

ORIGIN - DESTINATION ANALYSIS OF TAXI TRIPS IN CHICAGO

CLUSTERING WITH TIME-DEPENDENT TRIP STATISTICS



Nandhitha Ravindran, Pavithra CS, Yao Sutong
Mentor: Prof. Kam Tin Seong

School of
Information Systems

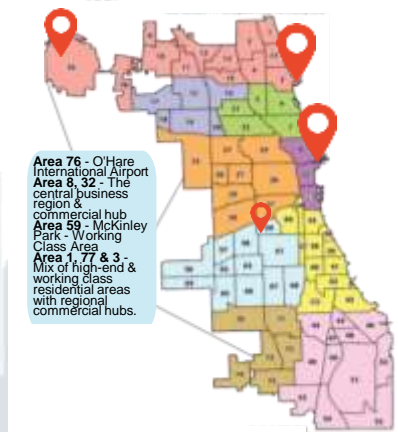
INTRODUCTION

ABSTRACT

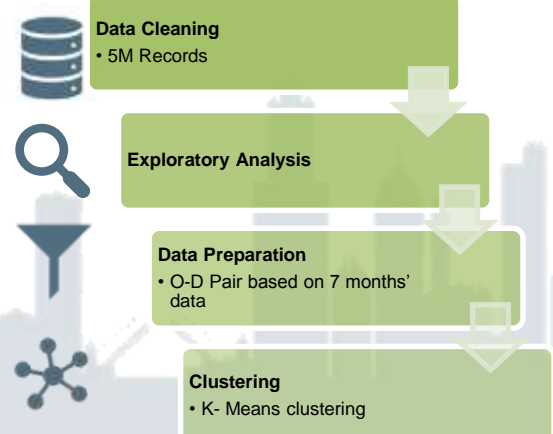
With growing population and rapidly evolving commuter patterns in urban areas, there is an increasing demand for transportation resources. Analyzing urban mobility behaviors can give us insight into factors that influence supply & demand. An integral part of today's urban commute are taxicabs and with the affluence of GPS technology, taxicab trips have become an extremely rich and informative source of data.

OBJECTIVE

To analyze the taxicabs data of Chicago and derive relationship statistics for each origin-destination pairs, which would give us insight into variations in demand of taxis, over different times periods of the day as well as within & across different regions in Chicago

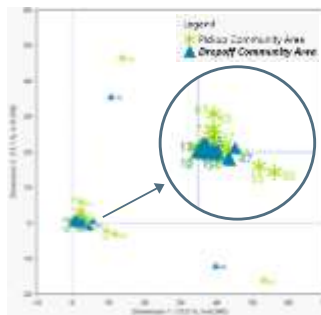
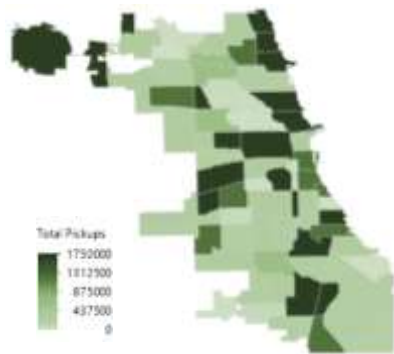


APPROACH



EXPLORATORY ANALYSIS

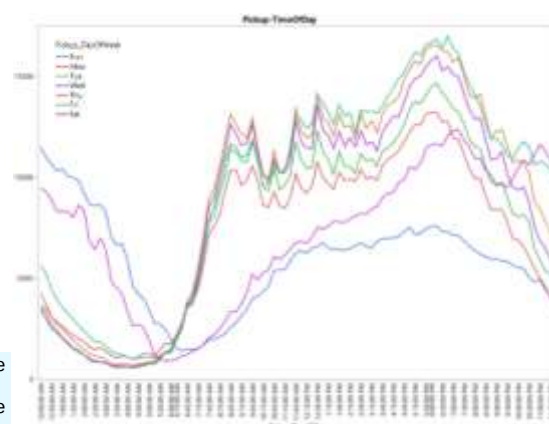
Choropleth Map of Chicago City



Correspondence Analysis

- To find the correlation between the pickup and drop-offs of the taxi trips
- Community Area 49 and 70 are distinct and very less correlated with the other areas
- Most drop-offs within the same region

Time Series Analysis Graph



The 24 hours in a day is divided into 5 buckets based on our exploratory analysis

- Peak Time – Evening
- Day-Wise Grouping to determine the pattern over the 7-day period
- 4 distinct Demand Periods:
 - Morning
 - 2 during Noon (Lunch Time)
 - Evening

Bin Name	Morning	Noon	Evening	Night	Late Night
Bin Division	05:00 AM - 10: 59 AM	11:00 AM - 03:59 PM	04:00 PM - 07:59 PM	08:00 PM - 10:59 PM	11:00 PM - 04:59 AM



77 COMMUNITY AREAS



18 HIGH DEMAND LOCATIONS

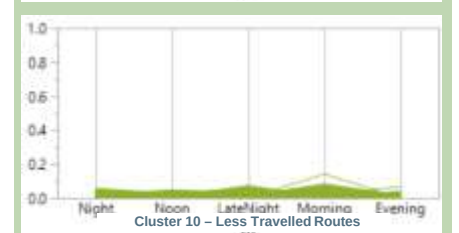
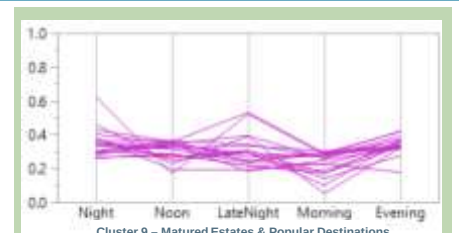
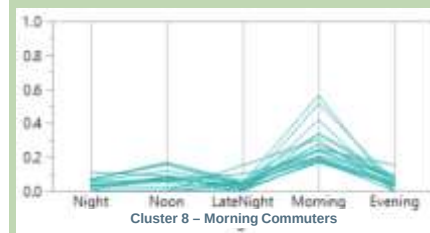
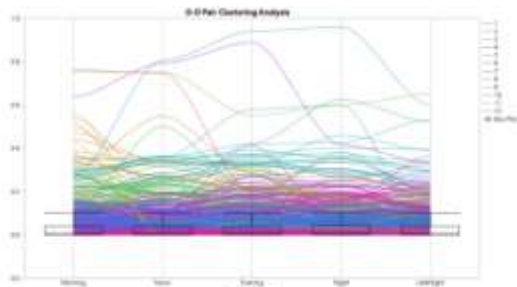


AIRPORT, BIZ DISTRICT & SCHOOLS

CLUSTERING ANALYSIS

Data Transformation before Clustering:

- Derived percentage of trips completed in each community area with a fixed pick-up area and period of the day.
- O-D pairs with less than 30 trips generated in the 7-month period were excluded from the analysis.
- Clustering 1,056 O-D pairs by k-mean method. 12 clusters were identified with distinct features and the most optimal CCC.



CONCLUSION

Origin - Destination interactions statistics derived from taxi trips details were used to conduct a clustering analysis of the travel pattern and needs of the residents of Chicago. Below are insights based on the analysis:

- Similar variation of daily volume is observed for taxi trips commenced through the week.
- A noticeable number of intra-region travels were observed for a variety of community areas, and with distinct daily travel patterns.

FUTURE SCOPE

- Incorporate more factors, e.g. day of the week, for derivation of O-D pair statistics
- Integrate demographic data into the O-D pair analysis to explore interesting correlations
- Conduct two-way analysis of the O-D pairs to further understand the taxi demand across various community areas