**Date:** 2 March 2018

**Time:** 0900-1000

**Venue:** SIS Level 5 Meeting Room 1

**Present**: Prof Kam, Instructor Minashi, Eunice, Prajval, Wen Jie

**Agenda:**

1. Interim Presentation
2. Submit Interim Report and Slides (Hardcopy)
3. Present Interim Components
4. Receive Feedback on how to move forward

|  |  |  |
| --- | --- | --- |
| S/N | Area of Discussion | Action to be taken  |
| 1 | General Feedback* Good presentation flow
* Need to include better screen shot choices
	+ For instance, we spoke of data cleaning and what we did. We should have included more screen shots of what data was removed
 | Include more evidence of data cleaning and removal of outliers. Such as the ones with 1k+ Page Views |
| 2 | Data Preparation* Unclear on how we treat our data sets
* We mentioned outliers but only had two pictures
* Sales Data issues: we mentioned issues but did not show evidence
 |  |
| 3 | Data Exploration* A lot of room for improvement
* Tableau is not THAT useful a way for data exploration
* When doing data exploration we should state our methodology for each exploration method appropriately
	+ i.e. barchart is used for…. To represent … for categorical data
* By listing each method for each scenario, it will be clearer for guiding our decisions
* We should be going to the overview before sharing about the specific visualisations
* Statistics methods should have been used before exploratory, when we did contingency analysis
 |  |
| 4 | Data Exploration Improvements: #9* (1) Should have been a scatter plot of the product sales quantity
* (2) Next could have done a Pareto Chart which plotted % Revenue against product category with a line on the dual axis that had the percentage frequency
* (3) Next could have done a Sales & Revenue graph with a Histogram and a Box Plot
	+ This gives us the distribution of sales and revenue to see if it is normally distributed or not
* (4) Next a scatter plot of Product revenue against Product sales could have been done. Thereafter a quartile analysis could have been achieved with this where we can group products into: High sales/High revenue or Low sales Low revenue
* (1)-(4) would’ve provided us more methodology for our binning methods!
 |  |
| 5 | Data Exploration Improvements: #14* Cross Tabulation on Contingency Analysis\*\*\*\*\*
* When we are doing categorical analysis, it is called contingency analysis
* Should have used a Mosaic Plot that is also available on Tableau
 |  |
| 6 | Data Exploration Improvements: * From categorical analysis, we should not use regression
* Should have used Chi-squared method
* Missing: Appropriateness of visualization choices
* Did not give conclusion to each visualization
* When showing statistics like average or median
	+ We should use box plots & histogram
* Total cost: Use Bar charts
 |  |
| 7 | Data Exploration Improvements: #26* Contingency Analylsis
	+ Should have used Mosaic Plot
	+ Most of them should have been mosaic plots
 |  |
| 8 | Data Exploration Improvements: #30* Should have used line graph instead of bar chart
	+ Split them into quarters if we are looking at seasonality
	+ Special
 |  |

Minutes taken by: Eunice Wang