Airline & Coca Cola Forecasting Assignment

Forecasting Assignment: Coca-Cola Prices and Airlines Passenger Data

Hello, class! We're excited to introduce our next assignment, which will give you hands-on experience with real-world time series forecasting. In this project, you'll be working with two distinct datasets: Coca-Cola prices and Airlines passenger data. This assignment is designed to challenge you and help you apply the concepts we've been discussing in class.

Assignment Objective

Your main task is to develop and compare various forecasting models for these two datasets. This exercise will deepen your understanding of time series analysis techniques and help you gain practical skills in predicting future values based on historical data.

What We Expect From You

1. **Thorough Data Exploration**: Begin by getting to know your data inside and out. Use descriptive statistics and visualizations to understand the patterns, trends, and potential seasonality in both datasets.
2. **Thoughtful Preprocessing**: Clean your data and handle any irregularities. We want to see how you approach issues like missing values or outliers. Also, think about creating meaningful features that could enhance your models.
3. **Rigorous Time Series Analysis**: Apply the techniques we've learned in class, such as creating lag plots and analyzing autocorrelation functions. We're particularly interested in how you assess and address stationarity in the data.
4. **Comprehensive Model Building**: You'll be implementing several models for each dataset, including linear, exponential, quadratic, and various seasonality models. We expect you to clearly explain the rationale behind each model choice.
5. **Critical Model Evaluation**: Don't just calculate the RMSE for each model - we want you to interpret these results. Why do you think certain models perform better than others for each dataset?
6. **Insightful Comparisons**: Compare the performance of your models across both datasets. Are there interesting differences in how the models perform on Coca-Cola prices versus Airlines passenger data? What might explain these differences?

7. **Clear Documentation**: Your code should be well-commented, and your analysis should be clearly explained. Remember, good data scientists need to communicate their findings effectively!
8. **Future Predictions**: Use your best-performing model to make future predictions for both datasets. Discuss the reliability and potential limitations of these forecasts.

Why This Assignment Matters

This project isn't just about crunching numbers. It's about developing your intuition for working with time series data. By the end of this assignment, you should have a better understanding of:

- How to approach different types of time series data
- The strengths and weaknesses of various forecasting models
- How to choose the most appropriate model for a given dataset
- The practical challenges of predicting future values

Remember, in the real world, you'll often be faced with messy, complex data. This assignment is designed to give you a taste of that complexity and help you develop strategies to tackle it.

Final Thoughts

We encourage you to be creative and thorough in your analysis. Don't hesitate to try approaches we haven't explicitly covered in class if you think they might yield interesting results. And as always, if you have any questions or run into any roadblocks, our door is open.

Good luck, and we're looking forward to seeing your innovative solutions!