# Data Generation Code Explanation

## Importing Libraries

import pandas as pd  
import numpy as np  
These lines import the necessary libraries for data manipulation and numerical operations. Pandas is used for data manipulation and analysis, while NumPy is used for handling large multi-dimensional arrays and matrices.

## Generating Performance Management Data

timestamp = pd.date\_range(start='2021-01-01', periods=200, freq='H')  
device\_ids = np.random.randint(1, 20, size=200)  
These lines generate a date range for the timestamps and random integers for device IDs. The date range starts from January 1, 2021, creating 200 hourly timestamps. Device IDs between 1 and 19 are randomly assigned to these timestamps.

Various network parameters like RSRP, RSRQ, SINR, user counts, traffic data, and geographical coordinates are generated using numpy's random functions. These parameters are stored in the 'performance\_data' DataFrame along with the timestamps and device IDs.

## Generating Fault Management Data

fault\_types = ['网络中断', '信号质量下降', '硬件故障', '连接丢失']  
Fault types such as network interruption, signal quality degradation, hardware failure, and connection loss are defined. Random fault types are assigned to each timestamp using numpy's random choice function.

Fault recovery times are also randomly generated and combined with the fault types, device IDs, and timestamps into the 'fault\_data' DataFrame.

## Generating Maintenance Data

Random operation times and temperatures are generated for the devices, along with randomly selected device types. This data is stored in the 'maintenance\_data' DataFrame.

## Saving Data

The dataframes 'performance\_data', 'fault\_data', and 'maintenance\_data' are saved to CSV files. The index=False parameter is used to prevent pandas from writing row indices into the CSV files.