Zipline Coding Exercises

In this notebook you will create your own pipeline and create custom factors and filters for it.

Install Packages

In [1]:
```python
import sys
!{sys.executable} -m pip install -r requirements.txt
```

```
Collecting zipline===1.3.0 (from -r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/be/59/8c5802a7897c1095fdc409fb557f04df8f75c37174e80d2ba58c8d8a6488/zipline-1.3.0.tar.gz (2.5MB)
    100% |████████████████████████████████| 2.5MB 11.0MB/s ta 0:00:01    91% |█████████████████████████████▏  | 2.3MB 21.9MB/s eta 0:00:01
Collecting graphviz==0.9 (from -r requirements.txt (line 2))
  Downloading https://files.pythonhosted.org/packages/47/87/313cd4ea4f75472826acb74c57f94fc83e04ba93e4ccf35656f6b7f502e2/graphviz-0.9-py2.py3-none-any.whl
Requirement already satisfied: pip>=7.1.0 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (18.1)
Requirement already satisfied: setuptools>18.0 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (38.4.0)
Collecting Logbook>=0.12.5 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/2f/d9/16ac346f7c0102835814cc9e5b684aaadea101560bb932a2403bd26b2320/Logbook-1.5.3.tar.gz (85kB)
    100% |████████████████████████████████| 92kB 19.4MB/s ta 0:00:01
Requirement already satisfied: pytz>=2016.4 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (2017.3)
Requirement already satisfied: numpy>=1.11.1 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (1.12.1)
Collecting requests-file>=1.4.1 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/23/9c/6e63c23c39e53d3df41c77a3d05a49a42c4e1383a6d2a5e3233161b89dbf/requests_file-1.4.3-py2.py3-none-any.whl
Requirement already satisfied: scipy>=0.17.1 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (0.19.1)
Collecting pandas<=0.22,>=0.18.1 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/da/c6/0936bc5814b429fddb5d6252566fe73a3e40372e6ceaf87de3dec1326f28/pandas-0.22.0-cp36-cp36m-manylinux1_x86_64.whl (26.2MB)
    100% |████████████████████████████████| 26.3MB 1.3MB/s eta 0:00:01  2% |▊                               | 624kB 25.5MB/s eta 0:00:02    34% |███████████                     | 9.1MB 32.8MB/s eta 0:00:01    40% |█████████████                   | 10.7MB 30.2MB/s eta 0:00:01    57% |██████████████████▌             | 15.2MB 32.0MB/s eta 0:00:01    63% |████████████████████▍           | 16.6MB 30.6MB/s eta 0:00:01    68% |██████████████████████          | 18.1MB 28.9MB/s eta 0:00:01    79% |█████████████████████████▍      | 20.9MB 24.9MB/s eta 0:00:01    89% |████████████████████████████▌   | 23.4MB 24.0MB/s eta 0:00:01
Collecting pandas-datareader>=0.2.1 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/14/52/accb990baebe0063977f26e02df36aa7eb4015ed4e86f828cd76273cd6f1/pandas_datareader-0.8.1-py2.py3-none-any.whl (107kB)
    100% |████████████████████████████████| 112kB 12.0MB/s ta 0:00:01
Requirement already satisfied: patsy>=0.4.0 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (0.4.1)
Requirement already satisfied: statsmodels>=0.6.1 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (0.8.0)
Requirement already satisfied: python-dateutil>=2.4.2 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (2.6.1)
Requirement already satisfied: six>=1.10.0 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (1.11.0)
Requirement already satisfied: requests>=2.9.1 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (2.18.4)
Requirement already satisfied: Cython>=0.25.2 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (0.29.7)
Collecting cyordereddict>=0.2.2 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/d1/1a/364cbfd927be1b743c7f0a985a7f1f7e8a51469619f9fefe4ee9240ba210/cyordereddict-1.0.0.tar.gz (138kB)
    100% |████████████████████████████████| 143kB 21.9MB/s ta 0:00:01
Collecting bottleneck>=1.0.0 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/62/d0/55bbb49f4fade3497de2399af70ec0a06432c786b8623c878b11e90d456/Bottleneck-1.3.1.tar.gz (88kB)
    100% |████████████████████████████████| 92kB 1.7MB/s ta 0:00:011
  Installing build dependencies ... done
Collecting contextlib2>=0.4.0 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/85/60/370352f7ef6aa96c52fb001831622f50f923c1d575427d021b8ab3311236/contextlib2-0.6.0.post1-py2.py3-none-any.whl
Requirement already satisfied: decorator>=4.0.0 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (4.0.11)
Requirement already satisfied: networkx<2.0,>=1.9.1 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (1.11)
Requirement already satisfied: numexpr>=2.6.1 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (2.6.4)
Collecting bcolz<1,>=0.12.1 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/6c/8b/1ffa01f872cac36173c5eb95b58c01040d8d25f1b242c48577f4104cd3ab/bcolz-0.12.1.tar.gz (622kB)
    100% |████████████████████████████████| 624kB 16.5MB/s ta 0:00:01    52% |████████████████▊               | 327kB 26.5MB/s eta 0:00:01
Requirement already satisfied: click>=4.0.0 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (6.7)
Requirement already satisfied: toolz>=0.8.2 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (0.8.2)
Collecting multipledispatch>=0.4.8 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/89/79/429ecef45fd5e4504f7474d4c3c3c4668c267be3370e4c2fd33e61506833/multipledispatch-0.6.0-py3-none-any.whl
Requirement already satisfied: MarkupSafe>=0.23 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (1.0)
Requirement already satisfied: Mako>=1.0.1 in /opt/conda/lib/python3.6/site-packages/Mako-1.0.7-py3.6.egg (from zipline===1.3.0->-r requirements.txt (line 1)) (1.0.7)
Requirement already satisfied: sqlalchemy>=1.0.8 in /opt/conda/lib/python3.6/site-packages (from zipline===1.3.0->-r requirements.txt (line 1)) (1.1.13)
Collecting alembic>=0.7.7 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/dc/6d/3c1411dfdcf089ec89ce5e2222deb2292f39b6b1a5911222e15af9fe5a92/alembic-1.3.2.tar.gz (1.1MB)
    100% |████████████████████████████████| 1.1MB 15.6MB/s ta 0:00:01
Collecting sortedcontainers>=1.4.4 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/13/f3/cf85f7c3a2dbd1a515d51e1f1676d971abe41bba6f4ab5443240d9a78e5b/sortedcontainers-2.1.0-py2.py3-none-any.whl
Collecting intervaltree>=2.1.0 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/e8/f9/76237755b2020cd74549e98667210b2dd54d3fb17c6f4a62631e61d31225/intervaltree-3.0.2.tar.gz
Collecting lru-dict>=1.1.4 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/00/a5/32ed6e10246cd341ca8cc205acea5d208e4053f48a4dced2b1b31d45ba3f/lru-dict-1.1.6.tar.gz
Collecting empyrical>=0.5.0 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/84/9e/9506e8b25464ff57ef93b5ba9092b464b44dc76b717695b126b3c93214a2/empyrical-0.5.3.tar.gz (50kB)
    100% |████████████████████████████████| 51kB 13.3MB/s ta 0:00:01
Collecting tables>=3.3.0 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/ed/c3/8fd9e3bb21872f9d69eb93b3014c86479864cca94e625fd03713ccacec80/tables-3.6.1-cp36-cp36m-manylinux1_x86_64.whl (4.3MB)
    100% |████████████████████████████████| 4.3MB 8.2MB/s eta 0:00:01    33% |██████████▋                     | 1.5MB 28.1MB/s eta 0:00:01
Collecting trading-calendars>=1.0.1 (from zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/a9/cc/c871c5f19bddf5c773acfa450de68de8ba3506a1d85b1927e53e127b03e7/trading_calendars-1.11.1.tar.gz (382kB)
    100% |████████████████████████████████| 389kB 19.5MB/s ta 0:00:01
Requirement already satisfied: lxml in /opt/conda/lib/python3.6/site-packages (from pandas-datareader>=0.2.1->zipline===1.3.0->-r requirements.txt (line 1)) (4.1.1)
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in /opt/conda/lib/python3.6/site-packages (from requests>=2.9.1->zipline===1.3.0->-r requirements.txt (line 1)) (3.0.4)
Requirement already satisfied: idna<2.7,>=2.5 in /opt/conda/lib/python3.6/site-packages (from requests>=2.9.1->zipline===1.3.0->-r requirements.txt (line 1)) (2.6)
Requirement already satisfied: urllib3<1.23,>=1.21.1 in /opt/conda/lib/python3.6/site-packages (from requests>=2.9.1->zipline===1.3.0->-r requirements.txt (line 1)) (1.22)
Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.6/site-packages (from requests>=2.9.1->zipline===1.3.0->-r requirements.txt (line 1)) (2019.6.16)
Collecting python-editor>=0.3 (from alembic>=0.7.7->zipline===1.3.0->-r requirements.txt (line 1))
  Downloading https://files.pythonhosted.org/packages/c6/d3/201fc3abe391bbae6606e6f1d598c15d367033332bd54352b12f35513717/python_editor-1.0.4-py3-none-any.whl
Building wheels for collected packages: zipline, Logbook, cyordereddict, bottleneck, bcolz, alembic, intervaltree, lru-dict, empyrical, trading-calendars
  Running setup.py bdist_wheel for zipline ... done
  Stored in directory: /root/.cache/pip/wheels/a4/d6/67/f303ab028b004bf8e00c05b5b04fba83d8ec238b6547becdb7
  Running setup.py bdist_wheel for Logbook ... done
  Stored in directory: /root/.cache/pip/wheels/d2/70/07/68b99a8e05dcd1ab194a8e0ccb9e4d0ac5dd6d8d139c7149b4
  Running setup.py bdist_wheel for cyordereddict ... done
  Stored in directory: /root/.cache/pip/wheels/0b/9d/8b/5bf3e22c1edd59b50f11bb19dec9dfcfe5a479fc7ace02b61f
  Running setup.py bdist_wheel for bottleneck ... done
  Stored in directory: /root/.cache/pip/wheels/31/36/8f/1ed7e6f1b3295499c8bbab934262f2494d0f6aebe0c5860754
  Running setup.py bdist_wheel for bcolz ... done
  Stored in directory: /root/.cache/pip/wheels/c5/cc/1b/2cf1f88959af5d7f4d449b7fc6c9452d0ecbd86fd61a9ee376
  Running setup.py bdist_wheel for alembic ... done
  Stored in directory: /root/.cache/pip/wheels/5c/66/53/e0633382ac8625ab1c099db6a290d1b6b24f849a4666a57105
  Running setup.py bdist_wheel for intervaltree ... done
  Stored in directory: /root/.cache/pip/wheels/08/99/c0/5a5942f5b9567c59c14aac76f95a70bf11dccc71240b91ebf5
  Running setup.py bdist_wheel for lru-dict ... done
  Stored in directory: /root/.cache/pip/wheels/b7/ef/06/fbdd555907a7d438fb33e4c8675f771f1cf41917284c51ebf
  Running setup.py bdist_wheel for empyrical ... done
  Stored in directory: /root/.cache/pip/wheels/10/a4/3b/951bd609878a82fd72b9ea23699daf1eaada4ff6f583152876
  Running setup.py bdist_wheel for trading-calendars ... done
  Stored in directory: /root/.cache/pip/wheels/e8/97/03/01e2eac187e7ad91f6502adeb910ff158005ee28ff0df7e34b
Successfully built zipline Logbook cyordereddict bottleneck bcolz alembic intervaltree lru-dict empyrical trading-calendars
Installing collected packages: Logbook, requests-file, pandas, pandas-datareader, cyordereddict, bottleneck, contextlib2, bcolz, multipledispatch, python-editor, alembic, sortedcontainers, intervaltree, lru-dict, empyrical, tables, trading-calendars, zipline, graphviz
  Found existing installation: pandas 0.23.3
    Uninstalling pandas-0.23.3:
      Successfully uninstalled pandas-0.23.3
Successfully installed Logbook-1.5.3 alembic-1.3.2 bcolz-0.12.1 bottleneck-1.3.1 contextlib2-0.6.0.post1 cyordereddict-1.0.0 empyrical-0.5.3 graphviz-0.9 intervaltree-3.0.2 lru-dict-1.1.6 multipledispatch-0.6.0 pandas-0.22.0 pandas-datareader-0.8.1 python-editor-1.0.4 requests-file-1.4.3 sortedcontainers-2.1.0 tables-3.6.1 trading-calendars-1.11.1 zipline-1.3.0
```

Load Data

In [27]:
```python
import os

from zipline.data import bundles
from zipline.data.bundles.csvdir import csvdir_equities


# Specify the bundle name
bundle_name = 'm4-quiz-eod-quotemedia'

# Create an ingest function
ingest_func = csvdir_equities(['daily'], bundle_name)

# Register the data bundle and its ingest function
bundles.register(bundle_name, ingest_func);

# Set environment variable 'ZIPLINE_ROOT' to the path where the most recent data is located
os.environ['ZIPLINE_ROOT'] = os.path.join(os.getcwd(), '..', '..', 'data', 'module_4_quizzes_eod')

# Load the data bundle
bundle_data = bundles.load(bundle_name)
```

```
/opt/conda/lib/python3.6/site-packages/ipykernel_launcher.py:14: UserWarning: Overwriting bundle with name 'm4-quiz-eod-quotemedia'
```

TODO: Build an Empty Pipeline with a Screen

In the code below create an empty with a screen that filters the pipeline output for stocks with a 60-day average dollar volume greater than $50,000.

In [28]:
```python
# import resources
from zipline.pipeline import Pipeline
from zipline.pipeline.factors import AverageDollarVolume

# Create a screen for our Pipeline
universe = AverageDollarVolume(window_length = 60) > 50000

# Create an empty Pipeline with the given screen
pipeline = Pipeline(screen = universe)
```

TODO: Create a Pipeline Engine

In the code below, create a data loader that loads our data bundle using Zipline's `USEquityPricingLoader`. Then create a function to be passed to the `get_loader` parameter in the pipeline engine. Using the NYSE trading calendar create a pipeline engine using Zipline's `SimplePipelineEngine`

In [29]:
```python
# import resources
from zipline.pipeline.loaders import USEquityPricingLoader
from zipline.utils.calendars import get_calendar
from zipline.pipeline.engine import SimplePipelineEngine

# Set the dataloader
pricing_loader = USEquityPricingLoader(bundle_data.equity_daily_bar_reader, bundle_data.adjustment_reader)

# Define the function for the get_loader parameter
def choose_loader(column):
    if column not in USEquityPricing.columns:
        raise Exception('Column not in USEquityPricing')
    return pricing_loader

# Set the trading calendar
trading_calendar = get_calendar('NYSE')

# Create a Pipeline engine
engine = SimplePipelineEngine(get_loader = choose_loader,
                              calendar = trading_calendar.all_sessions,
                              asset_finder = bundle_data.asset_finder)
```

TODO: Create Factors and Filters For The Pipeline

Factors

In the code below, create two factors:

1. **Custom Factor**: Create a custom factor, `percent_difference`, by combining a 30-day average closing price factor and a 60-average closing price factor. The `percent_difference` factor will calculate the difference between the 30-day and 60-average closing price factors, and will then normalize this difference by the 60-average closing price factor.

1. **Daily Returns Factor**: Create a factor that returns the daily percent change in closing price. To do this, use Zipline's built-in [DailyReturns](#) factor.

Both factors should use the closing prices in the `USEquityPricing` dataset as inputs.

Filters

In the code below, you will create a **Custom Filter**, `tradable_asset`, that returns `True` for the top 20 securities of average dollar volume in a 30 day-window that have a latest closing price above $30. To do this, you will create two filters first and them combine them. For the first filter, you can use Zipline's built-in factor `AverageDollarVolume()` and its method `.top()` to select the top 20 assets of average dollar volume in a 30 day-window. To get the latest close price of the `USEquityPricing` dataset you can use `USEquityPricing.close.latest`. Finally you can combine these two filters by using the logical operator `&`.

In [30]:
```python
# import resources
import pandas as pd
from zipline.pipeline.factors import SimpleMovingAverage
from zipline.pipeline.factors import DailyReturns
from zipline.pipeline.data import USEquityPricing

# Create the 30-day average closing price factor
avg_close_30 = SimpleMovingAverage(inputs = [USEquityPricing.close],
                                   window_length = 30)

# Create the 60-day average closing price factor
avg_close_60 = SimpleMovingAverage(inputs = [USEquityPricing.close],
                                   window_length = 60)

# Create the custom factor
percent_difference = (avg_close_30 - avg_close_60) / avg_close_60

# Create the daily returns factor
daily_returns = DailyReturns()

# Create a filter for the top 20 securities of average dollar volume in a 30 day-window
top_20_volume = AverageDollarVolume(window_length = 30).top(20)

# Create a filter for the latest closing price above $30
close_over_30 = USEquityPricing.close.latest > 30

# Create a custum filter
custom_filter = top_20_volume and close_over_30
```

TODO: Add Factors and Filters To The Pipeline

In the code below, add the factors and filters you created above to the pipeline

In [31]:
```python
# Add the custom factor to the pipeline
pipeline.add(percent_difference, 'Percent Difference (30 vs. 60 days)')

# Add the daily returns factor to the pipeline
pipeline.add(daily_returns, 'Daily Returns')

# Add the custom filter to the pipeline
pipeline.add(custom_filter, 'Tradable')
```

TODO: Run The Pipeline

In the code below, run the pipeline for the dates given

In [32]:
```python
# Set starting and end dates
start_date = pd.Timestamp('2014-01-06', tz='utc')
end_date = pd.Timestamp('2016-01-05', tz='utc')

# Run our pipeline for the given start and end dates
output = engine.run_pipeline(pipeline, start_date, end_date)

# Display the pipeline output
output.head()
```

Out[32]:

		Daily Returns	Percent Difference (30 vs. 60 days)	Tradable
2014-01-06 00:00:00+00:00	**Equity(0 [A])**	0.006951	0.025939	True
	Equity(1 [AAL])	0.046523	0.055887	False
	Equity(2 [AAP])	0.028611	0.051009	True
	Equity(3 [AAPL])	-0.021972	0.035906	True
	Equity(4 [ABBV])	0.006147	0.033018	True