Second Semester Examination
2017/2018 Academic Session

May / June 2018

MSG469 - Regression Analysis
(Analisis Regressi)

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of **NINETEEN (19)** pages of printed material before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi **SEMBILAN BELAS (19)** muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]*

Instructions: Answer **all four (4)** questions.

*[**Arahan:** Jawab **semua empat (4)** soalan.]*

In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah diguna pakai].

Question 1

(a) (i) Define the coefficient of determination, R^2 for a *k*-variable regression model. Explain the use of R^2 in regression model building.

(ii) Prove that for a *k*-variable regression model:

$$R^2 = \frac{(k-1)F}{(k-1)F + (n-k)}.$$

[20 marks]

(b) A collector of antique grandfather clocks sold at an auction believes that the price received for the clocks, *y* (dollars) depends on both the age of the clocks, x_1 (years) and the number of bidders at the auction, x_2. Thus, he hypothesized the following regression model relating the three variables:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \varepsilon .$$

A sample of 8 auction prices of the grandfather clocks, along with their ages and the number of bidders is given below.

Age	No. Bidders	Price
137	9	1297
113	9	946
137	15	1713
117	11	1024
137	8	1147
153	6	1092
117	13	1152
126	10	1336

(i) Find the matrix $\mathbf{X}'\mathbf{X}$, its inverse $(\mathbf{X}'\mathbf{X})^{-1}$ and the matrix $\mathbf{X}'\mathbf{y}$.

(ii) Estimate the parameters of the multiple linear regression model relating auction price to age and number of bidders.

(iii) Construct an ANOVA table and test for the overall significance of the regression model.

(iv) Calculate R^2 and $\bar{R}^2$.

(v) The estimated registration model is to be used to predict the mean auction price for two antique clocks of two different ages. The first clock of age 100 was auctioned when there were 12 bidders while the second clock is of age 155 and was auctioned when there were 18 bidders. Find the 95% confidence interval on the mean auction price of the two clocks. Which confidence interval gives better reliability in predicting the auction price? Explain your answers.

[80 marks]

...3/-

Soalan 1

(a) *(i)* *Takrifkan pekali penentuan R^2 bagi model regressi k-pembolehubah. Terangkan penggunaan R^2 dalam membangunkan model regressi.*

(ii) *Buktikan bahawa untuk model gressi dengan k-pembolehubah:*

$$R^2 = \frac{(k-1)F}{(k-1)F + (n-k)}$$

[20 markah]

(b) *Seorang pengumpul jam antik yang dijual di lelongan percaya bahawa harga yang diterima untuk jam, y (dolar) bergantung pada umur jam, x_1 (tahun) dan bilangan pembida di lelongan, x_2. Oleh itu, beliau membuat hipotesis model regresi berikut yang mengaitkan tiga pembolehubah tersebut:*

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \varepsilon$$

Suatu sampel lapan harga lelongan jam antik, bersama dengan umur jam dan bilangan pembida diberikan di bawah.

Umur	*Bil. Pembida*	*Harga*
137	9	1297
113	9	946
137	15	1713
117	11	1024
137	8	1147
153	6	1092
117	13	1152
126	10	1336

(i) *Dapatkan matriks* $\mathbf{X}'\mathbf{X}$, songsangannya $(\mathbf{X}'\mathbf{X})^{-1}$ *dan matriks* $\mathbf{X}'\mathbf{y}$.

(ii) *Anggarkan parameter bagi model regresi linear berganda yang mengaitkan harga lelongan dengan umur dan bilangan pembida.*

(iii) *Bina suatu jadual ANOVA dan uji kesignifikanan keseluruhan model regresi.*

(iv) *Hitung R^2 dan $\bar{R}^2$*

(v) *Model regressi yang dianggarkan akan digunakan untuk meramalkan harga lelongan purata untuk dua jam antik yang berbeza umur. Jam pertama dengan umur 100 dilelong apabila terdapat 12 pembida manakala jam kedua berumur 155 dilelong apabila terdapat 18 pembida. Cari selang keyakinan 95% bagi harga lelong purata dua jam tersebut. Selang keyakinan yang manakah memberikan kebolehpercayaan yang lebih baik dalam meramalkan harga lelong? Terangkan jawapan anda.*

[80 markah]

Question 2

(a) (i) State six major assumptions on a linear regression model.

(ii) Differentiate between leverage and influence points in regression analysis.

(iii) Discuss the difference among three measures of influence point; Cook's distance D, DFITS and DFBETAS.

[30 marks]

(b) Energy consumption (in Gigawatt-hours) of nineteen 4- and 5-star hotel buildings in Hainan Province, China were collected together with two of its attributes, area (in 1000 square meters) and occupancy rate (in 100 percent)

Consumption	Area	Occupancy
1.95	43.0	32.6
13.83	112.2	49.0
5.56	45.0	49.0
4.00	28.5	52.3
4.67	32.9	50.0
8.92	59.4	57.7
8.19	50.8	49.0
11.74	68.0	64.2
14.84	78.9	63.5
5.37	28.5	76.5
13.52	70.0	67.6
10.57	50.0	61.5

A multiple regression model of the following form is to be estimated:

$$\text{Cons} = \beta_0 + \beta_1 \text{Area} + \beta_2 \text{Occu} + \varepsilon$$

A multiple regression model was fitted to the above data with the following results:

$$\hat{\beta}_0 = -8.913, \qquad \hat{\beta}_1 = 0.147, \qquad \hat{\beta}_2 = 0.167,$$

$$\hat{\sigma}_\varepsilon^2 = MSE = 2.3331, \quad (\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 2.656350 & -0.008017 & -0.037937 \\ -0.008017 & 0.000156 & 0.000011 \\ -0.037937 & 0.000011 & 0.000688 \end{pmatrix}$$

(i) Interpret the coefficients of the model.

(ii) Calculate all the missing elements, A1 through to S2 in Table 1a, Table 1b and Table 1c in **Appendix 1**.

(iii) Investigate the presence of leverage and influence observations.
Is there any observation that degrades the overall model performance?

[70 marks]

Soalan 2

(a) *(i)* *Nyatakan enam andaian utama terhadap model regressi linear.*

(ii) *Bezakan antara titik leverej dan titik berpengaruh dalam analisis regresi.*

(iii) *Bincangkan perbezaan antara tiga ukuran titik berpengaruh; jarak Cook D, DFITS dan DFBETAS.*

[30 markah]

(b) *Penggunaan tenaga (dalam Gigawatt-jam) bagi sembilan belas bangunan hotel 4- dan 5-bintang di Wilayah Hainan, China dikumpulkan bersama dua sifatnya, keluasan (dalam 1000 meter persegi) dan kadar penghunian (dalam 100 peratus)*

Penggunaan	Keluasan	Penghunian
1.95	43.0	32.6
13.83	112.2	49.0
5.56	45.0	49.0
4.00	28.5	52.3
4.67	32.9	50.0
8.92	59.4	57.7
8.19	50.8	49.0
11.74	68.0	64.2
14.84	78.9	63.5
5.37	28.5	76.5
13.52	70.0	67.6
10.57	50.0	61.5

Model regresi berganda dalam bentuk seperti berikut akan dianggarkan:

$$\text{Gunaan} = \beta_0 + \beta_1 \text{Luas} + \beta_2 \text{Hunian} + \varepsilon$$

Model regresi berganda telah disuai bagi data di atas dengan keputusan berikut:

$$\hat{\beta}_0 = -8.913, \qquad \hat{\beta}_1 = 0.147, \qquad \hat{\beta}_2 = 0.167,$$

$$\hat{\sigma}_\varepsilon^2 = MSE = 2.3331 \quad \left(\mathbf{X}'\mathbf{X}\right)^{-1} = \begin{pmatrix} 2.656350 & -0.008017 & -0.037937 \\ -0.008017 & 0.000156 & 0.000011 \\ -0.037937 & 0.000011 & 0.000688 \end{pmatrix}$$

(i) *Tafsirkan pekali-pekali model.*

(ii) *Hitung semua elemen yang tidak lengkap, A1 hingga S2 dalam Jadual 1a, Jadual 1b dan Jadual 1c di* ***Lampiran 1****.*

(iii) *Siasat kehadiran cerapan leverej dan berpengaruh. Adakah terdapat cerapan yang merendahkan prestasi keseluruhan model?*

[70 markah]

Question 3

(a) (i) Consider a regression model: $y_i = \alpha_0 + \alpha_1 x_i + \varepsilon_i$. The variance of the estimated residuals is found to be proportional to x_i . Suggest an appropriate transformation and explain the procedure to deal with heteroscedasticity of such nature.

(ii) Explain the scenario for the application of a log-lin functional form in regression analysis.

(iii) Explain the use of Durbin-Watson test in detecting the presence of autocorrelation in the errors of a time series regression model.

[40 marks]

(b) An analyst is interested to investigate the factors affecting executive salary. Eight potential independent variables were considered from a sample of 100 executives. Variable selection techniques were conducted to decide which of the eight variables should be included in the final model.

(i) **Appendix 2a** show partial results (ANOVA table only) of the forward selection technique, conducted with $\alpha_{\text{IN}} = 0.10$ as well as the ANOVA tables for 4-variable models involving experience and gender. Explain why the number of employees supervised is the variable selected to enter at the third stage. Show all your calculations.

(ii) The ANOVA output for a 7-variable model is shown in **Appendix 2b**. By showing all your calculation, explain why age, profits and sales cannot be included in the model.

(iii) **Appendix 2c** shows the ANOVA tables of backward elimination technique, conducted with $\alpha_{\text{OUT}} = 0.10$ as well as the ANOVA tables for 6-variable models. By showing all your calculations, explain why sales is the variable to be eliminated at the fourth stage.

[60 marks]

Soalan 3

(a) *(i)* *Pertimbangkan suatu model regresi:* $y_i = \alpha_0 + \alpha_1 x_i + \varepsilon_i$ *. Varians bagi residuals anggaran didapati berkadar langsung dengan.* x_i *. Cadangkan suatu transformasi yang sesuai dan terangkan prosedur untuk menangani heteroskedastisiti yang bersifat sebegitu.*

(ii) *Jelaskan senario penggunaan bentuk fungsi log-lin dalam analisis regresi.*

(iii) *Jelaskan penggunaan ujian Durbin-Watson dalam mengesan kehadiran autokorelasi dalam ralat bagi model regresi siri masa.*

[40 markah]

(b) *Seorang penganalisa berminat untuk menyiasat faktor-faktor yang mempengaruhi gaji eksekutif. Lapan pembolehubah tak bersandar yang berpotensi telah dipertimbangkan daripada suatu sampel 100 eksekutif. Teknik pemilihan pembolehubah telah dijalankan untuk menentukan lapan pembolehubah yang perlu dimasukkan dalam model akhir.*

(i) ***Lampiran 2a*** *menunjukkan keputusan separa (jadual ANOVA sahaja) bagi teknik pemilihan ke hadapan, yang dijalankan dengan* $\alpha_{\text{IN}} = 0.10$ *serta jadual ANOVA untuk beberapa model 4-pembolehubah yang melibatkan pengalaman dan jantina. Jelaskan kenapa bilangan pekerja yang diselia adalah pembolehubah yang dipilih untuk dimasukkan pada peringkat ketiga. Tunjukkan semua pengiraan anda.*

(ii) *Output ANOVA untuk model dengan 7-pembolehubah ditunjukkan dalam* ***Lampiran 2b****. Dengan menunjukkan semua pengiraan anda, jelaskan mengapa usia, keuntungan dan jualan tidak boleh dimasukkan dalam model.*

(iii) ***Lampiran 2c*** *menunjukkan jadual ANOVA bagi teknik penghapusan ke belakang, yang dijalankan dengan* $\alpha_{\text{OUT}} = 0.10$ *, serta jadual ANOVA untuk model dengan 6-pembolehubah. Dengan menunjukkan semua pengiraan anda, jelaskan kenapa jualan adalah pembolehubah yang telah dihapuskan pada peringkat keempat.*

[60 markah]

Question 4

(a) Regression analysis was conducted to develop a model for relating mileage with physical and mechanical characteristics of automobile. The variables involved in the study are as follows:

y : miles/gallon	x_1 : displacement	x_2 : horsepower
x_3 : torque	x_4 : compression ratio	x_5 : rear axle ratio
x_6 : carburetor	x_7 : no. of transmission speed	
x_8 : overall length	x_9 : width	x_{10} : weight

(i) Correlation analysis and regression model involving all variables were conducted and the results are shown in **Appendix 3a**. Do the results provide any indication of multicollinearity?

(ii) A summary of analysis from auxiliary regression is given in **Appendix 3b**. Calculate the corresponding variance inflation factor (VIF). Which variables can be said to be the source of multicollinearity?

(iii) In an attempt to deal with multicollinearity, a few regression models were estimated and are shown in **Appendix 3c**. Explain the reason for the sequential steps taken by the analyst. Which model is free from multicollinearity?

(iv) **Appendix 3d** shows the results of fitting all possible regression model involving all ten independent variables. Explain why these results can be very misleading?

(v) **Appendix 3e** shows the output from all possible regression analysis after multicollinearity is eliminated.
Suggest two further simplified models that maybe selected by stepwise regression techniques to describe the factors affecting automobile mileage performance.

[70 marks]

(b) An equal-rights group claims that females are less likely to be hired than males with the same background, experience (x_1) and education level (x_2). Data were collected on 28 applicants of a particular firm and binary logistic regression was used for investigation. The response variable status, y is coded 1 if the applicant was "hired" while the variable gender (x_3) is coded 1 if the applicant was a male. The estimated logistic coefficients and the corresponding classification table are shown in **Appendix 4**.

(i) Interpret the coefficients and its corresponding odds ratios. Is there sufficient evidence for gender discrimination in the hiring practices of the selected firm?

(ii) Evaluate the effectiveness of fitted logistic regression model.

[30 marks]

Soalan 4

(a) *Analisis regresi telah dijalankan untuk membangunkan suatu model mengaitkan perbatuan dengan ciri fizikal dan mekanik kereta. Pembolehubah yang terlibat dalam kajian ini adalah seperti berikut:*

y *: batu/galon*	x_1 *: anjakan*	x_2 *: kuasa kuda*
x_3 *: tork*	x_4 *: nisbah mampatan*	x_5 *: nisbah gandar belakang*
x_6 *: karburetor*	x_7 *: bil. Kelajuan transmisi*	
x_8 *: panjang keseluruhan*	x_9 *: lebar*	x_{10} *: berat*

(i) *Analisis korelasi dan model regresi yang melibatkan semua pembolehubah telah dijalankan dan hasilnya ditunjukkan dalam **Lampiran 3a**. Adakah keputusan memberikan sebarang petunjuk tentang kehadiran multikolinearitI?*

(ii) *Ringkasan analisis daripada regresi tambahan diberikan dalam **Lampiran 3b**. Kira faktor inflasi variasi (VIF) yang sepadan. Pembolehubah manakah yang boleh dikatakan sebagai sumber multikolinearitI?*

(iii) *Dalam usaha untuk menangani multikolinearitI, beberapa model regresi telah dianggarkan dan ditunjukkan dalam **Lampiran 3c**. Terangkan sebab bagi langkah-langkah berurutan yang diambil oleh penganalisa. Model yang manakah bebas daripada multikolinearitI?*

(iv) ***Lampiran 3d** menunjukkan keputusan daripada penyuaian semua model regresi berkemingkinan melibatkan kesemua sepuluh pembolehubah bersandar. Terangkan mengapa keputusan ini boleh amat mengelirukan?*

(v) ***Lampiran 3e** menunjukkan output daripada analisis semua regresi yang mungkin selepas multikolinearitI telah dihapuskan*
Cadangkan dua model ringkas yang mungkin dipilih oleh teknik regresi langkah demi langkah untuk menerangkan faktor-faktor yang mempengaruhi prestasi perbatuan kereta.

[70 markah]

(b) *Satu kumpulan hak saksama mendakwa bahawa kemungkinan wanita diambil bekerja adalah lebih rendah berbanding lelaki dengan latar belakang, pengalaman dan tahap pendidikan yang sama. Data telah dikumpul daripada 28 pemohon suatu firma tertentu dan regresi logistik binari digunakan untuk penyiasatan. Pembolehubah bersandar status, y dikodkan 1 jika pemohon "diambil" sementara pembolehubah jantina dikodkan 1 jika pemohon adalah lelaki. Koefisien logistik yang dianggarkan dan jadual klasifikasi yang sepadan ditunjukkan di **Lampiran 4**.*

(i) *Tafsirkan pekali-pekali dan nisbah odds yang sepadan. Adakah terdapat bukti yang cukup wujudnya diskriminasi jantina dalam amalan pengambilan pekerja bagi firma yang dipilih?*

(ii) *Nilaikan keberkesanan model regresi logistik yang disuaikan.*

[30 markah]

APPENDIX 1/*LAMPIRAN 1*

Jadual 1a

Cons, y_i	Fitted, $\hat{y}_i$	Deleted Var, $S^2_{(i)}$	Hat h_{ii}	R matrix		
				$R_{1,i}$	$R_{2,i}$	$R_{3,i}$
1.95	2.8316	2.4705	0.4804	**D1**	**E1**	-0.0160
13.83	15.7145	**B1**	**C1**	-0.1020	-0.7839	-0.0055
5.56	**A1**	2.6182	0.1335	0.4367	-0.7944	**F1**
4.00	3.9933	**B2**	0.2051	**D2**	**E2**	-0.0023
4.67	4.2546	2.6108	0.1858	0.4958	-0.7962	-0.0039
8.92	**A2**	2.6083	0.0873	-0.0088	-0.7922	**F2**
8.19	6.7122	2.4734	0.1206	0.3902	-0.7935	-0.0048
11.74	11.7707	2.6247	0.1504	-0.3243	-0.7909	0.0054
14.84	13.2520	2.4127	0.2018	-0.3851	-0.7892	0.0048
5.37	8.0320	1.1229	0.4972	-0.4743	-0.7972	0.0144
13.52	12.6313	2.5581	**C2**	-0.4693	-0.7907	0.0078
10.57	8.6810	2.3838	0.1091	-0.0776	-0.7937	0.0038

Jadual 1b

Cons, y_i	Resi, e_i	Std. Resi d_i	Hat h_{ii}	Student Resi, r_i	Deleted Resi, $e_{(i)}$	Deleted Student, t_i
1.95	**G1**	-0.5772	0.4804	-0.8007	-1.6966	-0.7833
13.83	-1.8845	-1.2337	**C1**	-2.0163	**K1**	**L1**
5.56	**G2**	-0.1976	0.1335	-0.2122	-0.3483	-0.2006
4.00	0.0067	0.0044	0.2051	0.0049	0.0084	0.0046
4.67	0.4154	0.2719	0.1858	**J1**	0.5102	0.2856
8.92	-0.5050	-0.3306	0.0873	-0.3461	-0.5533	-0.3285
8.19	1.4778	**H1**	0.1206	1.0317	1.6804	1.0359
11.74	-0.0307	**H2**	0.1504	-0.0218	-0.0361	-0.0205
14.84	1.5880	1.0396	0.2018	1.1637	1.9895	1.1903
5.37	-2.6620	-1.7427	0.4972	**J2**	-5.2945	-4.0410
13.52	0.8887	0.5818	**C2**	0.6518	1.1153	**L2**
10.57	1.8890	1.2367	0.1091	1.3103	**K2**	1.3733

...11/-

Jadual 1c

y_i	D_i	$DFIT_i$	$DBETA_{1,i}$	$DBETA_{2,i}$	$DBETA_{3,i}$	$COVRAT_i$
1.95	0.1975	-0.7531	-0.7166	**Q1**	0.6630	**S1**
13.83	2.2641	-3.3184	0.2626	-2.9899	**R1**	0.6268
5.56	**M1**	-0.0788	-0.0577	0.0271	0.0390	1.6187
4.00	0.0000	**N1**	**P1**	-0.0017	-0.0005	**S2**
4.67	0.0069	0.1364	0.0963	-0.0879	-0.0473	1.6963
8.92	0.0038	-0.1016	0.0019	-0.0158	-0.0141	1.4985
8.19	0.0487	0.3836	0.2645	-0.0590	-0.2027	1.1099
11.74	0.0000	**N2**	0.0044	-0.0033	-0.0046	1.6756
14.84	0.1141	0.5985	-0.3148	0.3783	0.2460	1.0937
5.37	**M2**	-4.0185	**P2**	**Q2**	-3.1199	0.1007
13.52	0.0361	0.3179	-0.2031	0.1193	0.2088	1.5456
10.57	0.0701	0.4807	-0.0693	-0.1089	**R2**	0.8470

APPENDIX 2a / *LAMPIRAN 2a*

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.136	1	4.136	159.204	.000[b]
	Residual	2.546	98	.026		
	Total	6.683	99			
2	Regression	5.007	2	2.503	144.887	.000[c]
	Residual	1.676	97	.017		
	Total	6.683	99			
3	Regression	5.607	3	1.869	166.873	.000[d]
	Residual	1.075	96	.011		
	Total	6.683	99			
4	Regression	6.064	4	1.516	232.936	.000[e]
	Residual	.618	95	.007		
	Total	6.683	99			
5	Regression	6.152	5	1.230	218.061	.000[f]
	Residual	.530	94	.006		
	Total	6.683	99			

b. Predictors: (Constant), experience

c. Predictors: (Constant), experience, gender

d. Predictors: (Constant), experience, gender, employee

e. Predictors: (Constant), experience, gender, employee, education

f. Predictors: (Constant), experience, gender, employee, education, assets

...12/-

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.288	3	1.763	121.354	.000[b]
	Residual	1.394	96	.015		
	Total	6.683	99			

b. Predictors: (Constant), education, experience, gender

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.607	3	1.869	166.873	.000[b]
	Residual	1.075	96	.011		
	Total	6.683	99			

b. Predictors: (Constant), employee, experience, gender

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.199	3	1.733	112.099	.000[b]
	Residual	1.484	96	.015		
	Total	6.683	99			

b. Predictors: (Constant), assets, experience, gender

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.007	3	1.669	95.596	.000[b]
	Residual	1.676	96	.017		
	Total	6.683	99			

b. Predictors: (Constant), age, gender, experience

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.014	3	1.671	96.157	.000[b]
	Residual	1.669	96	.017		
	Total	6.683	99			

b. Predictors: (Constant), profit, experience, gender

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.008	3	1.669	95.700	.000[b]
	Residual	1.675	96	.017		
	Total	6.683	99			

b. Predictors: (Constant), sales, experience, gender

...13/-

APPENDIX 2b/*LAMPIRAN 2b*

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.152	6	1.025	179.788	.000[b]
	Residual	.530	93	.006		
	Total	6.683	99			

b. Predictors: (Constant), age, assets, education, gender, employee, experience

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.153	6	1.025	179.875	.000[b]
	Residual	.530	93	.006		
	Total	6.683	99			

b. Predictors: (Constant), profit, employee, experience, education, assets, gender

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.153	6	1.025	179.916	.000[b]
	Residual	.530	93	.006		
	Total	6.683	99			

b. Predictors: (Constant), sales, education, assets, experience, gender, employee

APPENDIX 2c / *LAMPIRAN 2c*

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.153	8	.769	132.104	.000[b]
	Residual	.530	91	.006		
	Total	6.683	99			
2	Regression	6.153	7	.879	152.635	.000[c]
	Residual	.530	92	.006		
	Total	6.683	99			
3	Regression	6.153	6	1.025	179.916	.000[d]
	Residual	.530	93	.006		
	Total	6.683	99			
4	Regression	6.152	5	1.230	218.061	.000[e]
	Residual	.530	94	.006		
	Total	6.683	99			

a. Dependent Variable: salary

b. Predictors: (Constant), sales, age, profit, employee, education, assets, gender, experience

c. Predictors: (Constant), sales, profit, employee, education, assets, gender, experience

d. Predictors: (Constant), sales, employee, education, assets, gender, experience

e. Predictors: (Constant), employee, education, assets, gender, experience

...14/-

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.152	5	1.230	218.061	.000[b]
	Residual	.530	94	.006		
	Total	6.683	99			

b. Predictors: (Constant), experience, education, assets, gender, employee

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.468	5	1.094	84.658	.000[b]
	Residual	1.214	94	.013		
	Total	6.683	99			

b. Predictors: (Constant), sales, education, assets, experience, gender

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.723	5	1.145	112.157	.000[b]
	Residual	.959	94	.010		
	Total	6.683	99			

b. Predictors: (Constant), employee, experience, assets, sales, gender

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.064	5	1.213	184.388	.000[b]
	Residual	.618	94	.007		
	Total	6.683	99			

b. Predictors: (Constant), education, experience, gender, sales, employee

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.136	5	1.027	62.421	.000[b]
	Residual	1.547	94	.016		
	Total	6.683	99			

b. Predictors: (Constant), assets, experience, education, sales, employee

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.304	5	.461	9.894	.000[b]
	Residual	4.378	94	.047		
	Total	6.683	99			

b. Predictors: (Constant), gender, education, assets, sales, employee

APPENDIX 3a / *LAMPIRAN 3a*

	miles	displacement	horsepower	torque	compression
displacement	-0.872				
	0.000				
horsepower	-0.671	0.793			
	0.000	0.000			
torque	-0.850	0.990	0.801		
	0.000	0.000	0.000		
compression	0.422	-0.350	-0.335	-0.326	
	0.020	0.058	0.070	0.079	
rearaxle	0.635	-0.671	-0.398	-0.673	0.414
	0.000	0.000	0.029	0.000	0.023
carburetor	-0.472	0.640	0.547	0.653	0.037
	0.008	0.000	0.002	0.000	0.844
transmission	0.708	-0.772	-0.513	-0.746	0.558
	0.000	0.000	0.004	0.000	0.001
length	-0.753	0.865	0.691	0.864	-0.304
	0.000	0.000	0.000	0.000	0.102
width	-0.763	0.800	0.616	0.788	-0.378
	0.000	0.000	0.000	0.000	0.039
weight	-0.826	0.923	0.854	0.909	-0.381
	0.000	0.000	0.000	0.000	0.038

	rearaxle	carburetor	transmission	length	width
carburetor	-0.220				
	0.244				
transmission	0.872	-0.276			
	0.000	0.140			
length	-0.561	0.422	-0.655		
	0.001	0.020	0.000		
width	-0.453	0.300	-0.655	0.883	
	0.012	0.107	0.000	0.000	
weight	-0.539	0.464	-0.680	0.935	0.882
	0.002	0.010	0.000	0.000	0.000

```
The regression equation is
miles = 17.0 - 0.0715 displacement + 0.0144 horsepower + 0.0530 torque
        + 2.86 compression + 5.11 rearaxle - 0.37 carburetor - 4.03 transmission
        + 0.133 length - 0.441 width - 0.00253 weight

Predictor          Coef   SE Coef      T      P
Constant          16.98     30.36   0.56  0.582
displacement   -0.07149   0.05728  -1.25  0.227
horsepower      0.01436   0.04255   0.34  0.740
torque          0.05298   0.06807   0.78  0.446
compression       2.859     2.863   1.00  0.331
rearaxle          5.108     3.204   1.59  0.127
carburetor       -0.369     1.169  -0.32  0.756
transmission     -4.033     2.910  -1.39  0.182
length           0.1332    0.1233   1.08  0.294
width           -0.4410    0.3294  -1.34  0.196
weight        -0.002530  0.005497  -0.46  0.651

S = 3.25785   R-Sq = 82.3%   R-Sq(adj) = 73.0%

Analysis of Variance
Source          DF       SS     MS     F      P
Regression      10   937.45  93.74  8.83  0.000
Residual Error  19   201.66  10.61
Total           29  1139.11
```

APPENDIX 3b / *LAMPIRAN 3b*

Dependent	R^2	Dependent	R^2
Displacement	99.2	Carburetor	76.7
Horsepower	91.7	Transmisson	90.1
Torque	98.9	Length	94.6
Compression	41.9	Width	89.5
Rearaxle	87.3	Weight	98.1

APPENDIX 3c / *LAMPIRAN 3c*

```
The regression equation is
miles = 3.3 + 0.0444 horsepower - 0.0215 torque + 2.66 compression
        + 4.63 rearaxle - 0.94 carburetor - 2.52 transmission + 0.207 length
        - 0.332 width - 0.00725 weight

Predictor          Coef   SE Coef      T      P     VIF
Constant           3.32     28.71   0.12  0.909
horsepower      0.04445   0.03554   1.25  0.226   8.146
torque         -0.02151   0.03319  -0.65  0.524  20.211
compression       2.663     2.899   0.92  0.369   1.717
rearaxle          4.627     3.225   1.43  0.167   7.738
carburetor       -0.942     1.090  -0.86  0.398   3.636
transmission     -2.518     2.682  -0.94  0.359   8.350
length           0.2068    0.1099   1.88  0.074  14.188
width           -0.3315    0.3219  -1.03  0.315   8.882
weight        -0.007250  0.004044  -1.79  0.088  39.222

S = 3.30296   R-Sq = 80.8%   R-Sq(adj) = 72.2%
```

...17/-

```
Analysis of Variance
Source          DF       SS      MS     F      P
Regression       9   920.91  102.32  9.38  0.000
Residual Error  20   218.19   10.91
Total           29  1139.11



The regression equation is
miles = 33.8 - 0.0199 horsepower + 2.59 compression + 5.39 rearaxle
        - 1.32 carburetor - 1.87 transmission + 0.0461 length - 0.678 width

Predictor         Coef  SE Coef      T      P    VIF
Constant         33.84    26.70   1.27  0.218
horsepower    -0.01992  0.02080  -0.96  0.348  2.467
compression      2.588    3.072   0.84  0.409  1.706
rearaxle         5.392    2.956   1.82  0.082  5.752
carburetor     -1.3171   0.7804  -1.69  0.106  1.648
transmission    -1.867    2.825  -0.66  0.516  8.195
length         0.04609  0.08260   0.56  0.582  7.097
width          -0.6784   0.2950  -2.30  0.031  6.599

S = 3.51157   R-Sq = 76.2%   R-Sq(adj) = 68.6%

Analysis of Variance
Source          DF       SS      MS      F      P
Regression       7   867.82  123.97  10.05  0.000
Residual Error  22   271.28   12.33
Total           29  1139.11



The regression equation is
miles = 20.6 - 0.0488 displacement + 0.0046 horsepower + 3.03 compression
        + 3.41 rearaxle + 0.210 carburetor - 3.19 transmission + 0.0942 length
        - 0.432 width

Predictor         Coef  SE Coef      T      P     VIF
Constant         20.60    25.19   0.82  0.423
displacement  -0.04876  0.02154  -2.26  0.034  17.439
horsepower     0.00464  0.02195   0.21  0.835   3.264
compression      3.026    2.825   1.07  0.296   1.714
rearaxle         3.415    2.850   1.20  0.244   6.349
carburetor      0.2099   0.9837   0.21  0.833   3.110
transmission    -3.186    2.657  -1.20  0.244   8.609
length         0.09416  0.07871   1.20  0.245   7.653
width          -0.4324   0.2917  -1.48  0.153   7.662

S = 3.22233   R-Sq = 80.9%   R-Sq(adj) = 73.6%

Analysis of Variance
Source          DF       SS      MS      F      P
Regression       8   921.05  115.13  11.09  0.000
Residual Error  21   218.05   10.38
Total           29  1139.11
```

APPENDIX 3d / *LAMPIRAN 3d*

Response is miles

Vars	R-Sq	R-Sq(adj)	Mallows Cp	S	displacement	horsepower	compression ratio	carburetor	transmission	overall length	wheelbase	weight
1	76.0	75.2	-0.3	3.1218	X							
1	72.2	71.2	3.9	3.3642		X						
1	68.2	67.1	8.1	3.5955								X
2	77.6	76.0	0.0	3.0729	X		X					
2	77.3	75.6	0.4	3.0943	X			X				
2	77.2	75.5	0.4	3.0996	X						X	
3	78.4	75.9	1.2	3.0766	X		X				X	
3	78.2	75.7	1.4	3.0918	X	X	X					
3	78.1	75.6	1.5	3.0953	X		X	X				
4	79.0	75.7	2.5	3.0906	X			X		X	X	
4	79.0	75.7	2.5	3.0925	X		X			X	X	
4	78.9	75.6	2.6	3.0980	X	X	X				X	
5	79.7	75.4	3.8	3.1072	X			X		X	X	X
5	79.7	75.4	3.8	3.1073	X		X		X		X	
5	79.6	75.4	3.8	3.1081	X		X			X	X	
6	81.0	76.0	4.4	3.0690	X		X		X	X	X	
6	80.9	75.9	4.5	3.0775	X		X	X	X		X	
6	80.8	75.8	4.6	3.0852	X			X	X	X	X	
7	82.1	76.4	5.2	3.0459	X		X	X	X	X	X	
7	81.3	75.3	6.1	3.1127	X		X		X	X	X	X
7	81.1	75.1	6.3	3.1276	X	X	X		X	X	X	
8	82.1	75.3	7.2	3.1119	X		X	X	X	X	X	X
8	82.1	75.3	7.2	3.1168	X		X	X	X	X	X	
8	82.1	75.3	7.2	3.1171	X	X	X	X	X	X	X	
9	82.2	74.2	9.1	3.1837	X	X	X	X	X	X	X	X
9	82.2	74.2	9.1	3.1849	X		X	X	X	X	X	X
9	82.1	74.0	9.2	3.1930	X	X	X	X	X	X	X	
10	82.3	73.0	11.0	3.2578	X	X	X	X	X	X	X	X

APPENDIX 3e / *LAMPIRAN 3e*

Response is miles

Vars	R-Sq	R-Sq(adj)	Mallows Cp	S	horsepower	compress	carburs	transmis	wideth
1	58.2	56.7	12.6	4.1229					X
1	56.7	55.1	14.0	4.1983				X	
2	68.7	66.4	4.9	3.6331		X			X
2	65.8	63.3	7.6	3.7991			X		X
3	73.7	70.7	2.3	3.3934		X	X		X
3	72.8	69.7	3.1	3.4516	X	X			X
4	74.9	70.9	3.2	3.3811	X	X	X		X
4	74.7	70.7	3.3	3.3928		X	X	X	X
5	75.5	70.4	4.6	3.4085	X	X	X	X	X
5	75.3	70.1	4.9	3.4266	X		X	X	X
6	75.8	69.5	6.3	3.4586	X	X	X	X	X
6	75.7	69.4	6.4	3.4683	X	X	X	X	X
7	76.2	68.6	8.0	3.5116	X	X	X	X	X

APPENDIX 4/*LAMPIRAN 4*

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
educ	1.155	.602	3.677	1	.055	3.174
expr	.910	.429	4.492	1	.034	2.484
gender(1)	5.604	2.603	4.635	1	.031	271.423
Constant	-14.248	6.081	5.491	1	.019	.000

a. Variable(s) entered on step 1: educ, expr, gender.

Classification Table

Observed		Predicted		
		status		Percentage Correct
		0	1	
status	0	18	1	94.7
	1	2	7	77.8
Overall Percentage				89.3

- oooOOooo -