Final report

Deliverable: A polished report (PDF or HTML via Quarto/R Markdown) of including figures and tables, plus an appendix with full code.

Objective: Select a real time series, build and compare **ARIMA**, **time series regression**, and **exponential smoothing (ETS)** models; evaluate using holdout and cross-validation; produce and discuss your forecasts.

1) Data Selection

- Choose a dataset with regular time intervals (e.g., daily, monthly, quarterly, or yearly).
- The dataset should contain at least 50 observations, your final dataset must include more than one time series variable, and you ideally you should look to forecast a non-seasonal variable.
- Briefly motivate the choice in your write-up.

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) Exploratory Analysis

Provide concise visuals and discuss:

- Time plot(s) with appropriate scales and transformations (if needed). Justify any transformations.
- Seasonal plots/subseries plots where applicable.
- ACF/PACF of the raw and differenced series; comment on trend/seasonality and stationarity cues.
- If seasonality is present, consider STL/classical decomposition to separate components and inform model choice.
- Suggested functions: autoplot(), ggseasonplot(), ggsubseriesplot(), STL()

3) Data Preparation & Transformations

- Check variance stability; apply log or Box–Cox transformation(s) as needed and justify.
- Clearly explain back-transformation of forecasts and intervals for interpretability.
- Document any data cleaning (missing values, calendar adjustments, outlier handling).

4) Evaluation Protocol

For the following model estimation steps, adhere to a two-part evaluation: - Fixed holdout: Reserve the last $h \times k$ periods (e.g., 12 months) as the test set; fit models on the training set; report RMSE/MAE/MAPE. - Cross-validation: Repeatedly generate multi-step forecasts at horizon h(e.g., $h = 1$ and $h = 6$ if relevant). Summarize average RMSE/MAE/MAPE across folds. - Report AICc for in-sample model comparisons.

5) Baseline Models

Fit and evaluate at least two baselines to benchmark gains:

- Mean, naïve (random walk), drift, seasonal naïve (as appropriate to frequency).
- Briefly explain why each baseline is plausible or implausible for your series.
- Relevant functions: MEAN(), NAIVE(), RW(y ~ drift()), SNAIVE().

6) ARIMA Models

Goal: Propose, estimate, and diagnose at least two reasonable ARIMA specifications.

- Identification: Use ACF/PACF on differenced series to motivate p,d,q (and P,D,Q if seasonal). Explain your differencing decisions.
- Estimation: Fit candidate models. Propose ARIMA models manually and justify your prpoposed models; don't just rely on auto-estimation.
- Diagnostics:
  - Residual plots and ACF of residuals.
  - Ljung–Box test for serial correlation (state hhh and degrees of freedom).
  - Parameter significance and invertibility/stationarity checks.
- Criteria & performance: Compare AICc and CV metrics.
- Forecasts: Produce forecasts with 80% & 95% intervals; comment on reasonableness.

7) Time Series Regression

Goal: Build a regression model that includes other predictors. - Specification: + Start with TSLM (trend and seasonal dummies as needed). + Add at least one external predictor with a clear causal or predictive rationale; consider lags as appropriate. - Autocorrelation handling: + Check residual ACF and discuss. - Diagnostics & evaluation: As in part 6, residual checks, information criteria (for the error model), and CV metrics. - Forecasts: Provide h-step forecasts and comment.

8) ETS Models

Goal: Fit at least two ETS variants and compare. - Specification: Consider additive vs. multiplicative error/trend/seasonal components (e.g., ETS(A,A,N), ETS(M,A,M)). - Selection: Use automatic ETS selection as one candidate and a hand-picked alternative based on decomposition results. - Diagnostics & evaluation: see 6 and 7. - Forecasts: Generate h-step forecasts with intervals and comment.

9) Comparison & Synthesis

- Present a summary table of in-sample (AICc) and out-of-sample metrics (holdout + CV).
- Compare interval performance.
- Discuss which model you would deploy and why, balancing accuracy, stability, interpretability, etc.
- Provide a sanity check: Are forecasts plausible relative to historical variability, seasonality, and context?