Task -2 Data Insights

Data Analytics approach

Sections

- Data exploration
- Model Development
- Interpretation and report

Data Exploration

- Understand the characteristics of given fields in the underlying data such as variable distributions ,whether the dataset is skewed towards a certain demographic and the data validity of the fields. For Example, a training dataset may be highly skewed towards the younger age bracket. If so, how will this impact your results when using it to predict over the remaining customer base.
- There are some limitations in the given datasets like some values are missing and some data types are different according to their value.

- Furthermore , transformations of required data so that it is an appropriate format for analysis. This may include the steps such as ensuring that the data types are appropriate ABS data at a geographical level to create additional variables.
- Document assumptions ,limitations and exclusions for the data; as well as how you would further improve in the next stage if there was additional time to address assumptions and remove limitations

Model Development

- First of all, we have to determine a hypothesis related to the business question that can be answered with the help of existing data.Perform statistical testing to determine if the hypothesis is valid or not.
- Create calculated fields based on existing data, for example, convert the D.O.B into an age bracket.
- Test the performance of the model using factors like residual deviance AIC, ROC, curves, R Squared).Appropriately according to the model performance, assumptions and limitations.

Interpretation and Report

- Visualization and presentation of findings. This may involve interpreting the significant variables and co-efficient from a business perspective.
- With the help of this Slide, we get an idea around the business issue and support our case with quantitative and qualitative observations.