**AIRMILES DATA CHALLENGE**

At Airmiles, customers earn rewards Miles when they shop with one of our sponsors or use their Airmiles credit card. In 2017 there was a slight decline with Airmiles earned by collectors/customers. The objective is to predict which Collectors are going to decline their base earn rate in the next quarter.

**Challenge Overview:**

You will be receiving two data set on the day of the competition. The first one will be composed of 30,000 Collectors, with aggregated transactional information, demographic information and a flag indicating if that Collector declined in the following quarter (1 or 0).

You will use the “Data Training set” to create a model and predict which of the Collectors in the test dataset are going to decline. The “Data test Set” is composed of 10,000 Collectors with the same information as the training, except the flag. **Feel free to use any object oriented language (i.e R, Python, SAS, Excel, etc..).**

The participants are going to take 55 minutes to model and assign the flag to the test dataset. After the 55 minutes, all participants must submit their code and their completed test dataset. After all submission are done, we are going to provide the participants the results, so they can calculate what is the precision of their models. **All students will complete the Test file that we provide and send it back to us in the same format through drop box, we will evaluate and select a winner.**

**Evaluation Process:**

The precision score is going to be calculated by adding 10 points for each Collector correctly assigned as a decliner and by subtracting 3 point for each Collector incorrectly assigned as a decliner.

**Data**

**Decliner Definition:** Customers that have earned 50% less MILES in Q1 2017 vs Q1 2016. The decliner flag has already been incorporated, and there is no need to calculate this field.

**Variables:**

ID: Customer ID

Q2016 [XX] \_SPONSORS: Number of different retailers where the Customer have earned Miles in the quarter

Q2016 [XX] \_Miles: Number of Miles earned by the customer in the quarter

Flag: flag indicating whether the customer has declined (1 for decliner)

MILES2016 [XX] GROCERY: Miles earned at a Grocery in the Quarter

MILES2016 [XX] CREDIT\_CARD: Miles earned using the AIRMILES credit card in the quarter

MILES2016 [XX] OTHER: Miles earned at other retailers in the quarter

PROGRAM\_TIER: Tier indicating the level of the customer within the AIRMILES program (B: Blue, G: Gold, O: Onyx)

PROVINCE: Province where customer resides

EMAIL\_OPTIN: whether the customer has agreed to receive promotional emails

DEC\_2016\_ACCOUNT\_BALANCE: Number of MILES on the customer’s account last day of 2016

**Deliverable:** Participants need to provide a CSV file containing the CUSTOMER ID and the predicted decliner flag (1 for decliner, 0 for non-decliner). Example below:

|  |  |
| --- | --- |
| CUSTOMER\_ID | DECLINER\_FLAG |
| 123 | 0 |
| 456 | 1 |
| 789 | 0 |
| 101 | 0 |

**Final Score Calculation:** Final score calculation is going to be according the following formula:

Score = 10\*(correct guesses) – 3\*(incorrect guesses)

\*A guess is defined as a value of “1” on the DECLINER\_FLAG.

**DROP BOX LINK TO SUBMIT ENTRY:** <https://www.dropbox.com/sh/0i2vwdo1lupq661/AABJIwoNZqy0McgMNJOAJa0Qa?dl=0>