Experiment No. 8

Aim:	To implement the concepts of Linear regression, Regression Analysis & K-Means clustering in python.

Experiment Number 8:

Aim: To implement & understand the concepts of Linear regression and Regression Analysis in python by predicting house prices using the number of rooms as the feature and concept of k-means clustering in python.

Theory:

Linear Regression:

Linear Regression is a fundamental statistical method used to model the relationship between two variables by fitting a linear equation to the observed data. It assumes a linear relationship between the independent variable (X) & the dependent variable (Y).

The equation for simple linear regression is:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

where, $X \equiv$ dependent variable

$Y \equiv$ Independent variable

$\beta_0 \equiv$ is the intercept, the value of Y when $X = 0$

$\beta_1 \equiv$ is the slope of regression line, representing the change in Y for a unit change in X

$\epsilon \equiv$ is the error term, representing the difference between observed & predicted values.

Thus the goal of the linear regression is to minimize the sum of squared errors (differences between observed & predicted values) to find the best fitting line through the data.

Regression Analysis :

Regression Analysis is a statistical technique used to investigate & model the relationships between a dependent variable and one or more independent variables. The analysis helps us understand how the typical value of the dependent variable changes when any one of the independent variables is varied.

There are mainly two types of Regression :

i) Simple Linear Regression : Involves one independent variable & one dependent variable.

ii) Multiple Linear Regression : Involves multiple independent variables predicting the dependent variable.

Regression Analysis provides insights into :

i) How much variation in the dependent variable can be explained by the independent variable.

ii) The strength & type of relationship between the variables.

iii) The significance of each predictor in the model.

Applications of Linear Regression :

i) Predicting the sales or prices based on historical data.

ii) Estimating the impact of changes in one var. on

K-means clustering is an unsupervised learning method for clustering data points. This algorithm iteratively divides data points into K clusters by minimizing the variance in each cluster.

Code:

```
#Implementation of Linear Regression:
import matplotlib.pyplot as plt.
x = [1,2,3,7,8,9,0,4,6]
y = [99,69,77,81,92,100,40,62,85]
plt.scatter(x,y)
plt.show()

#R for Relationship
from scipy import stats
x = [1,2,3,4,5,6,7,8,9,10]
y = [11,22,33,44,55,66,77,88,99]

slope, intercept, r, p, std_error = stats.linregress
                                    (x,y)
print(r)

#Implementation of K-means
from sklearn.cluster import KMeans
data = list(zip(x,y))
inertias = []
```

Conclusion: We have Successfully implemented Linear Regression & K-Means clustering method's in Python.

```
for i in range (1,11):
        KMeans = KMeans (n_clusters = i)
        KMeans.fit (data)
        inertias.en append (KMeans, inertias)

plt.plot(range (1,11), inertias, marker = 'o')
plt.title ('Elbow Method')
plt.xtitle('Number of Clusters')
plt.ytitle ('Inertia')
plt.show ()
```

Conclusion: We have successfully implemented Linear Regression & K-Means ~~in~~ clustering methods in python.