Exploratory Data Analysis

September 22, 2024

Functions (IGNORE)

Functions for Downloading Data

- 'get_data_from_start_to_end': Load in the data from YFinance API
- 'download_stock_data_for_periods': Look at different periods to understand how technical analysis changes throughout changes in macroeconomy
- 'get_adjusted_closed_price' : Get the dataframe of adjusted closed price with columns as ticker and row as date.

```python
import yfinance as yf
missing_data_tickers = [] # use this as a list of tickers with missing data

def get_data_from_start_to_end(ticker, start_date, end_date):
    global missing_data_tickers  # Use the global list to accumulate missing
 ↪tickers
    try:
        stock_data = yf.download(ticker, start=start_date, end=end_date)
        if stock_data.empty:
            missing_data_tickers.append(ticker)
            raise ValueError(f"Stock data for ticker {ticker} during the period
 ↪from {start_date} to {end_date} was not found.")
        return stock_data
    except Exception as e:
        print(f"An error occurred for ticker {ticker}: {e}")
        missing_data_tickers.append(ticker)
        return None
```

```python
# for a variety of periods load in different list of tickers
def download_stock_data_for_periods(tickers, periods):
    all_data = {}

    for period, (start_date, end_date) in periods.items():
        period_data = {}
        for ticker in tickers:
            data = get_data_from_start_to_end(ticker, start_date, end_date)
            if data is not None:
                period_data[ticker] = data
```

```
        all_data[period] = period_data

    return all_data
```

```python
import pandas as pd

# Get the adjusted close prices
adj_close_sector_etf = {}

# Create adjusted close price only listing of sector ETFs
def get_adjusted_closed_price(nested_dict, tickers, periods):
    for period in periods:
        stock_price_df = pd.DataFrame()  # Create a new DataFrame for each
 period
        for ticker in tickers:
            stock_price_df[ticker] = nested_dict[period][ticker]['Adj Close']

        adj_close_sector_etf[period] = stock_price_df  # Store the complete
 DataFrame for the period

    return adj_close_sector_etf
```

Descriptive Analytics

```python
# create a function to get a table of descriptive analysis of the different
 sector etfs in different time periods

# import pandas
import pandas as pd

def create_analysis(tickers,periods,nested_dict,col_name,function_name):
    """
    parameters:

    tickers: the list of tickers that will be looped through
    periods: the list of periods that will be looped through
    nested_dict: the dictionary which can be accessed through
 nested_dict[period][ticker]
    col_name: the column of the nested dictionary
    function_name: the function that you want to be applied

    returns:
    a dataframe with columns as time periods and rows as sector etfs with the
 column name within the nested dictionary with a descriptive statistic applied
    """
    df = pd.DataFrame(index=tickers,columns=periods)
    for period in periods:
```

```
        for ticker in tickers:
            data = nested_dict[period][ticker][col_name]
            df.at[ticker, period] = getattr(pd.Series(data), function_name)()
    return df
```

Visualising Data

```
[ ]: import plotly.graph_objects as go

     def visualize_candlestick(data, title="Candlestick Chart"):
         """
         Visualizes candlestick data using Plotly.

         Parameters:
         data (pd.DataFrame): The dataframe containing the stock's OHLC data.
                              It should contain 'Open', 'High', 'Low', and 'Close'␣
      ↪columns.
         title (str): The title of the plot.

         Returns:
         fig: The Plotly figure object.
         """
         fig = go.Figure(data=[go.Candlestick(x=data.index,
                                             open=data['Open'],
                                             high=data['High'],
                                             low=data['Low'],
                                             close=data['Close'])])

         # Set titles and labels
         fig.update_layout(title=title,
                           xaxis_title='Date',
                           yaxis_title='Price',
                           xaxis_rangeslider_visible=False)

         fig.show()

         return fig
```

Stock Trading Strategies

```
[ ]: import numpy as np

     def buy_hold_roi(nested_dict, periods, tickers, column):
         """
         Objective:
         To return the ROI when purchasing stocks during different economic time␣
      ↪periods.
```

```python
    Parameters:
    nested_dict: holds the historical data
    periods: a list of time periods to be used as columns
    tickers: a list of tickers to be used as rows

    Returns:
    DataFrame with the ROI using a buy and hold strategy.
    """
    # Create a DataFrame to give the ROI for different investments
    roi = pd.DataFrame(index=tickers, columns=periods)

    # Sort through each ticker and each period
    for ticker in tickers:
        for period in periods:
            # Get the DataFrame through the nested dictionary
            data = nested_dict.get(period, {}).get(ticker)

            # Make sure that the data is present
            if data is not None and column in data.columns:
                start_price = data[column].iloc[0]
                end_price = data[column].iloc[-1]

                start_date = data.index[0]
                end_date = data.index[-1]

                num_days = (pd.to_datetime(end_date) - pd.
 to_datetime(start_date)).days

                if num_days > 0:  # Ensure we don't divide by zero
                    # Calculate percentage difference
                    total_return = ((end_price - start_price) / start_price) *
 100
                    annualized_roi = ((1 + total_return / 100) ** (365 /
 num_days) - 1) * 100  # Annualized percentage
                    roi.at[ticker, period] = annualized_roi
                else:
                    roi.at[ticker, period] = np.nan
            else:
                roi.at[ticker, period] = np.nan

    return roi
```

```python
import random

def stochastic_modeling(nested_dict, tickers, periods,num_samples):
    # Store the returns in a nested dictionary
```

```python
    nested_dict_returns = {period: {ticker: [] for ticker in tickers} for␣
↪period in periods}

    # Go through each economic time period
    for period in periods:
        max_index = len(nested_dict[period]) - 30  # Ensure there's enough data␣
↪to calculate ROI

        # Generate random samples from the valid range
        random_dates = random.choices(range(max_index), k=num_samples)

        for ticker in tickers:
            for date_idx in random_dates:
                start_price = nested_dict[period][ticker].iloc[date_idx]
                end_price = nested_dict[period][ticker].iloc[date_idx + 30]

                # Get the return by the Holding Period Return
                roi = (((end_price - start_price) / start_price) * 100)

                nested_dict_returns[period][ticker].append(roi)

    return nested_dict_returns  # Return the nested dictionary with returns
```

```python
def stochasting_roi(tickers,periods,return_rates_list,analysis_type):
    df = pd.DataFrame(index=tickers,columns=periods)
    for period in periods:
        for ticker in tickers:
            data = pd.Series(return_rates_list[period][ticker])
            if analysis_type=='Mean':
                df.at[ticker,period] = data.mean()
            elif analysis_type=='Median':
                df.at[ticker,period] = data.median()
            elif analysis_type=='Std':
                df.at[ticker,period] = data.std()
            elif analysis_type=='Variance':
                df.at[ticker,period] = data.var()

    return df
```

1 EDA and Data Cleaning

Assignment: Exploratory Data Analysis (EDA) and Data Cleaning Report

1.1 Dataset Overview:

1.1.1 Yahoo Finance API

The data that will be used for my capstone can be accessed through the Yahoo Finance API.

The Yahoo Finance API is a free popular API that is used by developers to load in a variety of stock data. This includes financial, candlestick and other information that can be utilized to make informed investment decisions.

Yahoo Finance is free and very easy to set up. First pip install yfinance and then simply load it into your notebook.

```python
# load the API in
import yfinance as yf # yf is the usual name for the package
```

The YFinance API offers stock candlestick data, information about the stock as well as financial/balance sheets. This can be utilised to create a number of analysis such as fundamental and technical analysis. This project will have a concentration in technical analysis which looks at the price of the stock.

Candlestick Data Candlestick data is used for a variety of techniques in investing as it holds the crucial information associated with the price of the stock throughout the trading day. This can be used to create a deeper understanding of how the stock changes on a day to day basis.

The columns contain candlestick data which is:

- 'Open' : The price of the equity at the beginning of the trading day.
- 'High' : The highest price of the equity throughout the day.
- 'Low' : The lowest price of the equity throughout the day.
- 'Close' : The price of the equity at the end of the trading day
- 'Adjusted Close' : The came as closing price however it takes dividends, stock splits and new stock offerings into account
- 'Volume': The number of times that equity is traded within a day.

```python
# save the ticker information of 'MSFT' as a variable
msft = yf.Ticker('MSFT')

# download the candlestick data
msft_candlestick = yf.download('MSFT', start='2020-01-01', end='2024-01-01')

# display the dataframe
msft_candlestick
```

```
[*********************100%%**********************]  1 of 1 completed
```

```
                    Open        High         Low       Close     Adj Close  \
Date
2020-01-02    158.779999  160.729996  158.330002  160.619995  153.938217
2020-01-03    158.320007  159.949997  158.059998  158.619995  152.021423
2020-01-06    157.080002  159.100006  156.509995  159.029999  152.414383
2020-01-07    159.320007  159.669998  157.320007  157.580002  151.024689
2020-01-08    158.929993  160.800003  157.949997  160.089996  153.430298
...                  ...         ...         ...         ...         ...
2023-12-22    373.679993  375.179993  372.709991  374.579987  372.543945
2023-12-26    375.000000  376.940002  373.500000  374.660004  372.623505
```

```
2023-12-27  373.690002  375.059998  372.809998  374.070007  372.036713
2023-12-28  375.369995  376.459991  374.160004  375.279999  373.240112
2023-12-29  376.000000  377.160004  373.480011  376.040009  373.996002

               Volume
Date
2020-01-02   22622100
2020-01-03   21116200
2020-01-06   20813700
2020-01-07   21634100
2020-01-08   27746500
...               ...
2023-12-22   17091100
2023-12-26   12673100
2023-12-27   14905400
2023-12-28   14327000
2023-12-29   18723000

[1006 rows x 6 columns]
```

Stock Information The general information that is associated with a company can be a part of evaluating the current equity of a stock. This has essential information which can be used for analaysis but also to understand the parameters of the stock. This can be done to create subgroups of the demographics.

```python
# .info can be used for information about the company
print(msft.info)
```

{'address1': 'One Microsoft Way', 'city': 'Redmond', 'state': 'WA', 'zip': '98052-6399', 'country': 'United States', 'phone': '425 882 8080', 'website': 'https://www.microsoft.com', 'industry': 'Software - Infrastructure', 'industryKey': 'software-infrastructure', 'industryDisp': 'Software - Infrastructure', 'sector': 'Technology', 'sectorKey': 'technology', 'sectorDisp': 'Technology', 'longBusinessSummary': 'Microsoft Corporation develops and supports software, services, devices and solutions worldwide. The Productivity and Business Processes segment offers office, exchange, SharePoint, Microsoft Teams, office 365 Security and Compliance, Microsoft viva, and Microsoft 365 copilot; and office consumer services, such as Microsoft 365 consumer subscriptions, Office licensed on-premises, and other office services. This segment also provides LinkedIn; and dynamics business solutions, including Dynamics 365, a set of intelligent, cloud-based applications across ERP, CRM, power apps, and power automate; and on-premises ERP and CRM applications. The Intelligent Cloud segment offers server products and cloud services, such as azure and other cloud services; SQL and windows server, visual studio, system center, and related client access licenses, as well as nuance and GitHub; and enterprise services including enterprise support services, industry solutions, and nuance professional services. The More Personal Computing segment offers Windows, including windows OEM licensing and other non-volume licensing of the

Windows operating system; Windows commercial comprising volume licensing of the Windows operating system, windows cloud services, and other Windows commercial offerings; patent licensing; and windows Internet of Things; and devices, such as surface, HoloLens, and PC accessories. Additionally, this segment provides gaming, which includes Xbox hardware and content, and first- and third-party content; Xbox game pass and other subscriptions, cloud gaming, advertising, third-party disc royalties, and other cloud services; and search and news advertising, which includes Bing, Microsoft News and Edge, and third-party affiliates. The company sells its products through OEMs, distributors, and resellers; and directly through digital marketplaces, online, and retail stores. The company was founded in 1975 and is headquartered in Redmond, Washington.', 'fullTimeEmployees': 228000, 'companyOfficers': [{'maxAge': 1, 'name': 'Mr. Satya Nadella', 'age': 56, 'title': 'Chairman & CEO', 'yearBorn': 1967, 'fiscalYear': 2023, 'totalPay': 9276400, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Mr. Bradford L. Smith LCA', 'age': 64, 'title': 'President & Vice Chairman', 'yearBorn': 1959, 'fiscalYear': 2023, 'totalPay': 3591277, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Ms. Amy E. Hood', 'age': 51, 'title': 'Executive VP & CFO', 'yearBorn': 1972, 'fiscalYear': 2023, 'totalPay': 3452196, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Mr. Judson B. Althoff', 'age': 50, 'title': 'Executive VP & Chief Commercial Officer', 'yearBorn': 1973, 'fiscalYear': 2023, 'totalPay': 3355797, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Mr. Christopher David Young', 'age': 51, 'title': 'Executive Vice President of Business Development, Strategy & Ventures', 'yearBorn': 1972, 'fiscalYear': 2023, 'totalPay': 2460507, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Ms. Carolina Dybeck Happe', 'age': 51, 'title': 'Executive VP & COO', 'yearBorn': 1972, 'fiscalYear': 2023, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Ms. Alice L. Jolla', 'age': 57, 'title': 'Corporate VP & Chief Accounting Officer', 'yearBorn': 1966, 'fiscalYear': 2023, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Mr. James Kevin Scott', 'age': 51, 'title': 'Executive VP of AI & CTO', 'yearBorn': 1972, 'fiscalYear': 2023, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Brett Iversen', 'title': 'Vice President of Investor Relations', 'fiscalYear': 2023, 'exercisedValue': 0, 'unexercisedValue': 0}, {'maxAge': 1, 'name': 'Mr. Hossein Nowbar', 'title': 'Chief Legal Officer', 'fiscalYear': 2023, 'exercisedValue': 0, 'unexercisedValue': 0}], 'auditRisk': 3, 'boardRisk': 4, 'compensationRisk': 2, 'shareHolderRightsRisk': 2, 'overallRisk': 1, 'governanceEpochDate': 1726617600, 'compensationAsOfEpochDate': 1703980800, 'irWebsite': 'http://www.microsoft.com/investor/default.aspx', 'maxAge': 86400, 'priceHint': 2, 'previousClose': 438.69, 'open': 437.22, 'dayLow': 434.22, 'dayHigh': 439.21, 'regularMarketPreviousClose': 438.69, 'regularMarketOpen': 437.22, 'regularMarketDayLow': 434.22, 'regularMarketDayHigh': 439.21, 'dividendRate': 3.32, 'dividendYield': 0.0076, 'exDividendDate': 1732147200, 'payoutRatio': 0.2483, 'fiveYearAvgDividendYield': 0.9, 'beta': 0.896, 'trailingPE': 36.918575, 'forwardPE': 28.579777, 'volume': 55167106, 'regularMarketVolume': 55167106, 'averageVolume': 20156792, 'averageVolume10days': 21591850, 'averageDailyVolume10Day': 21591850, 'bid': 435.35, 'ask': 460.22, 'bidSize':

100, 'askSize': 100, 'marketCap': 3235379150848, 'fiftyTwoWeekLow': 309.45,
'fiftyTwoWeekHigh': 468.35, 'priceToSalesTrailing12Months': 13.199057,
'fiftyDayAverage': 421.6462, 'twoHundredDayAverage': 413.9077,
'trailingAnnualDividendRate': 3.0, 'trailingAnnualDividendYield': 0.006838542,
'currency': 'USD', 'enterpriseValue': 3257699663872, 'profitMargins': 0.35956,
'floatShares': 7422780788, 'sharesOutstanding': 7433039872, 'sharesShort':
65801156, 'sharesShortPriorMonth': 57774674, 'sharesShortPreviousMonthDate':
1722384000, 'dateShortInterest': 1724976000, 'sharesPercentSharesOut': 0.0089,
'heldPercentInsiders': 0.00054000004, 'heldPercentInstitutions': 0.73737997,
'shortRatio': 2.99, 'shortPercentOfFloat': 0.0089, 'impliedSharesOutstanding':
7433039872, 'bookValue': 36.115, 'priceToBook': 12.052332, 'lastFiscalYearEnd':
1719705600, 'nextFiscalYearEnd': 1751241600, 'mostRecentQuarter': 1719705600,
'earningsQuarterlyGrowth': 0.097, 'netIncomeToCommon': 88135999488,
'trailingEps': 11.79, 'forwardEps': 15.23, 'pegRatio': 2.3, 'lastSplitFactor':
'2:1', 'lastSplitDate': 1045526400, 'enterpriseToRevenue': 13.29,
'enterpriseToEbitda': 25.169, '52WeekChange': 0.3707564, 'SandP52WeekChange':
0.31472707, 'lastDividendValue': 0.75, 'lastDividendDate': 1723680000,
'exchange': 'NMS', 'quoteType': 'EQUITY', 'symbol': 'MSFT', 'underlyingSymbol':
'MSFT', 'shortName': 'Microsoft Corporation', 'longName': 'Microsoft
Corporation', 'firstTradeDateEpochUtc': 511108200, 'timeZoneFullName':
'America/New_York', 'timeZoneShortName': 'EDT', 'uuid': 'b004b3ec-
de24-385e-b2c1-923f10d3fb62', 'messageBoardId': 'finmb_21835',
'gmtOffSetMilliseconds': -14400000, 'currentPrice': 435.27, 'targetHighPrice':
600.0, 'targetLowPrice': 440.0, 'targetMeanPrice': 496.38, 'targetMedianPrice':
500.0, 'recommendationMean': 1.7, 'recommendationKey': 'buy',
'numberOfAnalystOpinions': 47, 'totalCash': 75531001856, 'totalCashPerShare':
10.162, 'ebitda': 129433001984, 'totalDebt': 97851998208, 'quickRatio': 1.141,
'currentRatio': 1.275, 'totalRevenue': 245122007040, 'debtToEquity': 36.447,
'revenuePerShare': 32.986, 'returnOnAssets': 0.14802, 'returnOnEquity': 0.37133,
'freeCashflow': 56705249280, 'operatingCashflow': 118547996672,
'earningsGrowth': 0.097, 'revenueGrowth': 0.152, 'grossMargins': 0.69764,
'ebitdaMargins': 0.52804, 'operatingMargins': 0.43143, 'financialCurrency':
'USD', 'trailingPegRatio': 2.3468}

Balance Sheets The financial and balance sheets of an equity are must be released so that investors can make educated decisions when looking at which stock to invest in.

```
# .get_balance_sheets returns the balance sheet information
print(msft.get_balance_sheet()[0:5])
```

	2024-06-30	2023-06-30	2022-06-30 \
OrdinarySharesNumber	7434000000.0	7432000000.0	7464000000.0
ShareIssued	7434000000.0	7432000000.0	7464000000.0
NetDebt	33315000000.0	12533000000.0	35850000000.0
TotalDebt	67127000000.0	59965000000.0	61270000000.0
TangibleBookValue	121660000000.0	128971000000.0	87720000000.0

	2021-06-30

```
OrdinarySharesNumber       7519000000.0
ShareIssued                7519000000.0
NetDebt                   43922000000.0
TotalDebt                 67775000000.0
TangibleBookValue         84477000000.0
```

1.1.2 Stock Data

There are thousands of equities to choose from on the stock market, to reduce the overload of data that will be used in training the trading techniques sector ETF's will be used instead.

Exchange Traded Funds (ETF) Electronically Traded Funds or ETF's are an investment equity which has a basket of securities that tracks or seeks to outperform indexes. It takes a smaller proportion of a large number of individual stocks so that the equity is a broader representation of a specific idea. For example if you want to invest in small-cap companies, there are a number of ETF's that can be used to take a smaller investment of a hundred companies rather than one. This looks to spread out investments to reduce risk.

Economic Time Periods The macroeconomic business cycle is the representation of overall economic activity within the economy. During different periods of economic activity the market sectors will react differently. The following are the typical components of the macroeconomic business cycle.

- Trough : This is the low point of the economy where supply is in excess and demand is in shortage. This can turn into a recession which is where there are two consecutive quarters where the GDP growth is negative.

- Expansion : Following a trough, there is a period of economic growth which leads to an increase in consumer and business confidence. This will result in an increase in wage growth and a decrease in unemployment. Expansion periods should be steady and not exponential to ensure prolonged prosperity.

- Peak : A peak occurs when an expansion has reached its limit and there ends up being downward pressure on the growth rate within economic activity. Businesses will then reevaluate their decision making to review budgets and other outlooks.

- Contraction : This is when growth slows, unemployment increases and business investment falls.

```python
# create time periods for where this takes place
economic_cycle_periods = {

    "trough": ("2008-10-01", "2009-06-01"),
    "expansion": ("2012-01-01", "2015-01-01"),
    "peak": ("2019-06-01", "2020-02-01"),
    "contraction": ("2007-12-01", "2008-10-01"),
}

economic_cycle_periods_list = ['trough','expansion','peak','contraction']
```

GICS Sectors According to Global Industry Classification Standard there are 11 major sectors at play within the economy. Each with their own industries and sub-industries but this capstone project there will be a concentration on the sectors.

For the capstone project a concentration on ETF's that represent the sectors will be used to help understand how different sectors will perform. The following ETF equities are going to be invested in based on their sector. More information about the equities can be found in the report.

- Materials : XLB
- Industrials : XLI
- Financials : XLF
- Information Technology : XLK
- Consumer Discretionary : XLY
- Consumer Staples : XLP
- Energy : XLE
- Healthcare : XLV
- Communication Services : VOX
- Utilities : XLU
- Real Estate : IYR

```python
# create etf tickers for sectors
sector_etf_tickers = [
    'XLB', # materials sector
    'XLI', # industrials sector
    'XLF', # financials
    'XLK', # information technology
    'XLY', # consumer discretionary
    'XLP', # consumer staples
    'XLE', # energy
    'XLV', # healthcare
    'VOX', # communication services
    'XLU', # utilities
    'IYR' # real estate
    ]
```

1.1.3 Nested Dictionary Storage of Relevant Data

Using the YFinance API load in candlestick data with the 11 industry sectors during the four different economic time periods. This can be used to evaluate sector etfs during different economic time periods using the daily adjusted closed price.

```python
# save nested dictionary data as a variable to be accessed.
sector_etf_data =␣
 ↪download_stock_data_for_periods(sector_etf_tickers,economic_cycle_periods)
```

```
[*********************100%%**********************]  1 of 1 completed
[*********************100%%**********************]  1 of 1 completed
[*********************100%%**********************]  1 of 1 completed
[*********************100%%**********************]  1 of 1 completed
[*********************100%%**********************]  1 of 1 completed
```

```
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
[********************100%%**********************]  1 of 1 completed
```

```python
# navigate the nested dictionary
sector_etf_data
```

```
{'trough': {'XLB':                 Open       High        Low      Close  Adj
   Close      Volume
   Date
   2008-10-01  32.759998  33.189999  32.130001  32.849998  23.222452  14639500
```

```
2008-10-02   31.540001   31.860001   29.930000   30.490000   21.554119   12581300
2008-10-03   30.190001   31.690001   29.780001   30.190001   21.342035   16770600
2008-10-06   29.510000   29.510000   26.889999   28.700001   20.288721   22512700
2008-10-07   29.160000   29.530001   27.049999   27.219999   19.242472   16004900

...             ...         ...         ...         ...         ...
2009-05-22   26.530001   26.660000   26.110001   26.299999   18.932842    8421500
2009-05-26   26.170000   26.969999   25.830000   26.930000   19.386372    7886900
2009-05-27   26.790001   26.850000   25.860001   25.920000   18.659296    7216600
2009-05-28   26.190001   26.440001   25.760000   26.379999   18.990435    8773400
2009-05-29   26.670000   27.200001   26.500000   27.170000   19.559145    7792200


[166 rows x 6 columns],
 'XLI':                 Open        High         Low       Close   Adj Close
Volume
  Date
2008-10-01   30.309999   30.309999   29.379999   30.070000   21.929304   21014600
2008-10-02   29.629999   29.629999   27.809999   28.270000   20.616604   14679300
2008-10-03   28.500000   28.930000   27.580000   27.820000   20.288429   18081900
2008-10-06   26.870001   27.209999   25.570000   27.090000   19.756060   20767300
2008-10-07   27.570000   27.780001   25.820000   26.190001   19.099714   15985000

...             ...         ...         ...         ...         ...
2009-05-22   21.750000   21.990000   21.570000   21.690001   16.125713    8858200
2009-05-26   21.840000   22.510000   21.510000   22.480000   16.713049   16082300
2009-05-27   22.480000   22.580000   21.750000   21.750000   16.170326   10835700
2009-05-28   21.950001   22.090000   21.469999   21.959999   16.326450   10851400
2009-05-29   22.160000   22.520000   21.930000   22.490000   16.720490   10999600


[166 rows x 6 columns],
 'XLF':                 Open        High         Low       Close   Adj Close
Volume
  Date
2008-10-01   16.133224   16.791227   15.913891   16.791227   12.458467   235842858
2008-10-02   16.645004   16.815598   15.913891   15.954509   11.837652   213501932
2008-10-03   16.401300   16.774981   15.255890   15.255890   11.319306   322736194
2008-10-06   14.638505   15.036556   13.996751   14.459789   10.728626   420354002
2008-10-07   14.711617   14.744110   12.932575   12.932575    9.595492   436457082

...             ...         ...         ...         ...         ...
2009-05-22    9.626320    9.650690    9.431357    9.471974    7.204327   118089215
2009-05-26    9.358245    9.821284    9.350122    9.780666    7.439115   185271163
2009-05-27    9.870024    9.902518    9.455727    9.488221    7.216683   211276284
2009-05-28    9.561332    9.788790    9.398863    9.756296    7.420582   228614303
2009-05-29    9.837530   10.032494    9.650690    9.935012    7.556509   231964224


[166 rows x 6 columns],
 'XLK':                 Open        High         Low       Close   Adj Close
Volume
  Date
```

```
2008-10-01   19.680000   19.770000   19.280001   19.670000   15.677793   13473700
2008-10-02   19.260000   19.379999   18.670000   18.900000   15.064079   10857900
2008-10-03   19.059999   19.530001   18.549999   18.629999   14.848871   13053800
2008-10-06   18.150000   18.270000   17.059999   17.600000   14.027914   20373400
2008-10-07   18.049999   18.070000   16.700001   16.700001   13.310583   14862500

…           …           …           …           …           …
2009-05-22   16.950001   17.040001   16.700001   16.940001   13.642543    5434700
2009-05-26   16.760000   17.420000   16.700001   17.379999   13.996900    5743400
2009-05-27   17.389999   17.600000   17.200001   17.209999   13.859987    7714900
2009-05-28   17.420000   17.540001   17.110001   17.459999   14.061328    6177900
2009-05-29   17.510000   17.680000   17.340000   17.650000   14.214342    5396800

[166 rows x 6 columns],
'XLY':            Open        High         Low       Close   Adj Close   Volume
Date
2008-10-01   28.850000   28.850000   27.480000   27.900000   22.636211    9123600
2008-10-02   27.900000   27.900000   26.610001   26.889999   21.816759    7379700
2008-10-03   27.260000   27.340000   25.910000   25.950001   21.054104    7869500
2008-10-06   25.370001   25.730000   23.740000   25.219999   20.461824   14043000
2008-10-07   25.650000   25.650000   23.270000   23.600000   19.147476    9430200

…           …           …           …           …           …
2009-05-22   22.540001   22.740000   22.150000   22.580000   18.558596    7362500
2009-05-26   22.389999   23.410000   22.110001   23.340000   19.183247   11048600
2009-05-27   23.250000   23.480000   22.730000   22.830000   18.764082    7775100
2009-05-28   23.090000   23.139999   22.350000   22.740000   18.690109    5382100
2009-05-29   22.900000   23.180000   22.600000   23.120001   19.002430    4351000

[166 rows x 6 columns],
'XLP':            Open        High         Low       Close   Adj Close   Volume
Date
2008-10-01   27.530001   28.080000   26.980000   27.799999   18.105207    8520100
2008-10-02   27.850000   27.850000   27.340000   27.450001   17.877258    5136400
2008-10-03   28.549999   28.750000   27.200001   27.230000   17.733984    5895100
2008-10-06   26.540001   27.020000   25.500000   26.440001   17.219488    9393600
2008-10-07   27.320000   27.320000   25.350000   25.600000   16.672422    6318900

…           …           …           …           …           …
2009-05-22   22.760000   23.129999   22.760000   22.930000   15.150266    6040400
2009-05-26   22.920000   23.250000   22.780001   23.219999   15.341876    5478700
2009-05-27   23.200001   23.260000   22.629999   22.629999   14.952051    6976400
2009-05-28   22.760000   23.000000   22.570000   22.870001   15.110631    7246000
2009-05-29   22.930000   23.180000   22.840000   23.120001   15.275807    6894700

[166 rows x 6 columns],
'XLE':            Open        High         Low       Close   Adj Close   Volume
```

```
Date
2008-10-01  62.700001  63.430000  60.000000  62.259998  37.702045  46615500
2008-10-02  61.299999  61.419998  57.619999  58.709999  35.552319  46997100
2008-10-03  58.500000  69.919998  57.490002  58.009998  35.128418  48539600
2008-10-06  55.759998  55.930000  50.980000  54.889999  33.239082  69545000
2008-10-07  56.540001  57.279999  51.540001  51.880001  31.416355  48025300

...             ...        ...        ...        ...        ...
2009-05-22  48.900002  49.459999  48.500000  48.610001  29.740948  13646000
2009-05-26  48.290001  49.779999  47.650002  49.639999  30.371140  26906900
2009-05-27  49.849998  50.500000  49.029999  49.099998  30.040745  20871300
2009-05-28  49.540001  51.209999  49.060001  50.730000  31.038040  25906900
2009-05-29  51.540001  51.889999  50.970001  51.680000  31.619278  22238700

[166 rows x 6 columns],
'XLV':                 Open       High        Low      Close    Adj Close
Volume
Date
2008-10-01  30.100000  30.480000  30.100000  30.250000  23.013002   6053600
2008-10-02  30.250000  30.590000  29.930000  30.299999  23.051035   6353400
2008-10-03  30.600000  30.600000  29.650000  29.650000  22.556545   6814400
2008-10-06  29.400000  29.879999  27.410000  28.540001  21.712105   8545000
2008-10-07  28.719999  28.780001  27.389999  27.850000  21.187176   5060200

...             ...        ...        ...        ...        ...
2009-05-22  25.280001  25.400000  25.070000  25.290001  19.476892   3655700
2009-05-26  25.190001  25.660000  24.889999  25.520000  19.654018   4412900
2009-05-27  25.549999  25.600000  25.219999  25.260000  19.453779   4591100
2009-05-28  25.209999  25.590000  25.139999  25.389999  19.553904   5720000
2009-05-29  25.360001  25.840000  25.290001  25.830000  19.892765   6549200

[166 rows x 6 columns],
'VOX':                 Open       High        Low      Close    Adj Close
Volume
Date
2008-10-01  53.180000  53.470001  52.439999  53.310001  36.781956     20600
2008-10-02  52.939999  53.650002  51.310001  51.730000  35.691811     14600
2008-10-03  51.930000  53.020000  50.970001  50.970001  35.167450     54800
2008-10-06  48.990002  50.259998  46.770000  48.580002  33.518425     39300
2008-10-07  48.009998  49.250000  46.500000  46.500000  32.083305    124800

...             ...        ...        ...        ...        ...
2009-05-22  46.830002  46.889999  46.360001  46.580002  33.158119     37500
2009-05-26  46.299999  48.360001  46.299999  48.250000  34.346920     32200
2009-05-27  48.250000  48.680000  47.759998  47.770000  34.005234     36200
2009-05-28  48.090000  48.560001  47.189999  48.380001  34.439465     29800
2009-05-29  48.619999  48.660000  47.869999  48.660000  34.638771     40800

[166 rows x 6 columns],
'XLU':                 Open       High        Low      Close    Adj Close
```

```
Volume
  Date
  2008-10-01  33.330002  33.410000  32.480000  33.090000  18.792999  6724500
  2008-10-02  32.799999  33.360001  32.200001  32.610001  18.520390  5241300
  2008-10-03  32.779999  33.279999  31.980000  32.119999  18.242102  8126700
  2008-10-06  31.280001  32.299999  29.500000  30.549999  17.350439  9640900
  2008-10-07  30.520000  31.120001  29.219999  29.299999  16.640511  9039900

  ...                ...           ...          ...          ...          ...          ...
  2009-05-22  25.559999  26.059999  25.559999  25.830000  15.019250  4892500
  2009-05-26  25.660000  26.580000  25.580000  26.580000  15.455345  7123400
  2009-05-27  26.629999  26.629999  25.980000  26.040001  15.141353  5936900
  2009-05-28  26.080000  26.709999  26.080000  26.580000  15.455345  6475000
  2009-05-29  26.629999  26.860001  26.350000  26.770000  15.565824  5180300


  [166 rows x 6 columns],
  'IYR':                 Open       High        Low        Close   Adj Close
Volume
  Date
  2008-10-01  61.169998  61.169998  59.950001  61.000000  34.248367  11929100
  2008-10-02  60.000000  60.320000  56.459999  56.959999  31.980118  15561600
  2008-10-03  57.990002  58.560001  53.919998  53.919998  30.273306  23607200
  2008-10-06  51.610001  53.000000  50.029999  52.930000  29.717470  23349100
  2008-10-07  53.240002  53.250000  48.419998  48.459999  27.207794  22293000

  ...                ...           ...          ...          ...          ...          ...
  2009-05-22  32.480000  32.880001  31.590000  31.709999  18.670605  24961700
  2009-05-26  31.400000  33.639999  31.250000  33.349998  19.636223  42247200
  2009-05-27  33.529999  33.529999  31.950001  32.240002  18.982670  36683800
  2009-05-28  32.619999  33.049999  31.730000  32.849998  19.341831  31800800
  2009-05-29  33.099998  33.770000  32.490002  33.750000  19.871742  28518600


  [166 rows x 6 columns]},
  'expansion': {'XLB':                 Open       High        Low        Close
Adj Close     Volume
  Date
  2012-01-03  34.240002  34.590000  34.230000  34.430000  26.606606  11213800
  2012-01-04  34.389999  34.750000  34.250000  34.720001  26.830727   8772400
  2012-01-05  34.419998  34.950001  34.320000  34.820000  26.907982   7783700
  2012-01-06  34.700001  34.889999  34.549999  34.770000  26.869352   7218700
  2012-01-09  34.820000  34.970001  34.680000  34.840000  26.923445   9204600

  ...                ...           ...          ...          ...          ...          ...
  2014-12-24  49.049999  49.240002  48.980000  49.060001  40.491211   1878800
  2014-12-26  49.360001  49.389999  49.160000  49.230000  40.631504   1297000
  2014-12-29  49.110001  49.369999  49.110001  49.230000  40.631504   2173200
  2014-12-30  49.119999  49.279999  49.040001  49.060001  40.491211   2167600
  2014-12-31  49.169998  49.169998  48.570000  48.580002  40.095047   2984400


  [754 rows x 6 columns],
```

```
'XLI':                   Open        High         Low       Close   Adj Close
Volume
  Date
  2012-01-03  34.410000  34.810001  34.410000  34.490002  27.187344  19704600
  2012-01-04  34.380001  34.689999  34.299999  34.650002  27.313467  10290300
  2012-01-05  34.430000  34.680000  34.150002  34.660000  27.321346   8828900
  2012-01-06  34.740002  34.740002  34.480000  34.590000  27.266167   8702400
  2012-01-09  34.689999  34.900002  34.540001  34.849998  27.471115  15523400

  …                …           …           …           …           …
  2014-12-24  57.470001  57.560001  57.349998  57.400002  48.113762   1958200
  2014-12-26  57.599998  57.599998  57.400002  57.400002  48.113762   5733900
  2014-12-29  57.330002  57.540001  56.910000  57.490002  48.189194   3566200
  2014-12-30  57.730000  57.730000  57.070000  57.150002  47.904190   3602800
  2014-12-31  57.380001  57.380001  56.549999  56.580002  47.426414   5572900

  [754 rows x 6 columns],
  'XLF':                   Open        High         Low       Close   Adj Close
Volume
  Date
  2012-01-03  10.787977  10.942323  10.779854  10.836718   8.567098  103362023
  2012-01-04  10.755483  10.844842  10.682372  10.804224   8.541410   69833891
  2012-01-05  10.706742  11.007311  10.658002  10.950447   8.657010   89935260
  2012-01-06  10.958570  10.958570  10.812348  10.885459   8.605632   83878617
  2012-01-09  10.926076  10.991064  10.869212  10.942323   8.650586   69189586

  …                …           …           …           …           …
  2014-12-24  20.365557  20.365557  20.259951  20.268076  16.865993    9773278
  2014-12-26  20.333063  20.341187  20.259951  20.268076  16.865993   18149002
  2014-12-29  20.251827  20.422421  20.211210  20.341187  16.926832   22126486
  2014-12-30  20.308693  20.365557  20.251827  20.316816  16.906549   20298451
  2014-12-31  20.389927  20.406174  20.064987  20.089357  16.717278   29869969

  [754 rows x 6 columns],
  'XLK':                   Open        High         Low       Close   Adj Close
Volume
  Date
  2012-01-03  25.870001  26.020000  25.790001  25.809999  21.661152  26839900
  2012-01-04  25.799999  25.930000  25.660000  25.879999  21.719894   4744300
  2012-01-05  25.850000  26.010000  25.730000  25.959999  21.787037   6756900
  2012-01-06  25.980000  26.030001  25.850000  25.969999  21.795429   4989200
  2012-01-09  26.059999  26.080000  25.860001  25.930000  21.761868   4086600

  …                …           …           …           …           …
  2014-12-24  42.279999  42.279999  42.130001  42.139999  37.335186   2655600
  2014-12-26  42.220001  42.400002  42.200001  42.320000  37.494671   2278500
  2014-12-29  42.160000  42.279999  42.070000  42.110001  37.308605   4638300
  2014-12-30  42.029999  42.090000  41.770000  41.810001  37.042824   5503400
  2014-12-31  41.950001  42.000000  41.310001  41.349998  36.635265   5955300
```

```
[754 rows x 6 columns],
 'XLY':                 Open       High        Low      Close   Adj Close
Volume
 Date
 2012-01-03  39.669998  39.810001  39.349998  39.349998  33.771416  12055500
 2012-01-04  39.279999  39.750000  39.279999  39.660000  34.037468   3118300
 2012-01-05  39.439999  39.990002  39.230000  39.990002  34.320690   4576300
 2012-01-06  40.009998  40.200001  39.799999  40.049999  34.372166   3092600
 2012-01-09  40.160000  40.160000  39.830002  39.990002  34.320690   3071400
 ...               ...        ...        ...        ...        ...       ...
 2014-12-24  72.150002  72.150002  71.820000  71.839996  64.390564   1687800
 2014-12-26  72.040001  72.379997  71.940002  72.230003  64.740112  14102300
 2014-12-29  72.190002  72.910004  72.029999  72.779999  65.233078  23013600
 2014-12-30  72.580002  72.760002  72.290001  72.440002  64.928329   6051200
 2014-12-31  72.669998  72.970001  72.120003  72.150002  64.668419   6664000

 [754 rows x 6 columns],
 'XLP':                 Open       High        Low      Close   Adj Close
Volume
 Date
 2012-01-03  32.689999  32.779999  32.450001  32.459999  23.236197  20757900
 2012-01-04  32.400002  32.450001  32.290001  32.389999  23.186098   6855200
 2012-01-05  32.330002  32.400002  32.169998  32.369999  23.171782   4563000
 2012-01-06  32.419998  32.419998  32.099998  32.160000  23.021448   6365900
 2012-01-09  32.240002  32.250000  32.119999  32.209999  23.057238   3304900
 ...               ...        ...        ...        ...        ...       ...
 2014-12-24  49.580002  49.630001  49.349998  49.380001  38.331535   1478000
 2014-12-26  49.430000  49.610001  49.410000  49.459999  38.393627   3302400
 2014-12-29  49.400002  49.450001  49.110001  49.310001  38.277206  13478700
 2014-12-30  49.240002  49.349998  49.070000  49.090000  38.106415   5094900
 2014-12-31  49.270000  49.270000  48.470001  48.490002  37.640667   5218700

 [754 rows x 6 columns],
 'XLE':                 Open       High        Low      Close   Adj Close
Volume
 Date
 2012-01-03  70.599998  71.400002  70.489998  71.050003  45.571899  24254800
 2012-01-04  70.779999  71.320000  70.360001  71.199997  45.668098  14018500
 2012-01-05  70.739998  70.910004  70.000000  70.860001  45.450035  12674700
 2012-01-06  70.989998  71.019997  70.290001  70.440002  45.180645   8893100
 2012-01-09  70.589996  70.940002  70.379997  70.779999  45.398720  12043200
 ...               ...        ...        ...        ...        ...       ...
 2014-12-24  80.250000  80.440002  79.199997  80.129997  54.443989  14047800
 2014-12-26  80.440002  80.809998  79.760002  80.110001  54.430386  14464600
 2014-12-29  80.290001  81.000000  79.919998  80.400002  54.627441  18224100
 2014-12-30  80.300003  80.470001  79.510002  79.760002  54.192585  19148100
 2014-12-31  79.290001  80.000000  78.690002  79.160004  53.784924  21541900
```

```
[754 rows x 6 columns],
 'XLV':                 Open       High        Low      Close   Adj Close
Volume
 Date
 2012-01-03  35.099998  35.259998  35.029999  35.139999  28.565081  18864300
 2012-01-04  35.110001  35.130001  34.849998  34.970001  28.426870   6174000
 2012-01-05  34.779999  35.110001  34.709999  35.070000  28.508163   5939700
 2012-01-06  35.119999  35.209999  34.959999  35.130001  28.556940   6762500
 2012-01-09  35.200001  35.220001  34.910000  35.189999  28.605717  10695300
 ...               ...        ...        ...        ...        ...       ...
 2014-12-24  68.190002  69.029999  68.129997  68.639999  58.791073   4519100
 2014-12-26  68.910004  69.320000  68.870003  69.129997  59.210770   4242800
 2014-12-29  69.059998  69.470001  68.980003  69.370003  59.416321   5225900
 2014-12-30  69.260002  69.419998  68.989998  69.070000  59.159389   4965100
 2014-12-31  69.269997  69.660004  68.370003  68.379997  58.568378   7863600

[754 rows x 6 columns],
 'VOX':                 Open       High        Low      Close   Adj Close
Volume
 Date
 2012-01-03  63.029999  63.029999  62.619999  62.630001  48.678196   131400
 2012-01-04  62.570000  62.680000  61.900002  61.959999  48.157440    19100
 2012-01-05  62.070000  62.070000  61.009998  61.430000  47.745514    39000
 2012-01-06  61.599998  61.840000  61.110001  61.340000  47.675571    24300
 2012-01-09  61.500000  61.580002  61.090000  61.439999  47.753292    22900
 ...               ...        ...        ...        ...        ...      ...
 2014-12-24  85.440002  85.669998  85.260002  85.550003  71.577934    51700
 2014-12-26  85.889999  86.209999  85.570000  86.110001  72.046455    58100
 2014-12-29  86.120003  86.230003  85.820000  86.129997  72.063187    34300
 2014-12-30  86.050003  86.050003  85.309998  85.510002  71.544472    51100
 2014-12-31  85.459999  85.540001  84.720001  84.720001  70.883453    62000

[754 rows x 6 columns],
 'XLU':                 Open       High        Low      Close   Adj Close
Volume
 Date
 2012-01-03  36.180000  36.250000  35.230000  35.340000  23.079243  19184900
 2012-01-04  35.240002  35.450001  35.130001  35.139999  22.948627   5775100
 2012-01-05  35.099998  35.320000  34.910000  35.209999  22.994341  12110000
 2012-01-06  35.270000  35.299999  34.980000  35.020000  22.870253   7305500
 2012-01-09  35.130001  35.200001  34.910000  35.070000  22.902910   4527300
 ...               ...        ...        ...        ...        ...       ...
 2014-12-24  47.000000  48.090000  47.000000  47.970001  35.108250  14412400
 2014-12-26  47.950001  48.759998  47.950001  48.549999  35.532734  19854100
 2014-12-29  48.580002  49.330002  48.549999  49.110001  35.942581  24520900
 2014-12-30  49.070000  49.080002  48.020000  48.090000  35.196064  13015400
```

```
    2014-12-31  48.209999  48.290001  47.160000  47.220001  34.559338  15156500

    [754 rows x 6 columns],
    'IYR':                    Open        High         Low       Close   Adj Close
Volume
    Date
    2012-01-03  57.740002  58.310001  56.779999  57.160000  37.618160   9054000
    2012-01-04  56.980000  57.040001  56.290001  56.520000  37.196949  12839200
    2012-01-05  56.340000  57.040001  56.060001  56.930000  37.466789   6432700
    2012-01-06  57.099998  57.250000  56.570000  56.830002  37.400986   4222200
    2012-01-09  56.930000  57.040001  56.529999  56.669998  37.295681   4740100
    ...                ...         ...         ...         ...         ...        ...
    2014-12-24  77.970001  78.260002  77.430000  77.480003  57.114262   4057500
    2014-12-26  77.650002  77.970001  77.629997  77.779999  57.335396   2903000
    2014-12-29  77.699997  78.440002  77.480003  78.120003  57.586056   3913300
    2014-12-30  77.889999  78.410004  77.830002  78.040001  57.527061   7674500
    2014-12-31  78.260002  78.620003  76.760002  76.839996  56.642494   9269300

    [754 rows x 6 columns]},
    'peak': {'XLB':               Open        High         Low       Close      Adj
Close      Volume
    Date
    2019-06-03  52.900002  54.430000  52.750000  54.410000  48.822552  13155400
    2019-06-04  55.000000  56.049999  54.680000  55.939999  50.195427  17777900
    2019-06-05  56.290001  56.480000  55.509998  56.349998  50.563332  19704000
    2019-06-06  56.630001  57.209999  56.430000  57.020000  51.164524  19911600
    2019-06-07  57.389999  57.740002  57.270000  57.529999  51.622154  23211100
    ...                ...         ...         ...         ...         ...        ...
    2020-01-27  58.680000  58.869999  58.450001  58.450001  53.318943   6934600
    2020-01-28  58.740002  59.299999  58.610001  59.110001  53.920998   4491100
    2020-01-29  59.310001  59.630001  59.270000  59.299999  54.094322   3569200
    2020-01-30  58.520000  59.009998  58.139999  58.930000  53.756805   8699100
    2020-01-31  58.389999  58.430000  57.439999  57.630001  52.570923  11020700

    [169 rows x 6 columns],
    'XLI':                    Open        High         Low       Close   Adj Close
Volume
    Date
    2019-06-03  71.930000  72.940002  71.910004  72.540001  66.340546  17515000
    2019-06-04  73.239998  74.300003  73.110001  74.269997  67.922676  16424700
    2019-06-05  74.660004  75.080002  74.400002  74.989998  68.581139  12865600
    2019-06-06  75.070000  75.320000  74.440002  75.120003  68.700050  11387300
    2019-06-07  75.400002  76.029999  75.349998  75.720001  69.248772   9234300
    ...                ...         ...         ...         ...         ...        ...
    2020-01-27  82.059998  82.519997  81.970001  81.970001  76.065933  15013800
    2020-01-28  82.239998  82.910004  81.860001  82.500000  76.557755  14918000
    2020-01-29  83.330002  83.570000  82.849998  82.940002  76.966064  10355100
```

```
2020-01-30  82.139999  83.120003  81.970001  83.029999  77.049583  13158200
2020-01-31  82.580002  82.699997  80.800003  81.110001  75.267868  19083600

[169 rows x 6 columns],
'XLF':                Open       High        Low      Close  Adj Close
Volume
 Date
2019-06-03  25.969999  26.290001  25.920000  26.180000  23.556055  63329300
2019-06-04  26.500000  26.920000  26.430000  26.889999  24.194893  45454400
2019-06-05  26.940001  27.080000  26.709999  27.059999  24.347853  46599700
2019-06-06  27.070000  27.270000  26.920000  27.160000  24.437830  37297600
2019-06-07  27.160000  27.250000  27.080000  27.120001  24.401840  33054800
...               ...        ...        ...        ...        ...       ...
2020-01-27  29.790001  30.059999  29.750000  29.920000  27.334768  50289300
2020-01-28  30.080000  30.400000  30.030001  30.270000  27.654528  34257900
2020-01-29  30.370001  30.480000  30.170000  30.219999  27.608847  41167500
2020-01-30  29.950001  30.620001  29.930000  30.600000  27.956013  51210000
2020-01-31  30.340000  30.379999  29.850000  29.959999  27.371309  58958800

[169 rows x 6 columns],
'XLK':                Open       High        Low      Close  Adj Close
Volume
 Date
2019-06-03  72.000000  72.400002  70.150002  70.629997  67.064896  19703700
2019-06-04  71.529999  72.989998  71.169998  72.989998  69.305786  15096200
2019-06-05  73.959999  74.110001  73.099998  73.940002  70.207840  14395800
2019-06-06  74.139999  74.949997  73.800003  74.800003  71.024452  11922500
2019-06-07  75.190002  76.639999  75.040001  76.190002  72.344292  11992100
...               ...        ...        ...        ...        ...       ...
2020-01-27  95.160004  95.860001  94.709999  95.110001  91.207321  17787300
2020-01-28  95.940002  97.190002  95.739998  96.930000  92.952614  15592200
2020-01-29  97.730003  97.870003  96.790001  97.180000  93.192352  10563300
2020-01-30  97.139999  98.110001  96.550003  98.019997  93.997894   9966800
2020-01-31  97.449997  97.500000  94.989998  95.330002  91.418282  13482300

[169 rows x 6 columns],
'XLY':                Open        High         Low       Close   Adj Close
\
 Date
2019-06-03  110.669998  111.059998  109.430000  110.010002  104.762810
2019-06-04  111.000000  113.029999  110.040001  112.989998  107.600685
2019-06-05  113.639999  113.779999  112.430000  113.550003  108.133972
2019-06-06  113.550003  114.410004  113.080002  114.129997  108.686302
2019-06-07  114.629997  116.129997  114.449997  115.779999  110.257584
...                ...         ...         ...         ...         ...
2020-01-27  123.230003  124.709999  122.989998  124.070000  119.345024
2020-01-28  124.669998  125.750000  124.190002  125.279999  120.508957
```

```
2020-01-29  125.690002  126.180000  125.120003  125.209999  120.441589
2020-01-30  124.230003  125.129997  123.980003  125.000000  120.239609
2020-01-31  127.559998  127.730003  125.059998  125.320000  120.547440

                 Volume
Date
2019-06-03  8065600
2019-06-04  5739500
2019-06-05  5258600
2019-06-06  3659800
2019-06-07  4044900
...              ...
2020-01-27  7877100
2020-01-28  5907700
2020-01-29  2930000
2020-01-30  5230500
2020-01-31  8422000

[169 rows x 6 columns],
 'XLP':                Open       High        Low      Close   Adj Close
Volume
 Date
 2019-06-03  55.689999  56.419998  55.680000  56.310001  48.905941  17060600
 2019-06-04  56.660000  56.830002  56.480000  56.820000  49.348885  16546800
 2019-06-05  57.080002  57.580002  56.860001  57.470001  49.913422  15069700
 2019-06-06  57.590000  58.189999  57.490002  58.070000  50.434528  20769600
 2019-06-07  58.270000  58.860001  58.270000  58.630001  50.920895  18860800
 ...              ...        ...        ...        ...        ...       ...
 2020-01-27  63.049999  63.689999  63.009998  63.389999  56.259918   8748400
 2020-01-28  63.430000  63.660000  63.369999  63.470001  56.330929   7899500
 2020-01-29  63.520000  63.570000  63.150002  63.150002  56.046917   8300800
 2020-01-30  63.180000  63.950001  63.139999  63.900002  56.712555  13824800
 2020-01-31  63.930000  64.209999  63.000000  63.180000  56.073544  15280000

[169 rows x 6 columns],
 'XLE':                Open       High        Low      Close   Adj Close
Volume
 Date
 2019-06-03  59.169998  59.810001  59.080002  59.520000  45.829723  18024700
 2019-06-04  60.080002  60.689999  59.919998  60.549999  46.622807  14743400
 2019-06-05  60.669998  60.689999  59.520000  59.880001  46.106918  15881900
 2019-06-06  60.000000  61.220001  60.000000  60.990002  46.961601  32393900
 2019-06-07  61.180000  61.660000  61.029999  61.270000  47.177197  15774100
 ...              ...        ...        ...        ...        ...       ...
 2020-01-27  55.480000  55.700001  54.980000  55.049999  44.890648  21682500
 2020-01-28  55.360001  55.639999  55.080002  55.299999  45.094505  14300100
 2020-01-29  55.610001  55.810001  54.650002  54.720001  44.621548  13745400
```

```
2020-01-30  54.150002   55.220001   54.009998    55.160000   44.980347   24430500
2020-01-31  54.230000   54.380001   53.160000    53.439999   43.577763   33430500

[169 rows x 6 columns],
 'XLV':                  Open         High          Low        Close   Adj Close
\
 Date
 2019-06-03   87.360001    88.239998    87.309998    87.489998   79.929428
 2019-06-04   88.220001    89.089996    88.110001    88.940002   81.254135
 2019-06-05   89.589996    89.820000    88.970001    89.610001   81.866249
 2019-06-06   89.750000    90.430000    89.540001    90.070000   82.286499
 2019-06-07   90.360001    91.230003    90.330002    91.010002   83.145264
 ...                ...          ...          ...          ...         ...
 2020-01-27  101.220001   102.070000   101.080002   101.680000   94.677200
 2020-01-28  101.980003   102.580002   101.699997   102.180000   95.142761
 2020-01-29  102.279999   102.639999   101.879997   101.900002   94.882050
 2020-01-30  101.150002   101.220001   100.290001   101.160004   94.193008
 2020-01-31  100.760002   100.839996    98.809998    99.139999   92.312126

                 Volume
 Date
 2019-06-03   19371000
 2019-06-04   15837200
 2019-06-05   24291700
 2019-06-06   18840900
 2019-06-07    9302700
 ...               ...
 2020-01-27   11762300
 2020-01-28    8079900
 2020-01-29    6435700
 2020-01-30   10989400
 2020-01-31   14476700

[169 rows x 6 columns],
 'VOX':                  Open         High          Low        Close   Adj Close
Volume
 Date
 2019-06-03   82.949997    82.949997    80.949997    81.389999   77.397781   484800
 2019-06-04   82.080002    83.139999    81.779999    83.089996   79.014389   132400
 2019-06-05   83.419998    83.470001    82.550003    83.150002   79.071434   224600
 2019-06-06   83.239998    83.660004    82.870003    83.430000   79.337700   148300
 2019-06-07   83.779999    84.760002    83.779999    84.470001   80.326683   129200
 ...                ...          ...          ...          ...         ...      ...
 2020-01-27   95.900002    96.410004    95.470001    96.050003   92.053963   152900
 2020-01-28   96.629997    97.440002    96.279999    97.110001   93.069862    68300
 2020-01-29   97.480003    97.500000    96.459999    96.620003   92.600250   196200
 2020-01-30   94.800003    95.970001    94.529999    95.919998   91.929375   146000
```

```
2020-01-31  95.919998  95.919998  94.250000  94.510002  90.578041  208100


[169 rows x 6 columns],
'XLU':                 Open       High        Low      Close  Adj Close
Volume
 Date
2019-06-03  58.439999  58.910000  58.139999  58.830002  49.740742  20335400
2019-06-04  58.889999  58.959999  57.790001  58.880001  49.783009  18528900
2019-06-05  59.099998  60.340000  58.889999  60.139999  50.848339  31755300
2019-06-06  60.220001  60.619999  60.049999  60.509998  51.161175  25511200
2019-06-07  60.820000  61.189999  60.029999  60.060001  50.780701  18993100
 ...               ...         ...         ...        ...        ...
2020-01-27  68.320000  68.800003  68.000000  68.190002  59.021690  11272200
2020-01-28  68.260002  68.760002  68.089996  68.470001  59.264042  17528400
2020-01-29  68.510002  68.750000  68.239998  68.650002  59.419838  10799800
2020-01-30  68.610001  69.330002  68.440002  69.290001  59.973797  14404300
2020-01-31  69.160004  69.440002  68.620003  68.980003  59.705467  20474500


[169 rows x 6 columns],
'IYR':                 Open       High        Low      Close  Adj Close
Volume
 Date
2019-06-03  87.150002  87.389999  86.570000  87.209999  75.486794   7063700
2019-06-04  87.199997  87.250000  86.080002  86.930000  75.244415   7322300
2019-06-05  87.379997  88.750000  87.129997  88.750000  76.819786  11722500
2019-06-06  88.849998  89.000000  88.190002  88.879997  76.932304   6501900
2019-06-07  89.379997  89.639999  89.010002  89.150002  77.166008   4795700
 ...               ...         ...         ...        ...        ...
2020-01-27  95.430000  95.760002  95.199997  95.500000  84.720749   4866600
2020-01-28  95.599998  96.169998  95.480003  95.910004  85.084480   6030700
2020-01-29  96.129997  96.309998  95.459999  95.610001  84.818344   4804700
2020-01-30  95.239998  95.849998  95.169998  95.730003  84.924797   7056700
2020-01-31  95.540001  95.699997  94.300003  94.570000  83.895729   7803900


[169 rows x 6 columns]},
'contraction': {'XLB':                 Open       High        Low      Close
Adj Close    Volume
 Date
2007-12-03  41.900002  41.900002  41.110001  41.439999  28.627274   6544500
2007-12-04  41.200001  41.439999  41.110001  41.189999  28.454569   7139100
2007-12-05  41.549999  41.950001  41.500000  41.950001  28.979591   5803900
2007-12-06  42.240002  42.770000  41.849998  42.619999  29.442436   5458700
2007-12-07  42.860001  43.220001  42.700001  43.020000  29.718761   7083000
 ...               ...         ...         ...        ...        ...
2008-09-24  36.759998  36.810001  35.799999  36.180000  25.576509   6640400
2008-09-25  36.130001  36.419998  35.619999  36.080002  25.505812   7455700
2008-09-26  35.150002  35.430000  34.580002  35.240002  24.911999   8788500
```

```
2008-09-29   34.380001   34.439999   31.690001   32.700001   23.116423   17620500
2008-09-30   32.500000   33.400002   32.169998   33.400002   23.611261    8561800

[209 rows x 6 columns],
 'XLI':                  Open        High        Low        Close   Adj Close
Volume
 Date
 2007-12-03   39.650002   39.650002   39.000000   39.000000   27.897243    3348700
 2007-12-04   38.889999   39.080002   38.799999   38.910000   27.832872    3114000
 2007-12-05   39.540001   39.560001   39.240002   39.549999   28.290670    2815800
 2007-12-06   39.700001   40.220001   39.419998   40.000000   28.612558    2220100
 2007-12-07   40.180000   40.369999   40.160000   40.299999   28.827137    3432400
 …            …           …           …           …           …
 2008-09-24   31.850000   31.879999   30.969999   31.450001   22.935705    8931100
 2008-09-25   31.370001   31.930000   31.000000   31.700001   23.118021   11687900
 2008-09-26   31.030001   31.610001   30.639999   31.570000   23.023211   14531000
 2008-09-29   31.040001   31.340000   29.290001   30.500000   22.242889   16947400
 2008-09-30   30.000000   30.799999   29.760000   30.799999   22.461670   10050300

[209 rows x 6 columns],
 'XLF':                  Open        High        Low        Close   Adj Close
Volume
 Date
 2007-12-03   25.052803   25.166531   24.784727   24.979691   17.888229    72568927
 2007-12-04   24.573517   24.695370   24.411049   24.516653   17.556635    75393703
 2007-12-05   24.898457   25.101543   24.606010   24.882210   17.818419   107066841
 2007-12-06   24.922827   25.686434   24.857840   25.686434   18.394337   122098951
 2007-12-07   25.735174   25.792040   25.345247   25.345247   18.150007   100382387
 …            …           …           …           …           …
 2008-09-24   16.799351   16.815598   16.206335   16.458164   12.211349   168258373
 2008-09-25   16.547522   17.043055   16.401300   16.913078   12.548880   196826560
 2008-09-26   16.125101   17.376118   16.019497   17.376118   12.892433   252887038
 2008-09-29   16.693745   16.799351   14.939074   15.085297   11.192731   422434145
 2008-09-30   15.353371   16.368807   15.166531   16.157595   11.988336   212701043

[209 rows x 6 columns],
 'XLK':                  Open        High        Low        Close   Adj Close
Volume
 Date
 2007-12-03   26.340000   26.370001   26.049999   26.100000   20.572214    2674700
 2007-12-04   26.090000   26.219999   25.799999   26.000000   20.493389    1971400
 2007-12-05   26.440001   26.650000   26.260000   26.600000   20.966311    3866200
 2007-12-06   26.790001   26.959999   26.549999   26.870001   21.179132    2272000
 2007-12-07   27.129999   27.129999   26.820000   26.900000   21.202776    2797300
 …            …           …           …           …           …
 2008-09-24   20.340000   20.490000   20.129999   20.350000   16.219780    7084300
 2008-09-25   20.469999   21.020000   20.370001   20.700001   16.498747   14856900
```

25

```
2008-09-26    20.170000    20.920000    20.139999    20.799999    16.578451    16681300
2008-09-29    20.320000    20.440001    18.959999    19.000000    15.143777    20520800
2008-09-30    19.180000    19.930000    19.180000    19.850000    15.821262    10347300

[209 rows x 6 columns],
'XLY':              Open          High          Low          Close    Adj Close
Volume
 Date
2007-12-03    34.869999    34.910000    34.470001    34.470001    27.534409    2419900
2007-12-04    34.549999    34.580002    34.230000    34.360001    27.446537    2988000
2007-12-05    34.549999    34.590000    34.130001    34.310001    27.406609    3764800
2007-12-06    34.509998    34.810001    34.099998    34.610001    27.646238    2617400
2007-12-07    34.759998    35.000000    34.750000    34.900002    27.877888    4098900

…               …             …             …            …             …
2008-09-24    28.770000    29.190001    28.190001    28.450001    23.082438    4654000
2008-09-25    28.790001    29.170000    28.299999    28.959999    23.496220    6353300
2008-09-26    28.559999    29.250000    28.299999    29.250000    23.731504    4583200
2008-09-29    29.150000    29.150000    26.790001    27.299999    22.149397    10763400
2008-09-30    27.900000    28.420000    27.200001    27.900000    22.636202    5225300

[209 rows x 6 columns],
'XLP':              Open          High          Low          Close    Adj Close
Volume
 Date
2007-12-03    29.370001    29.370001    29.030001    29.080000    18.534687    5158900
2007-12-04    29.049999    29.150000    29.000000    29.040001    18.509188    2187800
2007-12-05    29.389999    29.389999    29.090000    29.160000    18.585676    1632500
2007-12-06    29.330000    29.459999    29.150000    29.150000    18.579308    3177400
2007-12-07    29.400000    29.600000    29.309999    29.459999    18.776880    3010100

…               …             …             …            …             …
2008-09-24    27.209999    27.490000    27.049999    27.129999    17.668863    4509800
2008-09-25    27.330000    27.900000    27.240000    27.799999    18.105204    7749500
2008-09-26    27.450001    27.959999    27.450001    27.700001    18.040081    2529400
2008-09-29    27.889999    27.950001    26.660000    26.660000    17.362762    9894400
2008-09-30    27.010000    27.600000    26.980000    27.600000    17.974957    2583600

[209 rows x 6 columns],
'XLE':              Open          High          Low          Close    Adj Close
Volume
 Date
2007-12-03    72.879997    74.129997    72.879997    73.750000    44.157948    19546600
2007-12-04    73.599998    73.650002    73.000000    73.000000    43.708885    14566700
2007-12-05    74.000000    75.050003    73.870003    74.610001    44.672878    24490100
2007-12-06    74.639999    76.650002    74.330002    76.650002    45.894333    21135000
2007-12-07    76.300003    76.709999    75.709999    76.180000    45.612911    12377700

…               …             …             …            …             …
2008-09-24    68.699997    69.029999    67.620003    67.889999    41.111340    29585200
```

26

```
2008-09-25   67.919998   69.949997   67.900002   69.550003   42.116577   27722200
2008-09-26   67.589996   68.489998   66.019997   67.699997   40.996292   34620700
2008-09-29   65.989998   65.989998   59.130001   59.650002   36.121544   45861700
2008-09-30   61.919998   64.290001   61.430000   63.299999   38.331833   33200300

[209 rows x 6 columns],
'XLV':              Open        High        Low        Close   Adj Close
Volume
Date
2007-12-03   36.830002   36.830002   36.169998   36.240002   27.097029   1649900
2007-12-04   35.730000   36.380001   35.660000   36.230000   27.089571   1344500
2007-12-05   36.580002   36.599998   36.299999   36.520000   27.306400   1904000
2007-12-06   36.270000   36.799999   36.270000   36.770000   27.493319   2719100
2007-12-07   36.700001   36.820000   36.630001   36.750000   27.478361   3442900
...              ...         ...         ...         ...         ...        ...
2008-09-24   30.340000   30.670000   30.230000   30.350000   23.089081   4029600
2008-09-25   30.379999   31.280001   30.379999   30.900000   23.507494   9256200
2008-09-26   30.750000   31.219999   30.530001   30.840000   23.461855   5521100
2008-09-29   30.459999   31.059999   29.000000   30.150000   22.936928   8689400
2008-09-30   30.270000   30.450001   29.780001   30.049999   22.860847   3309800

[209 rows x 6 columns],
'VOX':              Open        High        Low        Close   Adj Close
Volume
Date
2007-12-03   77.629997   77.629997   76.169998   76.349998   51.317448     25200
2007-12-04   76.099998   76.099998   75.699997   75.699997   50.880585     75800
2007-12-05   75.870003   76.800003   75.870003   76.800003   51.619926     10700
2007-12-06   76.669998   77.440002   76.269997   77.440002   52.050087     24300
2007-12-07   77.529999   78.000000   77.410004   77.639999   52.184513     47100
...              ...         ...         ...         ...         ...        ...
2008-09-24   56.389999   56.389999   53.900002   54.840000   37.837605     30500
2008-09-25   56.599998   56.860001   54.919998   56.200001   38.775951     20000
2008-09-26   56.169998   56.439999   55.180000   56.220001   38.789742     18500
2008-09-29   55.400002   55.520000   51.840000   53.189999   36.699165     42800
2008-09-30   51.529999   53.779999   51.470001   53.349998   36.809566     34100

[209 rows x 6 columns],
'XLU':              Open        High        Low        Close   Adj Close
Volume
Date
2007-12-03   43.439999   43.660000   42.730000   43.150002   23.758991   3549800
2007-12-04   43.240002   43.770000   42.820000   43.459999   23.929691   7669100
2007-12-05   43.900002   44.240002   43.599998   44.209999   24.342651   6798800
2007-12-06   44.400002   44.459999   43.770000   43.950001   24.199488   4106400
2007-12-07   44.360001   44.470001   43.900002   44.180000   24.326134   6117300
...              ...         ...         ...         ...         ...        ...
```

```
2008-09-24   33.639999   33.849998   33.209999   33.849998   19.224617   3306600
2008-09-25   33.770000   34.889999   33.770000   34.520000   19.605141   4199500
2008-09-26   33.520000   34.799999   33.520000   34.529999   19.610825   3207700
2008-09-29   34.139999   34.139999   32.509998   32.720001   18.582861   7997900
2008-09-30   32.740002   33.259998   32.570000   33.230000   18.872509   3359200

[209 rows x 6 columns],
'IYR':                   Open        High         Low       Close   Adj Close
Volume
 Date
2007-12-03   69.309998   69.989998   68.730003   69.510002   37.294937    5246000
2007-12-04   68.519997   69.180000   67.580002   67.690002   36.318439    6202200
2007-12-05   68.370003   69.830002   68.260002   69.620003   37.353970    8667700
2007-12-06   69.900002   72.230003   69.070000   72.120003   38.695324    6237500
2007-12-07   72.080002   72.940002   71.720001   71.769997   38.507542   11116800
...              ...         ...         ...         ...         ...        ...
2008-09-24   62.000000   62.000000   59.880001   60.439999   33.933945   12940800
2008-09-25   61.700001   62.180000   60.080002   61.560001   34.562775   15335900
2008-09-26   60.200001   63.380001   59.639999   62.840000   35.281425   16272700
2008-09-29   61.509998   61.889999   58.570000   59.080002   33.170376   20029500
2008-09-30   60.650002   62.500000   58.980000   61.950001   34.781742   14512300

[209 rows x 6 columns]}}
```

Missing Data Values   'download_stock_data_for_periods' function checks to make sure if there are any missing dates for any of the stocks it will return a dataframe with those that are missing. However each sector etf seems to have all data so there is no missing data.

Accessing specific data   The following is an example of how you can access the healthcare 'XLV' sector ETF during the trough time period.

```
[ ]: # access healthcare sector during an expansion
     # first load in the trough data
     trough_sector_data = sector_etf_data['trough']
```

```
[ ]: # now access the healthcare sector etf 'XLV'
     trough_sector_data['XLV'].head()
```

```
[ ]:                Open        High         Low       Close   Adj Close    Volume
     Date
     2008-10-01   30.100000   30.480000   30.100000   30.250000   23.013002   6053600
     2008-10-02   30.250000   30.590000   29.930000   30.299999   23.051035   6353400
     2008-10-03   30.600000   30.600000   29.650000   29.650000   22.556545   6814400
     2008-10-06   29.400000   29.879999   27.410000   28.540001   21.712105   8545000
     2008-10-07   28.719999   28.780001   27.389999   27.850000   21.187176   5060200
```

This returns the candlestick data for the XLV equity during a trough which ranged from 2008-01-02 to 2009-6-30 due to the Global Financial Crises.

The same process can be completed in one step.

```
[ ]: # access 'XLV' equity in one line of code
     sector_etf_data['trough']['XLV'].head()
```

```
[ ]:                 Open       High       Low      Close  Adj Close   Volume
     Date
     2008-10-01  30.100000  30.480000  30.100000  30.250000  23.013002  6053600
     2008-10-02  30.250000  30.590000  29.930000  30.299999  23.051035  6353400
     2008-10-03  30.600000  30.600000  29.650000  29.650000  22.556545  6814400
     2008-10-06  29.400000  29.879999  27.410000  28.540001  21.712105  8545000
     2008-10-07  28.719999  28.780001  27.389999  27.850000  21.187176  5060200
```

This is extremely useful for technical analysis which is heavily reliant on historical data. There is now an easy process to access any stock data's candlestick data (in this case sector ETF's) during any time period (different macroeconomic time periods). This is going to be used to as a basis for the capstone project.

There are currently 6 columns however this will grow as different techniques require new columns to be used as features.

The size of the data currently stands at: 4 time periods which as - 166 (trough) - 754 (expansion) - 169 (peak) - 209

This is equal to 1298 trading days. Each trading day has 11 sector ETF's each with 6 columns which is 114,224 data points.

1.2 Initial Data Inspection / Univariate Analysis

The dataset is a nested dictionary with a number of time series. Using typical descriptive statistics might not be as relevant. Using the create_analysis function you are able to load in the nested dictionary and apply different descriptive statistical reports on the different sector etfs during different time periods as a table. This can be adjusted to fit a number of descriptive statistics and for different parameters within the dataframe of the nested dictionary.

1.2.1 Numerical Analysis

Initial Numerical Analysis Looking at the descriptive statistics of the different sector etfs during different time period can help observe how the data changes from the beginning to the end of the period. This will be helpful for initial understanding of how investments change during different macroeconomic cycles.

```
[ ]: # get the mean value of these stocks during the different time periods
     mean_data =␣
      ↪create_analysis(sector_etf_tickers,economic_cycle_periods_list,sector_etf_data,'Adj␣
      ↪Close','mean')
     mean_data
```

```
[ ]:        trough   expansion       peak  contraction
     XLB  16.686582  33.341455  52.917293   28.608563
```

```
XLI  15.852843   36.435797    72.431897    26.289616
XLF   6.886585   12.601652    25.941764    14.533609
XLK  12.620077   28.430873    79.864511    18.592997
XLY  16.963294   49.350607   116.163152     25.01046
XLP  14.870896   29.717787    53.480786    17.960932
XLE  28.421203    53.23693    47.020173    46.356162
XLV  19.386922   41.449804    87.094059    24.519837
VOX  31.093878   62.755677    85.677095     44.45534
XLU  15.769571   26.847841    53.767982    22.055004
IYR  19.002604   46.571497    80.521102    35.585493
```

The initial table can display the sector etfs and their mean values across the given time periods.

```
[ ]: # get the median value of these stocks
     median_data =␣
      ↪create_analysis(sector_etf_tickers,economic_cycle_periods_list,sector_etf_data,'Adj␣
      ↪Close','median')
     median_data
```

```
[ ]:          trough   expansion       peak  contraction
     XLB  16.347619   32.295517  52.799812    28.447906
     XLI  16.113979   35.817581  71.974792    26.422617
     XLF   6.943471    13.03797  25.566183     14.80596
     XLK   12.47098    27.19953  77.621544    18.339046
     XLY  16.923293   50.454159  116.36422     25.15683
     XLP  14.960625   30.248545  53.487617    18.002182
     XLE  28.328445   53.228485  46.976997    45.894333
     XLV  19.492291   41.068953  84.821175    24.091343
     VOX  30.996964   64.179016  84.958954    43.722988
     XLU  15.865283   26.341283  53.986691    22.073486
     IYR   18.73623   45.988676  80.861588    35.542835
```

The initial table can display the sector etfs and their median values across the given time periods.

```
[ ]: # get the difference between these two to decide if the mean and median are␣
      ↪close
     mean_data-median_data
```

```
[ ]:          trough  expansion        peak  contraction
     XLB   0.338963   1.045938   0.117481     0.160657
     XLI  -0.261136   0.618215   0.457105    -0.133001
     XLF  -0.056886  -0.436317   0.375581     -0.27235
     XLK   0.149097   1.231343   2.242967     0.253951
     XLY       0.04  -1.103552  -0.201068     -0.14637
     XLP  -0.089729  -0.530757  -0.006831     -0.04125
     XLE   0.092757   0.008445   0.043176     0.461829
     XLV   -0.10537   0.380852   2.272884     0.428494
     VOX   0.096914  -1.423339   0.718141     0.732352
```

```
XLU -0.095712  0.506559 -0.218709   -0.018482
IYR  0.266374  0.582821 -0.340485    0.042657
```

Notice that the mean and median values are very close together which means that the typical value of a stock is likely normally distributed where there are data values that above and below the average with common occurence.

```
[ ]: # how does the mean value change during the different states of the␣
     ↪macroeconomic cycle
     mean_data.mean()
```

```
[ ]: trough        17.959496
     expansion     38.249084
     peak          68.625438
     contraction   27.633456
     dtype: float64
```

The mean values of the sector etf's return what is to be expected. The trough has the lowest prices, the peak has the highest prices. Meanwhile the contraction is on the lower end and the expansion is on the upper end.

```
[ ]: # get the standard deviation
     create_analysis(sector_etf_tickers,economic_cycle_periods_list,sector_etf_data,'Adj␣
     ↪Close','std')
```

```
[ ]:        trough expansion      peak contraction
     XLB    1.8369  4.834144  1.567754    1.689659
     XLI  1.812336  6.592581  3.025957    1.449261
     XLF   1.60123  2.353624  1.424399    1.822995
     XLK  0.955407  4.112598  6.027887    1.242621
     XLY  1.718797   9.03084  3.309264    1.311233
     XLP  0.887215  4.054363  1.675483    0.437353
     XLE  2.229603  6.899179  1.668784    3.921771
     XLV  1.066656  9.149628  4.829238    1.223715
     VOX  2.207266  8.344901  3.730777    3.540011
     XLU  1.029368  3.061783  2.171691    1.239613
     IYR  3.596362  4.439796  2.271276    1.844816
```

This is heavily impacted by the time of which the stock is held on for. As expasnion goes for a longer period of time the standard deviation is higher, so this is not the most useful statistic (for now).

```
[ ]: # get the min for each stock
     create_analysis(sector_etf_tickers,economic_cycle_periods_list,sector_etf_data,'Adj␣
     ↪Close','min')
```

```
[ ]:         trough  expansion      peak contraction
     XLB  12.936712   25.42024 48.822552   23.116423
     XLI  11.310503  26.419207 66.340546   22.242889
```

```
XLF    3.780499     8.54141  23.495205   10.250643
XLK   10.512971  21.661152  67.064896   15.143777
XLY   13.169909  33.771416  104.76281   21.546104
XLP    12.74903  22.914076  48.905941   17.105051
XLE   23.203629  39.914341  43.396397   36.121544
XLV   16.746853   28.42687  79.929428   22.719442
VOX   24.479906  47.675571  77.397781   36.699165
XLU   13.078052  22.511068  49.740742    18.57753
IYR   12.788877  37.196949  75.244415   31.198677
```

The lowest price of the adjusted close

```
[ ]:  # get the max for each stock
      create_analysis(sector_etf_tickers,economic_cycle_periods_list,sector_etf_data,'Adj␣
        ↪Close','max')
```

```
[ ]:           trough   expansion        peak  contraction
      XLB   23.222452  41.566242   56.028214    32.200756
      XLI   21.929304  48.189194   78.330185    28.963055
      XLF   12.458467  16.926832   28.394539    18.533947
      XLK   15.677793  37.494671   93.997894    21.672701
      XLY   22.636211  65.233078  123.115753    27.893871
      XLP   18.105207  38.393627   56.881191    18.856041
      XLE   37.702045  68.046036   50.070969    54.416733
      XLV   23.051035  60.622314   97.517143    27.530714
      VOX   36.781956  76.104362   94.986649    53.203094
      XLU   18.792999  35.942581   59.973797    24.436258
      IYR   34.248367  57.641975    85.32399    39.473301
```

The highest price of the adjusted close

1.2.2   Visualisation

Creating a dashboard that can be used to adjust different time periods/stocks to visualise your stock data through candlestick data would be extremely useful. This will be done later. For now use the 'visualize_candlestick' function to plot candlestick data.

```
[ ]:  # create a sample visualisation of 'XLV' healthcare ETF from 2009-06 to␣
        ↪2010-01-01
      for period in economic_cycle_periods_list:
          visualize_candlestick(sector_etf_data[period]['XLV'],f'XLK {period}')
```

Observe the movement of the XLK healthcare sector etf and how it changes during different macroeconomic time periods. Good for initial inspection to averaging all of the stocks during the different macroeconomic cycles will be a better representation of the movement of the market.

1.3   Stock Investment: Buy and Hold

An extremely popular type of investment is a passive one, this is where individuals will see an opportunity for long term growth of an equity and will simply purchase the stock and hold onto it

for a varied period of time. As the beginning component of analysis, use a buy and hold strategy to look at the different stocks during different time periods to determine yearly return on investment

1.3.1 Return on Investment Simple Model

The simple model is going to pruchase the sector etf at the start of the time period, then hold onto it. This is extremely reliant on the time the stock is purchased and is simply down to luck. However, it is going to show a rough udnerstanding of how the price of the stock changes. The timing has a bit of wiggleroom so that it goes from the previous cycle to the current cycle to represent how it changes.

```
[ ]: buy_hold_roi(sector_etf_data,economic_cycle_periods_list,sector_etf_tickers,'Adj␣
     ↪Close')
```

```
[ ]:          trough  expansion        peak  contraction
     XLB  -22.978958   14.67684    11.80293   -20.770524
     XLI  -33.796257  20.419984   20.975941   -23.043271
     XLF  -53.252213  25.012766   25.408884   -38.349943
     XLK  -13.845861   19.18231      59.558   -27.193473
     XLY  -23.365403  24.228073   23.574825   -21.081112
     XLP  -22.774004  17.478538     22.9095    -3.638299
     XLE  -23.477594   5.689466   -7.317937   -15.718586
     XLV  -19.875651  27.096262   24.263915   -18.572862
     VOX   -8.725742   13.37127   26.766948   -33.074376
     XLU  -24.913935  14.433953   31.706786   -24.292124
     IYR  -56.301249  14.645414   17.268776    -8.086164
```

1.3.2 Stochastic Numerical Analysis

The downside of the initial numerical analysis is that it only takes in one event of where you buy at the start of the period and sell at the end of the period. This is going to make time period and time of investment be a parameter which impacts the expected return. To overcome this stochastic modeling can be performed.

By performing a stochastic model where different time periods of initial investment and holding onto the equity for a set period takes away the included parameters. This isolates the expected return more. It can also be repeated over a number of samples which can then increase the likeliness of it being a true representation of the data. A histogram can also be created to observe how different returns on investment are observed. This probabilistic value can be used to display the probability distribution function and a cumulative distribution function can be created.

Calculating Expected Returns The 'stochastic_modeling' function is going to go through each equity during each time period and take n samples of initial start dates and purchase one sector ETF. This is held for a month and then sold off. By repeating the model 1000 times you get a distribution of returns which can be used to understand expected returns and standard deviations which is essential for understanding which sector is best for different time periods and to understand how the data isdistributed.

```python
# create an adjusted close price sector etf dictionary which can changed based
 ↪on the macroeconomic period
adjusted_close_sector_etf =
 ↪get_adjusted_closed_price(sector_etf_data,sector_etf_tickers,economic_cycle_periods_list)
```

```python
# take 1000 samples of different initial purchase dates, hold the equity for 30
 ↪days then compare it with the start
expected_values_sector_etf_buy_hold =
 ↪stochastic_modeling(adjusted_close_sector_etf,sector_etf_tickers,economic_cycle_periods_lis
expected_values_sector_etf_buy_hold['trough']['XLK'][0:5]
```

```
[-2.233114541074024,
 -7.916431375114681,
 6.40207475775085,
 9.35615165782157,
 -20.50455454243595]
```

Visualising Expected Returns A dashboard is going to be used to create a visualisation of any sector etf during any time period however the following will display one sector etf during the different economic time periods.

```python
# example case of expected values by sampling

import matplotlib.pyplot as plt

for period in economic_cycle_periods_list:
        plt.hist(pd.
 ↪Series((expected_values_sector_etf_buy_hold[period]['XLK'])))
        plt.ylabel('Occurencences (/1000)')
        plt.xlabel('Return on Investment (%)')
        plt.title(f'Average ROI for XLK per month during a {period}')
        plt.show()
"""
print('Standard Deviation:')
print(pd.Series(expected_values_sector_etf_buy_hold['expansion']['XLK']).std())
print('Median:')
print(pd.Series(expected_values_sector_etf_buy_hold['expansion']['XLK']).
 ↪median())
print('Mean:')
print(pd.Series(expected_values_sector_etf_buy_hold['expansion']['XLK']).
 ↪mean())"""
```
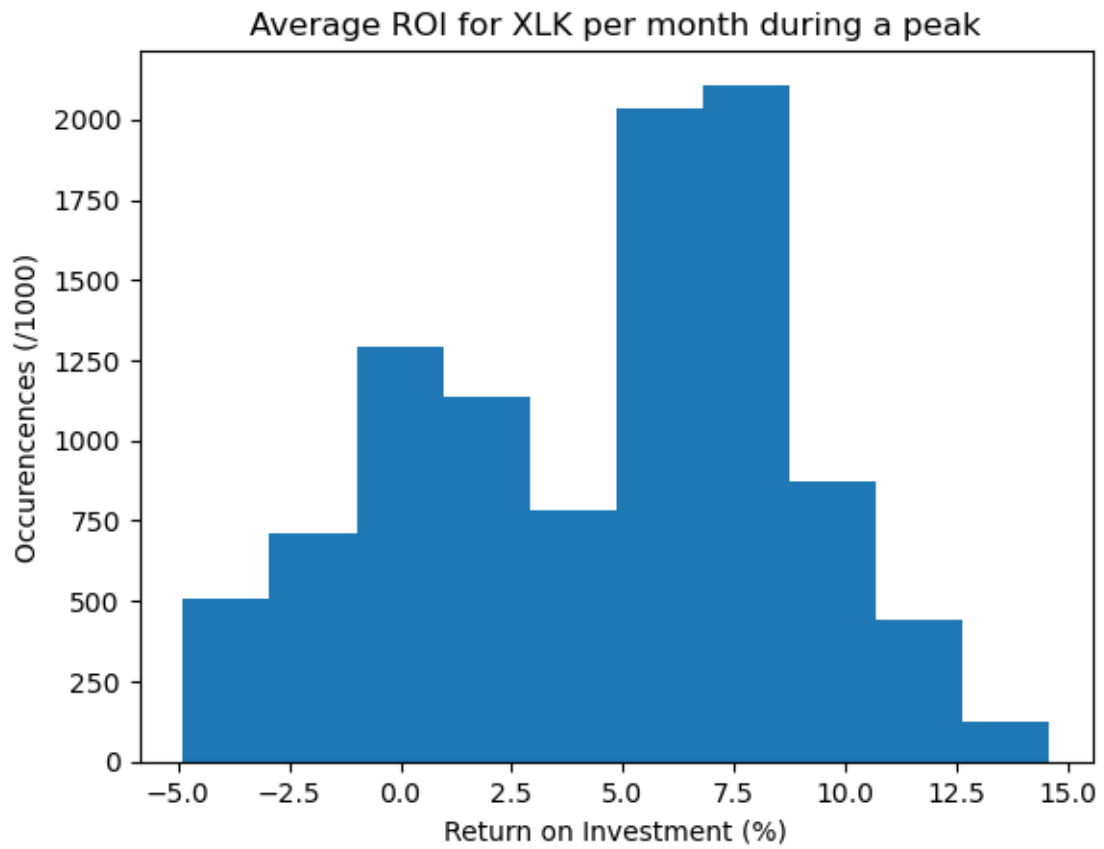
Average ROI for XLK per month during a trough

Average ROI for XLK per month during a expansion

Average ROI for XLK per month during a peak

Average ROI for XLK per month during a contraction
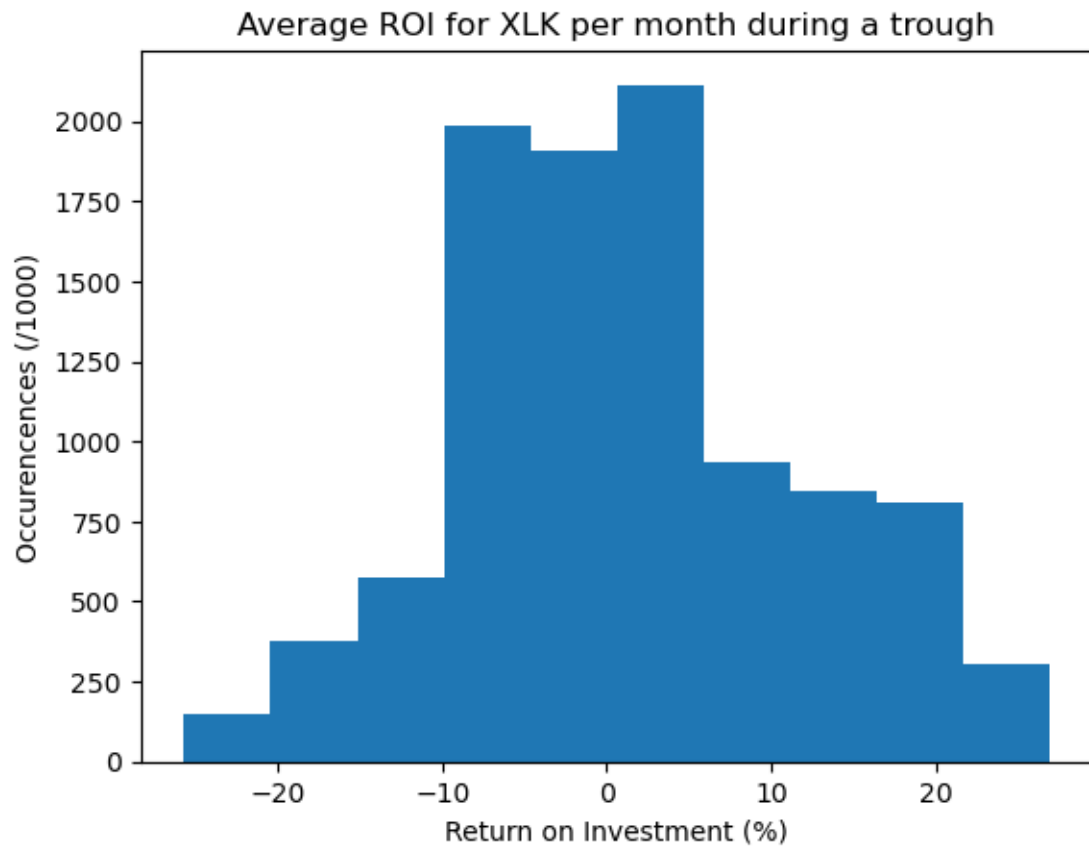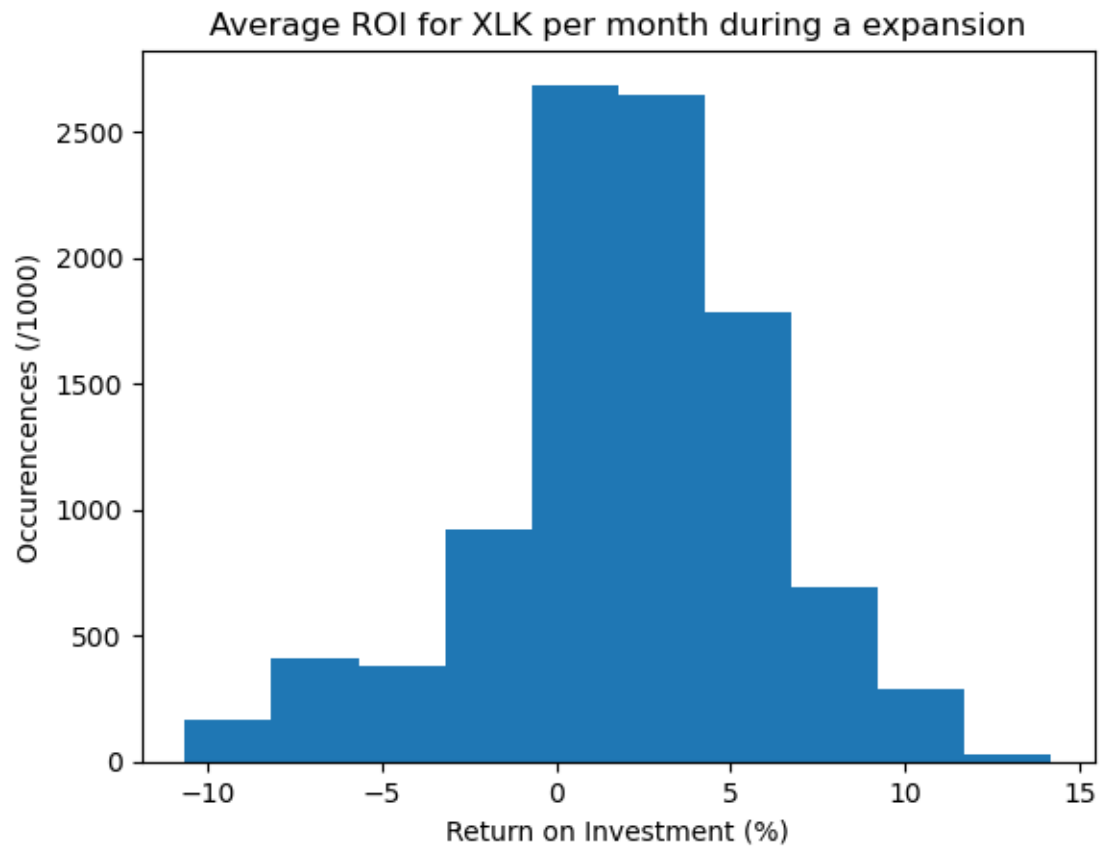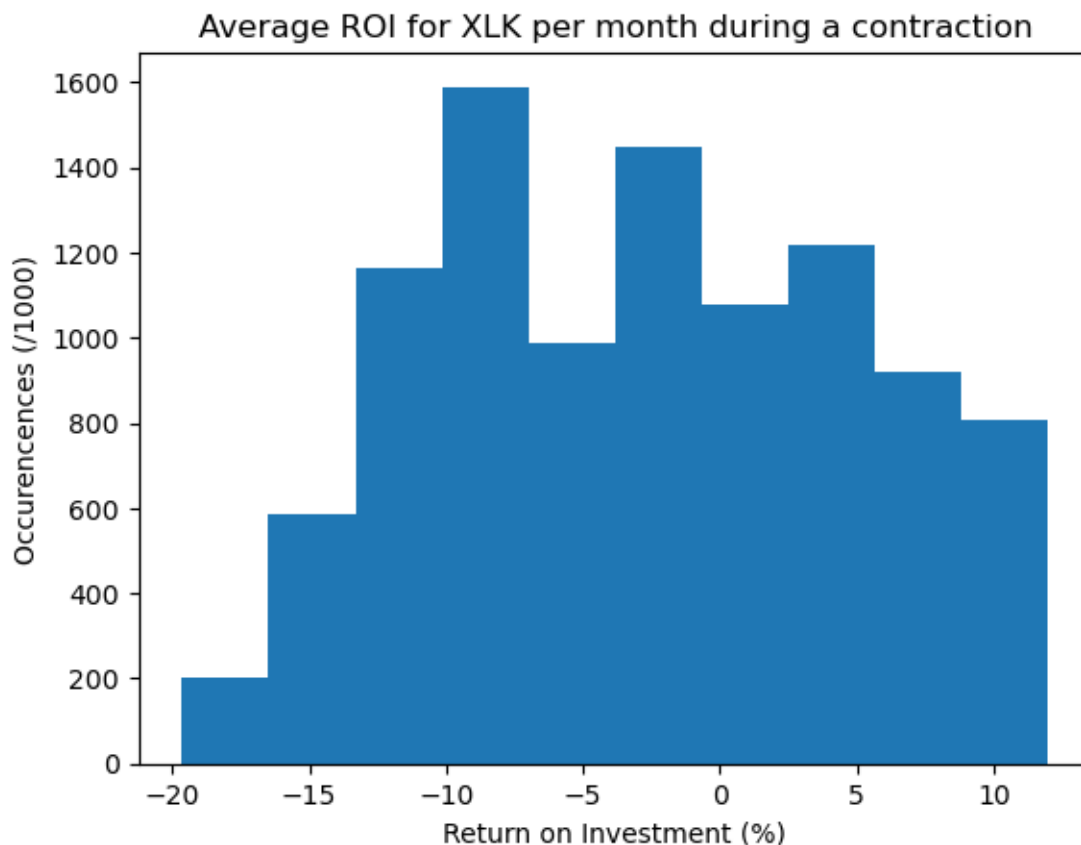
[ ]: "\nprint('Standard Deviation:')\nprint(pd.Series(expected_values_sector_etf_buy_
     hold['expansion']['XLK']).std())\nprint('Median:')\nprint(pd.Series(expected_val
     ues_sector_etf_buy_hold['expansion']['XLK']).median())\nprint('Mean:')\nprint(pd
     .Series(expected_values_sector_etf_buy_hold['expansion']['XLK']).mean())"

The standard deviation is only useful if the data has a normal distribution. None of the data looks
to be exactly normal but it looks close to normal.

Summarising Expected Results   Create a table for - Mean - Median - Standard Deviation -
Variation

```
[ ]: # find the average return
     stochasting_roi(sector_etf_tickers,economic_cycle_periods_list,expected_values_sector_etf_buy_
```

```
[ ]:      trough  expansion       peak  contraction
     XLB  1.198562   1.537088   1.24863    -1.222167
     XLI -0.981269   2.150231   2.100804   -2.038179
     XLF -2.421405   2.557886   2.832489   -4.809103
     XLK  1.437478   2.084359   4.555128   -2.816132
     XLY  2.199991   2.484361   1.485171   -1.040537
```

```
XLP -1.450495  2.018995    1.9574   -0.090087
XLE  0.120185  0.843974 -0.075263  -1.102912
XLV -1.130084  3.008388  2.775386  -1.762325
VOX  2.574735  1.707218  2.519156  -3.529541
XLU -1.254162  1.721219  2.302081  -2.681155
IYR -2.654339  1.606906  1.494454  -0.079736
```

[ ]: ```python
find the median return
stochasting_roi(sector_etf_tickers,economic_cycle_periods_list,expected_values_sector_etf_buy_
```

[ ]: 
```
        trough expansion   peak contraction
XLB -0.431482 1.618293 1.542702 -0.158259
XLI -4.297219   2.36529 2.170665 -1.142058
XLF -11.22692 2.959662 3.533226 -3.426746
XLK 0.657478 2.102444 5.544701 -3.179457
XLY -0.533992 3.084819 1.375891 -0.436236
XLP -1.367334   2.21635   2.07875   0.247268
XLE -0.087128 1.717592 1.581706 -0.767388
XLV   0.20568 3.360667   2.05419 -2.989357
VOX 2.094677 1.905289 3.153506 -5.083491
XLU 0.641228 2.160784 2.269188 -3.255732
IYR -7.419466 1.999812   1.41295   1.830996
```

[ ]: ```python
# find the standard deviation return
stochasting_roi(sector_etf_tickers,economic_cycle_periods_list,expected_values_sector_etf_buy_
```

[ ]:
```
         trough expansion      peak contraction
XLB  15.337694  3.990715  3.099894   6.71531
XLI  16.230812  3.785388  3.619344  6.043564
XLF  27.714892  4.294388  4.155808  8.910313
XLK   10.73694   3.97602   4.29604  7.680441
XLY  16.433959  3.745687  3.165938  7.297218
XLP   6.879296  3.026176  1.557041  3.952035
XLE  10.638283  5.359061  5.124742  9.482457
XLV   7.619162  3.222491  4.117268   4.81309
VOX  10.442552  3.962103  3.365846   7.90066
XLU    8.36837  3.976261   3.37901  5.444024
IYR  22.683168  4.279003  2.506886    7.4874
```

[ ]: ```python
find the variance return
stochasting_roi(sector_etf_tickers,economic_cycle_periods_list,expected_values_sector_etf_buy_
```

[ ]: 
```
          trough expansion   peak contraction
XLB 235.244869 15.925809   9.609342   45.095384
XLI 263.439267 14.329161 13.099652   36.524668
XLF 768.115261 18.441772 17.270737   79.393673
XLK 115.281878 15.808738 18.455962   58.989179
```

```
XLY  270.075004  14.030173  10.023163  53.249388
XLP   47.324718   9.157743   2.424378  15.618584
XLE  113.173075  28.719532  26.262979  89.916995
XLV   58.051632  10.384448  16.951897  23.165832
VOX  109.046886  15.698262  11.328918  62.420434
XLU   70.029609  15.810653  11.417707  29.637394
IYR  514.526102  18.309865   6.284478  56.061158
```

1.4 Bivariate Analysis

```
[ ]: # get the correlation for trough
     adjusted_close_sector_etf['trough'].corr()
```

```
[ ]:           XLB       XLI       XLF       XLK       XLY       XLP       XLE  \
     XLB  1.000000  0.810805  0.797901  0.926126  0.940809  0.713244  0.805334
     XLI  0.810805  1.000000  0.930429  0.623732  0.796610  0.955607  0.842349
     XLF  0.797901  0.930429  1.000000  0.607481  0.733800  0.909901  0.729371
     XLK  0.926126  0.623732  0.607481  1.000000  0.926797  0.500504  0.700050
     XLY  0.940809  0.796610  0.733800  0.926797  1.000000  0.684547  0.784799
     XLP  0.713244  0.955607  0.909901  0.500504  0.684547  1.000000  0.822337
     XLE  0.805334  0.842349  0.729371  0.700050  0.784799  0.822337  1.000000
     XLV  0.542796  0.742531  0.615790  0.426964  0.530545  0.773893  0.768463
     VOX  0.770734  0.490679  0.391562  0.882443  0.866890  0.385345  0.671205
     XLU  0.365713  0.742071  0.569087  0.222712  0.405506  0.795701  0.750064
     IYR  0.784803  0.916819  0.928703  0.608412  0.743235  0.896447  0.745821

               XLV       VOX       XLU       IYR
     XLB  0.542796  0.770734  0.365713  0.784803
     XLI  0.742531  0.490679  0.742071  0.916819
     XLF  0.615790  0.391562  0.569087  0.928703
     XLK  0.426964  0.882443  0.222712  0.608412
     XLY  0.530545  0.866890  0.405506  0.743235
     XLP  0.773893  0.385345  0.795701  0.896447
     XLE  0.768463  0.671205  0.750064  0.745821
     XLV  1.000000  0.379489  0.820143  0.729159
     VOX  0.379489  1.000000  0.253130  0.401263
     XLU  0.820143  0.253130  1.000000  0.619811
     IYR  0.729159  0.401263  0.619811  1.000000
```

```
[ ]: # get the correlation for expansion
     adjusted_close_sector_etf['expansion'].corr()
```

```
[ ]:           XLB       XLI       XLF       XLK       XLY       XLP       XLE  \
     XLB  1.000000  0.986982  0.964635  0.965615  0.962104  0.944093  0.936139
     XLI  0.986982  1.000000  0.984971  0.959310  0.985469  0.967985  0.912908
     XLF  0.964635  0.984971  1.000000  0.942745  0.993098  0.983776  0.894998
     XLK  0.965615  0.959310  0.942745  1.000000  0.935020  0.945932  0.851462
```

```
XLY   0.962104   0.985469   0.993098   0.935020   1.000000   0.975789   0.898844
XLP   0.944093   0.967985   0.983776   0.945932   0.975789   1.000000   0.860943
XLE   0.936139   0.912908   0.894998   0.851462   0.898844   0.860943   1.000000
XLV   0.970624   0.984149   0.985490   0.972424   0.979892   0.985053   0.867546
VOX   0.947038   0.959067   0.973338   0.919687   0.974767   0.971469   0.920449
XLU   0.909949   0.918728   0.917892   0.942792   0.898472   0.958003   0.792168
IYR   0.840137   0.836886   0.865614   0.873627   0.833315   0.910941   0.744439

            XLV        VOX        XLU        IYR
XLB   0.970624   0.947038   0.909949   0.840137
XLI   0.984149   0.959067   0.918728   0.836886
XLF   0.985490   0.973338   0.917892   0.865614
XLK   0.972424   0.919687   0.942792   0.873627
XLY   0.979892   0.974767   0.898472   0.833315
XLP   0.985053   0.971469   0.958003   0.910941
XLE   0.867546   0.920449   0.792168   0.744439
XLV   1.000000   0.960769   0.953500   0.885828
VOX   0.960769   1.000000   0.906095   0.862486
XLU   0.953500   0.906095   1.000000   0.943784
IYR   0.885828   0.862486   0.943784   1.000000
```

```python
# get the correlation for peak
adjusted_close_sector_etf['peak'].corr()
```

```
            XLB        XLI        XLF        XLK        XLY        XLP        XLE  \
XLB   1.000000   0.931061   0.933483   0.831381   0.790432   0.723403   0.534676
XLI   0.931061   1.000000   0.975108   0.933171   0.848105   0.831739   0.381272
XLF   0.933483   0.975108   1.000000   0.933766   0.830237   0.850624   0.362943
XLK   0.831381   0.933171   0.933766   1.000000   0.864236   0.922162   0.194119
XLY   0.790432   0.848105   0.830237   0.864236   1.000000   0.824515   0.430289
XLP   0.723403   0.831739   0.850624   0.922162   0.824515   1.000000   0.003893
XLE   0.534676   0.381272   0.362943   0.194119   0.430289   0.003893   1.000000
XLV   0.867563   0.907239   0.933462   0.944126   0.786762   0.830262   0.320859
VOX   0.845777   0.926502   0.918713   0.970180   0.911094   0.876923   0.323728
XLU   0.477936   0.665889   0.662305   0.792161   0.675354   0.891409  -0.199318
IYR   0.396946   0.591496   0.573272   0.692274   0.606475   0.833659  -0.297037

            XLV        VOX        XLU        IYR
XLB   0.867563   0.845777   0.477936   0.396946
XLI   0.907239   0.926502   0.665889   0.591496
XLF   0.933462   0.918713   0.662305   0.573272
XLK   0.944126   0.970180   0.792161   0.692274
XLY   0.786762   0.911094   0.675354   0.606475
XLP   0.830262   0.876923   0.891409   0.833659
XLE   0.320859   0.323728  -0.199318  -0.297037
XLV   1.000000   0.932945   0.648839   0.537455
VOX   0.932945   1.000000   0.717609   0.637115
```

```
XLU  0.648839  0.717609  1.000000  0.930607
IYR  0.537455  0.637115  0.930607  1.000000
```

[ ]: *# get the correlation for contraction*
     adjusted_close_sector_etf['contraction'].corr()

[ ]:
```
           XLB        XLI        XLF        XLK        XLY        XLP        XLE  \
XLB  1.000000   0.754209   0.369504   0.635813   0.515436   0.154849   0.880444
XLI  0.754209   1.000000   0.818386   0.830050   0.868215   0.371536   0.466680
XLF  0.369504   0.818386   1.000000   0.647745   0.826996   0.176189   0.051180
XLK  0.635813   0.830050   0.647745   1.000000   0.686525   0.502810   0.512872
XLY  0.515436   0.868215   0.826996   0.686525   1.000000   0.508089   0.166606
XLP  0.154849   0.371536   0.176189   0.502810   0.508089   1.000000   0.019872
XLE  0.880444   0.466680   0.051180   0.512872   0.166606   0.019872   1.000000
XLV -0.108733   0.384283   0.609219   0.563270   0.462638   0.419489  -0.285928
VOX  0.492346   0.797639   0.772419   0.939442   0.716776   0.429365   0.330297
XLU  0.655536   0.683622   0.557696   0.850614   0.436709   0.170406   0.637423
IYR  0.630365   0.707243   0.474916   0.507583   0.746499   0.493934   0.389375

           XLV        VOX        XLU        IYR
XLB -0.108733   0.492346   0.655536   0.630365
XLI  0.384283   0.797639   0.683622   0.707243
XLF  0.609219   0.772419   0.557696   0.474916
XLK  0.563270   0.939442   0.850614   0.507583
XLY  0.462638   0.716776   0.436709   0.746499
XLP  0.419489   0.429365   0.170406   0.493934
XLE -0.285928   0.330297   0.637423   0.389375
XLV  1.000000   0.712224   0.452431  -0.031447
VOX  0.712224   1.000000   0.841825   0.408048
XLU  0.452431   0.841825   1.000000   0.258034
IYR -0.031447   0.408048   0.258034   1.000000
```