



Analyzing Climate's Role in Urban Accidents in the U.S.

MVP 1: Initial Progress Presentation



Project Scope and Objectives

Objective: Quantify Weather-Accident Relationships

Approach: Time-Series Analysis

- Explore Seasonal Trends
- Pattern Recognition
- Forecasting

Data Sources

Traffic Data:

- U.S. Accidents (2016 - 2023) - Kaggle
- Crash Report Sampling System from NHTSA

The Weather Dataset: Meteostat from Kaggle

Cities:

- Madison , Wisconsin.
- Nashville, Tennessee.
- Boise, Idaho.





Preliminary Data Exploration & Preprocessing Strategy

- Initial Findings:
Summary Statistics & Distributions
 - Accidents Across Cities
 - Accidents Across Seasons
- Tentative Insights: Correlations
- Challenges:
Merging & Missing Data.
 - Time-Series Alignment
 - Missing Data Handling
- Solutions: Imputation & Alignment



Feature Engineering Approach

- **Features:** Temporal features, Lag Variables & Seasonality, Normalization.
- **Goals:** Capture Time-Dependent Patterns
- Short-term Trends, Long-term Trends



Initial Model Selection

- **Models:** ARIMA, SARIMA, (potentially) Prophet, and Recurrent Neural Network (RNN) like LSTM/GRU.
- **Criteria:** Accuracy & Interpretability.
- ARIMA (Autoregressive integrated moving average)/SARIMA (Seasonal AutoRegressive Integrated Moving Average) for Baseline.
- Prophet & LSTM (Long -short-term memory) and GRU (Gated Recurrent Unit) for Advanced Analysis.



Methodology Overview

- Methodology Phases:

- Data Collection
- Preprocessing
- Exploratory Analysis
- Feature Engineering
- Model Building
- Evaluation

- Evaluation Metrics:

- MAE
- RMSE
- MAPE



Next Steps

- **MVP 2 Goals:**
Feature Engineering & Initial Modeling
- **Final Goals:**
Model Evaluation & Academic Paper



Questions & Feedback



Presented by Sergio David

22.09.2023