Monash University: Assessment Cover Sheet

Student name	Khan	Mohib	
School/Campus	Malaysia	**Student's I.D. number**	33370311
Unit name	FIT3179 Data visualisation - S2 2023 MUM		
Lecturer's name	Ting Chai Wen	**Tutor's name**	Ting Chai Wen
Assignment name	Data Visualisation II Report	**Group Assignment: No** **Note, each student must attach a coversheet**	
Lab/Tute Class: Tutorial 03	**Lab/Tute Time:** 9AM-11AM	**Word Count:** 871 Words	
Due date: 15-10-2023	**Submit Date:** 17-10-2023	**Extension granted** ☑	

If an extension of work is granted, specify date and provide the signature of the lecturer/tutor. Alternatively, attach an email printout or handwritten and signed notice from your lecturer/tutor verifying an extension has been granted.

Extension granted until (date): 17/10/2023 Signature of lecturer/tutor: .....

Late submissions policy	**Days late**	**Penalty applied**
Penalties apply to late submissions and may vary between faculties. Please refer to your faculty's late assessment policy for details.		

Patient/client confidentiality: Where a patient/client case study is undertaken a signed Consent Form must be obtained.

Intentional plagiarism or collusion amounts to cheating under Part 7 of the Monash University (Council) Regulations

Plagiarism: Plagiarism means to take and use another person's ideas and or manner of expressing them and to pass these off as one's own by failing to give appropriate acknowledgement. This includes material from any source, staff, students or the Internet - published and unpublished works.

Collusion: Collusion means unauthorised collaboration on assessable written, oral or practical work with another person. Where there are reasonable grounds for believing that intentional plagiarism or collusion has occurred, this will be reported to the Associate Dean (Education) or nominee, who may disallow the work concerned by prohibiting assessment or refer the matter to the Faculty Discipline Panel for a hearing.

Student Statement:

- I have read the university's Student Academic Integrity Policy and Procedures
- I understand the consequences of engaging in plagiarism and collusion as described in Part 7 of the Monash University (Council) Regulations (academic misconduct).
- I have taken proper care to safeguard this work and made all reasonable efforts to ensure it could not be copied.
- No part of this assignment has been previously submitted as part of another unit/course.
- I acknowledge and agree that the assessor of this assignment may, for the purposes of assessment, reproduce the assignment and:

  i. provide it to another member of faculty and any external marker; and/or

  ii. submit to a text matching/originality checking software; and/or

  iii. submit it to a text matching/originality checking software which may then retain a copy of the assignment on its database for the purpose of future plagiarism checking.

- I certify that I have not plagiarised the work of others or participated in unauthorised collaboration or otherwise breached the academic integrity requirements in the Student Academic Integrity Policy.

Date: 17/10/2023 Signature: Mohib Khan *

Privacy Statement:

For information about how the University deals with your personal information go to http://privacy.monash.edu.au/guidelines/collection-personal-information.html#enrol

MONASH University

Mohib Khan <mkha0146@student.monash.edu>

Short extension application approved | ref #10736121 | FIT3179 | Data Visualisation II Report

Monash Student Global <noreply@f.e.monash.edu> 13 October 2023 at 16:59
To: "mkha0146@student.monash.edu" <mkha0146@student.monash.edu>

Hi Mohib,

We've approved your application for a short extension.

Unit: FIT3179 - Data visualisation

Assessment name: Data Visualisation II Report

Original due date: 15-Oct-2023

New due date: 17-Oct-2023 – your new due date has been updated in Moodle

If you need more help with your assessment, studies or support with your condition, take a look at the range of support that is available to you.

In the event that you need more time, you'll need to submit a new application with new supporting documents.

If you applied for an extension for more than one assessment, we'll let you know the outcome as soon as possible.

We wish you all the best with your studies.

Kind regards,

Special Consideration Team

Monash University

CRICOS Provider 00008C

FIT 3179

Data Visualisation 2

Link to my Repository

Link to my GitHub Page

Word count: 871 Words

Mohib Ali Khan 33370311

The Report

The metro systems domain encompasses intricate urban transportation networks characterised by underground rail systems, serving as lifelines in major cities worldwide. These systems efficiently transport millions of passengers daily.

Metro system data offers critical insights into urban infrastructure and public transportation, with data-driven decision-making becoming paramount in modern cities' development. These insights are invaluable to a wide spectrum of stakeholders, including urban planners, investors, government authorities, and the commuting public (NA,2022). The visualisations, such as donut charts for system count and choropleth maps for ridership, offer succinct yet comprehensive views, empowering efficient decision-making for a sustainable and convenient urban future.

The CSV datasets obtained from kaggle (EFIMPOLIANSKII,2021) serve as valuable sources of information for understanding global metro systems. The "metro_systems_countriescsv" dataset presents key details about metro systems, including their location, length, stations, and annual ridership. The "construction_progress_countries.csv" file tracks the development progress of these systems over time. Finally, the "metro_systems_cities.csv" dataset focuses on annual ridership by city, offering insights into usage trends. Together, these datasets provide a comprehensive foundation for analysing metro systems' characteristics, expansion, and their significance in urban transportation networks worldwide.

The choice of specific visualisation types in the dashboard was driven by their ability to effectively communicate the data and assist users in achieving their tasks. Here's the rationale for selecting each visualisation type and an explanation of special features:

Donut Chart: The donut chart was employed to illustrate the global distribution of metro systems. Donut charts excel at displaying parts of a whole, making them an ideal choice to convey the regional distribution of metro systems. Users can swiftly discern the total number of metro systems around the world.

Choropleth Map: Annual ridership by country is best represented through a choropleth map. This visualisation type uses colour intensity to highlight variations, enabling users to identify trends and discrepancies in ridership across countries.

Bar Charts: Bar charts were used to showcase the top 5 metro systems based on different attributes. They facilitate straightforward comparisons, helping users identify metro systems with the greatest lengths, annual ridership, and the most stations.

Gantt Chart: The Gantt chart effectively depicts the progress of metro system constructions. Its timeline view is well-suited for tracking project development over time, providing users with a clear understanding of construction milestones.

Heat Maps: For analysing correlations within the data, heat maps were chosen. They use colour gradients to visualise relationships between variables, aiding users in identifying patterns and connections in the attributes which contribute to Metro Systems.

Incorporating the well-known mantra "Overview first, zoom and filter, then details on demand" (Craft & Cairns,n.d) was a fundamental principle in structuring the dashboard's layout. The layout follows this user-centric approach by initially providing users with a high-level overview in the first row of visualisations. This is where the donut chart and choropleth map offer a global perspective on metro systems and annual ridership.

As users explore the dashboard, they can employ zoom and filter functions to refine their focus. For instance, they can interact with the bar charts to zoom in on specific details, such as the top 5 metro systems based on various criteria (length, annual ridership, and stations). This interactive capability allows users to tailor their exploration based on their interests and information needs.

Finally, the layout offers "details on demand" by including visualisations like the Gantt chart that illustrates the progress of metro system constructions. Users can access specific details by interacting with this chart, gaining insights into construction timelines. This approach ensures that users can progressively delve into the data, accessing more granular information as they desire.

The colour assignments were region-specific, ensuring that each region maintained a consistent colour scheme throughout the visualisations. In the heat map, colour selection adhered to the guidance from Week 4 pre-reading, which recommended using the 'viridis' colour scheme (FIT 3179,2023).

To establish a clear figure-ground relationship, opacity was strategically employed. This involved adjusting the transparency of text elements around the visualisations in the HTML and applying opacity settings within the donut chart.

Times New Roman for Title asTimes New Roman is a classic and formal typeface. It was chosen for the title to convey a sense of authority and seriousness, which is often expected in formal reports or documents. The title is a key element that needs to stand out and be easily readable, and Times New Roman's clarity and readability make it a suitable choice.

Poppins for Headings with Black Background as Poppins is a modern and versatile sans-serif typeface known for its clean and friendly appearance. The choice of Poppins for headings complements the formal tone set by Times New Roman in the title. The use of a black background with Poppins headings creates a strong visual contrast, making the headings pop and draw attention. This combination of a modern typeface and a bold background adds a touch of contemporary design while maintaining readability and visual impact.

The dashboard guides readers through the visualisations with concise and informative annotations. Each visualisation is accompanied by a brief title or description, and tooltips provide additional context when users interact with data points. This approach ensures that users can navigate the dashboard with a clear understanding of what they're viewing and its significance.

10 largest metro rails in the world. (2022, December 28). Retrieved on October 15, 2023 from The Business Standard website: https://www.tbsnews.net/world/10-largest-metro-rails-world-559090

FIT 3179 Week 4 (2023, August 17) Retrieved October 15, 2023, from Moodle https://d3cgwrxphz0fqu.cloudfront.net/e7/ec/e7ecd8003e9681b1968746d51a304fd01a27e6cb?response-content-disposition=inline%3Bfilename%3D%22FIT3179_Week04%20-%201%20-%20Colour.pdf%22&response-content-type=application%2Fpdf&Expires=1697562000&Signature=DdlnskTeYpIXTlc1crLhQI-wVyE9utacFDTYpH4UQkGC4Z~0EJDfyoLa6vEVRWeNKNQU7nCIMq1OukdJ-vdVJinkIR00WhL3RrOCySFyrtZU4UZRbYIjat0ChPIRC5bcRLNGKDN2Bl1XKbgLCwZyNlgO5wlqoiCFlZO6z8E~TOfvuCZV5cyBHmjALwYxxoFxu68yi8SyVa4jWXwu3Eg~qfqTqA3C4VUhyLdiupiQnYA0-00y6oFevansros03BByJM9rr8FxfLRYFyOzLCnwudJsgyxQWxCKybaqG4Mg-F8SKKqX6EVuz~w85Qr1MgQtcte1MiANk8FqCU7AHapdZQ__&Key-Pair-Id=APKAJRIEZFHR4FGFTJHA

Craft, B., & Cairns, P. (n.d.).Retrieved October 15, 2023 from Beyond Guidelines: What Can We Learn from the Visual Information Seeking Mantra? Retrieved from https://faculty.cc.gatech.edu/~john.stasko/8001/craft05.pdf

EFIMPOLIANSKII. (n.d.). 🚂 World Metro. Retrieved October 15, 2023, from www.kaggle.com website: https://www.kaggle.com/datasets/timmofeyy/-world-metro?select=metro_countries_total.csv

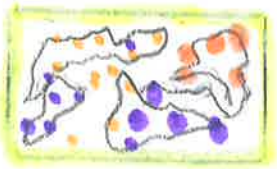
IDEAS

Regional Annual Ridership

- Sankey Diagram

- Choropleth Map

- Proportional Symbol Chart

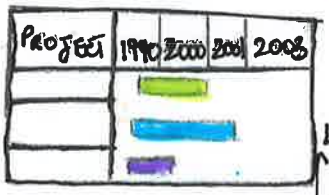


- Donut Chart

Global Metro Systems → Based on Countries

Upcoming Projects → Line Chart

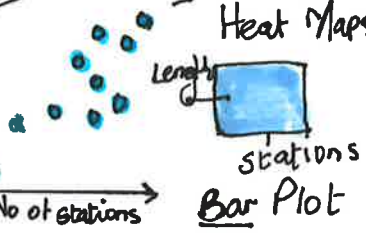
- GANIT Chart



- BUILT AND EXPANSION → Time Spiral

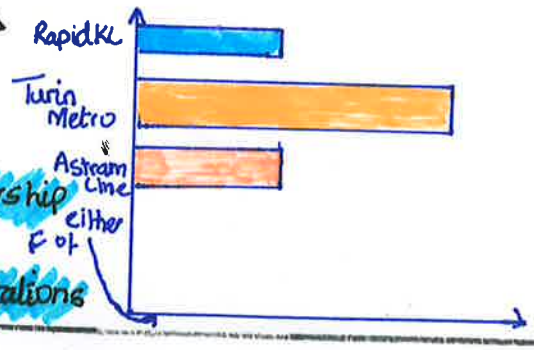
Correlation between Stations and length of metro lines

- Length
- Heat Maps



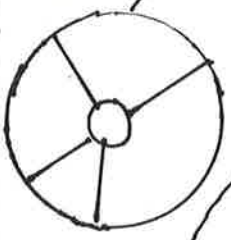
Bar Plot

Top Metro Systems based on 3 categories
- Annual Ridership
- Total length
- Number of Stations

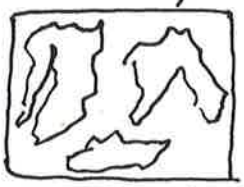


FILTER

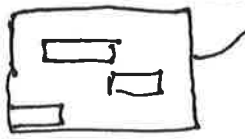
- Regional Ampetro system



- Choropleth Map for Country based Annual Ridership



- Progress of Metros



- Top 5 Metro System

- Stations
- Length

→ Correlation of Length and stations of Metro System

CATEGORIZE

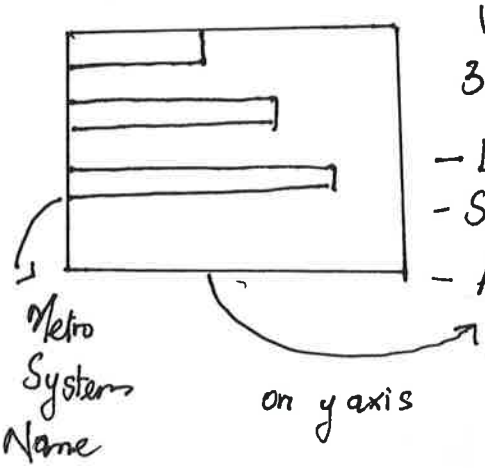
Regional Count of Metro Systems
↓
Country based Annual Ridership
↓
Top 5 Metro Systems
↓
Correlation
↓
Progress of Metro (conclusion)

COMBINE AND FOCUS

Top 5 Metros

Lets divide it into 3 categories

- Length
- Stations
- Annual Ridership

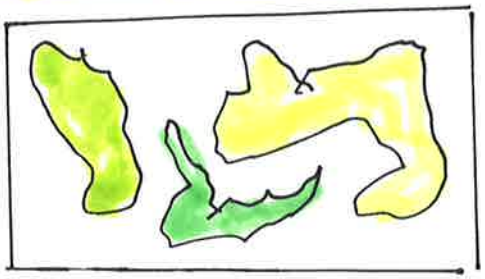


QUESTION

- We would need a correlation column?
- Are we allowed to make a new csv for correlation?
- Is Length and Stations even related?
- Lets add annual-ridership as a correlation too.

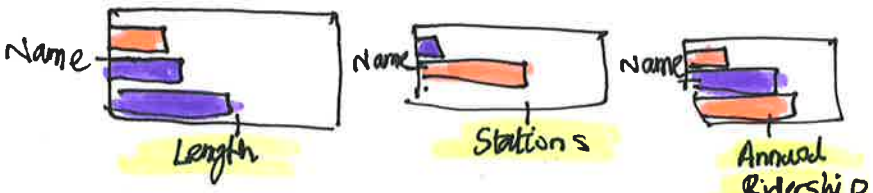
LAYOUT

Annual Ridership of Countries

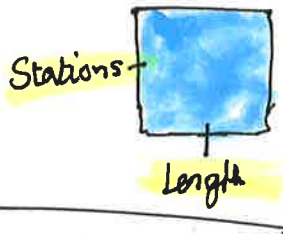


Total Metro System Regionally

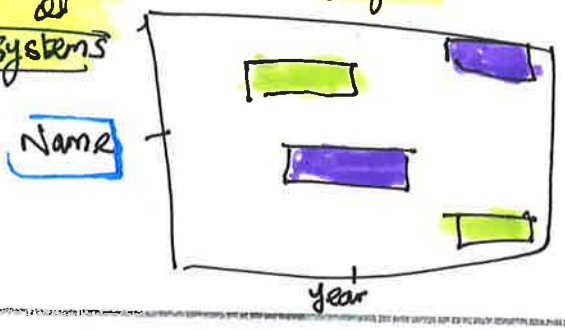
Top 5 Metro Systems



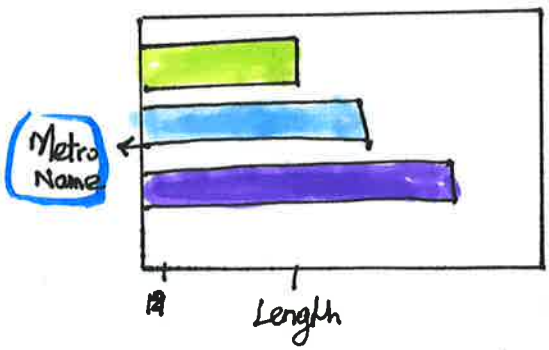
Correlation



Progress of Metro Systems



Focus



FILTER → to filter top 5 of each category

TITLE: Dashboard View

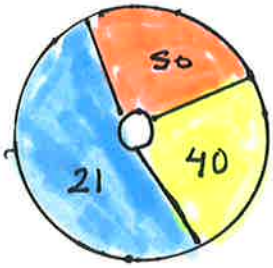
AUTHOR: Mohib Ali Khan

DATE: 1st October 23

SHEET: 02

TOPIC: World Metro Systems

OPERATION



A user hovers around the donut chart to highlight

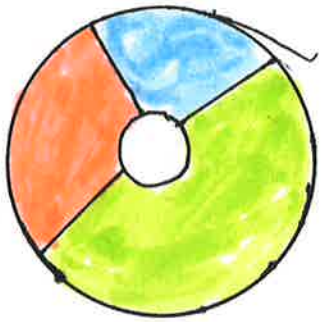
DISCUSSION

- To make a heatmap we would be using pandas
- To have a filter we would need to manually adjust the value
- Top 5 is tough to be calculated with "transform" in vega lite
- To implement the highlighting feature we need to invest a lot of time.

LAYOUT

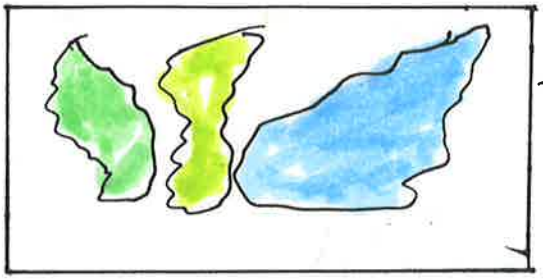
Metro Systems or any title

Text

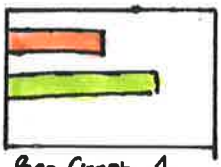


Text

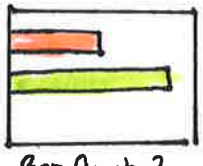
Text



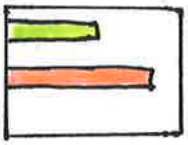
Text



Bar Graph 1

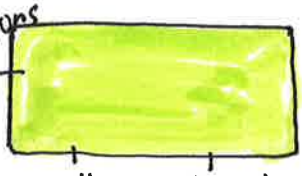


Bar Graph 2



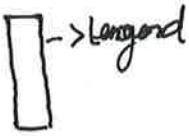
Bar Graph 3

Stations



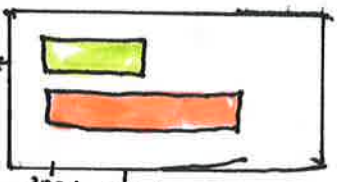
Length

Annual



->Longest

Text



Metro Name

2020

2021

Text

TITLE: DASHBOARD VIEW

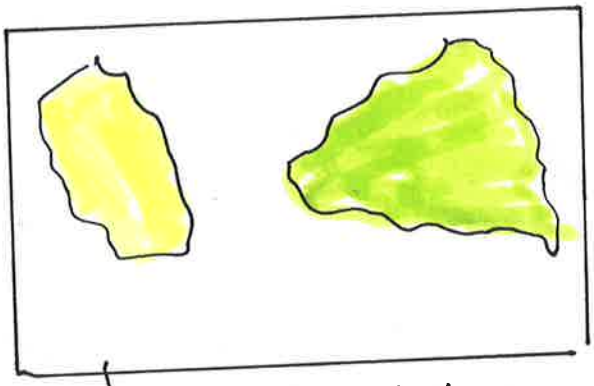
AUTHOR: Mohi'b Ali Khan

DATE: 2nd October 2023

SHEET: 03

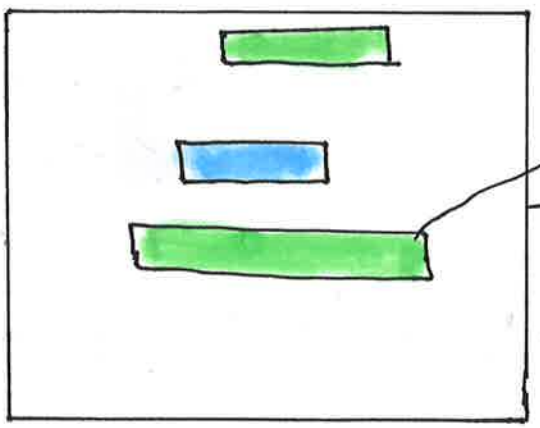
TOPIC: Metro Systems

OPERATION



Each portion of Map is highlighted when user hovers over it by adjusting opacity

FOCUS



Annotation for progress

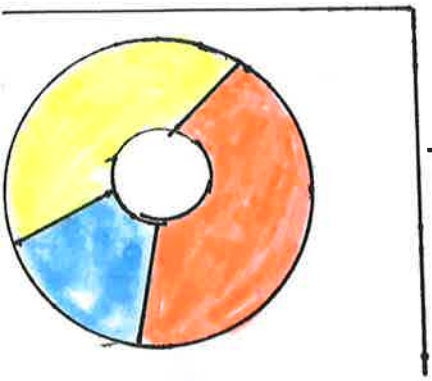
Metro Name

DISCUSSION

- Highlighting the Map is possible?
- The layout not well designed
- We would need regions for our heatmaps
- We need conclusion

LAYOUT

Title



Text

MAP

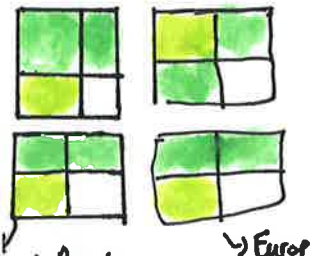
TEXT

Bar Graph 1 | Bar Graph 2 | Bar Graph 3

Text

FILTER → For Bar Graphs

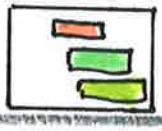
→ North America



Text

South America

→ Europ



Text

TITLE: DASHBOARD VIEW

AUTHOR: Mohib Ali Khan

DATE: 2nd October 23

SHEET: 04

TOPIC: Metro Systems

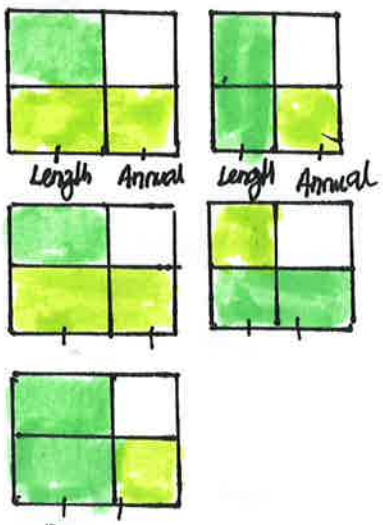
OPERATION

- Consistent Colour Pattern through Regions → Gantt Chart
  → Donut Chart.

-  → Bar Graph filter
  - Show all
  - Length
  - Stations
  - Annual Ridership

FOCUS

Heat Maps or Correlation



For Each Region

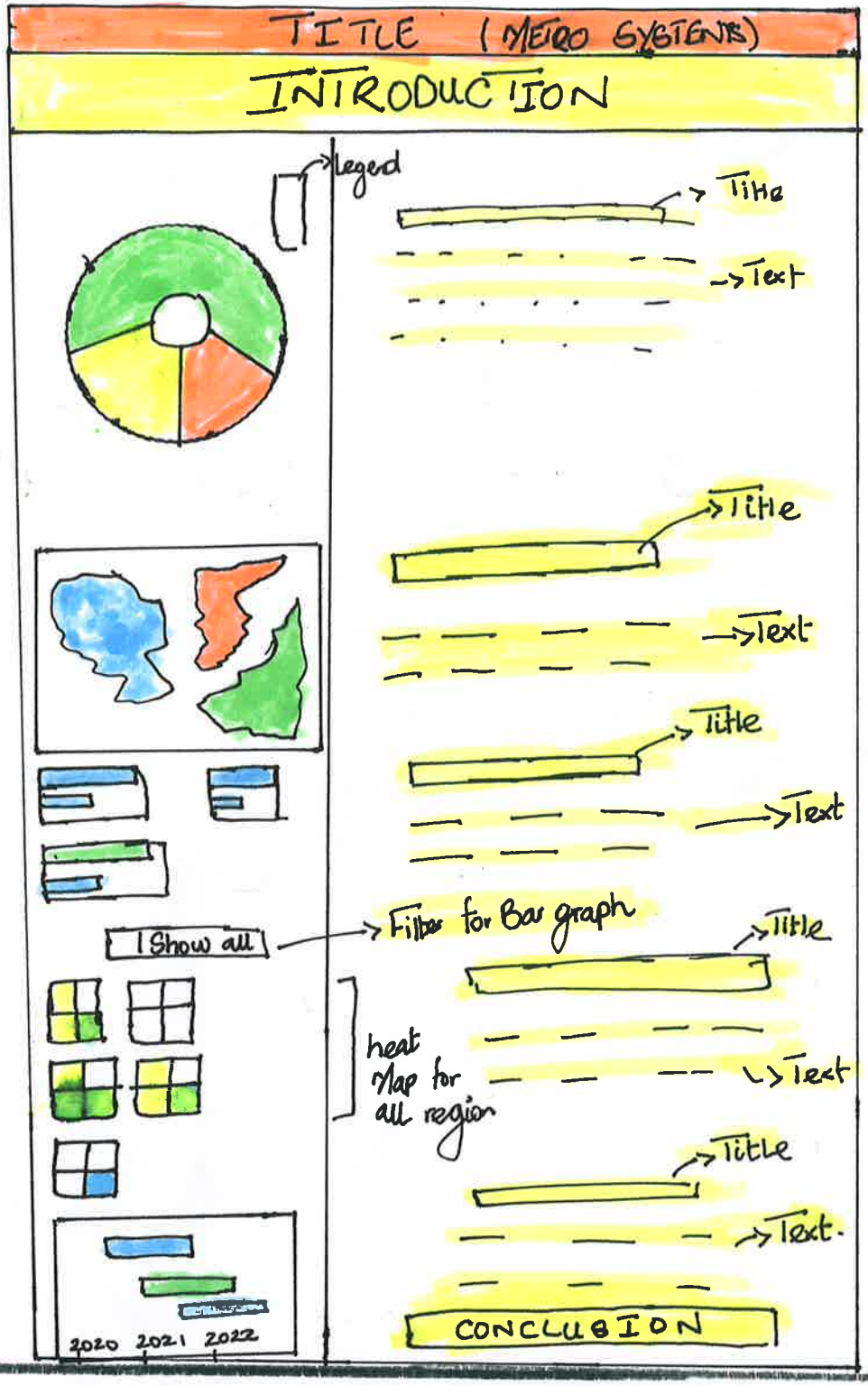
Text → Pattern described

DISCUSSION

- The correlation CSV must be created

- Colour consistency would really help user to distinctively identify regions

LAYOUT



TITLE : Dashboard View

AUTHOR: Mohib Ali Khan

DATE: 2nd October 23

SHEET: 05

TOPIC: Metro Systems

OPERATION

→ Filter for Barchart

→ Title and introduction with a shadow Border

Focus

All the Visualisation | All the text | to maintain the visual hierarchy

DISCUSSION

→ A very clean look

- Good sight line