Data Requirements for Predicting Sales of Protective Gear in Poland

A Comprehensive Guide

October 10, 2023

Contents

1.   Introduction

To build a machine learning model for predicting the sales of protective gear in Poland, it's crucial to gather and understand pertinent datasets. Given the provided datasets on sales data, inflation, and unemployment, this document outlines the data requirements.

2.   Data Requirements

2.1.   Historical Sales Data

- **Date/Time**: Monthly timestamps from January 2019 to September 2023.

- **Product ID**: A unique identifier for each product (if available).

- **Product Name**: Name and detailed description of the protective gear.

- **Classification and Subgroups**: Hierarchical categorization of the product.

- **Monthly Sales Quantity**: The number of items sold each month.

- **Total Sales Amount**: Cumulative sales for the period in the dataset.

2.2.   Economic Indicators

2.2.1   Inflation Rate

- **Date/Time**: Monthly timestamps from January 2019 to September 2023.

- **Inflation Rate**: Monthly inflation rate figures for Poland.

2.2.2   Unemployment Rate

- **Date/Time**: Monthly timestamps from January 2019 to September 2023.

- **Unemployment Rate**: Monthly unemployment rate figures for Poland.

3.   Additional Considerations

- **Data Consistency**: Ensure that the data across different sources is consistent in terms of units, scale, and format.

- **Data Integrity**: Ensure that the data is accurate, reliable, and free from errors or anomalies.

- **Frequency of Data Update**: For a more responsive model, more frequent data updates are preferable, but the provided data is monthly.

4.   Conclusion

The datasets provided will be crucial for building a machine learning model to predict sales of protective gear in Poland. The historical sales data will serve as the primary training data, while the economic indicators (inflation and unemployment rates) will be useful features to consider external economic influences on sales.