**Workers’ Compensation Claims Project**

**Objective**

Tasked with using data analytics to help a client organization to become efficient and responsive to claimants, their claim and transaction data needed to be merged and wrangled. Following, using the data, exploration and visualization was required to find interesting relationships between claim characteristics and outcomes.

**Main Steps in the project**

1. Data Exploration in SAS E Guide
2. Data Wrangling operations that involved understanding and cleaning the data, dealing with missing values, handling inappropriate ranges and outliers, recoding the columns via query builder etc. to obtain a cleaned merged data set.
3. As a part of Text analytics, generated word cloud on the text column to know the possible causes of injury to the workers.
4. Ran linear regression and logistic regression on the final dataset
5. By visualizing the dataset in Tableau, we found that claims that are related to mental illnesses or multiple body parts often take a longer time to process. In terms of the claim procedure, medical services only take up a relatively small portion of the whole claim processing time. Also, claims requiring greater payments require a significantly longer time to process the payment
6. Report the findings such as what are the possible causes of worker’s injury, process duration of claims data, data related to fatal incidents etc.
7. As a part of predictive Analysis, provide strategic recommendations

**Major Challenges**

The biggest data issue in all data files is missing data. The incomplete records hindered us from building a comprehensive view of the claim management of the client. We also encountered obstacles from the invalid ranges, such as negative ages and negative payments, which might be attributed to human error.