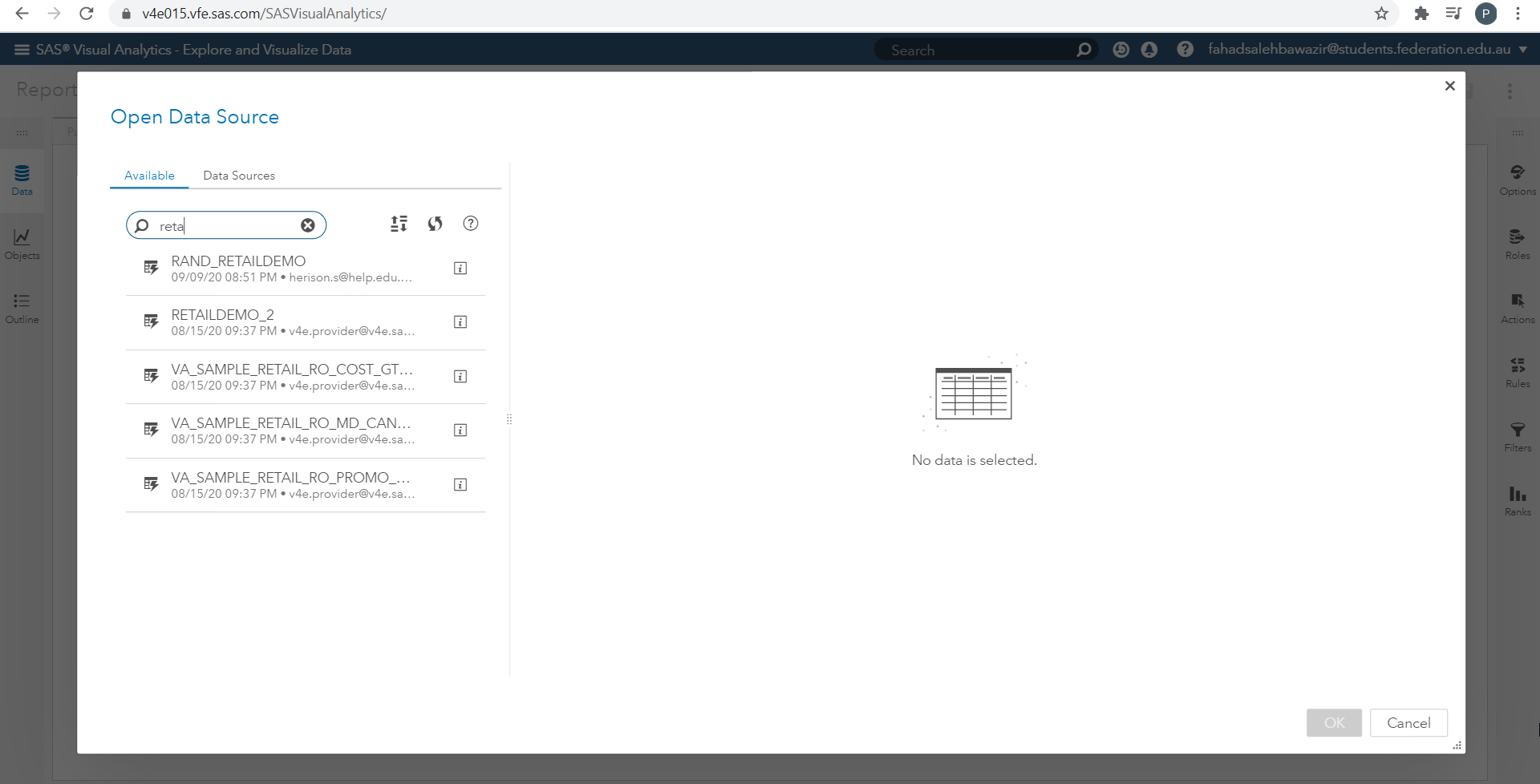
Data analytics using SAS Viya

Name of the Student

Name of the University

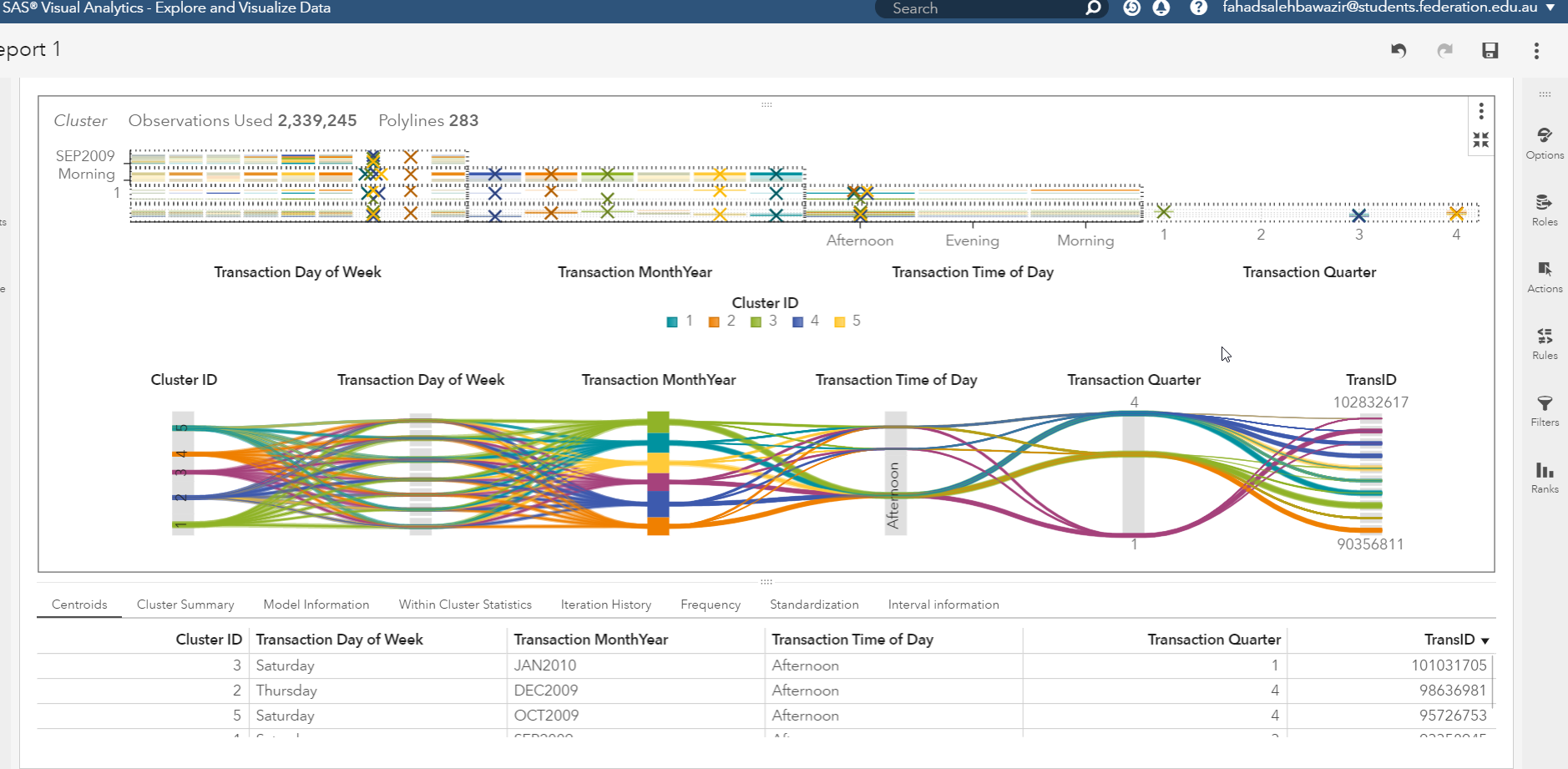
Authors note

For this project the RETAILDEMO2 dataset is utilized for generating the following plots. In the following sections the data set is analysed in order to plot 5 insightful plots from the available dataset.



Q16

For this question, the cluster analysis for the transaction related data.

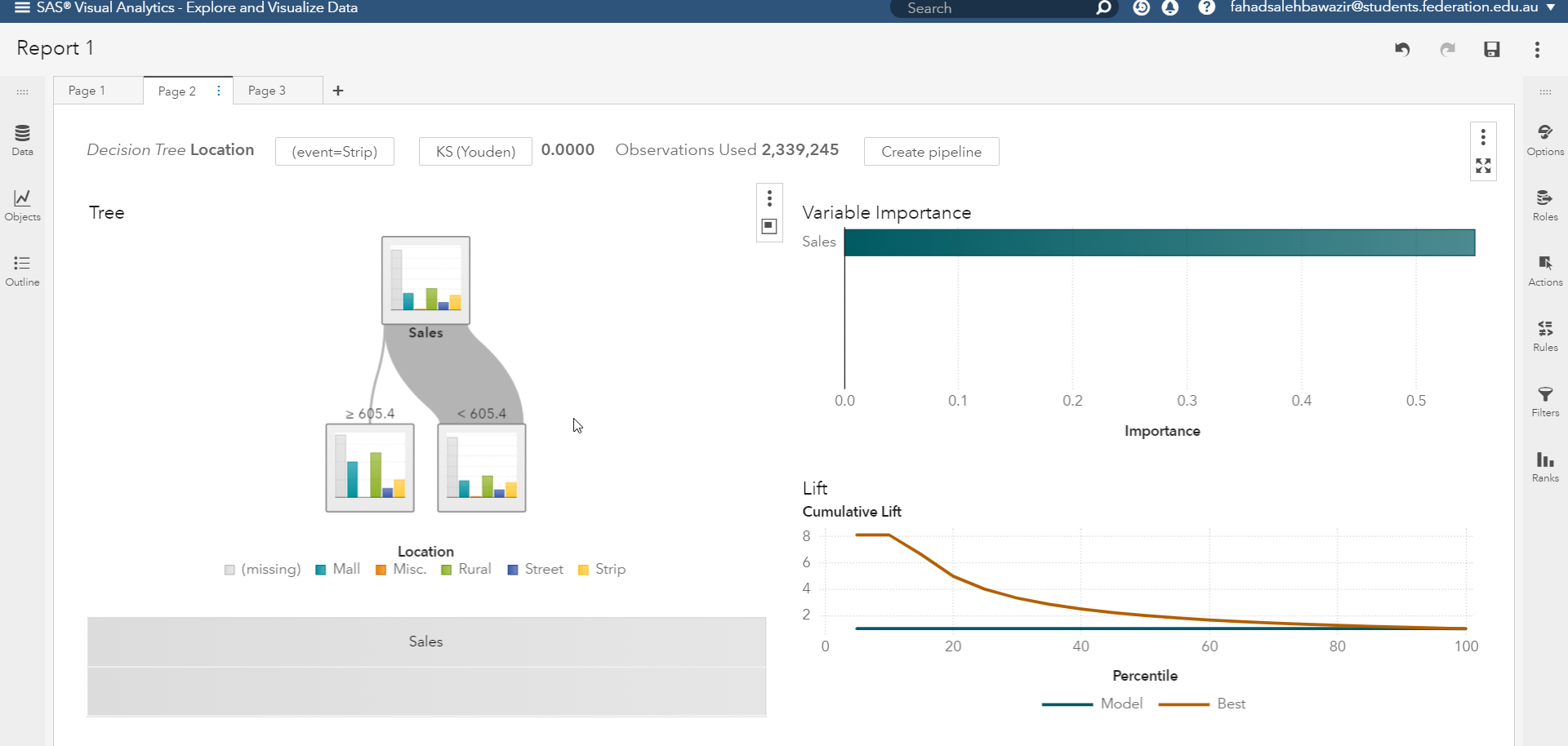


From the above analysis, it can be stated that there are more than 5 clusters for the transaction months, transaction day of the week. Most of the highest amount of sales are clustered in the afternoons of Thursdays and Saturdays. The months are January, December and October. The most sales are recorded in the first and 4th quarters.

The Cluster is considered as the group of objects or data rows that belongs to same specific classes that matches some criteria. In other words, it can be stated that similar type of objects are grouped in order to create a cluster whereas dissimilar objects can also grouped in order to create another cluster. The Clustering process can help the retail business owners or marketers in order to discover different distinct groups depending upon the characteristics from their available customer base. Depending upon the characteristics the organization can target the customer groups based on the purchasing patterns in order to improve the revenue.

Q17

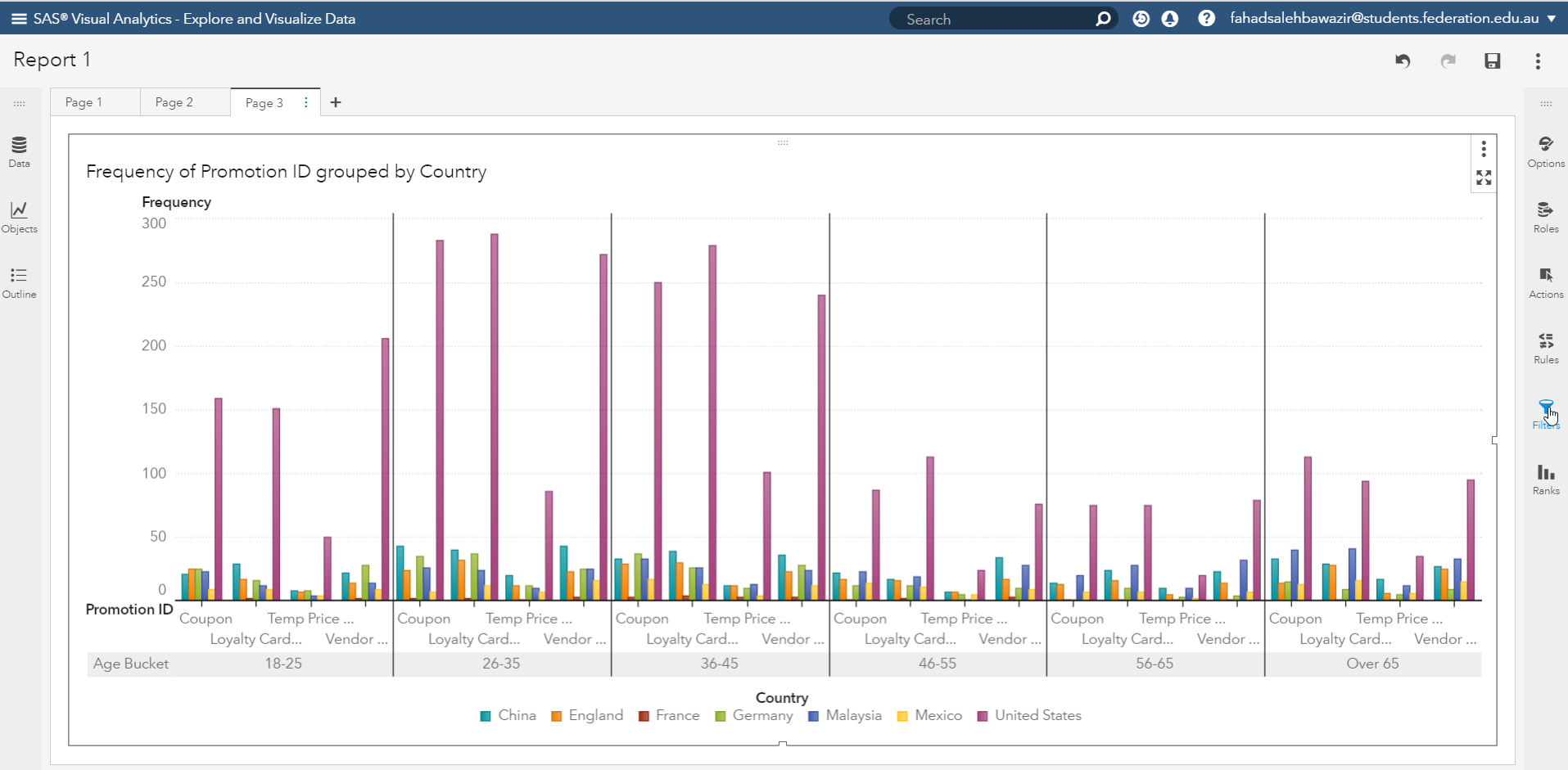
In the next plot the decision tree for the sales location is plotted.



From the above plot, it can be stated that sales amount is the best possible driver for determining the location of the sales. In case the sales amount is more than 605.4 Mall or Rural location. Where as in case the sales value is less than 605.4 then the location of sales can be one of street, strip or Rural.

Q18

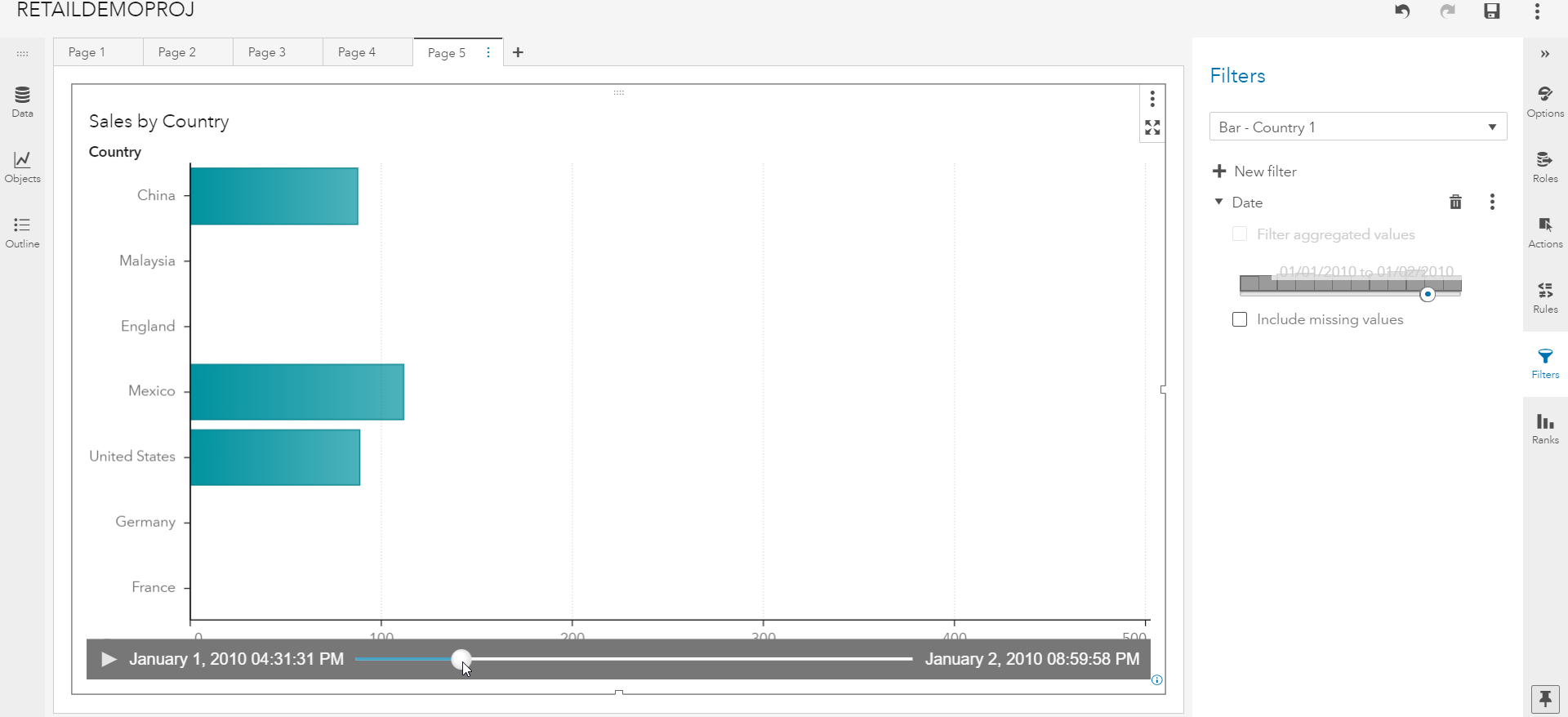
In order to find out the variation of promotion ID is across different age brackets in different countries, the following bar plot is generated using the SAS VIYA tool.



Here it can be stated that majority of the promotions are fruitful in case of united states. Majority of the promotions are helpful in attracting the customers belonging to the 26-35 and 36-45 age group. Minimal impact is on the 56-65 age group.

Q19

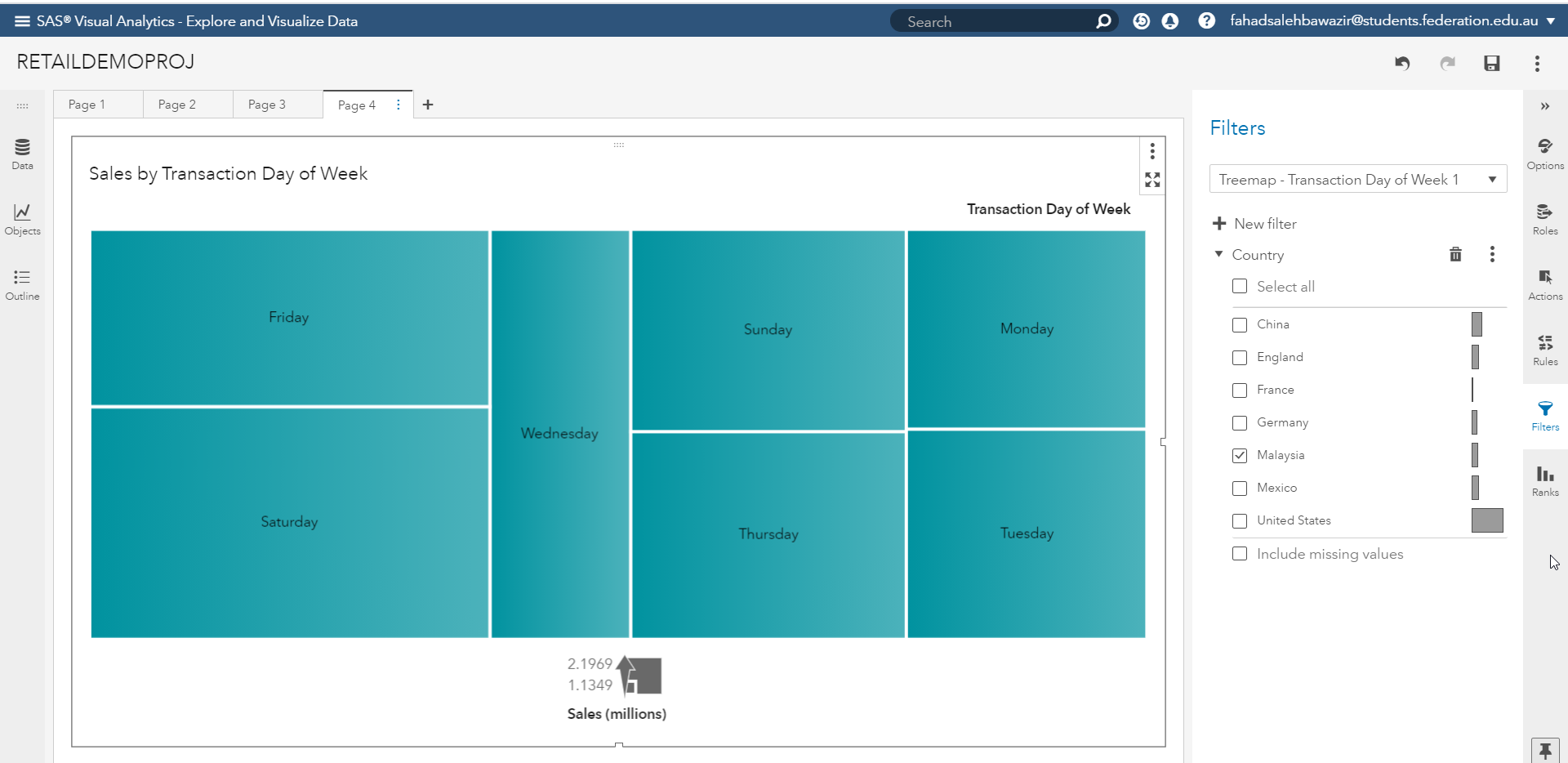
The next plot depicts the sales trend in the different countries 01/01/2010 to02/01/2010.



Here, it can be stated that the plot also includes the animation for the amount of sales depending on the time of the sales as depicted in the plot above for the same. This can be helpful in easily understanding the trend of sales. On the mentioned dates.

Q20

In the last plot the most and least popular day in terms of Sales for Malaysia is plotted using the tree plot object available in the SAS platform.



From the above tree plot it can be stated that most popular day in terms of sales is Saturday and least popular day is Tuesday. Here it can be stated that maximum amount of sales is recorded for the United states country and least sales is for France as recorded in the dataset retaildemo2 dataset available in SAS viya.