# Section 1: Basic info

Project title: Crime in Vancouver
Team Info: Bolin Wang (i3l2b), Michaux Sun (m5j2b), Kerry Zhou (q2k2b)

# Section 2: Overview

Safety is one of the most important factors to consider when moving into a new place. If we can understand the composition of crime in different areas, it is possible to respond to crime in a more targeted manner, thus enhancing security. To achieve this, we propose building a data visualisation that visually allows relevant government officials (i.e., police) and ordinary people to explore crime datasets. Our app will show the overall distribution of crime in Vancouver and enable users to analyse and compare the trend among several neighbourhoods by filtering the crime type, year, and neighbourhood, as well as daytime and nighttime.

# Section 3: Description of the data and data preprocessing

**Link:**
link for source from 2015 - 2020 + link for source in 2021 -> final dataset link
**Description**:
We will be visualising a dataset of **232439** crimes. Each crime has 18 attributes that include information about the crime id(CASE_ID), crime types (TYPE), the specific date and time the crime occurred (YEAR, MONTH, DAY, HOUR, MINUTE), name of the neighbourhood (NEIGHBOURHOOD) and their latitudes (X) and longitudes (Y). We also have aggregated attributes. The details are in the below, highlighted by [Aggregation].

| Attr. name | Attr. type | Cardinality for categorical attr. | Range for quantitative attr. |
|---|---|---|---|
| CASE_ID | categorical | 232439 | |
| TYPE | categorical | 11 | |
| YEAR | categorical | 7 | |
| MONTH | categorical | 12 | |
| DAY | categorical | 31 | |
| HOUR | categorical | 24 | |
| MINUTE | categorical | 60 | |
| NEIGHBOURHOOD | categorical | 24 | |
| X (latitudes) | quantitative | | [0.0, 498302.0] |
| Y (longitudes) | quantitative | | [0.0, 5462299.731] |
| NumPerNeighbourhood_all_yrs | quantitative | | [147, 67975] |
| NumPerNeighbourhood_2015 | quantitative | | [28, 8091] |
| NumPerNeighbourhood_2016 | quantitative | | [24, 9855] |
| NumPerNeighbourhood_2017 | quantitative | | [16, 9962] |
| NumPerNeighbourhood_2018 | quantitative | | [17, 10852] |
| NumPerNeighbourhood_2019 | quantitative | | [12, 12368] |

| | | | |
|---|---|---|---|
| NumPerNeighbourhood_2020 | quantitative | | [23, 7707] |
| NumPerNeighbourhood_2021 | quantitative | | [27, 9140] |

**Preprocessing**:

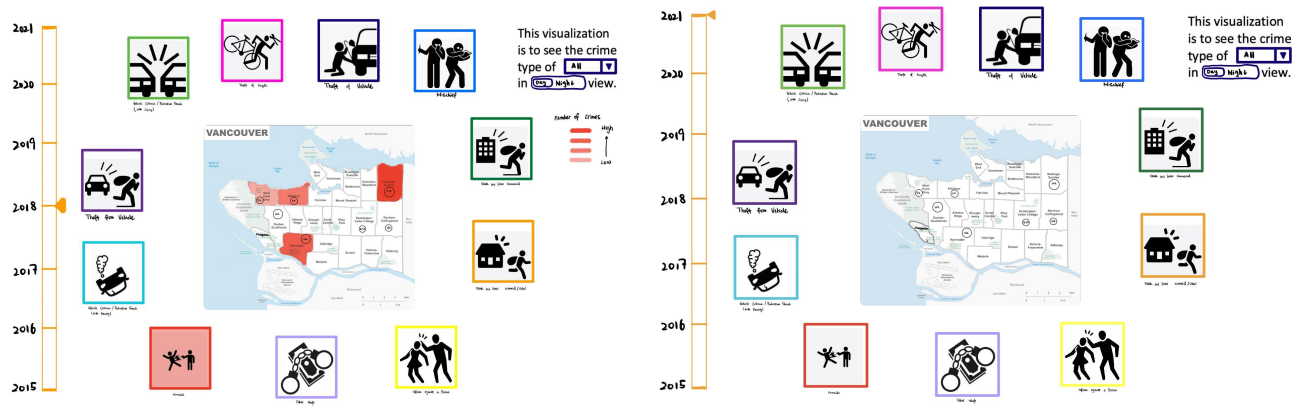
1. We noticed some missing values (NaN) in the column NEIGHBOURHOOD. We removed these crimes as their information is incomplete.
2. We extracted partial data (2015 - 2020) in the first source to shrink the number of points displayed in the chart.
3. In the first source, the data for 2021 is smaller than other years. So, we found another dataset for 2021. We ended up preprocessing and combining these two datasets.
4. We ignored the attribute HUNDRED_BLOCK as it is not helpful in our project. The HUNDRED_BLOCK is the street where the crime happened.
5. [**Aggregation: NumPerNeighbourhood_all_yrs**] A grouping by neighbourhood across 7 years will be generated to calculate the total number of crimes within each neighbourhood.
6. [**Aggregation: NumPerNeighbourhood_[2015, 2021]**] A grouping by neighbourhood and a certain year will be generated to calculate the total number of crimes within each neighbourhood and in that particular year.

# Section 4: Usage scenarios & tasks

Vancouver is a culturally diverse city and one of the most liveable cities globally, making it a great place to immigrate. As a Ukraine refugee, Maria wants to settle down in Vancouver, so she wants to [explore] numbers of crimes to [compare] the security level among neighbourhoods. She could also [identify] the most frequent crime types in a specific neighbourhood.
When she goes to our application, she will see the total number of all crime types, either in daytime or nighttime(by toggling the button), for each neighbourhood from 2015 to 2021.
Maria can't decide on Hasting and Point Grey. She first zooms in Hasting, sees all the cases in 2021 and checks Point Grey with the same steps. She then zooms out and clicks on these two regions. She is particularly aware of "homicide" and selects this crime type. The line chart below shows the total number of homicide cases in 2021. There are two lines in the chart, and each represents a neighbourhood. She looks at the line charts and compares the trends. She finds out that Point Grey has the lowest number of crimes, which goes down over the years. Therefore, she hypothesised that Point Grey is safe compared to Hasting and started looking for the rents in Point Grey.

# Section 5: Description of your visualisation & sketch

**Sketch 1:**




**Marks**: point, interlocking area
**Channels**: spatial region, colour hue, colour saturation
**Description**:
<u>M & C:</u>
Each point(icon) represents a crime type.
Each interlocking area represents a neighbourhood.

Interaction:

This is the zoom-out view of the choropleth map. In the centre is the map of Vancouver, in which each block is a neighbourhood. The right-hand-side text legend is a summary of the selection in the map.

**Icon filter:**

If no crime type is selected, the number in each block represents the sum of all types of crime.

A user can select a crime type by clicking on the corresponding icon (there are 11 of them circled the choropleth map), the chosen icon will be highlighted with its border colour, each neighbourhood will be colored with different saturation of the crime's colour, depends on their number of cases (less saturation means the small number of cases, more saturation means a large number of cases). The label that indicates crime types in the right-hand side legend will switch to the chosen crime name.

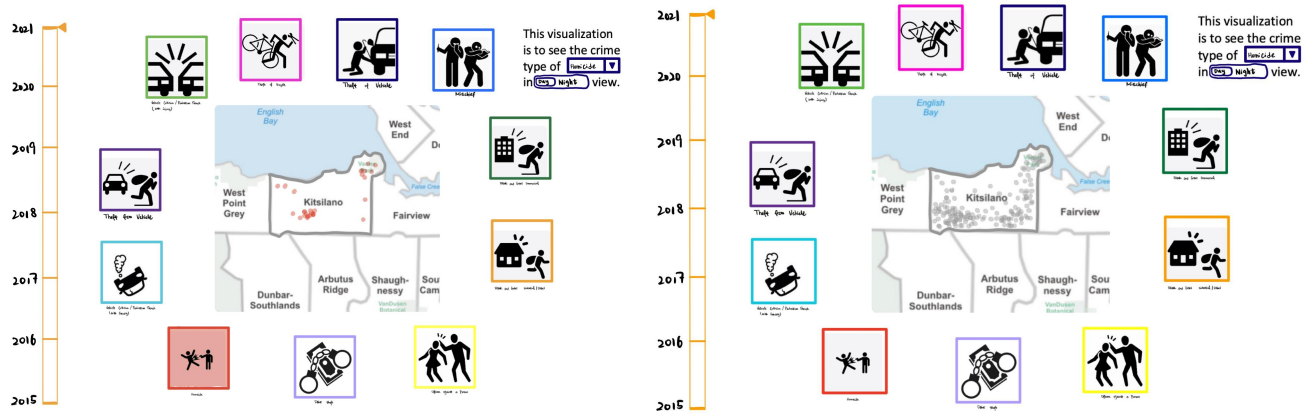
**Year bar:**

A user can scroll on the year bar to filter the data in a specific year. Let's say the user scroll to the year 2018, the number of cases shown in each neighbourhood will change to 2018, and so does the saturation.

**Day / Night switch:**

If a user switches to the Day toggle, the screen will show the UI with a set of bright palettes, and the data will be filtered to only show the crime that happens in the daytime. Similarly, a user can switch to night mode to see the crimes happening in the night time. The UI will also change to the dark background.

**Sketch 2:**




**Marks:** point, interlocking area

**Channels:** spatial position, colour hue, colour saturation

**Description:**

M & C:

The interlocking area encodes the neighbourhood that the user zooms in. Each point in the interlocking area encodes a crime case. The colour of the point is subjective to the icon's colour. The default colour is grey if no crime type was chosen. Each icon represents a type of crime, and each has a different colour.

Interaction:

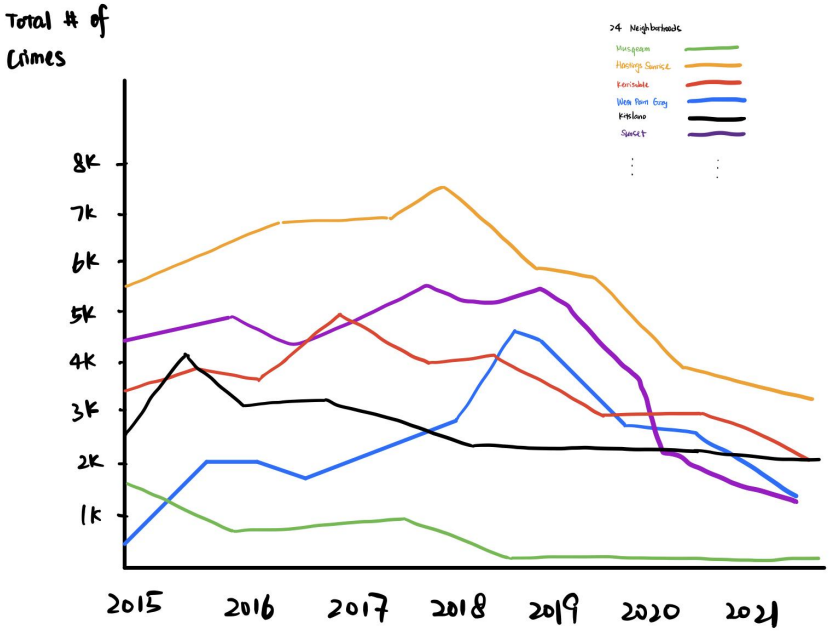
This is the zoom-in view for a specific neighbourhood (i.e., Kitsilano).

**Icon filter:**

If no crime type is selected, the chosen neighbourhood's boundary will be bolden. Each crime case is represented by a mark of point with colour grey (as shown in the right hand side of the sketch 2).

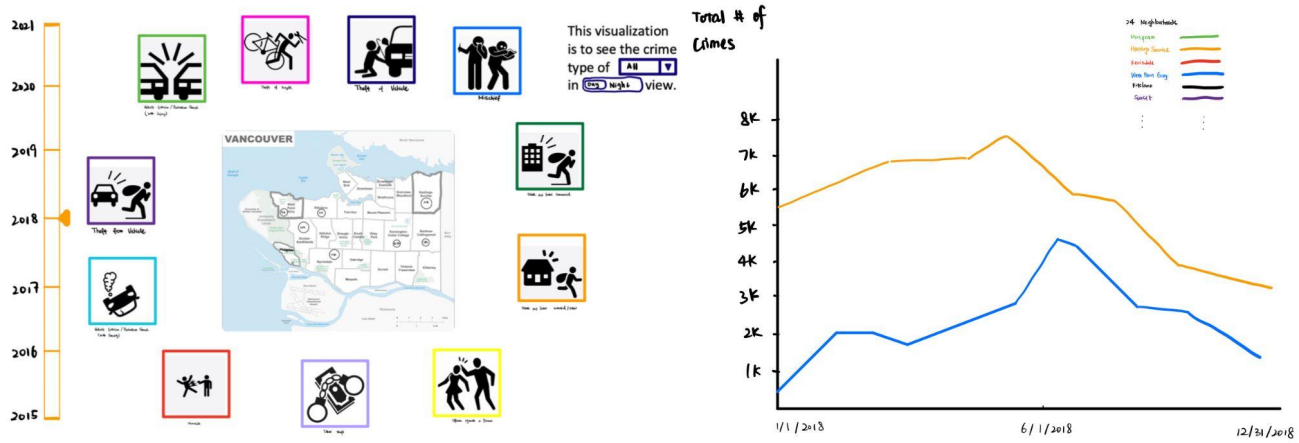
If a crime type is selected, the points will be highlighted by the icon colour (as shown in the left hand side of the sketch 2, points have colour red since "Homicide" was chosen), the points will only represent the selected crime type ("Homicide" in this case).

**Sketch 3:**




**Marks**: point
**Channels**: horizontal position, vertical position, colour hue
**Description**:
M & C:
Each point in the line represents a crime in this neighbourhood. The colours differentiate the neighbourhoods.
Interaction:
This line chart listens to the choropleth and the year bar.
The x-axis listens to the year bar. It shows the range of all the years by default.
The lines are the selected neighbourhoods from the choropleth. By default there will be 24 lines in the chart, in which each represents a neighbourhood.
Currently it shows the default view of the total number of crimes across all types from 01/01/2015 to 01/01/2021 of all 24 neighbourhoods. If a specific year 2015 is selected, the range of y-axis will be from 01/01/2015 to 12/31/2015.

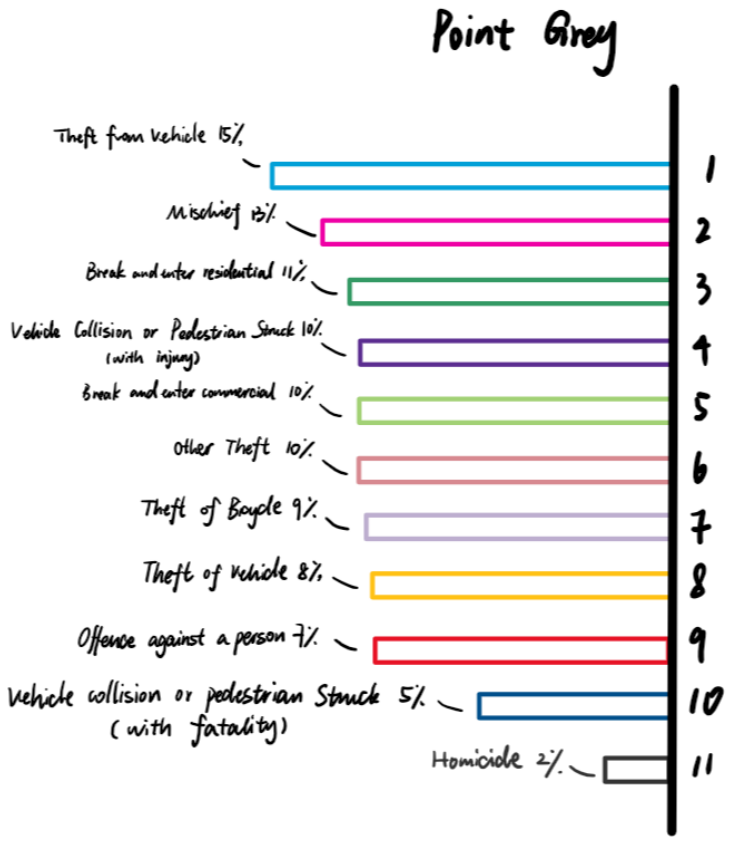
**Sketch 4: Selecting two neighbourhoods of 2018**




**Marks**: interlocking area, point
**Channels**:
Choropleth: colour hue, spatial region
Line chart: colour hue, spatial position
**Description**:

M & C:
Choropleth: each interlocking area represents a neighbourhood. The relative position of each interlocking area encodes the actual geographical position of neighbourhoods.
Line chart: each point represents a crime.
Interaction:
This sketch shows the changes on the line chart after a user chose a specific year from the year bar on the left-hand side of the choropleth map, and clicked to choose two neighbourhoods. Now that "West Point Grey" and "Hastings Sunrise" have been chosen, only two lines corresponding to the neighbourhoods get displayed on the line chart with the data from 2018. Similarly, a user can also click on different crime types on the choropleth map, and the line chart will display the total number of the chosen crime types.

**Sketch 5:**




**Marks**: line
**Channels**: length (1D size), vertical position, colour hue
**Description**:
M & C:
Each colour represents a type of crime.
The length of the line indicates the number of this type of crime in this neighbourhood.
The bars are displayed in descending order.
Interaction:
This bar chart always shows the details of the latest selected neighbourhood. If no neighbourhood is selected, the text "Please select a neighbourhood to see the detail" will be shown.

# Section 6: Work breakdown and schedule

| Task & Description | EST. comp. date | EST. time | asgmt. |
|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| Write up documentation (M1) | Design overall structure<br>Sketch charts<br>Breakdown work | Mar 11 | 6 hours for everyone | BW, MS, KZ |
| Set up project structure | Import dataset<br>Init project hierarchy and components' scripts<br>Init git branch | Mar 13 | 0.5 hours for everyone | BW, MS, KZ |
| Implement day & night toggle | Create initial static view of the day and night toggle, together with the text label | Mar 13 | 8.5 hour | BW |
| | Filter data by day time and night time<br>day time: (by default) 6 am - 6 pm<br>night time: 6 pm - 6 am (next day) | Mar 13 | 6.5 hour | BW |
| Implement year bar | Create initial static view of “2015 - 2021 vertical scrolling bar” | Mar 13 | 8.5 hour | MS |
| | Create initial static view of “all year” button | Mar 13 | 6 hour | MS |
| | Add interaction to the bar and button<br>i.e. when mouse click on, select the all years / a specific year | Mar 13 | 4.5 hour | MS |
| Implement icon filter (each icon represents a type of crime) | Create initial static view of the icon filter | Mar 18 | 4.5 hour | KZ |
| | Add interaction to the icons: when mouse click on, select the crime type | Mar 18 | 6 hour | KZ |
| Implement choropleth of zoom-out view | Create an initial static view<br>i.e. show the number of cases of all crime types for each neighbourhood, for all years / in a specific year | Mar 25 | 8 hour | BW |
| | Have the data items listen to the day & night toggle | Mar 25 | 4 hour | BW |
| | Have the data items listen to the year bar (all year and a specific year)<br>By default: all year | Mar 25 | 4 hour | MS |
| | Link colour hue of each region to the selected crime type | Mar 25 | 6 hour | MS |
| | Link saturation of each region to the number of cases of a specific crime type | Mar 25 | 6 hour | KZ |
| | Add interaction: when mouse hover on, show the tooltip of crime type introduction | Mar 25 | 6 hour | KZ |
| Implement choropleth of zoom-in view | Create an initial static view of the zoom-in view<br>i.e. each point represent a case | Mar 25 | 8 hour | KZ |
| | Link the view changes to the mouse event | Mar 25 | 4 hour | MS |
| | Have the cases related to this neighbourhood listen to the year bar (a specific year)<br>By default: 2021 | Mar 25 | 5 hour | BW |

| | | | | |
|---|---|---|---|---|
| | Have the cases related to this neighbourhood listen to the crime type filter | Mar 25 | 4 hour | BW |
| Implement line chart | Create an initial static view of the line chart | April 3 | 4.5 hour | MS |
| | Have x-axis listen to the the year bar (all year / a specific year), by default all year | April 3 | 4 hour | MS |
| | Have data items listen to the filtered crime type<br>By default: none | April 3 | 4 hour | KZ |
| | Have data items listen to the selected region(s)<br>By default: all the regions | April 3 | 4.5 hour | KZ |
| Implement bar chart | Create an initial static view of the bar chart | April 3 | 6.5 hour | BW |
| | Have data listen to the selected region | April 3 | 4 hour | BW |
| Add legend | Add legend / tips to the webpage | April 5 | 4 hour | MS |
| | Add text label for introduction in the header (optional) | April 5 | 4 hour | MS |
| Wrap up | Polish UI and harmonise colours | April 5 | 4 hour | BW |
| | Refactor and clean code | April 5 | 4 hour | KZ |
| Total: Approximate 53 hours/person | | | | |

# Appendix (subjective to change):

*Details for the views and elements*

**selectable / filterable data**:

1. year (horizontal scrolling bar with a widget)
2. neighbourhood (selected by clicking on the choropleth map)
3. crime type (filtered by the segmented semicircle embraced the choropleth map)
4. Day / Night (filtered by toggle)

**views**:

1. choropleth map with 11 types of crime embraced
    a. zoom out: 24 neighbourhoods (located by geographical location)
        i. by selecting the crime type, the **saturation** of each circle indicates the number of this type of crime happens in all the regions
        ii. select multiple neighbourhoods: for comparison
        iii. listen to all year button
        iv. can select a specific year
    b. zoom in: all the crime in **one** specific neighbourhood in one year
2. line chart (x-axis: all the year, y-axis: number of cases of a crime / all types of crime
    a. By default, all the neighbourhoods will be displayed, therefore we see the trend of all crimes throughout the entire Vancouver area (24 lines).
    b. If we select two neighbourhoods in the map, we see the trends of those two neighbourhoods (i.e. two lines) and if we click one of the crime types, the line chart will only show the trend of that specific crime type.
3. bar chart (latest selected neighbourhood)

    a. show the percentage of all types of crime in a year

**widgets**:

1. vertical year scrolling bar
    a. default year will be set to “all year”
    b. user scroll along the vertical direction to select different years, the corresponding data for the selected year will reflect on the choropleth map
2. toggle switching between day and night
3. icons for different type of crime
4. tooltips when hovering in the neighbourhood