# Data Analysis Exam 2 (STA 3001)
May 25, 2021

**Instruction**

- This exam aims to perform time series data analysis using computer and worth 50% of your course grade. For a given set of data, you are required to submit a report on the following questions. The clarity of your report will also be a part of credit.
- You are allowed to use any resources on your own, however, you are not permitted to discuss with anyone. The report should be your own work. Otherwise, it may subject to serious violation of the student code of conduct. By submitting your report online, it is assumed that you agree with the following pledge;

  **Pledge**: *I have neither given nor received any unauthorized aid during this exam.*
- The report should be written in English, and keep in mind that the clarity of your report is also a part of the credit. Remind that you can submit your report over icampus with pdf file format ONLY.

1. We will consider the (mysterious) daily time series data of length 1993 collected from Jan 1, 2005 to June 16, 2010. Perform the following analysis:
   (a) Time plot, correlograms (ACF) and discuss key features of the data.
   (b) Is it stationary? Include your evidence.
   (c) Find (your) best "regression + stationary errors" model. You need to include reasonings for your selection.
   (d) Find (your) best SARIMA model. You need to include reasonings for your selection.
   (e) Forecast the next 4 quarters with 95% prediction interval for **both** models (c) and (d) you selected. Use two decimal places (ex, 1.23) in your report. Report them as the table in the below:

| | | June 17, 2010 | June 18, 2010 | June 19, 2010 | June 20, 2010 |
|---|---|---|---|---|---|
| Model (c) | Point Forecast<br>95% PI | | | | |
| Model (d) | Point Forecast<br>95% PI | | | | |

   (f) Which one do you prefer (c) or (d), and why? If you have better model than models in (c) & (d), you can describe your own model here with your rational.
   (g) Write down summary (no longer than 1/2 page) on your data analysis result.
   (h) Attach R (or other softwares you used) code you have used in this analysis. Reduce font size/margin for the shortness sake. If the result cannot be reproduced with the attached code, the result cannot be trusted.

**Don't forget to write down your NAME, STUDENT ID in your report!**