# Data Science Take-Home Task Instructions

## Task Summary

Featurespace has been approached by prospective customer in the banking sector. The customer is an issuing bank seeking a machine learning solution to solve a growing fraud problem. The bank’s customers are having their cards defrauded and their money is being spent by fraudsters. This is causing customer dissatisfaction, so the bank is looking into introducing better transactional monitoring on activity on their customers’ cards.

They plan to have a small team of fraud analysts who review risky-looking purchases and decide whether to allow or block the transaction. This team will have the capacity to review 400 transactions a month. The scores from your model will be used to decide which transactions the fraud analysts should review. The bank is requesting that after working these alerts, as much fraud value as possible has been prevented.

The bank has provided 1 year of historical transactional data and fraud flags and asked you build a model which predicts the likelihood that a transaction is later marked as fraud. They have also provided a brief data dictionary, which describes some general payments terms.

## Format of Report

Please return the code used to generate the model and a 2-slide presentation pitched at the bank’s executive board, which describes the performance your model achieved on an appropriately chosen test set, and the uplift this would provide the bank as a business.

Please perform this task without the use of a ‘notebook’, like Jupyter, and instead submit your work in the form of a script.

## Assessment

You will be assessed on the following skills:

* Using data in a sensible way to solve the business problem.
* Extracting useful signals from the data.
* Choice and application of algorithms.
* General coding quality.
* Clarity of communication.

You will *not* be assessed on the following skills:

* Performance of your classifier (past a low threshold).
* Choice of programming language.