

COMP7507 Visualization and Visual Analytics

Project Report

Visual Analysis on the Covid-19 Pandemic Recovery of Hong Kong

Group 7

Yi Heng 3036195382

Zhong Hao 3036196984

Wang Yilin 3036197354

Zhou Xingjian 3036195148

Contents

OBJECTIVES	3
HIGHLIGHTS OF WORK	3
DATASETS AND VISUALIZATION TOOL	3
DATASETS	3
VISUALIZATION TOOL.....	4
OFFICE: VACANCY, RENT, AND BEYOND.....	5
DATA DETAILS	5
DESIGN	5
VISUALIZATION.....	6
ANALYSIS.....	7
TRANSPORTATION.....	8
DATA DETAILS	8
DESIGN	9
VISUALIZATION.....	9
ANALYSIS.....	11
RETAIL SALES	11
DATA DETAILS	11
DESIGN	12
VISUALIZATION.....	13
ANALYSIS.....	14
LIMITATIONS AND DIFFICULTIES	15
CONCLUSION.....	15
CONTRIBUTION & WORK PLAN.....	16
REFERENCE	17
APPENDIX.....	18

Objectives

COVID-19 pandemic has impacted diverse aspects in the Hong Kong society since 2019. Whether this city has recovered to its pre-pandemic status is unobvious and difficult to observe. The concept of recovery is abstract, thus more concrete evaluations are needed.

Inspired by Loh et al. [1][2], this report measures the recovery vitality of Hong Kong by three key indicators: **1) office price, rent and vacancy rates, 2) public transportation ridership, and 3) retail sales**. The outcome of our work is a package of data visualization workbooks with necessary interactions and analysis on these materials.

Highlights of Work

1. Measuring the targeted topic from analysis in three comprehensive aspects;
2. Plenty of interactable visualization work;
3. Well-designed webpage with clear navigation.

Datasets and Visualization Tool

In this part, we will briefly describe the composition of our datasets, giving an overview of the data structure of the three indicators above. Also, we will discuss the visualization tool and methodology we chose.

Datasets

1. Office data

Office buildings are the most common location for work and thus can reflect the health of economy. There are various data related to office, among which we care the most are as follows:

- Rent (/m² per month)
- Price (/m²)
- Vacancy rate

The data above relies on seven regions of Hong Kong where the office buildings are centralized. We keep this feature in case there are potential discrepancy among different regions.

2. Transportation

Public transportation is an indispensable part of citizens' daily life. By analyzing the

public transportation ridership, we could have an insight about whether people have got out of the impact of COVID-19. Specifically, we collect following data for analysis:

- Public transportation ridership (thousands)
- Cross Harbor transportation ridership (thousands)
- Control points transportation ridership (thousands)

The range of data is mostly from 2019 to 2023. And to better analyze the relation with COVID-19, corresponding pandemic data is displayed as well.

3. Retail Sales

Retail sales can reflect the level of consumer spending, and consumption drives economic growth. Therefore, we choose retail sales as an important indicator to evaluate the vitality of Hong Kong's economic recovery. In the vast dataset of the retail industry, we will target the following keywords to seek more visualization results:

- Total value of retail sales (HK\$ million)
- Total value of retail sales by type (HK\$ million)
- Total value of online retail sales (HK\$ million)

To analyze the comparison of retail sales before and after the COVID-19, above data is from January 2019 to September 2023 when the project was conducted.

Visualization Tool

1. Tableau

Initially, we considered Plotly to proceed our work. However, we find it hard to merge our contributions together. In that case, we finally chose Tableau Cloud to conduct our visualization project because it can be easily used for collaborative development.

The Tableau [3] is also good at displaying data on a map – in this case the map of Hong Kong. Although Tableau only provides geographic roles limited to province-level, we can reduce the granularity by importing Hong Kong's space file. Tableau is also good enough for realize user interaction with simple operations on the workbook.

2. Webpage

We found that Tableau is good at designing each single dashboard. However, it lacks the ability to combine all our works together. The navigation object in Tableau dashboard cannot show the sub-structure of each part, so that we used webpages for a more flexible design.

As we have mentioned, web application distinguished from its unique flexible design ability, but it exists its own drawbacks for:

- Not simple enough for quick design one single visual dashboard.
- Need server to run applications.

These drawbacks may carry unexpected burdens except pure visual applications. To overcome such shortcomings, we simply use *layui* as our UI framework and Go-live extension in vscode for a quick server build. (Appendix 3 Final Work Webpage)

Office: vacancy, rent and beyond

Data Details

Office vacancy is an intuitive indicator to reflect the demand for office space in a particular area, and the demand has close relationship to the economy status. Office vacancy rate is defined by the proportion of office vacancy (the area of non-used office) and stock (the total area of office). Another related figure is the take-up which is given by adding the completions to the vacancy figures at the beginning of the year, then subtracting the year's demolition and the year-end vacancy figures. The data can be found at the official website of Hong Kong government.

However, we can only collect year-based vacancy data, which is not detailed enough to explore the potential tendency of the demand for office space. To evaluate the tendency better, we also collect the change of rent and price of office from 2019 to 2023, of which the time granularity reaches to month. The data of rent and price comes from 7 different regions (Sheung Wan , Central, Wan Chai / Causeway Bay, North Point / Quarry Bay, Tsim Sha Tsui, Yau Ma Tei / Mong Kok, Kowloon Bay / Kwun Tong).

The raw data used the regions' name as the column name of price and rent in 7 areas, which makes it hard to create relations between different data tables (for example, in the table of rent, column *Sheung Wan* represents the rent in Sheung Wan, while in the table of price, the same column name represents the price). In that case, we reshape the data table, adding new columns named price, rent and region, and use region as the connection between rent and price.

The data above are divided into 3 groups – A, B and C – representing the grades of the office buildings defined by the government. Grade A is the most advanced which means it is more expensive, while grade C is the least.

Design

The best way to show the changing trends of a figure is line chart, so we built three line charts for vacancy, rent and price respectively at first. Considering that there may be unexplored features buried in different grades, we used grades as color labels and showed the data with 3 lines in different colors, each of which represented a grade. However, there were missing data that make the fluctuation of line charts have little meaning, while what we care is actually the tendency. We then rebuild the line chart of rent into scatter plot and draw three trend lines to show the changing tendency.

For the office price, missing data also result in great fluctuation. We have to sum up the data, looking from a whole perspective to reduce the fluctuation. But we would like to preserve the grade attributes, so we chose stacked bar chart to show the contributions

of different grades.

With new attribute region added, it is possible to show the data on the map in Tableau. However, we failed to find Hong Kong space file with areas small enough to show our 7 regions. Finally, we have to obtain latitude and longitude information from Google Map and add it manually.

The office vacancy rate can be easily showed by line chart. By contrast, the take-up data involves comparison to the last year, which inspires us to use a bar chart similar to the stock graph.

Visualization

We develop two dashboards to show the features above with some interactions.

Figure 1 is the dashboard of price and rent. The map on top left shows the month average price and rent of 7 regions respectively in a particular year, where the scale of the circle represents for the rent and shade represents for the price. The displayed year can be changed by dragging the slide block under the map. The right side shows the change of rent and price from 2019-2023. The change of a certain region can be observed by clicking the circles in the map.

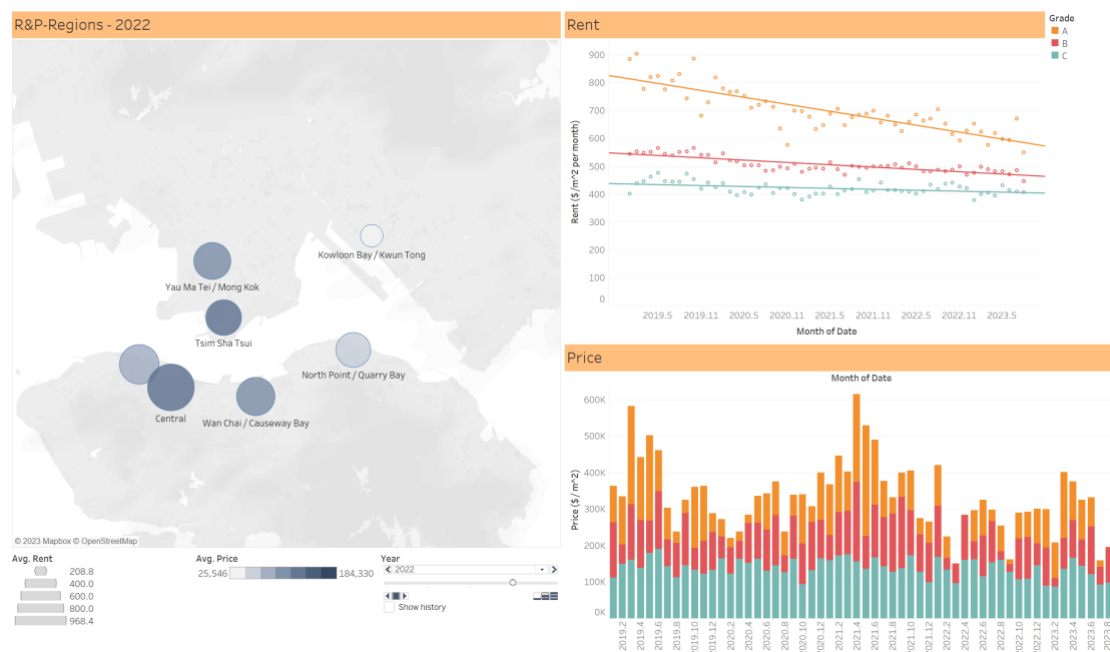


Figure 1. Dashboard of rent and price

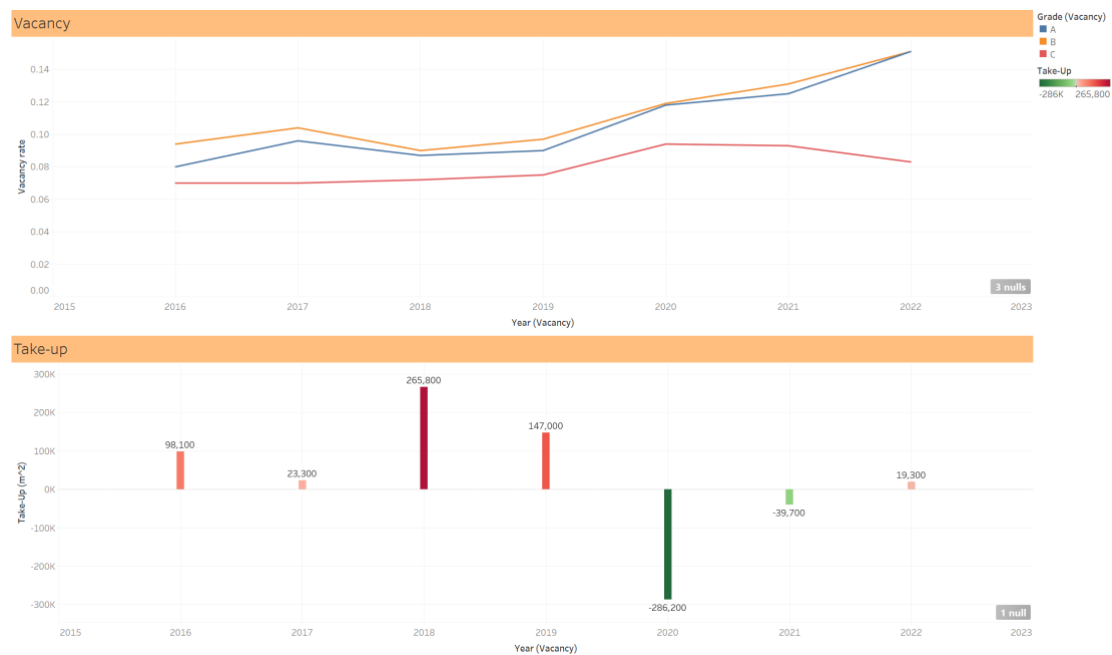


Figure 2. Dashboard of vacancy rate

Vacancy rate and take-up is displayed in Figure 2. The top line chart shows the vacancy rate of 3 grades, and the bottom bar chart shows the total take-up. The take-up of a certain grade can be shown by clicking on the lines from the top chart.

These two dashboards are published on the Tableau Public and can be viewed [here](#), or refer to the webpage covered in the project package.

Analysis

In Figure 1, the decreasing of rent during the pandemic period can be intuitively seen by the trend lines in the Rent-Date chart, especially for grade A, and the similar situation happens to all of the 7 regions (can be checked by clicking the circles). It seems that the impact of Covid-19 remains, indicating the economic downturn is still causing the drop of rent.

By contrast, the Price-Date chart does not show obvious relationship between price and the pandemic. One possible explanation is that the transaction of an office mainly relies on the market. That is, if the demand drops, the sellers are likely to not selling the office rather than lower the price, which agrees with the intuition.

The change of vacancy rate of grade A and B also agrees with the intuition (see Figure 2), but what seems abnormal is grade C. During the pandemic, the vacancy rate of A and B both increases, while the counterpart of C drops. In theory, the vacancy rate of all 3 grades should have increased, because the number of people who can go to work or get a job should have decreased. After analyzing, we get another way to observe this issue: people's demand for an office space cannot be easily removed, since grade A and

B has higher price and rent, people may tend to choose cheaper grade C. This is another proof of the relatively unhealthy economy.

On the other hand, the take-up's trend is quite obvious. It can be seen that it had an extremely sharp drop in 2020, when the pandemic broke out and spread fast. After 2020, the take-up does recover a little, but haven't return to its pre-pandemic status.

In conclusion, from the aspect of office, the impact of the pandemic still exists. All of the indicators in this part do not show recovery or do not recover to their initial conditions.

Transportation

Data Details

The data is provided by the Transportation Department of the Hong Kong Government, which could guarantee the accuracy and integrity of the data.

- **Public transportation ridership (thousands)**
In this section, we collect data of different transportation types to get a comprehensive insight about the situation. For easy understanding, three major transportation types are picked, which are railway, ferry and franchised bus. Since the total counts of these transportation types distribute in a quite wide range, the ridership is summarized in thousands to generate a rational axis scale.
- **Cross Harbor transportation ridership (thousands)**
For this part, transportation ridership of crossing harbors is classified into different transportation types to demonstrate details. The amount is also displayed in thousands.
- **Control points transportation ridership (thousands)**
For this part, we represent ridership of passing different control points, such as Shen Zhen Bay and Sha Tau Kok. Also, we collect ridership about arrival and departure at each control point to observe trend during the pandemic period.
- **COVID-19 confirmed case**
This data is collected to visualize the correlation with the transportation. We fetch data from the Department of Health. And the time range of the data is from 2020 to 2022.
- **Timespan**

To cover the pandemic period and compare the difference before and after COVID-19, the time range is generally from Jan. 2019 to Sep. 2023. Furthermore, the period is split into three sections to benefit the comparison:

- ◆ Jan. 2019 – Dec. 2019: Pre-pandemic period.
- ◆ Dec. 2019 – Dec. 2022: In-pandemic period. The landmark event is that several unknown pneumonia cases were reported from Wuhan in Dec. 2019.
- ◆ Jan. 2023 – Sep. 2023: Post-pandemic period. The landmark event is that the National health Commission announced that COVID-19 would be under Class B control from Jan. 2023.

For more details, the time scale is clustered quarterly to keep a balance between precision and workload. Due to delayed update and loss of historical documents, data at some specific time points is unreachable, we will try out most to mitigate such problem in the following analysis.

Design

After the data preparations, we should consider how to arrange and visualize the data to achieve our objectives. Given the content and the structure of the data, analyzing the data from the intra-city aspect and the inter-city aspect could be an acceptable choice.

The intra-city transportation indicates ridership among Hong Kong, which could directly reflect the impact of COVID-19 on people's daily activities. To visualize this sub-topic, publication transportation data and cross harbor transportation data are required.

The inter-city transportation indicates arrival as well as departure ridership of Hong Kong and it requires control points transportation data. As the center of finance and commerce in the world, the recovery of Hong Kong unavoidably concerns with the other cities and countries. Therefore, this part of analysis essential for the integrity of the conclusions.

Among diverse kinds of charts, line chart is appropriate to illustrate the trend of ridership according to quarters. Also, pie chart could compare the relative portion in pre-pandemic period and post-pandemic period to visualize the recovery status.

Visualization

With all preparations done, the visualization results are demonstrated as follows. Figure 3 demonstrate the situation of intra-city transportation. The line charts represent quarterly ridership in thousands. And different lines indicate different transportation types. Filters for transportation types and period are provided to go into details about the charts.

Visual Analysis on the Covid-19 Pandemic Recovery of Hong Kong

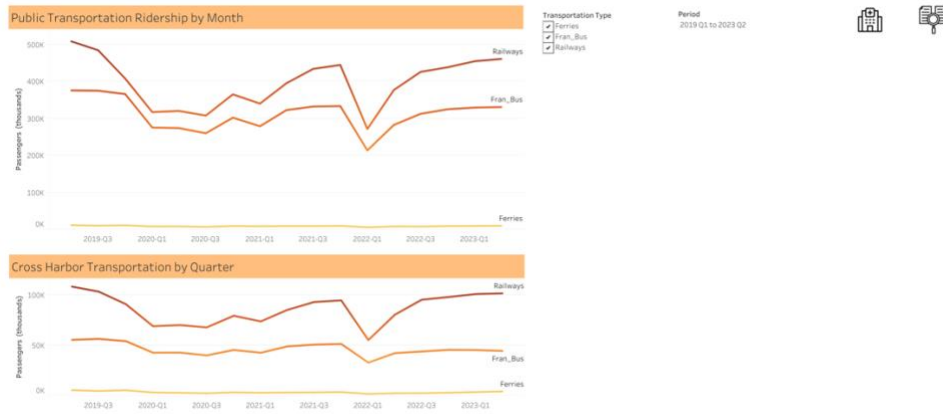


Figure 3. Dashboard of Transportation_intra1

By clicking the two icons in the upper side, the chart of COVID-19 confirmed cases and the chart of comparison between 2019 and 2023 are shown in the right side as Figure 4. These two charts are initially hidden to emphasize the content for different points at the story line.



Figure 4. Dashboard of Transportation_intra2

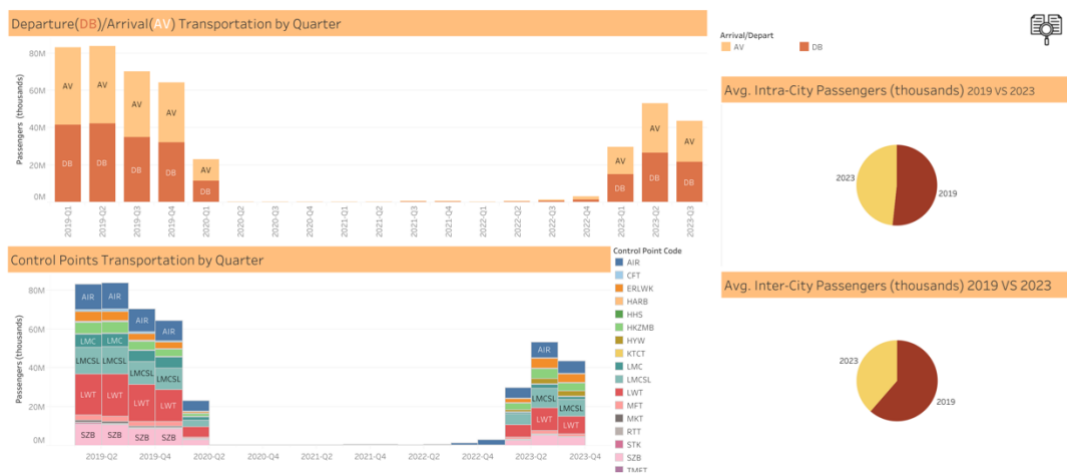


Figure 5. Dashboard of Transportation_inter

Then Figure 5 represents inter-city transportation from 2019/Q2 to 2023/Q4. Two charts in the left side show control points passengers. The pie charts for comparison are hidden at the beginning and are controlled by the icon in the top left corner.

Analysis

According to Figure 3, we could see that there is an obvious decrease in the first quarter in 2022. This could be explained by the outbreak of COVID-19. Data of confirmed cases increases in the corresponding period. Then the data decreases in the second quarter, followed by intra-city transportation ridership increasing from the same time point. Additionally, we could find that the relative numerical relationships of these three types of transportation are stable during the whole period.

With respect to the inter-city transportation, we could clearly see that there is almost no ridership between 2020 to 2022. This is because the Hong Kong government took a very strict policy about passing the port. Besides, the ratio of departure and arrival remains consistent during the period although the amount sharply decreased in the pandemic period. As for the ridership distribution, we could find that the proportions of some main control points such as Airport, Shenzhen Bay, Lok Ma Chau Spur Line remain almost unchanged.

Now back to the recovery status of Hong Kong, we would like to conclude that for the intra-city transportation, the ridership is slight lower than that before the pandemic. However, for the inter-city transportation, the data is obvious lower than that in 2019. The recovery is worse than that of the intra-city transportation. Considering the recovery of the whole world, the government is more cautious about worldwide communications.

Back to the complete process, the method of combining different charts achieves to build a bridge between transportation and COVID-19. The ridership trend during the pandemic is intuitively displayed and the comparison is effectively to illustrate the recovery of two aspects about ridership.

Retail Sales

Data Details

We got retail sales dataset from Census and Statistics Department of The Government of the Hong Kong Special Administrative Region, since the retail data provided on the official government website are of various types and complex content, for example, there are many key indicators in the original data set provided by the government, such

as operating expenses, employee compensation and other indicators that are not closely related to our topic. Therefore, the sets were processed and filtered to obtain the most useful information for our topic visualization, here are the main dataset tables:

Total retail sales:

- Total value of retail sales (HK\$ million)
- Date (Year/Month)

Online retail sales:

- Total value of online retail sales (HK\$ million)
- Date (Year/Month)

Online retail sales refer to the sales of goods to customers through computer networks specially operated by local retail establishments for the purpose of receiving or placing of orders. The goods are ordered by those methods, but the payment and the ultimate delivery of goods do not have to be conducted online.

Retail sales Category:

- Total value of retail sales (HK\$ million)
- Main type of retail outlet
- Detailed type of retail outlet
- Date (Year/Month)

The classification of retail establishments follows the Hong Kong Standard Industrial Classification Version 2.0[4] as from the reference month of January 2009. To facilitate analysis of the short-term business performance of the local retail sector, the industry classes of the retail sector have been grouped into the 19 retail outlet types (Appendix 1 Retail Outlet Types), taking into account their importance in the retail sector.

Design

Based on the above dataset, we design the retail section from three main visualization panels: the total retail sales, the total retail sales of online retail sales, and the total retail sales classified by specific type.

First, for the total retail sales, we would like to observe the overall change from 2019 to 2023 on the visualization results, and also get an overview on the special number of retail sales, so we choose a bar chart with different colors by different year.

While only observing the absolute value does not give us enough information, and 2019 as a year in which the epidemic has not yet occurred, which is used be a baseline to obtain the relative value for the next four years, so we can get the change of retail sales value in following four years due to the impact of the epidemic. To achieve this, it is better to use a line chart of the percentage difference corresponding to 2019 and use the same color for the fixed years to make the panel more harmonious.

Then it comes to the total value of online retail sales, we use a mix of line and bar charts, which gives us a clear view of both the specific numbers of online sales data and the year-to-year changes in online sales value.

The total retail sector can be divided to various economic sectors, so it is essential to get a more detailed view to further look at the change in sales in the segmented categories. Although pie charts can provide a good proportion of the display, the tree map can not only show the proportion, but also reflect the proportion of sub-categories under the main classification, besides it can better utilize the limited space of the dashboard, so we choose the tree map in this retail-type sales. We also add table and packed bubbles to the dashboard for the detailed type.

Visualization

After our design, we get the following visualization, as shown in Figure 6, this shows how overall sales have changed over the last five years. On the top left is a bar graph presentation, on the bottom left is a line graph of the percentage difference relative to 2019, and on the top right side, it is a pie chart of the retail sales value category share for a specific month following the change in the left side of the data, where you can clearly see how the retail sales value by different types has changed during the pandemic time. It is a mixed plot of the change in online retail sales on the bottom right.

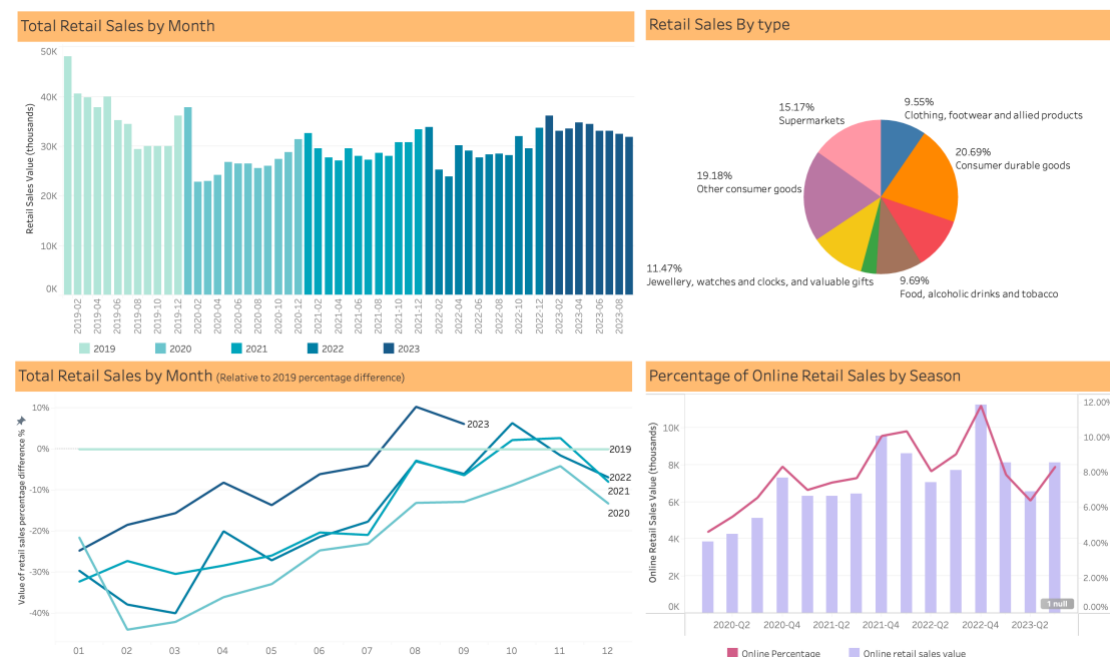


Figure 6. Dashboard of Total Retail Sales

In Figure 7, this shows the total retail sales value by detailed Type. First, the upper is a bar chart of year-to-year sales changes for different main categories, where we have divided the main categories with different unique colors. On the left side of the bottom is a tree map, where the main categories are differentiated by the same colors as on the

top, and different small chunks of the same colors indicate the percentage of the detailed sub-category. On the right side of the bottom is a combination of table and packed bubbles. The table shows the textual data of the detailed type, and the bubbles show the amount of data in the subcategories, where we use a different color from the main categories to avoid confusion.

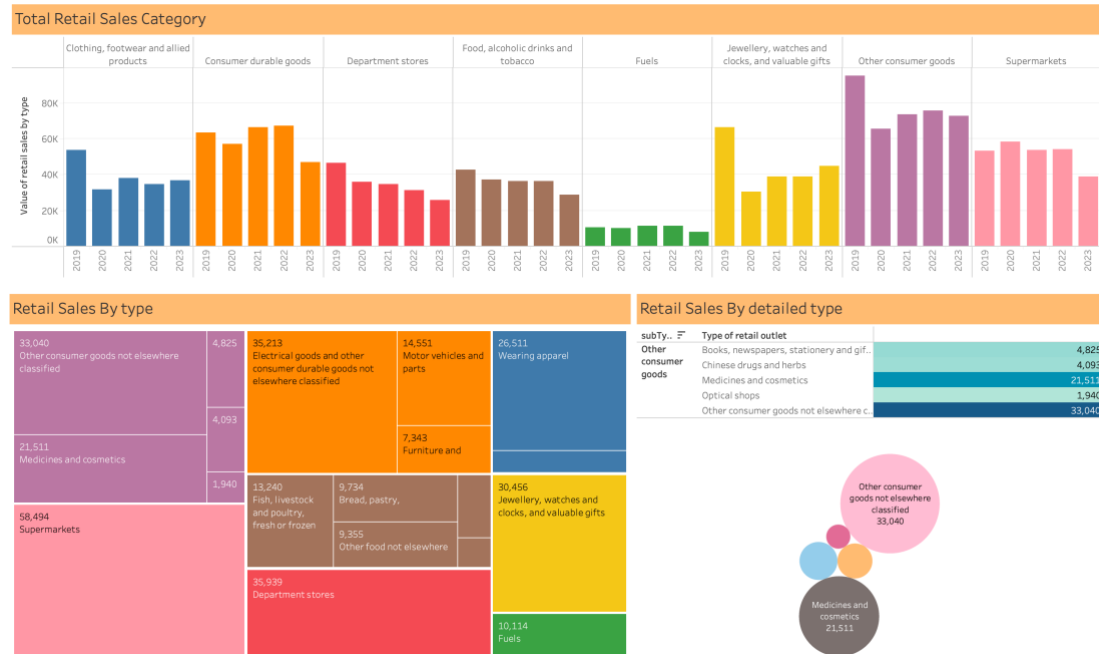


Figure 7. Dashboard of Total Retail Sales by Detailed Type

Analysis

From the visualization of Figure 6, we can clearly see that retail sales have experienced a downward trend from 2019 to 2021, and there is a small rebound to 2023 today. And for the percentage difference relative to 2019, and we can see that in 2020, the worst year of the covid-19 outbreak, total retail sales were all lower than 2019 for each month, and it is obvious that the total sales of 2023 has exceeded 2019 since August, based on a review of the last two years in 2021 and 2022, we can predict that this year's October November scenario will also be better than 2019. At the same time, we can look at retail sales by type, in the last four years of overall sales, the overall proportion of various categories has not changed much, and the most has been other consumer goods.

As for the online retail sales, from this chart, we got a small peak in online sales in the fourth quarter of 2022, which could be due to the epidemic opening and the development of online shopping Live-streaming.

Then for the dashboard in Figure 7, from the total retail sales category bar chart, we can find that the total sales of fuels have remained constant as it is a necessity of life, however, for the non-essential part like jewelry, watches, and clocks, people's enthusiasm for buying which dropped significantly during the worst of the epidemic.

Then we will want to analysis more detailed type. Other consumer goods consist of medicine, Chinese drugs and herbs, and other goods not classified, since it is a combination of lots of types and during the epidemic, it's no surprise that sales of other consumer goods get the highest percentage in the main type of retail sales.

In conclusion, in terms of this aspect of retail sales, it is true that in recent years there has been a significant downward trend in retail sales after the epidemic, however, since August 2023, there has been a gradual recovery and even surpassing of the year 2019.

Limitations and difficulties

As for the difficulties we have met during this project, incomplete data has hindered our process temporarily. For one thing, some earlier data is missed so that we have to shorten the time range to align the start point and the end point of the comparisons. For another thing, some data of recent months has not been updated yet. Thus, we switch to demonstrate the average value rather than the total value.

Besides, we attempt to eliminate the confusion caused by the abbreviations of harbors and control points. Finally, we decide to describe them briefly in the presentation and add labels beside charts.

Considering the limitations of the visualization tools we choose, one shortage of the Tableau Cloud is that, although we can share our dashboards, we cannot edit the same dashboard at the same time, like Git. The Tableau will update the whole dashboard, not the difference we have made. So, we schedule to edit our dashboards separately and merge them together.

In a brief, we have achieved our initial goals. However, if time permits, we would like to investigate the underlying relations among these three indicators to make the analysis more systemic.

Conclusion

In general, our conclusion is that Hong Kong has started to recover from the COVID-19, but still not reached the base line before the plague yet. For the aspect of vacancy, this indicator of those office buildings graded in “A” and “B” are increasing while the rate of buildings graded in “C” is decreasing during the COVID-19 period. This is probably because people may not afford the cost of those more advanced, expensive, and highly equipped buildings and tend to choose those building graded in “C”. In the transportation’s perspective, due to the traffic control policy, the inter transportation for Hong Kong nearly falls to one in a thousand compared to data of the pre-pandemic

period. Even we counted the data after COVID-19, the picture still shows a weak view compared to the pre-pandemic data. But it is a good signal that the data recovered in a rapid way, which may infer that the economy of Hong Kong has gained its motivation. As for retail sales, we set the total retail sales by month in 2019 as base line, and surprisingly found that even the data are below in the most of time during COVID-19, the data after August in 2023 are all above the base line. All these clues lead to our conclusion.

For the further work, we took three aspects for our project, it is obvious that there are extra indicators which are also suitable for evaluating the economy in Hong Kong. What is more, due to the time after Hong Kong government cancel the control policy is no more than 10 months, which may not be sufficient for our conclusion, and further work and data are necessary.

Contribution & Work Plan

	Division of Labour	Oct 8-15	Oct 16-23	Oct 24-31	Nov 1-8	Nov 9-16	Nov 17-24	Nov 25-30
Data Preparation	All	X	X					
Analysis of Office Vacancy Rates Rent and Price Vacancy	Yi Heng		X	X	X			
Analysis of Public Transportation Ridership Intra_city Transportation Inter_city Transportation	Wang Yilin		X	X	X			
Analysis of Retail Sales Total Retail Sales Total Retail Sales by detailed type	Zhong Hao		X	X	X			
Find out associations between the recovery rate and various factors	Zhou Xingjian			X	X	X		
Finalize the demo	All					X	X	
Finish the report	All					X	X	X

Reference

- [1] Brail, S. (2022) ‘Covid-19 and the future of Urban Policy and Planning’, *Current History*, 121(838), pp. 298–303. doi:10.1525/curh.2022.121.838.298.
- [2] Leong, M. et al. (2023) Can we save the downtown? examining pandemic recovery and polycentricity trajectories across 62 North American cities [Preprint]. doi:10.2139/ssrn.4456917.
- [3] Team, D. (2019) Top tableau case studies - JP morgan, Lenovo & Lufthansa, DataFlair. Available at: <https://data-flair.training/blogs/tableau-case-studies/> (Accessed: 08 October 2023).
- [4] Hong Kong Standard Industrial Classification (HSIC) Version 2.0. Available at: <https://www.censtatd.gov.hk/en/EIndexbySubject.html?pcode=B2XX0005&scode=452>

Appendix

Retail Outlet Types

<u>Type of retail outlet</u>	<u>Coverage</u>
Fish, livestock and poultry, fresh or frozen	Retail outlets selling fish, other sea products, livestock and poultry, fresh or frozen.
Fruits and vegetables, fresh	Retail outlets selling fresh fruits and vegetables.
Bread, pastry, confectionery and biscuits	Retail outlets selling bread and cakes, confectionery, biscuits, cookies and egg rolls.
Other food not elsewhere classified	Groceries of general provisions, groceries of Chinese provisions and retail outlets selling fish and other sea products, dried or preserved; meat, roasted, dried or preserved; preserved provisions and spices; rice; noodles and rice sticks; bean curds and bean products; eggs; non-alcoholic beverages in specialised stores; beverages (include alcoholic and non-alcoholic) and tea leaves and other retail outlets selling specialised food without seats.
Alcoholic drinks and tobacco	Retail outlets selling alcoholic beverages, tobacco products in specialised stores.
Supermarkets	Supermarkets and convenience stores.
Fuels	Petrol filling stations and retail outlets selling firewood, charcoal, coke and similar fuels, kerosene and L.P. gas.
Wearing apparel	Boutique shops and retail outlets selling garments, sportswear, evening dresses, fur clothing and accessories like gloves, hats and leather belts, etc.
Footwear, allied products and other clothing accessories	Retail outlets selling footwear, fabrics, tailoring accessories and other clothing, footwear and allied products.

Control Point Code

<u>Control Point Code</u>	<u>Description</u>
KTCT	Kai Tak Cruise Terminal
LMC	Lok Ma Chau Control Point
LMCSL	Lok Ma Chau Spur Line Control Point
LWT	Lo Wu Control Point
MFT	Macao Ferry Terminal
MKT	Man Kam To Control Point
RTT	River Trade Terminal
STK	Sha Tau Kok Control Point
SZB	Shenzhen Bay Control Point
TMFT	Tuen Mun Ferry Terminal

Final Work Webpage

