STAT 6800 HW3

Augustine Ennin

October 2025

Q1.(a)

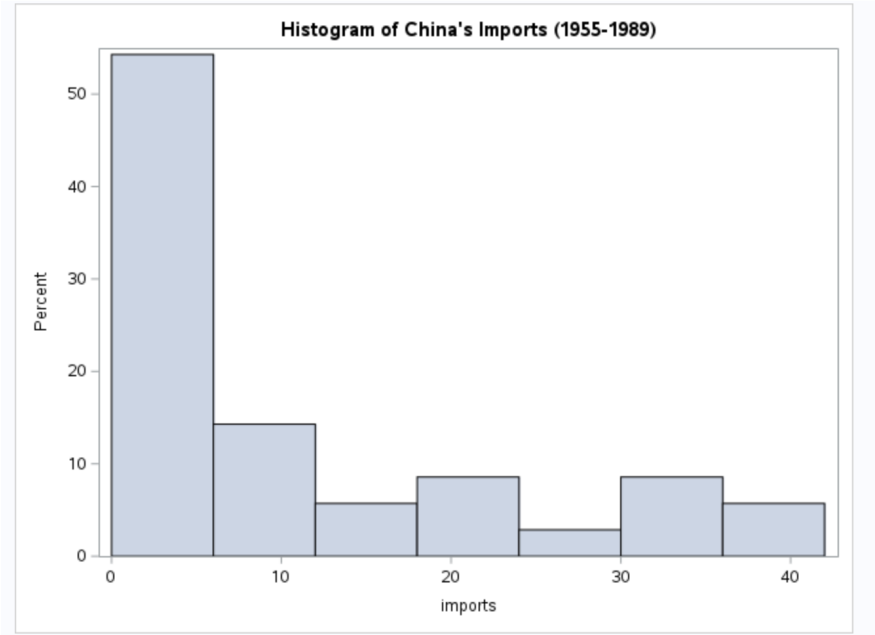
```
DATA china_trade;
INFILE "/home/u63997979/sasuser.v94/Elliott and Morrell/China#1.dat.txt";
INPUT year total exports imports;
RUN;
PROC PRINT DATA = china_trade(OBS=10);
TITLE "China Exports and Imports Data From 1955 to 1989";
RUN;
```

China Exports and Imports Data From 1955 to 1989

Obs	year	total	exports	imports
1	1955	3.15	1.41	1.73
2	1956	3.21	1.65	1.56
3	1957	3.10	1.60	1.51
4	1958	3.87	1.98	1.89
5	1959	4.38	2.26	2.12
6	1960	3.81	1.86	1.95
7	1961	2.94	1.49	1.45
8	1962	2.66	1.49	1.17
9	1963	2.92	1.65	1.27
10	1964	3.46	1.92	1.55

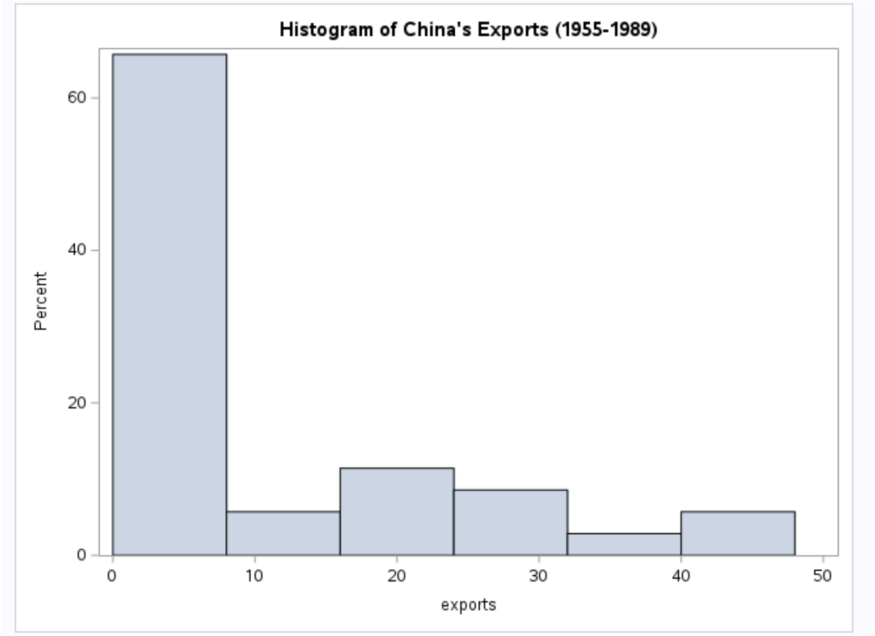
Q1.(b)

```
PROC SGPLOT DATA=china_trade;
HISTOGRAM imports;
TITLE "Histogram of China's Imports (1955-1989)";
RUN;
```



The histogram of China's imports from 1955 to 1989 exhibits a strongly right-skewed distribution, indicating that the vast majority of annual import values were relatively low, clustered in the 0-10 range, while only a few years experienced substantially higher import levels extending up to 40 units.

```
PROC SGPLOT DATA=china_trade;
HISTOGRAM exports;
TITLE "Histogram of China's Exports (1955-1989)";
RUN;
```

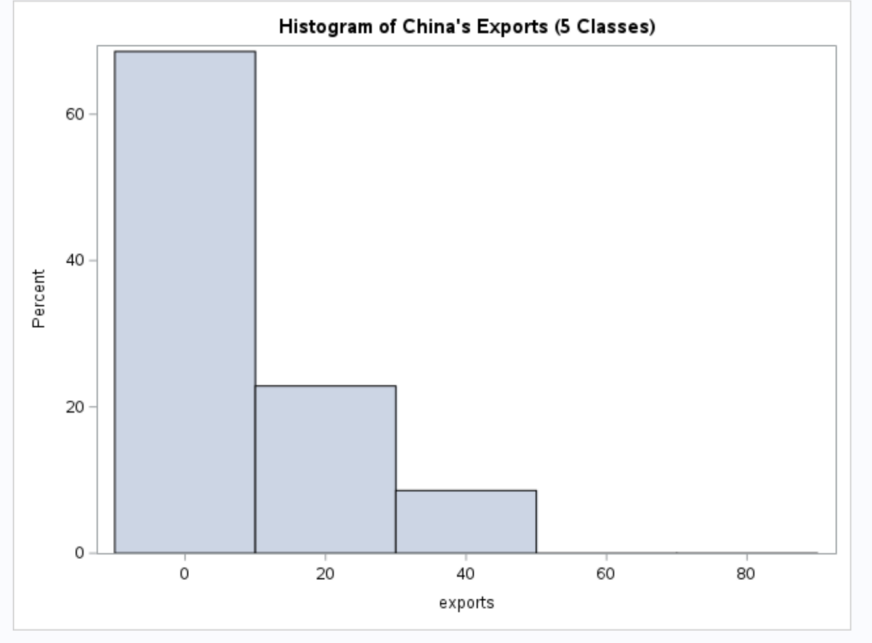


The histogram of China's exports from 1955 to 1989 is strongly right-skewed. A very large proportion of

years—as shown by the tall first bar—had export values at or near zero. The frequency of years drops sharply as export values increase, creating a long tail that extends toward higher values. This pattern indicates that for most of this period, China's export activity was minimal, with only occasional periods of significantly higher exports.

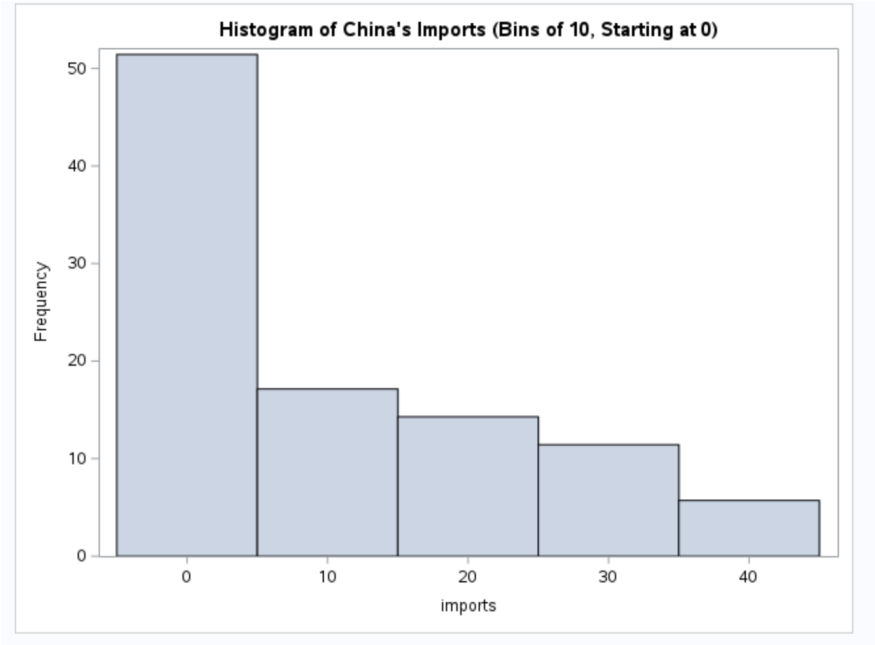
Q1.(c)

```
PROC SGPLOT DATA=china_trade;
HISTOGRAM exports / NBINS=5;
TITLE "Histogram of China's Exports (5 Classes)";
RUN;
```



Q1.(d)

```
PROC SGPLOT DATA=china_trade;
HISTOGRAM imports / binstart=0 binwidth=10;
YAXIS LABEL="Frequency";
TITLE "Histogram of China's Imports (Bins of 10, Starting at 0)";
RUN;
```



Q2.(a)

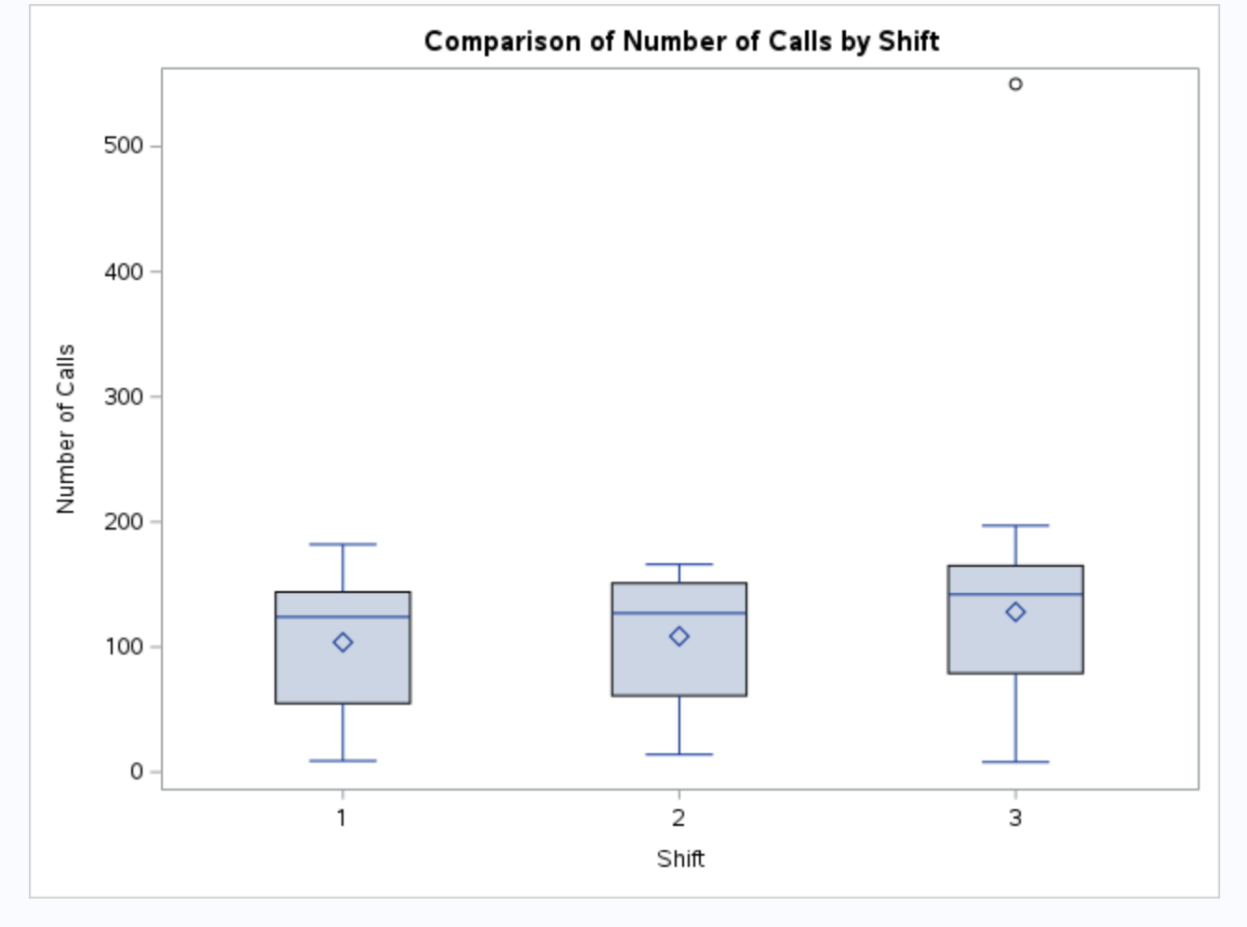
```
DATA calls;
INFILE "/home/u63997979/sasuser.v94/Elliott and Morrell/Calls.dat";
INPUT Week 1-2 Shift 4 Day $ 6-8 Number_of_Calls 10-12;
RUN;
PROC PRINT DATA=calls(OBS=10);
TITLE "First 10 Observartion of the Call Data Set";
RUN;
```

First 10 Observartion of the Call Data Set

Obs	Week	Shift	Day	Number_of_Calls
1	1	1	Mon	9
2	1	2	Mon	16
3	1	3	Mon	14
4	1	1	Tue	12
5	1	2	Tue	17
6	1	3	Tue	21
7	1	1	Wed	17
8	1	2	Wed	14
9	1	3	Wed	20
10	1	1	Thu	14

Q2.(b)

```
PROC SGPLOT DATA=calls;
VBOX Number_of_Calls / CATEGORY=Shift;
XAXIS LABEL="Shift";
YAXIS LABEL="Number of Calls";
TITLE "Comparison of Number of Calls by Shift";
RUN;
```



The distributions of the number of calls across all shifts appear similar and exhibit a left-skewed pattern. There is an outlier in Shift 3, with an approximate value of around 550 calls.

Q3.(a)

```
DATA utility;
INFILE "/home/u63997979/sasuser.v94/Elliott and Morrell/Utility.dat.txt";
INPUT Date $ 1-6 Telephone 9-15 Fuel 17-22 Electricity 25-29;
RUN;
PROC PRINT DATA=utility(OBS=10);
TITLE "First 10 Observartion of the Utility Data Set";
RUN;
```

First 10 Observartion of the Utility Data Set

Obs	Date	Telephone	Fuel	Electricity
1	Aug 88	100.02	41.61	36.93
2	Sep 88	80.62	24.48	45.73
3	Oct 88	62.55	23.90	50.95
4	Nov 88	69.35	48.67	48.93
5	Dec 88	90.79	120.60	56.61
6	Jan 89	40.27	151.23	50.44
7	Feb 89	49.29	144.29	44.50
8	Mar 89	91.50	72.75	40.67
9	Apr 89	93.71	49.63	36.04
10	May 89	46.64	33.22	39.79

Q3.(b)

H_0 : Average monthly telephone costs ≤ 65

H_1 : Average monthly telephone costs > 65

```
1 PROC TTEST DATA=utility H0=65 SIDES=U ALPHA=0.05;
2 VAR Telephone;
3 TITLE "One-Sided T-Test for Telephone Costs";
4 RUN;
```

One-Sided T-Test for Telephone Costs

The TTEST Procedure

Variable: Telephone

N	Mean	Std Dev	Std Err	Minimum	Maximum
55	73.3678	27.2047	3.6683	39.3900	145.2

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
73.3678	67.2287	Infty	27.2047	22.9027	33.5119

DF	t Value	Pr > t
54	2.28	0.0133

Since the p-value (0.0133) is less than the significance level of 0.05, we reject the null hypothesis. There is significant evidence to suggest that the true average monthly telephone cost is statistically higher than $65 per month.

Q3.(c)

H_0 : True average monthly electricity costs = 45
H_1 : True average monthly electricity costs $\neq$ 45

```
PROC TTEST DATA=utility H0=40 SIDES=2 ALPHA=0.05;
VAR Electricity;
TITLE "Two-Sided T-Test for Telephone Costs";
RUN;
```

Two-Sided T-Test for Telephone Costs

The TTEST Procedure

Variable: Electricity

N	Mean	Std Dev	Std Err	Minimum	Maximum
55	32.2964	8.4862	1.1443	18.6900	56.6100

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
32.2964	30.0022	34.5905	8.4862	7.1442	10.4536

DF	t Value	Pr > |t|
54	-6.73	<.0001

Since the p-value (<.0001) is less than the significance level of 0.05, we reject the null hypothesis. There is significant evidence to suggest that the true average monthly electricity cost is statistically different from $40 per month.

Q4.(a)

```
DATA BTT;
INFILE "/home/u63997979/sasuser.v94/Elliott and Morrell/btt.dat.txt";
INPUT
childid 1-4
sex 6
bweight 8-11
gestage 13-14
momage 16-17
parity 19
mdbp 21-23
msbp 25-27
momeduc 29
mmedaid 31
socio 33
dbp5 35-37
sbp5 39-41
ht5 43-47
wt5 49-52
hdl5 54-57
ldl5 59-62
```

```
trig5 64-67
smoke5 69
medaid 71
socio5 73;
bmi = wt5 / (ht5)**2;
RUN;
PROC PRINT DATA = BTT (OBS=10);
VAR socio momeduc;
TITLE "Birth to Ten Socioeconomic Status and Mother's Education first 10
    Observations";
RUN;
```

Birth to Ten Socioeconomic Status and Mother's Education first 10 Observations

Obs	socio	momeduc
1	1	3
2	1	1
3	1	3
4	2	3
5	1	4
6	1	3
7	1	3
8	1	4
9	1	3
10	1	2

Q4.(b)

$H0$: Mother's socioeconomic status at child birth is uniformly distributed accross all five categories

$H1$: Mother's socioeconomic status at child birth is not uniformly distributed accross all five categories

```
PROC FREQ DATA=BTT;
TABLE socio / CHISQ TESTP=(20 20 20 20 20);
TITLE "Chi-Square Goodness-of-Fit Test for Mom's Socioeconomic Status";
RUN;
```

Chi-Square Goodness-of-Fit Test for Mom's Socioeconomic Status

The FREQ Procedure

socio	Frequency	Percent	Test Percent	Cumulative Frequency	Cumulative Percent
0	7	3.23	20.00	7	3.23
1	163	75.12	20.00	170	78.34
2	41	18.89	20.00	211	97.24
3	5	2.30	20.00	216	99.54
4	1	0.46	20.00	217	100.00

Chi-Square Test for Specified Proportions	
Chi-Square	435.6498
DF	4
Pr > ChiSq	<.0001

Reject H0 since the p-value (< 0.0001) is less than the significance level of 0.05. Therefore, there is significant statistical evidence to suggest that the distribution of the mother's socioeconomic status is not uniformly distributed across the five categories.

Q4.(c)

$H0$: Mother's socioeconomic status at child birth is independent of mother's education at child birth
$H1$: Mother's socioeconomic status at child birth is not independent of mother's education at child birth

```
PROC FREQ DATA=BTT;
TABLE socio*momeduc / CHISQ;
TITLE "Chi-square Test of Independence: Mom's Socioeconomic Status vs. Mom's
    Education";
RUN;
```

Chi-square Test of Independence: Mom's Socioeconomic Status vs. Mom's Education

The FREQ Procedure

Frequency
Percent
Row Pct
Col Pct

Table of socio by momeduc

socio	momeduc 1	2	3	4	Total
0	2 0.92 28.57 7.14	1 0.46 14.29 2.38	4 1.84 57.14 3.70	0 0.00 0.00 0.00	7 3.23
1	21 9.68 12.88 75.00	32 14.75 19.63 76.19	79 36.41 48.47 73.15	31 14.29 19.02 79.49	163 75.12
2	5 2.30 12.20 17.86	8 3.69 19.51 19.05	20 9.22 48.78 18.52	8 3.69 19.51 20.51	41 18.89
3	0 0.00 0.00 0.00	1 0.46 20.00 2.38	4 1.84 80.00 3.70	0 0.00 0.00 0.00	5 2.30
4	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.46 100.00 0.93	0 0.00 0.00 0.00	1 0.46
Total	28 12.90	42 19.35	108 49.77	39 17.97	217 100.00

Statistics for Table of socio by momeduc

Statistic	DF	Value	Prob
Chi-Square	12	6.4715	0.8905
Likelihood Ratio Chi-Square	12	9.1731	0.6881
Mantel-Haenszel Chi-Square	1	0.6100	0.4348
Phi Coefficient		0.1727	
Contingency Coefficient		0.1702	
Cramer's V		0.0997	
WARNING: 60% of the cells have expected counts less than 5. Chi-Square may not be a valid test.			

Sample Size = 217

We fail to reject H0 since the p-value (0.8905) is greater than the significance level of 5%. Hence, there is no significant evidence to suggest that the mother's socioeconomic status is related to the mother's education level at childbirth. Therefore, the two variables appear to be independent.

Q5

```
DATA electric;
INFILE "/home/u63997979/sasuser.v94/Elliott and Morrell/Electric.dat";
INPUT houseSize 1-3 familyIncome 6-11 airconCapacity 14-16 applianceIndex 19-23
    familyMembers 26-28 peakHourLoad 31-35;
RUN;
PROC PRINT DATA=electric(OBS=10);
TITLE "First 10 Observation of Electric Data";
RUN;
```

First 10 Observation of Electric Data

Obs	houseSize	familyIncome	airconCapacity	applianceIndex	familyMembers	peakHourLoad
1	3.2	34.990	7.0	7.789	4.0	7.518
2	1.3	14.160	0.5	3.652	4.0	2.349
3	4.1	22.962	3.0	5.854	1.0	5.059
4	2.3	24.535	5.0	4.975	2.0	5.010
5	1.9	20.614	3.0	4.817	6.0	4.505
6	1.9	20.677	1.0	4.659	1.0	2.976
7	3.3	30.016	6.5	6.054	1.0	6.849
8	2.4	26.341	3.5	7.345	4.0	5.829
9	2.6	28.731	6.5	6.325	3.0	5.910
10	2.9	32.362	3.5	7.700	5.5	5.990

Q5(a)

```
PROC REG DATA=electric;
MODEL peakHourLoad = airconCapacity / CLM;
TITLE "Regression: Peak Hour Load vs. Airconditioning Capacity";
RUN;
```

Regression: Peak Hour Load vs. Airconditioning Capacity

The REG Procedure
Model: MODEL1
Dependent Variable: peakHourLoad

Number of Observations Read	60
Number of Observations Used	60

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	103.68699	103.68699	355.80	<.0001
Error	58	16.90217	0.29142		
Corrected Total	59	120.58915			

Root MSE	0.53983	R-Square	0.8598
Dependent Mean	4.63792	Adj R-Sq	0.8574
Coeff Var	11.63950		

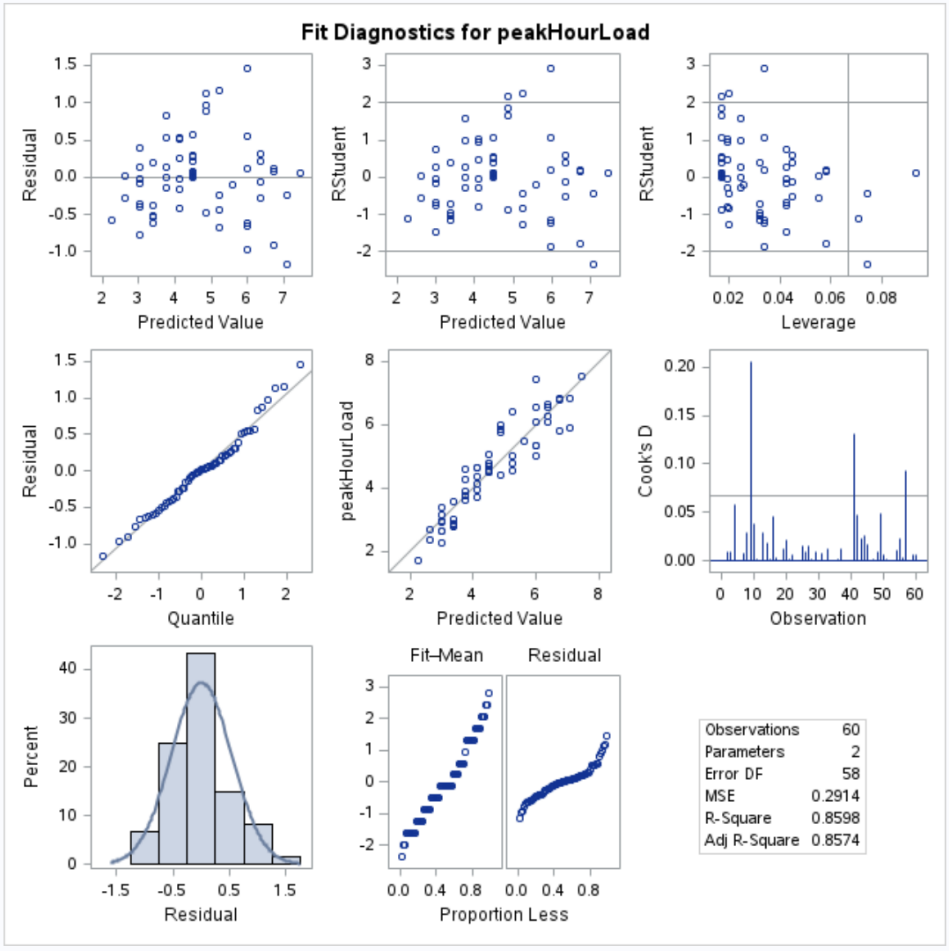
Parameter Estimates

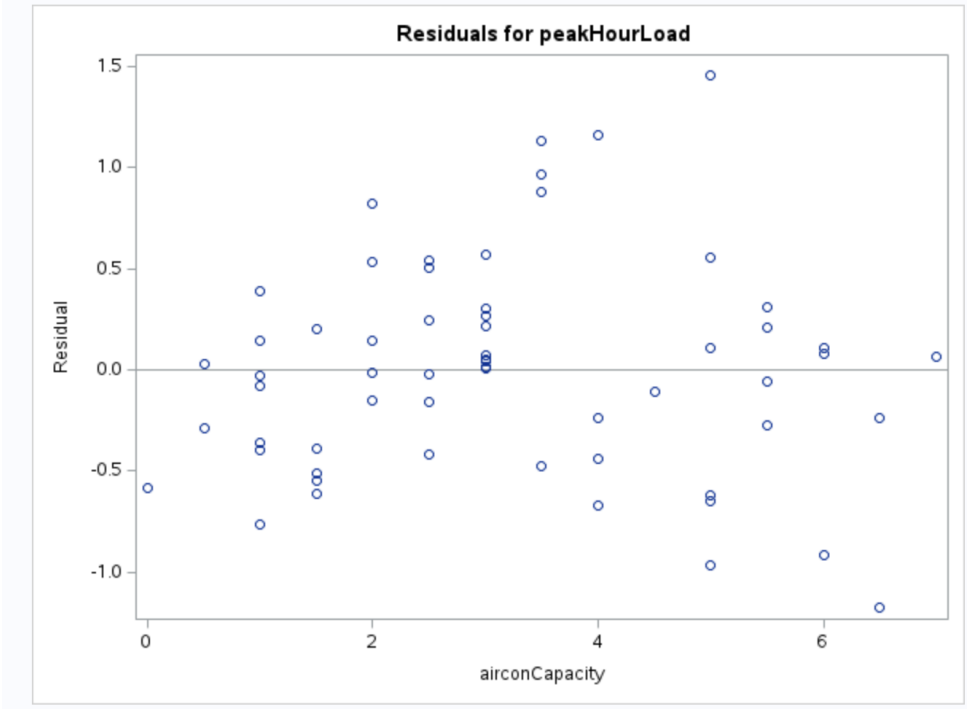
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	2.26523	0.14380	15.75	<.0001
airconCapacity	1	0.74147	0.03931	18.86	<.0001

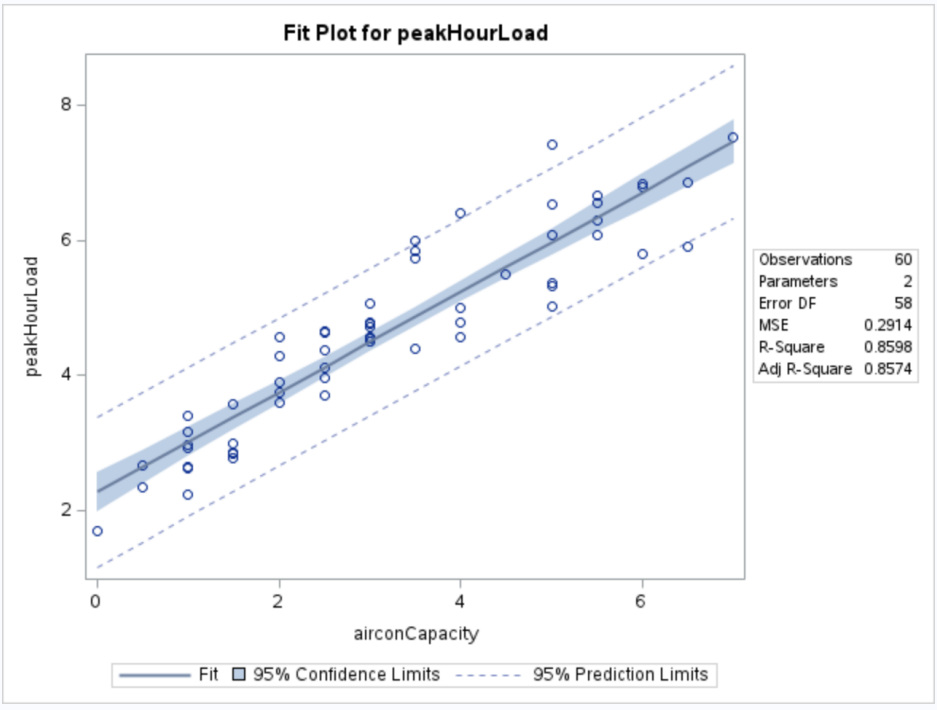
The REG Procedure
Model: MODEL1
Dependent Variable: peakHourLoad

Output Statistics						
Obs	Dependent Variable	Predicted Value	Std Error Mean Predict	95% CL Mean		Residual
1	7.52	7.4555	0.1648	7.1255	7.7854	0.0625
2	2.35	2.6360	0.1270	2.3818	2.8901	-0.2870
3	5.06	4.4896	0.0701	4.3492	4.6300	0.5694
4	5.01	5.9726	0.0993	5.7738	6.1714	-0.9626
5	4.51	4.4896	0.0701	4.3492	4.6300	0.0154
6	2.98	3.0067	0.1111	2.7844	3.2290	-0.0307
7	6.85	7.0848	0.1473	6.7900	7.3795	-0.2358
8	5.83	4.8604	0.0707	4.7189	5.0018	0.9686
9	5.91	7.0848	0.1473	6.7900	7.3795	-1.1748
10	5.99	4.8604	0.0707	4.7189	5.0018	1.1296
11	5.00	5.2311	0.0765	5.0780	5.3841	-0.2341
12	2.93	3.0067	0.1111	2.7844	3.2290	-0.0817
13	4.57	3.7482	0.0842	3.5797	3.9166	0.8228
14	6.53	5.9726	0.0993	5.7738	6.1714	0.5554
15	4.71	4.4896	0.0701	4.3492	4.6300	0.2174

Sum of Residuals	0
Sum of Squared Residuals	16.90217
Predicted Residual SS (PRESS)	18.10373


Fit Diagnostics for peakHourLoad
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
peakHourLoad
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 60
Parameters 2
Error DF 58
MSE 0.2914
R-Square 0.8598
Adj R-Square 0.8574





There is strong evidence of a positive linear relationship between air conditioning capacity and peak-hour electrical load. The high R-squared value of 0.8598 indicates that air conditioning capacity explains approximately 86% of the variability in peak-hour loads. The regression line fits the data very well, with narrow 95% confidence intervals suggesting precise estimation of mean peak loads across different air conditioning

capacities. This suggests that air conditioning capacity is a major determinant of household electricity demand during peak hours.

Q5(b)

```
PROC REG DATA=electric;
MODEL peakHourLoad = applianceIndex / CLM CLI;
TITLE "Regression: Peak Hour Load vs. Appliance Index";
RUN;
```

Regression: Peak Hour Load vs. Appliance Index

The REG Procedure
Model: MODEL1
Dependent Variable: peakHourLoad

Number of Observations Read	60
Number of Observations Used	60

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	94.66932	94.66932	211.84	<.0001
Error	58	25.91984	0.44689		
Corrected Total	59	120.58915			

Root MSE	0.66850	R-Square	0.7851
Dependent Mean	4.63792	Adj R-Sq	0.7814
Coeff Var	14.41382		

Parameter Estimates

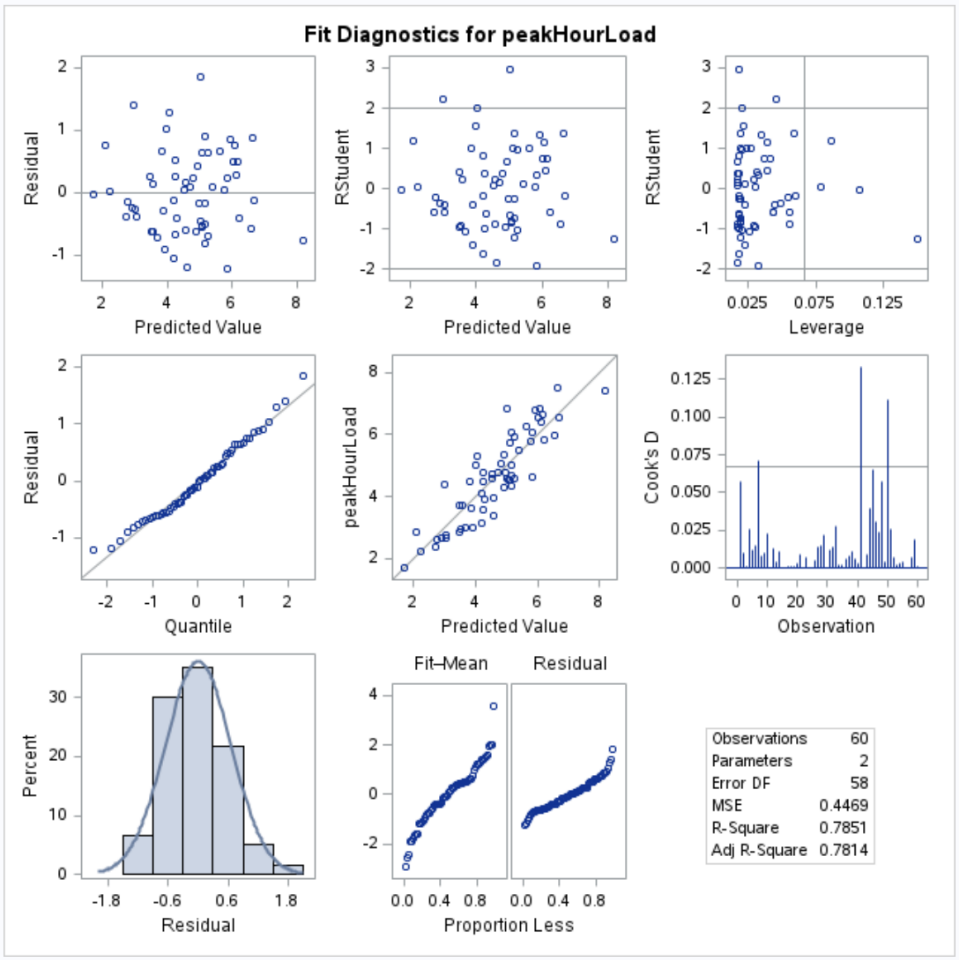
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-0.72880	0.37869	-1.92	0.0592
applianceIndex	1	0.94680	0.06505	14.55	<.0001

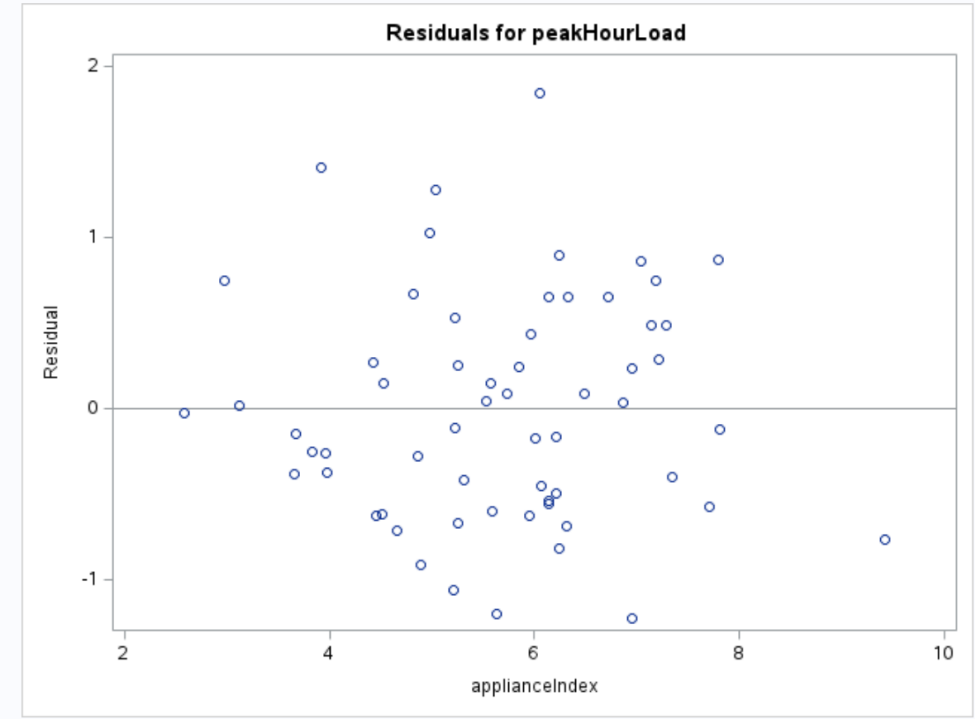
Regression: Peak Hour Load vs. Appliance Index

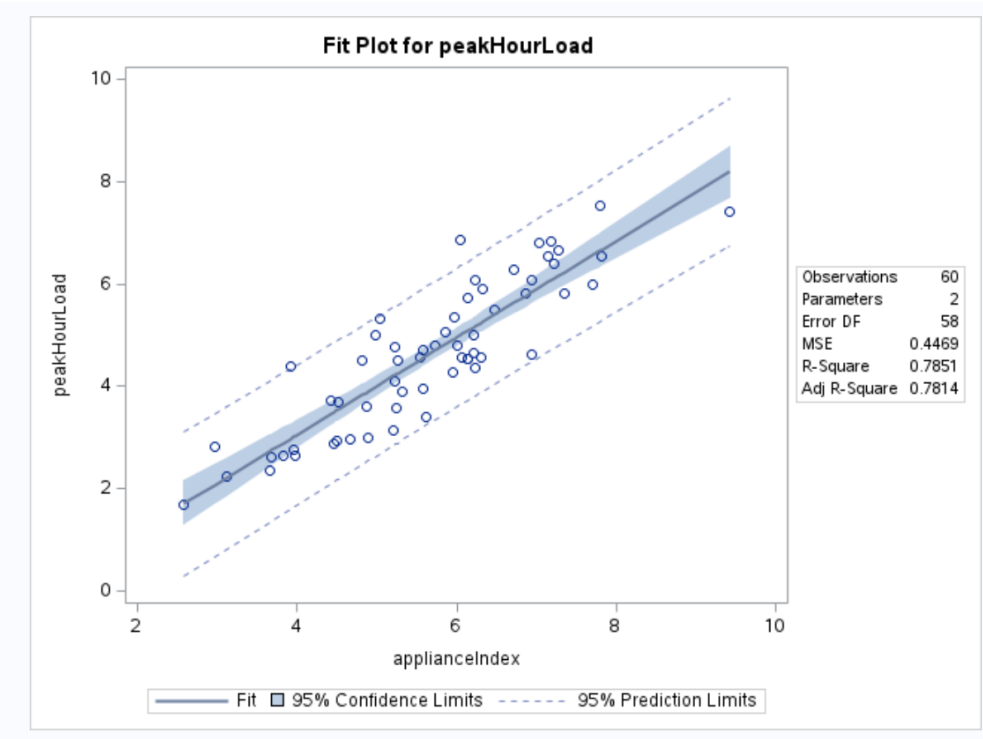
The REG Procedure
Model: MODEL1
Dependent Variable: peakHourLoad

Output Statistics								
Obs	Dependent Variable	Predicted Value	Std Error Mean Predict	95% CL Mean		95% CL Predict		Residual
1	7.52	6.6458	0.1627	6.3201	6.9715	5.2686	8.0230	0.8722
2	2.35	2.7289	0.1570	2.4146	3.0432	1.3543	4.1035	-0.3799
3	5.06	4.8138	0.0871	4.6393	4.9882	3.4643	6.1632	0.2452
4	5.01	3.9815	0.0974	3.7866	4.1764	2.6292	5.3338	1.0285
5	4.51	3.8319	0.1025	3.6267	4.0372	2.4781	5.1857	0.6731
6	2.98	3.6823	0.1084	3.4653	3.8994	2.3267	5.0380	-0.7063
7	6.85	5.0031	0.0899	4.8232	5.1830	3.6529	6.3533	1.8459
8	5.83	6.2254	0.1391	5.9470	6.5038	4.8586	7.5922	-0.3964
9	5.91	5.2597	0.0963	5.0669	5.4525	3.9077	6.6117	0.6503
10	5.99	6.5615	0.1578	6.2456	6.8775	5.1866	7.9365	-0.5715
11	5.00	5.1565	0.0934	4.9696	5.3434	3.8054	6.5076	-0.1595
12	2.93	3.5375	0.1147	3.3078	3.7671	2.1798	4.8952	-0.6125
13	4.57	5.0211	0.0902	4.8405	5.2017	3.6708	6.3714	-0.4501
14	6.53	6.0361	0.1291	5.7776	6.2946	4.6732	7.3990	0.4919
15	4.71	4.5534	0.0865	4.3802	4.7265	3.2041	5.9027	0.1536

Sum of Residuals	0
Sum of Squared Residuals	25.91984
Predicted Residual SS (PRESS)	27.63103


Fit Diagnostics for peakHourLoad
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
peakHourLoad
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 60
Parameters 2
Error DF 58
MSE 0.4469
R-Square 0.7851
Adj R-Square 0.7814





There is strong statistical evidence of a positive linear relationship between appliance index and peak-hour electrical load ($R^2 = 0.7851$, $p < 0.0001$). The appliance index explains approximately 78% of the variability in peak-hour loads, indicating it is a major determinant of household electricity demand during peak hours. The wider prediction intervals compared to confidence intervals appropriately reflect the greater uncertainty

in forecasting individual household consumption versus estimating average consumption for households with similar appliance indices.

Q5(c)

```
PROC REG DATA=electric;
MODEL peakHourLoad = familyMembers / CLM;
TITLE "Regression: Peak Hour Load vs. Family Members";
RUN;
```

Regression: Peak Hour Load vs. Family Members

The REG Procedure
Model: MODEL1
Dependent Variable: peakHourLoad

Number of Observations Read	60
Number of Observations Used	60

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.54044	0.54044	0.26	0.6113
Error	58	120.04871	2.06981		
Corrected Total	59	120.58915			

Root MSE	1.43868	R-Square	0.0045
Dependent Mean	4.63792	Adj R-Sq	-0.0127
Coeff Var	31.02000		

Parameter Estimates

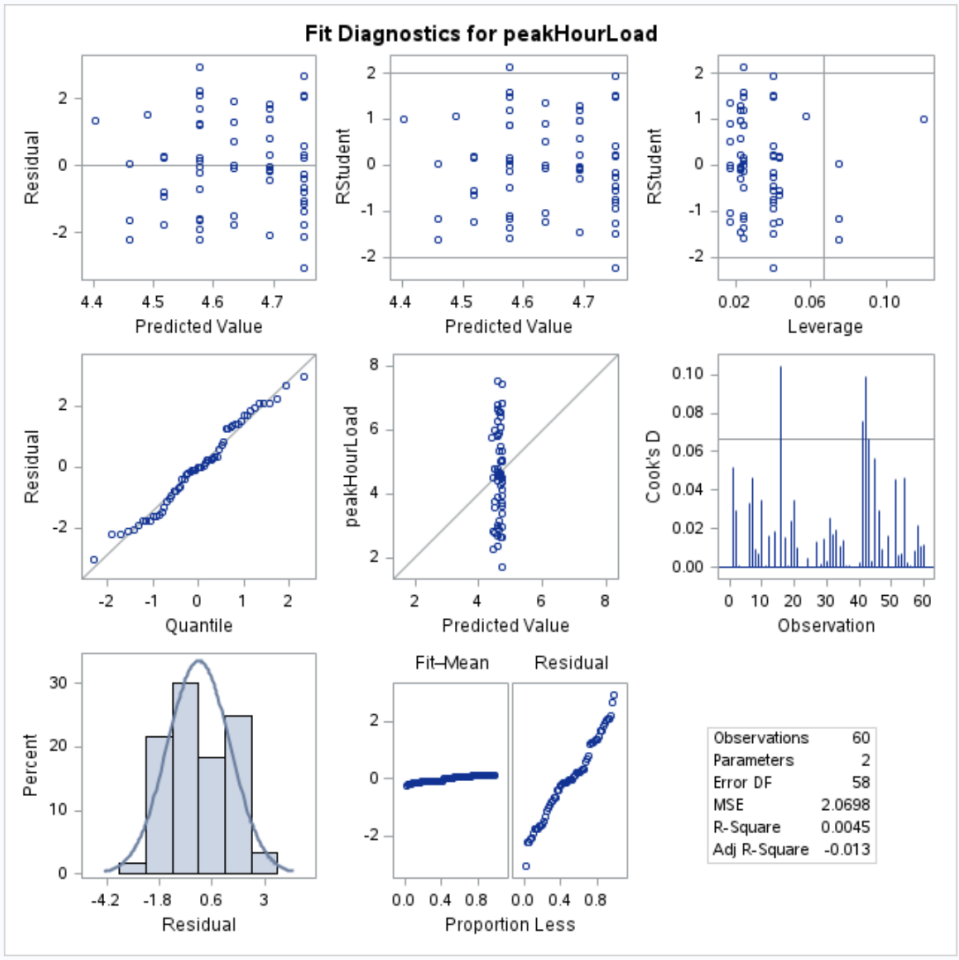
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	4.80886	0.38263	12.57	<.0001
familyMembers	1	-0.05811	0.11372	-0.51	0.6113

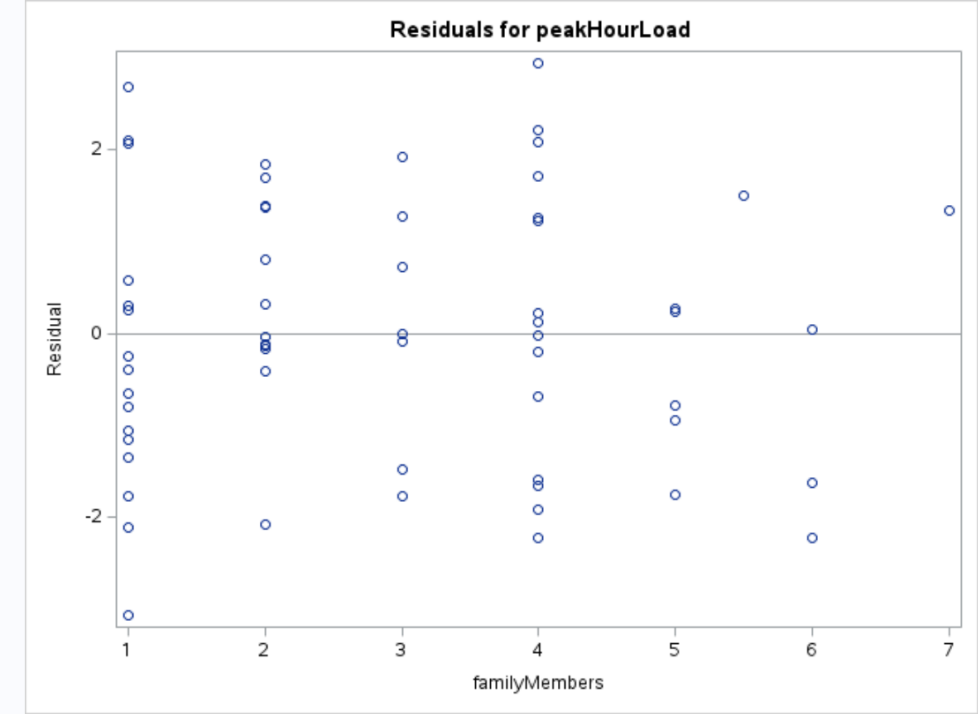
Regression: Peak Hour Load vs. Family Members

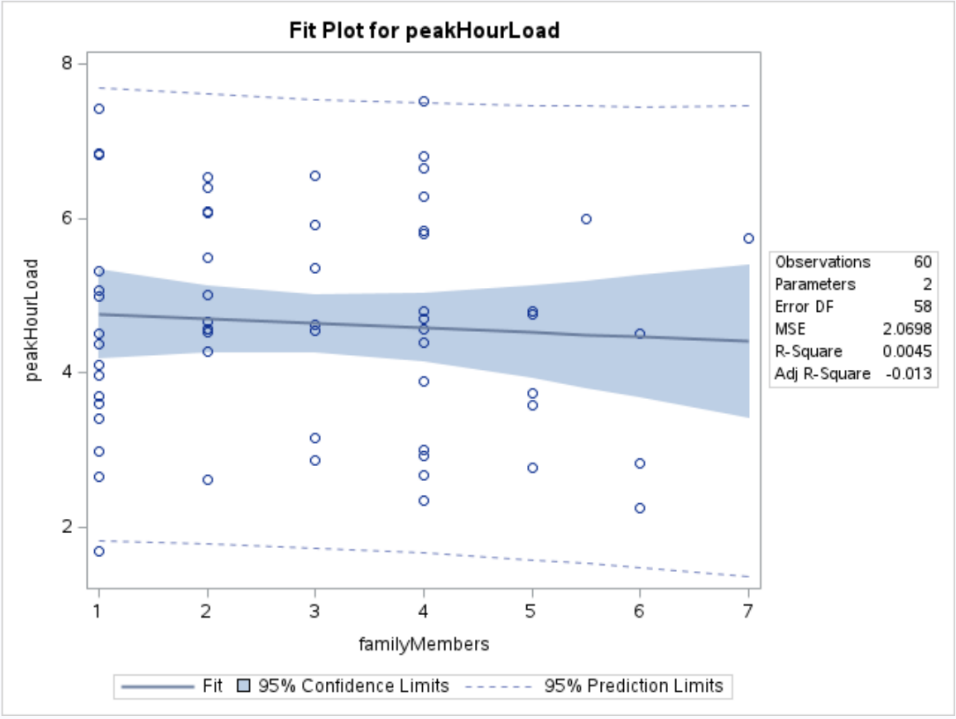
The REG Procedure
Model: MODEL1
Dependent Variable: peakHourLoad

Output Statistics						
Obs	Dependent Variable	Predicted Value	Std Error Mean Predict	95% CL Mean		Residual
1	7.52	4.5764	0.2213	4.1334	5.0194	2.9416
2	2.35	4.5764	0.2213	4.1334	5.0194	-2.2274
3	5.06	4.7507	0.2885	4.1732	5.3283	0.3083
4	5.01	4.6926	0.2144	4.2635	5.1218	0.3174
5	4.51	4.4602	0.3943	3.6709	5.2494	0.0448
6	2.98	4.7507	0.2885	4.1732	5.3283	-1.7747
7	6.85	4.7507	0.2885	4.1732	5.3283	2.0983
8	5.83	4.5764	0.2213	4.1334	5.0194	1.2526
9	5.91	4.6345	0.1859	4.2625	5.0065	1.2755
10	5.99	4.4893	0.3452	3.7983	5.1802	1.5007
11	5.00	4.7507	0.2885	4.1732	5.3283	0.2463
12	2.93	4.5764	0.2213	4.1334	5.0194	-1.6514
13	4.57	4.6926	0.2144	4.2635	5.1218	-0.1216
14	6.53	4.6926	0.2144	4.2635	5.1218	1.8354
15	4.71	4.5764	0.2213	4.1334	5.0194	0.1306

Sum of Residuals	0
Sum of Squared Residuals	120.04871
Predicted Residual SS (PRESS)	128.85905


Fit Diagnostics for peakHourLoad
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
peakHourLoad
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 60
Parameters 2
Error DF 58
MSE 2.0698
R-Square 0.0045
Adj R-Square -0.013





Based on the regression analysis, there is no statistically significant linear relationship between the number of family members and peak-hour electricity load ($p = 0.6113$). The null hypothesis that the slope equals zero cannot be rejected at the 0.05 significance level. The model also has a very low R-squared (0.0045), which means family members explain little to nothing of peak-hour electricity load. This indicates that

family size, by itself, is not a meaningful predictor of peak electricity demand in this population.

Q6

```
DATA gas;
INFILE "/home/u63997979/sasuser.v94/Elliott and Morrell/Gas.dat";
INPUT Displacement 1-5 Horsepower 7-9 Torque 11-13 transmissionSpeeds 25 carLength
     27-31 carWeight 38-41 gasMileage 45-48;
RUN;
PROC PRINT DATA=gas(OBS=10);
TITLE "First 10 Observation of the Gas Data";
RUN;
```

First 10 Observation of the Gas Data

Obs	Displacement	Horsepower	Torque	transmissionSpeeds	carLength	carWeight	gasMileage
1	318.0	140	255	3	215.3	4370	19.7
2	440.0	215	330	3	184.5	4215	11.2
3	351.0	143	255	3	199.9	3890	18.3
4	360.0	180	290	3	214.2	4250	21.5
5	140.0	83	109	4	168.8	2700	20.3
6	85.3	80	83	4	160.6	2009	36.5
7	350.0	165	260	3	200.3	3910	18.9
8	96.9	75	83	5	162.5	2320	30.4
9	351.0	148	243	3	215.5	4540	13.9
10	440.0	215	330	3	231.0	5185	14.9

Q6(a)

```
PROC REG DATA=gas;
MODEL gasMileage = Displacement Horsepower Torque transmissionSpeeds carWeight
    carLength / SELECTION=STEPWISE SLENTRY=0.10 SLSTAY=0.10;
TITLE "Stepwise Regression for Gas Mileage";
TITLE2 "Response: gasMileage, Predictors: All vehicle characteristics";
RUN;
```

Stepwise Regression for Gas Mileage
Response: gasMileage, Predictors: All vehicle characteristics

The REG Procedure
Model: MODEL1
Dependent Variable: gasMileage

Number of Observations Read	30
Number of Observations Used	30

Stepwise Selection: Step 1

Variable Displacement Entered: R-Square = 0.7601 and C(p) = 0.5804

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	864.52693	864.52693	88.70	<.0001
Error	28	272.90673	9.74667		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	33.48780	1.53711	4626.15275	474.64	<.0001
Displacement	-0.04706	0.00500	864.52693	88.70	<.0001

Bounds on condition number: 1, 1

All variables left in the model are significant at the 0.1000 level.

No other variable met the 0.1000 significance level for entry into the model.

Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	Displacement		1	0.7601	0.7601	0.5804	88.70	<.0001

The stepwise regression analysis identified engine displacement as the sole significant predictor of gas mileage in this dataset. The final model explains 76.01% of the variability in fuel efficiency ($R^2 = 0.7601$), indicating a very strong relationship. The Mallows' C(p) value of 0.5804 confirms this is an appropriately specified model without overfitting.

Q6(b)

```
PROC REG DATA=gas;
MODEL gasMileage = Displacement Horsepower Torque transmissionSpeeds carWeight
    carLength / SELECTION=Backward;
TITLE "Backward Regression for Gas Mileage";
TITLE2 "Response: gasMileage, Starting with All Predictors";
RUN;
```

Backward Regression for Gas Mileage
Response: gasMileage, Starting with All Predictors

The REG Procedure
Model: MODEL1
Dependent Variable: gasMileage

Number of Observations Read	30
Number of Observations Used	30

Backward Elimination: Step 0

All Variables Entered: R-Square = 0.7924 and C(p) = 7.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	901.28739	150.21457	14.63	<.0001
Error	23	236.14627	10.26723		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	13.86795	15.27070	8.46760	0.82	0.3732
Displacement	-0.03263	0.04831	4.68391	0.46	0.5061
Horsepower	-0.01435	0.06267	0.53832	0.05	0.8209
Torque	0.03694	0.07252	2.66416	0.26	0.6153
transmissionSpeeds	1.49297	1.71782	7.75537	0.76	0.3938
carWeight	-0.00684	0.00439	24.95056	2.43	0.1327
carLength	0.15213	0.11790	17.09463	1.66	0.2098

Bounds on condition number: 102.49, 1696.5

Backward Elimination: Step 1

Variable Horsepower Removed: R-Square = 0.7919 and C(p) = 5.0524

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	900.74907	180.14981	18.27	<.0001
Error	24	236.68460	9.86186		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	14.06809	14.94167	8.74241	0.89	0.3558
Displacement	-0.03123	0.04697	4.36100	0.44	0.5124
Torque	0.02551	0.05157	2.41364	0.24	0.6253
transmissionSpeeds	1.30900	1.48808	7.63115	0.77	0.3878
carWeight	-0.00679	0.00430	24.63398	2.50	0.1271
carLength	0.15393	0.11529	17.57923	1.78	0.1944

Bounds on condition number: 87.343, 1048.5

Backward Elimination: Step 2

Variable Torque Removed: R-Square = 0.7898 and C(p) = 3.2875

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	898.33543	224.58386	23.48	<.0001
Error	25	239.09823	9.56393		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	12.10912	14.18812	6.96646	0.73	0.4015
Displacement	-0.01074	0.02178	2.32311	0.24	0.6264
transmissionSpeeds	1.49445	1.41817	10.62045	1.11	0.3021
carWeight	-0.00731	0.00410	30.35908	3.17	0.0870
carLength	0.16916	0.10942	22.85604	2.39	0.1347

Bounds on condition number: 45.174, 333.83

Backward Elimination: Step 3

Variable Displacement Removed: R-Square = 0.7877 and C(p) = 1.5138

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	896.01232	298.67077	32.17	<.0001
Error	26	241.42134	9.28544		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	8.55178	12.03580	4.68776	0.50	0.4837
transmissionSpeeds	1.84319	1.21100	21.51062	2.32	0.1401
carWeight	-0.00901	0.00217	159.88565	17.22	0.0003
carLength	0.19781	0.09133	43.55770	4.69	0.0397

Bounds on condition number: 13.047, 80.123

Backward Elimination: Step 4

Variable transmissionSpeeds Removed: R-Square = 0.7688 and C(p) = 1.6089

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	874.50170	437.25085	44.90	<.0001
Error	27	262.93196	9.73822		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	17.05262	10.91829	23.75486	2.44	0.1300
carWeight	-0.01011	0.00210	225.75634	23.18	<.0001
carLength	0.20618	0.09336	47.49395	4.88	0.0359

Bounds on condition number: 11.618, 46.474

All variables left in the model are significant at the 0.1000 level.

Summary of Backward Elimination							
Step	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	Horsepower	5	0.0005	0.7919	5.0524	0.05	0.8209
2	Torque	4	0.0021	0.7898	3.2875	0.24	0.6253
3	Displacement	3	0.0020	0.7877	1.5138	0.24	0.6264
4	transmissionSpeeds	2	0.0189	0.7688	1.6089	2.32	0.1401

The backward elimination procedure resulted in a highly significant two-predictor model ($F = 44.90$, $p < 0.0001$) containing car weight and car length. This suggests that physical vehicle dimensions are more fundamental predictors of fuel efficiency than engine characteristics when all variables are considered together.

Q6(c)

```
PROC REG DATA=gas;
MODEL gasMileage = Displacement Horsepower Torque transmissionSpeeds carWeight
    carLength / SELECTION=Forward;
TITLE "Forward Regression: Selecting Predictors for Gas Mileage";
RUN;
```

Forward Regression: Selecting Predictors for Gas Mileage

The REG Procedure
Model: MODEL1
Dependent Variable: gasMileage

Number of Observations Read	30
Number of Observations Used	30

Forward Selection: Step 1

Variable Displacement Entered: R-Square = 0.7601 and C(p) = 0.5804

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	864.52693	864.52693	88.70	<.0001
Error	28	272.90673	9.74667		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	33.48780	1.53711	4626.15275	474.64	<.0001
Displacement	-0.04706	0.00500	864.52693	88.70	<.0001

Bounds on condition number: 1, 1

Forward Selection: Step 2

Variable Torque Entered: R-Square = 0.7687 and C(p) = 1.6208

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	874.37944	437.18972	44.87	<.0001
Error	27	263.05422	9.74275		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	32.73315	1.71024	3568.97494	366.32	<.0001
Displacement	-0.08159	0.03470	53.85627	5.53	0.0263
Torque	0.04874	0.04847	9.85251	1.01	0.3235

Bounds on condition number: 48.26, 193.04

Forward Selection: Step 3

Variable carWeight Entered: R-Square = 0.7740 and C(p) = 3.0397

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	880.34556	293.44852	29.68	<.0001
Error	26	257.08811	9.88800		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	34.98822	3.37591	1062.11341	107.41	<.0001
Displacement	-0.06947	0.03828	32.55604	3.29	0.0812
Torque	0.04903	0.04883	9.96657	1.01	0.3246
carWeight	-0.00159	0.00205	5.96612	0.60	0.4443

Bounds on condition number: 57.875, 351.22

Forward Selection: Step 4

Variable carLength Entered: R-Square = 0.7852 and C(p) = 3.7957

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	893.11792	223.27948	22.85	<.0001
Error	25	244.31574	9.77263		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	22.30521	11.59064	36.19174	3.70	0.0658
Displacement	-0.04981	0.04176	13.90131	1.42	0.2442
Torque	0.03694	0.04968	5.40294	0.55	0.4641
carWeight	-0.00561	0.00406	18.63753	1.91	0.1795
carLength	0.12621	0.11040	12.77237	1.31	0.2638

Bounds on condition number: 69.687, 719.19

Forward Selection: Step 5

Variable transmissionSpeeds Entered: R-Square = 0.7919 and C(p) = 5.0524

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	900.74907	180.14981	18.27	<.0001
Error	24	236.68460	9.86186		
Corrected Total	29	1137.43367			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	14.06809	14.94167	8.74241	0.89	0.3558
Displacement	-0.03123	0.04697	4.36100	0.44	0.5124
Torque	0.02551	0.05157	2.41364	0.24	0.6253
transmissionSpeeds	1.30900	1.48808	7.63115	0.77	0.3878
carWeight	-0.00679	0.00430	24.63398	2.50	0.1271
carLength	0.15393	0.11529	17.57923	1.78	0.1944

Bounds on condition number: 87.343, 1048.5

No other variable met the 0.5000 significance level for entry into the model.

Summary of Forward Selection							
Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	Displacement	1	0.7601	0.7601	0.5804	88.70	<.0001
2	Torque	2	0.0087	0.7687	1.6208	1.01	0.3235
3	carWeight	3	0.0052	0.7740	3.0397	0.60	0.4443
4	carLength	4	0.0112	0.7852	3.7957	1.31	0.2638
5	transmissionSpeeds	5	0.0067	0.7919	5.0524	0.77	0.3878

In forward selection, the model contained five predictors: Displacement, Torque, transmissionSpeeds, carWeight, and carLength. The model explains about 79% of gas mileage variation ($R^2 = 0.7919$) and appears to fit well by C(p). However, none of the predictors is individually statistically significant at the 0.05 level, most likely a result of multicollinearity between the predictors. No additional variables met the entry criterion, so the model stopped at five predictors.

Q7(a)

```
PROC LOGISTIC DATA=BTT;
CLASS sex (REF="1") / PARAM=REF;
MODEL sex(EVENT="2") = bweight gestage parity;
TITLE "Logistic Regression: Predicting Probability of Being a Girl";
RUN;
```

Logistic Regression: Predicting Probability of Being a Girl

The LOGISTIC Procedure

Model Information	
Data Set	WORK.BTT
Response Variable	sex
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	217
Number of Observations Used	217

Response Profile		
Ordered Value	sex	Total Frequency
1	1	114
2	2	103

Probability modeled is sex=2.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	302.268	299.969
SC	305.648	313.489
-2 Log L	300.268	291.969

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	8.2991	3	0.0402
Score	8.1041	3	0.0439
Wald	7.7930	3	0.0505

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.7936	3.9644	0.4966	0.4810
bweight	1	-0.00058	0.000328	3.1674	0.0751
gestage	1	0.1255	0.1095	1.3143	0.2516
parity	1	-0.2762	0.1285	4.6218	0.0316

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
bweight	0.999	0.999	1.000
gestage	1.134	0.915	1.405
parity	0.759	0.590	0.976

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	60.4	Somers' D	0.209
Percent Discordant	39.5	Gamma	0.209
Percent Tied	0.1	Tau-a	0.105
Pairs	11742	c	0.605

The probability of child being a girl is:

$$\hat{\pi} = \frac{e^{-2.79-0.00058\times bweight+0.13\times gestage-0.28\times parity}}{1+e^{-2.79-0.00058\times bweight+0.13\times gestage-0.28\times parity}} \tag{1}$$

Parity is a statistically significant predictor of a child being a girl, with higher parity associated with lower probability of having a girl. Birthweight shows a marginal negative relationship with the probability of being female, while gestational age does not appear to be a meaningful predictor. The model suggests that families with more previous births are less likely to have girls, independent of birthweight and gestational age.

Q7(b)

Stepwise procedure

```
PROC LOGISTIC DATA=BTT;
CLASS mmedaid (REF="2");
MODEL mmedaid (EVENT = "1") = bweight gestage momage parity mdbp msbp / SELECTION=
    STEPWISE;
TITLE "Model Building for Probability of Mother Having Medical Aid";
RUN;
```

Model Building for Probability of Mother Having Medical Aid

The LOGISTIC Procedure

Model Information	
Data Set	WORK.BTT
Response Variable	mmedaid
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	217
Number of Observations Used	217

Response Profile		
Ordered Value	mmedaid	Total Frequency
1	1	16
2	2	201

Probability modeled is mmedaid=1.

Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L	=	114.224

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
7.8940	6	0.2460

Step 1. Effect momage entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	116.224	114.493
SC	119.604	121.253
-2 Log L	114.224	110.493

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	3.7311	1	0.0534
Score	3.9714	1	0.0463
Wald	3.8114	1	0.0509

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
3.7225	5	0.5900

Step 2. Effect momage is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L	=	114.224

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
7.8940	6	0.2460

Note: Model building terminates because the last effect entered is removed by the Wald statistic criterion.

Summary of Stepwise Selection

Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	momage		1	1	3.9714		0.0463
2		momage	1	0		3.8114	0.0509

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.5307	0.2598	94.9168	<.0001

The stepwise selection procedure did not identify any statistically significant predictors of medical aid coverage among the variables tested (birth weight, gestational age, mother's age, parity, diastolic and systolic blood pressure). The final model contains only the intercept, indicating that none of these maternal or birth characteristics reliably predict whether a mother has medical aid coverage at childbirth.

The probability of a mother having a medical aid is:

$$\pi = \frac{e^{-2.5307}}{1 + e^{-2.5307}} \approx 0.074 \tag{2}$$

The estimated probability of having medical aid is about 7.4%, regardless of predictors.

Backward procedure

```
PROC LOGISTIC DATA=BTT;
CLASS mmedaid (REF="2");
MODEL mmedaid (EVENT = "1") = bweight gestage momage parity mdbp msbp / SELECTION=
    BACKWARD;
TITLE "Model Building for Probability of Mother Having Medical Aid";
RUN;
```

Model Building for Probability of Mother Having Medical Aid

The LOGISTIC Procedure

Model Information	
Data Set	WORK.BTT
Response Variable	mmedaid
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	217
Number of Observations Used	217

Response Profile		
Ordered Value	mmedaid	Total Frequency
1	1	16
2	2	201

Probability modeled is mmedaid=1.

Backward Elimination Procedure

Step 0. The following effects were entered:

Intercept bweight gestage momage parity mdbp msbp

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	116.224	121.011
SC	119.604	144.670
-2 Log L	114.224	107.011

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.2131	6	0.3016
Score	7.8940	6	0.2460
Wald	7.3195	6	0.2923

Step 1. Effect mdbp is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	116.224	119.023
SC	119.604	139.303
-2 Log L	114.224	107.023

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.2006	5	0.2061
Score	7.8908	5	0.1624
Wald	7.3175	5	0.1981

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
0.0125	1	0.9110

Step 2. Effect parity is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	116.224	117.071
SC	119.604	133.971
-2 Log L	114.224	107.071

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.1526	4	0.1280
Score	7.7207	4	0.1024
Wald	7.2249	4	0.1245

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
0.0605	2	0.9702

Step 3. Effect msbp is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	116.224	115.273
SC	119.604	128.793
-2 Log L	114.224	107.273

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.9506	3	0.0735
Score	7.3338	3	0.0620
Wald	6.9343	3	0.0740

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
0.2653	3	0.9664

Step 4. Effect bweight is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	116.224	113.733
SC	119.604	123.873
-2 Log L	114.224	107.733

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	6.4912	2	0.0389
Score	6.6189	2	0.0365
Wald	6.3616	2	0.0416

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
0.7259	4	0.9481

Step 5. Effect gestage is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	116.224	114.493
SC	119.604	121.253
-2 Log L	114.224	110.493

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	3.7311	1	0.0534
Score	3.9714	1	0.0463
Wald	3.8114	1	0.0509

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
3.7225	5	0.5900

Step 6. Effect momage is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L	=	114.224

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
7.8940	6	0.2460

Note: All effects have been removed from the model.

Summary of Backward Elimination					
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq
1	mdbp	1	5	0.0125	0.9110
2	parity	1	4	0.0481	0.8264
3	msbp	1	3	0.2015	0.6536
4	bweight	1	2	0.4548	0.5000
5	gestage	1	1	2.7813	0.0954
6	momage	1	0	3.8114	0.0509

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.5307	0.2598	94.9168	<.0001

Forward procedure

```
PROC LOGISTIC DATA=BTT;
CLASS mmedaid (REF="2");
MODEL mmedaid (EVENT = "1") = bweight gestage momage parity mdbp msbp / SELECTION=
    FORWARD;
TITLE "Model Building for Probability of Mother Having Medical Aid";
RUN;
```

Model Building for Probability of Mother Having Medical Aid

The LOGISTIC Procedure

Model Information	
Data Set	WORK.BTT
Response Variable	mmedaid
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	217
Number of Observations Used	217

Response Profile		
Ordered Value	mmedaid	Total Frequency
1	1	16
2	2	201

Probability modeled is mmedaid=1.

Forward Selection Procedure

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L	=	114.224

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
7.8940	6	0.2460

Step 1. Effect momage entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	116.224	114.493
SC	119.604	121.253
-2 Log L	114.224	110.493

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	3.7311	1	0.0534
Score	3.9714	1	0.0463
Wald	3.8114	1	0.0509

Residual Chi-Square Test		
Chi-Square	**DF**	**Pr > ChiSq**
3.7225	5	0.5900

Note: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Forward Selection					
Step	**Effect Entered**	**DF**	**Number In**	**Score Chi-Square**	**Pr > ChiSq**
1	**momage**	1	1	3.9714	0.0463

Analysis of Maximum Likelihood Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-4.6677	1.1768	15.7325	<.0001
momage	1	0.0820	0.0420	3.8114	0.0509

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
momage	1.085	1.000	1.179

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	63.4	Somers' D	0.306
Percent Discordant	32.8	Gamma	0.318
Percent Tied	3.8	Tau-a	0.042
Pairs	3216	c	0.653