

# Predicting Local Business Popularity during the Olympics through Geospatial Features

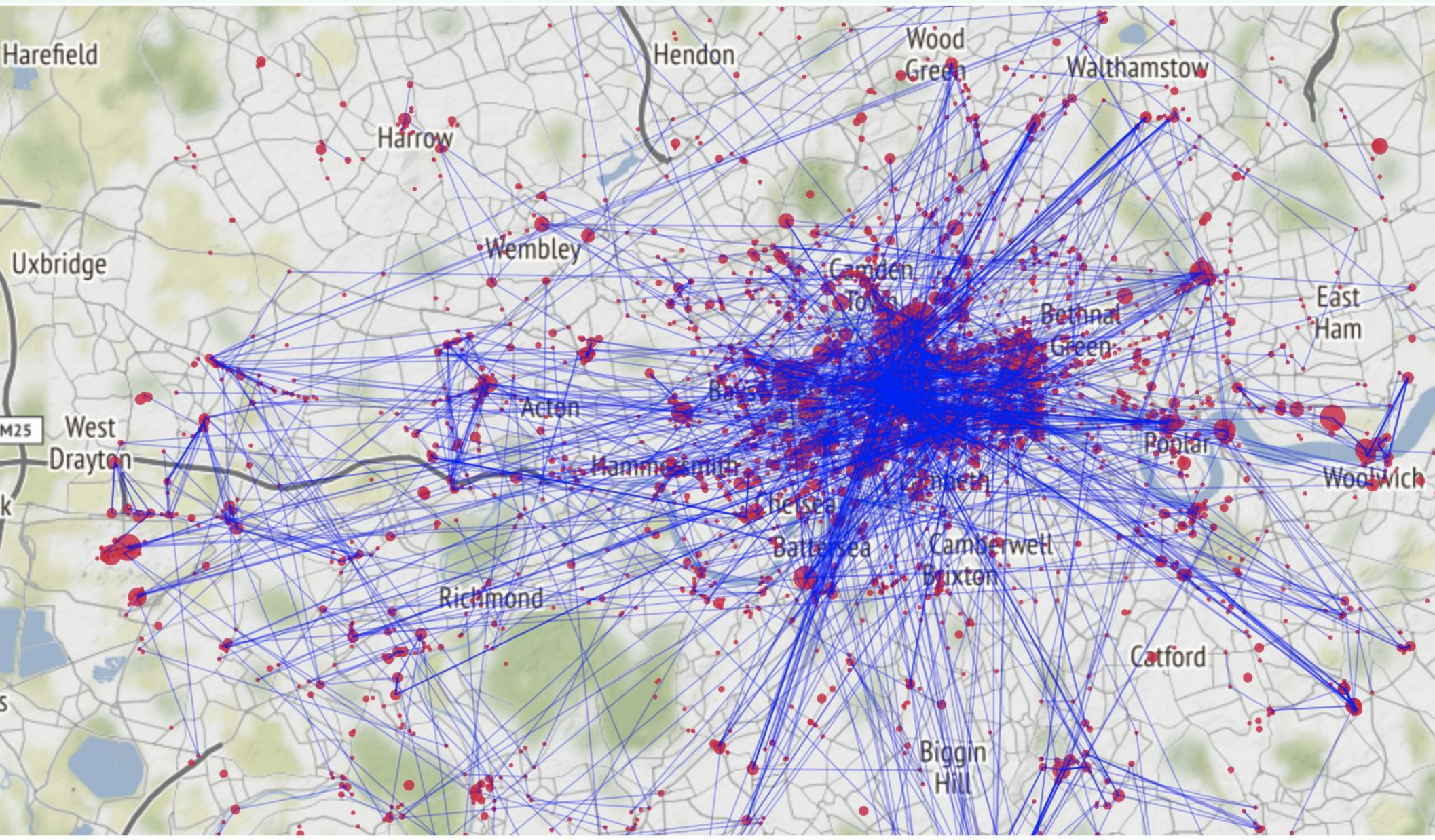
Shizhuo Duan, Vivek Gupta, Xu Han, Darshan Solanki

## Background

By understanding which factors impact a local business’ popularity during the Olympics, business owners can properly adapt to changes in customer flow during major event.

## Data

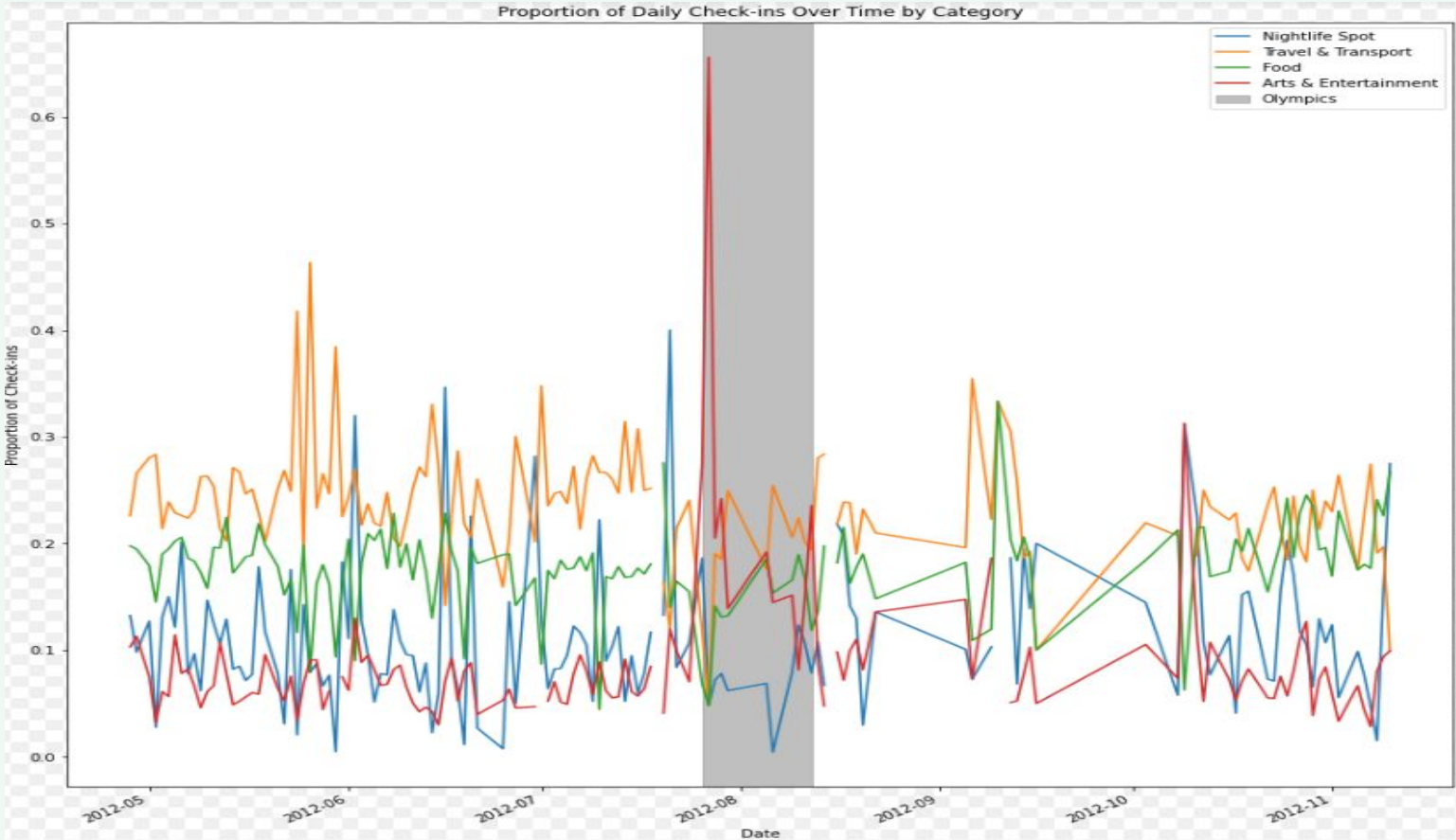
Foursquare checkin data for London local businesses (venues) in period around Olympics 2012. Also used Olympic events’ location data and London underground station data. To quantify the neighborhood quality of the venue we extracted features like Olympic Distance, Station Connectivity,neighborhood diversity and accessed the attractiveness of neighborhood using Jensen Quality Alpha. The plot below shows the transitions among venues before the Olympics. to determine effect on local businesses. The larger the dot, the more check-ins that venue had. Comparing the transitions before and during the game, we proposed four hypotheses.



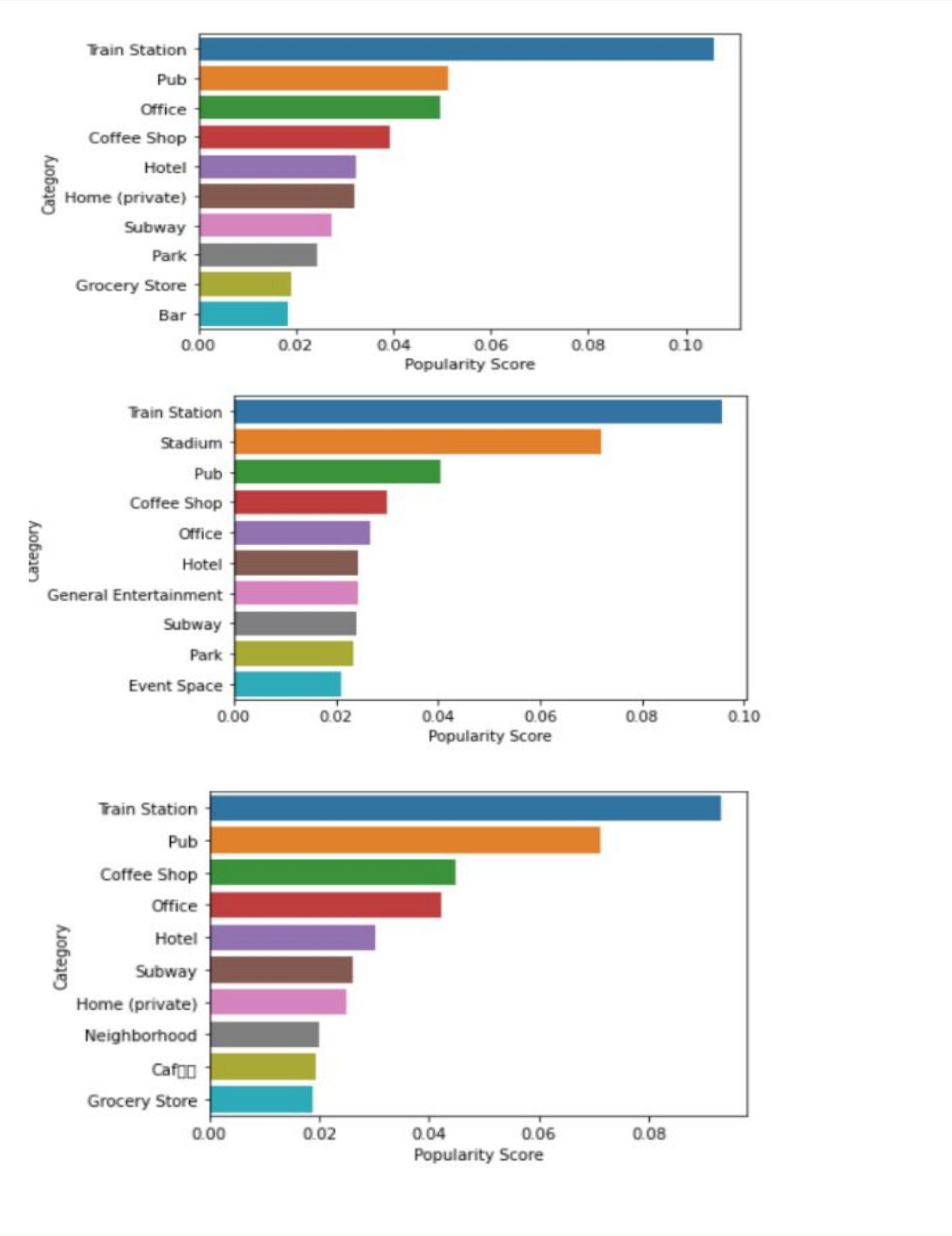
## Hypotheses

1. Close to travel station should increase popularity.
2. Close to Olympic event location should increase popularity.
3. Diversified neighborhood offers different activities and people will tend to stay longer and check more places.
4. The more attractive a venue is compared to its peers of same type, the more traffic it will gain from the event.

## Exploratory Data Analysis



People are engaging in entertainment activities during olympics. This thrives the impact on local business around those venues



Pre,During,Post Olympics Analysis for granular venues with relative popularity score calculation

## Model

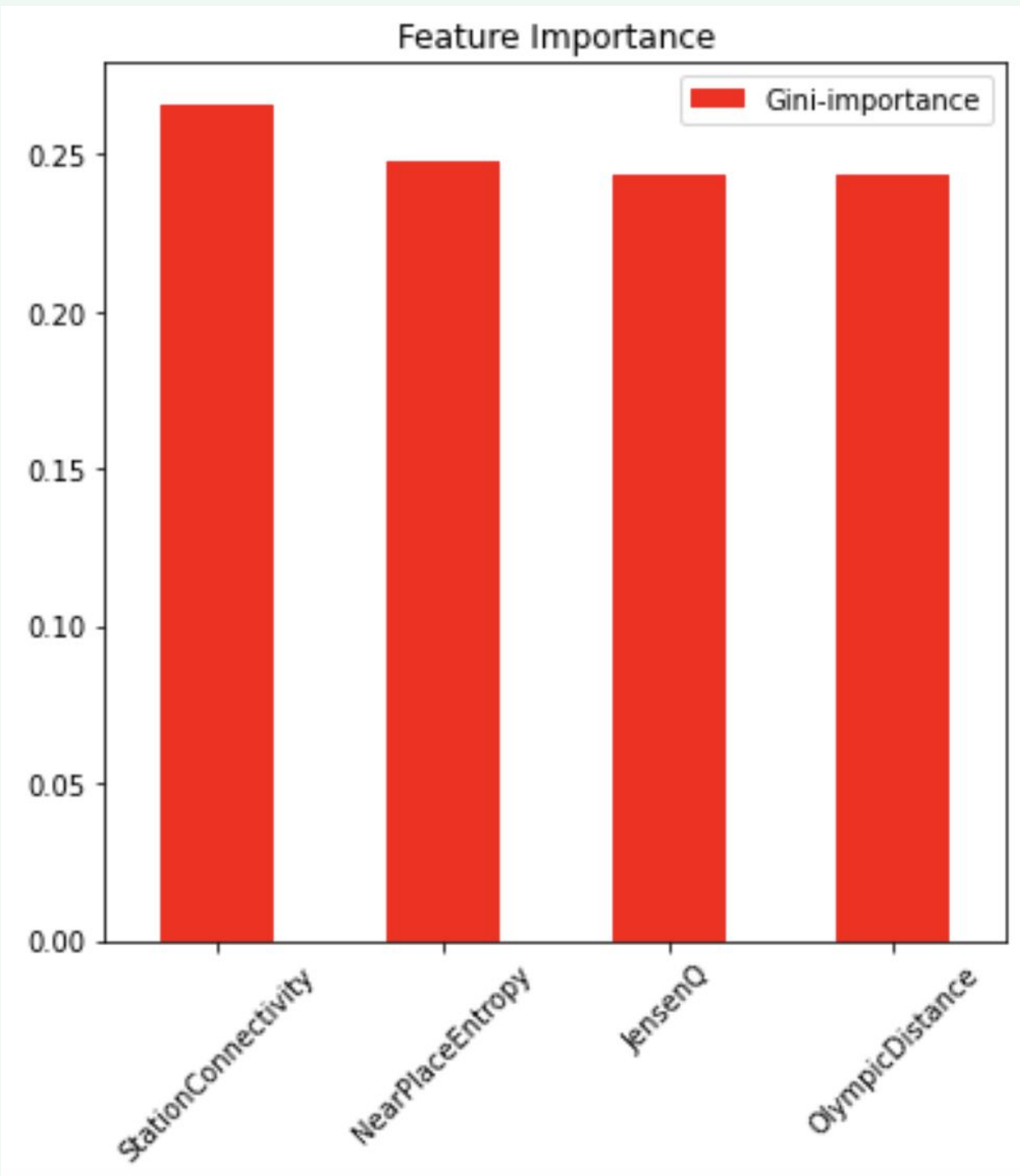
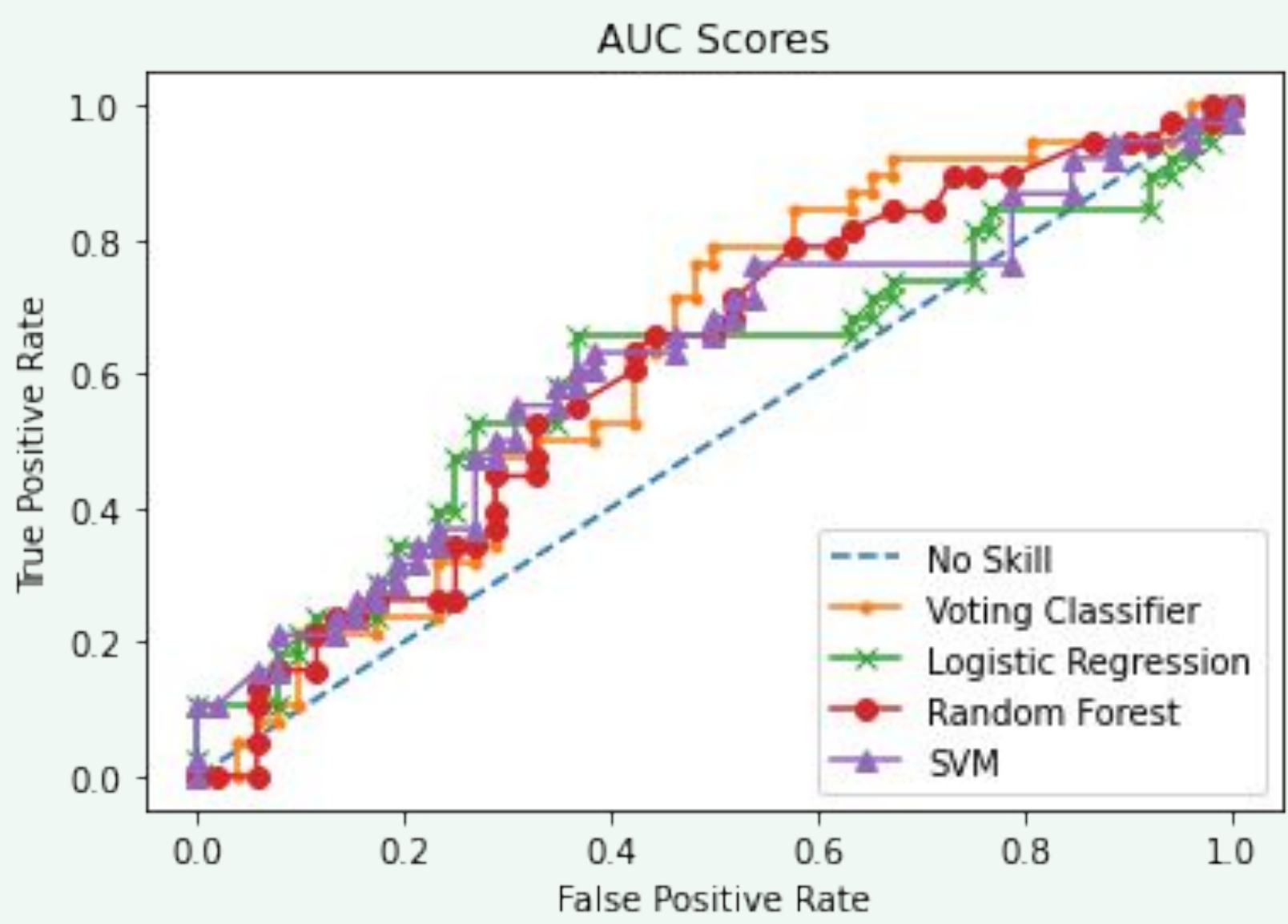
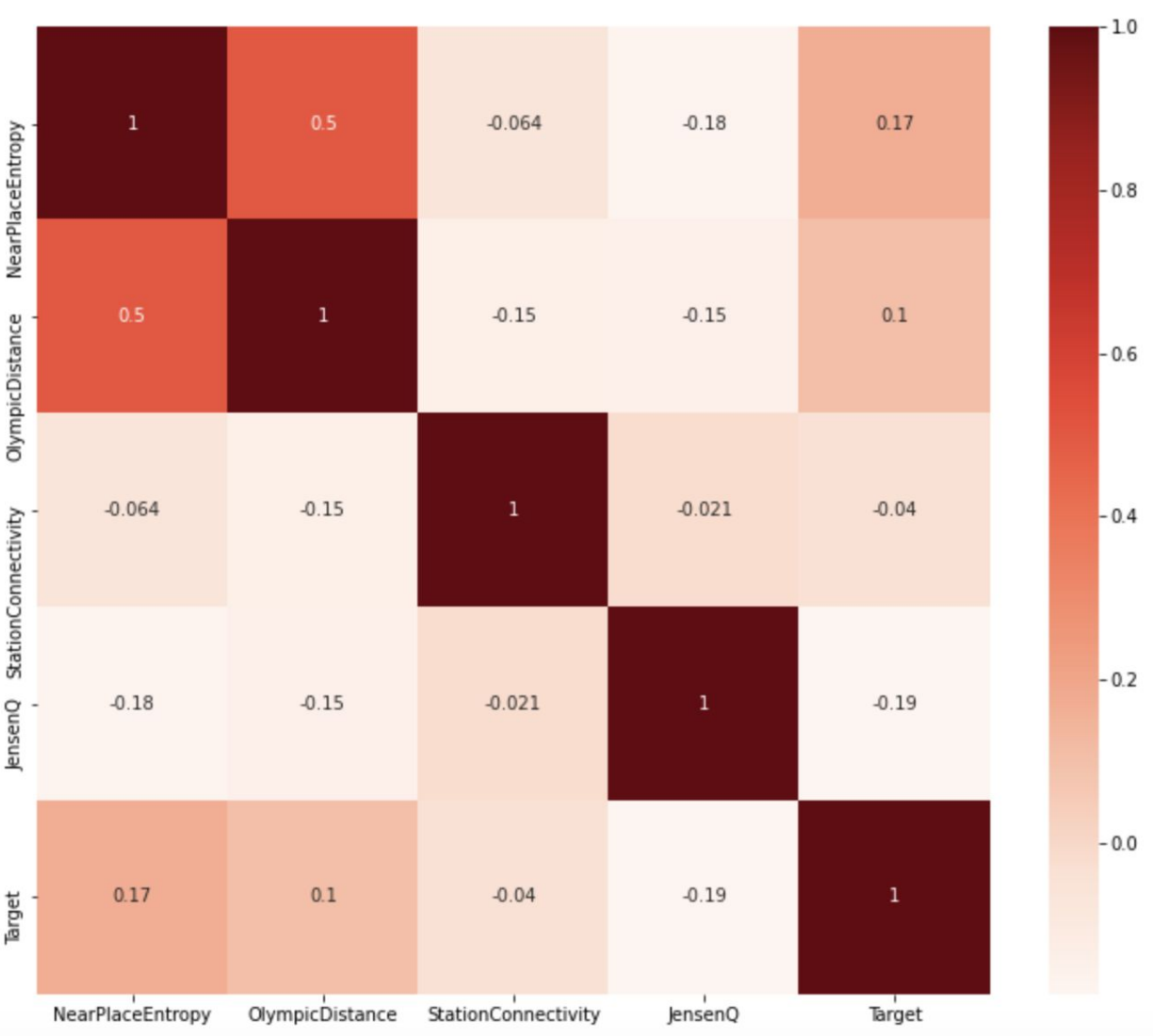
We perform a binary classification task. We label all food venues with 1 for an increase in popularity during Olympics and 0 otherwise. We experimented with SVM, Random Forest and Logistic Regression classifiers. Finally, we achieved the best AUC score using random forest. We prioritized the AUC score for analysis.

**During major event, better transport station connectivity may contribute more to foot traffic than short walking distance from event location, since we see more longer distance transitions. Besides spatial distance factors, neighborhood quality factors like diversity and venue attractiveness also contribute to change in popularity of a business.**

## Executive Summary

- Transport connectivity is the most important factor
- Diversity of neighborhood is the second most important factor
- Jensen Attractiveness alpha is the third most important factor
- Closeness to Olympic event locations is the least important factor

## Model Results



===== Test AUC Scores =====	
No Skill:	0.500
-----	
SVM:	0.539
Random Forest:	0.689
Logistic:	0.559
Voting:	0.664