

## STAT 4363: Course Project Guidelines

Students will select a time series dataset and apply the methods learned in the class to address important questions concerning that dataset. The dataset should not be one of the data examples considered and discussed during the course. The project will be a 4- to 7-page report and this will be due **Wednesday, May 3, 2023** on Canvas. Graphics and visualizations should be part of the paper and will count in the page count. Students are allowed to form groups of no more than two on this project.

## 1 Structuring your project report

The project should include the following sections:

- Project title and your name(s)
- Introduction (1/2-1 page)
  1. Explain the relevant background of the dataset including the data source.
  2. Provide a plot of the time series.
  3. Discuss an important question(s) concerning the data you wish to address.
- Statistical Analysis (3-4 pages)
  1. Describe the characteristics of the time series with relevant visual representations such as ACF, PACF, seasonal/trend component plot, periodogram, spectral density etc...
  2. Discuss any necessary data transforms.
  3. Describe what methods/techniques are appropriate for the data. Follow the notation from the notes/textbook when describing the models.
  4. Present the results and explain how the analysis helped in addressing the questions raised on this dataset.
- Conclusion (1/2 page)
  1. Summarize the dataset and the statistical analysis.
  2. Comment on how the analysis helped address the questions on the dataset that were raised.

## 2 Time series data sources

1. R packages such as `datasets`, `fma`, `forecast`, `expsmooth`, `fpp`, `AER`, `Ecdat`. This list is not exhaustive.
2. UCI Machine Learning Repository. See <https://archive.ics.uci.edu/ml/datasets.php>
3. Economic time series available through the `quantmod` package in R.

```
# quantmod example code #

# See these links for other economic variables #
# https://fred.stlouisfed.org/series/GFDEBTN
# https://data.nasdaq.com/data/FRED-federal-reserve-economic-data

# Federal debt (measured in millions of dollars)
# From the QUANTMOD R package #
econ.var = "GFDEBTN"
getSymbols( econ.var , src="FRED" )
```

```
xt = GFDEBTN
```

```
xt = xt["19900101/20210101"] # Subset to 1990-2021
```

4. Financial time series available through the `quantmod` package in R.

```
# quantmod example code #
```

```
# Closing prices and returns of Exxon Mobile #
```

```
library(quantmod)
```

```
library(astsa)
```

```
start.date = "2008-01-30"
```

```
end.date = "2009-05-30"
```

```
# company ticker
```

```
tkr.s = "XOM"
```

```
# For other company tickers see https://www.slickcharts.com/sp500
```

```
xt = getSymbols( tkr.s , from=start.date, to = end.date , src="yahoo" , env = NULL)
```

```
xt = xt[,6] # Column 6 for daily adjusted closing price
```

```
yt = diff(log(xt)) # financial returns time series
```

5. functional MRI time series data. See [preprocessed-connectomes-project.org/abide/](http://preprocessed-connectomes-project.org/abide/) for a description.
6. EEG time series from the Brain-Computer Interface Competition. See <http://www.bbci.de/competition/iv/> for a description.