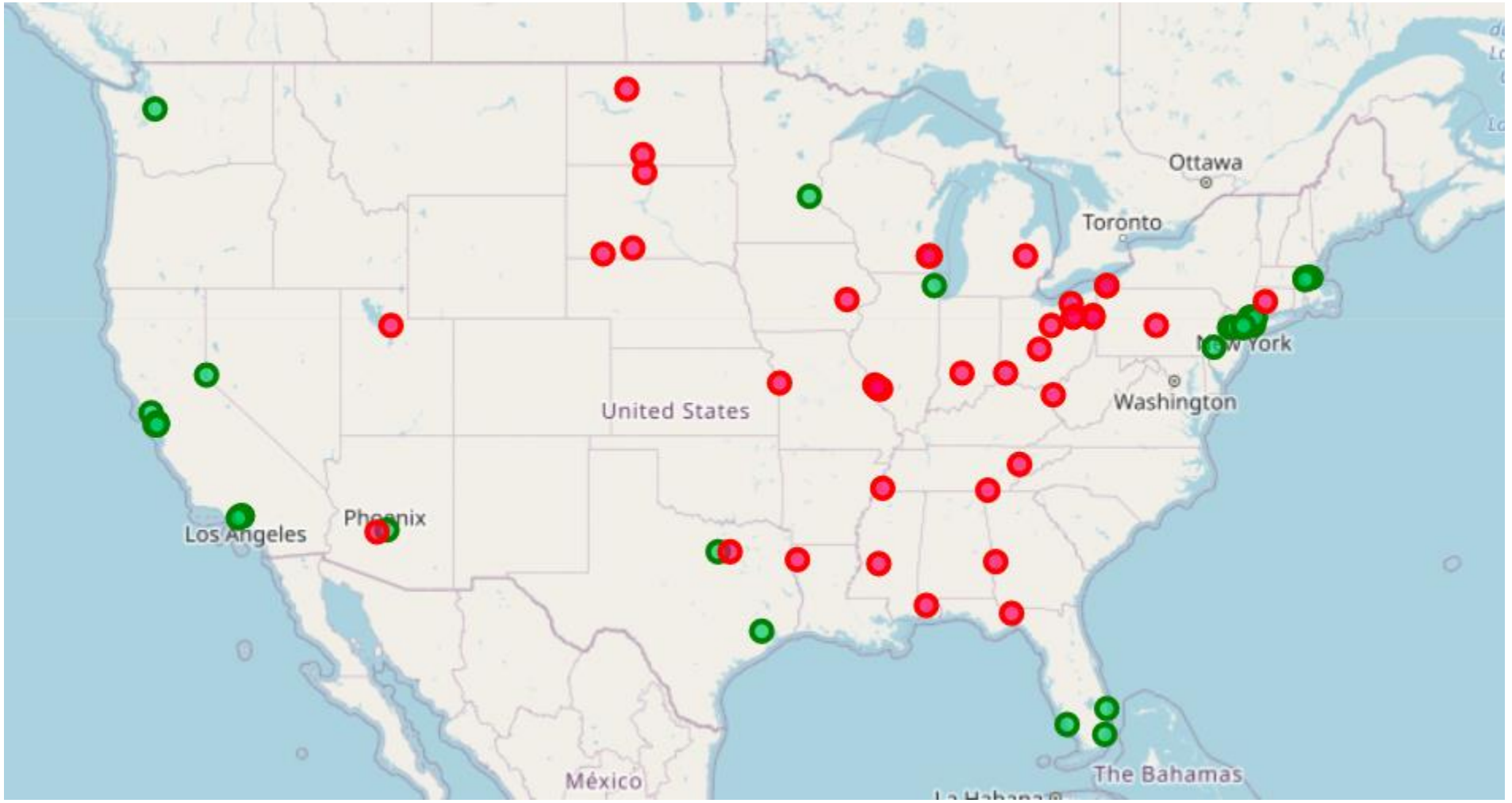
EXPLORATION OF HIGH INCOME AND LOW-INCOME AMENITIES USING IRS DATA AND FOURSQUARE

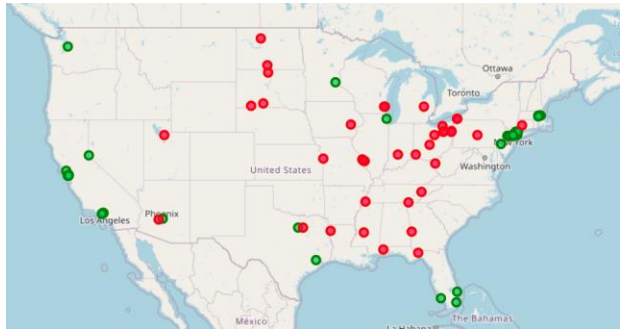
ERIC ROBERTS

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