30 Day Analysis of Medicare Mortality and Readmission from Heart Attacks

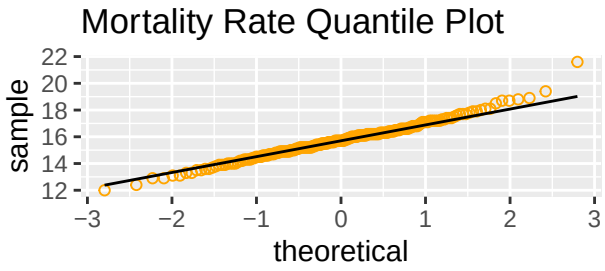
Group 2

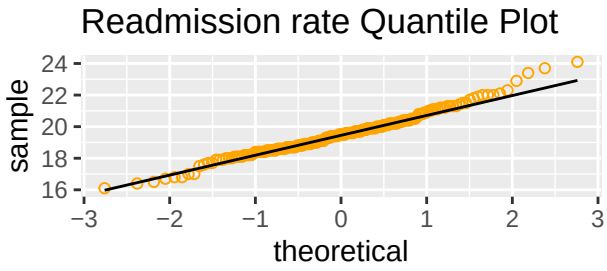
INTRODUCTION:

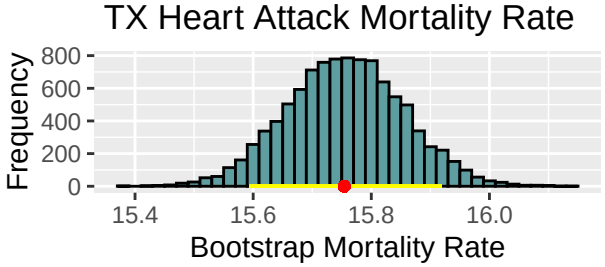
The following project analyzes hospital outcome of measures with data given from Medicare website of the United States of America (US). It specifically analyzes TX and United States on outcomes of measures such as mortality rates from heart attacks and readmission rates from heart attacks within a 30-day period. Particular outcomes of measures are determined by using methods such as the bootstrap percentile method and plotting using the quantile plot to analyze Texas mortality rates and readmission rates due to heart attack to determine the true mean. Methods also include Hypothesis Testing to make credible the difference in means between TX and US. Furthermore, permutation distribution to determine a precise result between the US and Texas mortality and readmission rates caused by heart attacks and using correlation hypothesis tests to check whether readmission to the hospital from a heart attack and death from a heart attack are in fact correlated. The project is significant because reducing hospital readmissions—especially those that result from poor inpatient or outpatient care—has long been a health policy goal because it represents an opportunity to lower health care costs, improve quality, and increase patient satisfaction.

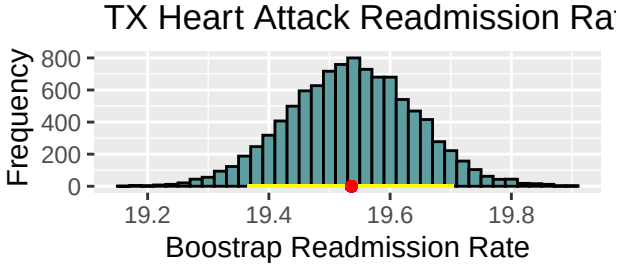
Phase 1 : Confidence Interval for Mean Mortality and Readmission Rate in TX

The classical t-test requires the data to come from a normally distributed population. The quantile plot reveals non-normally, therefore we will use the bootstrap percentile method to construct a 95% confidence interval for the true mean 30-day mortality and readmission rate from heart attack in Texas. After careful analysis we conclude with 95% confident that the true mean for 30-day mortality and readmission rate in Texas lies within the intervals (15.59381, 15.91755) and (19.36474, 19.70520) respectively. The confidence interval for each statistic is highlighted in yellow on the respective histograms and the observed statistic is represented by the red dots. Given the confidence intervals, we are interested in comparing the 30-day mortality rate from heart attack in Texas versus the entire United States of America.









Phase 2 - Hypothesis Testing - US vs TX Mortality Rate

Once again, a histogram and quantile plot of US mortality rate from heart attach reveal a slightly skewed distribution. To account for the non-symmetric nature of the data we utilized the permutation distribution of the difference of means to arrive to a more accurate result because the permutation test is robust for hypothesis testing and accounts for non-normal data. The permutation hypothesis test is useful to answer if our observed statistics is a common value if we assume that the null hypothesis is true and if the probability of getting this statistic due to chance is very small.

H0- true mean for 30-day mortality rates from heart attack is the same for Texas and US.

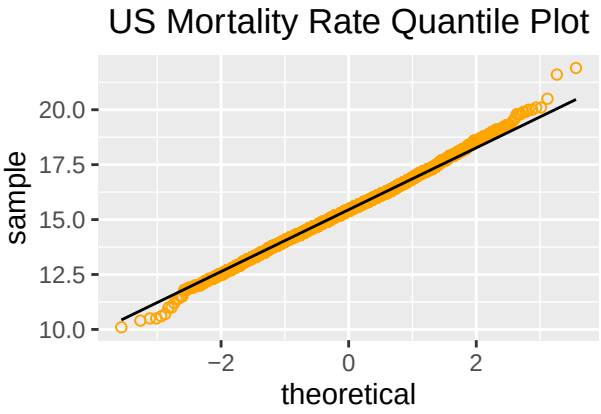
H1- true mean for 30-day mortality rates in US is less than for TX.

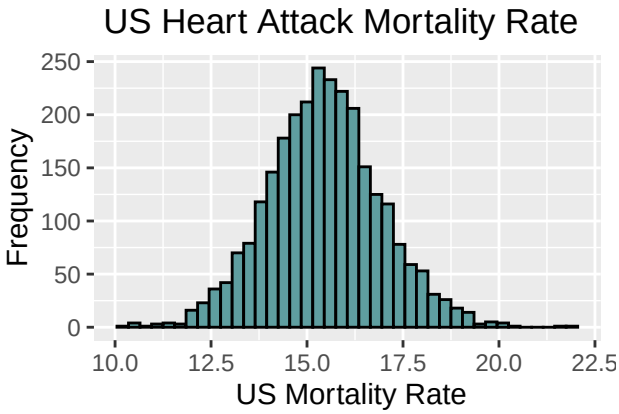
Significance level = 5% is the highest probability we allow for a type I error.

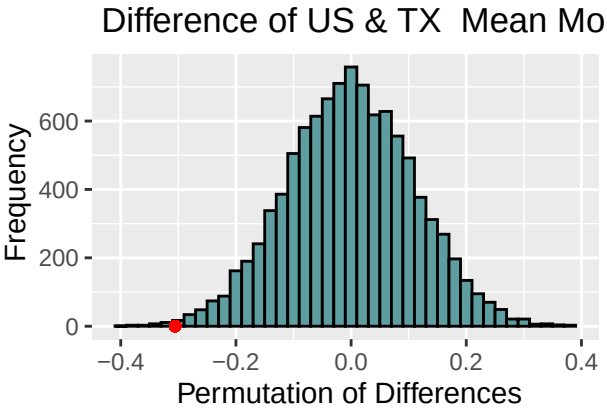
```
## [1] "Observed difference -0.305498848853958"
```

```
## [[1]]
## [1] "Decision: Fail to Reject the Null"
```

```
## [1] "Pvalue 0.0028997100289971"
```







Given the histogram for difference of means, we see clearly that the probability for getting the observed statistics due to chance is very small. In other words, the p-value of 0.002 is smaller than our significance level, 0.05; therefore, we reject the null and conclude that in fact the Texas mean 30-day mortality rate from heart attack is greater than US mean 30-day mortality rate. Our result is statistically significant.

Phase 3 - Correlation Hypothesis Test

Given the nature of the goal of outcome of care measurements, we ask the question of whether the number of readmitted patients is correlated with number of patient deaths of a specific treatment. In other words, are the two variables in question independent of each other? H0- correlation 0. H1 - x and y are correlated.

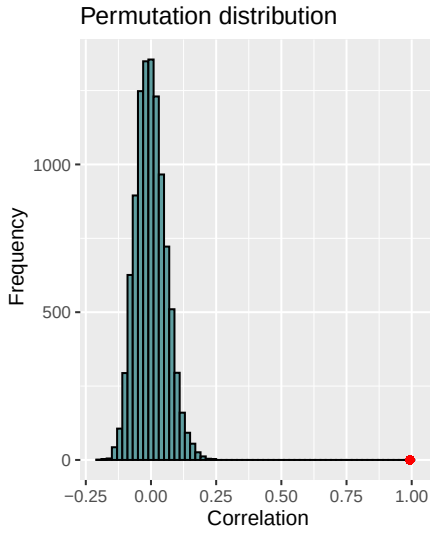
After analyzing the Texas data and testing the correlation between number of death and readmissions, we obtained an extremely small p-value, essentially zero. The probability of getting the observed correlation by change is 0.00019998. At 5% significance level we reject the null hypothesis and conclude that there is a real correlation between number of patients readmitted from heart attack and number of patients that were deceased from heart attacked.

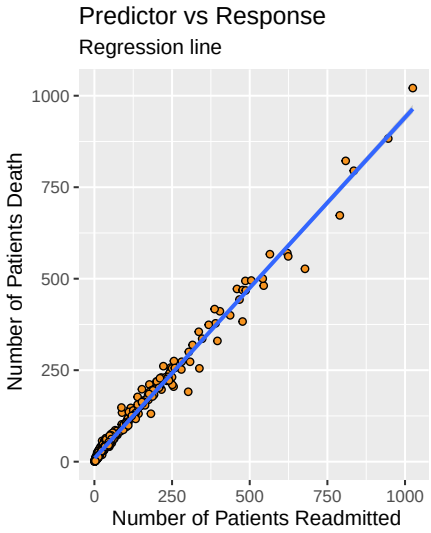
```
## [1] "Observed correlation"
```

```
## [1] 0.9933379
```

```
## [1] "pvalue"
```

```
## [1] 0.00019998
```

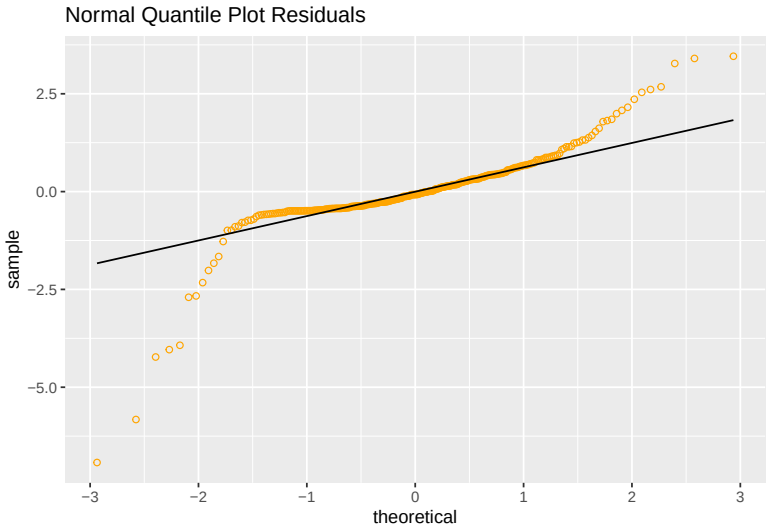


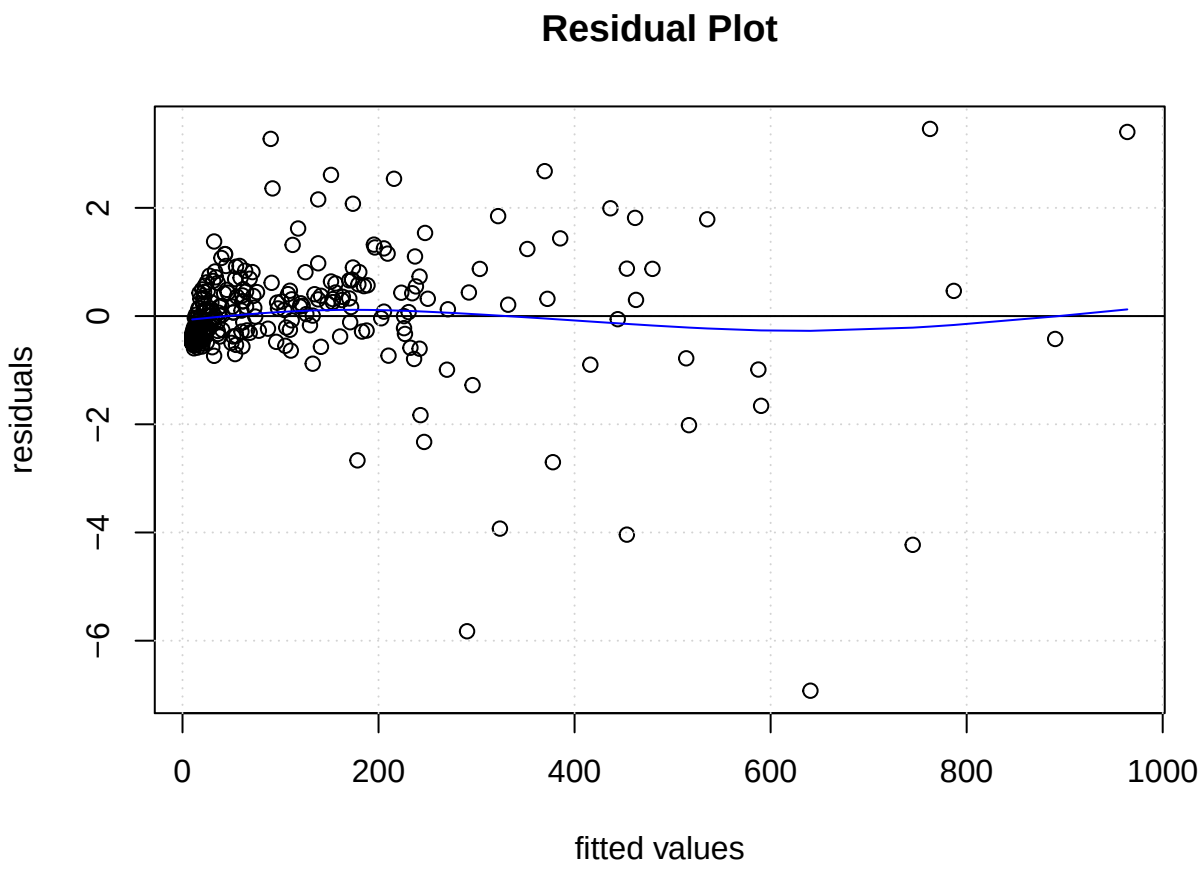


Phase 4 - Regression

Below is the model and analysis of the prediction of the number of patients who died from heart attack given the number of patients who were readmitted for heart attack.

$Model = 0.931672(Number of Patients Readmitted from Heart Attack))$





The residuals are non-normal and do not have constant variance.Furthermore the residual plot reveals a non-random shape which may suggest that a linear model is not appropriate for this instance. This suggests a need for further investigation or additional data. Meanwhile, renders the models unreliable to inform about the significance of the estimates which is a method that assumes that the residuals are normally distributed. For testing we may bootstrap to obtain a better confidance interval for the estimates and prediction for a specific predictor value. However we are still able to use the model to make predictions.

Conclusion:

Outcome of measures is crucial for clinical practices. Heart disease is the leading cause of death in the United States. It is imperative to reduce readmission rates in order to provide lower health care costs and increase theoverall health of society. Moreover, TX and the US was the prime focus of this analysis. Phase I proceededwith a visualization of the normality of the original data. Thus, we concluded that it is best to perform a bootstrap for the true mean 30-day mortality rate and readmission rate for the state of TX. With thisanalysis, we then proceeded to phase II-Hypothesis testing. Hypothesis testing led us to concluded that in fact, United States has a lower 30-day mortality rate than TX. This led us to further analyze Tx data. In phase III, we asked the question, “Is there a correlation between the number of readmission and deceased from heart attacks?” Hence, a Correlation Hypothesis Test was performed. Thus, it became evident that there was a real correlation. Furthermore, we proceeded with Phase IV-Regression. The regressor in this case did not meet all assumptions of normal residuals nor constant variance. However, the model may still be used to make predictions between death mortality and readmission rates from heart attacks. In conclusion, there is high correlation between number of patients deceased and their readmission. Thus, a comparison between states may hold key to determine discrepancies between the data.

References:

Blackwell, Willey. (2019). Why measuring outcomes is important in health care. J Vet Intern Med, Pages 356-362, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6430924/

Chihara, L., & Hesterberg, T. (2019).FMathematical statistics with resampling and R. Hoboken: John Wiley & Sons.

Tinker, Ann. (2018). The Top Seven Healthcare Outcome Measures and Three Measurement Essentials. Health Catalyst, https://www.healthcatalyst.com/insights/top-7-healthcare-outcome-measures

Dataset: https://data.medicare.gov/Hospital-Compare/Hospital-Outcome-Of-Care-Measures/47nj-iqgs/data