云实习实操测试题： 数据处理与因子研发

大鱼学院： fxdayu.com
视频网课： 网易云课堂搜索OFO

下载与安装资源：

安装大鱼团队扩展的jaqs版本：

在命令行输入 pip install git+https://github.com/xingetouzi/JAQS.git@fxdayu

下载链接: https://pan.baidu.com/s/1qZjQoGG

sz50.xslx(excel数据)

market(hd5数据文件)

Alpha1_Example(ipynb因子模板)

题目一： 下载并用pandas导入sz50.xlsx的所有股票，索引设置为datetime，将所有股票的keys打印出来。

输出如下：

```
print(data.keys())
```

```
dict_keys(['601288.XSHG', '601601.XSHG', '601988.XSHG', '601198.XSHG', '601766.XSHG', '601628.XSHG', '601318.XSHG',
'600104.XSHG', '600050.XSHG', '600000.XSHG', '600111.XSHG', '600837.XSHG', '600958.XSHG', '601668.XSHG', '601901.XSH
G', '601989.XSHG', '601800.XSHG', '601788.XSHG', '601166.XSHG', '600030.XSHG', '600999.XSHG', '601229.XSHG', '601818.
XSHG', '601328.XSHG', '600547.XSHG', '600028.XSHG', '600016.XSHG', '601006.XSHG', '601336.XSHG', '601390.XSHG', '6012
11.XSHG', '600029.XSHG', '600048.XSHG', '600100.XSHG', '600519.XSHG', '601169.XSHG', '600036.XSHG', '601688.XSHG', '6
01088.XSHG', '600887.XSHG', '601398.XSHG', '601985.XSHG', '601857.XSHG', '601186.XSHG', '600340.XSHG', '600919.XSHG',
'600606.XSHG', '601881.XSHG', '600518.XSHG'])
```

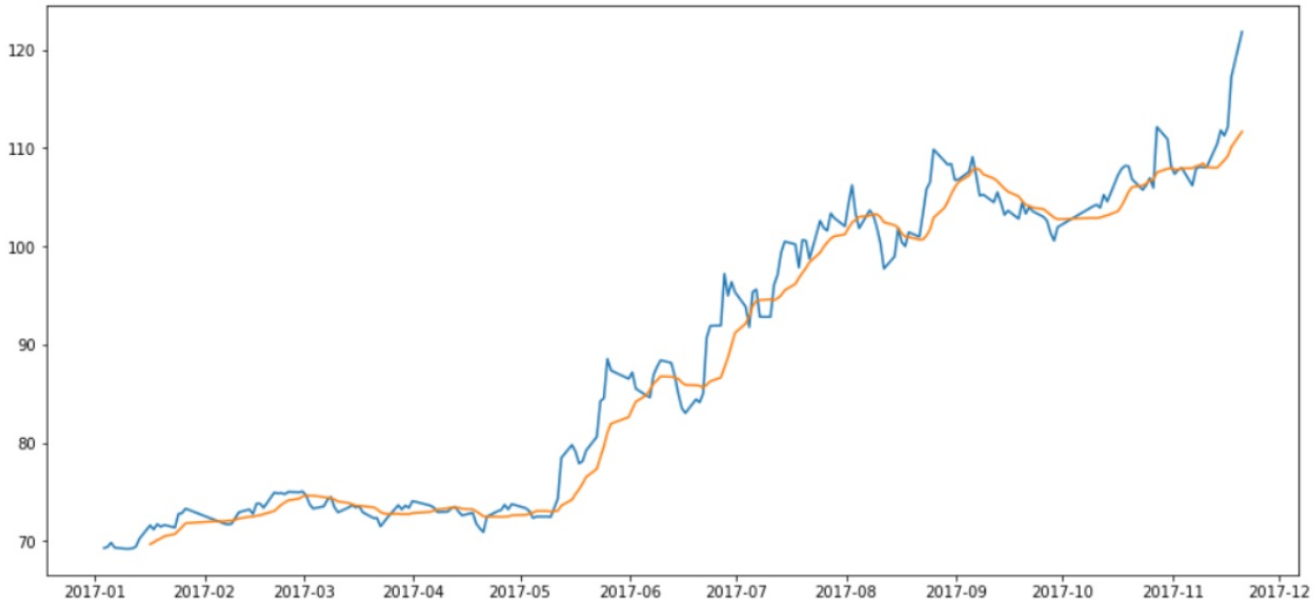
题目二：读取data里的600036这只股票的DataFrame,将其收盘价转换成用Numpy的Array格式，并用talib计算10日均线值，返回ndarray的最后五个值

返回如下

```
<class 'numpy.ndarray'>
[ 108.35   108.741  109.176  110.102  111.666]
```

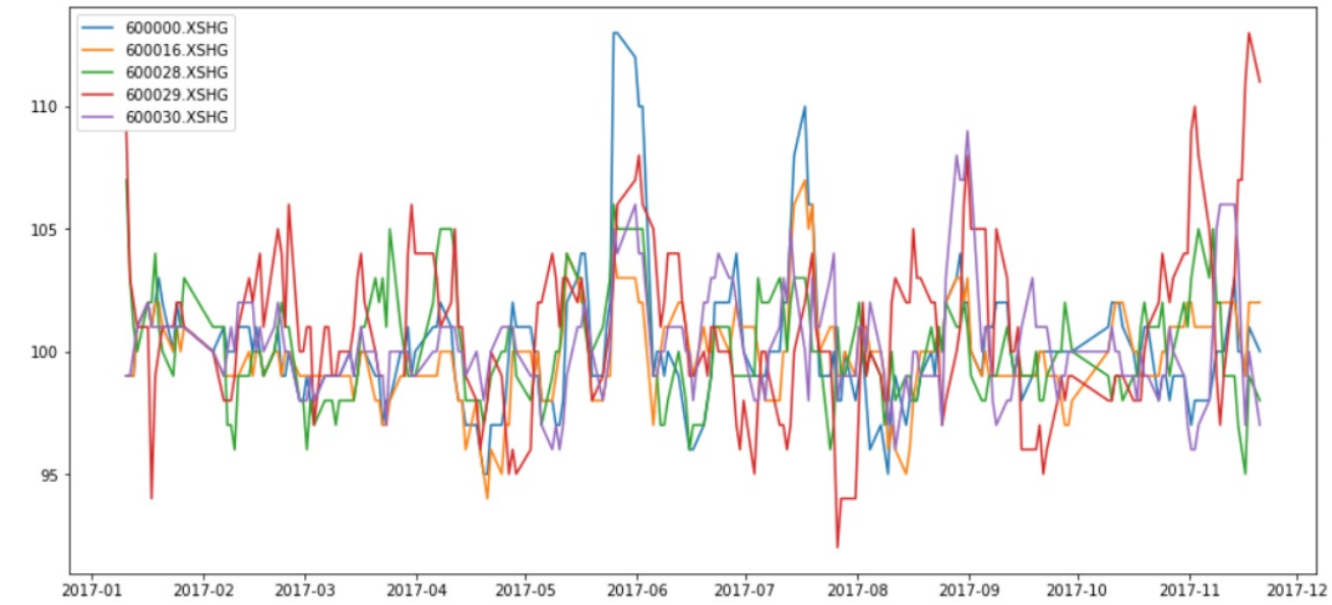
题目三：将MA的ndarray数据格式转换成Series格式，并加上datetime索引，最后将价格和MA值用Matplotlib展示出来

输出如下：



题目四：用talib计算50只股票的周期为5的ROCR100，生成Dataframe，并将前5只股票的ROCR100（参数timeperiod=20）用一张图显示出来

输出如下



题目五：用Panel来计算50只股票的MACD并且输出MACD的Panel的MultiIndex格式

输出如下

datetime	minor	macd	macdsignal	macdhist
2017-02-24 15:00:00	600000.XSHG	0.871114	1.197907	-0.326792
	600016.XSHG	-0.007077	0.265328	-0.272405
	600028.XSHG	0.028081	0.052015	-0.023934
	600029.XSHG	0.221787	0.132200	0.089587
	600030.XSHG	0.602873	0.620235	-0.017362
	600036.XSHG	1.291055	1.245643	0.045412
	600048.XSHG	2.316978	1.771997	0.544982
	600050.XSHG	-0.105248	-0.189278	0.084029
	600100.XSHG	0.453028	0.079491	0.373537
	600104.XSHG	2.271146	2.012658	0.258488
	600111.XSHG	2.429090	1.926561	0.502529
	600340.XSHG	8.025312	5.194384	2.830927
	600518.XSHG	1.014290	2.387216	-1.372926
	600519.XSHG	20.810882	7.146779	13.664103
	600547.XSHG	-0.041687	0.918715	-0.960402
	600606.XSHG	0.298737	0.267833	0.030904
	600837.XSHG	-1.406916	-0.673088	-0.733828
	600887.XSHG	7.150619	7.826700	-0.676082
	600919.XSHG	0.203864	0.155884	0.047980

题目六： 根据模板把以下算法实现

下载market的hd5数据与Alha1_Example因子模板

内置算法描述链接： https://github.com/quantOS-org/quantOSUserGuide/blob/master/jaqs/dataview_formula.md

内置算法源码路径（Parser）：from jaqs.data.py_expression_eval import Parser

根据模板实现的算法： RANK(SUM(CORR(RANK(VOLUME), RANK(VWAP)), 6), 2)
标注： VWAP = amount/volume

最终参考输出如下：

```
Best_Period                                                {5: 0.04}
Formula            -Correlation(Delta(Log(volume),{}),(close-open...
High_IC_Industry   [公用事业, 农林牧渔, 化工, 国防军工, 家用电器, 建筑装饰, 房地产, 电子, 食品饮料]
Low_IC_Industry                                    [休闲服务, 商业贸易]
classify                                                       sw1
description                          对数成交量的1天差与当天涨跌幅的过去6天相关系数
market                                                       hs300
name                                                        alpha1
parameter                                                   [1, 6]
type                                                          价量类
dtype: object
```