

11 SUSTAINABLE CITIES AND COMMUNITIES



Team KCR

Sai Gokul **K**anakala
Rohit **C**handramohan
Rahul Gadi

The Big Picture



[illegible]

SDG - How are we addressing?



Data description

- Uber Movement [1]- 600MB
 - Mean travel times
 - Hourly aggregated dataRows x Features - 1 million x 7
- MBTA Subway data [2] - 4GB
 - Subway routes and stop location, including real time status updates.
 - Details of entries and exits at each stop.Rows x Features - 1.5 million x 50



Data description

- Boston Neighborhood Data [3] - 250MB
 - Building location and other details
 - Rows x Features - 400,000 x 70
 - Areas of Interest including hospitals, schools, etc.
- Fire Incidents Reporting [3]- 80MB
 - Rows x Features - 250,000 x 22
- Crime Reporting [3] - 150MB
 - Rows x Features - 500,000 x 19
- Waste Collection Reports [3] - 10MB
 - Rows x Features - 50,000 x 6
- Water and Energy Consumption [3] - 50MB
 - Rows x Features - 50,000 x 35

Proposed Methodology

Data Processing

- Cleaning
- Sorting
- Aggregation

Transformation

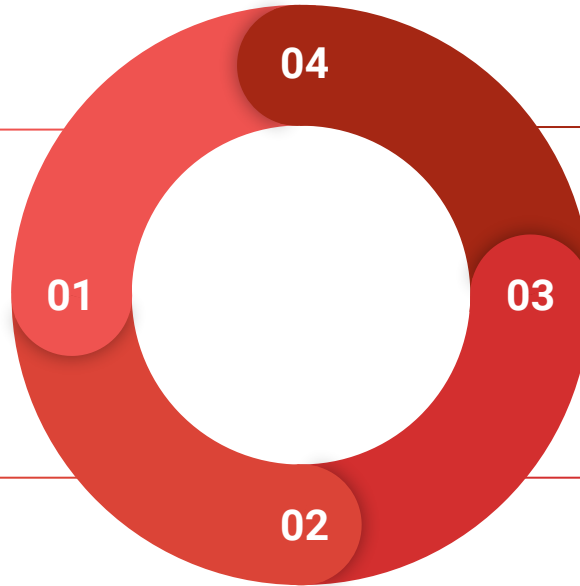
- Data discovery
- Data mapping
- Joins

Visualization

- Heat Maps
- Graphical plots

Analyses

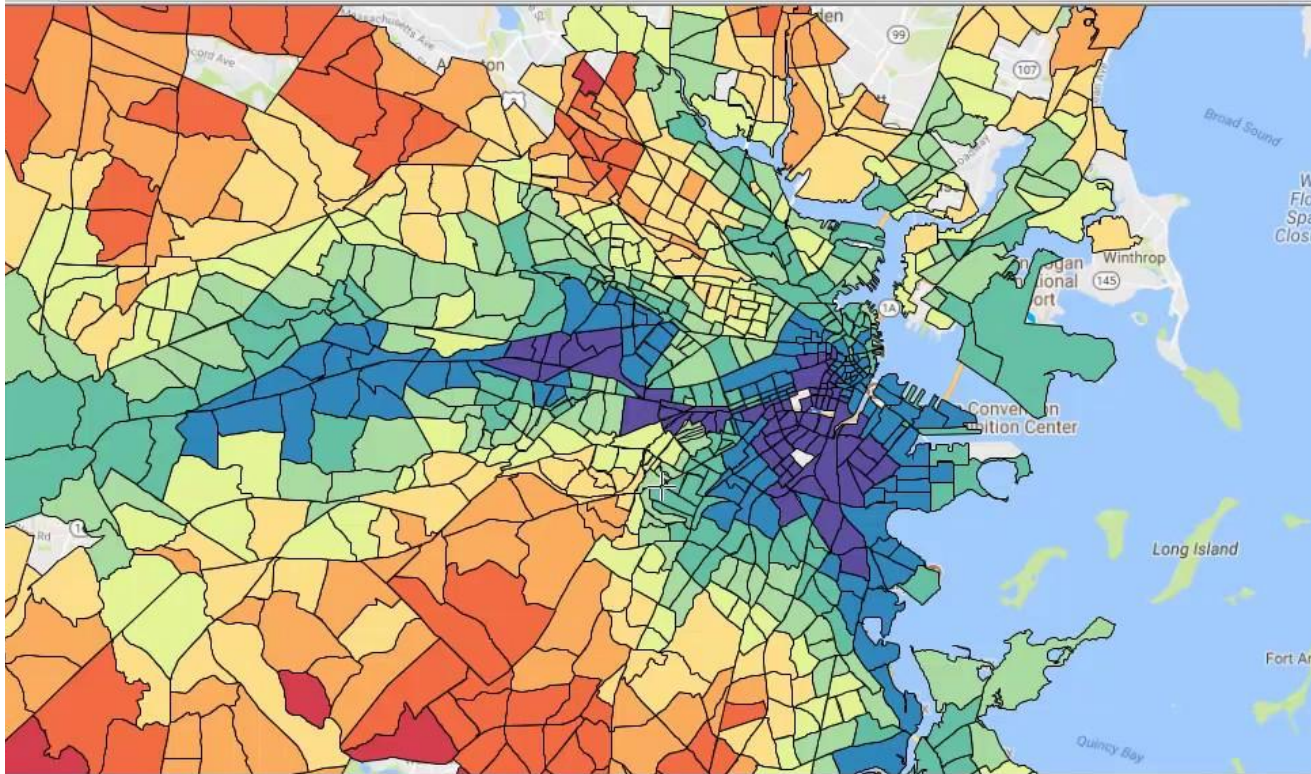
- Graphical & Geographical
- Route optimization
- Similarity scores
- Clustering
- Time series



Mock Output

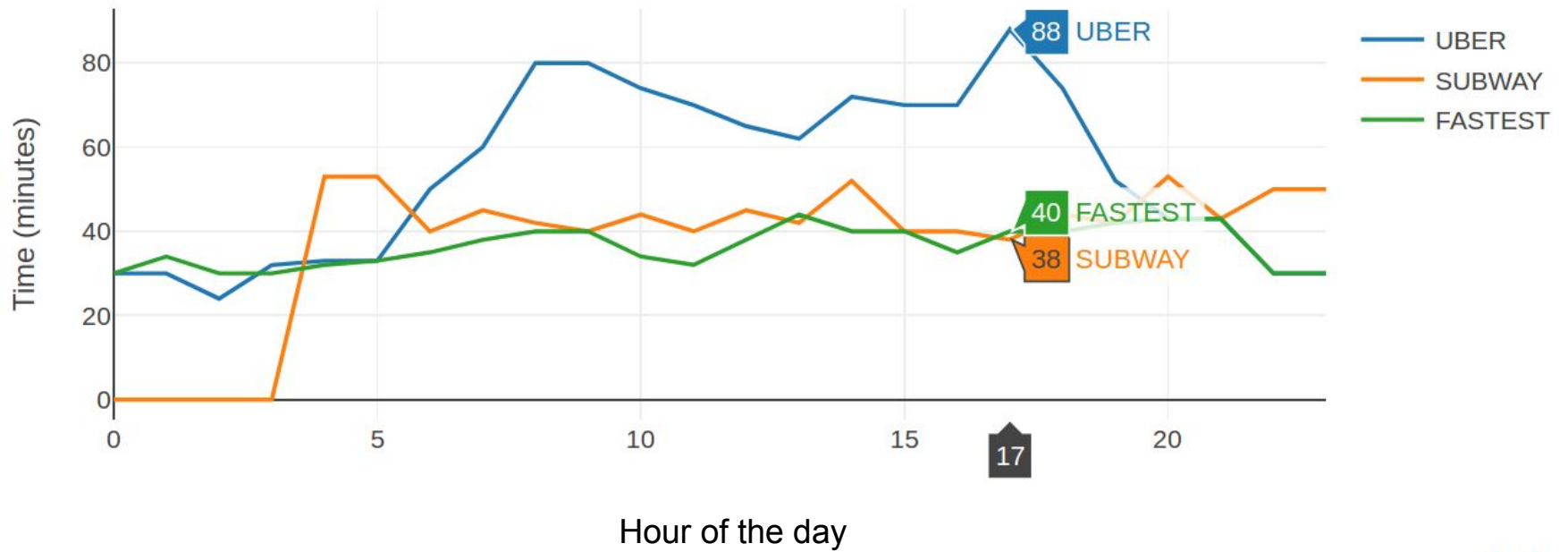
Day	Sunday - Saturday		
Type	Crime, Fire, Flood, None		
Target	Hospitals	Police Departments	Fire Stations
	Schools	Library	Gardens

Mock Outputs

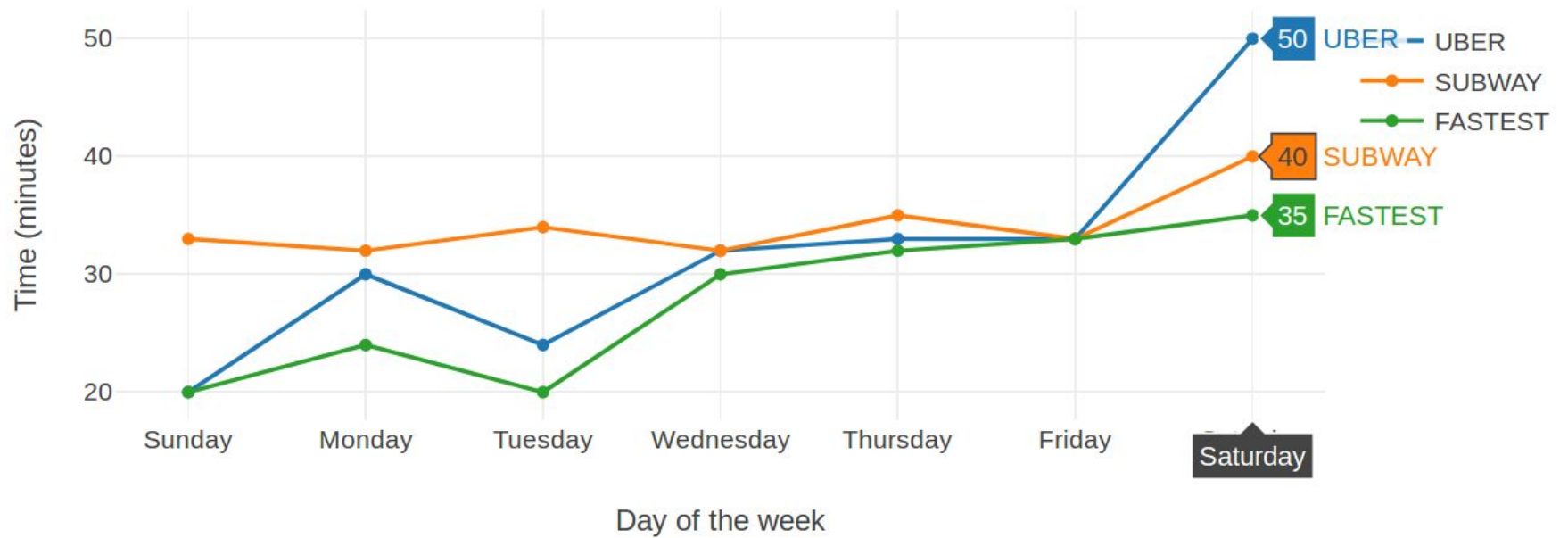


Boston Map Identifying Vulnerable Areas

Time taken to nearest Target Location



Time taken to nearest Target Location



Team Work

Data Processing

Gokul and Rohit

- Cleaning
- Sorting
- Aggregation

Transformation

Rahul and Rohit

- Data discovery
- Data mapping
- Joins

Visualization

Rohit and Gokul

- Heat Maps
- Graphical plots

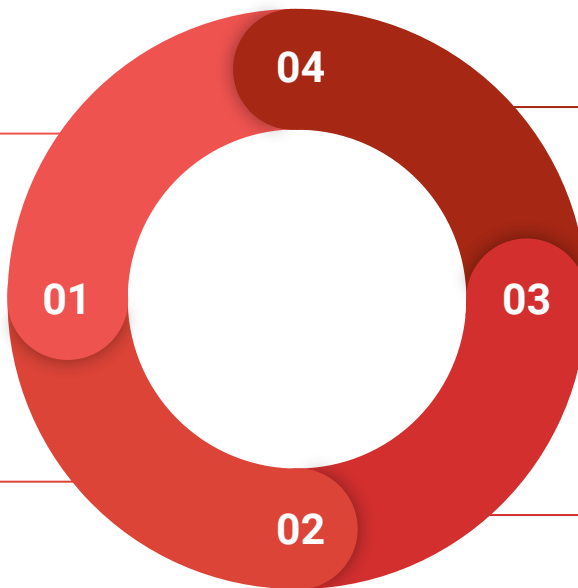
Analyses

Rahul and Rohit

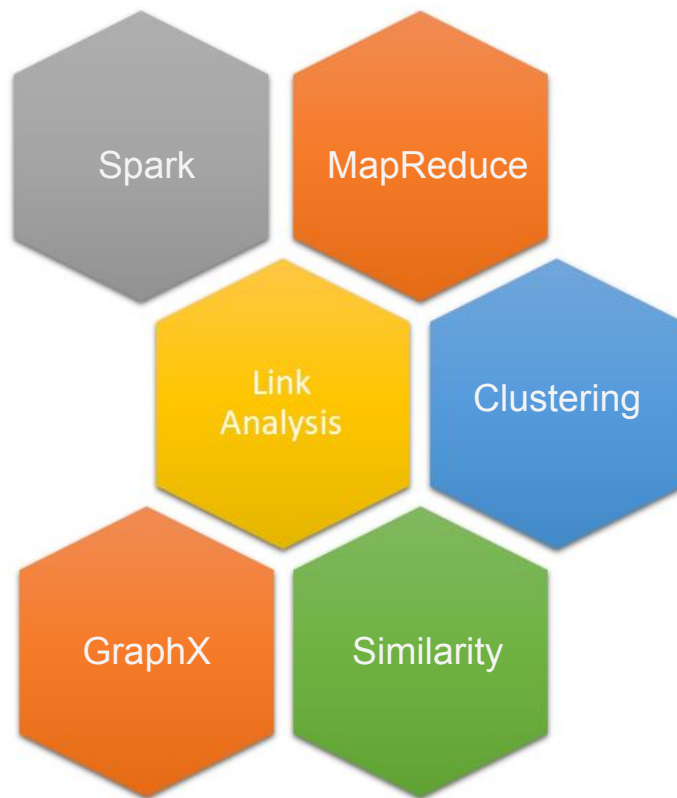
- Graphical & Geographical
- Route optimization
- Similarity scores

Gokul

- Clustering
- Time series



Course concepts



Conclusion

- Address the various issues in Boston and visualize the problems.
- Join multiple datasets to create a holistic view of the city.
- Identify vulnerable locations.
- Rank locations depending on access to public facilities
- Can be extended to other cities.
- Learn and implement concepts of big data analytics and work with large scale datasets.

References

1. <https://movement.uber.com/>
2. <http://realtime.mbtta.com/>
3. <https://data.boston.gov>