MODELING THE IMPACT OF AIR POLLUTION ON ASTHMA-RELATED HOSPITALIZATIONS: A TIME SERIES APPROACH

Executive Summary

This case study applies **time series analysis** to evaluate the association between **air pollution levels** and **asthma-related hospital admissions** in a metropolitan city. Using secondary data for PM2.5 concentrations and daily hospitalization counts over 2 years, students conduct correlation analysis, lag effect evaluation, and simple forecasting using ARIMA. The study equips learners with quantitative tools to support evidence-based environmental health policies.

1. Introduction

Air pollution is a major public health hazard, especially for individuals with respiratory conditions such as asthma. Monitoring its impact using real-world data helps policymakers issue timely warnings and allocate resources effectively. This case bridges epidemiology and environmental science using data analytics.

2. Data Overview

- **Data Source**:
    - PM2.5 levels: Government air quality monitors
    - Asthma hospitalizations: City Health Department
- **Time Period**: January 2022 – December 2023
- **Observations**: 730 (daily records)

Variable	Description
Date	Daily timestamp
PM2.5_Concentration	µg/m³ (airborne fine particles)
Hospital_Admissions	Number of asthma-related cases

3. Descriptive Statistics

Statistic	PM2.5 (µg/m³)	Admissions (per day)
Mean	71.3	23.5
Max	187.4	51
Min	22.6	6
Std. Deviation	26.1	8.4

4. Correlation Analysis

- **Pearson Correlation (Lag-0)**:

$r=0.62, p<0.001 r = 0.62, \quad p < 0.001$

- **Lag-1 Correlation** (Pollution effect delayed by one day):

$r=0.68, p<0.001 r = 0.68, \quad p < 0.001$

Strong positive correlation indicates that increases in PM2.5 often precede increases in asthma admissions by one day.

5. Time Series Visualization

Figure 1: Dual-Line Plot

- **X-axis**: Date
- **Y1-axis**: PM2.5 concentration
- **Y2-axis**: Hospital admissions

Seasonal peaks are visible in winter months with simultaneous spikes in both variables.

6. Forecasting with ARIMA

- **Model Used**: ARIMA(1,1,1)
- **Dependent Variable**: Daily asthma-related admissions
- **Exogenous Variable**: PM2.5 levels (Lagged by 1 day)
- **Forecast Horizon**: 30 days

Date	Forecasted Admissions
Jan 1	28
Jan 2	31
Jan 3	33
…	…
Jan 30	22

MAPE (Mean Absolute Percentage Error) = 6.7%

7. Threshold Identification

PM2.5 Range (µg/m³)	Average Admissions	Risk Level
0–50	14	Low
51–100	25	Moderate
101–150	33	High
151+	45	Very High

8. Public Health Implications

Action Area	Recommendation
Alert System	Issue “Asthma Red Alerts” when PM2.5 >100
Hospital Readiness	Schedule extra beds/staff in winter
Prevention Messaging	Send warnings to vulnerable populations
Urban Planning	Monitor industrial emissions daily

9. Learning Outcomes for Students

- Use time series data in epidemiological research
- Apply correlation and lag analysis to environmental health issues
- Build basic ARIMA models for forecasting health events
- Communicate statistical results into health policy insights

10. Suggested Assignments

- Conduct a Poisson regression instead of ARIMA to account for count data
- Build an interactive dashboard using Excel or R Shiny for daily health alerts
- Write a 1-page policy brief recommending a pollution-linked response system

11. References

- WHO (2021). *Air Pollution and Child Health: Prescribing Clean Air*
- US EPA (2023). *Integrated Science Assessment for Particulate Matter*
- Harvard T.H. Chan School of Public Health – Air Quality Research
- Box, G. & Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*