Sanchari Chowdhuri
sanchari@umd.edu

Is 150 the Maximum Number of People A Human Can Keep Track?

Robin Dunbar a famous anthropologist argued that humans only have the capacity to keep track of so many people at a time. He argued that the maximum number of people humans could keep track of is 150. This number is called Dunbar's number (Links to an external site.). Some have interpreted this to mean that any individual person has on average about 150 friends.

Some technologists have argued that social media and modern information technology allow individuals to keep track of more people. For example, Facebook algorithmically organizes and presents information about our Facebook friends in a centralized feed lowering the costs associated with keeping in contact with all these people.

Fortunately we can evaluate this claim because our colleague Prof. Vitak has collected data on Facebook usage among college students at a large university. Using this dataset named "Sample data" and the "codebook" which explains the variables evaluate the technologists's claim.

Question 1: Using the theoretical population mean and the sample provided conduct a one-sample t-test (with significance level alpha = 0.05) to determine whether individuals have more friends than would be expected by Dunbar's number.

a. Graph the total Facebook friend values for the sample. What is the shape?

b. State your null and alternative hypotheses.

c. Calculate your t-statistic

d. Look up the probability for the t-statistic

e. What is the effect size?

f. What would you conclude?

g. What are the limitations, if any?

Question 2: Are individuals who have 150 or more friends on Facebook using Facebook in a fundamentally different way from those that have fewer than 150 friends? Use two sample t-tests with a significance level of 0.05 to compare Facebook users who have 150 Facebook friends or more to those with fewer than 150 Facebook friends for the 6 items that make up question 28 "How likely are you to use Facebook to do the following things, either now or in

the future?". Feel free to do the calculations using t.test in R. Treat these Likert items as interval variables.

Hints: 1) Use the ifelse function in R to create a new variable binary variable to categorize users with 150 or more Facebook friends from those with fewer than 150 Facebook friends 2) You will need to run 6 t tests 3) for this problem you don't need to write out the null and alternative hypotheses (although you should know what these are!), 4) for this problem you don't need to calculate effect sizes (although if you have time and want the practice feel free to do so for an extra challenge).

a) Calculate 6 independent t-tests and organize results into a table with independent variable, dependent variable, and results of the statistical tests (i.e. t-statistic, degrees of freedom, and p-value)

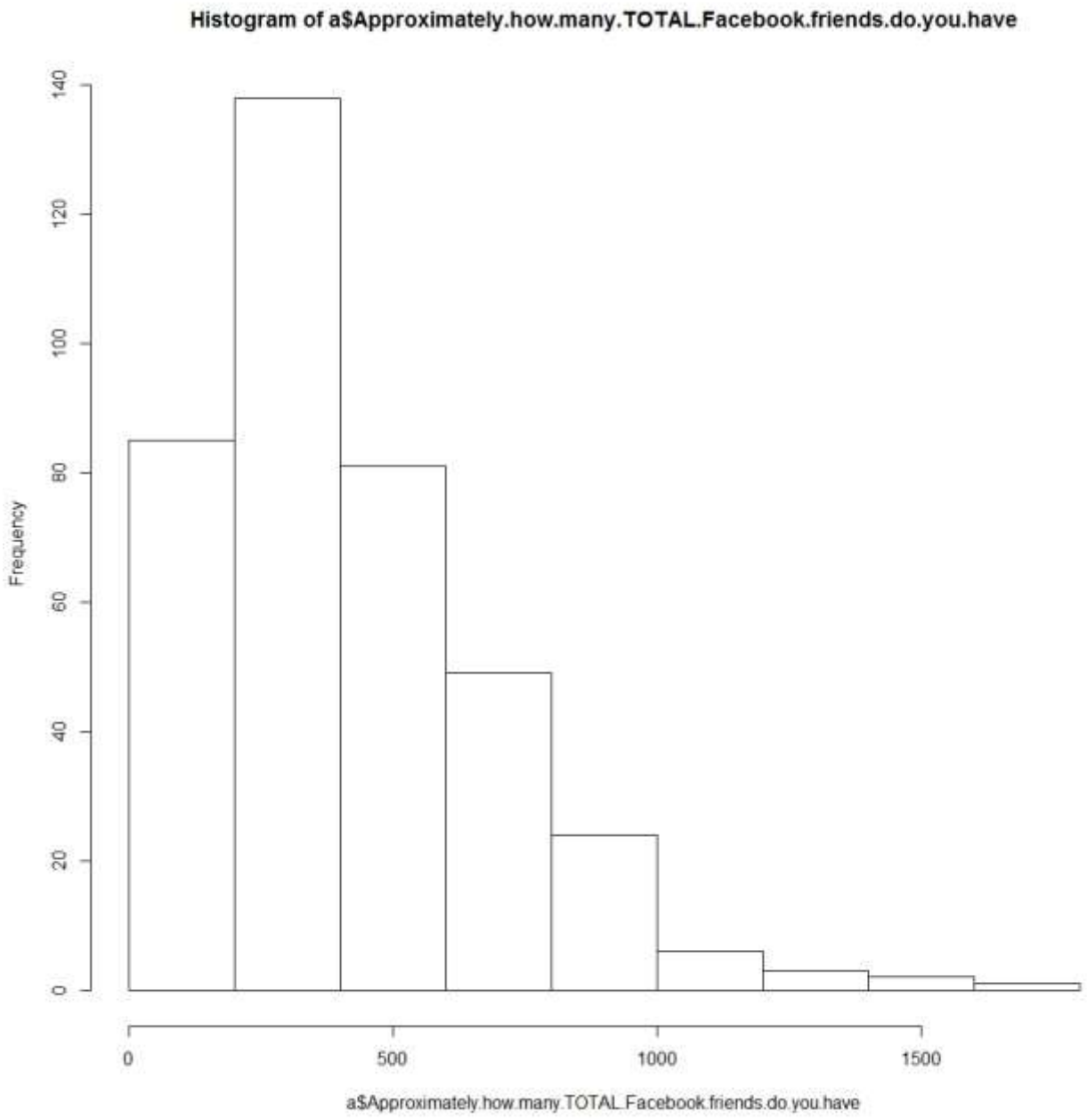
b) Explain any significant differences in how the two groups of users, those with more and fewer Facebook friends, are using Facebook by reporting the means and standard deviations for the two groups for the relevant type of use (e.g. "Those with more friends are more likely to use Facebook to find someone to date (Mean = X, SD = Y) than those who have fewer friends (Mean = Z, SD = W)").

c) What would you conclude?

d) What are the limitations, if any?

STATISTICAL ANALYSIS

Graph the total Facebook friend values for the sample is as follows



The shape of the curve is positively (right) skewed.

Establishing Hypothesis:

Null Hypothesis: No effect (individuals do not have more friends than would be expected by Dunbar's number.)

H_0 = 150 friends

Alternate Hypothesis: individuals have more friends than would be expected by Dunbar's number.

$H_0 \neq$ 150 friends

In the given sample students consider only a certain percentage of their "Facebook Friends" as actual friends. Hence under the assumption that people would like to keep track of only their actual friends and not unwanted people who happen to be there on the friends list. Actual friends can be obtained by (percentage of people considered as actual friends * Number of people on Facebook friend list)

To calculate T statistics

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{N}}}$$

$\bar{x}$ = Mean of actual friends from the sample,

To find this firstly we multiply the "*Approximately what percentage of your TOTAL Facebook friends do you consider to be actual friends*?" with "*Approximately how many TOTAL Facebook friends do you have?"* and then find mean $\bar{x}$ =211.587

$\mu = mean\ of\ population\ (Dunbar\ Number)$= 150

N=389

s= standard deviation of sample data (actual friends) =206.879

Substituting all these values in the above equation we get

t= 5.958

Next to find probability of T statistics. Following command is used in R

```
pt(abs(5.958),389,lower.tail=FALSE)*2
```

```
9.36e-09
```

Probability of obtaining a sample statistic as extreme as ours under the null hypothesis is extremely unlikely

Effect Size: It is calculated using Cohen's d

$$d = \frac{\bar{x} - \mu}{\sigma}$$

$$d = \frac{211.587 - 150}{206.879}$$

d=0.297

Comparing d with Cohen's Rule of Thumb:
|d| = 0.2 - 0.4 is a "small" effect
|d| = 0.4-0.8 is "medium" effect
|d| = > 0.8 is a "large" effect.

The effect size is **small**
Hence we reject the null hypothesis.

Conclusion: With the advent of technological benefits people can keep track of more number of people than the Dunbar Number however this effect is not a substantially large effect.

Limitations:

- The given dataset has many missing values of friends hence those values have been omitted result ing in a slightly smaller value of dataset which affects the result.
- I have assumed that for a given individual all his friends are on Facebook, however there may be some friends who are not on Facebook and he considers them close friends and does keep track of them. Thereby skewing the result.

To find Facebook usage variation of people with more than 150 friends and less than 150 friends we do a two sample t test. The p value obtained tells us how much there is a likelihood for that type of extreme statistical value to obtained.
Descriptive statistics are obtained by finding the sum and standard deviation of the two groups namely *less than 150 friends* and *greater than or equal to 150 friends.*

Following page consists of a tabular representation of two sample T test (t statistics, degree of freedom and p value) and descriptive statistics value

Reason to use Facebook	**Independent (IV) & Dependent variables (DV)**	**Descriptive statistics**	**Two sample t-tests values**
I use Facebook to find people to date.	IV- Number of friends on Facebook DV-Number of people who use Facebook to find people to date	$\mu_{mean\ of\ 150\ or\ greater} = 1.667$ $\sigma_{sd\ of\ 150\ or\ greater} = 0.927$ $\mu_{mean\ of\ less\ than\ 150} = 0.632$ $\sigma_{sd\ of\ less\ than\ 150} = 0.877$	t-statistics= -1.9131 degree of freedom=384 p-value=0.0564
I use Facebook to meet new people.	IV- Number of friends on Facebook. DV-Number of people who use Facebook to find new people to meet	$\mu_{mean\ of\ 150\ or\ greater} = 2.101$ $\sigma_{sd\ of\ 150\ or\ greater} = 1.165$ $\mu_{mean\ of\ less\ than\ 150} = 1.578$ $\sigma_{sd\ of\ less\ than\ 150} = 1.003$	t-statistics= -2.656 degree of freedom=381 p-value=0.00823
I use Facebook to learn more about other people in my classes.	IV- Number of friends on Facebook. DV- Number of people who use Facebook to learn more about other people in their classes.	$\mu_{mean\ of\ 150\ or\ greater} = 2.052$ $\sigma_{sd\ of\ 150\ or\ greater} = 1.107$ $\mu_{mean\ of\ less\ than\ 150} = 2.0769$ $\sigma_{sd\ of\ less\ than\ 150} = 1.178$	t-statistics= 0.1325 degree of freedom=383 p-value=0.8949
I use Facebook to learn more about other people living near me.	IV- Number of friends on Facebook DV- Number of people who use Facebook to learn more about other people living near them	$\mu_{mean\ of\ 150\ or\ greater}$ 2.988 $\sigma_{sd\ of\ 150\ or\ greater} = 1.190$ $\mu_{mean\ of\ less\ than\ 150} = 2.263$ $\sigma_{sd\ of\ less\ than\ 150} = 1.427$	t-statistics= -3.4261 degree of freedom=382 p-value=0.0006
I use Facebook to keep in touch with my old friends.	IV- Number of friends on Facebook. DV-Number of people who use Facebook to keep in tou ch with their old friends	$\mu_{mean\ of\ 150\ or\ greater} = 2.971$ $\sigma_{sd\ of\ 150\ or\ greater} = 1.202$ $\mu_{mean\ of\ less\ than\ 150} = 2.243$ $\sigma_{sd\ of\ less\ than\ 150} = 1.256$	t-statistics= -3.4195 degree of freedom=381 p-value=0.0006952
I use Facebook to find people to add to my "friends" list.	IV- Number of friends on Facebook. DV-Number of people who use Facebook to keep in touch with their old friends	$\mu_{mean\ of\ 150\ or\ greater} = 4.667$ $\sigma_{sd\ of\ 150\ or\ greater} = 0.661$ $\mu_{mean\ of\ less\ than\ 150} = 4.263$ $\sigma_{sd\ of\ less\ than\ 150} = 0.89$	t-statistics= -3.4465 degree of freedom=382 p-value=0.0006311

By looking at this table one does realize that not a huge difference is noted in trends of utilization of Facebook based on number of friends one has. The p value of "*I use Facebook to find people to date.*" and "*I use Facebook to learn more about other people in my classes*." is comparable to 0.05 of significance level. However the effect size is not very large. Also for the other variables we can easily reject the null hypothesis since value of p is very small compared to 0.05

Thereby we concluded that mostly there is not much of a difference in Facebook usage among groups of people with 150 or more friends and less than 150 friends. Facebook usage is alike among people irrespective of Dunbar's number.

Limitation:

- Some of the data values were missing hence they had to be deleted resulting in varied outputs.
- All the 6 variables basically aims towards answering if friends with more than 150 friends are more social outgoing or not based on their Facebook usage however its difficult to understand which of these variables are of more importance to decide.