# #DataWithDanny – Marketing Data Science Project

## Context

Congratulations – you have been hired by Banking Finance Global (BFG) to join the newly formed “AI Labs” team headed by Danny, the Chief Data Missionary ☺

BFG has been running a marketing campaigns to sell a Tesla Model X to their existing customers because Danny gets awesome kickbacks from Mr Elon for every car BFG sells.

In the past, campaigns were done by telephone calls only where Danny tried to hard sell existing banking customers the Model X by bribing them with a 10 BTC crypto cashback offer.

In this experiment, Danny selected a random sample of customers who were eligible to drive the Tesla (group A) – aka people who loved the environment ☺

From these environment lovers (group A) – a random subset of customers (Group B) were contacted by Danny because he only had so much BTC in his crypto-wallet.

At the end of the experiment – records were kept on how many of the total environment lovers actually bought a Tesla through BFG – whether they were bribed BTC bribes by Danny or not. This is the target variable in the provided dataset.

Assume that the price of the Model X was consistent across the entire experiment.

## The Assignment

Danny wants more kickbacks from Mr Elon so he wants to roll this experiment out to the entire environment lover customers. Obviously he wants the highest conversion rate but also wants to avoid bribing customers who would have purchased Tesla’s without his bribes too as he doesn’t want to waste his precious BTC.

Danny now asks you to help him use the machine learning to who to target to maximise the number of Tesla Model X’s sold but minimise the amount of BTC bribes he needs to give out.

Some questions he wants you to answer in your report include:

* Why did you choose your algorithm?
* What variables from the data will you use in your model?
* How did you setup your train/validation/test sets for the machine learning problem?
* How did you optimise the model? What metric did you optimise for?
* What are the top features?
* What validation steps did you take to make sure your model works?
* BONUS – Imagine that Danny wants to extend his call campaign to not just environment lovers, how would this change your existing solution (if any)?

# Data Dictionary

## Customer Data

|  |  |
| --- | --- |
| Banking Customer Data | Description |
| Age | How old is the customer |
| Job | What type of job did the customer have |
| Marital | Were they married |
| Education | What level of education did the customer have |
| Default | Has the customer default on a loan before |
| Housing | Does the customer have a housing loan |
| Loan | Does the customer have a personal loan |

## Campaign Data

|  |  |
| --- | --- |
| Campaign Data | Description |
| Campaign | Was the customer called by Danny or not |
| Contact | Which channel was used to contact the customer |
| Month | Month that the campaign was conducted |
| Day\_of\_week | Day of the week that the campaign was conducted |
| Duration | How long the phone call was in seconds |

## Other Data

|  |  |
| --- | --- |
| Random Data | Description |
| Previous | Number of times contacted before the actual BTC bribe campaign |
| Outcome | Did the previous campaigns work |
| BTC Spot Price | Periodic spot price |
| BTC Futures Index | Periodic futures index |
| Target Variable | Did the BTC bribe work and customer bought into the Tesla and helped Danny with his Elon kickbacks |