



**Job Market Analysis Report  
Part 2**

Daniel Spindler - s5238197  
Gabriela Almeida Monteiro - s5198626  
Julio Pimentel Albores – s5172620

Griffith University  
7030ICT Introduction to Big Data Analytics  
Lecturer: Dr Henry Nguyen  
Assessment item: Written Assignment

Word count: 2465  
Due date: 24/08/2020

# 1. Data Analysis and Interpretation

## 1.1. Analyse by comparison

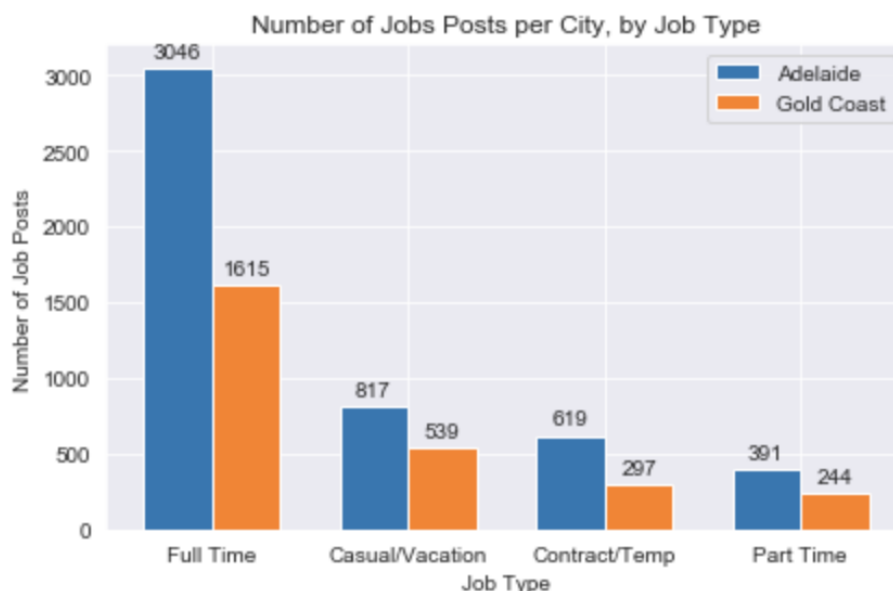
**Which city has more jobs? How many jobs each type (casual, fulltime, etc.) are there in each city?**

The bar chart below compares the cities of Adelaide and Gold Coast in terms of the number of job posts for four different job types. Overall, Adelaide has more job posts, and the ranking of job types is the same in both cities.

To begin, Adelaide is the city with the highest number of job posts, having 4,873 open positions. In contrast, the Gold Coast shows a total of 2,695 posts. The ranking of type of jobs in both cities follows the same pattern: full-time jobs occupy the first place in several job posts, followed by casual/vacation in the second place, contract/temp in the third place and part-time in the last position.

An explanation for the high number of Casual/Vacation jobs could be the time interval when the data was collected (October-November). It is the exact period when companies start to look for casual/vacation workers. Therefore, it could explain why this type of jobs occupies second place in the ranking.

It is essential to highlight that the analysis was done using the formatted and cleaned dataset from Assignment 2, part 1.



**In each city, which are top 5 job sectors? How many jobs are there in each sector?**

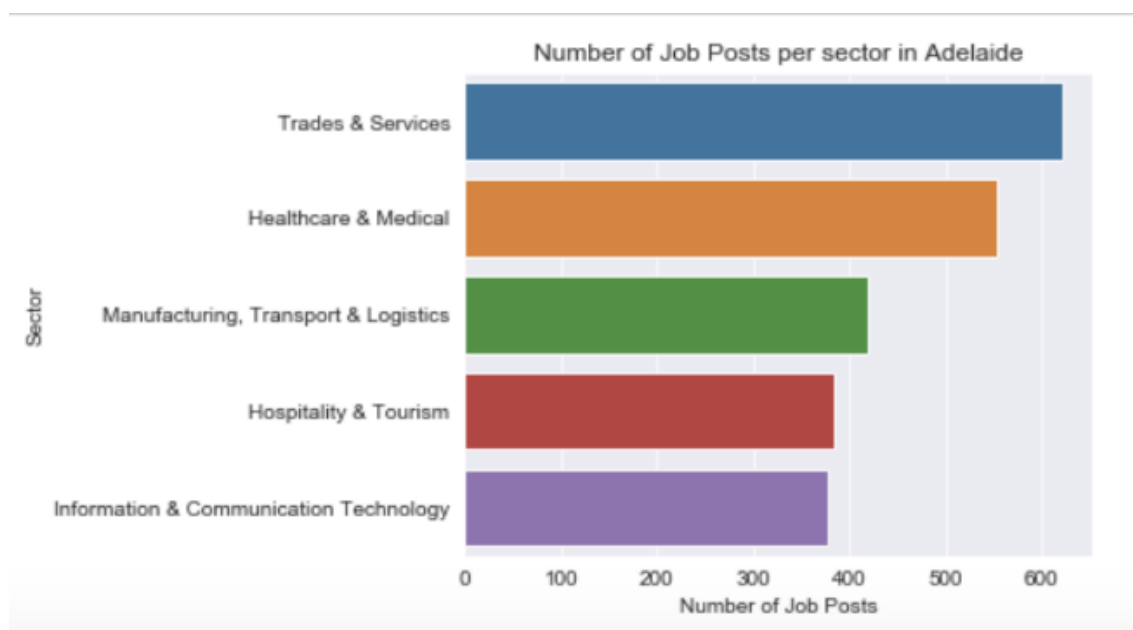
The tables below present the top five job sectors in Adelaide and Gold Coast and their respective number of job posts. Overall, it is possible to see that the top five sectors are very similar in both cities, with some slight variations.

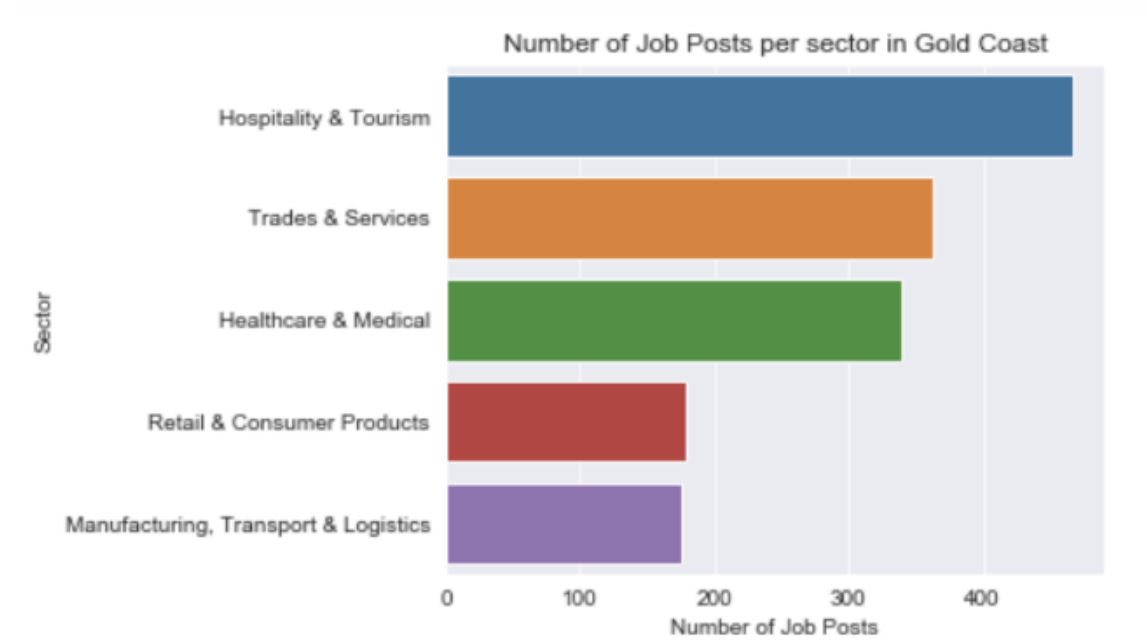
It is evident that “Trades & Services” and “Healthcare & Medical” are equally important in both locations, representing a good number of job posts. For instance, in Adelaide, “Trades & Services” and “Healthcare & Medical” occupy the first and the second positions of sectors with more job opportunities (621 and 554 job posts, respectively). In Gold Coast, the leading sector is “Hospitality & Tourism” (467 posts), but “Trades & Services” (363 posts) and “Healthcare & Medical” (339 posts) follow right after, occupying the second and the third places.

Despite the similarities, it calls attention that the only sectors that are not common to both locations are “Information & Technology” and “Retail and Consumer Products”. While “Information & Technology” is only present in the top five sectors in Adelaide, “Retail and Consumer Products” is only seen in the top five in Gold Coast.

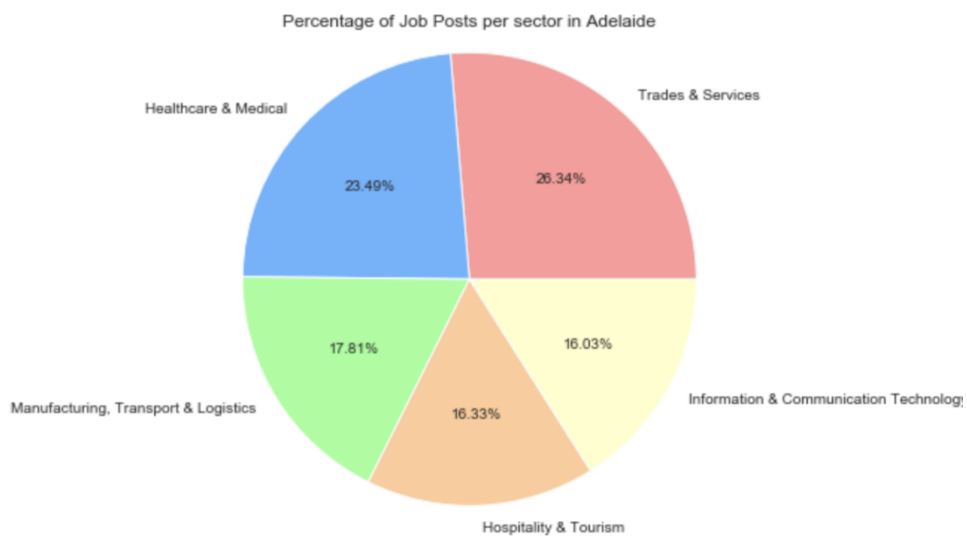
In Adelaide, “Manufacturing, Transport & Logistics” is the third sector with more job opportunities (420 posts), followed by “Hospitality & Tourism” in the fourth place (385 posts) and “Information & Communication Technology” in the last position (378 posts).

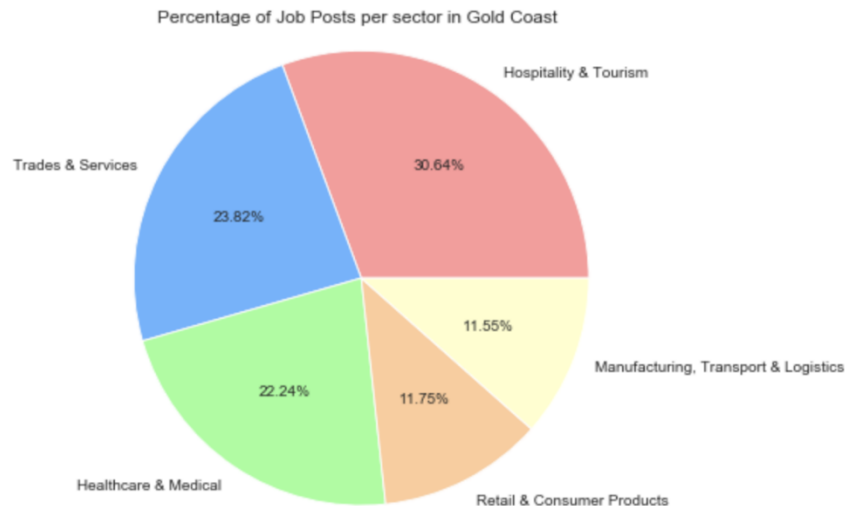
In the Gold Coast, “Retail & Consumer Products” occupies the fourth position (179 posts) and, finally, “Manufacturing Transport & Logistics” is in the last place (176 posts).





Visualise the top 5 job sectors in a pie chart for each city.





**In each city, list the job salary range with the corresponding number of jobs. Which city is more well-paid?**

Since it was required to list the job salary range with the corresponding number of jobs, the bin function was used. Additionally, a box plot is provided to visualise better the distribution of average salaries in Adelaide and Gold Coast.

Although in both locations, most of the job offers are for positions with low average salaries, it is evident that better salary options are available in Adelaide. As seen from the box plot, the median of the average salaries is higher in Adelaide, and the maximum salary is also higher. Therefore, it is possible to conclude that Adelaide is more well-paid than Gold Coast.

----- Adelaide -----

Number of Jobs by salary range in Adelaide

|                |      |
|----------------|------|
| (14.789, 50.0] | 2175 |
| (50.0, 85.0]   | 1467 |
| (85.0, 120.0]  | 788  |
| (155.0, 190.0] | 235  |
| (120.0, 155.0] | 212  |
| (190.0, 225.0] | 94   |

Name: AverageSalary, dtype: int64

----- Gold Coast -----

Number of Jobs by salary range in Gold Coast

|                |      |
|----------------|------|
| (14.789, 50.0] | 1524 |
| (50.0, 85.0]   | 778  |
| (85.0, 120.0]  | 261  |
| (155.0, 190.0] | 80   |
| (120.0, 155.0] | 69   |
| (190.0, 225.0] | 42   |

Name: AverageSalary, dtype: int64



### List top 5 companies in each city? Which sectors do they belong to?

It is assumed that the subclassifications refer to the type of job and not to the company itself. For this reason, it was found that one company can have multiple subclassifications. The top 5 companies in Adelaide and Gold Coast, as well as their respective subclassifications and number of job posts, can be seen in the table below.

| Adelaide  |                | Gold Coast  |                |
|---|----------------|---|----------------|
| Company/Classification  | Number of jobs | Company/Classification  | Number of Jobs |
| <b>Jora Local</b> <ul style="list-style-type: none"> <li>'Hospitality &amp; Tourism'</li> <li>'Retail &amp; Consumer Products'</li> <li>'Trades &amp; Services'</li> </ul>  | 109            | <b>Jora Local</b> <ul style="list-style-type: none"> <li>'Hospitality &amp; Tourism'</li> <li>'Retail &amp; Consumer Products'</li> <li>'Trades &amp; Services'</li> </ul>  | 139            |
| <b>S.A. Health</b> <ul style="list-style-type: none"> <li>'Healthcare &amp; Medical'</li> <li>'Government &amp; Defence'</li> <li>'Science &amp; Technology'</li> <li>'Administration &amp; Office Support'</li> <li>'Information &amp; Communication Technology'</li> <li>'CEO &amp; General Management'</li> </ul>  | 95             | <b>Gold Coast Hospital &amp; Health Service</b> <ul style="list-style-type: none"> <li>'Government &amp; Defence'</li> <li>'Administration &amp; Office Support'</li> <li>'Healthcare &amp; Medical'</li> <li>'Education &amp; Training'</li> <li>'Banking &amp; Financial Services'</li> <li>'Legal'</li> <li>'Hospitality &amp; Tourism'</li> </ul> | 64             |
| <b>Hays Talent Solutions</b> <ul style="list-style-type: none"> <li>'Banking &amp; Financial Services'</li> <li>'Retail &amp; Consumer Products'</li> <li>'Sales'</li> <li>'Call Centre &amp; Customer Service'</li> <li>'Insurance &amp; Superannuation'</li> <li>'Hospitality &amp; Tourism'</li> <li>'Mining, Resources &amp; Energy'</li> <li>'Accounting'</li> </ul> | 46             | <b>Hays Trades &amp; Labour</b> <ul style="list-style-type: none"> <li>'Trades &amp; Services'</li> <li>'Sport &amp; Recreation'</li> <li>'Construction'</li> <li>'Manufacturing, Transport &amp; Logistics'</li> </ul>   | 45             |

|   |    |   |    |
|---|----|---|----|
| <b>Excel Recruitment</b> <ul style="list-style-type: none"> <li>• 'Trades &amp; Services'</li> <li>• 'Sales' 'Manufacturing, Transport &amp; Logistics'</li> <li>• 'Information &amp; Communication Technology'</li> <li>• 'Construction'</li> <li>• 'Human Resources &amp; Recruitment'</li> <li>• 'Healthcare &amp; Medical'</li> </ul> | 41 | <b>Griffith University</b> <ul style="list-style-type: none"> <li>• 'Information &amp; Communication Technology'</li> <li>• 'Education &amp; Training'</li> <li>• 'Administration &amp; Office Support'</li> <li>• 'Healthcare &amp; Medical'</li> <li>• 'Banking &amp; Financial Services'</li> </ul>  | 28 |
| <b>Randstad - Industrial</b> <ul style="list-style-type: none"> <li>• 'Manufacturing, Transport &amp; Logistics'</li> <li>• 'Trades &amp; Services'</li> <li>• 'Government &amp; Defence'</li> <li>• 'Human Resources &amp; Recruitment' '</li> <li>• 'Accounting'</li> </ul>   | 36 | <b>City of Gold Coast</b> <ul style="list-style-type: none"> <li>• 'Government &amp; Defence'</li> <li>• 'Engineering'</li> <li>• 'Information &amp; Communication Technology'</li> <li>• 'Education &amp; Training'</li> <li>• 'Design &amp; Architecture'</li> <li>• 'CEO &amp; General Management'</li> <li>• 'Sport &amp; Recreation'</li> <li>• 'Consulting &amp; Strategy'</li> </ul> | 28 |

**Between 2 cities, which do you think is better for employees. Explain your choice.**

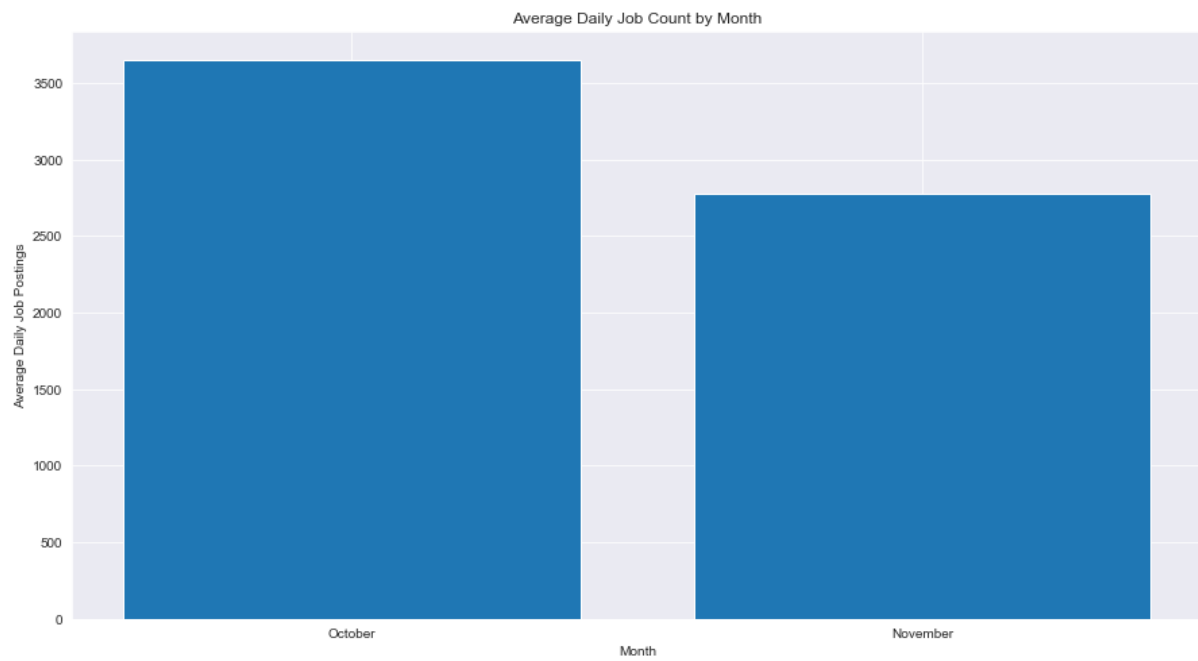
When analysing which city is better for employees, it is taken into consideration the probability of getting a job in each city. Therefore, the analysis observes the total number of job posts disregarding if the job is for white or blue-collar.

In general, because the city of Adelaide is bigger than Gold Coast, it is expected to have more job opportunities available and, also, more options for higher salaries. Therefore, if someone is looking for a place with more chances to get a job and be well remunerated, Adelaide is a better option than Gold Coast.

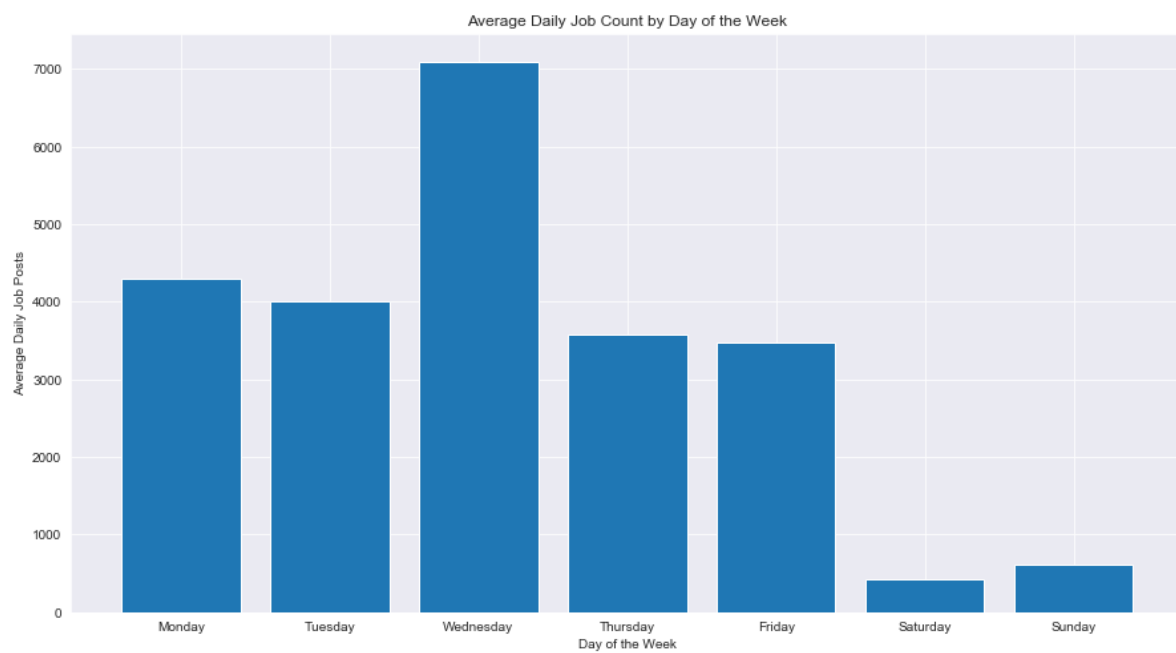
However, it is relevant to highlight that, when it comes to employability in the sector of "Hospitality & Tourism", Gold Coast can be a better choice than Adelaide due to the higher number of job posts for this sector. On the other hand, Adelaide offers far more opportunities for those pursuing a career in the sectors of "Trades & Services", "Healthcare & Medical", "Manufacturing, Transport & Logistics" and "Information & Communication Technology".

## 1.2. Analyse by time

### Visualise the number of job posts by month

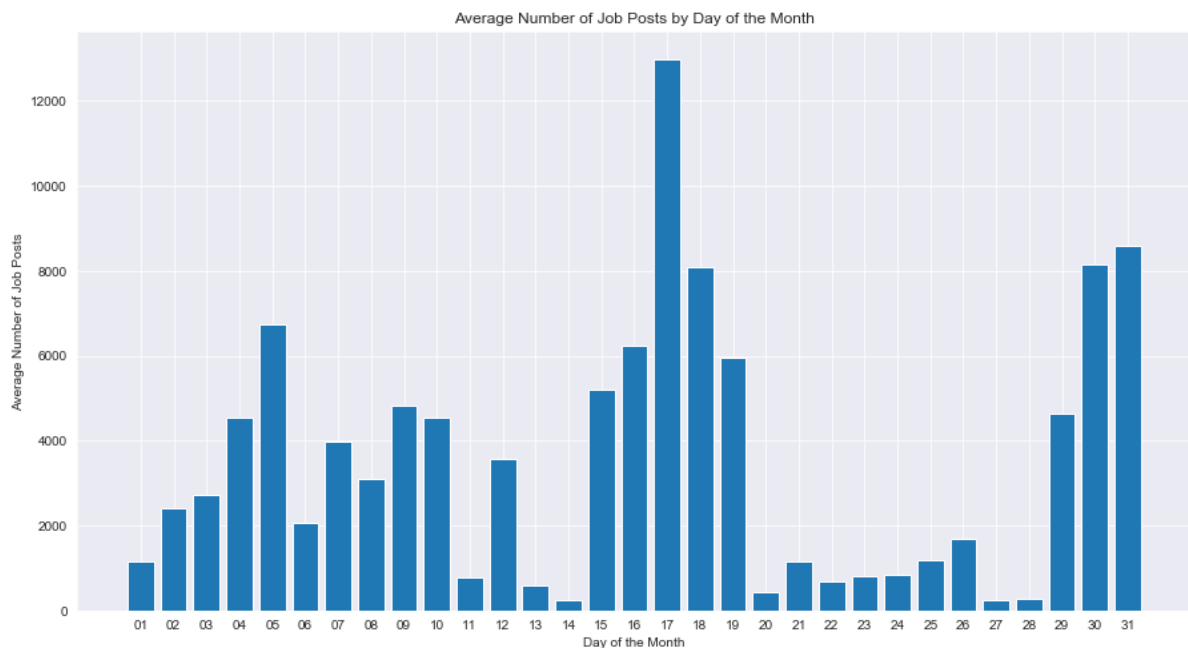


### Visualise the number of jobs posts by day of the week

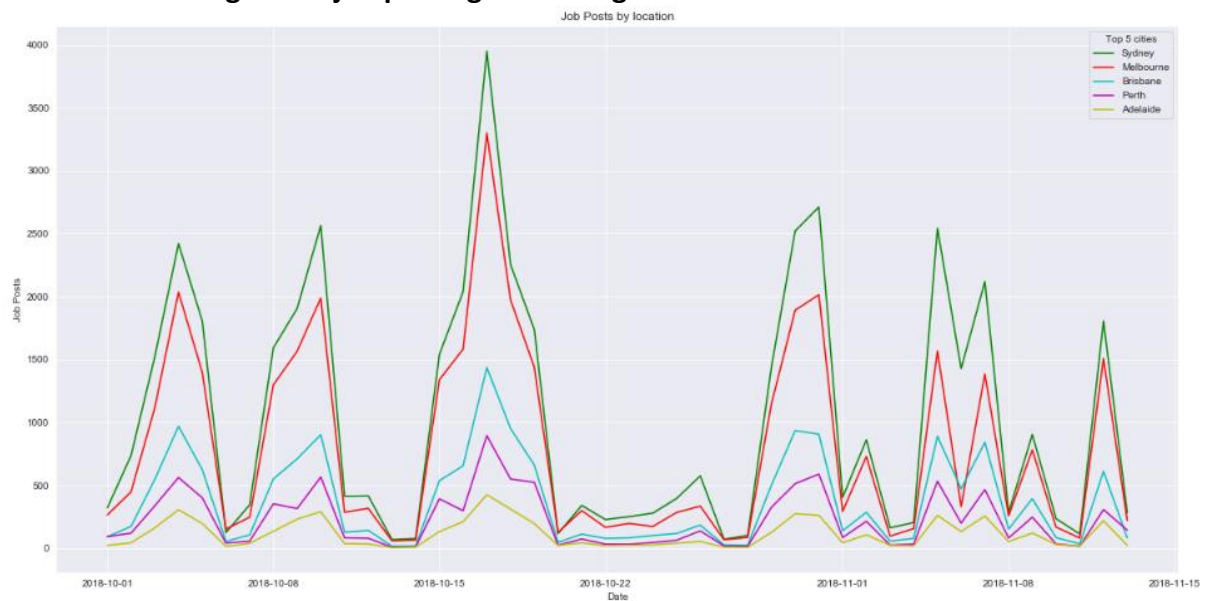




## Visualise the number of job posts by day of the month



## Visualise trending of the job posting for the big cities



**Based on the above charts, provide your observation about the number of job posts over time.**

The data covers a period of six weeks and one day, with data for the full month of October but only slightly less than half of November. With this in mind, a decision was made to average the job posts to a daily figure to compare across months, days of the week and day of the month to ensure meaningful insight.

The average daily job post count was significantly higher in October than November.

Most jobs are posted on a weekday, with Wednesday, on average, the most popular day to post a job, with the weekend showing very few job posts.

Averaging the daily job posts by day of the month shows a significant spike in the middle of the month, with a reasonably quiet period from the 20th to the 28<sup>th</sup> with very few job posts.

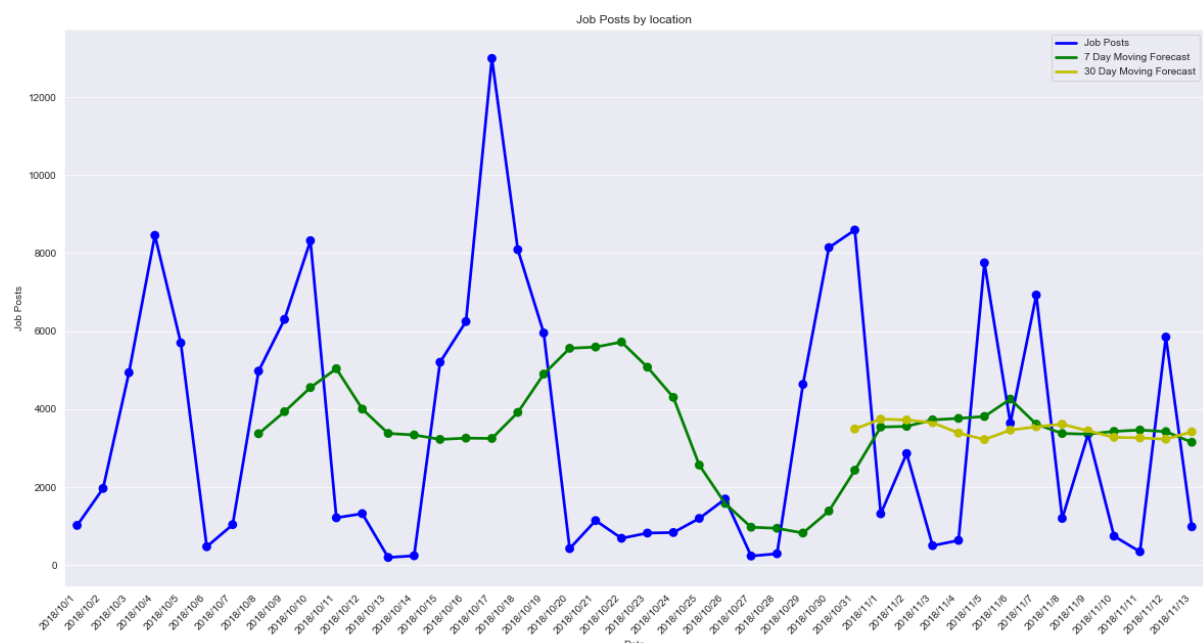
The top 5 cities follow the same trends regarding posting of jobs over time. However, there is a noticeable difference in the number of posts per city. Trend analysis further illustrates the low number of job posts, in the top 5 cities, in the latter part of October.

In conclusion, the most defining feature of job posts over time is the harmonious trend of all top 5 cities over the period and Wednesday being the most popular day of the week a job was posted.

### 1.3. Forecasting and skill extractions

**Using moving average for 7 days and 30 days to predict the number of job postings and visualise them in line chart. Which one creates a better prediction?**

The following figure presents daily plots of the job posts, moving average forecast for 7 and 30 days. Given that the dataset has 44 different dates, the moving average forecast for 30 days only presents 14 points. The moving average forecast for seven days predicts better than the 30 days. The moving average for 30 days does not show seasonal trends that the dataset has.



**Choose your favourite job sector/sub-sector, then use TF/IDF to extract important keywords. Visualise them in word cloud chart.**

Our group selected the variable “Information & Communication Technology” in the “Classification” column to make the TF/IDF analysis and extract relevant keywords on the “Title” column. A word cloud chart was created with the 40 most relevant words obtained from the “Title” column. Words such as “developer”, “analyst”, “senior”, “manager”, and “engineer” were the most relevant.



## 2. Discussion

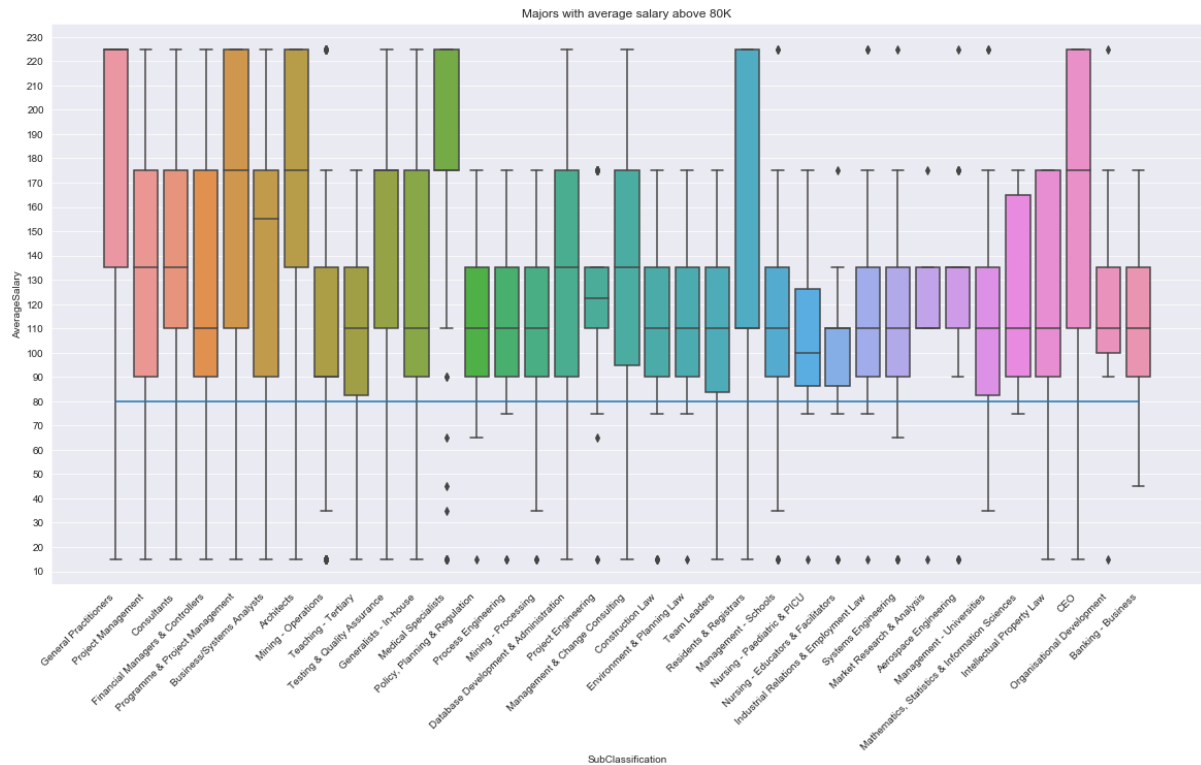
### 2.1. Scenario 1

**Tom, a grade 12 student in Brisbane, has good results in all of the subjects in his school. He is finding a major in the University which can guarantee a job in his state (Queensland) with a good income ( $\geq 80K$ ) in the future. Based on the current job market dataset, which major and the related subjects do you recommend him?**

Considering that Tom wants to find a job in the state of Queensland with a salary of 80K or more, it was necessary to establish filters for these two requirements. Therefore, the first step was to filter all job posts in the state of Queensland. After that, they were grouped by subclassification, and the 25% quantile of the average salary was calculated. In addition, the number of job posts was counted, and only those with more than five jobs posts and average salaries of 80K or more were retrieved. By doing that, it was possible to visualise which subclassifications had more job posts. In a sequence, a group by function was used to check the number of job posts per classification. As a result, it showed that the five sectors with more probability of finding a job are “Information & Communication Technology”, “Healthcare & Medical”, “Mining, Resource & Energy”, “Engineering” and “Construction”, in this order. These classifications can be understood as the types of courses that he could choose at the university.

When looking at the subclassifications, it is apparent that in “Information & Communication Technology” many job posts are “Programme & Project Management”, “Business/Systems Analysts”, “Architects”, “Testing & Quality Assurance” and “Consultants”. As for “Healthcare & Medical”, the subclassifications that have more demand are “General Practitioners” and “Medical Specialists”. On the other hand, “Mining, Resource & Energy” has a high demand for “Mining – Operations”, whereas “Engineering” and “Construction” have both high demand for “Project Management”. These subclassifications can be interpreted as the major or specialisation that Tom should look for during his academic studies.

Based on this analysis, it is evident that STEM courses are in high demand in the job market and that Tom would achieve his goals if aiming to study one of these types of courses (sectors) and majors (subclassifications) mentioned above. Our recommendation is based that given this dataset, these majors and related subjects have 75% of jobs above A\$80,000 annually with more than five job posts in 44 days.



#### Classification

|  |     |
|--|-----|
| Information & Communication Technology | 993 |
| Healthcare & Medical                   | 408 |
| Mining, Resources & Energy             | 336 |
| Engineering                            | 188 |
| Construction                           | 179 |

#### SubClassification

|                                |     |
|--------------------------------|-----|
| Mining - Operations            | 303 |
| Programme & Project Management | 279 |
| Business/Systems Analysts      | 268 |
| General Practitioners          | 241 |
| Project Management             | 239 |

## 2.2 Scenario 2

To gain a better reputation, Griffith University wants to attract more students by

enhancing the employability of the graduates. As a data analyst, you need to give the

recommendation for the management board for the changing of the admission numbers in each major. Which majors should be extended, and which majors should be reduced (Assume that we only need to use job market data, regardless of other factors such as social impact, economic impact, etc. )? Justify your recommendation.

An assumption has been made that employability means a higher salary while also obtaining stable work such as a full time or contract role. Therefore the data was filtered by these job types.

Each subclassification, has had the 25% quantile and the 75% quantile of the average salary calculated, along with the job count of that subclassification. Analysis can now be conducted on the average salary and the job count of subclassifications. We are assuming that subclassifications will align with majors. It should be noted that results with job posts less than 7 were filtered out.

To begin analysing majors for reduced intake we first sort by job post count to obtain the 50 bottom results. For this case we sort by 75% quantile of average salary, taking the top 15, results with low average salary and low job count should be noted for reduced intake. A boxplot is used to visually confirm the low average salary of the results.

Conversely, we move on to use the same logic to look for majors to extend. The top 50 job count results are used to then sort by 25% quantile of average salary and taking the top 15. These 15 subclassifications have a high job count and high salary and therefore should have extended intake for their majors. A boxplot is again used to visualize the average salary.

It was then decided to take weekly job count of classifications and apply linear regression to assess the job count trend over time. This would give an indication of job count increasing, decreasing or being stable. We should note that we only have 6 weeks of data and therefore trend analysis is limited. Regression was also applied to the total job count and showed a very small overall negative trend in job count. If we had a larger dataset we would give more weight to subclassifications that have a positive trend with job count, as more available jobs, result in more likelihood of a student gaining employment in that field. Whilst it was noted that many of the extended recommendations did have a negative trend, given the short timeframe of the data it was not given enough weight to change the decision.

We therefore conclude that the following subclassification and their associated majors are marked for increased intake:

| SubClassification                |
|----------------------------------|
| General Practitioners            |
| Architects                       |
| Programme & Project Management   |
| Financial Managers & Controllers |
| Consultants                      |
| Engineering - Software           |
| Developers/Programmers           |
| Business/Systems Analysts        |
| Project Management               |
| Foreperson/Supervisors           |

The following should be marked for reduced intake:

| SubClassification                |
|----------------------------------|
| Tailors & Dressmakers            |
| Airlines                         |
| Gaming                           |
| Nannies & Babysitters            |
| Photography                      |
| Performing Arts                  |
| Public Transport & Taxi Services |
| Locksmiths                       |
| Tour Guides                      |
| Promotions                       |