**Final Report of Project**

**Are people not being able to pay for their Electric bill is causing them to lose their place to live?**

Tashnim Khan Ishika and Md Istiaq Ahmed

**Introduction**

During Covid-19, our societal and economical aspects of life were both equally compromised but since then, we have very limited access to the mathematical perspective of the impact. Hence, we both chose the recent period of the Covid-19 survival war to demonstrate whether the effect is still relevant. We chose Percent Likelihood of Eviction or Foreclosure and Percent Unable to Pay Energy Bill in an attempt to prove a correlation among them as it is expected, generally. Then used statistical data and graphs to demonstrate the reality by using the central tendencies.

**Materials and methods**

The whole project was based on the data collected from this website: <https://www.census.gov/data-tools/demo/hhp/#/?measures=>

After the two variables: Percent Likelihood of Eviction or Foreclosure and Percent Unable to Pay Energy Bill, were selected, four exact weeks were selected for both the variables. Depending on the weeks which had data for both the variables, week 35, week 40, week 45 and week 50 were selected. All the data were selected for 50 states of USA excluding District of Columbia. Then depending on the data collected, using Excel, the following were calculated: summary statistics, histograms, comparative analysis, hypotheses, scatterplot and determined the correlation.

**Results and Discussion**

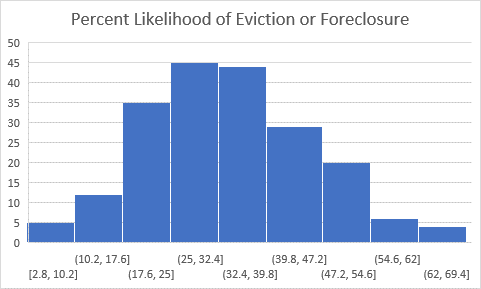
The goal of this project was to figure out a correlation between the two variables: Percent Likelihood of Eviction or Foreclosure and Percent Unable to Pay Energy Bill. We knew that only a scatterplot can identify the actual scenario. Hence, we slowly built up by calculating all the statistical standard graphs and table to prove the reliability on the data as well.

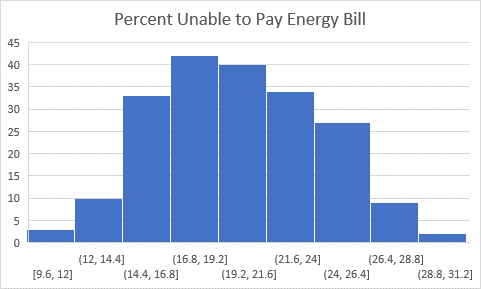
The following is a statistic summary / Five number summary and the central tendencies of both the variables for four weeks:

|  |  |  |
| --- | --- | --- |
|  | Percent Likelihood of Eviction or Foreclosure | Percent Unable to Pay Energy Bill |
| Mean | 33.752 | 20.026 |
| Std. Dev | 12.3729 | 4.047684 |
| Min | 2.8 | 9.6 |
| Q1 | 24.75 | 17.2 |
| Median | 33 | 19.65 |
| Q3 | 42.4 | 23.15 |
| Max | 67.5 | 30.6 |

These comparison side by side showed how the spread of the data as well as the mean is greater for Percent Likelihood of Eviction or Foreclosure than Percent Unable to Pay Energy Bill.

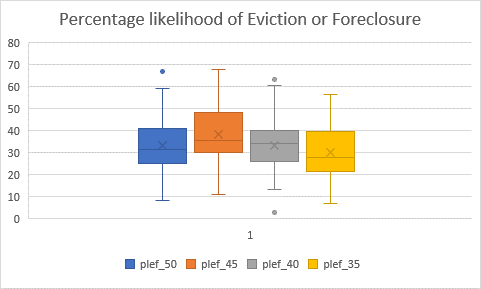
The following are histograms to show how the data are scattered defining whether it is a skewed or a symmetrical distribution.

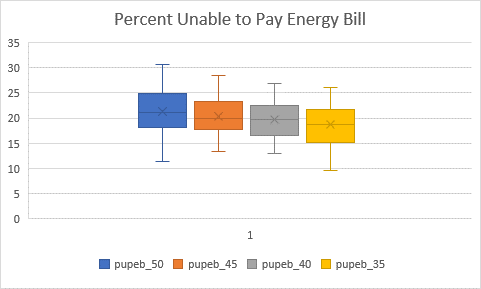




Both the graphs were proved to be slightly symmetrical with most of the data in the middle.

The following are the box plot of the four weeks selected for each variable to represent the difference in data.





The following are two tables used to do a hypothesis on two weeks’ of each variable separately, where both the null hypothesis were rejected proving that with increasing weeks, more people were affected in each case.

|  |  |  |
| --- | --- | --- |
| t-Test: Paired Two Sample for Means |  |  |
|  |  |  |
|  | plef\_45 | plef\_35 |
| Mean | 38.184 | 30.118 |
| Variance | 144.7315755 | 121.3835469 |
| Observations | 50 | 50 |
| Pearson Correlation | 0.172876064 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 49 |  |
| t Stat | 3.842803682 | Test statistic |
| P(T<=t) one-tail | 0.000175236 | P value |
| t Critical one-tail | 1.676550893 | Critical Value |
| P(T<=t) two-tail | 0.000350471 |  |
| t Critical two-tail | 2.009575237 |  |

|  |  |  |
| --- | --- | --- |
| t-Test: Paired Two Sample for Means |  |  |
|  |  |  |
|  | pupeb\_45 | pupeb\_35 |
| Mean | 20.284 | 18.672 |
| Variance | 12.96300408 | 16.27797551 |
| Observations | 50 | 50 |
| Pearson Correlation | 0.486646083 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 49 |  |
| t Stat | 2.933071129 | test statistic |
| P(T<=t) one-tail | 0.002545941 | P value |
| t Critical one-tail | 1.676550893 | Critical Value |
| P(T<=t) two-tail | 0.005091881 |  |
| t Critical two-tail | 2.009575237 |  |
|  |  |  |

Ultimately, the scatter plot proved there is no association as all the data were found to follow no pattern. Additionally, the correlation coefficient was close to zero, showing the association was reasonably weak.

**Conclusions**

The whole purpose of the project came to a halt when the scatterplot was used to demonstrate the data of both the variable. Irrespective of what we assumed through the initial data, we were stunned to identify that scatter plot of both the variables together and individually depending on the 4 weeks, were all showing no association. Additionally, the correlation coefficient was close to zero which proved that the association is significantly weak. Therefore, this drew the conclusion that even though there seems like a possible association between Percent Likelihood of Eviction or Foreclosure than Percent Unable to Pay Energy Bill due to the affect of pandemic on every household but there can also be several other issues that can drive the percent to a different direction than Covid-19.