Data Analytics Case Study 2

Part 2: Data Collection and Briefing Report

Prepared By: Group 3

Patel Gaurav (NF1005669)

Patel Harsh (NF1002720)

Patel Jimil (NF1004650)

Patel Rudra (NF1004756)

Course

Data Analytics Case Study 2

Term

Term II (Winter 2025)

Degree Program

Master of Data Analytics

Course Facilitator

Prof. Hany Osman

Dataset Overview

The dataset comprises **15,549 entries**, collected for detailed analysis of Logistics & Transportation under various PESTEL factors.

Research Objectives (Presented as Research Questions):

- **Political Factors**
  How do international trade regulations influence operational efficiency and costs in logistics and transportation?
- **Economic Factors**
  What is the impact of fuel price fluctuations on cost management within logistics and transportation operations?
- **Social Factors**
  How do shifts in consumer preferences influence the demand for logistics and transportation services?
- **Technological Factors**
  How does the adoption of AI and automation technologies affect operational efficiency in logistics and transportation?
- **Environmental Factors**
  What is the impact of logistics activities on environmental sustainability as indicated by the carbon footprint score?
- **Legal Factors**
  How do regulatory fines due to non-compliance impact financial stability and operational effectiveness in logistics and transportation?

Levels of Measurement and Rationale for Variable Selection in Logistics & Transportation

Factor Category	Variable Name	Source	Level of Measurement	Rationale
Political	Trade Regulations Impact	International trade reports and FedEx regulatory documentation	Nominal (High, Medium, Low)	Influence of trade policies on delivery capabilities
Political	Taxation Policy Impact	Governmental tax policy reports	Nominal (High, Medium, Low)	Financial impact on logistics operations

Political	Government Stability Score	World Bank & political risk assessment indexes	Interval (0-100)	Assessment of political stability
Economic	Fuel Price Impact	International/regional fuel price databases	Ratio (Numerical)	Affects operational costs and pricing strategies
Economic	Inflation Rate (%)	National economic reports & World Bank	Ratio (Numerical)	Indicator of economic stability affecting pricing and profits
Economic	Currency Exchange Rate	International financial markets	Ratio (Numerical)	Impacts international shipping costs
Social	Consumer Preference Shift	Market research & consumer trends	Nominal (Rising Demand, Stable, Declining Demand)	Influences service demand based on e-commerce trends
Social	Urbanization Level (%)	UN urbanization statistics and local census reports	Ratio (Numerical)	Impacts logistical complexity and infrastructure
Technological	Adoption of AI & Automation	Industry technology advancement reports	Ordinal (High, Medium, Low)	Competetive position in adopting new technologies
Technological	Internet Penetration Rate (%)	ITU and global internet usage reports	Ratio (Numerical)	Influences market growth potential via online purchase behaviors
Environmental	Carbon Footprint Score	Sustainability reports and third-party assessments	Ratio (Numerical)	Evaluates environmental impact and sustainability alignment
Environmental	Eco-Friendly Packaging Compliance	Internal compliance reports and sustainability audits	Nominal (Compliant, Non-Compliant)	Reflects adherence to environmental regulations

Legal	Compliance with Labor Laws	Labor regulatory bodies and internal compliance audits	Ordinal (High, Medium, Low)	Assessment of risk exposure related to labor regulations
Legal	Regulatory Fines ($)	Legal and regulatory filings	Ratio (Numerical)	Indicates efficiency and impact of regulatory compliance

Descriptive Statistics

Interpretation and Summary:

1. **Government Stability Score (Mean: 74.47, Median: 74, Range: 49, Std Dev: 14.48)**
    - Average political stability is relatively high (74.47), indicating stable political conditions in most regions where logistics and transportation services operate.
    - No outliers detected, showing consistent political stability across different areas.
2. **Fuel Price Impact (Mean: 3.24, Median: 3.24, Range: 3.50, Std Dev: 1.02)**
    - Fuel prices are moderately impactful on operations, with the average close to the median, suggesting predictable and stable fuel costs for logistics.
    - No extreme variability or outliers observed, indicating manageable fuel-related operational expenses.
3. **Inflation Rate (%) (Mean: 5.49, Median: 5.46, Range: 9, Std Dev: 2.58)**
    - Moderate inflation rates across markets with minimal variability imply predictable economic conditions affecting logistics and transportation.
    - Absence of outliers suggests uniform inflationary effects across regions.
4. **Currency Exchange Rate (Mean: 1.15, Median: 1.16, Range: 0.70, Std Dev: 0.20)**
    - Stable exchange rates with minimal fluctuations help mitigate currency-related risks in international logistics.
    - Uniform distribution without outliers indicates consistent international operational costs.
5. **Urbanization Level (%) (Mean: 72.42, Median: 72.43, Range: 45, Std Dev: 13.02)**
    - High urbanization levels indicate logistics and transportation services predominantly serve urbanized regions, simplifying infrastructure and operational planning.

- Consistent distribution without significant outliers implies stable infrastructure challenges.

6. **Internet Penetration Rate (%) (Mean: 79.69, Median: 79.76, Range: 39, Std Dev: 11.27)**
    - High average internet penetration indicates strong potential for growth in online logistics and transportation services.
    - Low variability highlights consistent market conditions favorable to digital and e-commerce logistics.
7. **Carbon Footprint Score (Mean: 110.38, Median: 110.92, Range: 179.95, Std Dev: 51.98)**
    - Moderate carbon footprint indicates there is considerable room for improvement in environmental sustainability within logistics operations.
    - Wide range highlights varied environmental impacts across regions, although no distinct outliers were identified.
8. **Regulatory Fines ($) (Mean: $24,995.40, Median: $25,056.95, Range: $49,991.04, Std Dev: $14,428.33)**
    - Regulatory fines indicate significant financial implications, pointing to varying compliance levels within the logistics sector.
    - High variability suggests substantial differences in compliance adherence across different operational regions, though extreme outliers were not identified.

Handling Missing Data

Upon thorough analysis of the dataset containing **15,549 records**, it was determined that there were **no missing values** across all variables. The dataset provided was already complete and pre-cleaned, eliminating the necessity for additional data-handling methods such as imputation or deletion.

Approach:

- Conducted a comprehensive missing value analysis using Pandas' isnull() function.
- Confirmed zero missing values for every variable included in the analysis.

Rationale:

- Since no missing data was present, no imputation methods or deletions were necessary. This ensures the integrity and reliability of the dataset and subsequent analyses.

Thus, the dataset is fully ready and optimal for robust analytical insights without any further data manipulation.

References

1. World Trade Organization. (n.d.). *Trade policies & tariffs impact on logistics operations*. Retrieved from https://www.wto.org
2. Organisation for Economic Co-operation and Development (OECD). (2025). *Inflation impact on transportation & logistics industry*. Retrieved from https://www.oecd.org
3. United Nations. (2025). *World urbanization prospects: Implications for transportation & infrastructure*. Retrieved from https://population.un.org/wup/
4. McKinsey & Company. (2024). *The rise of AI & automation in supply chain management*. Retrieved from https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights
5. World Customs Organization. (2024). *Trade regulations & compliance: Effects on cross-border logistics*. Retrieved from http://www.wcoomd.org

Data Source Citation:

Kamboj, P. (n.d.). *Logistics Data Containing Real World Data*. Kaggle.
Retrieved from https://www.kaggle.com/datasets/pushpitkamboj/logistics-data-containing-real-world-data

Google Collab Link (Data Cleaning & Preprocessing):

https://colab.research.google.com/drive/1zzus2DhYxP70X0V72qVRWth1liJaJfZh?usp=sharing

Conclusion

The analysis of the Logistics & Transportation sector across various PESTEL factors reveals a stable operating environment characterized by predictable political conditions, economic stability, and consistent urban infrastructure. However, the industry faces significant opportunities for improvement, particularly concerning environmental sustainability and regulatory compliance. Addressing these areas can enhance operational effectiveness, financial stability, and long-term competitiveness in the logistics and transportation industry.