

# Time-series Analysis Midterm

Supervisor: Ionut Florescu

Author: Yuxuan Xia

Email: yxia16@stevens.edu

I would highly recommend one to open the HTML file if ipython is not available in your PC, since it reserved all the merits of the ipython notebook.

## Notes:

---

- **About experiments:** Part i) ii) in problem 2(b) are not directly related to the question, however, it helps me to build up the basic sense of seasonal model. One can skip these two sections and directly switch to the part iii).
- **About seasonal modeling:** In section iii) of the problem 2(b), I tried to use SARIMA model in statsmodels library in Python. But it has some deadly defect (see explanation in my answer). When the seasonal period greater than 50, the computer just can't work it out. I tried both in my Windows system and Unix environment, sadly, all crashed. Then, I tried in R, surprisingly it performed well. It seems that this newly added model are not robust enough in Python. However, there are still 2 ways to bypass this problem.
  - Use a relatively small period, say seasonal\_period = 5. I implemented in this way in problem 2
  - Use the reduced ARIMA model. I first calculated the seasonal first difference of the raw data and then train it in ARIMA model. From the ARIMA model we calibrated, one can obtain the predictions under the seasonal difference. At last, I inverse the seasonal difference and got the prediction of the raw data (returns). In problem 3, I did in this way.
- **About Fitted QQ-plot:** In problem 4, I used a lot of qq-plots to illustrate the distribution. I find it has a fitting function and one can fit the data to the determined distribution (e.g. norm, student-t). It is so powerful and practical in data visualization. By this tool, I bypassed (fitted) the drift and sigma parameters, so the only thing mattered is the type of distribution itself. Please see my code and enjoy the beautiful(linear) qq-plots.