*Case Study 2:*

Load practice.csv file as a data-frame and perform following operations on the data-frame

1. Display all columns
2. create numerical and categorical columns list
3. display size of the data-frame
4. rename column MSSubClass -> SubClass, MSZoning -> Zones
5. display distinct values for Zoning, LotShape, LotConfig
6. display count of distinct values for Zoning, LotShape, LotConfig
7. max, min of column YearBuilt
8. create a new column “year\_diff’. This will be holding difference of current year and YearBuilt
9. display distinct MSZoning for each OverallQual
10. What is the maximum LotArea where BsmtExposure = Mn?
11. Sort dataframe based on following columns and orders: MSSubClass; ascending, YearBuilt; descending
12. What is average OverallQual.
13. convert column ‘YearBuilt’ into date type.
14. Group by YearBuilt and find maximum OverallQal
15. Load the practice.csv again with MSSubClass as new index
16. Convert LotArea as numpy array
17. In column MasVnrArea replace 0 with -1
18. Check if there is/are any Null values (NaN) in the data given
19. Display percentage of missing values in each column if any
20. Select records where LotConfig is Inside
21. Make a new dataframe with only numeric columns
22. Make a new dataframe with only factorial/string columns
23. Drop column ExterQual
24. Group data on LotShape and find average LotArea
25. Write code to get a pivot table as shown (average of **MSSubClass**)
26. 