Final Proposal (Group 14) - Advanced Track

Anishka Fernandopulle, Radhika Iyer, Sidharth Mirchandani, Hajra Ojha

Research Question

How does the density of free and low-cost food programs influence property and violent crime density in low-income communities in Vancouver?

Variables

Canadian Census, 2021

We will be using the 2021 Canadian census data-set to obtain the Population Density, Household Size, and Low-Income Density of neighborhoods in Vancouver. Then, we aim to establish a relationship between these factors and the crime rate within those specific communities. We will be considering the following variables:

- Population, 2021: v_CA21_1
  - This will provide us with a sample population to aid us in answering our research question. This numeric variable gives us quantitative values for the number of people within our sample along with population data specific to our chosen year. This enables us to make clearer comparisons.
- Population density per square kilometer: v_CA21_6
  - This will allow us to separate the population into specific locations to see the density within each neighborhood and provide context behind the crime rates in the area. It will also help us specifically pinpoint the locations of low-cost food programs. Based on population density some areas may experience different social dynamics that could influence crime. The unit for this variable represents the number of people.
- Average household size: v_CA21_452
  - This variable consists of quantitative data about the household composition which may influence the crime rate. This can also provide a brief estimate of income brackets within the neighborhood to pinpoint which locations consist of low-income earners who might need low-income food plans. Larger households might have different needs that may impact their food security as well as the crime rate.
- Prevalence of low income based on the Low-income measure, after tax (LIM-AT)(%): v_CA21_1040
  - This variable provides an indication of household financial vulnerability that helps examine a relationship between low income, access to food, and the crime rate. The data is represented in the form of percentages in a given area and is a relative measure in relation to median income.
- Prevalence of low income based on the Low-income based on the Low-income cut-offs, after tax (LICO-AT)(%): v_CA21_1085
  - This variable is a key indicator of financial strain that highlights the need for food programs as it takes into account living costs. It allows us to target areas where food insecurity may be high and food programs would be most likely needed. This is an absolute measure where the units are given in the form of percentages.
- Income statistics for private households: v_CA21_905

  - This variable will help identify low-income households to better understand the relationship between crime rates and low-income households. It provides information about income/earning levels.

Free and Low Cost Food Programs, Vancouver

From the Free and Low-Cost Food Programs data-set, we will be considering the following variables:

- Neighborhood the facility is located: local_areas
  - This displays the geographical distribution of food programs that can be geographically matched with crime rate locations to study the relationship between the two.
- Latitude and Longitude of facility: geom
  - This provides a more specific location of each neighborhood we are studying, increasing the accuracy of our findings. This provides more detailed data to analyze the proximity to crime incidents. We can analyze the proximity of crime occurrence to food program locations. It is given in terms of coordinates.
- Is the facility operational: program_status
  - This determines the availability of food resources to understand if the issue of food security is dealt with. This allows us to isolate the effect. Within this variable data is either 'open' or 'closed' allowing us to disregard 'closed' food programs as they do not contribute to solving food insecurity.

The data-set contains the location of free/low-cost food programs, food banks, grocery hampers, low-cost groceries, or grocery vouchers in Vancouver. These are valuable specifications of the low-food programs to better understand the reason for their location.

Vancouver Police Department Crime Data, 2021

For the VPD Crime data-set, we intend to calculate the total number of crimes committed in 2021. The location specifications can be matched with food program locations to better study their relationship through spatial analysis. We will be transforming the location data stored in the X and Y columns to match Census Dissemination Areas. To do this, we will be converting from the projection format of the crime data, to that of the census. We will then be calculating the density of crime within each dissemination area. By converting the location data to be projected the same way, we can compare across all three of our data-sets.

Summary Statistics

First we look at the summary statistics for our selected census variables:

```
kable(combo_statistics_census)
```

variable	mean	sd	max	min
v_CA21_1	226.62111	377.064103	8739.0	0.0
v_CA21_6	10590.88918	10733.159054	76474.4	0.0
v_CA21_449	636.31779	385.630529	8610.0	0.0
v_CA21_1040	13.28664	6.852265	61.0	2.4
v_CA21_1085	10.15816	4.968801	39.0	1.8
v_CA21_905	302.07648	238.095380	5060.0	75.0

Here we find the summary statistics for free and low-cost food centers:

```
kable(summary_statistics_food)
```

local_areas	mean	sd	max	min
Arbutus-Ridge	23.37398953 [1/m^2]	NA	23.37399 [1/m^2]	23.37399 [1/m^2]
Downtown	387.80580665 [1/m^2]	NA	387.80581 [1/m^2]	387.80581 [1/m^2]
Fairview	35.99613389 [1/m^2]	NA	35.99613 [1/m^2]	35.99613 [1/m^2]
Grandview-Woodland	141.97823835 [1/m^2]	NA	141.97824 [1/m^2]	141.97824 [1/m^2]
Hastings-Sunrise	55.34683860 [1/m^2]	NA	55.34684 [1/m^2]	55.34684 [1/m^2]
Kensington-Cedar Cottage	15.06040665 [1/m^2]	NA	15.06041 [1/m^2]	15.06041 [1/m^2]
Killarney	16.40175392 [1/m^2]	NA	16.40175 [1/m^2]	16.40175 [1/m^2]
Kitsilano	39.68890214 [1/m^2]	NA	39.68890 [1/m^2]	39.68890 [1/m^2]
Marpole	53.49618783 [1/m^2]	NA	53.49619 [1/m^2]	53.49619 [1/m^2]
Mount Pleasant	156.33418528 [1/m^2]	NA	156.33419 [1/m^2]	156.33419 [1/m^2]
Renfrew-Collingwood	28.50827662 [1/m^2]	NA	28.50828 [1/m^2]	28.50828 [1/m^2]
Riley Park	69.34252772 [1/m^2]	NA	69.34253 [1/m^2]	69.34253 [1/m^2]
Strathcona	217.10445409 [1/m^2]	NA	217.10445 [1/m^2]	217.10445 [1/m^2]
Sunset	14.94567027 [1/m^2]	NA	14.94567 [1/m^2]	14.94567 [1/m^2]
West End	13.88043555 [1/m^2]	NA	13.88044 [1/m^2]	13.88044 [1/m^2]
NA	0.03454083 [1/m^2]	1.113373	35.88792 [1/m^2]	0.00000 [1/m^2]

Finally, we have the summary statistics for our crime data-set:

```
kable(summary_statistics_crime)
```

neighbourhood	mean	sd	max	min
Arbutus Ridge	2629.61355 [1/m^2]	1596.30081	6451.22111 [1/m^2]	885.41254 [1/m^2]
Central Business District	130859.30386 [1/m^2]	123153.01540	981379.66507 [1/m^2]	1985.13828 [1/m^2]
Dunbar-Southlands	2189.40830 [1/m^2]	600.60127	3490.83516 [1/m^2]	47.74882 [1/m^2]
Fairview	20087.50380 [1/m^2]	13209.53731	93619.24353 [1/m^2]	4965.83133 [1/m^2]
Grandview-Woodland	19635.65553 [1/m^2]	13192.81611	69516.75117 [1/m^2]	1922.35121 [1/m^2]
Hastings-Sunrise	9077.31763 [1/m^2]	4329.75038	23639.65373 [1/m^2]	749.62766 [1/m^2]
Kensington-Cedar Cottage	16421.47205 [1/m^2]	5611.20020	41113.26573 [1/m^2]	4674.15799 [1/m^2]
Kerrisdale	2739.63742 [1/m^2]	1895.82687	9764.81469 [1/m^2]	257.99243 [1/m^2]
Killarney	3548.33630 [1/m^2]	2336.18848	13318.17239 [1/m^2]	470.92576 [1/m^2]
Kitsilano	21989.45448 [1/m^2]	13488.93363	80249.87806 [1/m^2]	4155.82560 [1/m^2]
Marpole	7522.41551 [1/m^2]	7051.19006	31559.39644 [1/m^2]	970.44255 [1/m^2]
Mount Pleasant	18905.97671 [1/m^2]	19634.36728	147527.98678 [1/m^2]	2307.49326 [1/m^2]
Musqueam	36.40338 [1/m^2]	25.80688	91.11212 [1/m^2]	19.48675 [1/m^2]
Oakridge	2310.80221 [1/m^2]	1341.49452	6925.25213 [1/m^2]	463.90297 [1/m^2]
Renfrew-Collingwood	16707.16813 [1/m^2]	18579.98242	126457.64301 [1/m^2]	3474.12683 [1/m^2]
Riley Park	7743.74079 [1/m^2]	3443.82556	23622.79162 [1/m^2]	991.61248 [1/m^2]
Shaughnessy	1343.91150 [1/m^2]	749.99081	4948.88938 [1/m^2]	669.65453 [1/m^2]
South Cambie	2221.17084 [1/m^2]	1393.19435	7661.00785 [1/m^2]	746.50013 [1/m^2]
Stanley Park	95.07740 [1/m^2]	708.86271	7730.04128 [1/m^2]	29.82130 [1/m^2]
Strathcona	30952.96158 [1/m^2]	26394.07452	122814.99322 [1/m^2]	2010.60983 [1/m^2]
Sunset	7840.52540 [1/m^2]	5372.17610	24069.12792 [1/m^2]	980.68169 [1/m^2]

neighbourhood	mean	sd	max	min
Victoria-Fraserview	5602.27469 [1/m^2]	1734.00240	11203.08562 [1/m^2]	1643.26607 [1/m^2]
West End	68853.89417 [1/m^2]	44788.96130	261276.64617 [1/m^2]	581.00992 [1/m^2]
West Point Grey	2535.45036 [1/m^2]	1138.79369	8579.99241 [1/m^2]	476.15477 [1/m^2]
NA	694.03939 [1/m^2]	651.76741	2823.22457 [1/m^2]	14.94633 [1/m^2]

Model

To answer this research question, we will use a multiple linear regression model that explores the relationship between food program density and crime rates in low-income communities, with the equation:

$$(\text{Crime Density})_i = \beta_0 + \beta_1(\text{Food Program Density} \cdot \text{Low Income})_i + \beta_2(\text{Food Program Density})_i + \beta_3(\text{Population Density})_i + \beta_4$$

In this regression equation, the dependent variable Crime Density$_i$ represents the density of crimes per area for each census dissemination area. Our key parameter of interest is β_1 which is the interaction term that allows us to isolate the effect of food programs within low-income neighbourhoods only, by multiplying Food Program Density$_i$ which measures the density of free or low-cost food programs in the area by the dummy variable Low Income$_i$ which takes on the value 0 or 1 depending on whether we classify it as a low-income neighbourhood. The covariates are Food Program Density$_i$, Population Density$_i$, Household Size$_i$, and Low Income$_i$, which are variables that may influence crime in a neighbourhood on their own. Including these covariates allows us to isolate the effect of food program density on crime density in low income areas by controlling these variables. The parameter β_0 is the intercept, which represents the expected value of the crime density when the independent variables are equal to zero. The other parameters, β_i, where $i \neq 0$, represents the change in the crime density for a one-unit increase in each variable, holding the other variables constant. This model assumes that the factors influence crime density independently and that the relationships are linear.

Works Cited

1. "Census Data 2021.", CensusMapper, 2021,

   https://censusmapper.ca/api/CA21.

2. "Free and Low Cost Food Programs." Open Data Portal, City of Vancouver, 2021,

   https://opendata.vancouver.ca/explore/dataset/free-and-low-cost-food-programs/information/.

3. "GeoDASH Crime Map.",

   https://geodash.vpd.ca/opendata/.