Gamification

Claire Kelling, Nikolas Siapoutis, Xiufan Yu, Aniruddha Rao

Nov 27, 2017

Introduction

Our Client:

- Yu Yan, PhD Candidate
- Department of Education
- STAT 501
- Data analysis stage of her study
- Interested in "gamification of education"

Introduction

Study details:

- interested to see whether gamification has a positive effect on student engagement in learning English
- conducted on students who took the English course where they had the option to use the English assessment system and were assigned to either the control group (no gamification) or treatment group (gamification)
- conducted over a period of 6 weeks, at the end of which they filled out a survey

Research Quesitons

There are two main parts to her Research Questions:

1. **Database part:** Is there any difference between the values of total time spent on the system (or total number of assessments or average time or average assessment) between the control and treatment group?
2. **Survey part:** Are students of the treatment group (gamification) more satisfied or more motivated than those that are not? How is their achievement/goal orientation different across groups (what are they motivated by)?

Data Description

Database:

- Various measures of engagement for each unique Student ID (340 students)
  - Total Time
  - Total Assessment
  - Average Assessment
  - Stay_30m (corrected average time)
- Gender
- Class and teacher ID
- Game (indicator)

Survey:

- Measures of motivation
  - MastApp (Mastery Approach)
  - MastAvo (Mastery Avoidance)
  - PerfApp (Performance Approach)
  - PerfAvo (Performance Avoidance)
- Questions about satisfaction
- Student ID (to pair with original dataset)

EDA-Original Data

We now have 340 samples, while 167 are in the gamification group and the rest 173 are in non-gamification (control) group.

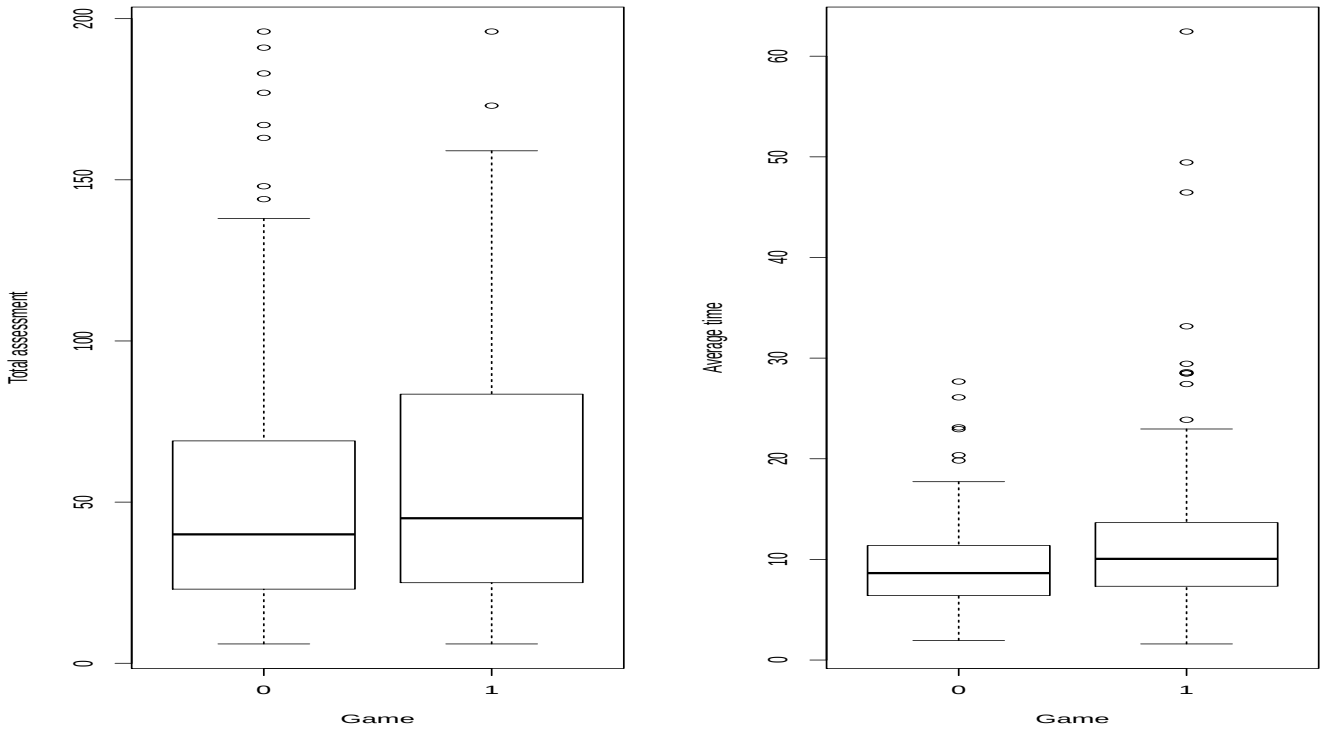

Figure: Boxplot for Total assessment and Average time vs Game

EDA-Survey Data

276 samples: 134 (gamification group) + 142 (non-gamification)

- 13 students: ONLY the evaluation questions, NO motivation info.
- a lot of missing values for problem #10 (118 out of 276)

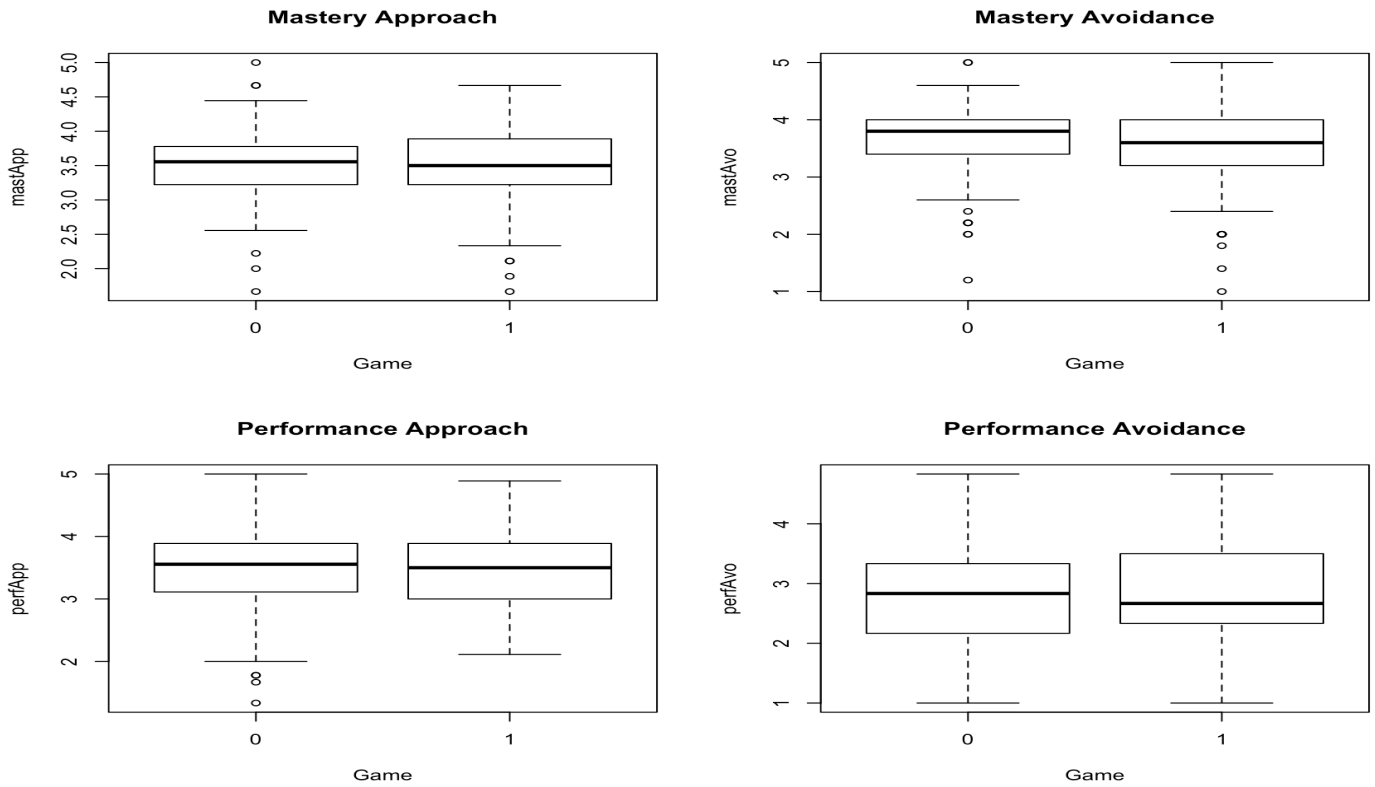

Figure: Boxplots for four aggregated measure variables: mastApp, mastAvo, perfApp, perfAvo

EDA-Survey Data

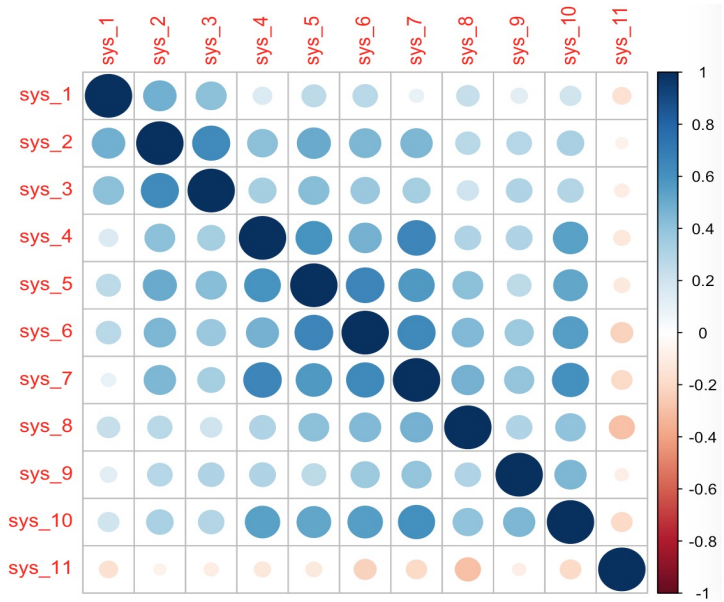

Figure: correlation plot for the eleven survey questions

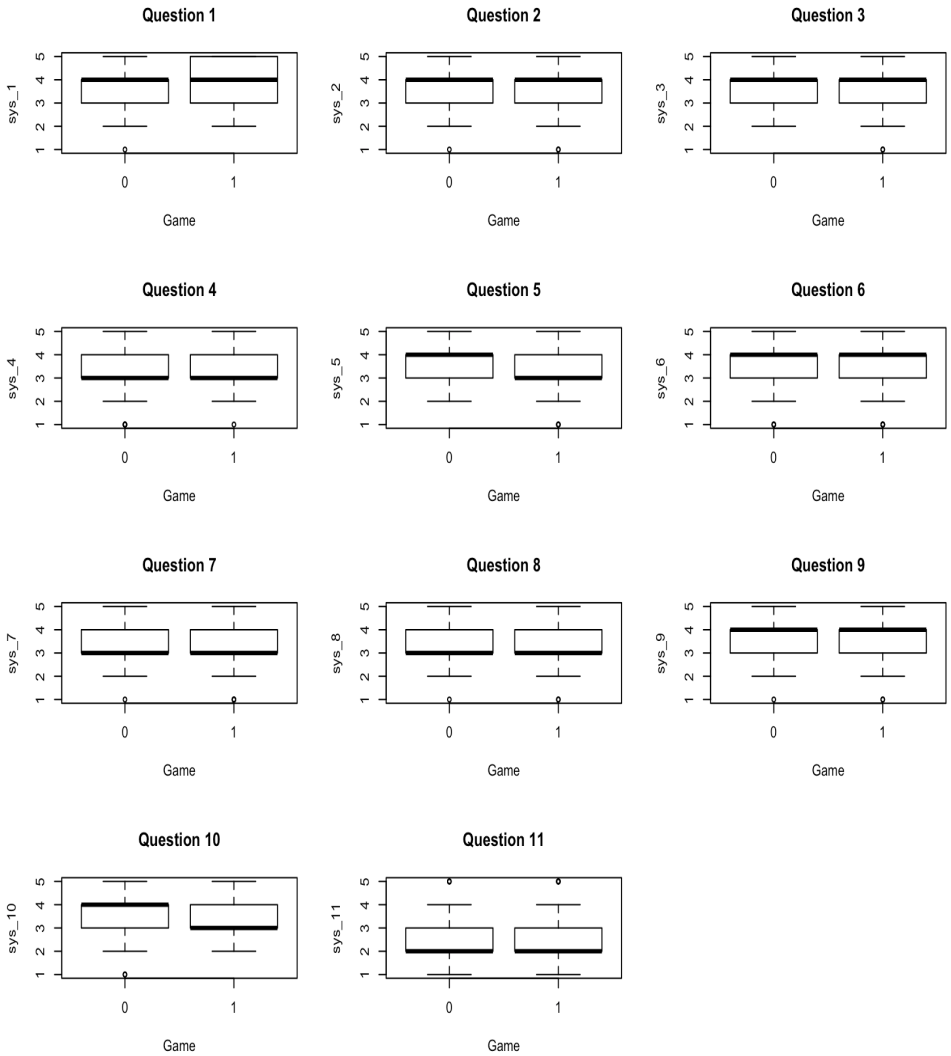

Figure: Boxplots for system questions

EDA-Missing Data

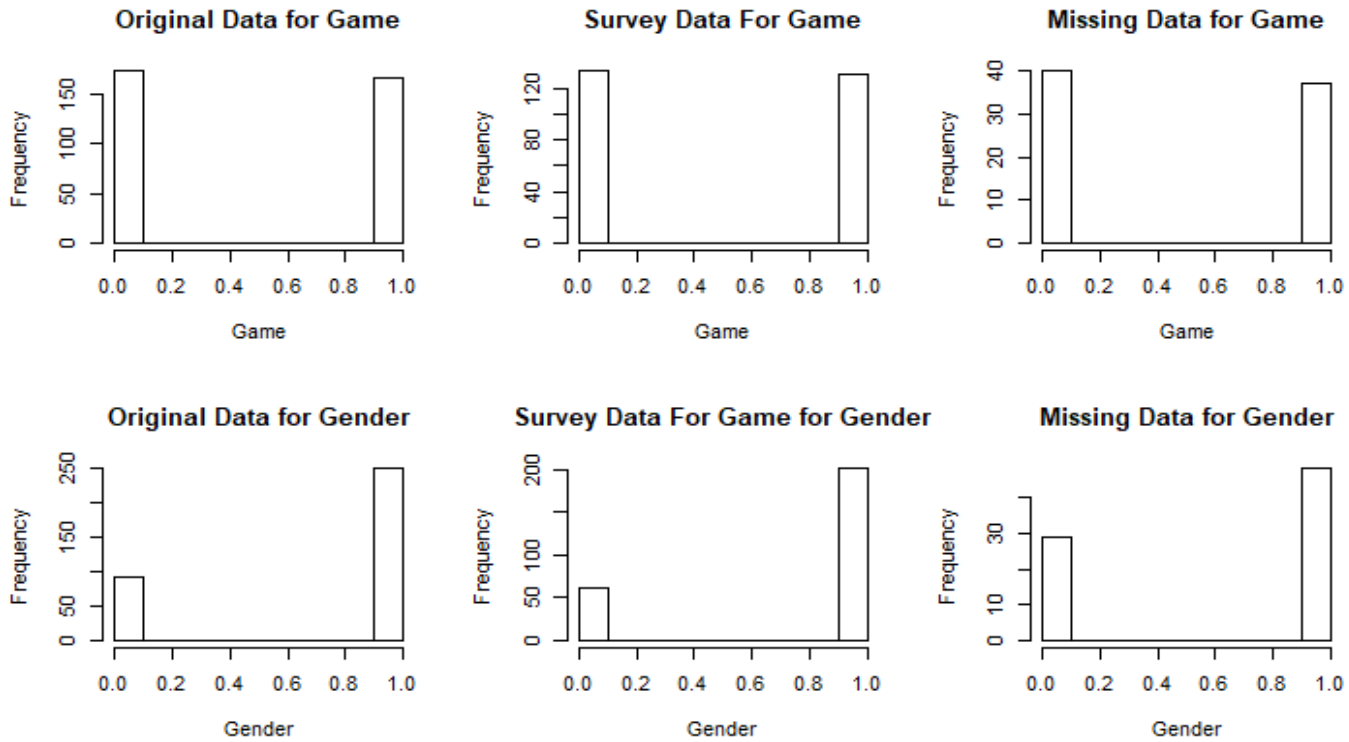

Figure: Histogram of Game and Gender on different Data set

Research Question 1

Is there a difference in engagement between groups?

Engagement

- We want to test the hypothesis:
  $H_0 : \mu_{TA,0} = \mu_{TA,1}$ $H_1 : \mu_{TA,0} < \mu_{TA,1}$
  where μ_{TA} is the average of total assessments spend on the system.
  $H_0 : \mu_{AT,0} = \mu_{AT,1}$ $H_1 : \mu_{AT,0} < \mu_{AT,1}$
  where μ_{AT} is the average of the average time spend on the system.
- Applying the **two sample t-test for heterogeneous variance**, we can conclude:

Table: Two sample t-test for heterogeneous variance

Variable	**p-value**	**Conclusion**
Total number of assessments	0.1260	Same
Average time	0.0003217	Treatment Higher

Engagement

- If the normality assumption made in t-test is in doubt, the nonparametric **Wilcoxon-Mann-Whitney test** is sometimes suggested. Applying this test in our data, we get the same results as the two sample t-test.
- We also check if there are **outliers** and if these affect our results. Without the outliers,the treatment group averages are greater than the control group for all the variables of interest.
- Similarly, t test can be done for variables like total time, average assessment, total number of visits, total course, max level (final level).

Research Question 2

Is there a difference in satisfaction between groups? What role does motivation play?

Satisfaction

First, consider the eleven questions together

- multivariate two sample t-test: p-value=0.036
- cronbach's alpha = 0.731

Next, consider the eleven questions separately

- run three t-tests for each question

Table: Alternative Hypotheses for three t-tests against $H_0 : \mu_{0,k} = \mu_{1,k},\ k = 1, 2, \cdots, 11$

Test 1: t-test (two.sided)	$H_1 : \mu_{0,k} \neq \mu_{1,k}$
Test 2: t-test (less)	$H_1 : \mu_{0,k} < \mu_{1,k}$
Test 3: t-test (greater)	$H_1 : \mu_{0,k} > \mu_{1,k}$

Satisfaction - Individually t-test

Table: Summary for individually t-test on 11 questions

Question(k)	1	2	3	4	5	6	7	8	9	10	11
mean-diff $(\mu_{0,k} - \mu_{1,k})$	-0.23	-0.12	-0.08	-0.03	0.14	-0.01	0.08	-0.07	0.15	0.11	-0.13
p-value (two.sided)	0.027	0.254	0.440	0.788	0.177	0.876	0.483	0.450	0.184	0.456	0.250
p-value (less)	0.014	0.127	0.220	0.394	0.911	0.438	0.758	0.225	0.908	0.772	0.125
p-value (greater)	0.986	0.873	0.780	0.606	0.089	0.562	0.242	0.775	0.092	0.228	0.875

Motivation

Research Question: What role does motivation play in determining engagement?

- Are certain kinds of students more apt to engage because of motivation?
- For example, are those that are highly motivated by exam scores more likely to be engaged in the system?

Approach:

- Include motivation measures as regression covariates
- Include interaction terms between game and motivation covariates

Motivation- ANOVA

Including outliers/Excluding outliers		Engagement (Dependent) Variable	
Independent Variable		Stay_30m (Average Time)	Total Assessment
	mastApp		
	mastAvo		0.003 / <0.001
	perfApp	0.008 / 0.023	
	perfAvo		
	teacher	<0.001 / <0.001	<0.001 / <0.001
	female		
	maxLevel	<0.001 / <0.001	<0.001 / <0.001
	game	0.038 / NA	

Figure: ANOVA fits for various measures of engagement, including significant variables

We see that some kinds of motivation are important, including mastAvo for Total Assessment, and perfApp when determining Average Time.

Response Variable: Final Exam Score

Table: Variable and Coefficient Estimate

Intercept	stay_30m	mastApp	mastAvo	perfApp	perfAvo	female	maxLevel
59.073	-0.102	2.838	-0.914	0.779	-1.010	6.521	1.427

Mixed Effects Model:

- Random effect for teacher, fixed effect for all others
- Tried various combinations but overall adding game into the model was not a significant positive contribution (LRT)

Recommendation

- Research Question 1:
  - Gamification helps in increasing the engagement of students in terms of Average time but not in terms of Total assessment.
- Research Question 2:
  - There exists significant difference on the satisfaction between gamification group and non-gamification group, especially for the survey question 1.
  - The performance approach and mastery avoidance are the two most important types of motivations when it comes to engagement with this online system.
  - Gamification does not change course performance.

Consideration

- The results after removing outliers are sometimes different
- The distribution of the Original Data and Survey Data is similar but there is variation wrt the Missing Data.
- Though the eleven survey questions are correlated, their cronbach's alpha value does not achieve the criterion, so we consider them separately rather than combine them by taking the average.

Thank you!
Questions?