#### **1. Team Members and Role Numbers**

* + **Abdul Moiz (20p-0017)**
  + **Syed Abdul Basit (21p-8044)**

#### **2. Approach and Execution**

* **Approach**:
  + The project aimed to analyze environmental factors (air quality and water pollution) in major cities of Pakistan using data analysis and visualization techniques.
  + Data was cleaned, filtered, and visualized to identify patterns and trends.
* **Execution**:
  + The team used Python with libraries like pandas, matplotlib, and seaborn for data analysis and visualization.
  + Key steps included:
    1. Cleaning column names and data entries.
    2. Filtering the dataset for specific cities of interest.
    3. Creating insightful visualizations to communicate findings.

#### **3. Explanation of the CSV File**

* **Dataset Content**:
  + The dataset includes information on air quality and water pollution in 35 cities of Pakistan, with attributes like region, population, and geographic coordinates.
  + It focuses on two key environmental metrics: **Air Quality** (numerical values representing pollution levels) and **Water Pollution**.
* **Purpose**:
  + To explore the relationship between air and water quality across cities and identify regions requiring intervention.

#### **4. Summary of the Code**

* **Main Logic**:
  + **Data Cleaning**:
    - Removed unnecessary characters and standardized column names.
    - Trimmed and formatted city names for consistency.
  + **Data Filtering**:
    - Focused on specific cities (e.g., Karachi, Lahore, Islamabad) to derive relevant insights.
  + **Visualization**:
    - Bar plots and heatmaps were created to compare air and water quality across cities.
* **Key Components**:
  + pandas for data manipulation.
  + matplotlib and seaborn for creating clear and informative visualizations.

#### **5. Insights and Resources**

* **Sources of Insights**:
  + Trends in air quality and water pollution were derived from city-specific data.
  + Cities with higher population densities showed significant pollution levels.
* **Resources Used**:
  + **Dataset**: Pakistan Water Pollution and Air Quality 2020 (provided as part of the project).
  + **Libraries**: Python-based data analysis and visualization libraries (pandas, matplotlib, seaborn).
  + **Articles**: Research on environmental impacts of urbanization in Pakistan.