INDIVIDUAL PROJECT
BSD1323 STORYTELLING AND DATA VISUALIZATION

PROJECT TITLE:
EMPLOYMENT IN MALAYSIA FROM YEAR 2001 TO 2020 BY ETHNIC GROUP, STATE, CITIZEN CATEGORY AND GENDER.

NO.	NAME	NO MATRIK
1.	NUR AINUL NASUHA BINTI MOHD AZREEN	SD22004

DUE DATE:
24 MAY – 13 JUNE 2023

LECTURER'S NAME:
DR. SITI ZANARIAH SATARI

1.0 TITLE AND MOTIVATION OF MY PROJECT TOPIC

Title: Employment in Malaysia from Year 2001 to 2020 by Ethnic Group, State, Citizen Category and Gender.

As a data analytics student, I have always loved data and numbers. So, my project revolves around experimenting with a data set with large numbers and categorical data. I am excited to explore the employment patterns in Malaysia over the past two decades, specifically focusing on ethnic diversity, state variations, citizenship status, and gender.

The purpose of my project is to analyse and provide an insight into the employment trends in Malaysia from 2001 to 2020 which highlights any significant shifts, challenges, or improvements in employment opportunities for different ethnic groups, states, citizen categories, and genders in Malaysia.

Through my project, I hope to find a clearer vision of the employment trends in Malaysia over the past two decades, specifically addressing opportunities for all ethnicity, state, citizenship, or gender.

2.0 DETAILS EXPLANATION OF MY DATASET

There are two data sets that I chose as data for my project. The data set that I chose is first, employed persons by ethnic group and state. Next, I chose a data set called employed persons by gender and state. Both of these data are from a website: www.data.gov.my. I chose to compile these data into an Excel file. I also cleaned and filtered the data before I resume with the dashboards.

	Ye	State	Citizen Category	Ethnic Group	Value in Thousar
2	2020	MALAYSIA	Total of employed	non-Malaysian	14956.7
3	2020	MALAYSIA	Malaysian citizens	Total Malaysian citizens	12865.9
4	2020	MALAYSIA	Malaysian citizens	Bumiputera	8540.2
5	2020	MALAYSIA	Malaysian citizens	Chinese	3281.6
6	2020	MALAYSIA	Malaysian citizens	Indians	948.7
7	2020	MALAYSIA	Malaysian citizens	Others	95.3
8	2020	MALAYSIA	Non Malaysian citizens	non-Malaysian	2090.9
9	2020	JOHOR	Total of employed	non-Malaysian	1775.6
10	2020	JOHOR	Malaysian citizens	Total Malaysian citizens	1553.6
11	2020	JOHOR	Malaysian citizens	Bumiputera	910.9
12	2020	JOHOR	Malaysian citizens	Chinese	527.8
13	2020	JOHOR	Malaysian citizens	Indians	108.6
14	2020	JOHOR	Malaysian citizens	Others	6.3
15	2020	JOHOR	Non Malaysian citizens	non-Malaysian	222
16	2020	KEDAH	Total of employed	non-Malaysian	914.7
17	2020	KEDAH	Malaysian citizens	Total Malaysian citizens	851.8
18	2020	KEDAH	Malaysian citizens	Bumiputera	668
19	2020	KEDAH	Malaysian citizens	Chinese	115
20	2020	KEDAH	Malaysian citizens	Indians	64.2
21	2020	KEDAH	Malaysian citizens	Others	4.6
22	2020	KEDAH	Non Malaysian citizens	non-Malaysian	62.9
23	2020	KELANTAN	Total of employed	non-Malaysian	680.8
24	2020	KELANTAN	Malaysian citizens	Total Malaysian citizens	649.3
25	2020	KELANTAN	Malaysian citizens	Bumiputera	620.5
26	2020	KELANTAN	Malaysian citizens	Chinese	23.3
27	2020	KELANTAN	Malaysian citizens	Indians	1.3
28	2020	KELANTAN	Malaysian citizens	Others	4.3
29	2020	KELANTAN	Non Malaysian citizens	non-Malaysian	31.4
30	2020	MELAKA	Total of employed	non-Malaysian	417.3
31	2020	MELAKA	Malaysian citizens	Total Malaysian citizens	384.9
32	2020	MELAKA	Malaysian citizens	Bumiputera	259.8
33	2020	MELAKA	Malaysian citizens	Chinese	100.4
34	2020	MELAKA	Malaysian citizens	Indians	23
35	2020	MELAKA	Malaysian citizens	Others	1.6
36	2020	MELAKA	Non Malaysian citizens	non-Malaysian	32.5
37	2020	N.SEMBILAN	Total of employed	non-Malaysian	474.1
38	2020	N.SEMBILAN	Malaysian citizens	Total Malaysian citizens	416.4
39	2020	N.SEMBILAN	Malaysian citizens	Bumiputera	252.5

The first data set consists of 2416 records. It has one date data type which is Year. Next, it has two categorical data types which are Citizen Category and Ethnic Group. Other than that, it has one geographic data type which is State. Lastly, it has a quantitative data type which is Value in Thousands.

This data has been cleaned in which I filtered the data according to years from 2001 to 2020. This data also sorted according to years descending. This data is in form of Microsoft Excel Worksheet (.xlsx).

This data set consists of four dimensions which are Citizen Category, Ethnic Group, State and Year as it also has one measure which is Value in Thousands.

	Yea	State	Sex	Employed Person in Thousands
2	2020	Malaysia	Male	9128.8
3	2020	Johor	Male	1173.8
4	2020	Kedah	Male	583.2
5	2020	Kelantan	Male	433.9
6	2020	Melaka	Male	245.3
7	2020	Negeri Sembilan	Male	284.5
8	2020	Pahang	Male	471.9
9	2020	Pulau Pinang	Male	490.2
10	2020	Perak	Male	639.8
11	2020	Perlis	Male	64.4
12	2020	Selangor	Male	1912.1
13	2020	Terengganu	Male	301.4
14	2020	Sabah	Male	1166.8
15	2020	Sarawak	Male	810.1
16	2020	W.P. Kuala Lump	Male	508.2
17	2020	W.P Labuan	Male	26.7
18	2020	W.P.Putrajaya	Male	16.5
19	2020	Malaysia	Femal	5827.9
20	2020	Johor	Femal	601.9
21	2020	Kedah	Femal	331.5
22	2020	Kelantan	Femal	246.9
23	2020	Melaka	Femal	172.1
24	2020	Negeri Sembilan	Femal	189.6
25	2020	Pahang	Femal	244
26	2020	Pulau Pinang	Femal	342.7
27	2020	Perak	Femal	403
28	2020	Perlis	Femal	41.2
29	2020	Selangor	Femal	1534.4
30	2020	Terengganu	Femal	163.6
31	2020	Sabah	Femal	702.8
32	2020	Sarawak	Femal	452.2
33	2020	W.P. Kuala Lump	Femal	366.3
34	2020	W.P Labuan	Femal	16.1
35	2020	W.P.Putrajaya	Femal	19.6
36	2019	Malaysia	Male	9202.4
37	2019	Johor	Male	1091
38	2019	Kedah	Male	576
39	2019	Kelantan	Male	431.5
40	2019	Melaka	Male	251.7
41	2019	Negeri Sembilan	Male	301

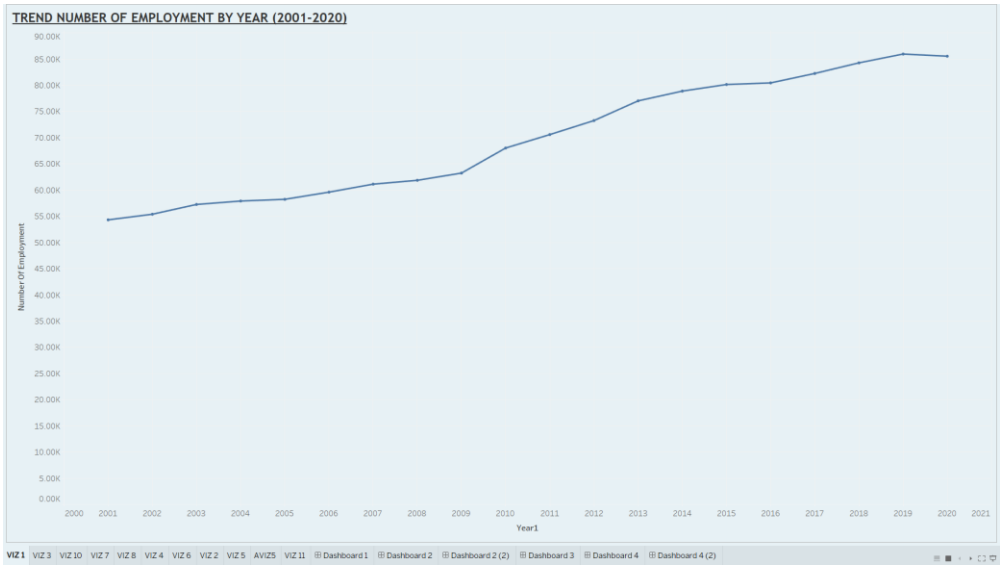
The second data set consists of 681 records. It has one date data type which is Year. Next, it has two categorical data types which is gender. Other than that, it has one geographical data type which is State. Lastly, it has one quantitative data type which is Employed Person in Thousands.

This data has been cleaned in which I filtered the data according to years from 2001 to 2020. This data also sorted according to years descending. This data is in form of Microsoft Excel Worksheet (.xlsx).

This data set consists of 3 dimensions which are Gender, State and Year as it also has one measure which is Employed Person in Thousands.

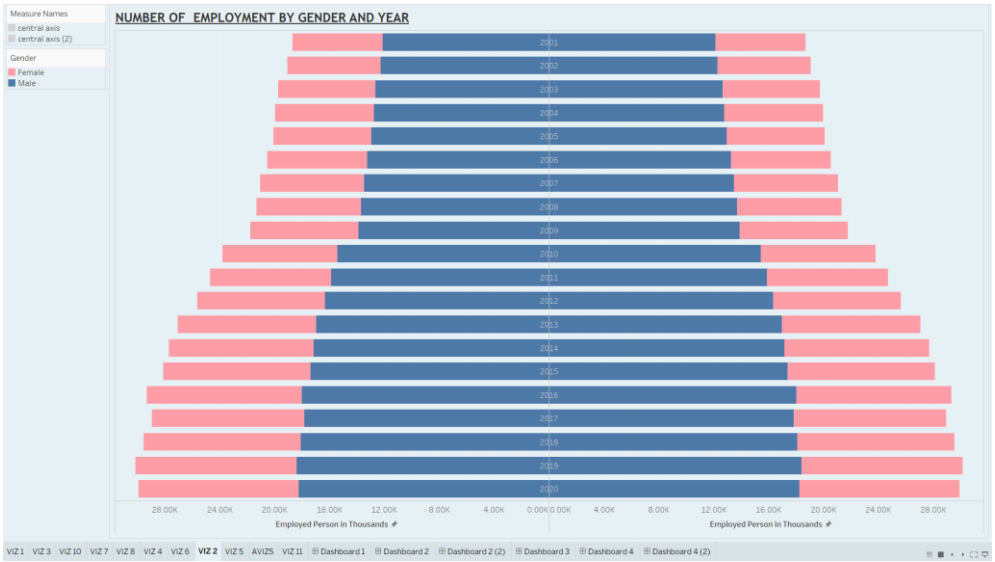
3.0 DETAILS ANALYSIS OF EACH VISUALIZATION AND DASHBOARD

VIZ 1(continuous line chart): Trends Number of Employment by Year (2001-2020)



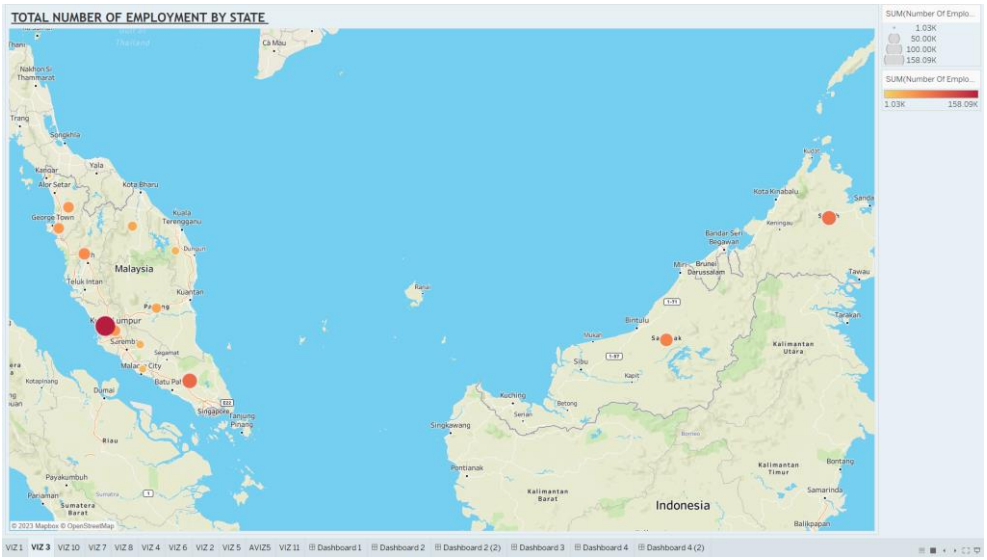
- This visualisation shows the trends for number of employments against year from year 2001 to 2020.
- There is an increasing trend of number of employments from 2001 to 2020.
- The number of employments for 20 years is totalled up by years.
- Year 2020 recorded highest total number of employments, while year 2001 shows the lowest.
- The increasing trends shows by the number of employments against year is the results of the economic positive growth. It tends to create more job opportunities such as industries expand, new businesses emerge, and existing businesses often hire more workers to meet increases demand.

VIZ 2(butterfly chart): Number of Employments by Gender and Year



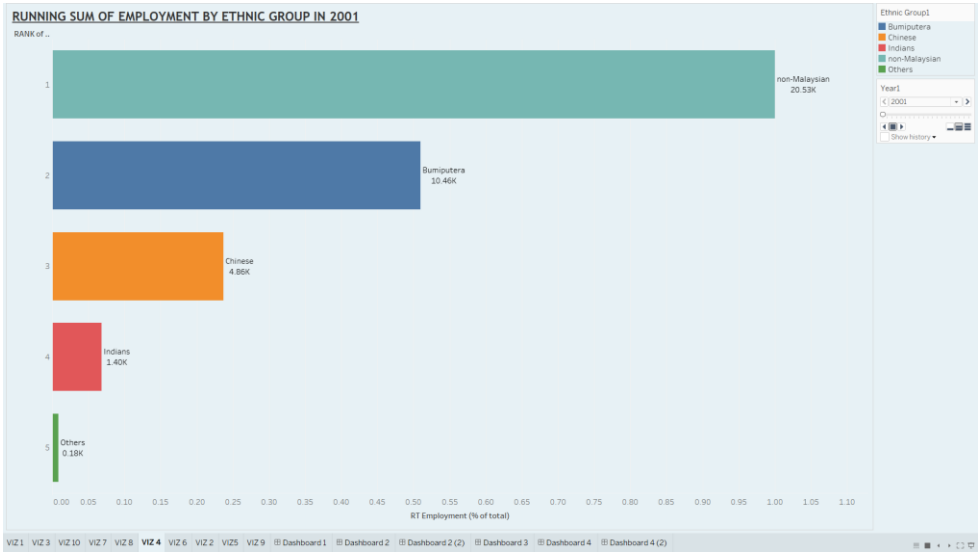
- This visualisation shows comparison of number of employments between gender and year.
- This chart was to shows how the number of employments trends over the year.
- Both male and female shows the same trends over the year which is increasing number of employments.
- But, total number of employments for male groups id higher compared to female group,
- For example, in year 2005 total number of employments for male is 12.94K while female is 7.15K.
- This is because of the gender stereotypes and bias that happened during the early years which limit females from pursuing certain fields or advancing in their careers, resulting in a lower overall number of female employees.

VIZ 3(symbol maps): Total Number of Employments by State



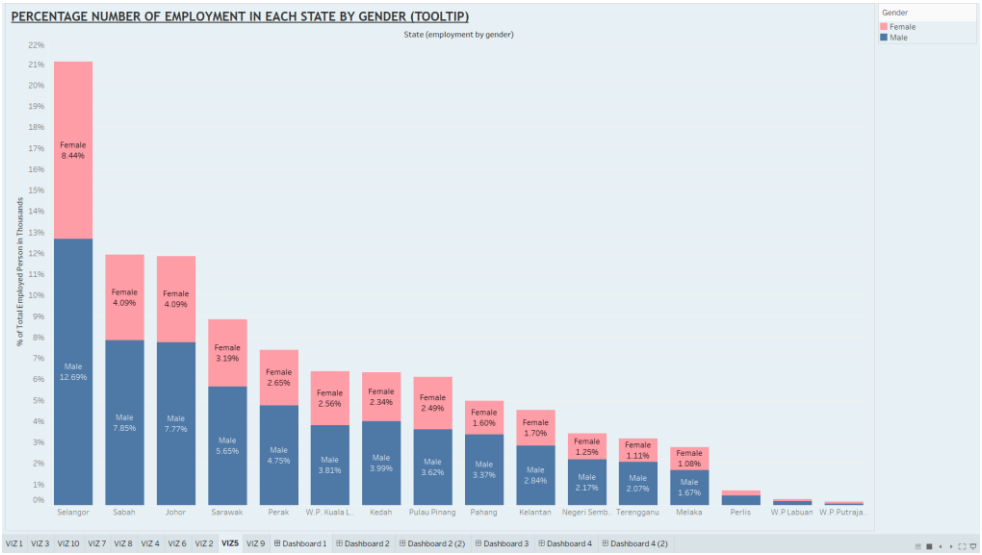
- This visualisation shows the sum number of employments for each state.
- The highest total number of employments is recorded in Selangor which is 158.09K.
- This is because Selangor has a significant urban population, with cities like Petaling Jaya, Shah Alam and Subang Jaya experiencing rapid growth. Urban areas often have higher employment opportunities due to the presence of businesses, industries and commercial centers.
- The lowest total number of employments is recorded in Perlis which is only 5.47K.
- This is because Perlis has a more rural and less urbanized landscape compared to other states. Urban areas tend to have a higher concentration of businesses, industries, and employment opportunities. The limited urbanization in Perlis may result in fewer employment options.

VIZ 4(Running bar chart): Running Sum of Employments by Ethnic Group <Page Name>



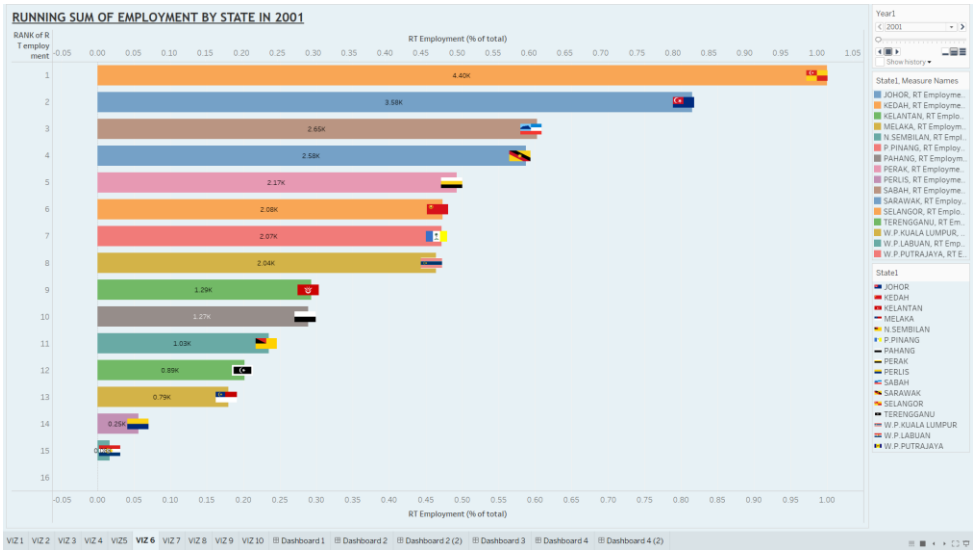
- This visualisation shows running bar chart of sum of employments by ethnic group for each year.
- The longest bar length represents an ethnic with the highest sum of employments for given year.
- For example, in year 2001 the longest bar represents non-Malaysian who have the highest sum of employments which is 20.5K.
- This is because Malaysia has historically relied on foreign employees to support various industries, such as construction, manufacturing, agriculture, and services. In the early 2000s, there may have been policies or practices in place that favoured the employment of non-Malaysians to meet employees demands in these sectors.

VIZ 5(Stacked bar chart): Percentage Number of Employments in Each State by Gender



- This visualisation shows the percentage distribution for total number of employments in each state by gender.
- The longest bar length represents state with the highest percentage sum of employments of all states.
- For example, Selangor has the longest bar length which represents the highest percentage sum number of employments for both genders which are 8.44% for female and 12.69% for male.
- This is because Selangor is the most developed state in Malaysia. The area serves as a major economic hub with a wide range of industries, including manufacturing, services, finance, technology, and tourism. The abundance of job opportunities and a robust business ecosystem in Selangor attracts a large number of employees, both male and female.

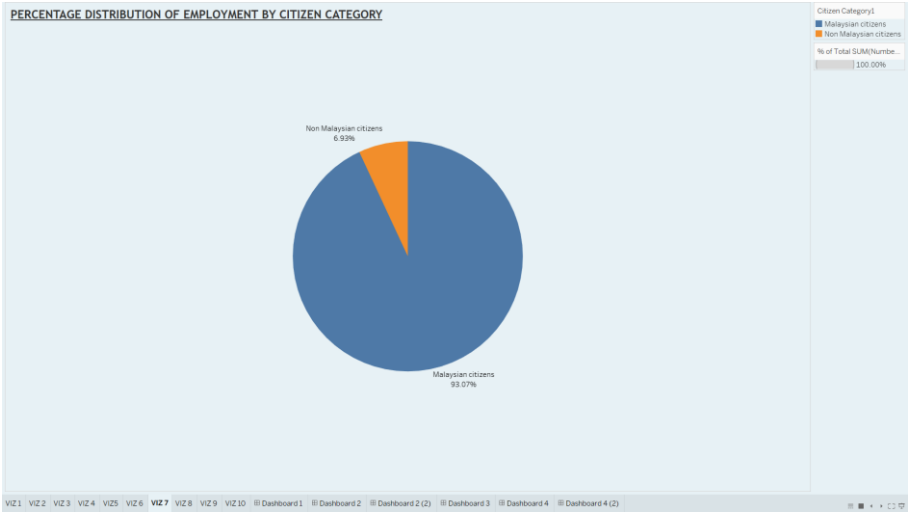
VIZ 6(Running bar chart): Running Sum of Employments by State <Page Name>



- This visualisation shows running bar chart of sum of employments by State for each year.
- The longest bar length represents a state with the highest sum of employments for given year.
- For example, in year 2001 the longest bar represents Selangor who have the highest sum of employments which is 4.40K.
- This is because Selangor boasts a diverse range of industries, including manufacturing, services, technology, logistics, and retail. This diversity can contribute to a higher number of employments as it accommodates various job sectors and skill sets.
- Another example, in year 2001 the shortest bar represents Wilayah Persekutuan Labuan who have the lowest sum of employments which is 0.08K.
- This is because compared to other states in Malaysia, Labuan has limited range of economic activities. The lack of diversification in industries and

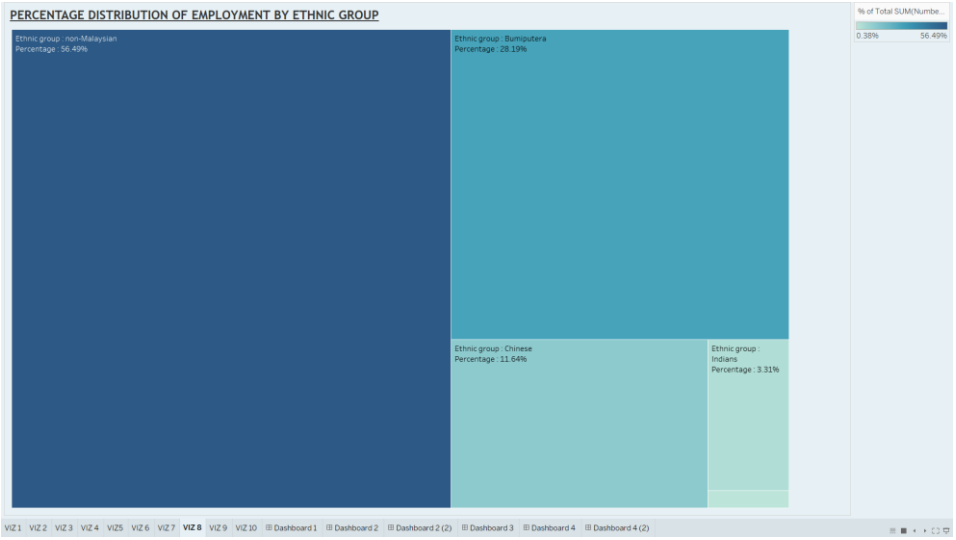
sectors may restrict employment opportunities, as there are fewer options available for job seekers.

VIZ 7(Pie chart): Percentage Distribution of Employment by Citizen Category



- This visualisation shows percentage distribution of employment by citizen category.
- This pie chart shows that 93.07% of percentage distribution of employment dominated by Malaysian Citizens.
- This is because The Malaysian government has regulations in place that prioritize the employment of Malaysian citizens. This is to reduce economic disparities and promote the participation of Malaysians in various sectors, including employment.
- Non-Malaysian citizens hold for about another 6.93%.

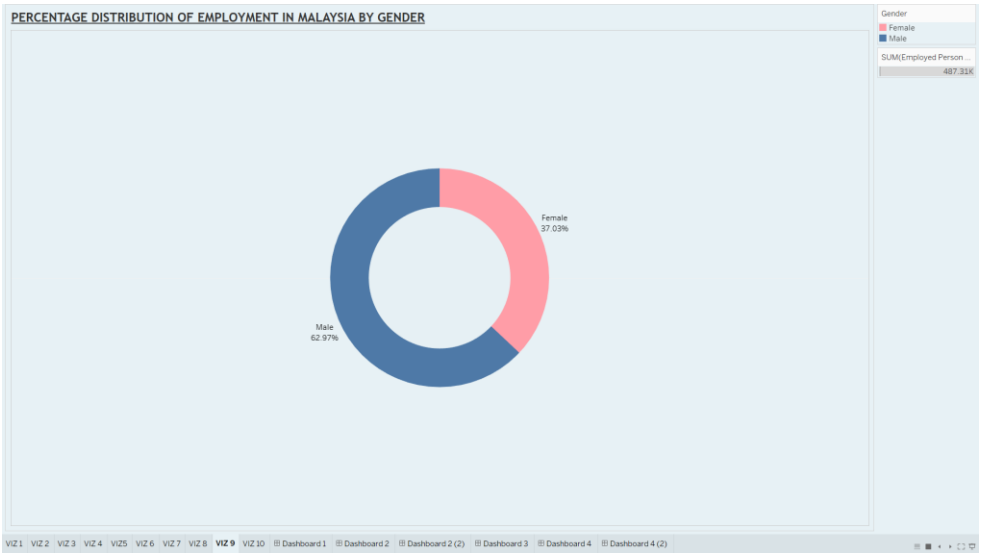
VIZ 8(tree maps): Percentage Distribution of Employment by Ethnic Group



- This visualisation shows the percentage distribution of employment by ethnic group.

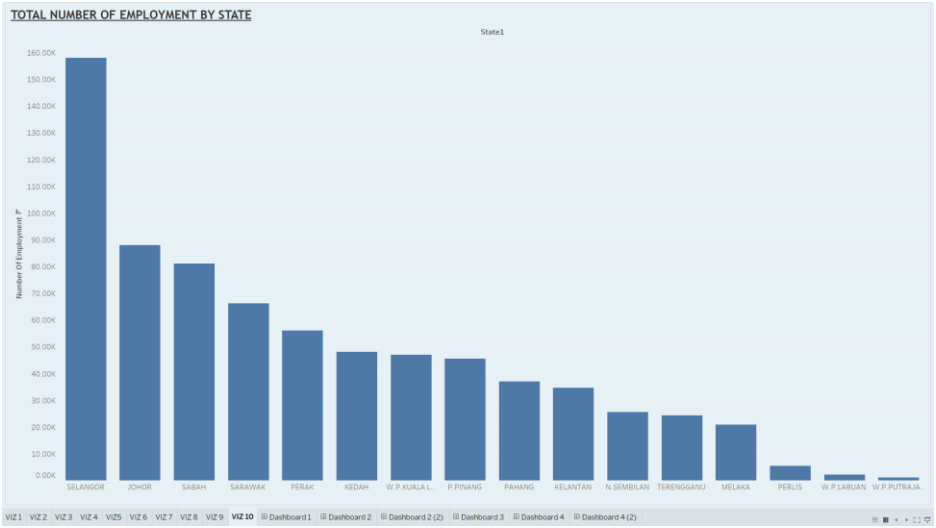
- The biggest area represents the highest percentage distribution of employment by ethnic group.
- The biggest area in this tree maps is non-Malaysian followed by Bumiputra, Chinese, Indian and others which is 56.49%.
- This is because year 2001 to 2020 may have marked a period of economic growth or specific industry expansions in Malaysia, leading to an increased need for employees. Non-Malaysians, particularly migrant workers from neighbour countries, may have been recruited to fill these job vacancies due to factors such as cost-effectiveness or availability.

VIZ 9(Donut chart): Percentage Distribution of Employments in Malaysia by Gender



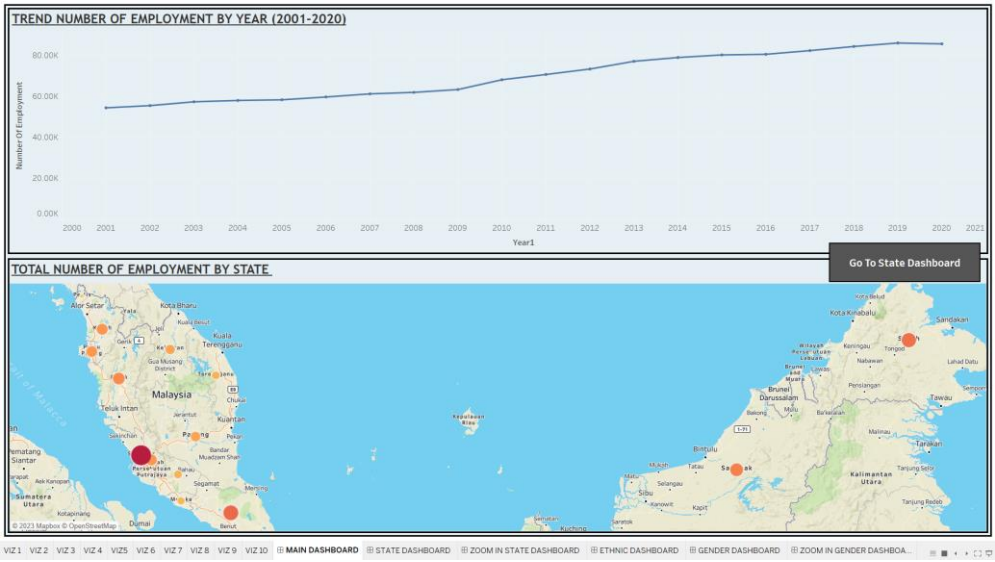
- This visualisation shows the percentage distribution of employments in Malaysia by gender.
- The donut chart shows that male hold higher percentage which is 62.97% while female is 37.03%.
- This is because the responsibilities and expectations related to family and caregiving can impact women's workforce participation rates. Balancing work and family responsibilities can be more challenging for women due to factors such as maternity leave, child-rearing, and societal expectations, which may result in lower overall employment numbers.

VIZ 10(Bar chart): Total Number of Employments by State

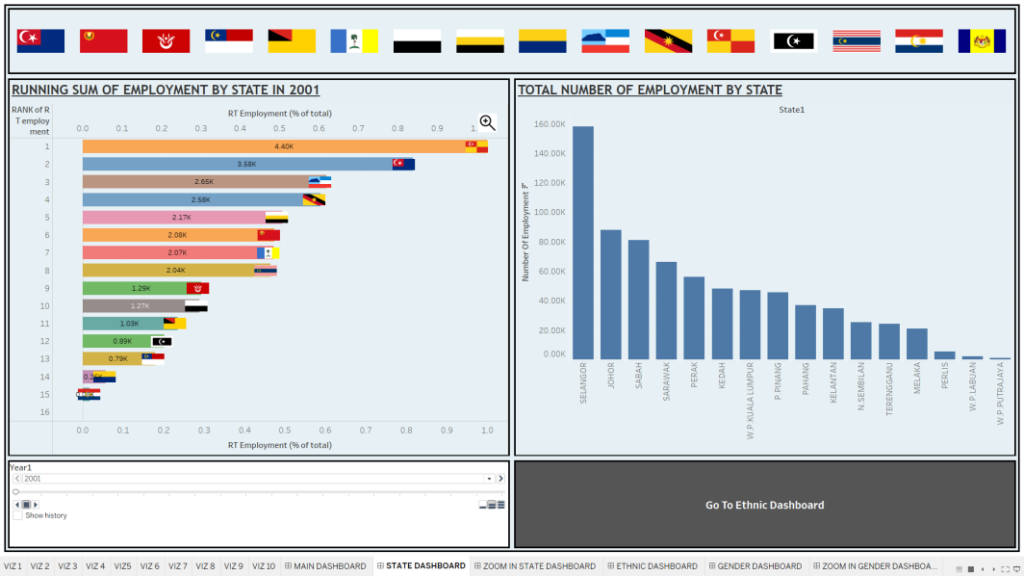


- This visualisation shows total number of employments by state.
- The bar chart indicates that Selangor hold the highest total number of employments compared to other state which is 158.09K.
- This is because Selangor benefits from a well-developed transportation infrastructure, including major highways, airports, and a comprehensive public transportation system. It also has extensive connectivity to other states and countries, which can attract businesses and create more job opportunities.
- However, Wilayah Persekutuan Putrajaya has the lowest total number of employments compared to other state.
- This is because Putrajaya has a relatively small residential population compared to other states in Malaysia. As an administrative capital, its primary purpose is to serve the government and public administration needs, rather than being a densely populated residential area. The smaller population base results in fewer employment opportunities within the city.

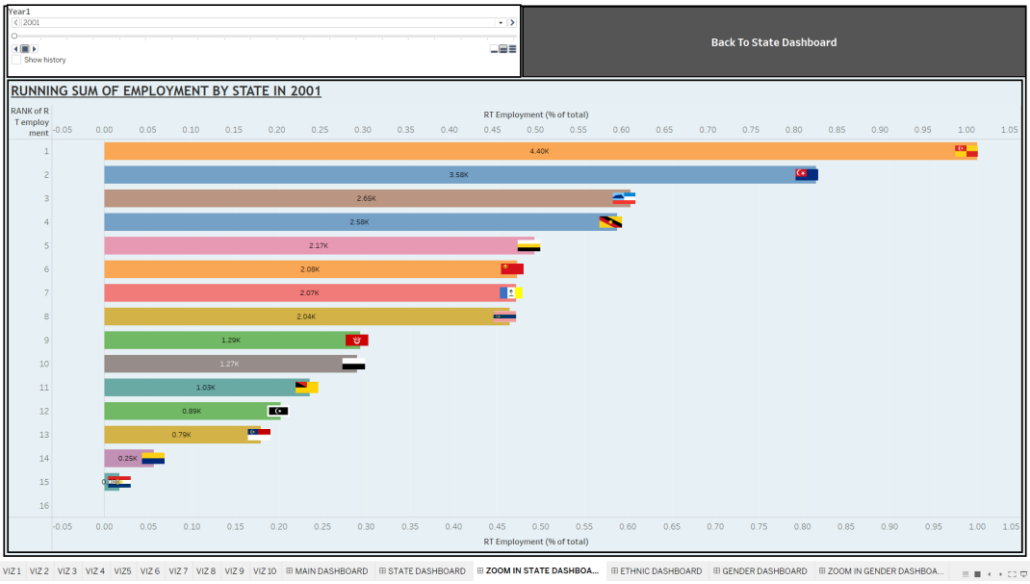
DASHBOARD 1: Main Dashboard



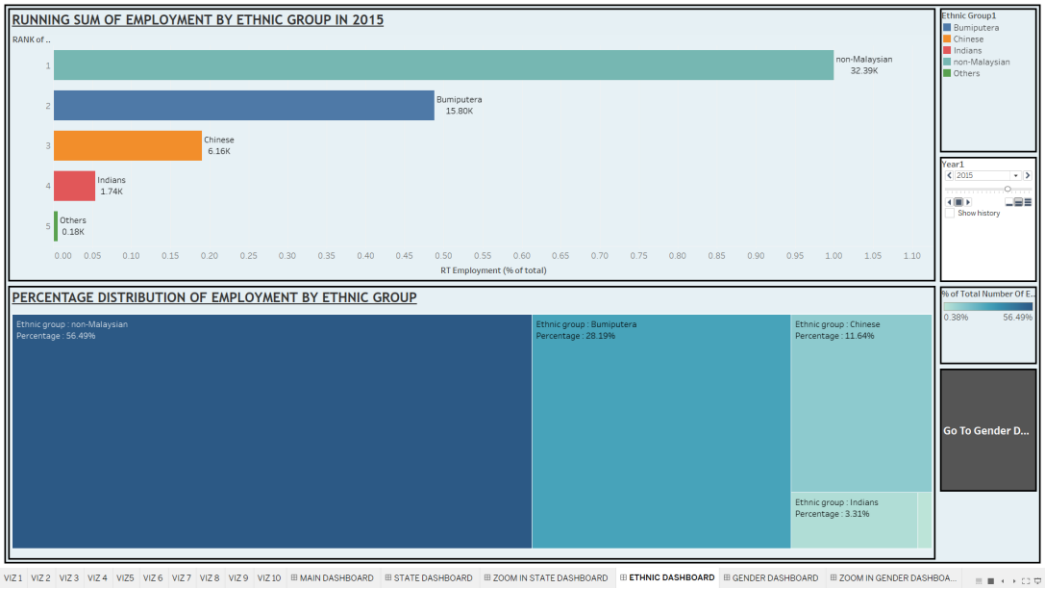
DASHBOARD 2: State Dashboard



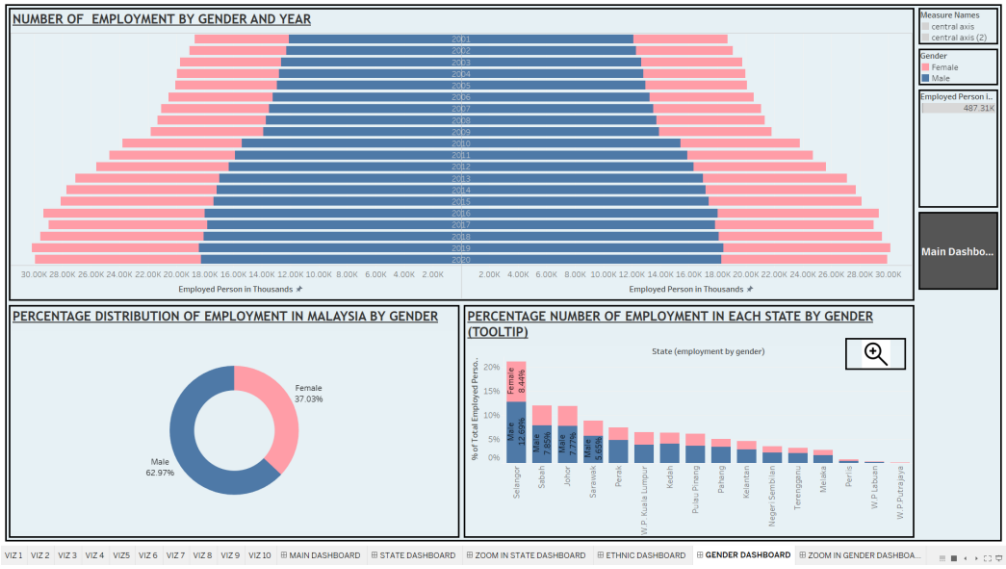
DASHBOARD 3: Zoom in State Dashboard



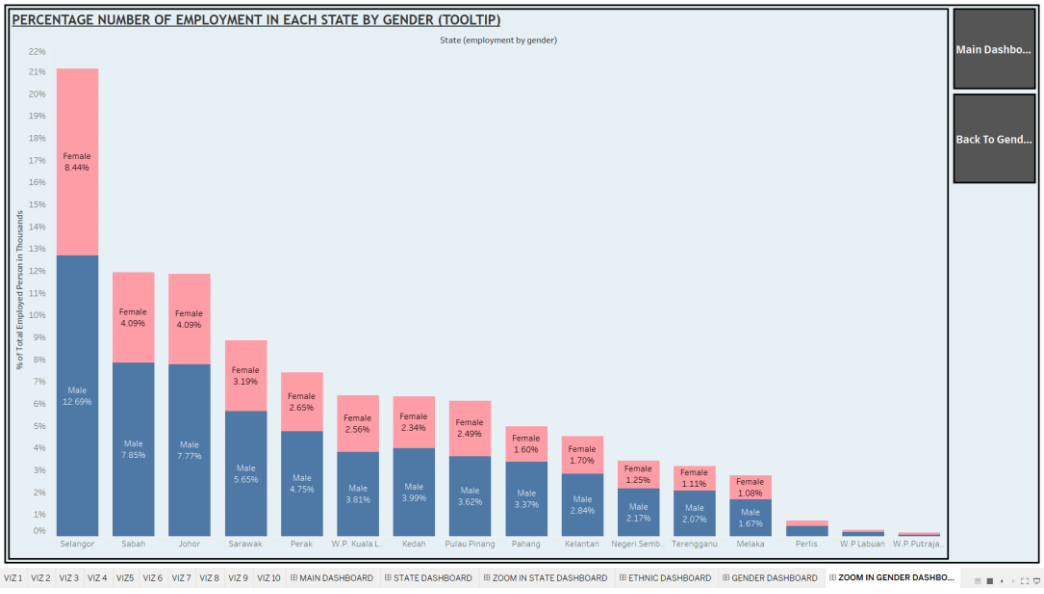
DASHBOARD 4: Ethnic Dashboard



DASHBOARD 5: Gender Dashboard



DASHBOARD 6: Zoom in Gender Dashboard



DASHBOARD	INTERPRETATION
1)Main Dashboard	• In this dashboard there are two visualisation which are Trend Number of Employment by Year (2001-2020) and Total Number of Employment by State. • There is also one action menu called Go To State Dashboard which can lead us to the second dashboard Called State Dashboard.
2) State Dashboard	• In this dashboard there are three visualisation which is the country flags, Running Sum of Employment by State in <Page Name> and Total Number of Employment by State. • There is also 2 action menu which are Zoom In which can lead us to zoom in the visualisation of Running Sum of Employment by State in <Page Name> to see clearer sight of the visualisation. Other than that, Go To Ethnic Dashboard which lead to the fourth dashboard called Ethnic Dashboard.
3) Zoom In State Dashboard	• In this dashboard we can see clearer visualisation of Running Sum of Employment by State in <Page Name>.
4)Ethnic Dashboard	• In this dashboard there are two visualisation which are Running sum of Employment by Ethnic Group in <Page Name> and Percentage Distribution of Employment by Ethnic Group. • There is also one action menu called Go To Gender Dashboard which can lead us to the fifth dashboard Called Gender Dashboard.
5) Gender Dashboard	• In this dashboard there are three visualisation which is the Number Of Employment by Gender and Year, Percentage Distribution of Employment in Malaysia by Gender and Percentage Number of Employment in Each State by Gender. • There is also 2 action menu which are Zoom In which can lead us to zoom in the visualisation of Percentage Number of Employment in Each State by Gender to see clearer sight of the visualisation. Other than that, Main Dashboard which lead back to the first dashboard called Main Dashboard.
6) Zoom In Gender Dashboard	• In this dashboard we can see clearer visualisation of Percentage Number Of Employment in Each State by Gender.

4.0 CONCLUDING REMARK

My project is to analyse and provide an insight into the employment trends in Malaysia from 2001 to 2020, focusing on ethnic diversity, state variations, citizenship status, and gender.

From these two data sets, I come out with 10 sheets of visualisation which are Trends Number of Employment by Year (2001-2020), Number of Employments by Gender and Year, Total Number of Employments by State, Running Sum of Employments by Ethnic Group <Page Name>, Percentage Number of Employments in Each State by Gender, Running Sum of Employments by State <Page Name>, Percentage Distribution of Employment by Citizen Category, Percentage Distribution of Employment by Ethnic Group, Percentage Distribution of Employments in Malaysia by Gender and Total Number of Employments by State.

Other than that, I also come out with 6 dashboards which are Main Dashboard, State Dashboard, Zoom in State Dashboard, Ethnic Dashboard, Gender Dashboard and Zoom in Gender Dashboard.

From all those visualisations, the results show that the sum of employments is increasing each year because Malaysia's population continues to grow, leading to an increase in the labour force. A larger population creates more demand for goods and services, which in turn drives job creation and employment opportunities across various sectors for both male and female. With this increasement, the number of employees which are non-Malaysian also expanding because certain industries in Malaysia, such as construction or plantation sectors, have traditionally relied heavily on non-Malaysian labour due to preferences or perceived skills associated with foreign workers especially in Selangor.

Hence, all the societies should strengthen the education system to equip individuals with relevant skills and knowledge that align with the evolving demands of industries. Emphasize technical and vocational training to bridge the gap between education and industry needs.

Last but not least, I would like to thanked Dr. Zanariah for her guidance and support throughout my individual project. I am also grateful to all individuals who have generously shared their time, insights, and experiences during the course of this project. Their contributions have been instrumental in enriching the research process and enhancing the validity of the findings.

Once again, I extend my deepest appreciation to everyone involved in this project. Your contributions, support, and commitment have been truly invaluable, and I am sincerely grateful for the opportunity to work with such exceptional individuals.

	SUBJECT: BSD1323 STORYTELLING AND DATA VISUALIZATION	**MARKS**: 60(15%)
	TOPIC: CHAPTER 1 to CHAPTER 8	
	INDIVIDUAL PROJECT **DUE DATE:** 27 May - 13 June 2023	
	ID: SD22004 **NAME: NUR AINUL NASUHA BINTI MOHD AZREEN**	**SECTION: 01G**

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**		
Motivation of project topic	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	
Details explanation of the dataset	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	
Details analysis of each dashboard	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	
Concluding remarks	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	

	TOTAL (20)	

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		
Theory/ Knowledge on data visualization	No theoretical knowledge on data visualizatio n observed.	Very little knowledge observed on data visualization or some information is incorrect.	Some knowledge or information on data visualizatio n 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	
Interactive Data Visualizatio n Techniques	Failed to demonstrat e the given task.	Inappropriate interactive data visualization techniques are demonstrated.	Partly correct interactive data visualizatio n techniques are demonstrat ed, with partly valid data.	Correct interactive data visualization techniques are demonstrated, with partly valid data.	Good interactive data visualization techniques are demonstrate d, with valid but not completely accurate data.	Competent interactive data visualization techniques are demonstrated, with valid and accurate data.	1	
Theory/ Knowledge on advanced dashboard	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	

Advanced Dashboard Techniques & Data Validation	Failed to demonstrat e the given task.	Inappropriate advanced dashboard techniques are demonstrated.	Partly correct advanced dashboard techniques are demonstrat ed, with partly valid data.	Correct advanced dashboard techniques are demonstrated, with partly valid data.	Good advanced dashboard techniques are demonstrate d, with valid but not completely accurate data.	Competent advanced dashboard techniques are demonstrated, with valid and accurate data.	2	
Efficiency/ Assembly/ Tidiness	Failed to demonstrat e the given task.	Not efficiently, effectively and neatly demonstrated the given task.	Partly efficient, but not effectively and neatly demonstrat ed the given task.	Efficiently, but not effectively and neatly demonstrated the given task.	Efficiently and effectively but not neatly demonstrate d the given task.	Efficiently, effectively and neatly demonstrated the given task.	1	
Final Results (the advanced dashboard)	Failed to demonstrat e 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	
						TOTAL (4)		