McDonald's Business Strategy in San Diego

Background Information + Questions

Motivation:

What does the distribution of McDonald's restaurants tell us about the markets they aim to occupy/their marketing strategy? Does McDonalds predominantly establish restaurants in low or high income areas? In areas where people have little education? Near schools or other point of interest?

Our research aims to address the following questions:

1) Does fast food restaurant like Mcdonald's is strategically positioned their stores?
2) Using demographic and market data, is it possible to predict where McDonald's restaurants open?
3) Compared to other fast-food chain establishments, are these factors unique to McDonalds?

Study Design

Data sources and variables:

- ArcGIS Data Gallery:
  - Fast Food Locations (Mcdonald's, Subway, Burger King)
  - Demographics by census tract (Income, race, age, education, gender)
  - School locations (Elementary and high Schools)
  - Tourist Attractions
  - Business sites (Supermarkets and grocery stores)

- SANDAG Open Data Portal:
  - Child Care Centers

Methodology

Reprojection the CRS to EPSG 3857

Spatial join restaurant location **within** each census tract and get demographic information

Calculate distance for each restaurant to the **nearest POI** (school, child care center, attraction, business site)

Plot **correlation matrix** to visualize relationships between demographic variables and distance indicators

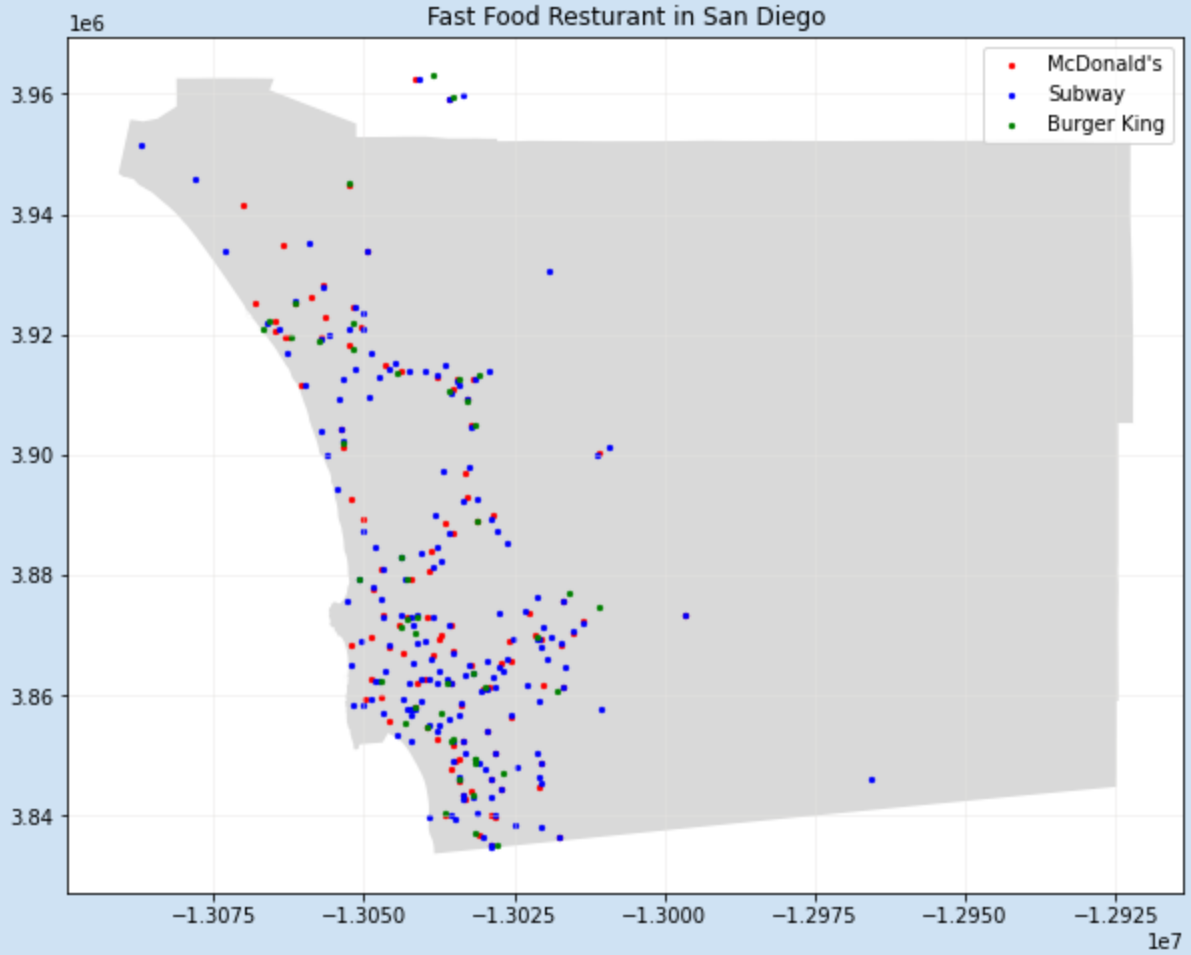
Split the data into train and test set, use **machine learning approach (Random Forest Regressor & Multi-output Random Forest Regressor)** to predict distances and locations based on multiple variables, ensuring the model is evaluated and fine-tuned for optimal performance

Fast Food in SD

106 Mcdonald’s

178 Subway

47 Burger King



Preliminary Findings (Mcdonald's)

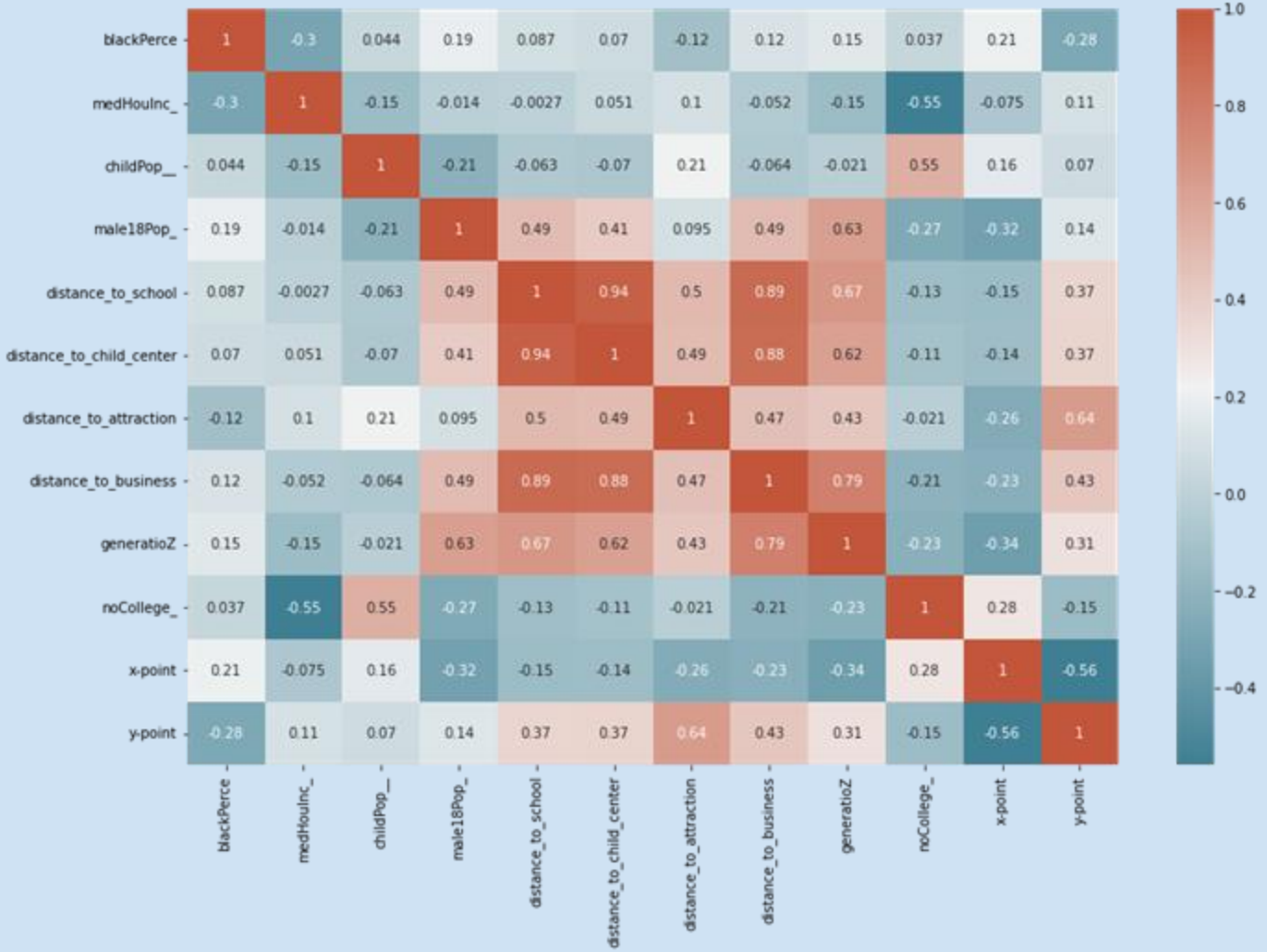
Distance to School Prediction:
R^2 on training set: 0.900
R^2 on test set: 0.752

Distance to Child Center Prediction:
R^2 on training set: 0.915
R^2 on test set: 0.677

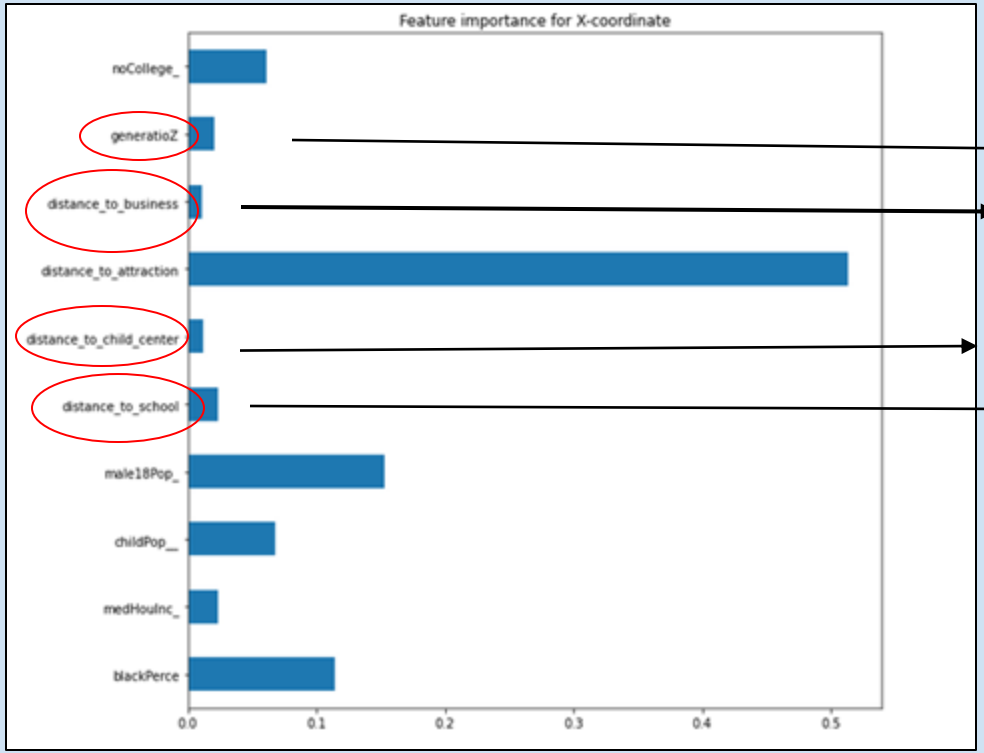
Distance to Attraction Prediction:
R^2 on training set: 0.959
R^2 on test set: 0.866

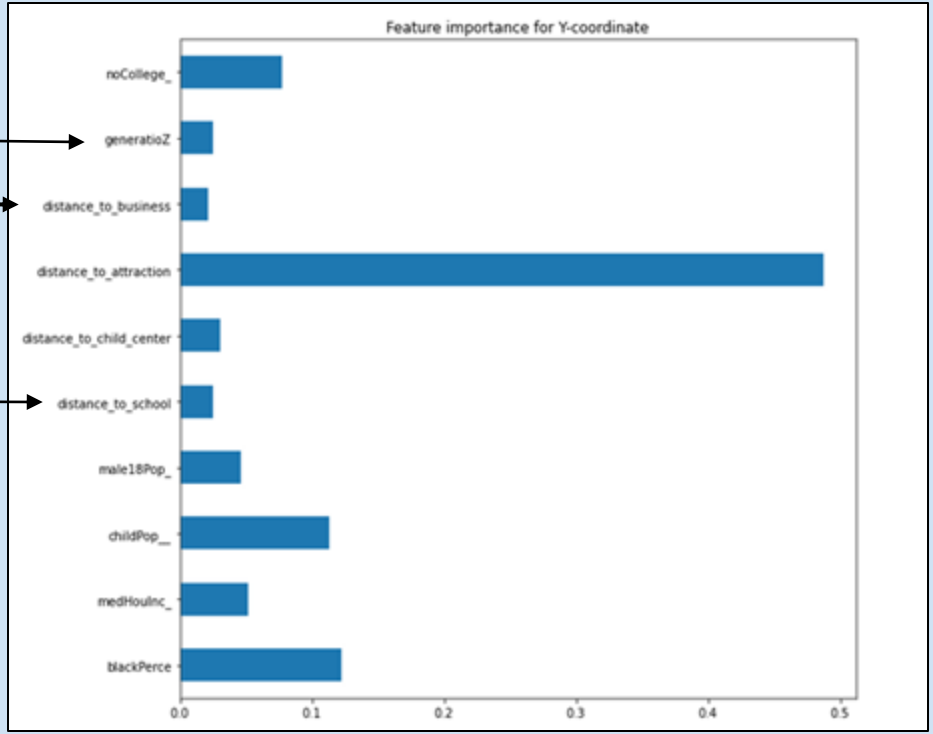
Distance to Business Prediction:
R^2 on training set: 0.914
R^2 on test set: 0.598

Location Prediction:
R^2 on training set: 0.906
R^2 on test set: 0.638



Variables of Low Importance





Results with Variables of High importance

With **low** importance variables:

the mean error of predicting x and y is

X: -0.0031%

Y: -0.026%

With **High** importance variables:

the mean error of predicting x and y is:

X: -0.0046%

Y: -0.043%

This means, removing low importance variables causes the RF model to produce worse coordinates

Preliminary Findings (Subway)

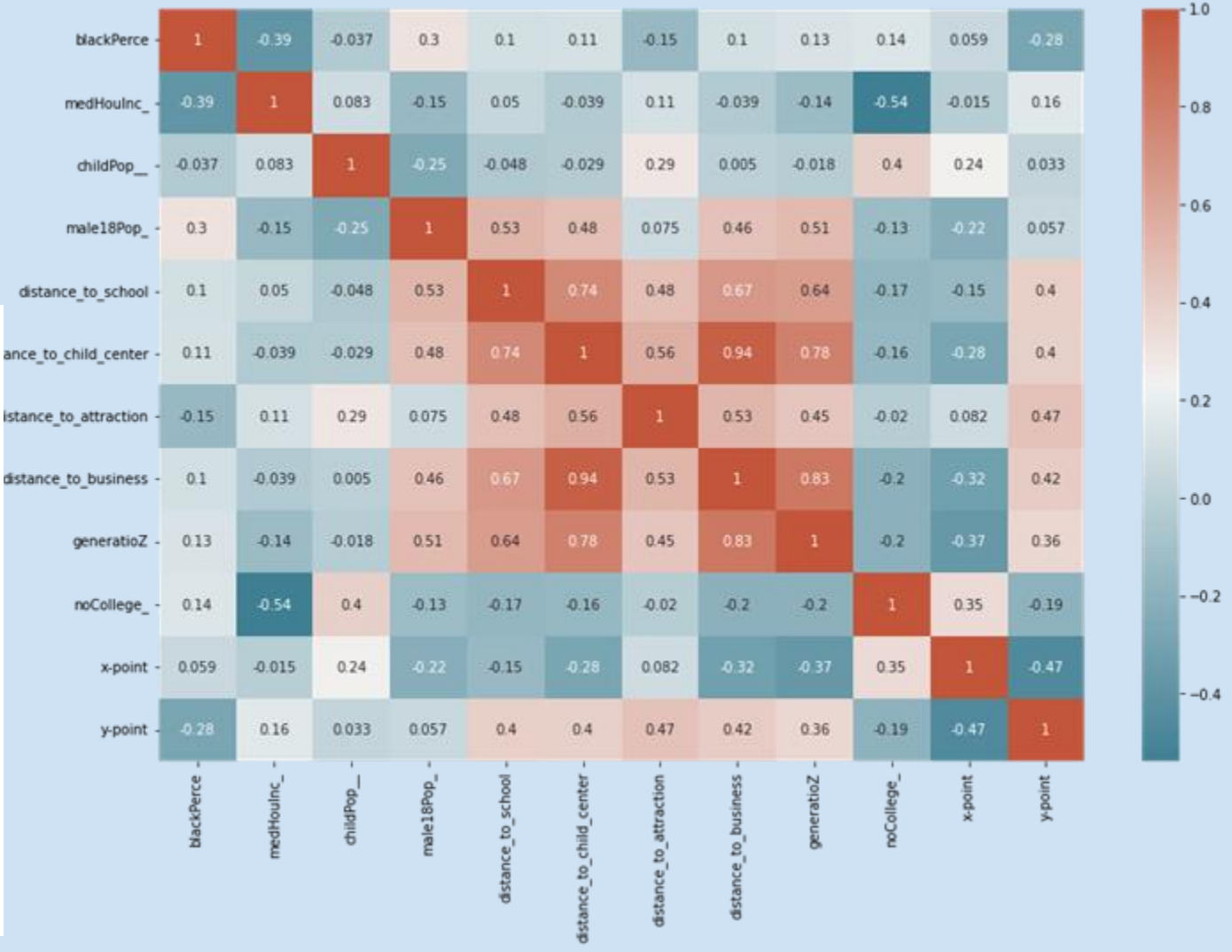
Distance to School Prediction:
R^2 on training set: 0.977
R^2 on test set: <0.1

Distance to Child Center Prediction:
R^2 on training set: 0.956
R^2 on test set: 0.521

Distance to Attraction Prediction:
R^2 on training set: 0.965
R^2 on test set: 0.386

Distance to Business Prediction:
R^2 on training set: 0.943
R^2 on test set: 0.116

Location Prediction:
R^2 on training set: 0.860
R^2 on test set: 0.116



Preliminary Findings (Burger King)

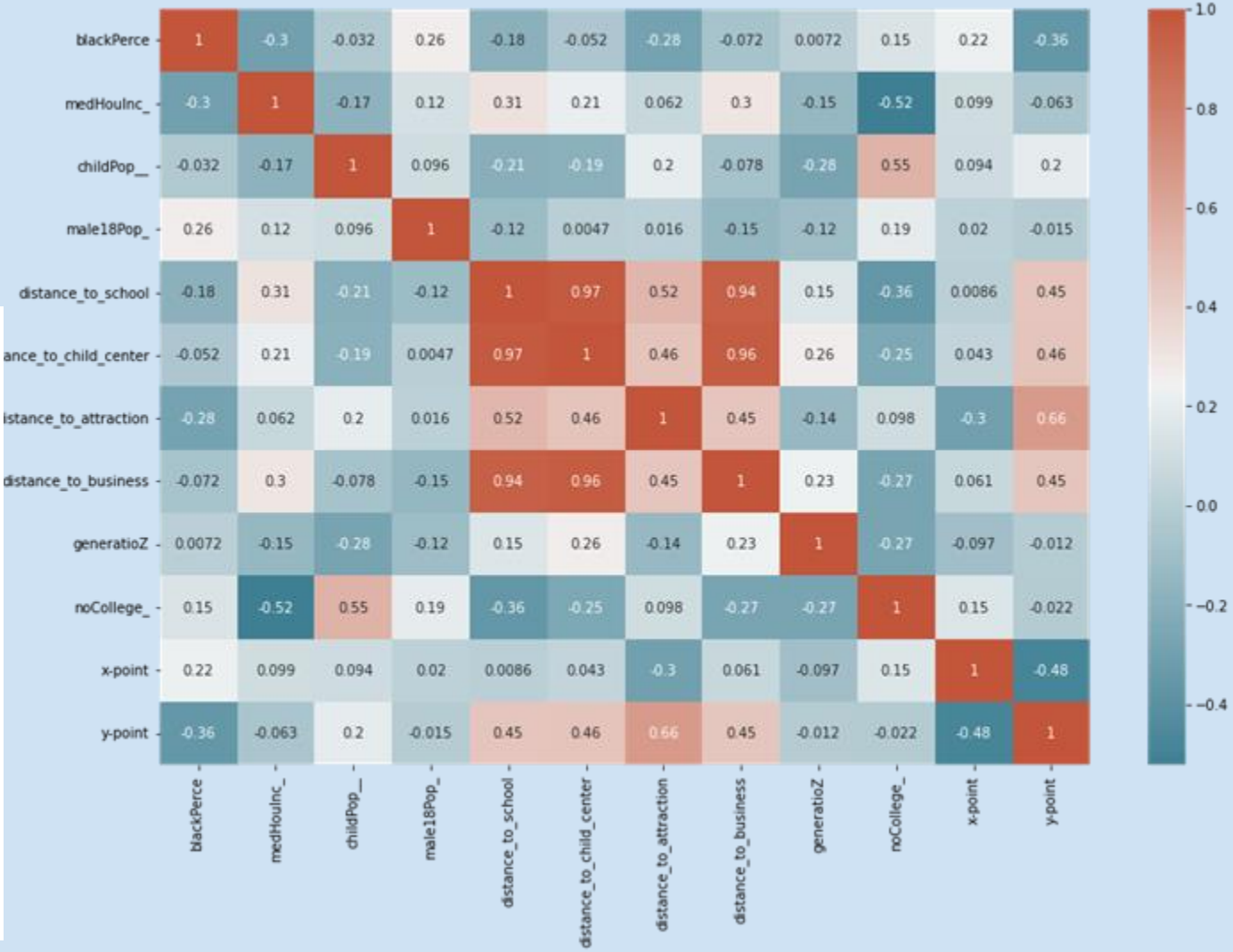
Distance to School Prediction:
R^2 on training set: 0.942
R^2 on test set: <0.1

Distance to Child Center Prediction:
R^2 on training set: 0.947
R^2 on test set: 0.243

Distance to Attraction Prediction:
R^2 on training set: 0.971
R^2 on test set: 0.886

Distance to Business Prediction:
R^2 on training set: 0.942
R^2 on test set: 0.676

Location Prediction:
R^2 on training set: 0.912
R^2 on test set: 0.404



RF Model for McDonald's vs Subway & BK

Subway Location Prediction:

R^2 on training set: 0.860

R^2 on test set: 0.116

Burger King Location Prediction:

R^2 on training set: 0.912

R^2 on test set: 0.404

McDonald's Location Prediction:

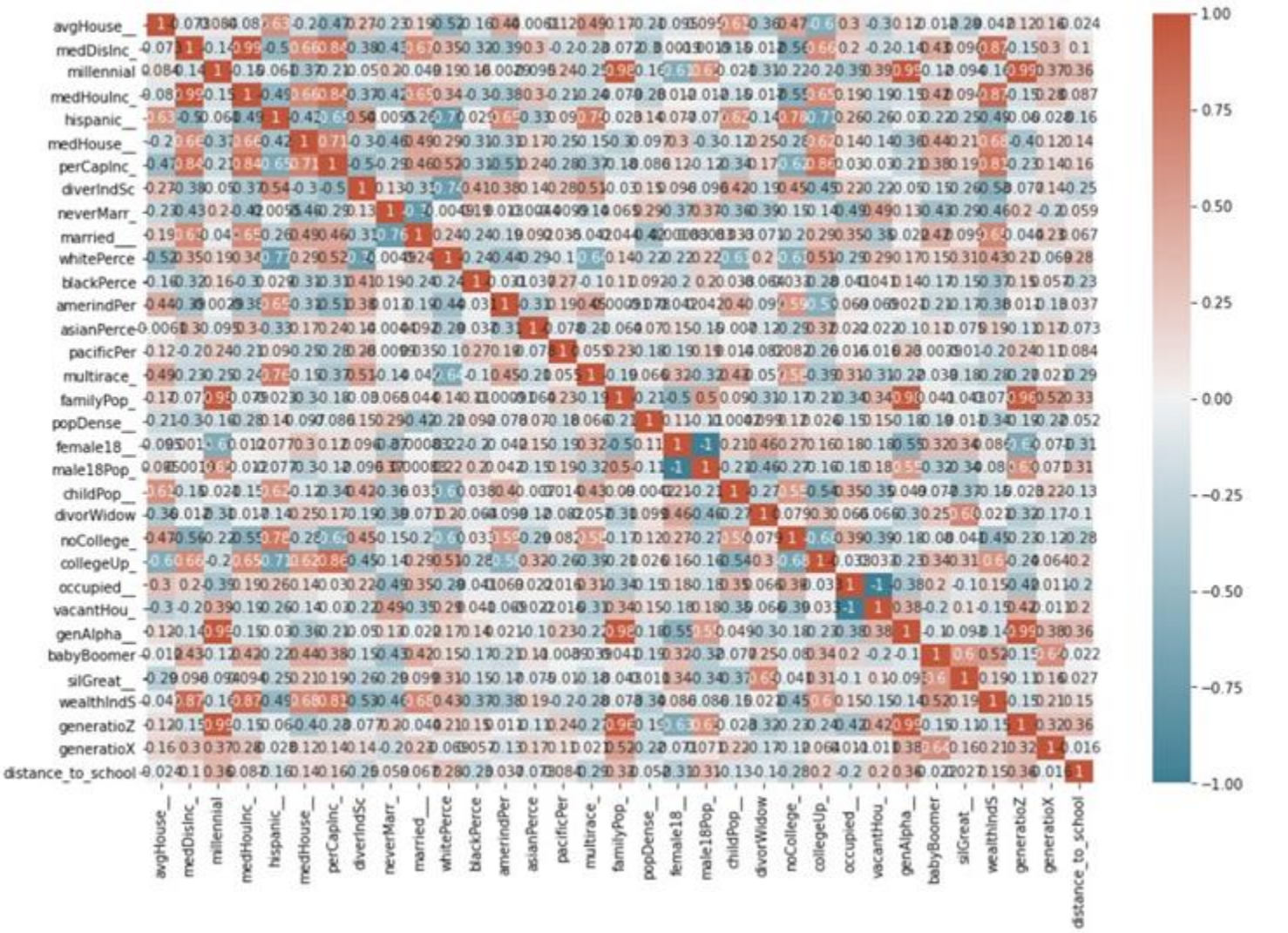
R^2 of training set: 0.921

R^2 of test set: 0.300

Add more variable?

R^2>0.8 in RF

More Explanatory Power



Key Insights

- **Correlation Matrix**
  - Distances to school, child center, attraction, business sites are **co-located** for all three
  - **GenZ ratio has positive correlations with distances** indicators for Mcdonald and Subway.
  - Mcdonald's and Burger King specifically **optimize their locations among schools, child center and business sites**, while Subway focus on places close to child care center
- **Random Forest Regression Results**
  - McDonald's model has a relatively **high R^2 on the test set**, suggesting that the model has good predictive power for data, and it can provide valuable information for strategic decision-making.
  - **High R^2 on the training set but low R^2 on the test set** across models for Subway and Burger King indicate potential overfitting.
- **Strategic Operation**
  - **Co-locate** near schools, child centers, and attractions.
  - Tailor campaigns to **Generation Z** in areas with schools and attractions
  - Focus on **mixed demographics information** of different areas to leverage both younger populations near schools and business professionals.

Next Steps/Challenges

- Limited by census tract data
- RF model for predicting location is still very unrefined
- Finding more variables and data for improving the model
- Small tweaks: *need* **scaling**, **normalizing**, double-checking variable
- Finding a way to plot predicted locations onto a map widget
- Experimenting with other Machine Learning models (KNN, Linear, etc.)