CDS6324 Data Visualization
Faculty of Computing and Informatics

TC1L Group 26

Project Documentation

Name	Student ID	Email
Lee Le Xuan	1211102976	1211102976@student.mmu.edu.my
Poh Ern Qi	1211101398	1211101398@student.mmu.edu.my
Choo Chee Choong	1221302637	1221302637@student.mmu.edu.my
Yap Weng Hong	1211103023	1211103023@student.mmu.edu.my

Data Sources

Source:
https://www.kaggle.com/datasets/safaeahb/car-sales-analysis-dashboard?resource=download&select=car+sales.csv

The dataset is a comprehensive **car sales dataset** sourced from Kaggle and consists of 23,906 records representing detailed car sales transactions from 2022 to 2023. It contains 17 attributes encompassing various dimensions of the car sales process, including customer demographics, vehicle specifications, sales pricing, and dealership details.

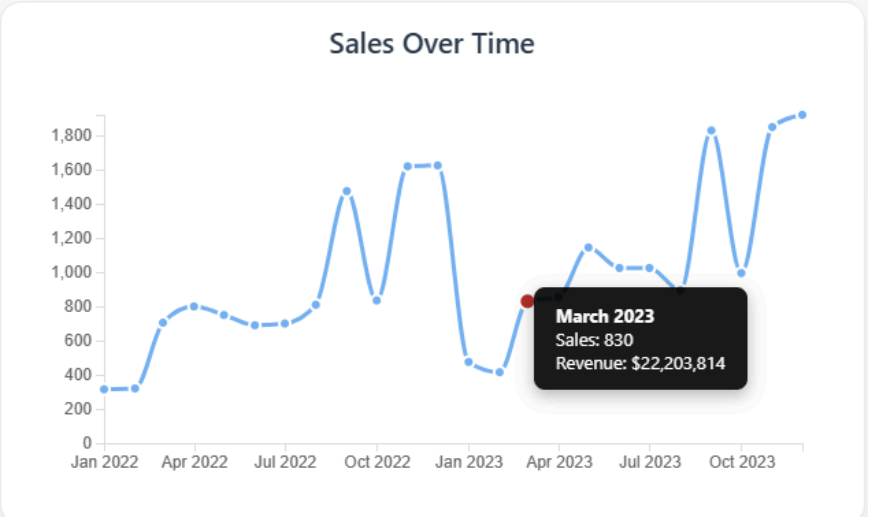
Attribute Name	Description
Car_id	Unique identifier for each car sold.
Date	Date of sale transaction.
Customer Name	Name of the customer.
Gender	Gender of the buyer.
Annual Income	Customer's annual income.
Dealer_Name	Name of the dealership where the sale occurred.
Company	The brand/manufacturer of the car (e.g., Ford, Toyota).
Model	Specific car models sold.
Engine	Engine type or configuration (e.g., V6, Electric).
Transmission	Gear system: Manual or Automatic.
Color	Exterior color of the car.
Price ($)	Selling price of the car.
Dealer_No	Unique identifier for each dealer.
Body Style	Style/category of car (e.g., SUV, Sedan, Hatchback).
Phone	Dealer’s contact number (excluded from visualizations).
Dealer_Region	Geographical sales region.

Visualizations

Dashboard 1

Charts

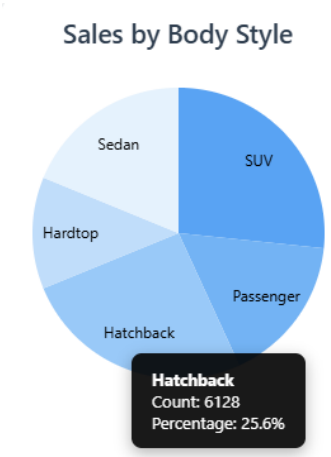
Sales Over Time



The line chart features an interactive visualization to show car sales trends by month. Users can hover over data points to view detailed tooltips with exact sales figures and revenue for the particular month. The purpose of this visualization is to help stakeholders monitor overall sales performance, detect seasonal trends, and assess the effectiveness of sales strategies or marketing efforts. Based on the chart, sales exhibit noticeable spikes and drops at several points throughout the timeline. In August to October, there is a sharp increase in sales followed by a noticeable decline. This may indicate a seasonal peak, possibly due to promotional events or year-end sales, followed by a natural drop-off. In January 2023, sales

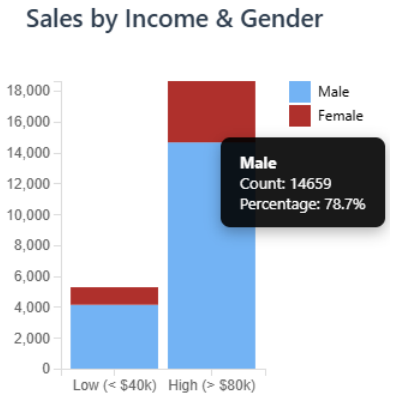
experienced a significant dip from 1625 to 475, which could suggest issues such as reduced consumer demand, post-holiday slowdowns or price adjustments in the new year. By October 2023, sales rise sharply again, showing another strong spike.

Sales by Body Style



This pie chart visualizes car sales distribution by body styles, allowing users to quickly grasp the relative popularity of different car categories. This allows for quick, high-level insights, making it especially useful for marketing teams aiming to tailor campaigns to customer preferences, as well as for inventory managers who need to ensure they are stocking the most popular vehicle body styles. Hovering it also shows sales figures and its sales percentage, with SUV showing the highest sales (26.7%) and HardTop showing the least sales (12.4%). To clear slow-moving inventory or balance stock, promoting least-selling body styles can help free up space in inventory, minimize losses from unsold vehicles, and attract niche buyers who may be more price-sensitive.

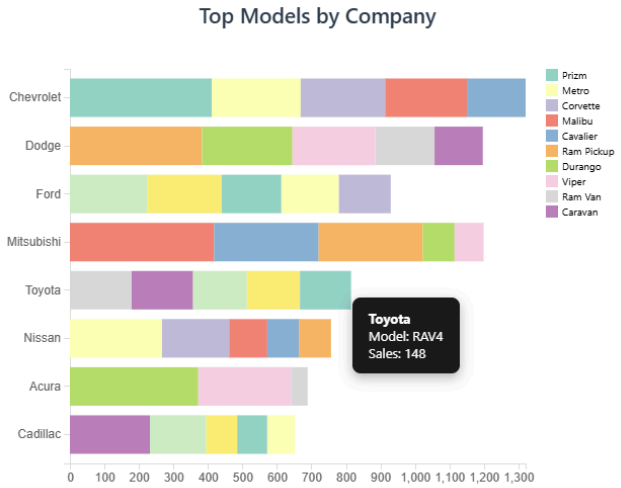
Sales by Income and Gender



The stacked bar chart shows car sales by income group and gender, highlighting how different demographics contribute to overall sales, which can show the purchasing behaviours between them. The high-income group contributes significantly more to overall car sales compared to the low-income group. Male customers make up the majority of car sales in both income groups, and high-income males are the most active buyers. Marketing

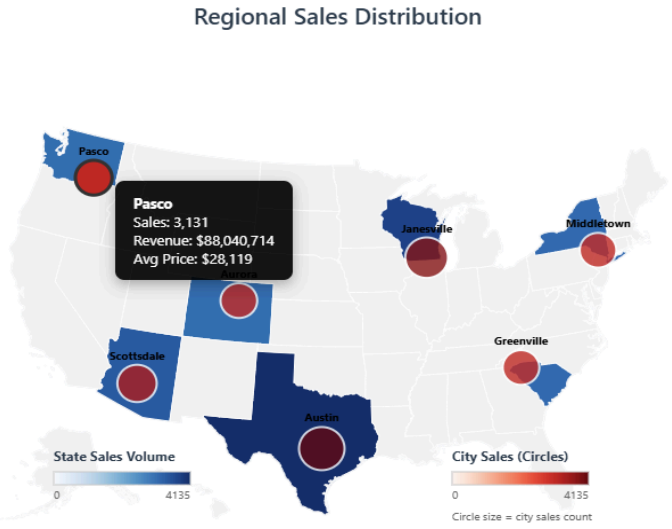
efforts should be strategically focused on the high-income group, as they contribute the most to overall car sales. Promotional strategies should also be tailored to appeal to male interests and buying behavior.

Top Models by company



The horizontal stack bar graph compares the sales performance of various car models across different manufacturers. This highlights which companies have the strongest-performing models in their lineup. Chevrolet has the longest bars, indicating they are the highest-selling models and Cadillac the lowest. Marketers should allocate more advertising budget to high-performing brands for brand reinforcement and special promotional campaigns for lower-selling brands to improve visibility and customer interest.

Regional Sales Distribution



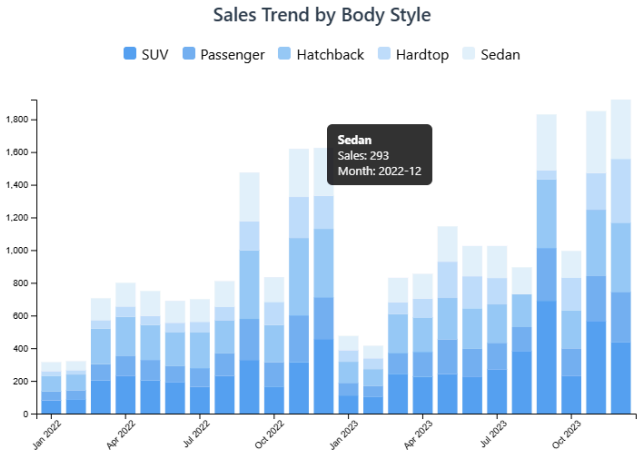
This geospatial visualization displays regional car sales breakdown across different dealership locations. This map color-code regions based on total sales volume, with circle size representing the city sales count. This allows users to spatially identify performance trends and regional demand patterns. A darker and larger red color on the region Austin

shows the highest sales volume and the blue color coded region represents the state it belongs to.

Dashboard 2

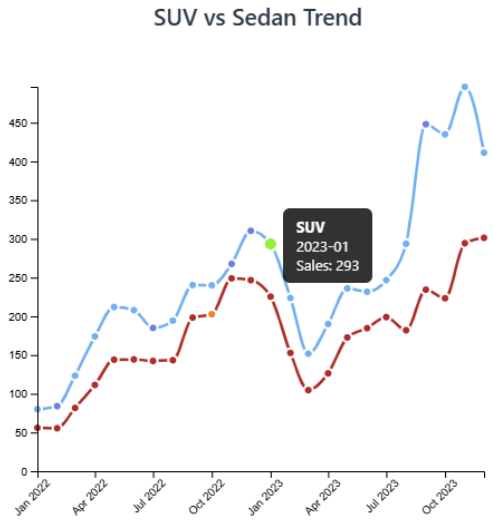
Charts

Sales Trend by Body Style



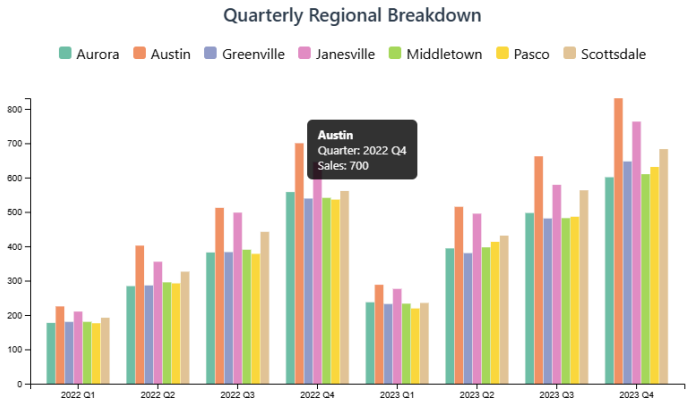
This stacked bar chart illustrates the monthly sales performance of different car body styles from 2022 to 2023. Each bar represents the total number of vehicles sold per month, segmented by body style (e.g., SUV, Sedan, Hatchback, etc.), with each style represented by a distinct color. This visualization helps identify shifting consumer preferences and seasonal buying behaviors. The chart reveals seasonal peaks, such as a surge in total sales around August, Q4 2022 and Q4 2023, possibly linked to marketing campaigns or end-of-year promotions. The stacking technique makes it easy to assess both overall volume and the relative popularity of each body style over time. This aids stakeholders in demand forecasting and refining inventory and marketing strategies based on evolving trends.

SUV vs SEDAN Trend



This dual-line chart displays the monthly sales trends of SUVs and Sedans from January 2022 to October 2023. Each line represents the smoothed sales volume for a respective body style, highlighting how consumer demand evolved over time. The chart uses distinct colors like blue for SUVs and orange for Sedans to clearly differentiate the two categories. The visualization reveals a consistent upward trend for both vehicle types, with SUVs maintaining a dominant lead in sales throughout the period. This pattern may indicate cyclical demand changes, possibly influenced by economic conditions, product launches, or seasonal factors. This comparison enables stakeholders to assess the relative popularity of each body style, monitor growth momentum, and adjust production or promotional strategies accordingly.

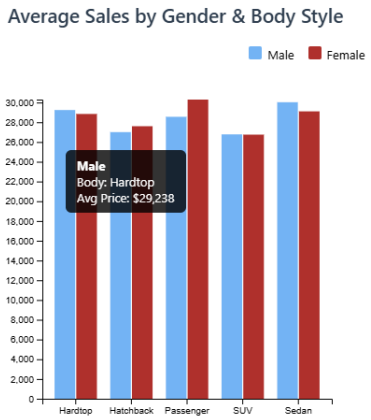
Quarterly Regional Breakdown



This graph visualizes car sales distribution across different regions on a quarterly basis. Each group of bars represents a specific quarter and within each group, individual bars denote the sales volume for each region. The chart effectively reveals trends and seasonal patterns by region. This graph shows how sales impact over time and identify regions that are growing or declining. We notice that the height of each region's bar steadily increases from Q1 to Q4 in 2022 and 2023. It may indicate strong performance or improved marketing in that area. The use of a grouped bar format ensures that regional comparisons are clear

within each quarter and over the full period. This makes it a valuable tool for strategic planning and resource allocation based on regional demand trends. Based on this graph, we found another key insight that Austin and Janesville regions outperform others in car sales across all quarters. Their bars are noticeably taller in each grouped quarter, indicating higher sales volumes. This suggests that these regions have either a larger customer base, more active dealerships, or stronger marketing and sales strategies compared to other regions.

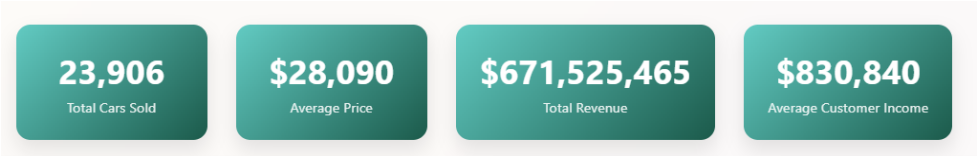
Average Sales by Gender & Body Style



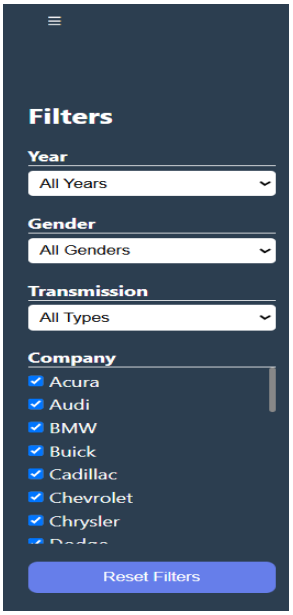
The graph compares average car sales by gender across different body styles. Overall, both male and female customers show comparable purchasing behavior, with slight variations by category. Passenger cars have the highest average price sales for female buyers indicating stronger interest or higher spending in this segment. Sedans also have the highest average price sales for male buyers.These insights highlight opportunities for targeted marketing—especially for Passenger cars towards female buyers and Sedan toward males.

Interactivity

The KPI (Key Performance Indicator) cards in both dashboards serve as vital interactive components that provide a real-time summary of critical business metrics, including Total Cars Sold, Average Price, Total Revenue, and Average Customer Income. These values dynamically update based on the filters applied such as year, gender, transmission type, and company, allowing users to instantly assess performance within specific segments. This interactivity enhances decision-making by offering clear, immediate insights into how different variables impact overall sales and customer demographics. The KPI cards are essential for highlighting high-level trends and facilitating quick evaluations, making them a valuable feature for stakeholders seeking to monitor business performance and identify opportunities for strategic growth.



In both dashboards, dropdowns for Year, Gender, and Transmission Type allow interactivity by enabling users to filter and analyze data based on specific criteria, making it easier to identify trends and patterns within selected segments. Users can also click on company checkboxes to filter based on company-specific performance reviews. Clicking Reset Filters instantly clears all selections and restores the dashboard to its original default view, allowing users to start a new analysis with a full, unfiltered dataset. The hide button on top of filters allows users to hide the filters from view.



Best data visualization practices

The Car Sales Analysis Dashboard follows Shaffer's 4Cs of Data Visualization Design Principles: Clear, Clean, Concise, and Captivating.

Clear

- Each visualization (line chart, bar charts, pie chart, map, etc.) is clearly labeled with titles like "Sales Over Time", "Sales by Body Style", and "Top Models by Company".
- Metrics at the top (e.g., Total Cars Sold, Average Price, Total Revenue) are highlighted with large, colored boxes for quick comprehension.
- Filters on the left (Year, Gender, Transmission, Company) allow users to refine data clearly.

Clean

- Uniform fonts, spacing, and color schemes contribute to a tidy, professional look.
- Consistent use of color coding across charts aids recognition without visual noise.
- Decorative elements are intentionally kept minimal so that the viewer's attention naturally goes to the important part: the data and what it's showing

Concise

- Dropdown filters are minimalistic and organized without clutter.

- No unnecessary elements (like grid lines, 3D effects, or harsh shadows) clutter the visuals.
- Adequate white space between charts, filters, and cards makes the layout breathable and not visually overwhelming.

Captivating

- It has a visually engaging layout, the dashboard combines multiple chart types (line, pie, bar, map) to provide variety and maintain user interest.
- Filters for Year, Gender, Transmission, and Company invite users to interact with the data, making the experience feel dynamic rather than static.
- Bright, attention-grabbing Metrics metrics at the top (Total Cars Sold, Average Price, Total Revenue, etc.) are placed in prominent cards that immediately draw the viewer's eye.
- The placement of visuals tells a cohesive story: it answers what is happening (ex: sales trends), who is buying (ex: income & gender), where it’s happening (ex: map), and what is popular (ex: top models).
- The use of well-balanced, harmonious colors adds to the aesthetic that attracts the user.

Challenges & Solutions

One of the key challenges encountered during the development of the dashboard was related to managing multiple D3.js charts on the same page, particularly when implementing animations and transitions. Initially, when an animation was applied to one chart, such as the line or bar chart, it unintentionally caused other charts or their legends to disappear or break. This issue stemmed from the way D3 handled element cleanup: our initial approach used d3.select(container).selectAll("*").remove(), which indiscriminately removed all DOM elements within the container, including shared or overlapping components.

This led to a major interference issue, especially in a dashboard where multiple charts were rendered and updated based on user input. In some cases, even unrelated charts lost labels or legends when another chart was re-rendered.

To resolve this, we refactored each chart's update function to specifically target and remove only its own SVG elements (e.g., d3.select(container).select("svg").remove()), ensuring that animations and updates were isolated to the individual chart. This approach preserved the integrity of all other components on the dashboard, allowing each visualization to animate and update smoothly without cross-interference.

Additionally, we faced minor challenges with filter synchronization and scrollable checkboxes for long lists such as car companies and regions. These were addressed by implementing overflow-y: auto with fixed height containers, improving both usability and layout stability. Ensuring dynamic KPI metrics reflected filtered data in real-time also required precise coordination between DOM updates and D3 data-binding. Through modular coding and iterative testing, we managed to overcome these technical hurdles and maintain a seamless user experience across all dashboard elements.

Development processes

The development of our car sales analysis dashboards was collaboratively contributed by four team members. There are two main dashboards and a total of nine interactive charts, all built using HTML, CSS, and D3.js. The work was divided based on dashboard ownership and each member's area of expertise to ensure efficiency and consistency throughout the project. The project required approximately 50 man-hours.

Name	**Work**
Lee Le Xuan	Line Chart, Stacked Bar Chart, Map, Interactivity and Integration
Poh Ern Qi	Pie Chart, Stacked Bar Chart and Color coding
Choo Chee Choong	Stacked Bar Chart, Line Chart and Scaling
Yap Weng Hong	Bar Charts, Interactivity and Integration

Dashboard 1 was handled by Lee Le Xuan and Poh Ern Qi, with a focus on high-level visual storytelling and regional insights. This dashboard includes five charts: Sales Over Time, Sales by Body Style, Sales by Income & Gender, Regional Sales Distribution, and Top Models by Company. Dashboard 2 was developed by Choo Chee Choong and Yap Weng Hong, emphasizing a filter-driven interface with scrollable checkboxes and dropdowns. This section contains four charts: Sales Trend by Body Style, SUV vs Sedan Trend, Quarterly Regional Breakdown, and Average Sales by Gender & Body Style.

The most time-intensive tasks across both dashboards were creating D3-based chart logic and ensuring the visualizations responded accurately to user filters. Particular attention was required for monthly and quarterly grouping, smoothing of line trends, and stacked bar compositions. Additional effort went into enhancing user experience by making long checkbox lists scrollable, refining tooltip behavior, and polishing the visual layout for both desktops and smaller screens. Overall, the clear division of responsibility allowed the team to work in parallel and deliver a visually appealing, functional, and fully interactive analytics dashboard.