Task 1 :

Linear regression is a method that trains our models which has dependent variable and one or more independent variables by the best fit curve bewteen them.

LinearRegression.fit() takes a set of input features (independent variables) and a target variable (dependent variable) and fits a linear equation to the data that minimizes the distance between the predicted values and the actual values of the target variable and can be used to make predictions on new data. This method uses least squares technique to estimate the coefficients of the linear equation.

Model = LinearRegression

Lreg = model.fit(X,Y) // Lreg.coeff_ gives the array of independent variables while Lreg.intercept_ gives the bias

Task 2 :

Gradient descent works iteratively by adjusting the values of the coefficients so that the curve / line moves towards the best fit curve/line for the data points given

In the case of one independent and one dependent variable we have the equation y = mx + c (where x is the independent variable and y is the dependent variable).In this scenario,we take random values for m and c and calculate the cost function by subtracting the predicted values from the data points and the actual values and taking the mean of these values after squaring them . Then we take the derivative of the cost function and update m and c by changing them in the direction opposite to the gradient with Learning_rate*gradient.These steps are repeated until we reach the minimum cost or a minimum no of iterations defined.

Task 3:

Degree	Bias	Variance
1	0.268809	0.00875452
2	0.0860908	0.00123017
3	0.0332896	0.00034148
4	0.0250848	0.000451292
5	0.0244016	0.000595884
6	0.0248141	0.000884395
7	0.0261302	0.00314885
8	0.0262656	0.00445863
9	0.0258572	0.00860399
10	0.027602	0.0688098
11	0.0498831	0.192034
12	0.0331709	0.134254
13	0.0340476	0.150915
14	0.157684	5.53852
15	0.230432	14.5806

THE DATA IS SPLITTED RANDOMLY INTO 20 PARTS

Here we can see that as the degree of the polynomial is increasing the bias is decreasing steadily while the variance is increasing. This is because our model is performing well on the training set as it is increasingly memorizing the training set values but it is performing bad on the test set as it has remembered the values too properly. The exact values in bias and variance are dependent on the data sets chosen and here we have divided randomly but the trends will moreover be same as bias will decrease and variance will increase

– Bias refers to error due to incorrect assumptions in learning algorithm

– Variance measures error due to small fluctuations in training set

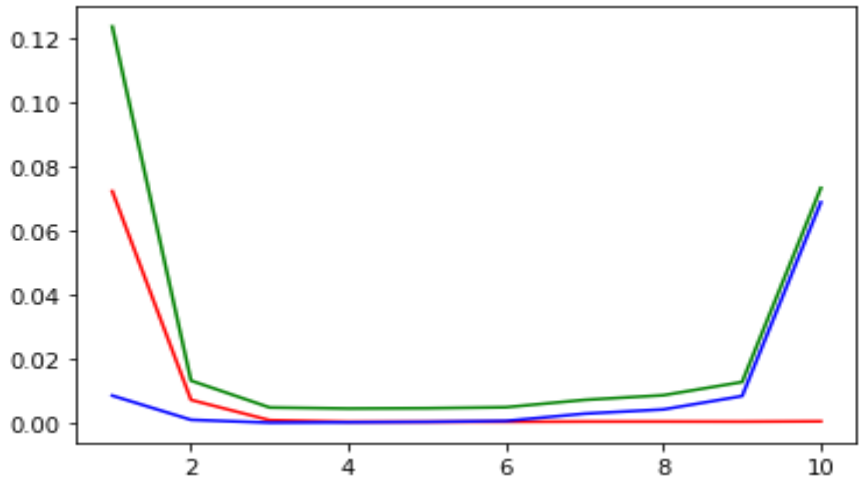
Task 4:

Degree	Bias	Variance	Irreducible Error
1	0.268809	0.00875452	1.97121e-16
2	0.0860908	0.00123017	-1.587e-18
3	0.0332896	0.00034148	-2.63453e-16
4	0.0250848	0.000451292	-3.86244e-16
5	0.0244016	0.000595884	-2.01645e-16
6	0.0248141	0.000884395	3.26372e-16
7	0.0261302	0.00314885	-6.25957e-16
8	0.0262656	0.00445863	-7.29939e-19
9	0.0258572	0.00860399	1.16927e-15
10	0.027602	0.0688098	8.06466e-16
11	0.0498831	0.192034	5.97314e-16
12	0.0331709	0.134254	-1.41803e-16
13	0.0340476	0.150915	9.53969e-16
14	0.157684	5.53852	5.9453e-16
15	0.230432	14.5806	-2.48197e-15

Irreducible error represents the inherent noise or variability in the data that cannot be explained by the features used in a model.

Here for some values irreducible error is coming out to be negative and the magnitude is of the order 10^-17 in almost all of them . The irreducible error is negative due to the precision error in python.

Task 5:



The red line denotes the bias(squared) while the blue lines denotes the variance . The green line denotes the MSE

Here we can see the bias2-variance graph upto degree 10 . We can see that as the degree increases bias falls steadily and variance rises giving the optimal model of degree 5 where the two lines intersect. This means that the models from 1 to 4 are underfitted models as they have high bias and high variance as seen from the figure. The models above 6 are overfitted models as they have low bias and high variance.At the Degree 5 we have low bias and low variance where the model is best fit . We can say that the data is well caliberated as the bias and variance are quite low at the fit region of the graph