A1. Is the mean outage time per week higher in customers who have churned within the last month?

A2. This is an important question because it could identify a source of churn. The Churn Data Consideration and Dictionary document states that churn can be as high as 25% and that acquiring new customers is much costlier than retaining current ones so figuring out what least to customer churn needs to be a major consideration.

If the key causes for customer churn can be identified then the company can put their resources toward those activities to better retain customers. Stakeholders will benefit because they will know where to direct their energy for the most impact and they can avoid the costly process of obtaining new customers which will improve the bottom line.

A3. The two features needed for answering the question are Outage_sec_perweek and Churn. The Outage_sec_perweek will be divided into two groups based on the Yes/No value of the Churn feature.

**********

B1. Please see attached code.
B2. Please see attached code.
B3. A t-test was chosen because it was an appropriate analysis to use in answering the question. This tests whether a statistically significant difference of means exists between the group that churned and the group that hasn't. The test allows for a simple comparison and the result can help guide the business decision-making process.

**********

C1. Please see attached graph.

**********

D1. Please see attached graph.

**********

E1. The null hypothesis is that the population means from the two groups are equal. The alternative hypothesis is that the means are not equal, with the mean of the churned group being higher than the non-churned group. The result of the test was a statistic very close to zero with about a 98% p-value so in this case the null hypothesis would be retained.

E2. Bias could be a major limitation in this analysis. For example, maybe there are external factors with churn such as a a higher churn rate during a certain time of year, an economic shift, or other telecommunication companies changing their policies. Another consideration is that the given data only shows churn data for the past month, analyzing data for more months could lead to discovering meaningful patterns.

E3. The recommended course of action is to analyze other causes of customer churn. The outage mean of customers who churned in the past month versus those who didn't failed to show a meaningful difference so it is likely there's an alternative reason for the churn. Another recommendation is to analyze customer churn based on more than just the past month's churn rates.

**********

References.

The only materials referenced were the library documentations.

NumPy: https://numpy.org/doc/stable/reference/index.html

Pandas: https://pandas.pydata.org/docs/

Matplotlib: https://matplotlib.org/3.3.3/contents.html

Seaborn: https://seaborn.pydata.org/

SciPy: https://docs.scipy.org/doc/scipy/reference/stats.html