ISCO630E

*Analysis Report on **Assignment 1***
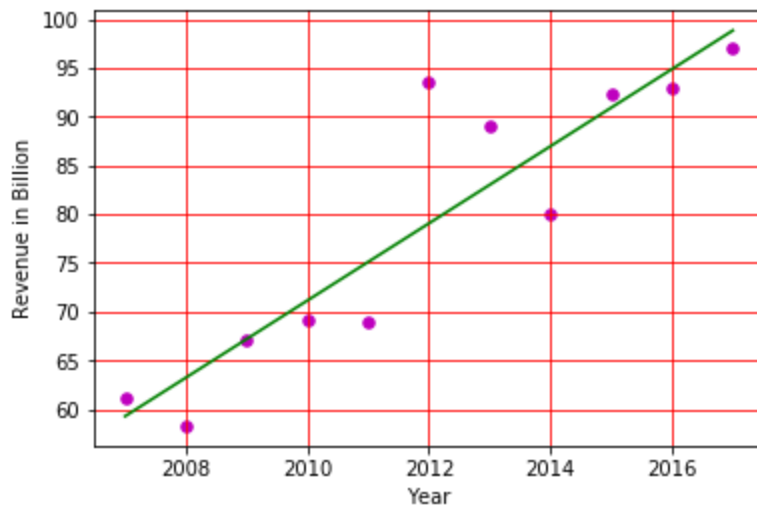
Group Members :
1. *Sashank Mishra - IIT2016515*
2. *Shreyansh Dwivedi - IWM2016501*
3. *Saurabh Kumar - IWM2016502*
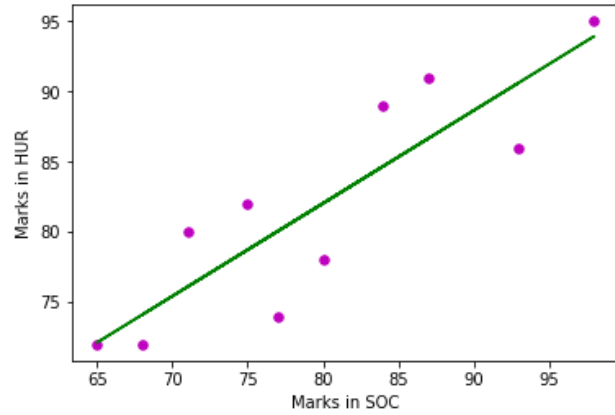4. *Shreyansh Chaudhary - IIT2016068*
5. *Pradeep Gangwar - IHM2016501*

QUESTION 1,2,3,4

Question 1, 2, 3 and 4 were simple calculations based on small given data. We had to calculate the values for some particular values of x after drawing the regression line.
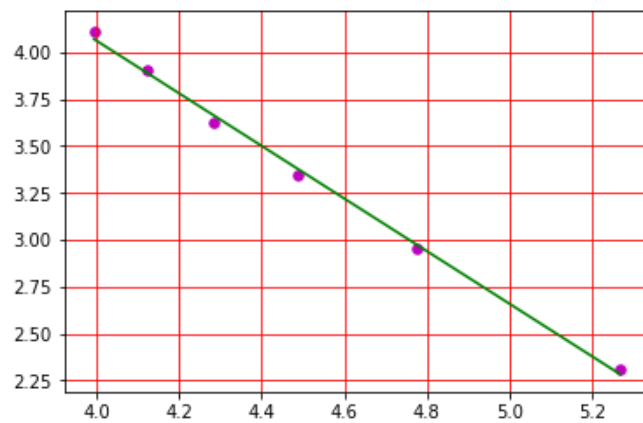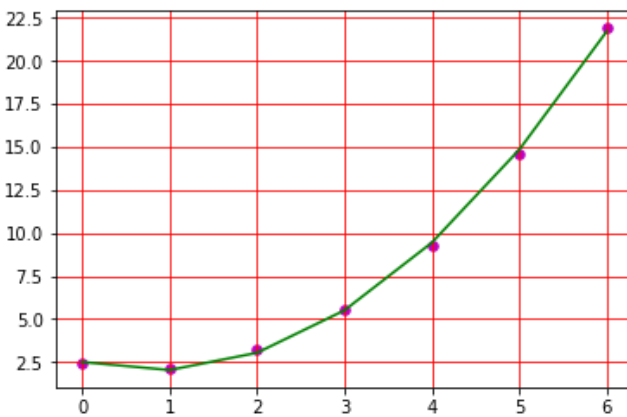
Question- 1

Question- 2

Question- 3

Question- 4

QUESTION 5

Dataset : Housing Price Prediction (in CSV format)
546 rows, 12 columns

price	lotsize	bedrooms	bathrms	stories	driveway	recroom	fullbase	gashw	airco	garagepl	prefarea
42000.0	5850	3	1	2	yes	no	yes	no	no	1	no
38500.0	4000	2	1	1	yes	no	no	no	no	0	no
49500.0	3060	3	1	1	yes	no	no	no	no	0	no
60500.0	6650	3	1	2	yes	yes	no	no	no	0	no
61000.0	6360	2	1	1	yes	no	no	no	no	0	no

Data Preprocessing :

- All the columns containing the option as 'yes' and 'no' are mapped to 0 and 1, respectively.
  The final data representation:

price	lotsize	bedrooms	bathrms	stories	driveway	recroom	fullbase	gashw	airco	garagepl	prefarea
42000.0	5850	3	1	2	1	0	1	0	0	1	0
38500.0	4000	2	1	1	1	0	0	0	0	0	0
49500.0	3060	3	1	1	1	0	0	0	0	0	0
60500.0	6650	3	1	2	1	1	0	0	0	0	0
61000.0	6360	2	1	1	1	0	0	0	0	0	0

- Data splitted into features and prediction(X and Y, respectively)
  Prediction variable here is the 'price'
- Data is normalized, and a Bias term added to feature set.

Analysis :

Initially, we applied the Regression fit without normalization. In that case, the normal equations was working fine, but the Gradient Descent gave values as 'NaN'.
Then, we did the normalization of the data, then both the algorithms worked correctly.

Final Cost obtained using Normal Equations - 0.1634

No. of iterations	Learning Rate	Final Cost
10000	0.0001	0.4585
10000	0.001	0.3559
10000	0.01	0.3561
1000	0.001	0.3587

The graph of cost function v/s iteration is obtained as follows: