Game Analytics: User Engagement and Performance

DGSA-1007 Final Presentation

Joseph Stansil, Doma Ghale, Jing Zhang, Chinelo Ibekwe

Updates and Outline

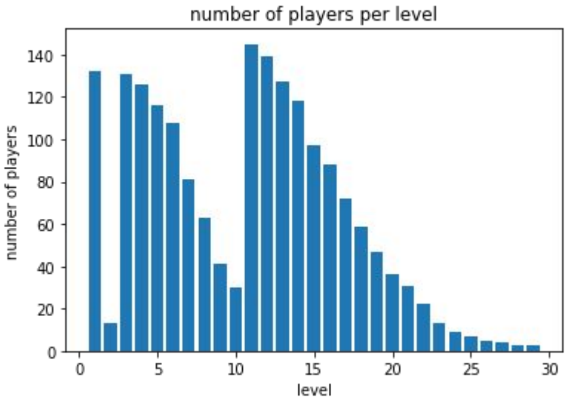
- Data comes from an experiment conducted in the fall of 2018 by the NYU CREATE Lab
    - Designed a game with the goal of training players' cognitive flexibility
- Our goals for the project included:
    - Addressing data formatting inconsistencies by cleaning, wrangling & merging
    - Assess player engagement
    - Assess player performance

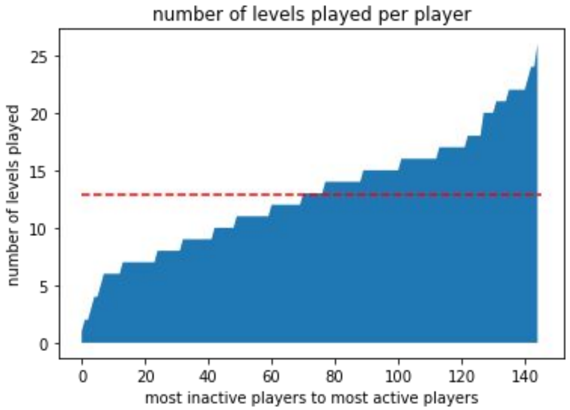
Data Preprocessing

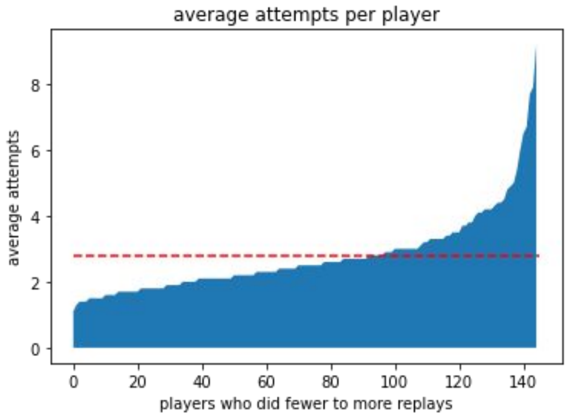
- **About the Data**
  - Consists of log data pulled from the dates of the experiment.
  - Timestamps corresponding to when the game server received input
  - Rows represent discrete events, i.e. hitting an alien
  - Example columns: reaction time, alien type, hit type, level, attempt, timestamp
  - 20 files w/ 5 unique variable sets
  - 1200k + rows, 58 columns (200k + rows, 29 columns after processing)
- **Challenges**
  - Separate files required merging
  - Duplicate records. Some events seem to have been logged twice. Need to find the correct key.
  - Inaccurate and inconsistent timestamps
  - Multiple keys including session ID, user ID, game level, attempt, game time...
  - Missing/incompletely logged levels (10% of data)

Data Analysis: Player Engagement (1)

- 50 levels total. Highest level reached was **#29**, while the most popular was **#11**.
- On average, each user played **12.9** levels.
- Levels are repeated by the player's choice. The most attempted level was **#3**.
- Players attempted each level an average of **2.8** times.



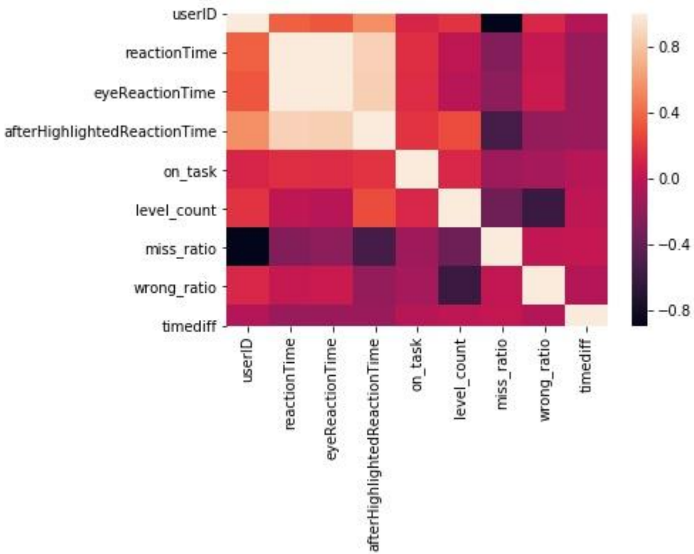




NYU

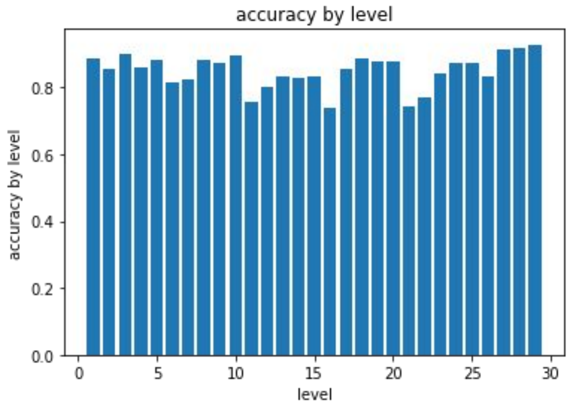
Data Analysis: Player Engagement (2)

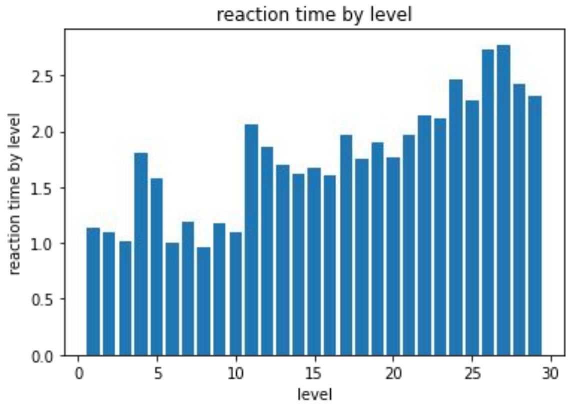
- **Off-Task Players**
  - Over the course of the study, 64 out of 145 participants were marked as ‘off-task’ by researchers
  - Attempted to measure in a number of ways
    - Few levels completed
    - Exceedingly high reaction times
    - Low average accuracy
    - Significant gaps between log timestamps
  - Modeled as a supervised learning task using researcher’s off-task list as the target
  - Few metrics correlated with their findings (<0.15 coeff.)
  - Logistic regression, decision trees, stochastic gradient descent and other ML methods proved ineffective



Data Analysis: Player Performance (1)

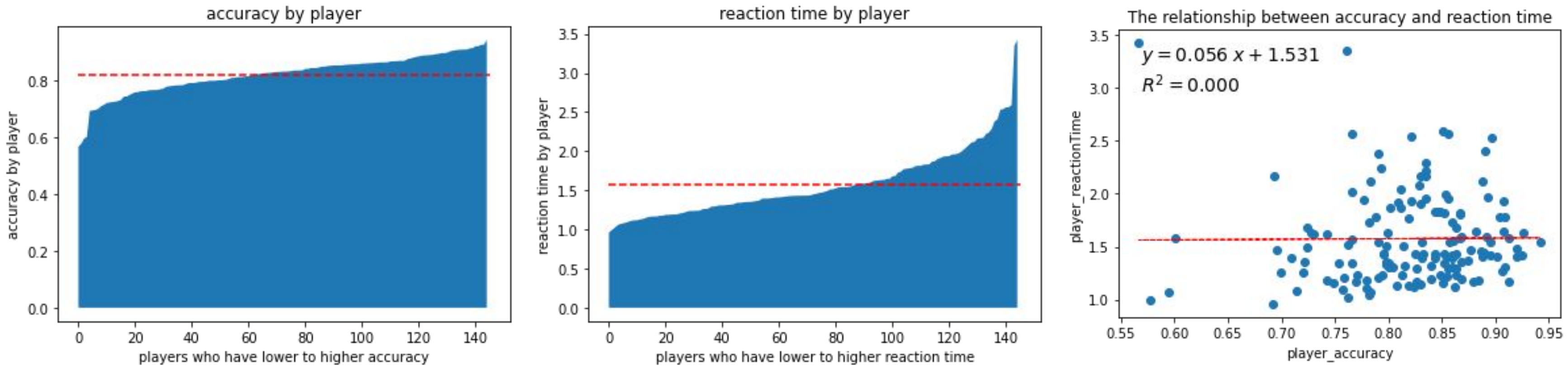
- The easiest level was **#29** (accuracy 92%). The most difficult level was **#16** (accuracy 73%)
- The fastest level was **#8** (0.95s). The slowest level was **#27** (2.77s)
- As levels get harder, reaction times increased while accuracy stayed almost the same.





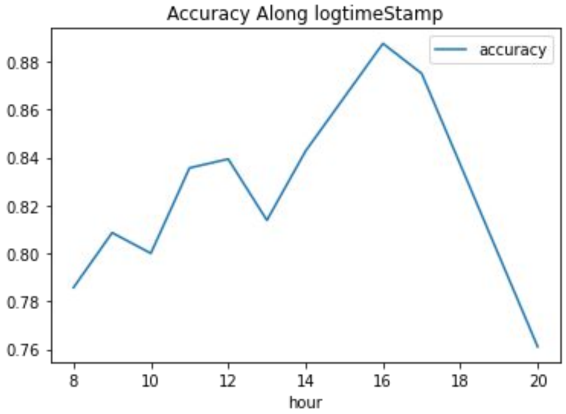
Data Analysis: Player Performance (2)

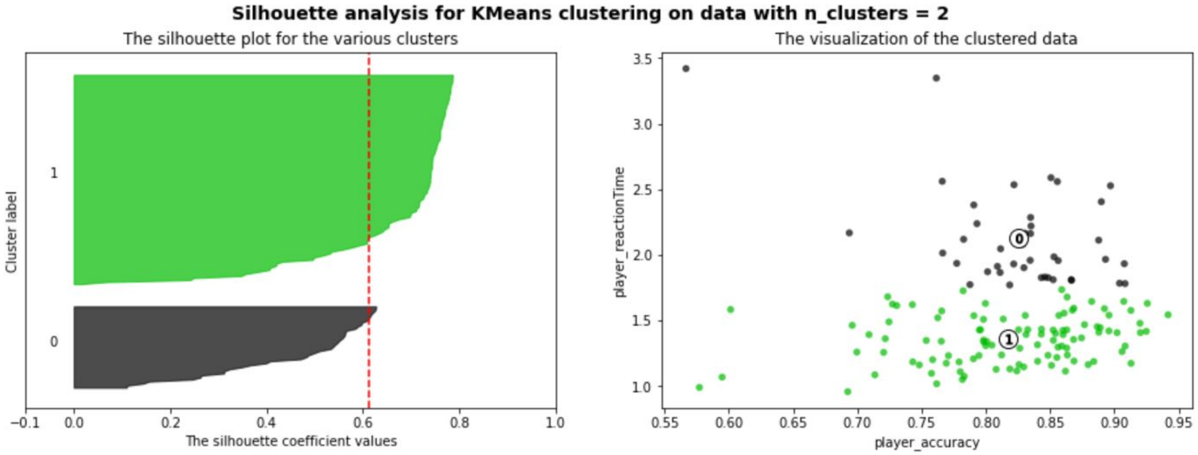
- Average accuracy per user was **82%**.
- Average reaction time was **1.58s.**
- No linear relationship between players' accuracy and reaction time.

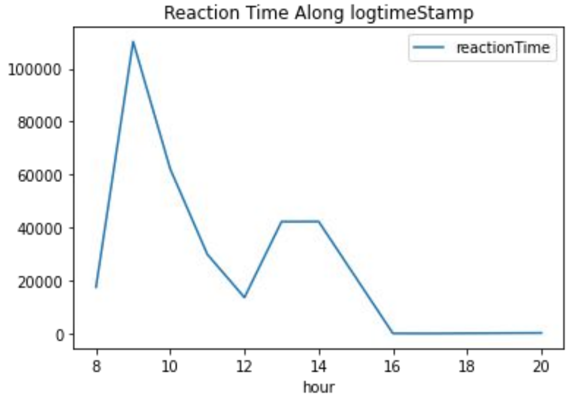


Data Analysis: Player Performance (3)

- Accuracy and reaction time were both higher when users played in the afternoon vs. the morning
- Cluster analysis revealed two player archetypes: slow & accurate (cautious players), and fast & slightly less accurate (balanced players).







Conclusion

- Levels are different in terms of average accuracy and average reaction time.
- Reaction times increased while accuracy stayed almost the same as levels get harder. This is players' strategy.
- Complicated off-task behavior.
- No linear relationship between players' accuracy and reaction time. (This is supported by theory.)
- Better player performance in the afternoon than in the morning.
- Two player groups: cautious players and balanced players.