w271: Homework 7 (Due: Week 8)

Professor Jeffrey Yau

Due: Before the Live Session of Week 8

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 convensation; 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 sumbit 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.

In this homework, you are asked to use `quantmod` to get a time sereis `HOUST` from the Federal website, conduct Time Series EDA, examine seasonality, develop a model that can capture both trend and seasonality in the series, and plot the observed vs fitted value, and use the model to make a 12-step ahead forecast.

```
#install.packages("quantmod") #if you have not installed quantmod
library(quantmod)
```

```
## Loading required package: xts
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: TTR
```

```
## Version 0.4-0 included new data defaults. See ?getSymbols.
```

```
# Use Quantmod to load data
HOUST = getSymbols('HOUST', src='FRED', auto.assign = F)
```

```
## 'getSymbols' currently uses auto.assign=TRUE by default, but will
## use auto.assign=FALSE in 0.5-0. You will still be able to use
## 'loadSymbols' to automatically load data. getOption("getSymbols.env")
## and getOption("getSymbols.auto.assign") will still be checked for
## alternate defaults.
##
## This message is shown once per session and may be disabled by setting
## options("getSymbols.warning4.0"=FALSE). See ?getSymbols for details.
```

```
## Warning in strptime(xx, f <- "%Y-%m-%d", tz = "GMT"): unknown timezone
## 'zone/tz/2018e.1.0/zoneinfo/America/New_York'
```

```
str(HOUST)
```

```
## An 'xts' object on 1959-01-01/2018-09-01 containing:
##   Data: int [1:717, 1] 1657 1667 1620 1590 1498 1503 1547 1430 1540 1355 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr "HOUST"
##   Indexed by objects of class: [Date] TZ: UTC
##   xts Attributes:
## List of 2
##  $ src    : chr "FRED"
##  $ updated: POSIXct[1:1], format: "2018-10-23 12:51:03"
```

```
head(HOUST)
```

```
##            HOUST
```

```
## 1959-01-01  1657
## 1959-02-01  1667
## 1959-03-01  1620
## 1959-04-01  1590
## 1959-05-01  1498
## 1959-06-01  1503
```

```r
tail(HOUST)
```

```
##            HOUST
## 2018-04-01  1276
## 2018-05-01  1329
## 2018-06-01  1177
## 2018-07-01  1184
## 2018-08-01  1268
## 2018-09-01  1201
```

```r
house.starts = ts(HOUST, frequency = 12, start = c(1959,1))
# subset your time series to one starting in Januagy 2010
#house.starts = window()

# Examine your data
# YOUR CODE TO BE HERE

# Examine seasonality
# YOUR CODE TO BE HERE

# Estimate a model with trend, seasonlity, or both
# YOUR CODE TO BE HERE

# Plot the observed and fitted values

# Make a 12-step ahead (out-of-sample) forecast
# That is, forecast from 2018-10-01 to 2019-09-01
```