Project Type	Pick an Option 1) Statistical Modelling **2) Machine learning** **3) Visualisation** 4) **Analysis/Insights**
Scope Definition	The proposed scope for this project is to analyse and visualise Vodafone's NPS and First Call Resolution (FCR) survey data with a primary focus on the free text responses. The purpose of this exercise is to develop actionable insights for Vodafone's Customer Care to improve NPS and First Call Resolution. This data set should be analysed in conjunction with accompanying customer demographic, and call features to provide more targeted insights. The proposed scope items are: a) Look at the raw statistical data and analyze the trends to see what insights can be extracted at a product/plan, agent skill/department, customer level b) Understand which products/plans/departments drive particular behaviors that have a specific NPS outcome and thus produce actionable insights. c) Analyse the free text responses customer gives. Create meaningful clustering or classification models to group themes from comments by volume relating to NPS promoters, detractors and passives and similar for Issue Resolution. The model developed should be able to be applied to future responses and classify/cluster with high precision and recall. d) From the text model, get a breakdown of top drivers of promoters and detractors as well as issue resolution, which will guide us as an organization on what to focus on to deliver higher NPS and FCR. e) Provide sentiment analysis model that can be re-applied to future surveys f) Create a repeatable reliable data model to be able to extract actionable insights from new surveys received (~4K per week) when required. g) Create some visualizations which can be use on a regular basis and refreshed with new data showing the main drivers of NPS promoters detractors and FCR.

Background	Vodafone takes ~100K Contacts per week and the primary purpose of those contacts is to service customers in a manner which will build and enhance the brand, reduce customer churn and deepen customer engagement and grow relationships as measured by NPS, therefore improving NPS is key
Expected Benefits	We would expect to see a data methodology set-up to enable quick weekly analysis of text comments in NPS surveys to be able to gain actionable insights and understand drivers behind NPS scores moving up and down. An overall analysis of NPS and FCR over the last 12 month of data to show insights as to the changes in drivers of NPS over the period of time to understand which outside factors drive NPS.
Risks & mitigation	Any NDA requirements – Yes, a confidentiality statement / agreement would be needed due to the sensitivity of the data (perceived or otherwise). What regulations and legal considerations were taken into consideration? – None at this time. What privacy considerations should be taken into account? – We would expect not to be giving any access to PII data.

Feasibility Factors	How would you like the data to be collected, verified, managed, stored and processed? No fixed or predetermined method, so would be subject to discussion and agreement with VHA. Is the data quality adequate for analysis and modelling? Is it authentic, trusted, validated, cleaned. Yes but there are several different sources of the data so will need to bring that data together, to be discussed with VHA. Underlying assumptions to be discussed.

Methodology & Approach	No fixed or predetermined method, so would be subject to discussion and agreement with VHA.
Potential Datasets	Reference datasets What data resources are currently available? All NPS data is available in a web repository and other data is available in other locations including Qlik. It can be made available in a way that would work for the students What data resources need to be discovered or created? None. Will the data be updated or refreshed regularly or is a static analysis fine to begin with? Static analysis is fine to begin with, but our online tNPS platform continuously receives new survey data.
Input	Actual metadata or data dictionary for the project. Can be provided or developed if it does not exist, either by Council or in collaboration with VHA. What are the specific data variables needed already included and what additional variables need to be collected? Variables would be usage during the period, start and end date of the period.
Key Stakeholders	
Outputs	Model Outputs Integration requirements (api, scoring) – None Software requirements (R, Python, SAS or other) – Python and appropriate TXT analytics ML tools, AWS Sagemaker or other cloud experience would be helpful in the development of the solution Hardware Requirements (Spark, Hadoop or in memory) – None Visualization link or Dashboard – Yes Insights Pack – Yes