

I needed to change my kaggle dataset a bit to avoid the potential OOM issue. If you use my kaggle dataset for the Numerai Signals, please see the following discussion to follow what has been changed:

[kaggle.com](https://www.kaggle.com)

### **YFinance Stock Price Data for Numerai Signals**

Daily Updates of Stock OHLCV (Close, High, Low, Open, Volume)

Presumably all you need to change in your pipeline is to use `pd.read_parquet`

instead of `pd.read_csv`

.