

# Weekly summary

## Monday, August 19

Worked at Sigma from 9AM to 5PM. I had a meeting with Johan and Markus regarding which data to use and how to progress. The following data is similar to the data that they usually encounter:

<https://github.com/yzhao062/anomaly-detection-resources#31-multivariate-data>

<https://github.com/khundman/telemanom>

Furthermore, I researched a GAN method for multivariate time series [Li2019a]. This needs to be investigated further and an implementation made.

Started to make the main document in Overleaf and sync it with Mendeley.

## Tuesday, August 20

Arrived at office at 9 and attended a meeting with the group. Before lunch I continued with making the Overleaf mechanism automated, I developed the note taking mechanism and synched Overleaf with Github for backup.

The note taking mechanism made will work like this. Weekly reports should be filed in PDF format while writing a daily report in the end of each working day. Ideas and interesting articles found can be linked to. Evaluated different programs for project management, such as making a GANTT-scheme and using task management. I will MS Project or Excel, although I had some problems getting MS Project to work. Finally got it installed after reinstalling with Click-to-go instead of Windows Installer.

Found article about anomaly detection using extreme value analysis. Look at later! Read about version control.

## Wednesday, August 21

Started working at 12 in Lund after training in the morning. MS Project seems to deliver too much details and Excel will work fine for the timeline. Developed a timeline based on GANTT in Excel and started to fill it up.

Worked on goal document. Still have to decide how if GAN should be included and if time-frequency analysis is relevant. ARIMA is out for the moment since it will be too much work.

Learned about conda. You should always use an environment for the current release. Never work in the root environment! Learned a bit about Git and cloned the repository to the local disc. Installed git from HomeBrew and run the Jupyter Notebook from the terminal. Everything is now synced and should be done on the Windows as well. To be done at that computer: 1. Install Anaconda correctly in Home folder (Already done?) 2. Install Git 3. Clone repository to local disc Reorganized the folders so that

everything is on GitHub. Three categories: 1. Code 2. Management (project plan, timeline and research diary) 3. Report

## Thursday, August 22

Worked from home in the morning. Cheat sheets were downloaded for all things that it could be used for. Telemanom from GitHub was added as a submodule to the repository, this is what you should do when adding others work.

Finally found out how to read find the data, it's of course in column 0. Read carefully and think! So, I should use this data to develop my methods. Another idea is to test my methods towards Numentas Anomaly Benchmark (NAB).

### What I've learnt:

Although .csv is used as industry standard for data files, it is more storage friendly and takes a lot less time to use .npz for data files.

<https://towardsdatascience.com/what-is-npy-files-and-why-you-should-use-them-603373c78883>

A temporal database stores data relating to time instances

## Friday, August 23

Use Jupyter as primary IDE for exploring, if you want to debug use Spyder and if you want to write ordinary python scripts use Atom!

Delivered email with goal document, data and example layout to Maria. Explored the data but seems to be hard to find different how they are related in time. Started to put things into the PDF. In the introduction things from [Aggarwal2013a] were put in.