Q1) Given a dataset (<https://raw.githubusercontent.com/databricks/LearningSparkV2/master/databricks-datasets/learning-spark-v2/iot-devices/iot_devices.json>) containing information about environmental sensors installed across various locations, with attributes including device ID, device name, location details, environmental readings, and timestamp, perform the following tasks using Spark API:

1. Display the count of environmental readings per location sorted in decreasing order.
2. Identify the location with the highest number of environmental readings in each temperature scale category.
3. An environmental researcher is interested in obtaining data from the sensor database regarding locations where the CO2 level exceeds 1400 ppm more than 10 times in a given timestamp range. Provide the list of such locations for further analysis.

Q2) (<https://www.kaggle.com/datasets/cankatsrc/medical-records-dataset>) this dataset contains simulated medical records for a fictional group of patients. The dataset was generated using the Python Faker library to create realistic but fake data.

1. The most common medical conditions among patients admitted to the hospital and their distribution by age group. (10-20, 20-30, and so on)
2. Number of unique types of allergies
3. Number of persons having surname “Davis”
4. The number of Male and Female Patients which are allergic to hospitals