Date: 8th March,2025

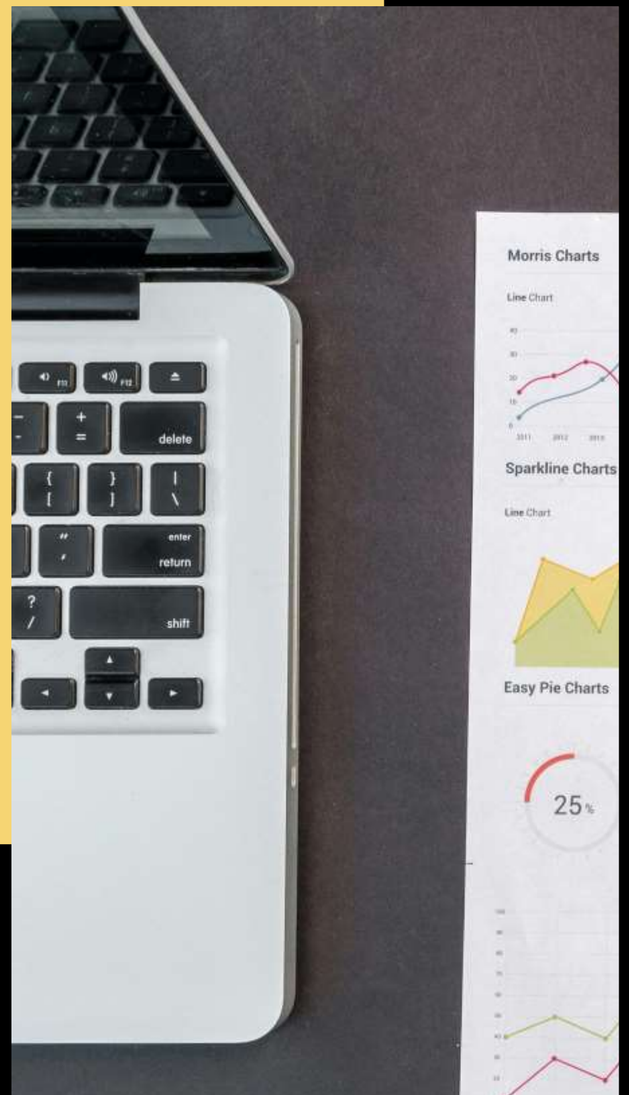
Visualizing cellular networks for transit infrastructure

NEW CANVAS

Presented by: Mohammed Syed

AGENDA

- Challenge Statement
- Objectives
- Data Preprocessing
- Exploratory Data Analysis
- Coverage Visualization
- Score Calculation
- Repeater Placement Algorithm
- Comparative Analysis



CHALLENGE STATEMENT

Challenge name:
Bus Division Cellular Coverage Study

Organization Name: GL Communications, Inc.

Challenge Partner: Dr. Kam Yee

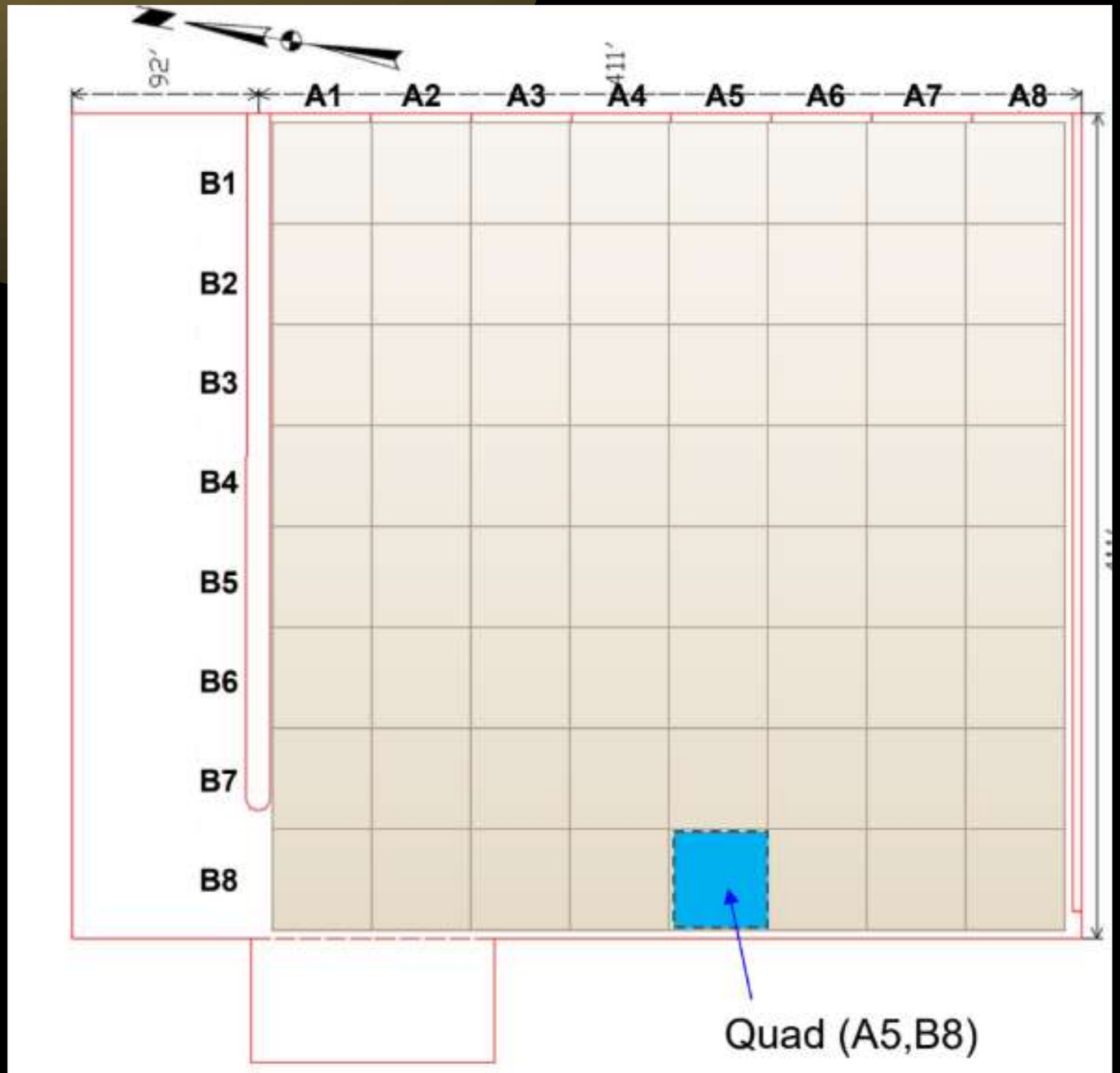
Challenge Summary

A certain transit agency is interested in evaluating and comparing robustness and reliability of cellular communication link from three different cellular service providers in one of its indoor bus maintenance facilities.

GL Communications collected data that measures the performance for coverage and data throughput. The objective is to compare cellular carriers' performance, verify adequate cellular data coverage, and identify any major coverage gaps.

CHALLENGE OBJECTIVES

Objective 1	Compare cellular carrier's performance
Objective 2	Verify cellular data coverage and identify any major coverage gaps
Objective 3	Provide recommendation on where to add cellular repeaters within the facility to close major coverage gaps



UNDERSTANDING THE DATA

The facility is subdivided into a grid of 8x8 Quads

8 Aisles across (A1-A8)

8 Blocks per aisle (B1-B8)

So, there are (8*8) 64 Quads in the facility

Dataset

Three datasets are provided, each containing seven columns:

- Aisle Number : Values ranging from A1 to A8
- Block Number: Values ranging from B1 to B8
- Timestamp – Date and time of test.
- RSSI, Received Signal Strength Indicator [dBm]
- TCP Upload Speed [Mbps]
- TCP Upload Quality of Service (QoS) (%)
- TCP Round Trip Time (RTT) [ms]

Key Performance Indicators (KPIs):

We have four KPIs available:

- RSSI (Received Signal Strength Indicator) [dBm]
- TCP Upload Speed (Data Transmission Rate)[Mbps]
- TCP Upload QoS (Quality of Service) [%]
- TCP RTT (Round Trip Time) [ms]

Data Preprocessing:

Data preprocessing is the process of cleaning, transforming, and organizing raw data to make it suitable for analysis or modeling.

Challenges:

- Null Values
- Duplicate Values
  - Average Approach
    - Log Scale
  - Maximum Approach
    - Normalization + Weights
    - Inverse Transformation
- Check for missing quads (<64)

Exploratory Data Analysis (EDA):

it is the process of analyzing and visualizing data to uncover patterns, detect anomalies, test hypotheses, and summarize key insights.

Challenges:

- Histogram for all KPIs using both approaches
- Scatterplot
- Statistical Analysis
  - All in one type of plot
    - Box Plot
    - Violin Plot
  - Comparision using violin plot

All of the plots are interactive

Histogram bins

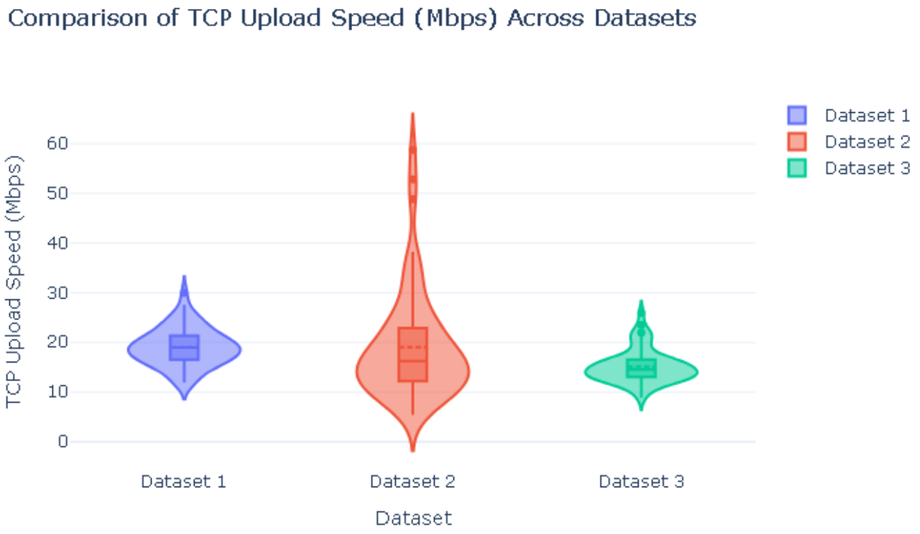
RSSI	
< -105 dBm	Poor
-105 to -96 dBm	Fair
-96 to -86 dBm	Fair
>= -86 dBm	Good

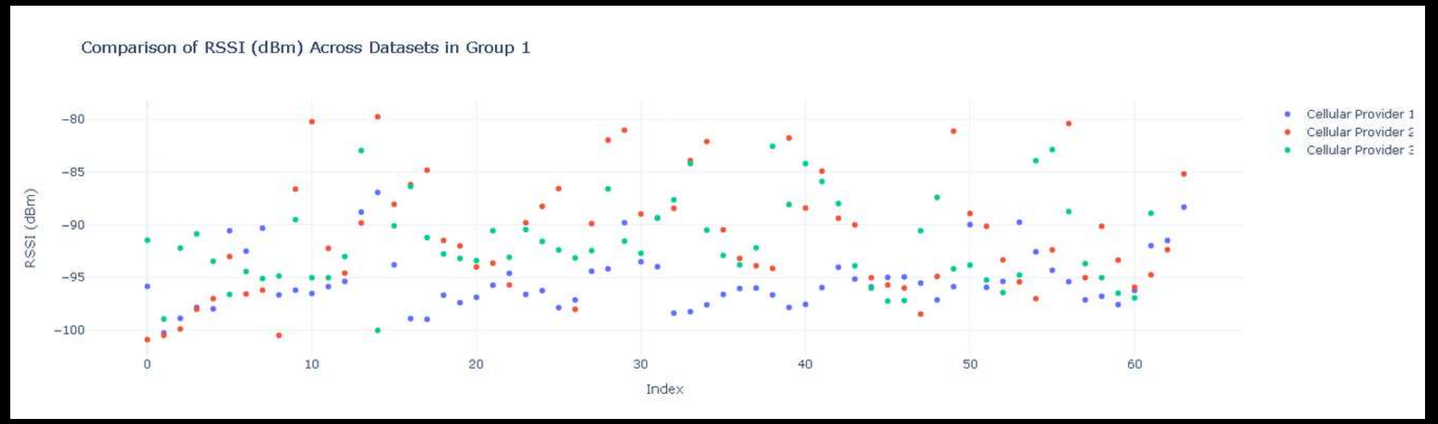
TCP Upload Speed	
< 2 Mbps	Poor
2 to 7 Mbps	Fair
7 to 15 Mbps	Fair
15 to 30 Mbps	Fair
>= 30 Mbps	Good

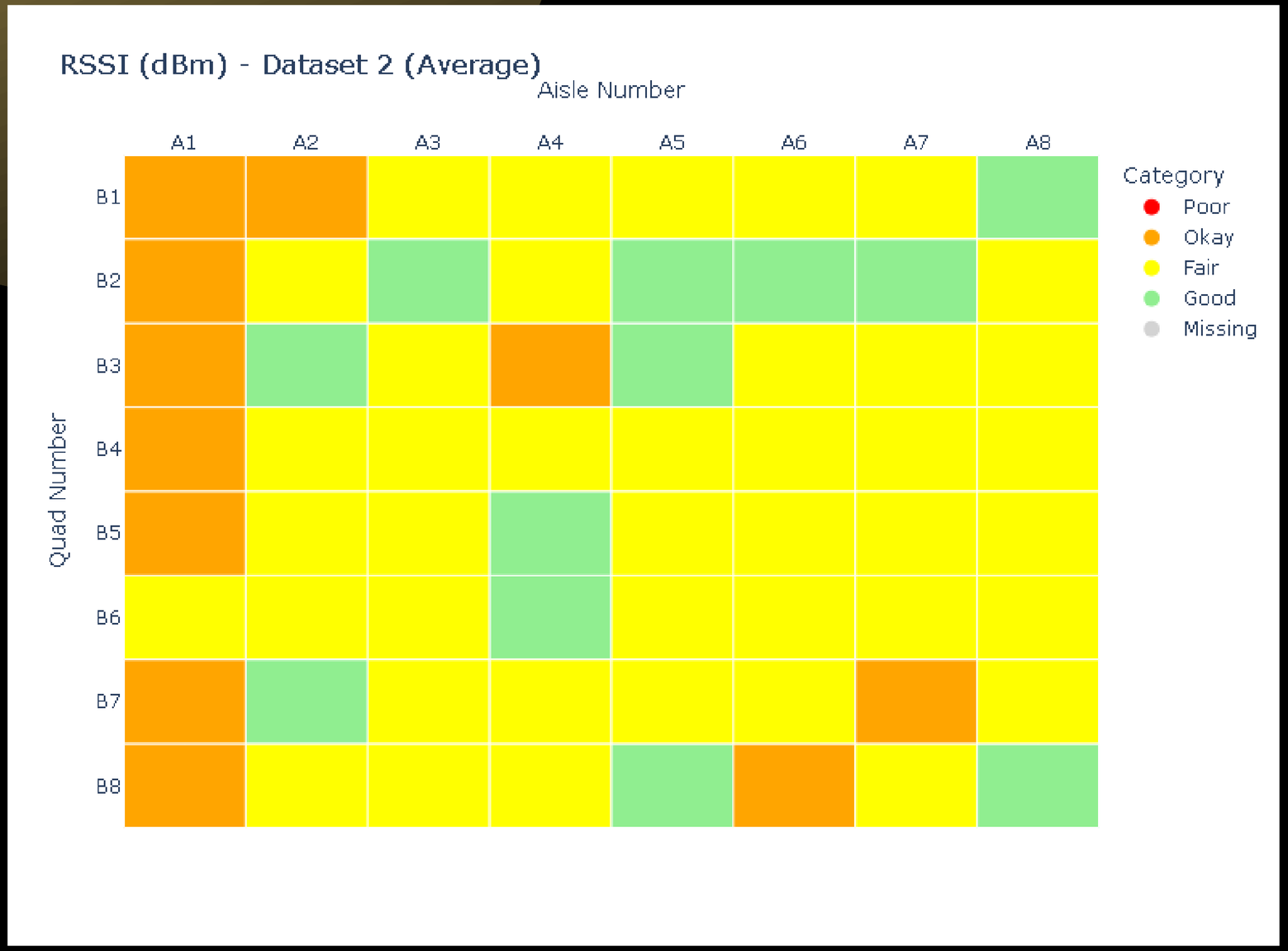
TCP Upload QoS	
< 30%	Poor
30 to 50%	Fair
50 to 70%	Fair
>= 70%	Good

TCP RTT	
< 50 ms	Good
50 to 100 ms	Fair
100 to 200 ms	Fair
>= 200 ms	Poor

EDA Insights







Coverage Visualization

Used an interactive heatmap to show the intensity of each KPI in both approaches across all three datasets.

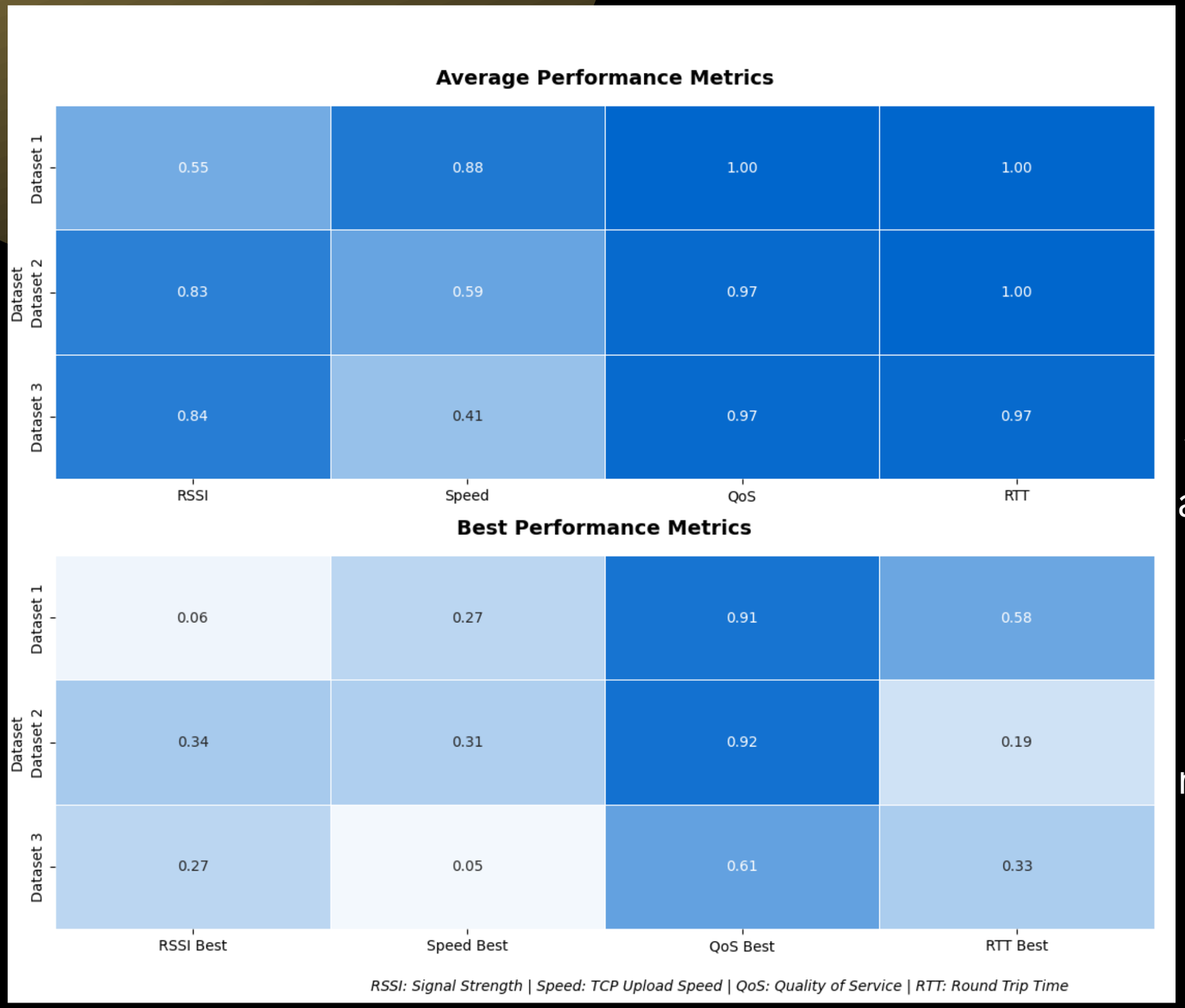
Score Calculation:

Score = (RSSI Fair+Good %) * 40% + (TCP Upload Speed Fair+Good %) * 40% + (TCP Upload QoS Fair+Good %) * 10% + (TCP RTT Fair+Good %) * 10%

Adjusted Score Calculation

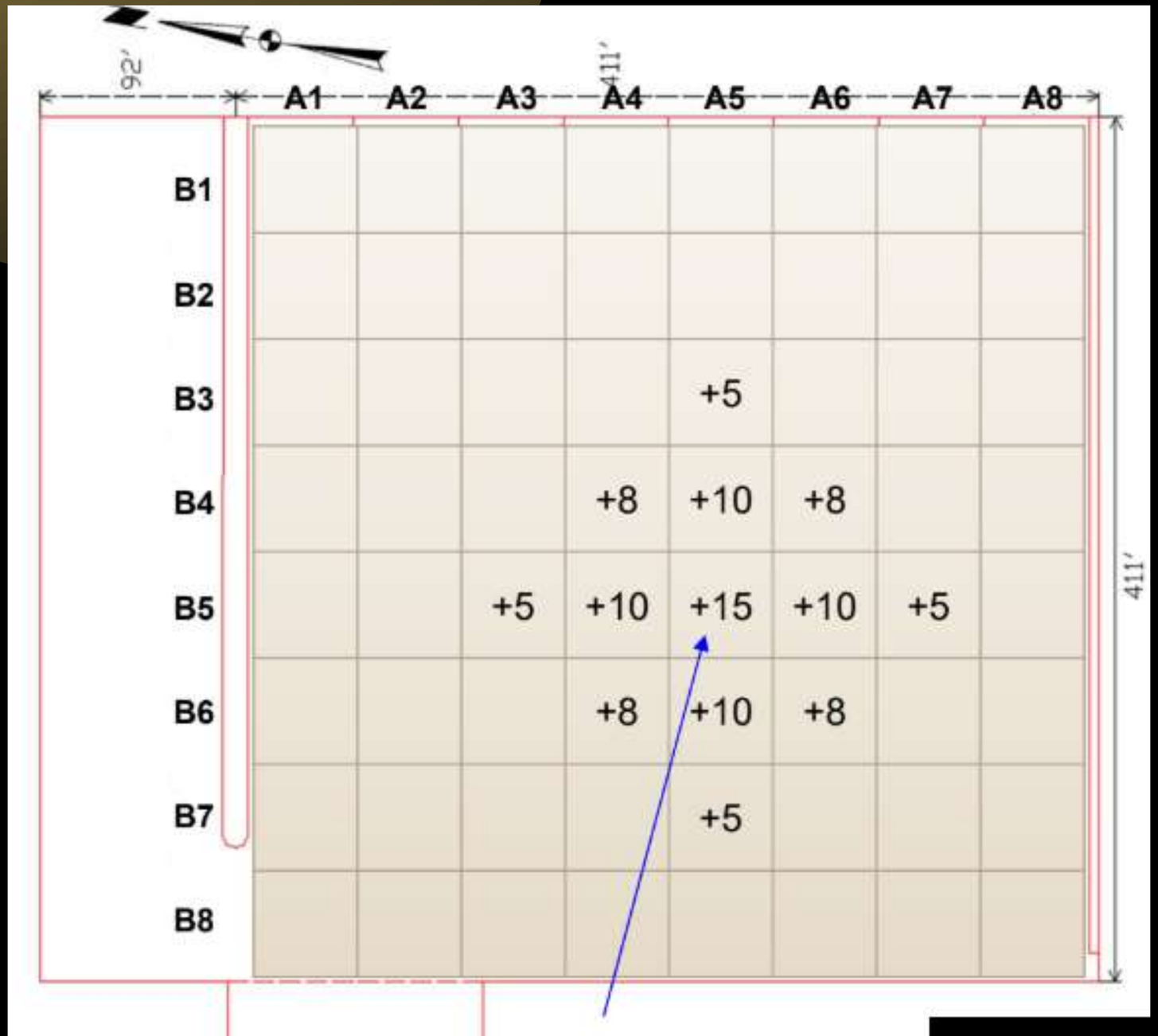
- Score = (RSSI Fair+Good %) * 40% + (TCP Upload Speed Fair+Good %) * 40% + (TCP Upload QoS Fair+Good %) * 10% + (TCP RTT Fair+Good %) * 10%
- Best Score = (RSSI Good %) * 40% + (TCP Upload Speed Good %) * 40% + (TCP Upload QoS Good %) * 10% + (TCP RTT Good %) * 10%

- *The metrics were revised as the dataset lacked "Not Poor" values and had multiple "Fair" categories. "Fair" now represents the second-best rating, and a "Best Score" metric was added to compare performance.*



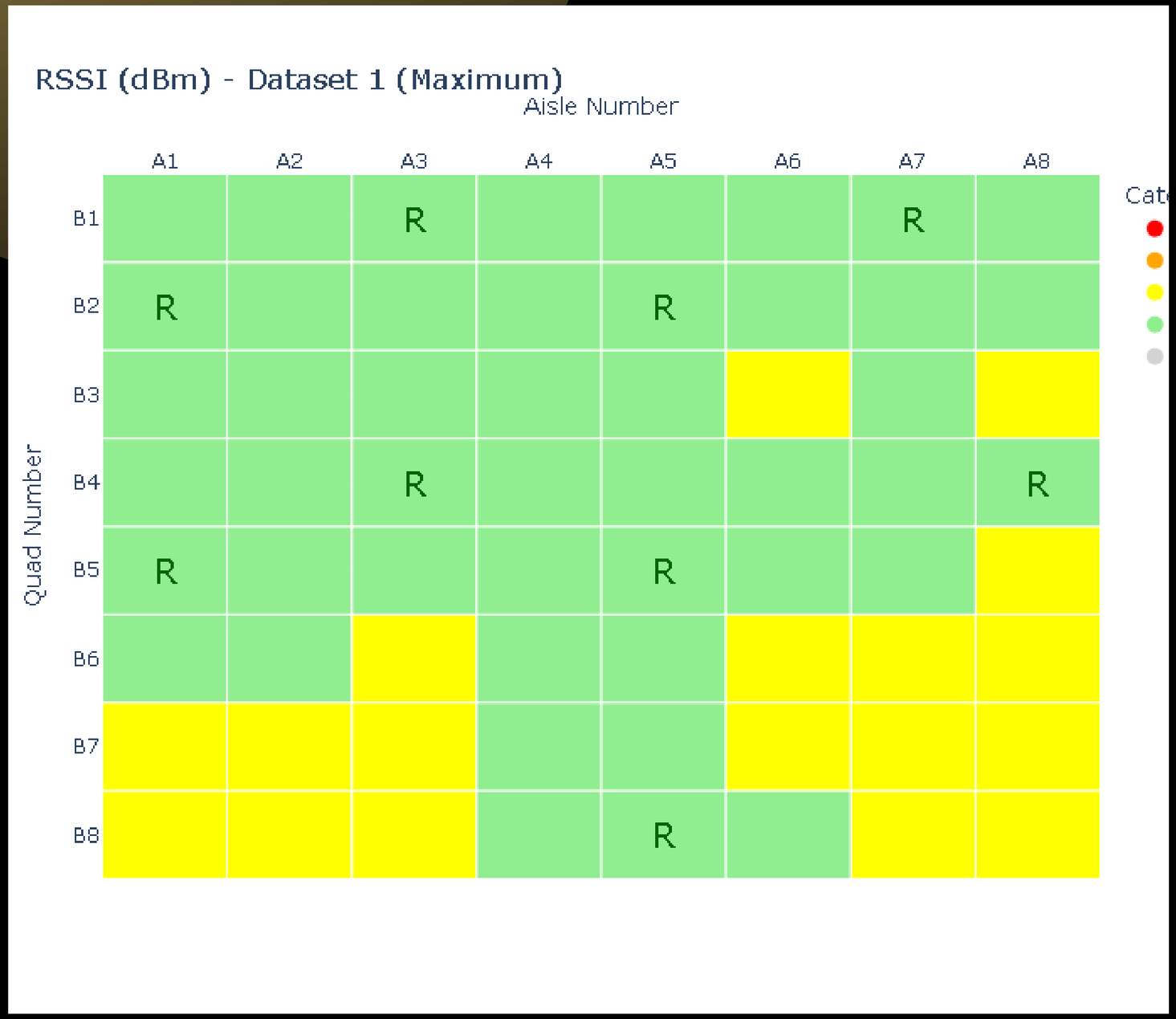
Comparative Analysis

Evaluated carriers using a dual-scoring system across signal strength, speed, QoS, and RTT metrics. The heatmap visualizations show Dataset 1 excels in QoS and RTT reliability, while Dataset 2 provides superior signal strength. Using normalized scores for both average and peak performance scenarios, I calculated weighted metrics that prioritize factors most critical for transit operations, creating a quantitative framework for carrier recommendation.



Cell Signal Booster

- Small antennas boost andredistribute the signal, as much as 32x stronger*
- Rough estimate of increase in dBm shown
- 32x = 15 dBm increase
- 10x = 10 dBm increase
- 6x = 8 dBm increase
- 3x = 5 dBm increase



Repeater Placement Algorithm

The *min_repeaters* function determines the minimum number of repeaters needed to improve weak RSSI signals in a grid. It identifies poor signal locations, places repeaters at the worst spots, and boosts nearby signals based on a predefined map. The function returns the number of repeaters used, their locations, and the updated RSSI values.

Date: 8th March, 2025

THANK YOU FOR YOUR ATTENTION

Presented by: Mohammed Syed