LEVERAGING SOCIAL MEDIA DATA TO IDENTIFY LATENT INDICATORS OF DIVERSITY, EQUITY, AND INCLUSION (DEI) OF TRANSPORTATION SYSTEM

Fariha Nazneen Rista, Md Rakibul Alam, Khondhaker Al Momin and Arif Mohaimin Sadri

School of Civil Engineering and Environmental Science, University of Oklahoma

Abstract

- The study aims to contribute to design and operation of an equitable transportation system utilizing the twitter data and census database.
- Analyzing data from twitter conversation of residents, the latent indicators of transportation inequity in the study area can be identified.
- Twitter sources can be correlated with demographic factors utilizing the geospatial information of the tweets.
- We plan to develop an agent-based modeling framework to assess system-wide performance and impacts due to unfair distributions of such indicators in the system in future.

Background

- Adoption of transportation policies that prioritizes highway expansion over public transportation has disproportionately impacted minorities and low-income people by restricting their access to social and economic opportunities and thus resulting in residential segregation.
- A diverse, equitable and inclusive (DEI) transportation planning has the potential to ensure accessibility for all classes of people and change the way residents experience urban space.
- U.S. government has enacted policies that require transportation agencies to implement DEI in all aspects.
- No comprehensive method has been found in the literature to assess system equity and to identify the optimum distribution of DEI indicators.
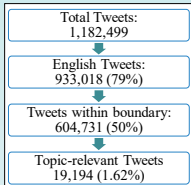
Methodology

Total Tweets: 1,182,499
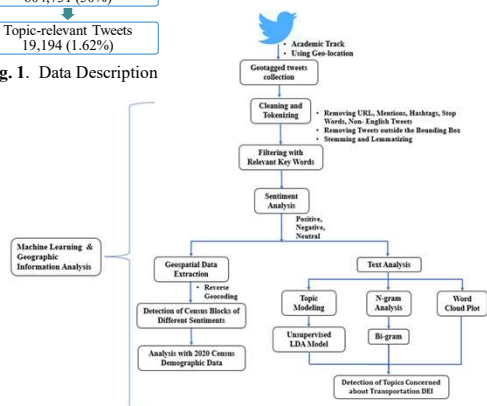English Tweets: 933,018 (79%)
Tweets within boundary: 604,731 (50%)
Topic-relevant Tweets 19,194 (1.62%)

Fig. 1. Data Description

Fig. 2. Framework for Data Collection, Preparation, and Analysis

Results and Discussions

Sentiment Analysis
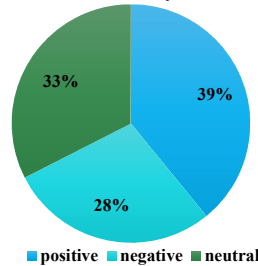

positive 39% negative 28% neutral 33%

Fig. 3. Sentiment of tweets analysis using Vader Lexicon Model

Fig. 4. Wordcloud of 50 most frequently discussed topic by netizens in DEI related tweets

Topic Mining

- Income & job — Poor, low, homeless, rich, income, job
- Gender — Gay, gender, women, language, deny
- Dependent group — child, young, wheelchair, disabled, drive
- Transportation modes — Bus, transit, train, flight, car
- Ride sharing — Driver, uber, time, dangerous, late, crime
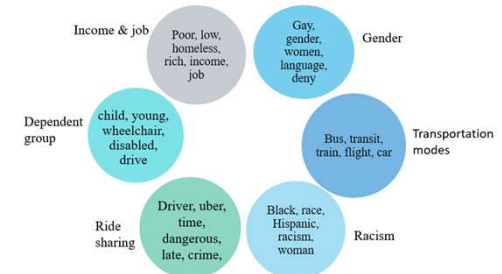- Racism — Black, race, Hispanic, racism, woman

Fig. 5. Broad classification of discussion topics using LDA model
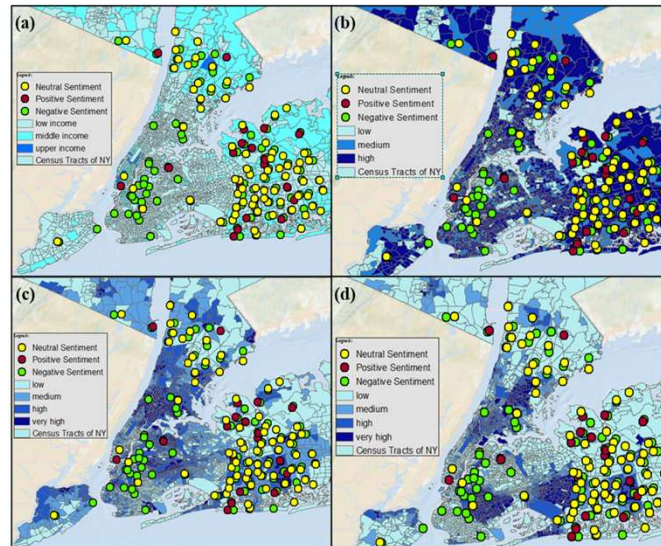
Demographic Factors Analysis

Fig. 8. Spatial distribution of tweeting activity across different socio-economic groups. (a) Per capita income,(b) Female population, (c) Hispanic/ Latino population, (d) Black/ African American population
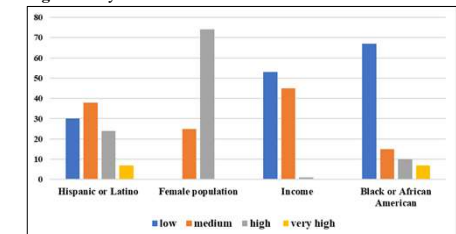

Fig. 6. Study Area

Fig. 7. Socio-demographic classification (by percentage) based on ethnicity, gender, income, and race in the study area

Conclusion and Key Insights

- People tweeting about DEI were more negative about issues like racism, income, unemployment, gender, ride dependency, transportation modes, and dependent group. It can be concluded that the area suffered from these issues most acutely.
- Locations where residents were more sensitive about DEIA had a high percentage of low-income people.
- People from locations with higher Latino and Hispanic population are more sensitive about DEIA of transportation system.
- People are more sensitive about inequity of transportation system where percentage of female population is high.
- Minority groups suffer more from inaccessibility and inequity.