



**BSD1323 STORYTELLING AND DATA VISUALIZATION**  
**INDIVIDUAL PROJECT REPORT**

**TOPIC: ANALYSIS OF EMPLOYEE IN MALAYSIA 1962-2021**

**LECTURER: DR. SITI ZANARIAH BINTI SATARI**

**NAME: ALMIRA DAMIA BINTI SYAHNIZAM**


**STUDENT ID: SD22002**

**SECTION: 01G**

**SUBMISSION DATE: 13 JUNE 2022**

## **TABLE OF CONTENTS**

<b>CONTENTS</b>	<b>PAGE</b>
<b>MARKING SCHEMES</b>	<b>3-5</b>
<b>1.0 MOTIVATION OF PROJECT</b>	<b>6</b>
	<b>7</b>
<b>2.0 EXPLANATION OF DATASETS</b>	
<b>3.0 VISUALISATION AND DASHBOARD ANALYSIS</b>	<b>8</b>
<b>3.1 ANALYSIS VISUALISATION</b>	<b>8</b>
3.1.1 MAP SUM OF EMPLOYEE BY INDUSTRY	8
3.1.2 Rank Sum of Employee by Industries	9
3.1.3 Average Number of Employee from 1982-2021	10-11
3.1.4: Average Number of Employee and Forecast for the next 5 years	11-12
3.1.5: Crosstab by Highlight Table of Marital Status	12-13
3.1.6: Donut Chart of Percentage of Employee by Age Group And Year	13
3.1.7: Number of Employee and its Cumulative by Age Group	14
<b>3.2 EXPLANATION OF DASHBOARD</b>	<b>15</b>
3.2.1 Main Dashboard	15
3.2.2: Industry Dashboard 1 and 2	16-17
3.2.3: Industry Dashboard 3	17-18
3.2.4: Marital Dashboard	18-19
3.2.5: Age Dashboard	19
<b>4.0 CONCLUSION</b>	<b>20</b>

 اونیورسیتی ملیسیا قهق UNIVERSITI MALAYSIA PAHANG PUSAT SAINS MATEMATIK	<b>SUBJECT: BSD1323 STORYTELLING AND DATA VISUALIZATION</b>		<b>MARKS:</b> 60(15%)
	<b>TOPIC: CHAPTER 1 to CHAPTER 8</b>		
	<b>INDIVIDUAL PROJECT</b>	<b>DUE DATE: 27 May - 13 June 2023</b>	
	<b>ID: SD22002    NAME: ALMIRA DAMIA BINTI SYAHNIZAM    SECTION: 01G</b>		

### INDIVIDUAL PROJECT: MARKING SCHEME

CLO	Description	PLO mapping	Percentage	Marks
CLO2	Demonstrate the data visualization skill using an effective storytelling.	PLO2: Cognitive Skills and Functional work skills with focus on Numeracy skills C3: Application	5%	20

CLO2 RUBRICS OF QUESTION 4								
CRITERIA	LEVEL OF ACHIEVEMENT						WEIGHTAGE	SCORE
	0	1 Inadequate	2 Emerging	3 Developing	4 Good	5 Excellent		
<b>Motivation of project topic</b>	No motivation of the project topic provided	Very little motivation of the project topic provided	Motivation of the project topic provided but missing all major points	Motivation of the project topic provided but unclear	Clear and good motivation of the project topic provided	Very clear and excellent motivation of the project topic provided	0.5	
<b>Details explanation of the dataset</b>	Failed to explain the dataset	Not Efficiently, effectively, and accurately explain the dataset	Partly accurate, but not effectively explain the dataset	Effectively explain the dataset but not accurate	Accurately and effectively but not efficiently explain the dataset	Accurately effectively, and efficiently explain the dataset	0.5	
<b>Details analysis of each dashboard</b>	Failed to analyse the dashboards	Not Efficiently, effectively, and accurately analyse the dashboards	Partly accurate, but not effectively analyse the dashboards	Effectively analyse the dashboards	Accurately and effectively but not efficiently analyse the dashboards	Accurately effectively, and efficiently analyse each dashboard	2	
<b>Concluding remarks</b>	No concluding remarks provided	Very little concluding remarks provided and inaccurate	Concluding remarks provided but unclear and inaccurate	Concluding remarks provided but partly inaccurate	Clear and good concluding remarks provided	Very clear and excellent concluding remarks provided	1	
<b>TOTAL (20)</b>								

CLO	Description	PLO mapping	Percentage	Marks
CLO3	Display a powerful data visualization, report, dashboard or stories in solving various applications using appropriate software.	PLO3: Functional work skills with focus on Practical, and Digital skills P4: Mechanism	10%	40

CRITERIA	LEVEL OF ACHIEVEMENT						WEIGHTAGE	SCORE
	0	1 Inadequate	2 Emerging	3 Developing	4 Good	5 Excellent		
<b>Theory/ Knowledge on data visualization</b>	No theoretical knowledge on data visualization observed.	Very little knowledge observed on data visualization or some information is incorrect.	Some knowledge or information on data visualization observed but missing all major points.	Some knowledge or information on data visualization observed but still missing some major points.	Good knowledge on data visualization observed, missing some minor points.	Excellent knowledge on data visualization observed; provides all necessary background principles.	1	
<b>Interactive Data Visualization Techniques</b>	Failed to demonstrate the given task.	Inappropriate interactive data visualization techniques are demonstrated.	Partly correct interactive data visualization techniques are demonstrated, with partly valid data.	Correct interactive data visualization techniques are demonstrated, with partly valid data.	Good interactive data visualization techniques are demonstrated, with valid but not completely accurate data.	Competent interactive data visualization techniques are demonstrated, with valid and accurate data.	1	
<b>Theory/ Knowledge on advanced dashboard</b>	No theoretical knowledge on advanced dashboard observed.	Very little knowledge observed on advanced dashboard or some information is incorrect.	Some knowledge or information on advanced dashboard observed but missing all major points.	Some knowledge or information on advanced dashboard observed but still missing some major points.	Good knowledge on advanced dashboard observed, missing some minor points.	Excellent knowledge on advanced dashboard observed; provides all necessary background principles.	1	
<b>Advanced Dashboard Techniques &amp; Data Validation</b>	Failed to demonstrate the given task.	Inappropriate advanced dashboard techniques are demonstrated.	Partly correct advanced dashboard techniques are demonstrated, with partly valid data.	Correct advanced dashboard techniques are demonstrated, with partly valid data.	Good advanced dashboard techniques are demonstrated, with valid but not completely accurate data.	Competent advanced dashboard techniques are demonstrated, with valid and accurate data.	2	

<b>Efficiency/ Assembly/ Tidiness</b>	Failed to demonstrate the given task.	Not efficiently, effectively and neatly demonstrated the given task.	Partly efficient, but not effectively and neatly demonstrated the given task.	Efficiently, but not effectively and neatly demonstrated the given task.	Efficiently and effectively but not neatly demonstrated the given task.	Efficiently, effectively and neatly demonstrated the given task.	1	
<b>Final Results (the advanced dashboard)</b>	Failed to demonstrate the given task.	Lack of results/zero readability of the result. Poor originality.	Partly complete result and less originality.	Result presented but at low readability/some result presented. Reader has to guess some of the missing information. Less originality.	Clear, neat presentation. All required results are presented. Readability. Complete with labels, title, axes, etc.	Very Clear, neat presentation. All required results are presented. High readability. Complete with labels, title, axes, etc.	2	
							<b>TOTAL (40)</b>	

## 1.0 MOTIVATION OF PROJECT

This project is about an analysis of employee focusing on Malaysia from 1982-2021. The number of employees in Malaysia are divided into several categories which are industry they work in, marital status and group age. This topic is chosen because, in the near future I will also be applying to one of the industries as an employee. Therefore, by conducting an analysis on this topic it can provide me the data representation of employee and its pattern of employment throughout the year stated. Other than that, being able see the impact of employment since Covid-19 happen in 2019 is one of the patterns I would like to focus on in the visualisation. Thus, to understand the relationship between Malaysian workforce with these factors throughout the past few years is another objective of this project.

The industry factor chosen will guide me in choosing which industry needs more employee also taking in consideration of the future economy, but as a data analyst I am aware that even I will be needed in many industries as we provide the data on prediction of future, explore the characteristics of a product or patterns, some industries take serious about their employee's background as it may affect their work. Marital status provides me insights on how many people with different backgrounds are employed throughout the years. While age group give me notes that at what aged people start or proceed to gain their experience whether it's for the monetary as of the increasing of economy or to gain more skills and acquaintance. Transforming the data collected into a graphic structure is the best way to discover patterns.

The data source is from a Malaysia statistic website which is <https://data.gov.my/>. Firstly, I connected the tableau and my data through Microsoft excel connection. Initially, the three categories of data are from three different excel files, so multiple connection has been used to connect the data together. In order to do the visualization, I joint the data using inner join for marital status and age group also for marital status and industry. The common data for both joins are year and the states in Malaysia. Data interpreter is used to get rid of unwanted rows that is in common among the fields. Now, the data is ready for further analysis by doing visualisation of the data.

## 2.0 EXPLANATION OF THE DATASETS

The foundation of my visualisation consists of several data that enables me to do compelling insights. The dataset is divided into types which is discrete dimension which are usually the blue fields and continuous measure which are usually green fields. There are 5 fields in the dimension shelf and 3 fields in the measure shelf excluding the auto generated values by tableau. The auto generated values are measure names, measure values, longitude and latitude and the count.

Dimension shelf is usually for the categorical or qualitative data, there are 5 data that I will explain. Firstly, the marital status, it consists of categories of marriage which are married, never married, widowed, divorced/permanently separated. It is a string data type under categorical data. Next, the industry it consists of several amount of industry over Malaysia. For example, manufacturing, education and many other industries, it is also a string data type under categorical data. The age group falls under the categorical data with string data type. It consists of age group from 15 to 64 with a width of 4 age for each group such as 15-19, 20-24 and others. The states are under geographical data where it consists of states all over Malaysia such as Pahang, Melaka, Selangor and others. Lastly, the year it is a numerical data type which is in whole number as it contains data from 1982 until 2021. Additionally, there are a few calculated values in this shelf that has been added throughout doing the visualisation such as return or next, go to and many more which are used for the navigation between dashboards.

On the other hand, Measure shelf contain data that are usually numerical data where there are 3 data in this shelf that are from the data source. The number of employees by marital status is in decimal number format, the number of employees by industries is in whole number format and number of employees by age group is in decimal number format. There are a few calculated values that I added in the process of making the visualisation. The calculated value in this shelf for example, age rank and top 5 age that are used for cumulative bar, running sum, running total (% of total) and rank of running total employee are used for the running bar chart, lastly the calculated field named 1 which is used for the pie chart visualisation.

### 3.0 EXPLANATION OF VISUALISATION AND DASHBOARD

#### 3.1 ANALYSIS FOR VISUALISATION

##### 3.1.1 Map Sum of Employee by Industry

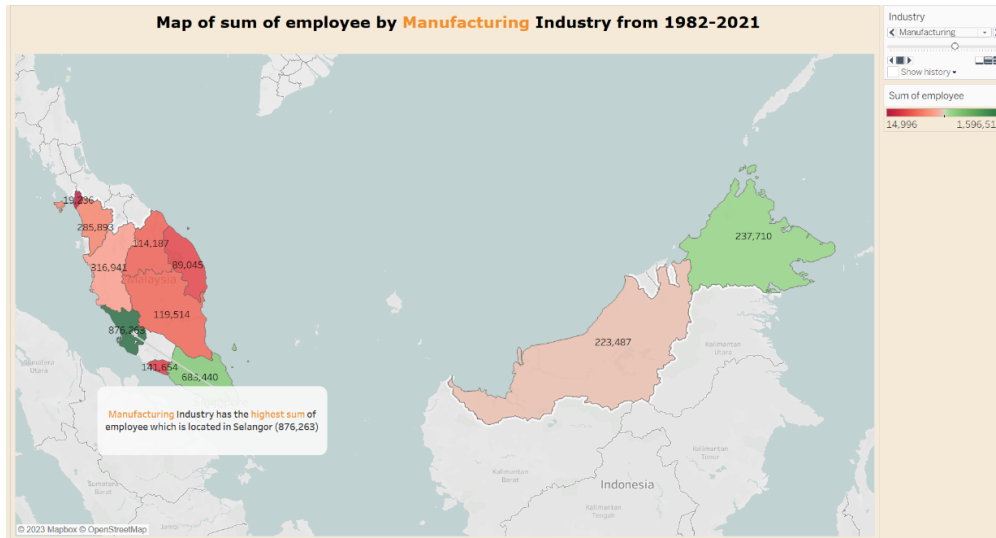


Figure 3.1.1: Visualisation 1

##### Analysis:

The highest sum of employment is in the manufacturing sector throughout the year from 1982-2021 which is located in Selangor. This can be seen from the heat map above as the Selangor state is in a dark green colour. The sum of employee is 876,263 people. Employed in the manufacturing industry is well-known to have high potential salary and also side benefits such as health insurance and retirement funds. The living cost in Selangor is quite high as it is the capital city of Malaysia. Thus, this industry gives the highest sum of employee for each state due to the reason mentioned above. Using the slider to change the industry type, it can be found that the lowest sum of employee is in the real estate activities which is located in Perlis. This is due to the small population of people in Perlis as it is said that Perlis is the smallest country in Malaysia. Thus, not so many people will have to buy houses which leads to the small number of employees in that state.



### 3.1.2 Rank Sum of Employee by Industries

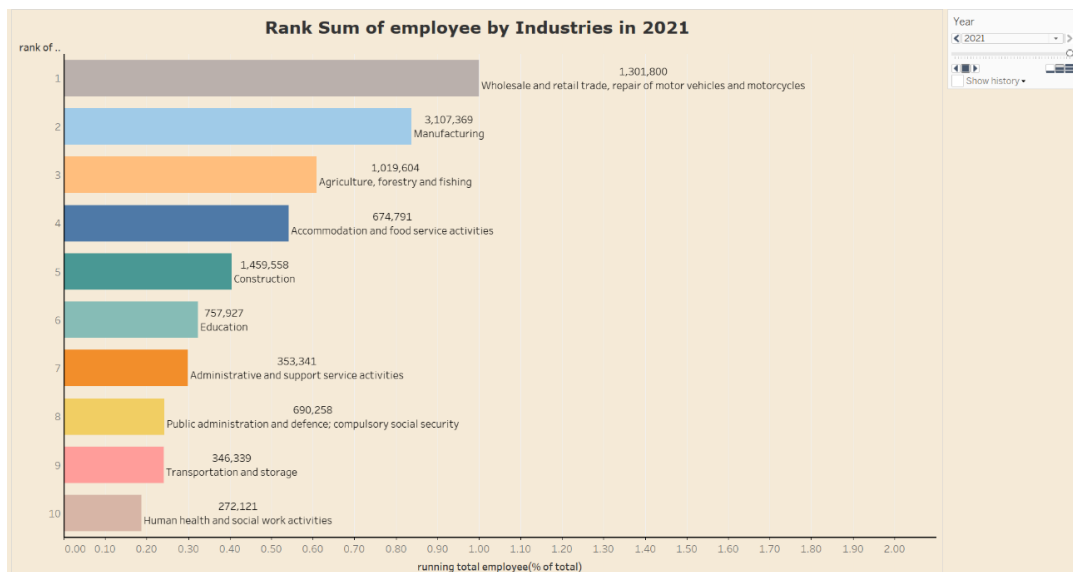


Figure 3.1.2: Visualisation 2

#### Analysis:

In the year 2021, the highest rank for sum of employee by industry according to the running total employee (percentage of total) is the wholesale and retail trade, repair of motor vehicles and motorcycles is in the highest rank. Even though, the sum number of employment in manufacturing are more than the sum of wholesale retail trade, but the percentage total is highest which is 1.0 where it is the percentage of all 2021 employee. This is due to, the recovery of Malaysia from the virus of Covid-19 where mostly people who work at factories become unemployed which explain why the percentage total is less in manufacturing industry. The lowest rank is the human health and social work activities. It holds the 10<sup>th</sup> place as, when the virus happened, people are advised to not go out unless for emergency reasons. This industry requires lots of physical meetings with people which in the year of 2021 we are still in the recovery situation cases sometimes begin to rise back at some point of the year.

### 3.1.3 Average Number of Employee from 1982-2021

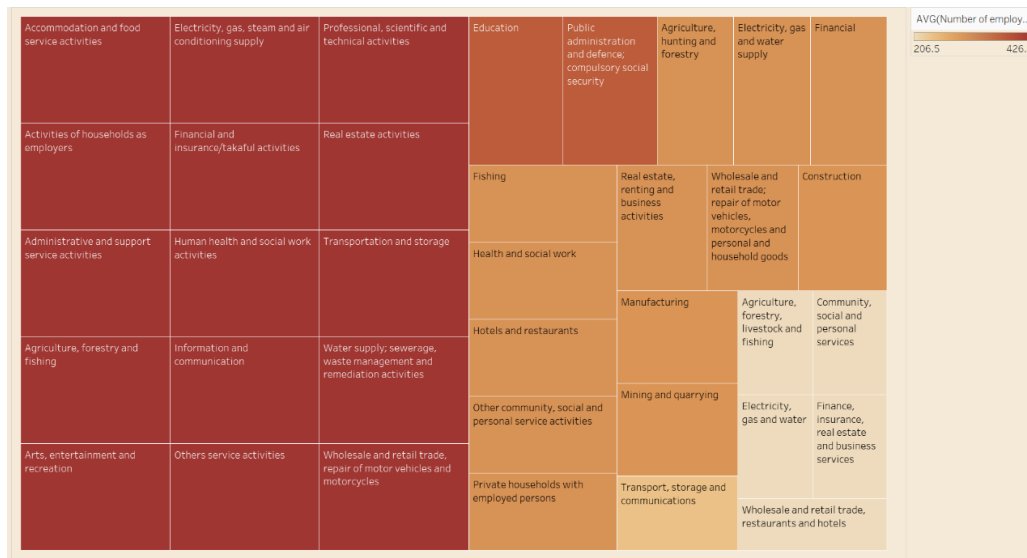


Figure 3.1.3: Visualisation 3

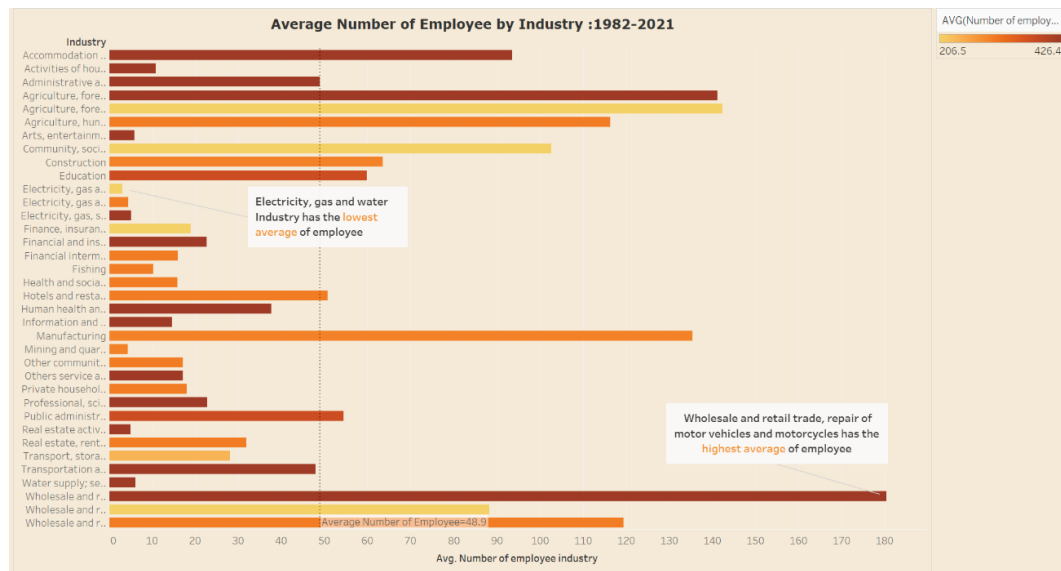


Figure 3.1.4: Visualisation 4

#### Analysis:

Visualisations 3 and 4 are the same as both of it shows the pattern of average number of employees in each industry throughout the year 1982-2021. The treemaps in figure 3.1.3 is used as an interactive way for the view to click on when it is implemented in the dashboard. The larger the square of the dimensions of the industry the higher the average number of employees. In figure 3.1.4, there are 13 industries that exceeds the average number of employees for overall industry. The highest average of employee is in the wholesale and retail trade industry, repair motor vehicles and motorcycles. Wholesale and retail trade are one of the most popular industries as it includes the business whether it its physical or online. The fact that it can be done through online many people during the Covid-19 joined this industry as they movement are restricted, and some might have lost their jobs due to the economy that has decrease for a while. This has become a job that

people can earn side income from home as it also can help promote work life balance. The lowest average number of employees is in the electricity, gas and water industry. This is due to the high barriers of entry where there is restriction for people to work there. People nowadays most likely prefer a free and clean environment to work in.

### 3.1.4: Average Number of Employee and Forecast for the next 5 years

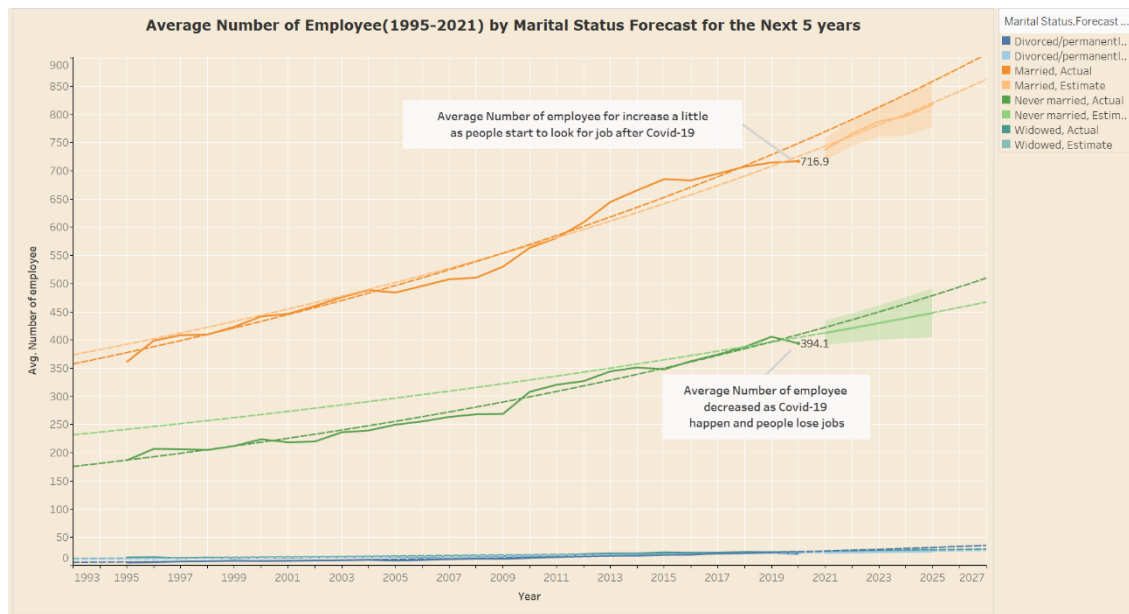


Figure 3.1.5: Visualisation 5

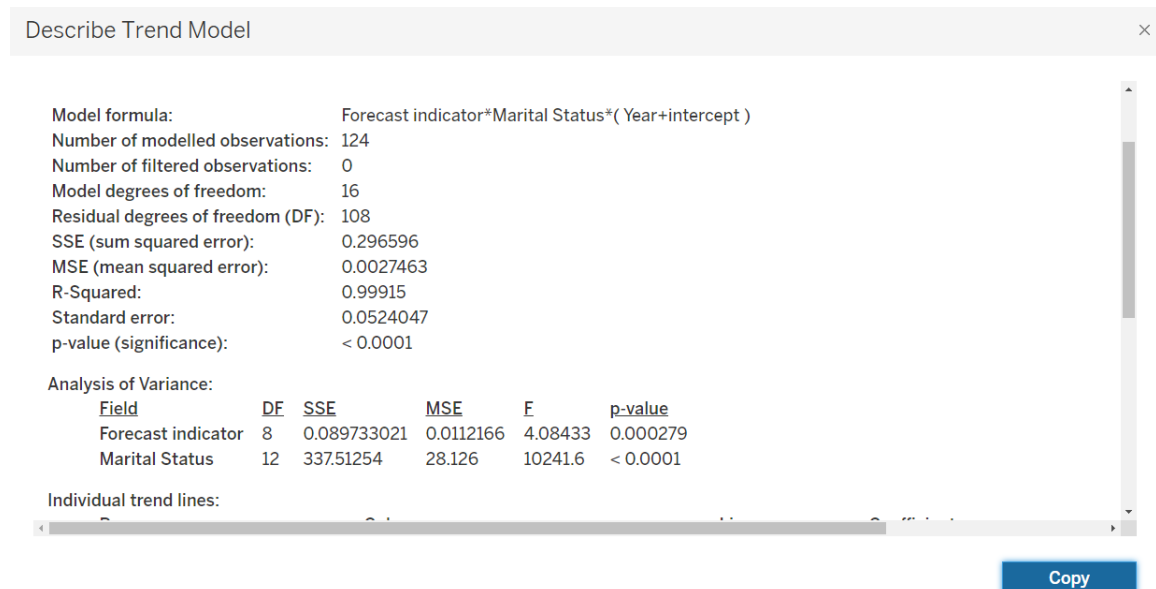


Figure 3.1.6: Description of Model Trend

### Analysis:

The visualisation in figure 3.1.5 is a time series chart with a forecast of 5 years. This chart shows the relationship of number of employees by marital status throughout the past few years. An exponential trend line is added to the chart where both of the trend line for marital status and forecast indicator are significance model as the p-value < 0.05 based

on figure 3.1.6. From 1982 to 2021 only two marital status that are gradually increasing which are the employees that are married and never married. This is due to the fact that, people with families will need to work hard as they need to pay off bills also tuition fees of their children. Not to forget that the increasing economy of this fast-paced world is encouraging the people who only need to support themselves to also be employed. The marked number for both categories is in the year of 2020 where the virus of Covid-19 has just entered the final phase of epidemic. People start to lose jobs as some businesses or workplace has to close due to the fact that the company cannot pay their staff. Additionally, the forecast is added to see how many people from each category will be employed so that I can know with whom I am competing to gain a job or whom I will be working with in the future. As for knowing the background may be an important point where it might interfere with the work performance.

### 3.1.5: Crosstab by Highlight Table of Marital Status

states	Divorced/permanently separated	Married	Marital Status Never married	Widowed	Grand Total
JOHOR	101,060	4,709,337	2,687,771	148,423	15,293,527
KEDAH	72,719	2,686,102	1,266,976	105,293	8,262,251
KELANTAN	71,484	2,014,118	836,184	90,335	6,024,406
MELAKA	29,427	1,105,657	621,675	37,973	3,589,392
PAHANG	42,305	2,041,815	1,045,853	77,271	6,414,557
PERAK	74,635	3,058,670	1,732,912	124,391	9,981,385
PERLIS	10,342	327,785	134,270	12,595	970,037
SABAH	124,190	4,562,565	2,562,187	143,244	14,784,561
SARAWAK	111,257	3,749,351	1,851,590	121,015	11,666,510
SELANGOR	162,733	8,162,533	4,483,252	215,517	26,048,707
TERENGGANU	40,661	1,333,719	660,916	53,737	4,178,184
Grand Total	840,812	33,751,652	17,883,587	1,129,795	107,213,516

SUM(Number of emplo...  
10,342 8,162,533

industry forecast marital total marital % employee age age cumulative goto industry goto marital goto age return next Main industry1 industry2 industry3 marital Age

Figure 3.1.6: Visualisation 6

#### Analysis:

The highest number of employees by marital status is the married employee while the lowest number of employees by marital status is widowed. This is due to the fact that the married employee will need to find extra support to buy basic needs of life for each member of the family while the widowed is the least as they may prefer to work alone then working with people in industries. The highest total of employee in the married status is located in Selangor while the lowest is located in Perlis. This is logical as these two states have very different size of area and different population of people. Some companies or industries prefer to take consideration of married men as they believe that men are the leader of the family so by that, they will have to have an income to support his family. This can be related to the fact that, in 2019 the virus started so some industries or companies will have to cut off some people, and the ones that are left are usually the experienced ones

and married as due to the monetary issue that will be needed at home. If widowed are employed, it may become mentally draining for them as they have lost someone that is important in their lives which may affect their work performance.

### 3.1.6: Donut Chart of Percentage of Employee by Age Group and Year

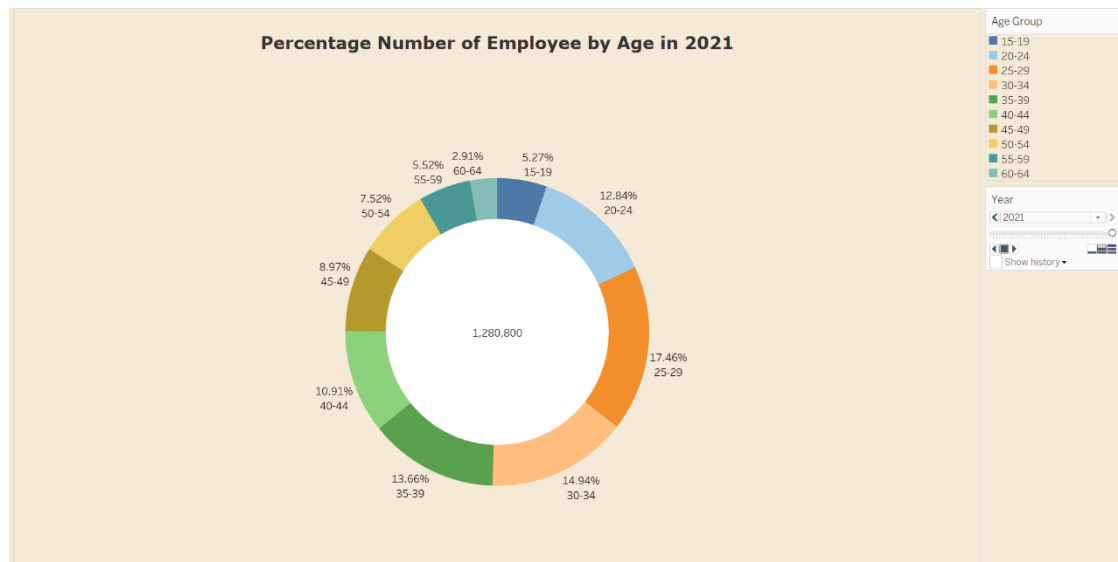


Figure 3.1.7: Visualisation 7

Analysis:

The highest percentage number of employee in the year of 2021 is 17.46% which is the age group of 25-29 years old. This is due to the students who has just graduated from university and got employed into the industry. Industry takes fresh graduates as they have the mind of the current generation, lots of energy to do work and the knowledge are still fresh that could be implemented to help grow the industry. The least percentage number of employee is 2.91% which is the age group of 60-64 years old. This group consists of the elderly where they are old and do not have enough energy and strength to do loads of work. Additionally, 2021 was the year where Covid-19 has start to recover and during the pandemic government advised to work from home where it is more convenient for the age group of early 20's, importantly the government constantly reminding so that the young children and elderly people do not go out often as they can caught the disease easily.

### 3.1.7: Number of Employee and its Cumulative by Age Group

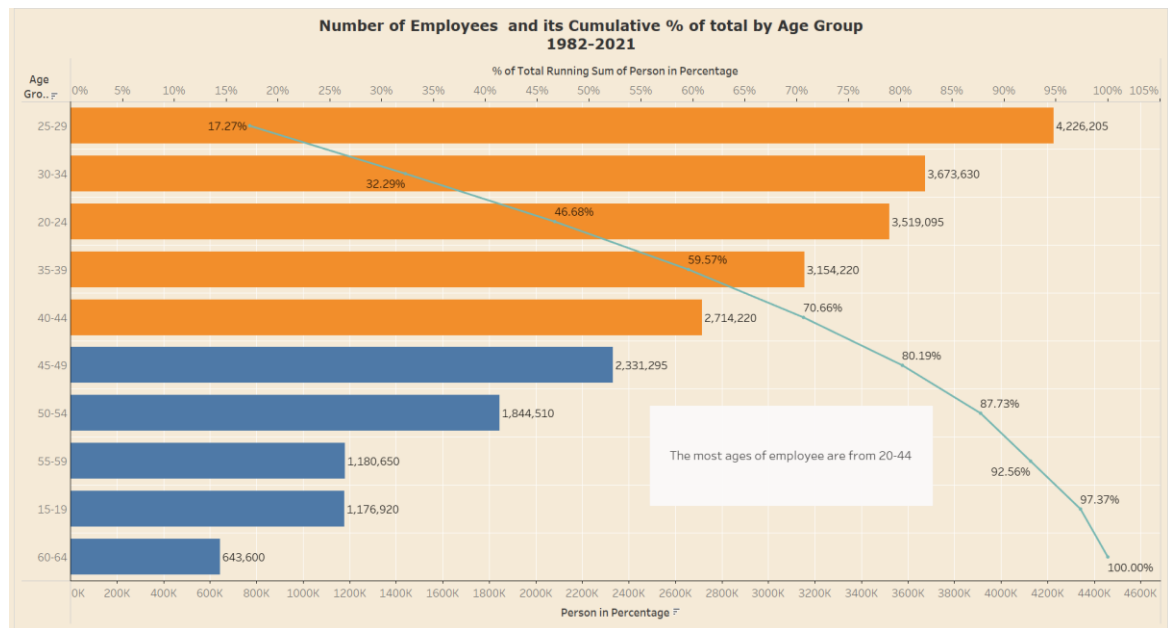


Figure 3.1.8: Visualisation 8

#### Analysis:

The figure 3.1.8 shows the top age group of employees and its cumulative percentage of total from 1982-2021. The top 5 age groups are employees from 20 to 44 years old. Nevertheless, only 4 age groups that exceed their cumulative percentage total which are the age group of 25-29, 30-34, 20-24 and 35-39. This shows that employee that reach the age of 40 starts to pension from their work as they would want to spend more time at home with their family members for the one that are married. From observing the graph, it is found that the top age group is 25-29 and the third age group is 20-24, this gap of position occurs due to the experience or skills that the person has. The more skills and experience you have the most likely you will be employed, moreover even you have just graduated basic skills should have been instil since the teenager phase of life to have a brighter future.

## 3.2 EXPLANATION OF DASHBOARD

### 3.2.1 Main Dashboard

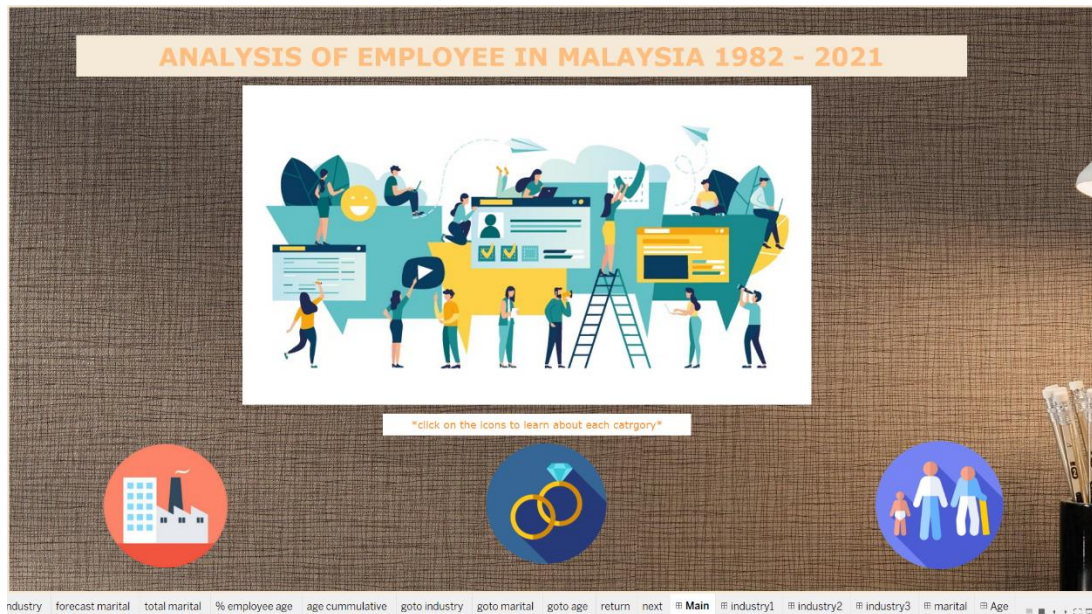


Figure 3.2.1: Dashboard 1

Analysis:

The above dashboard is the main dashboard, it contains of action buttons that will help to navigate and connect the main dashboard to other dashboards. Thus, the three icons at the bottom are used for to go to other dashboards. The first icon is a picture of industry as it will connect the main dashboard to the number of employees by industries, while the second icon that has picture of rings will connect to the dashboard with the analysis number of employee by marital status and the icon with people of different age will link to the dashboard that contains the analysis number of employee by age group.

### 3.2.2: Industry Dashboard 1 and 2

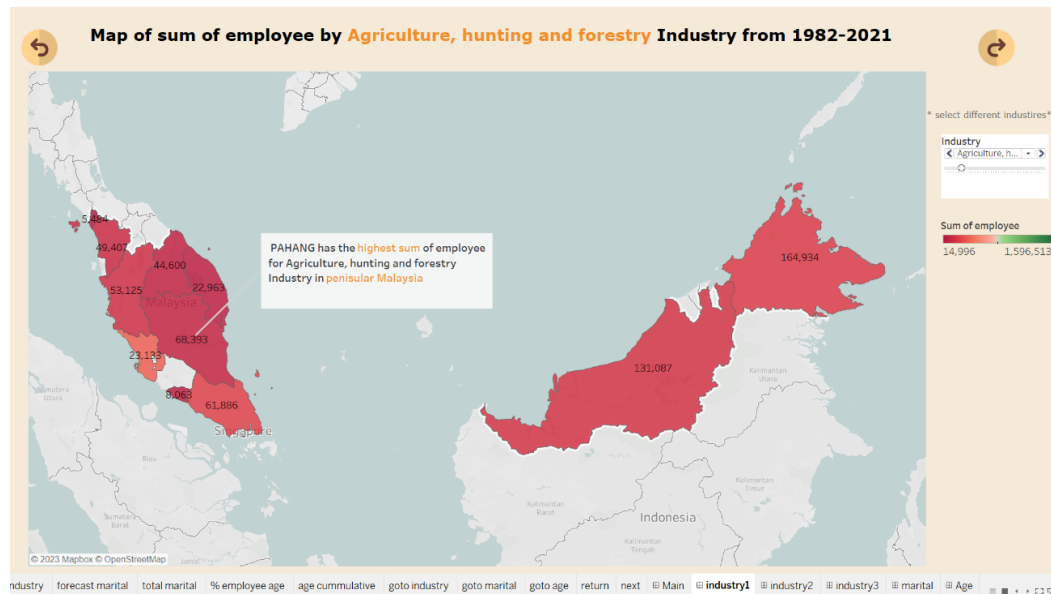


Figure 3.2.2: Dashboard 2

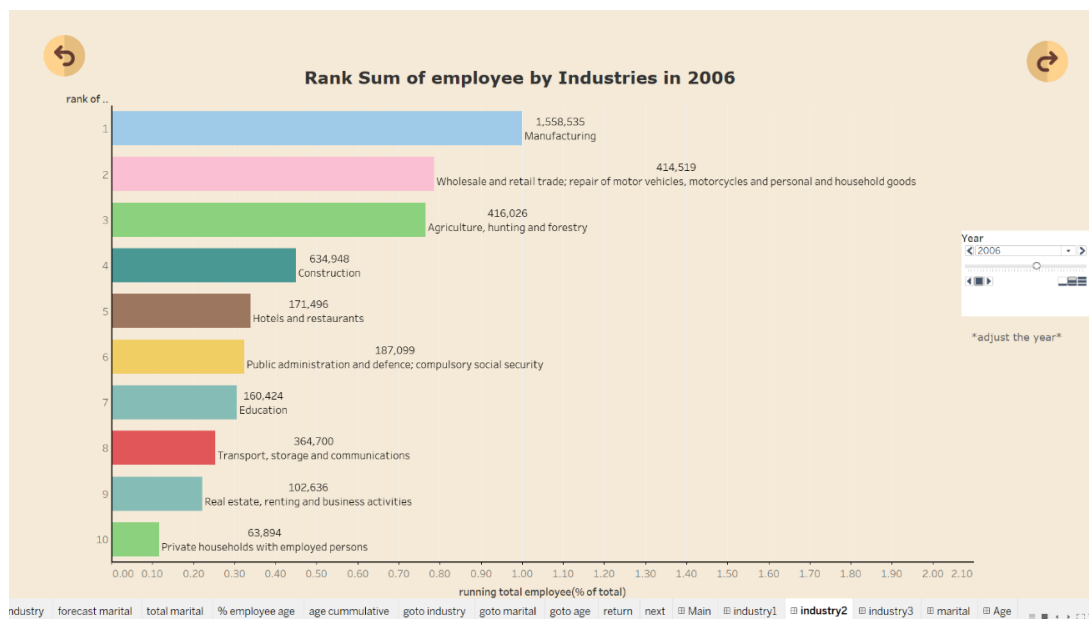


Figure 3.2.3: Dashboard 3

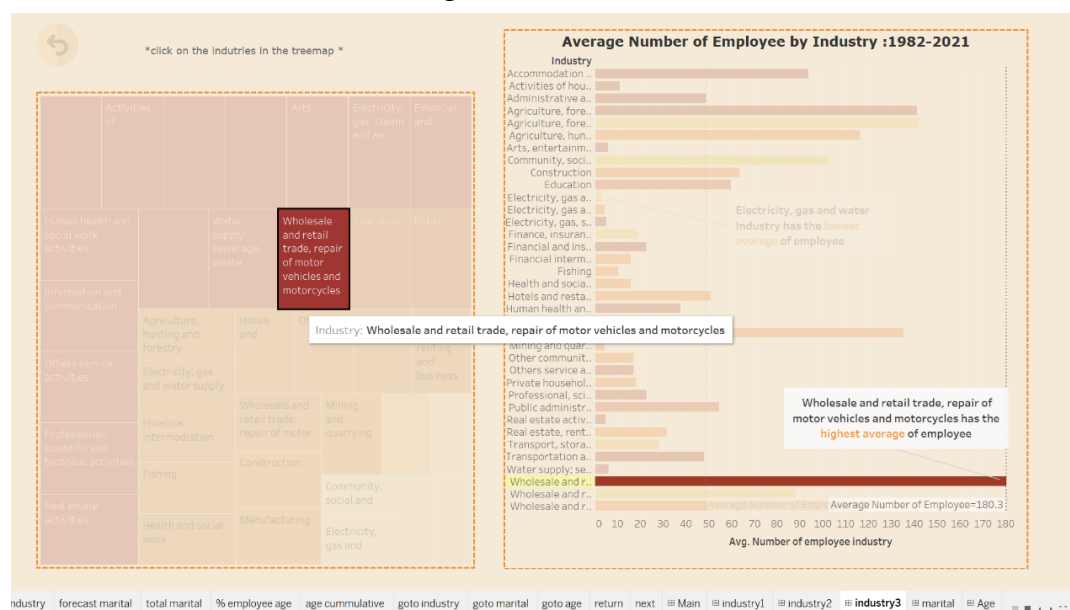
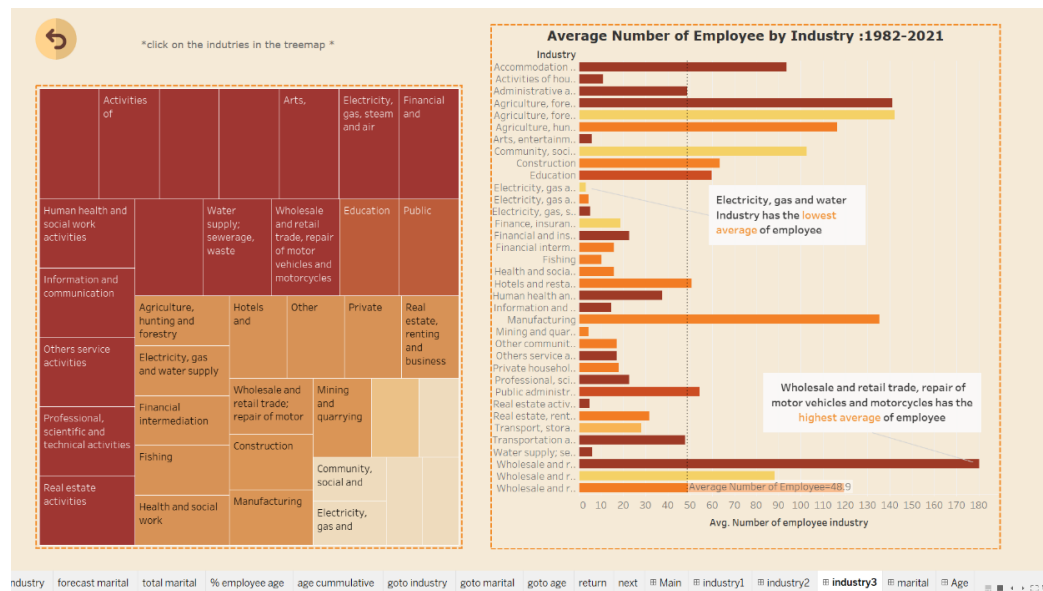
### Analysis:

Figure 3.2.2 is the first dashboard for the analysis number of employee by industries and state by the visualisation of the heat map. Therefore, the slider on the right helps in changing the industry type to see the changes of number and colour in each state. As you move the slider, some annotation will appear to describe the pattern of number for the stated industry. Meanwhile, for figure 3.2.3 it is the second dashboard for industry where it shows the rank of industry by year which could be change by moving the slider. The two visualisation allows the viewer to observe the relationship between the number of employees by state in an industry and the ranking of industry by year. By that, actions can be taken to improve the employment



at certain states for some industries. The return and next button are located on the top left and right corner, by clicking on the return button in figure 3.2.2 it will bring back to the main dashboard while in figure 3.2.3 it will bring back to dashboard 2 of industry but the next button links to the next visualisation for industry for both figures.

### 3.2.3: Industry Dashboard 3



### Analysis:

The dashboard in figure 3.2.4 is the third dashboard for analysis employee by industry. This dashboard focuses on the average number of employees in each of the industry. By clicking on one of the squares in the treemaps it will highlight the average number of employees from the bar chart beside it, as shown in figure 3.2.5. This action allows the viewer to read the

average number of employees in an interactive way. The return button on the left top corner, brings the viewer back to dashboard 2.

### 3.2.4: Marital Dashboard

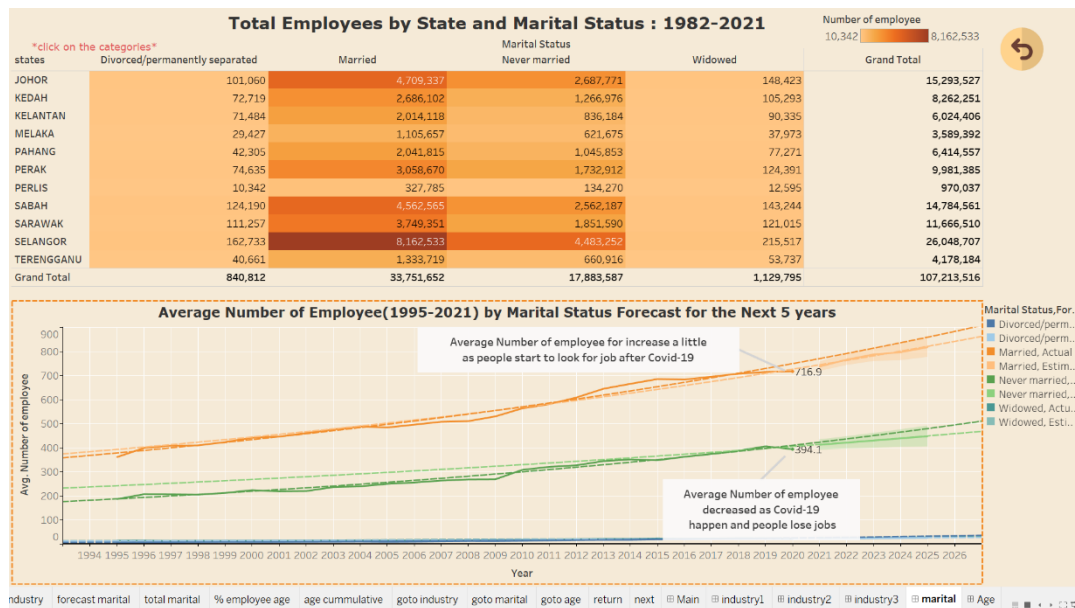


Figure 3.2.6: Dashboard 5

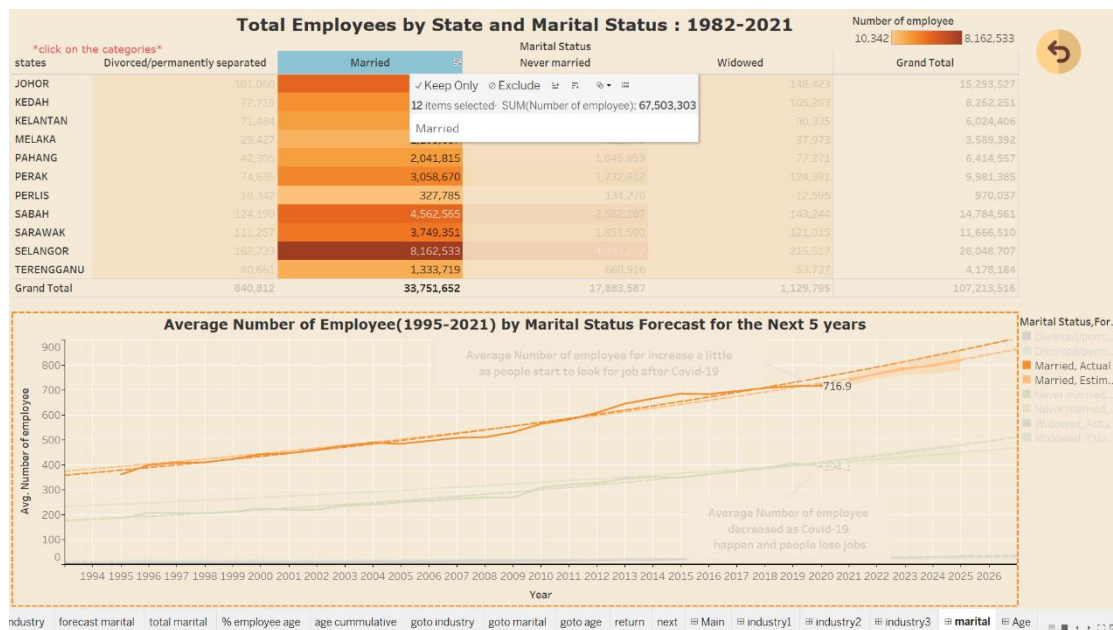


Figure 3.2.7: Highlights Action on Dashboard 5

Analysis:

Figure 3.2.6 shows the dashboard that will represent the visualisation relationship between the marital status and the number of employees. Highlight action is also applied in this dashboard to navigate through the visualisation, by clicking on one of the categories in the crosstab table with highlight it will call attention to the line chart of forecast employment based on the marital status. This can help to show the relationship of marital status in each state and the forecast for 5 years so that the viewer will know which state to influence the increase of

employee of year. The company may try to promote job vacancy at state that has least amount of employment throughout the year. The return button on the top right corner will connect back to the main dashboard.

### 3.2.5: Age Dashboard

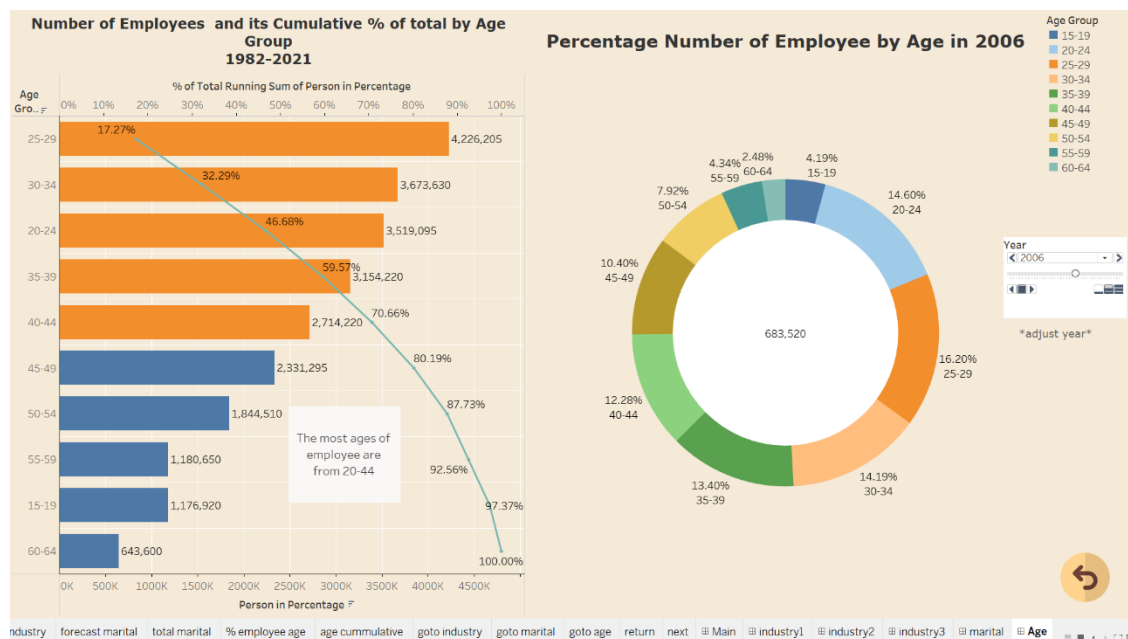


Figure 3.2.8: Dashboard 6

#### Analysis:

Figure above shows the dashboard for the relationship of age group. This dashboard consists of two chart which is a donut chart and a cumulative chart. The donut chart will help to show which year the top rank age group has the highest number of employees, so that the industry or company can do a strategy to gain back employee that are in the required rank after Covid-19 happened. The cumulative chart helps in showing which age group has exceed or did not exceed the required amount of employee throughout the past years and see the donut chart to give insights of which year contributed the most.

## 4.0 CONCLUSION

To sum up, the analysis for the number of employees in Malaysia from 1982 to 2021 it has given valuable insights into the country's labour market. Through the process of visualisation several findings are revealed. The most obvious pattern seen is that the gradual increased number of employees in Malaysia throughout the years due to several factors such as increase in population growth. It underlines the distribution of employees across different industries in each state. The impact of Covid-19 towards employee shows meaningful insights as some visualisations decrease and increase due to the pandemic that Malaysia has been through. The visualisation of age group allows me to know which time of our age industry like to hire also taking consideration of several factors such as health. Knowing that, the highest marital status of employment is married people highlights the point why it is hard to maintain a work life balance as some has pressure from work.

The visualisation offers meaningful insights that I will consider as my general knowledge that would be helpful in the future. The objective of this project has achieved to see the relationship between number of employees and the three factors. In general, the number of employees will increase yearly but in which industry depends on the situation of the country's economy. Even if there may be unforeseen circumstances, our basis for success is built on our understanding of employment trends and proactive thinking.