**Project 4 Proposal**

**Github link:**

https://github.com/DannyBruzz/Project\_4\_Capstone

**Aim:**

The main aim of this project is to create a model to predict whether a prospective customer from Perth will respond to a targeted ad campaign.

**Background:**

HQF (Health Quokka Fund) has received the below dataset of a recent ad campaign targeting existing policy holders with a recommended insurance product. They would like to know whether the campaign was successful, who was most likely to respond to them and whether the success of the campaign can be predicted.

**To Do:**

**Data Model Implementation:**

* Create & train model via pyspark & associated libraries
* Reach meaningful predictive power (>75% classification accuracy, >80 R-squared)
* Create script that predicts whether the campaign will be effective on potential consumer

**Data Model Optimization:**

* Show results of hyperparameter optimization

**Visualisations:**

* Create visualisations that show success/failure per category
* Can I use Tableau to create visualisations, then save to image file for later use?

**Presentation:**

* Create webpage hosted on github pages with the following tabs:
  + Home: brief description of project & link to form where input values will be fed
  + Data: page that displays visuals based on category
  + Result: Displays yes/no as well as relevant visuals based on characteristics
* Powerpoint presentation (10min):
  + Introduction: Danny
  + Background: Danny
  + Model creation: Danny
  + Visuals: Danny
  + Demonstration: Danny
  + Conclusion: Danny
  + Sources: Danny
* Github:
  + Readme
  + Github pages

**Data:**

Columns:

* ID 8973
* City\_Code 8973
* Region\_Code 8973
* Accomodation\_Type 8973
* Reco\_Insurance\_Type 8973
* Upper\_Age 8973
* Lower\_Age 8973
* Is\_Spouse 8973
* Health Indicator 6873
* Holding\_Policy\_Duration 5743
* Holding\_Policy\_Type 5743
* Reco\_Policy\_Cat 8973
* Reco\_Policy\_Premium 8973
* Response

**Sources:**

* Description of dataset: <https://www.kaggle.com/code/manabendrarout/lets-find-a-job-eda-visualize-fe-ensemble/notebook>
* Dataset Source: <https://www.kaggle.com/datasets/imsparsh/jobathon-analytics-vidhya>