

# **Data Visualization**

## **Learning Outcomes:**

1. Identify and discuss fundamental concepts related to visual perception and visual representation.
2. Design, implement and evaluate Interactive Visualization Systems.
3. Apply visualization tools and techniques to obtain insight from data sets.

## **Introduction to data visualization**

### ***Definition***

Data visualization is the representation of data and information by encoding it as graphical visual objects. These objects include but not limited to plots, lines or bars. The goal is to communicate information embedded within the data clearly and efficiently to users interacting with these objects.

There are a number of tools which provide an accessible way to view and understand trends, outliers, correlation, and patterns in data and our focus in this article is to present some of the tools and make conclusions based on observed patterns in the data.

### ***Advantages of effective data visualization tools***

The idea of visual has been a fundamental key in humans as our eyes are drawn to patterns (line length, shape, orientation), and colors (hue) in nature. This idea is referred to as pre-attentive processing, the idea that the subconscious accumulates information from the environment. This enables us to see red as different from blue and even identify art and beauty in things.

As such, data visualization plays the same role and it is a form of visual art that grabs our interest and keeps our eyes on the key concepts. When you have statistical data represented graphical, we can identify complex ideas communicated with clarity, precision, and efficiency.

### ***Quantitative messages***

They are seven and they differ in how separate values relate to each other. These messages always reveal relationships, and it's the relationships that conveys the deeper meaning that deserves attention. These messages include:

- Nominal comparison: A simple comparison of categorical subdivisions in no particular order (bar chart can be used) such as sales volume by product code.
- Deviation: Categorical subdivisions of a measure compared to a reference measure, expressed as the difference between them. A bar chart can show comparison of the actual vs. the reference amount.
- Frequency distribution: Counts of observations of a particular variable for a given interval, e.g. the number of students between the age 10-20, 21-30, 31-40 etc. Vertical bars can be used to emphasize individual values (histogram) and lines to emphasize overall pattern (frequency polygon)
- Correlation: Comparison between observations represented by two paired sets of variables to determine whether as one set moves up, the other moves down or the same direction as the former and how strong is the movement. Scatter plots can be used
- Part-to-whole: Measures the individual categorical subdivisions as ratios to the whole. Bar charts can only be used.

- Ranking: Categorical subdivisions of a measure ordered by size either ascending or descending. Bar chart may used
- Time-series: A single is captured over a period of time, such as the number of customers visiting a particular store.

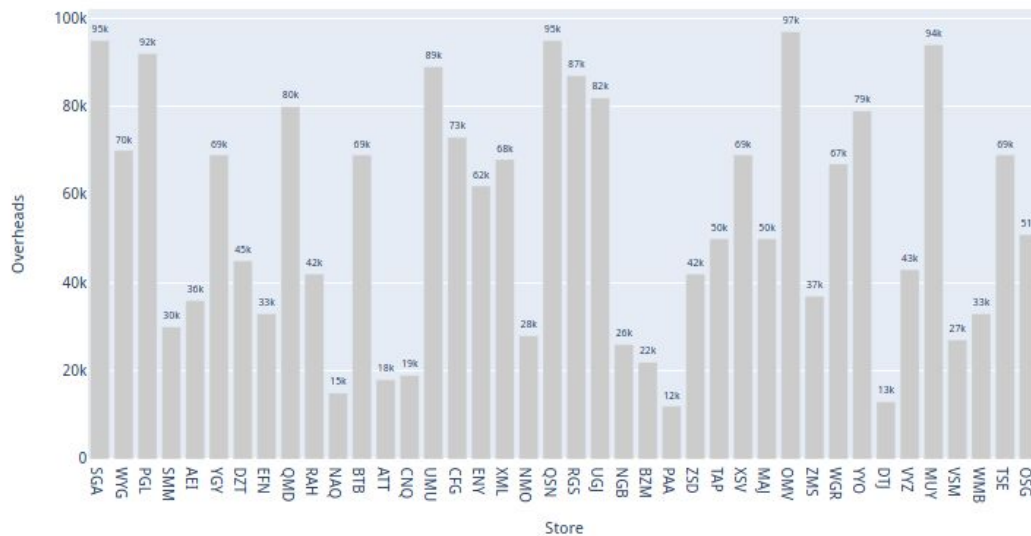
In this article, we will observe and use the above information to derive conclusions for the store/customer data for ChrisCo. a fictional company.

## Store Overheads

The following visualizes the annual cost of overheads for each store for Chrisco.

Overheads definition: [WIKIPEDIA] Refers to an ongoing expense of operating a business. Overheads are the expenditure which cannot be conveniently traced to or identified with any particular cost unit, unlike operating expenses such as raw material and labor. Therefore, the bar below gives the visual representation of this idea.

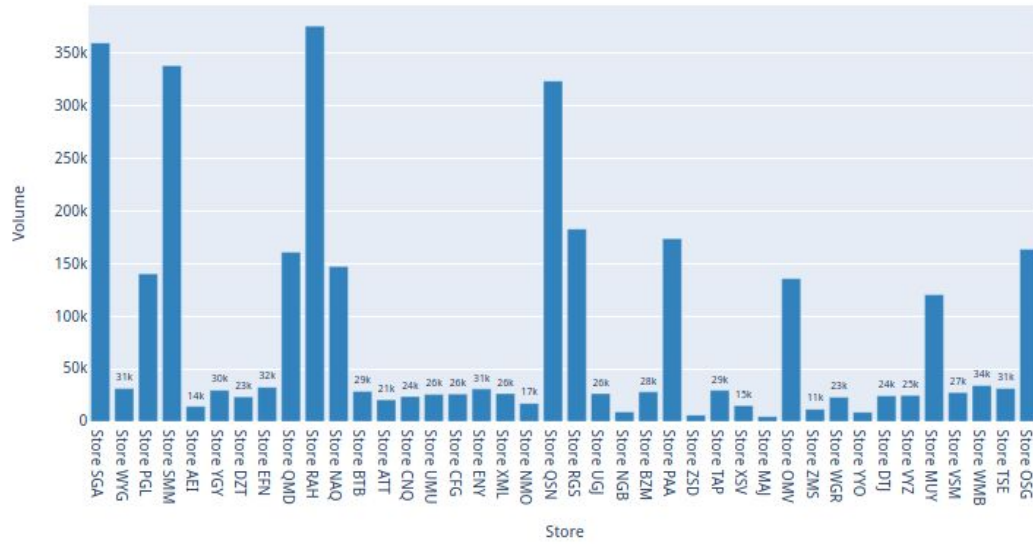
Figure 1: The total annual cost of overheads in (£) for each store



## Total Volume

The following plot shows a visual representation of the total volume of each store.

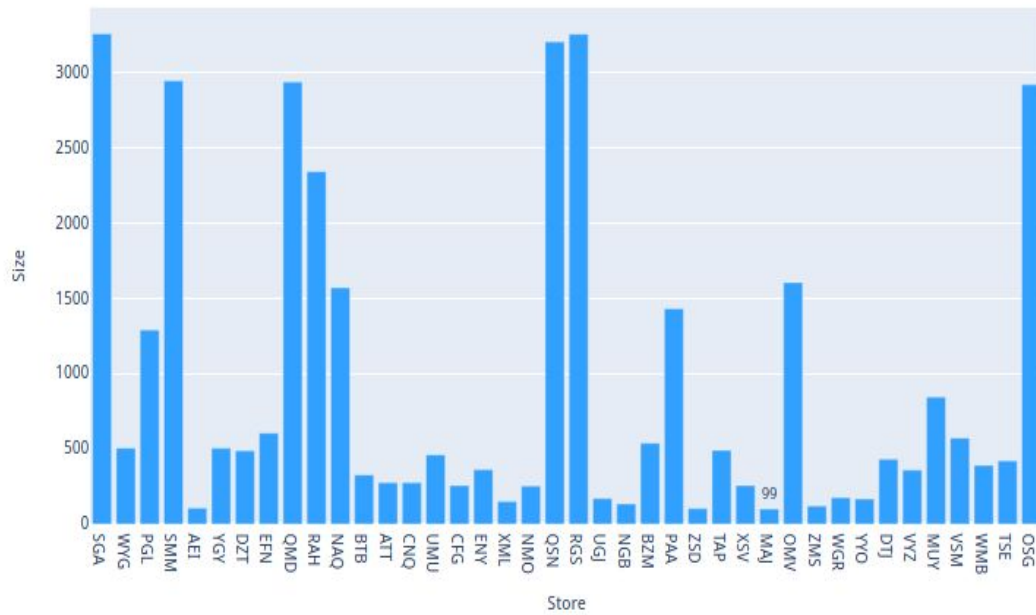
Figure 2: The total volume of each store



## Store size

The following plot shows a visual representation of the size of each store.

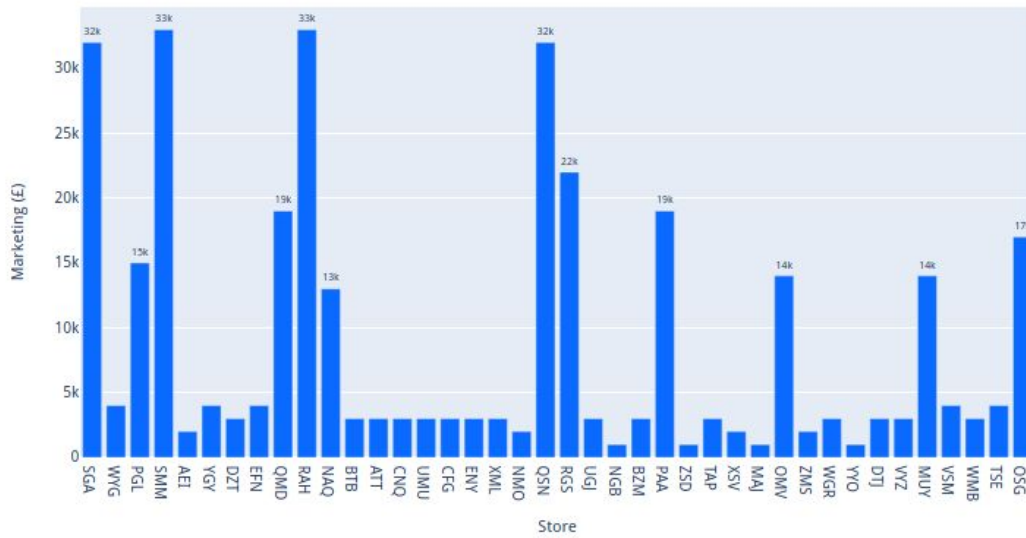
Figure 3: The store size (floor space) in metres squared for each store



## Store marketing

The following plot shows a visual representation of the local marketing for each store in (£).

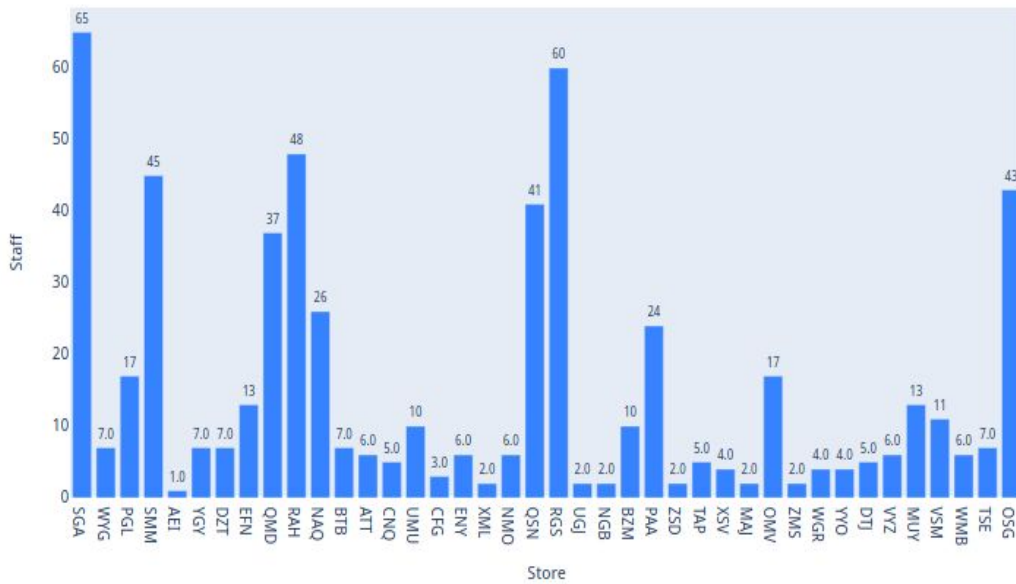
Figure 4: The total annual spend on local marketing for each store in (£)



## Store staff

The following plot shows a visual representation of the number of full-time staff employed at each store.

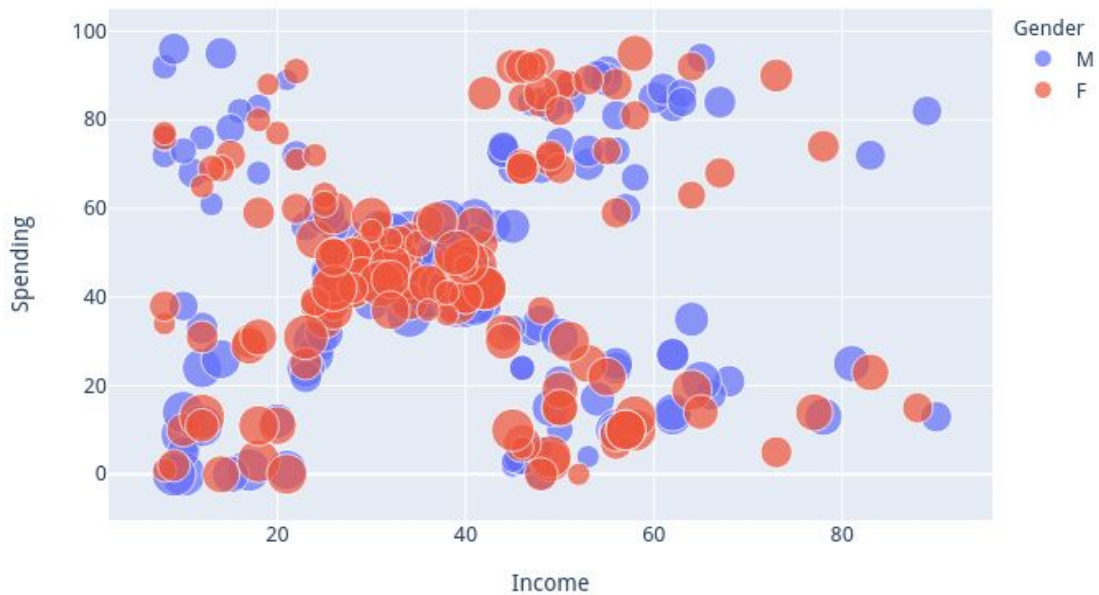
Figure 5: The total number of full-time staff employed at each store



## Customer profiles

The following plot shows a visual representation of the customers profiles which includes the age, gender, income and spendings.

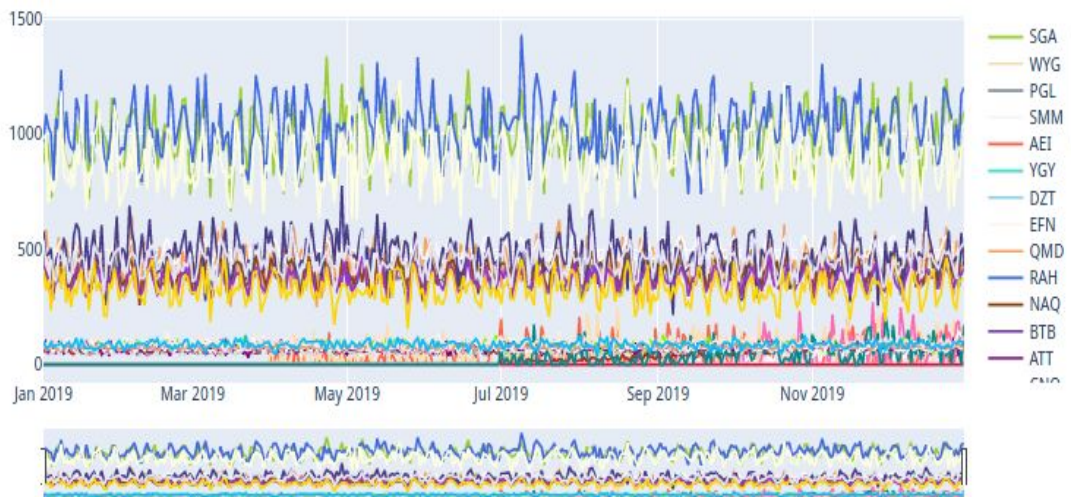
Figure 6: Customer Profiles



### Daily customers

The following plot shows a visual representation of the number of customers coming to shop at each of the stores per day.

Time Series with Rangeslider



### Findings

- We have observed that the number full-time employees a store has is correlated with the store size. This can be seen in Figure 3 and Figure 5 above.

- Another interesting observation we have made is that the number of overhead increases with the size of the store. Larger stores tend to have a higher overhead because of the bills they have to pay. These can be seen in Figure 1 and Figure 3.
- Another finding based on Figure 6 above is that customers who earn between 20 to 70 as income tend to shop more.
- We also see that larger stores receive more customers daily based on the Figure 7 above (Daily Customer time series data). Example stores are SGA, QSN, RAH, and SMM.

## **Work Review**

I have clearly illustrated the findings made by observing the graphical plots considering the quantitative messages I outlined at the start of this article. I've also labelled the plots for better user experience in trying to understand the information concealed in the data.

Humans are drawn to color and sizes, therefore I've made sure, my plots demonstrate this. To give a user better experience, I made the plots interactive in the sense that users can hover and get more information about the data and also change the graphical display.

## **References**

- [Wikipedia](#)
- Tapping the Power of Visual Perception by Stephen Few
- Selecting the Right Graph for Your Message by Stephen Few