**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](#_Toc152165088)

[﻿Highlights of Work 3](#_Toc152165089)

[Datasets and Visualization Tool 3](#_Toc152165090)

[Datasets 3](#_Toc152165091)

[Visualization Tool 4](#_Toc152165092)

[Office: vacancy, rent and beyond 4](#_Toc152165093)

[Data Details 4](#_Toc152165094)

[Design 4](#_Toc152165095)

[Visualization 4](#_Toc152165096)

[Analysis 4](#_Toc152165097)

[Limitations and diffculites 4](#_Toc152165098)

[Transportation 5](#_Toc152165099)

[Data Details 5](#_Toc152165100)

[Design 6](#_Toc152165101)

[Visualization 6](#_Toc152165102)

[Analysis 7](#_Toc152165103)

[Limitations and difficulties 7](#_Toc152165104)

[Retail Sales 8](#_Toc152165105)

[Data Details 8](#_Toc152165106)

[Design 8](#_Toc152165107)

[Visualization 8](#_Toc152165108)

[Analysis 8](#_Toc152165109)

[Limitations and diffculites 8](#_Toc152165110)

[Conclusion 9](#_Toc152165111)

[Contribution & Work Plan 9](#_Toc152165112)

[Reference 9](#_Toc152165113)

# ﻿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., 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. ﻿price, rent and vacancy rates,
2. public transportation ridership, and
3. retail ﻿sales

# 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.

1. 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.

1. Retail Sales

(Overview. Can follow the above for reference. Details can be given in the later part)

## 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 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.

1. Web…?

# Office: vacancy, rent and beyond

## Data Details

## Design

## Visualization

## Analysis

## Limitations and diffculites

* *Effectiveness of the* *visualizations (e.g., how you use them to develop hypotheses and understand/analyse data/problems)*
* *Different methods that you have tried and justification of your choices*
* *Difficulties that you have encountered, if any*
* *Anything you wanted to do but haven’t? Why?*
* *Any limitations to your visualizations and/or the tools that you used?*

# 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 harbor codes 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.

Chart 1 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.

图表

描述已自动生成

Chart 1

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 Chart 2 . These two charts are initially hidden to emphasize the content for different points at the story line.

图形用户界面

描述已自动生成

Chart 2

Then Chart 3 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 is controlled by the icon in the top left corner.

图表, 条形图, 瀑布图

描述已自动生成

Chart 3

## Analysis

According to Chart 2, 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 radio of departure and arrival remains consistent during the period although the amount sharply decreased in the pandemic period.

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.

## Limitations and difficulties

As for the difficulties we have met during this section, 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 amount 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 between transportation and other two indicators to make the analysis more systemic.

# Retail Sales

(Suggestion: develop this part from 4 subtitles: 1) detailed data structure (the columns), 2) design and how you came up with the design, 3) final outcomes (basically your tableau work, just put some images here and briefly describe them) and 4) analysis)

## Data Details

## Design

## Visualization

## Analysis

## Limitations and diffculites

* *Effectiveness of the visualizations (e.g., how you use them to develop hypotheses and understand/analyse data/problems)*
* *Different methods that you have tried and justification of your choices*
* *Difficulties that you have encountered, if any*
* *Anything you wanted to do but haven’t? Why?*
* *Any limitations to your visualizations and/or the tools that you used?*

# Conclusion

# Contribution & Work Plan

Overview of tasks completed by each of the team members (details fo tasks should go to the individual reports from members)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Oct**  **8-15** | **Oct**  **16-23** | **Oct**  **24-31** | **Nov**  **1-8** | **Nov**  **9-16** | **Nov**  **17-24** | **Nov**  **25-30** |
| Data Preparation | X | X |  |  |  |  |  |
| Analysis of Office Vacancy Rates |  | X | X | X |  |  |  |
| Analysis of Public Transportation Ridership |  | X | X | X |  |  |  |
| Analysis of Retail Sales |  | X | X | X |  |  |  |
| Find out associations between the recovery rate and various factors |  |  | X | X | X |  |  |
| Finalize the demo |  |  |  |  | X | X |  |
| Finish the report |  |  |  |  | X | X | X |

# Reference