Statistical Learning
IIT Madras

Assignment 1

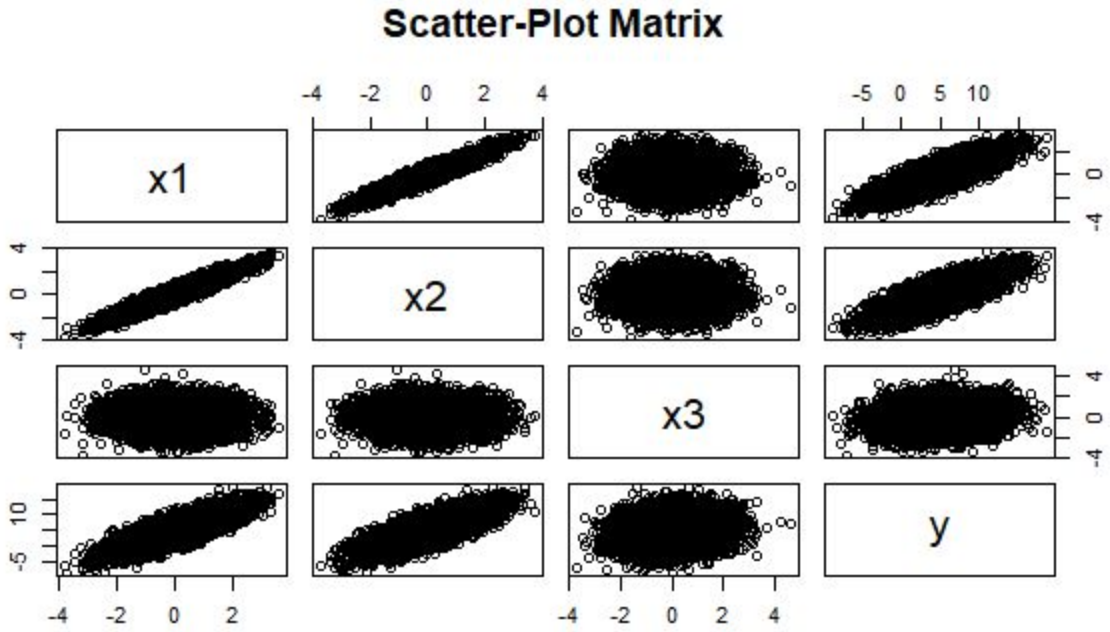
Submitted by:
Vivek Kumar Agrawal
Roll No.: CS17M094

Question 1. Correlation Assignment:

1.1 Correlation among predictors and y:

Correlation	x1	x2	x3	y
x1	1.000000e+00	9.500000e-01	-5.730157e-18	0.8235151
x2	9.500000e-01	1.000000e+00	-1.694362e-16	0.7816903
x3	-5.730157e-18	-1.694362e-16	1.000000e+00	0.1343833
y	0.8235151	0.7816903	0.1343833	1.000000e+00

1.2 Scatter-Plot Matrix



1.3 Predictors affecting y:

Predictors **x1** and **x2** has huge influence on **y** while the correlation between **x3** and **y** is very low that is **x3** won't affect **y** much.

1.4 Coefficients of fitted linear model to given data points:

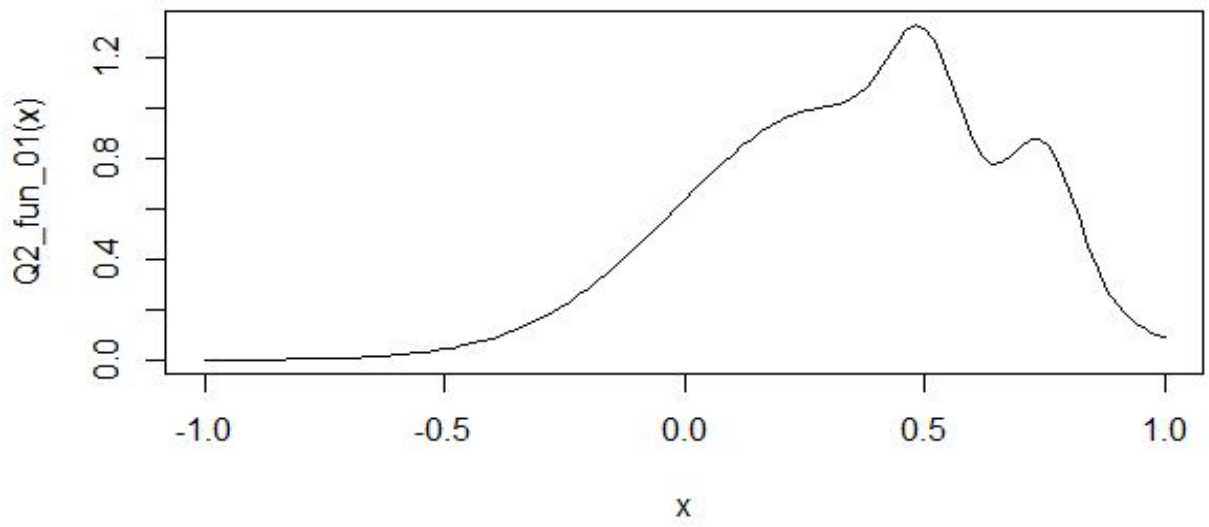
Intercept	**x1**	**x2**	**x3**
4.99599	3.03724	-0.02436	0.49185

The significant predictors are those which have higher coefficient values compared to other predictors. Therefore, here the most significant predictor is **x1** after that **x3** and then **x2.**

Question 2. Regression- Polynomial Fitting:

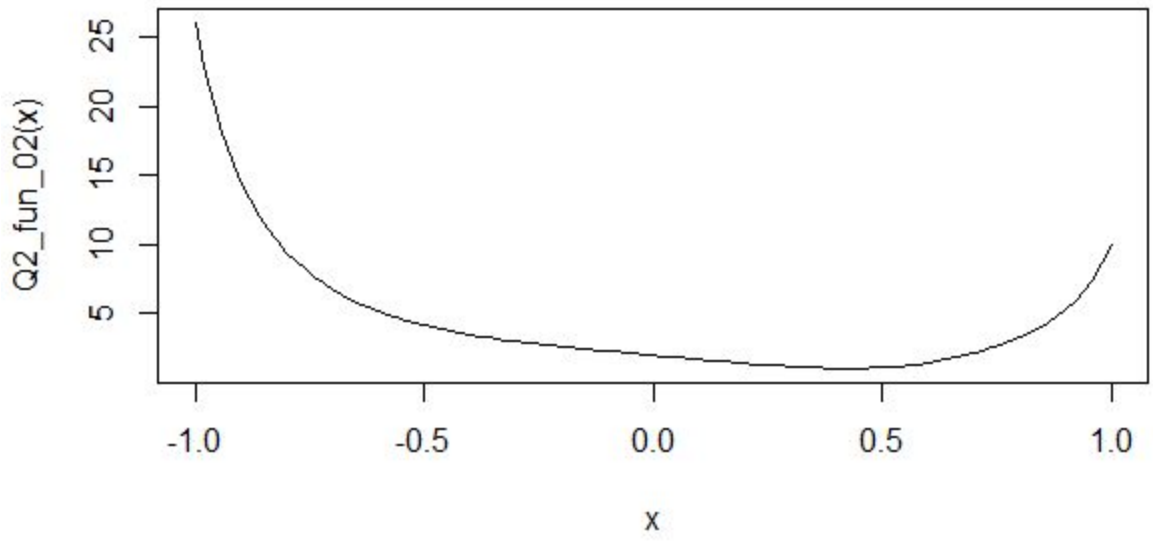
2.1 Plot of function Q2_fun_01:

```
f(x) = exp(-5*(x-0.3)^2)+0.5*exp(-100*(x-0.5)^2)+0.5*exp(-100*(x-0.75)^2)
```



Plot of function Q2_fun_02:

```
f(x) = 2-3*x+10*x^4-5*x^9+6*x^14
```

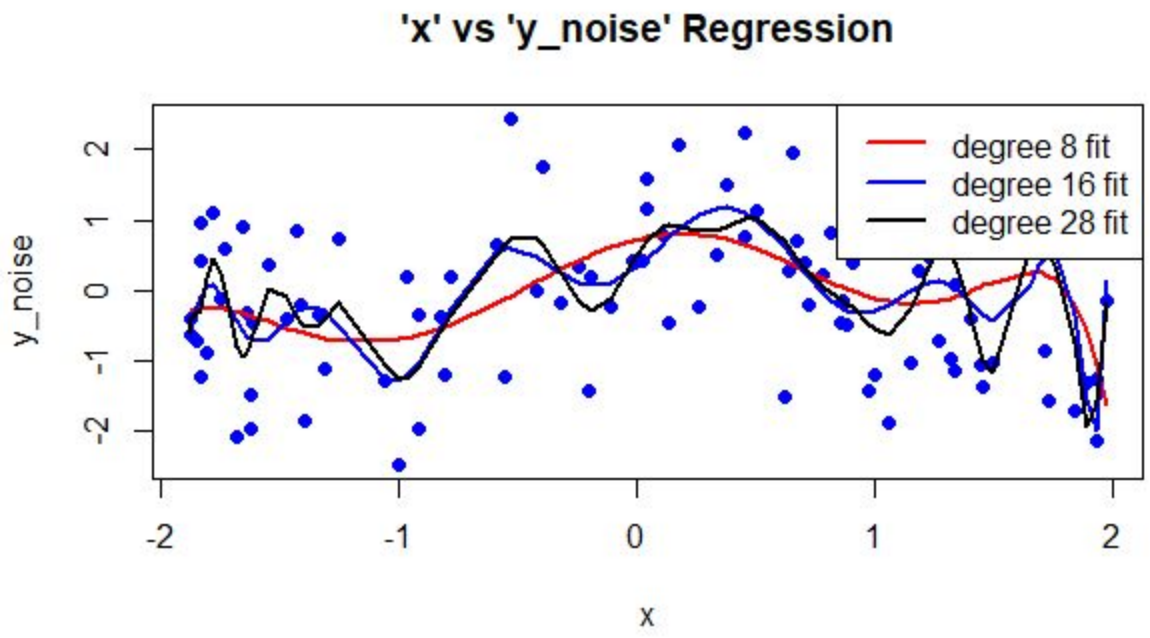


2.2 Below is the code snippet of generating 100 random points from function *Q2_fun_01* & *Q2_fun_02* and adding normal distributed noise to their respected *y.*

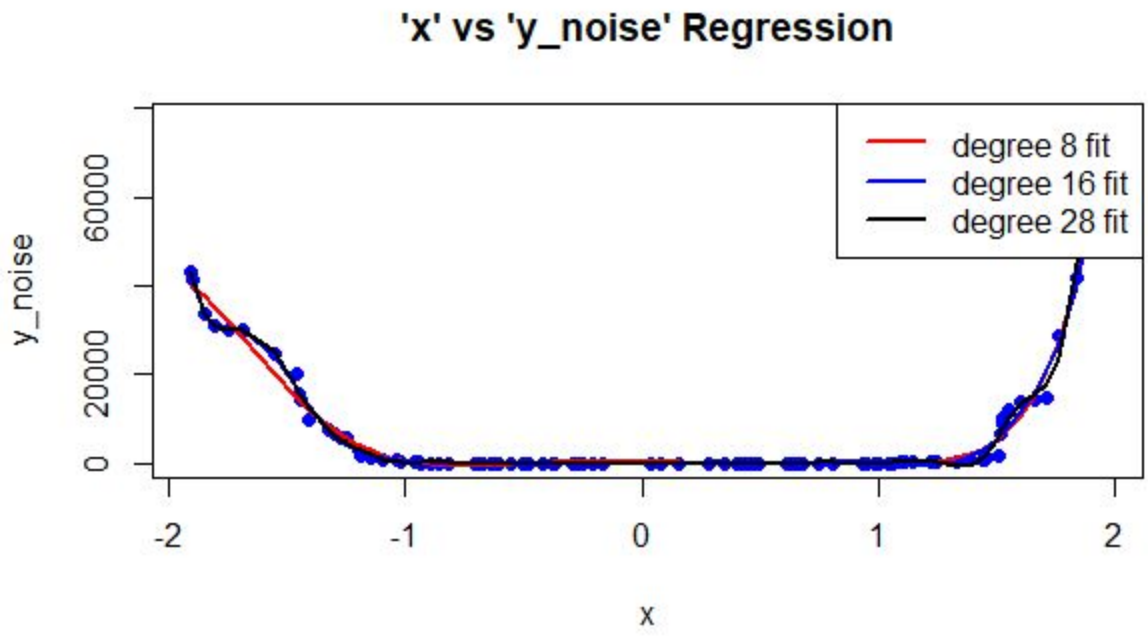
```
#Q2_fun_01
set.seed(123)
x1_points = runif(100)
y1_points = Q2_fun_01(x1_points)
y1_noisy = y1_points + rnorm(100)

#Q2_fun_02
set.seed(123)
x2_points = runif(100)
y2_points = Q2_fun_02(x2_points)
y2_noisy = y2_points + rnorm(100)
```

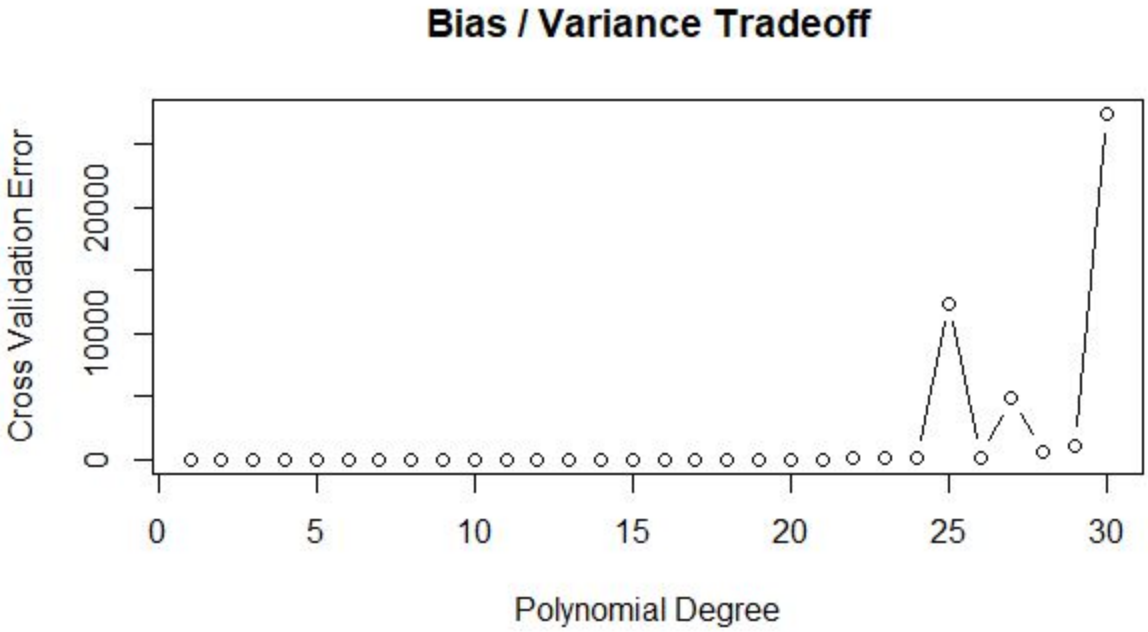
2.3 Fit wrt. Q2_fun_01:



Fit wrt. Q2_fun_02:



2.4-2.5 Bias_Variance Plot for Q2_fun_01:



Bias_Variance Plot for Q2_fun_02: