Report: forecasting with regression

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. Identify Potential Predictors

- Identify external variables (from your dataset or another source) that could help explain or predict your main time series.
- Justify your choices:
  - Why do you expect these variables to be predictive?
  - Is the relationship contemporaneous?

Hint: Use scatterplots or combined time plots to justify your reasoning.

---

4. Specify and Estimate a Time Series Regression Model

- Fit a time series regression model using your chosen predictors.
- Consider including trend and seasonality as regressors if not already captured.
- Check model diagnostics (residual plots, ACF of residuals, etc.).

Hint: Use `tslm()` from the `fpp3` package.

---

5. Forecasting

- Use your model to produce forecasts for a reasonable horizon (e.g., next 12 months).
- Plot the forecast and include prediction intervals.
- Discuss the forecast's plausibility and any limitations.

Hint: Use `forecast()` on your regression model.

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.
