

# **Where to open a coffee shop?**



Thomas Kohlborn  
21/03/2021

# Introduction

- **How can someone make a decision on where to open up a new venue?**
  - Before investing time and money into a new venture, one should ensure to understand what drives success?
  - Opening a coffee shop in Vancouver, Canada has always been my personal dream, so this is what this project will focus on.

# Data acquisition and cleaning

- **Understand current market**
  - Foursquare data to understand how many coffee shops are already established in each suburb
  - Census data to understand the demographic details of each of the suburbs that could influence the number of coffee shops in each suburb
    - Census data is publicly and freely available here:

<https://opendata.vancouver.ca/explore/dataset/census-local-area-profiles-2016/information/>

# Data acquisition and cleaning (cont.)

- **Subset of data**

- As the census data contains more than 5000 different variables, we will focus on the ones that could impact on the number of coffee shops in each suburb by common sense.
  - number of people living in each suburb ('population', the average age of the respective population in each suburb ('avg\_age'), as well as median age('med\_age'), average household size ('avg\_household\_size'), average income ('avg\_income'), median income ('med\_income'), employment rate ('employment\_rate') and the number of people commuting to other suburbs ('commuters')

# Data analysis

- **Gaining geo-spatial information**
  - Utilised GeoPy to get longitude and latitude for each suburb

	suburbs	latitude	longitude
0	Arbutus-Ridge	49.246	-123.160
1	Downtown	34.043	-118.248
2	Dunbar-Southlands	49.238	-123.184
3	Fairview	40.633	-90.164
4	Grandview-Woodland	49.276	-123.067
5	Hastings-Sunrise	49.279	-123.040
6	Kensington-Cedar Cottage	49.247	-123.073
7	Kerrisdale	49.221	-123.160
8	Killarney	52.060	-9.505
9	Kitsilano	49.269	-123.155
10	Marpole	49.209	-123.136

# Data analysis (continued)

- **Venues in the suburb**

- Over 419 venues in all of Vancouver's suburbs
- Users have provided information on 21 coffee shops with most coffee shops being in 'Grandview-Woodland' and 'South Cambie' (extract below)

	suburbs	Coffee Shop
0	Arbutus-Ridge	1
1	Downtown	3
2	Dunbar-Southlands	1
3	Fairview	0
4	Grandview-Woodland	4

# Data analysis (continued)

	Coffee Shop	population	avg_age	med_age	avg_household_size	avg_income	med_income	employment_rate	unemployment_rate	commuters
Coffee Shop	1.000	0.063	-0.385	-0.489	-0.383	0.041	0.365	0.454	-0.163	-0.038
population	0.063	1.000	-0.552	-0.689	-0.144	-0.414	-0.020	0.604	-0.370	0.951
avg_age	-0.385	-0.552	1.000	0.876	-0.016	0.067	-0.351	-0.753	0.572	-0.554
med_age	-0.489	-0.689	0.876	1.000	0.294	0.208	-0.370	-0.931	0.617	-0.651
avg_household_size	-0.383	-0.144	-0.016	0.294	1.000	-0.008	-0.330	-0.411	-0.026	0.028
avg_income	0.041	-0.414	0.067	0.208	-0.008	1.000	0.752	-0.105	-0.187	-0.523
med_income	0.365	-0.020	-0.351	-0.370	-0.330	0.752	1.000	0.494	-0.520	-0.146
employment_rate	0.454	0.604	-0.753	-0.931	-0.411	-0.105	0.494	1.000	-0.678	0.575
unemployment_rate	-0.163	-0.370	0.572	0.617	-0.026	-0.187	-0.520	-0.678	1.000	-0.378
commuters	-0.038	0.951	-0.554	-0.651	0.028	-0.523	-0.146	0.575	-0.378	1.000

- **Correlations between variables**

- Number of coffee shops negatively correlated with age, household size and unemployment rate
- Is it by chance?

# Data analysis (continued)

	Coffee Shop	population	avg_age	med_age	avg_household_size	avg_income	med_income	employment_rate	unemployment_rate	commuters
Coffee Shop	1.000	0.787	0.085	0.025	0.086	0.859	0.104	0.038	0.481	0.869
population	0.787	1.000	0.010	0.001	0.535	0.062	0.931	0.004	0.098	0.000
avg_age	0.085	0.010	1.000	0.000	0.947	0.773	0.119	0.000	0.007	0.009
med_age	0.025	0.001	0.000	1.000	0.196	0.365	0.099	0.000	0.003	0.001
avg_household_size	0.086	0.535	0.947	0.196	1.000	0.973	0.143	0.064	0.910	0.903
avg_income	0.859	0.062	0.773	0.365	0.973	1.000	0.000	0.650	0.416	0.015
med_income	0.104	0.931	0.119	0.099	0.143	0.000	1.000	0.023	0.016	0.529
employment_rate	0.038	0.004	0.000	0.000	0.064	0.650	0.023	1.000	0.001	0.006
unemployment_rate	0.481	0.098	0.007	0.003	0.910	0.416	0.016	0.001	1.000	0.091
commuters	0.869	0.000	0.009	0.001	0.903	0.015	0.529	0.006	0.091	1.000

- **p-values between variables**
  - Statistically significant relationship of  $p \leq .05$  between 'coffee shop' and 'median age' as well as 'employment rate'



# Results and discussion

- **Practical significance**

- The younger the median age of a suburb and the higher the employment rate, the more coffee shops would you expect to find in a suburb
- ‘Mount Pleasant’ has youngest median age and less coffee shops than other suburbs in top4

	suburbs	Coffee Shop	population	avg_age	med_age	avg_household_size	avg_income	med_income	employment_rate	unemployment_rate	commuters
11	Mount Pleasant	1	32955	38.300	35.500	1.800	54260	42362	77.900	4.700	4750
1	Downtown	3	62030	40.600	37.100	1.700	63251	41858	68.800	5.600	7010
9	Kitsilano	2	43045	40.600	37.700	1.900	63092	44084	71.100	5.200	5665
4	Grandview-Woodland	4	29175	40.200	38.100	1.900	42896	32438	69.900	5.300	4085
18	Sunset	0	36500	39.800	38.700	3.100	34212	25498	62.300	5.200	5635

# Conclusion and future direction

- 'Mount Pleasant' shortlisted as the suburb to be investigated further re suitability of opening a new coffee shop
- Other factors, such as cost, available properties etc. will play a role in the final decision
- Methodology applicable to other locations and venues