Final Project Proposal

2024-11-07

The group members (names and UNIs)

Tianci Zhu (tz2659)
Jingxi Wang (jw4690)
Avery Wang (xw3038)
Jianing Chen (jc6422)
Puyuan Zhang (pz2334)

The tentative project title:

Patterns and Impacts of Tornadoes in the U.S

The motivation for this project:

Severe tornadoes disrupt lives, damage property, and place considerable strain on local resources. By analyzing patterns in historical storm data, we can better understand where and when these storms are likely to happen, as well as their potential impacts. This information can support local governments, emergency services, and residents to prepare for future storms, reducing harm and costs. Our project aims to provide practical insights into storm patterns and impacts across different states, making essential information accessible to help communities become more resilient.

The intended final products:

1. A Detailed Report
2. An Interactive online Dashboard
3. A project website with all summaries and key takeaways
4. A two-minute video overview that walks viewers through the primary goals of the project, key findings, and highlights from the visualizations
5. GitHub Repository with all code and data processing scripts

The anticipated data sources:

https://www.spc.noaa.gov/wcm/
We will use the dataset "1950-2023_actual tornadoes.csv" from NOAA Storm Prediction Center Tornado Data, which provides data on individual tornado tracks without state segments or continuing county info. For this project, we will focus on data from 2000 to 2023.

The planned analyses / visualizations / coding challenges

Analysis:

Conduct statistical analyses such as tornado frequency, intensity, seasonal patterns, and impacts across states and years (from 2000 to 2023)

Visualizations:

a variety of plots such as interactive bar plots, line plots, scatterplots, box plots, and heatmaps will be implemented to explore trends, patterns, and relationships within the data, allowing for a clearer understanding such as tornado frequency, intensity, geographic distributions and impacts across different states and years

Coding challenges:

efficiently handling and cleaning large datasets with missing values and improper data types, along with data wrangling to prepare the data for appropriate analyses and visualizations. Handling large data volumes, geospatial data visualization, and advanced statistical analysis for impact metrics.

The planned timeline:

Week 1: Data collection and cleaning; initial EDA.
Week 2: Conduct trend and impact analysis and develop visualizations for storm types and frequencies.
Week 3: Complete mapping and geospatial analysis.
Week 4: Build the website. Prepare the final report and screencast. Review and polish all deliverables.