TASK - 2
DATA INSIGHTS

DATA ANALYTIC APPROACH

SECTION

1.Data exploration

2.Model development

3.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, transformation of required data so that it is in an appropriate format for analysis. This may include steps such as ensuring that the data types are appropriate and rolling data up to an aggregated level. Or, joining in already aggregated ABS data at a geographic 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

1.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.

2.Create calculated fields based on existing data, for example, convert the D.O.B into an age bracket

3. 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.

THANK YOU