

Regression Analysis Lab Exam-2023 (16th Batch) Report

Exam date: 21st May 2025

This report summarizes the results of the regression analysis performed using R, addressing each part of the provided exam questions.

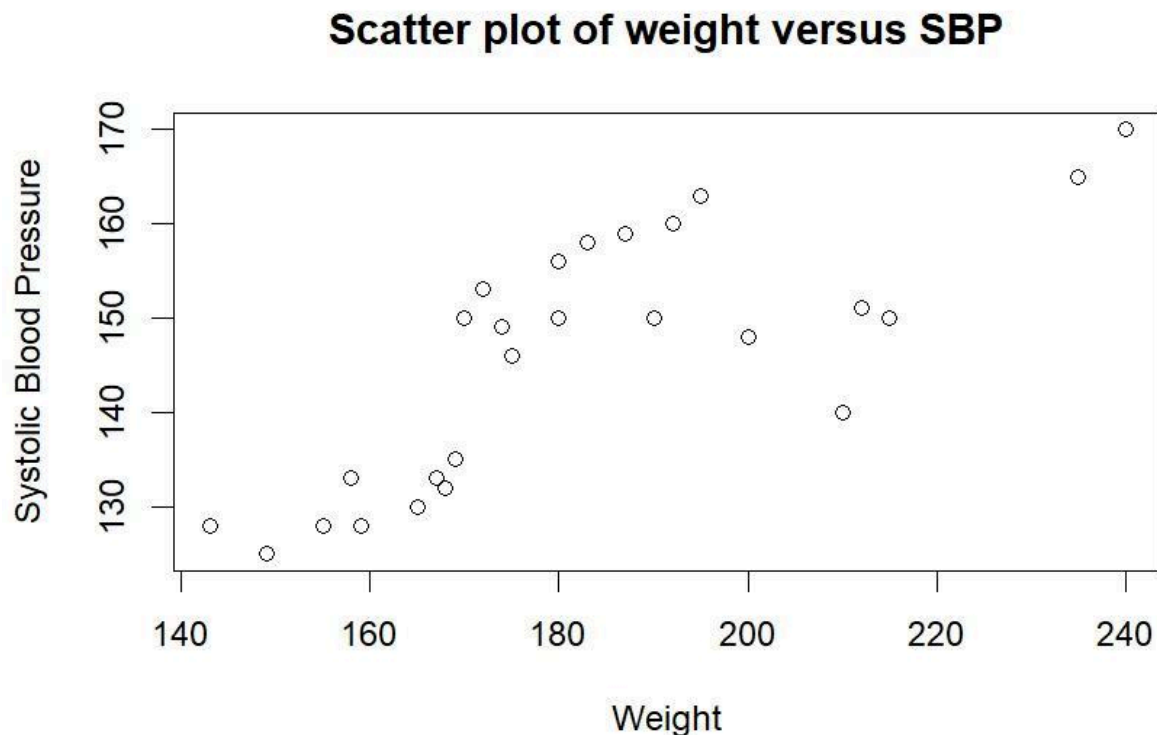
[Download R script](#)

Question 1: Simple Linear Regression (Weight vs. Systolic Blood Pressure)

This section analyzes the relationship between the weight (X) and systolic blood pressure (SBP) (Y) of 26 randomly selected males.

a. Plot weight versus SBP

A scatter plot was generated to visualize the relationship between Weight (X) and Systolic Blood Pressure (Y).



- **Observation:** The scatter plot shows a positive linear trend, indicating that as weight increases, systolic blood pressure tends to increase.

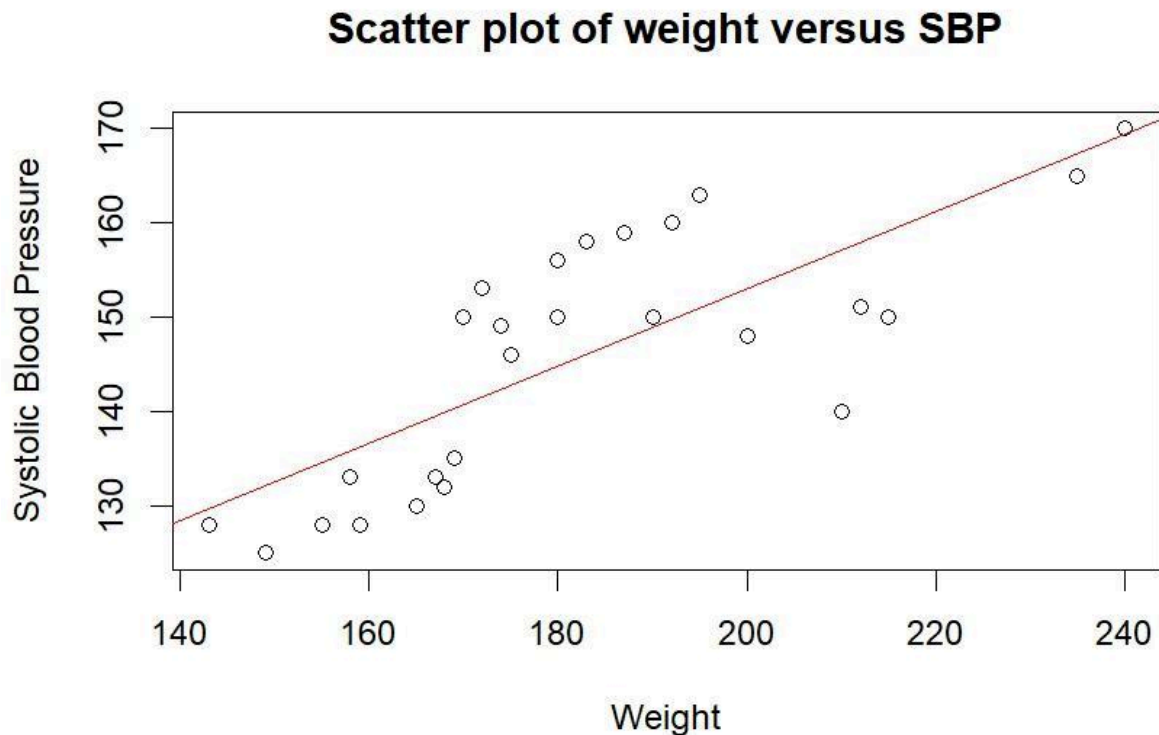
b. Fit $Y=\beta_0+\beta_1X+\epsilon$ by least squares and plot the fitted line

A simple linear regression model was fitted to the data using the least squares method. The fitted line was then plotted on the scatter diagram.

- Estimated Regression Equation:

The fitted model is:

$$\hat{Y}=71.13688+0.40912X$$



Where:

- \hat{Y} is the predicted Systolic Blood Pressure.
- X is the Weight.
- $\hat{\beta}_0 = 71.13688$ is the estimated intercept, representing the predicted SBP when weight is zero.
- $\hat{\beta}_1 = 0.40912$ is the estimated slope, indicating that for every one-unit increase in weight, the systolic blood pressure is predicted to increase by approximately 0.40912 units.

R output:

```
> summary(m1)

Call:
lm(formula = q1$Y ~ q1$X)

Residuals:
    Min       1Q   Median       3Q      Max
-17.051  -6.790  -1.960   8.654  12.085

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  71.13688    12.81381   5.552 1.04e-05 ***
q1$X          0.40912     0.06963   5.876 4.63e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.616 on 24 degrees of freedom
Multiple R-squared:  0.5899,    Adjusted R-squared:  0.5728
F-statistic: 34.52 on 1 and 24 DF,  p-value: 4.629e-06
```

c. Evaluate the residuals and check sum of residuals approximately zero

The residuals were calculated and rounded to two decimal places.

- **Rounded Residuals:**

-8.64, -6.46, 5.22, -6.55, -6.87, 3.27, 1.13, -17.05, -4.96, -7.10, -2.78,
-5.28, 9.31, 11.50, -8.19, -7.87, 6.68, 11.99, -9.10, 12.09, 11.22, -1.64,
0.68, -2.28, 10.31, 11.36

- **Sum of Rounded Residuals:**

The sum of the rounded residuals is approximately -0.01. This value is very close to zero, which is expected for ordinary least squares (OLS) regression models, as the sum of residuals is theoretically zero.

d. Find a 95% confidence interval on β_1 and test the hypothesis $H_0: \beta_1=0$

- **95% Confidence Interval for β_1 (Slope of Weight):**

The 95% confidence interval for β_1 is [0.2654099, 0.5528237]. This means we are 95% confident that the true population slope relating weight to SBP lies within this interval.

- **Hypothesis Test for $H_0: \beta_1=0$**

- **Null Hypothesis (H_0):** $\beta_1=0$ (There is no linear relationship between Weight and SBP).
- **Alternative Hypothesis (H_1):** $\beta_1 \neq 0$ (There is a linear relationship between

Weight and SBP).

○ **Test Statistics for q1\$X (Weight):**

- Estimate, $\hat{\beta}_1$: 0.4091168
- Standard Error: 0.06962885
- t value: 5.875679
- P-value: 4.629235×10^{-6}

- **Conclusion:** Since the p-value (4.629235×10^{-6}) is much smaller than the common significance level of $\alpha=0.05$, we reject the null hypothesis. This provides strong statistical evidence that there is a significant linear relationship between weight and systolic blood pressure.

e. Construct the analysis-of-variance table and test for significance of regression

The Analysis of Variance (ANOVA) table for the regression model is as follows:

Analysis of Variance Table

Response: q1\$Y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
q1\$X	1	2561.4	2561.45	34.516	4.629e-06 ***
Residuals	24	1781.3	74.22		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

● **Test for Significance of Regression:**

- **Null Hypothesis (H_0):** The regression model is not significant (i.e., all regression coefficients, excluding the intercept, are zero).
- **Alternative Hypothesis (H_1):** The regression model is significant (i.e., at least one regression coefficient is not zero).
- **F-statistic:** 34.52 on 1 and 24 degrees of freedom.
- **P-value:** 4.629×10^{-6}
- **Conclusion:** Since the p-value (4.629×10^{-6}) is extremely small (less than $\alpha=0.05$), we reject the null hypothesis. This indicates that the overall regression model is statistically significant, meaning that weight significantly explains the variation in systolic blood pressure.

f. Calculate R^2 and interpret

- **Multiple R-squared:** 0.5899091
- **Interpretation:** This means that approximately 58.99% of the total variation in systolic blood pressure (Y) can be explained by the linear relationship with weight (X). This suggests that weight is a moderately strong predictor of systolic blood

pressure.

Question 2: Multiple Linear Regression (Sealer Plate Temperature & Clearance vs. % Sealed Properly)

This section involves a multiple linear regression analysis to determine the effect of sealer plate temperature (X_2) and sealer plate clearance (X_1) on the percentage of items sealed properly (Y).

a. Obtain the solution $\hat{\beta} = (X'X)^{-1}X'Y$ using matrix manipulations

The estimated regression coefficients ($\hat{\beta}$) obtained through matrix manipulations are:

```
[,1]
[1,] -67.88435970 (Intercept)
[2,]  0.90608862 (Clearance, X1)
[3,] -0.06418911 (Temperature, X2)
```

- **Interpretation:**

- $\hat{\beta}_0 = -67.88436$: Estimated intercept.
- $\hat{\beta}_1 = 0.90609$: For a one-unit increase in sealer plate clearance, the percentage sealed properly is estimated to increase by 0.90609 units, holding temperature constant.
- $\hat{\beta}_2 = -0.06419$: For a one-unit increase in sealer plate temperature, the percentage sealed properly is estimated to decrease by 0.06419 units, holding clearance constant.

b. Find the variance-covariance matrix of $\hat{\beta}$

The variance-covariance matrix of the estimated regression coefficients ($\hat{\beta}$) is:

	(Intercept)	q2\$X1	q2\$X2
(Intercept)	1647.265746	-2.5277630715	-5.7495994936
q2\$X1	-2.527763	0.0152203720	0.0001528816
q2\$X2	-5.749599	0.0001528816	0.0268678743

- **Interpretation:**

- The diagonal elements represent the variances of the estimated coefficients (e.g., $\text{Var}(\hat{\beta}_0) = 1647.27$, $\text{Var}(\hat{\beta}_1) = 0.0152$, $\text{Var}(\hat{\beta}_2) = 0.0269$).
- The off-diagonal elements represent the covariances between the estimated

coefficients (e.g., $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -2.5278$).

c. Calculate the hat matrix H and I-H

- Hat Matrix (H):

The hat matrix (H) projects the observed response vector (Y) onto the column space of the design matrix (X) to obtain the fitted values ($\hat{Y} = HY$). It is a symmetric and idempotent matrix.

Output from R studio:

```
> H
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]
[1,] 0.31786633 0.20527414 0.226642257 -0.034364134 -0.026198343 0.042793362 0.114705760 -0.13807525
[2,] 0.20527414 0.36783804 0.091180712 0.118546272 -0.070388423 0.264621759 -0.017038764 -0.08531534
[3,] 0.22664226 0.09118071 0.185656409 -0.030592172 0.028790829 -0.009795544 0.126474156 -0.05644364
[4,] -0.03436413 0.11854627 -0.030592172 0.153095042 0.055439235 0.174130721 -0.010434292 0.12114783
[5,] -0.02619834 -0.07038842 0.028790829 0.055439235 0.124075757 -0.009382159 0.084527187 0.14534442
[6,] 0.04279336 0.26462176 -0.009795544 0.174130721 -0.009382159 0.265465361 -0.044289609 0.04426125
[7,] 0.11470576 -0.01703876 0.126474156 -0.010434292 0.084527187 -0.044289609 0.125593040 0.03869523
[8,] -0.13807525 -0.08531534 -0.056443643 0.121147835 0.145344417 0.044261248 0.038695234 0.22568355
[9,] 0.03173687 -0.18012806 0.105792979 -0.035992058 0.156451332 -0.148185187 0.164896660 0.12232380
[10,] -0.12147551 -0.04336823 -0.054912630 0.130814460 0.127512818 0.073854111 0.026339426 0.20747786
[11,] 0.08168505 0.17894468 0.045255471 0.106884618 0.016787288 0.160937082 0.015452146 0.03070738
[12,] 0.07080733 0.01407253 0.082687336 0.033414000 0.079217684 0.011698950 0.089785905 0.06437070
[13,] -0.07999107 0.03817642 -0.044572041 0.143593343 0.091550745 0.125801900 0.006687616 0.16566357
[14,] 0.01948818 0.10327399 0.016718478 0.110491654 0.053496548 0.127084134 0.022454443 0.08602893
[15,] 0.21999640 0.06507261 0.187649225 -0.039013894 0.039370238 -0.030446436 0.135911563 -0.04768139
[16,] 0.06910862 -0.05076234 0.099468177 0.002839369 0.103404848 -0.048549693 0.120239529 0.07581110

      [,9]      [,10]      [,11]      [,12]      [,13]      [,14]      [,15]      [,16]
[1,] 0.031736866 -0.12147551 0.081685055 0.07080733 -0.079991066 0.019488180 0.21999640 0.069108622
[2,] -0.180128055 -0.04336823 0.178944678 0.01407253 0.038176417 0.103273988 0.06507261 -0.050762340
[3,] 0.105792979 -0.05491263 0.045255471 0.08268734 -0.044572041 0.016718478 0.18764923 0.099468177
[4,] -0.035992058 0.13081446 0.106884618 0.03341400 0.143593343 0.110491654 -0.03901389 0.002839369
[5,] 0.156451332 0.12751282 0.016787288 0.07921768 0.091550745 0.053496548 0.03937024 0.103404848
[6,] -0.148185187 0.07385411 0.160937082 0.01169895 0.125801900 0.127084134 -0.03044644 -0.048549693
[7,] 0.164896660 0.02633943 0.015452146 0.08978590 0.006687616 0.022454443 0.13591156 0.120239529
[8,] 0.122323805 0.20747786 0.030707376 0.06437070 0.165663575 0.086028932 -0.04768139 0.075811101
[9,] 0.291549350 0.08767015 -0.045489648 0.11455126 0.024343801 0.002884265 0.12897756 0.178616915
[10,] 0.087670146 0.19508301 0.047268076 0.05722307 0.164417655 0.092622610 -0.04982599 0.059299114
[11,] -0.045489648 0.04726808 0.113500985 0.03824450 0.077738973 0.089656347 0.03425885 0.008168207
[12,] 0.114551263 0.05722307 0.038244496 0.07526015 0.044840391 0.045845737 0.08774094 0.090239519
[13,] 0.024343801 0.16441765 0.077738973 0.04484039 0.155589562 0.102407698 -0.04635899 0.030110421
[14,] 0.002884265 0.09262261 0.089656347 0.04584574 0.102407698 0.088376924 0.01140137 0.027768697
[15,] 0.128977564 -0.04982599 0.034258850 0.08774094 -0.046358991 0.011401365 0.19203846 0.110909482
[16,] 0.178616915 0.05929911 0.008168207 0.09023952 0.030110421 0.027768697 0.11090948 0.123328032
```

I-H Matrix:

The (I-H) matrix is used to obtain the residuals ($\hat{\epsilon} = (I-H)Y$). It is also symmetric and idempotent.

I-H =


```

> I_minus_H
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]
[1,]  0.68213367 -0.20527414 -0.226642257  0.034364134  0.026198343 -0.042793362 -0.114705760  0.13807525
[2,] -0.20527414  0.63216196 -0.091180712 -0.118546272  0.070388423 -0.264621759  0.017038764  0.08531534
[3,] -0.22664226 -0.09118071  0.814343591  0.030592172 -0.028790829  0.009795544 -0.126474156  0.05644364
[4,]  0.03436413 -0.11854627  0.030592172  0.846904958 -0.055439235 -0.174130721  0.010434292 -0.12114783
[5,]  0.02619834  0.07038842 -0.028790829 -0.055439235  0.875924243  0.009382159 -0.084527187 -0.14534442
[6,] -0.04279336 -0.26462176  0.009795544 -0.174130721  0.009382159  0.734534639  0.044289609 -0.04426125
[7,] -0.11470576  0.01703876 -0.126474156  0.010434292 -0.084527187  0.044289609  0.874406960 -0.03869523
[8,]  0.13807525  0.08531534  0.056443643 -0.121147835 -0.145344417 -0.044261248 -0.038695234  0.77431645
[9,] -0.03173687  0.18012806 -0.105792979  0.035992058 -0.156451332  0.148185187 -0.164896660 -0.12232380
[10,]  0.12147551  0.04336823  0.054912630 -0.130814460 -0.127512818 -0.073854111 -0.026339426 -0.20747786
[11,] -0.08168505 -0.17894468 -0.045255471 -0.106884618 -0.016787288 -0.160937082 -0.015452146 -0.03070738
[12,] -0.07080733 -0.01407253 -0.082687336 -0.033414000 -0.079217684 -0.011698950 -0.089785905 -0.06437070
[13,]  0.07999107 -0.03817642  0.044572041 -0.143593343 -0.091550745 -0.125801900 -0.006687616 -0.16566357
[14,] -0.01948818 -0.10327399 -0.016718478 -0.110491654 -0.053496548 -0.127084134 -0.022454443 -0.08602893
[15,] -0.21999640 -0.06507261 -0.187649225  0.039013894 -0.039370238  0.030446436 -0.135911563  0.04768139
[16,] -0.06910862  0.05076234 -0.099468177 -0.002839369 -0.103404848  0.048549693 -0.120239529 -0.07581110

      [,9]      [,10]      [,11]      [,12]      [,13]      [,14]      [,15]      [,16]
[1,] -0.031736866  0.12147551 -0.081685055 -0.07080733  0.079991066 -0.019488180 -0.21999640 -0.069108622
[2,]  0.180128055  0.04336823 -0.178944678 -0.01407253 -0.038176417 -0.103273988 -0.06507261  0.050762340
[3,] -0.105792979  0.05491263 -0.045255471 -0.08268734  0.044572041 -0.016718478 -0.18764923 -0.099468177
[4,]  0.035992058 -0.13081446 -0.106884618 -0.03341400 -0.143593343 -0.110491654  0.03901389 -0.002839369
[5,] -0.156451332 -0.12751282 -0.016787288 -0.07921768 -0.091550745 -0.053496548 -0.03937024 -0.103404848
[6,]  0.148185187 -0.07385411 -0.160937082 -0.01169895 -0.125801900 -0.127084134  0.03044644  0.048549693
[7,] -0.164896660 -0.02633943 -0.015452146 -0.08978590 -0.006687616 -0.022454443 -0.13591156 -0.120239529
[8,] -0.122323805 -0.20747786 -0.030707376 -0.06437070 -0.165663575 -0.086028932  0.04768139 -0.075811101
[9,]  0.708450650 -0.08767015  0.045489648 -0.11455126 -0.024343801 -0.002884265 -0.12897756 -0.178616915
[10,] -0.087670146  0.80491699 -0.047268076 -0.05722307 -0.164417655 -0.092622610  0.04982599 -0.059299114
[11,]  0.045489648 -0.04726808  0.886499015 -0.03824450 -0.077738973 -0.089656347 -0.03425885 -0.008168207
[12,] -0.114551263 -0.05722307 -0.038244496  0.92473985 -0.044840391 -0.045845737 -0.08774094 -0.090239519
[13,] -0.024343801 -0.16441765 -0.077738973 -0.04484039  0.844410438 -0.102407698  0.04635899 -0.030110421
[14,] -0.002884265 -0.09262261 -0.089656347 -0.04584574 -0.102407698  0.911623076 -0.01140137 -0.027768697
[15,] -0.128977564  0.04982599 -0.034258850 -0.08774094  0.046358991 -0.011401365  0.80796154 -0.110909482
[16,] -0.178616915 -0.05929911 -0.008168207 -0.09023952 -0.030110421 -0.027768697 -0.11090948  0.876671968

```

d. Calculate R^2 and adjusted R^2

- **Multiple R-squared:** 0.8063764
- **Adjusted R-squared:** 0.7765881
- **Interpretation:**
 - The R^2 value of 0.8064 indicates that approximately 80.64% of the variability in the percentage sealed properly (Y) can be explained by the linear model including sealer plate clearance (X_1) and sealer plate temperature (X_2). This suggests a strong fit.
 - The Adjusted R^2 of 0.7766 accounts for the number of predictors in the model and is a more conservative estimate of the model's explanatory power, especially useful when comparing models with different numbers of predictors.

Question 3: Hypothesis Testing with a Given Model

This section tests specific linear hypotheses about the regression coefficients (β) for the model $E(Y)=X\beta$. The model is $Y \sim X[2]+X[3]$ with $Y=(1,4,8,9,3,8,9)'$ and X and C matrices provided.

We test the hypothesis $H_0: C\beta=0$, which translates to three individual hypotheses:

Test 1: $H_0: \beta_1=0$ (equivalent to $C_1\beta=0$ where $C_1=[0,1,0]$)

- **Linear Hypothesis Test Results:**

Linear hypothesis test:

$$X[, 2] = 0$$

Model 1: restricted model

Model 2: $Y \sim X[, 2] + X[, 3]$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	5	12.1481				
2	4	8.1481	1	4	1.9636	0.2337

- **Conclusion:** With an F-statistic of 1.9636 and a p-value of 0.2337, which is greater than $\alpha=0.05$, we fail to reject the null hypothesis. There is no significant evidence to conclude that β_1 is different from zero.

Test 2: $H_0: \beta_1 - \beta_2 = 0$ (equivalent to $C_2\beta=0$ where $C_2=[0,1,-1]$)

- **Linear Hypothesis Test Results:**

Linear hypothesis test:

$$X[, 2] - X[, 3] = 0$$

Model 1: restricted model

Model 2: $Y \sim X[, 2] + X[, 3]$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	5	14.8293				
2	4	8.1481	1	6.6811	3.2798	0.1444

- **Conclusion:** With an F-statistic of 3.2798 and a p-value of 0.1444, which is greater than $\alpha=0.05$, we fail to reject the null hypothesis. There is no significant evidence to conclude that β_1 is different from β_2 .

Test 3: $H_0: 2\beta_1 - 3\beta_2 = 0$ (equivalent to $C_3\beta=0$ where $C_3=[0,2,-3]$)

- **Linear Hypothesis Test Results:**

Linear hypothesis test:

$$2 X[, 2] - 3 X[, 3] = 0$$

Model 1: restricted model

Model 2: $Y \sim X[, 2] + X[, 3]$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	5	23.5556				
2	4	8.1481	1	15.407	7.5636	0.05136 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- **Conclusion:** With an F-statistic of 7.5636 and a p-value of 0.05136, which is very close to $\alpha=0.05$ (and slightly above it), we would typically fail to reject the null hypothesis at the 0.05 significance level. However, the . significance code indicates marginal significance at the 0.1 level. This suggests weak evidence against the null hypothesis that $2\beta_1=3\beta_2$.

Question 4: Second Order Polynomial Model

This section fits and evaluates a second-order polynomial model for the given data.

a. Fit a second order polynomial model $Y=\beta_0+\beta_1X+\beta_2X^2$ to these data

A second-order polynomial model was fitted to the data.

- Estimated Regression Equation:

$$\hat{Y} = -4.4595 + 1.3837X + 1.4670X^2$$

- **Summary of the Model:**

Call:

lm(formula = q4\$Y ~ q4\$X + I((q4\$X)^2))

Residuals:

Min	1Q	Median	3Q	Max
-4.5729	0.1344	0.4641	0.7315	0.8495

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-4.4595	14.6343	-0.305	0.7675
q4\$X	1.3837	5.4971	0.252	0.8069
I((q4\$X)^2)	1.4670	0.4936	2.972	0.0156 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.657 on 9 degrees of freedom

Multiple R-squared: 0.9957, Adjusted R-squared: 0.9948

F-statistic: 1045 on 2 and 9 DF, p-value: 2.213e-11

- **Interpretation of Coefficients:**
 - The coefficient, β_2 for $I((q4$X)^2)$ (the quadratic term) is statistically significant at the 0.05 level (p-value = 0.0156), suggesting that the quadratic term significantly contributes to the model.

- The intercept and linear term $q4\$X$ are not individually significant, but their significance should be considered in the context of the overall model and the presence of the quadratic term.

b. Test the lack of fit and comment on the adequacy of the second-order model

A lack-of-fit test was performed by comparing the second-order polynomial model (Model 1) to a full model (Model 2) that treats X as a categorical variable to capture all possible variations.

- **Analysis of Variance Table (Lack of Fit Test):**

Analysis of Variance Table

Model 1: $q4\$Y \sim q4\$X + I((q4\$X)^2)$

Model 2: $q4\$Y \sim x_factor$

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	9	24.7202				
2	4	0.3995	5	24.321	48.7	0.001124 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- **Comment on Adequacy:**

- **Null Hypothesis (H_0):** The second-order model is adequate (no lack of fit).
- **Alternative Hypothesis (H_1):** The second-order model is not adequate (there is lack of fit).
- **F-statistic:** 48.7 on 5 and 4 degrees of freedom.
- **P-value:** 0.001124
- **Conclusion:** Since the p-value (0.001124) is less than $\alpha=0.05$, we reject the null hypothesis. This indicates that there is a significant lack of fit for the second-order model. In other words, the second-order polynomial model does not adequately capture the relationship between X and Y , and a more complex model (or a different functional form) might be needed.

c. Fit a second-order model to these data using orthogonal polynomials

A second-order polynomial model was fitted using orthogonal polynomials.

- **Summary of the Orthogonal Polynomial Model:**

Call:

`lm(formula = q4$Y ~ poly(q4$X, degree = 2))`

Residuals:

Min	1Q	Median	3Q	Max
-4.5729	0.1344	0.4641	0.7315	0.8495

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	54.4650	0.4784	113.842	1.58e-15 ***
poly(q4\$X, degree = 2)1	75.6062	1.6573	45.620	5.85e-12 ***
poly(q4\$X, degree = 2)2	4.9257	1.6573	2.972	0.0156 * ---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Residual standard error: 1.657 on 9 degrees of freedom

Multiple R-squared: 0.9957, Adjusted R-squared: 0.9948

F-statistic: 1045 on 2 and 9 DF, p-value: 2.213e-11

- Comparison:

The model fitted using orthogonal polynomials yields the same R-squared, adjusted R-squared, residual standard error, F-statistic, and p-value as the non-orthogonal polynomial model. The coefficients, however, are different because they represent the coefficients for the orthogonal polynomial terms, not the raw X and X² terms. Orthogonal polynomials are useful for reducing multicollinearity among polynomial terms and for testing the significance of each polynomial degree independently. The significance of the second-degree term (poly(q4\$X, degree = 2)2) is still evident (p-value = 0.0156).

Conclusion

This report detailed the step-by-step regression analysis for the given exam questions using R. We performed simple linear regression, multiple linear regression using matrix operations, various hypothesis tests, and polynomial regression. The analysis provided insights into the relationships between variables, model fit, and statistical significance of the predictors. The lack-of-fit test for the second-order polynomial model indicated that a more complex model might be required to adequately describe the data.