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:** 5	

Title: A company wants to predict the salary of employees based on their years of experience. The company has collected data on the salaries and expense of 100 employees. Use simple linear regression to model the relationship between salary and expense.

Code:

```
install.packages("ggplot2")

library(ggplot2)

data <- read.csv("C:\\Users\\Student\\Desktop\\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 :

```
> data <- read.csv("C:\\Users\\Student\\Desktop\\house_price_data.csv")
> head(data)
   Price SquareFootage
1 250000          1500
2 300000          1800
3 350000          2000
4 400000          2200
5 450000          2500
6 500000          2700
>
```

```
> summary(model)

Call:
lm(formula = Price ~ SquareFootage, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-9434.0 -4166.7   786.2  3852.2  8805.0

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -58490.566   7632.594  -7.663 5.94e-05 ***
SquareFootage    204.403      2.824  72.381 1.48e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6270 on 8 degrees of freedom
Multiple R-squared:  0.9985,    Adjusted R-squared:  0.9983
F-statistic:  5239 on 1 and 8 DF,  p-value: 1.479e-12
```

```
> cat("Mean Squared Error (MSE):", MSE, "\n")
Mean Squared Error (MSE): 31446541
> cat("R² Score:", R2, "\n")
R² Score: 0.9984753
>
```

```
House Price ) +   theme_minimal()
`geom_smooth()` using formula = 'y ~ x'
>
```