

w271: Homework 5 (Due: Week 6)

Professor Jeffrey Yau

Due: Before the Live Session of Week 6

Instructions (Please Read it Carefully!):

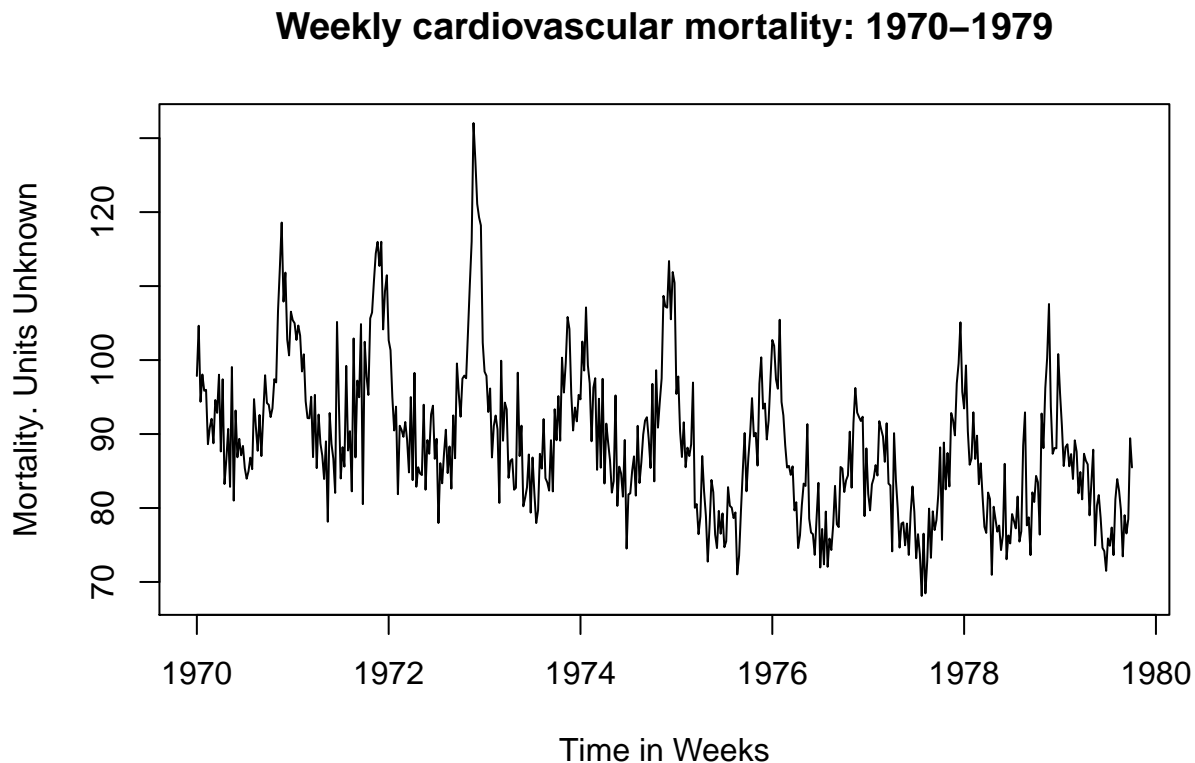
- **Page limit of the pdf report: None, but please be reasonable**
- Page setup:
- Use the following font size, margin, and linespace:
 - fontsize=11pt
 - margin=1in
 - line_spacing=single
- Submission:
 - Homework needs to be completed individually; this is not a group project.
 - Each student submits his/her homework to the course github repo by the deadline; submission and revision made after the deadline will not be graded
 - Submit 2 files:
 1. A pdf file that details your answers. Include all the R codes used to produce the answers. *Please do not suppress the codes in your pdf file.*
 2. R markdown file used to produce the pdf file
 - Use the following file-naming convention; fail to do so will receive 10% reduction in the grade:
 - * StudentFirstNameLastName_HWNumber.fileExtension
 - * For example, if the student's name is Kyle Cartman for homework 1, name your files as
 - KyleCartman_HW1.Rmd
 - KyleCartman_HW1.pdf
 - Although it sounds obvious, please write your name on page 1 of your pdf and Rmd files.
 - For statistical methods that we cover in this course, use only the R libraries and functions that are covered in this course. If you use libraries and functions for statistical modeling that we have not covered, you have to (1) provide an explanation of why such libraries and functions are used instead and (2) reference to the library documentation. **Lacking the explanation and reference to the documentation will result in a score of zero for the corresponding question.** For data wrangling and data visualization, you are free to use other libraries, such as dplyr, ggplot2, etc.

- For mathematical formulae, type them in your R markdown file. **Do not write them on a piece of paper, snap a photo, and either insert the image file or submit the image file separately. Doing so will receive a 0 for that whole question.**
- Students are expected to act with regards to UC Berkeley Academic Integrity.

```
rm(list = ls())  
library(astsa)
```

```
## Warning: package 'astsa' was built under R version 3.4.3
```

```
plot(cmort, xlab= "Time in Weeks", ylab="Mortality. Units Unknown")  
title(main="Weekly cardiovascular mortality: 1970-1979")
```



1. Conduct the EDA of the weekly cardiovascular mortality time series.
2. What features do you notice of the weekly cardiovascular mortality time-series plot?
3. Do you think that it is stationary in the mean? In the variance?
4. What pieces of information did you use from your EDA to arrive at your conclusion?
5. Do you find any evidence that there is a dependency structure in this time series data? Please explain.
6. What is the difference between strict and weak stationarity?
7. What is the difference between an acf and pacf plot?
8. (Open-ended question) Give two examples of questions people in your industry might ask that, based on what you learn in the async lecture, you think can be addressed using time-series analysis.