Report on Design Process

Contents

1. Introduction:

The regional sales data of a company across all regions of the United States of America and years **2018 - 2020** was analyzed in this project to establish correlations and review performance of the company over the years. This dashboard illustrates the information of the company's revenue generated across different attributes over time. It also gives an overview of how it performed against different categories such as states, customers, goods with revenue as measure. It also reviews performance of its sales teams, its stores, and its warehouses on based of their revenue generation. This dashboard is most useful for the executive management team and administration team of the company to get a categorical review of the company's performance and to make executive decisions based on the information.

2. Data:

The data file was obtained from website data. world, and it gave a very compressive views of orders and prices of different products across all the USA. The data was divided into multiple sheets and each sheet contained information about sales and orders, customers, store location, products sold and regions of sales and about the sales teams/ members. The data set contained almost 8000 entries of order placed across different states and locations such as warehouses and the price break down of the units, and number of units ordered and the shipment details and respective stores and customers and salesperson ID's. subsequent sheets contained the customer names and store location details, sales team details and region details.

3. Data preparation:

Since the data is spread out in different sheets and I wanted to use the data across all sheets to create the dashboard, I have used Tableau prep to clean, merge and sort data and tableau desktop to visualize the data.

First, I imported the data sheets into tableau Prep and then started joining the sheets to create the required flow of the data. I have used joints and mergers function to combine data on a common column. For example, I have used the left joint to combine sales/orders sheet with product sheet using the common column product ID. The final data consists of details of 7992 orders defined over 26 categorical fields.

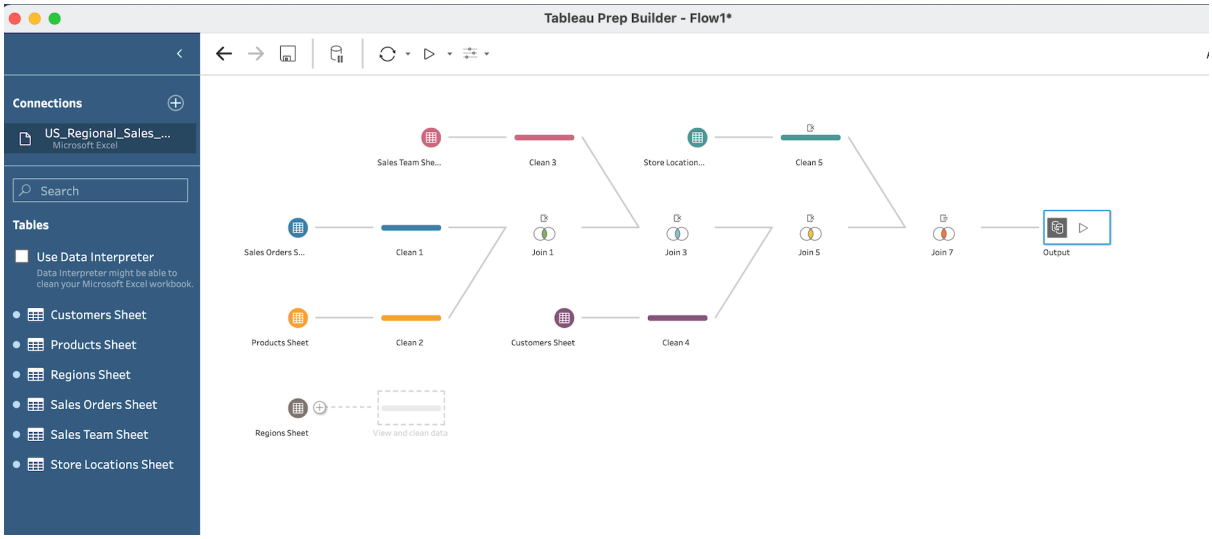

Fig 1: screenshot of data flow prep

4. Methodology:

This report, will explain the design choice of the dashboard created for the US sales regional dataset. The dashboard was designed using on the concept Munzner's four nested level model. The nested model has different threats to validity in each level. The project is validated accordingly using the four nested levels model.

The main purpose of this dashboard is to establish review company performance across the USA. To achieve that we will establish the relation between the following Revenue per state

- Revenue across time
- Revenue per state
- Revenue per region
- Revenue per Product type
- Revenue per sales team
- Revenue per customer

5. Design Approach:

The project is design using the concept Munzner's four nested level model: (Munzner, 2009)

Level 1: Domain Situation:

This top level block is used to observe target users and understand their specific interest. Here in this study our targeted users are the executive management of the company. We are using the specific data related to the company to give an overview of sales and revenue to the executive management of a company so they can make like where to invest in next financial cycle or decide if and how to boost sale in a particular area. This dashboard is a human centred design because the whole dashboard is refined with a specific target audience in mind.

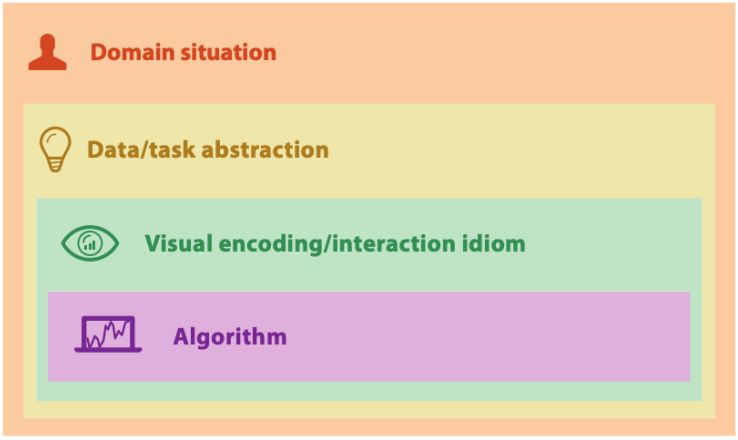

Fig2: The four nested levels of vis design (Munzner, 2009)

Level 2:Data and task abstraction

The abstraction level contain the questions what and why i.e., what is the data showing(abstraction of data) and whys is the user looking at it(abstraction of task). I identified the why first. The user wanted to know how the company is performing in the US. And we have data related to orders, customers, and stores to answer this question. I have analysed, searched and queried the dataset to put together the required data to answer the users questions. I have calculated 'revenue' from the columns 'unit cost' and 'unit price', 'discount' and 'order quantity'. I rated the customers, ordered items, sales teams, regions and stores which generated highest revenue.

Level 3 : Visual Encoding and Interaction Idiom

This step involves creating and manipulating visualizations based on abstraction to get the required output which answers the previous level question. Here I used a Dashboard containing Map, bar graph, donut graph and bubble plot to visually represent the revenue data needed by the user.

Level 4: Algorithm Design

In this step I create an algorithm for visualisations which are encoded and interactive. All the graph, maps and plots in the dashboard are interactive and tool tip is add to all the fields so if one hovers over the dashboards visualization a detailed description of the field pops up. And a drop named months gives the revenue over a specific time over the 3 years and dropdown of product gives revenue generated by a particular product.

6. Munzner's threats and four levels of validation of the project:

Domain Validation:

At domain level, I have a clear idea of what the executive management team is looking for from the data and the data I used is spread over so many attributes that it covers most of the questions related to how revenue is affected over time and over various categories.

Abstraction validation:

The threat at this level is the model fails to solve / answer the question of the targeted users. The only method to validate is to have the model tested by target users. (Munzner, 2009). We have collected good and valid data and enough relevant data to answers investors questions.

Encoding and Interaction idiom validation:

The threat at this level is that the chosen design is not communicating effectively to the target user regarding their problem. We have how ever used appropriate maps and plots to make the visualization

cater to the user needs and all the graphs and plots used are interactive and have appropriate filters applied to make using them very dynamic.

Algorithm validation:

The threat at this level is that the model my have computational issues, like lacking of enough memory.

The dashboard I simulated uses Tableau and works perfectly well to give a detailed analysis about the revenue of the company. The dashboard is easy to use, simple to understand and runs smoothly without any technical glitches.

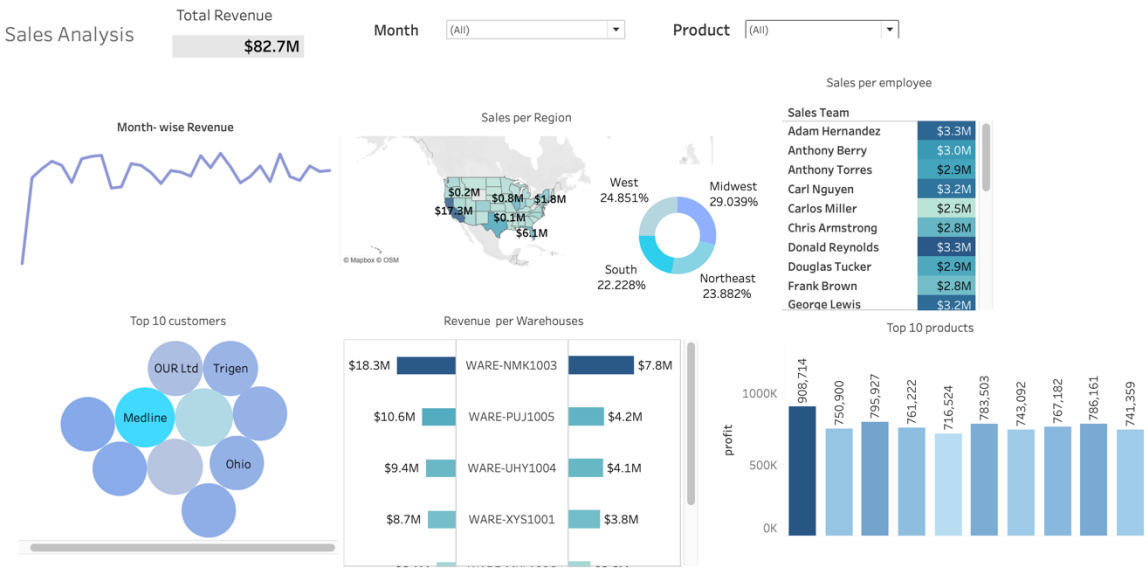

Fig 3: Screenshot of Dashboard

7. Evaluation

This dashboard is a sales analysis dashboard that is meant to be used by the particular company's executive management members. I use and validated Munzner's four nested level model for design of the project. It is an interactive dashboard and is a one stop solution for understanding company's revenue performance over the years 2018 – 2020 in the Us regional markets. The dashboard compares revenue between various attributes such as customers, time , stores, regions, product type. Such detailed review will give the target users a comprehensive idea of how their company is performing and how it can improve by focusing on a particular approach, like focusing on customer or a region, etc.

References

Munzner, T. (2009). A Nested Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6), 921–928. https://doi.org/10.1109/tvcg.2009.111

Step by Step resource guide to learn Tableau. (2019, July 26). Analytics Vidhya. https://www.analyticsvidhya.com/learning-paths-data-science-business-analytics-business-intelligence-big-data/tableau-learning-path/

US Regional Sales Data - dataset by dataman-udit. (n.d.). Data.world. https://data.world/dataman-udit/us-regional-sales-data