

Data Science Capstone project

JAI Borst, 19 november 2019

Business problem

Objective: Best location to open a new Italian restaurant on Curacao

Stakeholders:

- Tourism Board Curacao
- Entrepreneurs

Background:

-Neighborhoods Curacao highly variable in average income and % immigrants→ likelihood of succes

Data

1. Curacao Census 2011 (CSV), demographic data

FID	NAME	HH	pop	hhsz	sexratio	p0_14	p15_64	p65u	curborn
1	PARADIJS BIJ HEL	23	59	2.6	0	0.000	0.000	0.000	0.000
2	ABRAHAMSZ	220	573	2.5	85	0.162	0.651	0.187	0.713

2. GeoJson neighborhoods Curacao

3. Foursquare API



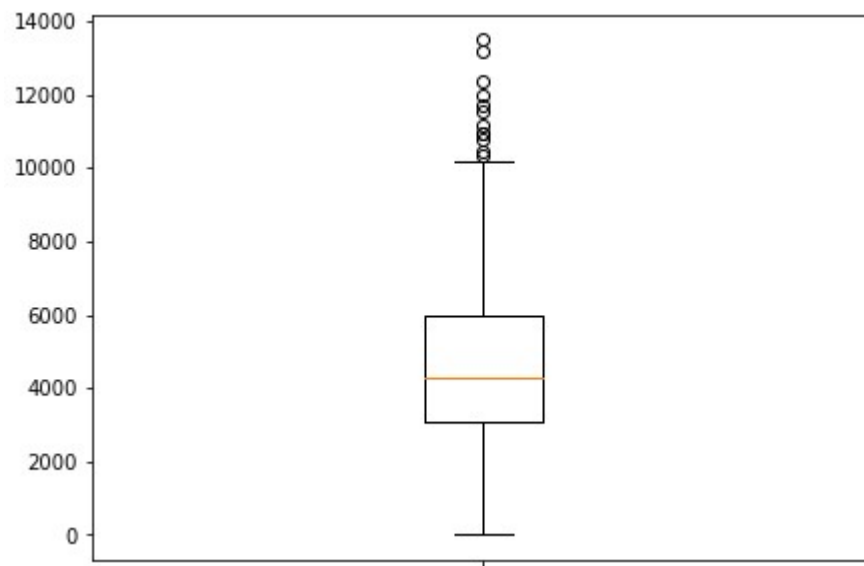
Methodology

A. Compare neighborhoods 'Census' to 'GeoJson'

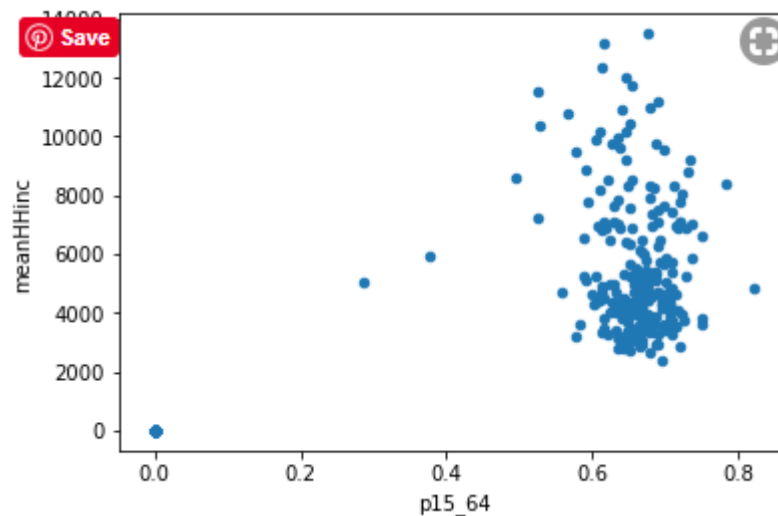
B. Describe Dataset 'Census'

	FID	HH	pop	hhsz	sexratio	p0_14	p15_64	p65u	curborn	forborn	...	airpolut
count	291.000000	291.000000	291.000000	291.000000	291.000000	291.000000	291.000000	291.000000	291.000000	291.000000	...	291.000000
mean	146.000000	188.783505	516.731959	2.582818	67.130584	0.153546	0.527997	0.119137	0.590570	0.210117	...	16.787973
std	84.148678	241.905377	681.345786	0.475228	34.668630	0.084837	0.268170	0.087312	0.310049	0.141562	...	19.718940
min	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000
25%	73.500000	44.000000	119.000000	2.500000	68.500000	0.136500	0.592500	0.060500	0.558000	0.121500	...	5.500000
50%	146.000000	98.000000	270.000000	2.600000	81.000000	0.180000	0.652000	0.126000	0.714000	0.215000	...	9.700000
75%	218.500000	233.000000	602.000000	2.800000	88.000000	0.207500	0.681500	0.171000	0.805000	0.313000	...	18.250000
max	291.000000	1455.000000	3892.000000	5.000000	113.000000	0.329000	0.822000	0.651000	0.953000	0.550000	...	100.000000

C. Average household income: 4408 ANG (local currency)



D. Scattering of mean household income for persons between 15 and 64 years of age.



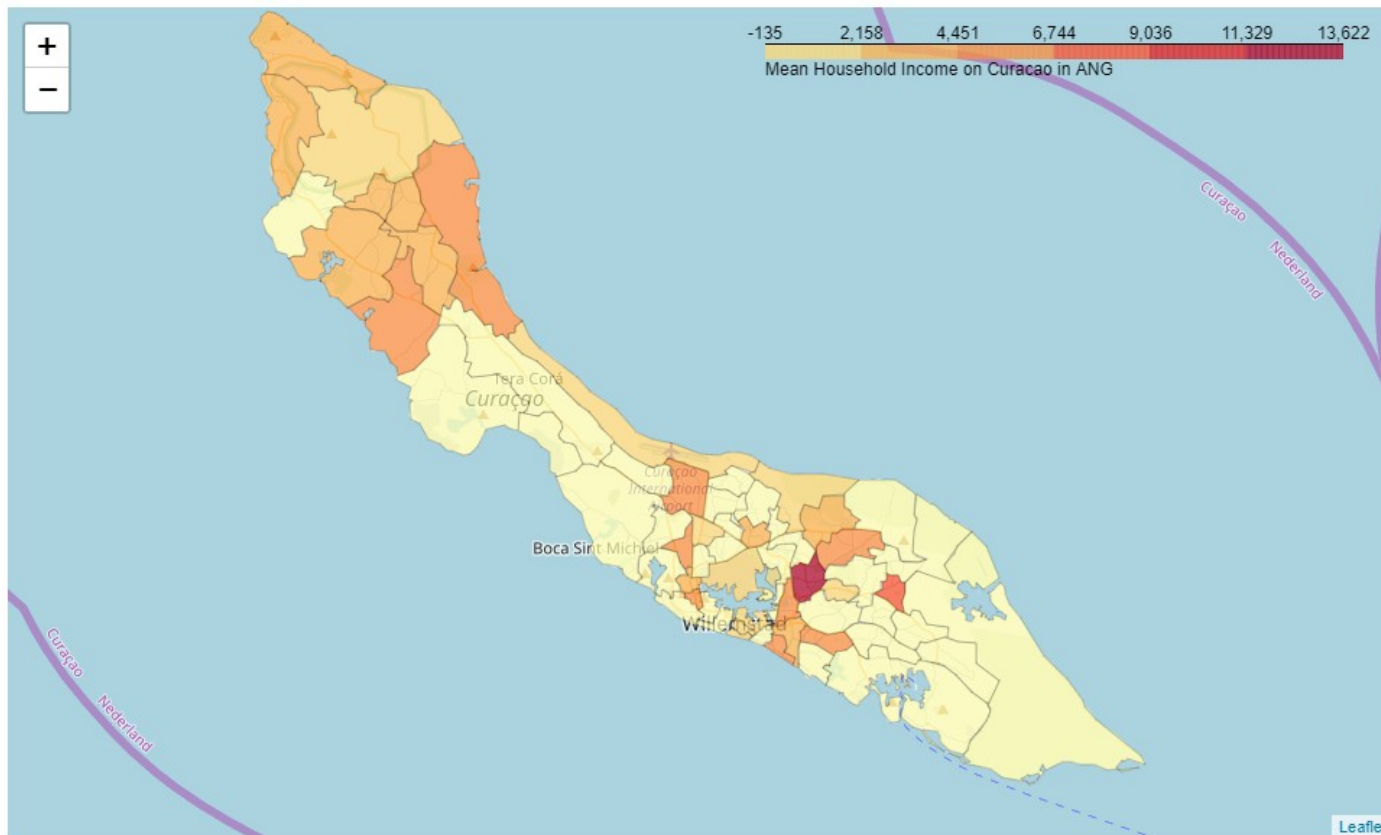
E. Neighborhood with max mean household income: ‘Mahaai’
with 12.358 ANG

F. Neighborhood with max % foreign born population: ‘Mahaai’
with nearly 49%

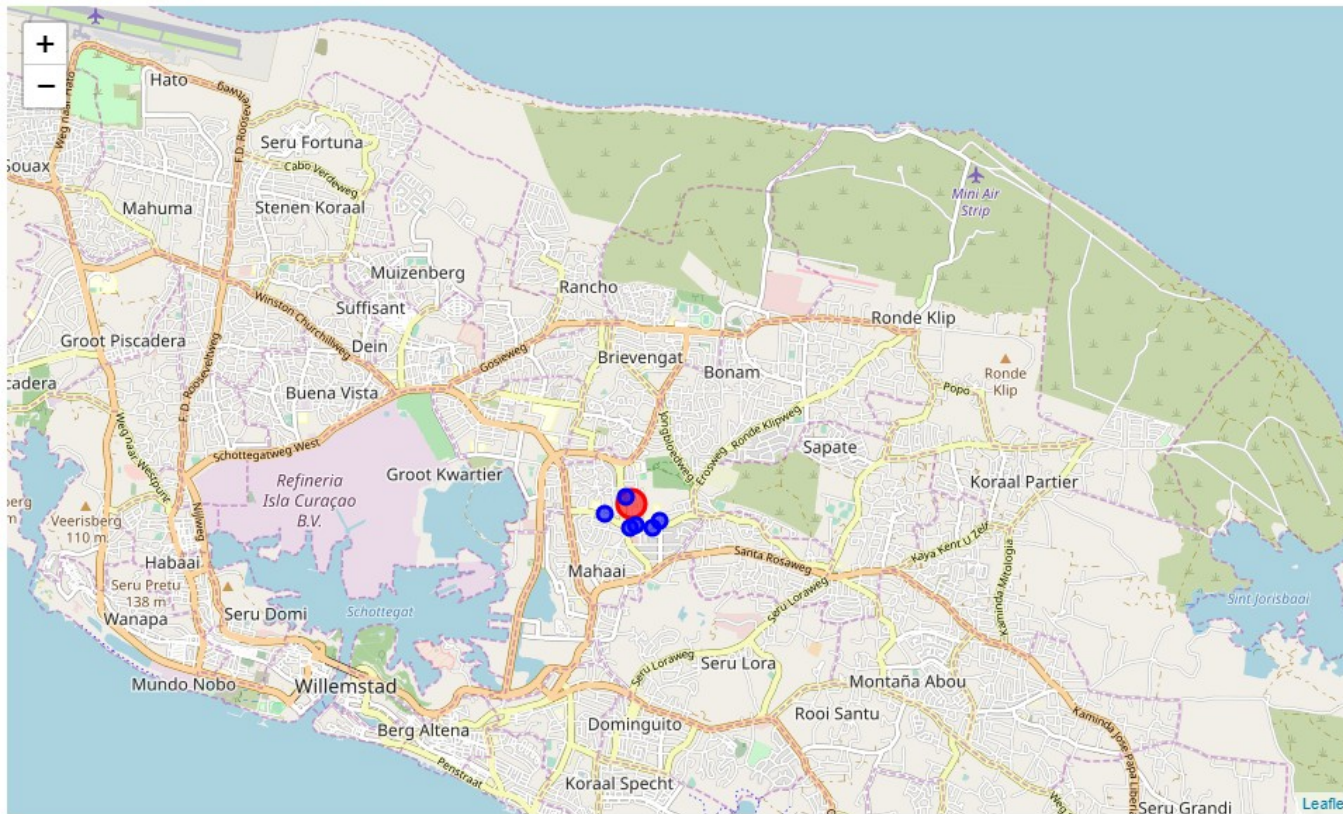
	NAME	HH	pop	hhsz	sexratio	p0_14	p15_64	p65u	curborn	forborn
237	Mahaai	152	389	2.6	80	0.211	0.612	0.177	0.513	0.487
251	Saliña	151	361	2.4	88	0.186	0.632	0.183	0.627	0.373
109	St. willibrordus	240	588	2.5	85	0.192	0.646	0.162	0.630	0.370
174	Zeelandia	88	135	1.5	39	0.067	0.378	0.556	0.630	0.370
171	Wishi	328	841	2.6	75	0.205	0.669	0.126	0.695	0.305

Results

Best location: 'Mahaai' in dark red color



No nearby Italian restaurants



Discussion

Missing values in dataset

→ not all neighborhoods match geographic outlay

→ suboptimal result

Conclusion

Objective: Best location to open a new Italian restaurant on Curacao

Neighborhood 'Mahaai'

→ Max avg income

→ Max foreigners