# Audience

The audience of this report would be the senior management which is interested in understanding the overall performance of the business, how different stores are performing, the wages of the employees, and other aspects related to the stores and business which need to be paid attention to. The dashboards and story will contain the performance of the business and the stores in addition to any specific thing we might come across which is not common to the rest of the data.

# Data Elements

The variable “Sales $m” contains the value of total sales in financial year in million and “Wages $m” is total wage and salary bill for financial year in million. “Sales $m” and “Wages $m” are numerical variable which are continuous in nature. “No. Staff” column contains the number of full-time employees which are employed on weekly basis and it is numerical variable which is Discrete in nature. “Location” tell us the kind of location the store is located in and it is a categorical variable and the scale is nominal. “State” tells us the state store is located in it is categorical and nominal. “Age (Yrs)” tells us the age of operation of the store and it is numerical and continuous in nature. “GrossProfit” is the gross profit of each store in financial year in millions. “Adv.$'000” is the money spent on promotional and advertisements in the entire financial year and it is in thousands. “GrossProfit” and “Adv.$'000” are both numerical variables which are continuous in nature. “Competitors” are the number of competitive stores in the area and it is numerical in nature and is discrete in nature. “HrsTrading” is the number of hours the store is open for business and “Sundays” tell us if a store is open on Sundays or not. But “HrsTrading” is numerical variable which has continuous nature and “Sundays” is a categorical variable which is nominal in nature. “Mng-Gender:” variable is indicator of gender of store manager and it is categorical and nominal, “Mng-Age” indicates age which is numerical variable and is continuous in nature, “Mng-Exp” is the experience store manager has in manager role and it is also numerical and continuous in nature, and “Mng-Train” is the count of management training courses taken while being under employment of Setco and even though it is numerical but it is discrete in nature. “Union%” is the percentage of workers who are in union. “Car Spaces” is the measure of number of parking spots available at the supermarket. “Online Channel” is the indicator of availability of online channel for each store. “Basket:2016” and “Basket:2017” has the value of cost of the basket food items in each store at 1st of June for both the years respectively. “OnlineSales$m” tells the online channel’s sale in the financial year in million. “Wastage” is the indicator of the stock which not sold. “Union%”, “Basket:2016”, “Basket:2017”, and “OnlineSales$m” are all numerical and continuous in nature but “Car Spaces” even though is numerical but is discrete in nature. “Online Channel” and “Wastage” are categorical variables but “Online Channel” is nominal and “Wastage” is ordinal.

# Types of Visualizations

In case of continuous variables bar graphs, line charts, and pie charts would represent the data well. Line charts or graphs might help us visualize the affect of the variable more. In case of discrete variables above charts in addition to the charts like stacked bar charts, spider charts. In essence bar graphs, line charts and pie charts would best represent the numerical data for uni-variate analysis. The same goes for categorical variables but the difference would be that we will be visualizing the counts of the different category under each variable. We have locations or states data as well so we can also make map and visualize different info accordingly. Heat maps can be made to visualize the frequency of the numerical variable or even categorical variable. Categorical and numerical variables can be visualized using box-plots and two numerical variables can also be visualized with the help of scatter plot.

# Ethics

We need to take care of security, privacy, and transparency while developing and sharing the visualizations considering the data provided to us by the company is fair and accurate, which means that the data provided to us fair and has no bias and it is accurate so that the accurate results can be produced. We need to take care of the security of the data and the visualizations we create whether or not we are making the visualizations public or not. This is to be done because there might be some sensitive data which the organization does not want to go public with or doesn’t want their competitors to know. In addition to that we need to take company’s consent before publically showing the visualizations related to the data they provided as some sensitive information might leak to the competitors which is also breach of privacy of the company and company’s data as visualization is nothing but visual representation of the data. Data transparency issue is not directly related to the person who creates visualizations but it should be made sure that the data business analyst is working on should have been collected using transparent means and the individuals involved should have knowledge about why or for what purpose the data was collected for.

# Dashboards and Stories

Task 1:

3. a. In this study, the senior management which is interested in understanding the overall performance of the business, how different stores are performing, the wages of the employees, and other aspects related to the stores and business which need to be paid attention to.

The data file for the study is called StoresThree.csv and contains number of stores, i.e., 150 stores dataset. There are 22 variables in the dataset such as total sales, number of staff, location of the shopping centre, etc. Variables in that file are:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Variable Name and Description: | | |  |  |  |  |  |
| Sales $m: | Total Sales revenue for each supermarket for the financial year. $ million. | | | | | | |
| Wages $m: | Total Wage and salary bill for the financial year. $ million. | | | | | |  |
| No. Staff: | The number of effective full-time staff employed on a weekly basis. | | | | | | |
| Location: | Strip: Or typical street shopping centre in a major town or city | | | | | | |
|  | Mall: Large complex or shopping center or mall in a major town or city | | | | | | |
|  | Country: Town of 20,000 or fewer inhabitants | | | | |  |  |
| State: | New South Wales (NSW) | | | Western Australia (WA) | | |  |
|  | Victoria (Vic) | |  | Tasmania (Tas) | |  |  |
|  | Queensland (Qld) | |  | Northern Territory (NT) | | |  |
|  | South Australia (SA) | |  | Australian Capital Territory (ACT) | | | |
| Age (Yrs): | The number of full years of operation of the store. | | | | | |  |
| GrossProfit: | Gross Profit for each supermarket for the financial year. $ million. | | | | | | |
| Adv.$'000: | Advertising and promotional expenses for the financial year, $'000 | | | | | | |
| Competitors: | The number of competing supermarkets in the consumer catchment area | | | | | | |
| HrsTrading: | The total number of hours open for trading per week | | | | | |  |
| Sundays: | Open on Sundays; Close on Sunday | | | |  |  |  |
| Mng-Gender: | Male store manager; Female store manager | | | | |  |  |
| Mng-Age: | Age of the store manager, years | | | |  |  |  |
| Mng-Exp: | No. of years of experience in some form of junior/senior management at Setco | | | | | | |
| Mng-Train: | No. of management training courses taken while employed at Setco | | | | | | |
| Union%: | The proportion of the staff at the supermarket which belongs to a union. | | | | | | |
| Car Spaces: | The number of parking spaces available to the supermarket. | | | | | | |
| Online Channel | Whether or not the supermarket has an online store channel. | | | | | | |
| Basket:2016 | Cost ($) of the basket of food items in each store at 1 June 2016 | | | | | | |
| Basket:2017 | Cost ($) of the basket of food items in each store at 1 June 2017 | | | |
| OnlineSales$m | Sales revenue from Online Sales Channel. $ million. | | | |
| Wastage | Levels of wastage (unsold stock): low, medium, high. |

b. There is no ERD in the file.