|  |  |  |  |
| --- | --- | --- | --- |
| Date | Task | Who / How | Status |
| Apr 15 (Today) | ✅ Fork + clone repo 🧼 Review raw data together | Madelyn created the repo /Charles forked + cloned — schedule 30 min sync to review variables, structure, and missingness. Charles, check and add the weather data to the data folder. It might be better to work using branches on Github (we can look into that). |  |
| Apr 16–17 | 🔧 **Data wrangling & cleaning** (split task) | Each person takes half the variables or one dataset section. Share progress via GitHub + meet Apr 17 to review changes. |  |
| Apr 18 | 📊 EDA together (histograms, correlations, outliers) | Both of us can work on EDA plots and compare |  |
| Apr 19 | 🏗️ Feature engineering | Brainstorm together; then one person implements, the other tests/reviews |  |
| Apr 20 | 🧪 Finalize clean dataset + split data | Pair program this step. Decide on train/test ratio, stratification, and seed. Save clean dataset to our “data” folder |  |
| Apr 21–23 | 🤖 Train multiple models (RF, LR, XGBoost, etc.) | Split: each person tries 1-2 models, documents results. Meet Apr 23 to compare and pick best |  |
| Apr 24 | 📈 Model evaluation | Break |  |
| Apr 25–26 | 💻 Build Shiny app | Add titles, explanations, styling. Test it. Pair review. Optional: deploy to shinyapps.io |  |
| Apr 27 | 🧽 Polish Shiny app |
| Apr 28 | 📝 Write Quarto file | Combine + render as .html |  |
| Apr 29 | ✅ Final review + submission | Full review meeting. Test app + Quarto. Push (Madelyn) everything cleanly to GitHub repo for submission |  |