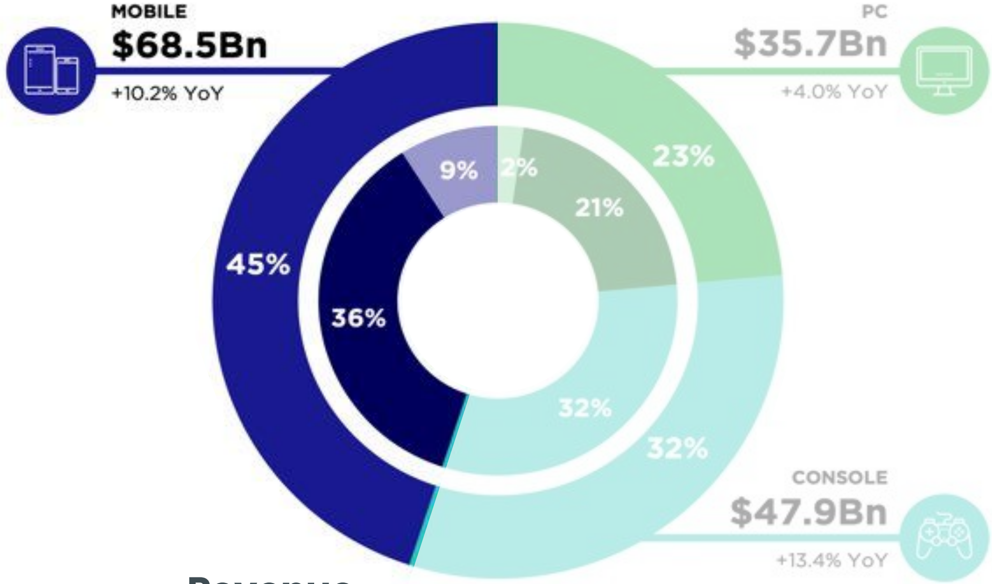
Mobile Game App Analysis

Customer & Social Analytics

0

Topic Intro

Mobile Gaming Share | Total Gaming Rev Worldwide



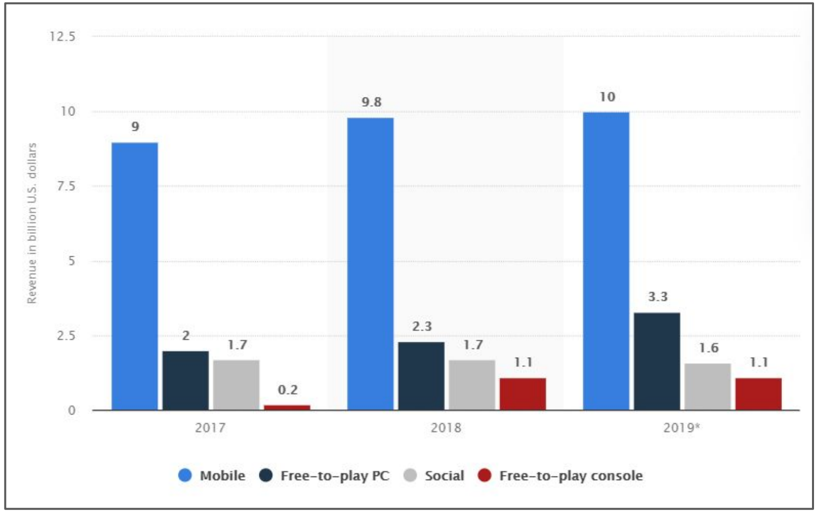
Revenue

Source: ©Newzoo | 2019 Global Games Market Report
newzoo.com/globalgamesreport

$68.5Bn

Mobile game revenues in 2019 will account for 45% of the global market.

Digital Games Revenue [North America]



Growth of Rev in Mobile

Proportion of Revenue is more than all other digital game platforms combined

Source: https://www.statista.com/topics/1906/mobile-gaming/

1

Intro & Objective

Table of Contents

Project Goals

Goal 1

Is **price** an important factor

Goal 2

Is **age rating** an important factor

Goal 3

Effect of **multilingual**

Goal 4

Effect of recent **1-yr updates**

Rating

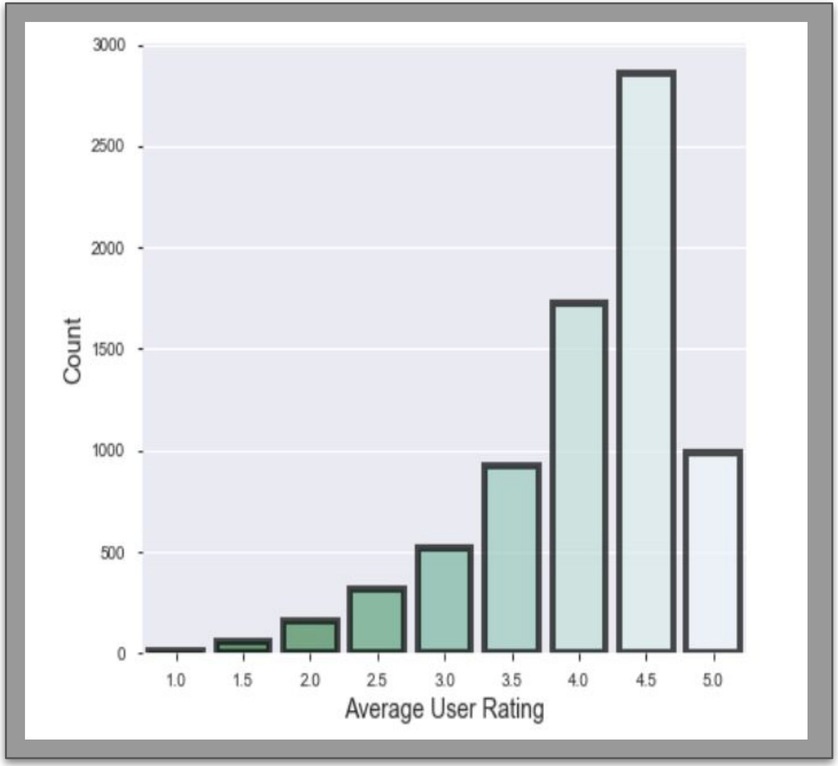
2

Data Review
EDA

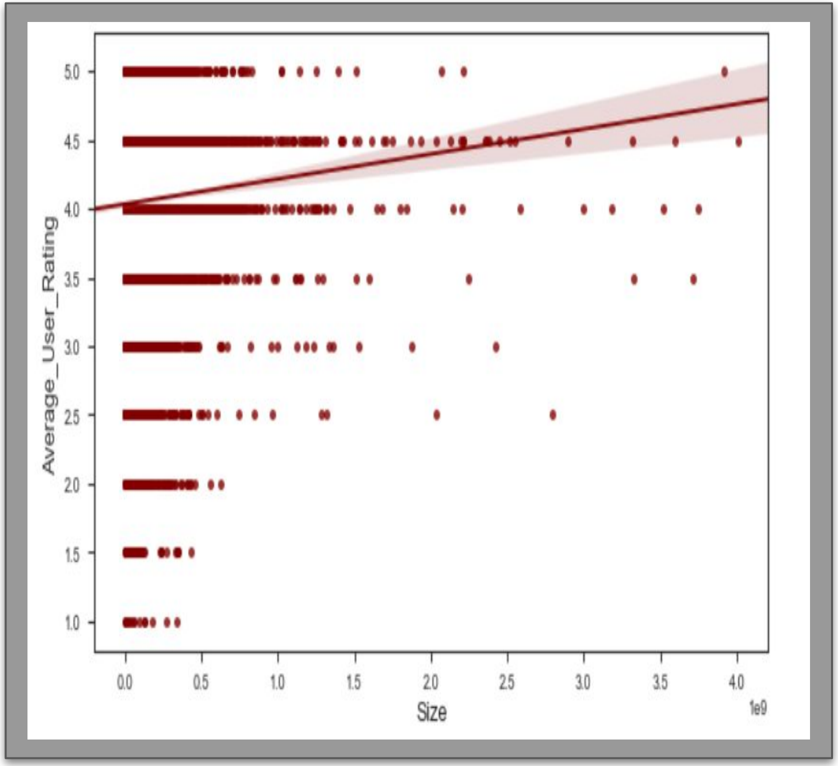
Data Overview

MOBILE STRATEGY GAMES FEATURE OVERVIEW	
Feature	**Feat. Desc.**
ID	The assigned ID
Name	The name
Subtitle	The secondary text under the name
Average User Rating	Rounded to nearest .5, requires at least 5 ratings
User Rating Count	No. of ratings internationally, null means it is below 5
Price	Price in USD
In-app Purchases	Prices of available in-app purchases
Description	App description
Developer	App developer
Age Rating	Either 4+, 9+, 12+ or 17+
Languages	ISO2A language codes
Size	Size of the app in bytes
Primary Genre	The main genre
Genres	Genres of the app
Original Release Date	Original Release Date
Current Version Release Date	Current Version Release Date
URL	The URL
Icon URL	512px x 512px jpg

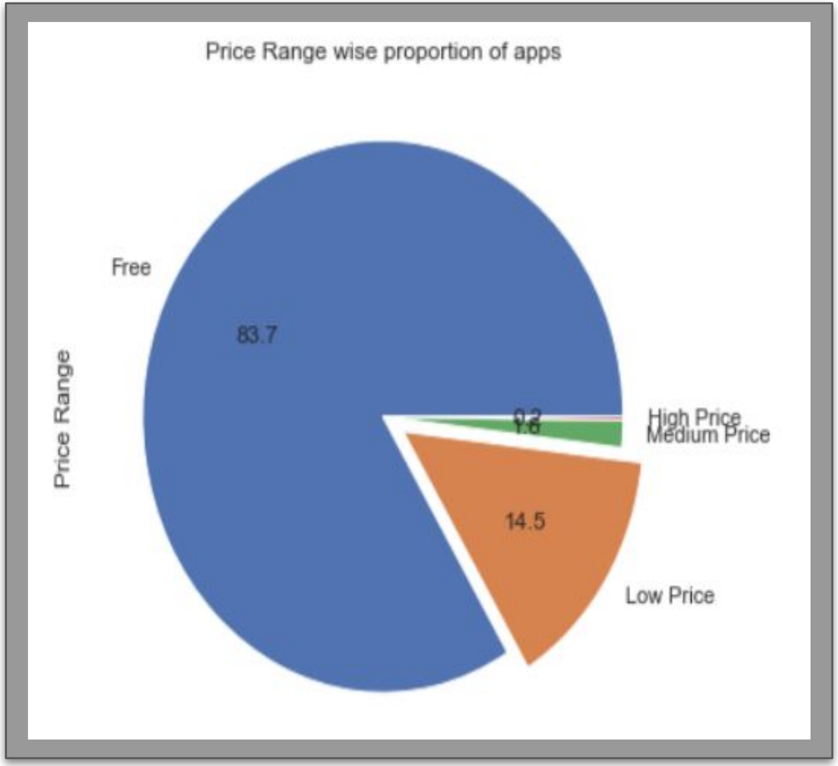
Average User Rating



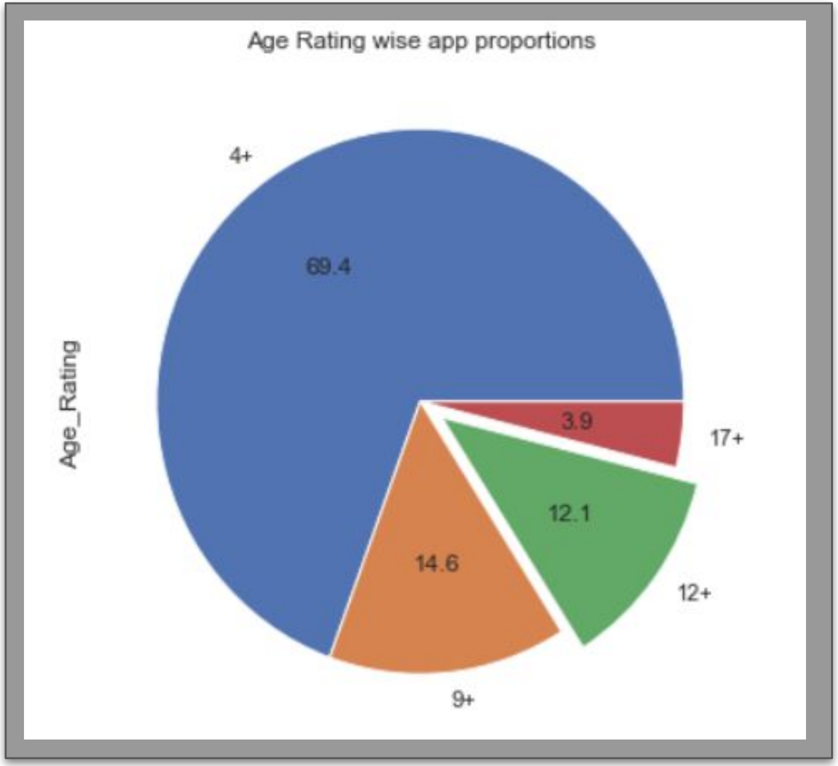
Average User Rating & Size Correlation



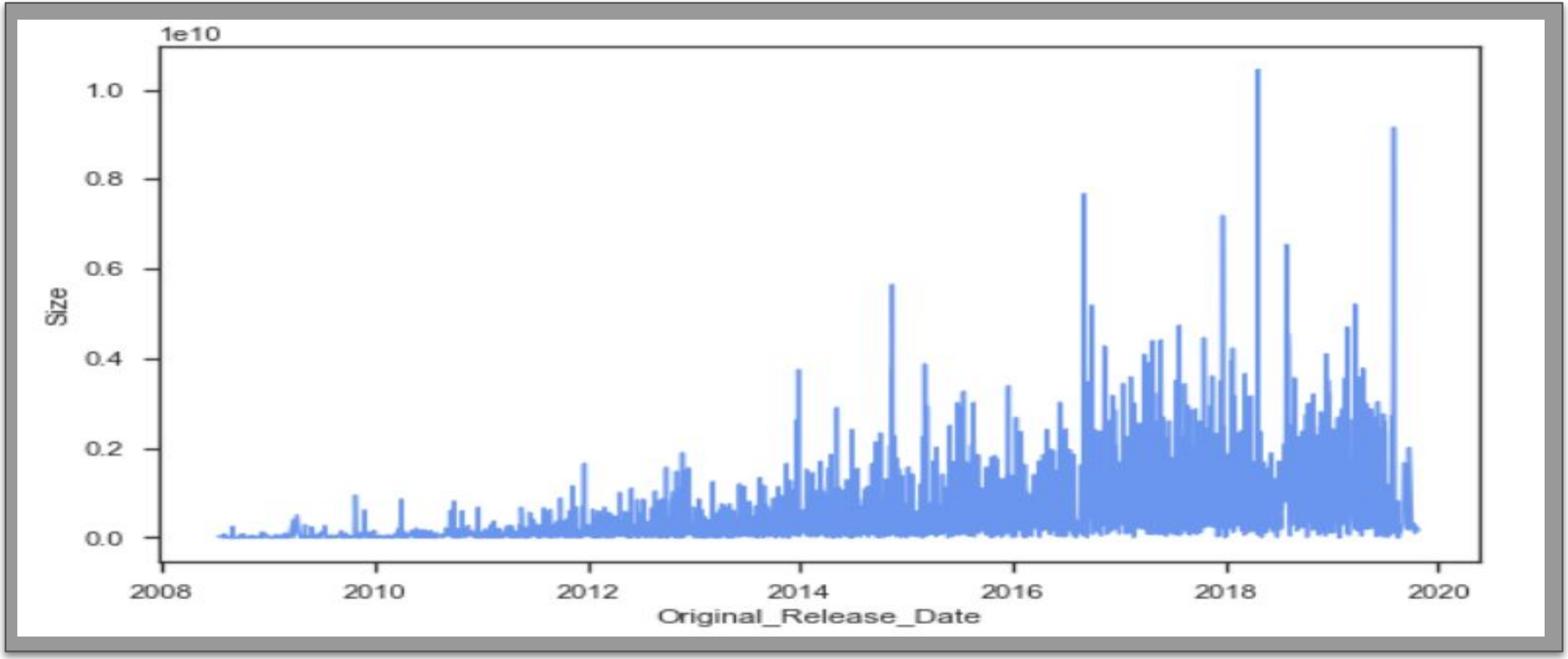
Price Range of Apps



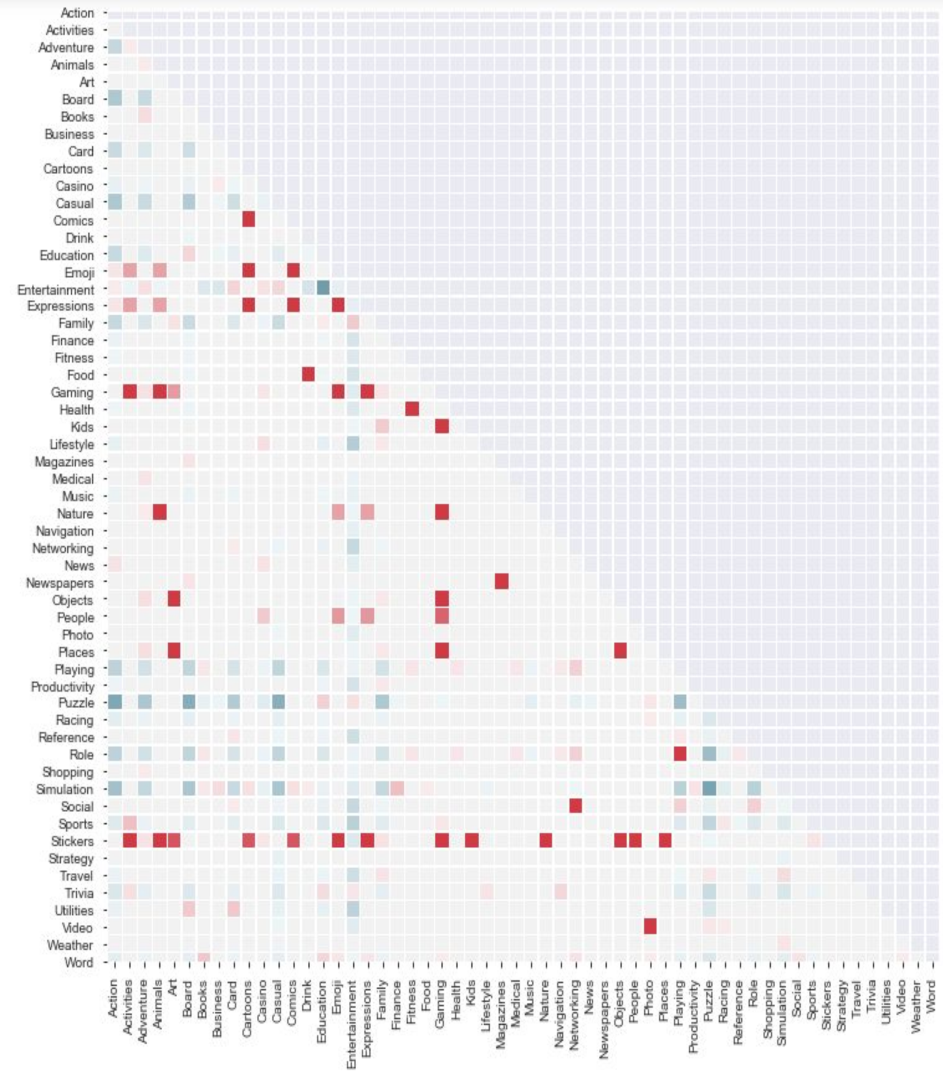
Age Rating wise app proportions

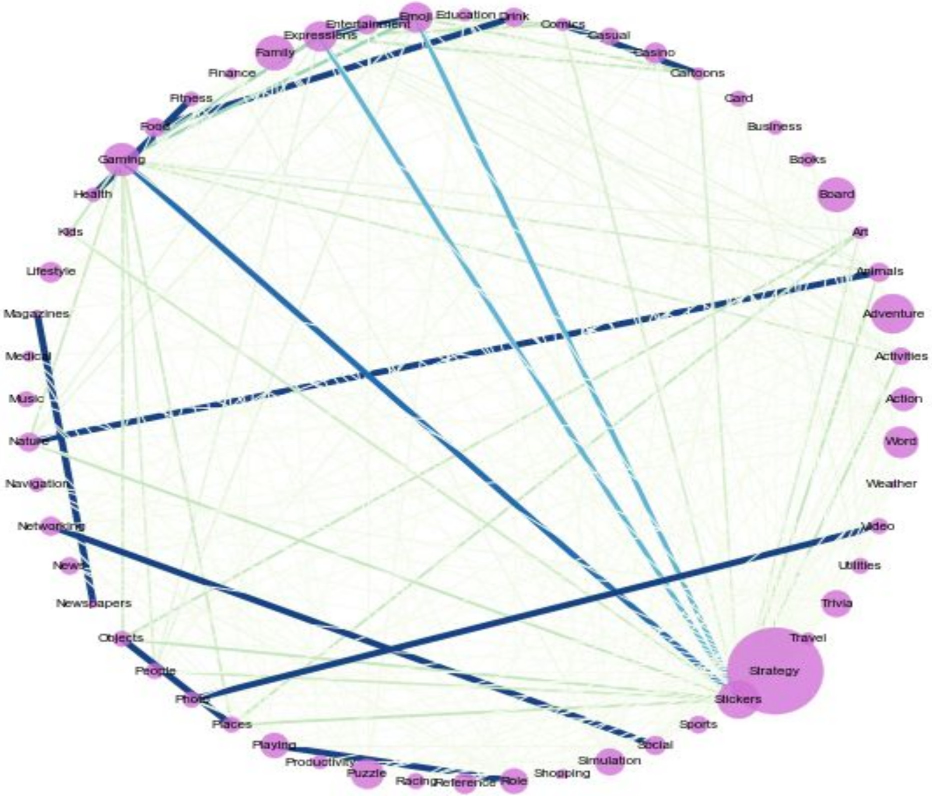


Effect of Size on Original Release Date

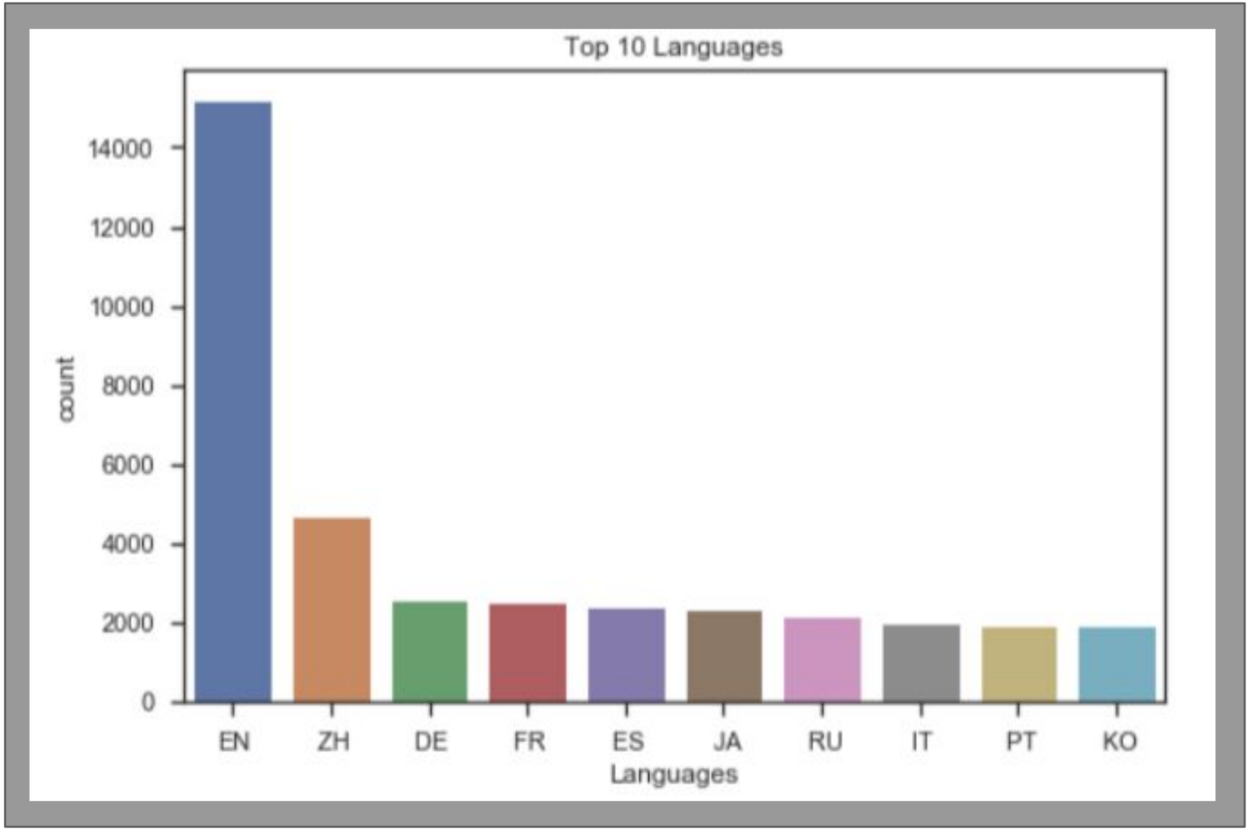


Correlation Plot for Genre List





Top 10 Game Languages

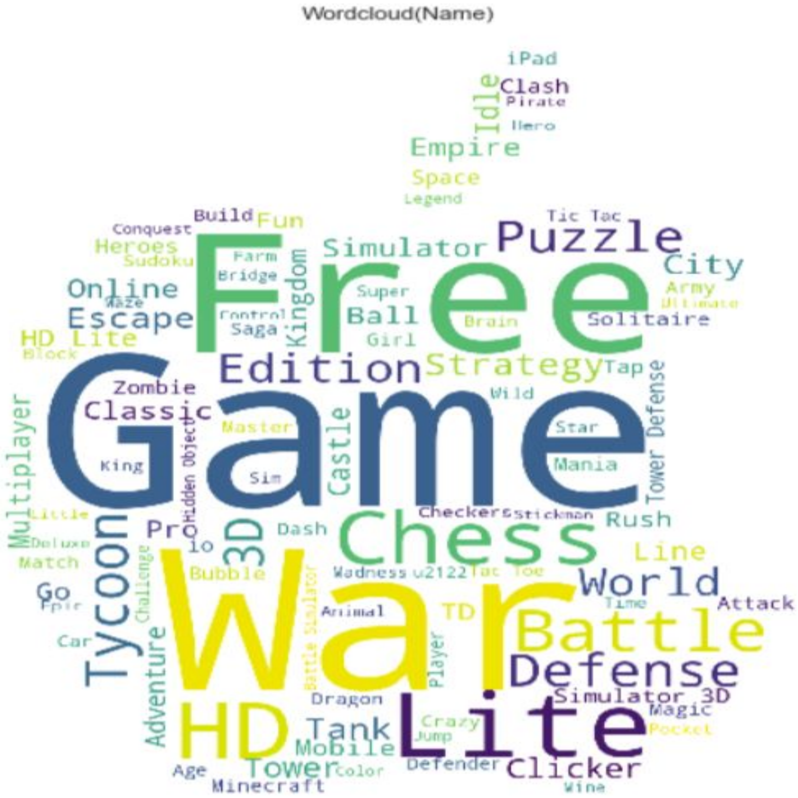


	Languages	Count
0	EN	16834
1	ZH	4989
2	DE	2599
3	FR	2526
4	ES	2420
5	JA	2364
6	RU	2179
7	IT	1996
8	PT	1921
9	KO	1919
10	TR	1357
11	NL	1283
12	PL	1152
13	SV	1127
14	CS	887
15	DA	749
16	TH	744
17	ID	688
18	NB	684
19	FI	668

WORDCLOUD NAME COMPARISON

Average User Rating > 4

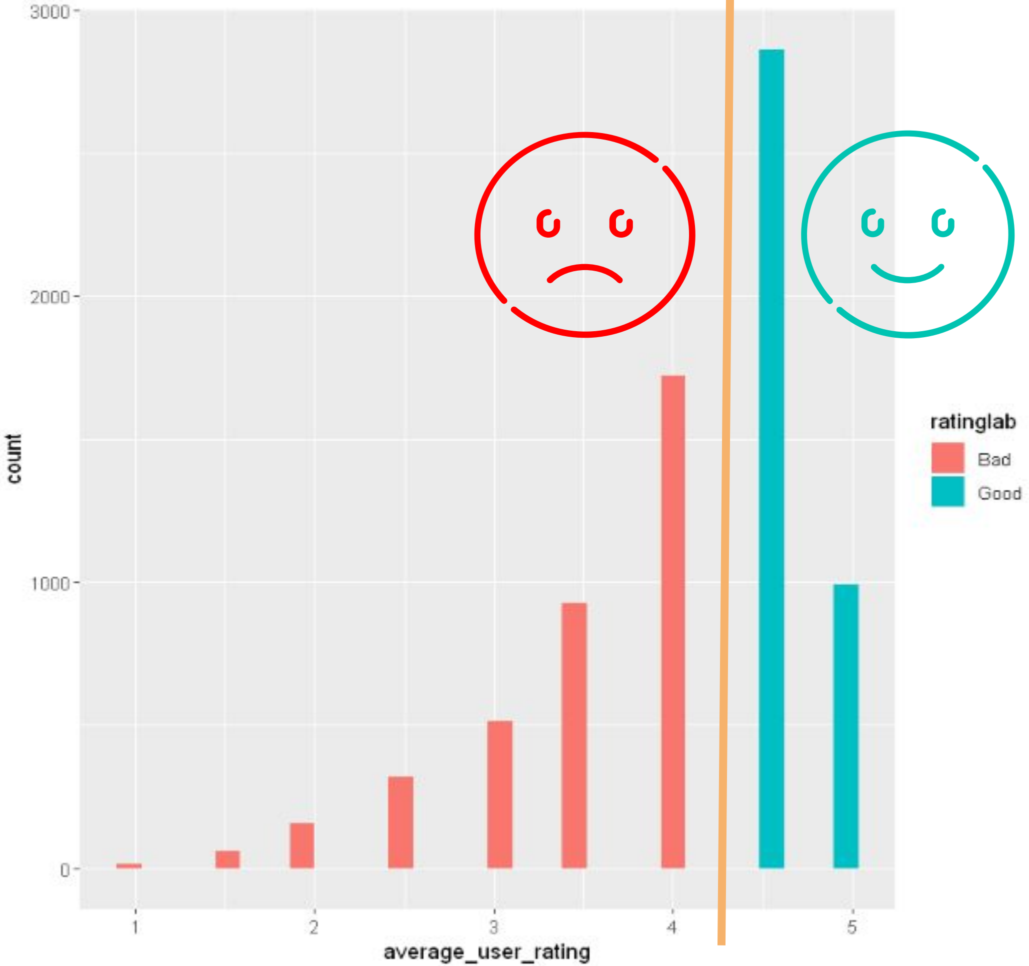
Average User Rating < 4

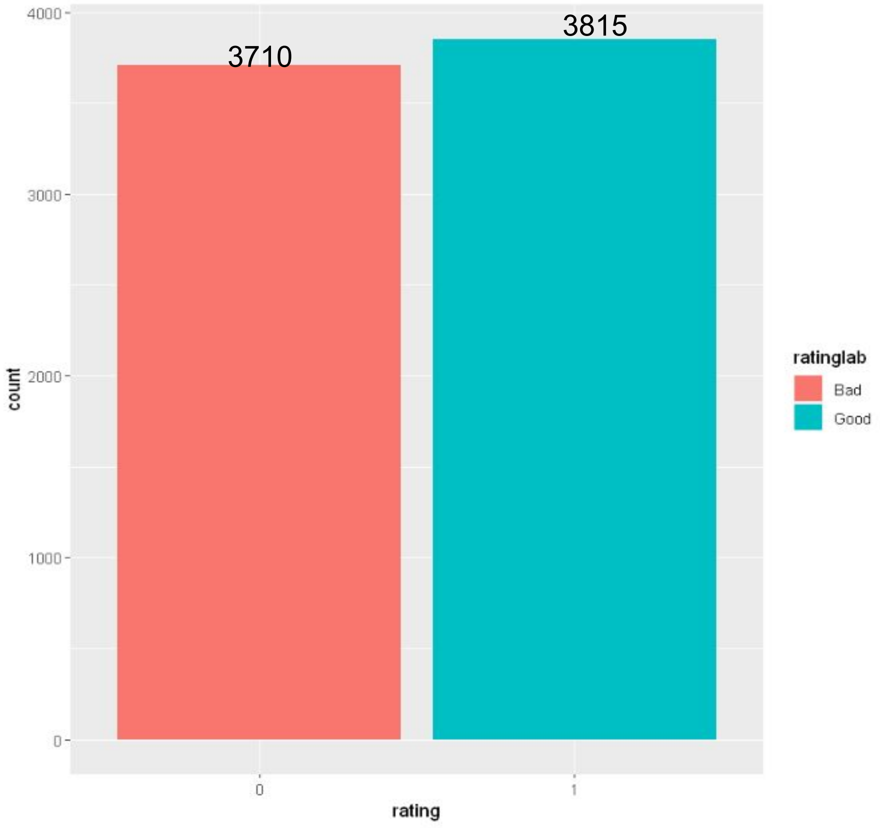


3

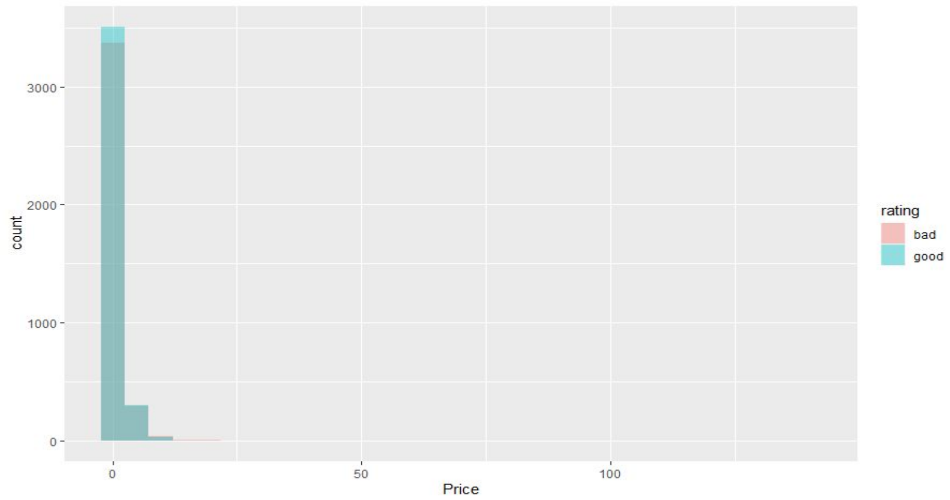
EDA | Cleaned Data

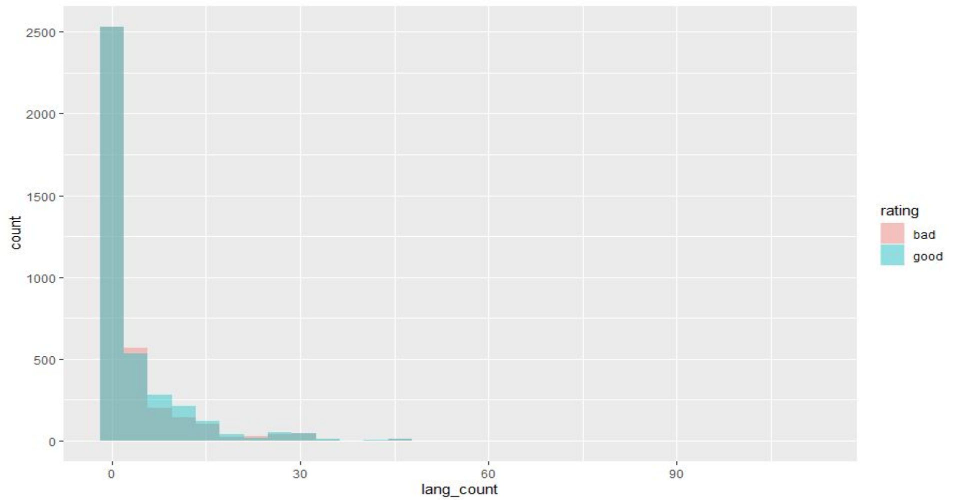
Average score Distribution

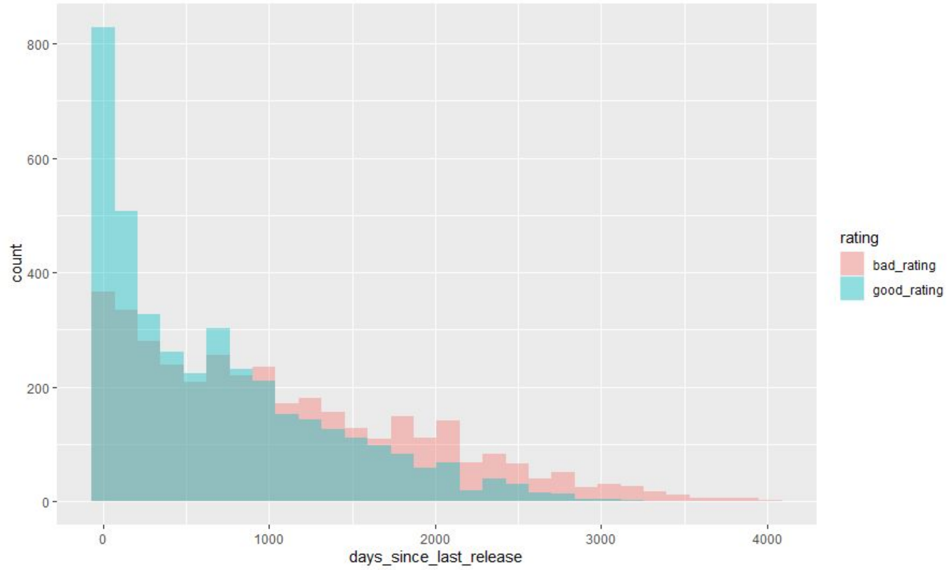


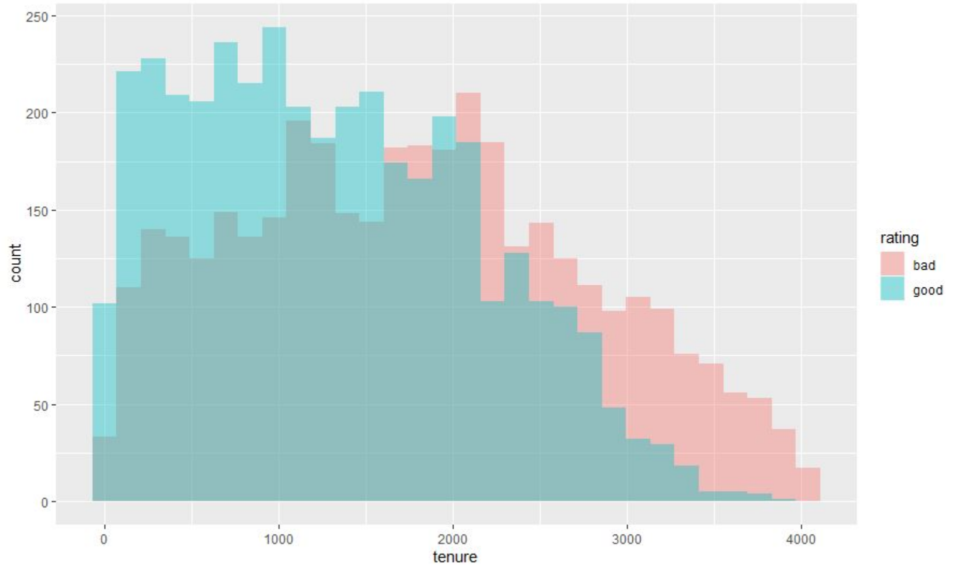


data$rating = ifelse(data$average_user_rating >= 4.5, 1, 0)

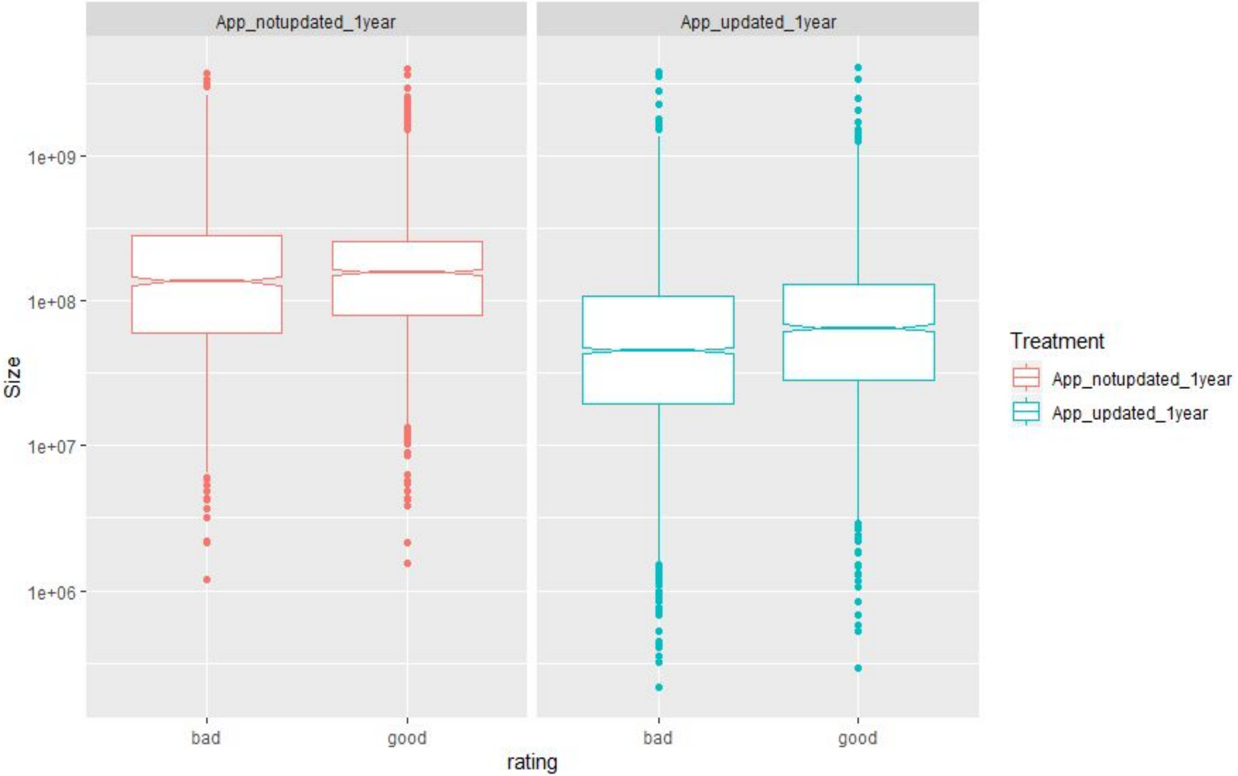

count
3000
2000
1000
0
0
50
100
Price
rating
bad
good


count
2500
2000
1500
1000
500
0
0
30
60
90
lang_count
rating
bad
good


count
800
600
400
200
0
0
1000
2000
3000
4000
days_since_last_release
rating
bad_rating
good_rating


count
250
200
150
100
50
0
0
1000
2000
3000
4000
tenure
rating
bad
good

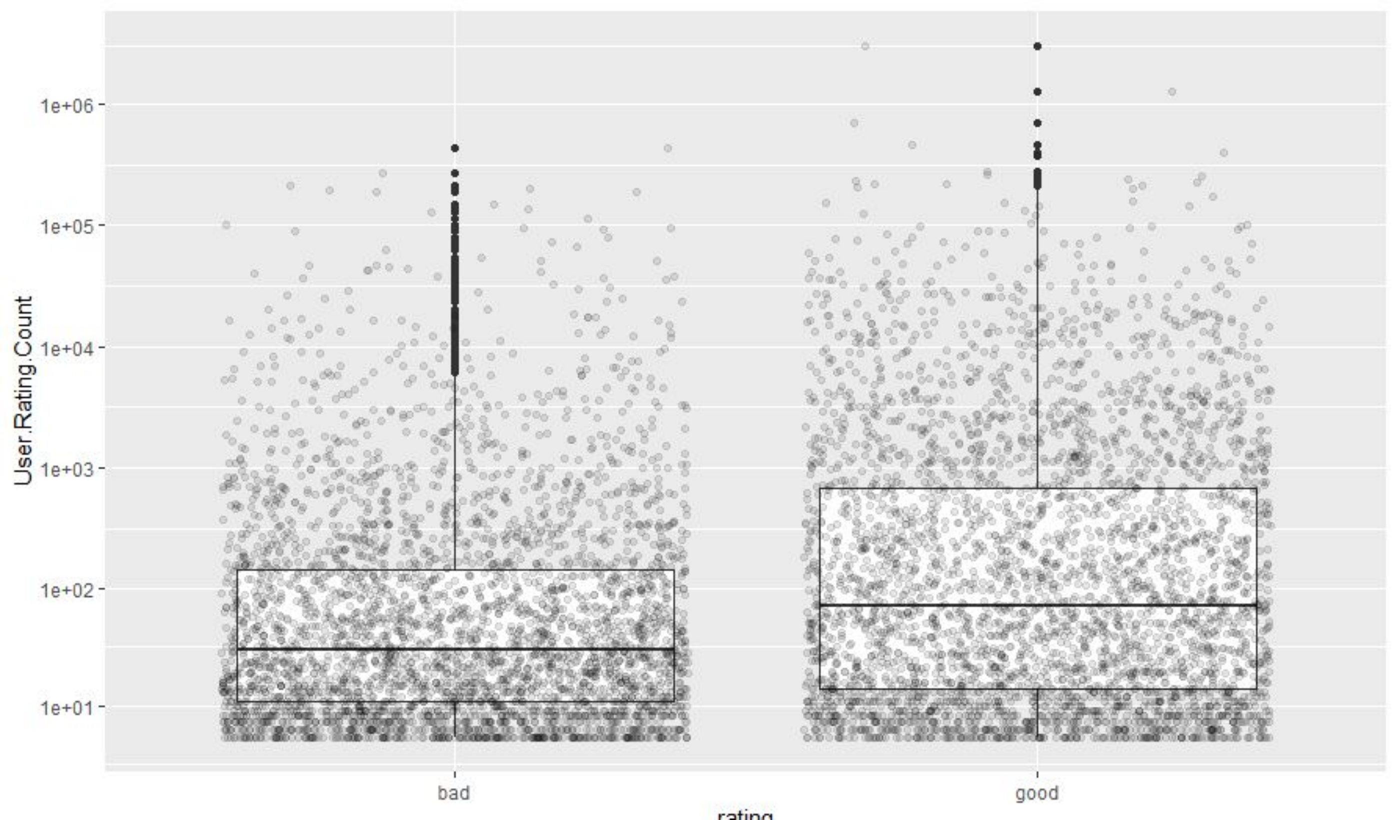
Updating Effect



df_1year$Treatment <- ifelse(df_1year$days_since_last_release <=365,1,0)


CLASH of CLANS

The relationship between rating and # rating



4

Modeling

Logistic | Propensity

Propensity Modelling

We decided to look **deeper** into the impact of apps with **multiple languages** and apps with a **recent updates**

Propensity

Apps with **multiple languages** as treatment

Propensity

Apps updated within **1-year** as treatment

Propensity – Multilingual | TTest

T-Test on Unmatched Data

```
> with(df1, t.test(rating ~ Treatment))

        Welch Two Sample t-test

data:  rating by Treatment
t = -2.3337, df = 4984.5, p-value = 0.01965
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.052439001 -0.004558589
sample estimates:
mean in group 0 mean in group 1
      0.4999012       0.5284000
```

Treatment = lang_count > 1 language

Propensity – Multilingual | TTest

T-Test Covariates

days_since_last_release

```
[[5]]

	Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = 1.3827, df = 4803, p-value = 0.1668
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -12.45115  72.05248
sample estimates:
mean in group 0 mean in group 1
       774.2719        744.4712
```

Adult

```
[[6]]

	Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = -0.41333, df = 4819.1, p-value = 0.6794
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.011900169  0.007755951
sample estimates:
mean in group 0 mean in group 1
     0.03025280      0.03232491
```

ln_mean_in_app_purchase

```
[[7]]

	Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = -1.0754, df = 4823.6, p-value = 0.2822
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.27609978  0.08049146
sample estimates:
mean in group 0 mean in group 1
       4.947795        5.045600
```

Ln_user.rating.count

```
	Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = -0.66119, df = 4824, p-value = 0.5085
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.17677644  0.08760867
sample estimates:
mean in group 0 mean in group 1
       4.932610        4.977194
```

ln_price

```
[[2]]

	Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = 0.3503, df = 4823.8, p-value = 0.7261
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.02624995  0.03767177
sample estimates:
mean in group 0 mean in group 1
      0.2443014       0.2385905
```

ln_size

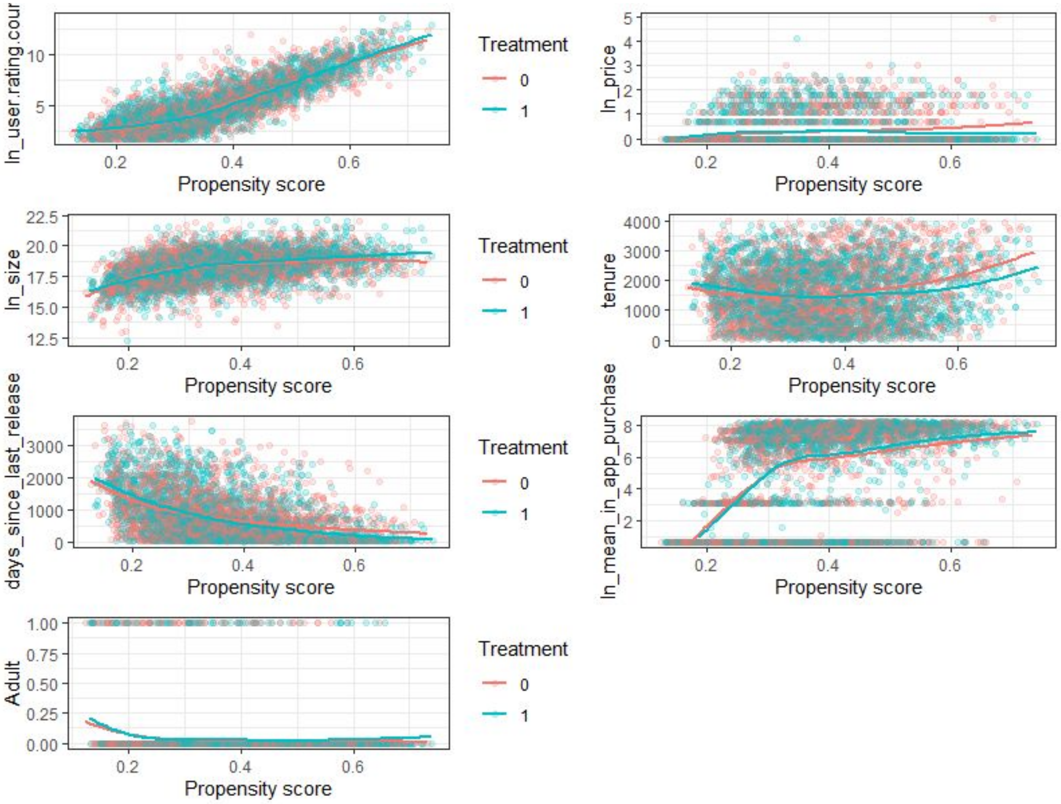
```
[[3]]

	Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = -1.3582, df = 4810.7, p-value = 0.1745
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.11879914  0.02156011
sample estimates:
mean in group 0 mean in group 1
       18.22278        18.27140
```

tenure

```
[[4]]

	Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = 1.4703, df = 4813.4, p-value = 0.1415
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -13.67674  95.73807
sample estimates:
mean in group 0 mean in group 1
       1602.697        1561.666
```

Propensity – Multilingual

Viz Matching Covariates



Propensity – Multilingual

T-Test on matched Data

```
        Welch Two Sample t-test

data:  rating by Treatment
t = 2.2233, df = 4823.9, p-value = 0.02624
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.003772112 0.060048858
sample estimates:
mean in group 0 mean in group 1
      0.5520099       0.5200995
```

Propensity - 1-Year Updates

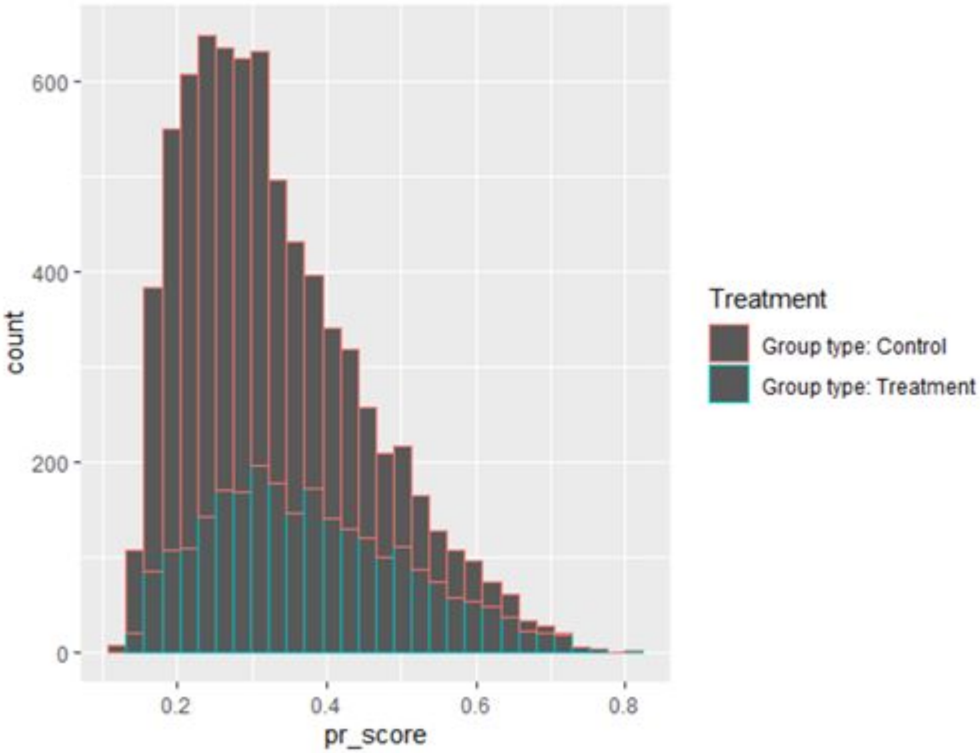
T-Test on Unmatched Data

```
> with(df_1year, t.test(rating ~ Treatment))

        Welch Two Sample t-test

data:  rating by Treatment
t = -15.519, df = 5732.5, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.2048160 -0.1588731
sample estimates:
mean in group 0 mean in group 1
      0.4441237       0.6259683
```

Treatment = days_since_last_release = *within last 365 days*

Propensity - Recent **1-Year** Updates

Match Distribution



Propensity - Recent **1-Year** Updates | TTest

T-Test Covariates

Tenure

```
[[5]]

        Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = 1.9225, df = 3240.1, p-value = 0.05463
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  -1.148843 116.831610
sample estimates:
mean in group 0 mean in group 1
       1400.534        1342.692
```

Adult

```
[[6]]

        Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = 0.16495, df = 3351.5, p-value = 0.869
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.01298339  0.01536860
sample estimates:
mean in group 0 mean in group 1
     0.04651163      0.04531902
```

Mean_in_app_purchase

```
[[7]]

        Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = 2.3382, df = 3351.4, p-value = 0.01944
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  15.84893 180.47903
sample estimates:
mean in group 0 mean in group 1
       1176.280        1078.116
```

User.Rating.Count

```
        Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = -1.6796, df = 1766.7, p-value = 0.09321
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -7498.5233   580.1596
sample estimates:
mean in group 0 mean in group 1
       1792.955        5252.137
```

Price

```
        Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = -0.29304, df = 3273.3, p-value = 0.7695
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1643180  0.1215869
sample estimates:
mean in group 0 mean in group 1
      0.6126058       0.6339714
```

lang_count

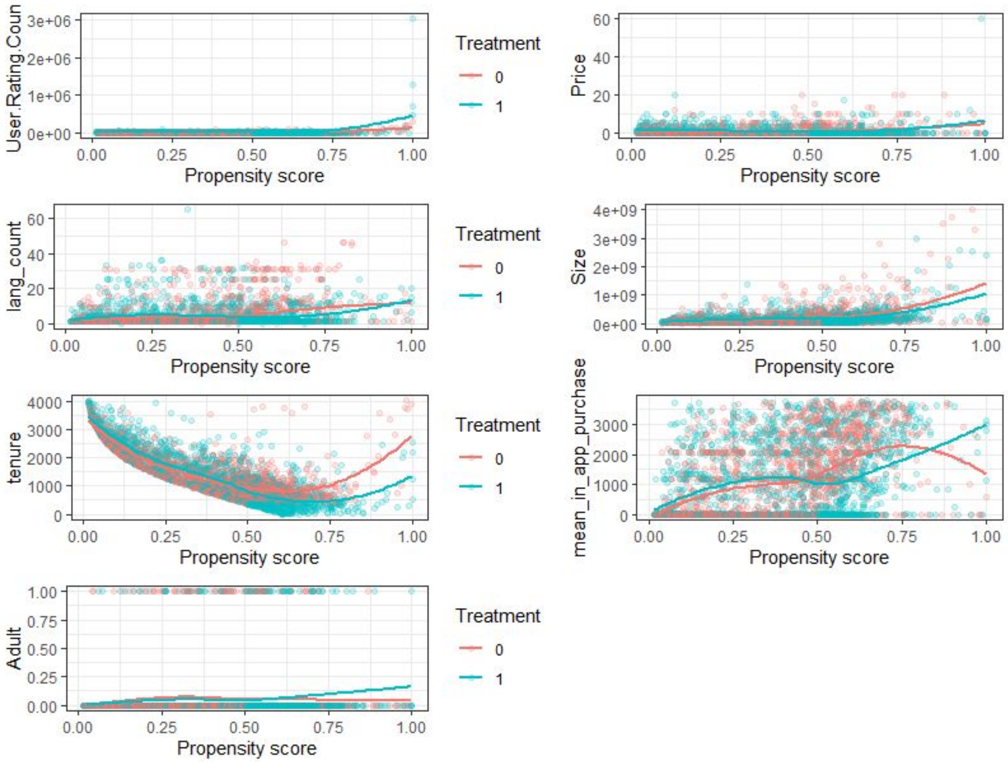
```
[[3]]

        Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = 1.4008, df = 3229.4, p-value = 0.1614
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1265507  0.7598244
sample estimates:
mean in group 0 mean in group 1
       4.112701        3.796064
```

Size

```
[[4]]

        Welch Two Sample t-test

data:  dta_m[, v] by dta_m$Treatment
t = 0.78383, df = 3047.1, p-value = 0.4332
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -10365606  24172604
sample estimates:
mean in group 0 mean in group 1
      166745315       159841816
```

Propensity - Recent 1-Year Updates

Viz Matching Covariates



Propensity - Recent 1-Year Updates

```
> with(dta_m, t.test(rating ~ Treatment))

        Welch Two Sample t-test

data:  rating by Treatment
t = -5.6279, df = 3351.6, p-value = 1.974e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.13025548 -0.06294667
sample estimates:
mean in group 0 mean in group 1
      0.4794275       0.5760286
```

Propensity - Recent 1-Year Updates

Logistics Regression matched data

```
Call:
glm(formula = rating ~ Treatment + User.Rating.Count + Size +
    tenure + Adult + mean_in_app_purchase, family = binomial(),
    data = dta_m)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.6864  -1.1852   0.8588   1.0939   2.0025

Coefficients:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)           4.033e-01  8.942e-02   4.510 6.49e-06 ***
Treatment             3.867e-01  7.109e-02   5.440 5.34e-08 ***
User.Rating.Count     7.036e-06  3.103e-06   2.267  0.02339 *
Size                 -4.879e-10  1.567e-10  -3.114  0.00185 **
tenure               -4.222e-04  4.314e-05  -9.787  < 2e-16 ***
Adult                -4.483e-01  1.707e-01  -2.627  0.00861 **
mean_in_app_purchase  1.598e-04  2.961e-05   5.396 6.81e-08 ***
```

Propensity – Recent 1-Year Updates | GLM-Log

Exp on Coefficients

```
> exp(coefficients(model))
        (Intercept)           Treatment   User.Rating.Count
          1.4966927           1.4721423           1.0000070
               Size              tenure               Adult
          1.0000000           0.9995779           0.6387005
mean_in_app_purchase
          1.0001598
```

Logistic | Rating Good

		Estimate	Std. Error	z value	Pr(>|z|)	
	(Rating > 4.0 = Good))	0.458	0.133	3.434	0.001	
Pos Odds	adultAll_age	0.541	0.126	4.304	**0.000**	Sig
	mean_in_app_purchase	0.006	0.003	2.200	0.**028**	
Pos Odds	price	0.002	0.010	0.180	0.**857**	Not Sig
	tenure	0.000	0.000	-13.053	0.000	
Pos Odds	days_since_last_release	0.000	0.000	-8.147	0.000	Sig
	size	0.000	0.000	-3.210	0.001	
	user_rating_count	0.000	0.000	5.198	0.**000**	
Neg Odds	lang_count	-0.001	0.004	-0.361	0.**718**	Not Sig

glm(formula = ratingGood ~ tenure + price + days_since_last_release + mean_in_app_purchase + user_rating_count + size + lang_count + adultAll_age, family = binomial, data = logicdata)

5

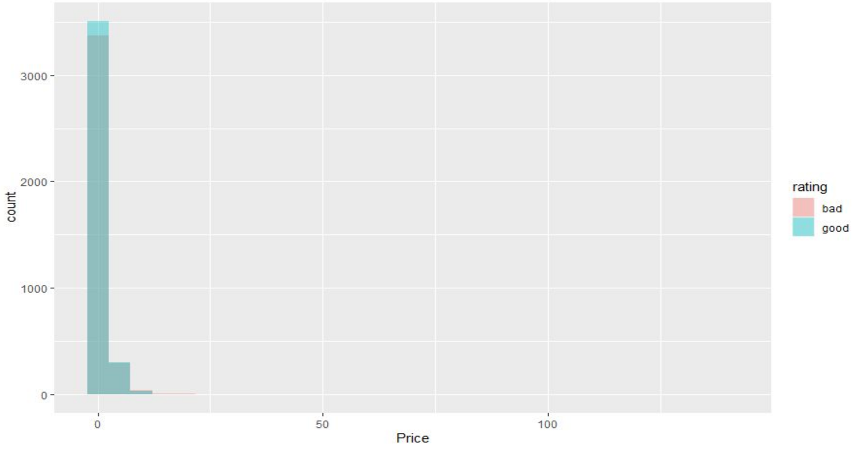
Findings

Project Goals

Goal 1

Is **price** an important factor on rating?

Given our dataset, **No**.



Price might be an important factor, but there is **not enough variance** of pricing in our data to determine this.

Project Goals

Goal 2

Is **age rating** an important factor on rating?

Yes

Age rating is **significant** in standard Logistic model.

Taking age = 17+ is also a **significant covariate** in our final propensity model.

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	4.033e-01	8.942e-02	4.510	6.49e-06	***
Treatment	3.867e-01	7.109e-02	5.440	5.34e-08	***
User.Rating.Count	7.036e-06	3.103e-06	2.267	0.02339	*
Size	-4.879e-10	1.567e-10	-3.114	0.00185	**
tenure	-4.222e-04	4.314e-05	-9.787	< 2e-16	***
Adult	**-4.483e-01**	**1.707e-01**	**-2.627**	**0.00861**	**
mean_in_app_purchase	1.598e-04	2.961e-05	5.396	6.81e-08	***

Project Goals

Goal 3

Does having an app with **multiple languages** have on effect on better ratings?

No

Project Goals

Goal 4

Does app **updating w/n 1-Yr** have an effect on better ratings?

Yes

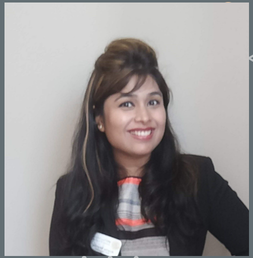
Updating an app within the last year is **proven significant** and has a **large positive effect** on app rating.

Treatment /1-year Update = **47.2%**

Our Team

Shuochen Xu

Geoffrey Nel



Arpita Sharda

Aneesh Kalag

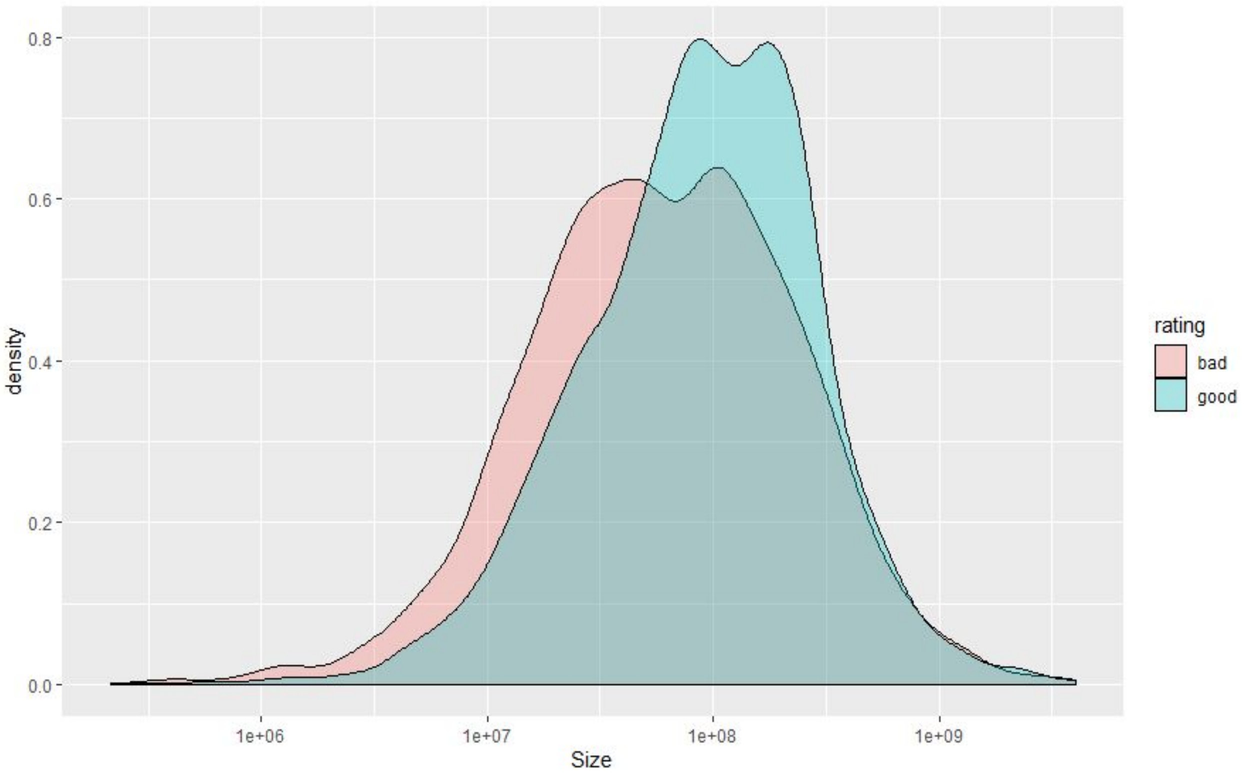
Anuja Dixit

END

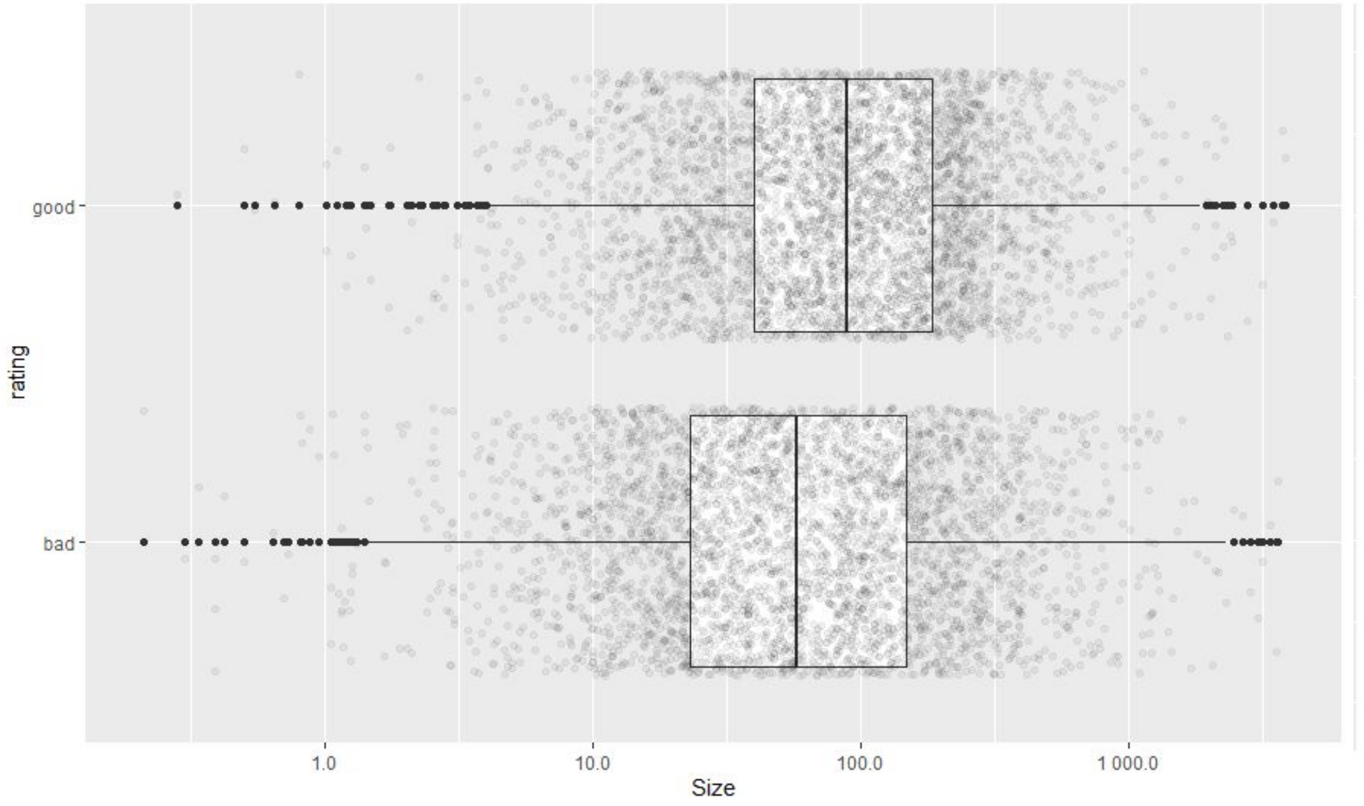
Thanks!

APPENDIX

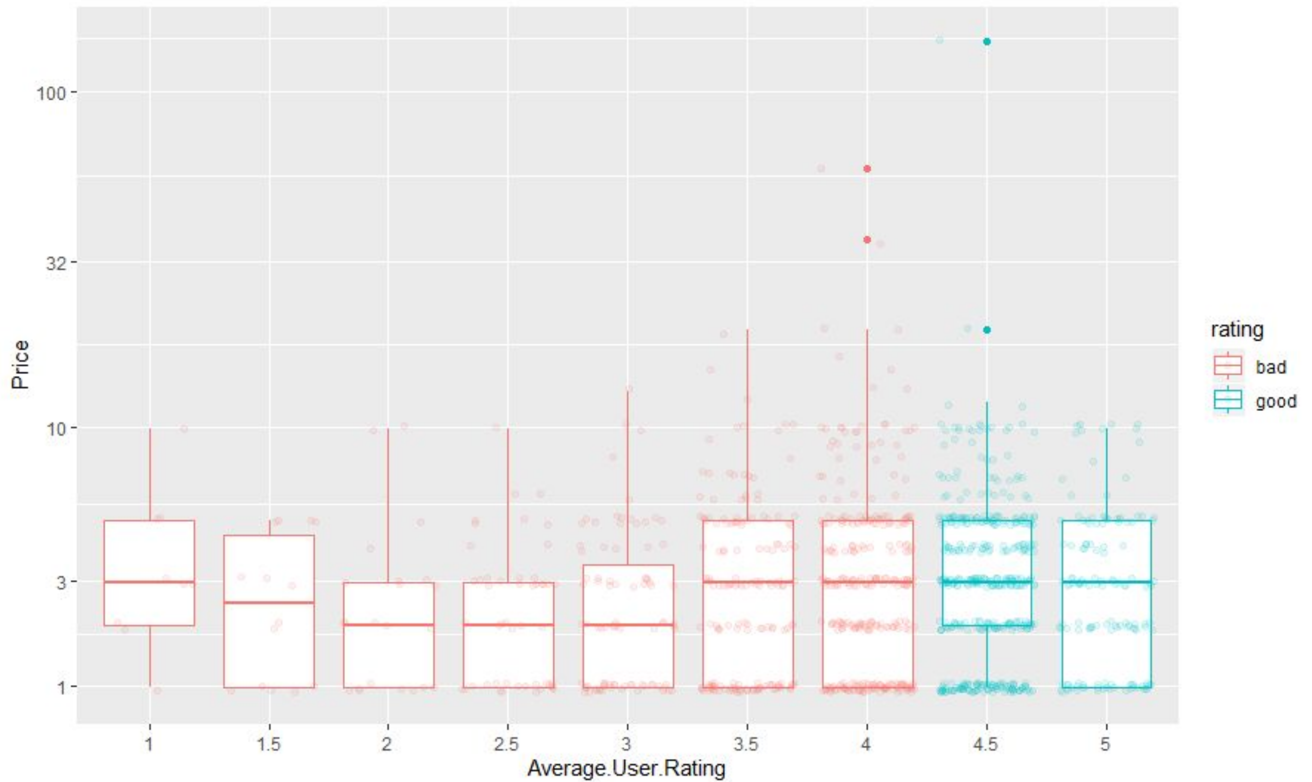
Size Density Graph



Size & Rating



The rating of non-free games



NLP – Lemmatizing the Description

```
['barrels', 'fun', 'simple', 'challenging', 'solitaire', 'game', 'match', 'barrels', 'clear', 'board', 'time', 'runs', 'faste
r', 'higher', 'score', 'featuring', 'difficulty', 'levels', 'complete', 'game', 'instructions', 'barrels', 'fun', 'sure', 'kee
p', 'entertained', 'bus', 'classes', 'whenever', 'wherever', 'go', 'challenge', 'friends', 'see', 'get', 'highest', 'score', 'g
et', 'look', 'gameplay', 'check', 'play', 'video', 'website']
```

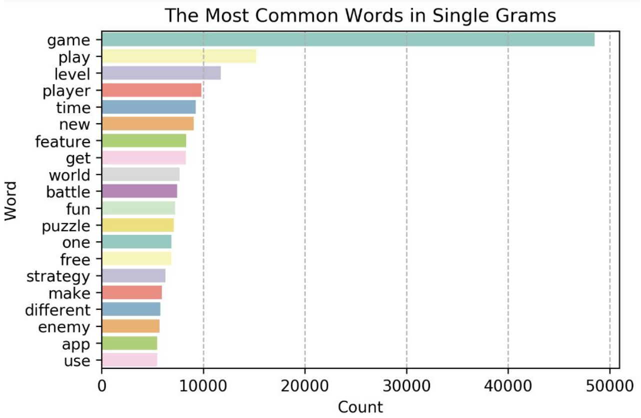
Before Lemmatizing

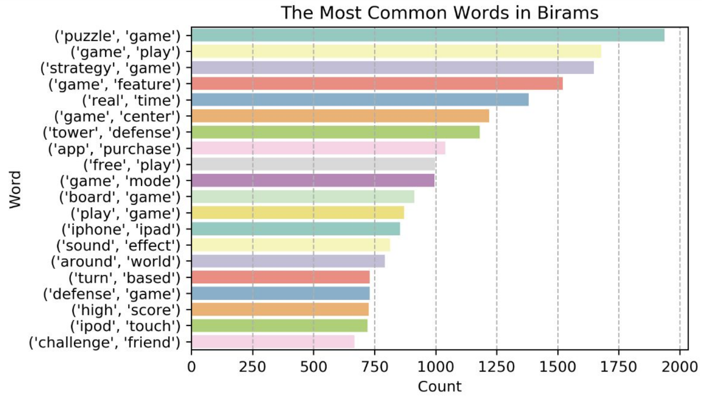
After Lemmatizing

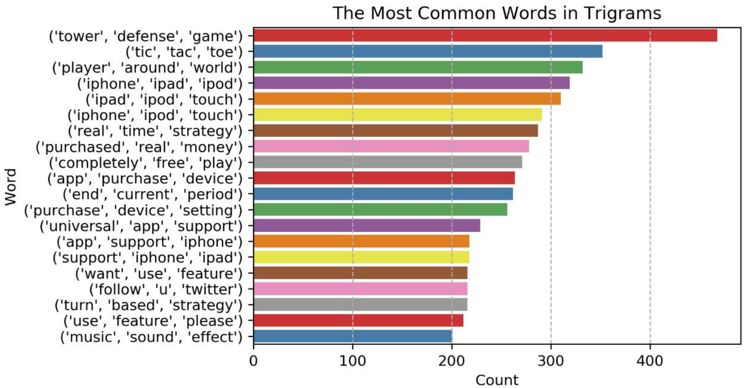
```
['barrel', 'fun', 'simple', 'challenging', 'solitaire', 'game', 'match', 'barrel', 'clear', 'board', 'time', 'run', 'faster',
'higher', 'score', 'featuring', 'difficulty', 'level', 'complete', 'game', 'instruction', 'barrel', 'fun', 'sure', 'keep', 'ent
ertained', 'bus', 'class', 'whenever', 'wherever', 'go', 'challenge', 'friend', 'see', 'get', 'highest', 'score', 'get', 'loo
k', 'gameplay', 'check', 'play', 'video', 'website']
```

```python
def tk(description):
    description = rm_lower(description)
    token = word_tokenize(description)

    stopwords_en = set(stopwords.words('english')) # set of stopwords

    token = [word for word in token if word not in stopwords_en]
    token = [WordNetLemmatizer().lemmatize(word) for word in token]

    return token
```

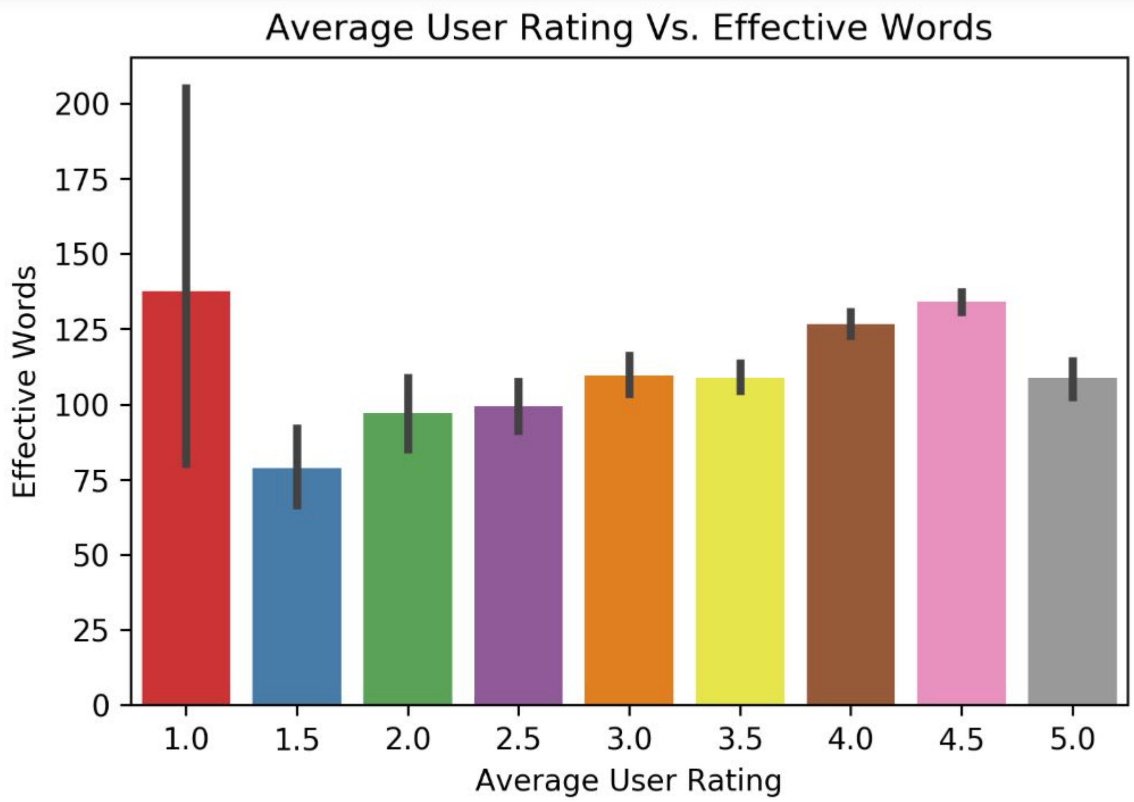
NLP – Birams & Trirams





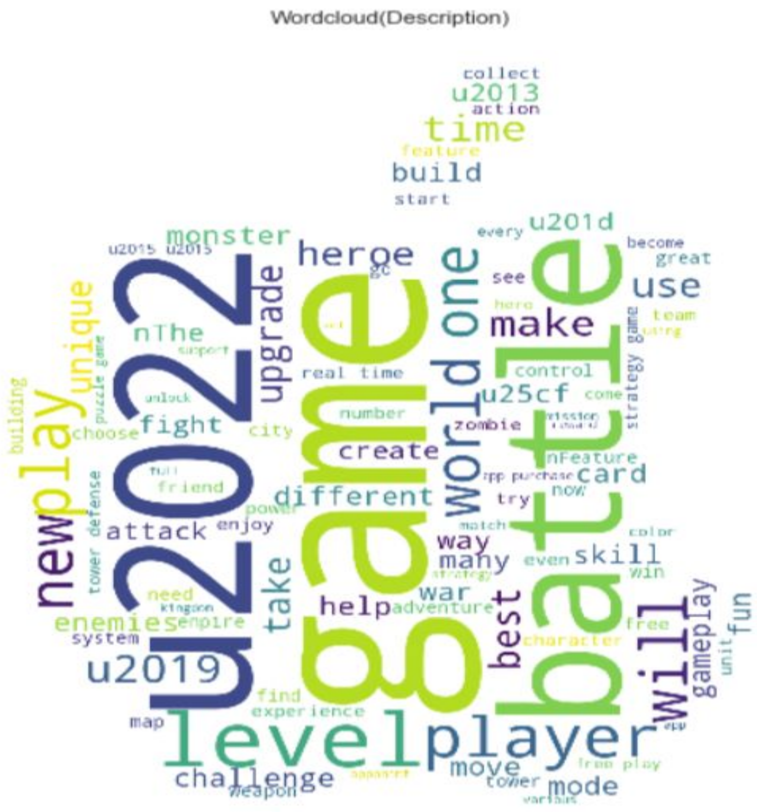


NLP – Impact of Effective Words



WORDCLOUD DESC COMPARISON

Average User Rating > 4



Average User Rating < 4

NLP – WordCloud after Lemmatizing