Algothon 2018 Quandl and Kaggle - Quick Start

Quandl

Quandl is kindly providing us with a premium package with varied information including core US fundamentals, their social media analytics and web sentiment datasets. Your aim is to use this data to create robust models able to trade using only the provided information. This model can later be tested in Quantopian.
Quandl supports Python, R and Excel and you can also download the data to your computer (even though this could take too much space). This guide will help you get started with Quandl and Python, if you want to use another programming language you will have to check the documentation at: https://docs.quandl.com.

First, install Quandl in your computer. If you are using python, go to your terminal or command line and type:
pip install quandl

or alternatively:
pip3 install quandl

If you are using anaconda you may want to type the following instead:
conda install quandl

Now you are ready to go, to set it up in your program or notebook, just import it using
import quandl

Then you will have to set you API key, so that you have access to the premium data that is provided for the competition. In order to do this, in your program write:
quandl.ApiConfig.api_key = "YOUR API KEY"

You will be given an API key once you sign up to quandl. Now you should be fully set up, it is extremely easy to access to the data now, as an example try the following:
data = quandl.get("WIKI/AAPL", rows = 5)

This will place in the variable data the last five days of Apple.inc stock price (Open, close, High, Low and some more details). We recommend using pandas as a way of treating the data. You can also get the data in a csv file using the API directly with pandas doing the following:
pd.read_csv("https://www.quandl.com/api/v3/datasets/WIKI/FB/data.csv?api_key=YOUR API KEY").head(5)

If the text above looks scary, don't worry, let me break it down for you. First, we are accessing the quandl website and using the api to retrieve information, that is why by the end of the URL you must give your API-key, to ensure you have access to that data. Then everything else is simple.

Replace "WIKI/FB" with any dataset that you may want (check your email for the datasets we're provided you for the Algothon). This example, as you might guess, will provide you with the daily ticker information of Facebook (hence FB), you will get the same information as you got last time with Apple, such as open, close, high and low that day. Finally, you can choose the format in which you want to receive the data, the three choices are CSV, JSON & XML, obviously in this example we are using csv (hence data.csv). The final statement "head(5)" is just to show the first five rows of the csv file.

All together the code is:

```
import quandl
import pandas as pd

quandl.ApiConfig.api_key = "YOUR API" #setting up your api-key
data = quandl.get("WIKI/AAPL", rows = 5) #accessing the data for apple using quandl

data2 = pd.read_csv("https://www.quandl.com/api/v3/datasets/WIKI/FB/data.csv?api_key=YOUR API KEY").head(5)
```

'data' and 'data2' both contain identical data pulled from the WIKI/FB dataset, showing 2 different methods for obtaining the same information.

Kaggle

Kaggle is a popular data science platform used for competitions, learning, and programming.

Our Kaggle competition page can be found at: www.kaggle.com/c/algothon-2018/
If you are taking part in the Reuters NLP challenge (predicting future behavior of the ticker we've provided you) this is where you will upload your prediction to see your accuracy.

To use Kaggle you should upload your prediction to the competition page, after which you will be provided with the error in your prediction compared to the actual behavior. This error is Root Mean Squared. You are limited to 20 uploads per day, this is in place to avoid trial and error techniques.

Your response should be in the .csv format with the columns 'time' and 'mid-price'. It should have hourly data with the date/time range of all trading days from 00:00 on the 1/1/2018 to 23:00 31/07/2018. An example submission file is available for download on the Kaggle page.