**UNIVERSITY OF COLORADO DENVER**

**Final Project Proposal**

ASSIGNMENT – 9

BANA - 6800

Avi Manawat (student id: 109638446)

**Problem Statement and Background**

As the' American Association of Suicidology ' tagline states, I strongly believe that suicide prevention is the business of everyone. Keeping in mind that anything can be resolved with the help of adequate support and lifestyle, the act of ending one's own life specifying the reasons for being depression, alcoholism, social reasons or any other mental illness in that matter is not a good idea.

We can choose to stand together in the face of a world that may often feel like a lonely and disconnected place, and we can choose to make a difference by making lives more livable for those who struggle to cope. Through this project, I look forward to identifying the trends in suicidal rates by region, gender, age and ethnicity, and to relate these trends to the possible reasons that lead to the drastic decision that could enable us to curb the thinking at the very start.

**The Data Sources intended to be used:**

The data is taken from Kaggle:

<https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016>

Data on suicides is deficient as there is a problem with the reliability of vital registration data in many countries. Also, there are problems with the accuracy of the official figures made available, since suicide registration is a complicated process involving several responsible authorities with medical and legal concerns.

Moreover, the illegality of suicidal behavior in some countries contributes to under reporting. I am looking to gather more data from different reliable resources in order to conform the accuracy of data.

**Description of Data Analysis Tools Planned to be used:**

Initially, a lot of data cleaning and preprocessing is to be done in order to deal with the missing and anomalies in the data. I will be using Excel for data cleaning and preparation. Next step would be to look out for outliers in the data which may impact visualization. I intend to use R for advance mathematical functions like clustering of countries based on different categories for a better understanding of their relations.

Finally, the visualization will be done using tableau because of its remarkable visualization capabilities and also the fact that it better supports geo spatial maps which would further help in presenting the information.