

Tutorial 08 Dashboards

This tutorial consists of just one data exploration and uses data that is relatively similar to the data in the lecture notes and examples.

I recommend you use Google Colab for this tutorial, for reasons explained in the lecture. So adapt the code in the Colab folder.

It is possible to use Jupyter **notebook** on your own machine (although not Jupyter **lab**), but I don't think you will be able to configure the libraries correctly to use Jupyter on the University computers.

Naming: For the exploration you should create a Python notebook and save it when you have finished. You should name the notebook Tut08.ipynb.

Structure: Every numbered item in the exploration should have a code section and a markdown section underneath where you discuss your findings. There should also be a code section at the top of the notebook with the imports.

Exploration

The company who supplied the Products data in the lecture notes also want a dashboard for their website. This should be relatively simple ... provided you are up to date with other tutorial material.

1. Copy the final dashboard example from today's download and modify the top section of code so that it reads the daily page hits data, rather than the daily product sales. Also modify the summary data so that it contains the other data about the webpages. Note: you can do all of this as a copy and paste from your previous tutorial work.
2. There are too many pages for the dashboard to handle easily so modify the Time Series section so that the user can switch between high volume and medium volume pages. [Hint: you will need to change which columns are selected from the data and also which values are shown on the dropdown.]
3. The entire dashboard should work now – test it by running all cells. However, you will find that the heatmap is illegible (it can't really cope with 167 x 167 page correlations). Also the bubble plot has a lot of very small values. To deal with this, restrict the entire dataset to just high and medium volume pages.

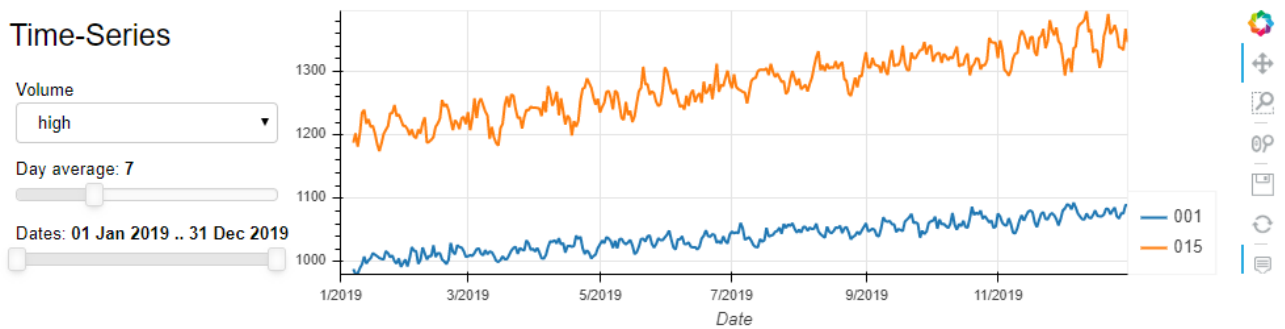
Hint: the place to do this is after all the data is read in and the summary data is constructed, so at the end of the first code section; once you have selected the column names you want, you can restrict the dataframes with

```
data = data[selected]
summary_data = summary_data.loc[selected]
```

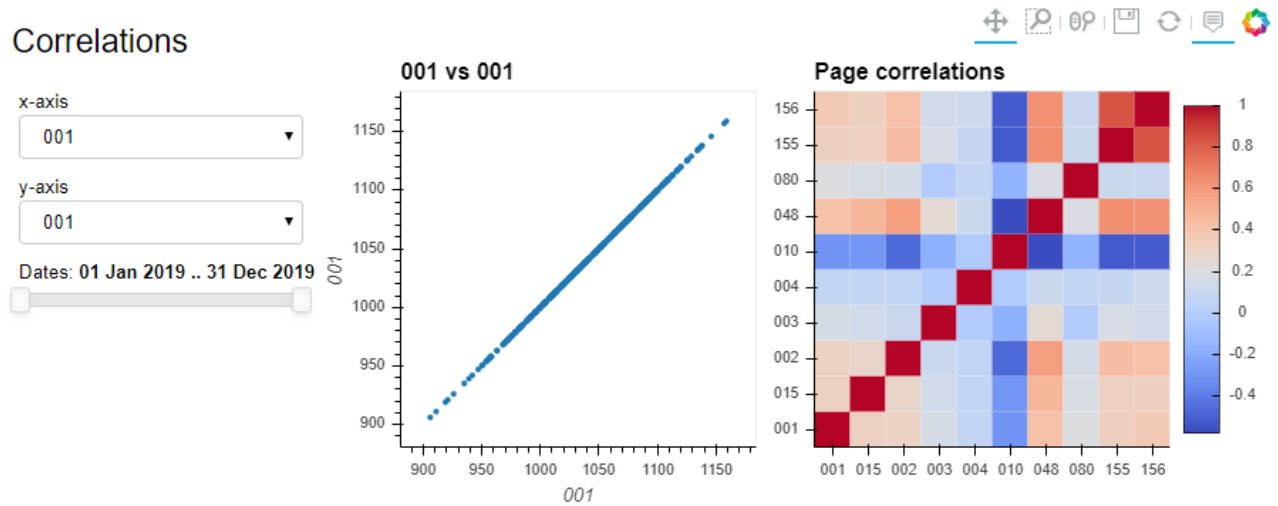
4. Finally test the dashboard and make sure all the titles, labels, etc make sense (it shouldn't say Products anywhere). Then write a short section in markdown about how you might modify it to include all pages rather than just high and medium volume – what extra widget(s) could you add to make this manageable?

Your completed dashboard should contain the following panels:

Time-Series



Correlations



Variables

