**SDU University**

Изображение выглядит как текст, Графика, графический дизайн, Шрифт

Автоматически созданное описание

**Final Project Documentation**

**Presenters: Daulet Baglan, Kongratbayev Yerzhan**

**Lecturer: Nazerke Sultanova**

**U.S car accidents dataset researching and vizualizing**

**1. Introduction**

Accidents on the road are an unfortunate and persistent issue that affect millions of people worldwide. Whether it’s minor fender benders or more serious incidents, accidents disrupt lives, cause injuries, and create financial burdens. Our project tackles this problem by using data analytics to predict where accidents are most likely to occur and understand the factors that contribute to their severity. By analyzing real-world data, our goal is to help authorities make informed decisions, potentially saving lives and reducing injuries on the road.

**2. Project Objective**

The aim of this project is simple but impactful: to use data to make our roads safer. Through a combination of accident data, weather conditions, and traffic patterns, we aim to identify:

* **Accident Hotspots:** Areas where accidents happen more frequently.
* **Accident Severity:** How factors like weather or traffic affect the severity of accidents.
* **Resource Allocation:** Helping authorities focus on the most dangerous areas to improve safety.
* **Predictive Insights:** Providing authorities with the ability to predict accidents and take proactive actions.

**3. Methodology**

To achieve our goals, we followed a structured approach that combined data collection, analysis, and machine learning techniques. Here's how we did it:

1. **Data Collection**  
   The first step was gathering relevant data. We used accident data from multiple cities, which included details like the location of the accident, the time it happened, the weather conditions, and the severity of the incident. Additional data on traffic volume, weather forecasts, and other factors were also incorporated to get a holistic view of the situation.
2. **Data Preprocessing**  
   Once we had the data, we cleaned it up. This meant getting rid of missing or incomplete data, correcting any errors, and ensuring all the information was in the right format. We also combined certain columns, like creating a “location” column to map city names into geospatial data, which allowed us to plot accidents on a map.
3. **Modeling and Analysis**  
   With our data ready, we used machine learning models to uncover patterns. We employed models like regression analysis and classification to determine what factors most influenced accident rates and severity. Weather conditions, traffic patterns, and time of day were all analyzed to predict where and when accidents were more likely to occur.
4. **Visualization**  
   Visualizing the data helped bring the findings to life. We used maps, heatmaps, bar charts, and scatter plots to show accident trends and to highlight areas with the highest risk. These visuals made it easier to understand the data and communicate the results clearly to stakeholders.

**4. Results and Insights**

After analyzing the data, we gained several important insights that help us better understand the dynamics of road accidents:

* **Accident Hotspots:**  
  Certain cities and intersections had higher accident rates than others. By using machine learning, we could pinpoint where accidents were most likely to occur, helping authorities focus on these areas for improvements.
* **The Impact of Weather:**  
  Weather conditions, particularly rain, fog, and snow, were found to significantly increase the likelihood of accidents. In fact, accidents that occurred in bad weather were often more severe, highlighting the importance of adjusting traffic control measures based on weather conditions.
* **Traffic Volume Correlation:**  
  We found a clear correlation between higher traffic volumes and more accidents. During rush hours or in congested areas, accidents were more frequent. This could be addressed by better managing traffic flows and introducing measures such as improved public transport or traffic diversion during peak hours.
* **Severity Predictors:**  
  The project also helped us understand why some accidents were more severe than others. Bad weather, high-speed zones, and certain time frames (like late-night or early-morning) were more likely to result in severe accidents.

**5. Social and Economic Impact**

The potential impact of this project goes beyond just the technical findings—it has tangible social and economic benefits:

* **Reducing Accidents:**  
  By identifying hotspots and understanding the factors contributing to accidents, the project can help reduce the number of incidents on the road. Fewer accidents mean fewer injuries and fatalities, which is a direct benefit to society.
* **Saving Costs:**  
  Accidents come with significant economic costs—emergency services, insurance claims, and road repairs are just a few examples. By preventing accidents, this project can save cities and governments substantial amounts of money.
* **Better Resource Management:**  
  With the insights generated, authorities can allocate resources (such as police presence, road maintenance, or infrastructure improvements) more effectively. This ensures that areas with higher risks receive the attention they need.

**6. Conclusion**

This project proves that data can be a powerful tool in making our roads safer. By using accident data, weather conditions, and machine learning, we've been able to predict accident hotspots, understand the factors that lead to severe accidents, and provide insights to help authorities take proactive measures. The potential to reduce road accidents, save lives, and optimize resources makes this project not just valuable, but necessary.

We’ve shown that with the right data, we can make better decisions, improve safety, and create smarter, more efficient systems for managing traffic and accidents. And this is just the beginning—this project has the potential to grow, evolve, and scale, contributing to safer cities around the world.