



EXPLORATION OF HIGH INCOME AND LOW-INCOME AMENITIES USING IRS DATA AND FOURSQUARE

ERIC ROBERTS

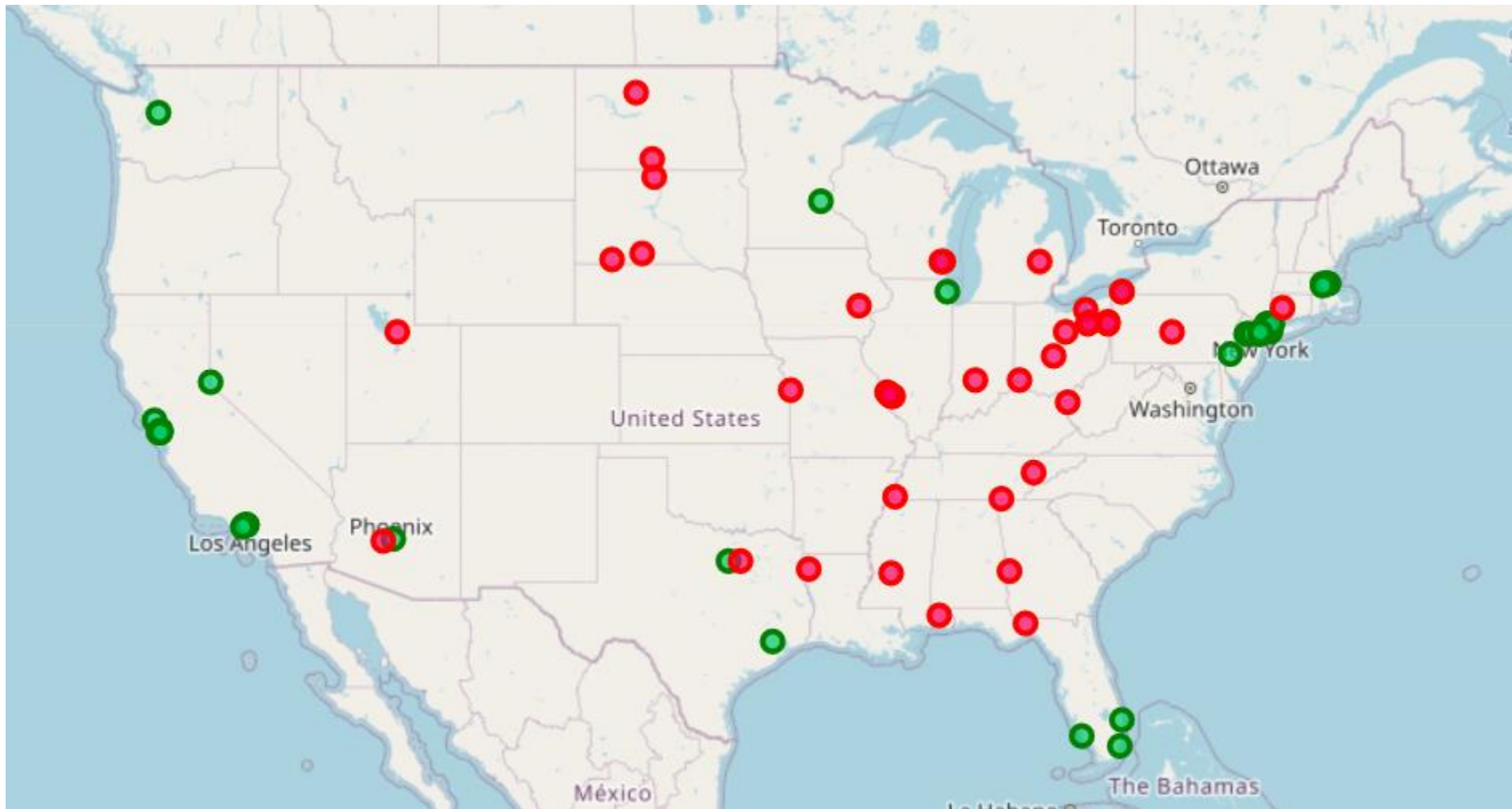


DO TOP INCOME ZIP CODES IN THE US SHARE THE SAME AMENITIES WITH LOWER-INCOME ZIP CODES?

- Top Income ZIP Codes in the US can be found using the IRS database.
- Their latitude and longitude can be found using another dataset. A merge yield the following.
- The same procedure can be used to get a list of the *lowest* income ZIP codes.

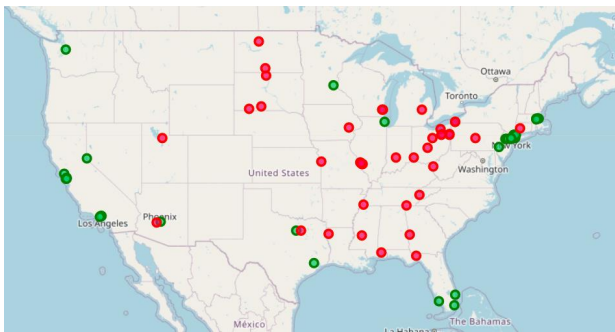
	zipcode	agi_stub	STATEFI PS	Average AGI	weighted _avg_p	NI	NI_SUM	income	City	State	Latitude	Longitud e	Timezon e	Daylight savings time flag	geopoint
8977	33109	21	72	2.198236e +06	1.0	280.0	1680.0	3.847401e +06	Miami Beach	FL	25.759474	-80.13907	-5	I	25.759474,-80.13907
25861	94027	21	36	1.718305e +06	1.0	3240.0	19440.0	3.813630e +06	Atherton	CA	37.458615	-122.20009	-8	I	37.458615,-122.20009
22395	76102	21	288	1.455423e +06	1.0	4340.0	26040.0	1.305899e +07	Fort Worth	TX	32.753880	-97.32987	-6	I	32.75388,-97.32987
9112	33480	21	72	1.369855e +06	1.0	5880.0	35280.0	3.754224e +06	Palm Beach	FL	26.715067	-80.03937	-5	I	26.715067,-80.03937
25884	94104	21	36	1.133140e +06	1.0	1290.0	7740.0	3.596518e +06	San Francisco	CA	37.791728	-122.40190	-8	I	37.791728,-122.4019

INCOME LEVELS ACROSS US



- Found using IRS AGI per ZIPCODE and then mapped
- Green = Top 40 Richest Zip Codes
- Red = Lowest 40 Poorest Zip Codes

WHAT DO THESE ZIPCODES HAVE AS LOCAL AMENITIES?



--poor----			--rich----		
venue	freq	common	venue	freq	common
0 Fast Food Restaurant	0.05	0	0 Coffee Shop	0.05	0.05
1 Pizza Place	0.05	0	1 Italian Restaurant	0.04	0
2 Coffee Shop	0.04	0.04	2 Hotel	0.04	0
3 Bar	0.04	0	3 American Restaurant	0.03	0.03
4 Sandwich Place	0.03	0.03	4 Park	0.03	0
5 American Restaurant	0.03	0.03	5 Steakhouse	0.02	0
6 Pharmacy	0.02	0	6 Bakery	0.02	0
7 Café	0.02	0.02	7 Restaurant	0.02	0.02
8 Convenience Store	0.02	0	8 Burger Joint	0.02	0
9 Diner	0.02	0	9 Café	0.02	0.02
10 Discount Store	0.02	0	10 Sushi Restaurant	0.02	0
11 Gas Station	0.02	0	11 Seafood Restaurant	0.02	0
12 Hookah Bar	0.01	0	12 Sandwich Place	0.02	0.02
13 Middle Eastern Restaurant	0.01	0	13 Dessert Shop	0.01	0
14 Restaurant	0.01	0.01	14 Deli / Bodega	0.01	0
	0.39	0.13		0.37	0.14

- Conclusions:

- The poor amenities are focused on basic needs: pharmacy, gas station, etc.
- There's about 3x the amenities in rich zip codes than poor ones
- Does everyone like coffee? Seems so. There's commonality there, American fare, and sandwich places.

NEXT STEPS?

- Using this data and the approach, a greater clustering analysis could be taken at middle income bands.
- The income information by ZIP code is useful for other analyses.
- All code is on GitHub: https://github.com/projekt888/Coursera_Capstone/tree/master/Final