Report: simple forecasting, decomposition, and evaluation

Please write a report on forecasting a variable (or multiple) of your choosing from a real dataset. You will select a dataset, perform some basic diagnostics and analysis, formulate a forecasting plan, and produce a forecast using time series regression techniques in R. The report should be around 5–8 pages, including figures.

---

1. Find a Dataset

- Choose a dataset with regular time intervals (e.g., daily, monthly, quarterly, or yearly).
- The dataset should contain at least 50 observations and ideally include additional variables that could be used as predictors in a regression model. Alternatively, you can merge in additional variables from an external source.
- Look for datasets related to topics that interest you!

Suggested Sources:

- FRED (Federal Reserve Economic Data)
- Kaggle
- TidyTuesday
- OECD Data
- World Bank Data
- US Bureau of Labor Statistics
- US Census
- US Government open data

2. Preliminary Analysis

- Create basic time series plots of your main variable(s).
- Generate diagnostic plots (e.g., ACF, seasonal plots, trend plots).
- Comment on any visible trends, seasonality, or irregularities.

Hint: Use `autoplot()`, `ggseasonplot()`, `ggsubseriesplot()` from the `fpp3` package.

3. Time Series Decomposition

- Apply classical decomposition or STL decomposition to your time series.
- Interpret the trend, seasonal, and remainder components.
- Discuss how decomposition informs your choice of forecasting method.

Hint: Use `decompose()` or `STL()` functions from fpp3.

4. Forecasting Methods

- Apply at least two different forecasting methods from the following:
  - Mean forecast
  - Naive forecast
  - Drift method
  - Seasonal naive forecast
- Justify your choice of methods based on the characteristics of your time series.
- Hint: Use `MEAN()`, `NAIVE()`, `SNAIVE()`, `RW(y ~ drift())`.

5. Transformations and Back-Transformations

- If your data exhibits non-constant variance or multiplicative seasonality, consider applying a transformation (e.g., log).
- Clearly explain the transformation and how you back-transform forecasts for interpretation.

Hint: Use `box_cox()`, `log()`, etc. as needed.

6. Prediction Intervals and Evaluation

- Include prediction intervals in your forecast plots.
- Evaluate forecast accuracy based on cross-validation using metrics such as:
  - RMSE
  - MAE
  - MAPE
- Compare the performance of your forecasting methods.

Hint: Use `accuracy()`.

Deliverables

Submit a report (PDF or HTML via R Markdown) that includes:

- A brief introduction to your dataset and forecasting goal.
- Clear plots and/or tables with captions.
  - Discussion of the point of each plot/table and reasoning behind analysis.
- Interpretation of results at each step.
- A conclusion summarizing your findings.