

**Department of Statistical Sciences**  
**Time Series Analysis**  
**STA457H1 S-LEC5101**  
**STA457H1/STA2202H S-LEC0101 & STA457H1S LEC2001**  
**Final Project, Winter 2022**  
**Due: April 17th, 2022 at 11.59 pm EST**

---

The project will involve analyzing time-series data and writing a brief report. First, you will need to choose a suitable time series. You can select a data set either from the “astsa” library that we used for this course <https://cran.r-project.org/web/packages/astsa/astsa.pdf> or any other resources. Choose a time series that

- interests you,
- is not too small it should have more than 100 observations,
- can easily be made approximately stationary
- is a different time series chosen from the lecture notes, quizzes, and assignments.

**Data Analytic Strategies:**

- Perform adequate exploratory analysis of the data and provide a concise presentation of the results:
  - exploring the time series plot and examining the properties of a stationary process
  - apply some transformations to convert as a stationary process
- Identifying the dependence orders of an ARIMA or SARIMA model and propose at least two models.
- The model building used to address the problem of interest:
  - estimate the parameters for the proposed models
  - provide precise interpretations of the parameters in your model in the context of the scientific problem
  - test the significance of the parameter estimates

- Perform all necessary diagnostics for the proposed models.
- Model Selection:
  - if you have two or more models that satisfy the models' assumptions, you can pick the best model based on model selection criteria.
- Forecasting:
  - after deciding on an appropriate model, forecast the data into the future ten-time periods ahead
  - give the 95% prediction intervals for each of the ten forecasts
  - comment
- Spectral Analysis:
  - perform a periodogram analysis to identify the first three predominant periods
  - obtain the confidence intervals for the identified periods
  - interpret your findings

**General Guidelines on the Report:(Total 50 pts)**

Your final analysis should be written in latex and presented in the PDF form of a brief report maximum of ten double-spaced pages (font size 12), including relevant tables and figures. Your report should be structured as follows:

- Abstract - A brief summary of your primary findings. **(5 pts)**
  - Background
  - Methods
  - Results
  - Conclusions
  - Keywords
- Introduction - A brief introduction/motivation to the problem at hand, relevant details about the data, additional relevant scientific information from searching the web, for example, and what is to be addressed. **(5 pts)**

- Motivation to the problem .
  - Relevant details about the data.
  - Cite relevant literature to support your explanation.
  - Purpose of the report.
  - Conclude the introductory paragraph with the report statement.
- Statistical Methods - A discussion and justification of the methods you have used to analyze the data and how you went about analyzing the data. Don't forget to describe in some detail how and why the particular model was selected. **(10 pts)**
    - Exploring the time series plot.
    - Examining the properties of a stationary process.
    - Apply some transformations or differencing to convert as a stationary process.
    - Identifying the dependence orders of an ARIMA or SARIMA model and propose at least two models.
- Results - A presentation of the results of your finding from data analytic strategies. **(25 pts)**
    - Estimate the parameters for the proposed models.
    - Provide precise interpretations of the parameters in your model in the context of the scientific problem.
    - Test the significance of the parameter estimates.
    - Perform all necessary diagnostics for the proposed models.
    - Model Selection.
    - Forecast the data into the future ten-time periods ahead and give the 95% prediction intervals for each of the ten forecasts.
    - Perform a spectral analysis to identify the first three predominant periods and obtain the confidence intervals for the identified periods.
- Discussion - A synopsis of your findings and discuss the limitations your model may suffer from. Present the conclusions in terms that non-statisticians will understand. Outline any limitations of your findings, and possibly suggest directions for further research. **(5 pts)**

**\*\*Graduate students will be required to submit a report and 5 minutes recorded presentation for their final project. (Presentation worth 20 pts)**

- Introduction/Opening (10 pts)
  - Objective and motivation of your project
  - What are you trying to accomplish in your project?
  - What interests you about your topic?
  - Why do you think your topic is important or interesting?
  - Give your listeners a road map of your presentation
- Body of Presentation (5 pts)
  - Statistical Methods
  - Results
- Conclusion/Closing (5 pts)
  - Summarize your major findings on your project
  - Outline any limitations of your findings, and possibly suggest directions for further research
- The written final report (maximum 10 pages including figures, summary tables, double spaced (font size 12) is due April 17th, 2021 at 11.59 pm EST. You can include Bibliography as an 11th page if you want.
- Upload your pdf version of the final project report by clicking the “Final Project” submission link under the week 13 module. Please don’t include any R code in the report, but you can upload the rmd version of the R code as a supporting document.
- Late project will be accepted but subject to a 20% penalty per day late. Late submissions will not be allowed beyond 48 hours of the due date.
- The written final report with more than ten pages will be accepted but subject to a 10% penalty per page.

- This final project must be completed independently. Please ensure you all follow academic integrity, including plagiarism, falsifying information, submitting work completed by someone else and unauthorized collaboration and sharing test questions or answers. If I confirmed any accused of cheating, the students involved would receive a 0% for the course (not just the Final Project).