**PROCESS BOOK**

**Date of entry: 2025-05-20**

*What I’ve worked on:* Fetched one year of BTC/USDT daily OHLCV data from Binance using the ccxt library and loaded it into a pandas DataFrame.

*What problems I encountered*: Had to troubleshoot installation and version conflicts for ccxt; raw timestamps came back in milliseconds and needed conversion.

*What I learned:* How to configure and call ccxt’s fetch\_ohlcv function; convert Unix-ms timestamps to pandas DateTime; forward-fill any missing trading days.

**Date of entry: 2025-05-21**

*What I’ve worked on:* Implemented two forecasting methods—Facebook Prophet and ARIMA(5,1,0) via statsmodels—and generated 30-day forecasts.

*What problems I encountered:* Prophet required renaming columns to “ds”/“y”; ARIMA order selection had convergence warnings until I chose (5,1,0).

*What I learned:* Prophet’s data schema and changepoint priors; reading ACF/PACF plots to justify ARIMA parameters.

*Which resources did I use:* Prophet official quick-start, statsmodels ARIMA tutorial, StackOverflow examples, ChatGPT advice.

**Date of entry: 2025-05-22**

*What I’ve worked on:* Added Holt–Winters exponential smoothing (additive trend & seasonality) and coded the Theta method for forecasting.

*What problems I encountered:* Determining the correct seasonal period parameter for ExponentialSmoothing; finding a reliable Theta implementation.

*What I learned:* How to call statsmodels.tsa.holtwinters.ExponentialSmoothing with seasonal\_periods=365/12; the theory behind Theta’s two-line decomposition.

**Date of entry: 2025-05-23**

*What I’ve worked on:* Organized and merged each model’s 30-day forecast into a single DataFrame aligned by date.

*What problems I encountered:* Inconsistent index alignment and NaNs when concatenating different forecast series.

*What I learned:* How to use pandas.concat with axis=1 and DataFrame.reindex to align dates; fill gaps safely.

**Date of entry: 2025-05-24**

*What I’ve worked on:* Built a simple ensemble by averaging the four model forecasts and visualized actual vs. each model vs. ensemble with matplotlib.

*What problems I encountered:* Date formatting on the x-axis was cluttered; legend overlapped the plot.

*What I learned:* Use matplotlib.dates for cleaner tick formatting; place legends outside the plot area.

**Date of entry: 2025-05-25**

*What I’ve worked on:* Computed evaluation metrics for all forecasts and the naïve baseline.

*What problems I encountered:* Got mixed up on RMSE’s denominator; needed to sample errors at regular 4-day steps.

*What I learned:* Formulas for RMSE; how to slice a pandas Series by index positions for interval error reporting.

*Resources I used:* Stackoverflow and ChatGPT