

## Part 1: Advertising and Sales

Store	TV Ad (\$k)	Radio Ad (\$k)	Social Media Ad (\$k)	Discounts (%)	Sales (\$k)
1	230	37	22	5	650
2	44	39	10	7	270
3	17	45	12	10	180
4	151	41	24	6	500
5	180	10	15	8	525
6	8	36	9	12	130
7	57	35	10	9	200
8	120	50	20	5	455
9	200	33	18	4	590
10	90	45	25	6	365

### Task 1.1

1. Fit a full multiple regression model of sales with all other variables as predictors.
2. Use stepwise selection (forward, backward, or both) to select the best subset of variables.
3. Write the final regression equation.
4. Interpret the coefficients of the selected variables.
5. Answer the following:
  - Which variables significantly affect Sales?
  - How does variable selection improve model parsimony?
  - Compare  $R^2$  and Adjusted  $R^2$  before and after selection.

### Task 1.2:

1. Check linearity, homoscedasticity, normality of residuals, and independence.
2. Plot residuals vs fitted values, Q-Q plot, and leverage plots.
3. Identify any outliers or influential points.
4. Answer the following:
  - Are the assumptions of multiple regression satisfied?
  - What remedial measures can be applied if assumptions are violated?

## Part 2: Monthly Sales and Advertising

Month	TV Ad (\$k)	Sales (\$k)
1	230	650
2	44	265
3	17	175
4	151	490
5	180	525
6	8	130
7	57	200
8	120	445
9	200	595
10	90	360
11	100	420
12	150	515

### Task 2.1:

1. Fit a time series regression model: Sales as a function of TV\_Ad and time trend.
2. Plot Sales vs time with regression line.
3. Compute Durbin-Watson statistic for autocorrelation of residuals.
4. Answer the following:
  - Is there evidence of autocorrelation in the residuals?
  - What does the Durbin-Watson value indicate about model validity?

### Task 2.2:

1. If autocorrelation exists, apply the Cochrane-Orcutt method to correct it.
2. Refit the regression model and compare coefficients,  $R^2$ , and residual plots before and after correction.
3. Answer the following:
  - How did the Cochrane-Orcutt method affect your model?
  - Explain the importance of correcting autocorrelation.

## Part 3: Sales with Heteroscedasticity

Store	Advertising (\$k)	Sales (\$k)
1	230	650
2	44	270
3	17	185
4	151	500
5	180	520
6	8	125
7	57	195
8	120	450
9	200	600
10	90	355

### Task 3.1:

1. Fit a simple regression: Sales  $\sim$  Advertising.
2. Plot residuals vs fitted values.
3. Conduct a test for heteroscedasticity.
4. Answer the following:
  - Is heteroscedasticity present?
  - How does it affect standard errors and hypothesis tests?

### Task 3.2:

1. Apply WLS regression to correct for heteroscedasticity.
2. Compare coefficients, standard errors, and model fit before and after WLS.
3. Answer the following:
  - How does WLS improve the model?
  - Why is weighting necessary in heteroscedastic data?

### Task 3.3

1. Using the final time series model (with or without Cochrane-Orcutt correction), forecast Sales for the next 3 months.
2. Compute 95% prediction intervals.
3. Plot historical Sales, forecasted values, and prediction intervals.
4. Answer the following:
  - How reliable are your forecasts?
  - Discuss practical decisions management can make based on the forecast and prediction intervals.

## Rubric

Criteria	Prestructural (1)	Unistructural (2)	Multistructural (3)	Relational (4)	Extended Abstract (5)
<b>1. Model Building &amp; Variable Selection</b>	Unable to fit regression model or misuses variables.	Fits a model with all predictors but does not attempt selection.	Correctly applies stepwise selection; reports regression equation.	Interprets model selection, compares $R^2$ vs Adjusted $R^2$ , explains parsimony.	Critically evaluates alternative models; discusses trade-offs between complexity and generalizability.
<b>2. Interpretation of Coefficients &amp; Significance</b>	No interpretation or incorrect conclusions.	Identifies at least one significant predictor but misinterprets its meaning.	Interprets coefficients for selected predictors with partial accuracy.	Provides clear, correct interpretations of all coefficients, significance tests, and variable effects.	Discusses effect sizes, practical implications, and sensitivity to variable inclusion/exclusion.
<b>3. Assumptions &amp; Diagnostics</b>	No diagnostic checks.	Mentions one assumption without evidence.	Performs standard diagnostics but interprets superficially.	Thoroughly evaluates diagnostics with residual, Q-Q, leverage plots, and explains violations.	Suggests and applies remedies, discussing impact on inference.
<b>4. Time Series Regression &amp; Autocorrelation</b>	Does not attempt time series model.	Fits regression with trend but omits residual autocorrelation.	Computes Durbin-Watson and reports statistic without interpretation.	Interprets DW value, applies Cochrane-Orcutt if needed, compares pre/post models.	Critically evaluates time series validity, alternative corrections, and management implications of corrected model.
<b>5. Handling Heteroscedasticity</b>	Ignores heteroscedasticity.	Identifies heteroscedasticity but does not act on it.	Conducts heteroscedasticity test and notes its presence.	Applies WLS, compares coefficients, SEs, and fit before/after correction.	Justifies weighting choice, reflects on efficiency gains, and explains broader consequences for inference.
<b>6. Forecasting &amp; Decision Implications</b>	No forecast attempted.	Produces forecast values without intervals.	Produces forecasts with 95% prediction intervals but	Plots forecasts with intervals,	Provides nuanced discussion of uncertainty, risk,

			limited explanation.	evaluates reliability, and links to model adequacy.	and practical decisions managers can make from forecast outcomes.
--	--	--	----------------------	---	---