Instructions: In this exercise, you will work on data from the National Oceanic and Atmospheric Administration (NOAA) that documents the concentration of atmospheric carbon dioxide (in micromols/mol) from three different monitoring sites across the globe. The file containing these data is called bWeek10data.csv and will be supplied by your instructor. Sites 1 and 3 are in Alaska: Site 1 is Cold Bay, AK (in the Aleutian islands) and site 3 is Barrow, AK on the Chukchi Sea. Previously, Barrow was ice locked for much of the year, but ice has gradually been diminishing since regular measurements began in the 1970s. On June 18, 2004 the coastal sea ice disappeared within 24 hours, signaling a major change in seasonal ice behavior. Site 2 is located at Mauna Loa in Hawaii, a tropical location with small seasonal variations in weather.

There are three major tasks for the breakout: 1) reading in and visualizing the data; 2) conducting diagnostics of cyclicity and stationarity; and 3) conducting substantive analyses of the time series data. The research question is this: Describe what is happening at these three sites and highlight similarities and differences among them.

**Import and Review data**: Read the data into R, inspect the data, and plot the data.

1. Use the “Import Dataset” dialog to import the data into R. It is recommend to use read\_csv(). Assign the result to a data frame called “co2data.” Run View(), summary(), and str() on the data. Paste in the results below and mention the data type of co2data.
2. Create a correlation matrix of the data. Comment on the correlations among the sites.
3. Plot the data for each site. You can force a plot to be a univariate time series by using plot.ts(). For example, plot.ts(co2data$site1) creates a time series plot of Cold Bay, AK. Describe what you see in a brief comment. Are three sites similar or different from one another? Mention any differences you see in a comment below.
4. Convert the raw data into a multivariate time series object with the following command:  
   *co2series <- ts(co2data[,4:6], start=c(1978,10), frequency=12)*  
   This start date was chosen by examining the first line of data. The frequency of 12 was selected because these data were generally collected monthly. Run this command:  
   *plot(co2series)*
5. Use the decompose() function to break each site’s time series into its components. For example, the following command decomposes the data for Mauna Loa, HI:  
   *dec2 <- decompose(co2series[,"site2"])*
6. Plot the decomposition objects using the plot() command. Write a comment describing what you observe. Include in your commentary anything you note about differences between the three sites.

**Evaluate Stationarity**: Visualize and test cyclicity and stationarity of the three time-series.

1. Check the stationarity of the *random component* of each of the time series. The following three lines put the random components into a data frame, remove missing data, and convert to a times series:  
   *sites <- data.frame(dec1$random, dec2$random, dec3$random)  
   sites <- sites[complete.cases(sites),]  
   sites <- ts(sites,start=c(1978,10),frequency=12)*  
   Run summary(), str(), and plot() on the data frame. Comment on any differences among the three time series.
2. Next, examine the auto-correlation function (ACF) graph for each of the random components from the previous problem. Remember that the ACF provides an informal way of reviewing the time structure of a time series, particularly cycles and stationarity.
3. Finally, run the Augmented Dickey-Fuller test on each time series. The function, adf.test() is in the tseries package, which you will have to install and library. Add a comment on what the test shows for each time series.

**Conduct Changepoint Analysis**: Everything we have done so far has focused on getting the time series data ready to analyze. In this phase conduct and interpret changepoint analysis.

1. Run cor() on the random component of the decomposition in the data frame called “sites”. Paste in the correlation matrix below and interpret it. Add a comment contrasting these results with your correlation matrix from the original data (before the decomposition process; see Question 2).
2. Conduct a change point analysis of the variability of each time series. The cpt.var() function that you will need is in the “changepoint” package. Save the output object. Report the results in a comment.
3. If/when the cpt.var() procedure finds a change point in the variance of a time series, it will report on the index of the change point in the time series. Run plot on the results to reveal where the change took place.
4. Write a brief paragraph integrating the results of the analyses you conducted for Questions 10-12. Answer the research question: Describe what is happening at these three sites and highlight similarities and differences among them.