

Практическое занятие № 1. Анализ котировок акций.

Анализ котировок акций является весьма увлекательным (и для кого-то прибыльным) занятием, которое в чем-то напоминает азартные игры. Однако нашей целью будет понять, как могут выглядеть данные и каким образом с ними можно работать.

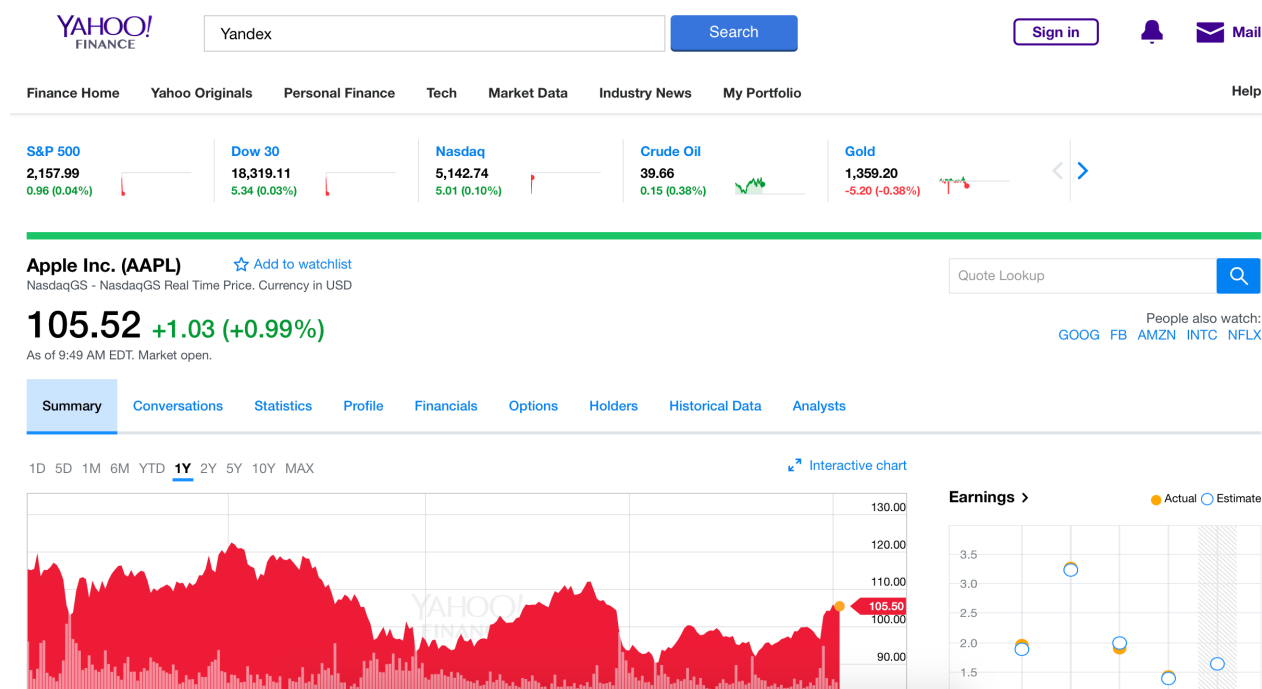
1. Цель занятия

- Найти, скачать и изучить набор данных;
- Вспомнить, как читать данные из файлов;
- Сгенерировать аналитические данные;
- Проанализировать данные: сколько мы могли бы заработать, если бы умели идеально играть на бирже;
- Визуализировать свои прибыли :)

2. Поиск набора данных

Сайт **Yahoo!Finance** является хорошим местом чтобы начать изучение исторических данных о торгах на биржах в США. Каждой компании на бирже присваивают короткий набор символов, который ее идентифицирует (по-другому этот набор еще называется тикер, ticker). Например, корпорация Apple имеет тикер AAPL, Microsoft — MSFT, Yandex — YNDX.

Задача: выберите любую корпорацию и найдите по ее тикеру посвященную ей страницу. Она выглядит примерно вот так:



На этой странице есть ссылка **Historical Data**, нажмите на нее и перед вами появятся ежедневные котировки акций выбранной компании.

Задача: скачайте всю историю ежедневных котировок выбранной компании к себе на диск. Чтобы выбрать весь период, когда акции компании торговались на бирже, нажмите на временной период и далее выберите **MAX**. Потом щелкните на синюю кнопку **Apply**, а потом выберите ссылку **Download Data**.

3. Чтение данных

Скачанный набор данных имеет формат .csv и если вы его откроете, то увидите, что это обыкновенный текстовый файл, который выглядит примерно так:

```
Date,Open,High,Low,Close,Volume,Adj Close
2016-08-02,106.050003,106.07,104.00,104.480003,33089600,104.480003
2016-08-01,104.410004,106.150002,104.410004,106.050003,38167900,106.050003
2016-07-29,104.190002,104.550003,103.68,104.209999,27733700,104.209999
2016-07-28,102.830002,104.449997,102.82,104.339996,39869800,104.339996
2016-07-27,104.269997,104.349998,102.75,102.949997,92344800,102.949997
2016-07-26,96.82,97.970001,96.419998,96.669998,56239800,96.669998
...
```

Первая строка в этом файле является заголовком и описывает приведенные данные. Остальные строки содержат данные. В каждой строке первой идет дата (Date), цена открытия (Open), наибольшая цена акций в этот день (High), наименьшая цена акций в этот день (Low), цена закрытия (Close), объем торгов, т.е. сколько акций было куплено/продано (Volume) и приведенная цена закрытия (Adj Close).

Чтобы прочитать этот файл можно использовать конструкцию with `open(...)` as `f`. При этом файлы читаются построчно при помощи цикла `for l in f`, где `l` — очередная строка файла. Можно также использовать метод `readlines()`.

4. Генерация аналитических данных

Из каждой строки надо удалить замыкающие переводы строк (метод `strip`) и разбить их на столбцы по запятым (метод `split(',')`).

Задача: сгенерируйте последовательность из приведенных цен на акции *AdjClose*.

По сути нам необходима только приведенная цена закрытия торгов (*Adj Close*), поэтому после разбиения данных на столбцы мы сохраним в списке цен акций только последнее значение из списка, получившегося после применения `split(...)`.

5. Анализ данных

Итак, у нас есть список цен закрытия торгов. Очевидно, что мы можем что-то заработать, если купим акции дешевле, а продадим дороже. Допустим у нас есть \$100, которые мы вложили в акции в первый день торгов. Если на следующий день акции дорожают, то мы их продаем и получаем прибыль иначе не вкладываем деньги в акции. Промоделируйте этот процесс на списке цен.

Сравните вашу стратегию с чисто случайной стратегией, когда перед решением, покупать акции или нет в очередной день, вы подкидываете монетку и если выпал орел, то покупаете, а если решка, то нет (здесь вам поможет модуль `random`).