PYTHON PROJECT PROPOSAL

Introduction:
Patients are very frequently suffering from various diseases, which makes it difficult to analyse the characteristics of combination of diseases. To understand and infer the useful insights from disease combination patterns, it is very important to perform Data analysis on the dataset. Similarly, our goal is to perform such operations on the survey data of New York's patients to understand data in the best possible way and deduce useful inferences. The purpose of this project is to analyse this data and bring out valuable information related to it.

Data Description:
The data informs about the compliance with federal reporting requirements, planning process for local governments with the mental health services and distribution of funding. It is popularly known as The Patient's Characteristics Survey (PCS) conducted every two years, and collects demographic, clinical, and service-related information for each person who receives a public mental health service during a specified one-week period. The data consists of 76 columns which includes Survey Year - Dates for 2022 Patient Characteristic Survey are between 03/21/2022 and 03/27/2022, Age Group - 'Child', 'Adult', 'Unknown', Living Situation - 'Private Residence', 'Institutional Setting', 'Other Living Situation', or 'Unknown', Veteran Status - 'Yes', 'No', or 'Unknown', Mental Illness - 'Yes', 'No', or 'Unknown', so on.

Proposed Analysis:
The analysis of this project will mainly include the work towards achieving any inference from the Patient dataset of New York . Data will help us with information in text and visualization format. Data Visualization is an integral part of Data Analysis. Data Visualization for this project will involve Heat map, Box plot, Count plot etc. Some of the data analysis part will include performing data cleaning and clustering approaches to better understand and analyse data. Further steps on the steps will depend as we move forward with the dataset.

Analysis Methods to be used:
Mainly the Exploratory Data Analysis technique will be used for analysing the data.

Milestones:
The result towards performing Data Analysis on Patient dataset is to infer useful information from Patterns of Patient data having different diseases. The end goal milestone will be achieved with various data processing and visualization techniques.

References:
The data has been considered from the below link –
https://data.ny.gov/Human-Services/Patient-Characteristics-Survey-PCS-2022-Persons-Se/w8eu-45mn

Group members name:
Achal Vats - achalvat@buffalo.edu
Amanjot Singh Sachar - asachar@buffalo.edu
Debparna Bhattacharya - debparna@buffalo.edu
Ekamjot Kaur Khaira - ekamjotk@buffalo.edu