Objectives

AT&T U-verse offers an internet plan titled Internet 45, which boasts download speeds of 45 megabits per second. Many households have been noticing a significant decrease in internet speeds as people stay home and practice self-isolation with their families. The goal of this project is to determine whether AT&T’s promise of 45 mbps is accurate in a time when household internet usage is at unprecedented levels, as a large portion of the workforce in this area practices work-from-home.

Data Description

Data was collected using a Unix tool created by Ookla (the makers of speedtest.net) called speedtest-cli to calculate download and upload speeds. Data was analyzed using a python script that parses through the outputted csv file from the speedtest tool and calculates the p-value for a one sample t-test on the download speeds of 50 samples. The following shell script was executed 5 times throughout a single day (11 am, 2 pm, 5 pm, 7 pm, 9 pm).

for i in {1..10}

do

speedtest-cli --no-upload --csv >> speedStats.csv

done

Data

Here is the raw data from the speedStats.csv file outputted from the shell script. Unrelated information such as ping and distance was trimmed off. Download speeds were recorded in bits per second.

|  |  |
| --- | --- |
| **Timestamp** | **Download (bps)** |
| 2020-04-24T18:17:23.808643Z | 10370202.13 |
| 2020-04-24T18:17:36.115610Z | 14805146.16 |
| 2020-04-24T18:17:48.424280Z | 18410995.06 |
| 2020-04-24T18:18:00.191904Z | 12230830.69 |
| 2020-04-24T18:18:13.767435Z | 12551821.59 |
| 2020-04-24T18:18:26.427796Z | 5944065.169 |
| 2020-04-24T18:18:50.137505Z | 10066436.26 |
| 2020-04-24T18:19:03.266016Z | 10333252.47 |
| 2020-04-24T18:19:15.857288Z | 14000215.94 |
| 2020-04-24T18:19:29.109994Z | 18397860.3 |
| 2020-04-24T21:19:46.665242Z | 18532443.17 |
| 2020-04-24T21:19:58.616773Z | 17089326.96 |
| 2020-04-24T21:20:11.287880Z | 19308881.22 |
| 2020-04-24T21:20:22.876583Z | 19840289.98 |
| 2020-04-24T21:20:34.419973Z | 19872376.54 |
| 2020-04-24T21:20:45.950749Z | 16836229.36 |
| 2020-04-24T21:20:57.690813Z | 15597989.14 |
| 2020-04-24T21:21:09.670197Z | 17624830.92 |
| 2020-04-24T21:21:21.251671Z | 17266926.86 |
| 2020-04-24T21:21:32.765640Z | 17182134.2 |
| 2020-04-24T22:50:22.521213Z | 17285878.24 |
| 2020-04-24T22:50:34.333806Z | 10743184.25 |
| 2020-04-24T22:50:46.186957Z | 9901959.903 |
| 2020-04-24T22:50:58.357951Z | 15628006.7 |
| 2020-04-24T22:51:11.522687Z | 21933161.9 |
| 2020-04-24T22:51:25.333538Z | 18972925.56 |
| 2020-04-24T22:51:37.025196Z | 20928316.56 |
| 2020-04-24T22:51:48.569462Z | 12834168.47 |
| 2020-04-24T22:52:00.178642Z | 18898097.92 |
| 2020-04-24T22:52:11.729411Z | 18247448.24 |
| 2020-04-25T02:10:25.622405Z | 12603808.42 |
| 2020-04-25T02:10:37.861037Z | 13097988.12 |
| 2020-04-25T02:10:49.812087Z | 13221658.95 |
| 2020-04-25T02:11:01.669098Z | 13012313.33 |
| 2020-04-25T02:11:14.306134Z | 11101281.97 |
| 2020-04-25T02:11:27.299841Z | 13643028.17 |
| 2020-04-25T02:11:39.679588Z | 11842617.13 |
| 2020-04-25T02:11:53.437022Z | 14445414.54 |
| 2020-04-25T02:12:05.940415Z | 11321228.16 |
| 2020-04-25T02:12:18.329567Z | 13817300.83 |
| 2020-04-25T04:47:55.154824Z | 23415187.32 |
| 2020-04-25T04:48:07.209838Z | 21814081.98 |
| 2020-04-25T04:48:19.829953Z | 14296264.6 |
| 2020-04-25T04:48:31.749259Z | 20953066.12 |
| 2020-04-25T04:48:43.793606Z | 22136297.18 |
| 2020-04-25T04:48:55.570767Z | 21792430.53 |
| 2020-04-25T04:49:07.383210Z | 22393780.3 |
| 2020-04-25T04:49:19.000851Z | 22861141.62 |
| 2020-04-25T04:49:31.115455Z | 24247248.23 |
| 2020-04-25T04:49:42.771535Z | 23280613.83 |

Null and Alternative Hypotheses

The goal of this project is to determine whether internet speeds collected during a data of normal data usage is less than AT&T U-verse’s claim of 45 mbps. Therefore, the null hypothesis for this test was and the alternative hypothesis was . The shell script recorded download speeds in bits per second, so to keep the units consistent, the hypotheses were also defined in bits per second. The null hypothesis is that the distribution of download speeds recorded during a single day has a mean equal to 45,000,000 bits per second, which is what AT&T U-verse claims. The alternative hypothesis is that the sample mean is less than 45,000,000 bits per second, which means the internet provider was underdelivering at the time of the test.

Test Statistic, Conclusion, P-value

The following python program parses through the csv and calculates the mean, standard deviation, test statistic and p-values. It utilizes the python library scipy to calculate the p-value for a one sample t-test. It reads from the outputted csv file using the pandas library and performs a t-test on the ‘Download’ column of the file.

from scipy import stats

import pandas as pd

def main():

speedTTest(45000000.00, 0.05)

def speedTTest(popMean, alpha):

dataframe = pd.read\_csv('speedStats.csv')

downloads = dataframe['Download'].values.tolist()

testStat = (sampMean popMean)/(np.sqrt(dev\*dev/len(

downloads)))

print("t =", testStat)

p = stats.ttest\_1samp(downloads, popMean)[1]

reject = p < alpha

print("p-value =", p)

if (reject):

print("At a ", (1-alpha)\*100, "% confidence level, we reject the null hypothesis that the sample mean is equal to ", popMean)

else:

print("At a ", (1-alpha)\*100, "% confidence level, we fail to reject the null hypothesis that the sample mean is equal to ", popMean)

main()

The sample mean for the data was 16,338,643.06 bits per second (about 16.34 mpbs) and the sample standard deviation was 4,425,383.20 bits per second (about 4.43 mbps). Using the formula for a one sample t-test, the resulting test statistic was The p-value generated was 6.96 x 10-42. I chose an alpha value of 0.05. Below is the t distribution of the data with the p-value plotted on it. Since the p-value was significantly small, the online graphic utility used to generate this distribution labels the p-value as 0.

A close up of a map

Description automatically generated

Because the p-value was significantly less than the significance level of 0.05, we have strong evidence to reject the null hypothesis. Therefore, this test concludes that from data collected throughout a full weekday, there is significant evidence to reject AT&T U-verse’s claim that it provides internet speeds of 45 megabits per second.

References

Johnston, N. (n.d.). StatDistributions.com. Retrieved from http://www.statdistributions.com/t/