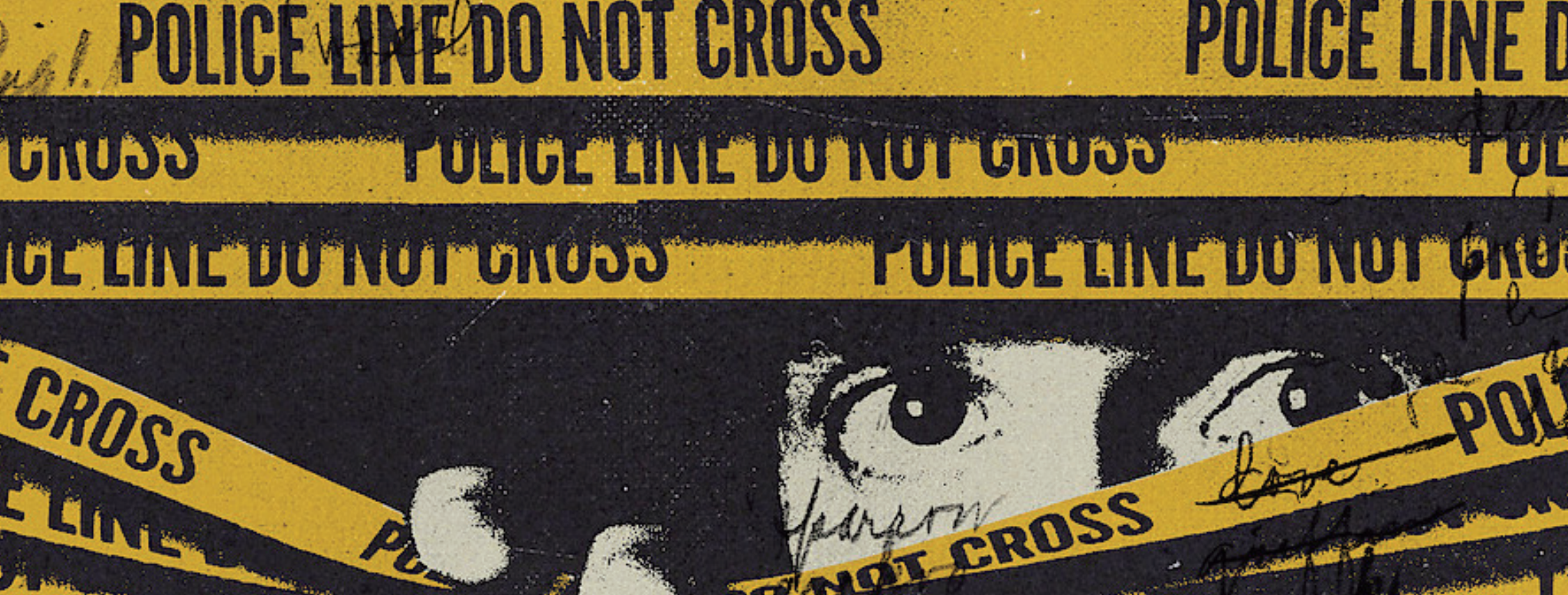


COMP7507 Project Final Report - GROUP 09

The Data Visualization Analysis of Hong Kong's 21st-Century Crime and Unnatural Death

CONTENT:

Shi Bofan
Sun Zhengxiao
Wei Lan
Zhang Kaiwen
Zhang Zhuan

A. Introduction

1. Project Title and Objectives:

Project Title: <The Data Visualization Analysis of Hong Kong's 21st-Century Crime and Unnatural Death Pattern>

Hong Kong is known to have a higher level of security in daily life, but crime never stops. Different from the countries that have legalized gun ownership, we rarely see malicious incidents reported in the news. Most of the crimes in Hong Kong happen quietly and unobtrusively. The project contains a series of data visualization analyses of Hong Kong's 21st-century crime, which study the sociological patterns of crime.

The project aims at digging into the social factors related to crime rate change in 21 century; identifying the sex and age ratio of offenders; distinguishing violent and non-violent crime types; and mapping the unnatural death on Hong Kong's map. Our group wants to know more about Hong Kong's overall performance in crime, the demographic characteristics of the society, and the differential between Hong Kong's 18 districts.

The visualization analysis consists of three chapters:

Chapter1: Hong Kong's crime rate changes in the 21st century and its relevant homochromous social factors are analyzed. A horizontal comparison between Hong Kong and other metropolitans is also included.

Chapter2: The offense types' category pattern and demographic characteristics of offenders

The sex and age ratio of offenders can be calculated and compared to Hong Kong's demographic census datasets. The dataset marks the offense types of the arrested person, which can develop the crime type preference of Hong Kong's population.

Chapter3: Haunted houses map with real estate development data

Unnatural death cases resulting from violent crime or suicide in residence are visualized on the map to reveal any potential spatial pattern. Real estate development data of Hong Kong's districts are collected for geographic pattern analysis.

The three chapters' scopes evolve from a worldwide scale to a Hong Kong scale to a district scale.

2. Project Highlights:

- Chapter1: In a worldwide comparison, HK's crime rate, and suicide rate are low. The overall trend of HK's crime rate is down, while there has been a pronounced fluctuation since 2008. The trends in crime and suicide rates are highly fitted in general. Whereas, HK's GNI data and police resource data show a relatively low correlation to crime rate trends.

- Chapter2: The total number of arrests shows a downward trend, while the change acts more obvious in the male group. Compared to census data, teenage and male groups are more inclined to commit crimes. The proportion of violent crimes accounted for about 22% of overall crimes.

- Chapter3: Kwun Tong, Sha Tin, and Yuan Long are the top three districts in Hong Kong, that have the most haunted houses that result from violent crime and suicide. However, when we take the population and overall housing supply into account, the proportion of haunted houses is a more accurate indicator showing the social stability of the districts. Yau Tsim Mong, Kwun Tong, and Southern are the most alarming districts.

3. Group Members' Contribution:

Group09	Student Number	Contribution
Shi Bofan		Data sourcing, cleaning, and visualization of chapter3; final report and demo of chapter3
Sun Zhengxiao		Data sourcing, cleaning, and visualization of chapter2; final report and demo of chapter2
Wei Lan		Data sourcing, cleaning, and visualization of chapter1; final report and demo of chapter1
Zhang Kaiwen		Data mining, cleaning, and visualization of chapter3; final report and demo of chapter3; submission combination
Zhang Zhuan		Data sourcing, cleaning, and visualization of chapter 1,2; final report and demo of overall structure and chapter2

B. Data Visualization Outcomes and The Stories Revealed

1. The Datasets Supports the Analysis:

1.1 Chapter1

- Crime and murder rates in Hong Kong[(1)]
- Comparison of crime rates with other countries(All)[(2)]
- Total number of crimes[(3)]
- World crime rate 2020[(4)]
- Suicide rate in Hong Kong[(5)]
- Overview of suicide rates from 1985 to 2016[(6)]
- Mental health and suicide rate[(7)]
- Human and financial resources of Hong Kong Police Force[(8)]
- Hong Kong's gross national income as economic development indicator[(9)]

1.2 Chapter2

- Persons arrested for crime by type of offence, age group and sex[(10)]
- Population by sex and age group 2011, 2016, 2022[(11)]

1.3 Chapter3

- Hong Kong Monthly Digest of Statistics from Census and Statistics Department of the Government of Hong Kong SAR[(12)]
- Haunted House Dataset[(13)]
- Hong Kong Government Census Dataset[(14)]

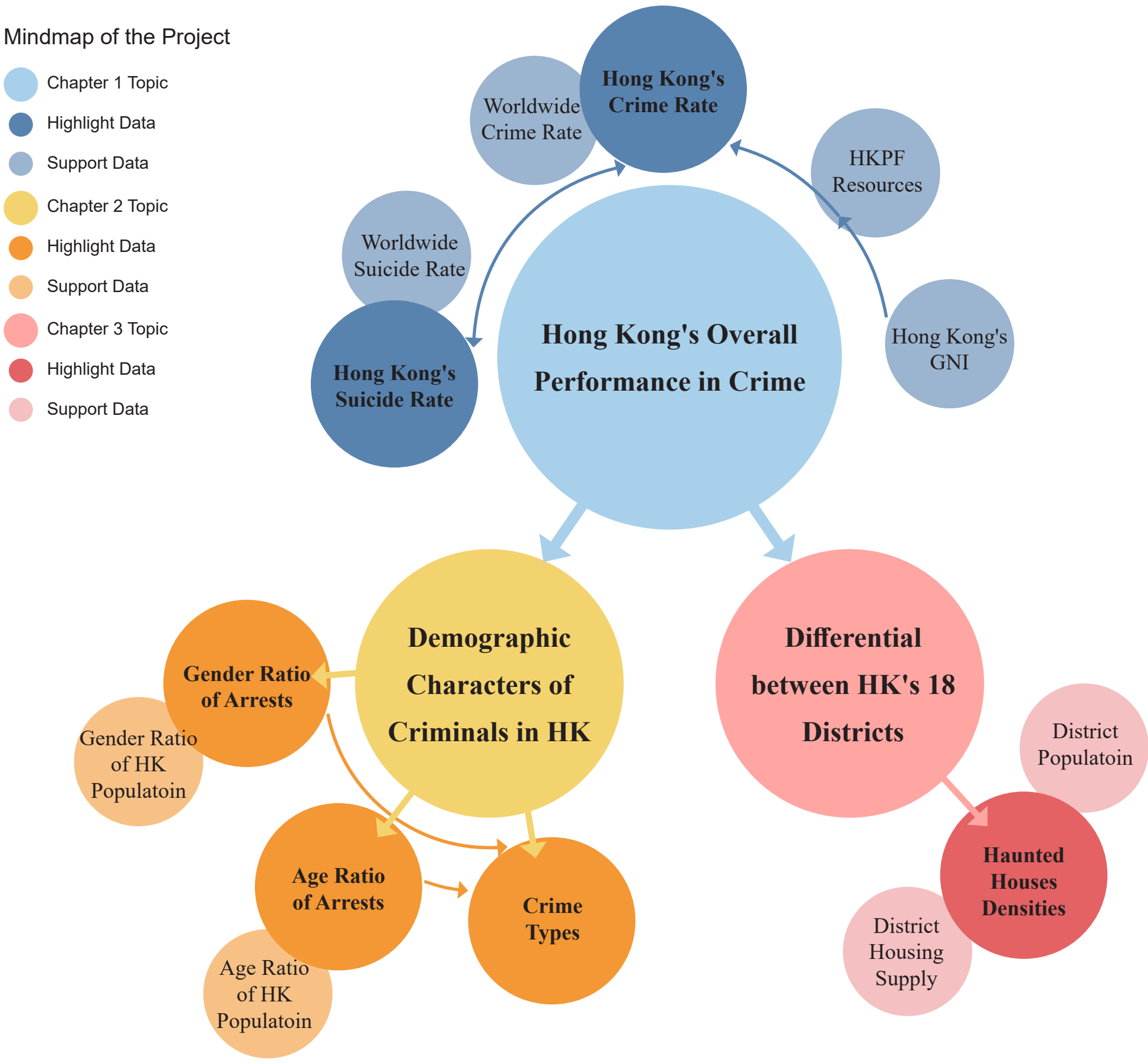
2. Overall Outcomes and Logic Structure of the Project:

The project starts with HK's 21st-century crime rate record. It is visualized in a chronological line chart and an annual percentage change bar chart. HK's crime rate is compared with worldwide data to determine Hong Kong's level. We also make a cross-sectional comparison of crime rate and suicide rate to validate the hypothesis: these two datasets should share similar trends, as they are both indicators of social-emotional stability. Then we want to know whether the crime trend has a relationship with HK's economic development and Hong Kong Police Force's resources.

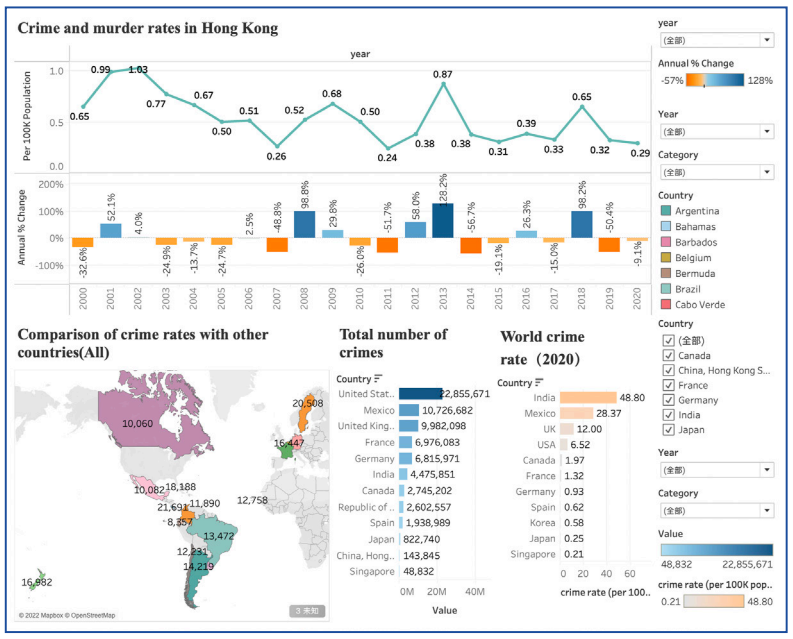
As chapter 1 has studied the crime rate on the scale of the country, we decide to explore the criminal population in Hong Kong. The keywords of chapter 2 are sex ratio, age ratio, and crime types. In chapter 2, the project focuses on the population pattern of arrested offenders to determine the tendency of the crime of different groups and count the percentage of violent and non-violent crime types.

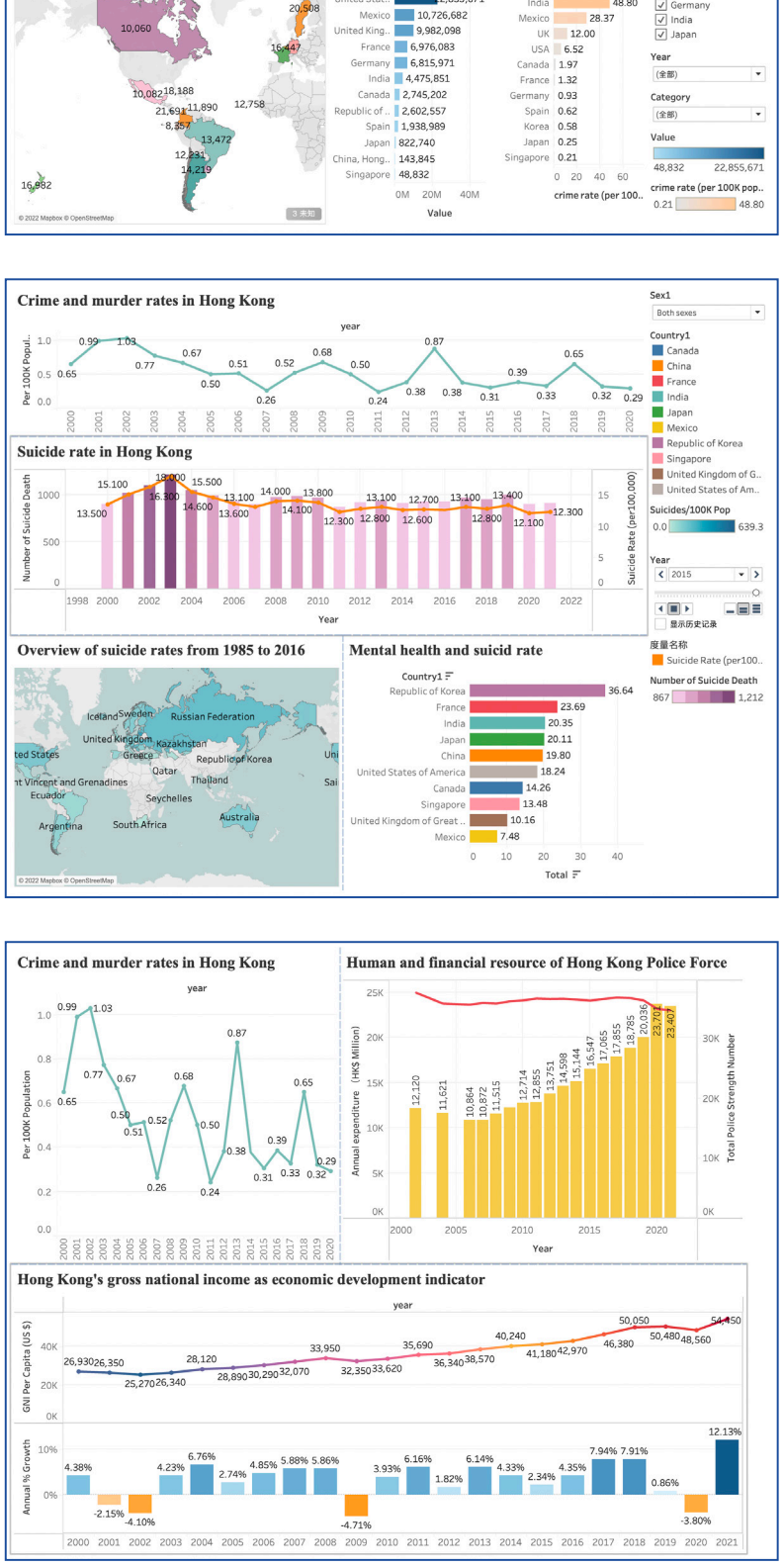
For chapter 3, we take a closer look at Hong Kong's 18 districts. The essential step is to do the data mining work for haunted houses' location data. Then the population data and housing supply data are compared to the number of haunted houses amount of each district. The frequency of unnatural deaths in each district is revealed

In the 7 dashboards of the project, multiple charts are applied according to the focusing aspect of the data. Line chart draws the tendency along time; bar chart identifies the data amount; pie chart reveals the percentage among data; density map shows each district's condition visually. A stacked bar chart and scatter plot can show multiple-layer data in one graph.

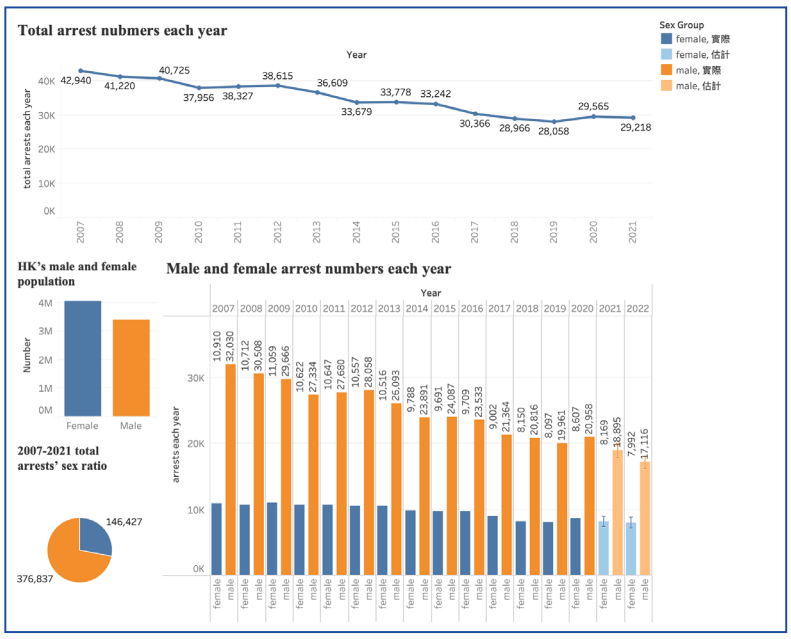


Chapter 1 - 3 dashboards 11 graphs

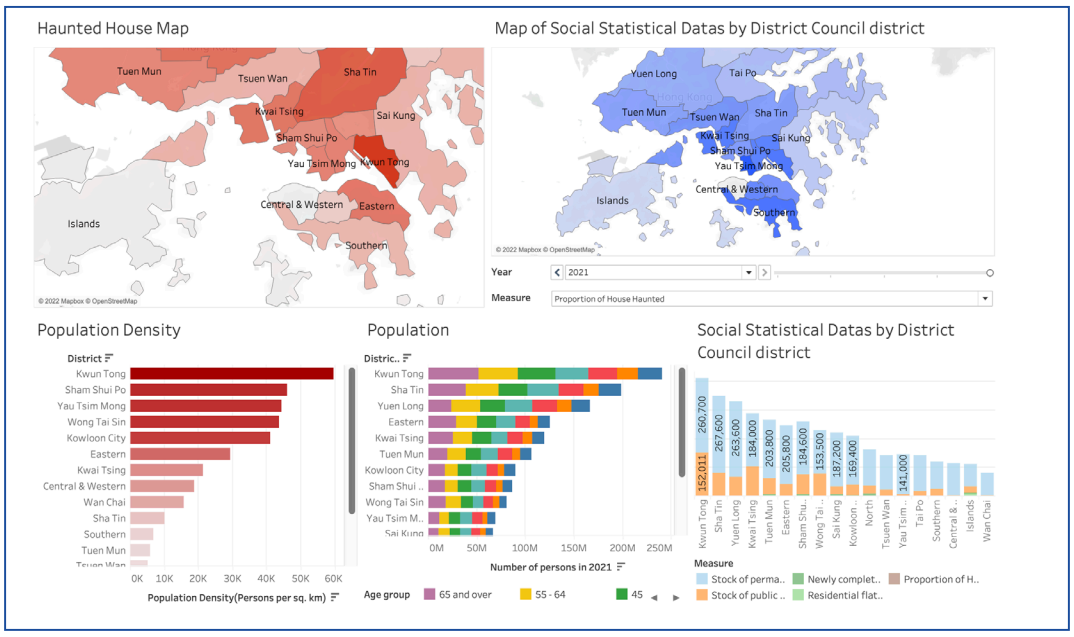




Chapter 2 - 3 dashboards 10 graphs



Chapter 3 - 1 dashboards 5 graphs



3. The Effectiveness of the Visualization and Methods Chosen:

CHAPTER 1 Hong Kong's crime rate changes in the 21st-century

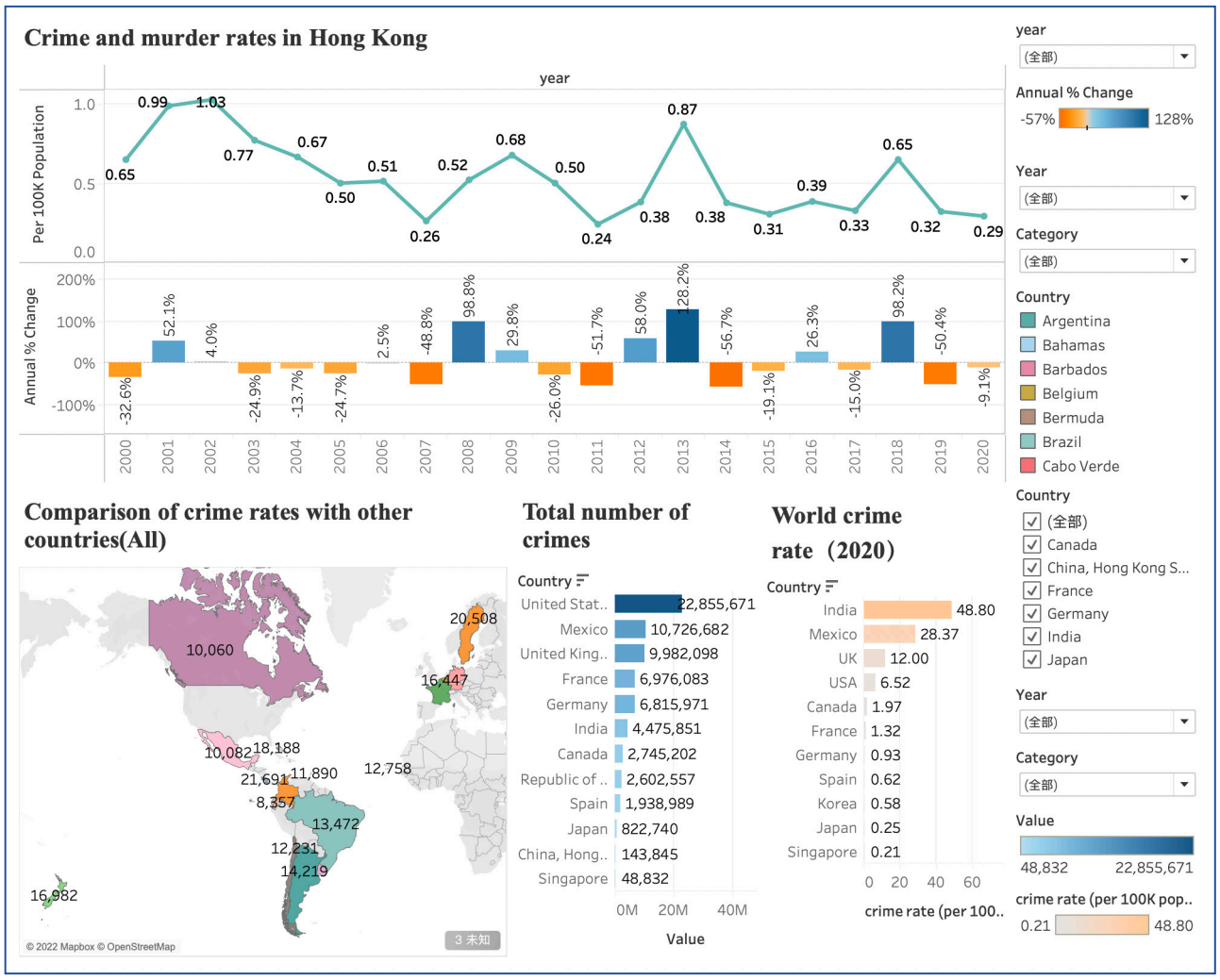
This project is based on the crime rate in Hong Kong in the 21st century. Aims to find out the factors that can reflect the local criminal characters and tendencies. The statistic of crime rate and unnatural death in the past 2 decades are visualized in Chapter 1.

Graph1: Crime and murder rates in Hong Kong[(1)]

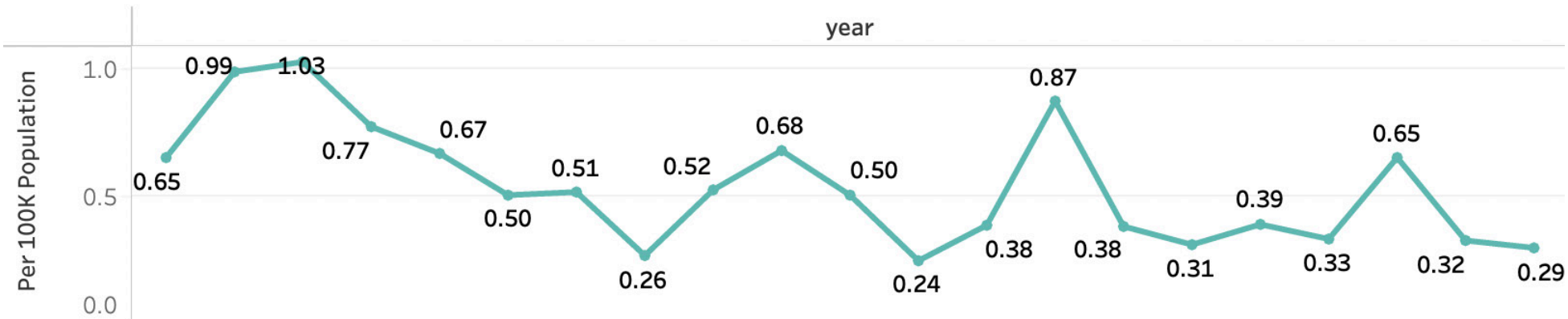
Graph2: Comparison of crime rates with other countries (All)[(2)]

Graph3: Total number of crimes[(3)]

Graph4: World crime rate (2020)[(4)]



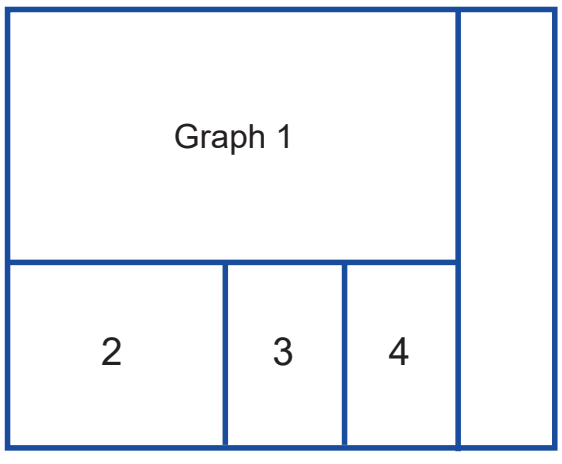
Chapter 1 Dashboard1



Dashboard 1 contains 4 main graphs, of which the frame structure is designed as shown in the frame diagram above, Graph1 shows the trend growth graph of the crime population base (number of crimes per 100K people) and growth rate in Hong Kong in the past 20 years, we can see that the crime population base declined more obviously during 2002-2007, probably because the administration and legal education work was popularized from 2008 to 2013, the trend of growth started again, reaching a maximum value of 0.87, and then gradually declined after 2013, remaining at a stable level of about 0.3.

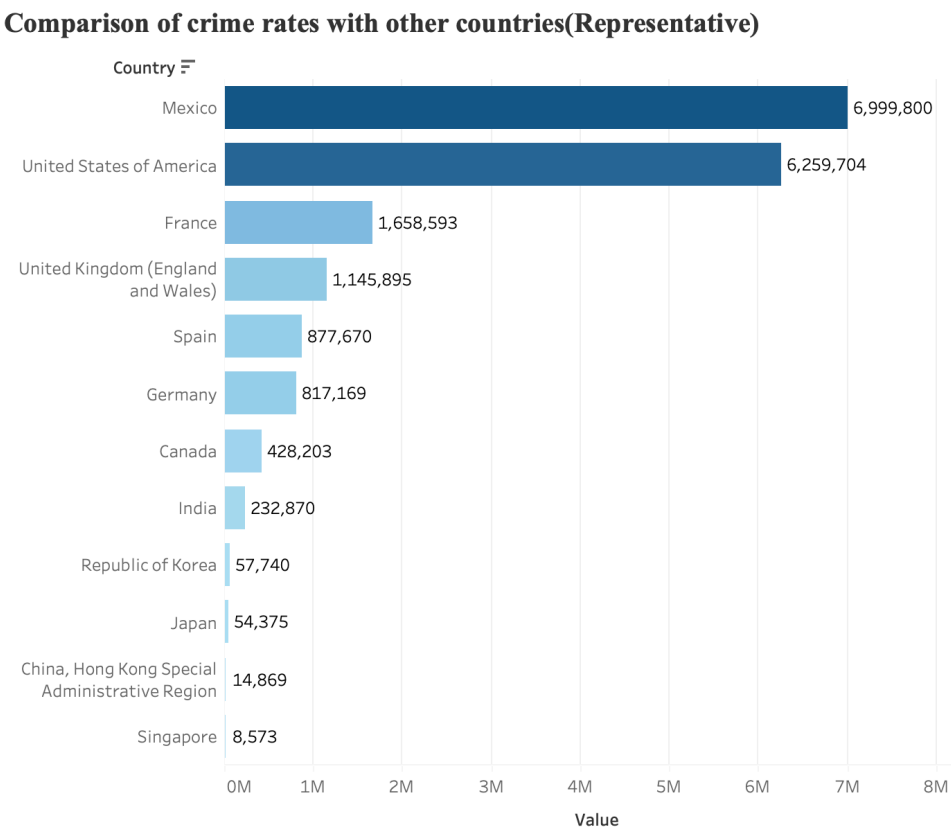
The fundamental conclusion of HK's 21st-century crime rate is: the trend goes down in general, though entering fluctuation since 2008.

——— *Conclusion 1*

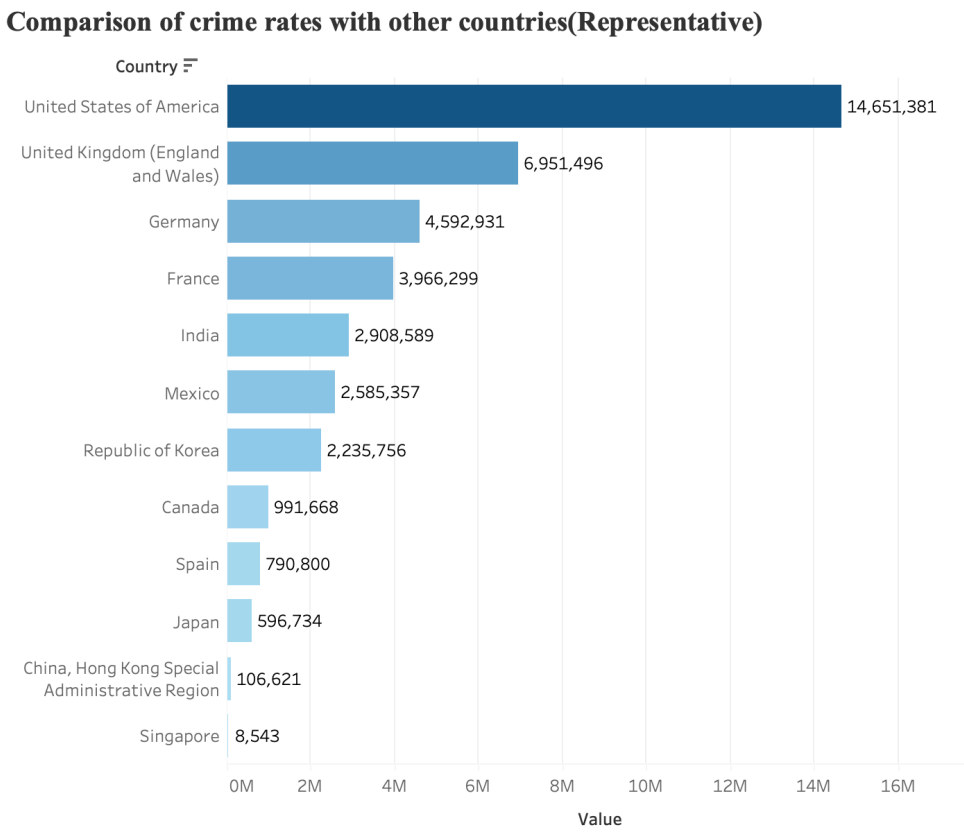


Graph 2 in dashboard 1 is a map comparison of the number of crimes in different countries for different crime types, which can be filtered and interacted with by the Year and Category filters on the right. By filtering Year, you can see how a certain crime type has changed over time and see if it has increased or decreased, and by filtering Category, you can see how the number of crimes in different countries compares for a particular crime type. For example, if you filter to Sexual Violence, you can see that Australia, Russia, and some European countries are more prominent in this type of crime base. This graph can help users get a better understanding of worldwide crimes.

CHAPTER 1



Data of Robbery

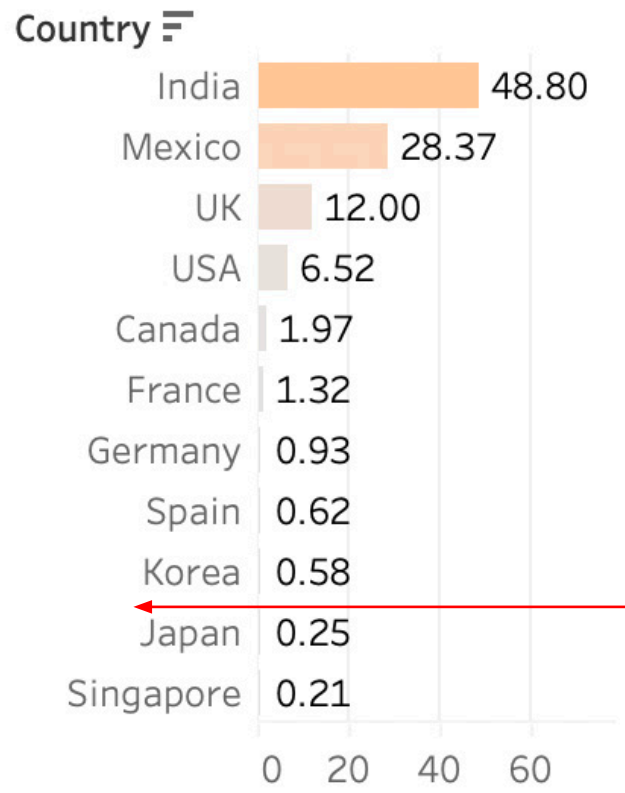


Data of Serious assault

Filter of Graph 3

Graph3 is a comparison bar chart of the total number of crimes based on the map crime base (number of crimes per 100K population), sorted by shades of color to get the ranking of TOP-N countries under a certain crime type, e.g. by choosing the Robbery crime type, you can see that in the Americas, e.g. Mexico, USA, their robbery crime numbers are much higher than other countries. HK's bar is almost at the bottom. This data can be considered as a sign that HK has much less crime than most of the regions in the world. However, HK has less population than the other countries on the list. To be precise, we also need to check the crime rate of these countries.

World crime rate（2020）



Graph 4 is a global comparison based on the crime rate per 100K people in 2020. From the graph we can see that India's crime rate is the highest in the world, can reach 48.8, followed by Mexico, which also can reach 28.37, ranked third and fourth are the United Kingdom and the United States. The 2020 crime rate of HK is 0.29, which places it third from the bottom of the list, a bit higher than Japan and Singapore. India is 168 times higher than HK.

Our conclusion from this dashboard is that: Asia has lower crime rates than other continents, and Hong Kong follows this basic observation. Although Hong Kong is not the region with the lowest crime rate in Asia, the overall performance of HK's 21st-century crime rate indicates Hong Kong's remarkable superiority in social security.

——————— *Conclusion 2*

CHAPTER 1

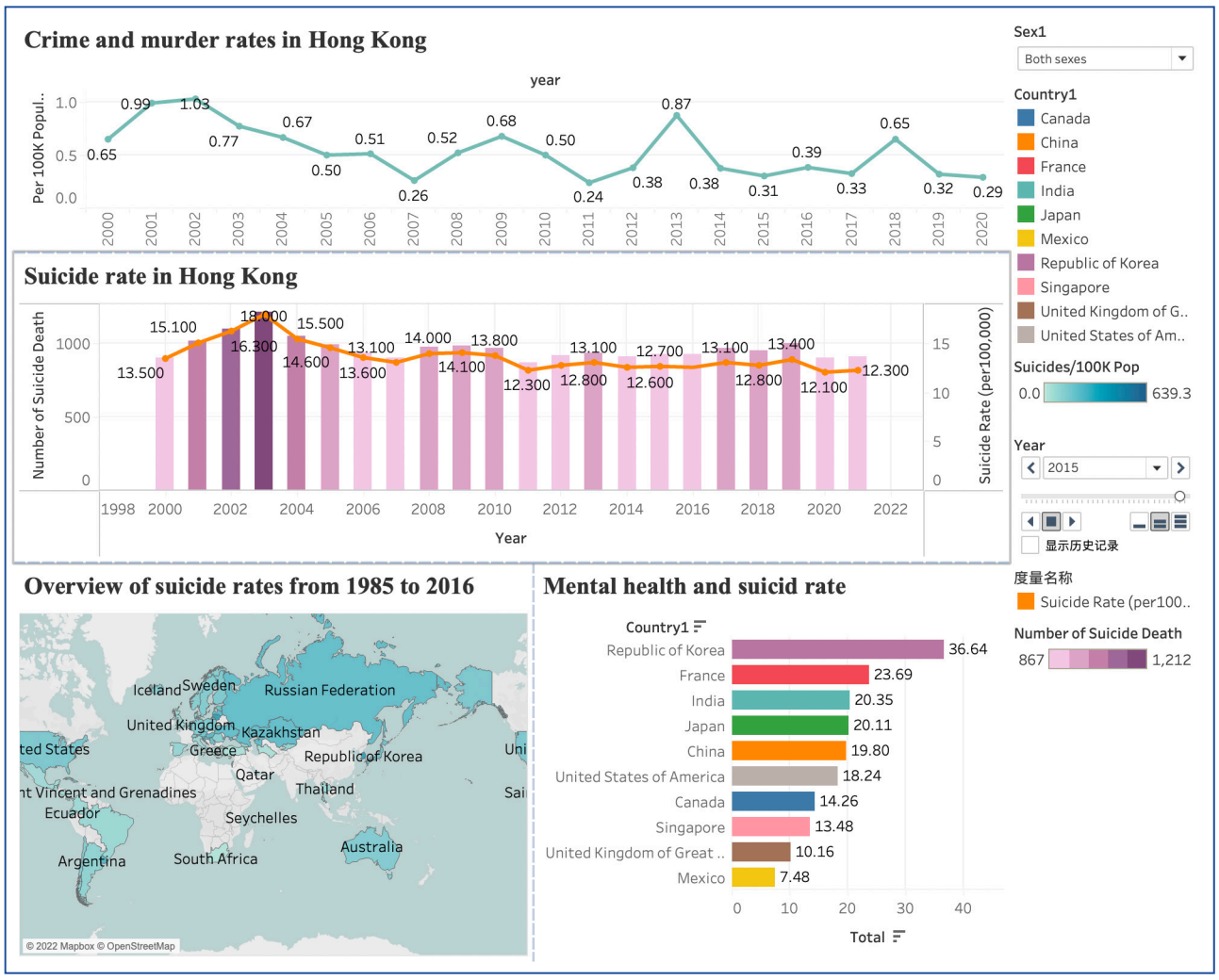
Dashboard 2 depicts HK's and worldwide suicide rates' relationship with HK's crime rate. ***We have a hypothesis that the suicide rate and crime rate are both indicators of social-emotional stability. There should be a positive correlation between the two rates.***

Graph1: Crime and murder rates in Hong Kong[(1)]

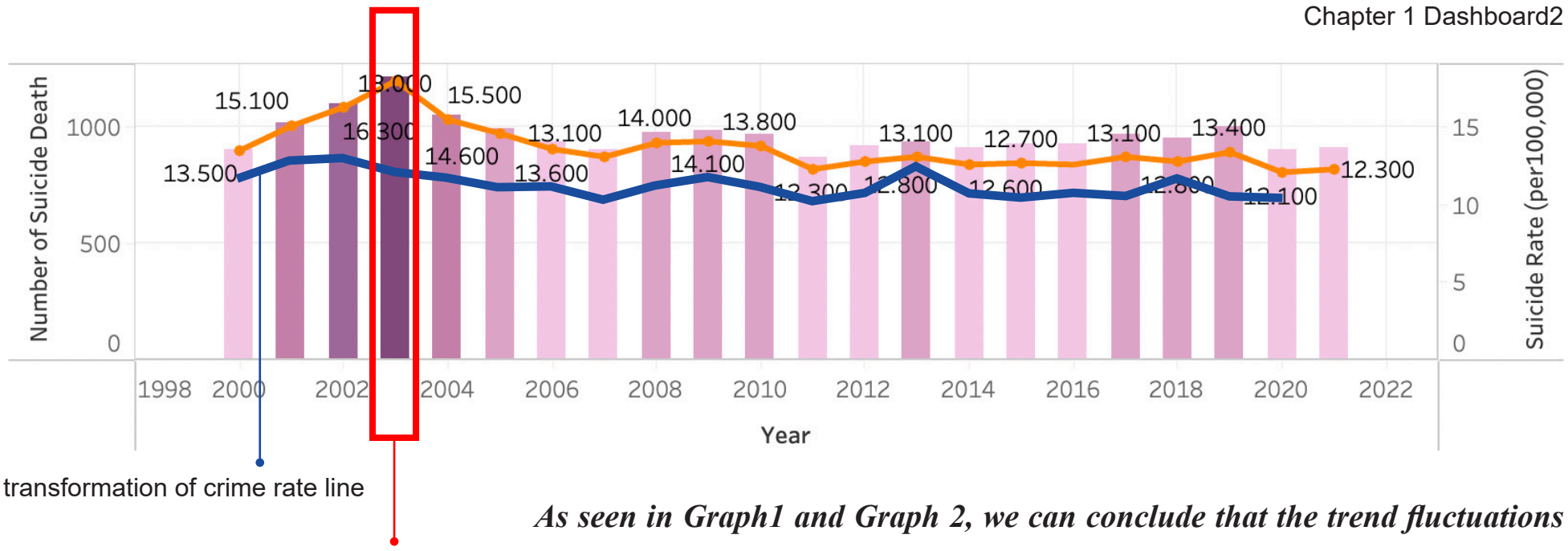
Graph2: Suicide rate in Hong Kong[(5)]

Graph3: Overview of suicide rates from 1985 to 2016[(6)]

Graph4: Mental health and suicide rate[(7)]



Chapter 1 Dashboard2



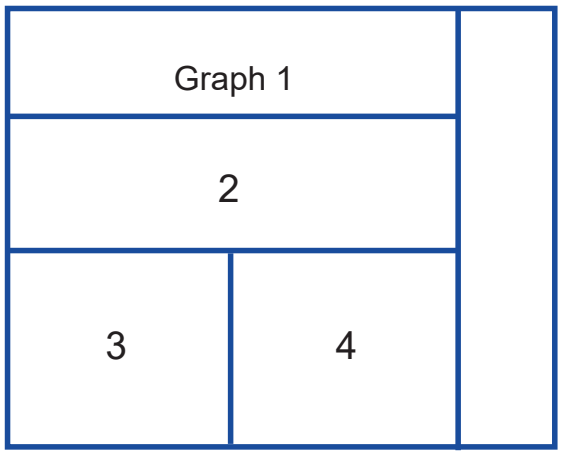
As seen in Graph1 and Graph 2, we can conclude that the trend fluctuations of crime and suicide rates show a clear positive and highly fitted correlation, which is accurately consistent with our hypothesis that the two indicators are demonstrations of the degree of social-emotional stability.

——— *Conclusion 3*

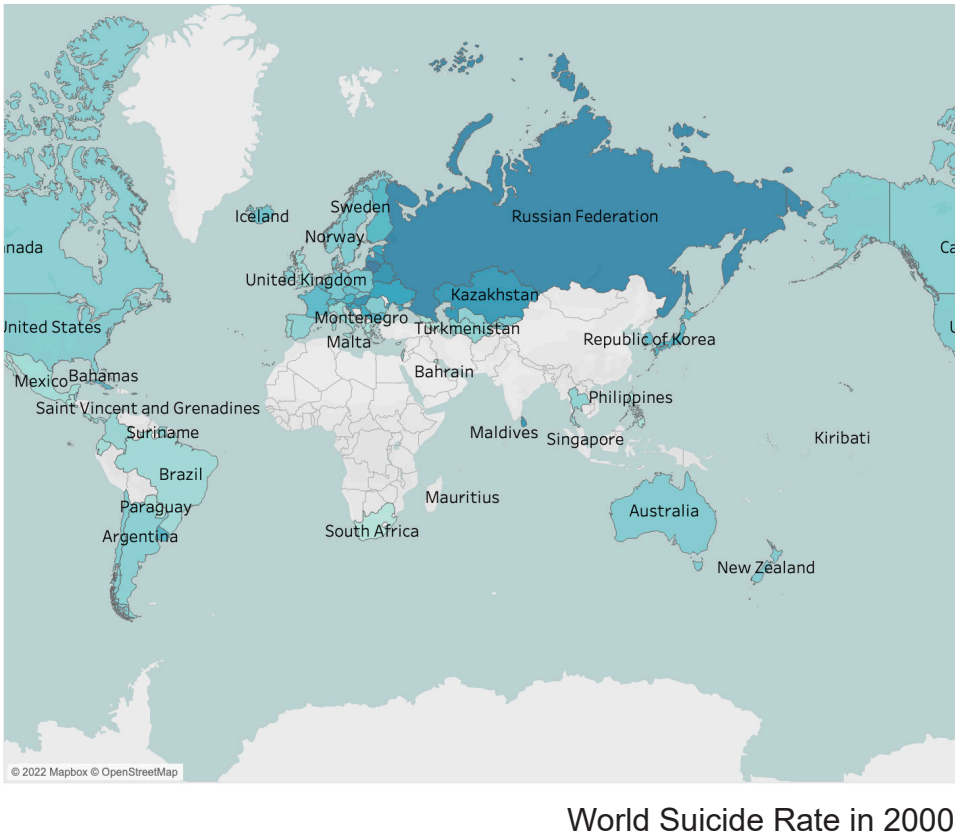
The trend of the two graphs does not fit consistently only in 2003, because we found that the SARS outbreak in 2003 was an extremely social event, both in terms of epidemic and economic and social damage, which depressed the mood of the people, even to the point of fear.

——— *Conclusion 4*

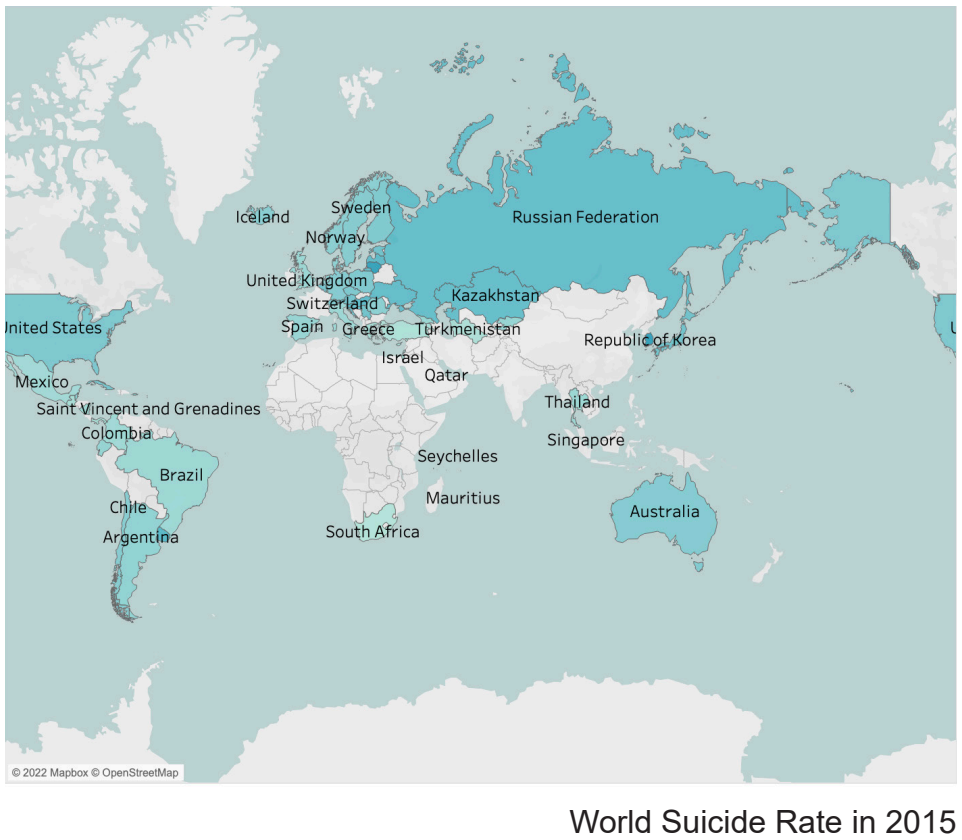
The exceptional state of coping with the epidemic reduced the activity of social activities and to some extent had a dampening effect on crime, but it did not alleviate the social mood and the suicide rate instead reached an all-time high.



CHAPTER 1

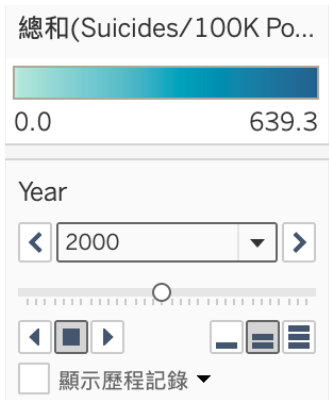


World Suicide Rate in 2000



World Suicide Rate in 2015

Graph3 is a global color map of suicide rates (per 100K base) from 1985 to 2016, the darker the color, the larger the suicide base in that country, in addition to the Year progress bar added to this graph, which can dynamically play the trend of the map, from the color heat map we can see that Russia, the United States, and other countries are relatively high global suicide rates countries. The map helps the users get a better understanding of worldwide suicide rate distribution.

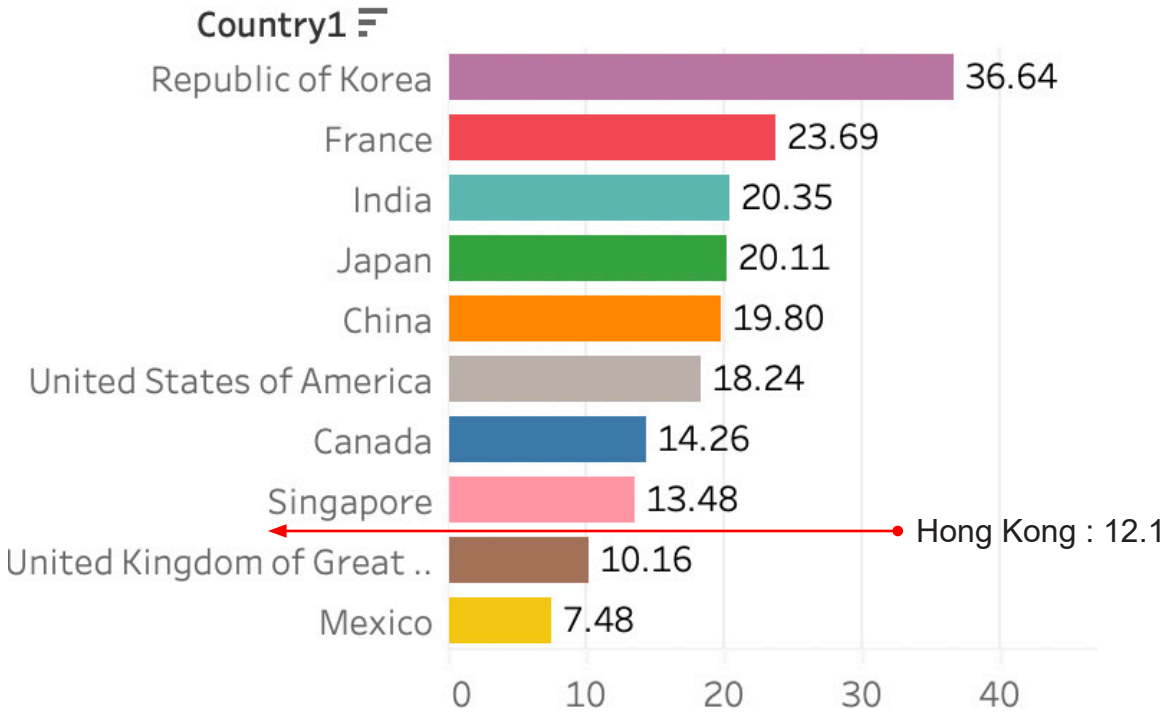


Legend and Filter of Graph 3

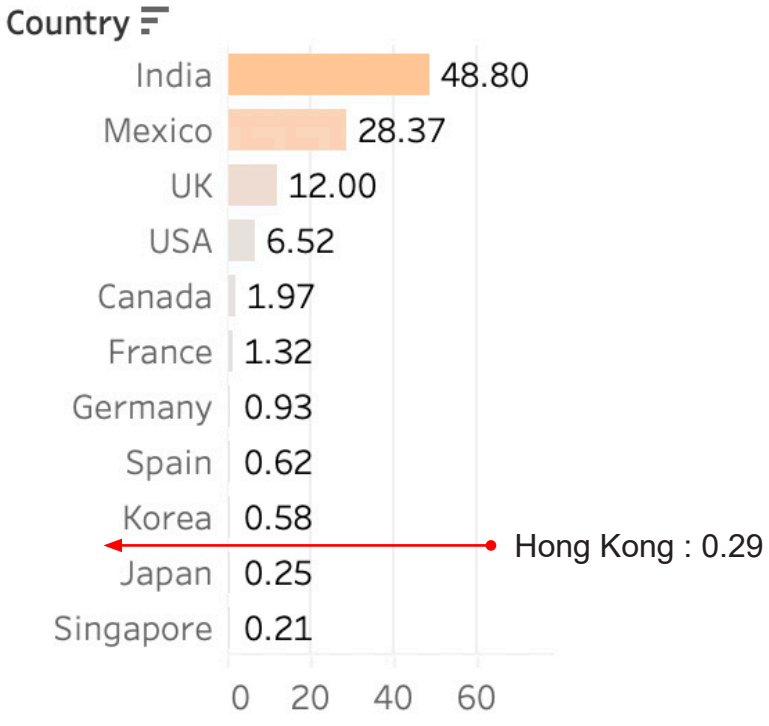
Graph4 is a global comparison of physical and mental health and suicide rates in different countries, we can see the country comparison of suicide rates for people of different genders by adding gender filters, we can see that from the perspective of male suicide, Korea, France, the United States, and Japan have relatively high suicide rates, while from the perspective of women, Korea is still ranked first, but China and India are ranked second and the third place. If we take HK's 2020 suicide rate: of 12.1 as a reference, it places HK third from the bottom of the list again. However, the countries with the low crime rate in Asia jump to the top tier in the suicide rate list.

The conclusion of comparing dashboard2 to dashboard1 is that: Hong Kong has both a low crime rate and a low suicide rate that represents a relative health social status, while the iconic countries chosen in the list cannot outperform Hong Kong.

Conclusion 5



Iconic Countries' Suicide Rates in 2020



Iconic Countries' Crime Rates in 2020

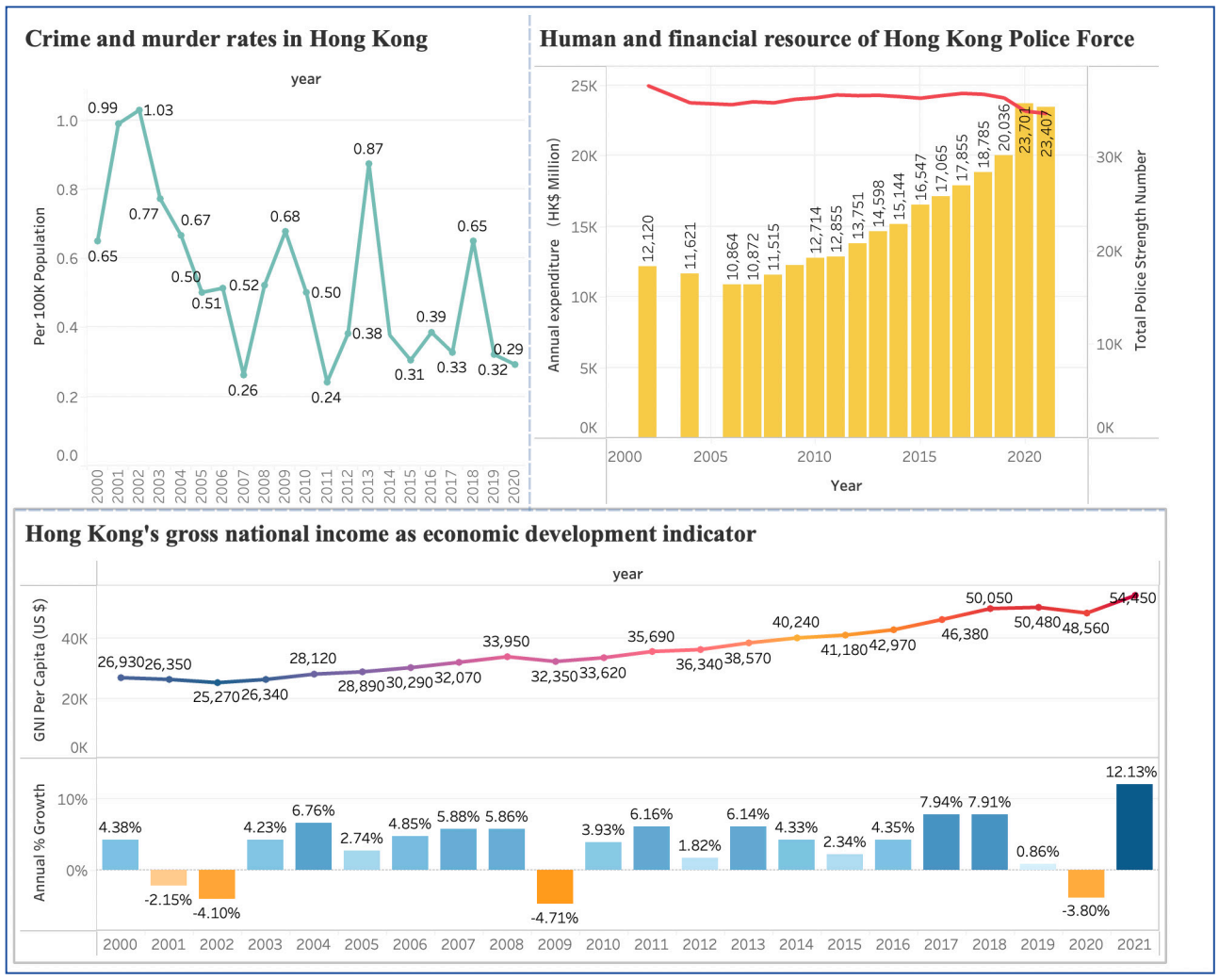
CHAPTER 1

Dashboard 3 tries to study whether the crime trend has a relationship with HK's economic development and Hong Kong Police Force's resources. ***We have two hypothesis that :1. better economic development can refrain from crime; 2. fewer resources of HKPF lead to increase in crime, vice-versa***

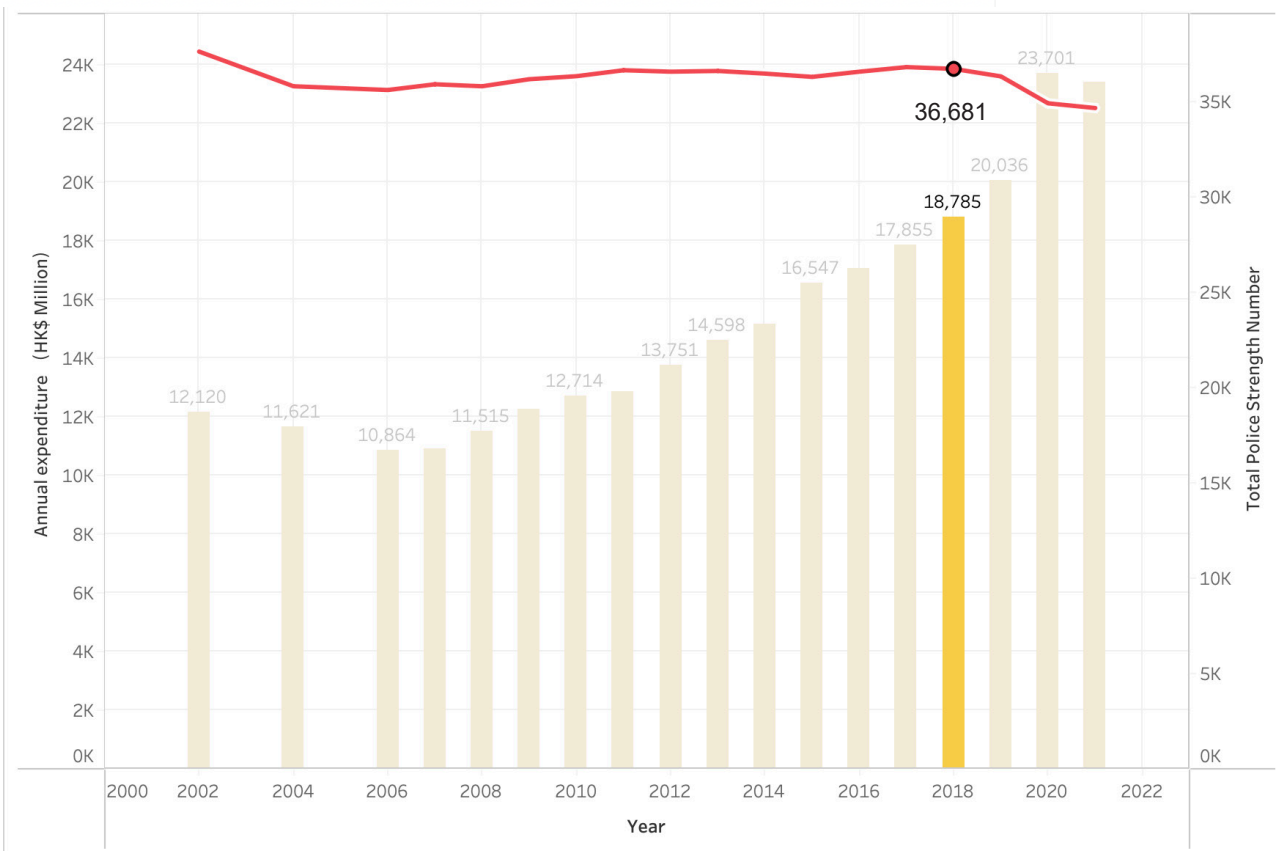
Graph1: Crime and murder rates in Hong Kong[(1)]

Graph2: Human and financial resources of Hong Kong Police Force[(8)]

Graph3: Hong Kong's gross national income as economic development indicator[(9)]

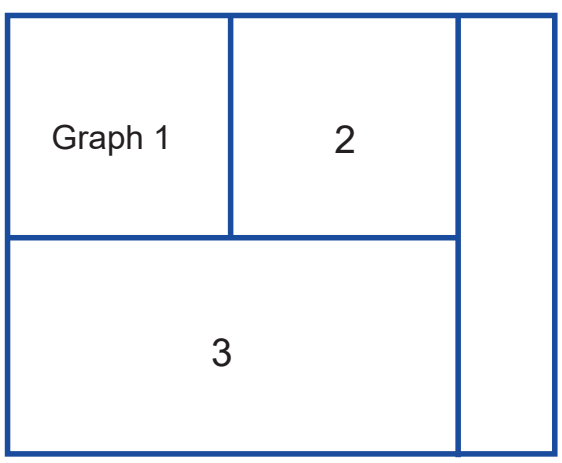


Chapter 1 Dashboard3



Data of HKPF resources in 2018 shows nothing special

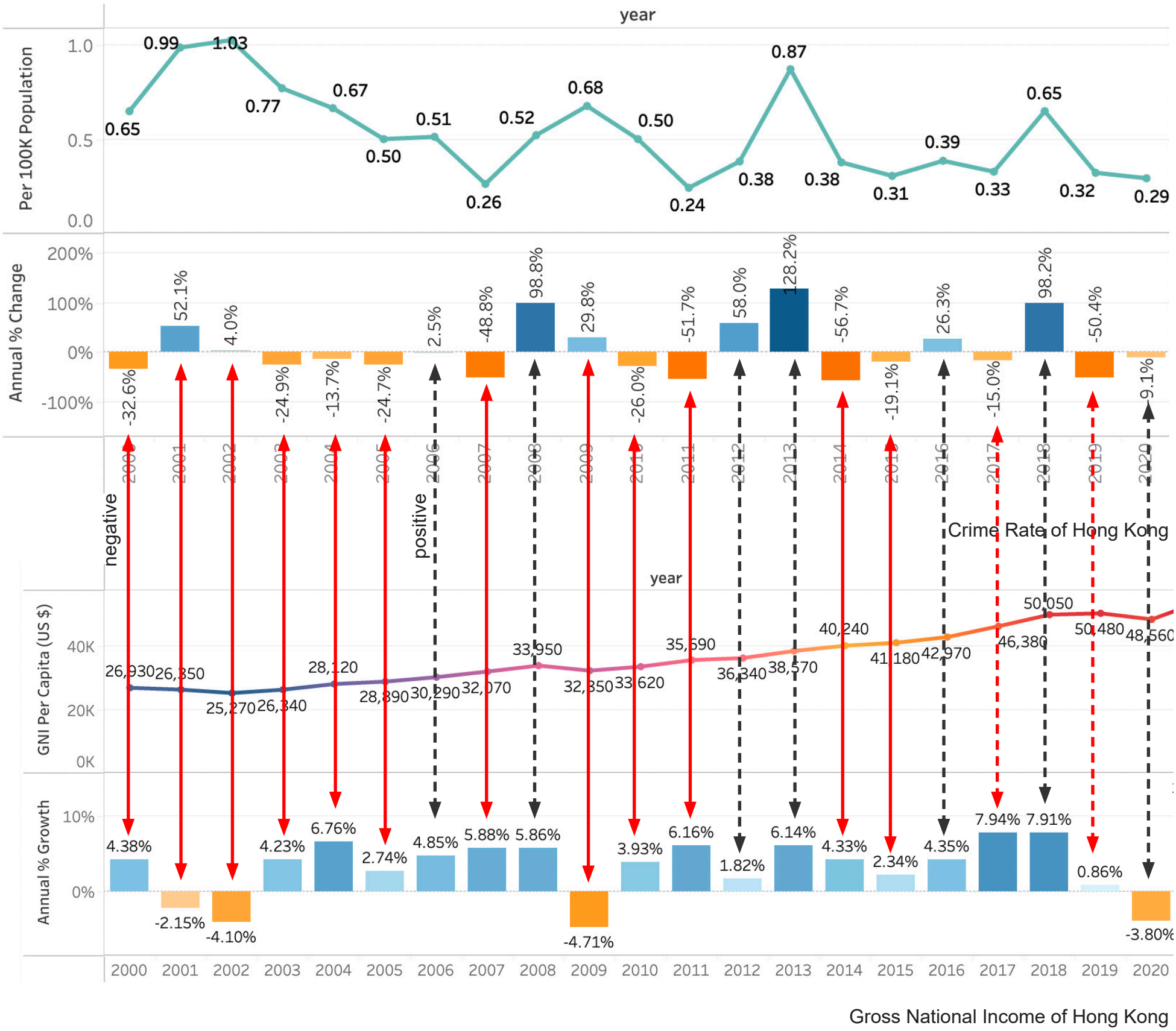
Graph2 in Dashboard 3 is a biplot of bar graphs and line graphs depicting the data trend of the number of police officers and fiscal revenue in Hong Kong during 2000-2020, which shows a clear growth trend of fiscal revenue in Hong Kong, from 12,000 in 2000 to 23,000 in 2020, almost a doubling of the level. However, from the red line graph of the number of police officers, we can see that the number of police officers in Hong Kong is on a downward trend, dropping to a level of more than 34,000 by 2020. For instance, the human and financial resource of HKPF in 2018 cannot explain the sudden rising in crime rate.



There is no clear relationship between police resources and trends in crime rates in Hong Kong, which does not align with our hypothesis that fewer resources lead to more crimes. The growth of the annual expenditure of HKPF should be the result of increasing the salaries of officers and purchasing police vehicles and equipment.

——————— *Conclusion 6*

CHAPTER 1



Graph 3 in Dashboard 3 is the trend of Hong Kong's gross national income and growth rate. From the growth of the GNI index, we can see that GNI grows from 26,000 in 2000 to 54,000 in 2020, a doubling of the level. And from the rate value of the growth rate, there are fewer years of negative growth, most of them are positive growth levels, especially in 2021, the growth rate reaches the highest, and the rate value can reach 12% positive growth.

From the comparison of graph1 and 2 in dashboard 3, it can be found that the trend of crime rate and GNI show a clear negative correlation until 2005. The overall trend of GNI is increasing, while the overall trend of crime rate is decreasing, and the negative growth of GNI in 2001, 2002, and 2009 can correspond to the positive growth of the crime rate. However, after 2011, the economic indicator GNI remained stable and increasing, but the crime rate trend fell into fluctuation. The growth of the crime rate in 2013 and 2018 cannot correspond to any clear signal in GNI data.

Economic development has less influence on crime rates nowadays than at the start of the 21st century.

——— *Conclusion 7*

CHAPTER 2 The offense types' category pattern and demographic characteristics of offenders

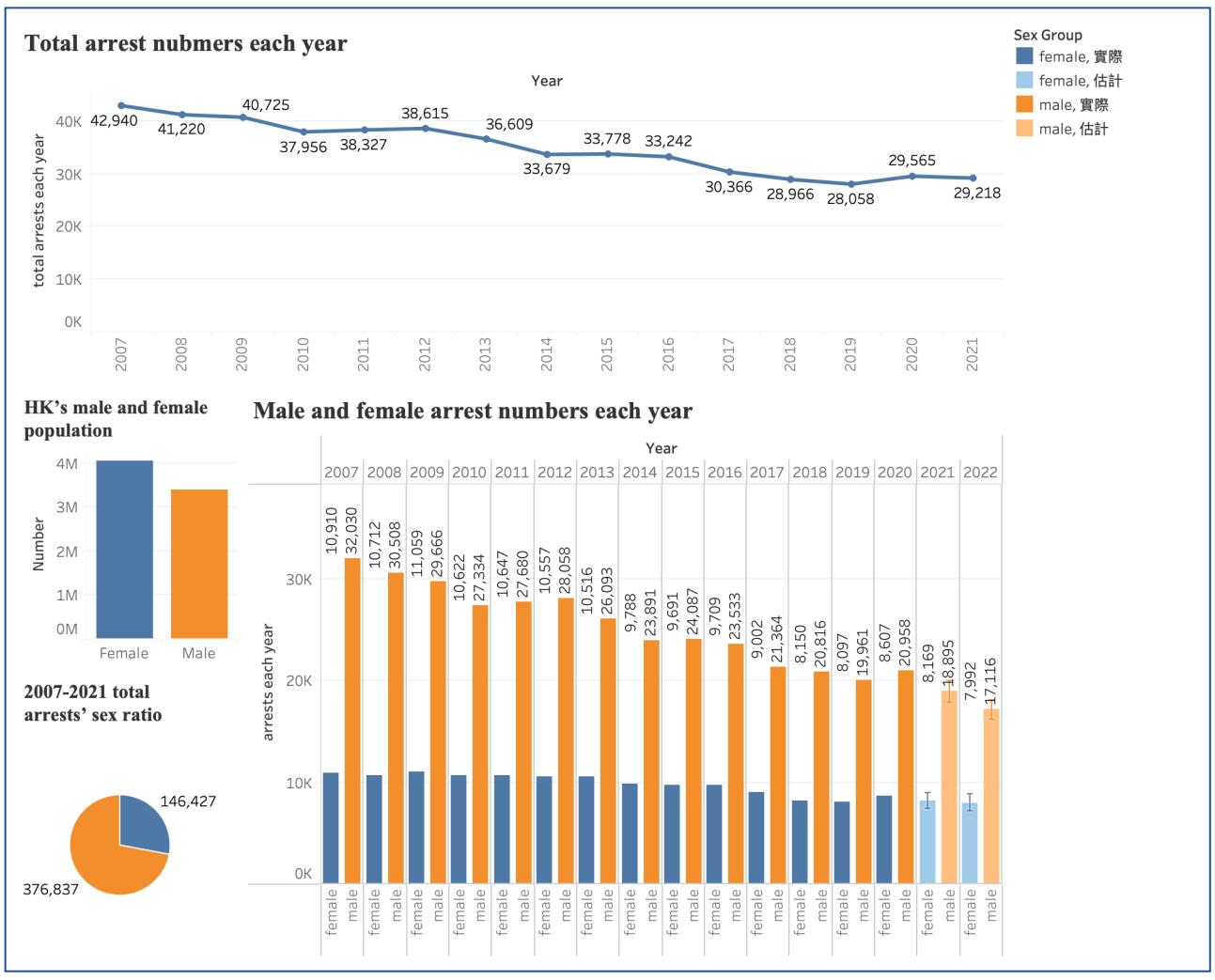
The first dashboard in chapter2 focus on trends in total arrest numbers and the gender groups' behavior. It is a generalized social experience that men are more likely to commit crimes than women. But to what extent is the disparity in crime rates between men and women?

Graph1: Total arrest numbers each year[10]

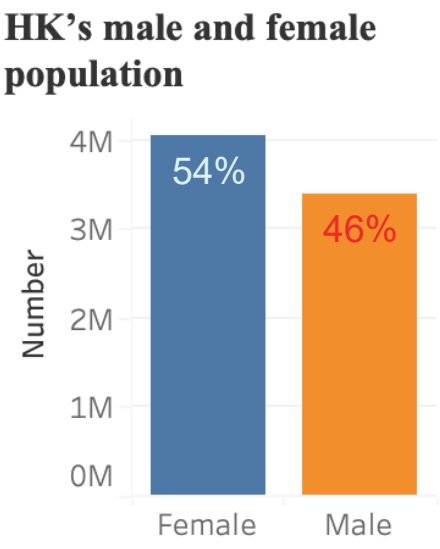
Graph2: HK's male and female population[11]

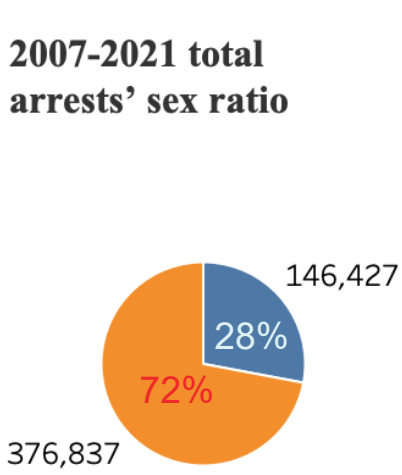
Graph3: 2007-2021 total arrests' gender ratio[10]

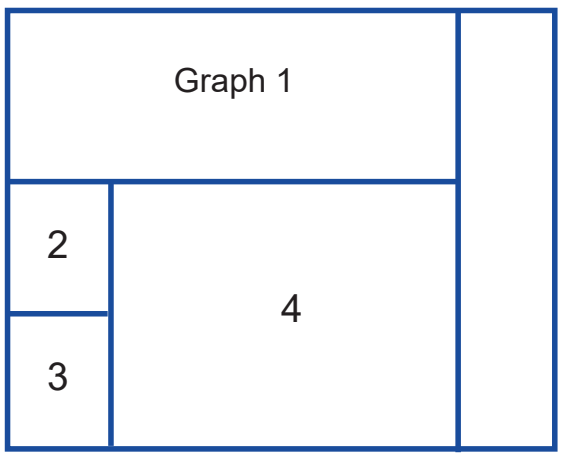
Graph4: Male and female arrest number each year[10]



Chapter 2 Dashboard1







The graph1 shows a clear downward trend in the total number of arrests from 2007 to 2021. The total number of arrests in 2020 is only 69% of the 2007's figure. Due to limitations in the data provided by the government, the analysis of the number of arrests does not include data before 2007.

Then we study the gender ratio of the arrested offenders. Firstly, when comparing graph2 and graph3, we can see that 54% of Hong Kong's population is female. While only 28% of the total arrests are female.

Male groups are 3 times more likely to commit crimes than female groups.

——— *Conclusion 8*

Secondly, the bar chart of graph 4 shows that the declining proportion of male groups is much higher than that of female groups. In 2007, male vs. female was 2.94 to 1. In 2020, male vs. female was 2.43 to 1. The number of male offenders arrested in 2007 was 32,030, while the figure is 20,958 in 2020, only 65% of 2007. Tableau can forecast data for 2021 and 2022 based on data up to 2020.

We can conclude here: the number of male arrests shows a significant drop in the total number compared to the female data.

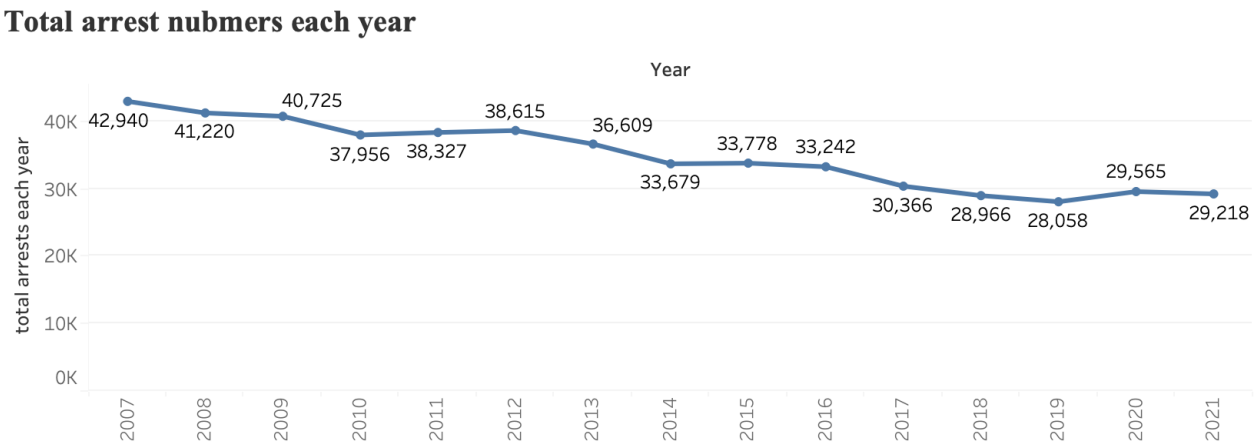
——— *Conclusion 9*

CHAPTER 2

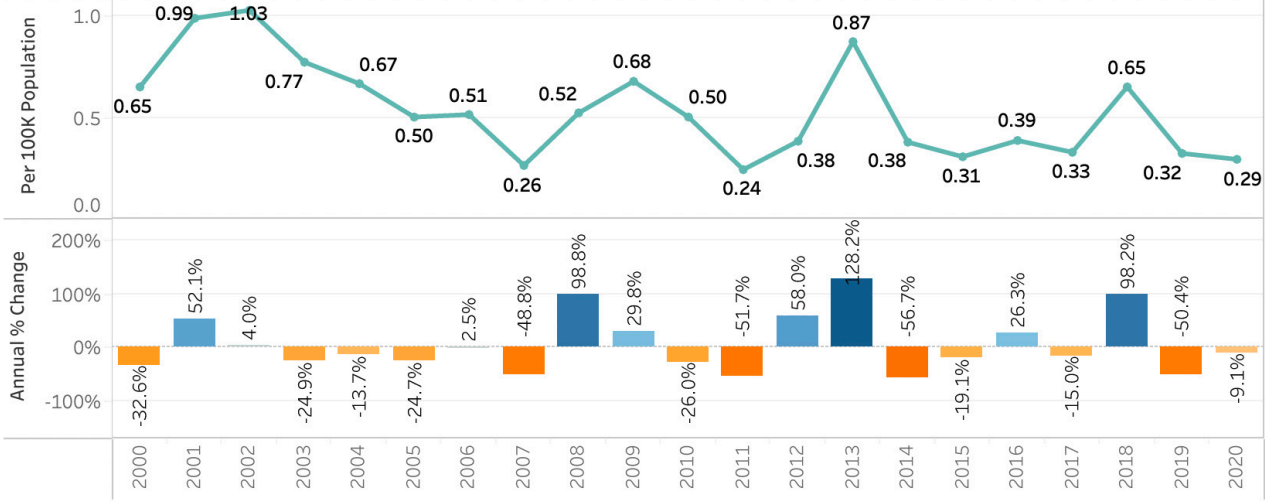
According to the crime rate data in chapter 1, the crime rate in 2007 and 2020 in HK are respectively 0.26 and 0.29 per 100,000 population. Hong Kong's total population increased from 6.9 million in 2007 to 7.4 million by 2020. In general, the arrest number should show the same fluctuation as the crime rate, whereas it does not.

Our first suspicion is that the data collection standard of crime rate from the World Bank is not aligned with that of the arrests data from the Hong Kong Census and Statistics Department. It is difficult to verify whether the data itself is right or wrong.

So we followed up with the second suspicion that declining detection rates lead to fewer arrests. The data published by the Hong Kong Police Force verified our suspicions. The detection rates of overall crime and violent crime in 2007 are respectively 45.6% and 65.5%. Nevertheless, the detection rates of overall crime and violent crime in 2020 are respectively 37.8% and 52.4%.



Chapter 2 Dashboard1



Chapter 1 Dashboard1

罪案 CRIMES	2007
整體罪案 Overall crime	80 796
破案率 (%) Detection rate (%)	45.6
每十萬名人口計算的整體罪案比率 Overall crime per 100 000 population	1 166.6
暴力罪案 Violent crime	14 934
破案率 (%) Detection rate (%)	65.5
每十萬名人口計算的暴力罪案比率 Violent crime per 100 000 population	215.6

HKPF Published Data

罪案	2020
整體罪案	63 232
破案率 (%)	37.8
每十萬名人口計算的整體罪案比率	845
暴力罪案	9 391
破案率 (%)	52.4
每十萬名人口計算的暴力罪案比率	126

HKPF Published Data

Declined detection rate leads to the differential between the crime rate trend and arrests number trend.

——— *Conclusion 10*

If we recall conclusion 2, there are questions arising. ***Does the significant drop in male numbers result from a lower male crime rate? Or it is because more males escaped from being arrested?*** The questions need more data analytics and sociological research to answer.

CHAPTER 2

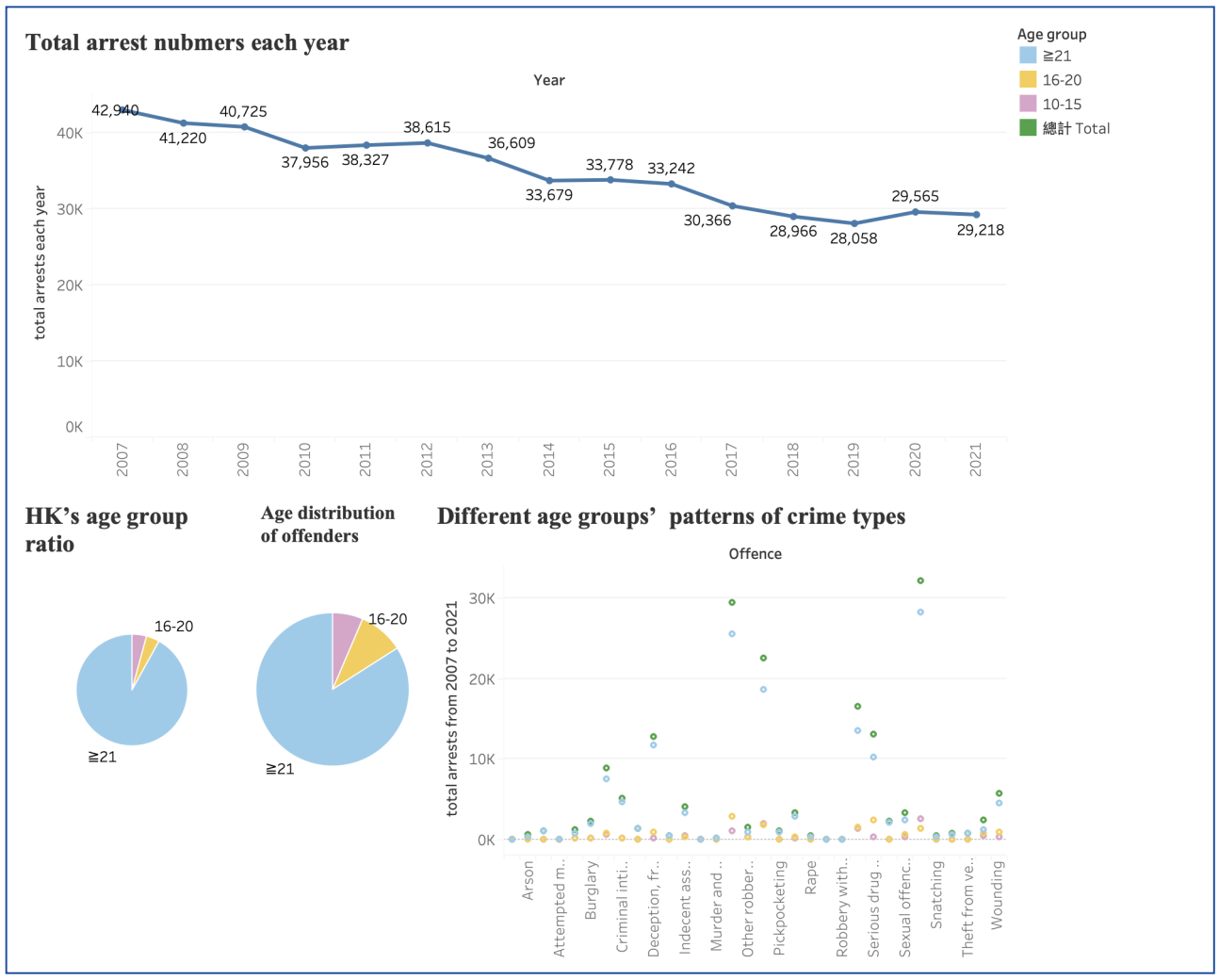
The second dashboard in chapter2 focuses on age groups' performance. Will adults commit more crime than teenagers? Are there any differences among the crime types that are committed by each age group?

Graph1: Total arrest numbers each year(10)

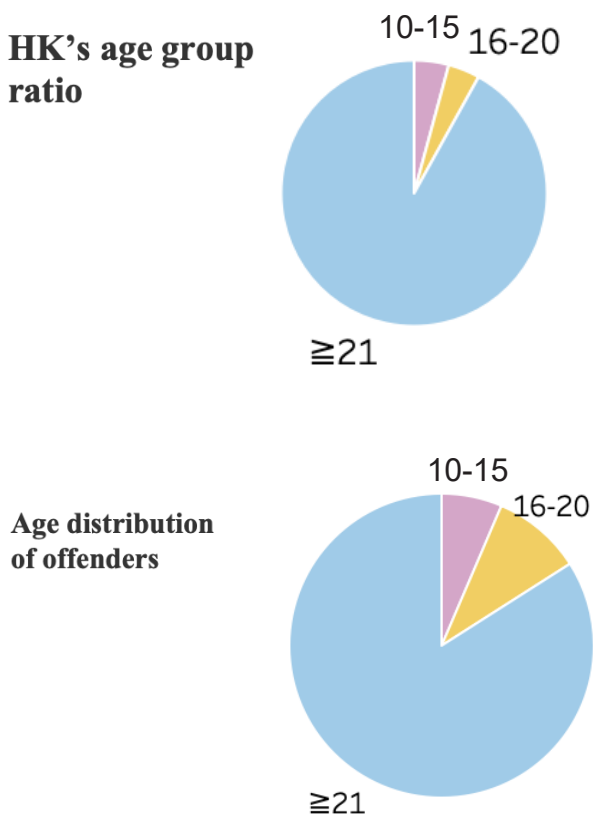
Graph2: HK's age group ratio(11)

Graph3: Age distribution of offenders(10)

Graph4: Different age groups' patterns of crime types(10)



Chapter 2 Dashboard2

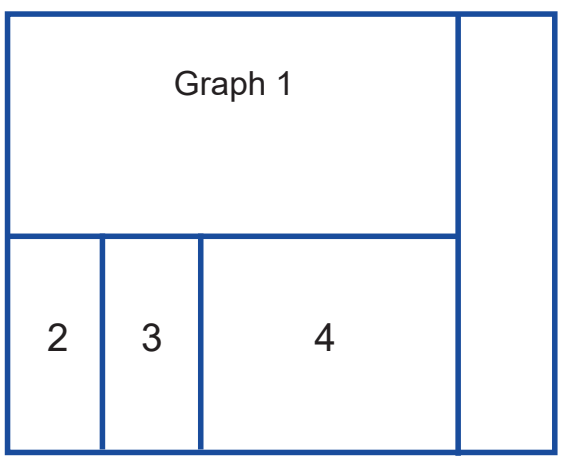


In the Graph2 pie chart, 10-15 occupies 4.2%,16-20 occupies 3.84%, and over 21 occupies 91.96%. While, in Graph 3, 10-15 occupies 6.41%, 16-20 occupies 9.53%, and over 21 occupies 84.06%. When we compared the census data with the total arrests, the percentage of 10-20 teenagers raises from 8.04% to 15.94%. Meanwhile, the percentage of groups over 21 years old decreased from 91.96% to 84.06%.

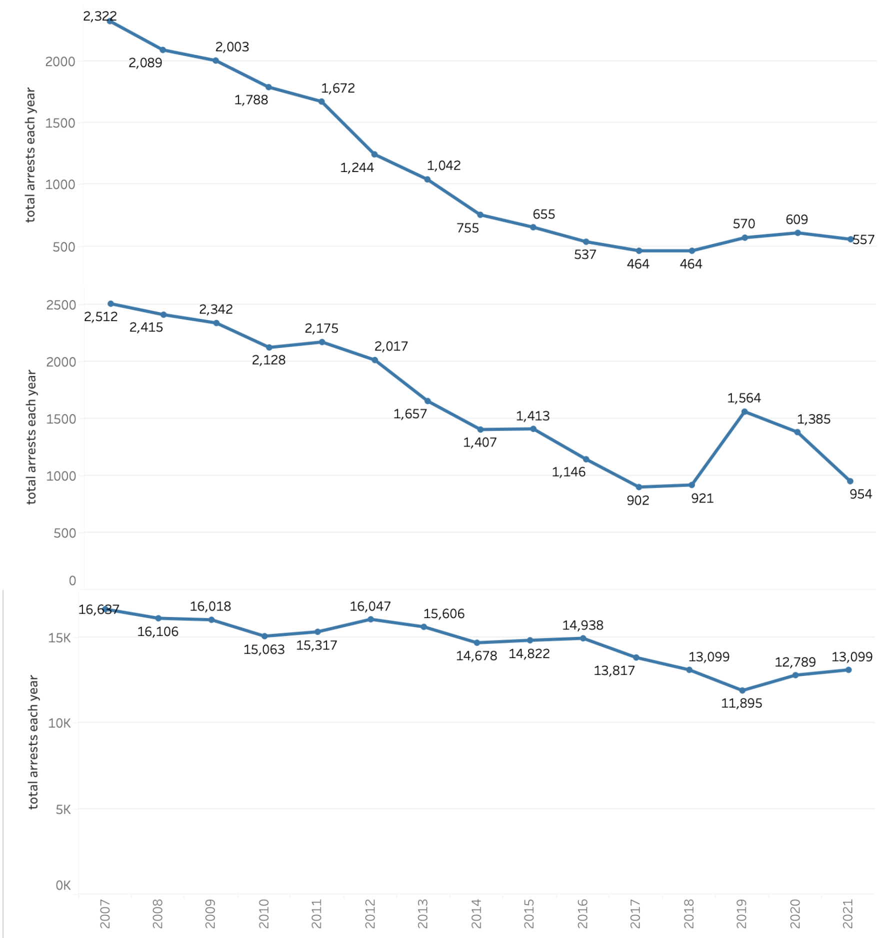
Based on the visual look of the data, one can simply conclude: the teenage population has a higher proportion of offenders and they are 2.2 times more likely to commit crime than adults.

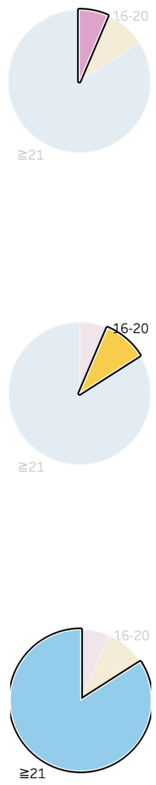
——— *Conclusion 11*

Whereas, we can observe there are hidden unfair data treatments between the two datasets. The data provided by the government has no arrests data lay in the range 0-9. The reason is obvious, 0-9-year-old children have little ability in committing a crime. Therefore, all the data in Graph 3 are limited to the range of groups who have the criminal ability. The data in Graph 2 does not share the same standards. Hong Kong is an aging city, which has a huge group of senior citizens. Most senior citizens lose their criminal ability as body function declines. It is difficult to strip this part of the data from the overall data as there is no standard of the upper age limit for commiting crime. ***To revise the conclusion above, more information on the numbers of senior offenders can help to identify a more accurate dataset in this case.***



CHAPTER 2





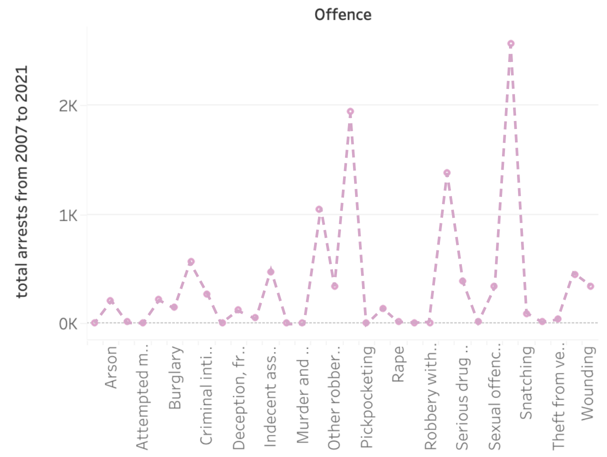
When we filter the total arrests data with age groups, the trends show different patterns. In the age group of 10-15, data in 2021 is only 23.9% of the data in 2007. In the age group of 16-20, data in 2021 is 26% of the data in 2007. In the age group of over 21, data in 2021 is 78.6% of the data in 2007.

The significant reduction in juvenile crime in Hong Kong is attributed to the Hong Kong government's legal literacy and prevention efforts.

——————— *Conclusion 12*

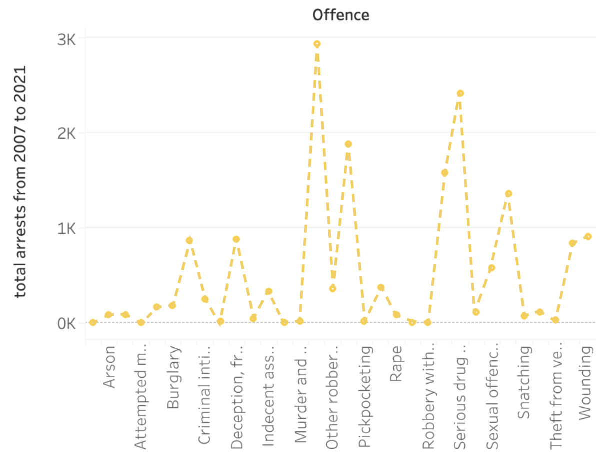
Age group
10-15

Different age groups' patterns of crime types



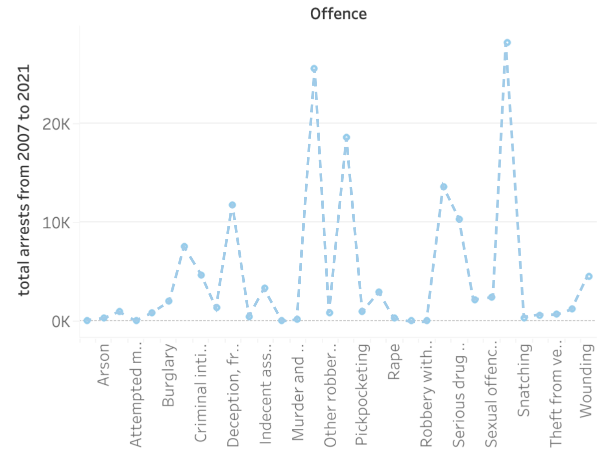
Age group
16-20

Different age groups' patterns of crime types



Age group
≧21

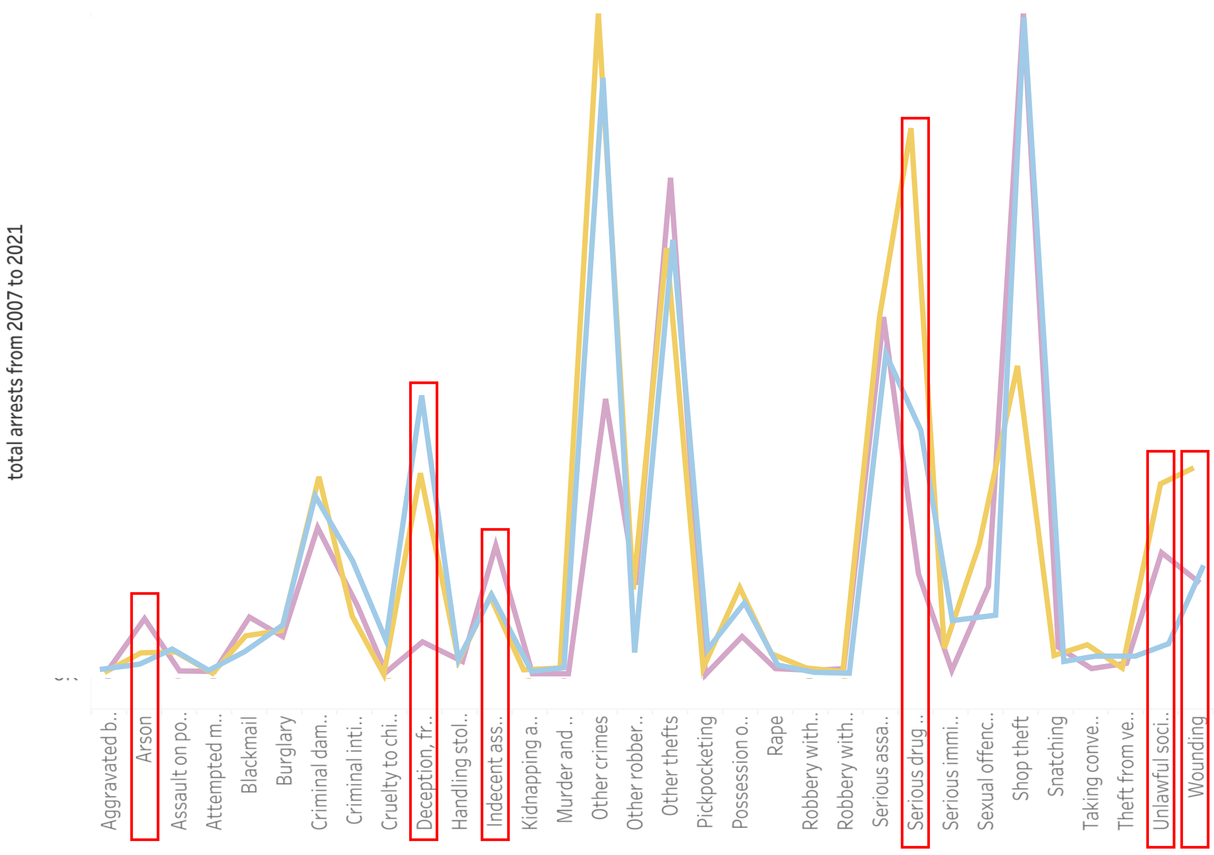
Different age groups' patterns of crime types



When comparing the filtered crime types patterns of the three age groups (the three groups' patterns share the same upper and lower lines in graphy), we can find some highlighted types that show obvious age preference differences.

10-15 group shows preference in arson and indecent assault.
16-20 group shows preference in serious drug offenses, unlawful society offenses and wounding.
Over 21 group show preference in deception, fraud and forgery.

——————— *Conclusion 13*



CHAPTER 2

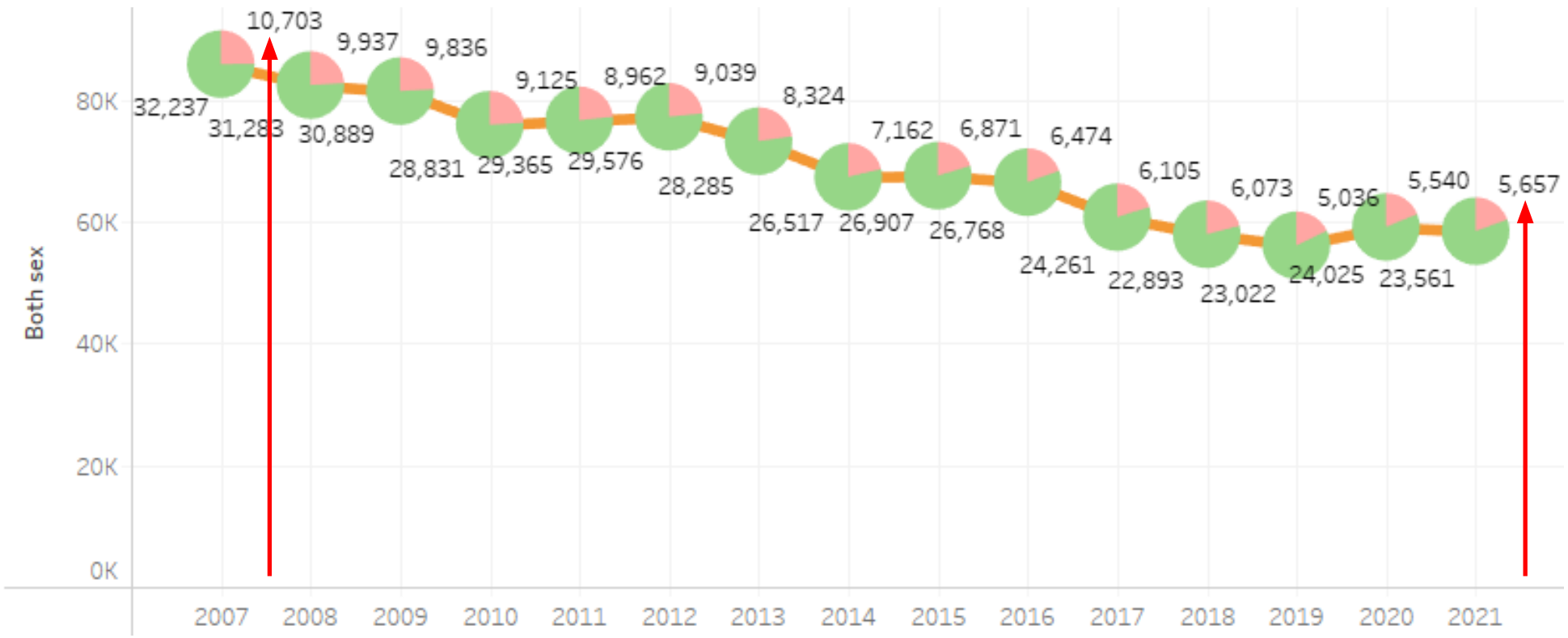
The last dashboard focuses on the types of crime and the percentage of men and women who commit crime in Hong Kong. What are the trends in the total number of crimes in recent years and how have the types of crime changed?

Graph1: Overall number of crimes and the proportion of violent and non-violent crimes(10)

Graph2: The number of crimes committed by each gender and the proportion of violent non-violent crimes(10)

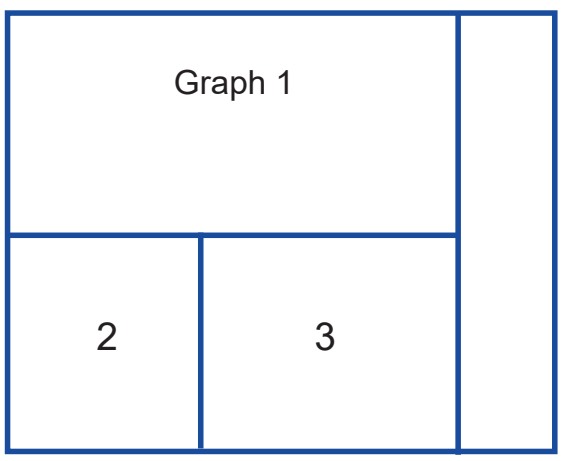
Graph3: Breakdown of crime types and numbers(10)

Chapter 2 Dashboard3

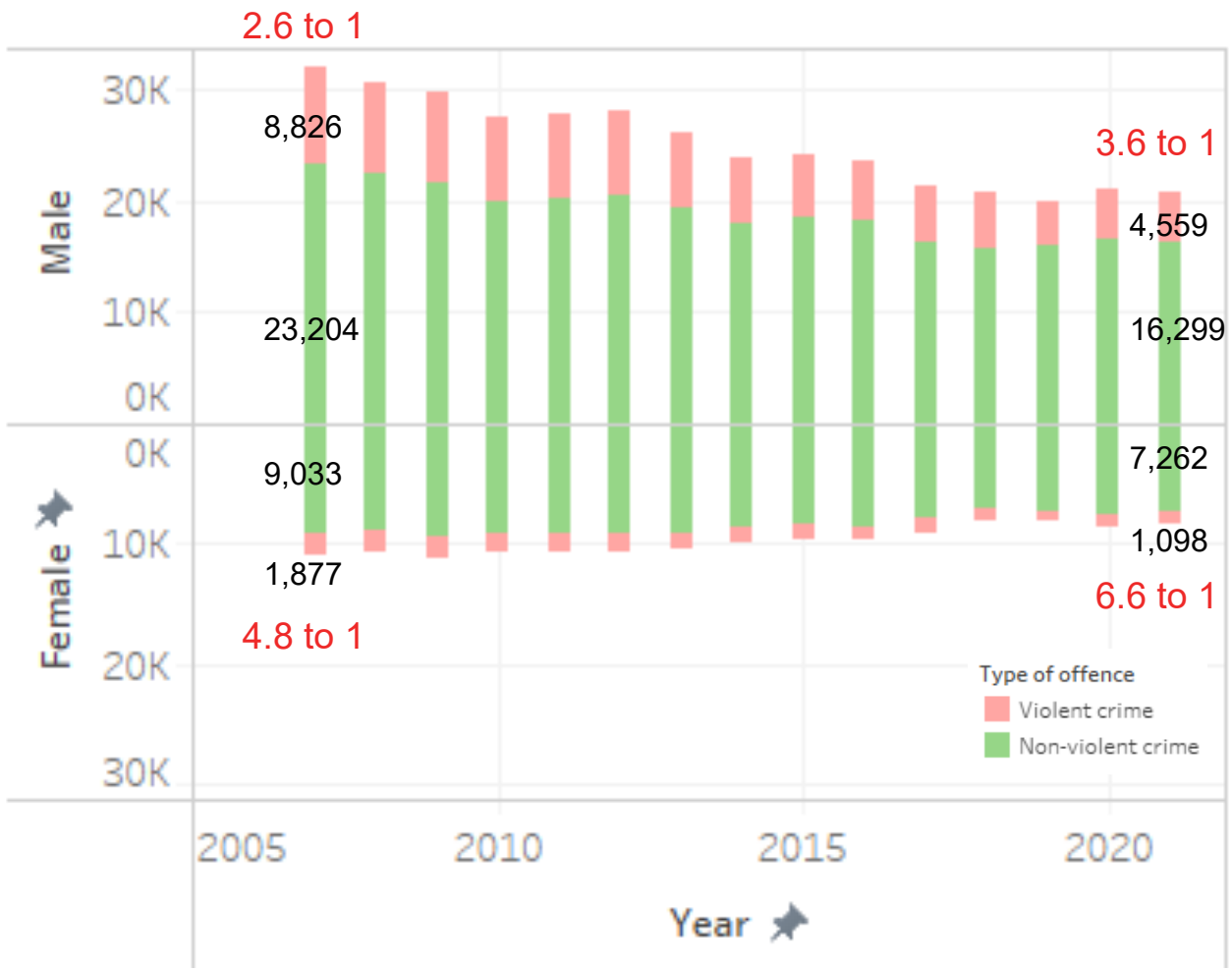


The Graph1 shows the total number of crimes per year for all genders and the respective proportions of violent and non-violent crimes. The graph shows that the overall number of crimes has decreased slightly, from a total of 42,940 in 2007 to 29,218 in 2021. At the same time the proportion of violent crime is gradually decreasing. It has decreased from 10,703 to 5,657. ***From the graph we can conclude that the total number of crimes and the proportion of violent crimes are gradually decreasing.***

Conclusion 14



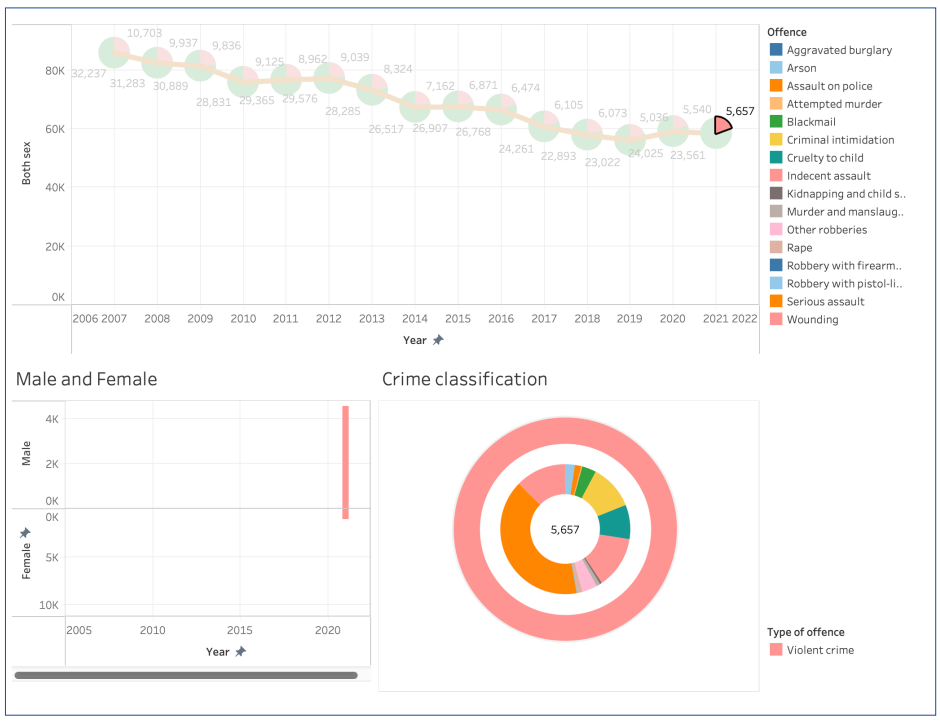
CHAPTER 2



The bar chart of Graph 2 shows the proportion of violent and non-violent crimes for men and women respectively, where the ratio of violent to non-violent is 23:9 for men and 9:2 for women, with the total number of crimes for men being much greater than for women, about three times greater. We can see from the bar chart that the proportion of violent crime for males compared to females has decreased significantly over time, from 8,826 to 4,559. ***We can therefore conclude that the male group are more likely to be involved in violent crime comparing to the female group, though both gender groups tend to commit less violent crimes recent years than 2007.***

——— *Conclusion 15*

Violent Crime

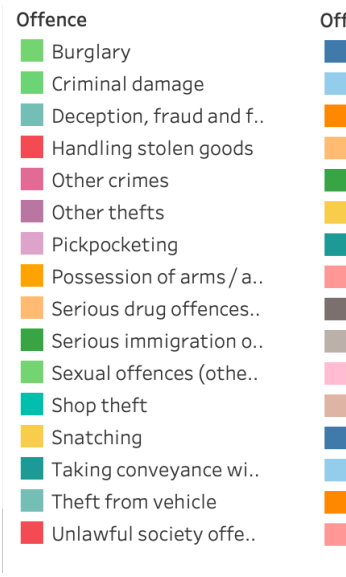


Non-Violent Crime

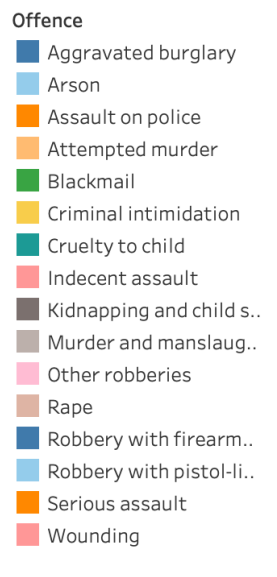
And Graph 3 are two rings. The large outer ring shows the proportion of violent and non-violent crime, while the small inner ring shows the specific types of crime and their proportions, but the exact figures change when we click on the category in the outer ring or select the small pie chart in the graph above.

When we select the small pie chart above we can see that the percentage of violent crime has dropped from 24% in 2007 to 19% in 2021. When we select individual violent or non-violent crimes, we can see that the top three violent crimes, which account for the largest proportion of violent crimes, change from serious assault, drugs, and other crimes to serious assault, wounding, and other crimes, while shoplifting and cheating consistently account for the largest proportion of non-violent crimes. We can see that drug offenses are much lower in 2021 compared to 2007, while the other major categories of crime still account for a large share

——— *Conclusion 16*



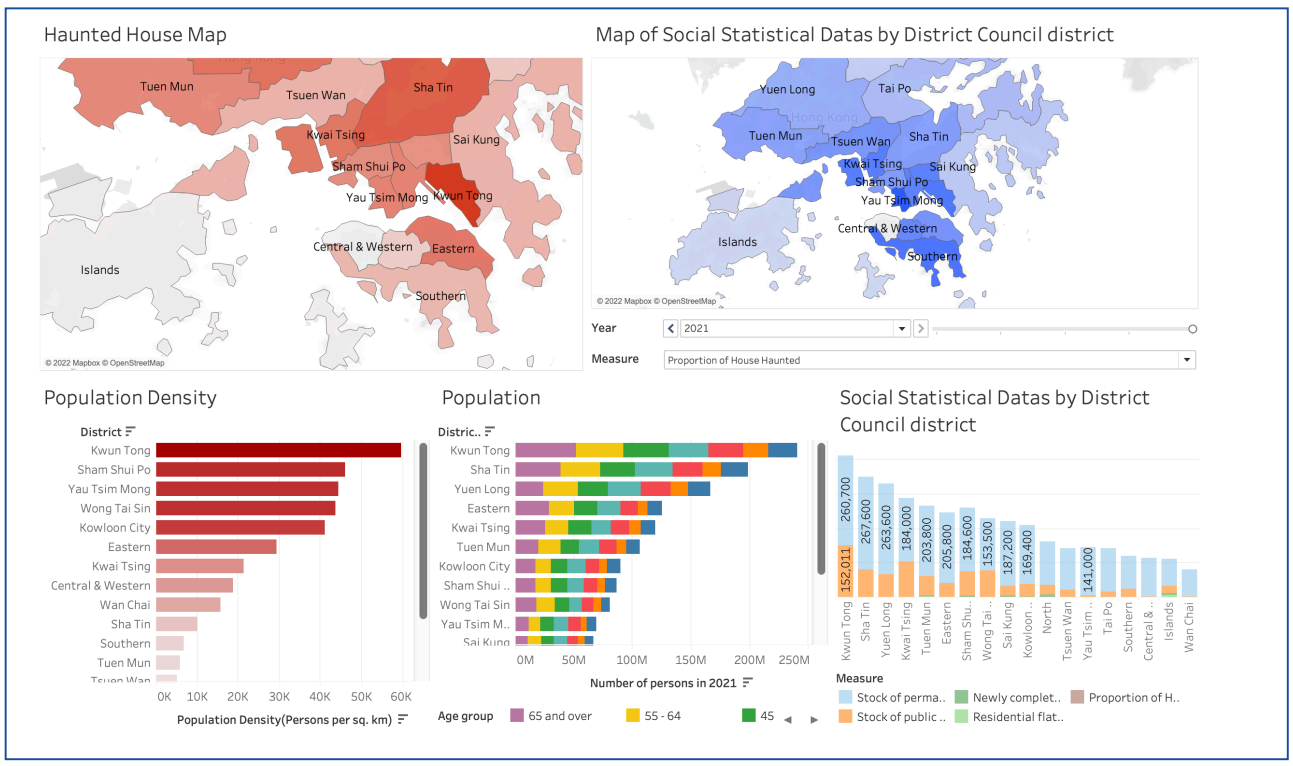
Violent Crime



Non-Violent Crime

CHAPTER 3 Haunted houses map with real estate development data

In this chapter, we are focusing on the number of haunted houses, population and house types by districts in Hong Kong.



Chapter 3 Dashboard1

Graph1: The Haunted House Map[13] plot at the top left illustrates the amount of haunted houses among the 18 districts of Hong Kong. As the color becomes more dark red, the amount of haunted houses in that district becomes larger. By just looking at the Map plot, we can see that Yau Tsim Mong has the most haunted houses and Islands has the least haunted houses.

Graph2: The Social Statistical Data Map[12] plot at the top right shows the real estate development in the 18 districts. By changing the year and the type of house supply filter, we are able to see the change of the number of houses across years.

Graph3: The Social Statistical Data Bar chart[12] at the bottom right shows the house supply data in the Social Statistical Map from another angle. As we change the filter in the Social Statistical Data Map, this bar chart changes accordingly, allowing us to get a more clear sense of the quantity differences among different types of houses. Moreover, this chart shows the proportion of haunted houses and total houses by districts, which is a more accurate indicator as it takes both haunted houses and other types of houses into consideration.

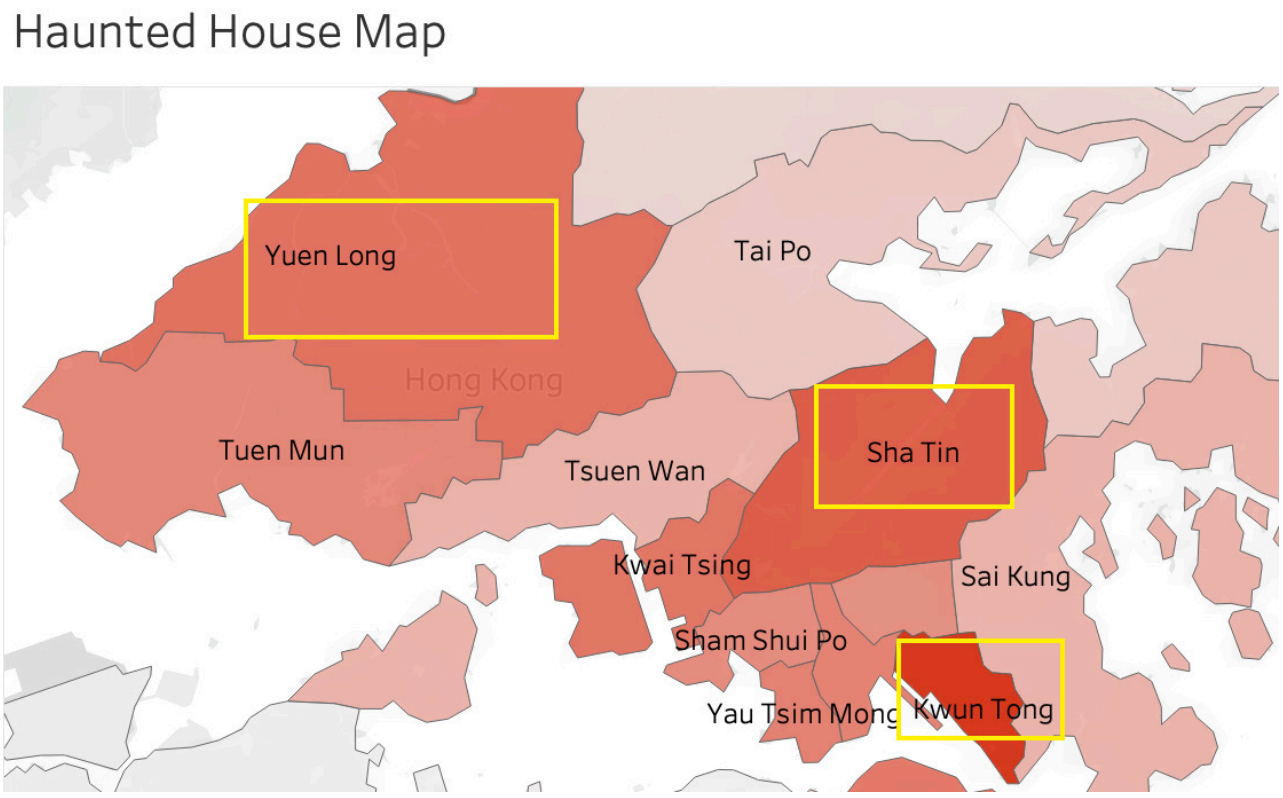
Graph4: The Population Density Bar chart at the bottom left shows the population density by districts. Based on the plot, we can see that Kwun Tong has the most population density with approximately one third higher than Sham Shui Po which is the second place. We can also see from the plot that Sham Shui Po, Yau Tsim Mong and Wong Tai Sin have roughly the same population density.

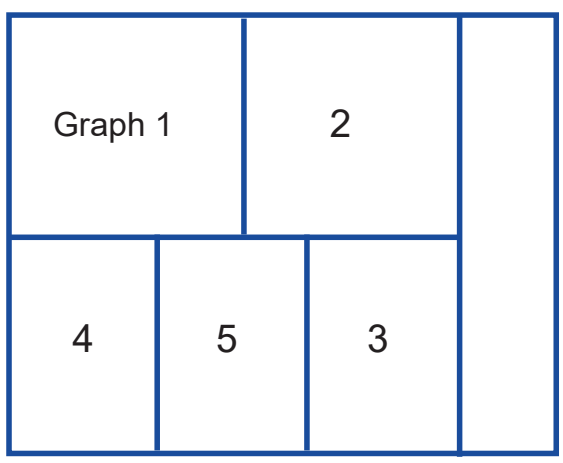
Graph5: The Population Bar chart at the middle bottom illustrates the population of different age groups by districts. Kwun Tong, Sha Tin and Yuen Long are the top three districts with the largest population. Compared with the Population Density Bar chart on the bottom left, we can see that although Sha Tin and Yuen Long have the second and third place in population, their population density ranking is nearly at the end.

From the Haunted House Map plot, we can see that the top three districts that have the largest number of haunted houses are Kwun Tong, Sha Tin and Yuen Long. By looking at the Population Bar chart, we find that those three districts have the most population as well. However, in the Population Density chart, the top three districts are Kwun Tong, Sham Shui Po and Yau Tsim Mong, which are different from the three districts that have the most population.

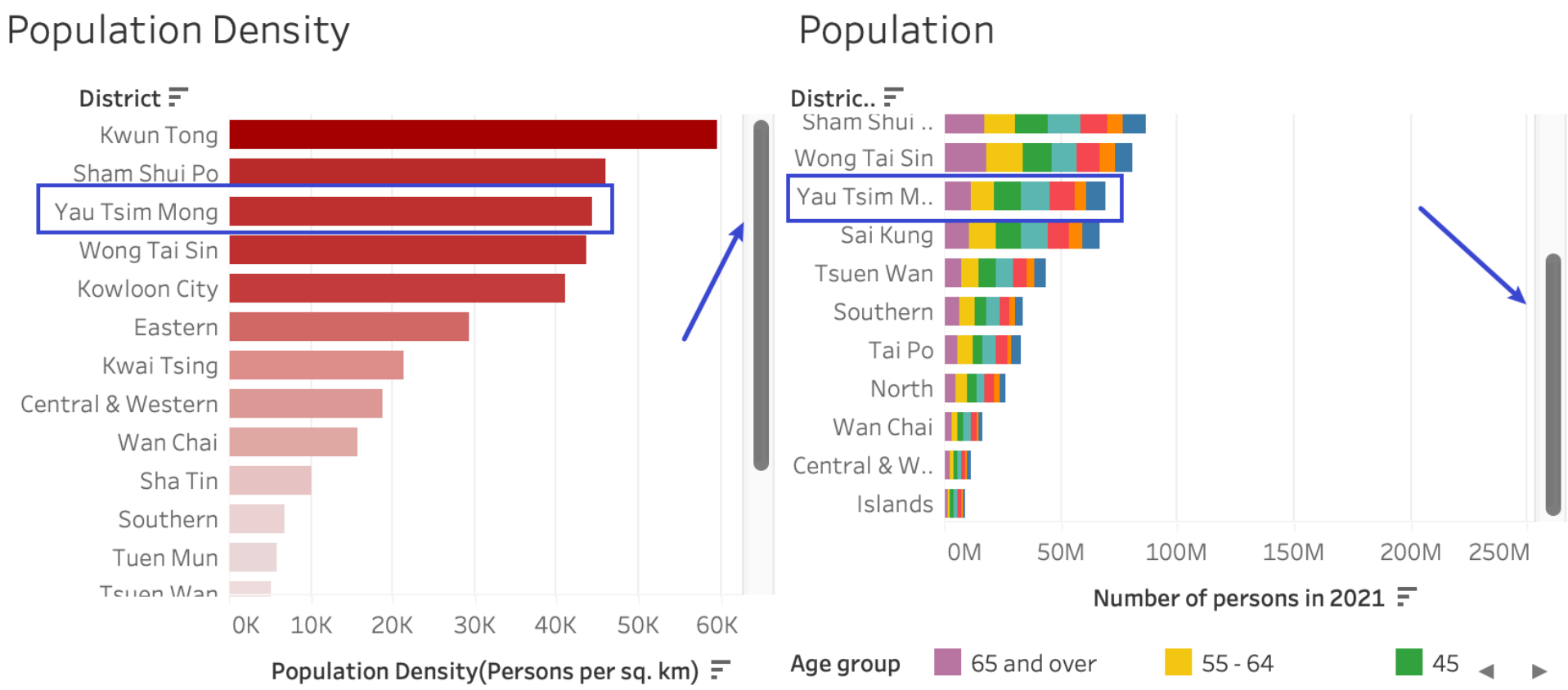
From there we can conclude that population does not necessarily have a positive correlation between population density.

Conclusion 17

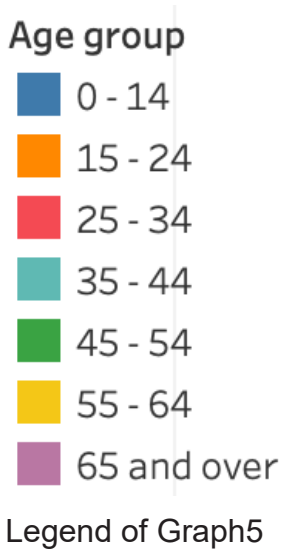




CHAPTER 3



Population and Population Density



Legend of Graph5

By looking at the Population, Population Density, and Proportion of House Haunted together, we find that although the population in Yau Tsim Mong is less than half of the district that has the most population, it ranks third in population density and ranks first in proportion of house haunted. Moreover, it is worth noting that the area of Yau Tsim Mong is the least among all districts[(15)] but ranks tenth among the eighteen districts.

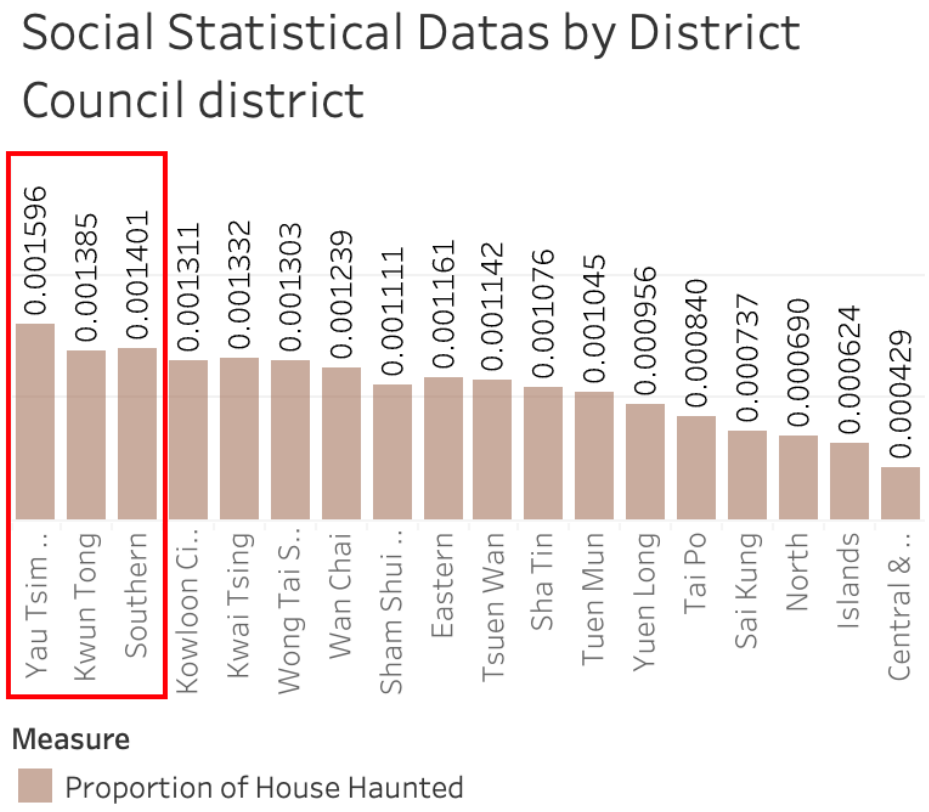
This means that the possibility of encountering a haunted house when walking around in Yau Tsim Mong is the highest among all districts.

——— *Conclusion 18*

The top three districts that have the most haunted houses and the top three districts that have the largest proportion of haunted houses are different. We think that the reason why they are not the same is because the proportion of haunted houses takes house supply into consideration. ***It illustrates that the proportion of haunted houses is probably a more accurate indicator of accessing public security of districts in Hong Kong.***

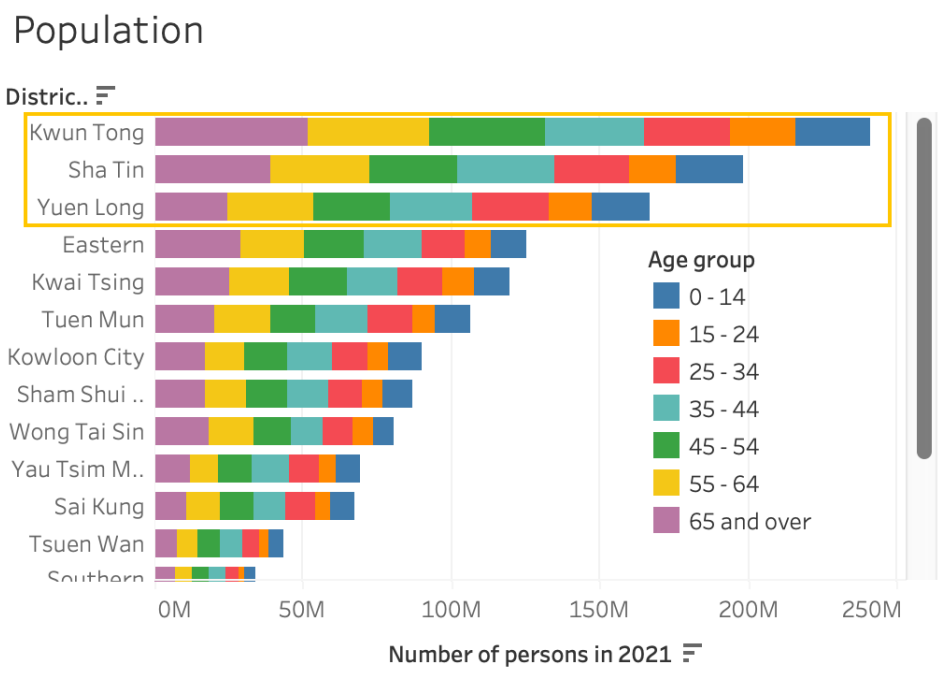
——— *Conclusion 19*

To further utilize our finding, we can say that from a customer point of view, to avoid the possibility of purchasing or renting a haunted house, we recommend the customer not to only pay attention to the houses in Kwun Tong, Sha Tin, Yuen Long, but also to pay close attention to the houses in Yau Tsim Mong as well. Moreover, from the peace officer point of view, we recommend adding more police forces to Kwun Tong and Yau Tsim Mong, as one district has the most population and the largest number of haunted houses, and the latter district has the most proportion of haunted houses.



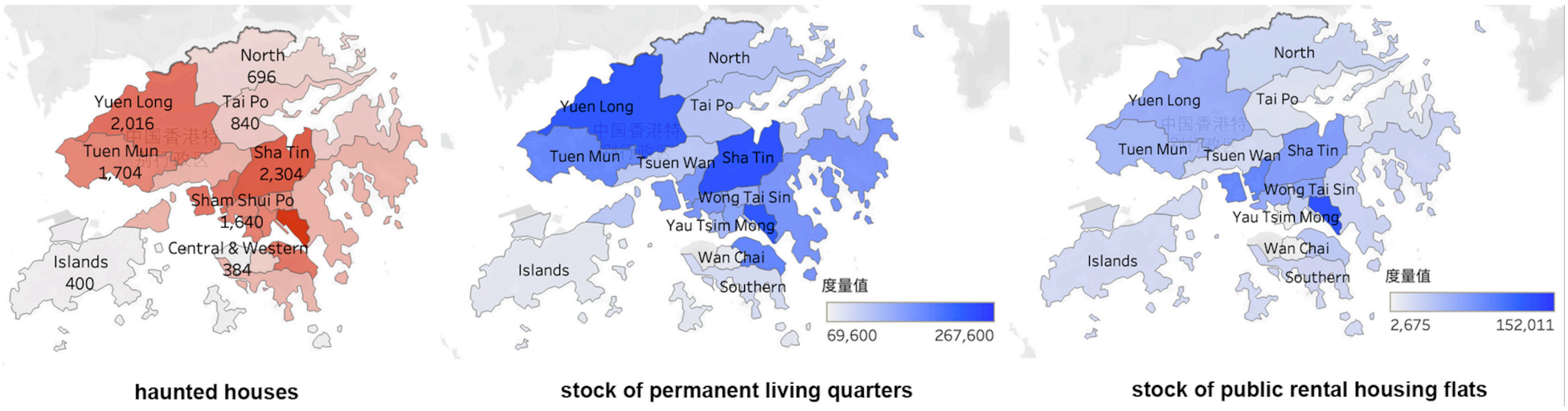
Haunted Houses Densities of 18 Districts

CHAPTER 3



By comparing the account of haunted houses and the age composition of population, we can find that the three districts with highest number of haunted houses also have the top three population sizes. Meanwhile, we can find that the number of haunted houses declines as the proportion of elderly people declines, ***from which we may infer that the social security status has certain relationship with the extent of population aging in the district .***

——— *Conclusion 20*



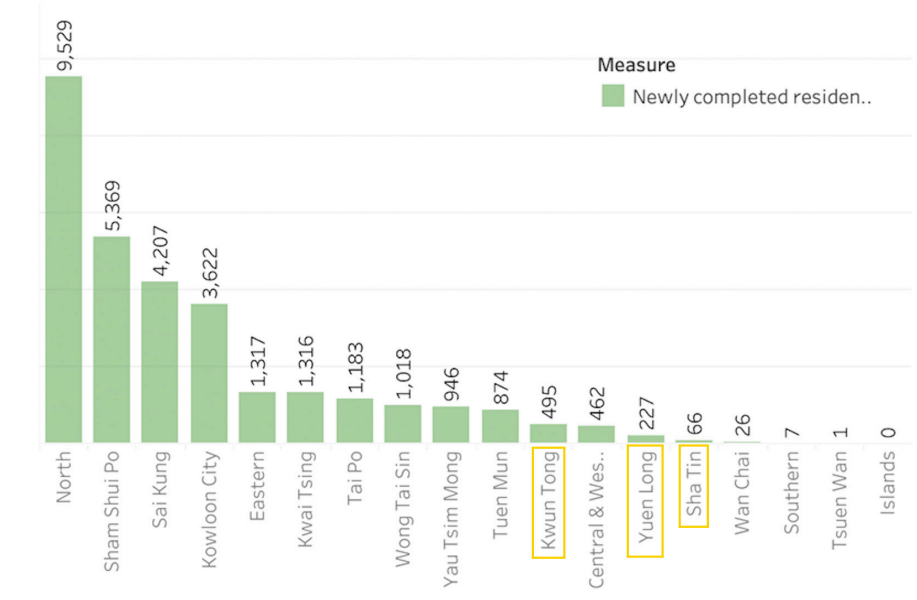
Density Maps of Haunted, Permanent, Rent Houses

With common sense, we assume that areas with a higher concentration of rentals will be less secure. ***By looking at the Social Statistical Data Map, we can see the stock of permanent living quarters is in consistency with the account of haunted house, but the pattern of stock of public rental housing flats is not consistent with our hypothesis.***

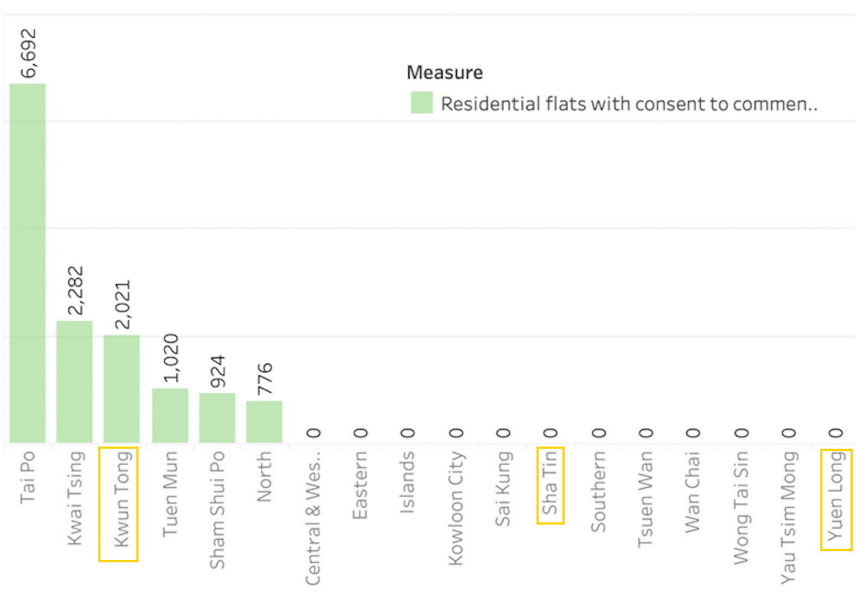
——— *Conclusion 21*

So there may not be a strong correlation between renting and policing. We suspect there is other factors. Or conversely, perhaps it is because the security is better, there are fewer haunted houses and people are willing to rent there.

When considering the new housing, we refer to the Newly completed residential flats and Residential flats with consent to commence work. In the years of the data coverage, even though Kwun Tong, Sha Tin and Yuen Long have the largest base in number of housing, they don't rank very high in the two indicators. This may have contributed to more indoor crime, as we would assume that more newly built houses may reflect better economic situation of the residents and districts, and that governments are more willing to solve people's housing problems there.



Newly Completed Residential Units



Under Developing Residential Units

4. Difficulties, Further Improvement and Limitation of the Analysis:

4.1 Difficulties that we have encountered:

- The topic we have chosen to analyse is one of the more complex social issues and there are very many potential influences, such as the ups and downs of property prices, population density, the level of well-being of the population and the stress of living.
- There are some problems in the analysis process where the data are difficult to distinguish and collate, for example, it is difficult to distinguish through figures whether the haunted houses are caused by suicide or violent crime; also some violent crimes or suicides occurring outdoors, not in buildings, are difficult to count. If we want to deepen the analysis, We may need to exploit web crawlers and natural language processing.
- There are some data that are not well correlated and do not meet our hypothetical expectations, for example, the relationship between the pattern of police officer numbers and crime rates in Hong Kong is not clear
- No suitable comparator was found and the comparison using the country and Hong Kong (region) lacks comparability, the data of metropolitans would be more suitable for comparison eg. Tokyo, New York.
- As for the data of people's livelihood, there are few similar studies and only limited government data is available from publications. It is difficult to preprocess and format data. Especially, because of the different granularity, it is difficult to make appropriate horizontal comparison between different types of data.

4.2 Anything we wanted to do but haven’t

- We tried to correlate census data with house price data and murder house data to infer the effect of house prices on the occurrence of murder houses, but abandoned this task because of the limited data available and the impossibility of aligning the period of house prices with the time of occurrence of murder houses.
- We expected to observe some noteworthy patterns from the data on people's livelihood, but due to the lack of data scale and time span, we could only provide a visual framework with simple analysis and further work is needed to draw more specific conclusions.

4.3 Any limitations to our visualizations

- We had intended to load all the dashboards onto a home-grown HTML site and broadcast it to the World Wide Web, but the interaction on the tableau became so laggy after the projection that we couldn't fix the problem.
- For chapter 3, we had intended to combine two maps into one and switch between them with the click of a button, but the fused data was too complex to relate, so we kept the two side by side maps.

D. Reference

(1) Crime and murder rates in Hong Kong, Hong Kong Crime Rate & Statistics 1990-2022, macrotrends.net. (https://www.macrotrends.net/countries/HKG/hong-kong/crime-rate-statistics)

(2) Comparison of crime rates with other countries, Crime Stats: compare key data on Hong Kong, nationmaster.com.(https://www.nationmaster.com/country-info/compare/Hong-Kong/United-States/Crime)

(3) Total number of crimes, nationmaster.com. (https://www.nationmaster.com/country-info/compare/Hong-Kong/United-States/Crime)

(4) World crime rate(2020), Google,(https://www.google.com)

(5) Suicide rate in Hong Kong, wikipedia. (https://en.wikipedia.org/wiki/Suicide_in_Hong_Kong)

(6) Overview of suicide rates from 1985 to 2016□ Kaggle. (https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016□

(7) Mental health and suicide rate□ Kaggle. (https://www.kaggle.com/twinkle0705/mental-health-and-suicide-rates?select=Crude%2Bsuicide%2Brates.csv)

(8) Human and financial resource of Hong Kong Police Force, macrotrends.net. (https://www.macrotrends.net/countries/HKG/hong-kong/gni-per-capita)

(9) Hong Kong's gross national income as economic development indicator, Hong Kong police office. (https://www.police.gov.hk/ppp_tc/09_statistics/)

(10) Persons arrested for crime by type of offence, age group and sex, Hong Kong police office. (https://www.police.gov.hk/ppp_en/09_statistics/csc.html)

(11) Population by gender and age group 2011, 2016, 2021, Census and Statics department. (https://www.censtatd.gov.hk/en/EIndexbySubject.html?scode=180&pcode=B1130303)

(12) Hong Kong Monthly Digest of Statistics from Census and Statistics Department of the Government of Hong Kong SAR, Census and Statics department. (https://www.censtatd.gov.hk/tc/EIndexbySubject.html?pcode=B1010002&scode=160)

(13) Haunted House Dataset, 852 House.(https://852.house/zh/haunted-houses?page=1)

(14) Hong Kong Government Census Dataset, data.gov.hk. (https://data.gov.hk/en-data/dataset/hk-censtatd-census21c-2021-population-census-statistical-reports)

(15) Hong Kong Land Area By District Council District, CEIC. (https://www.ceicdata.com/zh-hans/hong-kong/land-area-by-district-council-district)