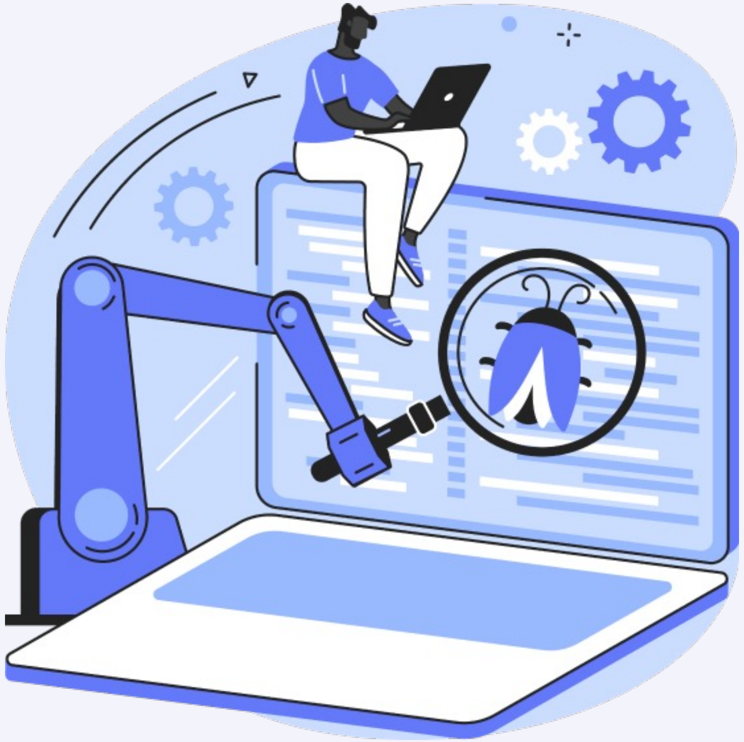
BOREALIS AI

Let's SOLVE it

ShelterSight Toronto



Optimizing Homeless Shelter Operations in Toronto

By Nida Copty, Emily Nguyen, Tom Nguyen, India Tory

Table of Contents

BOREALIS AI

Let's SOLVE it

Introduction

25,000 to 35,000

people are homeless on any given night in Canada

The annual cost of homelessness to society is

$ 7B

The Canadian Government is investing

$ 2.2B

over 10 years to expand funding for homelessness

Problem Statement

To enhance the efficiency of Toronto's homeless shelter operations through the development of a machine learning model designed to accurately anticipate fluctuating shelter demands driven by dynamic urban challenges.

Goals

Bridging AI and Decision-Making in Shelter Management

To develop a user-friendly, reliable forecasting model that integrates seamlessly into the decision-making processes of shelter operations.

Empowering Shelter Operations with Predictive Analytics

To equip shelter operators with predictive tools that can anticipate periods of high demand and adjust resource allocation accordingly.

Our Team

Nida Copty

Thomas Nguyen

Emily Nguyen

India Tory

Methods

Data

Data

- Daily Shelter & Overnight Service Occupancy & Capacity
- Government of Canada Historical Climate Data
- New Housing Price Index (Monthly)
- Person in Crisis Calls for Service Attended Open Data
- Unemployment Data

Dataset Selection

Daily Shelter & Overnight Service Occupancy & Capacitance by the City of Toronto

- Program ID
- Occupancy Date
- Organization ID
- Location ID
- Location Postal Code
- Location City
- Sector
- Program Model
- Program Area
- Capacity Type
- Unavailable Beds/Rooms
- Occupancy Rate Beds/Rooms

BOREALIS AI

Let's SOLVE it

Data Visualization

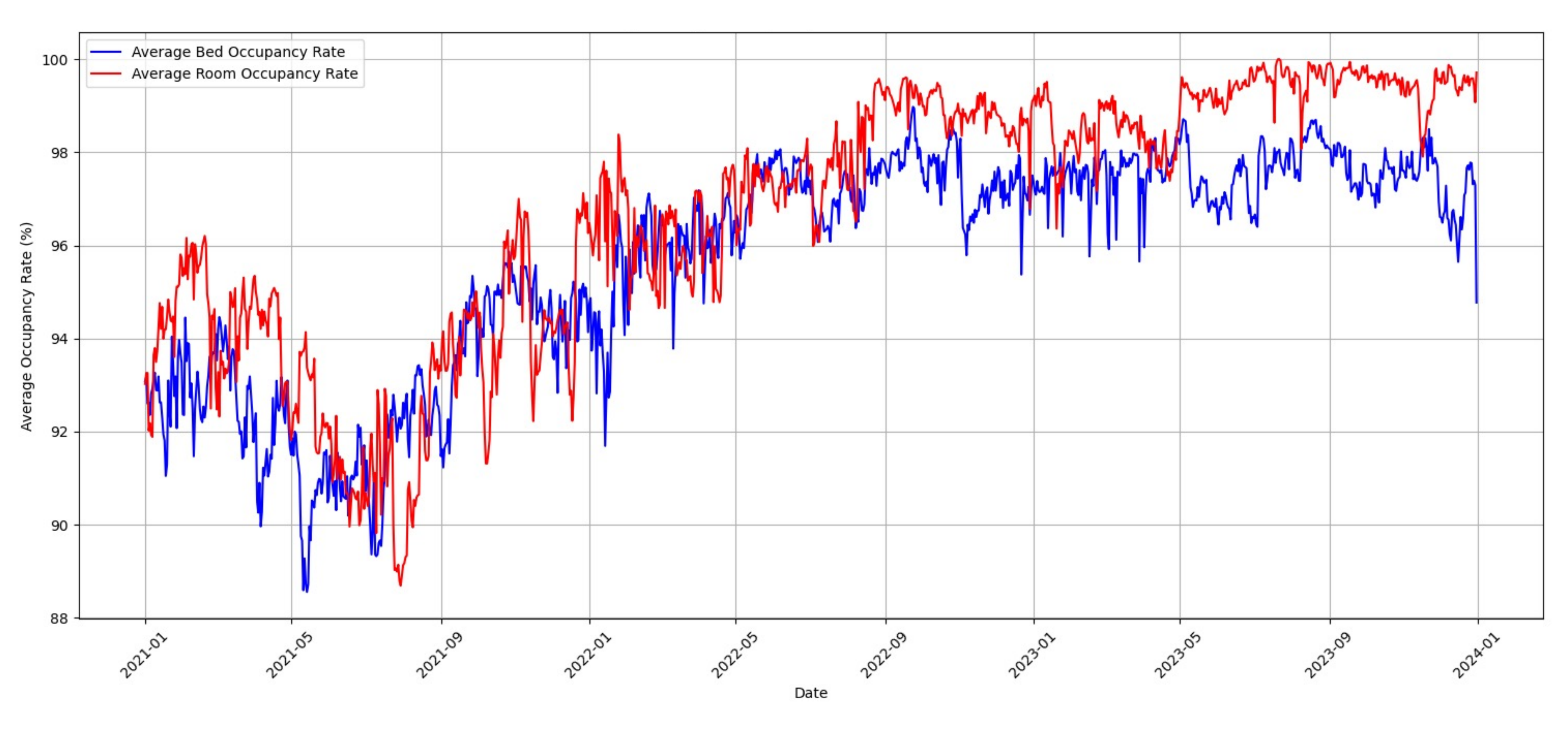


Figure 1: Daily Average Occupancy Rate Across All Shelters

Data Visualization

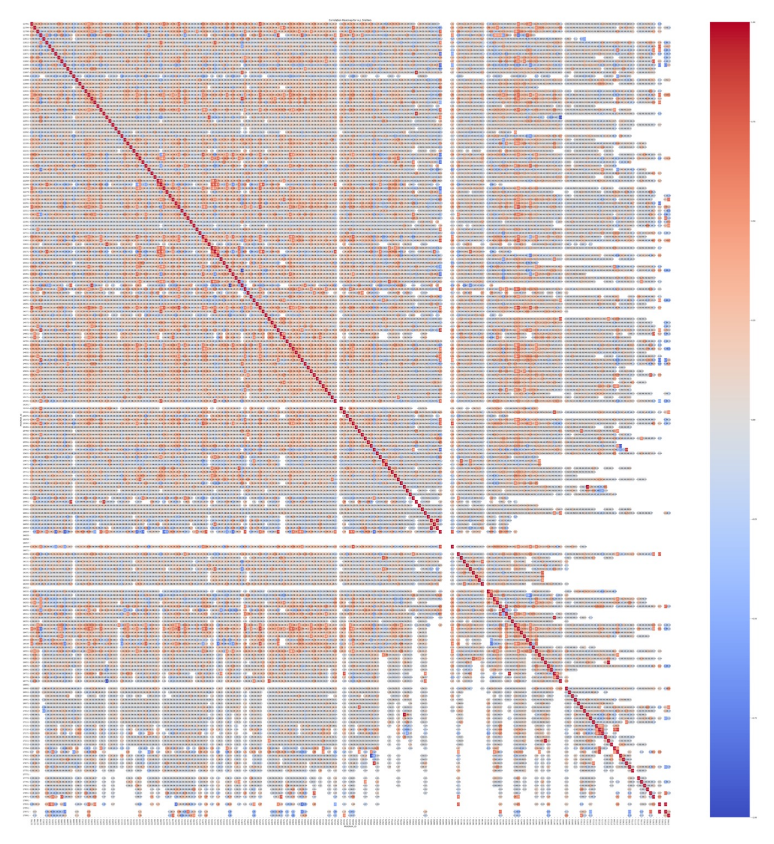


Figure 3: Correlation Heatmap for All Shelters

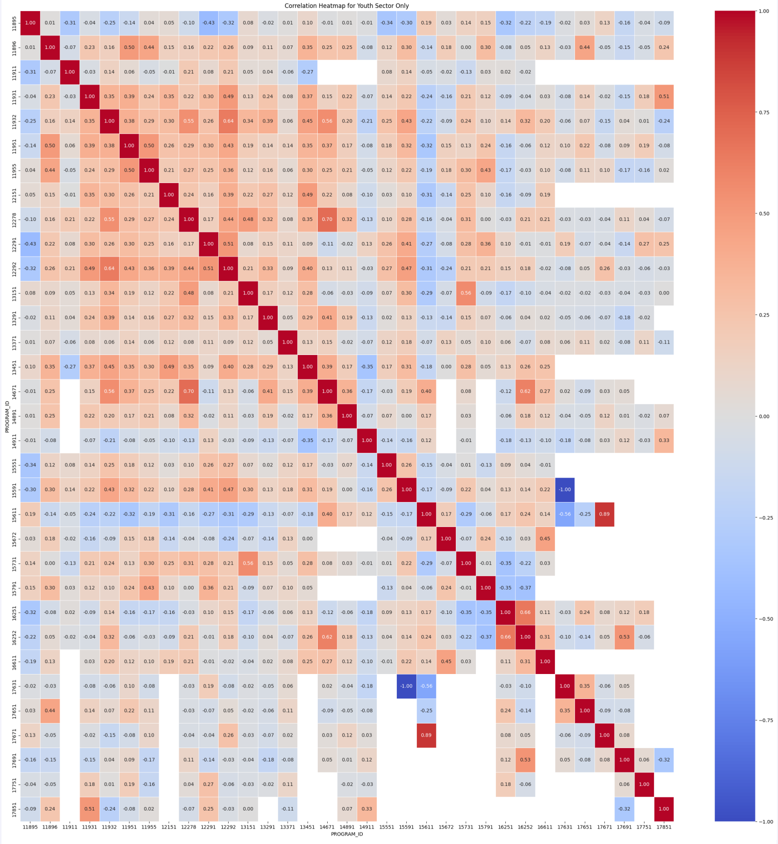


Figure 4: Correlation Heatmap for Only "Youth" Shelters

Baseline Models

Baseline Model Performance

Random Forest Regression Model

Root Mean Squared Error: 7.04

Mean Absolute Error: 2.59

Autoregressive Model

Root Mean Squared Error: 4.12

Mean Absolute Error: 16.96

Prophet Model (5-Day Horizon)

Root Mean Squared Error: 13.66

Mean Absolute Error: 13.44

Model Overview

Two-Step Methodology

1

Long Short-Term Memory for Initial Demand Forecasting

2

Forecast Adjustment Based on Data Features and Correlation

Long Short-Term Memory

What is LSTM?

- LSTM is an advanced recurrent neural network (RNN) architecture
- Uses a unique structure of gates that regulate the flow of information

Why LSTM for Forecasting Shelter Demand?

- Handling sequential data
- Adaptability to time series forecasting
- Robustness to data variability
- Dynamic response to external factors

Model Construction

Model Workflow

1. Data Collection
2. Data Pre-processing
3. Clustering by Location
4. LSTM Model Training
5. Feature-Based Adjustment
6. Prediction Output

LSTM Iterations

Univariate LSTM
for Combined City Average Forecast

Multivariate LSTM
for Combined City Average Forecast

Geospatial Multivariate LSTM
for Individual Shelter Predictions

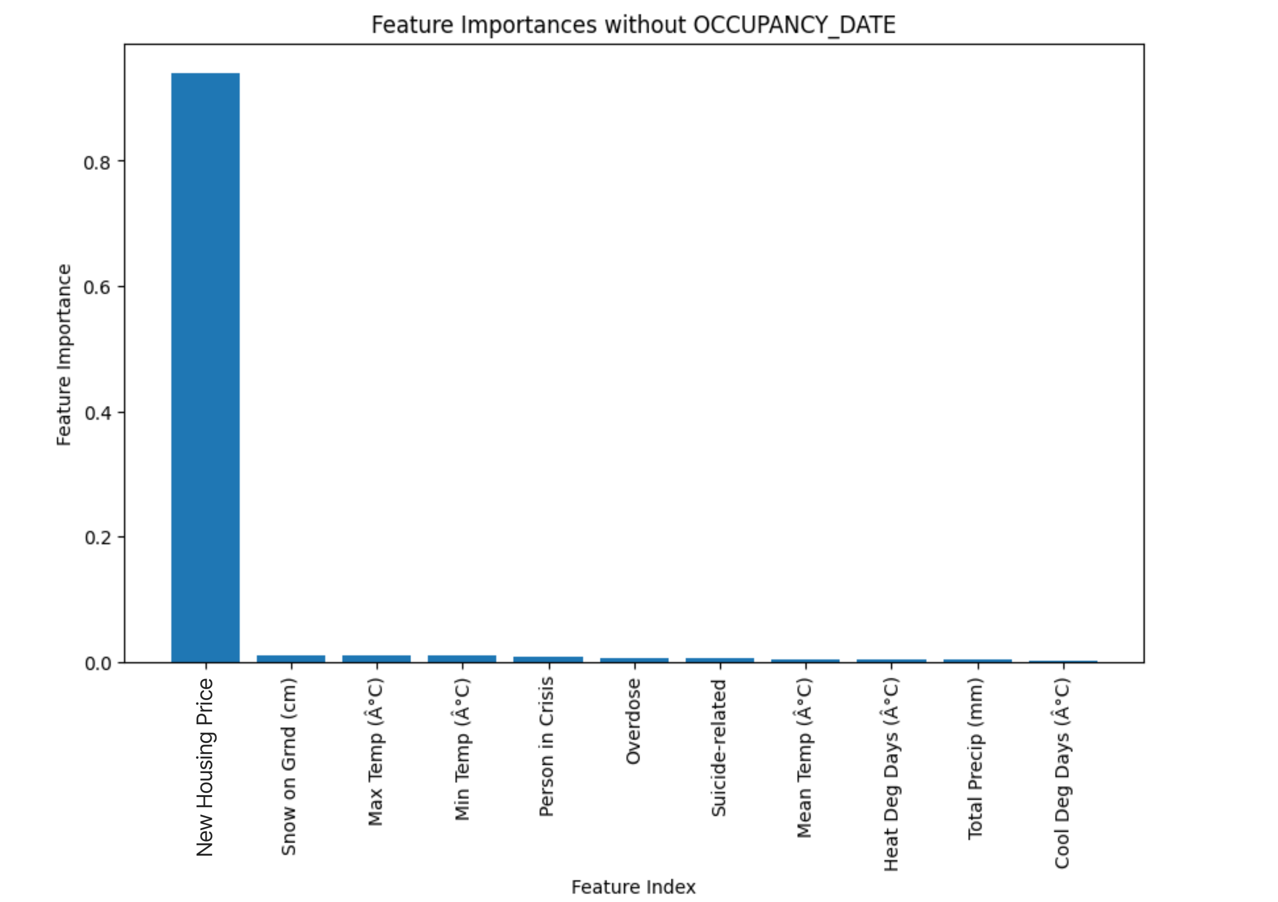
- Moderate forecasting performance for nightly city average
- Weak forecasting performance for individual shelters

BOREALIS AI

Let's SOLVE it

Feature Importance

- Used random forest regression model
- Determined Gini importance of each feature
- Narrowed down to 4 features to compare against



Geospatial Multivariate LSTM

Clustering

- **Feature Grouping:** apply clustering algorithms to categorize shelters
- **Dimensionality Reduction:** use clustering to reduce the complexity of the data
- **Enhanced Accuracy:** improve accuracy and specificity of forecasts

Model

- **Architecture:** utilizes multiple stacked LSTM layers and fully connected layers
- **Sequential Data Processing:** batch-first approach to predict based on historical context
- **Flexible and Scalable:** adaptable for varying input sizes and grouping schemes

1. Understanding Shelter Geography

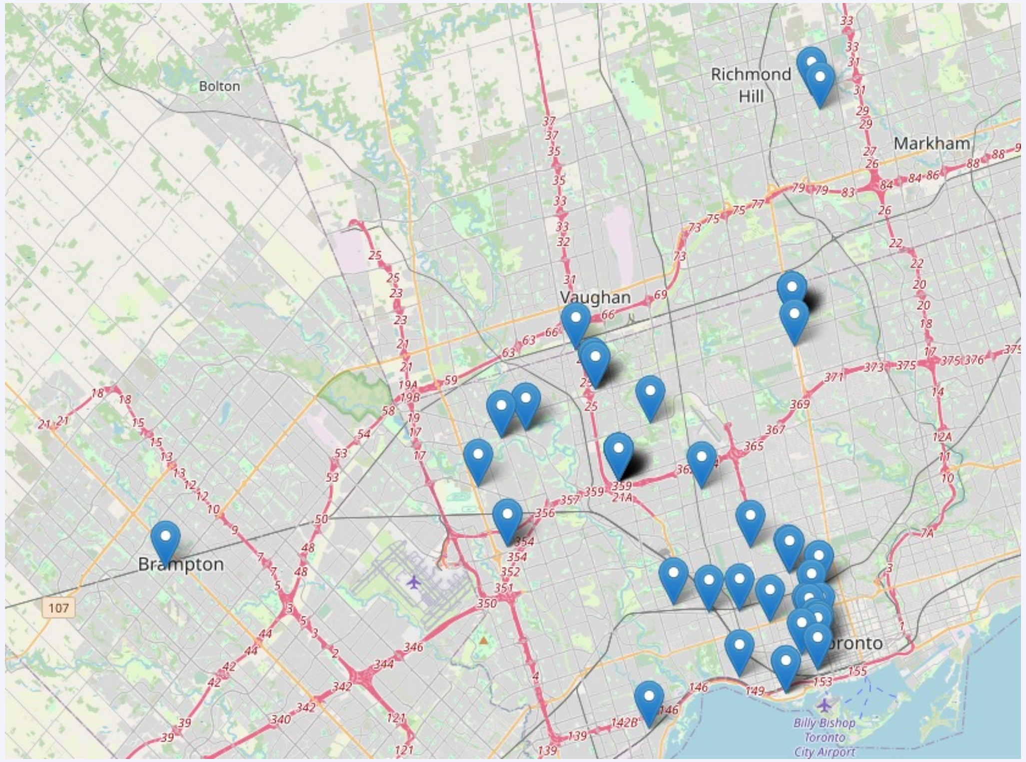


Figure 6: Map of Toronto Shelter Locations

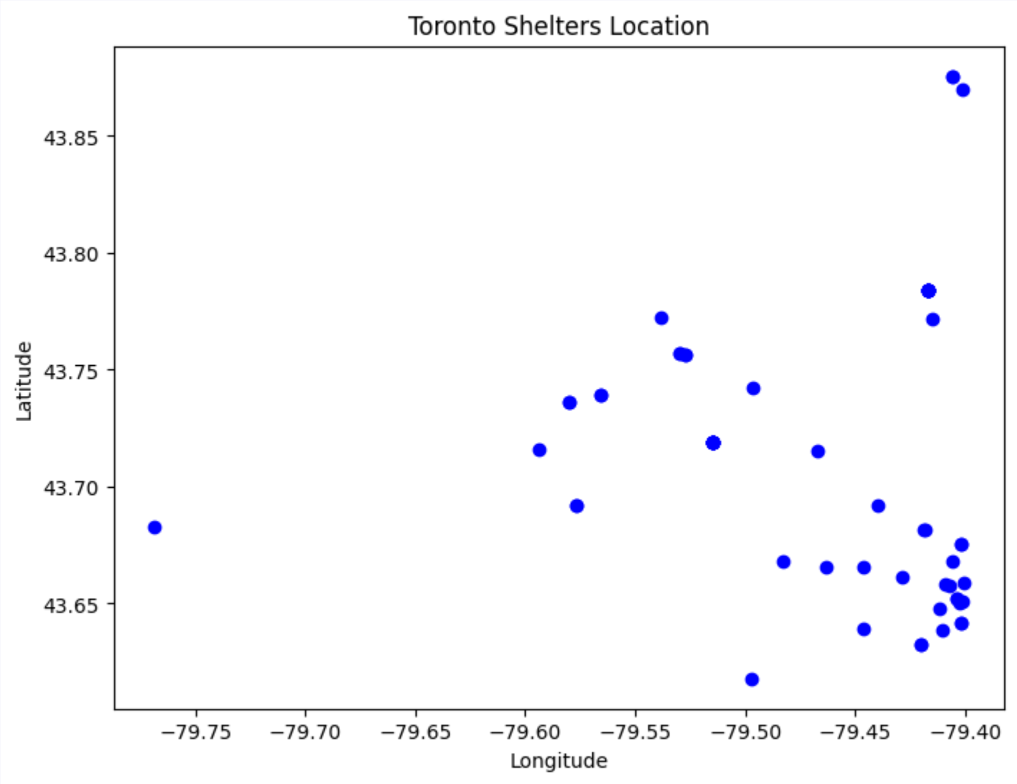


Figure 7: Longitude vs Latitude Plot of Shelter Locations

2. Determining Clusters with Distortion Analysis

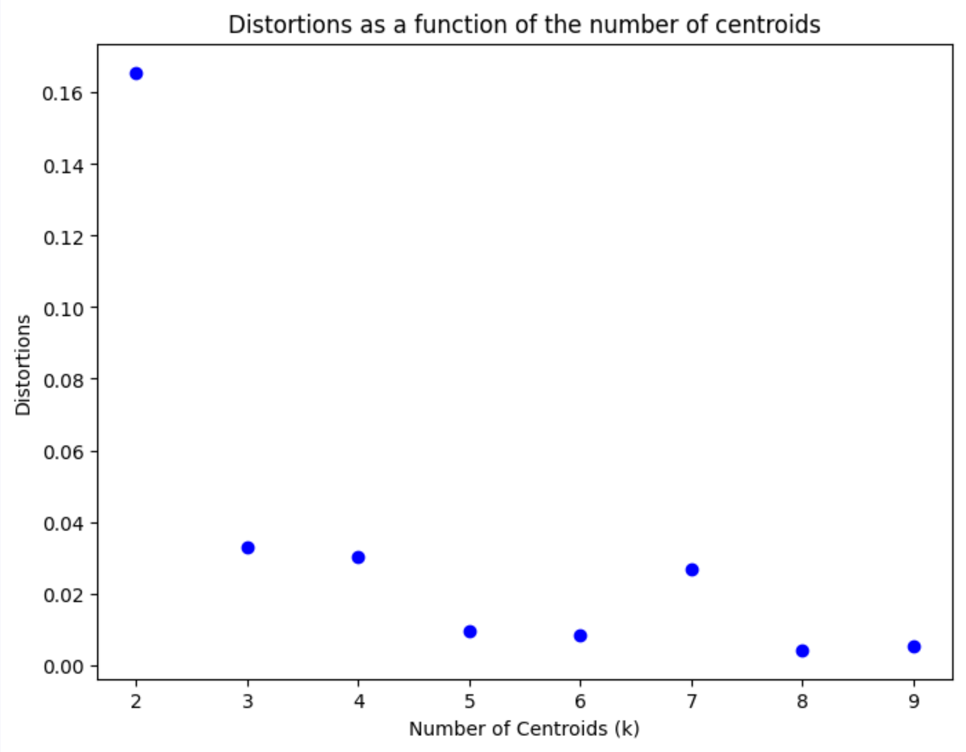


Figure 8: Distortions as a Function of the Number of Centroids

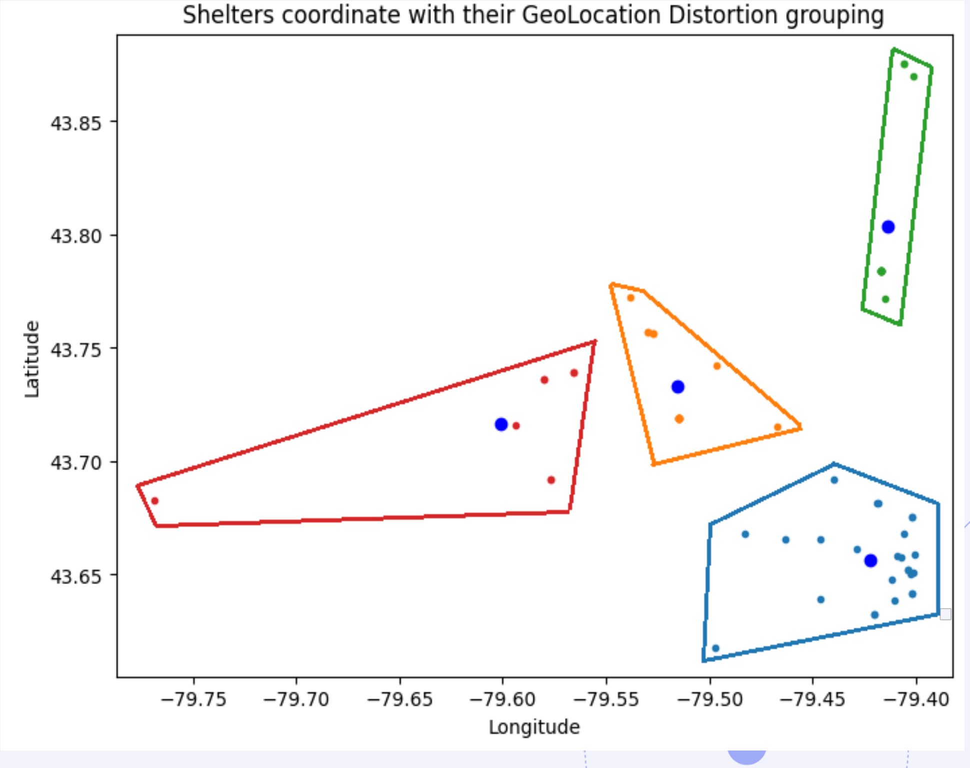


Figure 9: Plot of Shelter Locations and Resulting Groupings

2. Model Training with Distortion Analysis

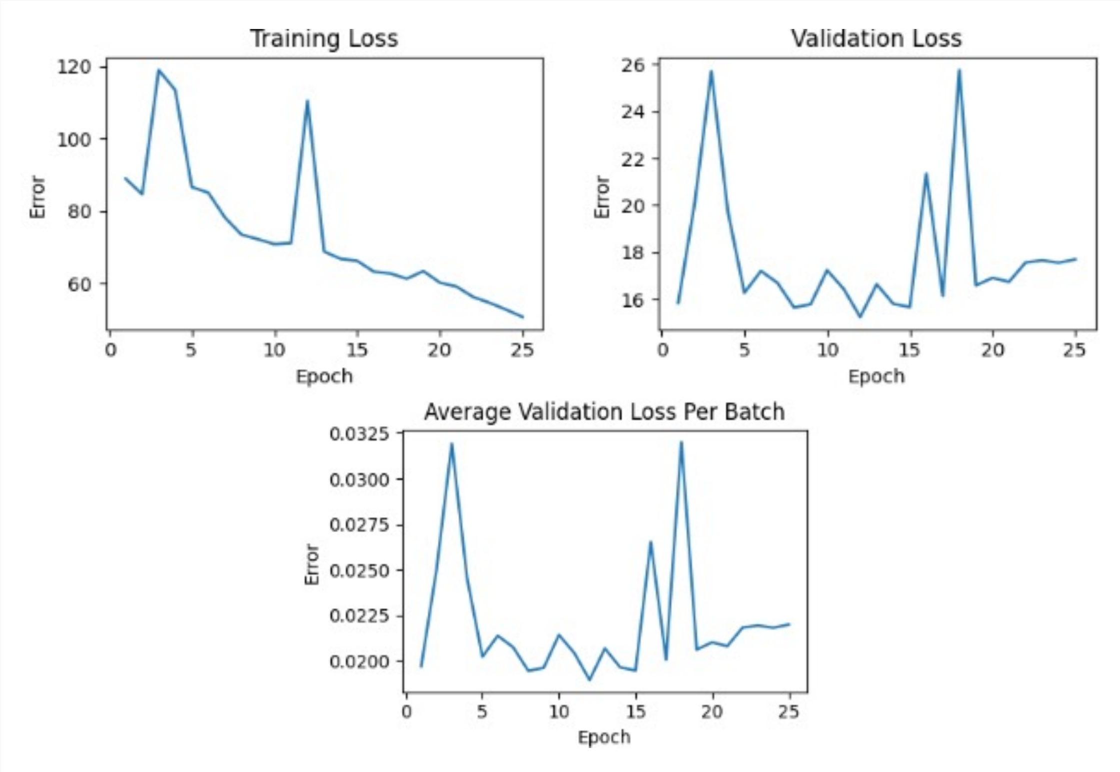


Figure 10: Training and Validation Loss

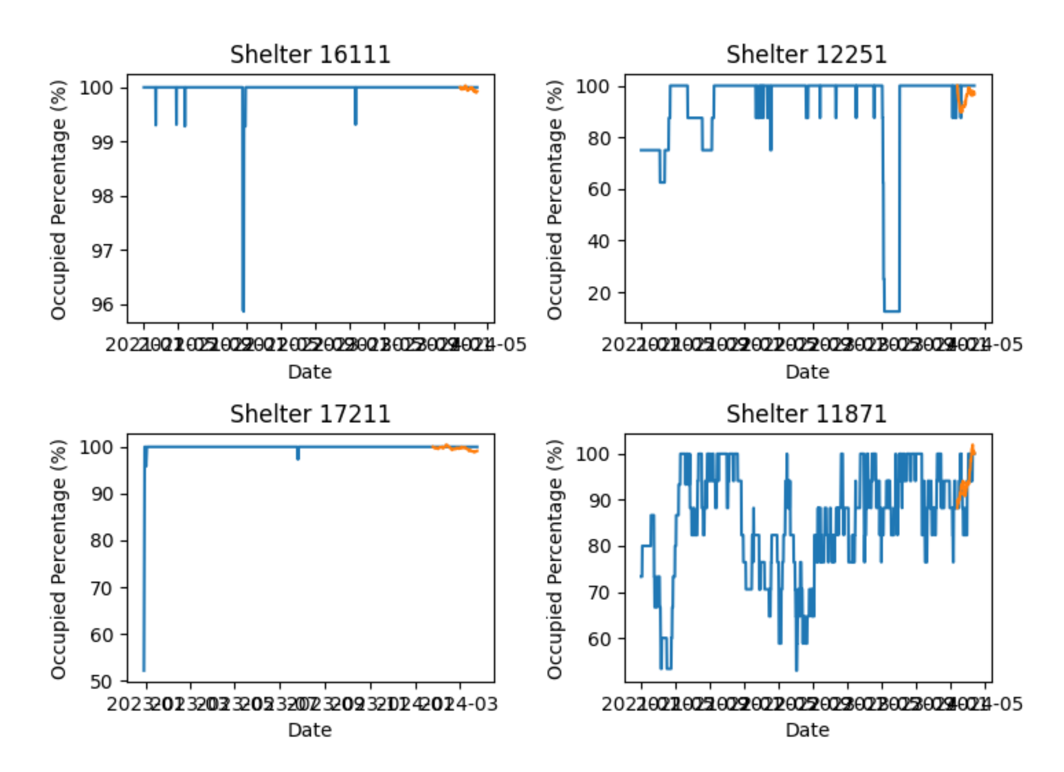


Figure 11: Model Prediction on Random Shelters

3. Determining Clusters with Correlation Analysis

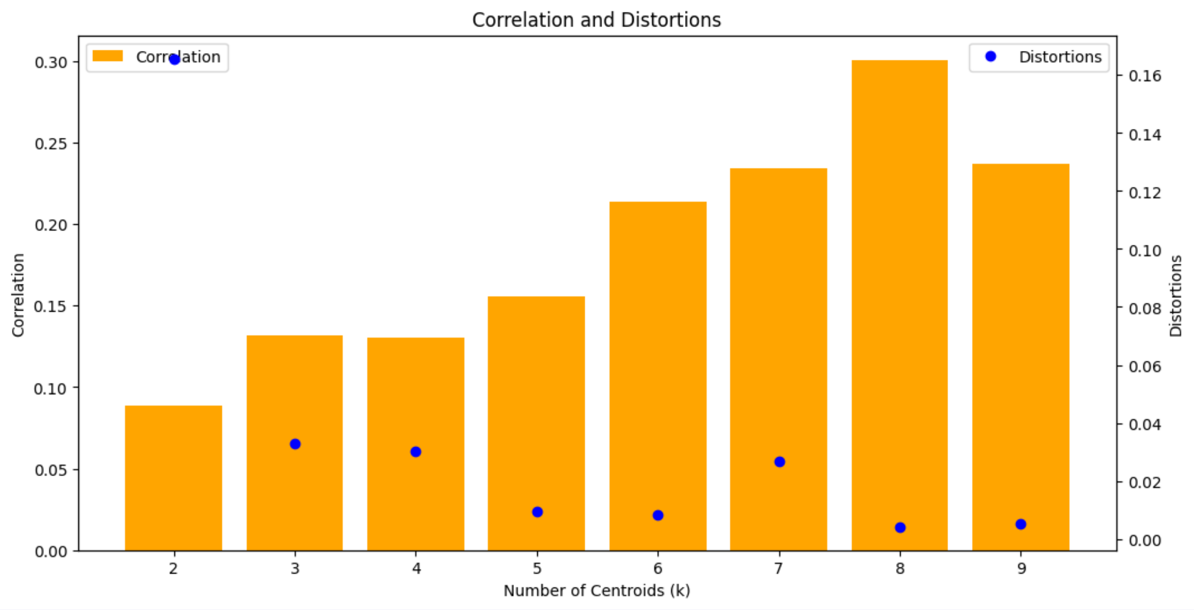


Figure 12: Plot of Correlation vs Number of Centroids

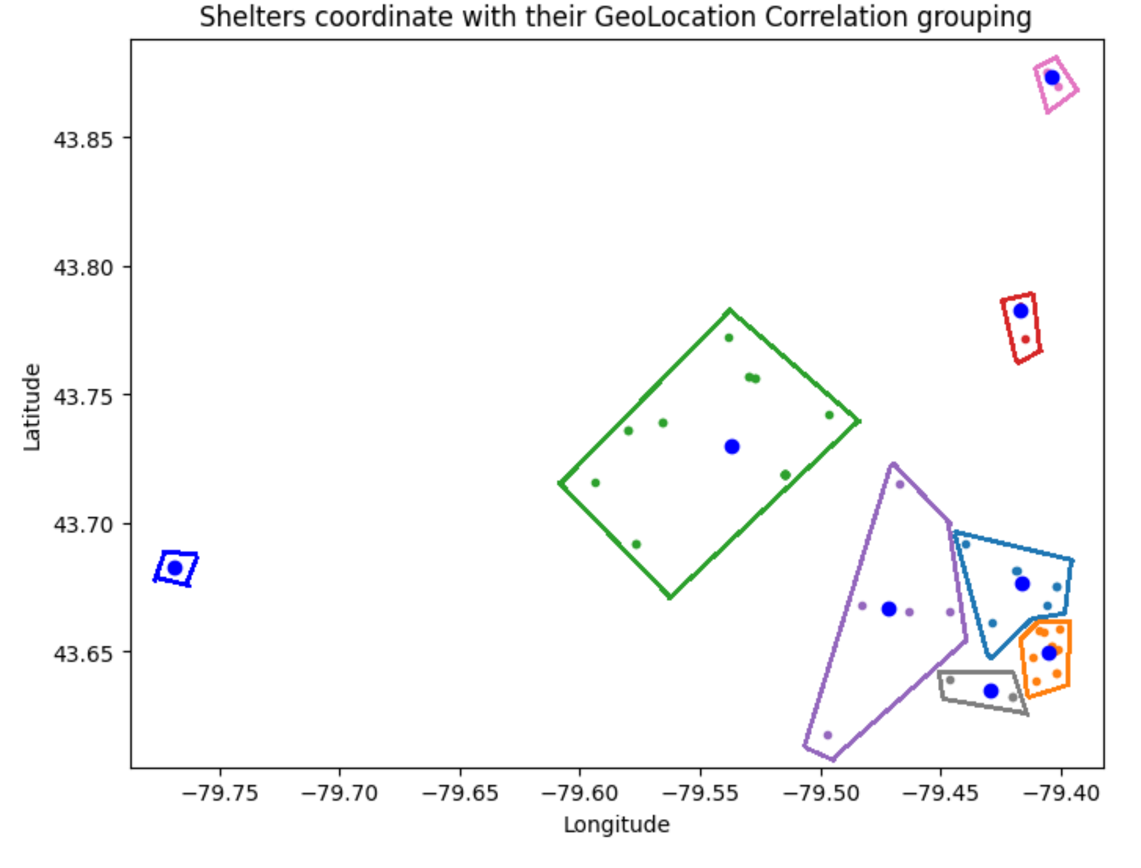


Figure 13: Plot of Shelter Locations and Resulting Centroids

3. Model Training with Correlation Analysis

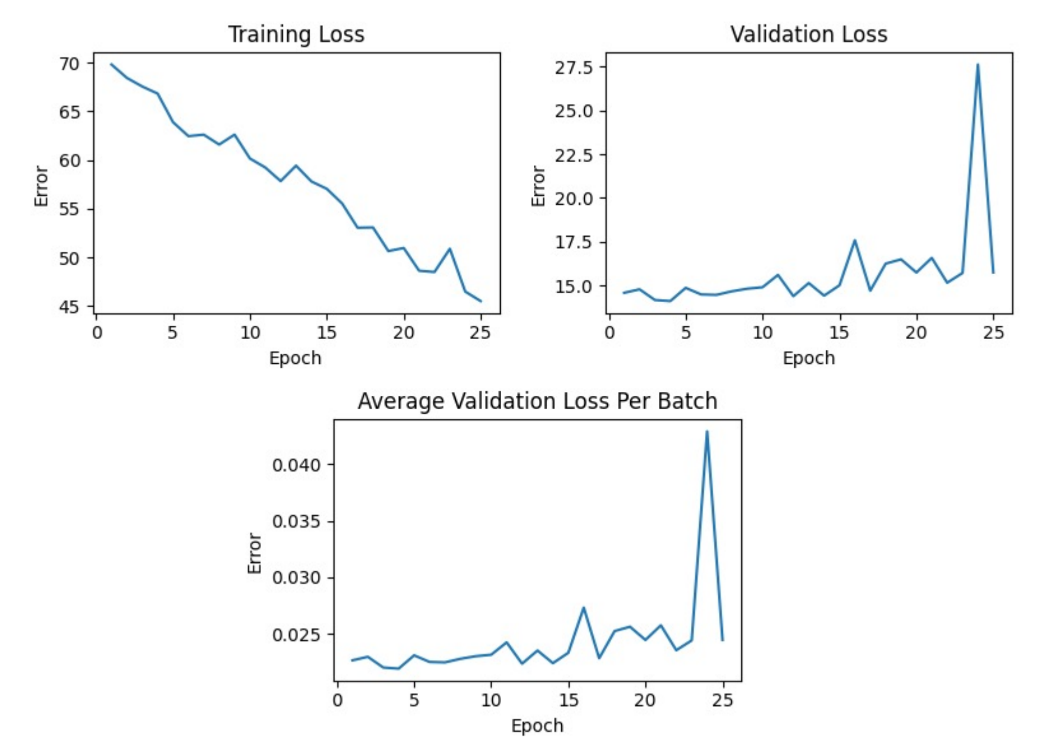


Figure 14: Training and Validation Loss

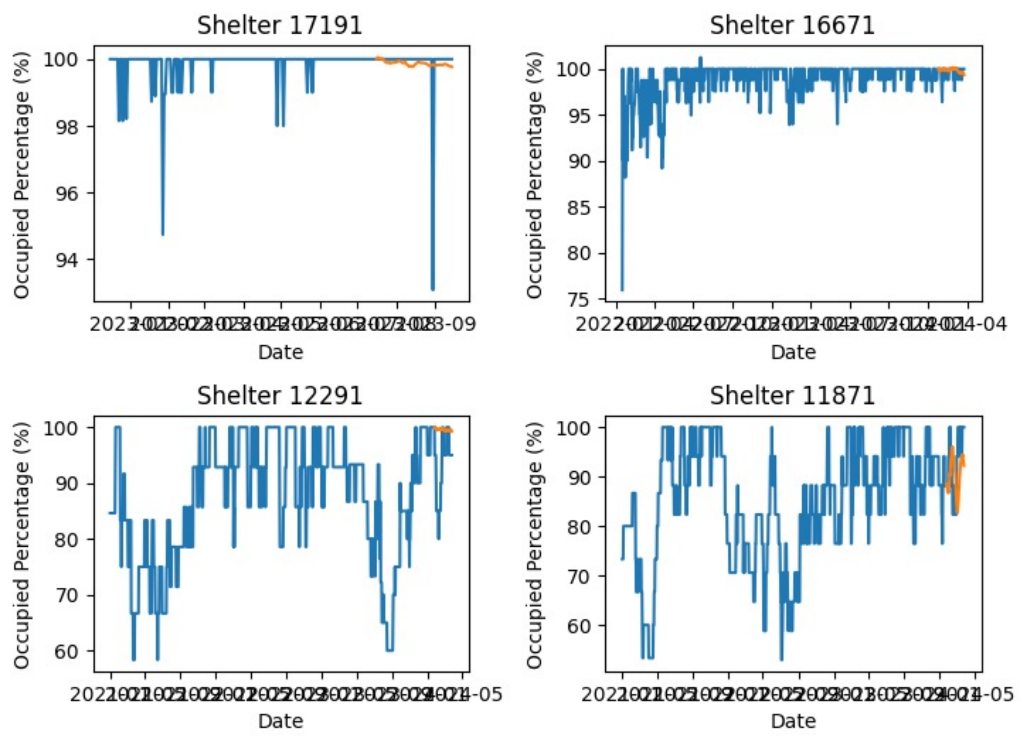


Figure 15: Model Prediction on Random Shelters

Incorporating Feature Occupancy Differences

- Initial correlation analysis highlighted trends in data linked to features that were not included in LSTM
- Decided to explore this further and understand how it could be incorporated into our LSTM Model

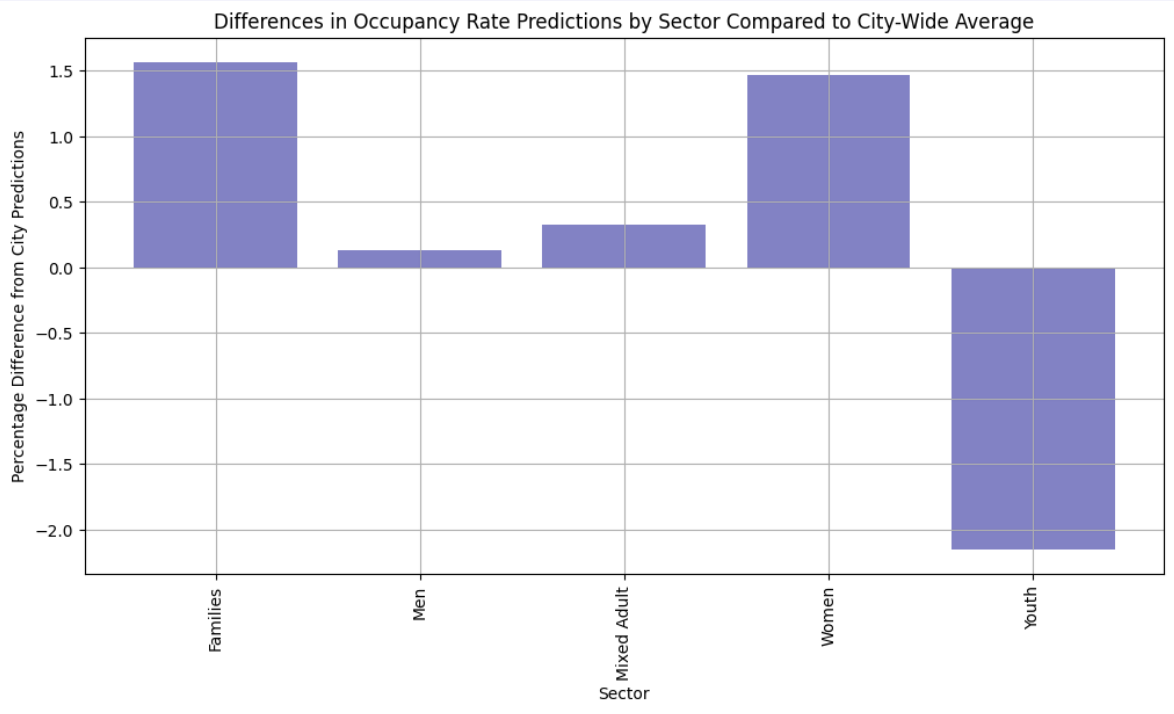


Figure 16: Plot of Average Difference Between City Average and Sector Average

Results

Model Accuracy

Model Accuracy: Adjusted vs. Non-adjusted

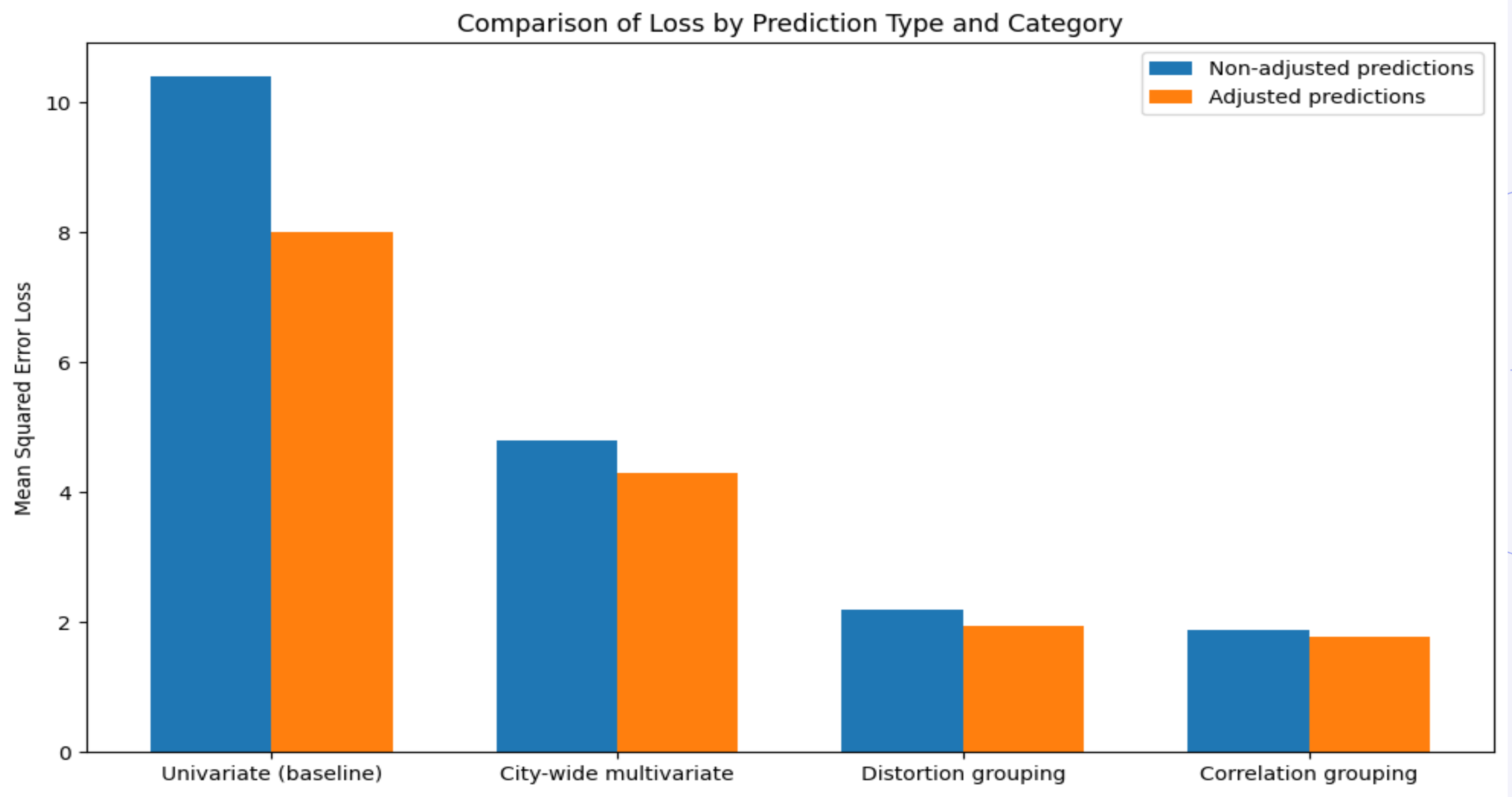


Figure 17: Mean Squared Error for Different Models

Performance Compared to Baseline Models

Random Forest Regression

Root Mean Squared Error: 7.04

Mean Absolute Error: 2.59

Autoregressive Model

Root Mean Squared Error: 4.12

Mean Absolute Error: 16.96

Prophet (5-Day Horizon)

Root Mean Squared Error: 13.66

Mean Absolute Error: 13.44

Adjusted Geospatial Multivariate LSTM

Distortions Grouping:

Root Mean Squared Error: 1.88

Mean Absolute Error: 2.22

Correlation Grouping:

Root Mean Squared Error: 1.88

Mean Absolute Error: 1.96

Final Product


ShelterSight
BOREALIS AI
DEMO

Discussion

Accomplishments and Learnings

Application of Academic Knowledge

Successfully transitioned theoretical knowledge from an academic setting to solve real-life challenges.

Advancing Technical Skills

Gained experience in specifically in building and tuning LSTM models for time-series forecasting.

Future Discussions

Geographic Expansion
Scale the model to other cities and regions, adapting the approach to different urban environments and regulatory contexts to widen the impact of the research

Real-Time Implementation
Develop a real-time forecasting system that dynamically updates predictions based on live data stream

Impact Analysis
Conduct longitudinal studies to measure the long-term impact of optimized shelter operations on homelessness alleviation

Conclusion

Questions

Thank You