

Unemployment In Queensland

DATA7001 GROUP 7

Ashama Plipat s4692549, **Jia Ji** s4720608, **Yanran Li** s4722469,
Jingyun Cheng s4718965, **Shilin Vinayak Agre** s4709182

We give consent for this report to be used as a teaching resource.

EXECUTIVE SUMMARY

The report details how Queensland's unemployment rate has changed during the COVID-19 pandemic. In the whole process of the survey, the main stakeholders are the people working or seeking jobs in Queensland, because they are directly affected by changes in national policies and their own conditions. There are other stakeholders, such as institutions that issue national epidemic control and industries that are subject to job changes.

In section 1, we describe the motivation for this research, the main questions raised, and the people we hope to help. Section 2 describes the data sets we used and where they came from. In Section 3 we specifically tuned our data sets to make them available to us. Section 4 analyzes the dataset we used and creates models for it. Section 5 includes answers to our questions about the unemployment rate in Queensland and recommendations that can be made to stakeholders.

The project was largely successful because we followed data science processes and design thinking to answer the questions we asked. What we found was that what really affected Queensland's unemployment rate during the pandemic was not COVID-19, but national policy. In addition, we also provide useful suggestions for stakeholders through the analysis of the data set. For example, women tend to have the largest percentage in health care and social assistance and the next industry is the education and training; while for men, the most popular one is construction and the following one is manufacturing. One limitation of this project is that we cannot quantify the impact of COVID-19 on stakeholders. Such as some national policies are issued just because of COVID-19, and the number of job seekers will also be affected by the epidemic.

TABLE OF CONTENTS

| | |
|--|----|
| EXECUTIVE SUMMARY | 3 |
| 1 Introduction | 5 |
| 1.1 Why employment is so important | 5 |
| 1.2 The main factors affecting employment | 5 |
| 1.3 Our targets | 5 |
| 1.4 Key stakeholders | 5 |
| 2 Getting data I need | 6 |
| 2.1 Data sources | 6 |
| 3 Making the data fit for use..... | 7 |
| 3.1 Data preprocessing..... | 7 |
| 3.2 Data exploration | 7 |
| 3.2.1 Exploring the employment factor..... | 7 |
| 3.2.2. Exploring the covid-19 factor | 8 |
| 3.2.3. Exploring the education factor..... | 8 |
| 3.2.4 Exploring the age factor | 9 |
| 4 Make the data confess | 10 |
| 4.1 Relationship between covid19 and unemployment..... | 10 |
| 4.2 Relation between employment and Education | 11 |
| 4.3 Relation between employment and Age | 11 |
| 4.4 Relation between employment and Family Relationship | 12 |
| 4.5 Connection between covid-19 and unemployment..... | 13 |
| 4.6 Connection between other aspects and employment rate | 14 |
| 4.6.1 Education level..... | 14 |
| 4.6.2 Age..... | 15 |
| 5 Management of Change | 16 |
| 6 Recommendations | 17 |
| 7 References | 19 |

1 Introduction

1.1 Why employment is so important

From three aspects:

1. Social stability needs employment to develop. Only when the people live and work in peace and contentment can society develop stably.
2. Family happiness needs employment support. Employment is the foundation of people's livelihood. Full employment can guarantee the economic source of the family and thus improve the happiness of family members.
3. Personal value needs employment to realize. Only through employment can one bring his personal value into full play.

1.2 The main factors affecting employment

In order to analyze employment, we must first find the factors that affect employment.

There are two key factors:

- National policy factors.
- Personal factors of job seekers.

The adjustment of the economic structure or the adjustment of the industrial structure by national policies will affect the employment situation to a certain extent. In addition, a series of policies proposed for the special period will also have an impact on employment. The knowledge, skills and other conditions of job seekers themselves, as well as their family background, will have an impact on their employment.

1.3 Our targets

We mainly focus on the two questions below and we followed data science processes and design thinking steps to investigate the questions we ask:

- Impact of Covid-19 on unemployment of Queensland?
- Other aspects affect employment rate?

1.4 Key stakeholders

Through our analysis of datasets, we also want to get benefits for stakeholders as following:

Table 1: Key stakeholders of the project

| Stakeholders | Interest |
|--------------|--|
| Job hunters | People who is looking for a job could get useful information. |
| Students | Students who is preparing for career can get suggestions for further decision. |
| Government | The government should improve and provide necessary policies for citizens. |

2 Getting data I need

2.1 Data sources

To take account of time changes and various factors that may affect Queensland's unemployment rate, we need to use several different data sets at the same time. See [References](#) for a table outlining the file names for the below datasets.

1. [Detailed monthly and quarterly Labor Force Survey data during 2015-2020\[1\]](#)
From Australian Bureau of Statistics, we get detailed monthly and quarterly labor force survey data, including hours, regions, families, job duration, casual, industry and occupation.
2. [Statistical releases relating to labor and employment in Queensland\[2\]](#)
We get more statistics about relation between labor and employment in Queensland during the pandemic years from Queensland Government website.
3. [Tourism region data profiles include data across July 2018 to December 2020\[3\]](#)
This supplementary source lists all current and historical NOC codes with country name.

3 Making the data fit for use

3.1 Data preprocessing

The datasets were imported to the R and then we inspected them. Fortunately, we did not find anything suspicious or unclear in dataset. All attributes with missing data were inspected using the 'is.na' function. Some data was missing, so we either removed those values or replaced it with '0'.

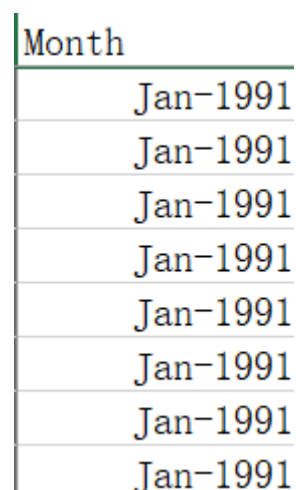
The data that we had was quite huge, so we had to trim it down according to our requirement.

3.2 Data exploration

In this section we will drill down to the major features of this analysis and identify any potential issues.

3.2.1 Exploring the employment factor

As can be seen from the below graph that the data was in textual format for months which could be quite difficult for mapping, so we changed the date format from months name to months number.



| Month |
|----------|
| Jan-1991 |
| Jan-1991 |
| Jan-1991 |
| Jan-1991 |
| Jan-1991 |
| Jan-1991 |
| Jan-1991 |
| Jan-1991 |

Figure 3.1: Date format for the employment data

Also, the data was quite big in amount and had the entries from 1990 to 2022. However, we were not interested in the data before 2010 as we wanted to compare the employment rate from 2019 to the covid-19 time. So we only used the data from 2019 to 2022[4].

3.2.2. Exploring the covid-19 factor

We had the dataset for number of covid-19 cases everyday which is evident from the figure 3.2.

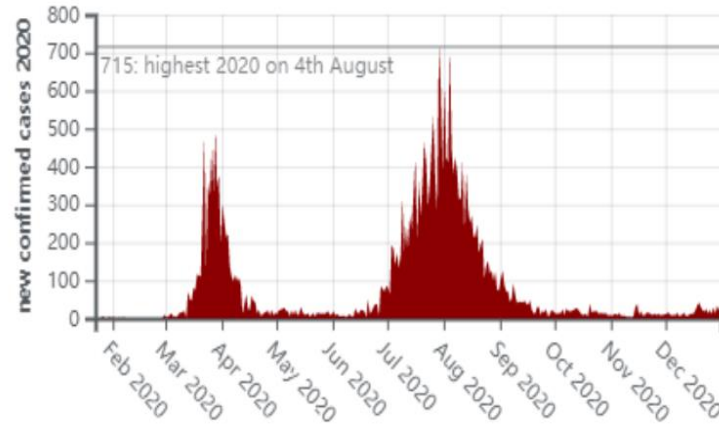


Figure 3.2: Number of confirmed covid-19 cases

Our analysis was taking place on a monthly basis so we transformed this daily covid cases to monthly covid cases. So it would become easy for us to compare it with the other data. Also, one day timeframe is not appropriate to show a smooth trend on any kind of analysis as the possibility of sudden increase or decrease very high unless we are monitoring something which is occurring and changing on hourly or smaller measure unit. That is the reason that we transformed it to a monthly graph which turned out to be figure 3.3.

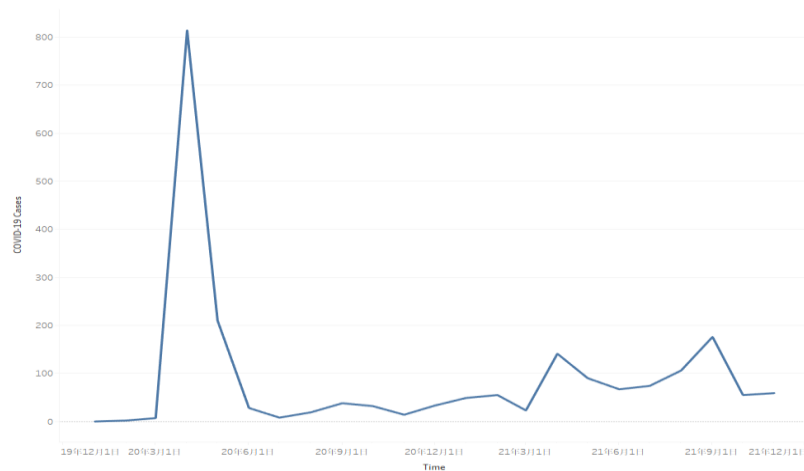


Figure 3.3: Monthly increasing number of covid-19

3.2.3. Exploring the education factor

Education dataset was simple which had education qualification for the people who were employed and unemployed. However, the educational qualifications were stored as strings such as 'bachelors', 'masters', etc. and as we mentioned earlier that working with string data is

difficult and there could be a chance that we might interpret something wrong which could make out whole analysis wrong. So, to overcome this problem we transformed the education levels with a specific number which can be seen in the figure 3.4. Which makes it far easier for us to perform the analysis.

```

Diploma or above
Postgraduate Degree-> 7
Graduate Diploma/Graduate Certificate-> 7
Bachelor Degree-> 6
Advanced Diploma/Diploma-> 6

Certificate III/IV
Certificate III/IV-> 5

Year 12 or below/No educational attainment
Year 12 or equivalent-> 4
Year 11 or equivalent-> 3
Year 10 or equivalent-> 2
Below Year 10/No educational attainment -> 1

Level not determined
Level not determined -> 0

```

Figure 3.4: The breakdown of the highest education of job seekers

3.2.4 Exploring the age factor

Age was a little bit complicated for us as we there was some missing data which we removed. Apart from that, the age column did not have a specific age number but the age bracket for each entry which can be seen in the figure 3.5. To solve this problem, we used median age for a given age bracket. For example, for a age bracket of 15 to 19 we used mean of 15 and 19 which is 17 and introduced a new age column with the age number. We also made one exclusion that the people who are not 17 may not have reached his/her highest degree yet so, we removed the entries with the age being 17.

| Age | Age |
|-------------|-----|
| 15-19 years | 17 |
| 15-19 years | 17 |
| 15-19 years | 17 |
| 15-19 years | 17 |
| 15-19 years | 17 |
| 15-19 years | 17 |
| 15-19 years | 17 |
| 20-24 years | 17 |

Figure 3.5: Age bracket column

4 Make the data confess

Based on the previous analysis, we focus on the following aspects:

Impact of Covid-19 on unemployment of Queensland?

Other aspects affect employment rate.

4.1 Relationship between covid19 and unemployment

```
Call:
lm(formula = d5$Unemployed.persons ~ d5$COVID.19.Cases)

Residuals:
    1     2     3     4     5     6     7
-15479 -31921 -28486  -4237  13310  19121  47692

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  185498.75   13497.44   13.74 3.66e-05 ***
d5$COVID.19.Cases    10.19     42.45    0.24  0.82
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 31320 on 5 degrees of freedom
Multiple R-squared:  0.01139,    Adjusted R-squared:  -0.1863
F-statistic: 0.05762 on 1 and 5 DF,  p-value: 0.8198
```

Figure 4.1 Employment and covid19

As we can see in the graph, we did a linear analysis of the number of covid19 and unemployed people.

And the p-value is 0.82, which is substantially higher than 0.05, demonstrating that our suspicions were incorrect. Furthermore, the r-squared is near zero. Our model isn't quite right.

As a result, based on our current level of evidence, we believe that the growth in the number of persons quitting their jobs is unrelated to covid19.

4.2 Relation between employment and Education

```
Call:
lm(formula = EmployedFull ~ Education)

Residuals:
    Min       1Q   Median       3Q      Max
-57.389 -23.495  -8.087   9.434 178.826

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  14.8802     1.0589   14.05  <2e-16 ***
Education     6.0727     0.2248   27.02  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 40.97 on 5867 degrees of freedom
Multiple R-squared:  0.1107,    Adjusted R-squared:  0.1105
F-statistic: 730.1 on 1 and 5867 DF,  p-value: < 2.2e-16
```

Figure 4.2 Education and Employed

As shown in the figure, we conducted a linear analysis of the relationship between education level and employment.

Our r-squared is 0.11, although this does not have a completely linear relationship. But the p-value is small enough to prove that our guess is accurate: people with higher levels of education tend to be more likely to be hired when looking for a job.

We plotted employment and Education and the fitted regression line based on the findings, as shown in photo 1, left panel below.

4.3 Relation between employment and Age

```
Call:
lm(formula = EmployedFull ~ Age)

Residuals:
    Min       1Q   Median       3Q      Max
-53.29 -30.53 -13.81  15.21 178.78

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  61.75221     2.19290   28.160  < 2e-16 ***
Age          -0.38462     0.04991   -7.707 1.55e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 44.92 on 4858 degrees of freedom
Multiple R-squared:  0.01208,    Adjusted R-squared:  0.01187
F-statistic: 59.39 on 1 and 4858 DF,  p-value: 1.555e-14
```

Figure 4.3 Age and Employed

Like what we did before for Education and employment, we also performed a linear analysis of the relationship between Age and employment.

Our r-squared is 0.011. Even though this does not reflect an utterly linear relationship. Even though this does not reflect an utterly linear relationship, the p-value is small enough to prove that our guess is accurate: older people tend to be had lower employment rate.

We plotted employment and Age and the fitted regression line based on the conclusion, as shown in photo 1 below.

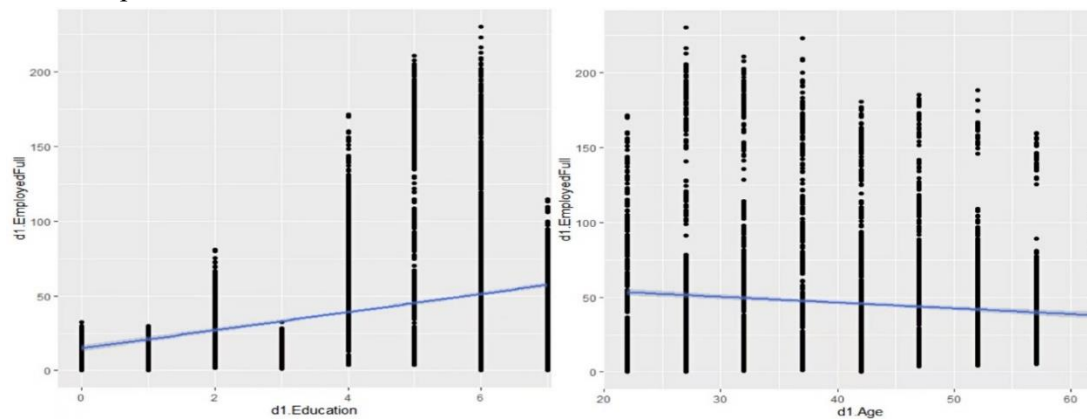


Figure 4.4 Regression line of age and education on employed

4.4 Relation between employment and Family Relationship

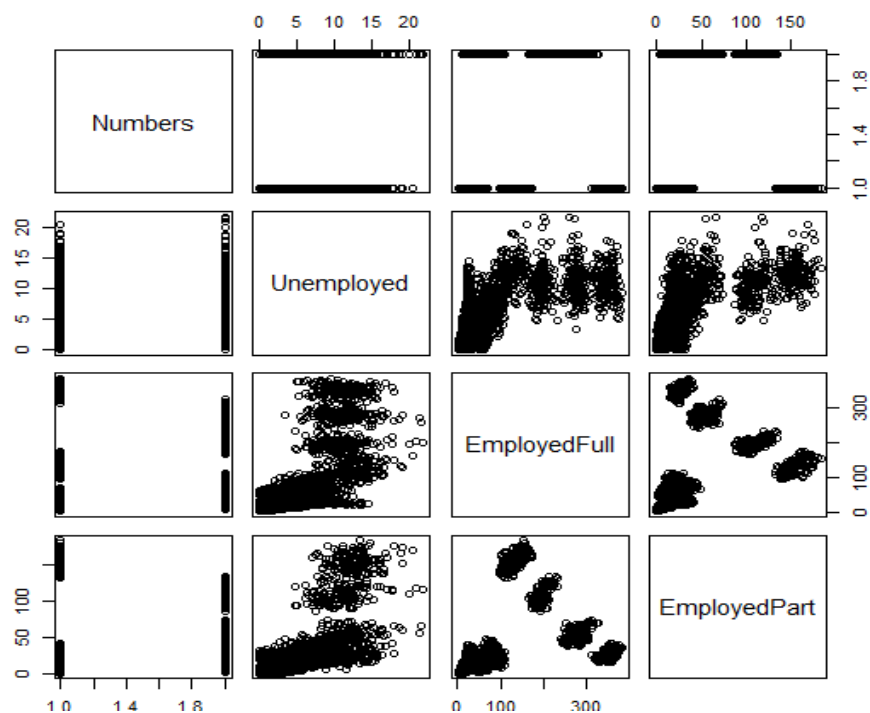


Figure 4.5 Family relationship and employment status

```
Call:
glm(formula = Numbers ~ Unemployed + EmployedFull + EmployedPart,
     family = "binomial", data = d2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3553  -0.8104  -0.6972   1.1238   2.1827

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.2673422   0.0610243  -20.768  < 2e-16 ***
Unemployed   0.3064274   0.0173628   17.649  < 2e-16 ***
EmployedFull -0.0031436   0.0006098   -5.155 2.53e-07 ***
EmployedPart -0.0144447   0.0014773   -9.778  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 4.6 Family relationship and employment status

This section analyzed the Relationship between family relationships and full employment time, employment part-time, and unemployment by logistic regression. The family relationship is divided into 1 for those with children and 0 for those without children. We can also conclude that workers with children tend to be employed full time from the analysis.

4.5 Connection between covid-19 and unemployment

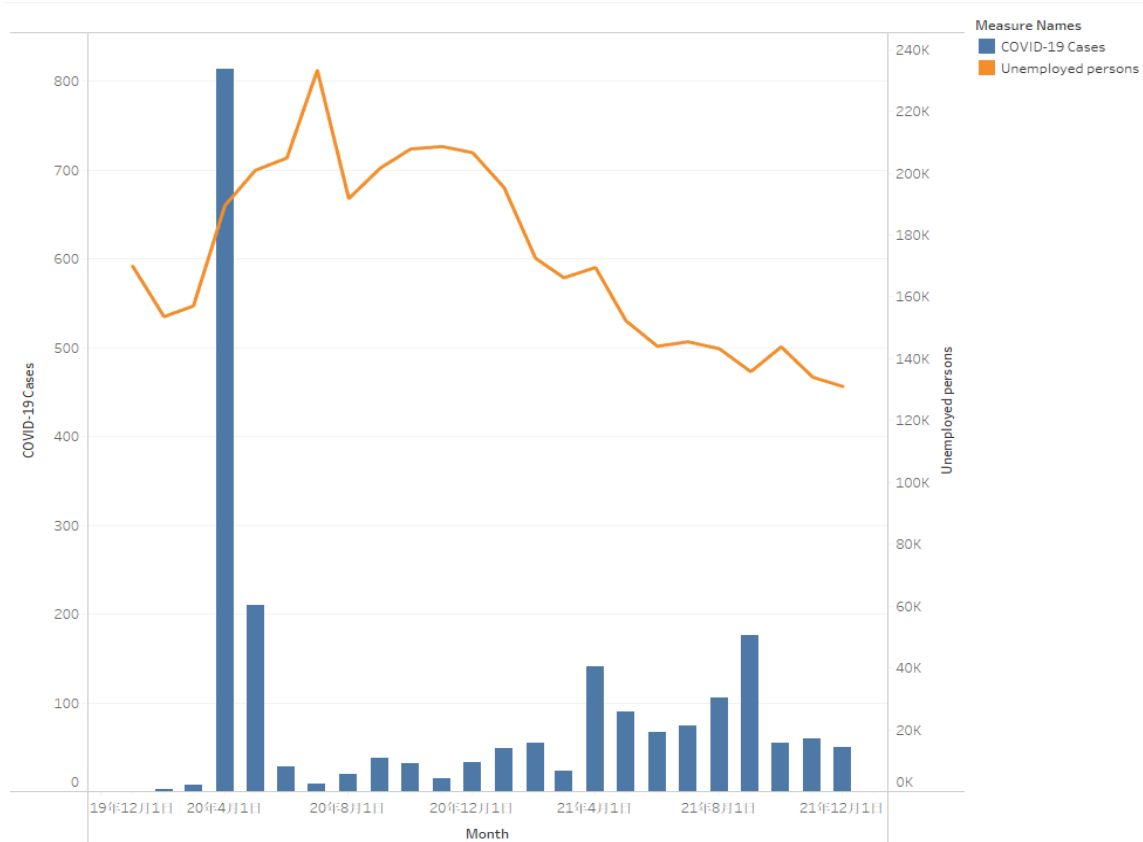


Figure 4.7 Graph1-Covid19 cases and unemployed persons in Queensland from 2019-2021

From the graph above we may think about there might exist some connection between Covid-19 and unemployment because the number of unemployment people has a rapidly increasing during the beginning outbreak year of the pandemic. But after our data analysis about the datasets of covid-19 and unemployment for a long time, we still couldn't find the definite connection between the pandemic and the rise in the number of unemployed people in Queensland. So, we planned to find more elements that might have impact on the employment.

4.6 Connection between other aspects and employment rate

Our group mainly analyzed from two aspects which are education level and age to figure out if they had any connection with our employment rate in Queensland.

4.6.1 Education level

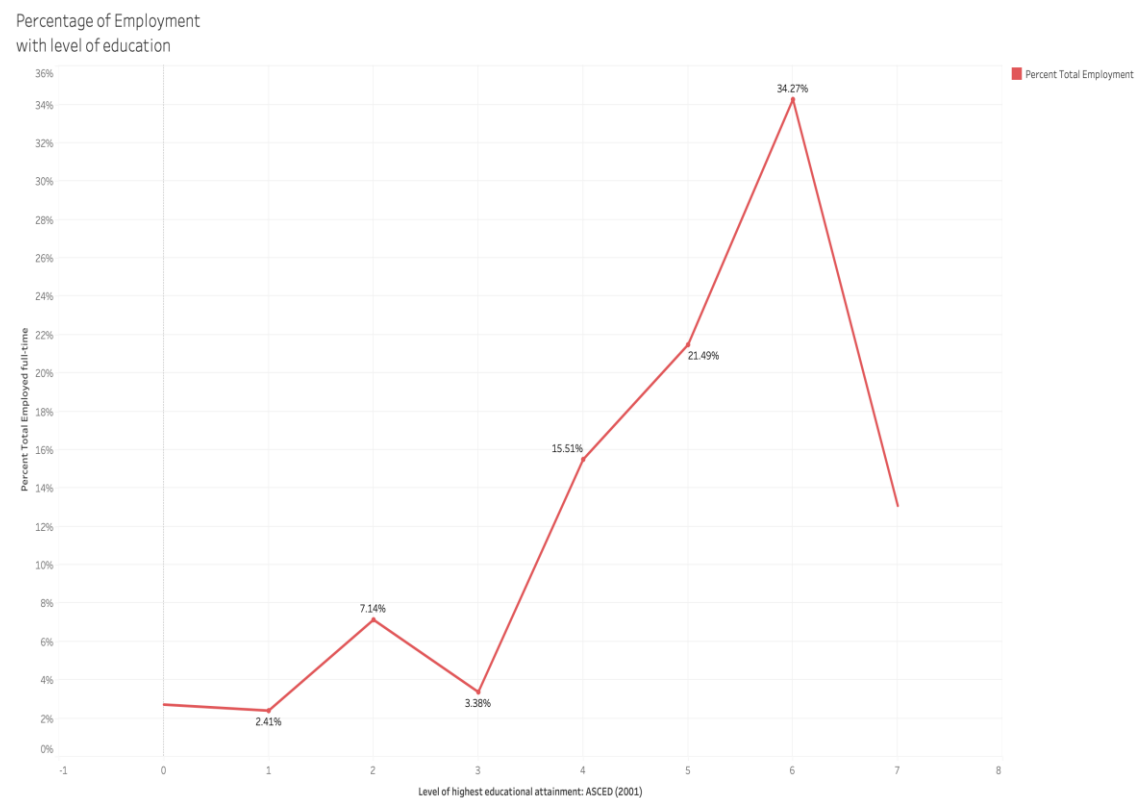


Figure 4.8 Graph2-Education level and percentage of total full-time employment in Queensland

The graph above is about percentage of employment and level of education containing both genders. We can see from it: even though there is some fluctuation during low levels, higher education tends to have much more percentage of employment generally which means you could have more opportunities to get employed if you got a higher degree. Besides, we can see from graph that for people having education of level 6, they have the highest chance of full-time employment.

In general, there is an increasing trend of employment in Queensland according to people having higher level education. There exists connections between education level and employment rate.

4.6.2 Age

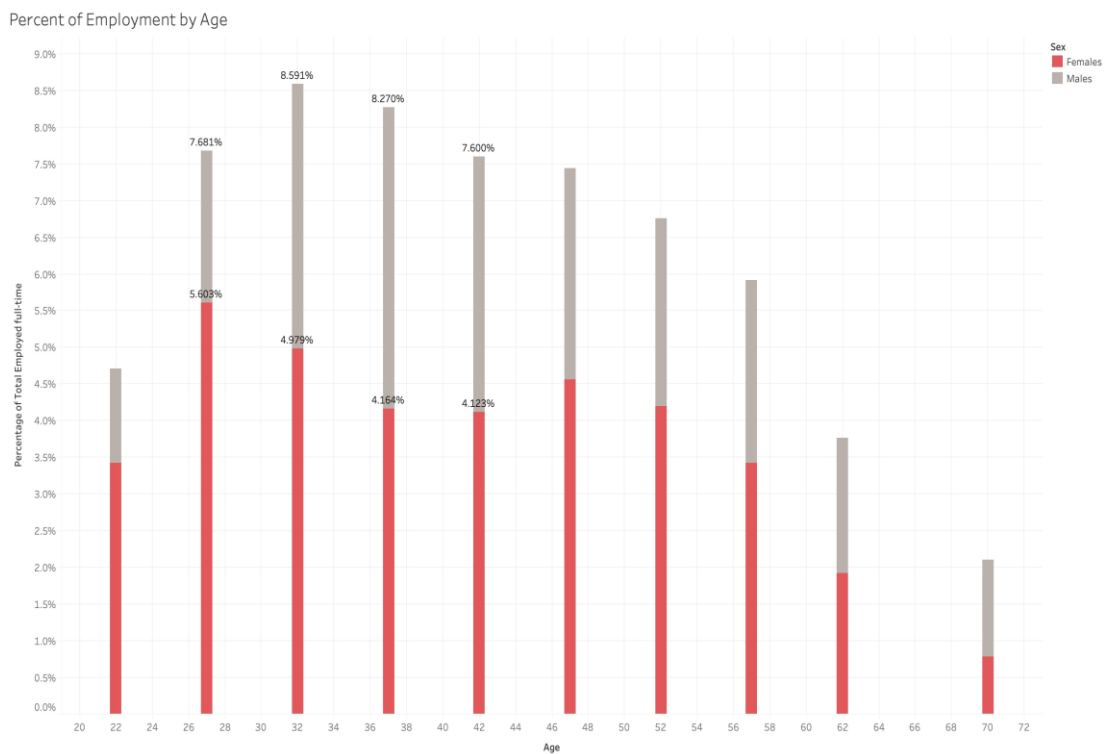


Figure 4.9 Graph3-Age and percentage of total full-time employment in Queensland

For age, we only consider about the range of age which have the opportunity to get full-time jobs.

From graph 3, we can figure out that people age between 20-32 tend to have an increasing number of employment but after that it shows a continuing decreasing trend which indicate us that the older age, the lower employment rate. So, we can get the conclusion that people after 32 years old have fewer employment chances by the growth of age.

5 Management of Change

Since the beginning of the project, we have received feedback from professors, staffs and classmates for improving the project. For the first feedback, it was from professors and we were informed that our problem formulation was too broad. Therefore, we decided to narrow the problems down so we could focus more about unemployment with tourism. Besides, our collected data might not be adequate to make an analysis then, we decided to research more and collect more related data. Apart from that, we improved data confession by indicate more about solutions of questions that we have. Moreover, we had planned to offer recommendations that we got from the analysis. However, at this stage, we could not find much insight information that could answer our problems clearly.

Before the final stage of this project, at trial presentation, we received feedback from our classmates and staffs. For problem solving, we were told that there were not clear aims of our goals and there was unrelated topic included in the story as well. So, after investigating through our project, we fixed unrelated topics and improved for some missing references. Besides, we examined the data and get more insightful from the dataset. In addition, for storytelling data, we were informed that our visualization was not clear enough and unrelated thus, we decided to change some graphs into another type of graph to make it clearer for example, we excluded heat map and included histogram instead. Furthermore, we improved our recommendations and make it easier to understand.

6 Recommendations

Industry main job-Employment by gender

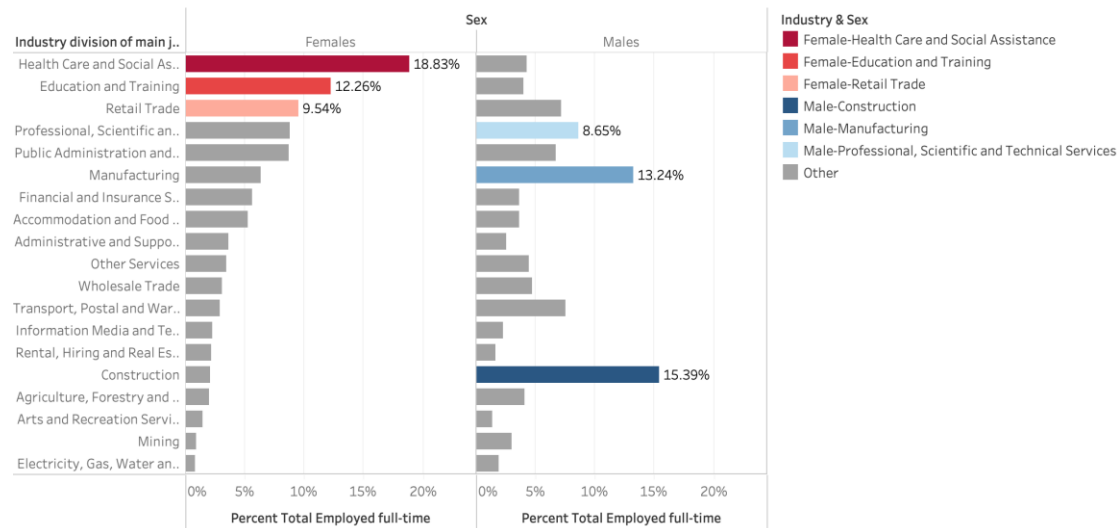


Figure 6.1 Graph4-Industry main job employment in Queensland

Industry - Employment

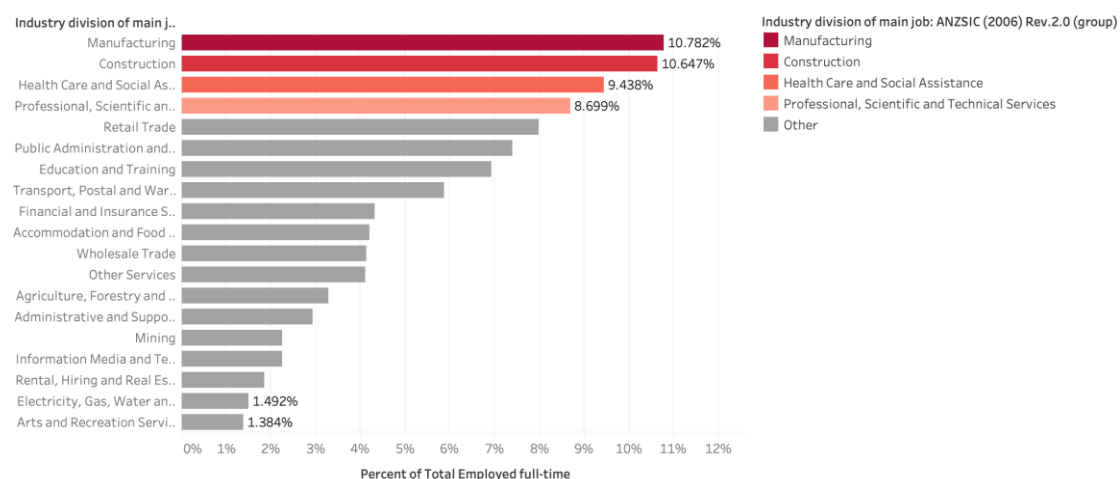


Figure 6.2 Graph5- Industry division of main jobs in Queensland

According to our research questions, we can answer that through an analysis.

- 1) Does Covid-19 have impact on unemployment of Queensland?
- 2) Are there other aspects that affect employment rate?

From the analysis, it shows that unemployment rate in Queensland is not affected by COVID-19 directly but from national policies. For another question, the graphs above can indicate that women tend to have the largest percentage in health care and social assistance and the next industry is the education and training. While for men, the most popular one is construction and the following one is manufacturing. In general, manufacturing has the most job opportunities which is almost 10.8% of all the industries. The second and third ones are construction and health care separately. Arts and recreation service has the least job chances.

From this information, job hunters could find more opportunities in different industries. Besides, students could use this information for considering academic degree based on interest or job employment.

Lastly, Government relevant departments could provide more job opportunities to unpopular industries which could improve our job-seek environment to be a more balanced one.

7 References

- [1] Queensland Government, Queensland COVID-19 case line list – location & source of infection, <https://www.data.qld.gov.au/dataset/queensland-covid-19-case-line-list-location-source-of-infection/resource/1dbae506-d73c-4c19-b727-e8654b8be95a>
- [2] Australia Bureau of Statistic, Labour Force, Australia, <https://www.abs.gov.au/statistics/labour/employment-and-unemployment/labour-force-australia-detailed/latest-release>
- [3] Tourism region profiles data tables, <https://www.tra.gov.au/regional/tourism-regional-profiles>
- [4] Wikipedia, COVID-19 Pandemic in Australia, https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Australia