Phase 1: Problem Definition and Design Thinking

Problem Definition:

Introduction:

The problem at hand revolves around evaluating the efficacy of public health awareness campaigns, with a specific focus on mental health, within the tech workplace. We have been entrusted with a valuable dataset sourced from a 2014 survey. This dataset provides insights into attitudes towards mental health and the frequency of mental health disorders among professionals in the tech industry. Our primary objective is to glean meaningful insights from this dataset to gauge the impact of these campaigns and to chart a path forward for future strategies.

Specific Problem Statements:

1. Assess Mental Health Prevalence:
We aim to quantify the prevalence of mental health disorders among employees in the tech workplace based on the dataset.

2. Identify Influential Factors:
We intend to identify demographic and workplace-related factors that significantly influence attitudes towards mental health and treatment-seeking behaviour.

3. Evaluate Campaign Effectiveness:
We need to determine the effectiveness of previous public health awareness campaigns, including their impact on awareness levels, treatment-seeking behaviour, and overall workplace atmosphere.

4. Provide Actionable Insights:
Our ultimate goal is to generate actionable insights that will guide the development of future strategies for mental health awareness initiatives within the tech industry.

Design Thinking:

Step 1: Data Acquisition and Exploration

➢ **Data Collection:** We will begin by accessing the dataset, which contains a wide range of attributes related to mental health, workplace conditions, and individual demographics.
➢ **Data Exploration:** Initial data exploration will be essential to understand the dataset's structure. We will perform checks for missing values, outliers, and assess data quality.

Step 2: Data Preprocessing

➢ **Data Cleaning:** To ensure data quality, we will address missing values through imputation or appropriate handling methods. Any inconsistencies in data formatting will be resolved.
➢ **Feature Engineering**: If necessary, we will create new features or apply transformations to enhance the analysis.
➢ **Data Encoding:** Categorical variables will be encoded into a suitable format for analysis.

Step 3: Data Analysis and Visualization

➢ **Descriptive Statistics:** We will calculate summary statistics for key variables to gain insights into their distributions.
➢ **Visualization:** IBM Cognos will be employed to create an array of visualizations including histograms, bar charts, heatmaps, and geographical plots. These visuals will be instrumental in identifying patterns and relationships within the data.

Step 4: Statistical Analysis and Machine Learning

➢ **Hypothesis Testing:** We will conduct hypothesis tests to ascertain statistically significant differences in attitudes towards mental health across demographic and workplace segments.
➢ **Regression Analysis:** If applicable, regression analysis will be employed to determine predictors of mental health conditions or attitudes.
➢ **Machine Learning (Innovation):** As part of the innovation phase, we may explore machine learning models to predict the success of future campaigns based on historical data.

Step 5: Insights and Recommendations

➢ **Key Insights:** Our analysis will culminate in the identification of key insights pertaining to mental health prevalence, awareness, and campaign effectiveness.
➢ **Recommendations:** We will provide well-informed recommendations for refining and enhancing public health awareness strategies within the tech workplace.

Step 6: Documentation and Reporting

➢ **Detailed Documentation:** Every step of the analysis, including code explanations, data preprocessing, modelling, and findings, will be meticulously documented.
➢ **Comprehensive Reporting:** Our final report will be a comprehensive document incorporating clear and informative dashboards, reports, and visualizations to effectively convey our findings.
➢ **GitHub Repository:** All project files, documents, and code will be neatly organized in our GitHub repository, adhering to the prescribed naming convention "AI_Phase1."

By adhering to this structured approach, we aim to deliver a thorough and comprehensive analysis of public health awareness in the tech workplace, addressing the problem statement in a detailed and effective manner. This report provides the foundation for our subsequent project phases, leading to data-driven insights that will be instrumental in shaping future strategies.