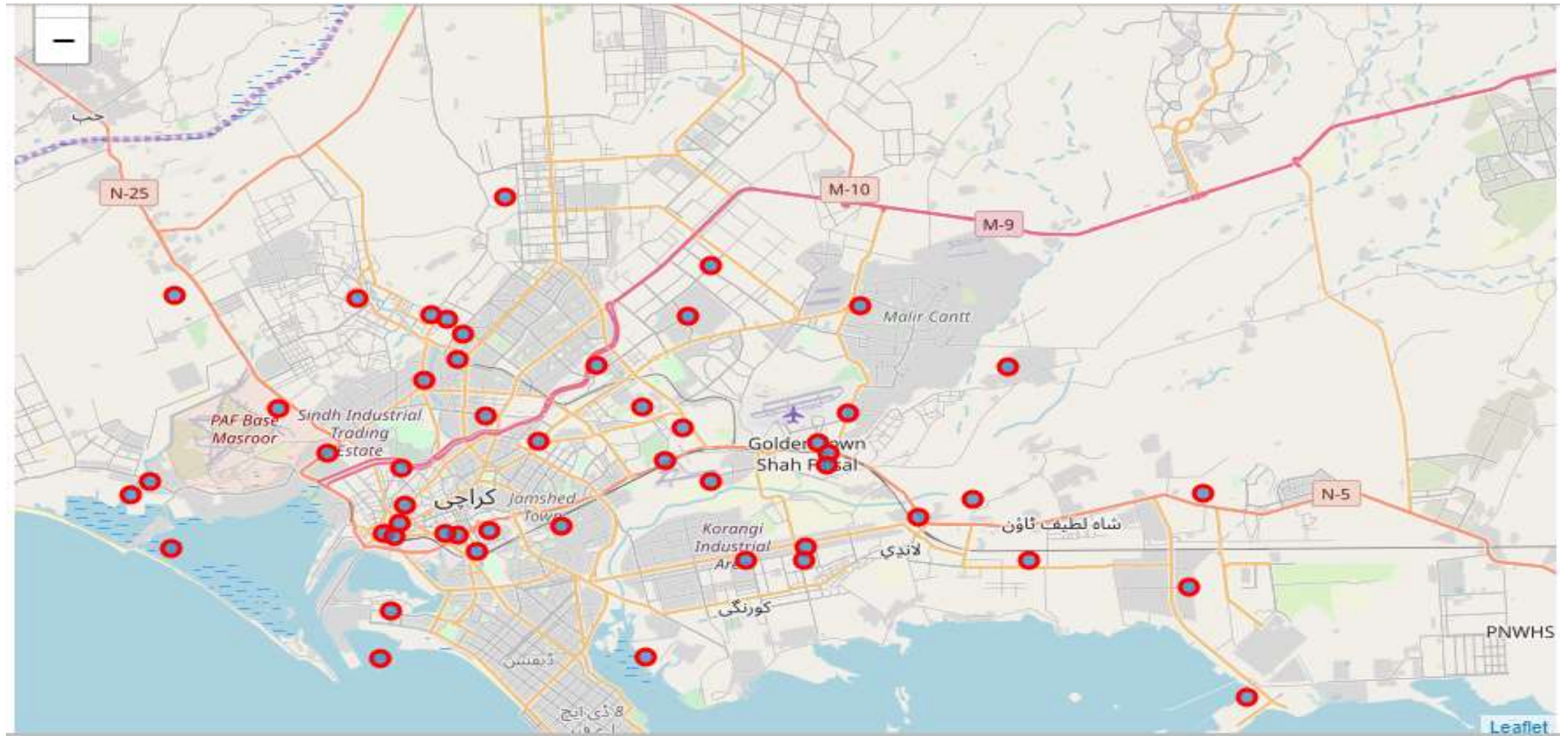


IBM Applied Data Science Capstone

ANALYZING TOWNS IN KARACHI AND IDENTIFYING POTENTIAL
BUSINESS OPPORTUNITIES



Business Problem

- Karachi, despite being the biggest city and commercial hub of the country, very minimal amount of data available for it and what available is neither compact nor decision ready
- Existing and new both businesses of any type are having problem in finding a place which is both economically viable and suit their business model
- Several business and startups in the city have to research from scratch to find optimum locations or neighborhood for their product/services success with reasonable property valuation

Target Audience

- This project can be utilized by several groups, mainly business owners and new startups also tourist and citizens can use this project to find places of any category near where they are
- Government can also utilize this project to track a certain town growth and what measure the decision makers can take to further improve particular towns conditions and attractions

Data

Following data has been used to work on this plan

- Postal Codes of all areas in Karachi
- Population Statistics of each area
- Property estimated Value of each area

Methodology

Includes:

- Web scraping for relevant data
- Data wrangling and cleaning (Tidy data format)
- Use of Foursquare API
- Data visualization using Folium, Waffle and WordCloud.
- Applying unsupervised machine learning algorithm (DBSCAN)

Results

- District Central is dominated by Fast food and Bakery, having the third highest population and No. of houses.
- District Korangi is favored by Fast Food and Burger Joints with second highest domination in both population and No. of houses.
- District West has most beaches as venues and is the most populated district both in term of population and houses
- District South has most market and Hotels and has opportunity for future developments for both houses and population.
- District East also dominated by Fast Food and Pizza and is moderately populated and has opportunity for future growth
- District Malir has most Fast Food and Farm and has opportunity for future developments for both houses and population.

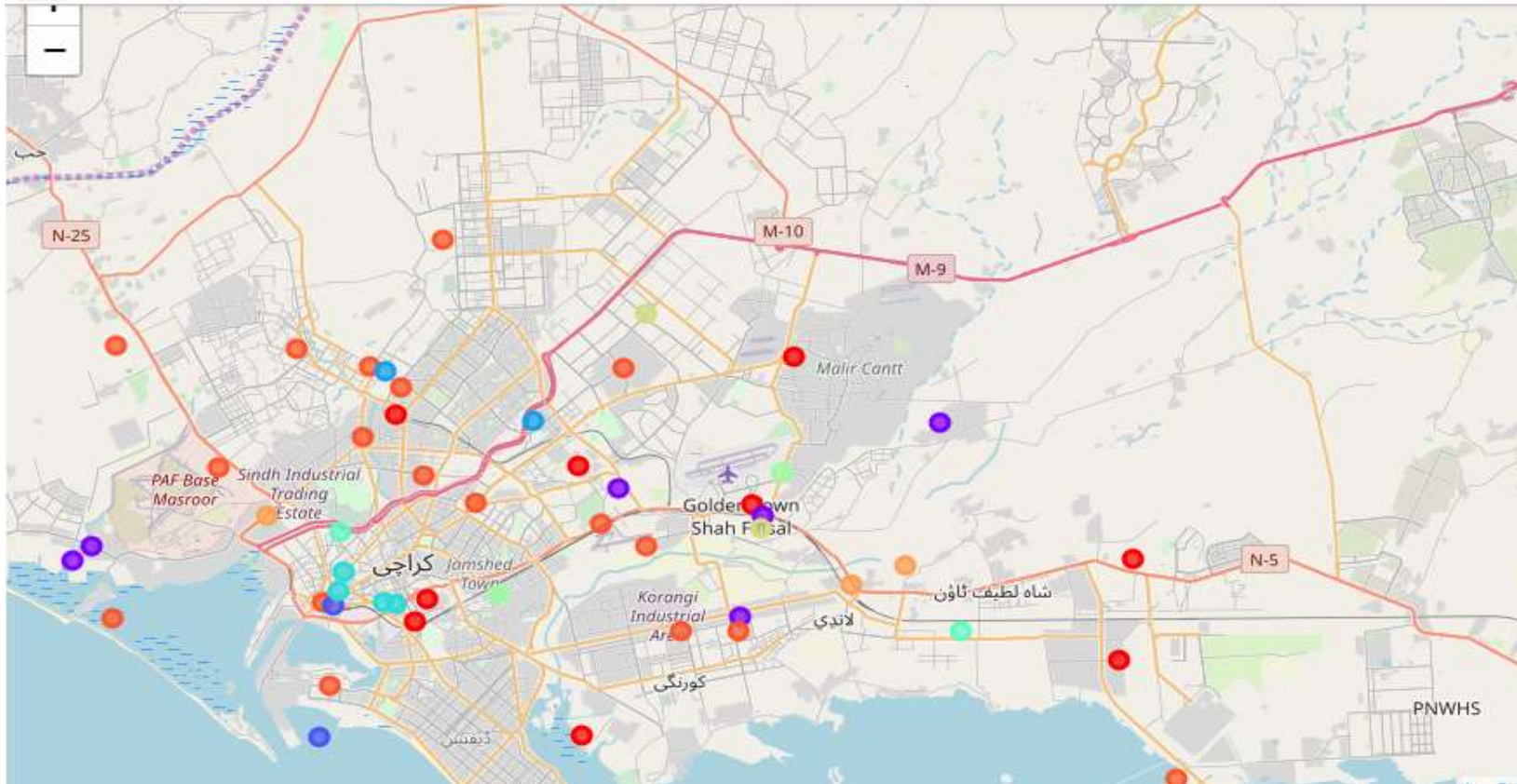




Discussion

- It is been evident that restaurant is the most common venue of the city.
- The result for each district and cluster wise provide insights that each district have different venue proportions.
- The limitations in this report is that forsquare has only provided data points of major chains and popular venues and real number of venues might be much greater, for further research google api can be used for much broader analysis
- Moreover, the unavailability of data for the city is biggest limitation as most of the time consumed is to find and put together the data

DBSCAN OUTPUT



Recommendations and Future Prospects

- This Project can be further enhanced by adding real time data collaborating several companies
- Also, government stake holders to check area wise real time monitoring of circulation of wealth, economics and potential business opportunity for any kind of business.

Conclusion

To conclude, Karachi's each district has its own essence in terms of venues and by combining their analysis with DBSCAN it is indicated that what will be the feasible opportunity based on current venues in a cluster

Thankyou

