**Which topic did you choose to apply the data science methodology to? (2 marks)**

I choose to focus on the use of data science in the credit card industry for this assignment. This decision was made science it related to my father bussiness.

**Next, you will play the role of the client and the data scientist.**

**Using the topic that you selected, complete the Business Understanding stage by coming up with a problem that you would like to solve and phrasing it in the form of a question that you will use data to answer. (3 marks)**

You are required to:

1. Describe the problem, related to the topic you selected.
2. Phrase the problem as a question to be answered using data.

For example, using the food recipes use case discussed in the labs, the question that we defined was, "Can we automatically determine the cuisine of a given dish based on its ingredients?".

Therefore, bank's primary challenge with credit cards is that they must develop a model to determine who qualifies for the. some clients won't be viable because they lack the financial means to support this service. so, we would ask, "can we automatically access whether a client qualifiers for a credit card?"

**Briefly explain how you would complete each of the following stages for the problem that you described in the Business Understanding stage, so that you are ultimately able to answer the question that you came up with. (5 marks):**

1. Analytic Approach
2. Data Requirements
3. Data Collection
4. Data Understanding and Preparation
5. Modeling and Evaluation

You can always refer to the labs as a reference with describing how you would complete each stage for your problem.

Analytical strategy: Gives that the problem requires a binary( Yes/ No) response, our analytical strategy will involve applying a classification model designed to categorize data into distinct classes or categories, making them suitable for this type of problem. Data requirements: to build an effective classification model, we will require data on the bank's customers. This data should include personal information about the clients, such as demographics(age, gender, location), financial history(income, credit score), and any other relevant features. Additionally, it is crucial to have information on whether each client has defaulted on their credit card payments or if they have successfully made their payments. this labeled data will serve as the target variable for training the classification model. Data collection: In this phase, we will employ techniques such as descriptive statistics and data evaluation to ensure that we gather relevant and meaningful data for our model. descriptive statistics can help us understand the characteristics and distributions of the available data, while data evaluation involves assessing the quality, completeness, and suitability of the collected data. this evaluation process may include checking for missing values, handling outliers, and verifying that the collected data aligns with our requirements for the classification model. Data understanding and preparation: once the data is collected, we will proceed to analyze and prepare it for the classification model. This stage involves tasks such as data cleaning, where we address any inconsistencies, errors, or missing values in the dataset. additionally, we may perform, feature engineering, transforming, or creating new features that could enhance the predictive power of the model. Modeling and Evaluation: with the prepared data, we will select an appropriate classification algorithm and train the model. The dataset will be divided into training and testing subsets to access the model's performance. During the modeling phase, we may tine a hyperparameter, select relevant features, or employ techniques like cross-validation to improve the model's accuracy and generalization capabilities.