DATA VISUALIZATION PROJECT

Name: Aman Roy, **Student ID:** 35011203, **Teaching Associate:** Maxim Srour.

Project Title: Comparison of fuel consumption between planes and cars

Contents

Table of figures

Section 1: Introduction

Problem Description

Recently, there has been a strong push towards developing electric vehicles to lower fuel consumption and emissions. While electric vehicles are a positive step, many other modes of transportation consume significantly more fuel and emit higher CO_2 levels than a single car could in a month. Planes, in particular, have fuel consumption and emissions so high that with the same amount of fuel, cars could operate for months. This disparity motivates me to compare cars and planes across various parameters to highlight these differences.

Motivation

I have a strong interest in cars and closely follow various manufacturers. In recent years, many car manufacturers have phased out their internal combustion engine vehicles. While the shift is largely focused on four-wheeled vehicles, other modes like planes—which consume vast amounts of fuel—have not received the same attention. This drives my interest in comparing cars and planes based on different parameters.

Questions

1. How much fuel does a single flight use compared to a car, and which mode of transport is more efficient when considering factors like cost, fuel consumption and co2 emissions?
2. What is the typical lifetime distance travelled by planes versus cars, and how does this compare when looking at the total fuel consumption and CO_2 emissions of all cars versus all planes worldwide? Which mode is more economical?

Intended Audience

The primary audience for this visualization includes environmentally conscious citizens and policymakers focused on understanding and reducing the environmental impacts of various transportation modes.

1. **Environmentally Conscious Citizens**
   This audience is interested in making choices that lower their carbon footprint. By exploring the differences in fuel consumption and CO_2 emissions between cars and planes, they can make informed decisions about travel—choosing to drive or fly based on sustainability. The visualization is designed with engaging, interactive visuals and clear metrics, such as cost per seat and emissions per kilometer, allowing these users to quickly assess the environmental impact of different vehicles and understand the larger implications of their choices.
2. **Policymakers**
   Policy makers, who shape regulations and policies influencing transportation practices, are also a key audience. This visualization offers them a direct comparison between cars and planes in terms of emissions and fuel consumption, highlighting the environmental costs of each mode. The insights provided can support decisions related to fuel standards, investments in public transportation infrastructure, and incentives for greener options. By presenting the fuel and emission disparities visually, this project aims to give policymakers a data-driven basis for developing environmental legislation.

Section 2: Design Process

Using the **5 Design Sheet methodology**, I was able to generate, combine, and refine ideas for my project. I explored various visualization approaches throughout this process. For example, I initially considered using a donut chart in Sheet 4 to represent all planes and cars in a single large circle. However, this approach was not feasible due to the high number of car models—around 1,000—which would make the chart overly complex and difficult to interpret. I also experimented with a dot chart in Sheet 2 as a possible solution, and tried using dual-axis line charts in Sheet 3 to illustrate differences between cars and planes. Ultimately, I decided on a stacked bar chart for the first part of the project to effectively compare cars and planes, as it could convey multiple parameters at a glance while maintaining clarity.

For the second part of the project, I chose a **map to display different routes**, with line graphs below to show how various parameters compare between planes and cars over specific travel distances. This route map allows users to select different cities and car or plane models, making it highly interactive and engaging. Through this setup, users can observe how different vehicles perform on parameters like fuel consumption and emissions, helping them understand which models are more environmentally friendly.

To illustrate clusters of cars based on similar features, I selected a **scatter plot**. This chart shows clusters based on two parameters, making it easy to visually distinguish between different car groups. For the final section on **lifetime fuel consumption**, I opted for a simple bar chart due to its straightforward readability and effectiveness in showing clear comparisons. This section includes two bar charts: one displaying total consumption across parameters, and another comparing the cumulative impact of all cars to that of all planes worldwide. Given the substantial difference in consumption and emissions, the bar chart format is well-suited to communicate this contrast in an impactful way.

In my **narrative visualization**, I chose a **scrollytelling approach**. This modern technique integrates text, visuals, and interactivity as users scroll through the content, guiding them through a structured and compelling story. Scrollytelling allows users to engage with data in a linear, narrative-driven way, which aligns with my preference for a smooth and immersive user experience.

My project focuses on two main entities—**planes and cars**—which allowed me to maintain a minimal color scheme. I used two distinct colors for each category to ensure clarity and visual contrast. Each color was carefully chosen to represent different metrics, maintaining coherence across all visualizations. Specifically:

- **Blue** for fuel consumption in line charts.
- **Green** for cost comparisons.
- **Red** for CO_2 emissions.
- **Cyan (#00C0C0)** for cars and **coral (#FF6F61)** for planes in bar charts and maps.

The overall theme uses **neon green on a black background** for readability and emphasis, providing a high-contrast, tech-focused aesthetic. Neon green is applied to text, buttons, and interface highlights, enhancing both visual appeal and readability on digital screens, particularly for an audience interested in environmental data.

Typography: I chose a modern, sans-serif font that complements the clean, tech-driven design. Sans-serif fonts improve readability on digital platforms, and paired with the neon color scheme, they ensure clarity across different devices.

Visual Variables: I utilized various visual variables to enhance clarity and engagement:

- **Color**: Each metric and category has distinct colors to prevent confusion and reinforce the visual hierarchy. Neon green is consistently used for UI elements, creating a cohesive user experience.
- **Position**: The content follows a logical, top-to-bottom flow, with each section building on the previous one, guiding users from introductory context to detailed comparisons and concluding insights.
- **Size**: Key charts like the stacked bar chart and clustering plot are larger to emphasize their importance, while smaller elements like buttons and controls are compactly positioned to reduce clutter.
- **Shape**: Different shapes (bars, lines, points) correspond to data types—bars for categorical comparisons, lines for time-based or sequential data, and points in scatter plots for clustering—enhancing the user's ability to interpret the data.
- **Orientation**: I used a horizontal layout for the fuel consumption chart to improve readability, especially for long model names, while vertical orientations in clustering and global comparison charts help separate categories visually.

Interactivity: I incorporated several interactive features to enhance the user experience:

- **Tooltips**: Hover-based tooltips provide additional details, such as fuel consumption or cost per model, allowing users to interact directly with the data without cluttering the visuals.
- **Controls for Search, Filter, and Pagination**: These features empower users to customize their view, aligning with a user-centered design approach and facilitating in-depth exploration of specific data points.

Munzner's What-Why-How Framework:

- **What**: This project includes a mix of quantitative, categorical, and geospatial data, structured to enable meaningful comparisons.
- **Why**: The primary goal is to support exploration, comparison, and analysis, allowing users to identify patterns and make informed decisions regarding environmental impact.
- **How**: Through a combination of color, shape, size, and interaction, complex data is presented in an accessible and engaging way, addressing the needs of environmentally conscious users and policymakers.

By applying **Munzner's framework**, I ensured that my project remains purpose-driven and user-centered, effectively using data to tell a compelling story.

In my data visualization project, I applied several **Gestalt laws**, either explicitly or implicitly, to help users interpret and interact with the visual elements more intuitively. Here's a look at some of the key Gestalt principles I used:

1. **Proximity**
    - **Application**: I grouped related elements, such as radio buttons, sliders, and search boxes, closely together in sections. This helps users recognize which controls or inputs belong to each chart or visualization, reinforcing their functional connections.
    - **Effect**: Proximity reduces cognitive load and clarifies which inputs affect specific parts of the visualization. Users can easily find and adjust controls associated with specific sections, like selecting a vehicle type or metric.
2. **Similarity**
    - **Application**: I consistently used similar colors, font styles, and input designs across different sections, such as buttons with the same styling or neon-colored headings.
    - **Effect**: Similarity allows users to quickly understand that similar elements perform related functions. For example, buttons with the same color and styling suggest they perform similar actions, like navigating or clearing inputs, making the interface feel cohesive.
3. **Figure-Ground**
    - **Application**: I used a dark background with neon-colored text and highlighted sections to make key content stand out, guiding users' attention to focus areas.
    - **Effect**: This contrast makes each section, like "Fuel Consumption Comparison" or "Car Clustering Analysis," more readable and visually prominent against the background, helping users quickly identify main sections and interactive elements.
4. **Continuity**
    - **Application**: The layout follows a clear vertical flow from the introduction to each section, leading the user's eye naturally down the page.
    - **Effect**: Continuity ensures that users follow the narrative in the intended order, starting with high-level information and then exploring specific comparisons and insights. It also encourages users to explore each section sequentially, supporting a logical flow.
5. **Closure**
    - **Application**: Boxes and borders around charts, maps, and control sections give each part a distinct visual "closure," setting boundaries around each group of related elements.
    - **Effect**: Closure helps users understand the grouping of related information, like charts and their associated controls. Each chart appears within a defined box, creating a clear separation from other sections and making it easier to interpret each visualization individually.

6. **Symmetry**
    - **Application**: I used a balanced layout, with centered elements and symmetrical spacing in charts, controls, and text, to create a sense of order.
    - **Effect**: Symmetry makes the interface visually pleasing and reduces confusion. Users can rely on predictable locations for controls and focus more on the content and analysis without distractions.
7. **Common Fate**
    - **Application**: Elements that interact together, like control buttons that modify the charts, are aligned and grouped. The interactive nature, such as buttons responding to user inputs or sliders updating charts, signals that they are part of a cohesive action set.
    - **Effect**: When users see that certain controls modify specific visualizations, they understand that those elements are functionally linked, enhancing usability and the interactive experience.

Hybrid Approach: The project combines both author-driven and user-driven elements, creating a balanced hybrid style. The author-driven aspect is evident in the **sequential narrative flow**, where scrollytelling leads users step-by-step through the story, with each section focusing on specific comparisons, like cost, emissions, and fuel consumption. This structure conveys a clear perspective, guiding users to understand the environmental impact of various transport options. Meanwhile, the user-driven aspect comes from the interactive exploration tools—**search, filter, pagination, and route comparisons**—which allow users to focus on models, routes, or metrics that interest them. Additionally, clustering and map customization give users control over specific data points. This hybrid approach supports my objective of informing environmentally conscious users and policymakers by providing a structured story while allowing them to delve deeper into particular insights.

For my final design, I selected a **stacked bar chart** to compare different car models across various parameters, using the same approach for planes. Stacked bar charts are easy to interpret and effectively highlight differences between attributes, making them suitable for both specialists and general viewers—an ideal choice for the first chart on the website. For the second visualization, I wanted to compare the travel distance of cars and planes between cities, so a **route map** was a natural fit. This map allows viewers to select different cities and car or plane models, offering an interactive way to explore the data. Based on their selection, the graphs below dynamically update, showing the user differences in parameters such as fuel consumption, cost, and CO_2 emissions. This feature helps users understand which models are better for the environment and which ones to watch out for.

To illustrate car clusters, I chose a **scatter plot implementing k-means clustering**, where clusters based on two parameters are visually distinct, allowing users to quickly identify patterns. For the final section on **lifetime fuel consumption**, I opted for a **simple bar chart** due to its straightforward readability. This section includes two bar charts: one showing total consumption across parameters and another comparing all cars to all planes worldwide. Given the significant difference between them, the bar chart effectively conveys this contrast in a clear and impactful way.

Section 3: Implementation

3.1 Technical implementation

High-Level Implementation Description

This project uses **R** for developing an interactive data visualization application. The main libraries utilized in this implementation include:

- **Shiny**: To create a responsive and dynamic user interface and server logic.
- **Plotly**: For generating interactive plots, enabling users to explore data visually.
- **dplyr**: To facilitate data manipulation and wrangling for filtering and summarizing data based on user input.
- **leaflet**: For rendering interactive maps to show route comparisons between cities, particularly useful for visualizing distances in the context of fuel consumption and CO2 emissions.
- **shinythemes**: For applying a consistent theme across the app, ensuring a modern and cohesive aesthetic.
- **geosphere**: Used for calculating distances and creating route-based comparisons for the maps.
- **shinyjs**: To handle specific UI elements, such as disabling/enabling buttons based on user interactions.

Challenges and Complexity

The project presented several challenges:

1. **Data Wrangling**: Extensive data wrangling was required, particularly in handling datasets with varying units (e.g., CO2 emissions for cars in grams vs. planes in kilograms) and normalizing values for a fair comparison.
2. **Interactive Plot Adjustments**: Ensuring that hover information in Plotly did not display redundant data fields (e.g., duplicate labels) required customization within the Plotly structure.
3. **UI Customization**: Applying consistent and visually appealing neon themes across the application required custom CSS adjustments and frequent testing to ensure readability, especially on a dark background.
4. **Responsive Layout**: The need for a highly interactive and responsive UI led to the inclusion of flexible layout elements, aligning inputs, buttons, and outputs across devices.

Difference between 5 design sheet and the final webpage:

Map’s part having three extra line graph which change according to the selected car and plane models comparing them.

In my 5-design sheet I did not show my clustering diagram which I felt was necessary to add in my final project after seeing some examples of previous project implementations shown in the applied lectures

In the end I decided to show tow bar graphs instead of one because of my narrative story, I was originally planning to show both of them together but then I was not able to create a proper flow which would justify them being together and that's when I decided to separate them from each other

In my 5-design sheet it was not clearly explained that my layout shows the user should scroll to get to the next graph but those arrows mean that the user is scrolling to get to the next graph.

3.2 Interactive Narrative Visualisation Implementation

In my first chart I used a stacked bar-chart with buttons to switch between cars and planes, the user has the option to search for a particular car or plane of their liking, they can select the different parameters to view and sort the table in ascending or descending order, with the help of the slider the user can decide to show how many bins they what to see, the graph also has a hover effect which shows the exact value for that car of the parameter selected, the page buttons below are added because there are around a 1000 car models who's information can be looked up and I could not fit them all in a single slide as it would make the text unreadable, that's why I created multiple pages for it.

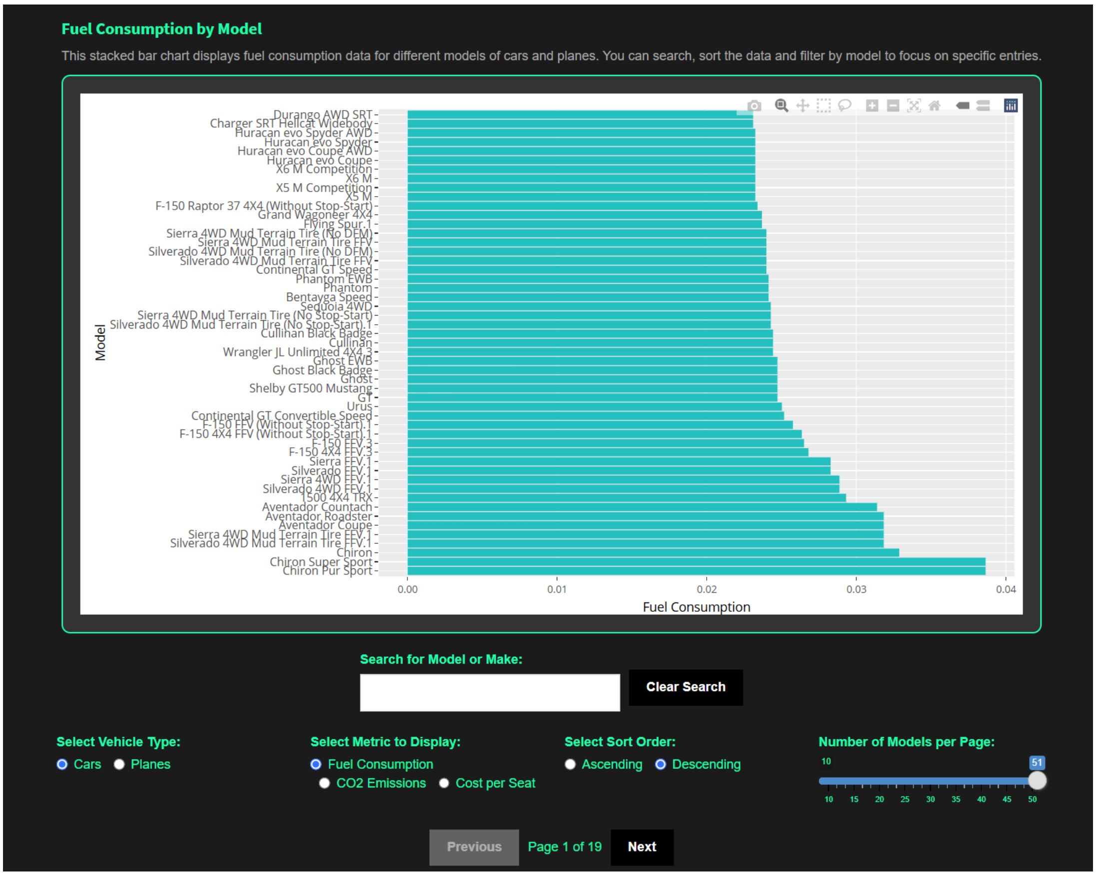

Figure 1

My second chart is the map which helps the user to select a source city and a destination city from a dropdown menu and similarly they can select or search for a specific car or plane mode to compare when you hover over the lines you can see the car and plane model you selected and based on these selections the line graph below also change their values, hovering over the line graph points will show the exact value calculated for the vehicles.

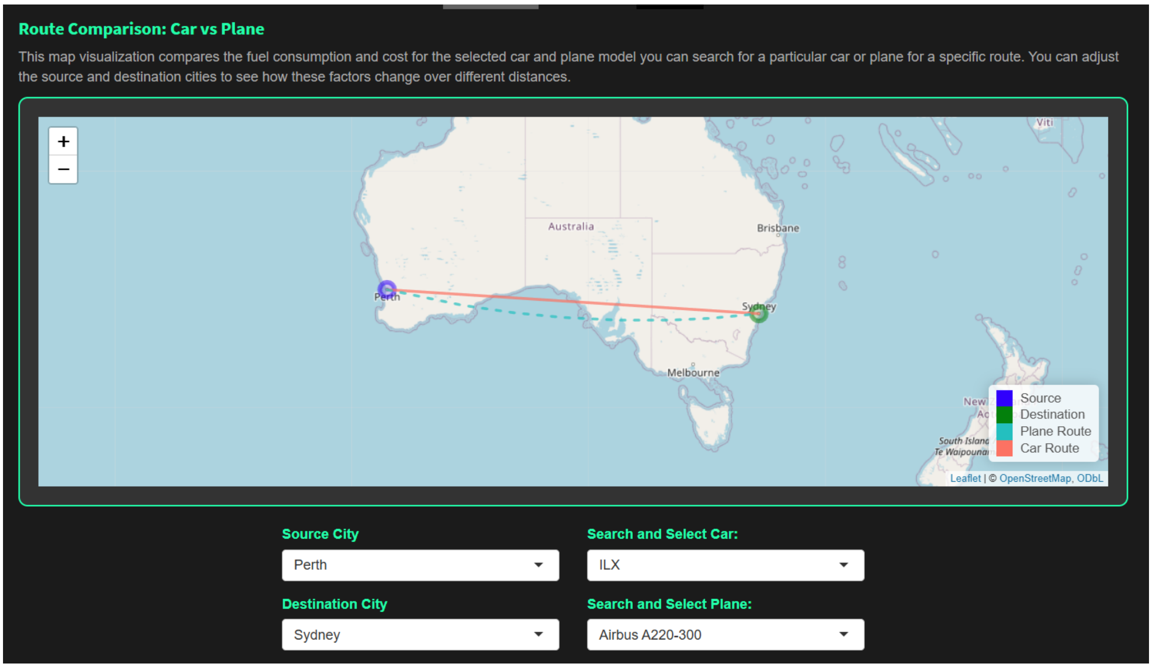

Figure 2

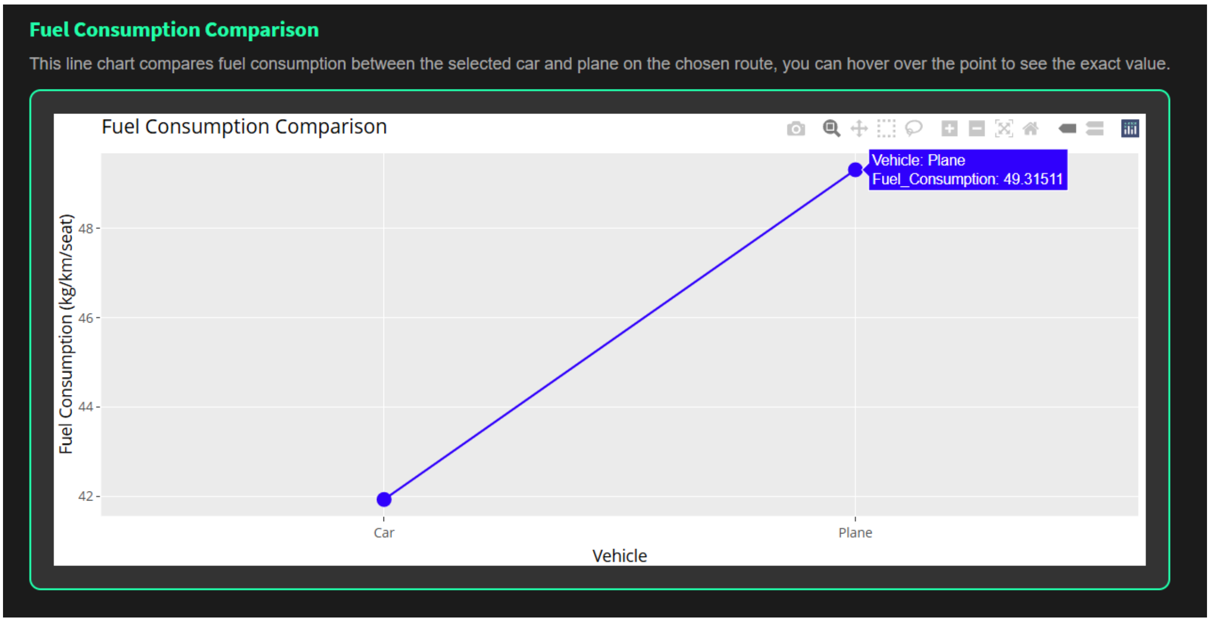

Figure 3

In the cluster parameter you can check the groups which are formed when you select different parameters, you can also select the number of clusters that you want to form via a slider and the site prompts you to select at least two parameters to view the clusters.

Figure 4

The last two charts for my website are just bar charts which are easy to interpret by any user, they have the option to change the parameters and when hovering over the bars you can see the exact value of that bar.

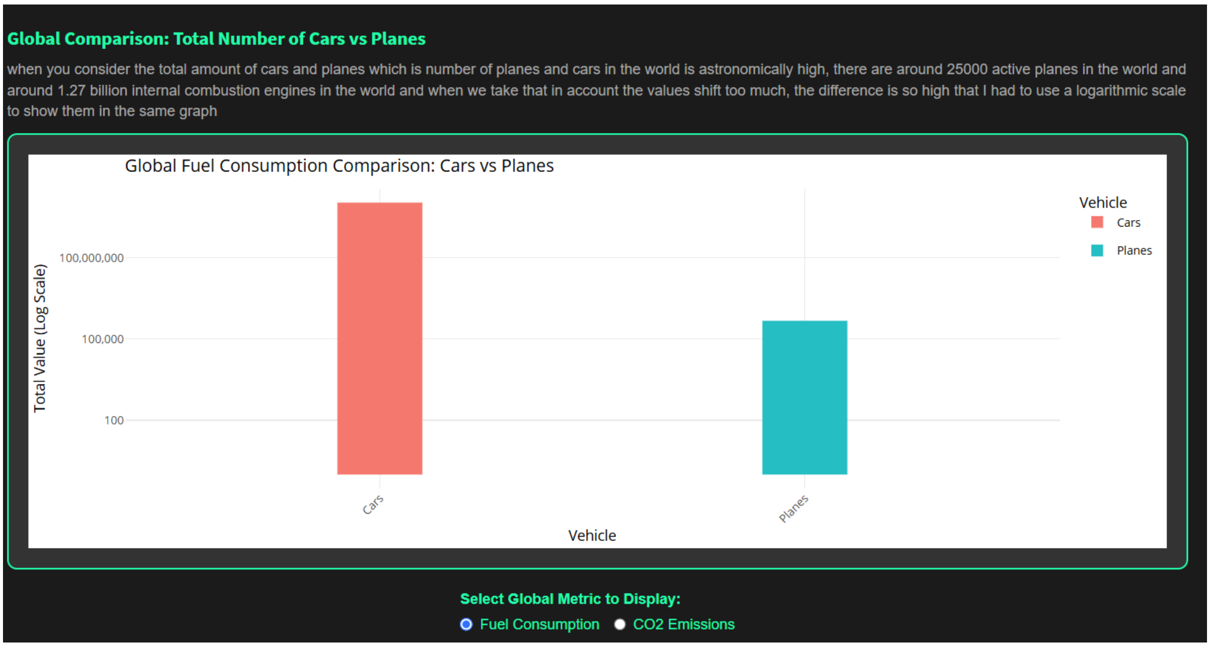

Figure 5

The last two sections are conclusion and data sources.

3.3 Using the Implementation

To run the code file you just need to install the appropriate libraries mentioned in section 3.1

The things that a viewer may miss in my project are:

Filter and Sorting Options: I've included controls for filtering and sorting vehicle types, metrics, and ordering, although these might not be immediately apparent to all users. It's important to note that users can search and filter by specific car or plane models to customize the visualizations.

Pagination and Search for Models: In the fuel consumption chart, I've added pagination to allow browsing through all models. There's also a search function to filter by specific model names, making analysis easier, though this feature could be easily missed.

Route Map Comparison: The map lets users compare fuel, CO_2 emissions, and cost for selected car and plane models on specific routes. Adjusting source and destination cities provides unique insights, and users should be aware that they can search and select vehicles for this comparison.

Clustering Analysis: In the clustering section, users can select different features for clustering analysis, which affects how data points are grouped visually.

Global Comparison and Log Scale: The global comparison chart uses a log scale to emphasize the vast differences in fuel consumption and emissions between cars and planes worldwide, a detail that might not be immediately obvious.

Section 4: Conclusion

Through this narrative visualization project, I have successfully illustrated the differences in fuel consumption, cost, and CO_2 emissions between cars and planes on both a per-seat basis and a global scale. Key findings include the significantly higher fuel consumption and emissions of planes per kilometer compared to cars, although the immense volume of cars in operation worldwide results in a considerable cumulative environmental impact. Additionally, clustering analyses provided insights into car models based on fuel efficiency and emissions, highlighting patterns that could guide user choices or policy considerations.

Reflecting on this project, I've learned a lot about handling complex datasets, using R Shiny for interactive visualizations, and designing an intuitive, visually engaging narrative experience. Balancing the technical elements with user experience and clarity was challenging but rewarding. In hindsight, I would have focused earlier on optimizing the data wrangling process and simplifying the UI, as some initial complexities in the dataset caused delays.

For future work, I'd like to expand the project by integrating additional environmental impact metrics, such as lifetime maintenance emissions for vehicles, and by incorporating predictive modelling for future emissions trends. I would also like to add data from cargo ships and trains, as they are popular modes of transport and would provide a more comprehensive comparison.

Section 5: Bibliography

Lawlor, L. C. (Laurence C. (1978). *Usage patterns of urban cars: their effect on fuel consumption and emissions*. Australian Govt. Pub. Service.

Shiny - Using sliders. (2017, June 28). Shiny.posit.co. https://shiny.posit.co/r/articles/build/sliders/

Shiny - Kmeans example. (2014, July 29). Shiny.posit.co. https://shiny.posit.co/r/gallery/start-simple/kmeans-example/

How Many Miles Can a Car Last? (n.d.). Www.progressive.com. https://www.progressive.com/answers/how-many-miles-does-a-car-last/

TODAY, J. C., special for USA. (n.d.). *Ask the Captain: How far does a jet fly during its lifetime?* USA TODAY. https://www.usatoday.com/story/travel/columnist/cox/2012/11/19/ask-the-captain-how-far-does-a-jet-fly-during-its-lifetime/1712269/

freeCodeCamp.org. (2021). R Shiny for Data Science Tutorial – Build Interactive Data-Driven Web Apps. In *YouTube*. https://www.youtube.com/watch?v=9uFQECk30kA

Section 6: Appendix

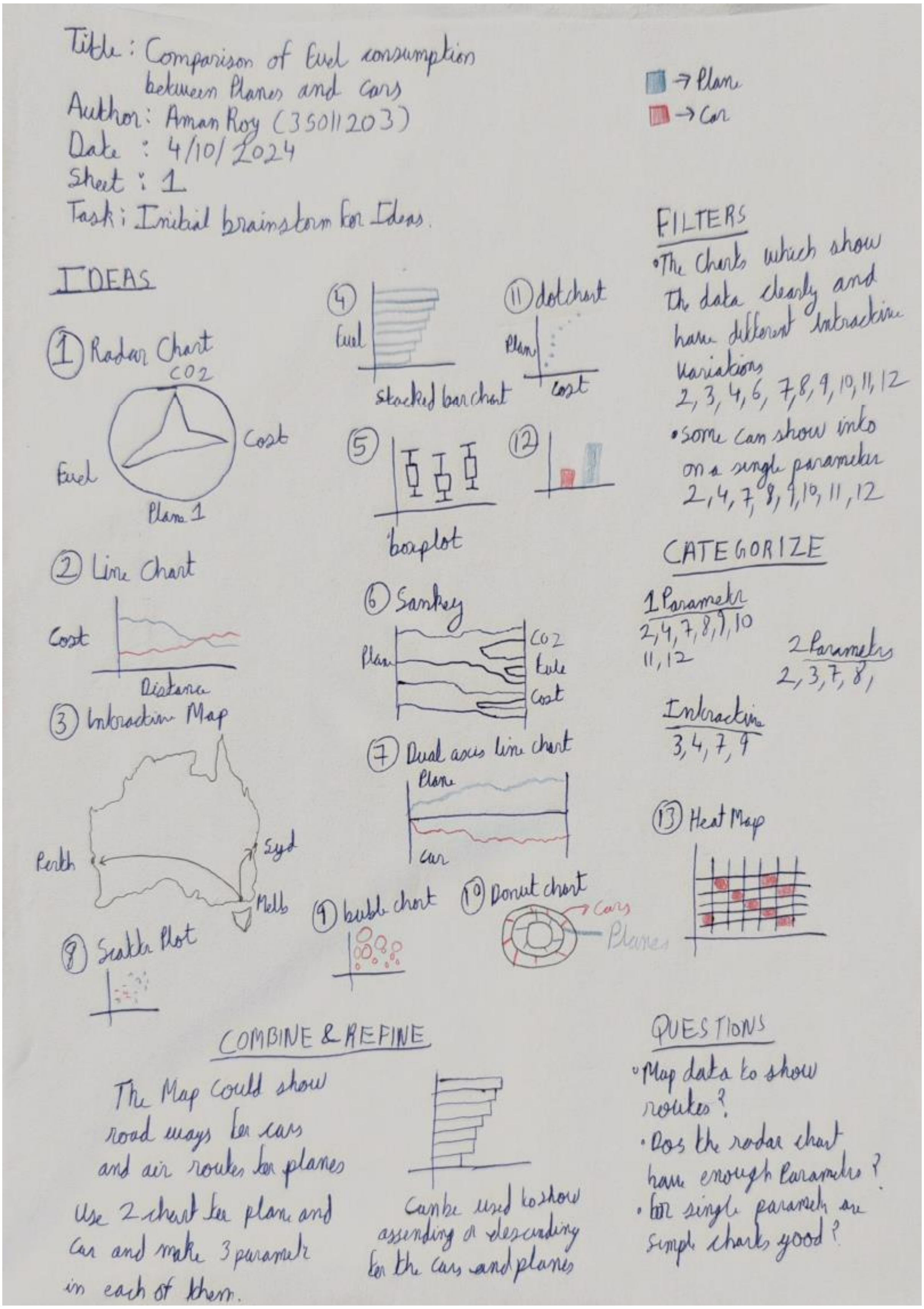

Figure 6

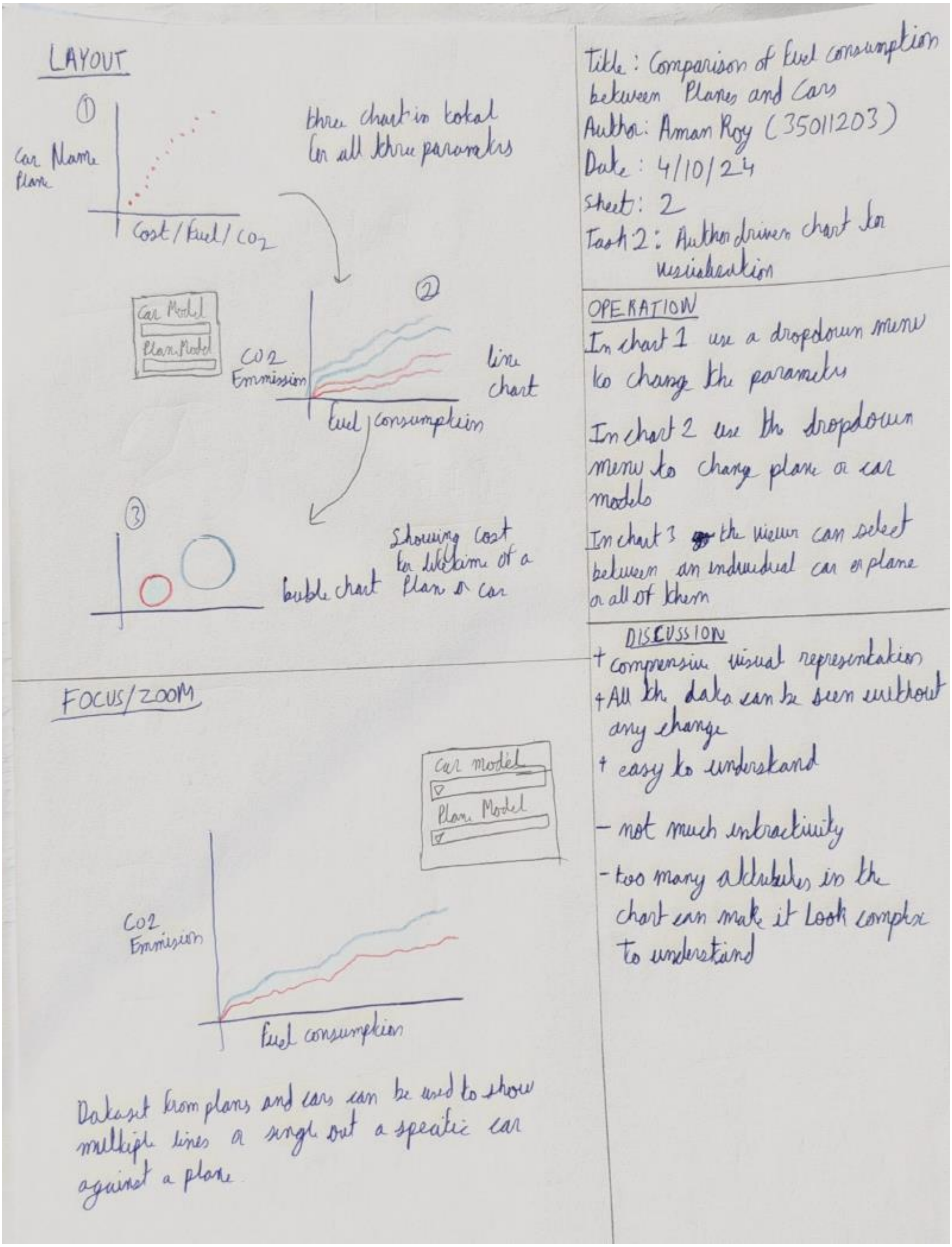
LAYOUT
①
Car Name
Plane
Cost / Fuel / CO2
three chart in total
for all three parameters
Car Model
Plane Model
②
CO2
Emmission
line
chart
Fuel consumption
③
Showing Cost
for lifetime of a
buble chart Plane or car
FOCUS/ZOOM
Car model
Plane Model
CO2
Emmision
Fuel consumption
Dataset from planes and cars can be used to show
multiple lines or single out a specific car
against a plane
Title: Comparison of fuel consumption
between Planes and Cars
Author: Aman Roy (35011203)
Date: 4/10/24
Sheet: 2
Task 2: Author driven chart for
visualisation
OPERATION
In chart 1 use a dropdown menu
to change the parameters
In chart 2 use the dropdown
menu to change plane or car
models
In chart 3 the viewer can select
between an individual car or plane
or all of them
DISCUSSION
+ comprensive visual representation
+ All the data can be seen without
any change
+ easy to understand
- not much interactivity
- too many attributes in the
chart can make it look complex
to understand

Figure 7

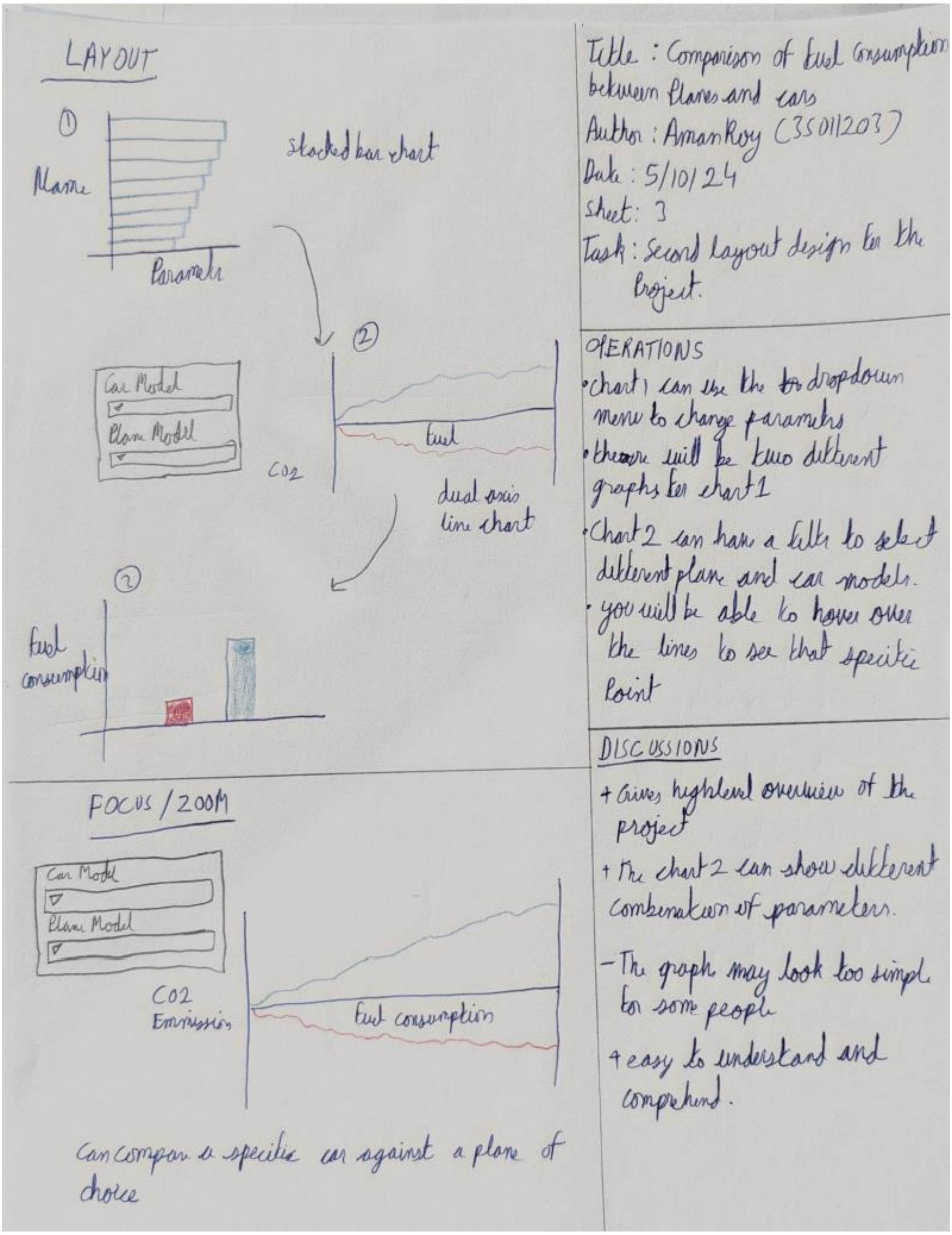
LAYOUT
Name
Stacked bar chart
Parameter
Car Model
Plane Model
CO2
fuel
dual axis line chart
fuel consumption
FOCUS / ZOOM
Car Model
Plane Model
CO2 Emmission
fuel consumption
Can compare a specific car against a plane of choice
Title : Comparison of fuel consumption between planes and cars
Author : Aman Roy (35011203)
Date : 5/10/24
Sheet: 3
Task : Second layout design for the project.
OPERATIONS
• chart 1 can use the ~~to~~ dropdown menu to change parameters
• there will be two different graphs for chart 1
• Chart 2 can have a filter to select different plane and car models.
• you will be able to hover over the lines to see that specific point
DISCUSSIONS
+ Gives highlevel overview of the project
+ The chart 2 can show different combination of parameters.
– The graph may look too simple for some people
+ easy to understand and comprehend.

Figure 8

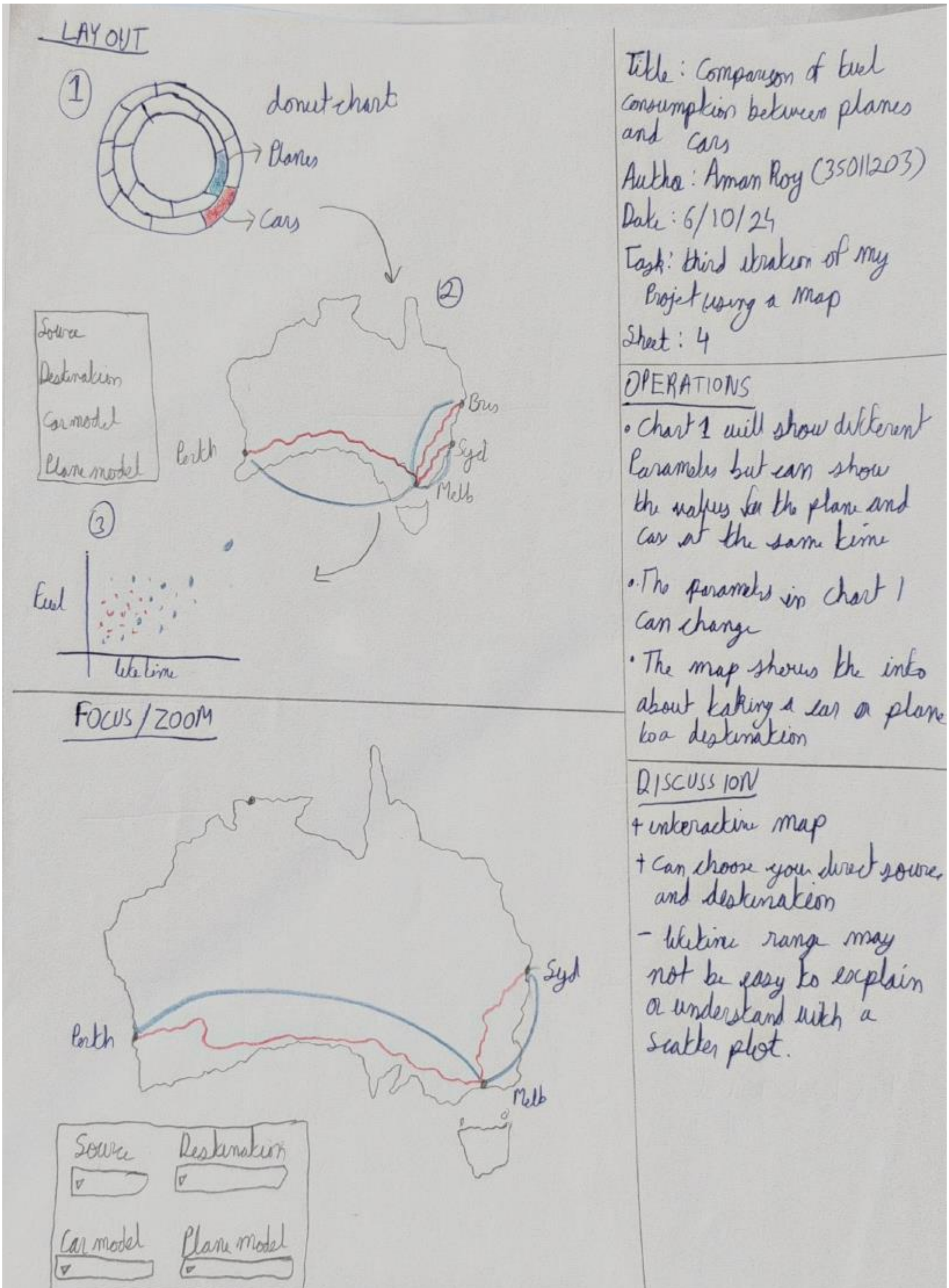
LAYOUT
1
donut chart
Planes
Cars
2
Source
Destination
Car model
Plane model
Perth
Bris
Syd
Melb
3
Fuel
Lifetime
FOCUS / ZOOM
Syd
Perth
Melb
Source
Destination
Car model
Plane model
Title: Comparison of fuel consumption between planes and cars
Author: Aman Roy (35011203)
Date: 6/10/24
Task: third iteration of my Project using a map
Sheet: 4
OPERATIONS
• Chart 1 will show different Parameters but can show the values for the plane and car at the same time
• The parameters in chart 1 can change
• The map shows the info about taking a car or plane to a destination
DISCUSSION
+ interactive map
+ Can choose your direct source and destination
- lifetime range may not be easy to explain or understand with a scatter plot.

Figure 9

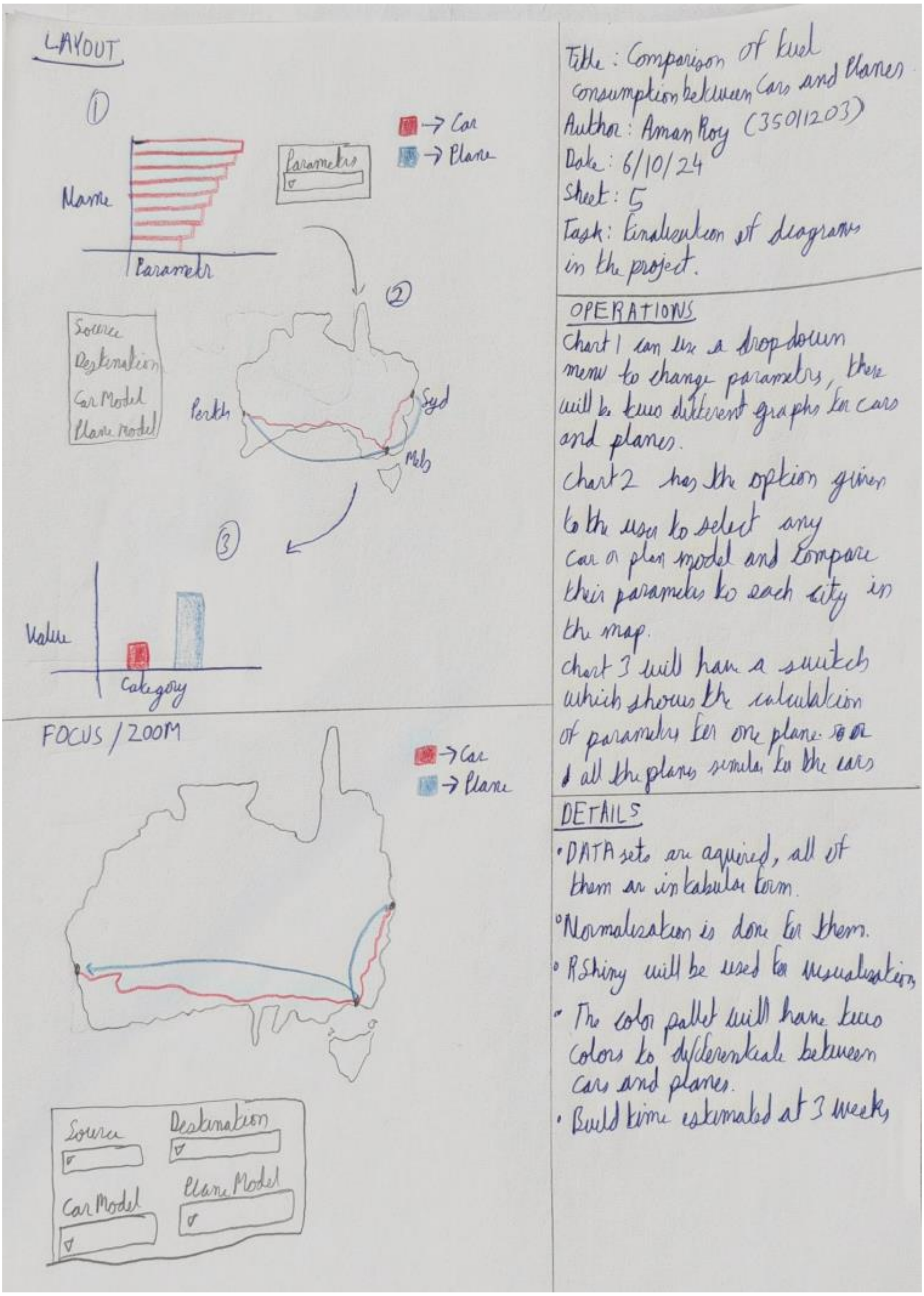
LAYOUT
1
Car
Plane
Parameters
Name
Parameter
2
Source
Destination
Car Model
Plane model
Perth
Syd
Melb
3
Value
Category
FOCUS / ZOOM
Car
Plane
Source
Destination
Car Model
Plane Model
Title: Comparison of fuel consumption between Cars and Planes
Author: Aman Roy (35011203)
Date: 6/10/24
Sheet: 5
Task: Finalization of diagrams in the project.
OPERATIONS
Chart 1 can use a drop down menu to change parameters, there will be two different graphs for cars and planes.
chart 2 has the option given to the user to select any car or plan model and compare their parameters to each city in the map.
chart 3 will have a switch which shows the calculation of parameters for one plane & all the planes similar for the cars
DETAILS
• DATA sets are aquired, all of them are in tabular form.
• Normalization is done for them.
• RShiny will be used for visualization
• The color pallet will have two colors to differentiate between cars and planes.
• Build time estimated at 3 weeks

Figure 10