

**E-COMMERCE SALES**

**DATA ANALYSIS**

**-BY**

**Detailed Project Report**

**Adil Anwar**



E-COMM SALES ANALYSIS



**OBJECTIVE:**

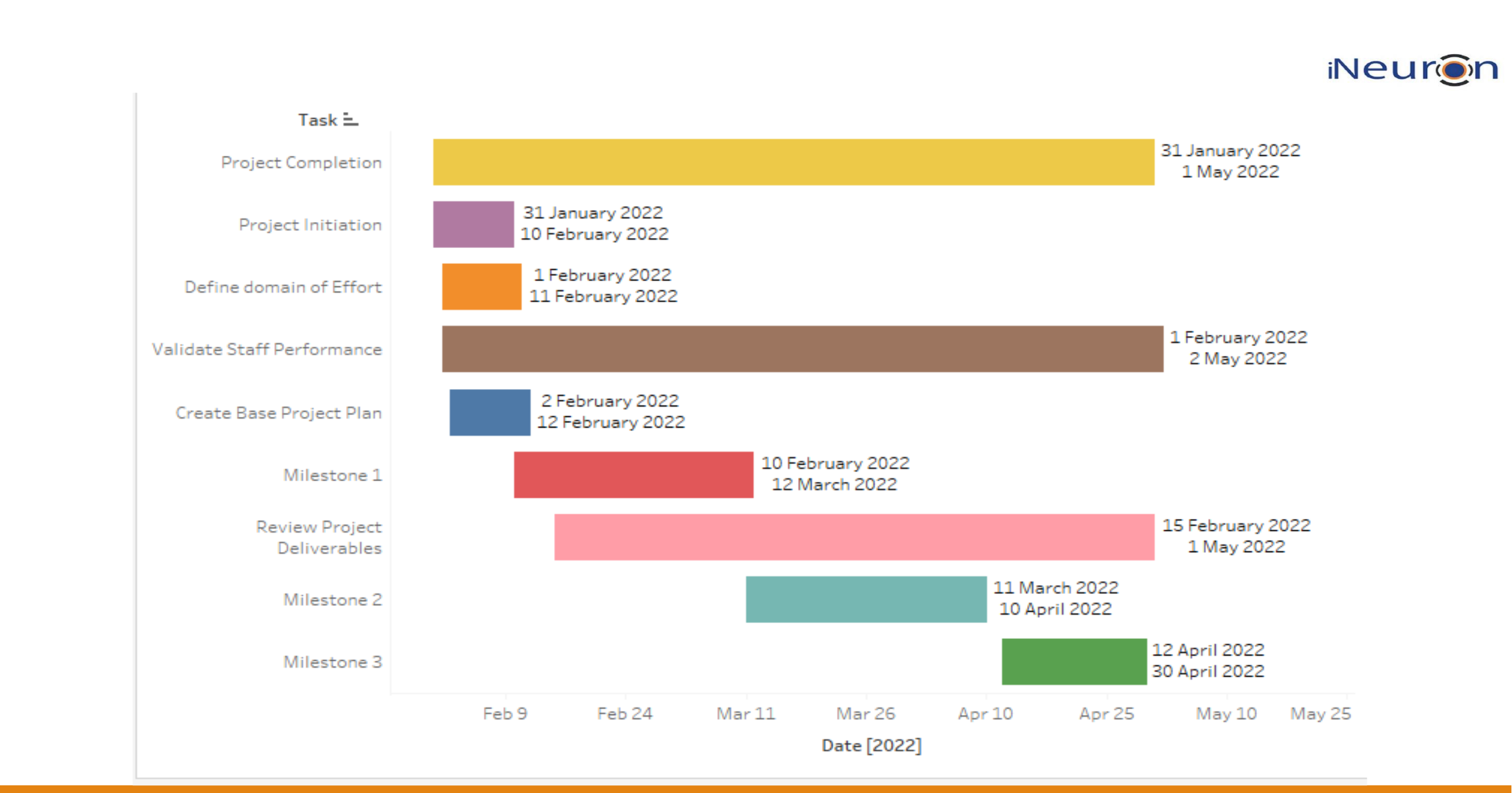
Finding key metrics and factors and showing the meaningful relationships between attributes that can help the business understand the insights better and take crucial decisions accordingly.

**BENEFITS:**

* Important Key Metrics will be identified
* Will give a better insight of the Sales data
* Will help the business take crucial decisions accordingly
* The chances of making wrong decisions will be reduced



E-COMM SALES ANALYSIS

**PROJECT TRACKER**



E-COMM SALES ANALYSIS

**DATA SHARING AGREEMENT:**

Sample file name : SALESDATA.xls

Length of date stamp (8 digits)

Length of time stamp (6 digits)

Number of Columns: Mentioned in LLD

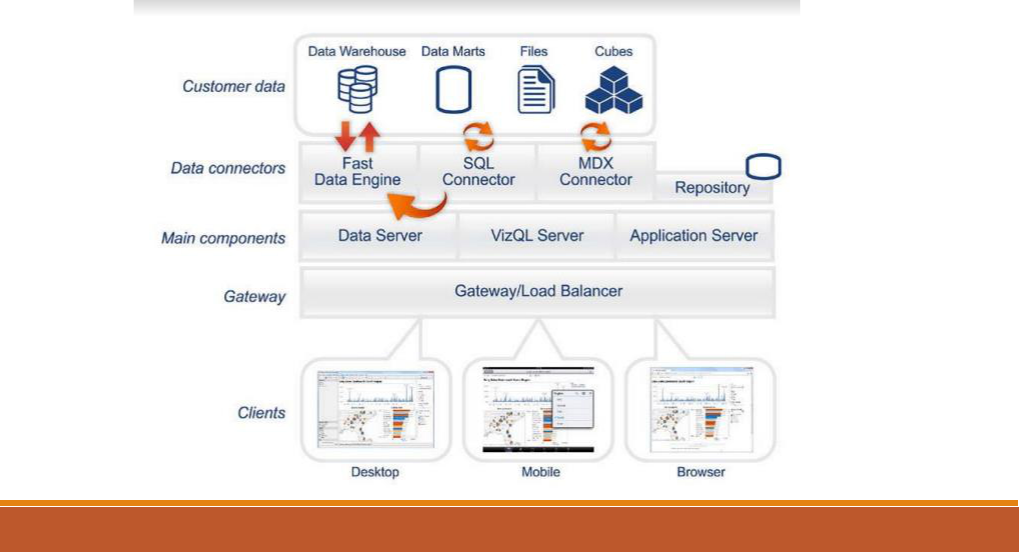
Column names: Mentioned in LLD

Column data type: Mentioned in Tableau file



E-COMM SALES ANALYSIS

**ARCHITECTURE**



E-COMM SALES ANALYSIS

**Data Validation and Data Transformation :**

**Name Validation** - Validation of files name as per the DSA. We have created a regex pattern for validation. After it checks for date format and time format if these requirements are satisfied, we move such files to "Good\_Data\_Folder" else "Bad\_Data\_Folder.“

**Number of Columns** – Validation of number of columns present in the files, and if it doesn't match then the file is moved to "Bad\_Data\_Folder.“

**Name of Columns** - The name of the columns is validated and should be the same as given in the schema file. If not, then the file is moved to "Bad\_Data\_Folder".

**Data type of columns** - The data type of columns is given in the schema file. It is validated when we insert the files into Database. If the datatype is wrong, then the file is moved to "Bad\_Data\_Folder".

**Null values in columns** - If any of the columns in a file have all the values as NULL or missing, we discard such a file and move it to "Bad\_Data\_Folder".



E-COMM SALES ANALYSIS



**Data Insertion in Database:**

**Table creation :-**

Table name “Sales\_Cleaned" is created in the database for inserting the files. If the table is already present then new files are inserted in the same table.

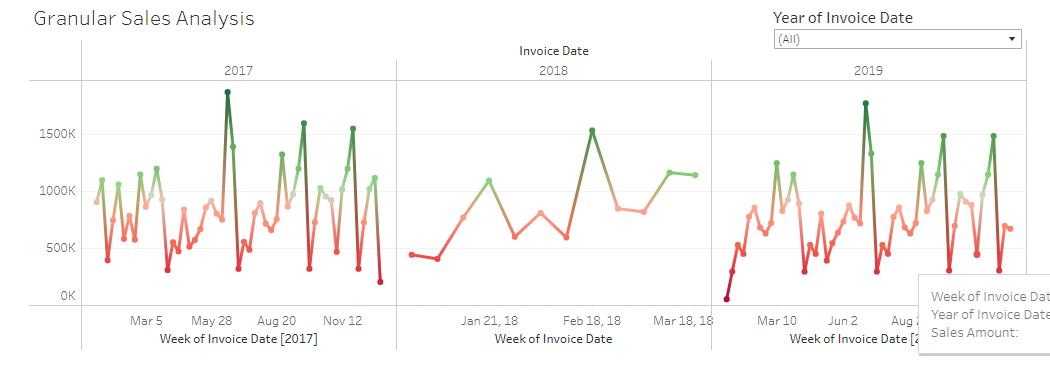
**Insertion of files in the table :–**

All the files in the "Good\_Data\_Folder" are inserted in the above-created table. If any file has invalid data type in any of the columns, the file is not loaded in the table



E-COMM SALES ANALYSIS

**DATA VISUALIZATION – SALES ANALYSIS**



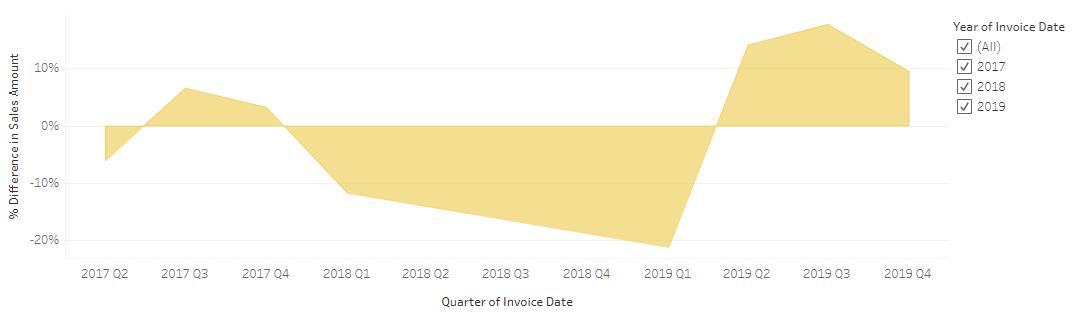
A dynamic graph that can be used to

analyse the Yearwise -> monthwise -> Weekwise sales with all the high-low values

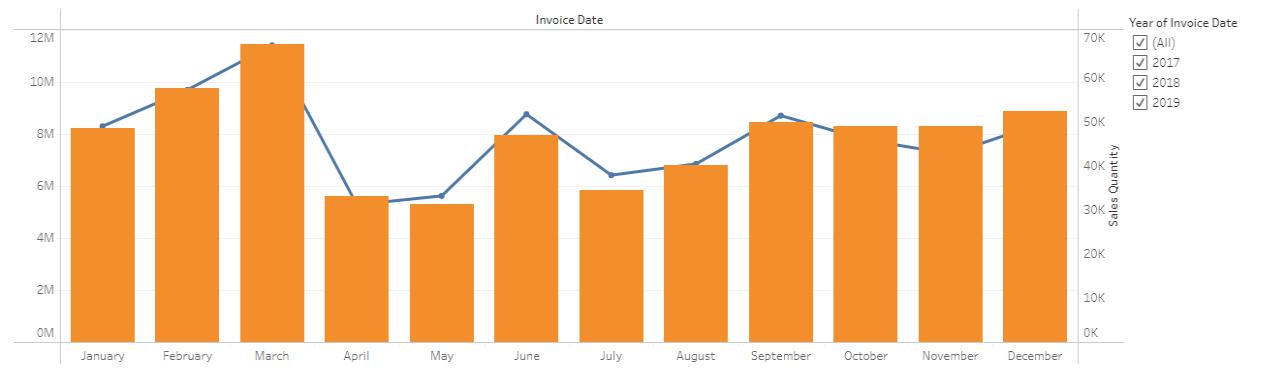


E-COMM SALES ANALYSIS

**DATA VISUALIZATION – SALES ANALYSIS**



The area under the graph calculates the % difference (15.3% ) in sales amount across the table. The data can be filtered by the year of sales. Q3 2019 showed the highest % difference than the other quarters.



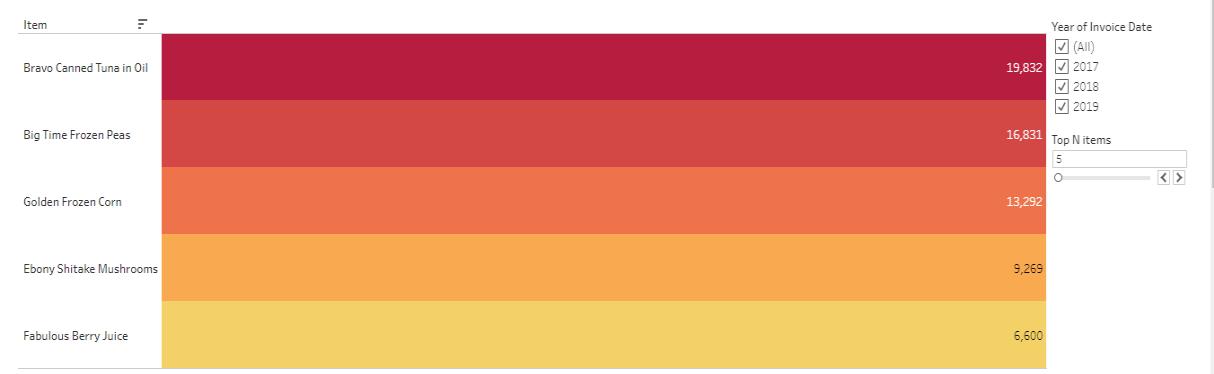
A Pareto Chart that shows the quantities sold every month. The data can be filtered by the year of sales



E-COMM SALES ANALYSIS

**DATA VISUALIZATION:**

**PRODUCTS WITH HIGHEST MARGIN**



As can be seen, of all the times, the products - Bravo Canned Tuna oil, Bg Time Frozen Peas etc are the products with Highest margins



E-COMM SALES ANALYSIS

**DATA VISUALIZATION**

**BEST SELLING PRODUCTS**



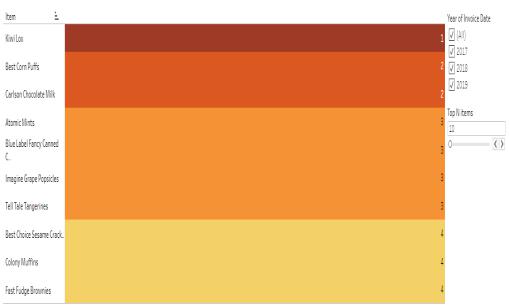
As can be seen in the table, Top 15 best selling products have been highlighted. A filter can be applied to view the top selling products yearwise as well.



E-COMM SALES ANALYSIS

**DATA VISUALIZATION**

**LEAST SOLD PRODUCTS**

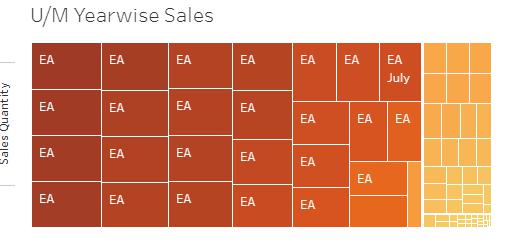


As can be seen from the table, the least sold products have been shown. The year can be filtered to view the data for every year.



E-COMM SALES ANALYSIS

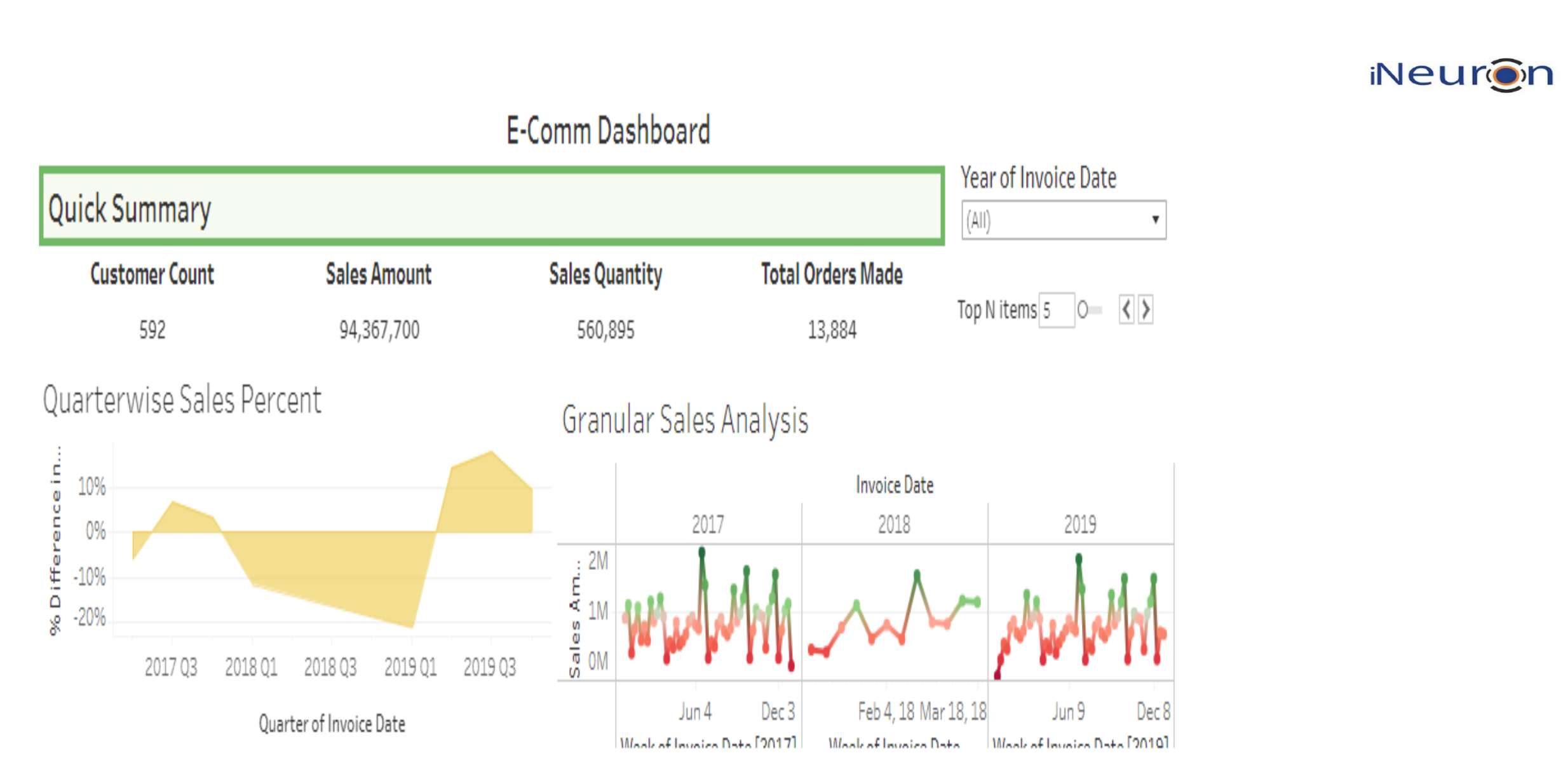
**DATA VISUALIZATION**



As can be seen from the treemap, the area wise sales distribution has been analysed. EA occupies maximum customers.



E-COMM SALES ANALYSIS



• The quick summary

section shows the

important KPIs like -

Customer Count,

Sales Amt, Sales

Quantity & Total

Orders Made based

on the year

selected.

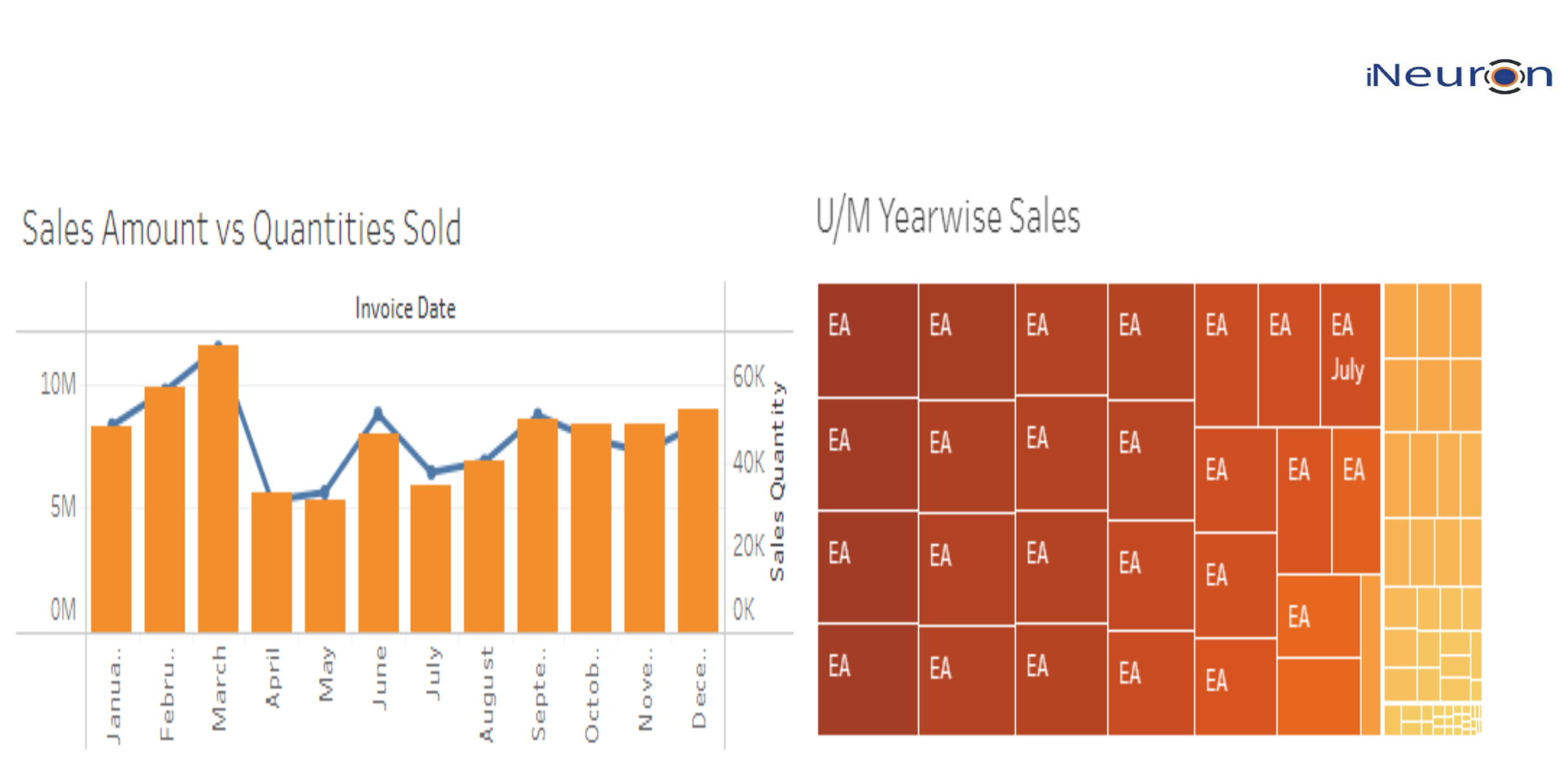
• The entire dashboard gives a complete sales information

Dashboard Link:

**Tableau Public Link**



E-COMM SALES ANALYSIS

**E-COMM DASHBOARD**

Dashboard Link:

**Tableau Public Link**



E-COMM SALES ANALYSIS

**E-COMM DASHBOARD**



Dashboard Link:

**Tableau Public Link**



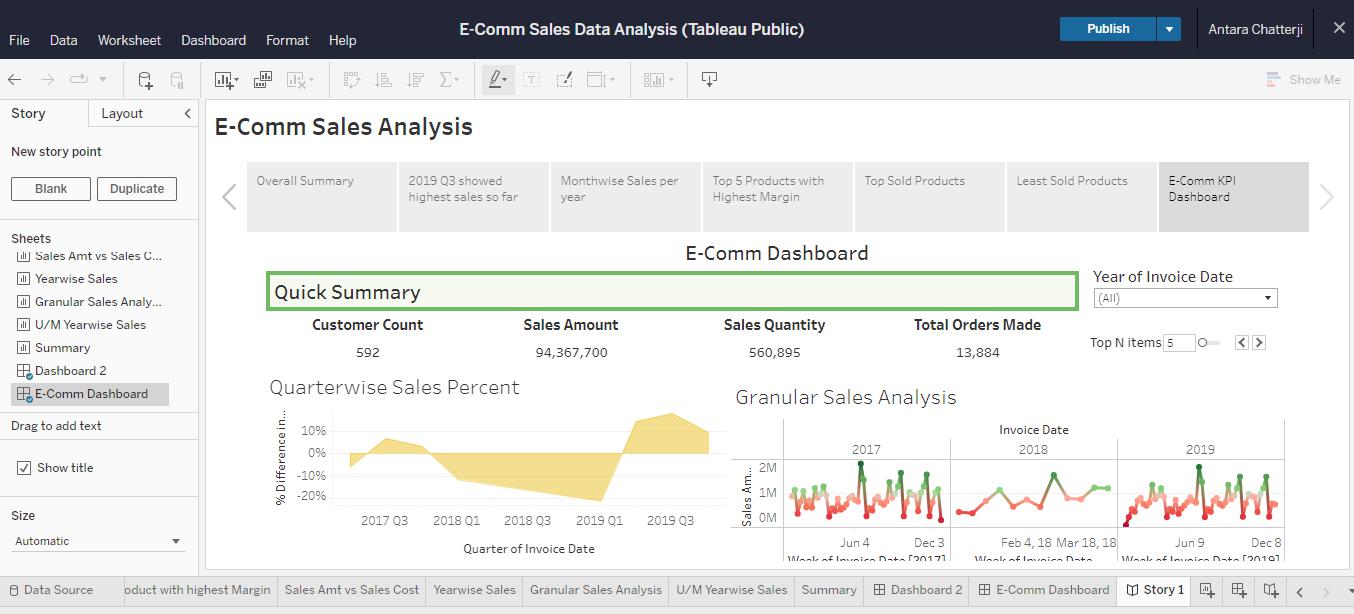
E-COMM SALES ANALYSIS

|  |  |  |  |
| --- | --- | --- | --- |
|  | **UNIT TESTING** | |  |
|  |  |  |  |
| **TEST CASE DESCRIPTION** |  | **EXPECTED RESULTS** |  |
|  |  |  |  |
| **Year of Invoice Dropdown** |  | When clicked a dropdown should occur which |  |
|  | has all the years listed |  |
|  |  |  |
| **Top N Items** |  | The parameter should return the specified |  |
|  | number of items |  |
|  |  |  |
| **Granular Sales Analysis** |  | This is one of the most important graph to |  |
|  | perform the LOD analysis of the Sales data |  |
|  |  |  |
|  |  |  |  |
| **Quarter Graph** |  | The area under the graph shows the sales for |  |
|  | that Quarter |  |
|  |  |  |
| **Sales Amount vs Sales Quantity** |  | An analysis of Quantities sold and the sales |  |
|  | earned |  |
|  |  |  |
| **Products with Highest Margins** |  | Parameter lists the product names with highest |  |
|  | margins |  |
|  |  |  |
| **Top Sold Products** |  | Parameter lists all the products that are top |  |
|  | selling overall (and/or) yearwise |  |
|  |  |  |
| **Least selling products** |  | Parameter lists all the products that are least |  |
|  | selling overall (and/or) yearwise |  |
|  |  |  |
| **E-comm Dashboard** |  | Contains required KPIs to understand Sales |  |
|  | distribution throughout 3 years |  |
|  |  |  |
|  |  |  |  |



E-COMM SALES ANALYSIS

**DASHBOARD DEPLOYMENT**



Dashboard Link:

**Tableau Public Link**



E-COMM SALES ANALYSIS

**FINAL RECOMMENDATIONS**

* The year 2018 recorded the sales only for the first quarter, the next 3 quarters received 0 orders. (Refer Slide 9)
* Q3 2019 recorded the highest sales of 15.3% of the total sales so far. The sales strategies should be carefully monitored and used for future campaigns as well. (Refer Slide 9)
* The U/M area, EA has very high sales as compared to other parts. More campaigns can be targeted in these areas to increase sales. (Refer slide 13)
* The U/M area, PR recorded the least sales so far. The sales and marketing team should design campaigns and discounts accordingly to increase sales in such areas. (Refer slide 13)
* The product with highest margins can be targeted to be sold more in the market (Refer slide 10)
* The most sold products should be stored in sufficient amount in the inventories in order to avoid exhaustion of these items. (Refer slide 11)
* The least selling products can be offered with reasonable discounts to increase sales. (Refer slide 12)



E-COMM SALES ANALYSIS