Progressive Education Society's
Modern College of Engineering, Shivajinagar, Pune-05.
(An Autonomous Institute Affiliated to Savitribai Phule Pune University)
Department of MCA

PRACTICAL SUBMISSION RECORD- A.Y. 2025-26

Class: SYMCA Division : A **Semester: III**	**Course Code:** MCA01604 **Course Name:** Data Science Laboratory	**Batch:** S2
Name: Pranav Raju Malwatkar	**Roll No:** 52037	
CO No: CO605.1	**Assignment No:** 6	

Title : Use a dataset with features like square footage vs. house price. Implement a simple linear regression model to predict house prices. Plot the regression line and calculate MSE and R2 score.

Code:

```
install.packages("ggplot2")
library(ggplot2)

data <-
read.csv("house_price_data.csv")
head(data)

model <- lm(Price ~ SquareFootage, data =
data) summary(model)
predicted <- predict(model, newdata =

data) MSE <- mean((data$Price -

predicted)^2)

R2 <- summary(model)$r.squared
cat("Mean Squared Error (MSE):", MSE,
"\n") cat("R² Score:", R2, "\n")

ggplot(data, aes(x = SquareFootage, y =
  Price)) + geom_point(color = "blue") +
  geom_smooth(method = "lm", col = "red")
```

```
+
labs(title = "House Price Prediction using Linear
    Regression", x = "Square Footage",
    y = "House Price") +
theme_minimal()
```

Output :

```
> head(data)
  Experience   Salary
1   3.588198 49207.58
2   8.094746 70331.00
3   4.680792 53189.61
4   8.947157 81578.79
5   9.464206 76192.17
6   1.410008 44632.40
>
```

```
> model <- lm(Salary ~ Experience, data = data)
> summary(model)

Call:
lm(formula = Salary ~ Experience, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-11189.9  -3066.1    -98.7   2981.7  11086.1

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  30005.1     1149.2   26.11   <2e-16 ***
Experience    4950.1      189.9   26.07   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4846 on 98 degrees of freedom
Multiple R-squared:  0.874,     Adjusted R-squared:  0.8727
F-statistic: 679.5 on 1 and 98 DF,  p-value: < 2.2e-16
```

```
> R2 <- summary(model)$r.squared
> cat("Mean Squared Error (MSE):", MSE, "\n")
Mean Squared Error (MSE): NaN
> cat("R² Score:", R2, "\n")
R² Score: 0.873951
```

```
y = "Salary") +   theme_minimal()
`geom_smooth()` using formula = 'y ~ x'
>
```

```
> print(predicted_value)
       1
54755.58
>
```