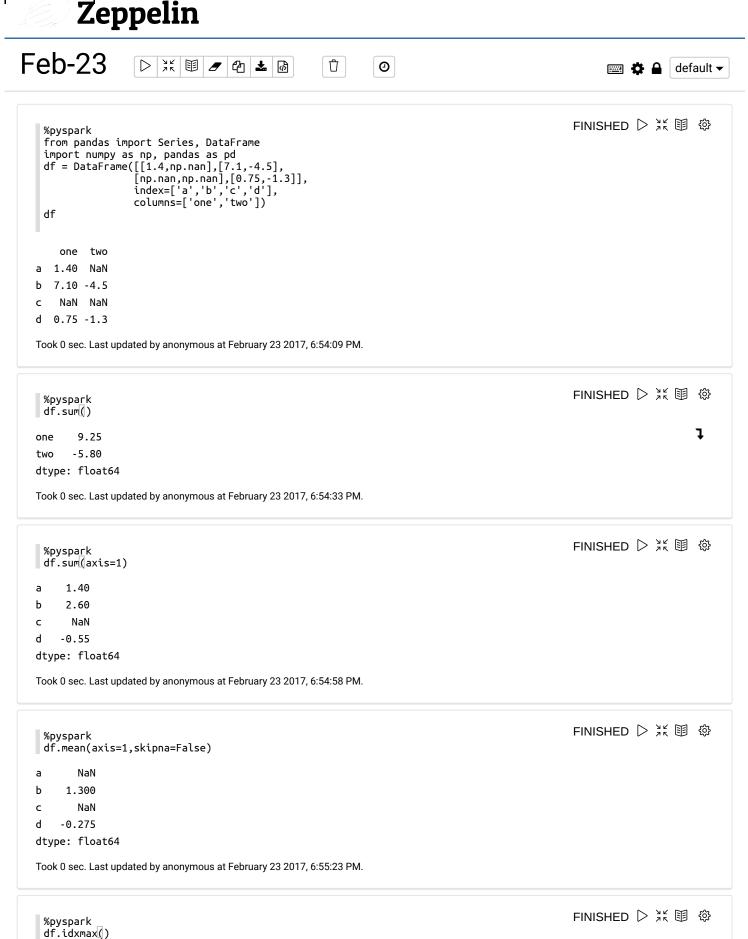
Untitled Unt



and the same that are

一人婦 つり

localhost:8080/#/notebook/2CBZQ1ME4 Untitled Unt Took 0 sec. Last updated by anonymous at February 23 2017, 6:55:44 PM \* I Ů ( default ▼ %pyspark df.describe() l one two count 3.000000 2.000000 3.083333 -2.900000 mean std 3.493685 2.262742 min 0.750000 -4.500000 1.075000 -3.700000 25% 50% 1.400000 -2.900000 75% 4.250000 -2.100000 7.100000 -1.300000 max Took 0 sec. Last updated by anonymous at February 23 2017, 6:56:29 PM. FINISHED ▷ ¾ ■ ፟ %pyspark obj = Series(['a','a','b','c'] \* 4) obj obj.describe() count 16 3 unique top а freq 8 dtype: object Took 0 sec. Last updated by anonymous at February 23 2017, 6:56:53 PM. FINISHED ▷ ※ 圓 � %pyspark from pandas\_datareader import data, wb  $all_data = {}$ for ticker in ['AAPL','IBM','MSFT','GOOG']: all\_data[ticker] = web.get\_data\_yahoo(ticker)
price = DataFrame({tic: data['Adj Close']} for tic, data in all\_data.items()}) volume = DataFrame({tic: data['Volume'] for tic, data in all\_data.items()}) returns = price.pct\_change() returns.tail() ļ AAPL IBM **MSFT** GOOG Date 2017-02-15 0.003629 -0.001792 0.008605 -0.000619 2017-02-16 -0.001181 0.006325 -0.001376 -0.000155 2017-02-17 0.002734 0.004744 -0.004189 0.001550 2017-02-21 0.007221 0.004335 -0.002269 -0.002012 2017-02-22 0.002999 -0.001082 0.004937 -0.002016 Took 1 sec. Last updated by anonymous at February 23 2017, 6:53:12 PM. FINISHED ▷ 兆 圓 ۞ %pyspark returns.MSFT.corr(returns.IBM) ı 0.49515377802280924 Took 1 sec. Last updated by anonymous at February 23 2017, 6:53:53 PM.

