Data Analysis of The tuda/ubicompzis dataset (v. 2020-04-08)

Link to dataset : http://crawdad.org/tuda/ubicompzis/20200408/
Link to code on github : https://github.com/JARACH-209/DataAnalytics.git

Data Background and Structure

The collected context information is used to evaluate the performance of five zero-interaction pairing or authentication schemes.
The data is present in the form of 3 Trace sets : Raw, Processed and Result set.
Raw data is in the form of compressed json files which is almost 1 TB in size.
The dataset is split into data that is based on the recorded audio, and those based on sensor readings.
These trace sets are from 3 different scenarios which are :

1. Office (Static)
2. Car (Moving)
3. Office (Static and Moving)

The data contains information of 24 sensors and each sensor records the data of audio restricted access, Wi-Fi and Bluetooth Low Energy (BLE) beacons, barometric pressure, humidity, luminosity, temperature, accelerometer, gyroscope, and magnetometer. The Sensor data is present in the form of two intervals 10seconds and 30 seconds for all sensor pairs along with the Summary, Summary of nights, Summary of Weekdays and Summary of Weekends.

DataAttributes :

On exploration of Office experiment -> sensors results -> wfi -> sensor1 and summary data

1. Summary.json data for each sensor in 10 seconds interval contains:

   Indexes :
   a. Feature
   b. Sensor
   c. time_interval
   d. Value
      i. Sensor Result values have further attributes which store the readings :
         1. Euclidean
         2. Jaccard
         3. Mean_exp
         4. Mean_hamming
         5. sum_squared_ranks
   e. co_located
   f. non_colocated
      And sensor pair data for each sensor i from 24-i upto sensor24

   Columns :
   Meta Data and Results

2. Sensor.json data :
    a. Each Sensor.json file contains metadata of
        i. 'created_on', 'generator_script', 'generator_version', 'parameters', 'processing_end', 'processing_start', 'source_files' and 'system',

Data Explored So Far !