**Project Title:**

Mental Health in Tech Analysis

**Description:**

Surveying Mental Health Trends within the Technology Industry, United States compared to rest of the world

**Research questions:**  
(1) How is mental health in tech affected by employer size (column B)?  
(2) How is mental health in tech affected by age (column BD)?  
(3) How is mental health in tech affected by gender (column BE)?  
(4) How is mental health in tech affected by the perception of consequences for discussing mental health topics (column K)?

(5) Did you hear of or observe negative consequences for co-workers with mental health issues in your previous workplaces?(Column AJ)

(6)Remote Work?(Years: 2016 Column BK / )

**Datasets: (And links to datasets)**

2014:

<https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey>

2016:

<https://www.kaggle.com/datasets/osmi/mental-health-in-tech-2016>

2017:

<https://www.kaggle.com/datasets/osmihelp/osmi-mental-health-in-tech-survey-2017>

2018:

<https://www.kaggle.com/datasets/osmihelp/osmi-mental-health-in-tech-survey-2018>

2019:

<https://www.kaggle.com/datasets/osmihelp/osmi-mental-health-in-tech-survey-2019>

**Tasks:**

**Main Repo:**

Tyler Potts

**Data Cleaning and Formatting:**

**Alex Wiley & Thierno Diallo**

Utilize Pandas to clean and format your dataset(s). This involves tasks like handling missing values, converting data types, and ensuring consistency in formatting.

**Jupyter Notebook for Data Exploration and Cleanup:**

**Alex Wiley & Thierno**

Create a Jupyter notebook documenting the data exploration and cleanup process. This should include code cells with explanations of the steps taken to clean and format the data.

**Jupyter Notebook for Final Data Analysis:**

**Tyler Potts & Kayli Smith**

Develop another Jupyter notebook demonstrating the final data analysis. This notebook should contain code cells showcasing the analysis performed on the cleaned data.

**Visualization:**

**Sylvia Turner, Jeremiah Mergenthaler, Kayli Smith**

Use Matplotlib to generate 6 to 8 visualizations of the data. Aim for at least 2 visualizations per "question" posed to the data. These visualizations should help in interpreting and understanding the dataset.

**Save PNG Images of Visualizations:**

**Sylvia Turner, Jeremiah Mergenthaler, Kayli Smith**

Save PNG images of your visualizations to share with the class and instructional team. These images will also be included in your presentation.

**Write-Up of Major Findings: Sylvia Turner & Kayli Smith**

Create a write-up summarizing your major findings. This should include headings for each "question" asked to the data, along with a brief description of your findings and relevant plots.

**Integration of API Data: Tyler Potts & Alex Wiley**

Incorporate data from at least one API, if relevant to your primary research questions. This could enrich your analysis and provide additional insights into the dataset.

**Dashboard**

**Jeremiah Mergenthaler & Alex Wiley**

Points of Contact:

Thierno Diallo; [thamiru@live.com](mailto:thamiru@live.com)

Jeremiah Mergenthaler; [jeremiah.mergenthaler@gmail.com](mailto:jeremiah.mergenthaler@gmail.com)

Tyler Potts; [twpotts11@gmail.com](mailto:twpotts11@gmail.com)

Kayli Smith; [alexiskayli8@gmail.com](mailto:alexiskayli8@gmail.com)

Sylvia Turner; [sturner6207@gmail.com](mailto:sturner6207@gmail.com)

Alex Wiley; [alessandro.wiley@gmail.com](mailto:alessandro.wiley@gmail.com)