Data Analytics for Performance Management
Group Assignment
Instructions and Tips

Sasan Mansouri & Rauf Assad & Qiong Tang & Can Nalci

Case Assignment: Performance Management for an Ice Cream Van

Introduction

Lily and Suzy are two sisters who operate a family-owned ice cream van. Lily manages the van on **odd-numbered days of the month** (e.g., 1st, 3rd, 5th), while Suzy takes over on **even-numbered days** (e.g., 2nd, 4th, 6th). Each sister is responsible for daily operations, including deciding on procurement based on their sales estimates and remaining stock from the previous day. They are seeking a data-driven approach to optimize procurement and establish a fair profit-sharing mechanism.

As data analysts, your group is tasked with assisting Lily and Suzy in forecasting their procurement needs and evaluating their sales performance to determine an equitable profit-sharing strategy. Additionally, you will assess how varying profit margins between drinks and ice creams impact the profit distribution.

Objectives

1. **Procurement Forecasting**: Develop a predictive model using regression analysis to forecast the daily quantity of each product that needs to be stocked in the ice cream van.

2. **Performance Analysis**: Determine who outperforms in sales between Lily and Suzy, considering both total and average sales.

3. **Key Performance Indicators (KPIs)**: Propose ideal performance KPIs based on the external factors affecting sales.

4. **Profit-Sharing Proposal**: Propose a fair profit-sharing mechanism based on sales performance and product profitability.

5. **Profit Margin Impact**: Analyze how different profit margins for drinks and ice creams affect the profit-sharing model and procurement strategy.

Dataset Overview ([Link to the Dataset]())

You are provided with two datasets covering a complete year:

- **Sales Transactions Dataset**:

    - Individual scanned sales transactions for 10 products sold from the ice cream van.
    - Each transaction includes:
        * Date of sale
        * Product ID and name
        * Quantity sold
        * Price per unit

- **Weather and Temperature Dataset**:

  - Daily weather conditions and temperature ranges corresponding to the sales data.
  - Each record includes:
    * Date
    * Weather condition (e.g., sunny, cloudy, rainy)
    * Temperature range (e.g., low, medium, high)

- Alternatively, an **already-joined 'main' dataset** is available, combining sales transactions with weather and temperature data, and grouping weather and temperature into categories.

Business Context

Ice cream sales are influenced by:

- **Day of the Week**: Weekdays vs. weekends.

- **Month of the Year**: Seasonal variations.

- **Weather Conditions**: Sunny, cloudy, rainy days.

- **Temperature Ranges**: Low, medium, high temperatures.

Lily and Suzy need accurate procurement forecasts to prevent stockouts or overstocking, which can lead to lost sales or increased operational costs. They also want to establish a fair method to share profits based on individual performance and product profitability.

Assignment Tasks

Task 1: Data Exploration and Preprocessing

- **Data Integration**:

  - Merge the sales transactions with weather and temperature data based on the date.

- **Data Cleaning**:

  - Identify and handle missing or inconsistent data.

- **Feature Engineering**:

  - Convert categorical variables into numerical formats suitable for modeling.
  - Create a binary variable indicating who managed the van on each day (Lily or Suzy).

Task 2: Exploratory Data Analysis (EDA)

- Compute summary statistics for sales quantities of each product.

- Visualize sales trends over time and compare sales on weekdays vs. weekends.

- Examine correlations between sales and weather/temperature.

- Identify patterns and anomalies in the sales data.

Task 3: Procurement Forecasting Model

- **Model Selection**:
  - Use multiple linear regression to predict daily sales quantities for each product.
  - (Optional) Explore more sophisticated models if desired.

- **Model Development**:
  - Define the dependent variable (quantity sold) and independent variables (day of the week, month, weather conditions, temperature ranges).

- **Model Evaluation**:
  - Use metrics like R^2, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).
  - Perform residual analysis to check model assumptions.

- **Forecasting**:
  - Generate procurement forecasts for future scenarios based on the model.

Task 4: Sales Performance Analysis

- Calculate total and average daily sales for Lily and Suzy.
- Assess sales efficiency using relevant metrics.
- Conduct statistical tests (e.g., t-test) to compare performance.
- Interpret results and discuss possible reasons for differences.

Task 5: Key Performance Indicators (KPIs)

- Identify external factors affecting sales (e.g., weather, temperature).
- Propose ideal KPIs that Lily and Suzy can use to monitor and improve their performance.
- Explain how these KPIs are influenced by the external factors.

Task 6: Profit-Sharing Proposal

- Propose a profit-sharing ratio based on total sales contributions.
- Adjust the ratio considering efficiency metrics and KPIs.
- Provide recommendations to ensure fairness and motivation.

Task 7: Impact of Profit Margins

- Assume drinks have a larger profit margin than ice creams.
- Recalculate profit shares factoring in profit margins.
- Discuss how profit margins influence procurement and sales strategies.

Deliverables

1. **Submission Deadline**: Sunday 08/12/2024, 23:59

2. **Data Analytics Report**:

   - A technical report in the form of a Jupyter Notebook converted to PDF[1].
   - Include all data analysis, code, and visualizations.

3. **Presentation**:

   - An 8-10 minute presentation summarizing your findings.
   - Prepare slides to accompany your presentation.

4. **Appendices**:

   - Detailed calculations, statistical tests, and additional charts.
   - Code snippets or screenshots if necessary.

5. **Data Files**:

   - Any modified datasets used for analysis.
   - Scripts or notebooks containing your code.

Group Work Guidelines

- You will work in a **group of 4 or 5** members. Groups will self-organize via BrightSpace.

- Group members should divide the work equally.

- All group members will receive the same grade unless some members are reported as not contributing sufficiently. If some group members are found to have contributed less than their fair share, they might receive a lower grade than the rest of the group.

Presentation Schedule

- Presentation of the group assignment will take place in Week 5:

  - **Monday, 09/12/2024**
  - **Friday, 13/12/2024**

- Each group will have 8-10 minutes to present their findings in class.

Guidelines

- **Assumptions**: Clearly state any assumptions made during analysis.

- **Referencing**: Cite any external sources or references used.

- **Clarity**: Present findings in a clear, concise, and logical manner.

- **Professionalism**: Use professional language suitable for a business context.

---

[1]To create a PDF from a Jupyter Notebook Use Google Colab's Built-in Print to PDF. Open your notebook in Colab. After running the whole notebook, go to `File > Print` or press `Ctrl+P` (`Cmd+P` on Mac). In the print dialog, change the destination to `Save as PDF`. Click `Save` to download the PDF. **Alternatively,** when using your own PC instead of Colab, use the command: `jupyter nbconvert --to pdf your_notebook.ipynb`. Note that this process requires a LaTeX installation (e.g., TeX Live or MiKTeX). If LaTeX is not available, you can first convert the notebook to HTML using `jupyter nbconvert --to html your_notebook.ipynb` and then print the HTML to PDF.

Evaluation Criteria

- **Analytical Rigor**: Depth and correctness of analysis.

- **Model Accuracy**: Effectiveness of the forecasting model and justification of model choice.

- **Insights and Interpretation**: Ability to draw meaningful conclusions from the data.

- **Practical Recommendations**: Feasibility and usefulness of proposed strategies.

- **Communication**: Clarity of writing and quality of visualizations.

- **Technical Proficiency**: Appropriate use of analytical tools and techniques.

- **Teamwork**: Evidence of equal contribution from all group members.

Additional Information

- **Tools**: You may use any statistical software or programming language (e.g., Python with libraries like pandas, Seaborn, scikit-learn).

- **Collaboration**: This assignment should be completed as a group effort.

- **Scenario Enhancements (Optional)**:

    - Incorporate data on holidays or local events that may affect sales.
    - Analyze hypothetical customer satisfaction data if available.
    - Factor in operational costs related to transportation and refrigeration.
    - Consider industry trends in ice cream and beverage sales.

Conclusion

By completing this assignment, your group will demonstrate the ability to apply regression analysis techniques to real-world business problems involving forecasting, performance management, and strategic decision-making. Developing ideal KPIs based on external factors will enhance your understanding of how such factors influence business performance. The collaborative effort will also enhance teamwork skills and the ability to present complex data insights effectively.