

Comprehensive Sales Data Report and Forecast

This report provides a comprehensive analysis of the cleaned sales data, covering key performance indicators, an exploratory data analysis, and a sales forecast for the upcoming year.

1. Dataset Overview

- The dataset contains **254,082 records** and **18 columns**.
 - The data spans a three-year period from **January 2018 to December 2020**.
 - There are a few missing values in the `Sales per Unit` and `Profit Margin (est)` columns.
-

2. Exploratory Data Analysis (EDA) & Key Findings

Sales Performance Over Time

Total sales show a clear **upward trend** over the three-year period, with noticeable seasonal peaks, indicating consistent growth.

Performance by Key Segments

- **Country:** **Germany** is the top-performing country, followed by Poland and Austria.
- **Channel:** The **Hospital** and **Pharmacy** channels dominate sales, with the Hospital channel having a slight edge.
- **Product Class:**
 - **Analgesics** generate the highest sales, totaling approximately **\$2.37 billion**.
 - **Antiseptics** are the second-most popular class, with sales of about **\$2.24 billion**.
 - The remaining product classes, in descending order of sales, are Mood Stabilizers, Antipyretics, Antibiotics, and Antimalarial.
- **Sales Team & Manager:**
 - **Delta**, managed by **Britanny Bold**, is the highest-performing sales team with total sales of approximately **\$3.64 billion**.
 - **Charlie**, managed by **Alisha Cordwell**, is the second-best team with sales of approximately **\$2.82 billion**.
 - **Bravo** (managed by **Tracy Banks**) and **Alfa** (managed by **James Goodwill**) follow, with total sales of **\$2.76 billion** and **\$2.58 billion**, respectively.

Key Correlations

- **Sales** are highly correlated with **Quantity**, indicating that higher sales figures are strongly linked to larger order quantities.
- A moderate correlation exists between **Sales** and **Price**.

3. Sales Forecast

A sales forecast was created for the next 12 months using a **SARIMA (Seasonal Autoregressive Integrated Moving Average)** model, which is well-suited for time series data with trends and seasonality. The sales data was decomposed into trend, seasonality, and residual components to inform the model.

Forecast for the Next 12 Months

The forecast continues the upward trend observed in the historical data, with an expected increase in sales over the next year.

Month	Forecasted Sales
2021-01-31	\$5,745,432.48
2021-02-28	\$6,323,668.49
2021-03-31	\$8,913,382.25
2021-04-30	\$7,725,123.51
2021-05-31	\$7,639,521.82
2021-06-30	\$8,764,249.77
2021-07-31	\$8,056,234.22
2021-08-31	\$8,111,353.79
2021-09-30	\$8,622,468.96
2021-10-31	\$9,076,822.45
2021-11-30	\$10,214,354.34
2021-12-31	\$9,923,431.13