Marwadi University	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Obtain the best fit non-linear line using polynomial regression in R**	
Experiment: 11	**Date: 18/04/2023**	**Enrollment No: 92200133030**

Aim: Obtain the best fit non-linear line using polynomial regression in R

IDE: R Studio

Theory:

Polynomial regression is a form of Linear regression where only due to the Non-linear relationship between dependent and independent variables, we add some polynomial terms to linear regression to convert it into Polynomial regression. In polynomial regression, the relationship between the dependent variable and the independent variable is modeled as an nth-degree polynomial function. When the polynomial is of degree 2, it is called a quadratic model; when the degree of a polynomial is 3, it is called a cubic model, and so on. Suppose we have a dataset where variable X represents the independent data and Y is the dependent data. Before feeding data to a mode in the preprocessing stage, we convert the input variables into polynomial terms using some degree.

Consider an example my input value is 35, and the degree of a polynomial is 2, so I will find 35 power 0, 35 power 1, and 35 power 2 this helps to interpret the non-linear relationship in data. The equation of polynomials becomes something like this.

$$y = a_0 + a_1x_1 + a_2x_1^2 + \ldots + a_nx_1^n$$

The degree of order which to use is a Hyperparameter, and we need to choose it wisely. But using a high degree of polynomial tries to overfit the data, and for smaller values of degree, the model tries to underfit, so we need to find the optimum value of a degree. Polynomial Regression models are usually fitted with the method of least squares. The least square method minimizes the variance of the coefficients under the Gauss-Markov Theorem.

Why Is Polynomial Regression Called Polynomial Linear Regression?
If you see the equation of polynomial regression carefully, then we can see that we are trying to estimate the relationship between coefficients and y. And the values of x and y are already given to us, only we need to determine coefficients, and the degree of coefficient here is 1 only, and degree one represents simple linear regression Hence, Polynomial Regression is also known as Polynomial Linear Regression as it has a polynomial equation and this is only the simple concept behind this. I hope you got the point right.

Linear Regression Vs. Polynomial Regression
Now we know how polynomial regression works and helps to build a model over non-linear data. Let's compare both algorithms practically and see the results.
First, we will generate the data using some equation ax^2 + bx + c, and then apply simple linear regression to it to form a linear equation. Then we will apply polynomial regression on top of it, which will make an easy comparison between the practical performance of both algorithms.
Initially, we will try it with only one input column and one output column. After having a brief understanding we will try it on high-dimensional data.

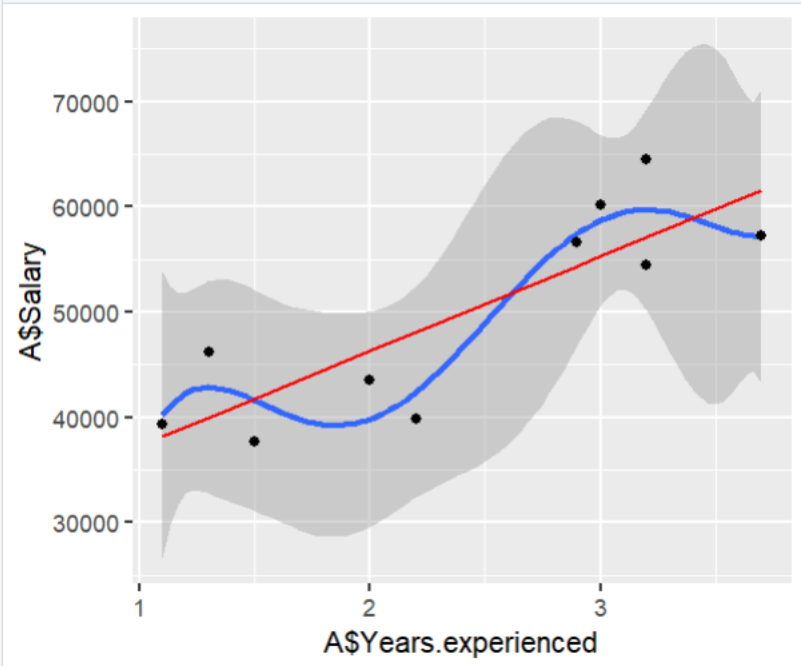
Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Introduction to R and R Studio (01CT0106)	Aim: Obtain the best fit non-linear line using polynomial regression in R	
Experiment: 11	Date: 18/04/2023	Enrollment No: 92200133030

Program:

```
poly_model = lm(formula =  Salary~poly(A$Years.experienced,5,raw = TRUE),data = A)

ggplot(data = A,aes(A$Years.experienced,A$Salary)) + geom_point() +
  stat_smooth(method = lm,formula = y~poly(x,5,raw=TRUE)) +
  geom_point(aes(x = A$Years.experienced , y = A$Salary)) +
  geom_line(aes(x = A$Years.experienced, y = predict(Linear_Reg,newdata = A)),color = "red")
```

Output:



Observation and Learnings:

_____

_____

_____

_____.