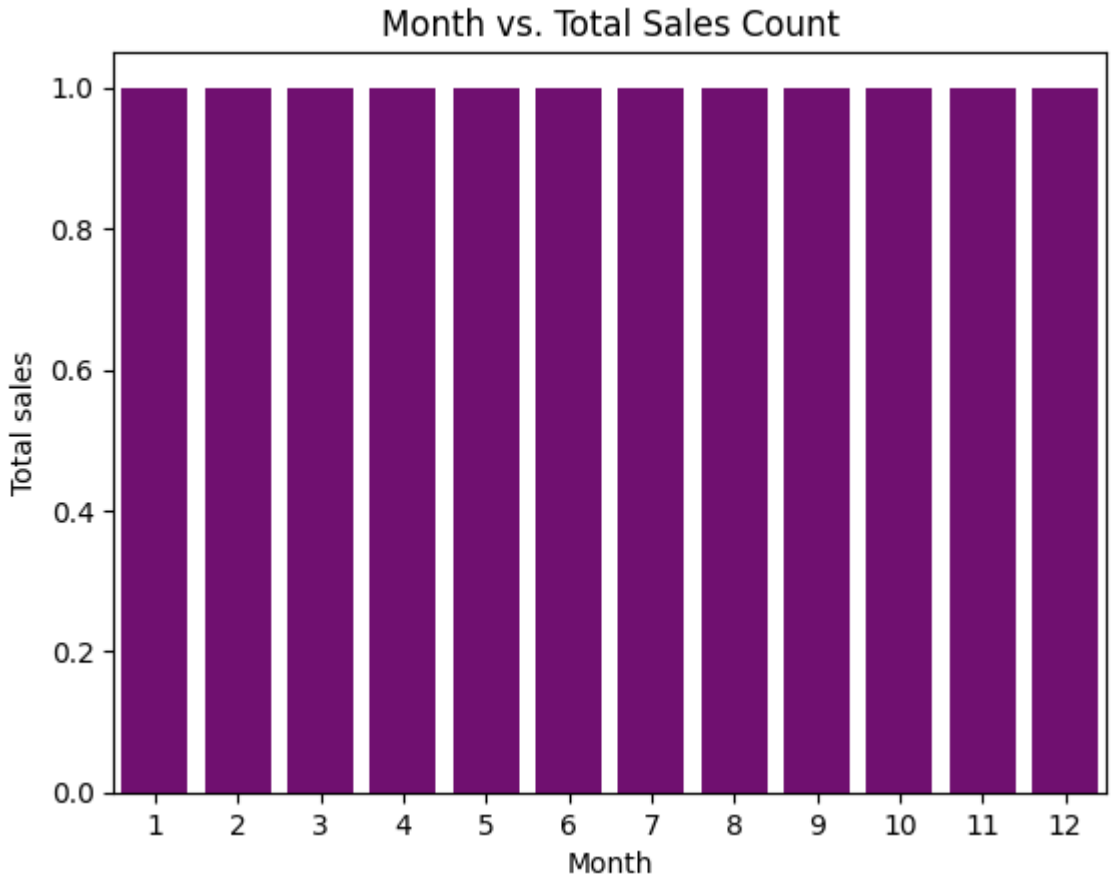
WORD CLOUD B/W SALE TYPE AND PRODUCT

- This plot between the order and total sales
- In this plot the higest total sales is Order no #1080
- Lowest sales is Order no #1078



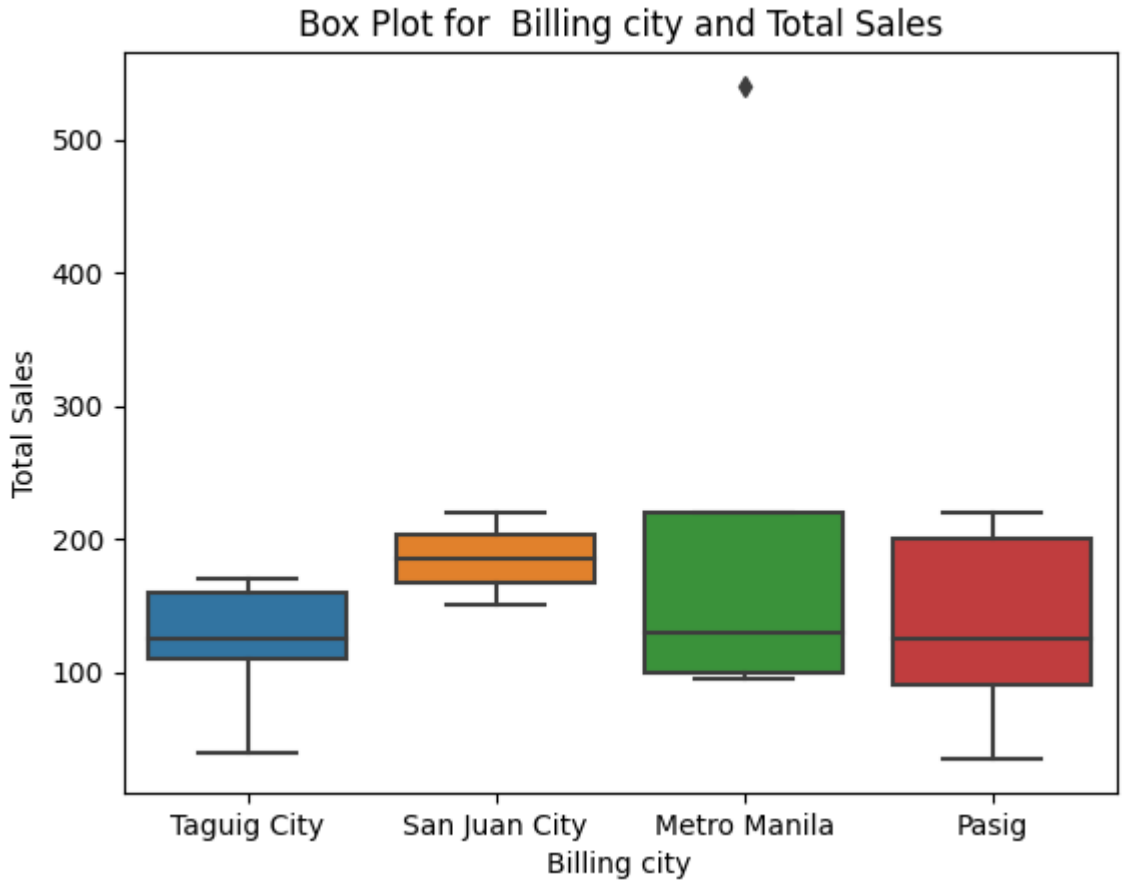
MONTH VS TOTAL SALES COUNT

- This Bar plot between the Month and Total sales
- In this plot we observe from JAN TO DEC Total sales volumes are same



BOX PLOT FOR BILLING CITY AND TOTAL SALES

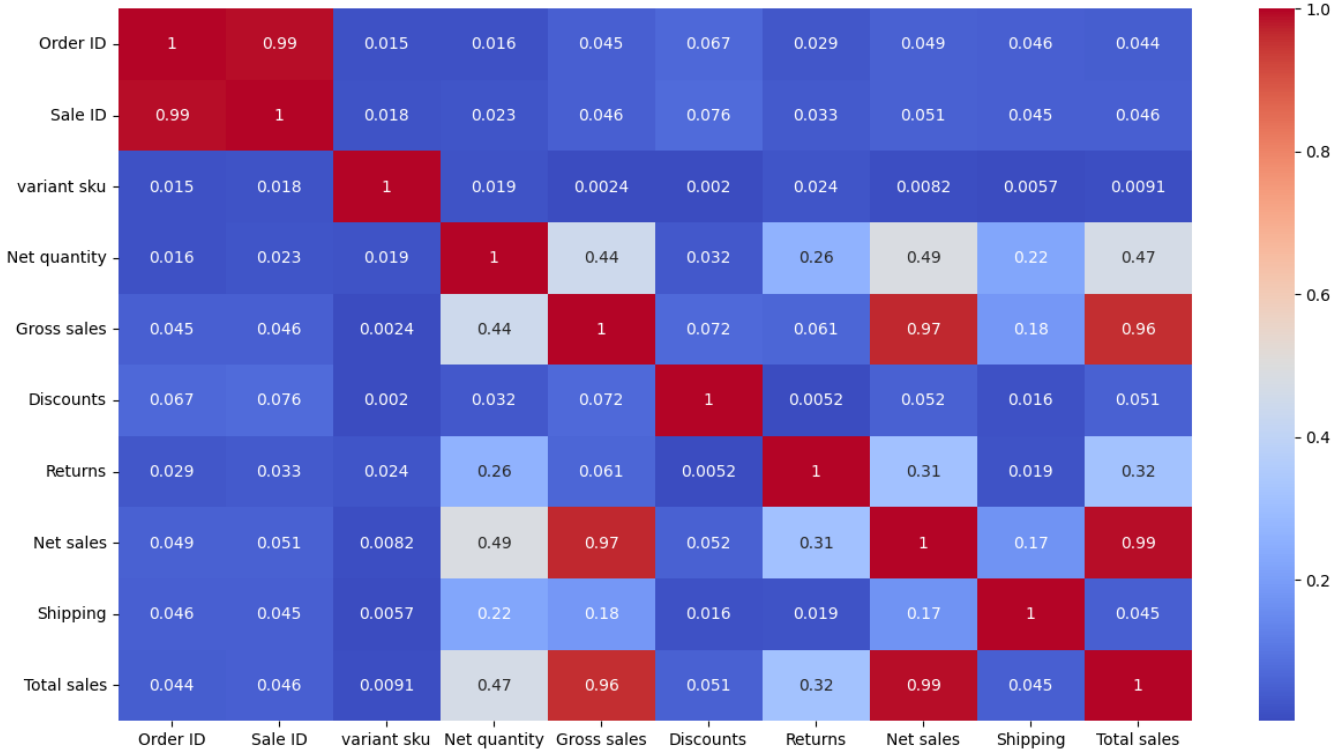
- This Box plot between the Billing city and Total sales
- In this plot the higest total sales volumes from Metro Naila city
- Lowest total sales volumes from Taguig City



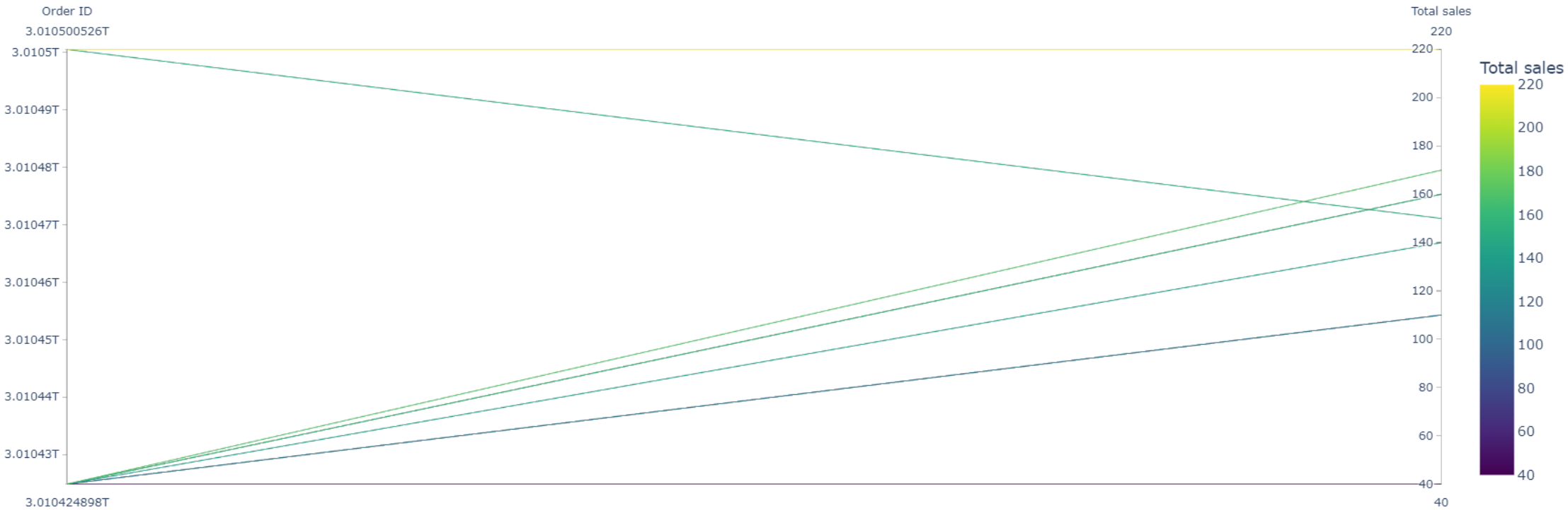
edureka! LEARNING CENTER

HEATMAP FOR ALL COLUMNS CORRELATION

- This is the Heat map in this plot co-relation between all columns in the data set
- The highest correlation between Order id and Sale ID IS 0.99
- Gross sales and Net sales co-relation is 0.97
- Remaining columns have least correlation.



PARALLEL COORDINATES PLOT BETWEEN TOTAL SALES AND ORDER ID

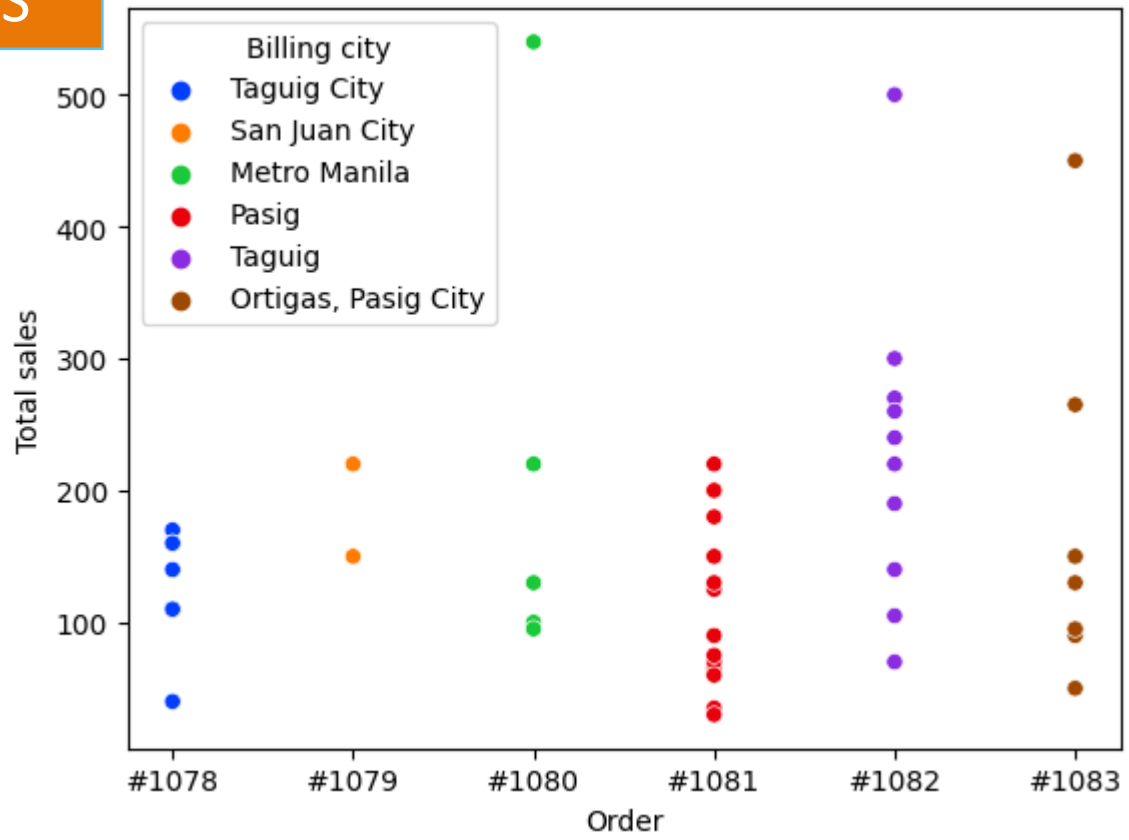


- This is parallel coordinates plot between Total sales and Order ID
- From this Plot the order ID = 3010500526229 having the Total sales 220 and 150
- Order ID = 3010424897685 has more lines between Order ID and Total sales, this Order ID has highest sales than 3010500526229

Code Text

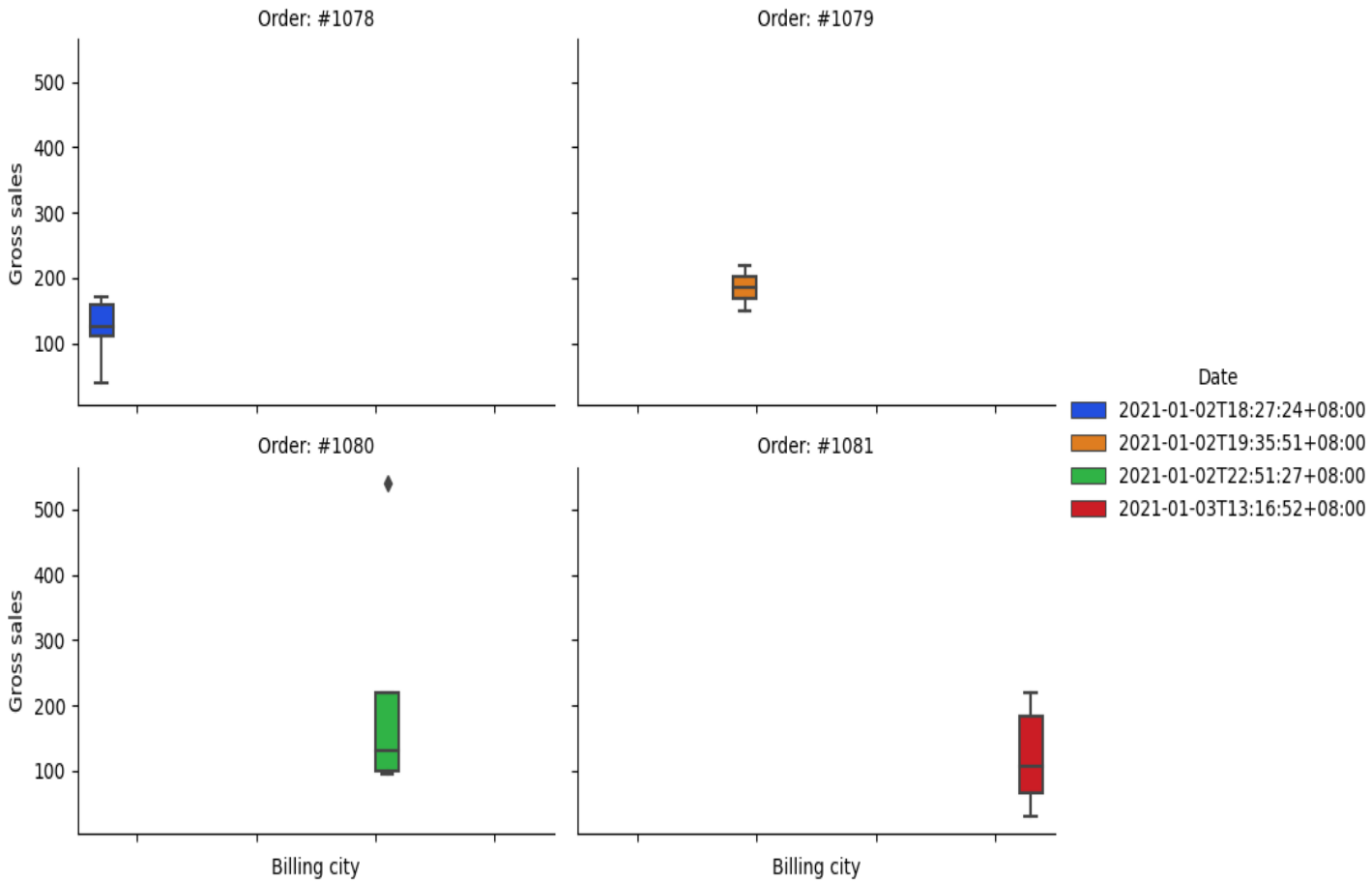
SCATTER PLOT B/W BILLING CITY,ORDER,TOTAL SALES

- This plot betwee Order,Billing city and Total slaes
- From this Plot Order = #1080 has the highest sales and Billing city is Metro Manila
- Lowest sales from San Juan City and Order = #1079



BOX PLOT B/W DATE,ORDER,GROSS SALES

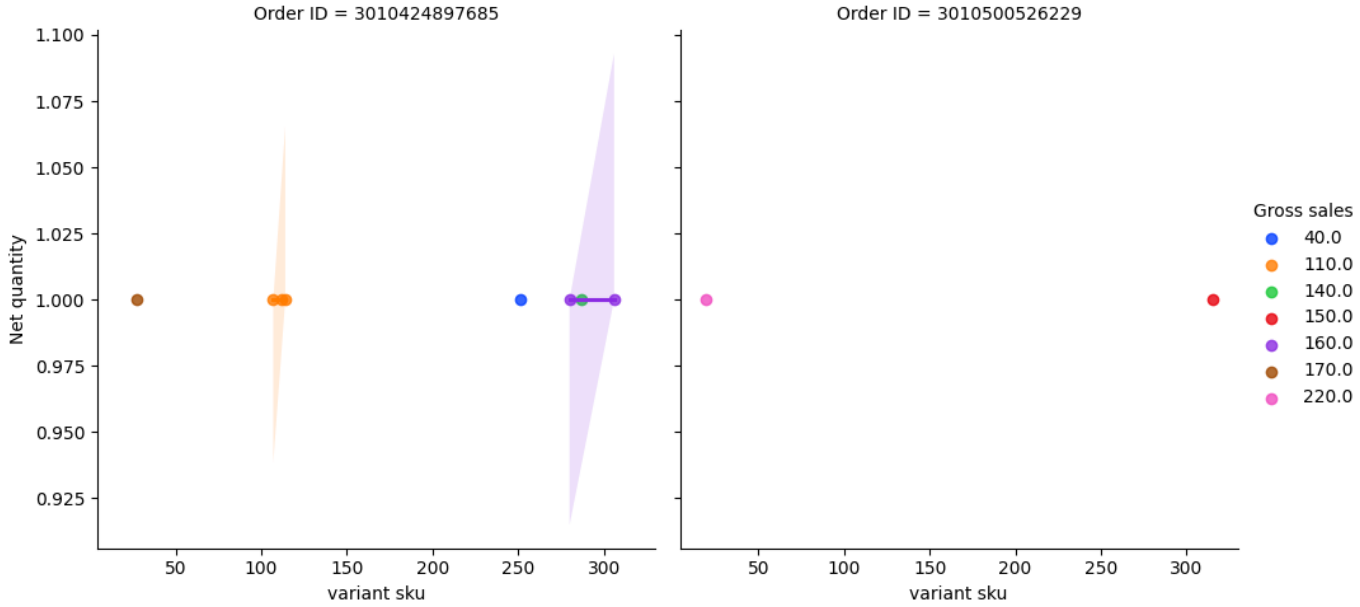
- This Box plot between billing city,Order ,Date and Gross sales
- On 2021-01-02 and Order #1080 has highest sales, On this date Order = #1079 has lowest sales



IMPLOT B/W DATE,ORDER,GROSS SALES

- This Implot between Variant Sku,Order ID and Net Quantity

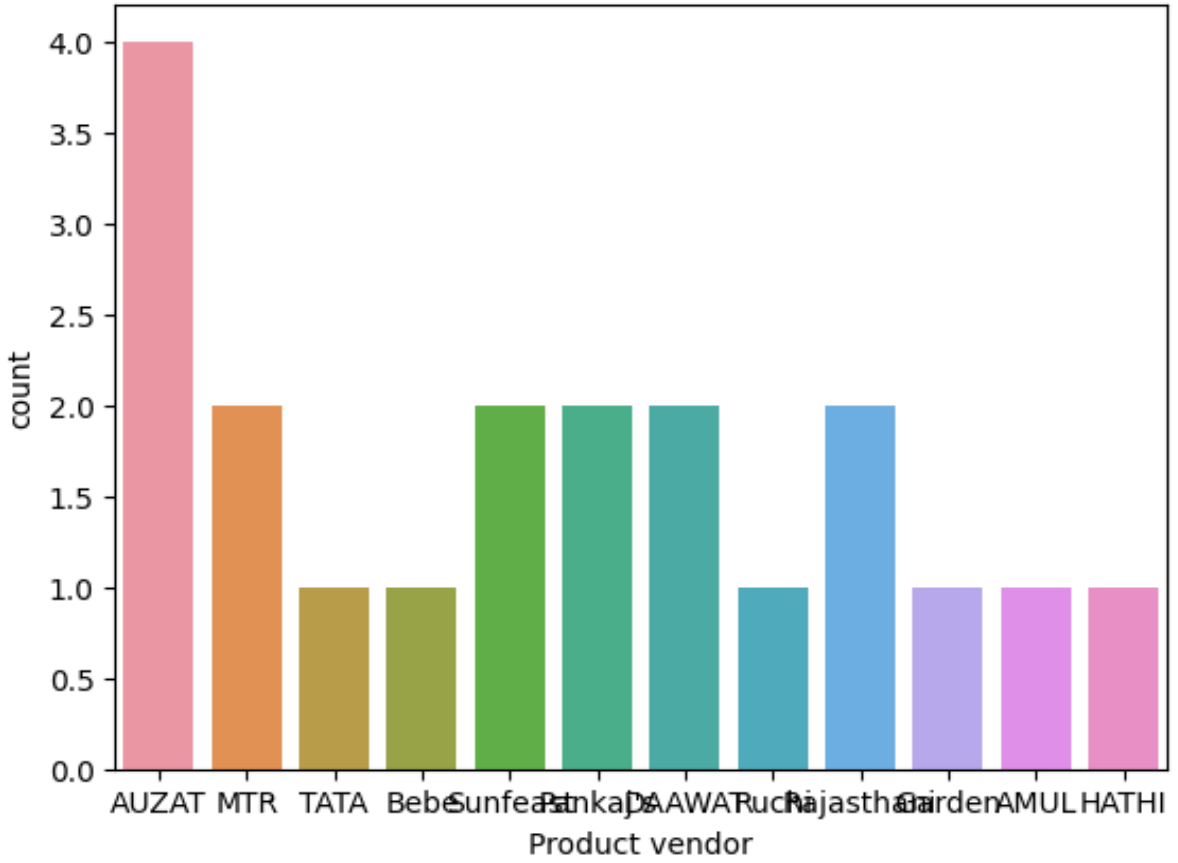
- Order ID = 3010424897685 has the highest Gross sales and Order ID = 3010500526229 has lowest Gross sales



HIGHEST PRODUCT VOLUMES

- This Bar shows highest Product volumes provided by the AUZAT vendor

- Second highest Product volumes Provided by MTR,Sunfest,DAAWAT



Conclusion

From this analysis

- AUZAT Company have the highest product volumes in sales
- TATA Company have the lowest product volumes in sales
- Metro Manila city is highest billing city
- San Juan city is lowest billing city

My Suggetions

From this analysis

- Focus on TATA Company sales
- My suggestion to increase TATA Company sales introduce all products along with the salt item as already competitors sell all items

Q&A

edureka!
LEARNING CENTER