**Introduction:**

For this dataset I focused on the homelessness population. The research and data collection were completed by the city of Austin, as well as California State. They were conducted to gauge funding needs and to help guide mitigation strategies for the homeless population within the state of California as well as the city of Austin. The homeless population within these government agencies localities have steadily increased, which made this research relevant and necessary. The two research were conducted by two separate entities. This means that research acquired different data points.

Within this report, I reviewed the two data sets. I gauged what was similar, as well as gauged what was different between the two datasets. One of the data sets focused specifically on the homeless population in Austin (where individuals were followed to see if within 2 years after being homeless and receiving services if they would return to being homeless), the other data set focused on the homeless population in various counties in California (where it measured the emergency room visits of homeless individuals).

**The Subject-Matter:**

Homelessness

**Industry:**

Government

**Objective:**

Within this research report we aim to answer the following questions:

1. What age are most affected by homelessness?
2. What gender are typically homeless?
3. What race is at a high risk of homelessness

**Data Sets:**

City of Austin’s “Strategic Measure Number of Return to Homelessness”

California Counties “Homeless ip and ed by facility”

These data sets are important as they can tell hospital trends as well as the likely hood for a previous homeless individual to return to homelessness. Knowing this data will allow government agencies to readily mitigate homelessness by having a running record of what works and what doesn’t work. Knowing the demographic of people that homelessness impacts can help create a target approach when attempting to eradicate homelessness.

**Answers:**

1. It seems that age 18-30 is the ages where people are the most likely to be homeless.
2. Within the data set, it was said that Asian were more likely to be homeless in Austin, this may be due to data transformation error on my part.
3. Within both research studies, it seems as though, men are more likely to be homeless, than women.

**Tools:**

In my Power Bi dashboard, I included different tables to show visuals of the data collected. I also was able to utilize line charts. All the information was inserted through the sqlite database from two excel files that originated from [www.data.gov](http://www.data.gov).

The homeless population is vastly increasing. It is an epidemic. As the rich get richer, the amount of people in poverty continually increases.

**Implication:**

This dataset can help localities make decision as far as budgets and programs to help mitigate the homeless population. Part of solving the problem is knowing about the problems. Knowing who it affects, at what age and which gender will have a great impact when it comes to removing people off the streets and placing them into homes. When it comes to analyzing the data, Power Bi allows for the visualization and allowed for readers to easily understand the two csv files that contained so many records. It made the data readable and readily accessible.

***Graduate Essay on Data Engineering, ETL and ELT Data integration technique***

Within this class we focused on using Power Bi as a tool to analyze data, as well as a tool to help advise business decisions. We know that Power Bi helps us develop and deliver data to the businesses that is in pursuit of guidance by data. Before focusing on developing and delivering data for any business we must understand what data engineering means, as well as the idea of Extracting, transforming, and loading data or the idea of extracting, loading then transforming data. Within this class we dove into data engineering and understanding the fundamentals of preparing data. We needed to think critically on how we would help businesses utilize the data they send us, as well as the data we sent them.

We have learned that data engineering is a practice where we can design as well as build different systems to collect, store, and analyze data. At times data engineer is compared to data science. Although both can deal with large amounts of data, how the data is utilized is what differentiate the two methods. Data engineering is quite different from data science. Whereas a science tells us why something is the way it is, engineering tells us how something is the way it is. This means that when we are speaking of data engineering, the data is guiding us to understand how the business operates, which can in turn inform the business what may help guide their business towards profitability and lessen customers churn rates.

As the people who will help inform businesses decisions by help them decipher what their data means we must be able to extract, transform and load the data that they collect from consumers. This means that the data may be moved as well as combined with other data within a data warehouse. Data engineers make sure that the data is readily available for dissecting. The data engineers basically prepare the data for analysis for the businesses. Extracting, transforming, and loading the data (ETL) allows for raw data to be collected for different sources.

On a similar scale data engineer at times are task with extracting, loading then transforming the data (ELT). It is very similar in nature to ETL, except transforming the data occurs last. This means that data is still initially read. Loading the data means that the data has already been extracted to the data database second, where as in ETL the data is loaded to the data base last. In the ELT data integration process, the data is transformed last. An example of the data being transformed could be, replacing certain codes with specific values. Another could be converting data types, as well as applying mathematical formulas.

All in all, within this class we learned the process that data engineer uses to analyze and prepare data. Power Bi is a powerful that is used by data engineers to help guide business decisions. The data engineers utilize data integration processes such as ELT and ETL. Within these two processes data can either be loaded after extraction or transformed after extraction. Either method is used by data engineers, specifically based on the need of the company.