**ASSIGNMENT -1 REPORT**

Name : Vasantha Kavya Karanam ( 22021986 )

Data Source : <https://landregistry.data.gov.uk/app/ukhpi/browse?from=2020-01-01&location=http%3A%2F%2Flandregistry.data.gov.uk%2Fid%2Fregion%2Funited-kingdom&to=2022-10-01&lang=en>

**Summary of assignment:**

This project main theme is to analyze the process of data visualization that has been implemented by using Spyder IDE software. Three different approaches have been implemented while executing the assignment. With the help of this project, an overview related to UK house prices data can be understand.

**Intro:**

In order to visualize the data, it has been required to collect a set of data based on which the overall execution has been made. In case of doing that, the application of a secondary data collection method has been chosen. In terms of being specific, it can be said that the data has been collected from internet sources. Apart from that, it can also be added that in this project, a data on the UK house price index has been selected. This dataset has kept the records of the house prices that have changed throughout the months and years. Along with that, this dataset has also kept records of the Percentage change (yearly) Detached houses, and other related variables and their respective dates as well. Apart from that it has also been instructed that the task of data visualization is required to be executed by following three different methods.

**Discussion:**

This section of the assignment has been executed in terms of discussing the outputs that have been obtained while implementing the data visualization assignment. In order to execute this assignment application of python programming language has been required to be used.

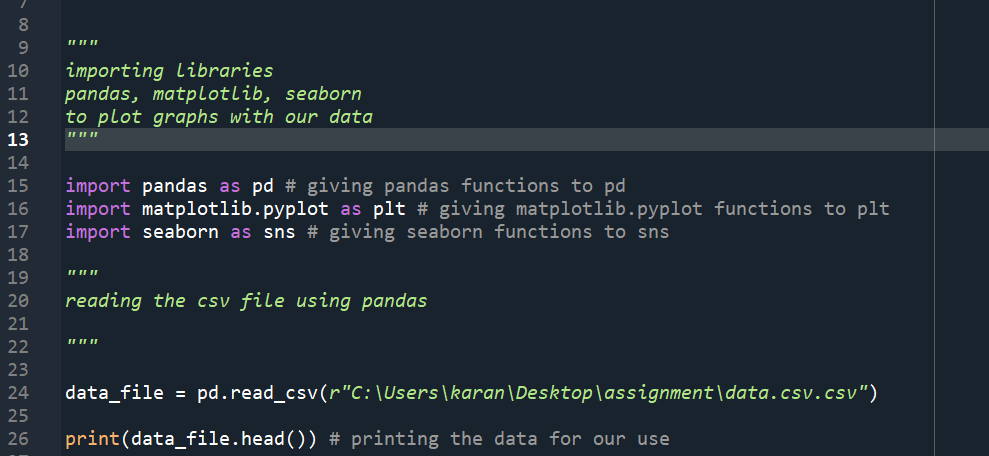


Figure 1: Importing libraries and data

(Source: Self-created)

The above-attached diagram is the reflection of the action that has been performed in terms of reading the libraries and the data that has been imported in terms of executing the task accordingly. It has been instructed that the data is required to be read by using the pandas' library and in case of doing that Pandas library has been imported initially. Once the library has been imported then the selected dataset has been imported by using the pandas' library. Apart from that, it can also be noticed that the selected data has been defined in a certain data frame that has been named data\_file.

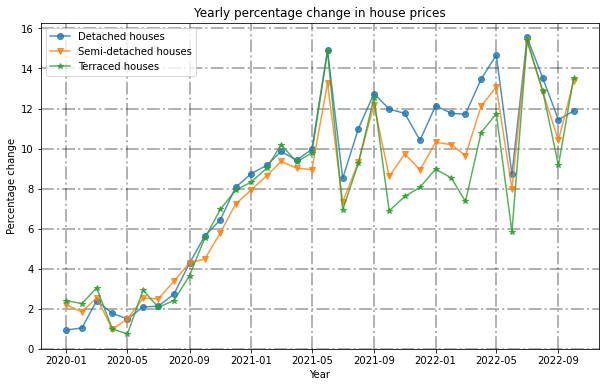


Figure 2: Line plot

(Source: Self-created)

A specific requirement has been given in order to do this assignment that says a line plot is required to be plotted according to the values inside the data. It has also been instructed that three different methods need to be implemented in case of having a successful outcome. As a reflection of that, it can be observed that the line plot has been plotted in this section. In the case of plotting the line plot, the matplotlib library has been used. With the help of the above-mentioned methods, a line plot of the Percentage change (yearly) of three different type houses, namely detached, semi-detached, terraced have been displayed that has varied over the different time periods. In order to be precise, it can be said that the graph has mostly differed during the time span of 01-2020 to 10-2022.

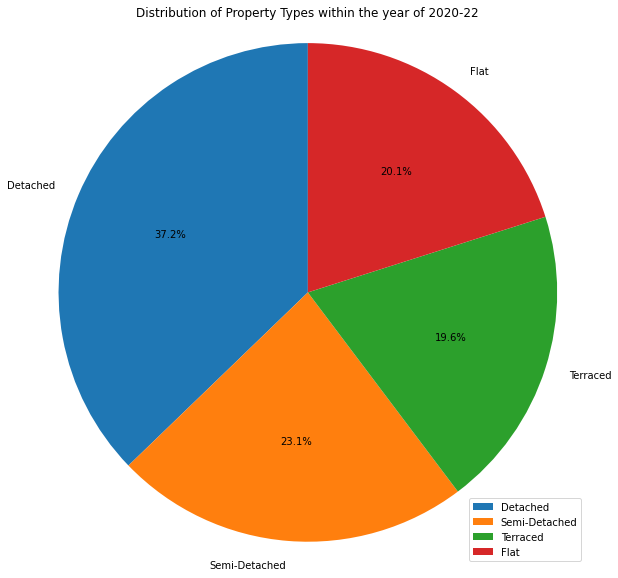


Figure 3: Pie plot

(Source: Self-created)

As per the image that has been attached above it can be noticed that a pie plot has been plotted during the execution of the assignment. However, in this case, a different method has been selected while plotting the above pasted image. In that case, it can be noticed that the plotly library has been used. Based on the pie plot that has been pasted in the above figure it can be said that this graph represents the data of price distribution of houses over the time period.

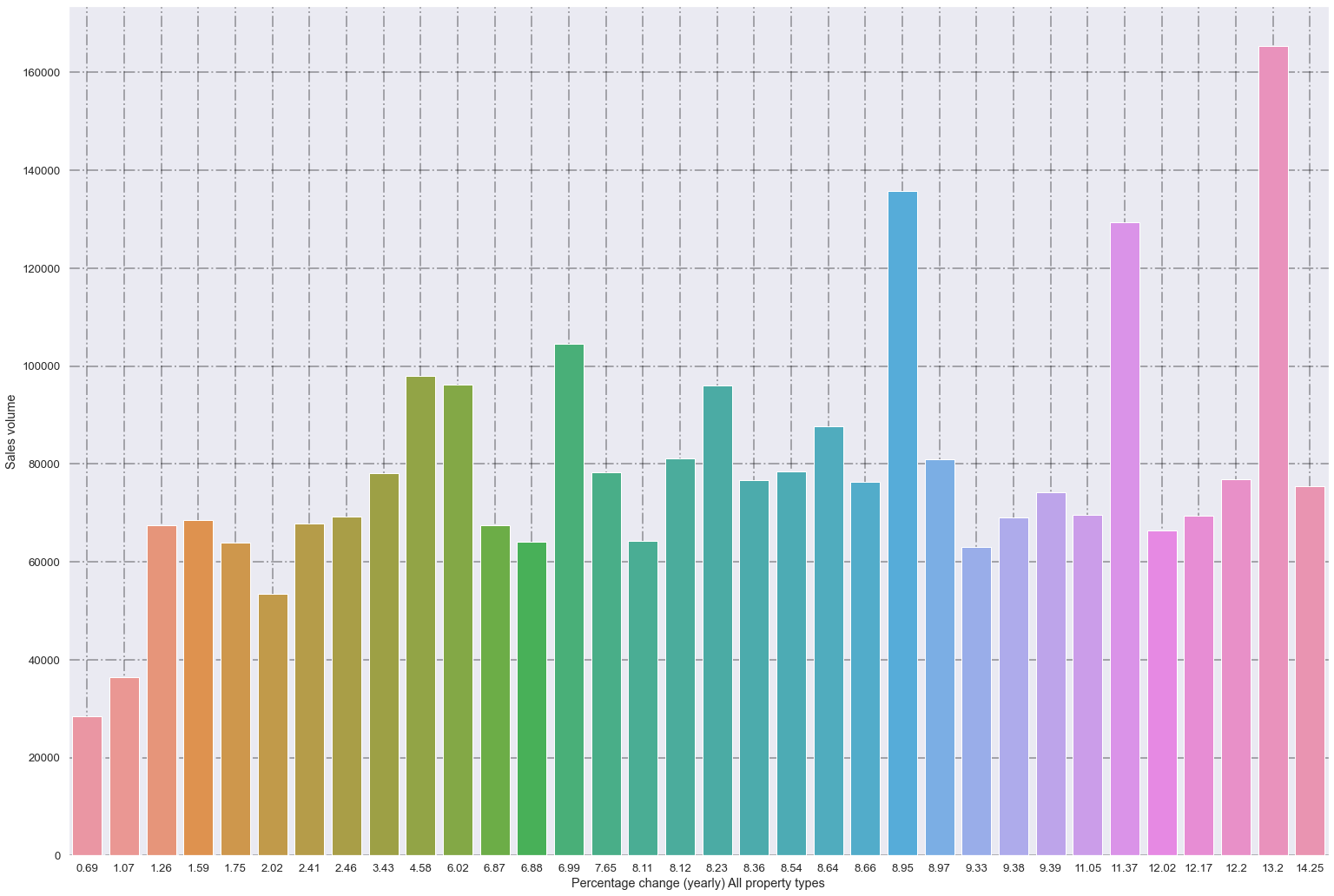


Figure 4: Bar plot

(Source: Self-created)

According to the instruction another graph has been displayed that can be noticed from the image that has been attached above. In this case, the bar plot has been plotted by using the seaborn library. In most of the cases, this library is used to visualize the random distribution within the data. In order to be precise, it can be said that the above plotted graph has been plotted in terms of displaying the distributions between the sales volume and Percentage change (yearly) of all property types variable. Based on the graph that has been plotted above that the percentage changes ratio of 13.2 has the highest distribution of sales volume.

**Conclusion:**

In conclusion, it can be said that this project discusses the process of data visualization that has been executed with the help of the python programming language. This project has been beneficial in terms of displaying the relationships, and the distributions of the data. In terms of displaying certain data related to the UK house price index been considered that has been by following a