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

- SANDAG Open Data Portal:
 - o Child Care Centers

Methodology

Reprojection the CRS to EPSG 3857

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

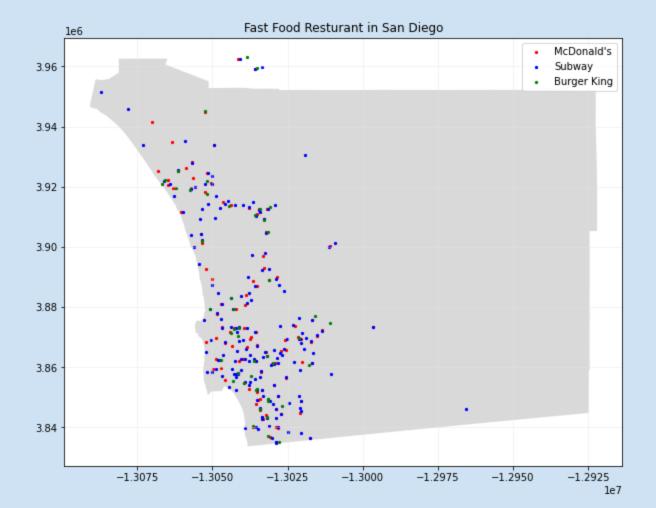
Calculate distance for each restaurant to the nearest POI (school, child care center, attraction, business site)

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's178 Subway47 Burger King



Preliminary Findings (Mcdonald's)

Distance to School Prediction:

R² on training set: 0.900 R² on test set: 0.752

Distance to Child Center Prediction:

R² on training set: 0.915 R² on test set: 0.677

Distance to Attraction Prediction:

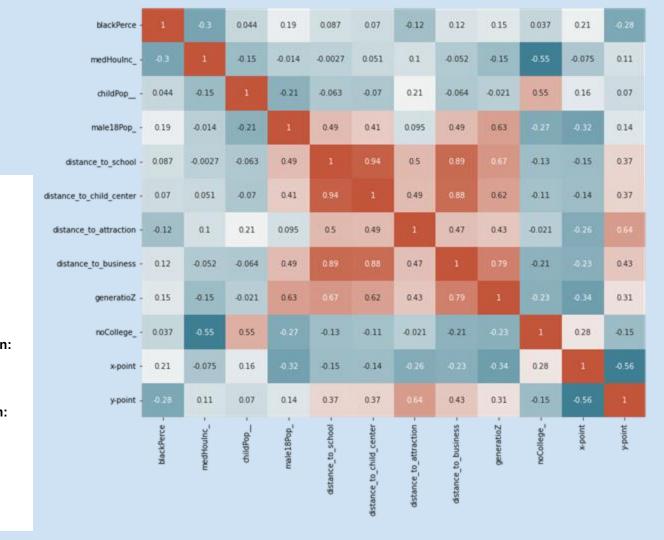
R² on training set: 0.959 R² on test set: 0.866

Distance to Business Prediction:

R² on training set: 0.914 R² on test set: 0.598

Location Prediction:

R² on training set: 0.906 R² on test set: 0.638



-08

- 0.6

- 0.4

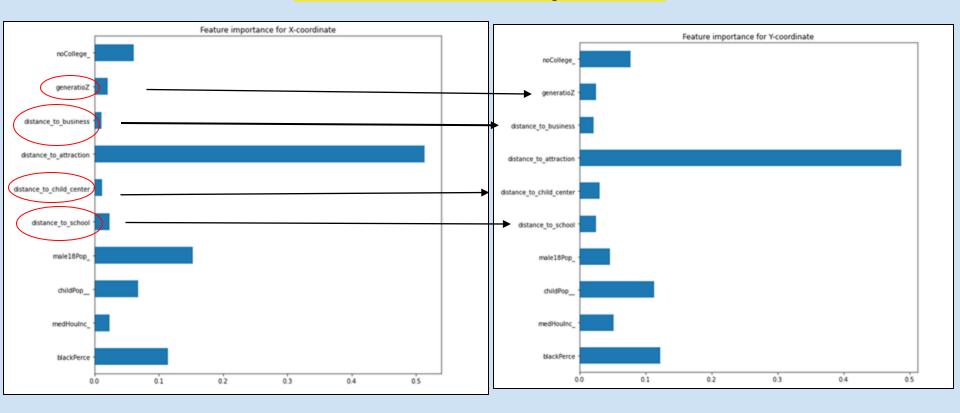
- 0.2

- 0.0

--0.2

--0.4

Variables of Low Importance



Results with Variables of High importance

With <u>low</u> importance variables:

With **<u>High</u>** importance variables:

the mean error of predicting x and y is

X: -0.0031%

Y: -0.026%

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 <u>worse</u> coordinates

Preliminar Findings (Subway) **Distance to School Pre**

D 11 1	blackPerce -		-0.39	-0.037	0.3	0.1	0.11	-0.15	0.1	0.13	0.14	0.059	-0.28
Preliminary medHoulnc Findings childPop		0.39		0.083	0.15	0.05	-0.039	0.11	-0.039	-0.14	-0.54	-0.015	0.16
		-0.037	0.083	1	-0.25	-0.048	-0.029	0.29	0.005	-0.018	0.4	0.24	0.033
(Subway)	male18Pop	0.3	0.15			0.53	0.48	0.075	0.46	0.51	-0.13		0.057
(distance_to_school -	0.1	0.05	-0.048	0.53			0.48		0.64	-0.17	-0.15	0.4
Distance to School Prediction: R² on training set: 0.977 R² on test set: <0.1	ance_to_child_center -	0.11	-0.039	-0.029	0.48	0.74	1	0.56	0.94	0.78	-0.16	-0.28	0.4
	istance_to_attraction -	-0.15	0.11	0.29	0.075	0.48	0.56		0.53	0.45	-0.02	0.082	0.47
Distance to Child Center Prediction:	distance_to_business -	0.1	-0.039	0.005	0.46		0.94	0.53	1	0.83	-0.2	-0.32	0.42
R ² on training set: 0.956 R ² on test set: 0.521	generatioZ -	0.13	-0.14	-0.018	0.51	0.64	0.78	0.45	0.83		-0.2	-0.37	0.36
Distance to Attraction Prediction: R ² on training set: 0.965 R ² on test set: 0.386	noCollege	0.14	-0.54	0.4	-0.13	-0.17	0.16	-0.02	-0.2	-0.2		0.35	-0.19
	x-point -	0.059	-0.015	0.24		-0.15	-0.28	0.082	-0.32	-0.37	0.35	1	-0.47
Distance to Business Prediction: R ² on training set: 0.943 R ² on test set: 0.116	y-point -	-0.28	0.16	0.033	0.057	0.4	0.4	0.47	0.42	0.36	-0.19	-0.47	1
		Perce -	- Julio	do	- dog	- lood	enter -	ction -	ness -	atioZ -	lege -	point -	point -

- 0.8

- 0.6

- 0.4

- 0.2

- 0.0

- -0.2

Distance to Business P R² on training set: 0.943 R² on test set: 0.116 **Location Prediction:** R² on training set: 0.860 R² on test set: 0.116

D 1: :	blackPerce -	1	-0.3	-0.032	0.26	-0.18	-0.052	-0.28	-0.072	0.0072	0.15	0.22	-0.36
Preliminary	medHoulnc	-0.3		-0.17	0.12	0.31	0.21	0.062	0.3	-0.15	-0.52	0.099	-0.063
Findings (Burger King) (Burger King) (Burger King)		-0.032	-0.17	1	0.096	-0.21	0.19	0.2	-0.078	-0.28	0.55	0.094	0.2
		0.26	0.12	0.096	1	-0.12	0.0047	0.016	-0.15	-0.12	0.19	0.02	-0.015
		0.18	0.31	-0 21	-0.12	1	0.97	0.52	0.94	0.15	0.36	0.0086	0.45
Distance to School Prediction: R² on training set: 0.942 R² on test set: <0.1	ance_to_child_center -	-0.052	0.21		0.0047	0.97		0.46	0.96	0.26		0.043	0.46
	istance_to_attraction -	-0.28	0.062	0.2	0.016	0.52	0.46	1	0.45	-0.14	0.098	-0.3	0.66
Distance to Child Center Prediction:	distance_to_business -	-0.072	0.3	-0.078	-0.15	0.94	0.96	0.45	1	0.23	-0.27	0.061	0.45
R ² on training set: 0.947 R ² on test set: 0.243	generatioZ -	0.0072	0.15	-0.28	-0.12	0.15	0.26	-0.14	0.23	1	0.27	-0.097	-0.012
Distance to Attraction Prediction: R ² on training set: 0.971	noCollege	0.15	0.52	0.55	0.19	-0.36	-0.25	0.098	-0.27	-0.27	1	0.15	-0.022
R ² on test set: 0.886	x-point -	0.22	0.099	0.094	0.02	0.0086	0.043	-0.3	0.061	-0.097	0.15	1	-0.48
Distance to Business Prediction: R ² on training set: 0.942	y-point -	-0.36	-0.063	0.2	-0.015	0.45	0.46		0.45	-0.012	-0.022	-0.48	1
R ² on test set: 0.676		dackPerce -	medHoulnc	drildPop	nale18Pop	school -	center -	raction -	usiness -	generatioZ -	noCollege	x-point -	y-point -
Location Prediction: R² on training set: 0.912 R² on test set: 0.404		blac	MedH	6	male	distance_to_school	distance_to_child_	distance_to_attraction	distance_to_business	aua6	noc		

- 0.8

- 0.6

- 0.4

- 0.2

- 0.0

--0.2

- -0.4

RF Model for McDonald's vs Subway & BK

Subway Location Prediction:

R² on training set: 0.860

R² on test set: 0.116

Burger King Location Prediction:

R² on training set: 0.912

R² on test set: 0.404

McDonald's Location Prediction:

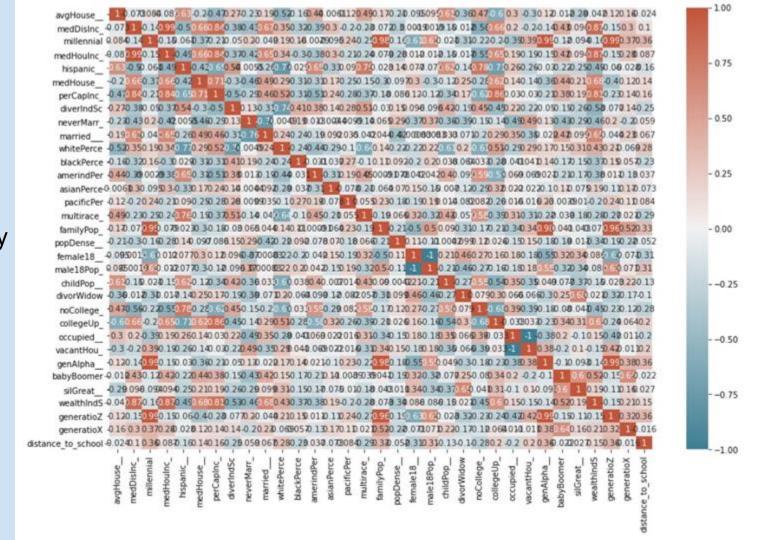
R² of training set: 0.921

R² 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 R² 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 R² on the training set but low R² 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.)