MASTER'S DEGREE PROGRAM IN DATA SCIENCE AND BUSINESS ANALYTICS

Data Visualization

By

Shanjida Roman - 20221395

Jaime Duarte - 20220675

Yousef Ebrahimi - 20221382

João Guilherme Mateus Gameiro - 20221364

April 2023

Dashboard link:
https://public.tableau.com/views/DefectsAnalysisDashboardv2/Dashboard

Index

Introduction

The purpose of this report is to outline the use of visualization concepts and techniques in transforming data into a meaningful interactive visualization, as required by the objective of the project assigned to our group. Our project was implemented through data exploration with Python and Plotly, and visualization/story using Tableau software. We made sure to frame it using the data visualization techniques and concepts explored during the course. Our goal was to create a visualization that matched our professional/academic interests while meeting the project requirements. The users of this dashboard get different types of interactivities, such as filtering by regions/states, time, and sites, as well as parameters for the previous and present years.

Data

The inspiration for this data visualization project was the need to address a critical issue faced by OneStop Solutions, a soft services provider company in the UK, whose defects KPI was not meeting the contract standard resulting in substantial penalties. The senior operations management team and the head of the contract initiated this project to analyze all defects data from 2022, gain insights into underlying causes, develop effective strategies, and ultimately improve the company's KPI and reputation. Utilizing data visualization techniques and tools such as Tableau, the team was able to transform complex data into meaningful insights and take appropriate measures to improve the company's service level agreement (SLA).

Our dataset is based on this company, and it has 1755 records during 2022. The description of each variable followed by:

- Work Type - Type of defects reported in the system. Three kinds of defects 1, complaint 2. Recall 3. Service failure
- WO Number - Request number. Unique ID of defects.
- Raised Date - Date and time when defects were reported.
- Completion Target - Completion target date according to SLA.
- Completion Status -Current completion status of the defect

- WO Status - Status of the defect in terms of priority
- Description - Description of the defect
- Building - Building name where the defect occurred
- County - County name where the defect occurred.
- Post - the first part of postal code to use in mapping.
- Operational Business Unit - Business unit information
- Region- Region name
- Country - Country name
- Short Description - Short description of metadata
- WO Complete - Completion date
- Penalty Rate - Penalty rate per defect

Business Questions

To kick off this analysis the team established important questions to be answered, to get insights precisely and clearly:

1) How many penalties are we paying?
2) What is the total number by each criterion?
3) Which area of business going wrong?
4) Which are the worst-performing areas?
5) Why are these sites performing low?
6) Why complaints numbers are so high?
7) What can be done geographically?
8) What can be done operationally?
9) What can be done to improve communication?
10) What would be the trend if the business grows by 6.5% next year (2023)?
11) Is this a scalable tool?
12) Can it be used to analyze future data?

These questions are answered in the following chapters.

Methodology

To achieve our objective, we began by selecting a dataset that aligned with our interests and was suitable for visualization. We used Python programming language and Tableau to create interactive visualizations, incorporating data visualization techniques and concepts explored during the course. The best part of our Tableau dashboard was the heat map of the geo-location of each region in the UK, which showed the distribution of different types of defects reported in the system (complaints, recalls, and service failures). This interactive heat map allowed users to change the view and focus on specific regions, providing a deeper understanding of the distribution of defects in each region.

We made sure to consider the principles of effective data visualization, including simplicity, clarity, and accuracy. We also incorporated interactivity, which allowed users to interact with the visualizations and gain insights into the data. We paid close attention to the

selection of colours, typography, and layout, as these elements can affect the user's perception and understanding of the visualization.

Results

To achieve our objective, we selected a dataset that contained information about the cost of operations for different buildings (sites) in various regions. We used Plotly with Tableau software to create interactive visualizations, incorporating data visualization techniques and concepts explored during the course. We created a dropdown menu that allowed users to select individual buildings (sites) to display their respective costs. This feature allowed for a more detailed analysis of the data and provided a clearer understanding of the cost breakdown for each building.

To visualize the regional cost distribution, we used a pie chart. This chart enabled users to quickly identify the regions with higher costs and compare them to regions with lower costs.

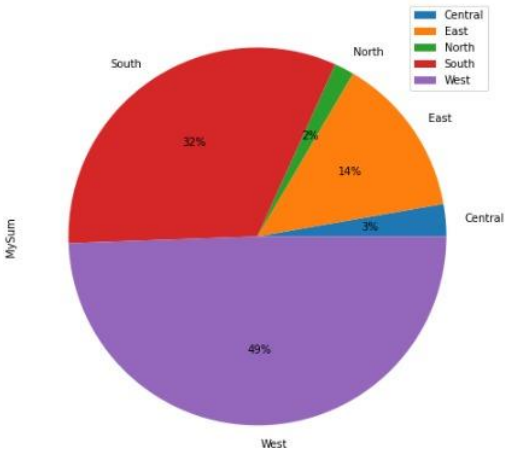

Figure 1: Total penalty by region from plotly visualisation

We also used a bar chart to compare the cost of each operation unit, which provided insight into the most expensive and least expensive units. This visualization allowed users to identify potential areas for cost optimization. To review the cost of penalties for each month during the year, we presented a line chart. This chart enabled users to track the cost of penalties over time and identify trends or patterns. This visualization was useful for identifying months with the highest penalties and finding ways to reduce or avoid these costs in the future.

Overall, our interactive visualization allowed for a deeper understanding of the cost data and provided insight into potential areas for cost optimization. The use of different chart types and interactivity enhanced the user's ability to explore and analyze the data.

Conclusion

Through our data visualization project, we were able to gain insights into the distribution of defects and penalties in the UK and identify patterns and trends in the data. One key finding was that the north and east regions of the UK had a higher rate of penalties per defect compared to other regions. This suggests that there may be underlying factors contributing to this trend, such as differences in regulations or enforcement.

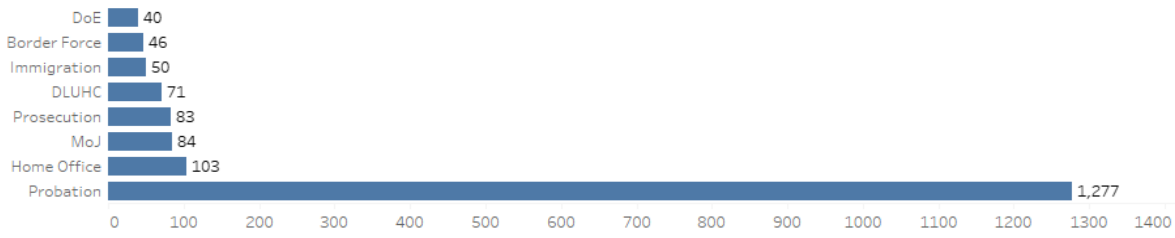

Figure 2: The number of defects based on the Operational Business Unit

We also discovered that the Probation in Operational Business Unit part had the highest number of defects and penalties, followed by the Home Office and Moj. This highlights the need for further investigation and improvements in these areas to reduce the number of defects and associated penalties. Also, we found that the number of defects decreased steadily throughout 2022, from 246 in January to 98 in December.

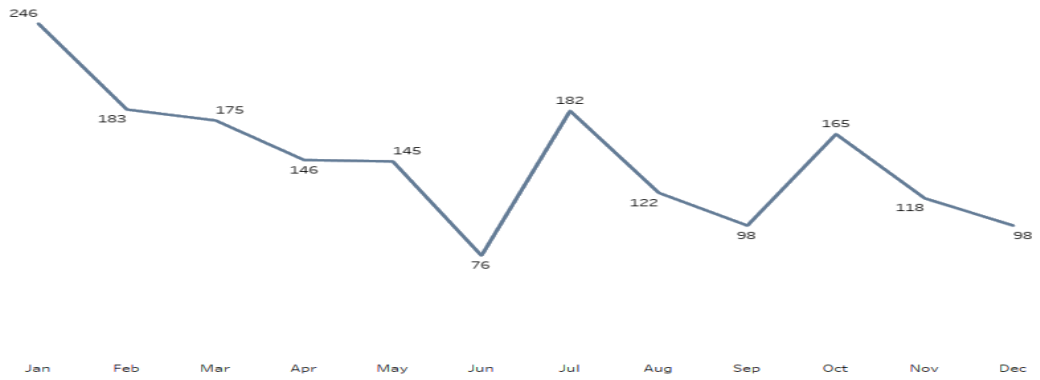

Figure 3: The number of defects during 2022

Finally, we identified a region that has a high number of defect records, indicating a need for increased attention and remediation efforts in this area. The total amount of the penalties for the government was £762,744 in 2022 with 1,754 records. Overall, our data visualization project provided valuable insights into the distribution of defects and penalties in the UK and highlighted areas for improvement and attention. By utilizing effective data visualization techniques and tools such as Tableau, we were able to transform complex data into meaningful insights and support informed decision-making.

Figure 4: The number of defects on the UK map based on the regions

References

Munzner, T. (2014). Visualization Analysis & Design. Visualization Analysis and Design (pp. 150). CRC Press. https://doi.org/10.1201/b17511

Munzner, T. (2014). Visualization Analysis & Design. Visualization Analysis and Design (pp. 168). CRC Press. https://doi.org/10.1201/b17511

Plotly, Low-code data app development (no date) Plotly. Available at: https://plotly.com/ (Accessed: March 22, 2023).

It supports and solutions, security, VoIP systems, MS 365 (no date) Onestop IT. Available at: https://onestopit.com/ (Accessed: March 22, 2023).

Create a Dashboard (no date) Online Help. Available at: https://help.tableau.com/current/pro/desktop/en-us/dashboards_create.htm (Accessed: March 19, 2023).