\*9.4;

**data** hands;

infile '\\apporto.com\dfs\LOYOLA\23SPStatisticalComputingST710W01\Elliott\handinj.dat';

input injury $**7**-**11** lostwork **13**-**14** cost **16**-**19**;

**run**;

\*plot cost vs days lost;

**proc** **gplot** data=hands;

plot lostwork\*cost;

title Plot of work days lost vs cost;

**run**;

9.4 plot discussion: The plot shows that generally, as the number of days lost increases, the cost also increases. In come cases the cost varies per day; this could be due to the type of injury sustained.

\*run a correlation test on cost vs days lost;

**proc** **corr** data=hands;

var lostwork;

with cost;

title Correlation of work days lost vs cost;

**run**;

9.4 correlation discussion: There is a strong correlation (0.908) between the cost and number of workdays lost. Because the probability is less than 5% (.005%), we can reject the null hypothesis and conclude that the relationship is statistically significant.

\*15.1;

**data** electric;

infile '\\apporto.com\dfs\LOYOLA\23SPStatisticalComputingST710W01\Elliott\electric.dat';

input sqft **1**-**3** income **6**-**11** accap **14**-**16** appliancecap **19**-**23** famnum **26**-**28** peakload **31**-**35**;

**run**;

\*test H0 of rho=0;

**proc** **corr** data=electric nosimple;

var sqft appliancecap famnum income;

title Correlation of house size, appliance index, family size, and income;

**run**;

15.1a discussion of house size vs appliance index: Because the probability is less than 5% (.005%), we can reject the null hypothesis and we can conclude that the relationship is statistically significant.

15.1b discussion of family size vs appliance index: Because the probability is greater than 5% (17.74%), we cannot reject the null hypothesis, but we can conclude that the relationship is NOT statistically significant.

15.1c discussion of house size vs income: Because the probability is less than 5% (.005%), we can reject the null hypothesis and we can conclude that the relationship is statistically significant.

\*test corr of peak load vs air capacity;

**proc** **corr** data=electric nosimple;

var peakload;

with accap;

title Correlation of peak loak vs air conditioning capacity;

**run**;

15.1 peak load vs air capacity correlation discussion: There is a strong correlation (0.92727) between the peak load and the air capacity. Because the probability is less than 5% (.005%), we can reject the null hypothesis and we can conclude that the relationship is statistically significant.

\*16.2;

**proc** **reg** data=electric;

model peakload = accap;

plot peakload\*accap;

plot r.\*p.;

plot r.\*accap;

output out=electric\_stats p=yhat r=resid student=std\_resid;

title Estimated regression lines of peak load with air capacity as the explanatory variable;

**run**;

**quit**;

16.2 discussion: Based off the graphs, the line does fit the data well.

\*17.2 check for normality for 16.2;

**proc** **univariate** data = electric\_stats;

var resid;

histogram resid/normal;

probplot resid/normal (mu=est sigma=est) square; \*check module17 code output;

title normality check for residuals;

**run**;

17.2 discussion: The test on residuals for normality affirms that the distribution is normal as the p-value is 50%. The fit diagnostics shows that the distribution of residuals is random. They are close to the line, and the histogram shows a normal curve. It can be concluded that as the ac capacity increases, the peak load will increase positively.