The Precognition Store

# Regression Problem, 100pts

In the time of digital commerce, the best way to assure survival against competition; e.g., Amazon, Walmart, it is necessary to be one step ahead and understand your customer.

As a data scientist, you should predict the rating given to a set of products by a given number of customers.

OBJECTIVE:

Define an algorithm that will correctly predict the rating of a product given by a customer.

TRAIN SET:

There are 800,000 ratings coming from 6000 shoppers and 4000 products.

TEST SET:

The test set consists of 50,000 combinations of shopper-product that have not been rated products.

LINK TO DATA:

DELIVERABLES:

* Summary (must): Explain your approach to solve the problem to any hackathon supporter of the Data Science challenge.
* Regression: Share your regression results of the test set. Include only the ID\_PRODUCT, ID\_CUSTOMER and RATING columns. You will get back your R^2 score\*

**\*** Your R^2 score represents the number of points you will obtain. Zero being 0 points and One being 100 points. You can share your regression as many times as you wish, your highest value will be reported as the number of points obtained.