Hong Kong Labour Market Data Visualization

Fong Chun Him (3035377115), Wu Szu Han (3035562617)

November 17, 2020

1 Introduction

Labour market is an important determinant of economic development. In a context of complex and dynamic labour market, gathering intelligence on labour market data, including but not limited to employment and vacancies statistics, labour treatment, and skills needs, can identify and understand issues in the labour market, inform decisions about employment policies, evaluate the benefits and costs of policies, match supply to current demand for skills and help enterprises and workers adapt to change.

This study discusses various concerns of school graduates entering the labour market, policy makers, and educational institutes. We provide data visualization on Hong Kong labour market data to provide clear pictures of the current labour market situation and answer the following questions:

- What is the best industry and occupation for one to enter as a fresh graduate, based on his/her skills and interests, as well as the market conditions (saturation level and salary)?
- How would one select a program of study in school that can meet the market demand and his/her personal interests?
- Which industries and occupations have the brightest future and, therefore, requiring educational institutes to direct more teaching resources and to encourage student enrollments in?
- Which industries and occupations have the highest potentials and can be boosted further with policies to attract companies and talents?
- Which demographic groups (age, education level, sex) are the most disadvantaged and will need government assistance on improving the industry outlook and plans for re-training of the current labor force?
- What types of training programs are most beneficial to the workforce and students?

The contribution of this study is to create exploratory visualizations for school graduates entering the labour market, policy makers, and educational institutes to discover labour market trends and make decisions responsively.

2 Methodology

2.1 Data and Analysis

1. Employment statistics and vacancies of different industries provide insights into job demands. One can infer from the trends of the past years to see which industries are expanding and vanishing. Recruiters, employees and students can learn more about the job opportunities to make decisions about human resources or job seeking, and adjust their expectations on salaries.

   Graduate statistics of different broad programmes provide views into the labor supply of different industries. Supply and demand can be compared to provide valuable insights. For example, students can search for industries with growing demands that are not yet saturated. The government can support and develop industries with growing demand. Institutions can launch new courses to meet the needs.

2. Employment earning statistics, including monthly wage of employees by educational attainment, sex, age group, occupational group and industry section. Students will see the wages of different industries and the effect of educational attainment, to help decide on pursuing further studies. For employers, the statistics are key information to benchmark salary to attract and maintain their labour force.

3. Hours of work is an important factor for students to decide their career path. It is also a key reference for government to set policies to help overworked employees.

4. Training needs of economically active people in different industries gives insight on the most important types of training and languages for each industry. This is crucial for students to understand which skills are in demand for the industries. Government and institutions can use this to determine the trainings to offer.

2.2 Choices of Visualization and Justification

1. An animated squarified treemap where it divides the industries recursively by industry section and industry division and rectangle size is the number of people engaged in an industry. Each node summarizes all the employment statistics, such as headcount and vacancies of the corresponding industry. Treemap is chosen so that a user can browse the relevant data of an industry easily. He/she can then see a full summary conveniently without investigating multiple charts. Moreover, rather than using multiple side-by-side treemaps, an animated treemap is used to visualize the trend of employment composition much more concisely. In addition, the size of the outermost rectangle is fixed so that the size of each inner rectangle represents the proportion of the industry section or industry division to all other industries. As we move along the timeline, we can see the changes in proportion of different industries at once. Note that squarified tiling method is used to maximize the visibility of small rectangles.

   A side-by-side bar chart grouped by industries and each group contains a bar for the industry headcount and a bar for UGC-funded programme graduate headcount. Whether graduate supply grows with actual demand can then be seen. It helps educational institutions to evaluate their programmes. A slide bar is provided to dynamically change data by year.

A centered stacked bar chart to show headcount by industries; two genders compose a bar. The chart is then sorted by percentage of male. In this way, industries in which females are significantly more than males or males are significantly more than females can be highlighted easily. For instance, if the pink bar representing females is of larger proportion, females are more than males. This visualization helps understand the work environment of different industries.

A side-by-side stacked bar chart to show the vacancies of different industries, where different industries compose a bar and bars of different years are put aside. The stacked bars are internally sorted by number of vacancies. This method of visualization allows users to easily identify which industries have the most job opportunities and how total job opportunities differ by years.

2. Simple line charts to show the median monthly wage by sex, age group, educational attainment, occupational group and industry section respectively, where different lines represent different groups and x-axis represents years. Since the number of groups of the charts for sex, age group, educational attainment and occupational group is small, a simple line chart is concise and intuitive. For occupational group, it is believed that there will not be complicated patterns in many industries and therefore using a line chart consistently is still clear.

3. A heatmap to show the working hours, where x-axis represents different industries, y-axis represents the years and different pages contain data for both gender, male and female respectively. Cells are coloured by hours of work. The data indicated that the working hours have improved since the past years for some industry sections such as retail, accommodation and food. The grids will become cooler in more recent years. Since a heatmap reveals gradual patterns quickly and concisely, it is used for the sake of showing gradual patterns. We successfully show patterns along the x-axis and along y-axis.

4. A packed bubble chart to show the proportion of employed persons with training needs in different occupations. The bubbles are in different colours, corresponding to employed persons with different training experiences arranged by employers. Size of bubbles is the number of employed persons who cited and considered the training to be the top 5 important. This shows the importance of training in each occupation easily. The ratio between total training attendance and number of all employed persons in a particular occupation is also shown. For example, it can indicate high training resources demand if there's a high percentage of people needing training in a large occupation.

3 Source of data

All data are prepared and provided by the Census and Statistics Department of Hong Kong Government. All data are well-formatted in csv files, clean, annotated and almost complete with only a few missing data entries. Hundreds of excel tables are inside our scopes and will likely be grouped together and included in our visualization.

4 Results

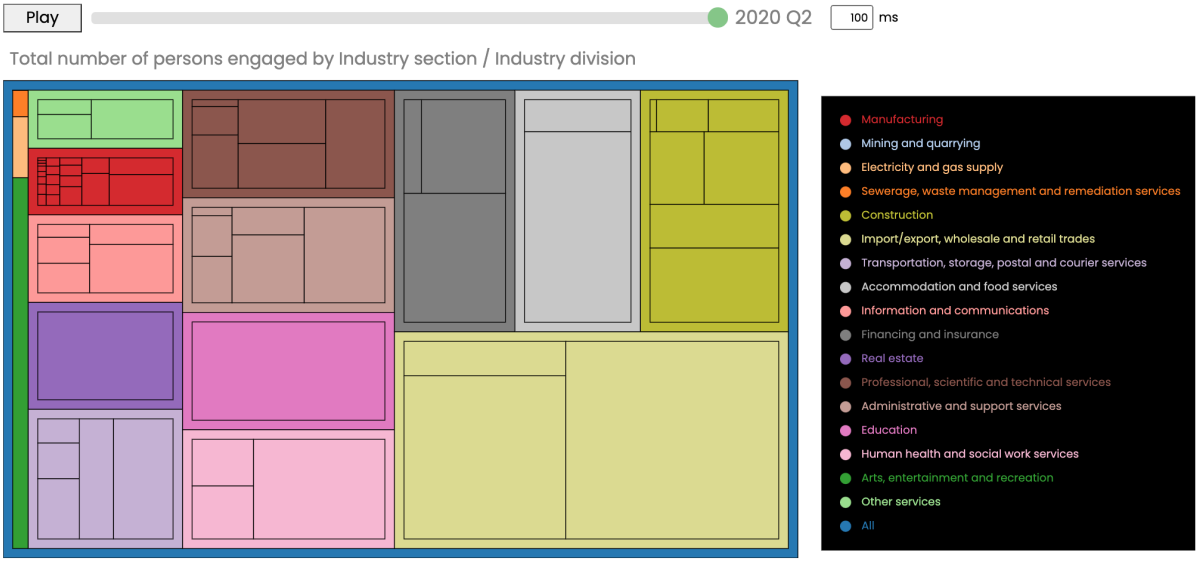

Figure 1: Employment and Vacancies by Industry Treemap

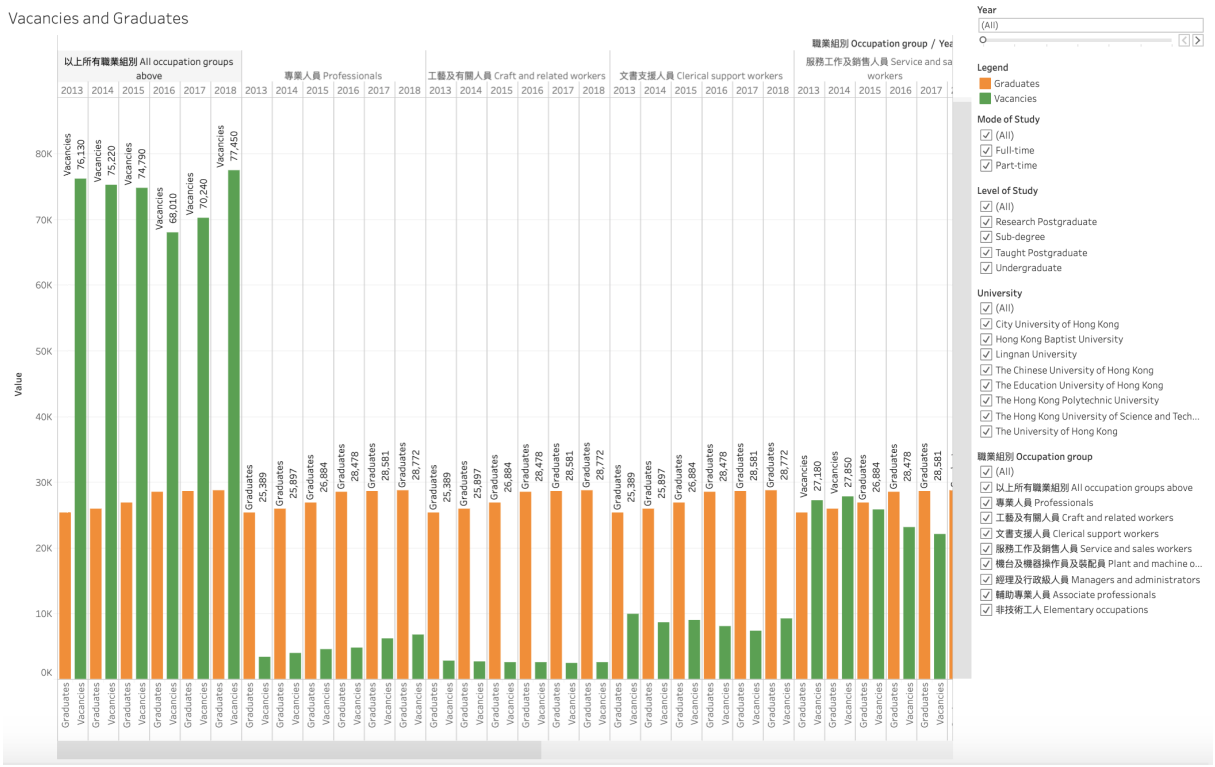

Figure 2: Vacancies by Occupation and Graduates Side-by-Side Barchart - All Years

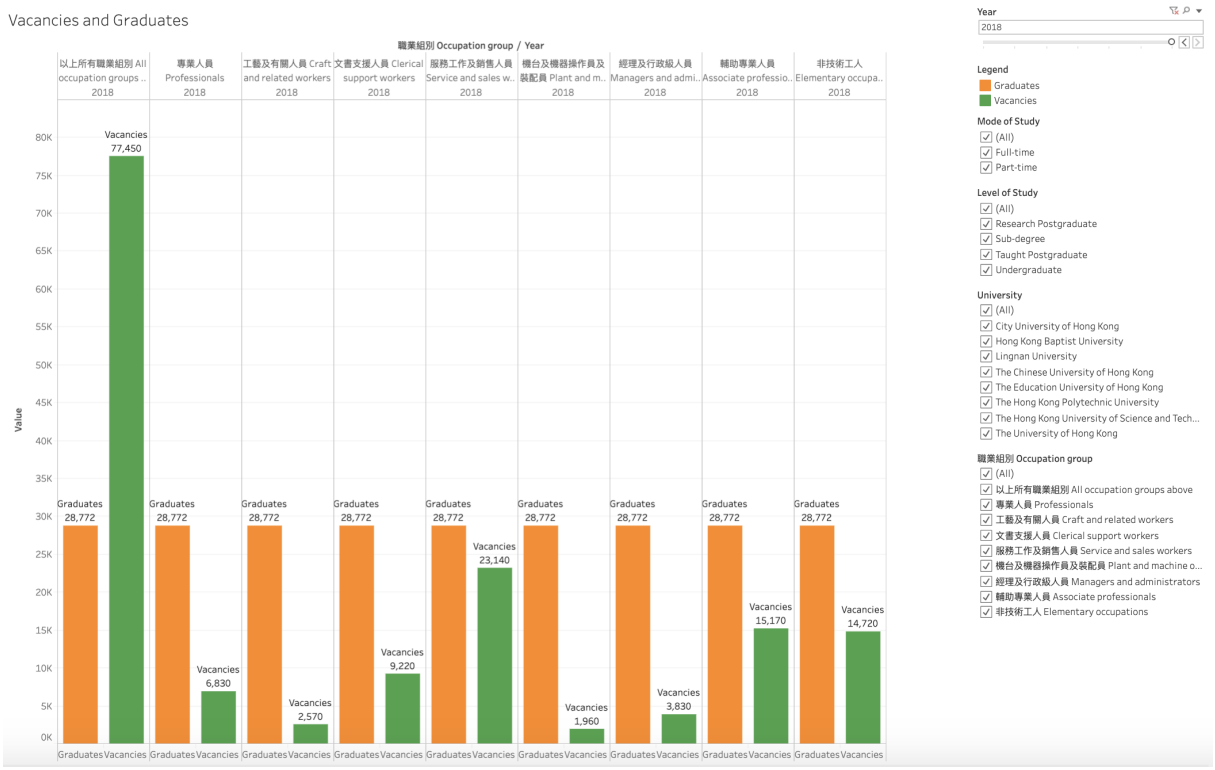

Figure 3: Vacancies by Occupation and Graduates Side-by-Side Barchart - Latest Year

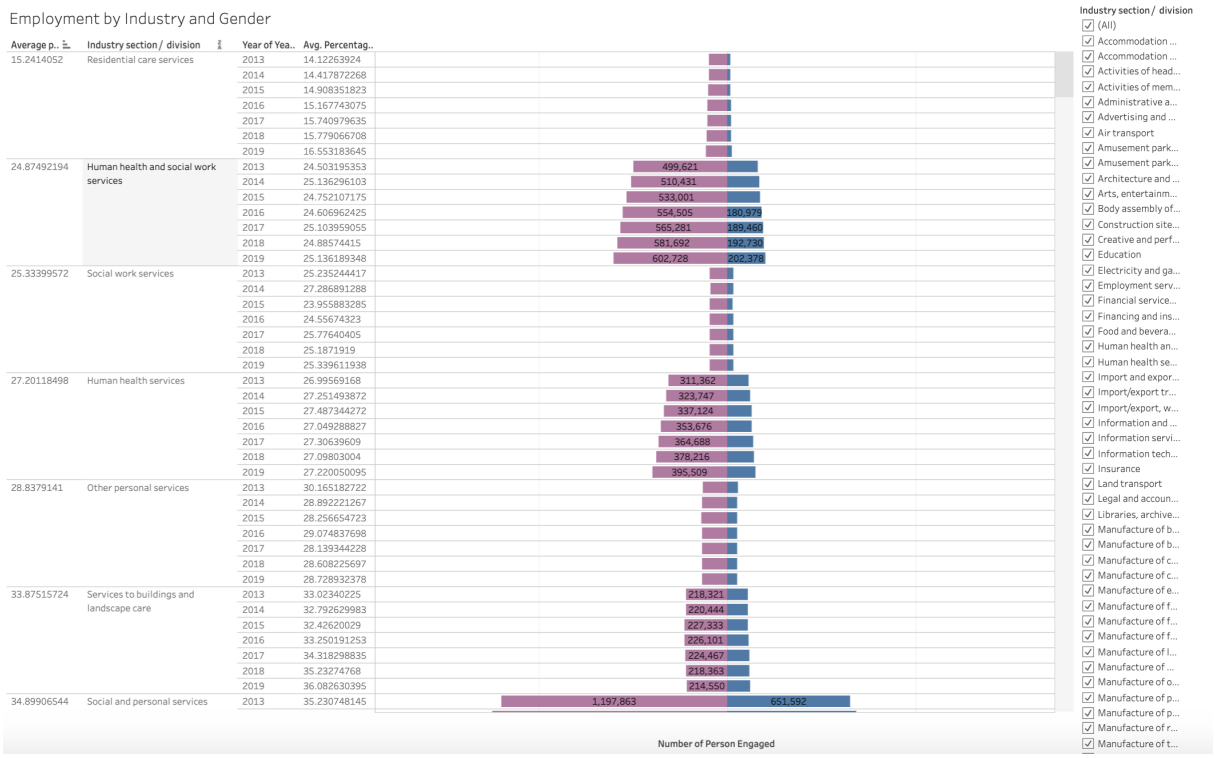

Figure 4: Employment by Industry and Gender Centered Stacked Barchart

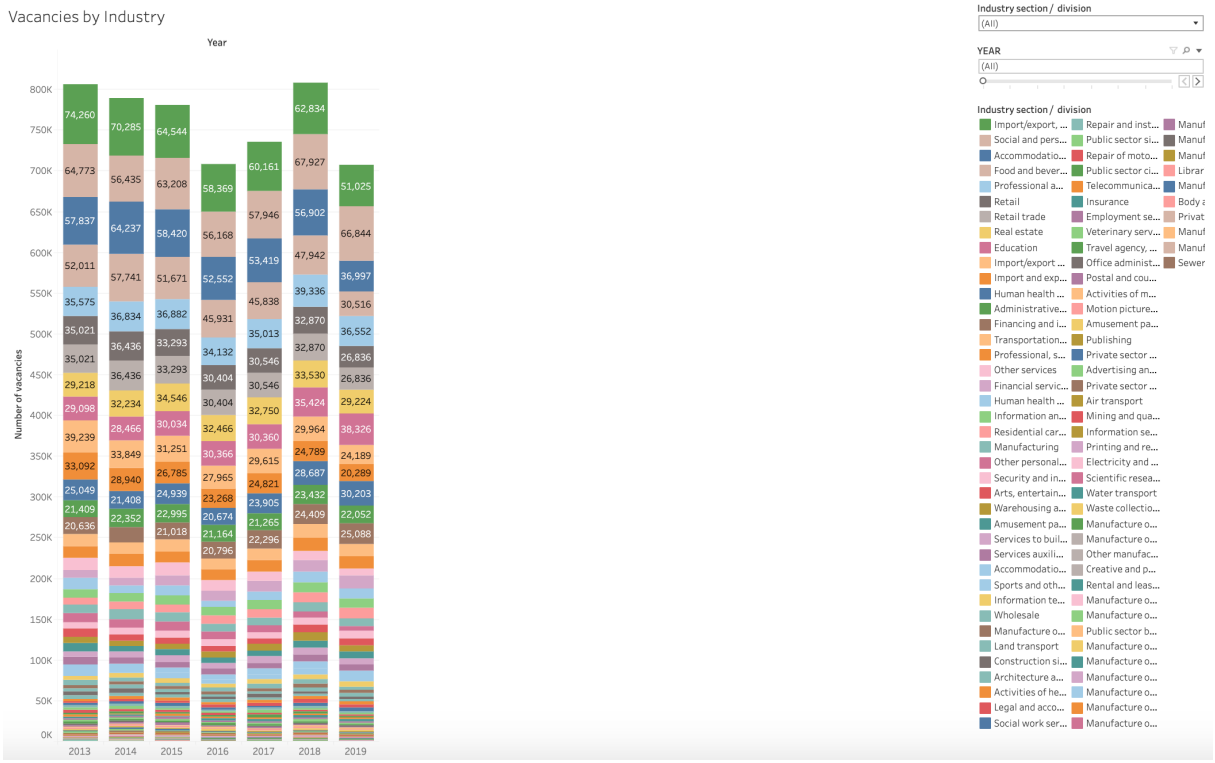

Figure 5: Vacancies by Industry

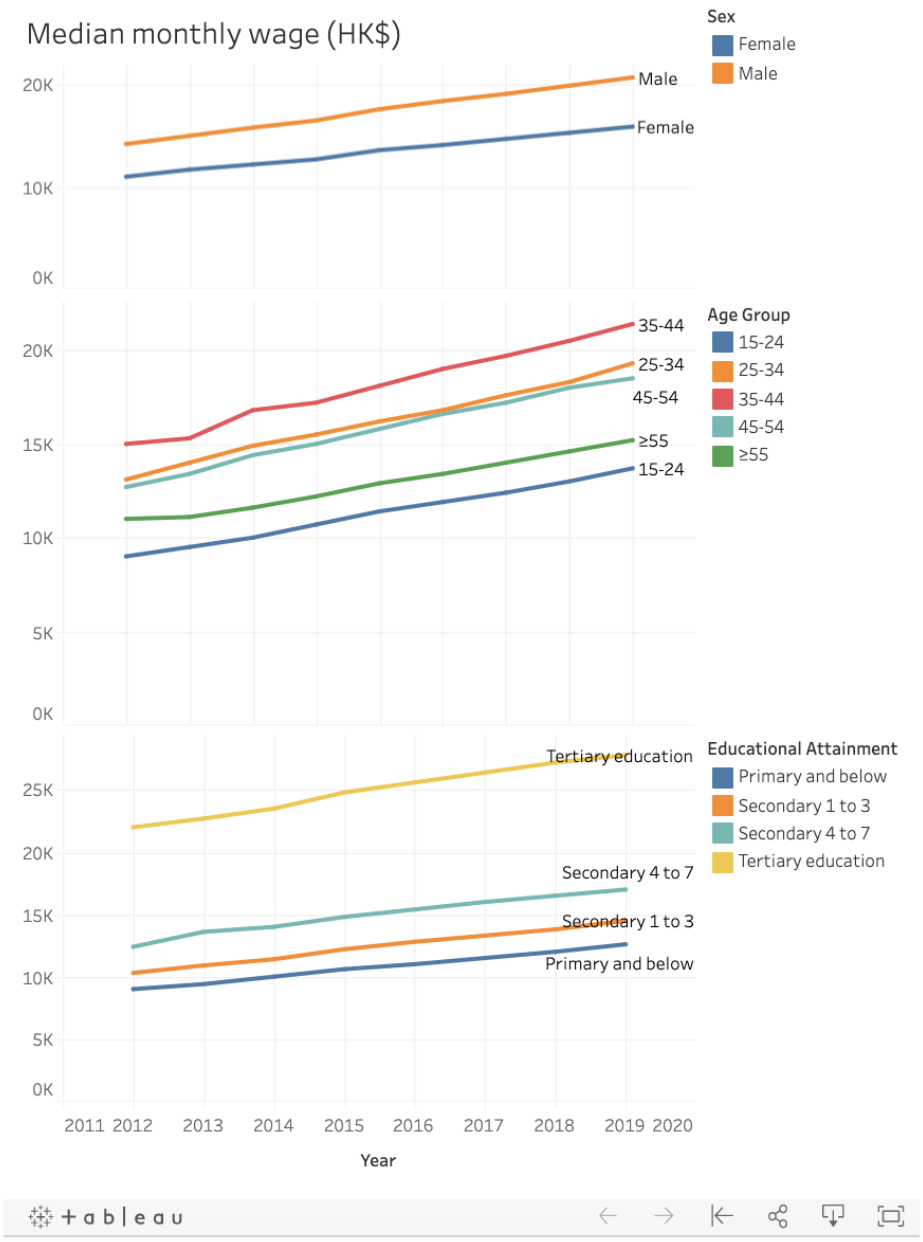

Figure 6: Median Monthly Wage by Gender, Age Group and Educational Attainment

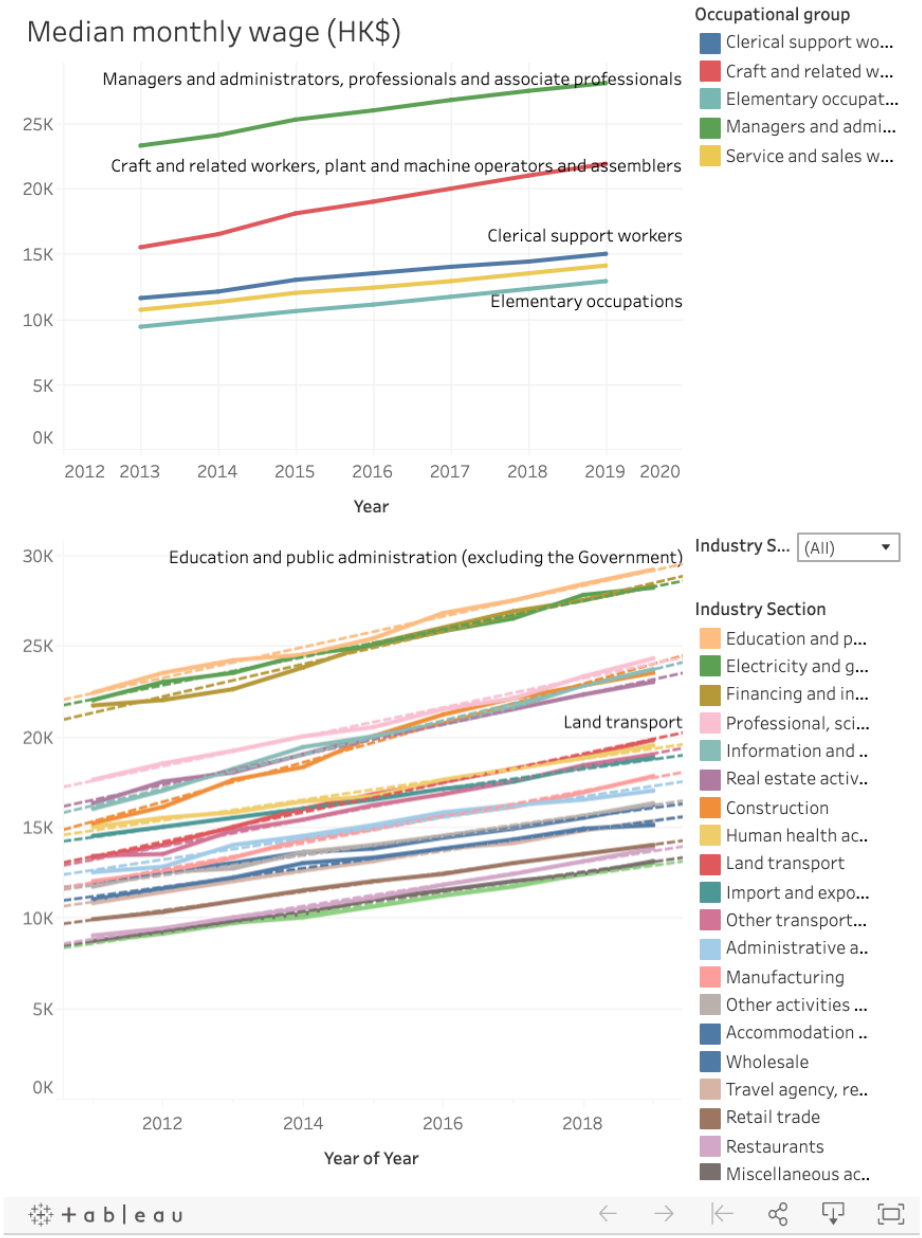

Figure 7: Median Monthly Wage by Occupational Group and Wage

Hours of work

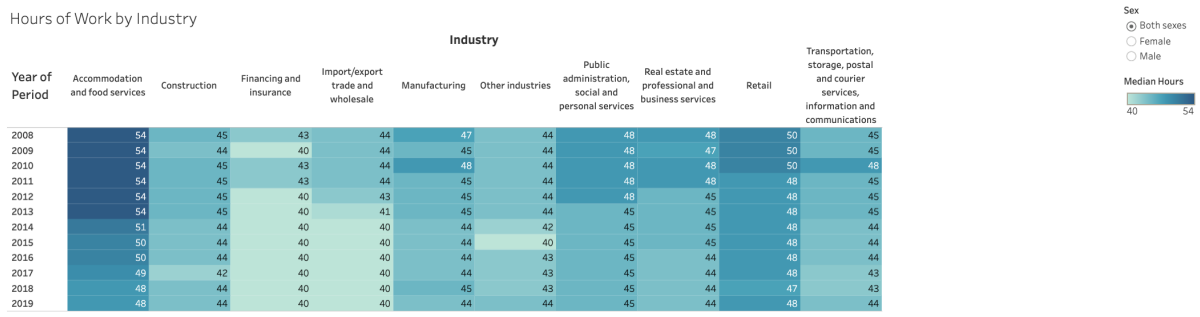

Figure 8: Hours of work by Industry

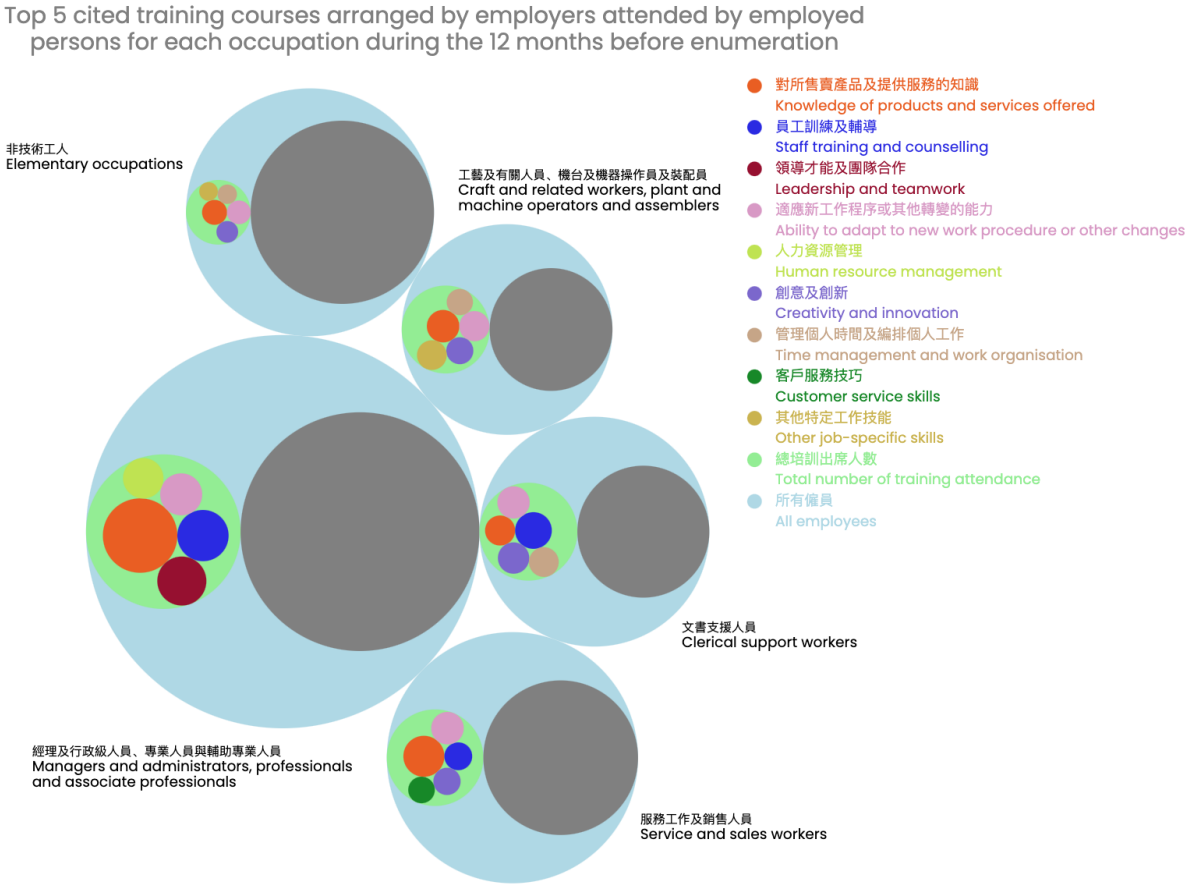

Figure 9: Training needs by Occupation

5 Discussion

5.1 Interpretation of the results

5.1.1 Animated squarified treemap: Employment and Vacancies by Industry

When the treemap is interpreted, industries can be ranked by their total number of persons engaged. An unambiguous ranking is shown since the data are already sorted. Key observations are shown below:

- Arts, entertainment and recreation industry grows very sharply, its total number of persons engaged increase from 2379 in 2000 Q1 to 47259 in 2020 Q2. This is a 1890% increase.
- Manufacturing industry diminishes rapidly, its total number of persons engaged drops from 215560 in 2000 Q1 to 84187 in 2020. This is a 60.9% drop.
- Although import/export, wholesale and retail trades industry is the largest industry along the timeline, its proportion to all industries keeps decreasing. Its total number of persons engaged does not significantly grow with Hong Kong population.

- Construction industry is the second largest industry from 2000 Q1 to 2002 Q2. Its rank drops to the lowest in 2009 Q3. From 2009 Q4 onwards, its rank and total number of persons engaged increase rapidly. It has become the second largest industry again since 2015 Q4.
- In the treemap, the areas of Professional, scientific and technical services industry, Administrative and support services industry, Education industry and Human health and social work services industry increase greatly. These industries' increase in total number of persons engaged from 2000 Q1 to 2020 Q2 ranges from 76.1% to 113% whereas total number of persons engaged of all surveyed industries only increase by 20.6%. Therefore, it is obvious that these four industries are ranking up and gaining in proportion significantly along the timeline.
- Financing and insurance industry, Real estate industry and Information and communications industry are ranking up and gaining in proportion whereas Transportation, storage, postal and courier services industry is ranking down and losing in proportion along the timeline.

5.1.2 Side-by-Side Barchart: Vacancies and Graduates by Occupation Group

Seeing vacancies next to graduates in one chart allows one to visualize the supply and demand of the job market. In 2018:

- Top 3 occupations with highest vacancies: service and sales, associate professionals, then elementary occupations.
- Bottom 3 occupations with lowest vacancies: plant and machine operators and assemblers, craft and related workers, and managers and administrators.
- Analysis for potential job application pool can be done by selecting mode of study, level of study, and university and seeing how the graduates numbers change.

5.1.3 Stacked barchart: Vacancies by Industry

When evaluating industries and occupations, vacancies are important as well as employment. Observations:

- In 2019, top 3 industries are: social and professional services, trading, and education.
- Top industries do not fluctuate much over the years, from 2013 to 2019.

5.1.4 Centered Stacked Barchart: Employment by Industry and Gender

This barchart can be used to answer how equally distributed are the two genders across industries. Observations:

- Top 3 industries with the highest average percentage of male: public sector civil engineering sites, private sector civil engineering sites, public sector sites.

- Bottom 3 industries with the lowest average percentage of male: residential care services, human health and social work services, social work services.
- Most evenly distributed industries in terms of gender: Services auxiliary to financial and insurance services, manufacture of textiles, sports and other entertainment services.

5.1.5 Line charts: Median Monthly Wage by Gender, Education, Age Group, Occupation and Industry

5 charts to provide insight into wage by demographics over the years. Observations:

- Top 3 wages by industries: education and public administration, financing and insurance, and electricity and gas supplies.
- Bottom wages by industries: estate management, security and cleaning services, miscellaneous activities and restaurants.
- All industries have growing wages from 2011 to 2019.
- Top wage by occupation: managers and professionals (HK$28,100 in 2019).
- Bottom wage by occupations: elementary occupations (HK$12,900 in 2019).
- Wage by gender reveals that male average wage is consistently higher than female, with over 30% difference in 2019.
- Wage by age group shows that 35-44 group have the highest wage, followed by 25-34, then 45-54 group. This is aligned with the expectation that wages are highest with over 10 years of work experience and decreases after over 20 years of work.
- Wage by educational attainment shows tertiary education group has a much higher wage than the next group, secondary 4 to 7 (63.52% higher in 2019). Differences between the rest of the 3 groups are more evenly distributed.

5.1.6 Heatmap: Hours of Work by Industry

Color-coded data to easily spot trends and changes over the years across different industries. Observations:

- Top number of hours of work industry: accommodation and food services, longest working hours consistently between 2008 and 2019 (48-54 hours per week).
- Lowest number of hours work industry: trade and wholesale have the shortest working hours (40-44 hours).
- Common trend of working hours reducing in recent years across most industries.
- Males have longer working hours in general. However, female's working hours have less reductions over the years compared to male's.

5.1.7 Packed bubble chart: Training needs by occupation

The bubbles are in different colours, corresponding to employed persons with different training experiences. The innermost bubbles count the number of training attendance of different training types. The light green bubbles count the total number of training attendance. Note that each person can attend multiple types of trainings. The light blue bubbles, which represent the number of employed persons in different occupations, act as references for examining how relatively large the numbers of trainings are. Since area of bubbles is proportional to number of employed persons only for bubbles at the same depth in a packed bubble chart, grey bubbles are important for visualizing the proportion of total number of training attendance to the number of all employed persons. There are some key observations:

- Training about knowledge of products and services offered is the most attended training since it is cited in all 5 occupations and it has the largest coverage among all types of trainings.
- The proportion of total number of training attendance to the number of all employed persons of elementary occupations is obviously smaller than that of other occupations. This proportion is only 6.59% whereas other 4 proportions range from 17.2% to 24.3%.
- As total number of training attendance is less than one-fourth of number of all employed persons, most of the employed persons do not have training needs since their employers have not arranged any trainings for them.
- Actual training resources needed for different occupations can be obtained by investigating the sizes of light green bubbles. For example, managers and administrators, professionals and associate professionals have the most training needs.
- Occupation-specific trainings which may not be important in general can be highlighted easily. For instance, dark red bubble, which is leadership and teamwork, is particularly important for managers and administrators, professionals and associate professionals whereas green bubble, which is customer service skills, is particularly important for service and sales workers.

5.2 Answers to the research questions

5.2.1 What is the best industry and occupation for one to enter as a fresh graduate, based on his/her skills and interests, as well as the market conditions (saturation level and salary)?

Personalized answers can be derived using the visualizations. One would look at the industry of interest based on his/her education, skills and career preferences.

As an example, a female HKU computer science masters student can start with the industry treemap and find the following useful insights on promising industries to enter:

- The top 3 largest industries by employee numbers are trading (import/export, wholesale and retail), construction, and food and beverage services.

- Arts, entertainment and recreation industry grew rapidly in the last 2 decades (1890% increase).
- Professional, scientific and technical services industry is another growing industry and is highly relevant for computer science graduates.

Another factor to consider when deciding on the most promising industry, besides the employment numbers, is vacancies. If an industry has both a large number of employees and vacancies, it signals that it is both a substantial industry and it has a strong need of hiring talents. Looking at the vacancies by industry stacked barchart, the following can be deduced:

- The top 3 industries in 2019 are: social and professional services (66,844), trading (import/export, wholesale and retail) industry (51,025) and education (38,326).
- Professional services and trading industries are also promising industries from the industry treemap. Education is another potential industry to consider that is revealed from using the vacancies barchart.
- The top industries did not change much over the years, when inspecting 2013 to 2019 barcharts together.

Further examining the vacancies and graduates side-by-side bar chart, she would see that:

- Professionals has the 5th highest vacancies out of 8 occupation groups in 2018, signaling a potentially competitive occupation.
- Selecting all 6 years' data on professionals reveal that the vacancies have been increasing every year.
- Assuming that the comparable candidates for the type of professional roles that she'd pursue are graduate level students, by filtering out undergraduate and sub-degree graduates, the potential candidates go down dramatically from 28,772 to 4,503 in 2018. Compared to the vacancies of 6,830, it appears to be a promising occupation.
- Further analysis can also be done by comparing the number of graduates from HKU and other universities, using the selections offered in the visualization, to gauge the competitiveness of the job market.

Next, looking at the 5 high potential industries from the industry treemap, using the employment by gender centered stacked barchart, the masters student notices that trading and food & beverage services industries have higher female employees than the other industries. This can be useful for making decisions in terms of fair representations in the industry by gender.

Financial compensations from the chosen occupation can be further understood from the Hours of Work by Industry heatmap and Wage charts.

Hours of Work:

- Accommodation and food services have the longest working hours consistently between 2008 and 2019 (48-54 hours per week).
- Trade and wholesale have the shortest working hours (40-44 hours).
- Professional services have working hours in between the two industries above (44-48 hours).
- The heatmap also makes it easy to see that there is a common trend of working hours reducing in recent years.
- Working hour can also be viewed by gender, should further insight be desired.

Wage:

- Education and public administration has the highest median monthly wage (HK$29,200 in 2019).
- Out of the high potential industries deduced from industry treemap, professional, scientific and technical activities industry have the highest wage (HK$24,300 in 2019).
- All industries have an upward trend, with the education industry having the highest growth out of the industries of interest.
- The different industries and occupations' wage ranking had remained the same over the 9 years, for the industries of interest.
- In terms of occupations, managers and professionals have the highest average wages (HK$28,100 in 2019).

Viewing the results of studying these visualizations as a whole, the female computer science master's student may arrive at the decision to focus on professional, scientific and technical activities and education industries for the best combination of job prospects, wage and working hours.

5.2.2 How would one select a program of study in school that can meet the market demand and his/her personal interests?

For new entrants to universities, personalized answers can be derived in a similar manner to the first question, using demands by industry and occupation. Top industries in terms of employment size and vacancies by gender can be understood from vacancies and graduates side-by-side barcharts, employment by industry and gender, and vacancies by industry charts.

Wages and hours of work can be explored with the corresponding charts as well, to give ideas of whether the jobs and industries are aligned with the students' preferences and needs.

After studying these areas, students can decide which program of study would meet the demand of the desired occupation and industry.

5.2.3 Which industries and occupations have the brightest future and, therefore, requiring educational institutes to direct more teaching resources and to encourage student enrollments in?

For educational institutes the most important visualizations to consider are employment, vacancies and wages. The institutes can go through these visualizations, taking the standpoint of an overall view, and determine the most promising industries and occupations. They can then review the current enrollments and availability of teaching resources to plan for recruiting students and teachers to optimize the futures of the school and society as a whole.

From the vacancies and graduates barchart, service and sales workers have the highest vacancies, followed by associate professionals, then elementary occupations. These are occupations that the schools can compare the enrollment levels of relevant majors to determine if encouragement of increased enrollments are necessary. On the other hand, plant and machine operators and assemblers have the lowest vacancies and would probably not be an occupation to encourage students to enter into.

In terms of wages, education and public administration, financing and insurance, and electricity and gas supplies have the highest median monthly wage, while estate management, security and cleaning services, miscellaneous activities and restaurants have the lowest wages. These factors shall be considered when deciding on the skills to be taught at the schools.

5.2.4 Which industries and occupations have the highest potentials and can be boosted further with policies to attract companies and talents?

Government agencies can use the visualizations to determine industries and occupation outlooks and direct resources to attract foreign and domestic companies, as well as talents, to enter the right industries and occupations.

In particular, from the treemap, industries with rapid gaining of proportion and headcount such as Professional, scientific and technical services industry, Administrative and support services industry, Education industry and Human health and social work services industry can be highlighted easily. These industries have good prospects and are gaining significance.

In addition, understanding the vacancies and graduates statistics helps the government design and implement policies.

5.2.5 Which demographic groups (age, education level, sex) are the most disadvantaged and will need government assistance on improving the industry outlook and plans for re-training of the current labor force?

Visualizations on hours of work, employment, vacancies and wage differences by demographics equip government agencies with the insight to help determine areas needing most assistance and improvement.

Industries with the least number of vacancies are areas that may require further investigation to understand if assistance is needed: sewerage, manufacture of paper, furniture, leather, body

assembly of motor vehicles.

From hours of work heat map, it can be seen that the accommodation and food services industry has the highest working hours, followed by retail. Efforts can be devoted to these industries to help improve the working hours.

Wage by gender, age group, educational attainment, occupational group, and industry section, government charts reveal these key facts:

- Gender gap in wages is significant, with a 30.2% difference in 2019.
- Wages stay relatively robust until the age group of 45-54. Starting with the age group above 55, wage drops significantly by 21.7% in 2019 and this group may potentially require further government attention. Another low-income group is 15-24. However, this group includes a large population of students, which are expected to have little to no income.
- There's a huge gap in wage between tertiary education level and all other lower education level groups. Wage dropped by 63.52% from tertiary education to secondary 4 to 7 education level. Assistance to groups with education less than tertiary should be a focus of the government, as well as increasing education levels to maximize tertiary education recipients.
- In terms of industry, estate management, security and cleaning services, miscellaneous activities and restaurants have the lowest wages and will benefit from attention from the government.
- For the occupations, elementary occupations, service and sales workers, and clerical support workers standout as significantly lower wage earners that the government can help to improve wages in these industries, or help workers transition to other industries.

5.2.6 What types of training programs are most beneficial to the workforce and students?

When certain skill sets are necessary and valuable for operations, it is likely for employers to provide more trainings about these skill sets. Therefore, size of bubbles reflects training needs.

Training about knowledge of products and services offered is the most needed training since it is cited in all 5 occupations and it has the largest coverage among all types of trainings. Therefore, it is in general beneficial for employees, job seekers and students to train their domain knowledge and understand the products and services of the interested industries.

Moreover, if students aim at becoming managers and professionals, knowledge of products and services offered, leadership and teamwork and staff training and counselling are especially important and targeted for them. They are advised to evaluate themselves and seek for suitable trainings.

The government, educational institutions and training units can allocate their resources better to different occupations based on the size of training needs reflected by the bubble sizes.

5.3 Evaluation of this study

5.3.1 Difficulties encountered

- Data from the Census and Statistics Department of Hong Kong Government, while easy for humans to read, was not structured as suitable input for visualization tools such as Tableau. Required significant amount of work to convert to the desired format, before visualization can be started.

- Determining which data are most relevant to use from a large amount of data on employment, vacancies, wages, hours of work, etc. took a significant amount of time. Especially with the challenges of being unable to visualize the data easily before they are loaded into visualization tools.

5.3.2 Something planned but not done

- One area that we could look into is data on different fields of studies in universities and the correlations to the industries and occupations. For example, it would be useful to know the percentage of students of one major going into certain industries. We can then map the graduates to the industries more precisely. We did not do this because the data on which industries each student goes to upon graduation would not be totally accurate as it is only based on students who volunteered the information. Also, the industry can change if a student changes job soon after joining a company. This type of information is not part of the Hong Kong Census and Statistics Department data, making it more difficult to obtain, verify and visualize.

5.3.3 Limitations to the existing tools

- Existing tools could not display all the information on the visualizations clearly in some cases. For example, industry names as text in treemap and wage line charts, since the industry names are long and will cause the words to overlap on screen. As a tradeoff, the names are removed from the graphs and would appear when hovered over. This is an example of how interactivity can solve the limitations of fixed displays. However, it is less intuitive to process than being able to see all the information with one glance.

6 Conclusion

A healthy and efficient labor market is crucial to the economy of any city. Having insight into issues in the labor market, making the right decisions on employment policies, matching labor supply to market demands and helping enterprises and workers to adapt to technological and societal changes by picking up the right skills are incredibly important to the well-being of citizens living in the city.

Building visualizations using data from the Census and Statistics Department of Hong Kong Government, the following crucial questions can be answered with certain confidence, based on the support of the data and the insight gained through interpreting the data:

- What industries should a fresh graduate choose to enter?
- What program of study should a new student select?
- Which industries and occupations should educational institutes direct more teaching and recruiting resources to?
- What policies and assistances should governments provide to businesses and workers?
- What demographic groups should government direct resources to support?
- What type of trainings should the promoted by the government and businesses?

Visualizations of employment and educational data are powerful tools for providing insight to contribute to the well-being of not only the labor market, but also the society as a whole.

7 Division of Labour

- Data searching: Fong Chun Him
- Create detailed ideas and drafts of project objectives and methodology: Fong Chun Him
- Review detailed ideas and drafts of project objectives and methodology, and then finalize the proposal in formal and professional writing: Wu Szu Han
- Animated squarified treemap and packed bubble chart visualizations and interpretations: Fong Chun Him
- Other visualizations and interpretations: Wu Szu Han
- Frontend: Fong Chun Him
- 3-min highlight video creation: Wu Szu Han
- Final report introduction, methodology, division of labour: Fong Chun Him
- Final report other sections: Wu Szu Han
- Review and typeset final report: Fong Chun Him
- 5-min presentation: Wu Szu Han