**Project Topic:**

**New dataset:** Health Insurance Prediction Costs in USA

**Purpose:** To determine the effect of several factors on insurance costs & predict the cost of health insurance based on factors that influence.

**Analytic Model:** Likely model to be used will be, decision tree regressor forecasting

**Analysis that could be done:**

* Age group analysis
* Sex (male/female)
* BMI
* Smoking

**Questions to ask:**

* Finding correlations between different categories, e.g. bmi vs charges, smokers vs charges, children vs charges
* What categories contributed more to the health cost, if so & why?
* What are the effects of medical/health cost between male & female, cigarettes usage, BMI effects?

**Bullet Points & other misc:**

* Explain the choice of models used for the analysis, why one model was preferred over the other.
* If you tried one model and it failed, why did it fail

**Dashboard presentation (Tableau preferably)**