Name :

Student ID:

Table of Contents

**Table of Figures2**

**Data Choice3**

**Background Information3**

**Data Description4**

**Data Processing7**

**Data Mining8**

**Discussion of Findings11**

**References11**

**Appendices12**

Table of Figures

[Figure 1: Top cities based on numbers 4](#_Toc104808282)

[Figure 2 : Bottom cities based on numbers 5](#_Toc104808283)

[Figure 3 : Every Brand’s Transaction over a month 5](#_Toc104808284)

[Figure 4 : Top 5 classes based on number of transactions 6](#_Toc104808285)

[Figure 5 : Bottom 5 classes based on number of transactions 6](#_Toc104808286)

[Figure 6 : Prominent brand in terms of frequency in many departments in various nations 7](#_Toc104808287)

[Figure 7 : Month-by-Month transaction trend for all nations 7](#_Toc104808288)

[Figure 8 : which brands of goods Americans prefer to buy at GRAND 8](#_Toc104808289)

[Figure 9 : In Asia, each brand's market share fluctuated 8](file:///C:\Users\vbara\Downloads\Name.docx#_Toc104808290)

[Figure 10 : data trends at regular intervals 9](file:///C:\Users\vbara\Downloads\Name.docx#_Toc104808291)

[Figure 11 : comparisons across different categories 9](#_Toc104808292)

[Figure 12 : a set of data series: connected data points 10](file:///C:\Users\vbara\Downloads\Name.docx#_Toc104808293)

[Figure 13 : Clustering produces a group of data objects 10](file:///C:\Users\vbara\Downloads\Name.docx#_Toc104808294)

1. Data Choice

The data set RETAILDEMO 2 was chosen for this analytical report (Mike, 2021). There are a total of fifty seven column along with 2.3 million rows in this data set. The amount of data is enormous, and analyzing and filtering it with a normal relational-database is tough. In light of this, SAS Viya will be used for analyzing and visualizing in this report. We can derive some measures from either the provided data set, and we can further filter and sort the primary cluster data based on these measures. The primary cluster can be filtered, re-divided, and ordered based on a single measure in the real visualization study procedure. The primary cluster can also be joined, and first dataset could be taken as a foundation for categorizing data from another set. These activities can assist us in establishing the visual analysis that is required. We can generate required analytical reports naturally and rapidly by mixing the manifestations of numerous charts. Below following are the report's principal findings:

1. The firm should boost its involvement in large cities while decreasing its Investment in others.

2. Research how or when to prevent the sales decrease of other brands by focusing to the brand Maple, which total sales is increasing against the trend.

3. High-volume categories, including such bath body, should raise buy quantity to boost the amount of transactions while lowering the buy quantity of the segment with the greatest sales volume.

4. The Pines brand is the most popular in practically every country.

5. Each country's sales volume fluctuates over time.

6. The most popular way for Americans to purchase Pine brand products is through the GRAND chain of stores.

7. In Asia, Oak's market share has been dropping, while other brands' market share has been increasing.

1. Background Information

A data set (also known as a dataset) is a collection of data, usually in the form of a table. Each column corresponds to a different variable. Every corresponding to a query in the data set of a member. This lists the values for each variable as a random number, such as an object's length and weight. A data sheet is the name given to each value. The data set's data may have one or more members, depending on the number of rows. The collection and analysis of measurement data is used in a lot of scientific research. A fascinating trend may be observed here: data sets (often in conjunction using modeling and factors) are becoming increasingly essential. (2006, Dekker) An analytical report is a sort of business report that analyses and evaluates a business strategy or process using qualitative and quantitative firm data while encouraging workers to create info driven decisions based on facts and analytics and mined data (Greiner, 2011). "What would be an analysis report?" we posed the question. As a result, now is the time to think about the advantages of document analytics for business growth. Business intelligence is accessed through analysis reports (BI). And, as your company evolves smarter, you will gain an advantage over the competition. You will be able to make educated and accurate judgments by using analytical reports, as well as solve problems and react to new in a careful way.

1. Data Description
   1. Top 5 and Bottom 5 cities based on number of customers

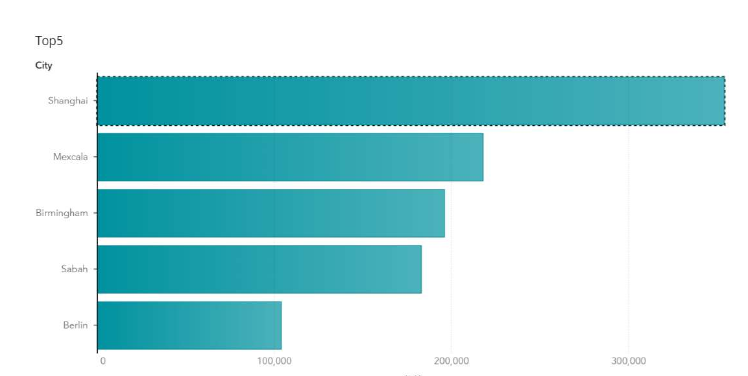


Figure 1: Top cities based on numbers

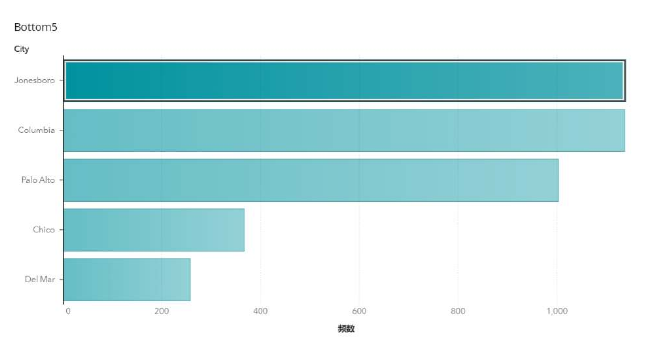


Figure 2 : Bottom cities based on numbers

* 1. Every Brand’s Transaction over a month

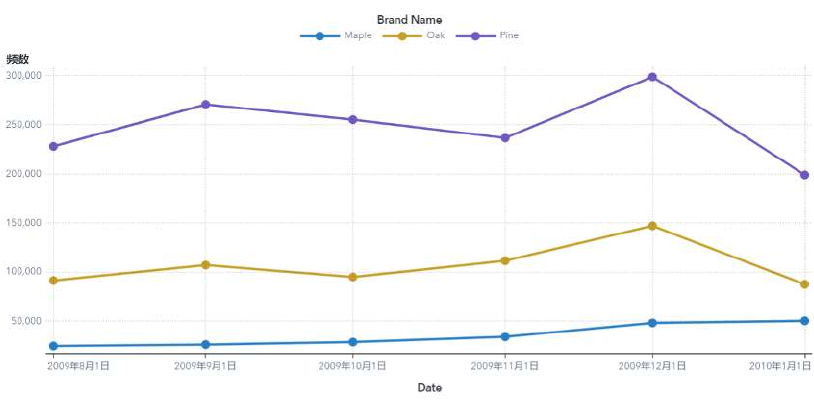


Figure 3 : Every Brand’s Transaction over a month

* 1. Top 5 and Bottom 5 classes based on number of transactions

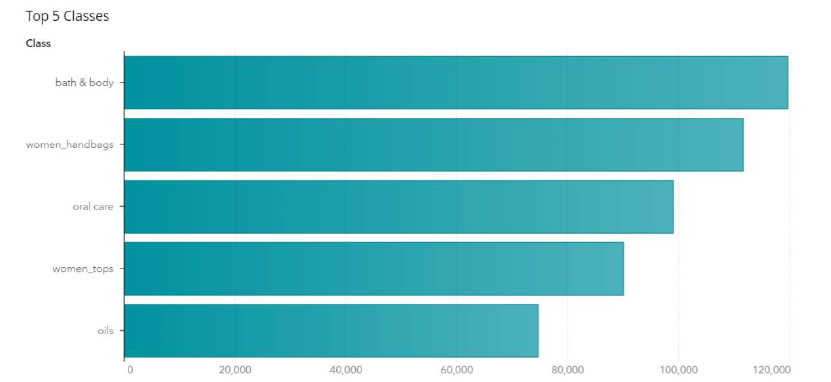


Figure 4 : Top 5 classes based on number of transactions

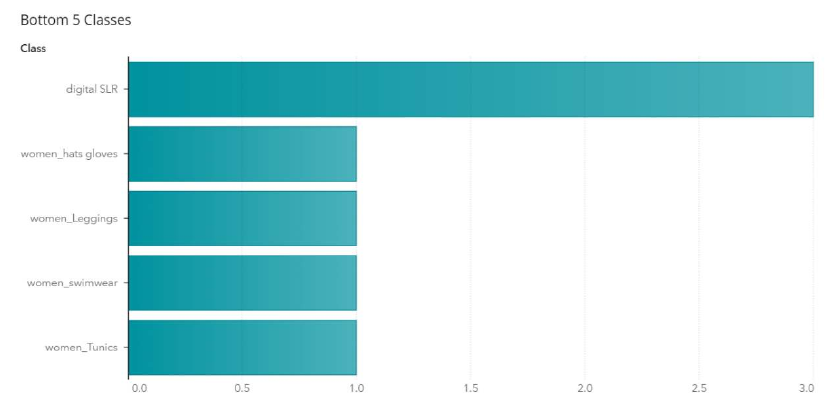


Figure 5 : Bottom 5 classes based on number of transactions

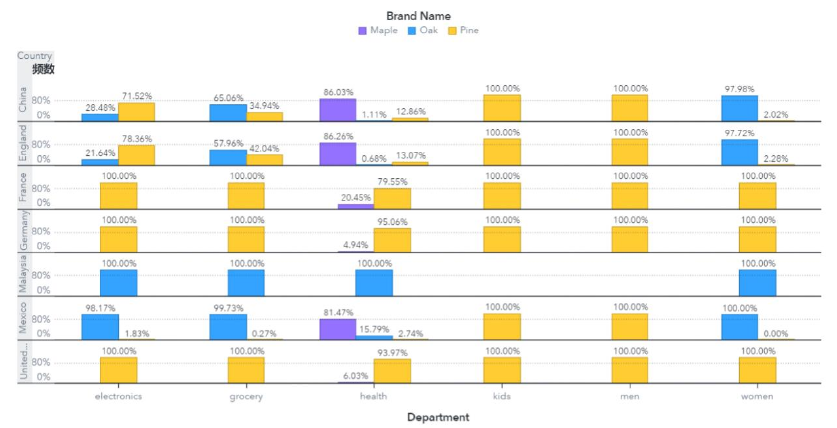
* 1. Prominent brand in terms of frequency in many departments in various nations

Figure 6 : Prominent brand in terms of frequency in many departments in various nations

* 1. Month-by-Month transaction trend for all nations

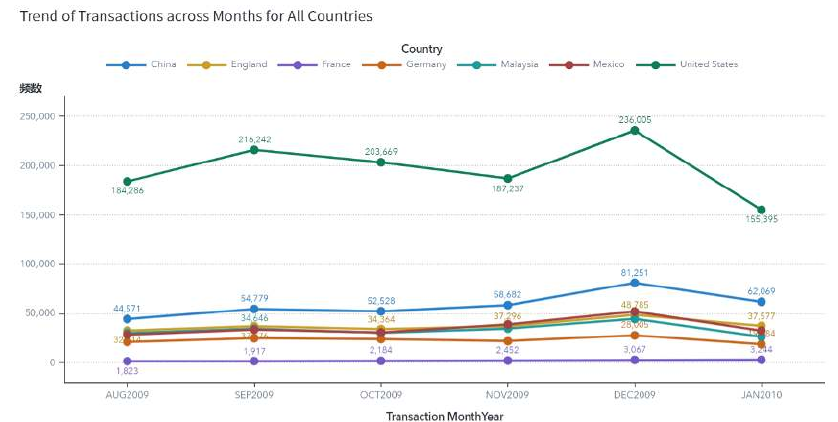


Figure 7 : Month-by-Month transaction trend for all nations

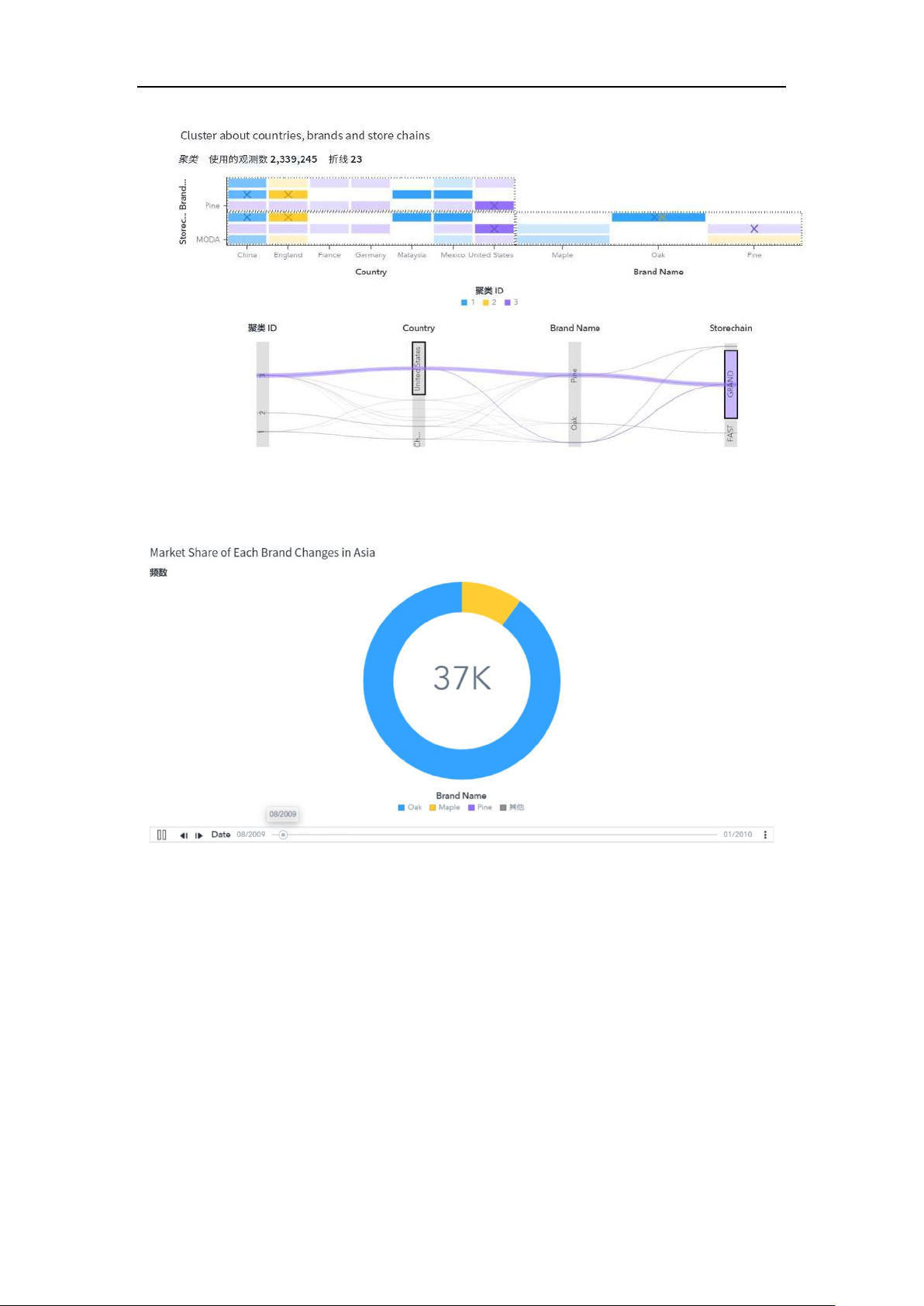
1. Data Preprocessing
   1. Check to see which brands of goods Americans prefer to buy at GRAND. 

Figure 8 : which brands of goods Americans prefer to buy at GRAND

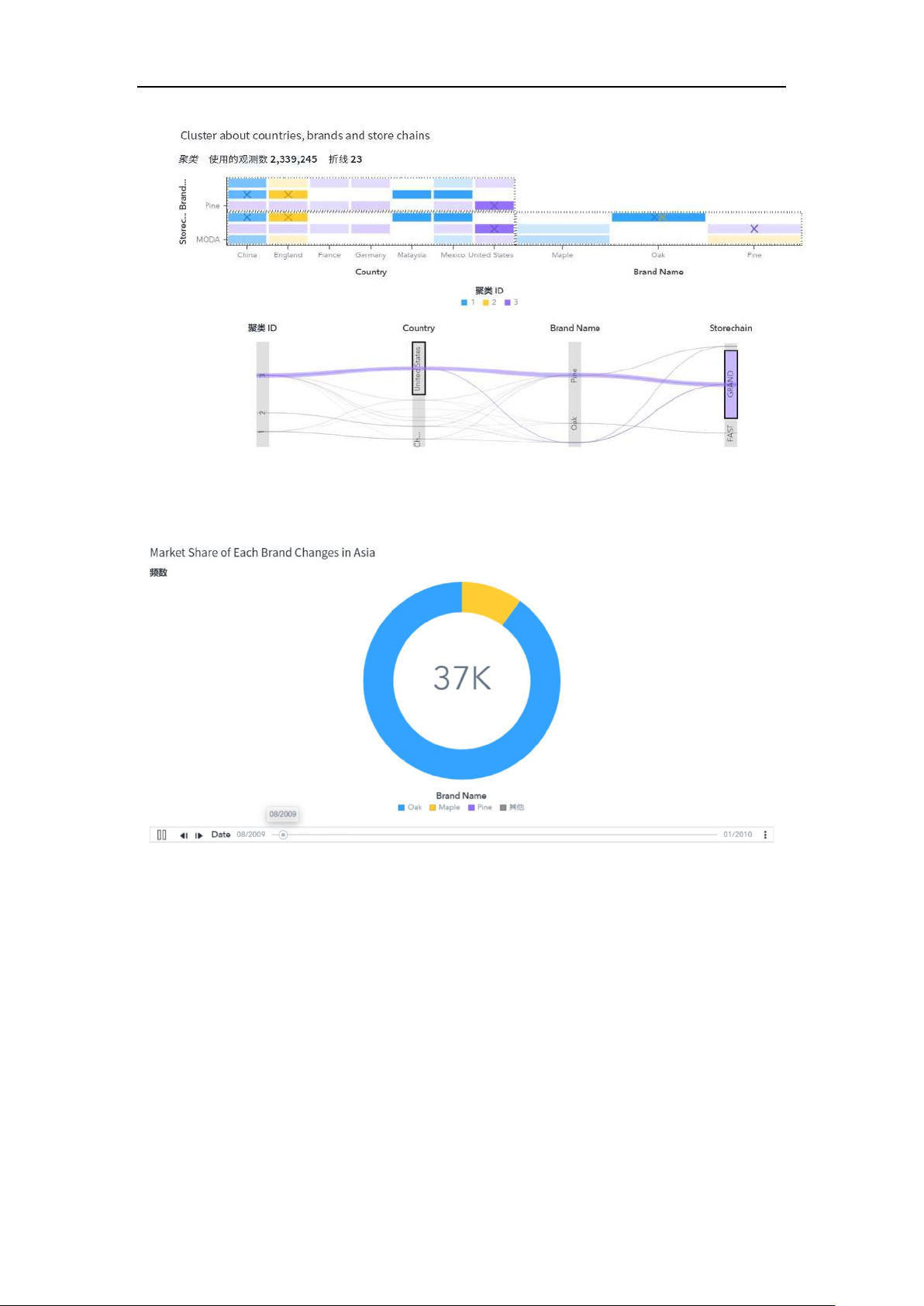
* 1. In Asia, each brand's market share fluctuated.

Figure 9 : In Asia, each brand's market share fluctuated

1. Data Mining

Bar graphs are used because they may directly indicate the quantitative difference between elements which have to be evaluated. Furthermore, we can simply achieve other visualization tasks such as the most or least selection, as well as comparisons across different categories, by employing filtering and sorting tools.



Figure 10 : data trends at regular intervals

Figure 11 : comparisons across different categories

I've decided on a line chart. The data from the worksheet's columns or rows can be turned into a line chart. The line chart may show continuous data that varies over time, making it ideal for displaying data trends at regular intervals. A line chart has category data divided evenly all along horizontal axis as well as all value data distributed evenly all along vertical axis. As a result, it is ideal for activities such as demonstrating how something changed through time.

I utilize cluster for further visualization. Clustering is the technique of grouping a group of tangible or abstract objects into numerous classes made up of comparable elements. Clustering produces a group of data objects that are comparable to those in the single cluster but different from those in other clusters. "Things are categorized, and individuals are sorted into groups." There are numerous classification issues in both natural and social sciences. Cluster analysis is a statistical analytic method for examining classifications problems, often known as cluster analysis.

I use a pie chart for further visualization. The pie chart depicts a set of data (data series: connected data points displayed inside this chart, those data are generated out from columns or rows of the data table). In the chart's legend, each data series has its own color or pattern. Represents the proportion of each item's size to the sum of all items in). The pie chart's data points are represented as a proportion of the overall pie chart.

Figure 12 : a set of data series: connected data points

Figure 13 : Clustering produces a group of data objects

1. Discussion of Findings
   1. The top five cities and the worst five cities can be found in Figures 1-and 2. That means we should concentrate our efforts on high-profile cities such as Shanghai and Mescal, while reducing our efforts in Del Mar and Chico.
   2. Figure 3 shows three lines of varying colors, each representing a brand, and the point on the line denotes their transactions. The graph shows that Oaks and Pine were generally constant between August and November 2009, then rose and fell afterwards. However, Maple continues to rise.
   3. Figures 4 and 5 show the top and bottom of five classes. That suggests we should invest more in top classes like bath body and woman handbag, and less in woman swimwear and other similar categories.
   4. Diagram 6 depicts the attractiveness of brands in different countries based on the frequency of each section. The brand Pine controlled the majority of classes and counties, as seen in the diagram. However, in other regions and categories, Oak takes the crown of sales.
   5. Diagram 7 depicts the monthly transactions in those countries. The number above the dots reflects the total number of transactions, while lines connect the dots with colors that represent different countries. The line graph with animations depicts how the number of transactions evolved over time.
   6. The association between Country Brand name and Store chain is depicted in figure 8 as a cluster. By selecting the block we were familiar with, the line between them would appear automatically; by analyzing them, analysts would examine the association between them and deduce a clear clew. Whenever we select the United States, the diagram below depicts what happens.
   7. Figure 9 depicts the three brands' market share in Asia, with the amount in the center representing total sales for the brand names in Asia. This pie-chart is composed of three layers by different colors, each representing a brand, and the area of each section representing the portion of the market.
2. References

# References

Greiner, L. (2011, January 7). *dbta*. Retrieved from dbta: https://www.dbta.com/Editorial/Trends-and-Applications/What-is-Data-Analysis-and-Data-Mining-73503.aspx

Mike. (2021, January 26). *Whatagraph*. Retrieved from Whatagraph: https://whatagraph.com/blog/articles/analytical-report

1. Appendices
   1. Appendix A

According to a recent global analysis of ACI Worldwide (NASDAQ: ACIW) and Global Data, upwards of 70.3 billion real time transactions were conducted globally in 2020, up 41% from the previous year, as the COVID-19 epidemic drove trends away from cash and cheque and toward greater dependence on real-time and digital payments. The second edition of 'Prime-Time for Real-Time,' which was initially released in 2020, examines worldwide real time, account-to-account payments volumes and predictions in 48 global markets. It forecasts a 23.6 percent compound annual growth rate (CAGR) for real-time payments from 2020 to 2025.

* 1. Appendix B

"Real-time payments are still in their infancy around the world, with many countries focusing on the obvious use-case of P2P payments," said Samuel Mirant, Global Data’s lead analyst for payments. "However, the epidemic has created an opportunity to accelerate these instruments' growth trajectory." Consumers will naturally switch to using P2P payments for e-commerce over the slower and less comfortable procedure of using cards online as they become accustomed to the speed of real-time settlement. Once enough consumers recognize real-time payment brands and the user base is large enough to provide sufficient value to businesses, there is potential to go into in-store payments."