

1 Introduction

A typical 4G network consists of several nodes connected to each other, where each node serves users in the surrounding area. While users are accessing network services, their mobile phones record Key Performance Indicators (KPI's) which can help network operators in assessing their quality of service.

There are many KPI's which evaluate different aspects of the network. For example, **RSRP** (Reference Signals Received Power) is a KPI which measures the network coverage in the user's location. **Traffic Volume** is another KPI which measures how much data has been consumed by the user.

2 Problem Statement

For this challenge, you will be provided with **two crowdsourced datasets**: RSRP and Traffic Volume. Each dataset has the corresponding KPI measurements collected from mobile phones of different users over a week. It also includes the user's location, operator, phone model and other information. The detailed description of each field can be found in **DataDescription.xlsx**.

Please note that the data is in its raw format, so you need to filter the it from invalid values before working with it.

3 Task

3.1 Visualizing Data

Create a dashboard to evaluate the 4G network performance of different operators using the provided datasets.

The dashboard **MUST** include these 4 charts:

1. A time-lapse of a density map showing how users move during the hours of the day. You should see the density on the map changing according to people's activities every hour. Use the locations in **RSRP** dataset.
2. A heat map of user locations for a selected operator colored by **RSRP** value with a dropdown menu to choose the operator.
3. A bubble map showing the areas with high downlink traffic for each operator. Each bubble represents an area on the map and the area covered by this bubble can be controlled with a slider. In this chart, the bubble size depends on the sum of the **Traffic Volume** of all users in this area and it's colored per operator. **Hint**: consider using **hex bins** for this task
4. A Bar chart of **RSRP** per device type per operator with an option to choose the aggregation method of RSRP (avg, min, max and 90th Percentile). Add a slider to control the minimum number of users of each device.

Now that you have made a great-looking dashboard, What can you see from these dashboards? Present your findings, Highlight observations and key insights that you think the reader should take away from the dashboard.

Examples of these findings can be:

- Areas having good or poor coverage compared to competitors
- Certain mobile phones that suffer from poor performance
- Traffic hotspots for operators

You will find many interesting things in the data. **The goal** is to show your understanding of the data and communicate your message in a clear way.

3.2 Applying ML !

Time to do some machine learning magic!

Imagine you are running a conversation with one of the operators whose data is in the dataset provided. The discussion is referring to any polygon of your choice that's **at least** 10km x 10km size and **has enough data samples**. Can you help the operator to answer the questions below?

1. Assuming that the coverage next week will improve in that polygon compared to competition (i.e., RSRP will get better than other operators). What would be the impact on downlink and uplink traffic volumes?
2. Samsung devices are the main handsets in our network. Can you predict the traffic volume growth, uplink and downlink, over time for these devices and compare it to the competitors?