VISUALIZING SEATTLE TRAFFIC ACCIDENT HOTSPOTS AND PATTERNS

1. Executive Summary

This data visualization project focuses on analyzing traffic accidents in Seattle to better understand the patterns and factors contributing to their frequency. Using the SDOT collision dataset[1], which includes data from 2015 to 2024, we explored how conditions like weather, time of day, road conditions, and whether it's a weekday or weekend affect the occurrence of collisions. The aim is to identify accident hotspots and provide insights that could help improve road safety in the city.

The motivation for this study comes from the need to address urban traffic accidents with evidence-based strategies. By analyzing this data, we hope to help city planners and officials take preventive actions like adjusting speed limits, installing warning signs, or improving street lighting in areas with frequent collisions. The focus is on turning raw data into useful information that can ultimately save lives.

The project's main outcome is an interactive Tableau dashboard. This includes a geospatial heatmap of Seattle to highlight accident hotspots, along with bar charts and time series visualizations that showcase how different factors influence collision rates. The dashboard allows users to explore trends and patterns more intuitively, making it a practical tool for understanding and addressing traffic safety issues.

The visualization provides valuable insights into Seattle's traffic accidents, offering city planners and officials a foundation for making informed decisions. By visualizing the data in an accessible way, we hope this project will contribute to ongoing efforts to create safer streets and reduce the number of accidents in the city.

View our dashboard here: Visualizing Seattle Traffic Accident Hotspots and Patterns.

2. Concept Background

2.1 Ideation:

The ideation process for this project began with identifying the primary goal: to explore traffic accident hotspots in Seattle and provide actionable insights to improve road safety. The challenge was to design a project that not only highlighted accident hotspots but also uncovered the underlying factors contributing to these incidents. To address this, we aimed to develop a clear and actionable framework for analyzing critical aspects such as weather conditions, light conditions, day of the week, and month of the year.
One of the key decisions during ideation was determining the timeframe for analysis. While the SDOT collision dataset[1] offers data from 2003 to 2024, We opted to focus on data from the past decade (2015–2024) to ensure our findings reflect current trends and patterns. This timeframe captures the effects of recent infrastructure changes, population growth, and evolving traffic behaviors, providing a comprehensive foundation for effective urban planning and safety interventions. Using a decade's worth of data allows for more accurate identification of long-term trends and patterns compared to older or shorter timeframes.

2.2 Data Sourcing:

The data for this project was sourced from the Seattle Department of Transportation (SDOT) collision dataset[1], a comprehensive repository that documents traffic accidents occurring within the city. This dataset spans from 2003 to 2024, providing detailed records of collisions, including variables such as location, severity, type of collision, weather and road conditions, light conditions, and whether speeding was involved. It serves as a rich source of information for understanding the various factors contributing to accidents in Seattle.

For the purposes of this project, we focused on the data from 2015 to 2024, narrowing down the scope to reflect more recent trends and patterns over the past decade. The dataset consists of 51 variables, offering a diverse range of attributes that are critical for the analysis. By filtering and cleaning the data, we selected the most relevant fields, such as Inckey, location, Incdate, Incdttm, weather, Roadcond, lightcond, X, Y to align with the project's objectives. This iterative approach allowed us to refine the dataset to ensure it effectively supports the exploration of accident hotspots and the identification of patterns influencing collision frequency. For the purpose of visualization, we have selected the following attributes:

1. **Inckey**: A unique identifier for each collision incident, ensuring data integrity and facilitating easy referencing.
2. **Location**: Describes the general area where the collision occurred, helping to identify spatial patterns in incidents.
3. **Incdate**: Records the date of the incident, allowing for temporal analysis of collision trends over time.
4. **Incdttm**: Combines the date and time of the collision, enabling detailed analysis of patterns based on specific hours.
5. **Weather**: Captures the weather conditions during the collision, providing insight into the environmental factors influencing accidents.
6. **Roadcond**: Reflects the condition of the road surface (e.g., dry, wet, icy), a critical factor in understanding collision causes.
7. **Lightcond**: Specifies the lighting conditions (e.g., daylight, darkness, street lights), revealing the role of visibility in traffic incidents.
8. **X and Y**: Represent coordinates using the Washington State Plane Coordinate System (SPCS)[2], a standard geographic reference system. SPCS allows precise mapping of incidents within the state's boundary, making it easier to visualize and analyze spatial patterns.

2.3 Data Profiling and Cleaning:

STEP 1: Filtering the Dataset by Timeframe

The first step in our data profiling and cleaning process was filtering the dataset to match the required timeframe for analysis, which is 2015 to 2024. To achieve this, we utilized the attribute Incdate, which records the date of each incident. In Tableau, we converted the Incdate attribute to a year format, making it easier to isolate and filter the data to include only the incidents that occurred within the chosen decade. This transformation streamlined the process of preparing the dataset for further analysis by ensuring we worked exclusively with the relevant timeframe, enhancing the focus and accuracy of our study.

STEP 2: Handling Missing Time Data

The second step in our data profiling and cleaning process involved handling the Incdttm attribute, which contains the date and time of each incident. We noticed that some entries in this column contained only date information without a time value. Since the time is recorded in the hh:mm:ss format, we identified that entries lacking time information do not have a colon (:). Using Python, we labeled these entries as NaN to signify missing time data. Once this was done, we removed all NaN values in the Incdttm column and other columns, ensuring that the dataset was free from null values. Finally, we converted the Incdttm attribute from a string format to a proper date-time format, preparing it for accurate temporal analysis.

STEP 3: Converting X and Y Coordinates to Latitude and Longitude

Next, we focused on the X and Y attributes, which represent the coordinates of each incident in the Washington State Plane Coordinate System (SPCS)[2]. Using Python, we converted these coordinates into latitude and longitude. This transformation was particularly important for creating a heatmap of Seattle, as it allowed us to accurately plot the spatial distribution of incidents on a map. By converting the coordinates into a universally recognized format, we ensured that the data was compatible with geographic visualization tools, enhancing our ability to analyze collision patterns across the city.

To transform the X and Y attributes into proper geographical latitude and longitude data for our heatmap, we reached out to SDOT for clarification about the coordinate system used. They informed us that all the city's geospatial data is based on the NAD83 HARN reference datum, projected to the State Plane Coordinate System, Washington North Zone[2] (EPSG: 2285).

Using this information, we implemented a solution in Python with the pyproj[3] library to convert the coordinates. We defined the source Coordinate Reference System (CRS) as EPSG: 2285 (State Plane, Washington North Zone)[2] and the target CRS as EPSG: 4326 (WGS84, representing latitude and longitude).

STEP 4: Grouping and Merging Data for Simplification

To prepare the dataset for meaningful analysis, we performed grouping and merging of similar values across several key attributes using Tableau. This process simplified the dataset while preserving critical details and ensuring its usability for deriving insights.

For the road condition attribute, we grouped related values into broader categories. Conditions such as Wet, Ice, Oil, Snow, Slush, Standing Water, Sand, Mud, and Dirt were grouped under "Slippery," while "Dry" was left as its own category. This allowed us to better represent the road conditions' overall impact on traffic safety.

Similarly, we refined the weather conditions attribute to reduce complexity and focus on meaningful patterns. For example, "Partly Cloudy" and "Overcast" were combined into "Cloudy," as they share moderate visibility effects. Wind-related conditions like "Severe Crosswind" and "Blowing Sand/Dirt" were grouped under "Windy," highlighting their common hazards to visibility and stability. All precipitation-based conditions, including "Raining" and "Sleet/Hail/Freezing Rain," were consolidated into "Rain," as they share challenges like wet surfaces and reduced traction. Additionally, "Snowing," "Fog/Smog/Smoke," and "Blowing Snow" were grouped under "Snow" due to their shared impact on visibility and road conditions. This logical categorization ensured that the analysis remained focused on the most significant weather factors influencing collisions.

We transformed the date attribute into a day-of-the-week format in Tableau and then grouped the days into two categories: 'Weekend' for Saturday and Sunday, and 'Weekday' for Monday through Friday. This allowed us to explore how collision patterns varied based on the type of day.

In the light condition attribute, we categorized the data into logical and intuitive groups: "Dawn," "Daylight," "Dusk," "Dark (Street Lights On)," and "Dark (Street Lights Off/No Street Lights)." These categories reflected the varying degrees of visibility and lighting infrastructure available during incidents, making it easier to analyze the role of lighting conditions in collision frequency and severity.

By grouping and merging these values, we streamlined the dataset, making it more manageable and conducive to uncovering meaningful patterns while ensuring that critical details were preserved. This process was a crucial step in preparing the data for effective visualization and analysis.

STEP 5: Removing "Unknown" or "Other" Entries

After grouping and merging the values in the dataset to create more meaningful categories, we proceeded to clean the data further by removing all entries labeled as "Unknown" or "Other." These values lacked specific information and added noise to the dataset, potentially skewing the analysis. By eliminating these ambiguous entries, we ensured that the dataset remained reliable and focused only on well-defined, actionable data. This step was essential in maintaining the integrity of the analysis and enabling more accurate insights.

3. Process Description

3.1 Highlights of Usability test results:

Usability testing was a critical step in evaluating the functionality and intuitiveness of our visualization. To achieve this, we developed a low-fidelity functional prototype featuring a heatmap and accompanying graphs. These visual elements were complemented by interactive sliders that allowed users to adjust variables such as “time of day” and “weather conditions.” The primary goal of this testing phase was to determine whether users could easily navigate the interface and extract meaningful insights from the data. Users were tasked to explore accident patterns and hotspots in Seattle, addressing specific questions such as:

1. What are the accident patterns in Seattle for a particular year?
2. How do “time of day” and “weather conditions” affect the number of accidents?
3. Under what conditions do most accidents occur?

The testing was conducted with both classmates and target users outside the classroom. Participants were closely observed as they interacted with the prototype, and their thought processes were simulated through a cognitive walkthrough to identify areas where the design might fall short of user expectations. One of the most appreciated features was the slider functionality, which was intuitive and allowed users to explore accident hotspots under specific conditions with ease. Users particularly liked how the sliders worked together, revealing nuanced patterns by combining variables such as weather and time of day. This feature significantly enhanced the user experience and validated its inclusion in the design.

Based on user observations and feedback, several refinements were made to the prototype. One major adjustment involved the layout of the visualizations. Initially, the heatmap, bar charts and time series visualizations were presented together, but users felt this arrangement made it difficult to focus on individual stories each visualization was telling. To address this, the layout was redesigned to separate the heatmap and pattern visualizations (bar graphs and time series) into distinct dashboards. This change not only improved clarity but also reduced cognitive load, allowing users to focus on the insights presented by each visualization without unnecessary overlap. Additionally, users suggested updating the headings for greater clarity. For example, the heading “Patterns” was considered vague and was replaced with “Seattle Accident Hotspots and Patterns” to provide a more accurate description of the content. These changes made the dashboard more user-friendly and informative.

The entire process of usability testing underscored the importance of user-centered design. Observing how participants interacted with the prototype and incorporating their feedback was invaluable in refining the visualization. By addressing their concerns and suggestions, we were able to develop a dashboard that not only conveyed accident patterns effectively but also provided an engaging and intuitive user experience. Ultimately, this iterative approach to design ensured that the final product met the needs of its users while maintaining a focus on clarity, interactivity, and accessibility.

3.2 Challenges encountered and modifications to the plan:

The first challenge we faced was the absence of latitude and longitude data in the dataset, which is essential for overlaying a heat map on a geographical map in Tableau. Initially, we attempted to derive coordinates based on location descriptions, but this proved unfeasible as the dataset provided ambiguous addresses like "between street A and street B" without area codes. To resolve this, we contacted the SDOT authorities, who clarified that the dataset's X and Y columns represented coordinates in the Washington State Plane Coordinate System[2]. Using this information, we converted the X and Y values into latitudes and longitudes.

Another challenge arose in the patterns dashboard, where we plotted the number of incidents against the day type (weekday or weekend). The results initially showed significantly higher incidents on weekdays compared to weekends, raising concerns about accuracy. Upon reflection, we realized the imbalance was due to weekdays constituting five days and weekends only two. To address this, we normalized the data by dividing weekday incident counts by 5 and weekend counts by 2. Similarly, Seattle's predominant rainy weather required normalization for meaningful comparisons. We adjusted the data based on annual averages for weather conditions: 75 clear days, 210 cloudy days, 155 rainy days, and 7 snowy days, ensuring accurate representation across conditions.

To enhance the heat map's usability, we added road conditions as a filter, enabling identification of accident-prone locations with slippery roads. This insight is valuable for better urban planning and improving travel safety. Additionally in the patterns dashboard, we incorporated two time series graphs—one plotting the number of accidents over the years (2015–2024) and another showing incidents across months of the year. These visualizations revealed long-term trends and seasonal patterns, providing deeper insights for analysis and decision-making.

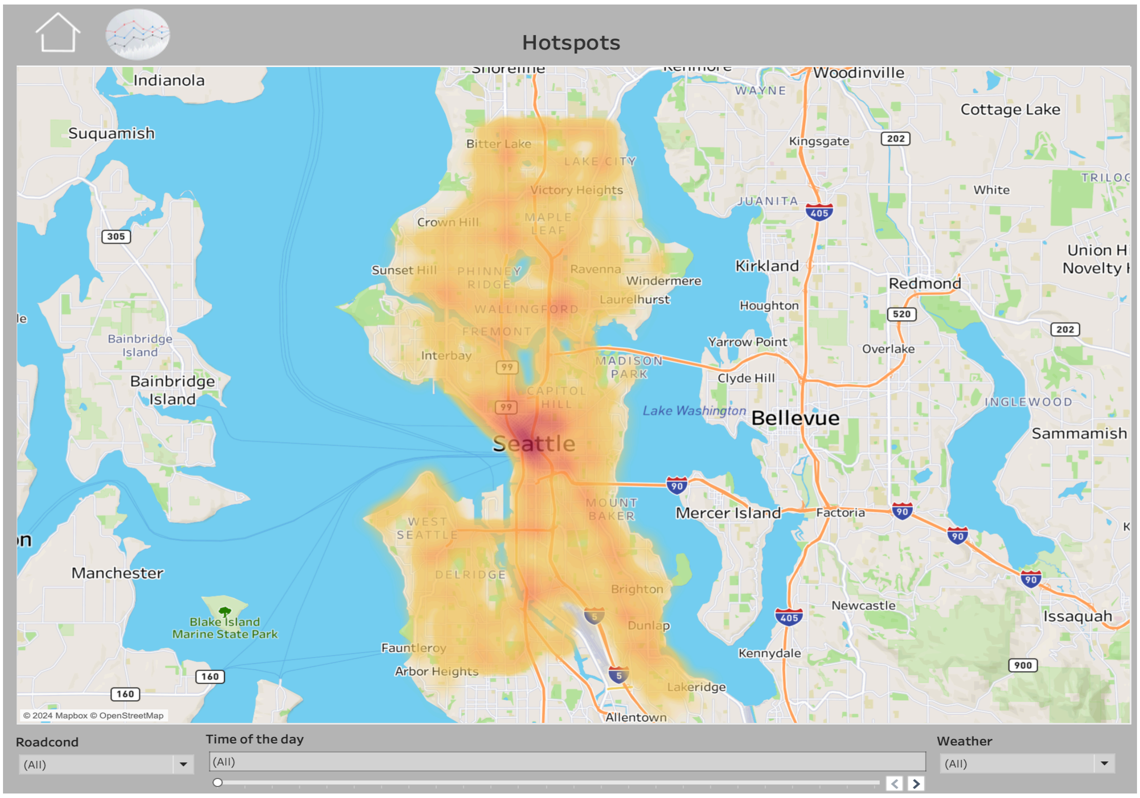
During the class demo presentation, users provided valuable feedback regarding the "Patterns" page. They noted that the close proximity of the five plots caused them to blend together, reducing clarity. To address this, we adjusted the background color to visually separate the elements, improving readability and helping users interpret the data more effectively. Users also suggested incorporating better color variability to enhance the visual hierarchy and guide attention. Based on this feedback, we revised the color palette[5], introducing contrasting colors to highlight key data points while maintaining a cohesive design across the dashboard.

In conclusion, addressing these challenges and incorporating user feedback significantly enhanced the dashboard's accuracy, usability, and visual appeal. These modifications not only improved the presentation quality but also ensured the insights derived were meaningful and actionable.

3.3 Final visualization:

In the final visualization, we integrated comprehensive feedback from development, testing, and in-class prototype presentations to create a user-friendly and insightful dashboard. Our design comprises three main components: a landing page, a "Hotspots" dashboard, and a "Patterns" dashboard, each tailored to provide distinct perspectives on accident data.

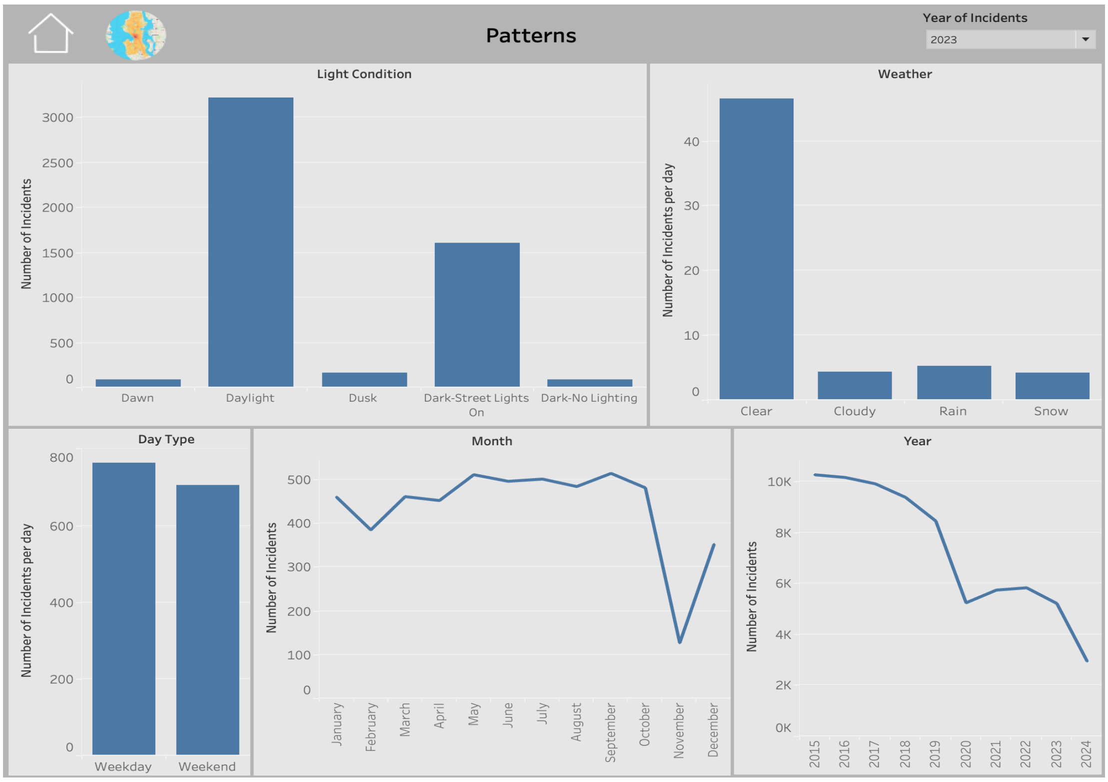
The landing page serves as an intuitive entry point, offering users the option to explore accident data geographically or analyze patterns. The Hotspots dashboard leverages a dynamic heatmap, overlaid on a geospatial map of Seattle, to visualize accident-prone areas. Users can interact with the filters to adjust the display based on "time of day", "road conditions" and "weather conditions," tailoring the heatmap to their specific interests. Additional enhancements, such as zoom functionality and detailed tooltips displaying street-level data, allow users to pinpoint exact accident locations with precision. To complement the heatmap, we included a refined layout and contrasting colors, addressing earlier feedback about overlapping elements and improving clarity.



Hotspots Dashboard

The graphs in the Patterns dashboard provide insights into incident patterns across various factors. The Light Condition Bar Graph depicts incidents under different light conditions. The x-axis lists the conditions—Dawn, Daylight, Dusk, Dark-Street Lights On, and Dark-No Lighting—while the y-axis shows the number of incidents recorded under each condition. Daylight exhibits the highest number of incidents, likely due to increased activity and traffic during daytime hours. The Weather Condition Bar Graph focuses on incidents per day during various weather conditions, with the x-axis showing Clear, Cloudy, Rain, and Snow, and the y-axis indicating the normalized number of incidents per day. Clear weather shows the highest number of incidents per day, which may be attributed to increased travel and higher speeds during favorable weather. The Day Type Bar Graph compares incidents per day by day type, distinguishing between Weekdays and Weekends on the x-axis, with the y-axis representing the normalized number of incidents per day.

The Month Time Series focuses on incidents by month, with the x-axis showing the months and the y-axis indicating total incidents. The Year Time Series examines incidents by year, with the x-axis spanning 2015 to 2024 and the y-axis showing total incidents per year. Additionally, there is a filter available to refine the graphs based on the selected year, allowing users to focus on specific time periods for more detailed analysis.



Patterns Dashboard

Across all dashboards, we included interactive features, such as synchronized charts, tooltips, and drill-down capabilities, allowing users to navigate seamlessly through the data and derive meaningful insights. By refining the aesthetics and usability based on continuous feedback, the final visualization offers a comprehensive, clear, and engaging representation of accident data, aligning with the project's objectives and user expectations.

3.4 Insights and breakthroughs:

The visualizations provided several meaningful insights into accident trends in Seattle, particularly through the heatmap and patterns dashboards.

The heatmap, overlaid on a geospatial map of Seattle, highlights areas with the highest accident intensity. Users can zoom in to street-level detail, with tooltips providing precise information about accident locations. By incorporating filters for road conditions, weather conditions, and time of day, the heatmap enables a detailed exploration of accidents under various scenarios. These insights can inform urban development, road infrastructure improvements, and targeted traffic management strategies in high-risk areas. For instance, identifying accident-prone streets during adverse weather conditions could lead to better drainage systems or additional signage.

The patterns dashboard reveals several insights. From the light conditions bar chart, Daylight has the highest number of incidents, likely due to increased activity and traffic during the day. Incidents under Dark-Street Lights On are also significant, suggesting that while lighting helps, it may not entirely prevent accidents. Conditions such as Dawn, Dusk, and Dark-No Lighting report fewer incidents, possibly because of reduced traffic volumes. In the weather conditions bar chart, normalized using Seattle's weather data[3], Clear weather has the highest number of incidents per day. This could be attributed to increased travel and higher speeds in favorable conditions. Conversely, conditions like Rain, Cloudy, and Snow show fewer incidents per day, likely due to reduced travel or more cautious driving. The day type bar chart shows that Weekdays have slightly more incidents per day compared to Weekends, reflecting higher traffic volumes during workweek and school commutes. From the time series charts, the month-wise data indicates a steady number of incidents from February to September, with a slight peak during the summer months, likely due to increased outdoor activities and travel. There is a sharp drop in November (winter), possibly due to seasonal changes in traffic and reduced outdoor activities. The year-wise chart shows a consistent decline in incidents over the years, potentially due to improved road safety measures. There is a sharp drop in 2020, attributed to reduced traffic during the pandemic.

4. Conclusion and Evaluation

Choosing a topic like this highlights the importance of addressing real-world challenges through data visualization. We selected this topic to explore accident-prone areas, providing insights that can aid in improving traffic management, urban planning, and public safety. This project allowed us to apply classroom concepts to create meaningful visualizations with practical applications. We developed two main dashboards for this project. The first focused on a heatmap, overlaid on a geospatial map of Seattle, showcasing accident density across different regions. Using principles of data representation, we incorporated color gradients and size encoding to effectively convey accident hotspots, making it easier to identify high-risk areas. The second dashboard examined patterns in accident data, allowing users to explore trends based on variables such as light conditions, weather conditions, day type, and month of the year. This provided a deeper understanding of when and where accidents are most likely to occur.

For the dashboards, we applied several key concepts and best practices learned in class. In the first dashboard, we implemented the widely recognized visualization mantra[6], known as the information-seeking methodology conceived by Ben Shneiderman, which advocates 'overview first, zoom and filter, then details on demand.' This approach enabled users to get a broad view of accident data and interactively explore specific areas of interest. By incorporating an interactive heatmap, the dashboard became more user-friendly and effective in visualizing accident hotspots.

We used color gradients and size encoding in the heatmap to represent accident density. Following Mackinlay's ranking for encoding data[7], we prioritized color intensity to convey information about accident frequency, ensuring that the visualization was both intuitive and accurate. This approach helped make the data more accessible and easier to interpret at a glance. We also incorporated a user-centered approach, as explored in the R3 assignment[8], by leveraging feedback from class demos and C3 sessions to refine the design and enhance the user experience. This emphasis on designing for the user ensured that our dashboards were not only functional but also aligned with users' needs and expectations.

Finally, we referred to data visualization tips from Tableau[9] and class lectures to design effective worksheets and dashboards. This included creating interactive filters, selecting appropriate graph types, and organizing layouts and legends to enhance clarity and functionality. These best practices helped ensure that the dashboards were aesthetically pleasing, insightful, exploratory, and easy to navigate, while maintaining a professional and polished appearance.

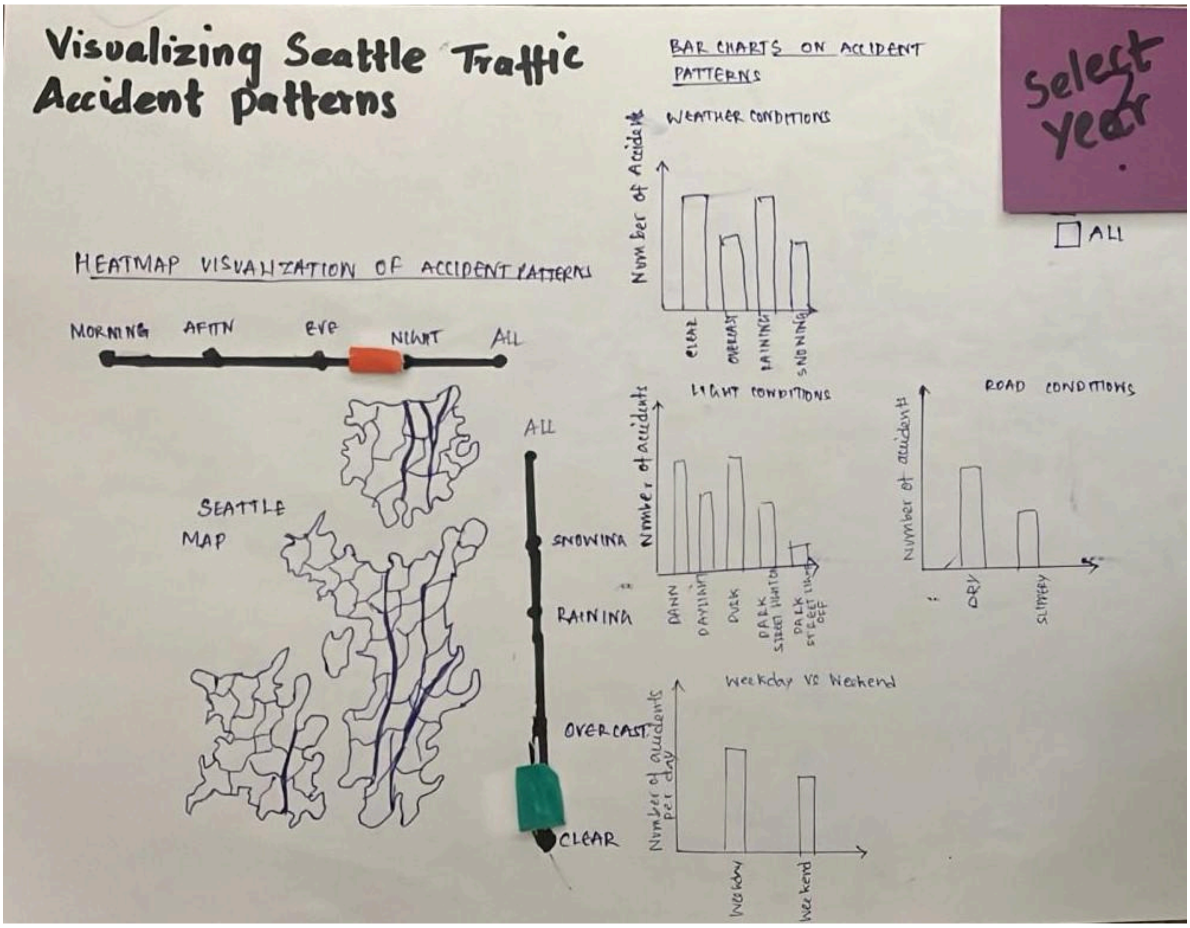
Through this project, we gained experience in using data visualization to derive actionable insights. The dashboards highlight areas and patterns most susceptible to accidents, offering a foundation for authorities to implement targeted interventions. By leveraging these insights, communities can improve traffic management, road infrastructure, and urban planning, ultimately reducing accidents and fostering safer environments.

5. References

1. SDOT Collision Dataset Link
   https://data-seattlecitygis.opendata.arcgis.com/datasets/SeattleCityGIS::sdot-collisions-all-years/about
2. Washington State Plane Coordinate System Link
   https://epsg.io/2285#:~:text=EPSG%3A2285%20Projected%20coordinate%20system%20for%20United%20States%20%28USA%29,Oreille%3B%20San%20Juan%3B%20Skagit%3B%20Snohomish%3B%20Spokane%3B%20Stevens%3B%20Whatcm.
3. Working with Geospatial data
   https://www.geeksforgeeks.org/working-with-geospatial-data-in-python/
4. Seattle Weather Data Link
   https://www.currentresults.com/Weather/Washington/Places/seattle-weather-averages.php
5. What Color to consider for Visualization
   What to consider when choosing colors for data visualization - Datawrapper Academy
6. J. Heer and B. Shneiderman, "Interactive dynamics for visual analysis", Communications of the ACM, vol. 55. (2012)
7. Mickael Causse. “The Mackinlay ranking of perceptual tasks.” researchgate.net
   https://www.researchgate.net/figure/The-Mackinlay-ranking-of-perceptual-task_fig2_221098028
8. R3: Fry (2008). Visualizing Data Chapter1
   https://canvas.uw.edu/courses/1746707/discussion_topics/9157847
9. Data visualization tips:Tableau
   https://www.tableau.com/learn/articles/data-visualization-tips
10. GitHub Repo of all Files Used
    https://github.com/ReddyKhajaValluru/Visualizing-Seattle-Traffic-Accident-Hotspots-And-Patterns

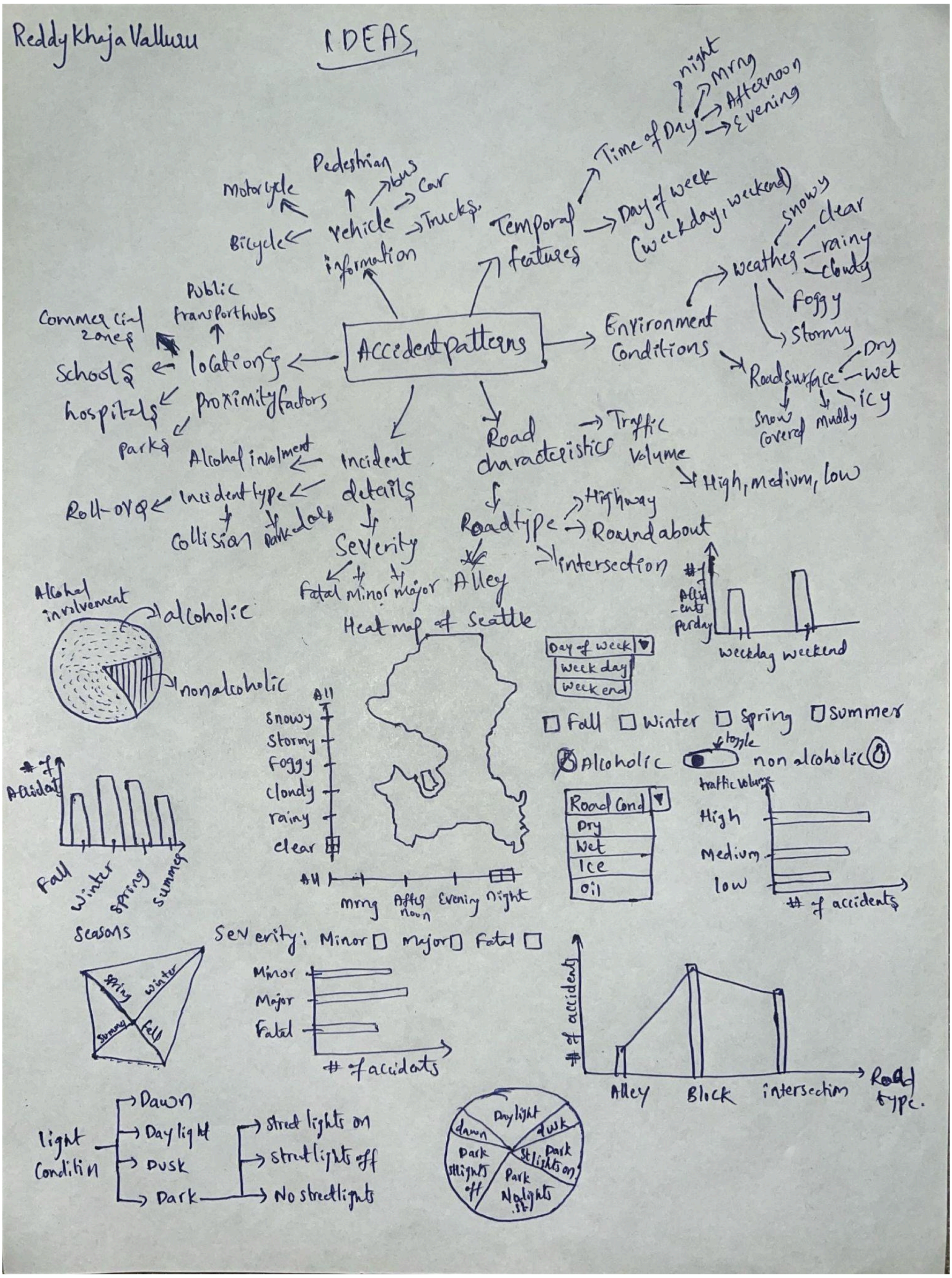
Appendix:

1. Prototype Development (C3) - Design Sketches

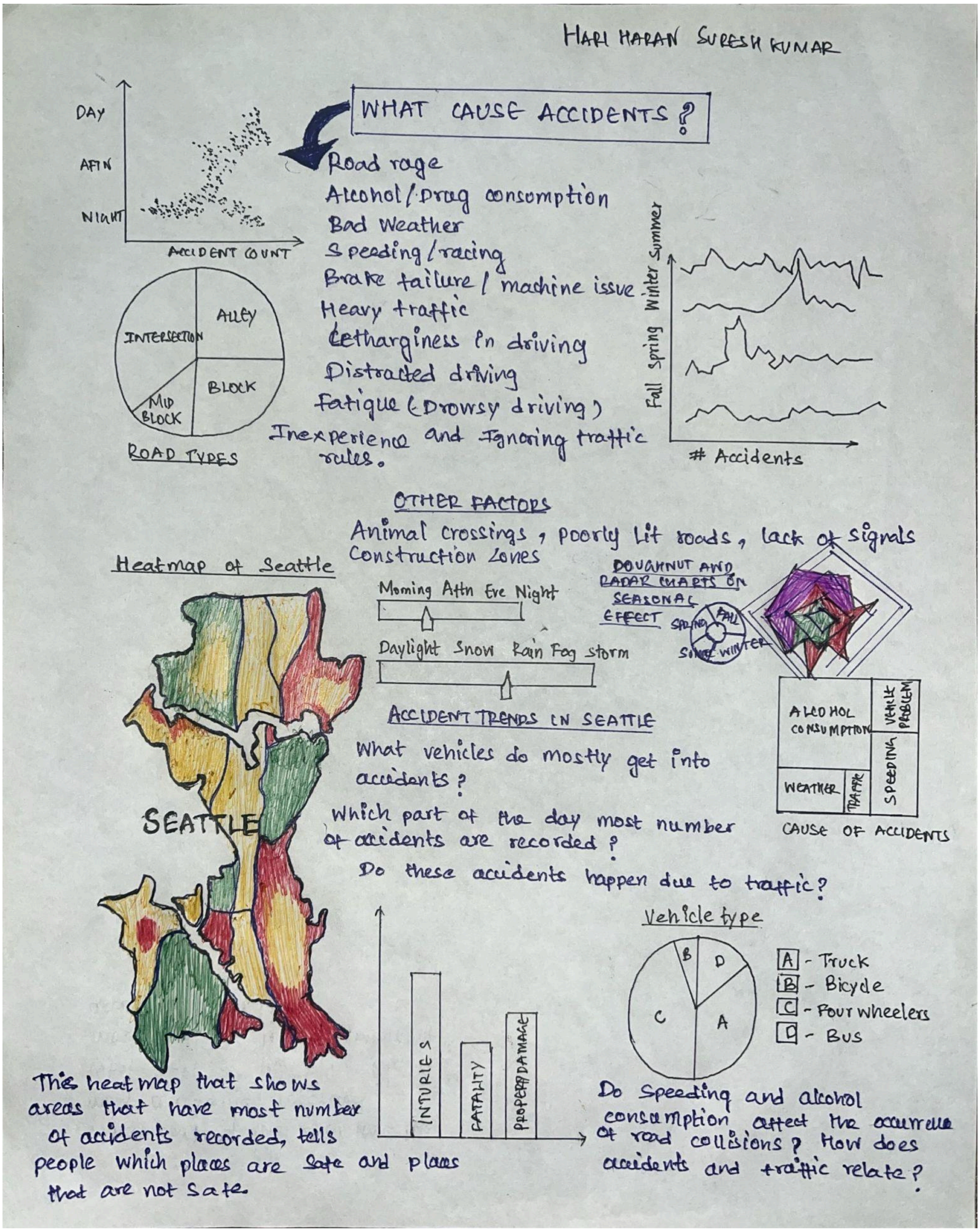


Design sketch

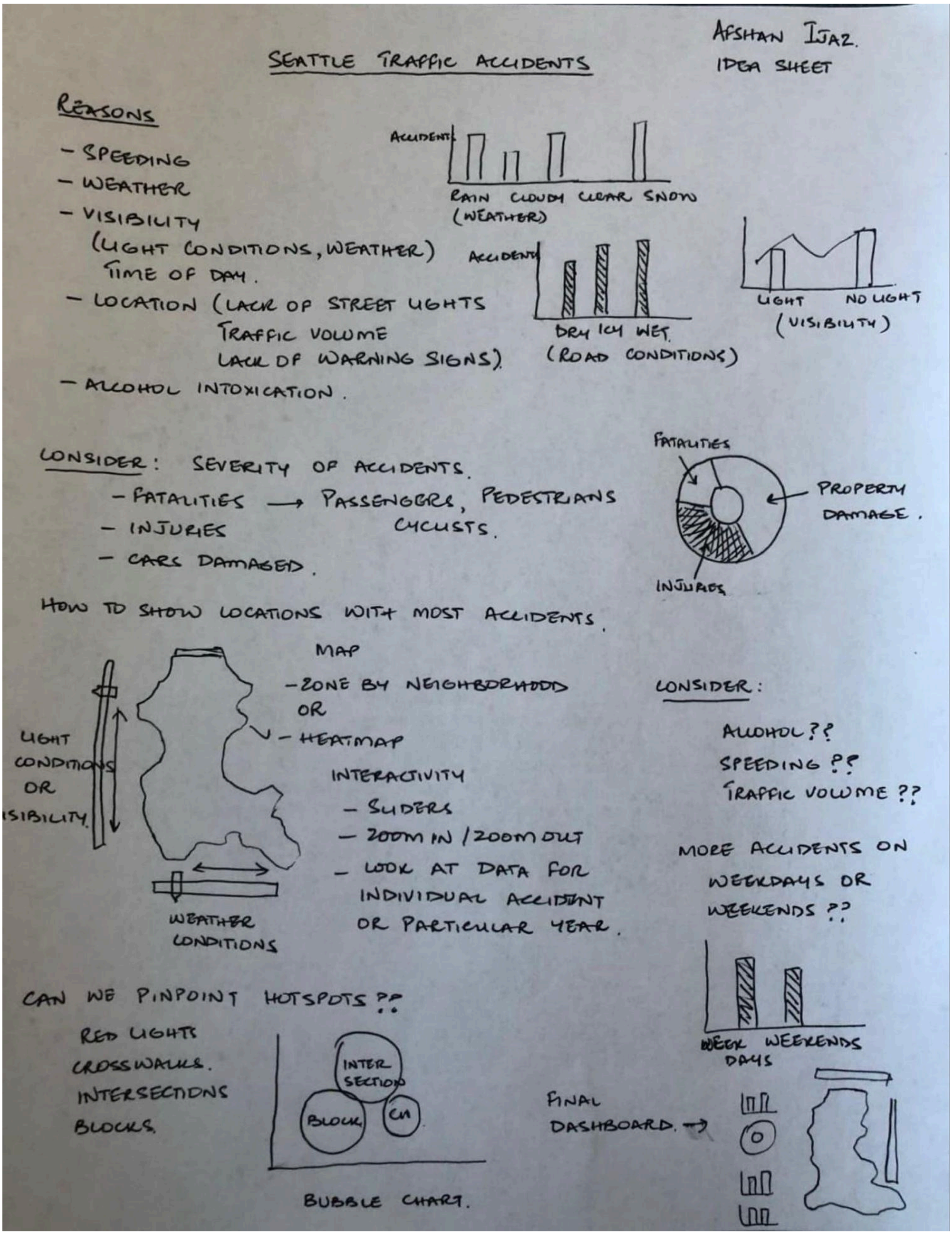
2. Design Ideation (C2) - Five Design Sheets (FDS)



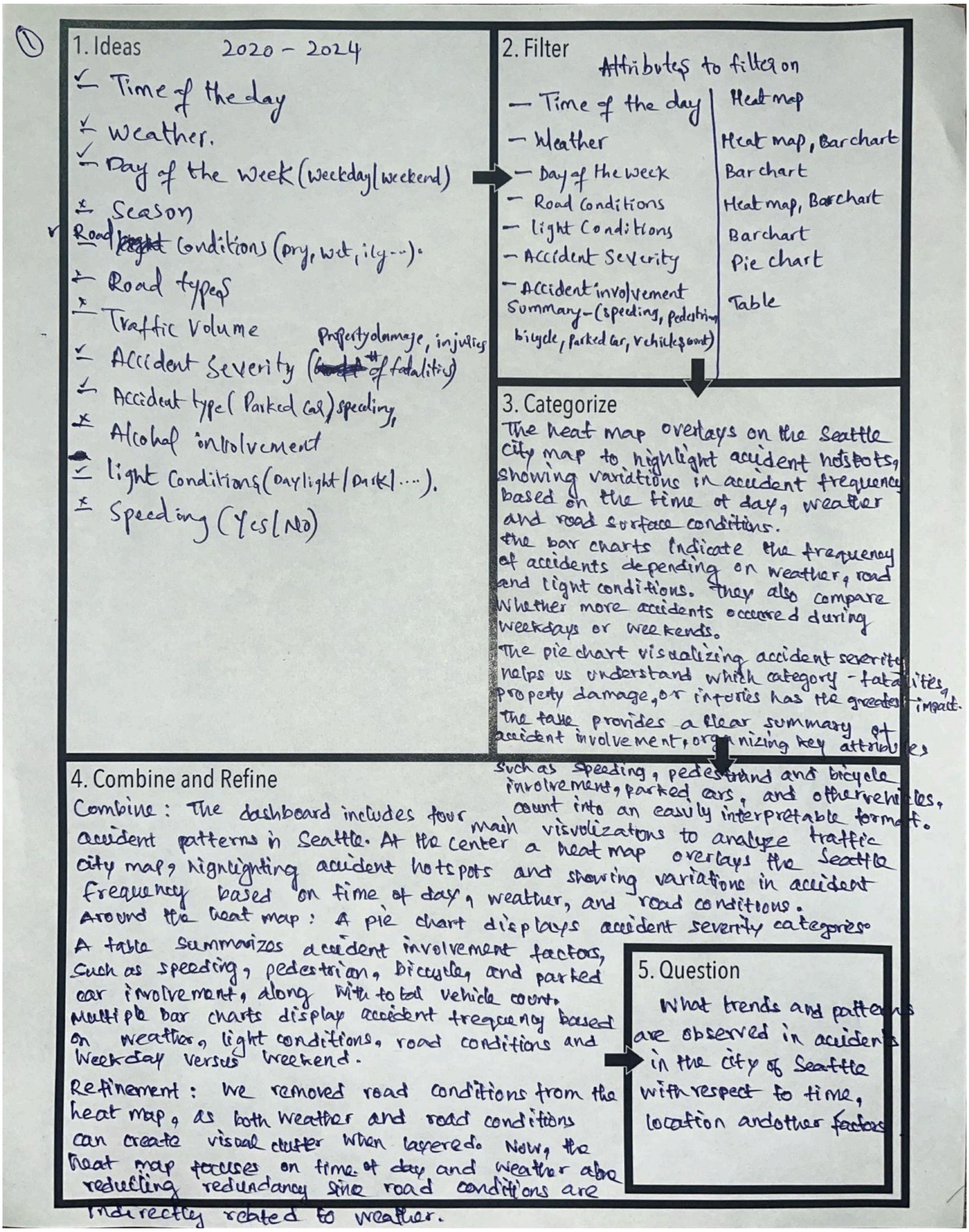
Brainstorming sheet 1



Brainstorming sheet 2



Brainstorming sheet 3



1

1. Ideas 2020 – 2024
- Time of the day
- Weather.
- Day of the week (weekday | weekend)
- Season
- Road ~~light~~ conditions (Dry, wet, icy...).
- Road types
- Traffic Volume
- Accident Severity (~~death~~ # of fatalities) Property damage, injuries
- Accident type (Parked car) speeding,
- Alcohol involvement
- light Conditions (Daylight | Dark | ....).
- Speeding (Yes | No)

2. Filter

Attributes to filter on

— Time of the day	Heat map
— Weather	Heat map, Barchart
— Day of the week	Bar chart
— Road Conditions	Heat map, Barchart
— light Conditions	Barchart
— Accident Severity	Pie chart
— Accident involvement Summary – (speeding, pedestrian bicycle, Parked car, vehicle count)	Table

3. Categorize

The heat map overlays on the Seattle city map to highlight accident hotspots, showing variations in accident frequency based on the time of day, weather and road surface conditions.
The bar charts indicate the frequency of accidents depending on weather, road and light conditions. They also compare whether more accidents occurred during weekdays or weekends.
The pie chart visualizing accident severity helps us understand which category – fatalities, property damage, or injuries has the greatest impact.
The table provides a clear summary of accident involvement, organizing key attributes such as speeding, pedestrian and bicycle involvement, parked cars, and other vehicles count into an easily interpretable format.

4. Combine and Refine

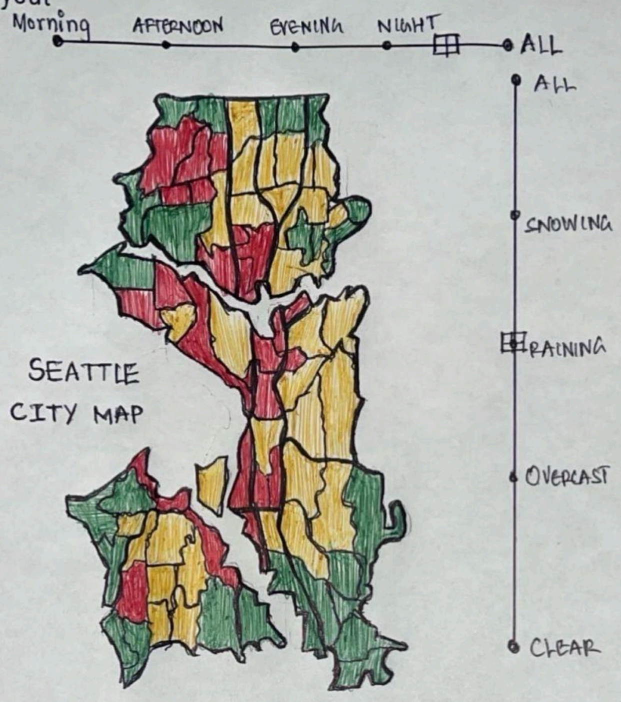
Combine: The dashboard includes four main visualizations to analyze traffic accident patterns in Seattle. At the center a heat map overlays the Seattle city map, highlighting accident hotspots and showing variations in accident frequency based on time of day, weather, and road conditions.
Around the heat map: A pie chart displays accident severity categories.
A table summarizes accident involvement factors, such as speeding, pedestrian, bicycle, and parked car involvement, along with total vehicle count.
Multiple bar charts display accident frequency based on weather, light conditions, road conditions and weekday versus weekend.
Refinement: We removed road conditions from the heat map, as both weather and road conditions can create visual clutter when layered. Now, the heat map focuses on time of day and weather also reducing redundancy since road conditions are indirectly related to weather.

5. Question

What trends and patterns are observed in accidents in the city of Seattle with respect to time, location and other factors

Design Sheet 1

Layout



Title: Traffic Accident Visualization by Weather and Time of Day

Auther: Afshan, Khaja, Hari

Date: 05/11/12

Sheet: Sheet 02

Task: Develop a heatmap visualization to explore the relationship between traffic accidents, weather conditions, and time of day across Seattle.

Operations

The central heat map of Seattle displays accident hotspots, with interactive filters for time of day and weather conditions through two sliders beneath the map.

1. Weather Condition Slider: This slider allows users to select between different weather conditions: Clear, Overcast, Raining, Snowing, or All. When a specific weather condition is selected, the heat map dynamically updates to show accident hotspots under the chosen weather, highlighting areas most affected by accidents during that condition. Choosing All displays the combined data across all weather conditions, providing a comprehensive view of accident hotspots without weather-specific filtering.

2. Time of Day Slider: This slider lets users filter by Morning, Afternoon, Evening, Night, or All. When users select a specific time period, the heat map adjusts to display accident hotspots relevant to that time of day. Selecting All includes data from all times of day, giving an overview of general accident patterns across Seattle, unaffected by time-specific data.

These sliders give users control over the map's data display, helping them focus on specific conditions to identify trends and patterns in accident frequency.

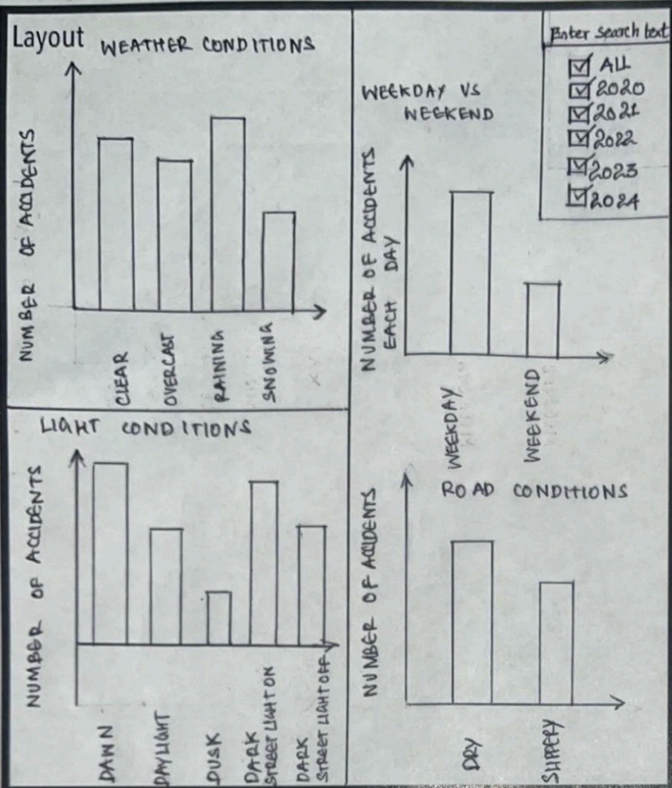
Focus

This visualization aims to explore links between traffic accidents in Seattle and conditions under which they occurred. The core visualization is a heatmap which will show accident locations superimposed on a geospatial map of Seattle. There will be interactive features for the user to explore how weather and time of day (both factors affecting visibility) impact the number of accidents. By shifting the sliders, we can see how the frequency of accidents changes when the weather is clear, rainy, or overcast or when it is daytime or dark. The user will also be able to zoom in and see in which part of the city the accident happened and if it was in an accident prone area. An option will be available to see all accidents which occurred under all conditions.

Discussion

The heatmap will effectively show the multivariate representations that we are trying to convey to the audience. We discussed how to set up our sliders for the weather and time of day variables. There were 13 categories of the weather variable (null, blowing sand/dirt, blowing snow, clear, fog/smog/smoke, other, overcast, partly cloudy, raining, severe crosswind, sleet/hail/freezing rain, snowing, and unknown). Some categories did not have meaningful data so they will be dropped. These would include null, other, and blowing sand/dirt. The slider will be graded according to visibility so some categories could be merged (overcast would include partly cloudy, snowing would include blowing snow etc.) The final scale would grade weather conditions as: clear à overcast à raining à snowing. Similarly, the time of day scale would have four stages: morning (5:00 am – 12:00 pm), Afternoon (12:00 pm – 5:00 pm), Evening (5:00 pm – 9:00 pm), Night (9:00 pm to 5:00 am). Both scales will culminate in showing all the data i.e. all accidents under all conditions.

Design Sheet 2



Title: Accident Conditions and Frequency Dashboard
Auther: Afshan, Khaja, Hari
Date: 05/11/12
Sheet: Sheet 03
Task: Examine how various conditions—such as weather, light, and road conditions—affect the frequency of accidents.

Operations

The 4 bar charts provide insights into the frequency of accidents under various conditions and days of the week, with interactive elements for a tailored analysis.

1. Year Filter Dropdown:
 • Action: Users select a specific year (2020, 2021, 2022, 2023, 2024, or All) from the dropdown menu.
 • Result: The bar charts dynamically update to display accident data only for the selected year, allowing users to analyze how accidents vary over time. Choosing All displays the combined accident data across all years.

2. Hover Tooltips:
 • Action: Users hover over a bar within any of the four charts (Weather, Light Conditions, Road Conditions, Weekday vs. Weekend).
 • Result: A tooltip appears, displaying the exact count of accidents, percentage of total accidents, and any other relevant details. This interaction gives users immediate access to data specifics without additional clicks, aiding quick, detailed analysis.

These interactions ensure that users can effectively filter, interpret, and compare accident data across different conditions and timeframes, enhancing the overall utility of the bar charts in the dashboard.

Focus

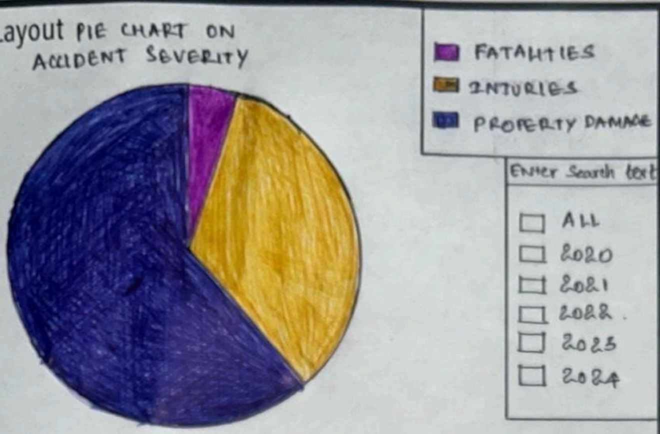
There will be 4 bar charts which will show accident frequency under different weather conditions, road conditions, light conditions, and whether the accident occurred on a weekday or the weekend. These bar charts will provide the user with instant quantitative information about which conditions have the most accidents without having to discern that information from the heatmap. This will reduce the cognitive load for the users as they will be able to glean insights from the data immediately. A common drop-down menu will also enable the users to track accidents through the years (2020 – 2024) with conditions updated for each year simultaneously.

Discussion

The first three bar charts will have the count of accidents on the y-axis. The first will compare accidents under different weather conditions (clear, overcast, raining, snowing) on the x-axis. The second bar graph will have road conditions on the x-axis with dry and slippery as the two independent variables. Since the datapoints for icy conditions and others like "oil" etc were too few and happened infrequently through the year, it was more meaningful to compare dry and slippery. The third bar graph would compare traffic accidents under light conditions including dawn, daylight, dusk, dark (streetlights on), dark (streetlights off/no streetlights). The fourth bar chart will display the average accident counts per day on weekdays vs. weekends to highlight differences in accident frequency between these two periods. Using a per-day average helps account for the difference in the number of weekdays versus weekend days, providing a fair comparison of accident frequency.

Design Sheet 3

Layout PIE CHART ON ACCIDENT SEVERITY



ACCIDENT INVOLVEMENT SUMMARY

INVOLVEMENT TYPE	SPEEDING	PEDESTRIAN INVOLVED	BICYCLE INVOLVED	PARKED CAR INVOLVED	VEHICLES INVOLVED
ACCIDENT COUNT (OUT OF TOTAL NUMBER OF ACCIDENTS)	[COUNT]	[COUNT]	[COUNT]	[COUNT]	[COUNT]

Title: Accident Severity Analysis
Auther: Afshan, Khaja, Hari
Date: 06/11/12
Sheet: Sheet 04
Task: Explore accident severity by fatalities, injuries, and property damage through a pie chart, with a summary of accident involvement types in a table.

Operations

Pie Chart Interactions

1. Year Filter Dropdown:
 * Action: Users select a specific year from the dropdown to view data for that year only.
 * Result: The pie chart updates to display the breakdown of accident severity (fatalities, injuries, property damage) specifically for the selected year, allowing users to analyze how accident severity distribution changes over time.
2. Hover to Display Actual Numbers:
 * Action: Users hover over each category in the pie chart.
 * Result: While the pie chart normally displays percentage values for each severity category, hovering over a category reveals the actual count, such as "8,000 fatality collisions out of 35,000 accidents." This provides a more precise understanding of accident severity proportions.

Table Interaction

1. Year Filter Dropdown:
 • Action: Users select a specific year from the dropdown to filter table data.
 • Result: The table updates to show accident involvement details (speeding, pedestrian, bicycle, parked car, and vehicle involvement) only for the selected year, making it easy to see trends in accident involvement types over time.

Focus

The data for accident severity will be shown on a pie chart. The categories include accidents resulting in fatalities, injuries, and property damage. The pie chart will be labelled with percentages, but the user would be able to hover on a specific category and see information about the total proportion of that category, for example, total fatalities in total number of accidents. This would give an added perspective on how many accidents result in loss of life, injury, or damage to cars or other property. There will be a drop down option to filter for a specific year. A text table will present facts of general interest to the user from the entire dataset. It would list the number of accidents in which speeding was as issue, in which vehicles, pedestrians, and cyclists were involved, and in which parked cars were hit. This table would also have an option to filter for a specific year.

Discussion

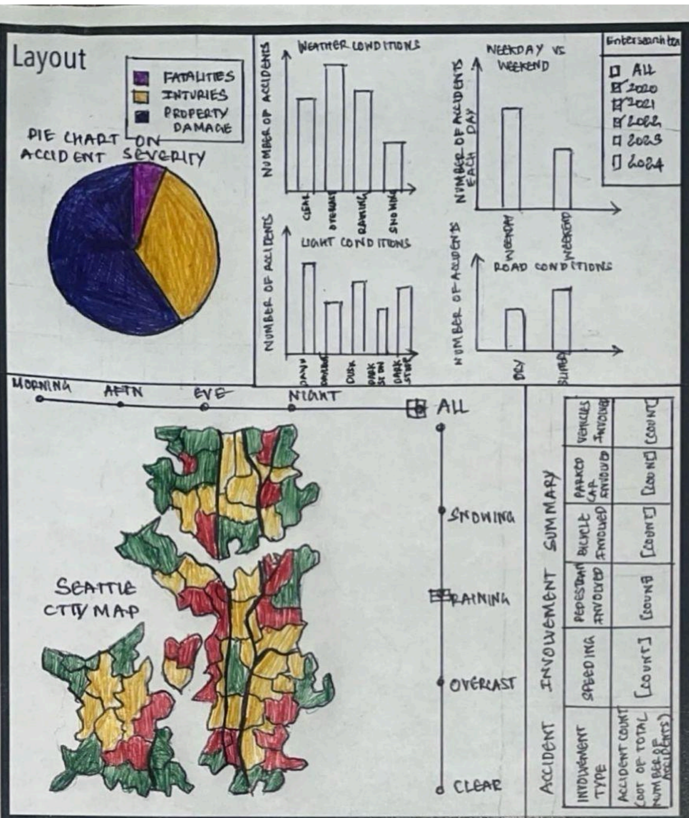
The accident severity variable includes three distinct categories: fatalities, injuries, and property damage. Since the dataset treats each category separately and does not account for any overlap (e.g., accidents involving both fatalities and injuries), we used a pie chart to represent the proportion of accidents in each category. This visualization effectively conveys which category most accidents fall into at a glance.

For the text table, we focused on five key involvement types—speeding, pedestrian involvement, bicycle involvement, parked car involvement, and vehicle involvement—because these factors provide insight into the primary contributors to accident scenarios. Each involvement type highlights a different dimension of road safety without implying direct relationships between th

Design Sheet 4

5

Layout



Title: Visualizing Seattle Traffic Accident Patterns

Auther: Afshan, Khaja, Hari

Date: 06/11/12

Sheet: Sheet 05

Task: Combine visualizations from sheets 2, 3, and 4 to create an interactive Tableau dashboard visualizing Seattle traffic accident patterns from 2020 to 2024.

Operations

1. Heat Map:
 • Weather Condition Slider: Allows user to track accidents through different weather conditions ranging from clear, Overcast, Raining, or Snowing. Selecting All displays combined data across all weather conditions
 • Time of the Day Slider: Allows user to track accidents across all times of the day starting from Morning and going through afternoon, evening, and night. Selecting All displays combined data from all times of the day.
2. Bar and Pie Chart:
 • Year Filter Dropdown: User can filter data for specific year
 • Hover tooltips: Display actual numbers.
3. Text Table:
 • Year Filter Dropdown: User can filter data for specific year.

Focus

The final dashboard brings all the visualizations together to tell a story. The heat map is the main feature where the user can explore how the weather conditions and time of the day impact the number of accidents in Seattle. The interactive options would provide the user control over filtering for a specific year, and use sliders to track the impact of weather conditions and time of the day on the number of accidents. If interested in more quantitative information, the user can glean information from the bar and pie charts and the text table. Here again, the user will be able to filter over specific years and hover to look at exact figures. Overall, the aim is to educate under which conditions most accidents occur and if there are specific hotspots for accidents in the city.

Detail

• Tableau Desktop is used for dashboard design, with Tableau Prep handling data preparation, including:
1. Retaining essential columns: location (latitude and longitude), time of accident (date and time), weather condition, light condition, road condition, accident severity, and involvement details (speeding, pedestrian, bicycle).
2. Transforming columns, e.g., extracting the year from the date for filtering by 2020–2024.
Data is cleaned in Tableau Prep, then imported into Tableau Desktop for visualization. The dashboard can be published to Tableau Server or Tableau Online for wider access. Tableau's interactive features enhance data integration, manipulation, and visualization.

Visualization Methods and Algorithms:
 • Heat Map: Displays accident hotspots on a Seattle city map using color intensity to indicate accident frequency based on location, time of day, weather, and road conditions.
 • Bar Charts: Visualize accident data across conditions (weather, light, road, weekday vs. weekend) with aggregated data presented as bars.
 • Pie Chart: Shows accident severity (fatalities, injuries, property damage) as a percentage of the total, with each category as a slice of the pie.
 • Table: Provides details on accident involvement (speeding, pedestrian, bicycle, parked car, vehicle) by counting occurrences and displaying them in tabular format.

Design Sheet 5