Assignment: Data Analysis and Preprocessing with Pandas

Objective:
Analyse and preprocess a dataset using the Pandas library.

Instructions:

1. **Select a Dataset** (CSV, Excel, Kaggle).
2. **Create a Jupyter Notebook** with the following:
    - **Filename**: Name your notebook as `Student_Name_Assignment.ipynb`.
    - **Student Name DataFrame**:
    - `import pandas as pd`
    - `student_name =pd.read_csv(r"C:\Users\Desktop\AirQualityUCI.csv")`
3. **Explore Dataset**:
    - `student_name.head()`
    - `student_name.tail()`
    - `student_name.info()`
    - `student_name.shape`
    - `student_name.describe()`
    - `student_name.isnull().sum()`
    - `student_name.corr()`
    - `student_name.columns`
    - `student_name.dtypes`
    - `student_name['column_name'].value_counts()`
    - `student_name.groupby('column_name').mean()`
    - `student_name.sort_values(by='column_name')`
    - `student_name.dropna()`
    - `student_name.fillna(0)`
    - `student_name['column_name'].unique()`
    - `student_name['column_name'].apply(lambda x: x.upper())`
    - `student_name.sample(5)`
    - `student_name.pivot_table(values='column_name', index='column_name', aggfunc='mean')`
    - `etc.`
4. **Preprocess Data**:
    - Handle missing data: `fillna()` or `dropna()`.
    - Remove duplicates: `drop_duplicates()`.
    - Convert data types: `pd.to_datetime()`.
    - Encode categories: `pd.get_dummies()`.
5. **Save and Submit**:
   Save the notebook as **HTML**: **File > Download as > HTML**.