**Information Visualisation and Big Data**

**Coursework**

**Deadline Date: 09/12/2020**

Course Teacher: Christopher Walshaw

Written by Nasra Khalif

**PART C: Information Visualisation [40 marks]**

The following information is regarding the dataset;

* There are two tables; Data and Summary\_data
* In the table Data, the venue names are the columns, and the dates are the rows. The dates are from 01/01/2019 till 31/12/2019. This table contains the daily visitors at these venues.
* In the table Summary\_data, it contains the venues and other information such as the average age, distance travelled, duration, gender and spend of the visitors.

**Visualisation 1**

Chart, histogram

Description automatically generated

Figure 1 Bar chart showing the total visitors of venues

* This bar chart shows the total number of visitors in 40 venues in 2019, in descending order.
* The venue with the biggest volume of visitors is RDA with 189,234 visitors.
* The venue with the least visitors is AEQ with 5676 visitors.
* This bar chart shows the venues can be grouped into 3 groups; highest is >80,000, medium is between 30,000 and 80,000, and the lowest is below 50,000.

**Visualisation 2**

Chart

Description automatically generated

Figure 3 This line plot shows the daily visitors in all the venues

* Figure 3 shows the daily visitors during 2019 of all the venues.
* Most venues seem to be oscillating throughout the year.
* It is also evident that not all the venues had visitors throughout the year, the venues with lowest number of visitors did not have visitors in the beginning of the year. This could be because these venues were closed or people were not aware they existed.

**Visualisation 2**

**Chart, histogram

Description automatically generated**

Figure 4 This is graph shows the line plot, 14 day moving average and trendline of the least visited venues

* Figure 4 shows the venue ZPL (green line) was open in the beginning of 2019, however, the trendline shows there is a decrease in visitors in this venue. However, there is a sudden drop in the number of visitors in July and the numbers have not recovered. This is suggestive of the venue closing.
* It also shows that ZBG and BQV started to have visitors in April 2019, BKI and YDI started to have visitors in July. YVW and AEQ started to have visitors in September.
* Although, these new venues have the least number of visitors, the trendline is positive, and this indicates they are increasing in the number of visitors.

**Visualisation 4**

Chart, histogram

Description automatically generated

Figure 5 This is graph shows the line plot, 14 day moving average and trendline of the highest visited venues

* Figure 5 shows the venues that have had the most visitors in 2019. However, unlike the least visited venues, it does not show they have increasing number of venues.
* The venue PXI has a downward trend and shows there are less visitors in the end of 2019 than in the beginning.

**Visualisation 5**

Chart, radar chart

Description automatically generated

Figure 6 shows radar subplots for highest visited venues.

* Figure 6 shows the radar subplots for the highest visited venues.
* In all the venues the average duration spent is not common.
* The venue RDA is the most visited, it is also the venue most travelled for. The proportion of women and the average age visiting this venue is also high. This may because of its reputation, or it is well known. However, the average time spent here is low and the amount of money people spend is also less.
* In contrast, the venue SJU do not have visitor’s travel most travelling long distances, but the proportion of women is still high. The average time spent in this venue is low, but the amount of money spend s high.
* The subplots show that the venues with the more visitors does equal high spending. The time visitors spend in these venues also show that they are not interested very much.

**Visualisation 6**

Chart, radar chart

Description automatically generated

Figure 7 shows radar subplots for least visited venues

* Figure 7 shows the radar subplots of the venues least visited.
* The venue SJU is a popular venue, the average time visitors spend is more than the highest visited venue. This could be because it is new, and people are more interested in exploring this venue. The average age of the visitors is older, and they spend a lot more money on this venue on average.
* The average age of visitors in venue SPF is young. They are majority female, and they travel long distances to visit this venue. However, they do not spend much in this venue.
* The time visitors spend at these locations are generally more than the highest visited places.