**MAKERERE     UNIVERSITY**

**COLLEGE OF COMPUTING AND INFORMATICS**

**SYSTEMS**

**DEPARTMENT OF NETWORKS.**

**GROUP7**

**SOFTWARE REQUIREMENT**

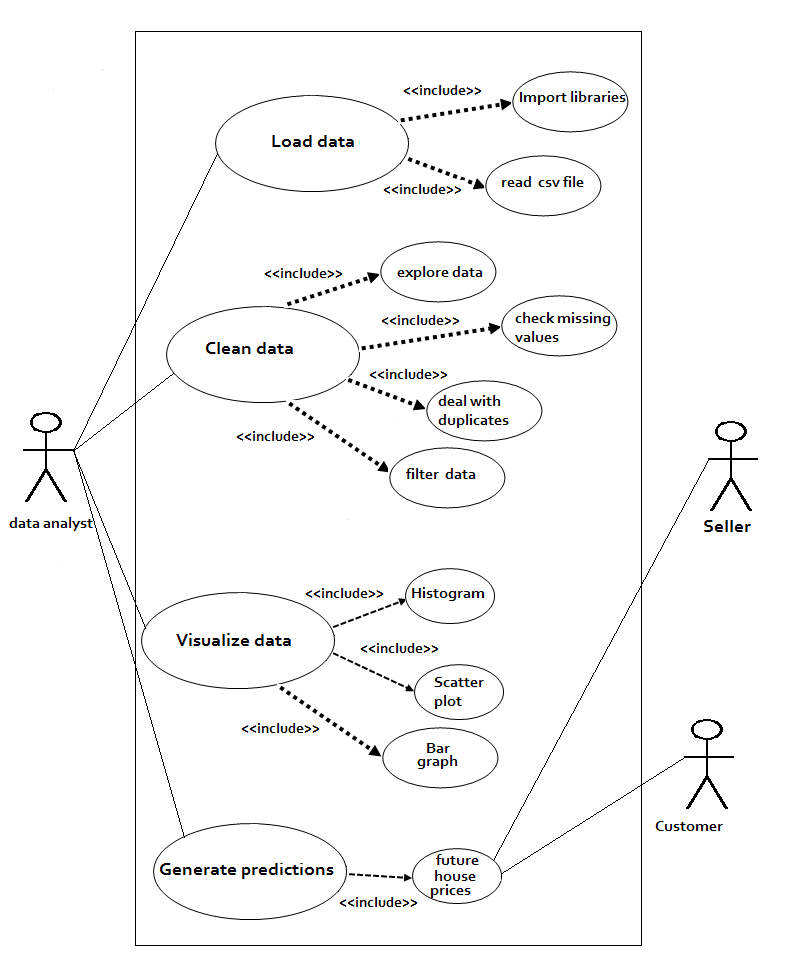
**SPECIFICATION (SRS)**

*GROUP7 MEMBERS:*

|  |  |  |  |
| --- | --- | --- | --- |
| **NO** | **NAMES** | **Reg No** | **Stud No** |
| 1 | Niwandiinda John Martin | 17/U/9138/PS | 217010696 |
| 2 | Miiro Henry | 17/U/6112/EVE | 217009093 |
| 3 | Onyang Joshua | 17/U/9702/PS | 217016961 |
| 4 | Odongo Abraham | 17/ U/9494/PS | 217002258 |

**SOFTWARE REQUIREMENT SPECIFICATION FOR HOUSE PRICE PREDICTION SYSTEM.**

This system will be used by the three categories of people that is data analyst, customer (buyer) and the seller as shown in the use case below.

x

*The breakdown of the use case.*

*Load data*

|  |  |  |
| --- | --- | --- |
| Use case name: | Load data | |
| Scenario: | Load the house\_prices dataset | |
| Brief description: | When the data analyst loads the dataset, its format changes from csv to data frame by pandas. | |
| Actors: | Pandas library. | |
| Related use cases: | Includes: import libraries, read csv file. | |
| Pre-condition: | Python libraries that is pandas must be imported | |
| Postconditions: | Data must be in a table / data frame format | |
| Flow of Events: | **Actor** | System |
| Pandas manipulates the csv file into table format. | Loads data |

***Clean data***

|  |  |  |
| --- | --- | --- |
| Use case name: | Clean data | |
| Scenario: | Clean the loaded data | |
| Brief description: | After loading the data, data analyst checks for the missing values. If any, deal with them. Also checks for the duplicates, deals with them and filter the data. | |
| Actors: | Pandas and NumPy libraries. | |
| Related use cases: | Includes: explore data, check missing vales, deal with duplicates, filter data. | |
| Pre-condition: | The data must be explored. | |
| Postconditions: | Data must be clean and ready for visualization | |
| Flow of Events: | **Actor** | System |
| Pandas manipulates the csv file into table format.  NumPy provides high performance multidimensional array object and tools for working with these arrays. | Cleans data |

***Visualize data***

|  |  |  |
| --- | --- | --- |
| Use case name: | Visualize data. | |
| Scenario: | Visualize the cleaned data. | |
| Brief description: | The data analyst visualizes the cleaned data by plotting various graphs including histograms, bar graphs and scatter plots. | |
| Actors: | Matplotlib and seaborn libraries. | |
| Related use cases: | Includes: histogram, scatter plot, bar graph. | |
| Pre-condition: | Libraries i.e. Matplotlib and seaborn must be imported | |
| Postconditions: | The insight must be drawn from the graphs. | |
| Flow of Events: | **Actor** | System |
| 1. Matplotlib creates 2D graphs and plots by using python scripts. 2. Seaborn provides a high-level interface for drawing attractive and informative statistical graphics. | Displays graphs. |

*Generate prediction*

|  |  |  |
| --- | --- | --- |
| Use case name: | Generate predictions | |
| Scenario: | Generate house price predicted | |
| Brief description: | * The data analyst applies linear regression to generate the predicted house prices. * The seller checks the generated prediction and determines the factors that lead to an increase or decrease in house prices which include renovation, bed space, location, living space, grade, etc. * The customer checks the future price predictions and determines whether to buy now or not. | |
| Actors: | Scikit learn library. | |
| Related use cases: | Includes: future house prices. | |
| Pre-condition: | * Scikit learn must be imported * Data must be loaded, cleaned and visualized | |
| Postconditions: | Future house prices must be generated. | |
| Flow of Events: | **Actor** | System |
| Scikit learn provides a number of algorithms which generate predictions. | Generates house price predictions. |