**Steps to run the application to get the correlation calculations and Plots**

Below steps should be implemented to run the application

1. RealEstateCorrelation is the main class that calls the methods for correlation calculation and plotting.
2. The zillow\_home\_index and sp500cspdsdata are other classes of this application and they need to be imported into RealEstateCorrelation class and present in the folder for this application to run.
3. The class needs to have the below packages imported (by default they are imported):

**import** pandas **as** pd  
**import** seaborn **as** sns  
**import** matplotlib.pyplot **as** plt  
**import** sp500cspdsdata  
**import** zillow\_home\_index

1. The first step is to ensure that in the class ZillowHomeIndex, the dataset is read for the dataframe creation. This involves the path of the csv files to be included in the class.
2. In the method, get\_hIndex\_fr\_year\_melt\_later(), please give the correct location of the dataset (csv files). In the zillow\_home\_index.py, on line #79, #82, #85, #88, #91, please change the path of the csv files.

**Ex.** The highlighted portion needs to be changed to the location where the respective csv file is located. /Users/abhijeet/Desktop/MSDS/CSSemProject/Zip\_zhvi\_bdrmcnt\_2\_uc\_sfrcondo\_tier\_0.33\_0.67\_sm\_sa\_mon.csv

1. Similarly, the Sp500CspdsData.py needs to have the location of SP500.csv. In the constructor, the line #17 needs to be modified to include location of SP500.csv.

**Ex.** The location of the SP500.csv needs to be put in the highlighted portion.

self.df\_sp500 = pd.read\_csv("/Users/abhijeet/Desktop/MSDS/CSSemProject/SP500.csv")

1. The line # 24 needs to be modified to include the location of CSPDS.csv by changing the highlighted portion to the location of the csv file pd.read\_csv("/Users/abhijeet/Desktop/MSDS/CSSemProject/CSPDS.csv")
2. The application can be run after these changes with default values for **NY and 2021**.
3. To run the application for the desired **state, county, zip and year** to get the correlation calculations and the required plots following input needs to be provided in the class **RealEstateCorrelation** on the lines from **#328** through line **#332**
4. The input details are as follows:

* **All** pertains to 🡪 All types of bedroom (1 BR, 2 BR, 3BR, 4BR, 5BR). If specific bedroom type analysis is to be conducted, provide the input as “3” or “2” etc.
* **State** pertains to🡪 based on State. It can take values -- **‘State’, ‘Region’, ‘City’, ‘County’**
* **NY** pertains to 🡪 The name of the State or the name of the region or city or county if they are selected. If the Zip is selected instead of State, then zip code needs to be given.
* **State** pertains to 🡪 based on State. It can take values -- **‘State’, ‘Region’, ‘City’, ‘County’**
* **New York** pertains to 🡪 The name of the State or the name of the region or city or county if they are selected. If the Zip is selected instead of State, then zip code needs to be given. This is the full form of the name of the State like Virginia, California.
* **2011** pertains to 🡪 **Start year** from which the correlation needs to be calculated.
* **2021** pertains to 🡪 **End year** till which the correlation needs to be calculated.

1. The lines from #329 till #332 take the inputs provided above.

|  |
| --- |
| **Line #328 through Line #332**  **#328** - cspds\_df = real\_estate\_corr.get\_hindex\_cspds\_corr(**'All', 'State', 'NY', 'State', 'New York', '2011', '2021'**)  **#329** - corr\_lst, num\_bed\_corr\_lst\_dict, hIndex\_df\_dict, sp\_df = real\_estate\_corr.get\_hindex\_sp\_corr\_fr\_yrs(**"All", 'State', 'NY', '2021'**)  **#330** - hIndex\_df\_dict\_on\_year = real\_estate\_corr.get\_raw\_hindex(**'All', 'State', 'NY', '2021'**)  **#331** - sp\_data\_on\_year = real\_estate\_corr.sp\_cspds\_data.get\_sp\_data(**'2011', '2021'**)  **#332**- real\_estate\_corr.plot\_trend\_on\_yr\_multi\_plot(**'NY'**, hIndex\_df\_dict\_on\_year, sp\_data\_on\_year, corr\_lst, num\_bed\_corr\_lst\_dict, hIndex\_df\_dict, sp\_df, cspds\_df, **'2021', '2011', '2021'**) |