**Power BI – Retail Sales Analysis Case Study**

**About the customer**

ABC is the customer who are operating in US for retail industry. They wanted to build a datawarehouse to forecast the sales by inline with the customer needs and in that way can improve their business among the competitive market world.

**Objective**:

Retail sources customer data from multiple sources and their transaction information from the

Accounting System. Requirement has come to the warehouse team to build the retail mart which needs the below information:

1. To Analyse key aspects to improve sale of retail stores and to know what customer buys and in which season.
2. To know about which product is needed by what type of customer and what are the key questions of retail business
3. To carry out the analysis of retail store in different state like how much is the revenue generation, amount of product sold in which month

**Source Files:**

* Global Super Store.xls
* Seasons.xlsx
* State Revenue.xlsx

**Problem Statement:**

1. Load the below source files available in the eLearning portal (under Input Files)

[**https://elearning.hclets.com/course/view.php?id=831**](https://elearning.hclets.com/course/view.php?id=831)

* Retail\_Data.xls

(Import only Orders and Returns data)

* Retail\_Region.csv

1. Transform the data as follows

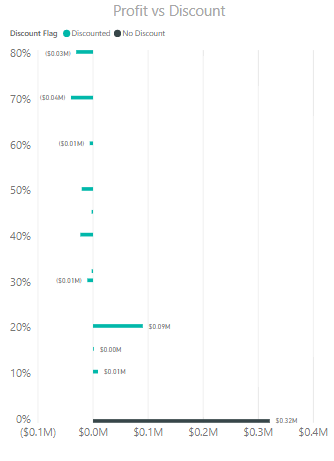
* Make the row which has logical business names as header
* Duplicate Orders should be removed
* Retain only the needed columns. Postal Code column is not used in reports
* Datatype should be proper as per the data. Datatype of date data should be date and decimal data should be decimal
* Replace NA data with Others in Ship Mode
* Create new field named Sub category from category
* Add a column named profit Category If Profit > 2000 then High, If it is less than 100 then it is low otherwise normal
* Join with region data to extract region value
* Only the newly joined table and Returns table should be visible in the modelling area
* Organize queries

1. Model the data as follows,

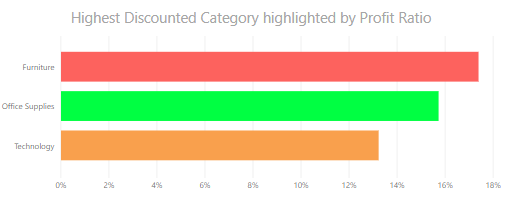
* Join Orders and Returns table
* Format date column in MM/DD/YYYY format
* Create year and Month Year Columns and order by calendar wise
* Add a new field, If there is discount then “Discounted” else it should be “No Discount”
* Discount multiplied by sales will give new field Discounted value
* Calculate no of products
* Discount should be displayed in % and other currency fields should be displayed in $
* Add synonym “item” to product and “no of units” to quantity
* Create region hierarchy with region, state and city

1. Visualize and answer the following business questions with appropriate titles, clear fonts and neat layout

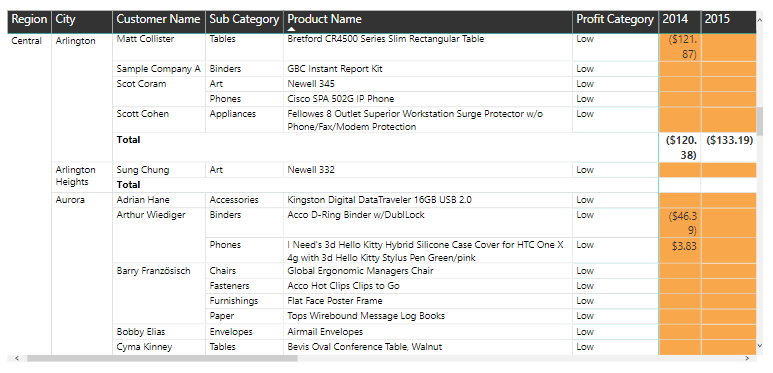
* Are discounted orders more or less profitable than non-discounted orders? Highlight category as discounted and non-discounted and showcase profit vs discount



* Which category got highest discount? Highlight the profit ratio in diverging colours with red being low profit, Orange medium and green being high.

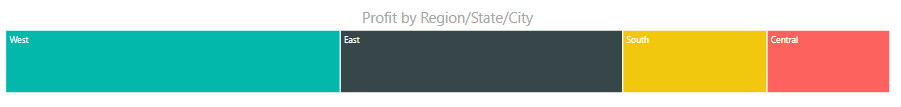


* Display detailed report with grouped region, city, customer name, sub category, product customer bought and profit category. For each year, display the profit in $. Highlight the background of the profit data where red being low, orange being medium and green being high (status colours).

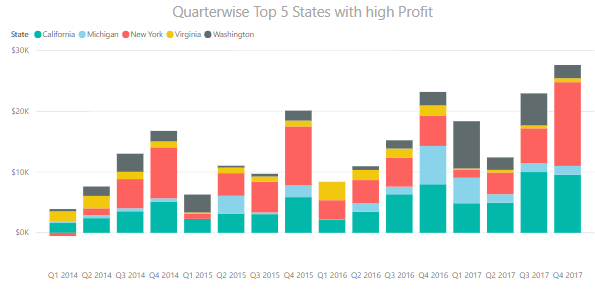


* Display the performance of different regions based on profit, going down further to state and city wise profit. Profitable and non-profitable regions should be highlighted by rectangle size.

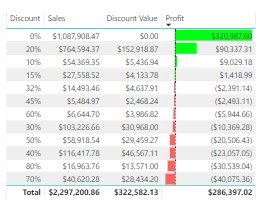
Also, when selecting a region, detailed report for that region should be displayed



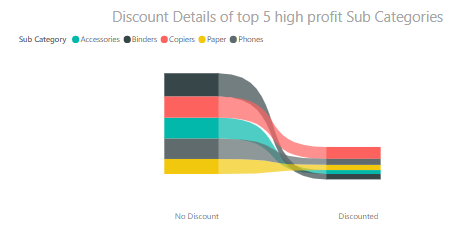
* Display quarterly report all years with quarters sorted as per calendar and display top 5 states which has high profit in different colours for each quarter



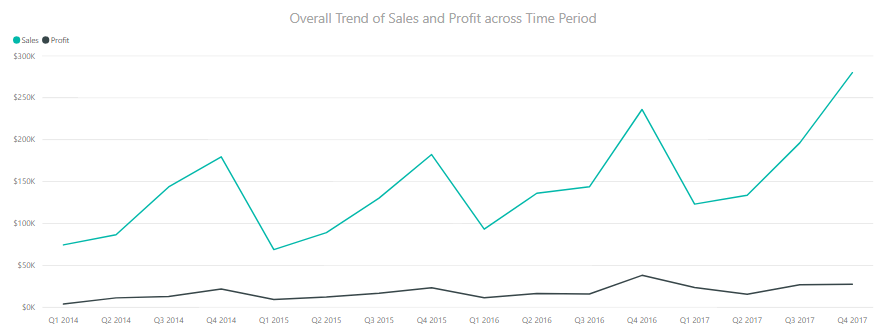
* Display net statistics – Amount discounted along with sales against each discount percentage. Highlight status colours as bar for profit data ordered by profit



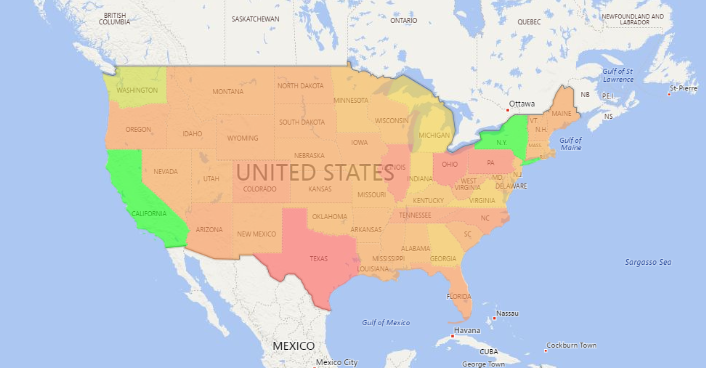
* Show top 5 high profit sub categories which should be categorized as discounted and non-discounted



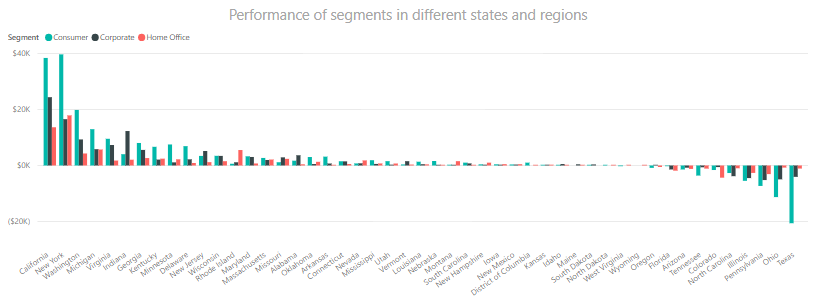
* Display the overall trend of sales and profit across time period across Year/Quarter-Year/Month-Year which is sorted as per calendar. Analyse and explain the increase between 2015 and 2016 and add that to report.



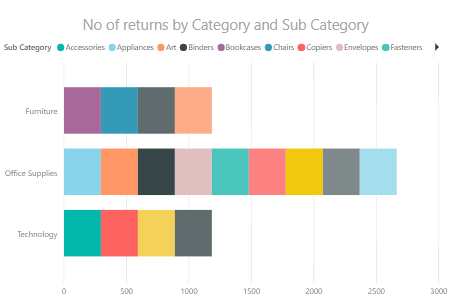
* Use appropriate visual and show the performance of each state with RAG status on states based on profit



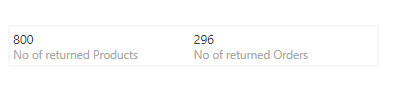
* Display the Performance of segments in different regions and states.



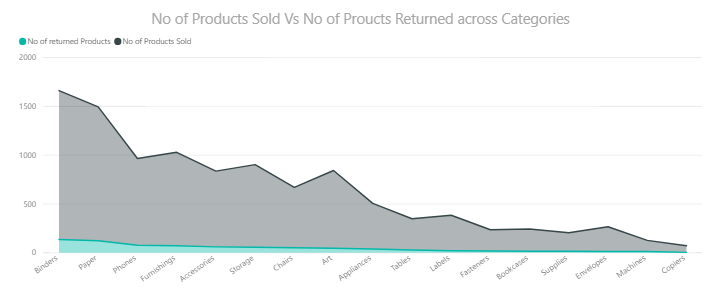
* Showcase categories and sub categories with number of products returned.



* How many products ordered are returned? Display along with no of orders returned.



* No of Products Sold Vs No of Products Returned across Categories further down to sub categories.



1. Create a dashboard titled as “Retail Sales” with neat font and alignment which includes the below,

* Include the below visuals
* Highest Discounted Category highlighted by Profit Ratio
* No of Products Sold Vs No of Proucts Returned across Categories
* Quarter wise Top 5 States with high Profit
* Net statistics
* Profit by Region/State/City
* Add a slicer on category
* Visual with “Highest Discounted Category highlighted by Profit Ratio” should not be changed on slicer selection and other visuals should be changed on slicer selection
* Add the below featured Q/A questions
* Display the Performance of segments in different regions and states
* Category and state wise profit displaying highest profit state in descending order
* In the QA question on category,
* Add sales data to the result
* Convert the visual to clustered column chart
* Change the title as “State & Category wise Profit and Sales”
* Display data for top 10 countries which gives high profit
* Add to the dashboard
* Share the dashboard to the email id ([Keerthiga.b@hcl.com](mailto:Keerthiga.b@hcl.com)) and make sure the recipient cannot share it with others
* Make the created dashboard as default dashboard
* In phone view, “No of Products Sold Vs No of Products Returned across Categories” should be available

**Support Questions:**

* Handle change in the column name of the source
* Handle change in the data type of the source column
* Handle change the source file path
* Handle change in DB password
* Handle error data
* Handle null columns
* Handle duplicate data