Using pandas Dataframes With Excel

lesson_4_2_l

Description

You are applying to be a junior-level data analyst for a Financial Planner's office. The senior-level data analyst has given you a file named `dow_jones_index.data` and a machine that has Anaconda installed. He asks that you produce an Excel file with worksheets for each of three stock ticker names: GE, IBM, and KRFT.

Determine the data structure of the provided file

```
# open file for reading
f = open('dow_jones_index.data')

# print the first two lines
print(f.readline())
print(f.readline())

f.close()
```

It appears the file is CSV. Read the file into a dataframe.

```
import pandas as pd

stock_df = pd.read_csv('dow_jones_index.data')

stock_df.head()
```

Create a dataframe for each of the requested stocks

```
ge_df = stock_df[stock_df.stock=='GE']
ibm_df = stock_df[stock_df.stock=='IBM']
krft_df = stock_df[stock_df.stock=='KRFT']
```

Write the Excel file

```

```

```
with pd.ExcelWriter('stocks.xlsx') as writer:
    ge_df.to_excel(writer, sheet_name='GE')
    ibm_df.to_excel(writer, sheet_name='IBM')
    krft_df.to_excel(writer, sheet_name='KRFT')
```

Load the file into an ordered dict dataframe and then check that each worksheet is populated.

```
my_stock_df = pd.read_excel('stocks.xlsx', sheet_name=None)

my_stock_df.keys()
```

```
my_stock_df['GE']
```

```
my_stock_df['IBM']
```

```
my_stock_df['KRFT']
```