Executive Summary and Project Report

Executive Summary

This project delivers a modular, interpretable forecasting pipeline tailored for a B2B dental supply company. Using anonymized SAP sales data, it applies time series models and regression analysis to forecast revenue at multiple levels: by year, month, product category, and customer. The objective is to provide data-driven insights for inventory planning, sales forecasting, and strategic decision-making.

Background & Business Motivation

In the dental supply industry, revenue trends are influenced by seasonality, product mix, and customer behavior. Accurate forecasting is essential to:

- Align inventory and production with expected demand
- Set informed sales and marketing targets
- Prevent stockouts and reduce inventory holding costs

The primary business goal is to establish a scalable forecasting tool that leverages historical data to improve operational efficiency and forward planning.

Data Overview

The project uses exports from SAP Business One, primarily:

- ORDR - Sales order headers
- RDR1 - Sales order line details
- ODLN - Delivery records
- OITM - Item master data

Merged datasets include:

- Transactional revenue (line totals)
- Order and delivery dates
- Product category information
- Customer identifiers (fully anonymized)

Note: All customer information (e.g., brand names, account codes) has been anonymized to comply with data privacy standards.

Methodology

The project employs several modeling approaches:

1. Time Series Modeling
- ARIMA and SARIMAX used for univariate and multivariate forecasting
- Differencing and stationarity checks via Augmented Dickey-Fuller (ADF) test
- Includes seasonal decomposition and confidence intervals

2. Linear and Regularized Regression
- OLS regression to fit annual and monthly revenue trends
- Polynomial Lasso and Ridge regression to capture nonlinear effects
- Feature engineering includes order volume, SKU count, and customer count

3. Segment-Based Forecasting
- Forecast revenue by product category and customer brand
- Includes Pareto chart logic to identify top-contributing categories
- Models are trained per group to capture segment-specific dynamics

All modeling steps are modular and reproducible, using functions defined in the src directory.

Outputs

- Forecasted annual revenue for future years (e.g., 2025)
- Monthly revenue predictions and diagnostics
- Segment-level forecasts by category and customer
- Visualizations of actual vs. predicted values with confidence bands
- Cumulative contribution plots (Pareto-style)

Business Value

The forecasting pipeline adds value in the following ways:

- Supports demand-driven inventory planning and procurement
- Enables dynamic sales planning and performance tracking
- Provides transparency for stakeholders through interpretable outputs
- Facilitates scenario analysis and budgeting with forward-looking estimates

By integrating domain knowledge with statistical rigor, the system helps guide data-informed decisions across operations and finance teams.