**Name:** Your name here
**Due:** 2024/10/07

# Homework 4

Be sure to submit **both** the .pdf and .qmd file to Canvas by Monday, October 7th at 11:59 pm.

0. [5 pt] Please complete this Google Form to provide feedback on how the semester is going so far. The form is anonymous, so I cannot verify that you have actually completed the form. Please do it :( It is for your benefit as well as my own!

1. [1 pt] With whom did you work on this assignment?

2. [17 pt] We will focus on a data set describing weekly avocado sales volume and price in the United States between 2015 and 2018 for this question.

   a) [1 pt] Read the data in (naming it `avocado`) and filter to sales of conventional avocados in `Albany`.

   b) [2 pt] Create a `ts` object with the `AveragePrice` vector (called `avo_ts`) and plot the series.

   > **Tip**
   >
   > Be sure to pay attention to how the data set is arranged with respect to date. Additionally, **do not** specify an end date. This is one way to handle the fact that there are 53 Sundays in 2017 (omitting an end date forces R to treat the week that begins on 12/31/2017 as the first week of 2018).

   c) [2 pt] Describe the series in terms of trend and seasonality. Decomposing the series might help.

   d) [1 pt] Create a reduced version of the `avo_ts` time series, called `avo_ts_red`, that only spans 2015 to 2017.

   e) [1 pt] Fit an additive Holt-Winters model, called `avo_hw1`, to the reduced time series and allow R to estimate the smoothing parameters.

   f) [4 pt] Create an object, called `avo_hw1_pred`, that predicts the first 13 weeks of 2018 and include prediction intervals. Plot the original time series, fitted series, and forecasted series on the same plot. How well does the forecast predict the avocado prices from the complete time series?

g) [1 pt] Calculate the sum of squared differences between the forecasted time series and true value of the time series in 2018 (and print that value).

h) [5 pt] Find values for the smoothing parameters that result in better forecasts of avocado prices in 2018. Recreate the figure from part f for the new model, and print out the sum of squared differences between the forecasted and observed time series for the new model. What do you notice about the variability of the forecasted prices for the new model relative to the old model?

3. [6 pt] In class, we saw that the estimate of the non-stationary mean of the exponential smoother is

$$a_t = \alpha x_t + (1-\alpha)a_{t-1}$$

for $0 < \alpha < 1$. This formula is defined *recursively*, meaning that each term is a function of the previous term. We will encounter many time series models that are defined recursively, so it is helpful to fully understand what it is implied by this kind of model.

a) [1 pt] Show that $a_t$ may also be defined as $a_t = \alpha(x_t - a_{t-1}) + a_{t-1}$.

b) [4 pt] Show that $a_t = \alpha x_t + \alpha(1-\alpha)x_{t-1} + \alpha(1-\alpha)^2 x_{t-2} + ....$

c) [1 pt] Recall that $0 < \alpha < 1$. Comment on what the result in part b implies about how the weight associated with recent observations changes as we move backwards in time.

4. [5 pt] Show that

$$\text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \text{Cov}(X_i, Y_j)$$

.

You may use any (read: all) of the following results in your proof:

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$$

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

$$\sum_{i=1}^{n} X_i \sum_{j=1}^{m} Y_j = \sum_{i=1}^{n}\sum_{j=1}^{m} X_i Y_j$$