Team Error: Insufficient Values

Anna Wang, Donny Lou, Jim Liu, Nancy Yan, Roger Gao

Exploration: Data Cleaning & Trends Discovering

Our Goals:

Play2Prevent: to reduce risking-taking behavior

Method: access skill powers among players

Filter Variables:

01 **WHAT & WHY to filter?**

Outside factors: race, age, gender, school

Skills related variables: old_skill_point, oew_skill_point, minigame_id, minigame_level

02 **HOW to clean those variables?**

Remove NAs. Group_by, Create new metrics and variables

Match targeting variables to each player
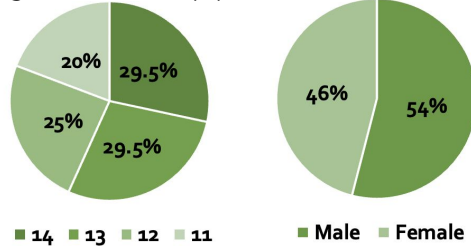
Define New Variables:

Group_by player_id and skill_id

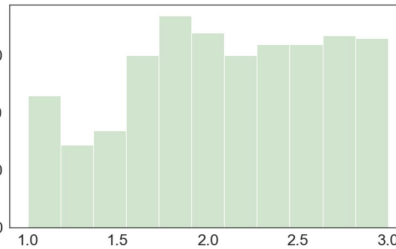Newskill: select when old_skill_point = 0, calculate the mean.

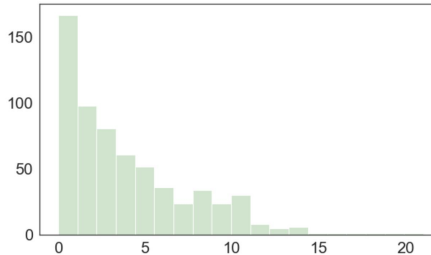Oldskill: count how many time old_skill_point ≠ 0
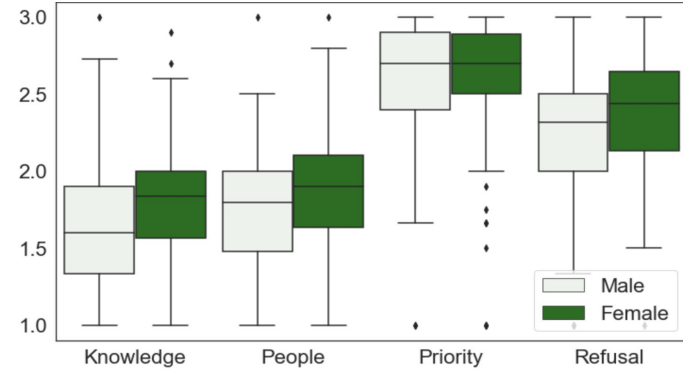
Visual Presentations:

Age Distribution(%)

29.5%, 20%, 25%, 29.5%

■ 14 ■ 13 ■ 12 ■ 11

Gender Distribution(%)

46%, 54%

■ **Male** ■ Female

Histogram for new_skill

Histogram for old_skill

Initial Skill Points mean vs. Skill Ids (By **GENDER**)

Male, Female

Knowledge, People, Priority, Refusal

Initial Skill Points mean vs. Skill Ids (By **RACE**)

Hispanic, African American, Caucasian

Knowledge, People, Priority, Refusal
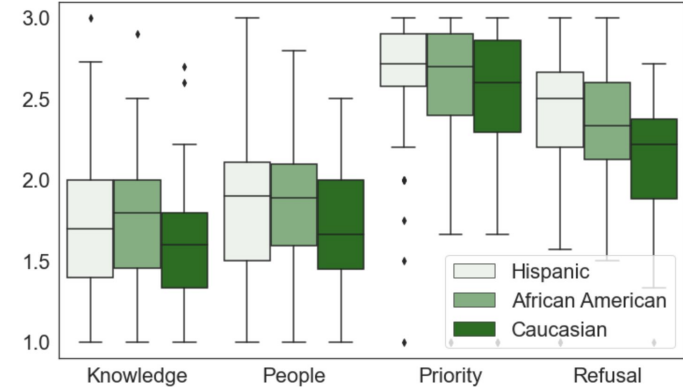
Interpretation: Modeling & Insights

## Multi-variable Regression Model:

Started with: **7** Predictors + **3** Response variables

Final Model: Newskill ~ skill_id + gender + minigame_level

```
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.21605    0.07493   16.229  < 2e-16 ***
skill_idPeople    0.11175    0.04537    2.463   0.0141 *
skill_idPriority  0.87008    0.04543   19.150  < 2e-16 ***
skill_idRefusal   0.56949    0.04636   12.285  < 2e-16 ***
genderMale       -0.08032    0.03246   -2.474   0.0136 *
minigame_level    0.06609    0.00791    8.356 4.21e-16 ***
---
Multiple R-squared:  0.4739,    Adjusted R-squared:  0.4697
F-statistic: 112.8 on 5 and 626 DF,  p-value: < 2.2e-16
```
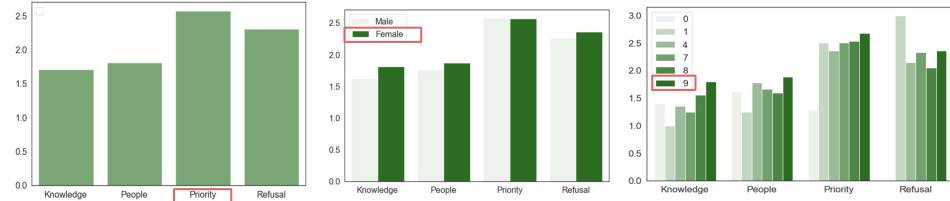

Residuals vs Fitted

Normal Q-Q

Scale-Location

Residuals vs Leverage
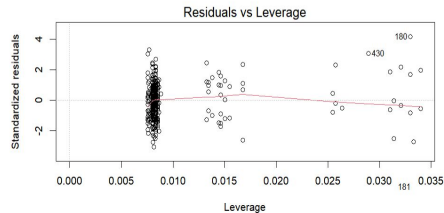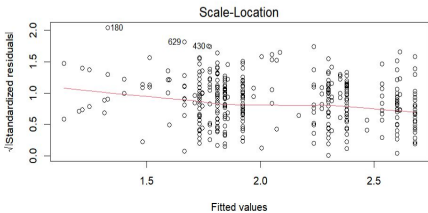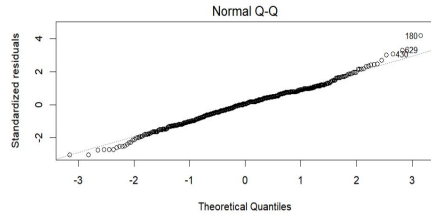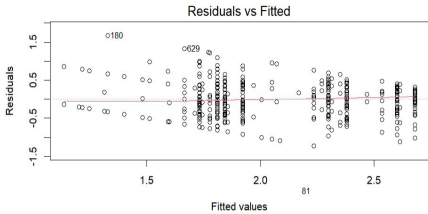
Insights:

Three Bar-plots Grouped by Skill_ids, Genders, Minigame_levels

Cor(game retry times ~ master of refusal power after) = **0.079**

Cor(original refusal power ~ progress of refusal power) = **-0.136**

Students can be educated through refusal category games, but the effectiveness of this game is limited.

Recommendations:

01 Game should focus more on knowledge sense.

02 Educators may pay more attention on boys' health and behavior education

03 Game can improve user engagement