Exercise 1

Simon A. Broda

1. (a) Open the file `maunaloa.csv`; this is a famous data set used in machine learning. Make a time series plot.

   (b) Estimate a linear trend by regressing the `CO2` series on an intercept and a time trend. Hint: you can add a time trend to a dataframe `df` as follows.

   ```
   import statsmodels.tsa.api as tsa
   df = tsa.add_trend(df, trend="t")
   ```

   (c) Plot the data together with the estimated linear trend, and the residuals. An easy way to do this is

   ```
   import statsmodels.api as sm
   sm.graphics.plot_regress_exog(model, "trend");
   ```

   What do you notice?

   (d) Produce a forecast for Sept 1st, 2004 (one month after the sample ends), first manually using the fitted model

   $$\widehat{Y_t} = \widehat{\beta_0} + \widehat{\beta_1} t,$$

   then using Python.

   (e) Repeat Questions 1b through 1d, but using a quadratic trend.

   (f) Repeat Questions 1b through 1d, but using an exponential trend.

2. (a) Compute the 3rd order moving average of the `CO2` series for Feb 1st, 1964, both by hand and using Python. Hint: use a `rolling` object.

   (b) Estimate the trend with a 12 month moving average (12 months are necessary to cover a full cycle). Then plot the resulting trend estimate and the data together in a time series plot.

3. (a) Estimate a model with a linear trend and 12 monthly dummies (and no intercept) for the `CO2` series. Then, produce an (in-sample) forecast for August 1st, 2004, both by hand and using Python. Plot the data together with the estimated linear trend, and the residuals.

   (b) Same, but include an intercept. This will automatically remove one dummy.