



SF EDUCATION

Образовательный и карьерный менторинг

ПРОЕКТ КУРСА **DATA SCIENCE** **ACADEMY**

ЦЕЛЬ ПРОЕКТА

Выбрать наиболее выгодный портфель акций из предложенных и найти наиболее верное с точки зрения финансовой теории распределение акций в портфеле.

В основу анализа ложатся данные за 2019 год.

ФАКТИЧЕСКИЙ РЕЗУЛЬТАТ

Проект начинается с возможного портфеля акций торгующихся на российских и американских биржах.

Финальный результат проекта заключается в создании Excel файла с обозначенными акциями, их историческими показателями, показателями возможного риска и возврата.

Пример финального и начального файлы предоставлены в соответствующем архиве.

ПОРЯДОК ВЫПОЛНЕНИЯ ЗАДАНИЯ

1

После урока по Excel убедиться, что вы можете прочитать входные файлы и разделить их на русские и американские акции.

2

После урока по работе с удаленными данным убедиться, что вы можете взять информацию из входного файла по акциям, скачать последний год каждой акции из соответствующего источника, используя `pandas_datareader`, и сохранить информацию в отдельный лист в финальном файле.

Для получения русских акций рекомендую работать с каждой из акций по отдельности и делать `inner join` к американским акциям с изменением имени колонки на тикер акции.

Вам нужно будет исключить значения из `boardID` RPUA, RPEU, RPEO.

Для конвертирования в доллар возьмите значения колонки “CLOSE” и поделите на 64.

3

После урока по Монте-Карло

Взять входящую информацию по историческим данным, используя метод перебора (минимум 10,000 вариантов) с разными весами каждой акции в портфеле и при помощи коэффициента Шарпа найти наиболее оптимальное соотношение риска/возврата. Пример вариантов перебора, риска, возврата и коэффициента Шарпа можно посмотреть в соответствующем файле.

The Sharpe Ratio

$$= \frac{\bar{r}_p - r_f}{\sigma_p}$$

Where:

\bar{r}_p = Expected portfolio return

r_f = Risk free rate

σ_p = Portfolio standard deviation

Для expected portfolio return выберите среднее значение дневного процентного изменения x 253.

Другими словами (ненастоящий код)

```
aveReturn = df.pct_change().mean()*253
```

```
expReturn = weight*aveReturn
```

Для Risk Free Rate возьмите 0

Для portfolio standard deviation вам нужно взять covariance дневного процента изменения x 253.

Другими словами (ненастоящий код)

```
aveReturn = df.pct_change().cov()*253
```

4

Провести визуализацию вариантов различного веса в портфеле, как на финальном файле.

Сохранить всю информацию в финальном файле в необходимом формате. Все элементы, включая изображения должны быть вставлены автоматически.

Сохранять данные в отдельном листе

```
#####
#
# An example of writing multiple dataframes to worksheets using Pandas and
# XlsxWriter.
#
# Copyright 2013-2020, John McNamara, jmcnamara@cpan.org
#

import pandas as pd

# Create some Pandas dataframes from some data.
df1 = pd.DataFrame({'Data': [11, 12, 13, 14]})
df2 = pd.DataFrame({'Data': [21, 22, 23, 24]})
df3 = pd.DataFrame({'Data': [31, 32, 33, 34]})

# Create a Pandas Excel writer using XlsxWriter as the engine.
writer = pd.ExcelWriter('pandas_multiple.xlsx', engine='xlsxwriter')

# Write each dataframe to a different worksheet.
df1.to_excel(writer, sheet_name='Sheet1')
df2.to_excel(writer, sheet_name='Sheet2')
df3.to_excel(writer, sheet_name='Sheet3')

# Close the Pandas Excel writer and output the Excel file.
writer.save()
```

Сохранять изображения

```
plt.savefig('books_read.png')
```

Вставлять Изображения

```
import openpyxl

wb = openpyxl.load_workbook('input.xlsx')
ws = wb.active

img = openpyxl.drawing.image.Image('myplot.png')
ws.add_image(ws.cell('A1'))

ws.save('output.xlsx')
```