Chloe Quinto
CS581
HW 9
I pledge my honor that I have abided by the Stevens Honor System - Chloe Quinto

For this assignment, I wanted to explore the theme of Dunbar's Number from the article "Dunbar's number, Why can we only maintain 150 relationships" (Ro, 2019). The theory of Dunbar's number holds that we can only really maintain about 150 connections at once. But is the rule true for today's world of social media?
I begin my program by showing the demographics of this data set. For example,

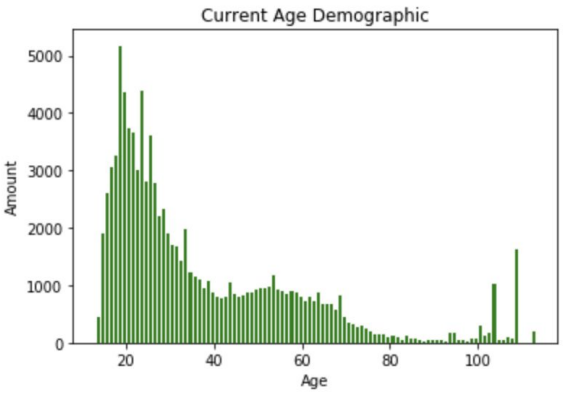
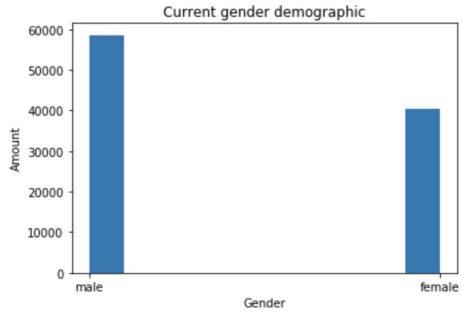


Figure 1




```
Amount of females:
male      58574
female    40252
Name: gender, dtype: int64
```

Figure 2

Our data set shows that the majority of our users are between the age of 10 and 30 years old. Dunbar's own research shows generational differences in terms of social networks. In the article, they mention that those between the ages of 18-24 have a much larger online social network than those aged 55 and above in which *Figure 1* reflects this idea clearly. In terms of gender, the majority of our users are male with around 58,574 and females of about 40,252.

Throughout the notebook, I explore different ideas such as the Number of Friends (Figure 3), Correlation between Number of Friends and Likes Received (Figure 4), If you are more likely to initiate friendships, are you more likely to receive likes? (Figure 5), Do females or males get more likes (Figure 6).

Number of Friends

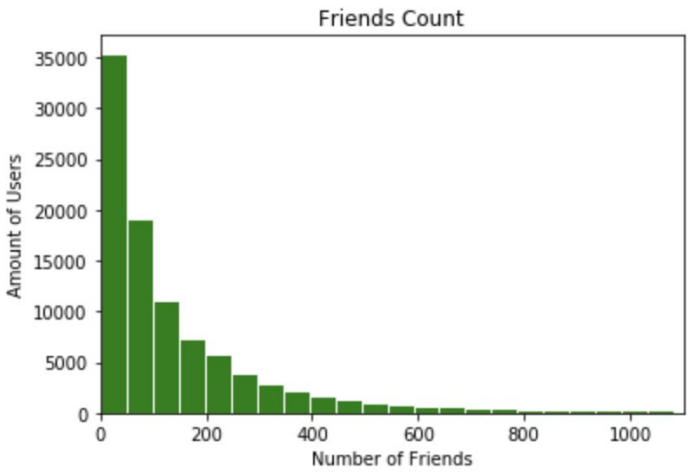


Figure 3

Most users in the dataset have less than 200 friends which is around Dunbar's number. Furthermore, I wanted to see if there's a correlation between the number of friends that one has to the number of likes they receive. The motivation behind this question is that my friend group and I do not use Facebook to post content and like content. As a matter of fact, I know that my age group mostly use Facebook messenger to connect with friends. I do not expect to see a strong correlation between the number of friends you have and the amount of likes you receive.

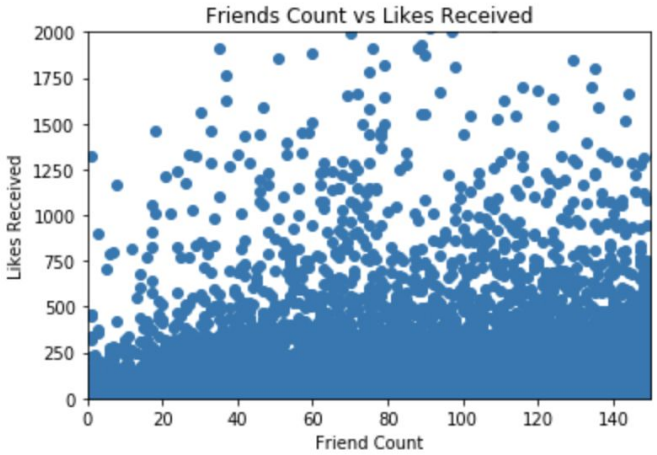


Figure 4

I applied Stats Model Library - Ordinary Least Squares Regression to the graph.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:          likes_received   R-squared:                       0.056
Model:                            OLS   Adj. R-squared:                  0.056
Method:                 Least Squares   F-statistic:                     5854.
Date:                Thu, 09 Apr 2020   Prob (F-statistic):               0.00
Time:                        16:01:17   Log-Likelihood:            -8.5252e+05
No. Observations:               98826   AIC:                         1.705e+06
Df Residuals:                   98824   BIC:                         1.705e+06
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const        -23.8062      4.813     -4.946      0.000     -33.240     -14.373
friend_count   0.8477      0.011     76.509      0.000       0.826       0.869
==============================================================================
Omnibus:                   414718.512   Durbin-Watson:                   1.751
Prob(Omnibus):                  0.000   Jarque-Bera (JB):    1511916308714.215
Skew:                         119.318   Prob(JB):                         0.00
Kurtosis:                   19163.207   Cond. No.                         487.
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

```
print("y="+str(results.params[1]) + "*x + " + str(results.params[0]))
```

```
y=0.8477273439786176*x + -23.806203658600637
```

```
print("R2: ", results.rsquared)
```

```
R2:  0.0559210921220894
```

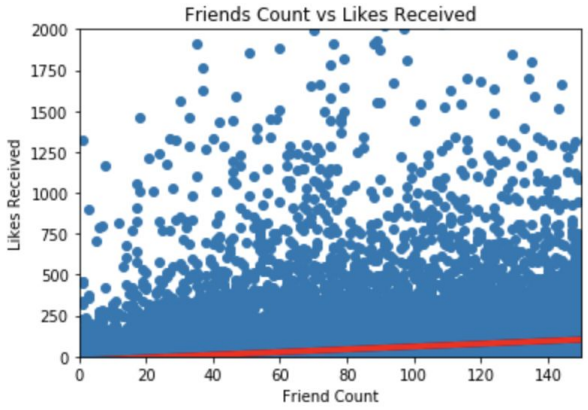


Figure 5

It seems that there is a weak correlation between the number of friends and the likes you receive. We can't confidently say that the more friends you have, the more likely you are to receive likes. This may be due to the fact that a lot of Facebook users are "lukers", those who do not post, comment or like. Lurkers just read posts and comments and do not react. Another possible reason is the nature of Facebook is not one that is focused on giving and receiving likes and comments (unlike Instagram).

Furthermore, I wanted to see if you are more likely to initiate friendships, are you more likely to receive likes?

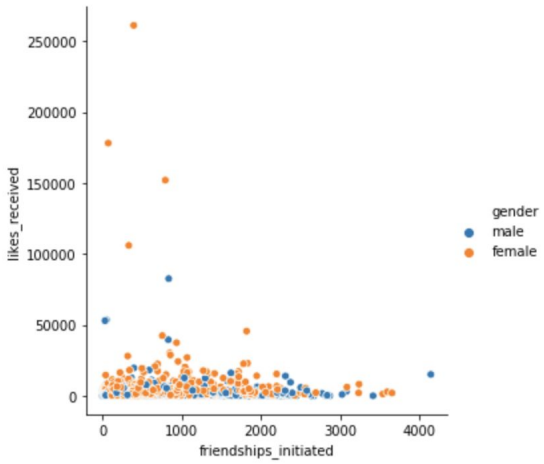


Figure 6

In Figure 6, it looks like the more friends you try to initiate, the less likely you are to receive likes on your post. Dunbar's theory says that you should have 500 acquaintances and 1500 people you recognize. Therefore, it's not surprising that if you have connections of more than 1500, they are very much weak connections.

I wanted to verify this statistically:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:         likes_received   R-squared (uncentered):                   0.041
Model:                            OLS   Adj. R-squared (uncentered):              0.041
Method:                 Least Squares   F-statistic:                              4201.
Date:                Thu, 09 Apr 2020   Prob (F-statistic):                        0.00
Time:                        17:51:40   Log-Likelihood:                     -8.5383e+05
No. Observations:               98826   AIC:                                  1.708e+06
Df Residuals:                   98825   BIC:                                  1.708e+06
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                             coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
friendships_initiated     1.2976      0.020     64.817      0.000       1.258       1.337
==============================================================================
Omnibus:                   411466.435   Durbin-Watson:                   1.735
Prob(Omnibus):                  0.000   Jarque-Bera (JB):    1386787522729.464
Skew:                         116.311   Prob(JB):                         0.00
Kurtosis:                   18353.170   Cond. No.                         1.00
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

```
: print('R2: ', results.rsquared)

  R2:  0.040778686475535064
```

Surprisingly, I was wrong. There is a very weak correlation - 4% variation in the likes received can be explained by the friendship initiated.

Lastly, I wanted to go beyond Dunbar's number and verify a concept that females generally will get more likes on posts than males.

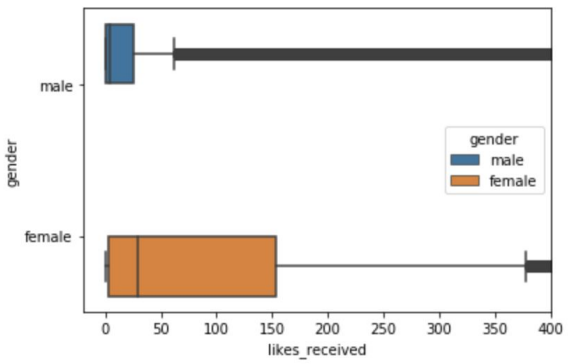


Figure 7

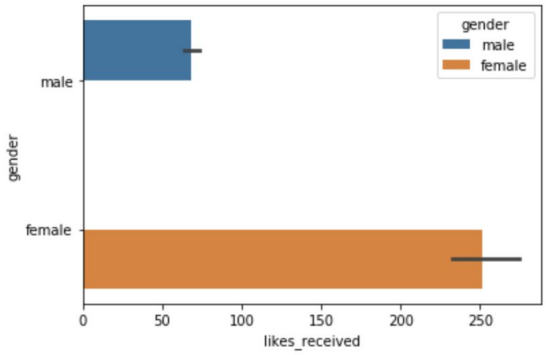


Figure 8

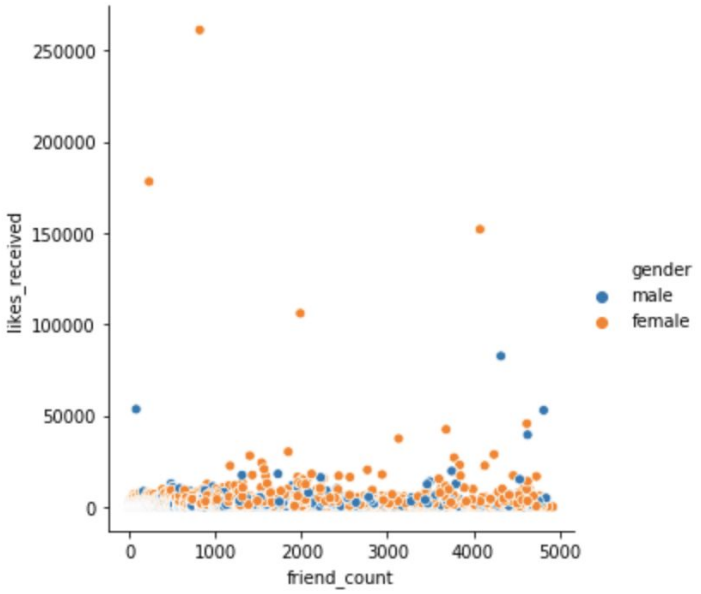


Figure 9

I haven't found any concrete research on this topic as to why women get more likes than men on Facebook. My opinion is that females are more inclined to compliment others irrespective of gender whereas males are less likely to compliment other males due to some other social construct.

In conclusion, the overall motivation for showing these analyses is to get an understanding of Dunbar's number. I found Dunbar's number interesting ever since it was brought up in a discussion on canvas. It was such a specific number (150) that I thought it was not accurate. Therefore, I thought it was a perfect opportunity to verify his numbers. I conducted all of my analysis on a jupyter notebook called main.ipynb.

Works Cited

Ro, Christine. “Dunbar's Number: Why We Can Only Maintain 150 Relationships.” *BBC Future*, BBC, 9 Oct. 2019, www.bbc.com/future/article/20191001-dunbars-number-why-we-can-only-maintain-150-relationships.