



BCN Analytics & Social Point 2nd Data Hackathon

October 2016



The Challenge

THE CHALLENGE

➤ The challenge

- The goal of this analysis is to predict the monster (1 monster) the user will like and doesn't already have.

➤ The data

- The data you have contains information of Monster Legends players, the monsters they have and also some features of those monsters.
- You would be provided with two datasets: the training set (users_train.csv) and the test set (users_test.csv).
 - Training set: contains the full register of monsters of the users.
 - Test set: contains partial register of monsters of the users (one monster has been randomly deleted).
- Additionally, two more datasets are provided with detailed information of the users and the features of each monster.

IMPORTANT NOTES ON THE GAME

- **The game is free-to-play**
 - There are no required purchases, so if a user has paid in the game, that is a voluntary payment
- **The first level of the game is tutorial**
 - The user follows the instructions of 'Pandalf'.
- **The purpose of the game is to collect Monsters and use them to fight**
 - You obtain a monster by breeding two monster or directly by buying the monster in-game with gems. For this analysis, we only took into account the monsters bought in the store with their full price (without any offer).
 - There are three different battle modes: Single player mode and dungeons (spm), player vs player (pvp), and wars.
- **The game is multi-platform**
 - That means you can play the game on an Apple device or Android or Facebook, or all three if you wish
- Here is a short video of the [game](#)

IMPORTANT NOTES ON THE DATA

- **There are four different csv files. Those files contain information about the users and monsters:**
 - users_train: complete register of which monster has each user.
 - users_test: partial register of which monster has each user.
 - Table 2_Monsters: information about the monsters.
 - Table 3_Users: information about the users.

User_Train and User_Test

- The following variables are in the train and test files:
 - **user_id**: a unique identifier for each user.
 - **monster_id**: a unique identifier for each monster.
 - **count_monster**: the number of times the users has that monster.

TABLE 2_Monsters Information

➤ The following variables are in the dataset:

- **monster_id**: a unique identifier for each monster.
- **monster_name**: name of the monster
- **Attributes**: the elements of the monster. The monster can have more than one element and there are 10 different elements:
 - f= fire, e= earth, w = water, mt = Metal, d = Dark, m = Magic, n = Nature, t = Electric, l = light, e = legendary
- **Rarity**: How rare is the monster. It ranges from 1 to 5.
- **Breedable**: This variable indicates if the monster can be obtained by breeding two monsters.
- **Price**: price of the monster in the store in gems.
- **Lvl**: the user needs to be in this level (or higher ones) to be able to buy the monster in the store.

TABLE 3_Users Information

- The following variables are in the dataset:
- **user_id:** a unique identifier for each Monster Legends user.
 - **level:** current level of the user in the game.
 - **country:** the two digit country code of the user.
 - **date_register:** the first time the user downloaded the game (this could be on either an Apple or Android phone, or on Facebook).
 - **days_since_register_playing:** the first time the user downloaded the game onto an Apple device.
 - **sex:** sex of the user.
 - **fb_link:** indicates if we have their facebook link.

TABLE 3_Users Information

- The following variables are in the dataset (Continued):
- **register_platform:** this variable tells us the platform where the users first downloaded they game. It can be 'android', 'ios' or 'canvas' (facebook).
 - **payer_category:** this variable indicates how much money the user has spent in the game. '0' is a non-payer user, '1' is a low spender user and '2' is a high spender user.
 - **num_battles_spm:** total number of single player or dungeon battles played.
 - **num_battles_spm_won:** number of single player or dungeon battles won.
 - **num_battles_pvp:** total number of player vs player battles played.
 - **num_battles_pvp_won:** number of player vs player battles won.