**Paper 3885-2019**

# Using Simulation Games to Teach Analytics

Jean-François Plante, HEC Montréal; Pierre-Majorique Léger, HEC Montréal;
Jean-François Michon, HEC Montréal; Forough Karimi-Alaghehband, HEC Montréal;
Marc Fredette, HEC Montréal; Michael Bliemel, Ontario Tech University

## ABSTRACT

Literature shows that gamification and experiential learning improve the attitude, retention and engagement of students. We describe and showcase a simulation game developed in the field of analytics that uses SAS Enterprise Miner®. The game scenario puts the students in charge of targeted marketing actions for the fundraising campaign of a charitable foundation. Students have to decide which members to contact for a possible donation based on data from one million members. The game is designed to help participants apply and consolidate different data analytics concepts in a realistic setting. It provides pedagogical flexibility so that it could be used in classes of different levels (e.g., UG, MBA) and with students with varying levels of statistical knowledge and competency.

## INTRODUCTION

Gamification has been defined as "the use of game design elements in non-game contexts" (Deterding et al., 2011, p.10) with the main objective being to increase students' engagement. A systematic literature review by Subhash and Cudney (2018) shows that "improved student attitude, engagement, and performance were observed as the most significant benefits of gamification and game-based learning in higher education" (p.204).

While serious games are becoming widespread tools for learning and teaching various business and management concepts (Legner et al., 2013), developing and accessing realistic, engaging, and dynamic games remain a challenge for many schools and academic departments (Léger et al., 2014). While management fields such as business process management, project management and supply chain have been the main topics for game development and use for the past decades (e.g., ERPsim, Albasim, Beer game), more recent fields are yet to benefit the same level of focus. One such field is business analytics.

This article presents a simulation game developed in the field of analytics that could be used across different levels (e.g., Undergraduate, Masters, MBA) with participants with varying levels of statistical knowledge and competency. The simulation game is designed to help participants apply and consolidate different data analytics concepts in a realistic setting. It also helps them develop skills by gaining hands-on experience with tools commonly used in the industry. We describe a two-round version of the game that is suitable for undergraduate and MBA level classes where technical training is not one of the main objectives of the course.

The simulation game has been tested in a North American business school and obtained the following main outcomes: i) students' engagement due to the dynamism and competitive nature of the game, ii) interaction among students, iii) desire to know more and to explore different options to perform better in the game.

## THE GAME

SAS® has partnered with expert game developer from ERPsim Lab at HEC Montréal (https://erpsim.hec.ca/) to create an Analytics Simulation Game. Much like an airline flight simulator, the objective of this game is to provide an active learning opportunity for aspiring data scientists to learn and apply analytics.

Using SAS Enterprise Miner® students work with a special simulated data set – derived from US Census Data and other sources – which features one million members of a major charity. Students are placed in the role of a data scientist for this charity, where they must unlock the value in members' data to target which individuals should be called to increase the funds raised for the charity. They are provided with a cost schedule and must decide how many and which members to call.

Student submit their decisions to a live web-based scoreboard (as shown in Figure 1), where they can see how much they raised with their models and how they compare with other classmates. Students can use SAS Enterprise Miner via an AWS cloud instance, on their own PCs or laboratories. The SAS Global Forum will mark the official release of the game in time for professors and educators to use it in their spring or fall 2019 curriculum.

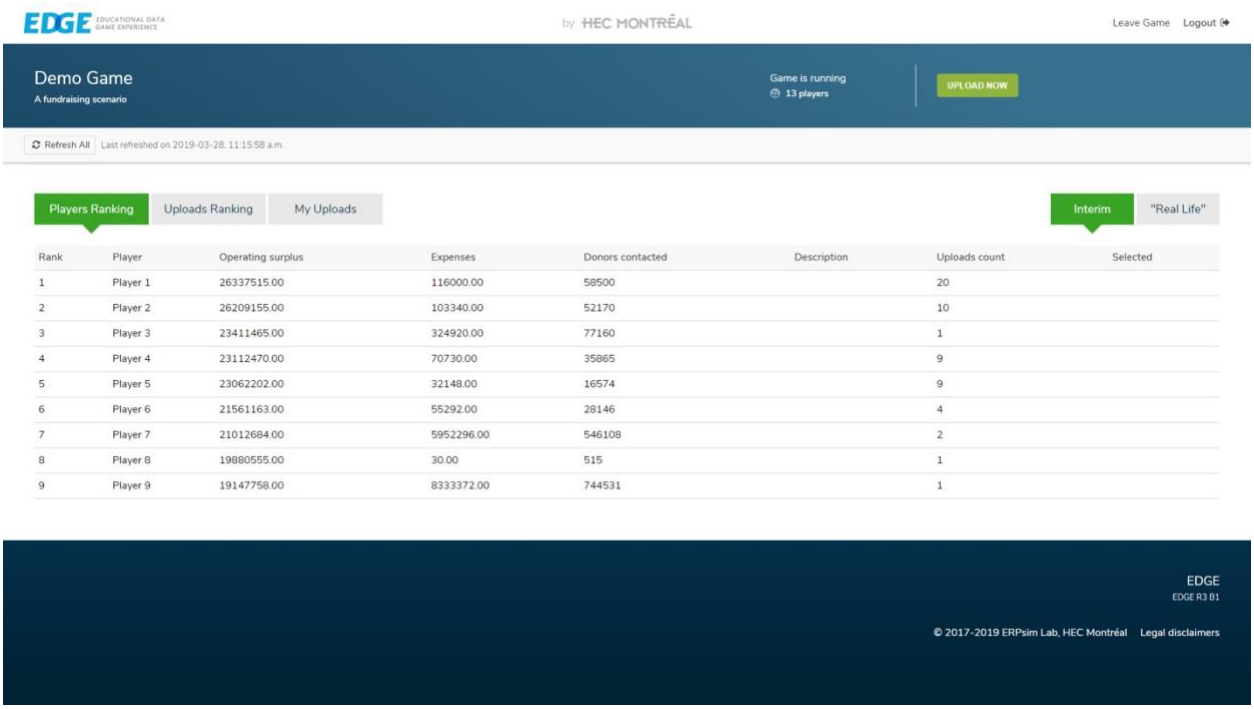



**Figure 1. The Live Web-based Scoreboard Interface**

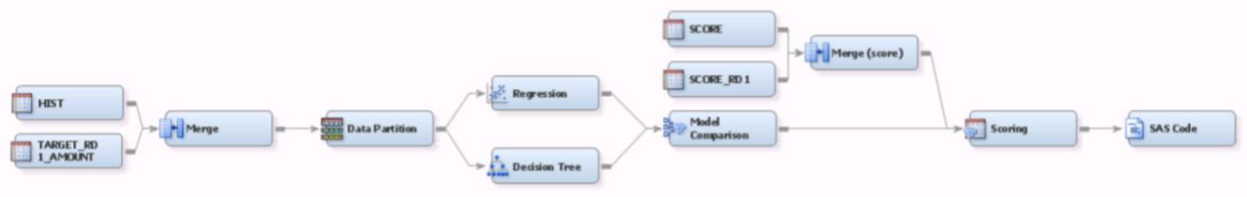
A two-round version of the game is designed to provide experimental learning to students in a class where they learn about using predictive analytics without focusing on technical details. To achieve this, we built SAS Enterprise Miner diagrams in advance such that the students obtain predictions within a few minutes. Their role is then to try new models, explore the data, while discovering the possibilities of SAS Enterprise Miner. In the first round, students predict the amount each individual may give using a model that may not be well specified yet provides value. In the second round, we introduce two-stage modelling as well as uplift modelling (e.g., Hansotia and Rukstales, 2002) and provide prebuilt diagrams

that allow students to estimate the expected amount they may receive from each individual if they contact them and if they do not.

## ROUND 1

Display 1 shows the default diagram that allows students to predict the donation amount. All data sources and libraries are pre-defined; hence students do not need to go through these tedious steps initially. The diagram includes Regression and Decision tree as two initial models. Students are encouraged to improve the two initial models by modifying the parameters of the nodes, and to add more models to the diagram to improve the predictions. A written document summarizing useful nodes of SAS Enterprise Miner is provided to guide them.

The output predictions are written in an Excel file from which the students can select the ID of the members that must be called. This decision is saved as a csv file and uploaded on the game leaderboard for immediate scoring and feedback.

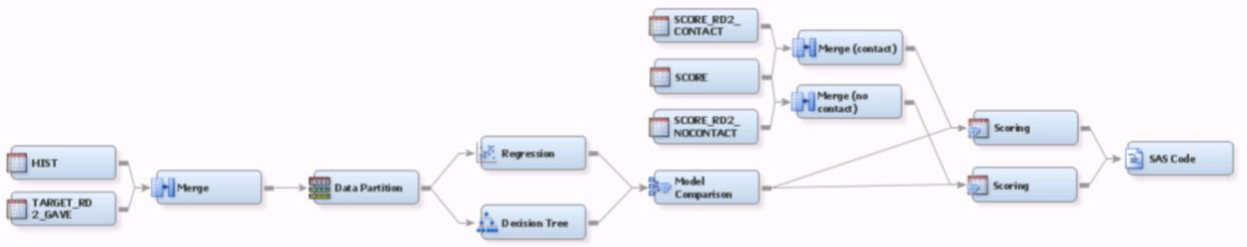



**Display 1. Pre-built SAS Diagram for Round 1 – Predict Amount**

## ROUND 2

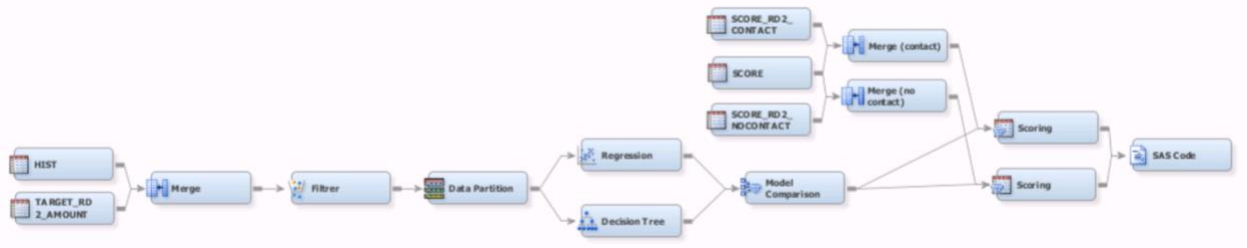
Round 2 presents two new elements: two-stage modelling and uplift modelling. There is in fact two diagrams for Round 2, one that gives the probability that a member will give (see Display 2), and a second one that provides the conditional amount, given that a member gives (see Display 3). Multiplying those two values yields the expected donation for each member.

Uplift modelling implies some planning in data collection where a randomly selected group of people have received the treatment (here a phone call). The variable "contact" indicates for each individual if they were contacted or not. There are many modelling strategies for uplift modelling, but we put forward a simple solution where the selected model is evaluated with contact=0 and contact=1.




**Display 2. Pre-built SAS Diagram for Round 2 (Stage 1) – Probability of Giving**

Note that the two diagrams in Round 2 look similar, but the second diagram uses a Filter node to select all individuals who made donations in the most recent year. As for Round 1, predictions are written to Excel files, one per diagram. Students need to combine the outputs from the two diagrams and compute the expected values and uplift. This reinforces their understanding of the required operations. The decision of how many people to call is also simplified as the students now have an estimate of the gain from calling each member, hence having a measure of the value of their action. Again, the decision is to be uploaded as a one-column csv file on the game's platform.




**Display 3. Pre-built SAS Diagram for Round 2 (Stage 2) – Conditional Amount**

## PEDOGOGICAL NOTES

### A FELIXIBLE GAME

We described a version of the game suitable for introductory types of courses. However, the game's scenario allows for pedagogical flexibility so that instructors could adjust the game's level of difficulty according to the specific needs/competency of their students. For participants with higher levels of statistical competencies, such as master students in data science, the instructor could use the data and the game platform, but decide to provide no pre-built diagrams, possibly no hint as to what variable should be predicted and let the students figure it out. Although we describe a two-round version, the game could be played with more rounds, each lasting one week or two, for instance, where students can try the new modelling techniques that they acquired in class through an academic term.

There are several models available in SAS Enterprise Miner that students could use during their analysis such as regression, decision trees, random forests, neural networks, and gradient boosting. Depending on the course's length, learning objectives, as well as students' competency in statistics, instructors could decide on the number of models to introduce or the depth of the training provided for each model.

The game is accompanied by pedagogical material including participants guide, teaching notes and instructional videos.

### A GAMIFIED AND EXPERIENTIAL APPROACH

The benefits of gamification are well documented. Figure 2 demonstrates the most cited benefits of gamification and game-based learning according to the systematic literature review by Subhash and Cudney (2018).

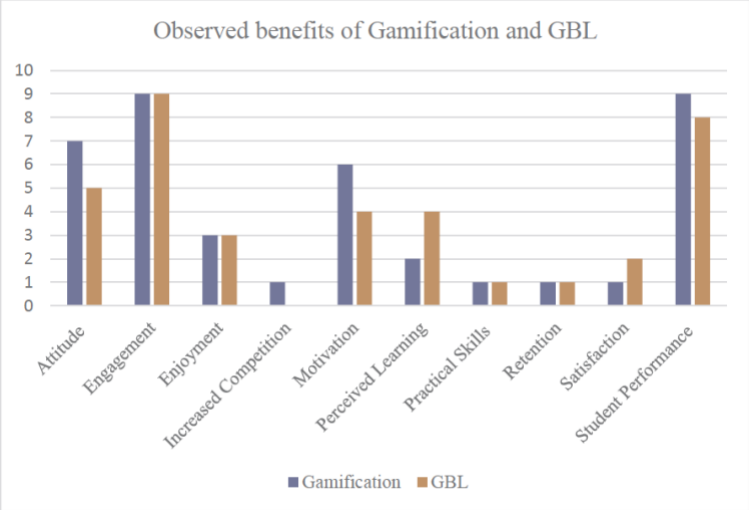



**Figure 2. Benefits of Gamification and Game-based Learning (Adopted from Subhash and Cudney 2018, p.205)**

The same study also highlights the importance of "points, badges, leaderboards, levels, missions/quests, and feedback were identified through the systematic literature review as important game elements" (p.205).

Reexamining the elements of the developed game, 'Leaderboards' and 'Feedback' are the two main elements embedded in the game, which contributes to students' engagement and eventually their performance. The game has two leaderboards: Interim and Real-Life. The Interim leaderboards provide immediate feedback and ranks all the players based on their uploads –each player is allowed multiple uploads. Students must select one of their uploads as their final decision. The Real-Life leaderboard prevails for the results of the game. Having an alternative, hidden version of the results provides an opportunity to discuss overfitting further and prevents rewarding it.

An experiential approach defines learning as "the process whereby knowledge is created through the transformation of experience" (Kolb 2014, p.49). We believe that the developed game encourages an experiential learning cycle as shown in Figure 3.

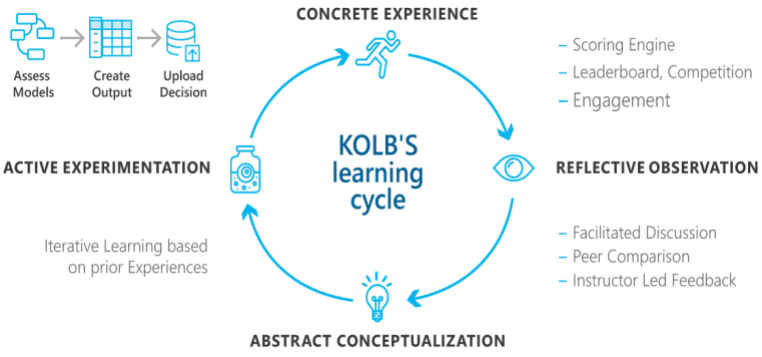



**Figure 3. The Game and Kolb's Learning Cycle (Based on Kolb 2014, p.50)**

The game's iterative nature contributes to the learning cycle of grasping – i.e., concrete experience and conceptualization - and transformation – i.e., observation and experimentation. The game's design allows students to try first and immediately engage with the game– i.e., to grasp concrete experience - and improve their models and performance in later iterations. The leaderboard rankings paired with class discussion and debriefing facilitates reflective observation. This is where students realize why they performed the way they performed, and they develop ideas – i.e., abstract conceptualization. They can now apply the new ideas and their experiential learning in the next iteration of assessing/improving models, creating outputs, and uploading decisions.

## CONCLUSION

ERPsim Lab has been creating and promoting gamification for the last 15 years, with proved and documented benefits (see e.g., Labonte-LeMoyne et al., 2017). The purpose of this simulation is to allow participants to understand and apply key concepts of data analytics. The game reinforces theoretical concepts (e.g., overfitting, fair assessment of models) by giving participants the opportunity to apply their knowledge in a practical scenario of testing, validating and refining techniques to solve a realistic business problem. In doing so, the simulation evaluates the effectiveness of data modelling decisions made by the participants comparing to other participants in the class.

The simulation engine allows the participants and instructors to immediately observe the outcome that would otherwise have taken weeks to realize. This immediate feedback feature facilitates participants learning. It also enhances instructors' ability to coach their participants.

The two-stage version of the game that we described in this paper was successfully used in introductory level classes and allowed students to experience predictive analytics. Although some of them had no quantitative background, they could use ensemble models and neural networks to make informed real-life decisions. They also got to understand practical constraints, for instance that a complicated model might take hours to run. Most importantly, the gamification and the experiential learning improved the learning outcomes of the students.

## REFERENCES – UPDATE


Deterding, S., Dixon, D., Khaled, R., and Nacke, L. 2011. "From game design elements to gamefulness: Defining gamification". *MindTrek 11 proceedings of the 15th international academic MindTrek conference: Envisioning future media environments:* 9–15

Kolb, D.A. 2014, *Experiential Learning: Experience as the Source of Learning and Development*, 2nd ed. NJ: Pearson Education LTD

Hansotia, B. and Rukstales, B. 2002, "Incremental value modeling," *Journal of Interactive Marketing*, (16:3): 35-46.

Labonte-LeMoyne, E., Léger, P.M., Robert, J., Babin, G., Charland, P., and Michon, J.F. 2017 "Business intelligence serious game participatory development: lessons from ERPsim for big data," *Business Process Management Journal* (23:3): 493-505.

Léger, P.M., Cronan, T.P., Charland, P., Pellerin, R., Babin, G., and Robert, J. 2012. "Authentic OM problem solving in ERP context," *International Journal of Production and Operations Management* (32:12): 1375-1394.

Legner C., Estier T., Avdiji H., and Boillat, T. 2013 "Designing Capstone Courses in Management Education: Knowledge Activation and Integration Using an ERP-based

Simulation Game." In *Proceedings of the 34th International Conference on Information Systems*, Baskerville R, Chau M (eds.), Milan, Italy: 1–19.

Subhash, S. & Cudney, E.A. 2018, "Gamified learning in higher education: A systematic review of the literature", *Computers in Human Behavior,* 87: 192-206.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Jean-Francois Plante
HEC Montreal
jfplante@hec.ca