



## KABAM / SAUDER VARSITY CHALLENGE 2019

### Submissions guideline

#### A MESSAGE FROM KABAM

Our initial intent for this challenge was to provide a real world experience for students to develop their skills. It is not intended to be solely a coding exercise. Kabam partnered with Sauder Business School because in addition to the data analytics component this challenge reflects a business problem that will require quantitative thinking as well as business acumen and communication. Data analysts encounter such situations regularly.

The challenge was presented as an open management question: to devise a model that predicts the value of a user. At the kick-off presentation we discussed what is value and what are our challenges. A valuable player is one is engaged with game, for an engaged player will be an advocate to attract more players, and a strong community is a key to success for a game. Then, a player who makes an in-app purchase is valuable because as a business we require revenue to be a sustainable entity that can continue entertaining our fans.

We did not want to stifle creativity by being overly prescriptive.

Yet, feedback indicates that participants would appreciate more prescribed direction, so we have pivoted slightly. There is still a subjective element to the challenge and an opportunity to distinguish your team by being creative, but we have clearly identified objective measures for scoring.

#### PROCESS

The competition culminates in the Sauder Hackathon held Monday April 8 to Tuesday April 9. Kabam will have a team on site to conduct scoring.

However, teams may send first submission prior to Monday April 8 so we can start the Hackathon with a leaderboard so teams can start the Hackathon with the goal of improving their submissions. Please submit only one submission prior to 9AM on Friday, April 5 to allow our team sufficient time to score. Submissions after 9AM Friday, April 5 may be scored at the Hackathon.

Please send your submissions to [varsitychallenge@kabaminc.com](mailto:varsitychallenge@kabaminc.com) . See below for details of submission content.



Kabam' team will score submissions during the competition. However, they may conduct scoring intermittently if they have to tend to corporate tasks. Though, we plan to block time to score as follows:

**Monday April 8**

12:00PM

3:00PM

**Tuesday, April 9**

12:00PM

5:00PM

Kabam will maintain a leaderboard that will show the top 5 teams. The leaderboard will be revealed on Monday April 8 at 9AM, the start of the Hackathon. This leaderboard is not automatic, but manually updated by Kabam when there is a change to the standings. It will include a timestamp indicating the last update.

Due to the volume of assessments and the potential for teams to conduct trial-and-error to achieve optimization, the number of submissions is limited to three per team over the course of the Hackathon. We will only accept one submission at a time, and the highest scoring submission will be the team's score.

The last submission is accepted at 5PM on Tuesday April 9. While we hope to announce the finalists at the after-party, if the volume of submissions prohibits doing so we will announce via Slack on Wednesday, April 10.

## **SUBMISSIONS REQUIREMENT**

Contestants will predict the quality of players of Kabam games.

To quantify the quality of a given player, it is necessary to define a metric. For the purpose of this competition, this metric is called the player quality score (PQS). Previously, Kabam asked the contestant to create their own PQS based on the data that has been released. Given the feedback, we have decided to provide more guidance on defining PQS.

In this competition, PQS is a metric based on **lt\_spend**, **total\_logins**, **power** and **num\_friends**.

For the purpose of this challenge, teams will predict the values of these four variables as closely as possible. Then, based upon these four variables, the contestant will then design, define, and calculate their version of PQS.

We expect that teams will consider third-party data sources. If you use a third-party data source it needs to be available to all contestants. Contact [stapp@kabaminc.com](mailto:stapp@kabaminc.com) and Kabam will ensure it is communicated to all



teams via Slack.

For the Varsity Challenge, the following outputs are expected:

- R or Python scripts used to produce the predictions in this submission.
  - For Python users, please use Python 3.6.5 and be prepared to provide a requirements.txt for packages used. For R users, please be prepared to provide self-contained scripts that will install any necessary packages. Example format of the submission will be provided at a later time.
  - We will be running your scripts against our local copy of data. So, you should setup the code such that it reads the Kabam data in from a local directory.
- A comma separated file (CSV) containing predictions of lt\_spend, total\_logins, power and num\_friends as well as your version of the PQS.
  - The submitted CSV file will be in the following format. There will be six columns: uid\_i, lt\_spend, total\_logins, power and num\_friends and PQS. Each row of file will consist of predictions corresponding to the uid\_i and the file will be sorted based on uid\_i.
- A document explaining your rationale behind PQS design.
  - In a Google Document no longer than 300 words, please detail the rationale behind the design of your PQS score. In your email submission, please provide a link to the document, and ensure it is accessible by our team.

## SCORING

Submissions are scored on objective components as well as subjective components.

### ***Scripts***

Completion and code quality.

### ***Output (CSV)***

Based on the accuracy of the predicted variables: lt\_spend, total\_logins, power and num\_friends.

### ***Design rationale***

Based upon the PQS and supporting documents. Utility and creativity are the primary criteria.



**ADVICE**

Plan how you will use your three submissions, as you may only have one submission in the queue at a time.

Remember, data analytics is not just about discovery of systemic patterns in data. A data analyst must be able to put these discoveries in an operational context for it to be useful information. So, the objective score is not a sole determinant of assessment.

For this reason, we require not only the model but a written document explaining the decisions made. And finalists will make a presentation to the judges panel to explain the business context and logic of the model.