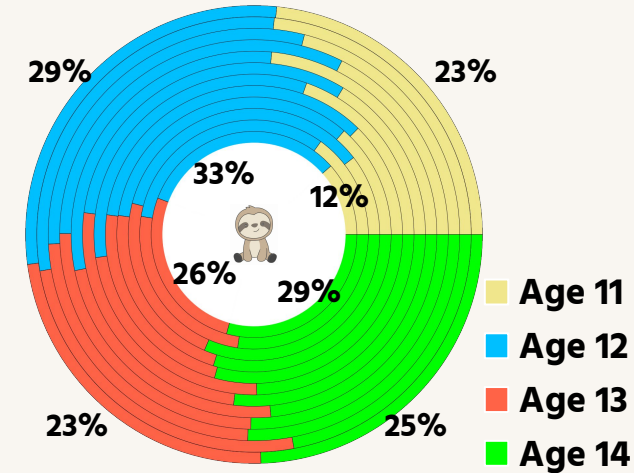# 2022 Datafest
Elm City Stories

Team : Sloth

Team Member: Junxi Feng, Hongxiang Fu Hengyuan Qi, Songyu Sun, Harry Zhang
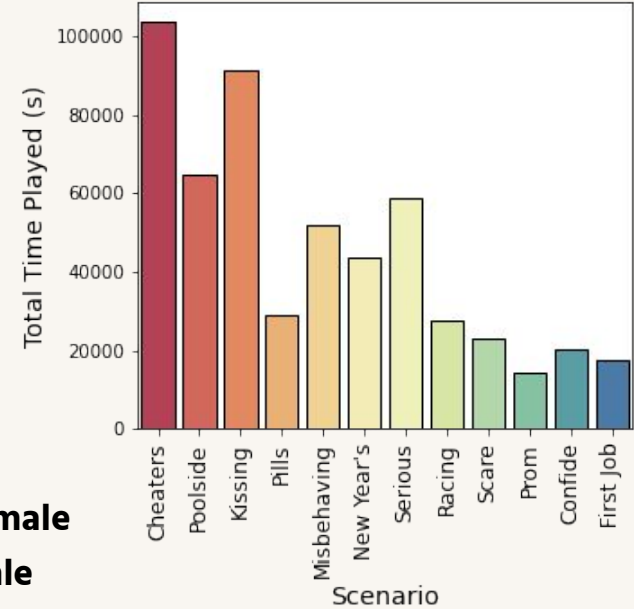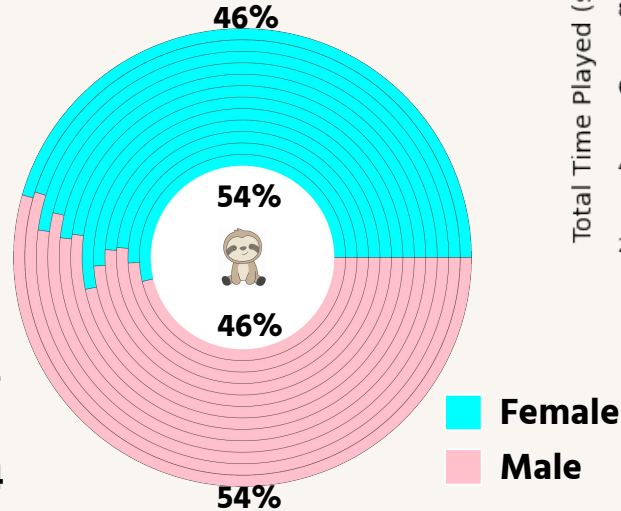
Data Trend Discovery

Total Time Played in Different Scenarios

Age Scenarios Donut Plot

29%
23%
33%
12%
26%
29%
23%
25%

- Age 11
- Age 12
- Age 13
- Age 14

Gender Scenarios Donut Plot

46%
54%
46%
54%

- Female
- Male

- Outer donut to inner donut goes from scenario **Cheaters** to **First job** (same as the bar plot on the right)
- **Age 11** percentage **decreases** as game goes on
- **Male** percentage **decreases** as game goes on

- **Cheaters** and **Kissing** scenarios are the most popular
- **Prom** scenario is the least popular

Modeling and Insights

Ordinary Least Square Multivariate Regression

Hierarchical Clustering
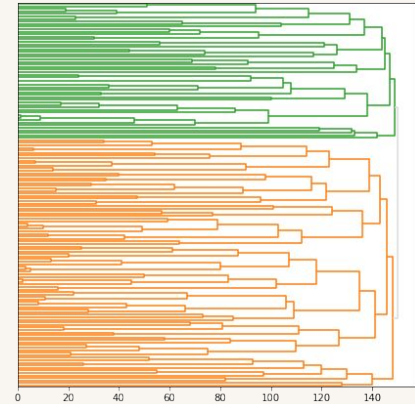
- Principal Component Analysis

- Leave One Out Cross Validation

```
                        OLS Regression Results
==============================================================================
Dep. Variable:     Number of Stacks Completed   R-squared:              0.567
Model:                                    OLS   Adj. R-squared:         0.546
Method:                         Least Squares   F-statistic:            26.93
Date:                        Sat, 30 Apr 2022   Prob (F-statistic):  2.36e-23
Time:                                22:25:20   Log-Likelihood:       -46.374
No. Observations:                         152   AIC:                    108.7
Df Residuals:                             144   BIC:                    132.9
Df Model:                                   7
Covariance Type:                    nonrobust
==============================================================================
                                  coef   std err       t    P>|t|    [0.025    0.975]
------------------------------------------------------------------------------
const                           3.7337     0.438    8.530    0.000     2.869     4.599
Failure rate of finding objects -0.2093    0.031   -6.671    0.000    -0.271    -0.147
KNOW SENSE Failure rate         -0.2018    0.041   -4.912    0.000    -0.283    -0.121
KNOW SENSE Average Playing Time -0.1148    0.034   -3.354    0.001    -0.182    -0.047
PRIORITY SENSE Average Playing Time -0.0199 0.029  -0.695    0.488    -0.077     0.037
PEOPLE SENSE Average Playing Time 0.0135   0.027    0.498    0.619    -0.040     0.067
REFUSAL POWER Average Playing Time -0.0292 0.014   -2.033    0.044    -0.058    -0.001
Age                             -0.0101    0.028   -0.367    0.714    -0.065     0.044
==============================================================================
Omnibus:                       38.671   Durbin-Watson:              2.094
Prob(Omnibus):                  0.000   Jarque-Bera (JB):          65.267
Skew:                          -1.244   Prob(JB):                6.72e-15
Kurtosis:                       5.030   Cond. No.                    242.
==============================================================================
```

Logistic Regression Confusion Matrix

- As **failure rate** increases, **number of completed stacks** decreases
- **Reducing difficulty level** leads to a more thorough game experience
- Use more predictors to find the similarity between players