Code:You Capstone Project Plan

Abstract:

The purpose of this project is to provide a risk assessment narrative for the continental United States by analyzing historical natural disaster data. Using NOAA’s natural hazards data and FEMA’s disaster declarations, the project will map high-risk and low-risk regions based on previous disaster occurrences. This map aims to offer insight into which areas are most vulnerable or resistant to natural disasters, helping inform risk awareness and planning.

Problem Statement:

Problem: Communities and planners across the U.S. often lack a clear, data-backed narrative of disaster risk by region. Without accessible risk maps derived from comprehensive historical data, there’s a gap in awareness that can hinder effective decision-making for disaster preparedness and mitigation.

Opportunity: By combining NOAA and FEMA data, this project provides a consolidated view of historical disaster patterns, helping highlight regional vulnerabilities. This insight could benefit community leaders, businesses, and residents in understanding regional risks and making informed decisions.

Goals and Objectives:

Primary Goal: Create a risk assessment map that visually communicates natural disaster vulnerability across the continental United States, based on historical trends.

Additional Goal: Provide a narrative through data analysis that highlights which regions are relatively resilient or prone to natural disasters, making this information accessible and actionable.

Features:

Core Features:

Data Merging and Cleaning: Integrate NOAA’s and FEMA’s historical data into DuckDB, ensuring data accuracy and consistency for analysis.

Risk Analysis and Pattern Identification: Use DuckDB to run queries identifying high-risk vs. low-risk regions.

Interactive Risk Map in Tableau: Build an interactive map in Tableau that enables users to explore disaster risk across regions, with filtering options for disaster types and time frames.

Stretch Goals:

Seasonality and Trend Analysis: Visualize seasonal disaster patterns to provide additional insight into timing-related risks.

Interactive Filtering by Disaster Type: Allow users to filter data by specific disaster types (such as floods, hurricanes etc.) to customize their risk assessments.

Technologies:

Data Processing and Querying: Python & DuckDB

Data Analysis and Preparation: pandas for initial data cleaning and manipulation, DuckDB for in-depth queries.

Visualization and Mapping: Tableau Data Sources: NOAA’s Natural Hazards Data, FEMA’s Disaster Declaration data

Architecture:

The architecture will involve ingesting and merging NOAA and FEMA datasets in DuckDB, performing analytical queries to categorize regions by risk level, and then exporting summarized data to Tableau for visualization.

Risk and Mitigation:

Data Integration Challenges: NOAA and FEMA datasets may have differing structures.

Mitigation: Standardize fields during preprocessing for consistency.

Dashboard Responsiveness: Ensuring the Tableau map responds to user interaction.

Mitigation: Manage dataset size and responsiveness.