**Day-1**

**Project Scope Definition:**

**Title: Healthcare Department Data Processing**

1. **Analyzing 2011 census data:** The purpose of this part is to analyze the 2011 census data of healthcare department in order to get relevant data. However this large dataset contains a lot of information which is irrelevant and can be ignored. Hence the objectives includes:

* Cleaning the data by removing irrelevant columns
* To form a uniform nomenclature across various datasets to ensure consistency
* To handle the changes in names of new states and union territories in order to match the current data
* To handle missing values by using information from other fields

1. **Extract Information from the housing data along with the census data:** The purpose of this part is to analyze housing data along with census data to extract additional insights for healthcare department. The objectives includes:

* Integrating housing data with census data using common fields
* Cleaning and processing the data
* To explore the correlations ,trends between the datasets

1. **Clean and gain Insights from the Hospital Data:** The purpose of this part is to analyze the hospital data to gain insights of hospital beds availability in different states and union territories. The objectives include:

* Cleaning and processing the data by fixing the header and ensuring uniformity in states/UT names.
* Analyzing and visualizing the number of hospital beds in each state or union territory.

1. **Government Hospitals Requirement Analysis:** The purpose of this part is to analyze government hospitals data to recommend for establishment of new hospital in a state or union territory. The objectives include:

* Fix the header, state/UT names and standardize the date formats in the hospital data for uniformity.
* Identify three states or union territories with least number of hospital beds per population and recommend the establishment of new government hospital.

1. **World Health Organization Standards:** The purpose of this part is to analyze the hospital data to determine if different regions in India meet the WHO standards for hospital bed availability. The Objectives include:

* To calculate the expected number of hospital beds required based on WHO standards.
* Visualize the gap between expected and available number of beds in each region.
* Calculate the number of government hospitals required to fill the gap in order to meet WHO standards.

**Tools:**

1. Pandas for data processing

2. Matplotlib and Sea born for visualization