

Assignment 4: Correlation

Sukanya Mukherjee

128041217

Statistics for Analytics

BAN100ZBB

Samaneh Gholami



Given Data

- **Sample:** Fitness data Various measures of heart and pulse rate were taken on men in a physical fitness course at N.C. State Univ.
- **age:** age of the man
- **weight:** weight of the man
- **runtim:** running time of the man
- **runpulse:** pulse rate of the man while running
- **maxpulse:** maximum pulse rate of the man
- **rstpulse:** Resting pulse rate of the man
- **oxy:** oxygen consumption of the man
- **Objective:**
 - Describe the relationships between Variables.
 - Generate scatter plots and correlations for the variables.
 - Interpret the results.
 - Also check the limitations of correlation coefficient for the Fitness data.

Given Code



```
data fitness;
  input age weight oxy runtime rstpulse runpulse maxpulse;
  case = _n_;
datalines;
44 89.47 44.609 11.37 62 178 182
40 75.07 45.313 10.07 62 185 185
44 85.84 54.297 8.65 45 156 168
42 68.15 59.571 8.17 40 166 172
38 89.02 49.874 9.22 55 178 180
47 77.45 44.811 11.63 58 176 176
40 75.98 45.681 11.95 70 176 180
43 81.19 49.091 10.85 64 162 170
44 81.42 39.442 13.08 63 174 176
38 81.87 60.055 8.63 48 170 186
44 73.03 50.541 10.13 45 168 168
45 87.66 37.388 14.03 56 186 192
45 66.45 44.754 11.12 51 176 176
47 79.15 47.273 10.60 47 162 164
54 83.12 51.855 10.33 50 166 170
49 81.42 49.156 8.95 44 180 185
```

```
51 69.63 40.836 10.95 57 168 172
51 77.91 46.672 10.00 48 162 168
48 91.63 46.774 10.25 48 162 164
49 73.3750.388 10.08 67 168 168
57 73.37 39.407 12.63 58 174 176
54 79.38 46.080 11.17 62 156 165
52 76.32 45.441 9.63 48 164 166
50 70.87 54.625 8.92 48 146 155
51 67.25 45.118 11.08 48 172 172
54 91.63 39.203 12.88 44 168 172
51 73.71 45.790 10.47 59 186 188
57 59.08 50.545 9.93 49 148 155
49 76.32 48.673 9.40 56 186 188
48 61.24 47.920 11.50 52 170 176
52 82.78 47.467 10.50 53 170 172
;
run;
```

```
proc contents data=fitness varnum;  
run;
```

Contents of
the dataset

Variables in Creation Order

#	Variable	Type	Len
1	age	Num	8
2	weight	Num	8
3	oxy	Num	8
4	runtime	Num	8
5	rstpulse	Num	8
6	runpulse	Num	8
7	maxpulse	Num	8
8	case	Num	8

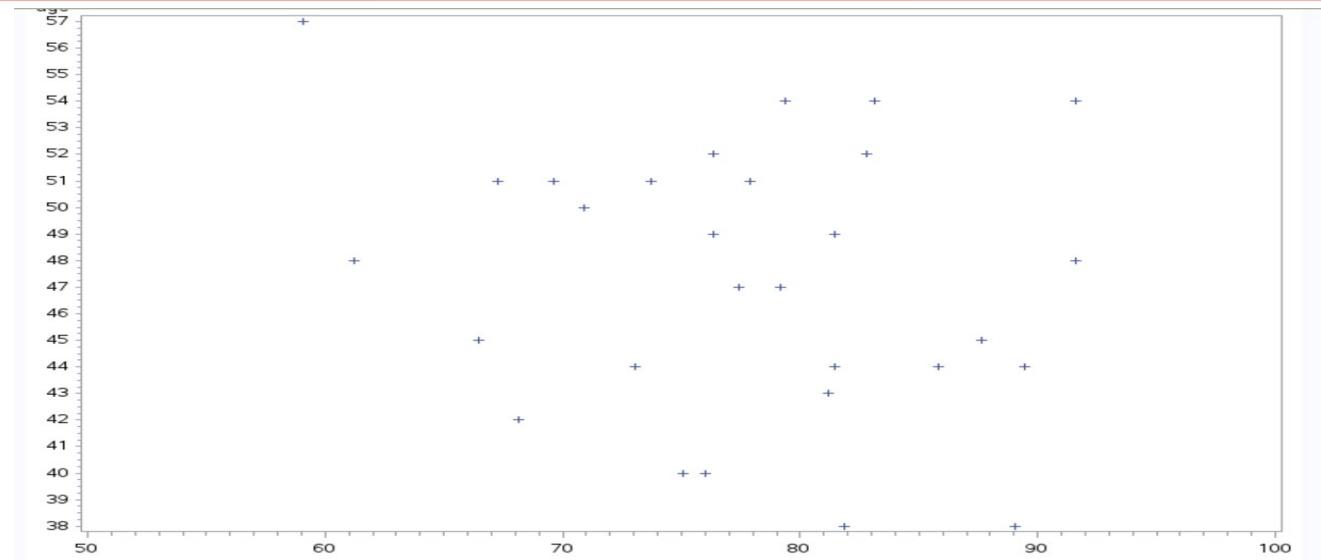
Pearson Test of Independence Hypothesis (Age and Weight)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var age weight;  
run;  
proc gplot data=fitness;  
plot age * weight;  
run;
```

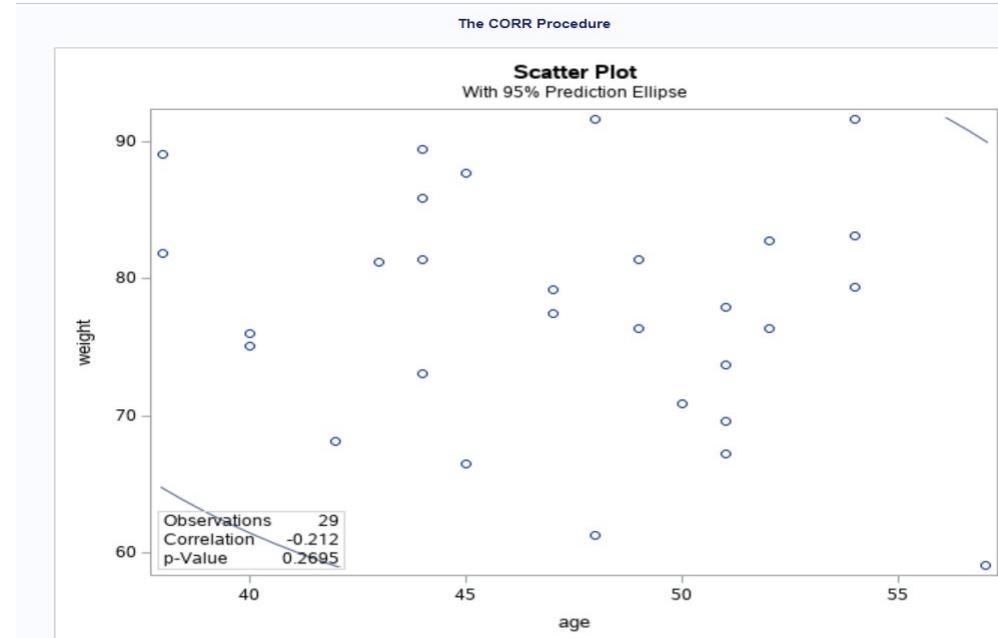
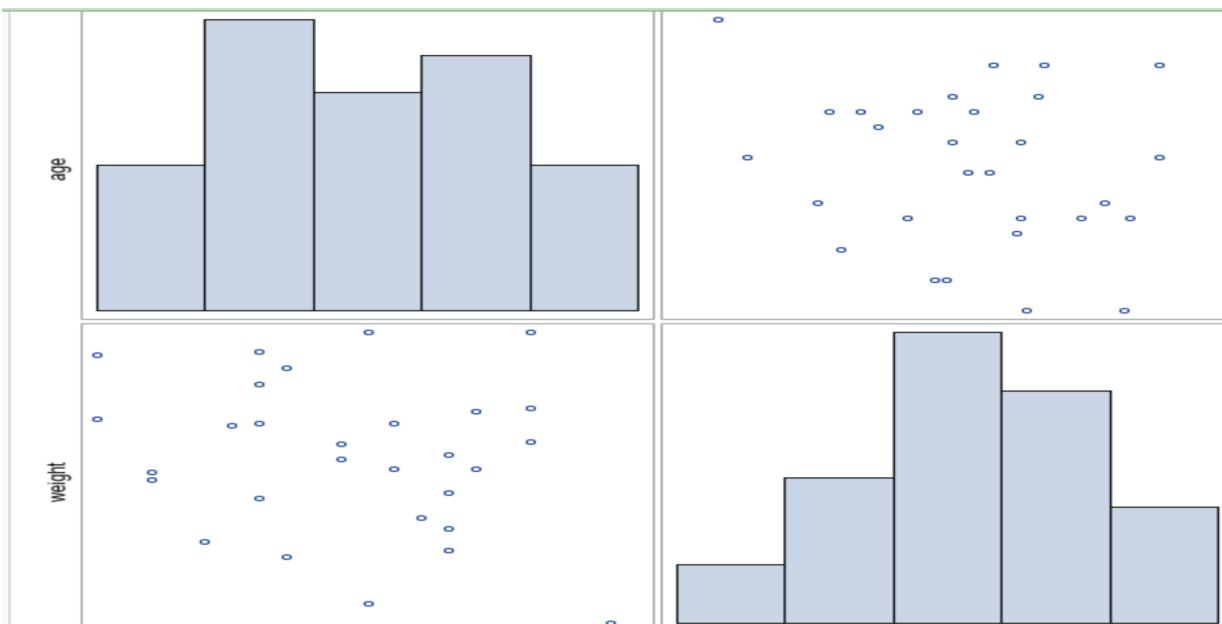
Ho (Null Hypothesis):
Age is independent of Weight

H1 (Alternate Hypothesis): Age is not independent of Weight

Pearson Correlation Coefficients, N = 29		
	age	weight
age	1.00000	-0.21205 0.2695
weight	-0.21205 0.2695	1.00000



Scatter Plot Matrix



Conclusion

Observations	29
Correlation	-0.212
p-Value	0.2695

- Since p-value (0.2) > alpha (0.05), we fail to reject null hypothesis.
- This means Age is independent of Weight.
- Since correlation is -0.21 (between 0 and -0.25) , this means there is an insignificant negative association.

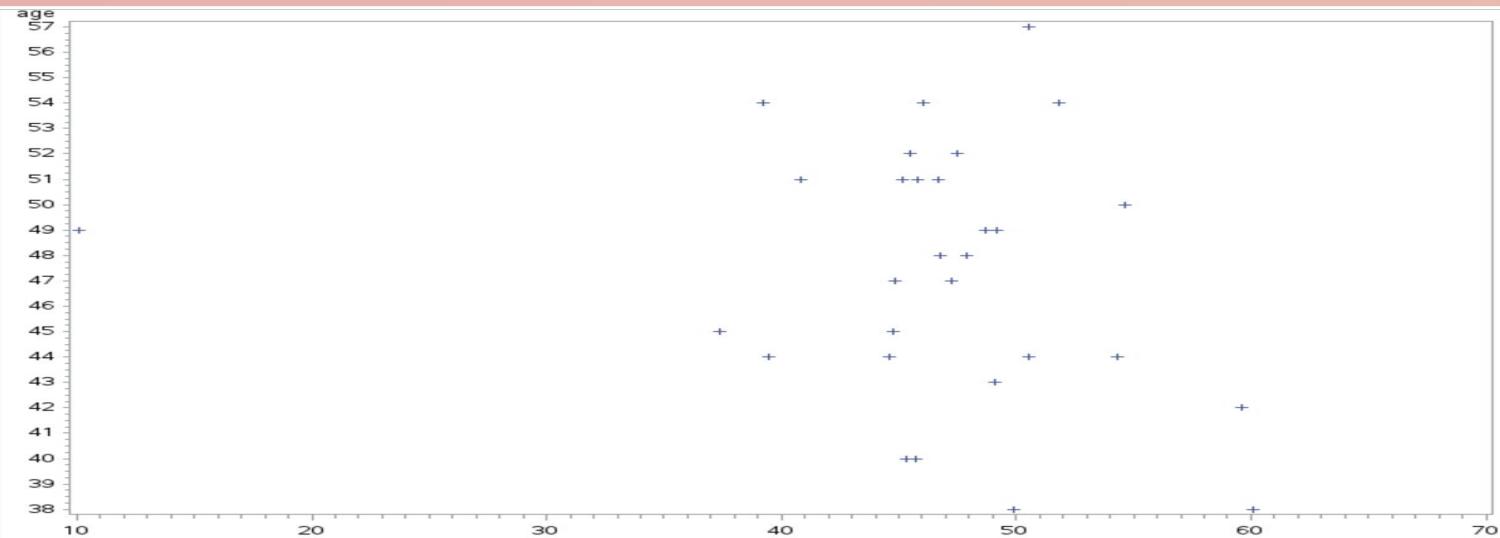
Pearson Test of Independence Hypothesis (Age and Oxy)

```
proc corr data=fitness nomiss  
  plots=matrix(histogram)  
  plots=scatter(nvar=2 alpha=0.05);  
  var age oxy;  
run;  
proc gplot data=fitness;  
  plot age * oxy;  
run;
```

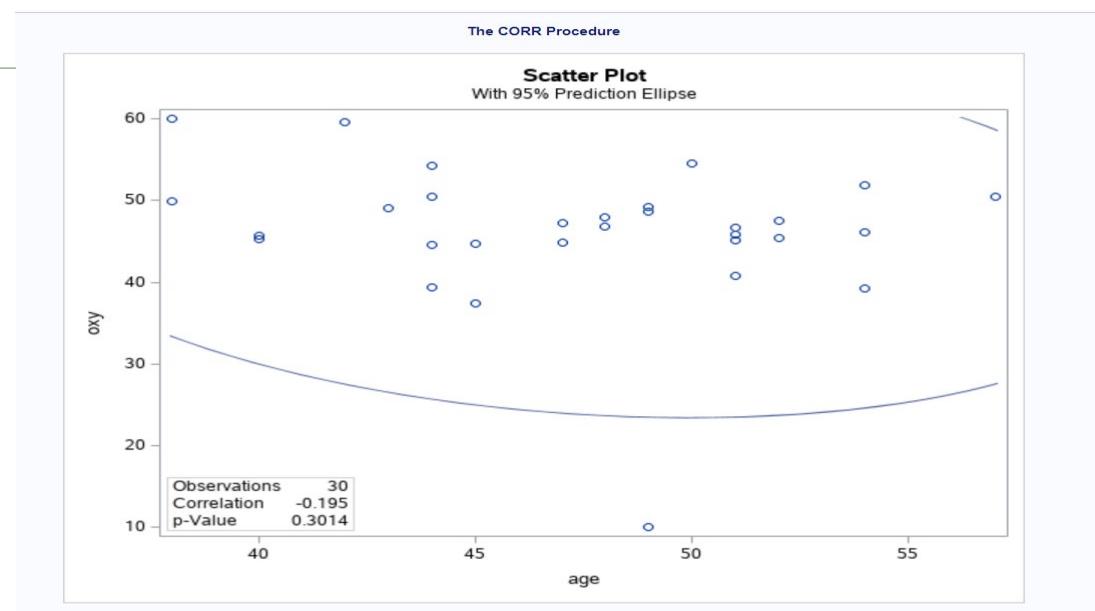
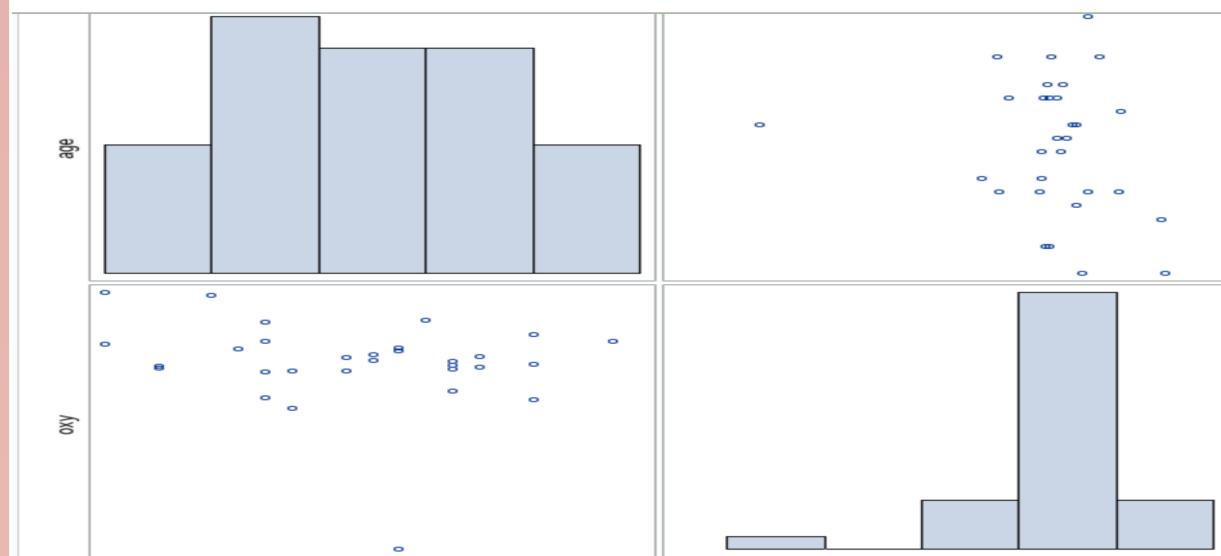
Ho (Null Hypothesis):
Age is independent of Oxy

H1 (Alternate Hypothesis): Age is not independent of Oxy

Pearson Correlation Coefficients, N = 30		
	age	oxy
age	1.00000	-0.19513 0.3014
oxy	-0.19513 0.3014	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.195
p-Value	0.3014

- Since p-value (0.3) > alpha (0.05), we fail to reject null hypothesis.
- This means Age is independent of Oxy.
- Since correlation is -0.19 (between 0 and -0.25) , this means there is an insignificant negative association.

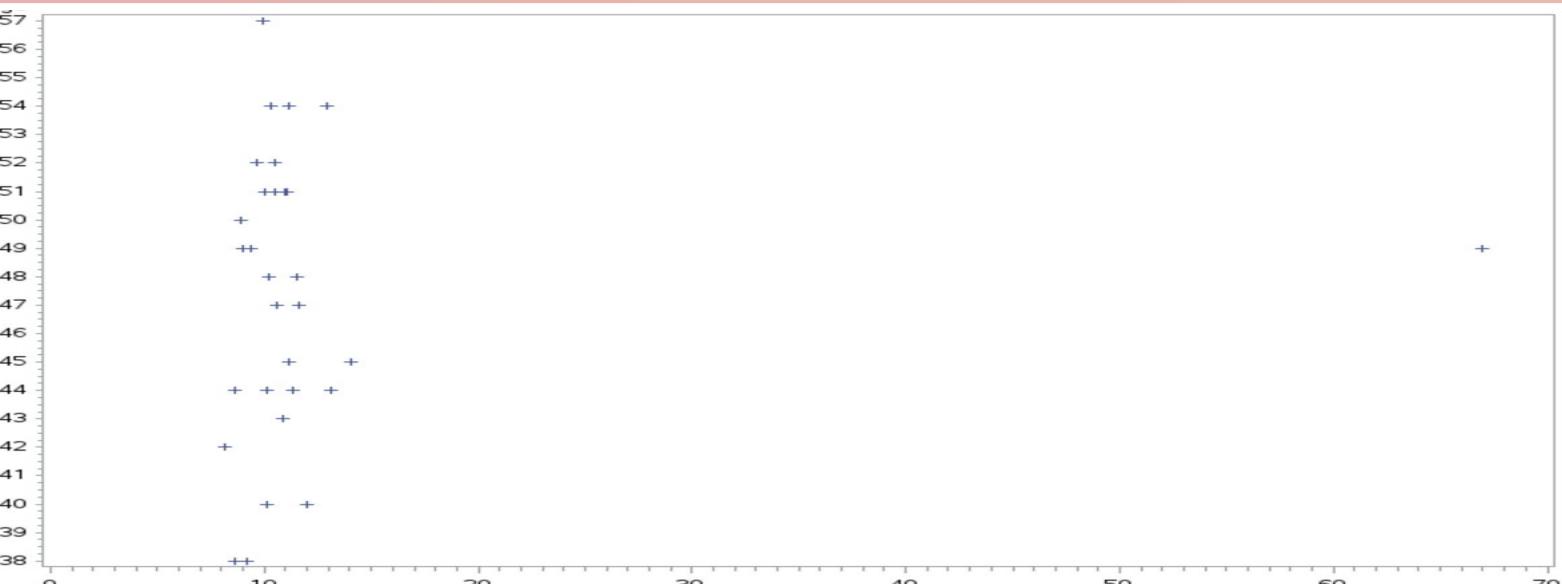
Pearson Test of Independence Hypothesis (Age and Runtime)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var age runtime;  
run;  
proc gplot data=fitness;  
plot age * runtime;  
run;
```

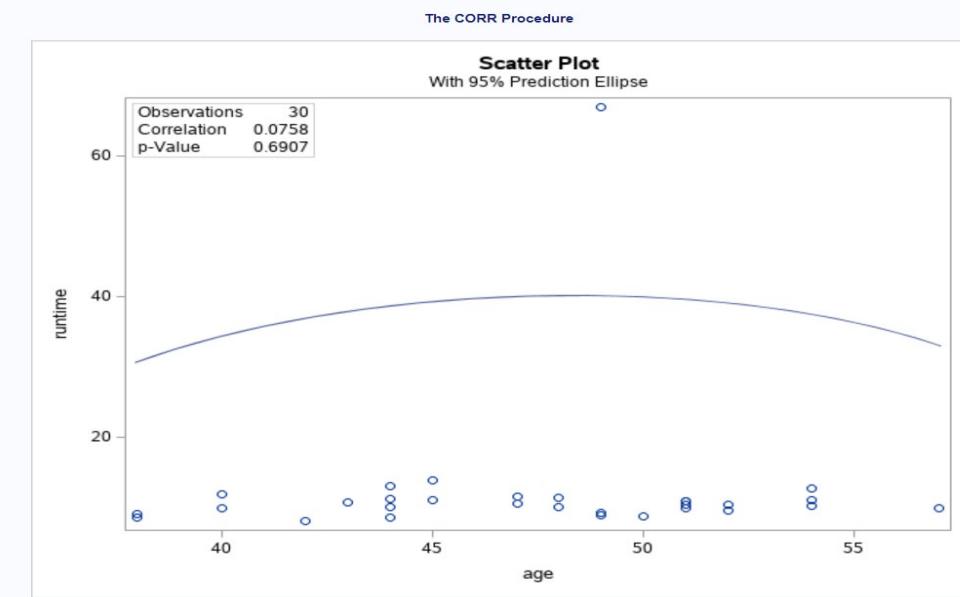
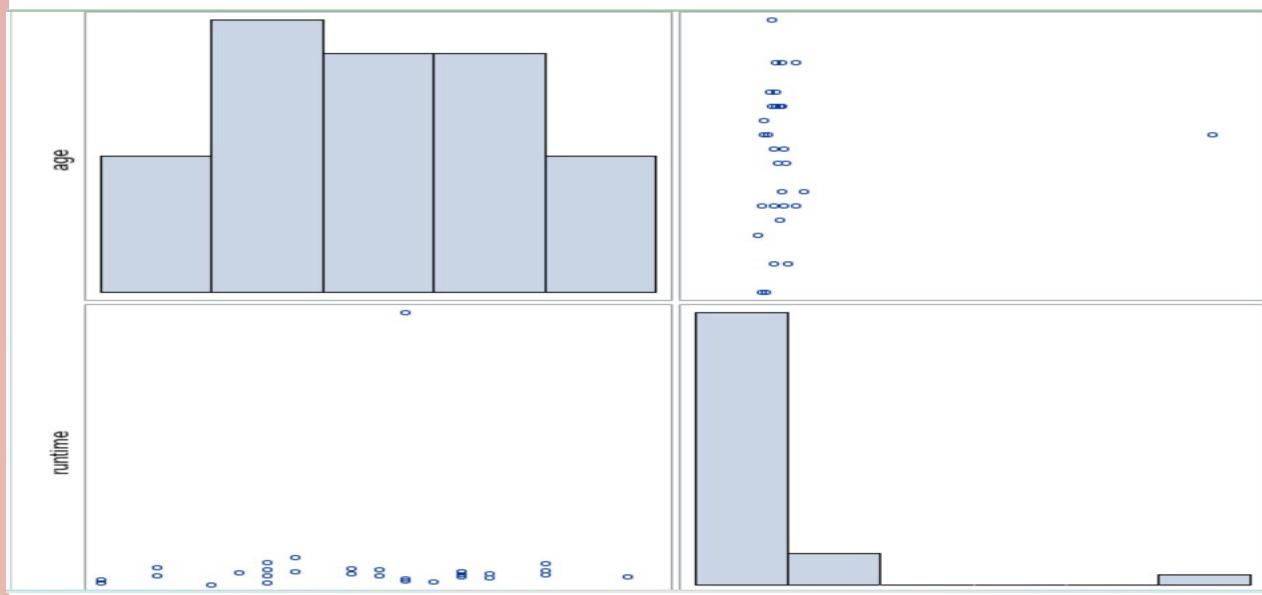
Ho (Null Hypothesis):
Age is independent of Runtime

H1 (Alternate Hypothesis): Age is not independent of Runtime

Pearson Correlation Coefficients, N = 30		
	age	runtime
age	1.00000	0.07576 0.6907
runtime	0.07576 0.6907	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	0.0758
p-Value	0.6907

- Since p-value ($0.69 > 0.05$), we fail to reject null hypothesis.
- This means Age is independent of Runtime.
- Since correlation is 0.07 (between 0 and 0.25) , this means there is a negligible positive association.

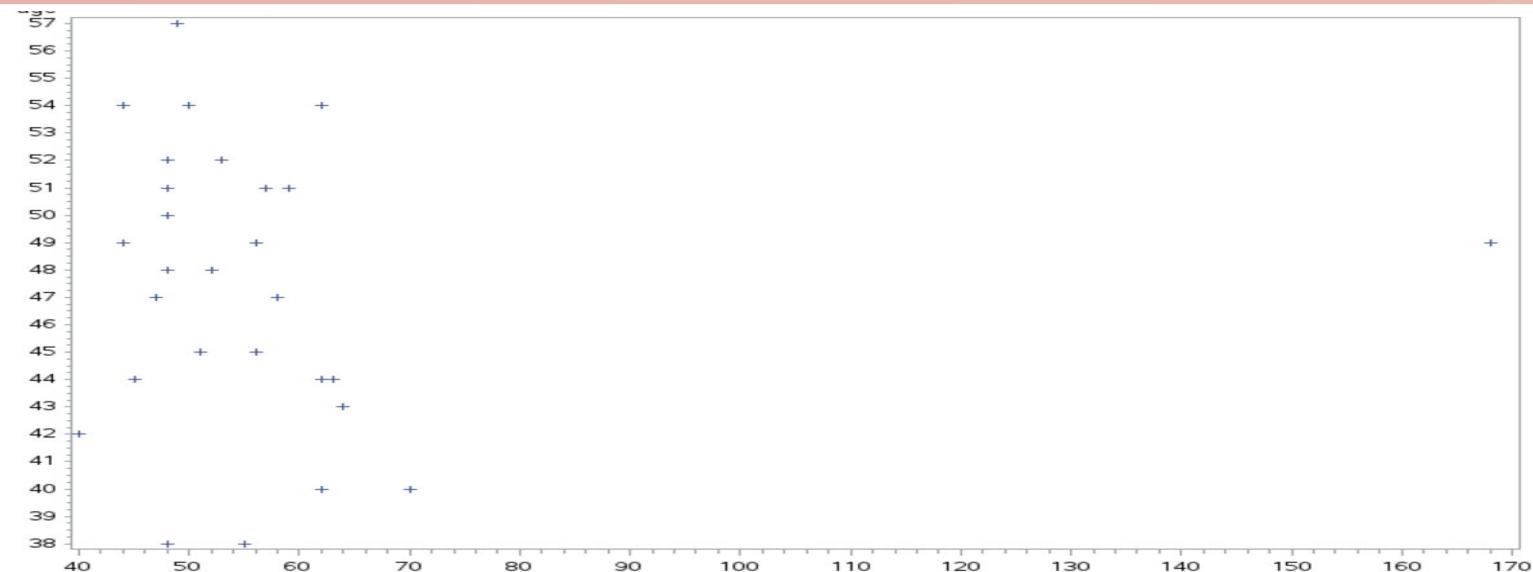
Pearson Test of Independence Hypothesis (Age and Rstpulse)

```
proc corr data=fitness nomiss  
  plots=matrix(histogram)  
  plots=scatter(nvar=2 alpha=0.05);  
  var age rstpulse;  
run;  
proc gplot data=fitness;  
  plot age * rstpulse;  
run;
```

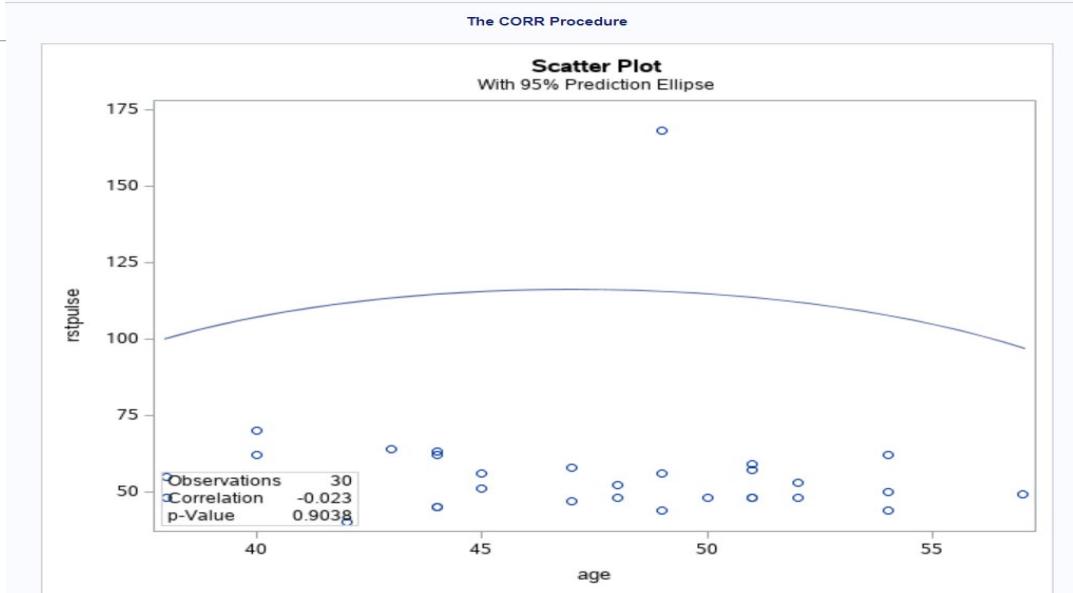
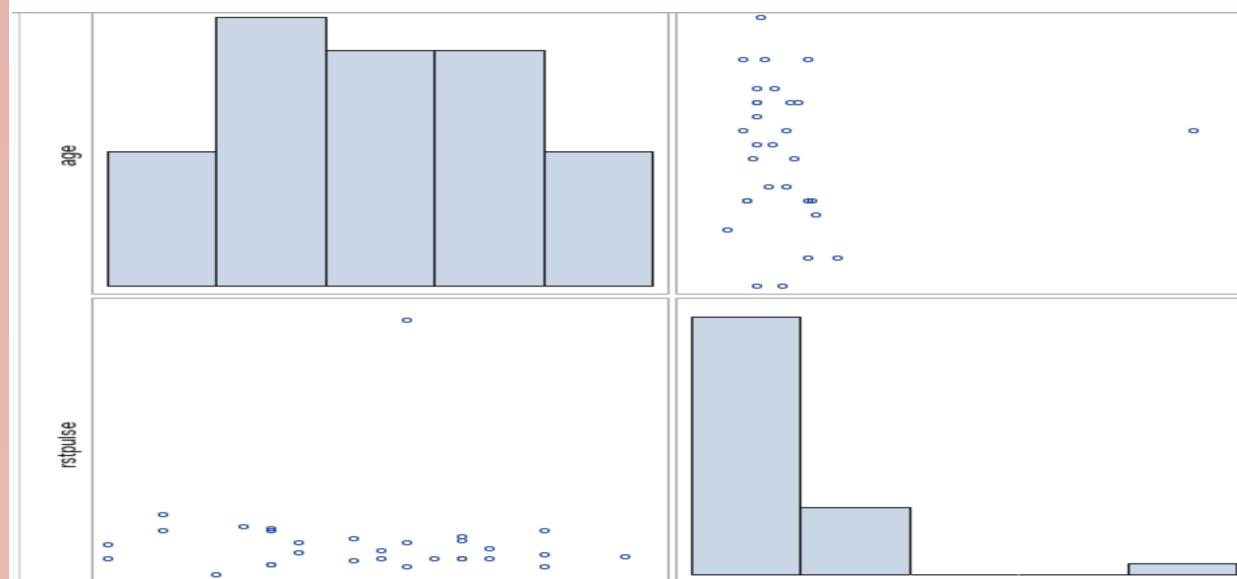
Ho (Null Hypothesis):
Age is independent of
Rstpulse

H1 (Alternate Hypothesis): Age is not independent of Rstpulse

Pearson Correlation Coefficients, N = 30		
	age	rstpulse
age	1.00000	-0.02305 0.9038
rstpulse	-0.02305 0.9038	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.023
p-Value	0.9038

- Since p-value ($0.90 > 0.05$), we fail to reject null hypothesis.
- This means Age is independent of Rstpulse.
- Since correlation is -0.02 (between 0 and -0.25) , this means there is an insignificant negative association.

Pearson Test of Independence Hypothesis (Age and Runpulse)

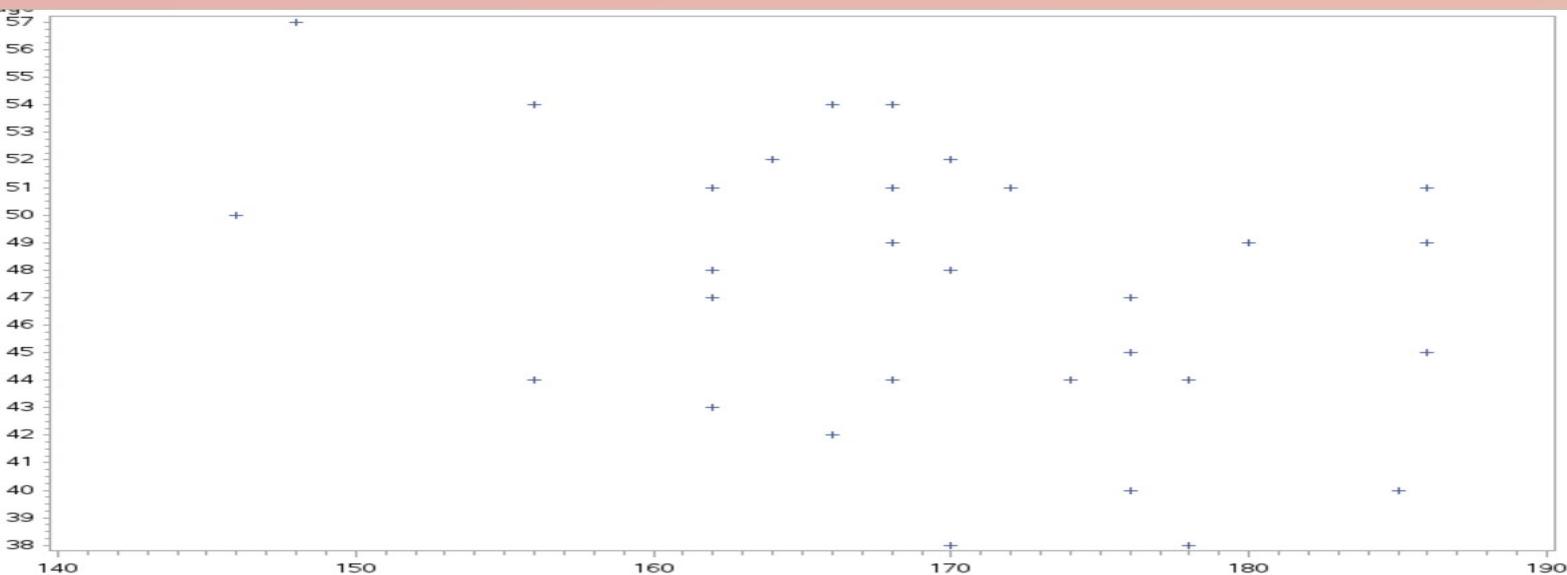
```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var age runpulse;  
run;  
proc gplot data=fitness;  
plot age * runpulse;  
run;
```

Ho (Null Hypothesis):
Age is independent of Runpulse

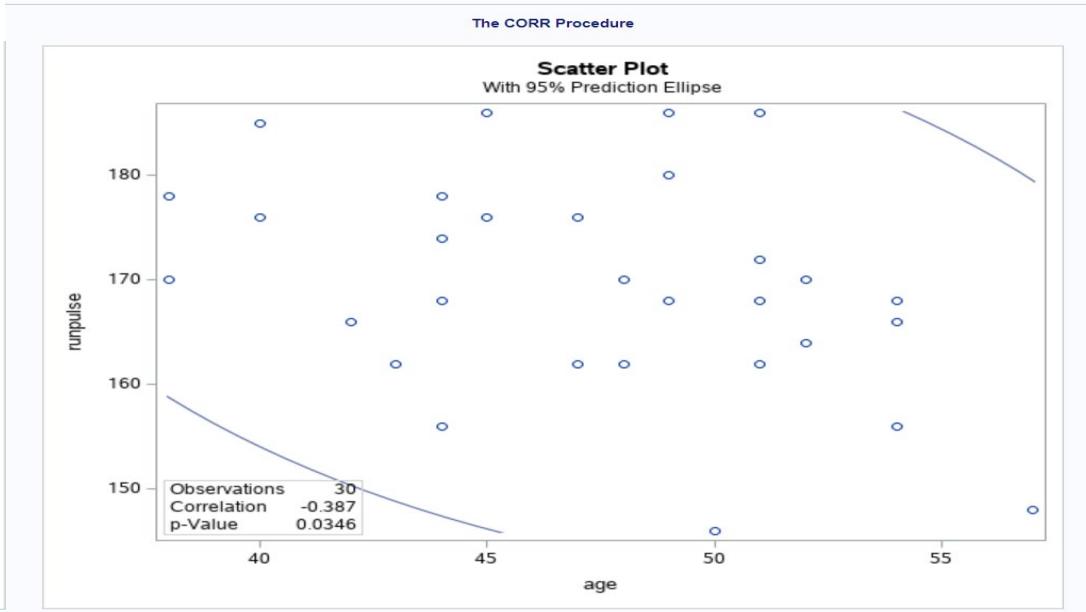
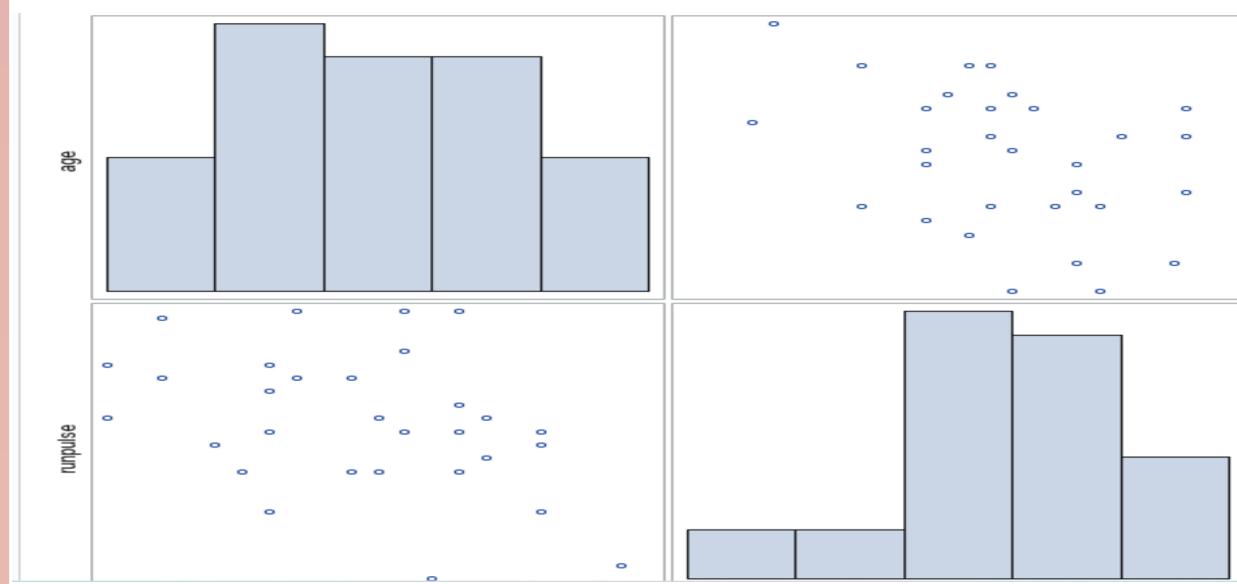
H1 (Alternate Hypothesis): Age is not independent of Runpulse

Pearson Correlation Coefficients, N = 30
 Prob > |r| under H0: Rho=0

	age	runpulse
age	1.00000	-0.38714 0.0346
runpulse	-0.38714 0.0346	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.387
p-Value	0.0346

- Since p-value (0.03) < alpha (0.05), we fail to reject null hypothesis.
- This means Age is not independent of Runpulse.
- Since correlation is -0.38 (between -0.25 and -0.5), this means there is a weak negative association.

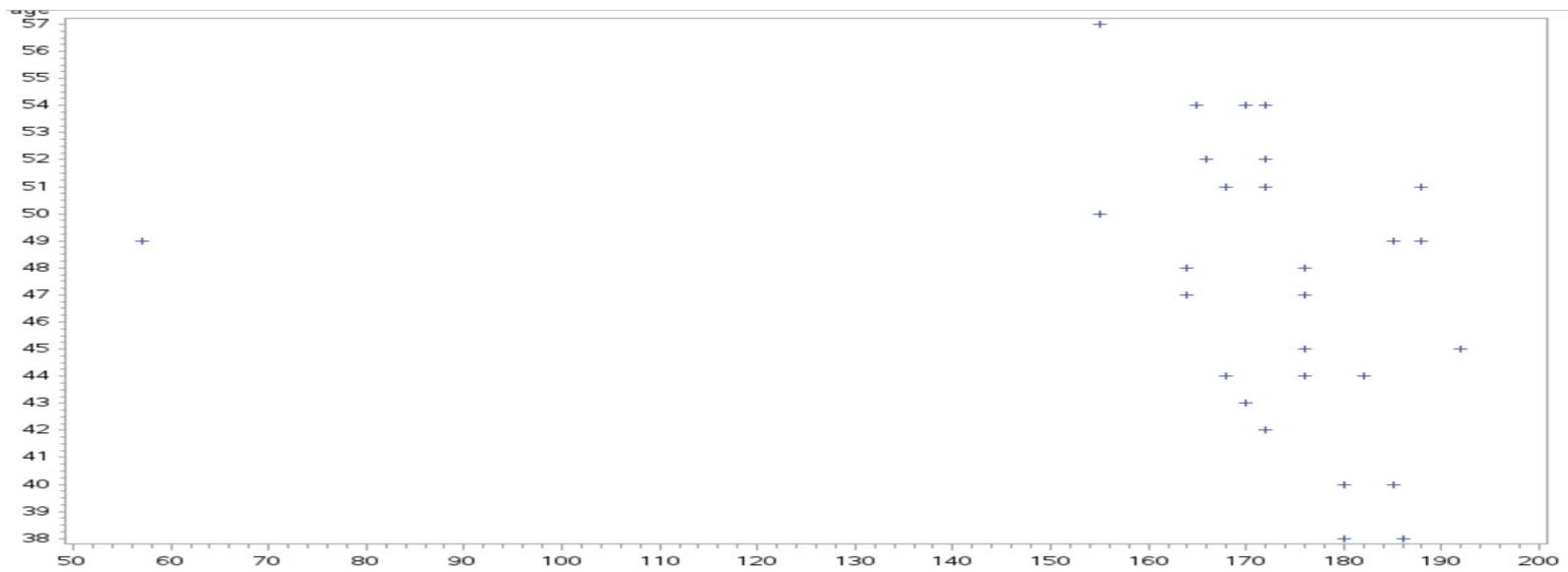
Pearson Test of Independence Hypothesis (Age and Maxpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var age maxpulse;  
run;  
proc gplot data=fitness;  
plot age * maxpulse;  
run;
```

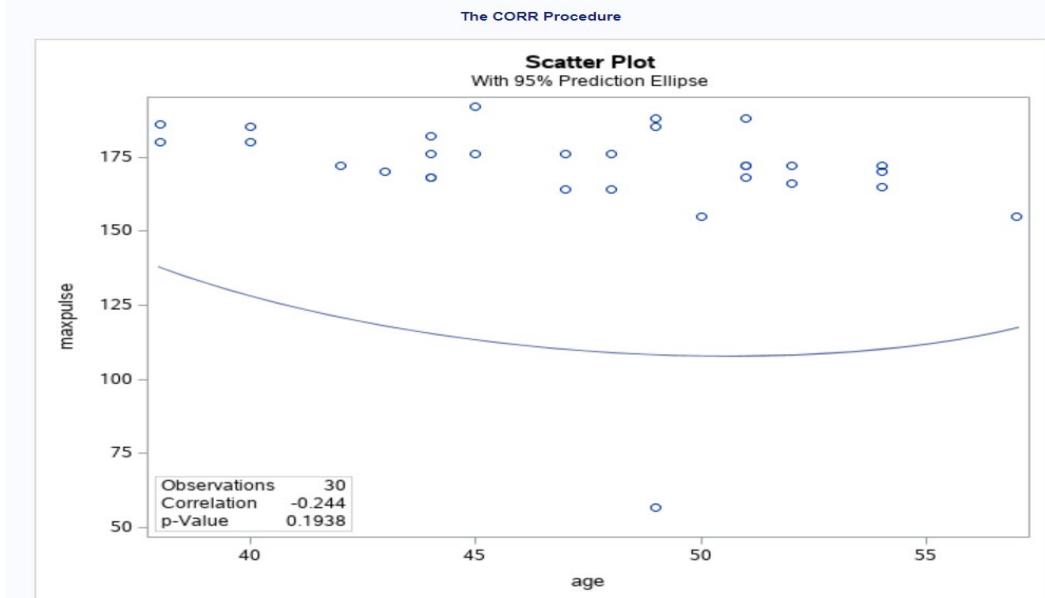
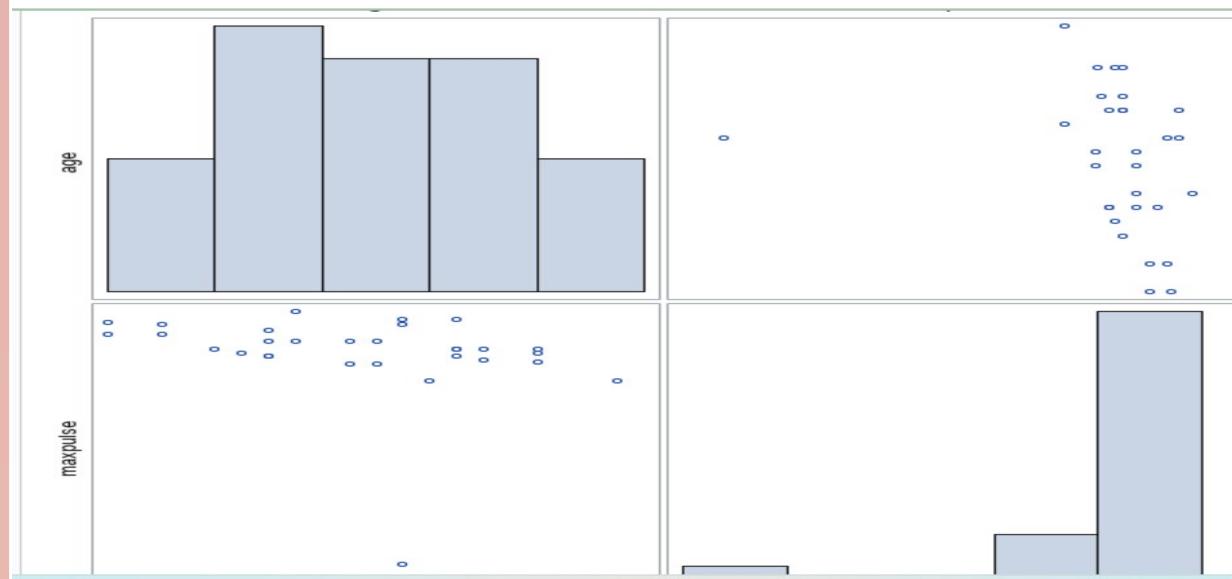
Ho (Null Hypothesis): Age is independent of Maxpulse

H1 (Alternate Hypothesis): Age is not independent of Maxpulse

Pearson Correlation Coefficients, N = 30		
Prob > r under H0: Rho=0		
	age	maxpulse
age	1.00000	-0.24402 0.1938
maxpulse	-0.24402 0.1938	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.244
p-Value	0.1938

- Since p-value (0.19) > alpha (0.05), we fail to reject null hypothesis.
- This means Age is independent of Maxpulse.
- Since correlation is -0.24 (between 0 and -0.25) , this means there is an insignificant negative association.

Pearson Test of Independence Hypothesis (Weight and Oxy)

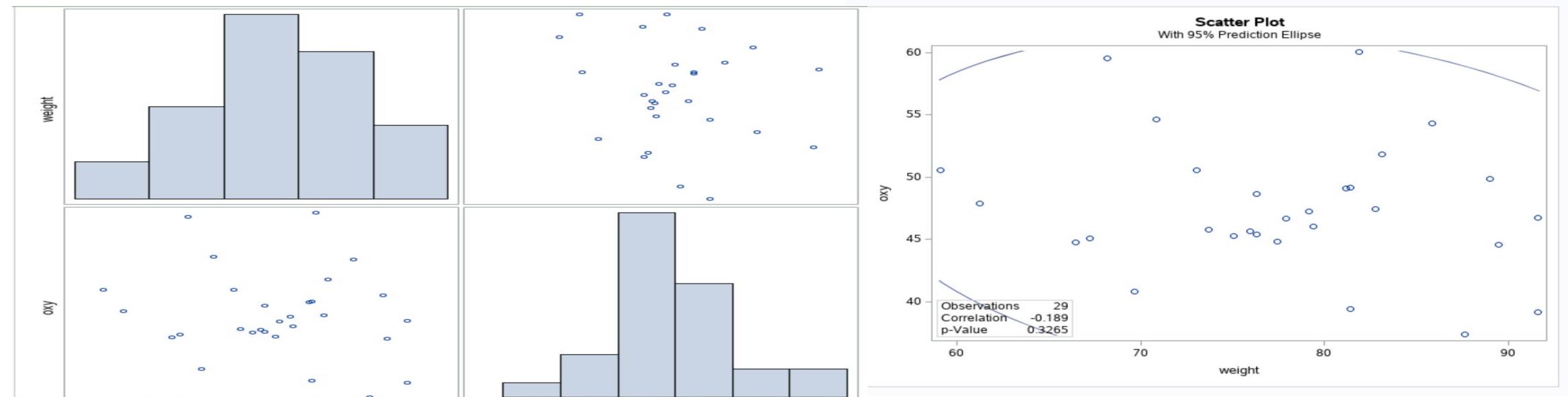
```
proc corr data=fitness nomiss  
    plots=matrix(histogram)  
    plots=scatter(nvar=2 alpha=0.05);  
    var weight oxy;  
run;  
  
proc gplot data=fitness;  
plot weight * oxy;  
run;
```

Ho (Null Hypothesis):
Weight is independent
of Oxy

H1 (Alternate
Hypothesis): Weight is
not independent of Oxy

Pearson Correlation Coefficients, N = 29		
Prob > r under H0: Rho=0		
	weight	oxy
weight	1.00000	-0.18884 0.3265
oxy	-0.18884 0.3265	1.00000

Scatter Plot Matrix



Conclusion

Observations	29
Correlation	-0.189
p-Value	0.3265

- Since p-value (0.32) > alpha (0.05), we fail to reject null hypothesis.
- This means Weight is independent of Oxy.
- Since correlation is -0.18 (between 0 and -0.25) , this means there is an insignificant negative association.

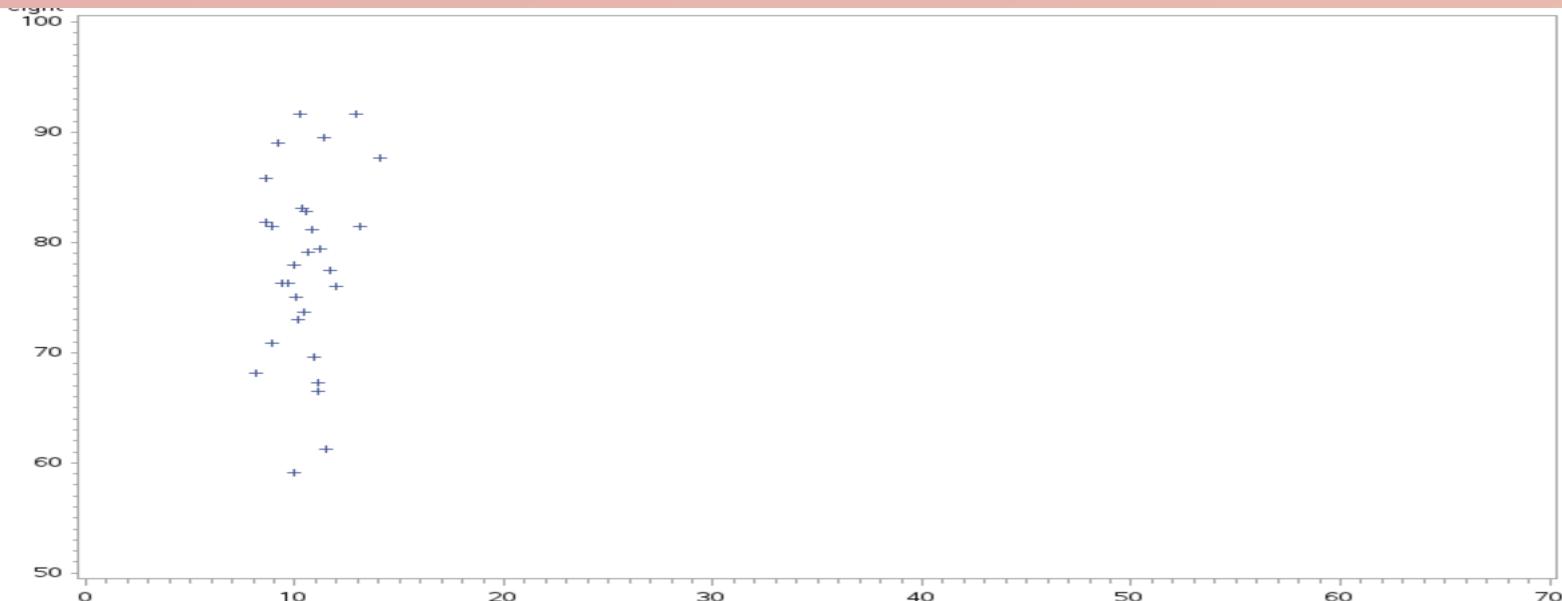
Pearson Test of Independence Hypothesis (Weight and Runtime)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var weight runtime;  
run;  
  
proc gplot data=fitness;  
plot weight * runtime;  
run;
```

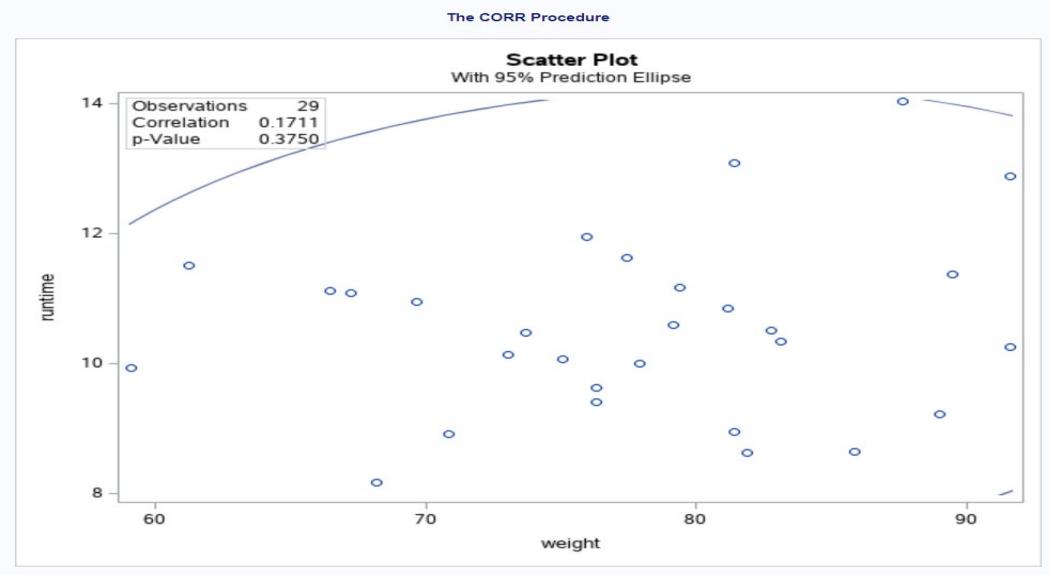
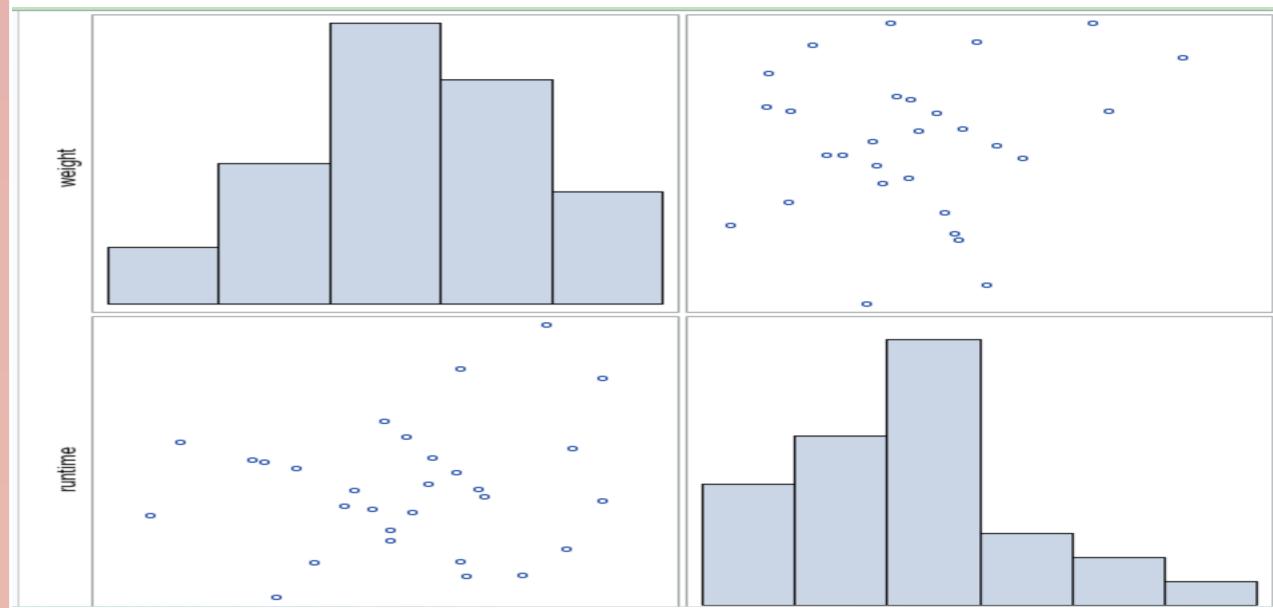
Ho (Null Hypothesis):
Weight is independent of Runtime

H1 (Alternate Hypothesis): Weight is not independent of Runtime

Pearson Correlation Coefficients, N = 29		
	weight	runtime
weight	1.00000	0.17106 0.3750
runtime	0.17106 0.3750	1.00000



Scatter Plot Matrix



Conclusion

Observations	29
Correlation	0.1711
p-Value	0.3750

- Since p-value ($0.37 > 0.05$), we fail to reject null hypothesis.
- This means Weight is independent of Runtime.
- Since correlation is 0.17 (between 0 and 0.25) , this means there is a negligible positive association.

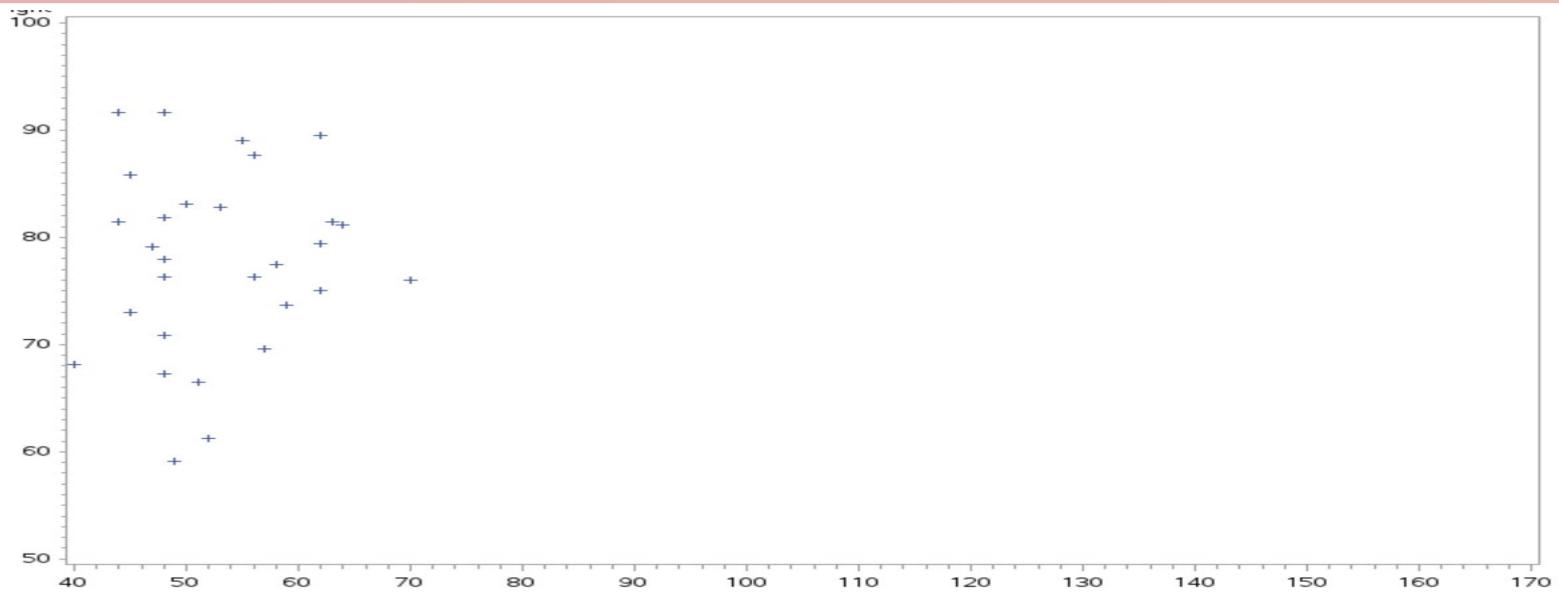
Pearson Test of Independence Hypothesis (Weight and Rstpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var weight rstpulse;  
run;  
proc gplot data=fitness;  
plot weight * rstpulse;  
run;
```

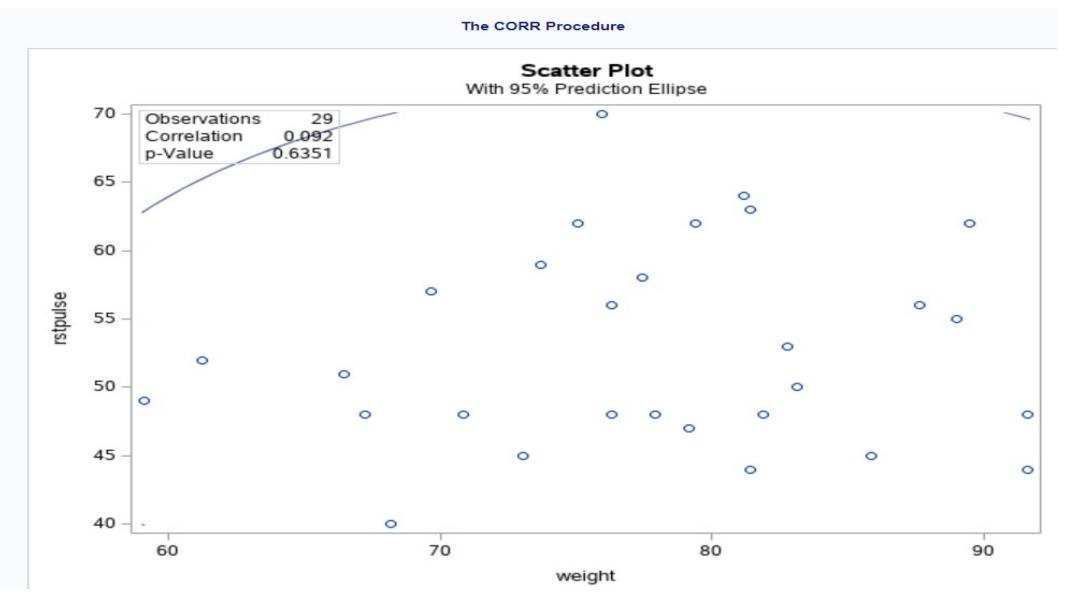
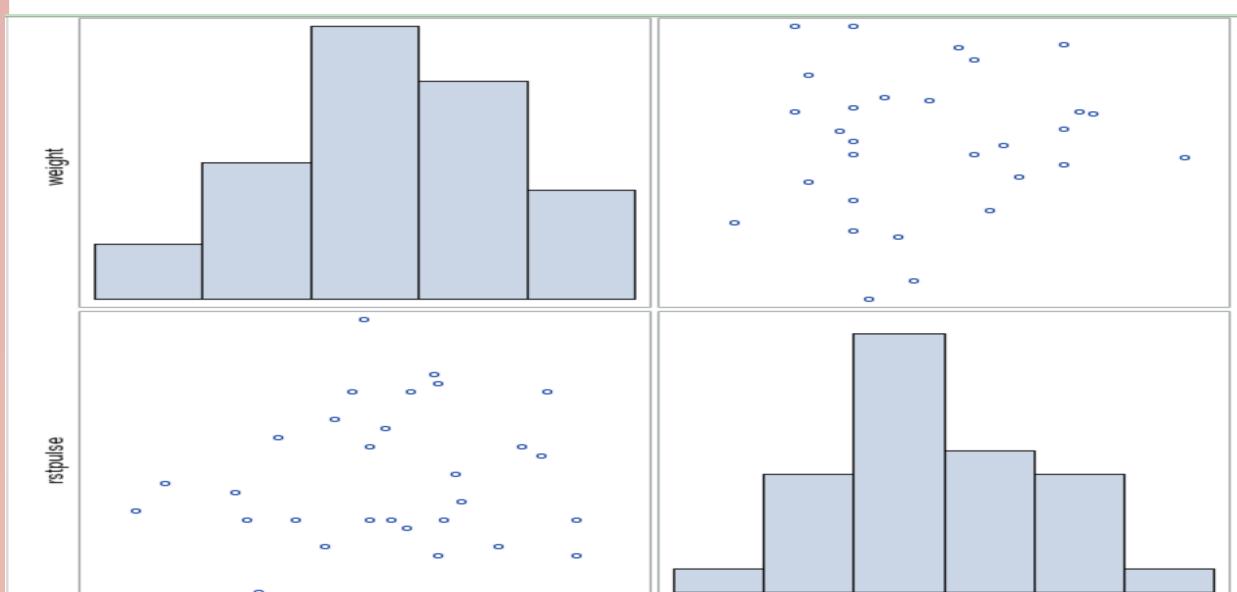
Ho (Null Hypothesis):
Weight is independent of
Rstpulse

H1 (Alternate
Hypothesis): Weight is not
independent of Rstpulse

Pearson Correlation Coefficients, N = 29		
	weight	rstpulse
weight	1.00000	0.09199 0.6351
rstpulse	0.09199 0.6351	1.00000



Scatter Plot Matrix



Conclusion

Observations	29
Correlation	0.092
p-Value	0.6351

- Since p-value ($0.63 > 0.05$), we fail to reject null hypothesis.
- This means Weight is independent of Rstpulse.
- Since correlation is 0.09 (between 0 and 0.25), this means there is a negligible positive association.

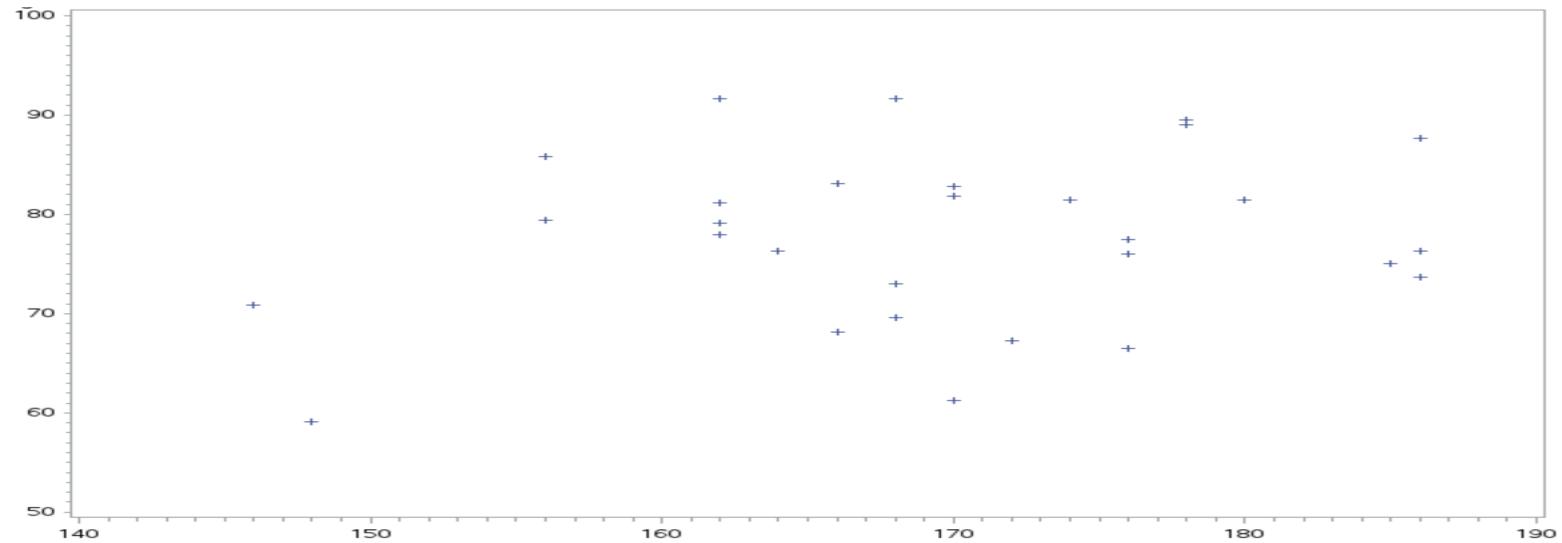
Pearson Test of Independence Hypothesis (Weight and Runpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var weight runpulse;  
run;  
proc gplot data=fitness;  
plot weight * runpulse;  
run;
```

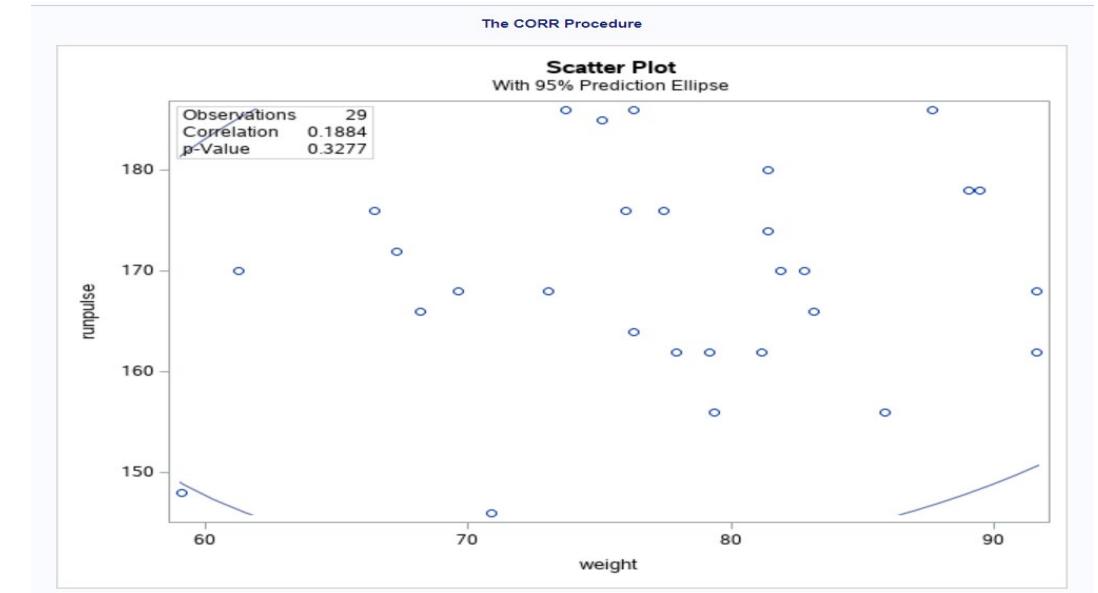
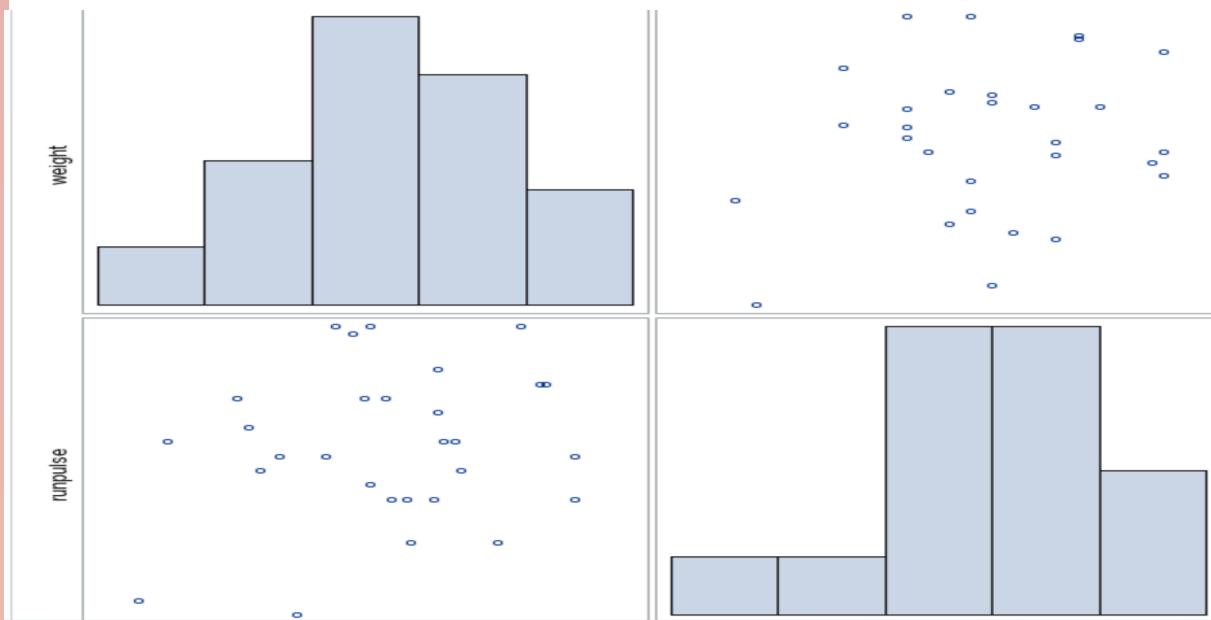
Ho (Null Hypothesis):
Weight is independent of
Runpulse

H1 (Alternate
Hypothesis): Weight is not
independent of Runpulse

Pearson Correlation Coefficients, N = 29		
	weight	runpulse
weight	1.00000	0.18839 0.3277
runpulse	0.18839 0.3277	1.00000



Scatter Plot Matrix



Conclusion

Observations	29
Correlation	0.1884
p-Value	0.3277

- Since p-value ($0.32 > 0.05$), we fail to reject null hypothesis.
- This means Weight is independent of Runtime.
- Since correlation is 0.18 (between 0 and 0.25), this means there is a negligible positive association.

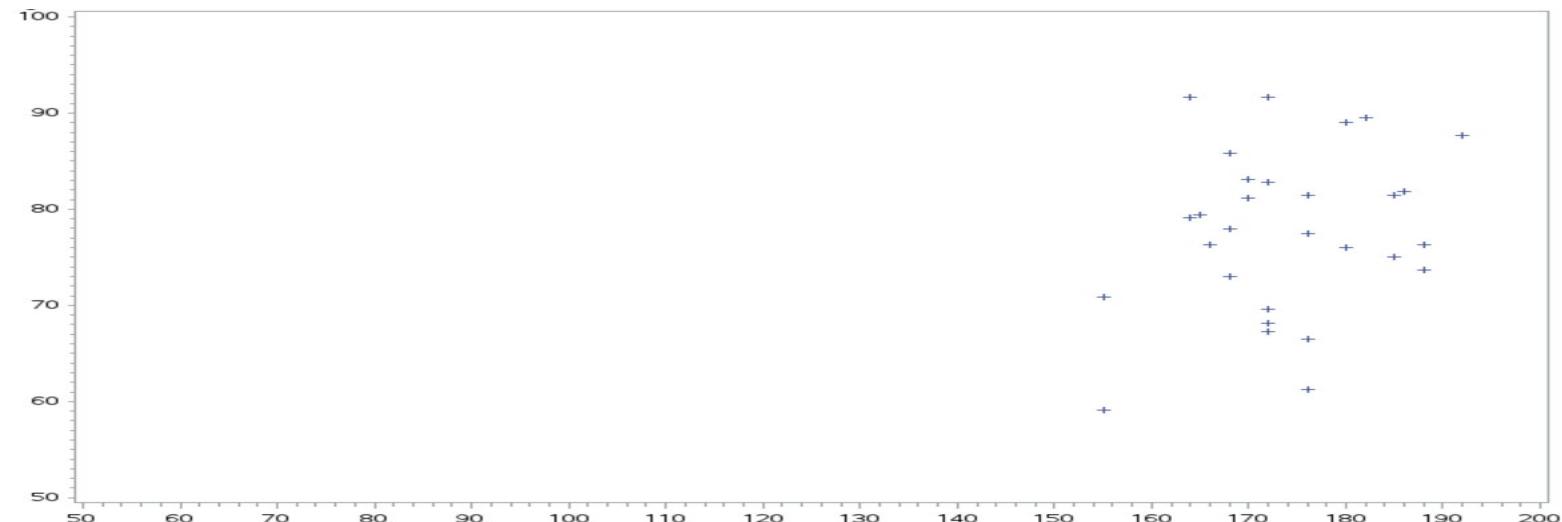
Pearson Test of Independence Hypothesis (Weight and Maxpulse)

```
proc corr data=fitness nomiss  
  plots=matrix(histogram)  
  plots=scatter(nvar=2 alpha=0.05);  
  var weight maxpulse;  
run;  
proc gplot data=fitness;  
plot weight * maxpulse;  
run;
```

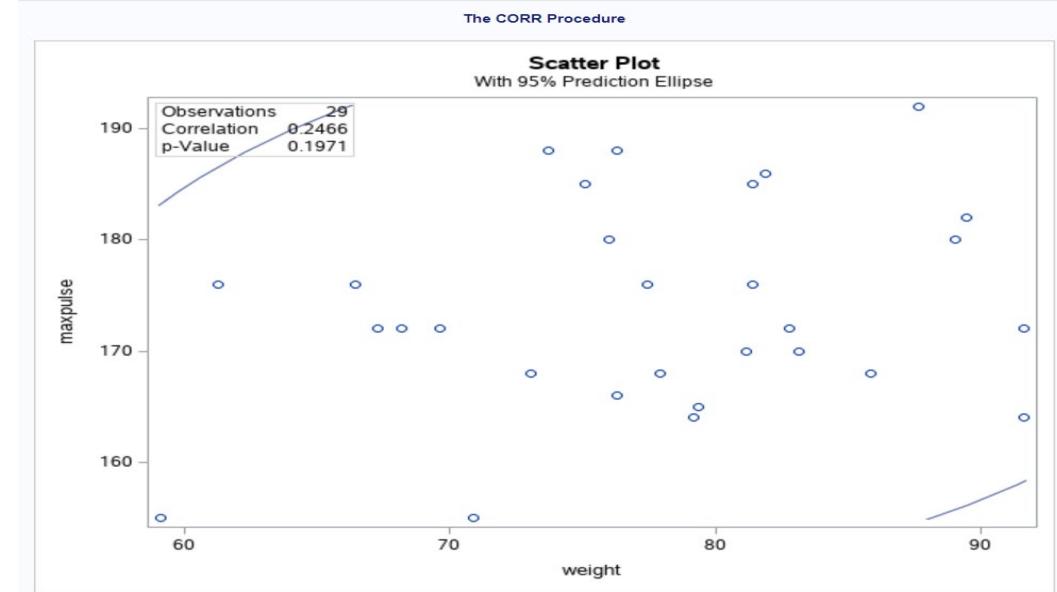
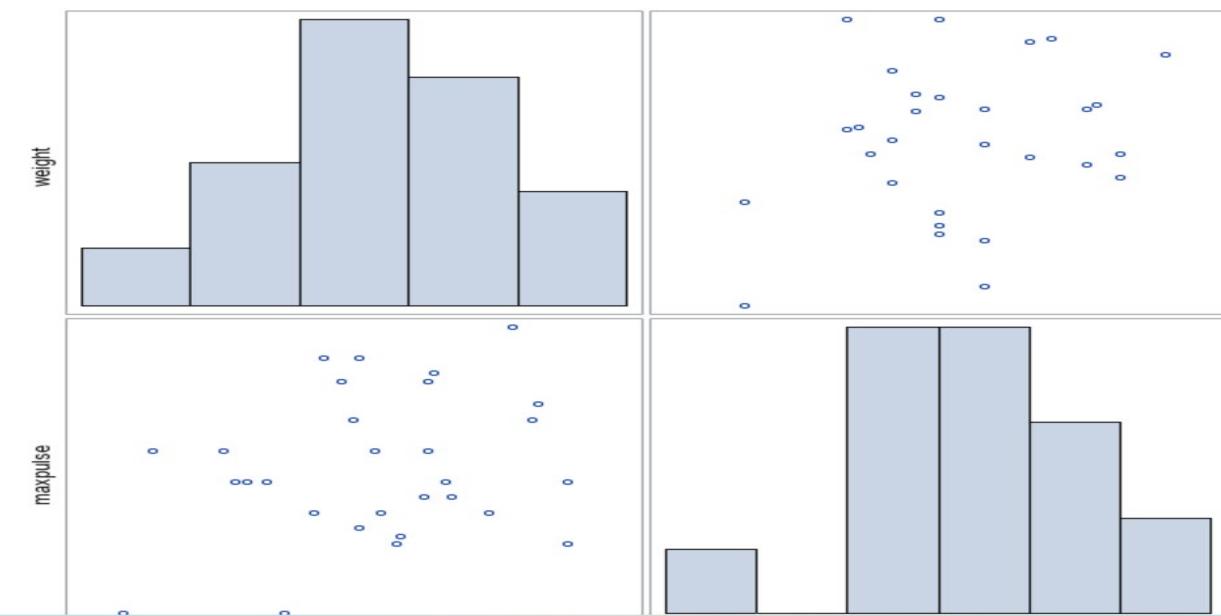
Ho (Null Hypothesis):
Weight is independent of
Maxpulse

H1 (Alternate
Hypothesis): Weight is not
independent of Maxpulse

Pearson Correlation Coefficients, N = 29		
	weight	maxpulse
weight	1.00000	0.24663 0.1971
maxpulse	0.24663 0.1971	1.00000



Scatter Plot Matrix



Conclusion

Observations	29
Correlation	0.2466
p-Value	0.1971

- Since p-value ($0.19 > 0.05$), we fail to reject null hypothesis.
- This means Weight is independent of Maxpulse.
- Since correlation is 0.24 (between 0 and 0.25), this means there is a negligible positive association.

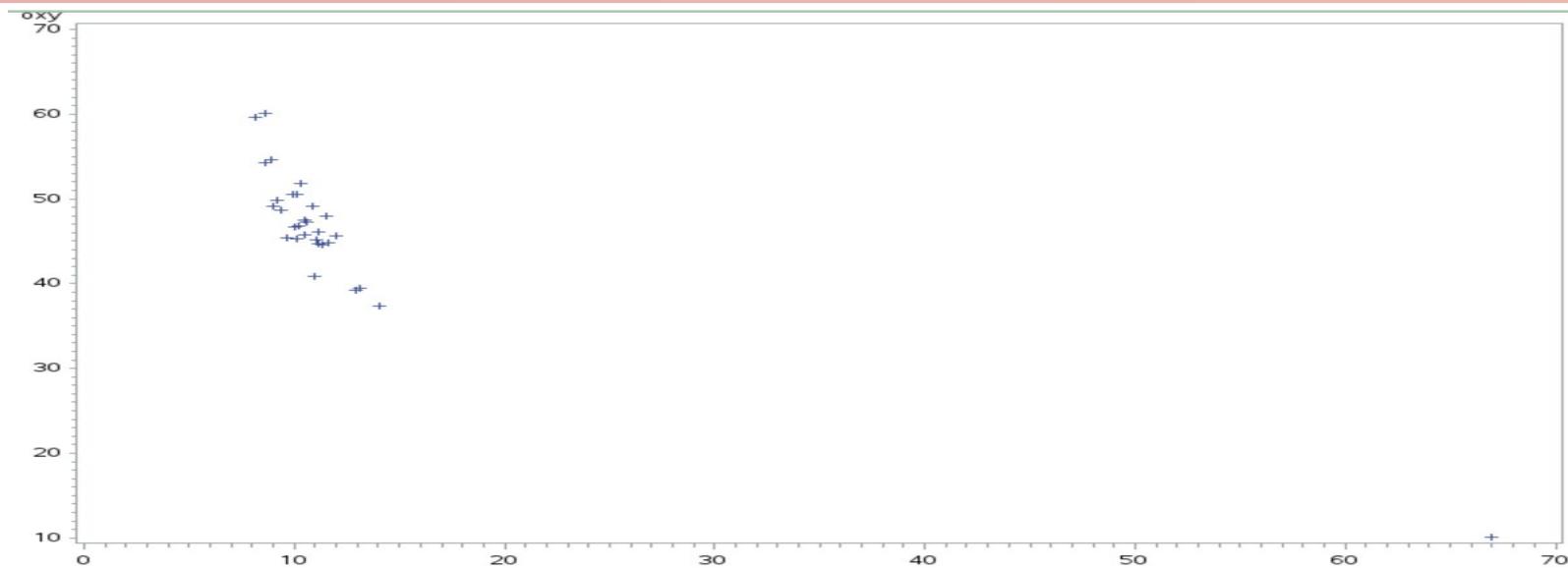
Pearson Test of Independence Hypothesis (Oxy and Runtime)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var oxy runtime;  
run;  
proc gplot data=fitness;  
plot oxy * runtime;  
run;
```

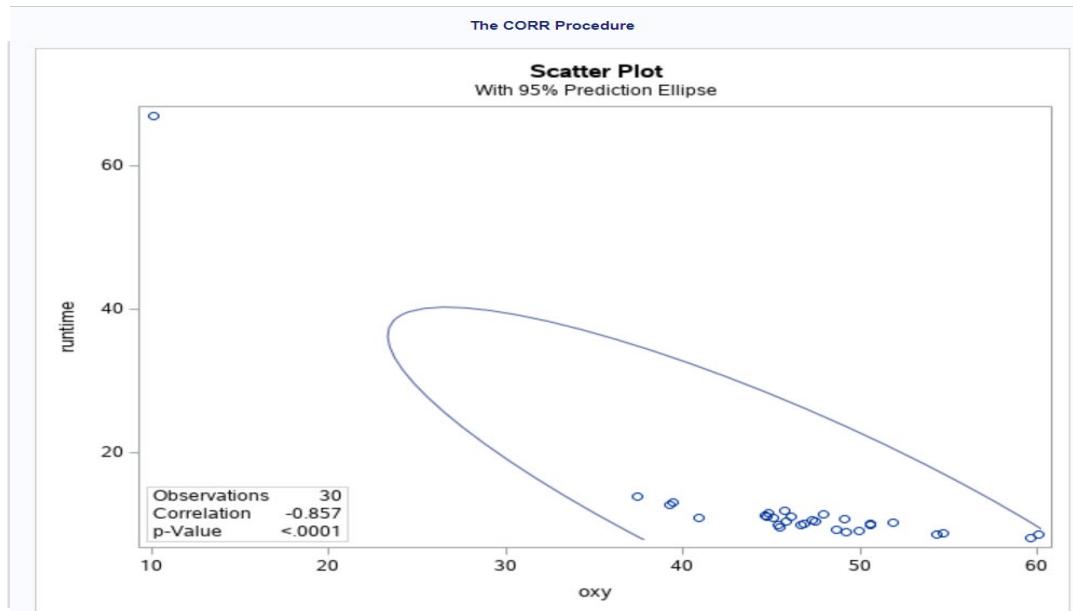
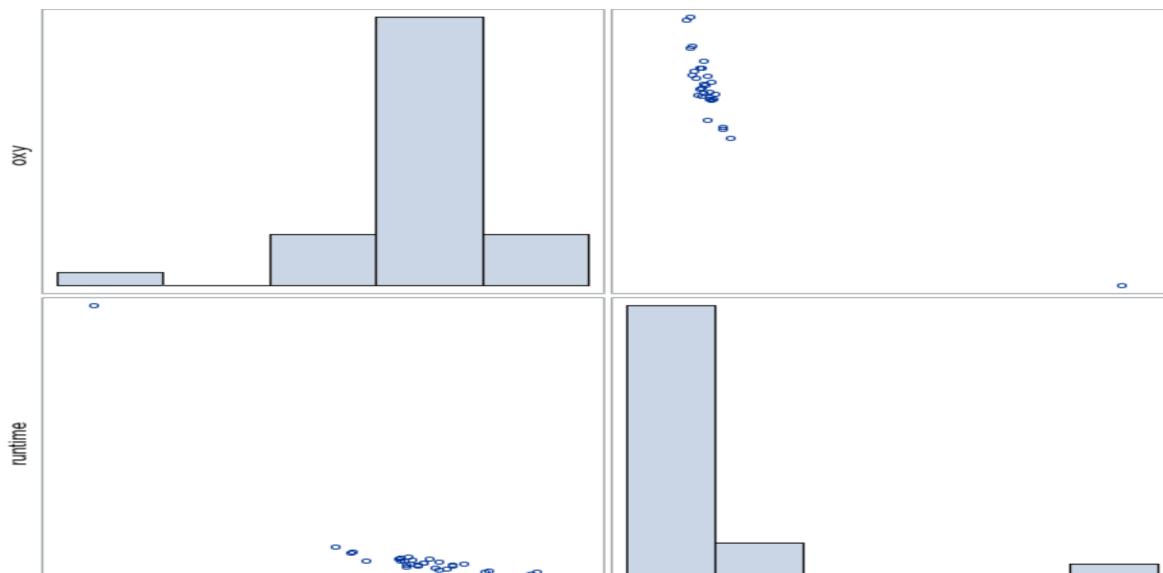
Ho (Null Hypothesis):
Oxy is independent of Runtime

H1 (Alternate Hypothesis): Oxy is not independent of Runtime

Pearson Correlation Coefficients, N = 30		
	oxy	runtime
oxy	1.00000	-0.85736 <.0001
runtime	-0.85736 <.0001	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.857
p-Value	<.0001

- Since p-value (<0.0001) $<$ alpha (0.05), reject null hypothesis.
- This means Oxy is not independent of Runtime.
- Since correlation is -0.8 (between -0.75 and -1) , this means there is a strong negative association.

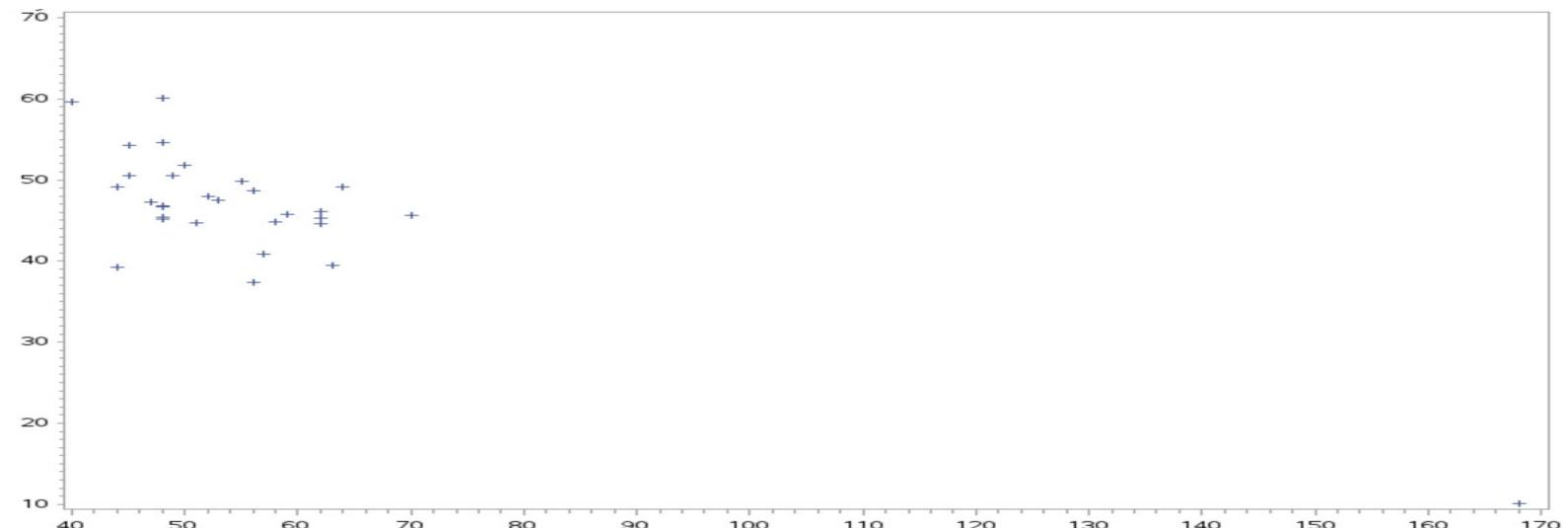
Pearson Test of Independence Hypothesis (Oxy and Rstpulse)

```
proc corr data=fitness nomiss  
  plots=matrix(histogram)  
  plots=scatter(nvar=2 alpha=0.05);  
  var oxy rstpulse;  
run;  
  
proc gplot data=fitness;  
  plot oxy * rstpulse;  
run;
```

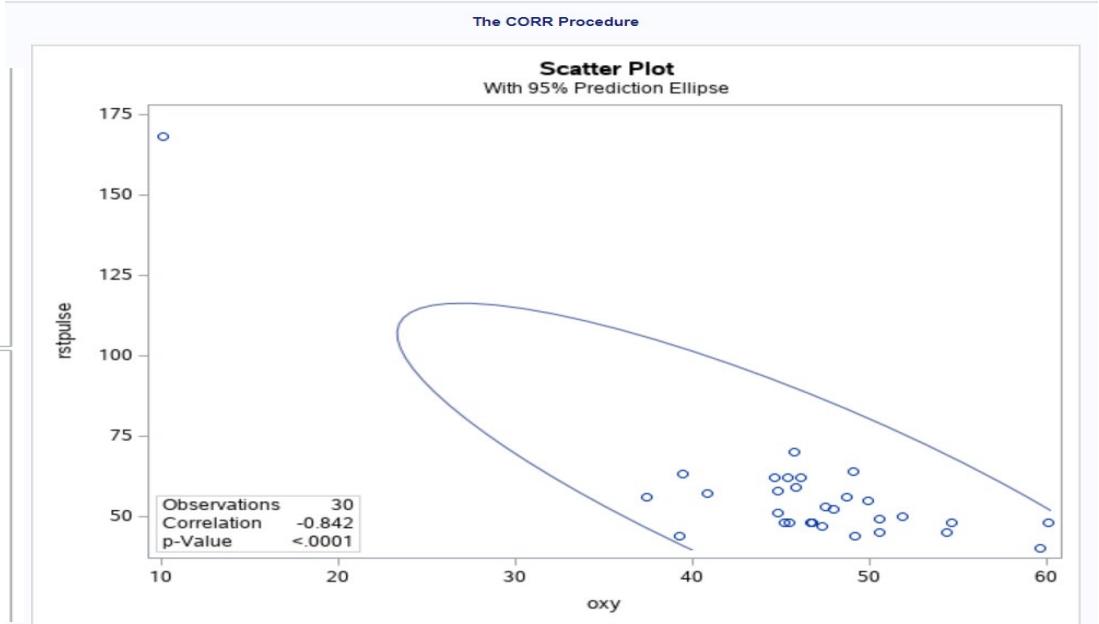
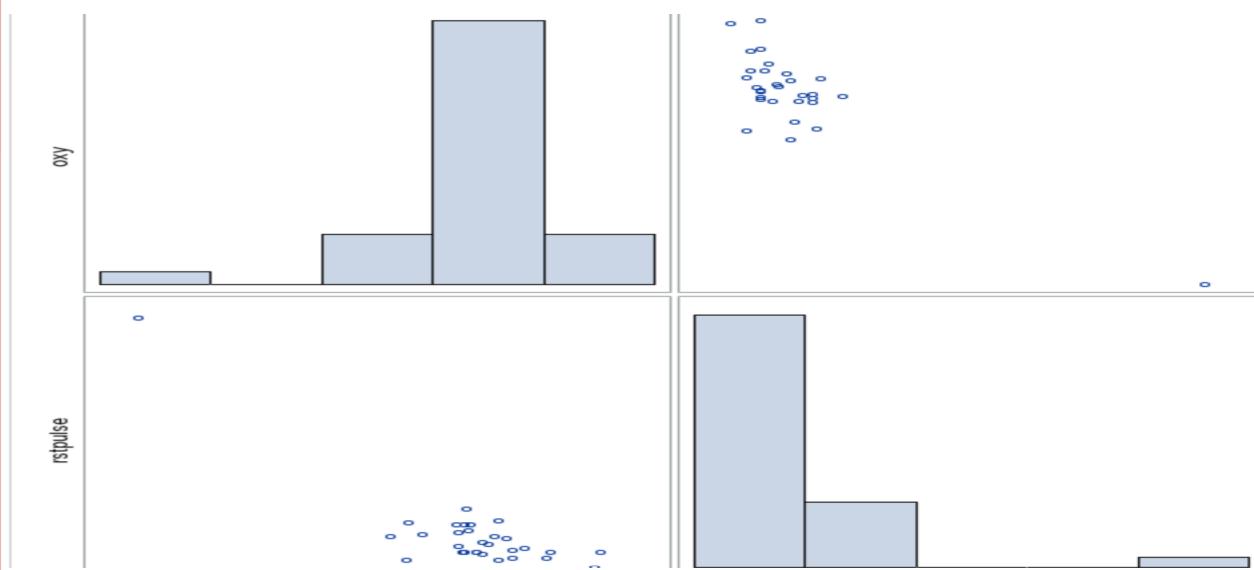
Ho (Null Hypothesis):
Oxy is independent of
Rstpulse

H1 (Alternate Hypothesis): Oxy is not independent of Rstpulse

Pearson Correlation Coefficients, N = 30		
	oxy	rstpulse
oxy	1.00000	-0.84179 <.0001
rstpulse	-0.84179 <.0001	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.842
p-Value	<.0001

- Since p-value (<0.0001) $<$ alpha (0.05), reject null hypothesis.
- This means Oxy is not independent of Rstpulse.
- Since correlation is -0.8 (between -0.75 and -1) , this means there is a strong negative association.

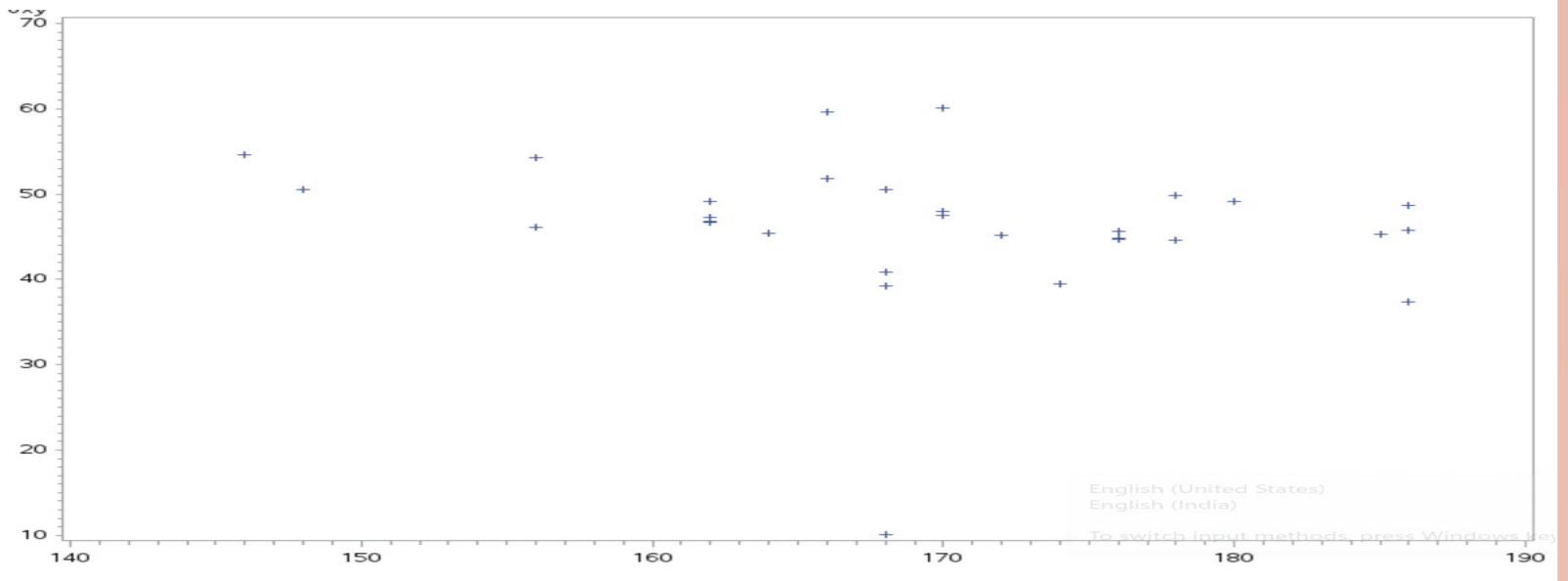
Pearson Test of Independence Hypothesis (Oxy and Runpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var oxy runpulse;  
run;  
proc gplot data=fitness;  
plot oxy * runpulse;  
run;
```

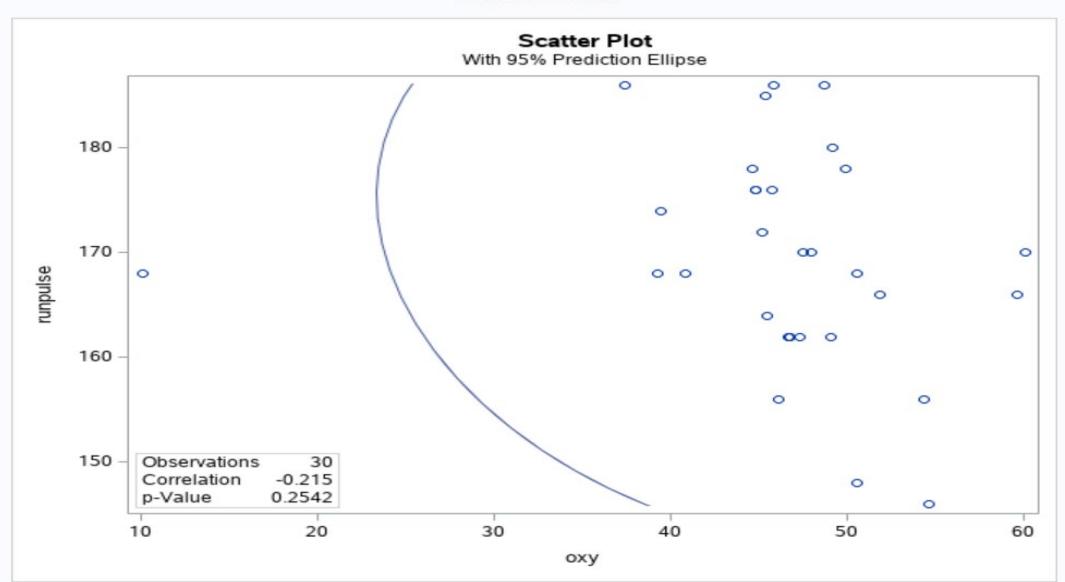
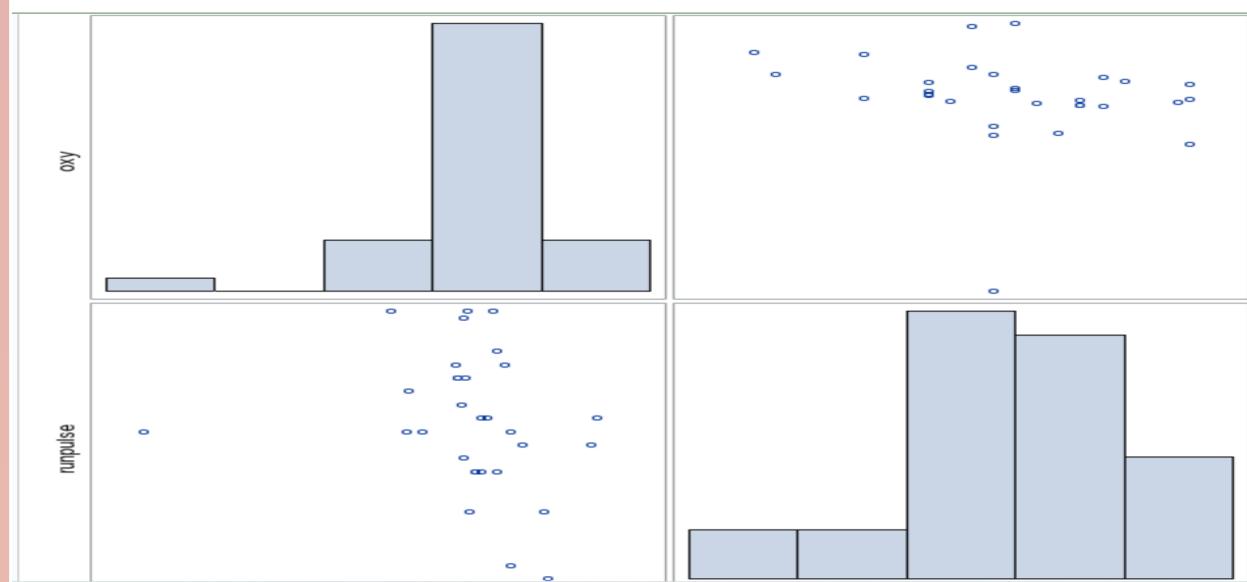
Ho (Null Hypothesis):
Oxy is independent of
Runpulse

H1 (Alternate Hypothesis): Oxy is not independent of Runpulse

Pearson Correlation Coefficients, N = 30		
	oxy	runpulse
oxy	1.00000	-0.21486 0.2542
runpulse	-0.21486 0.2542	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.215
p-Value	0.2542

- Since p-value (0.2) > alpha (0.05), we fail to reject null hypothesis.
- This means Oxy is independent of Runpulse.
- Since correlation is -0.21 (between 0 and -0.25) , this means there is an insignificant negative association.

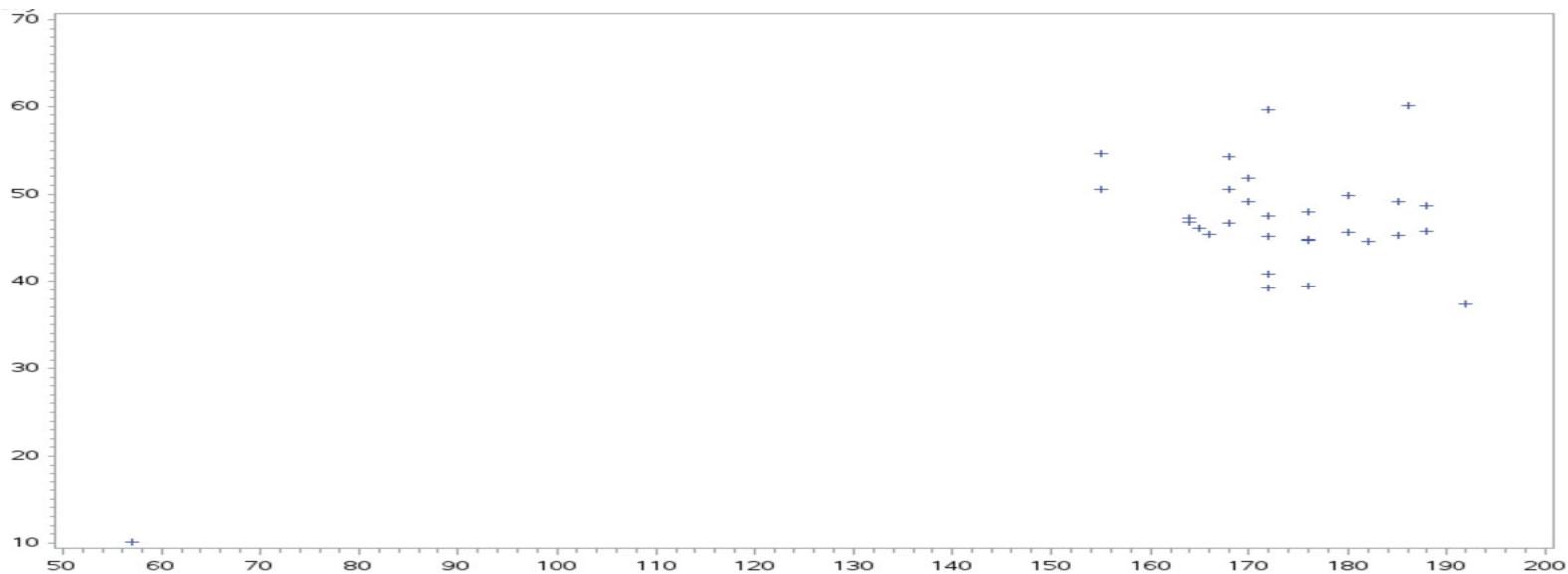
Pearson Test of Independence Hypothesis (Oxy and Maxpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var oxy maxpulse;  
run;  
proc gplot data=fitness;  
plot oxy * maxpulse;  
run;
```

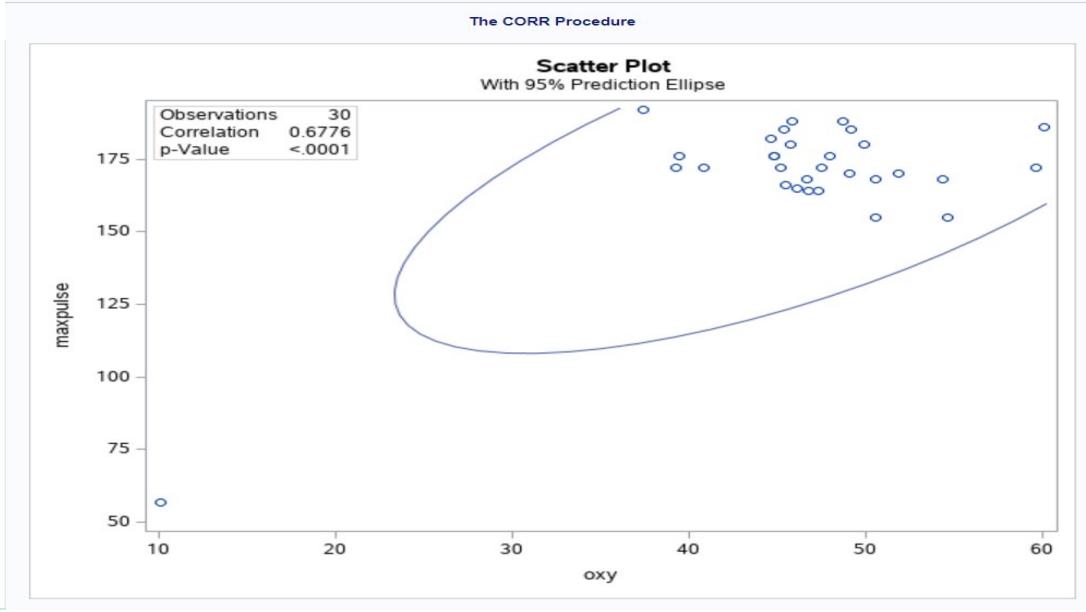
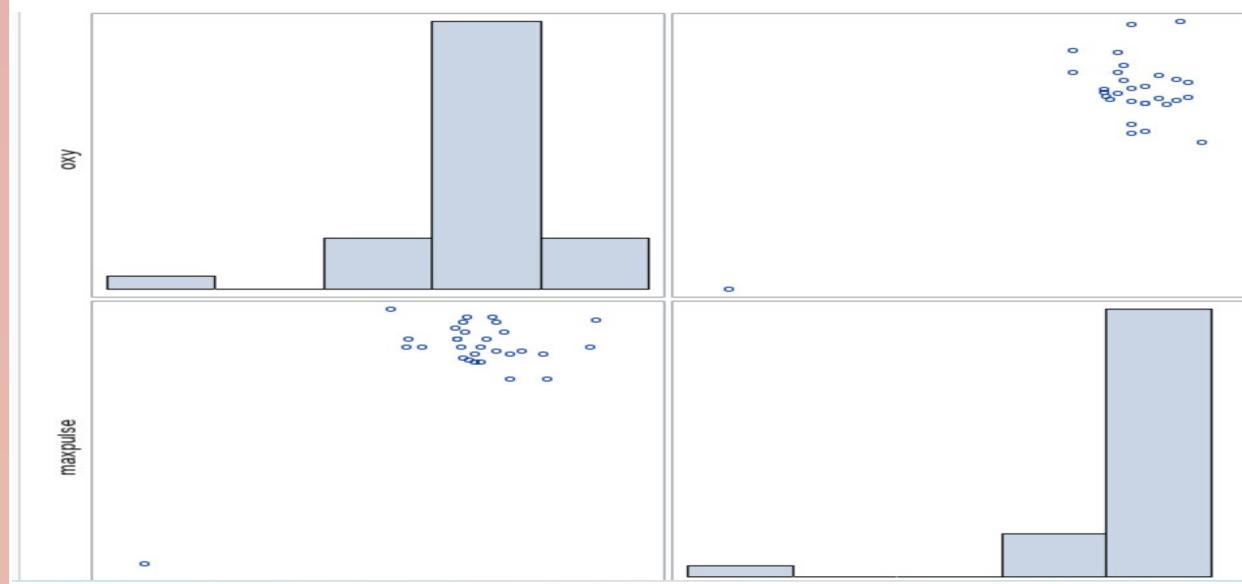
Ho (Null Hypothesis): Oxy is independent of Maxpulse

H1 (Alternate Hypothesis): Oxy is not independent of Maxpulse

Pearson Correlation Coefficients, N = 30		
	oxy	maxpulse
oxy	1.00000	0.67757 <.0001
maxpulse	0.67757 <.0001	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	0.6776
p-Value	<.0001

- Since p-value (<0.0001) $<$ alpha (0.05), reject null hypothesis.
- This means Oxy is not independent of Maxpulse.
- Since correlation is 0.6 (between 0.5 and 0.75) , this means there is a moderate positive association.

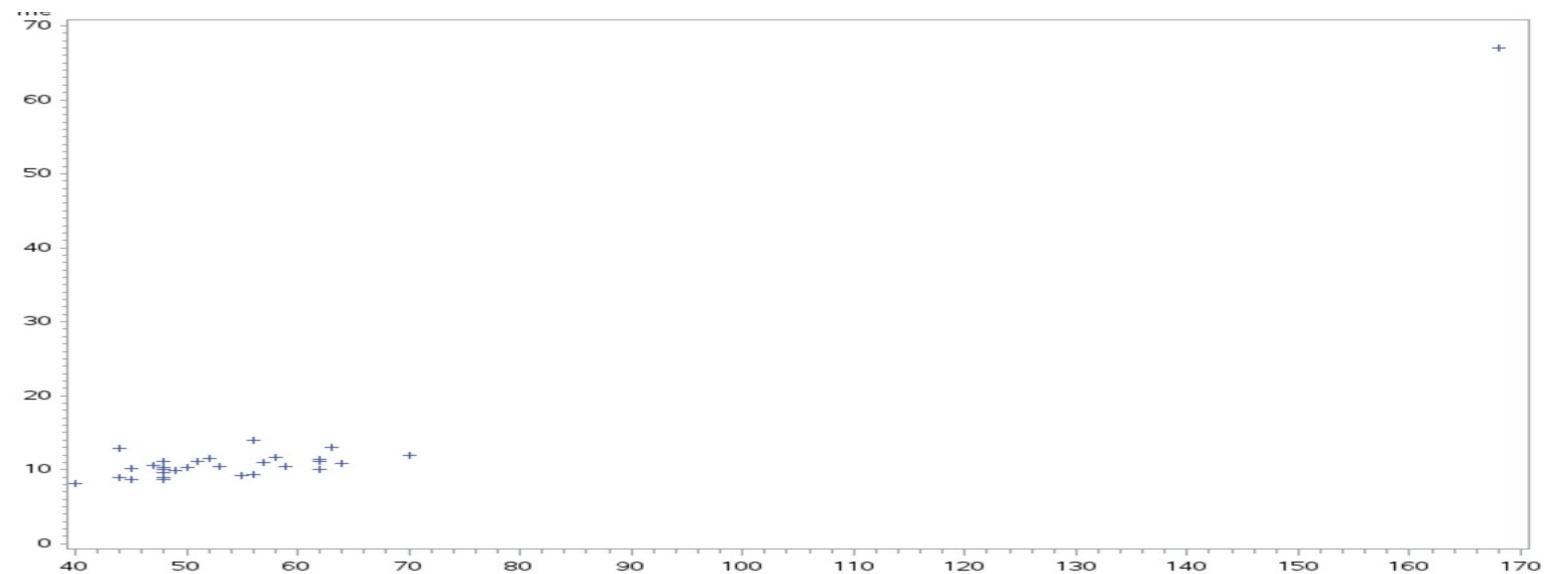
Pearson Test of Independence Hypothesis (Runtime and Rstpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var runtime rstpulse;  
run;  
proc gplot data=fitness;  
plot runtime * rstpulse;  
run;
```

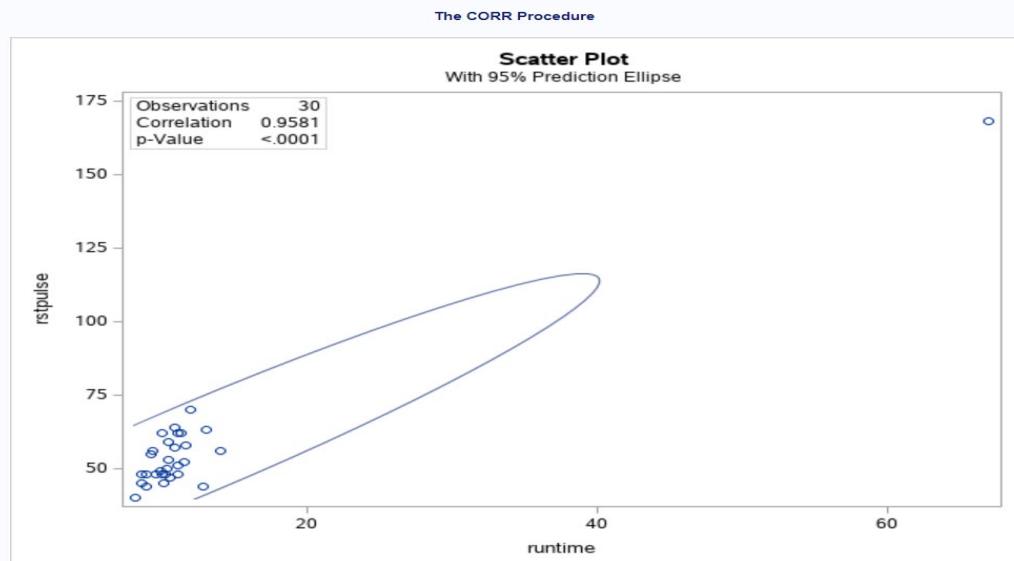
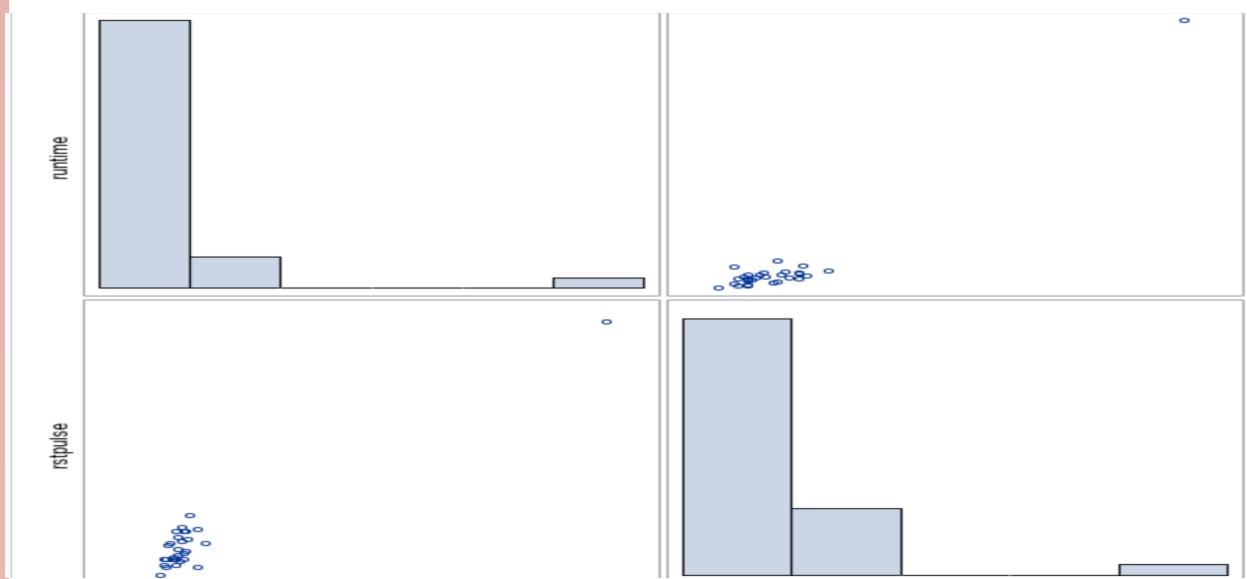
Ho (Null Hypothesis):
Runtime is independent of
Rstpulse

H1 (Alternate Hypothesis):
Runtime is not
independent of Rstpulse

Pearson Correlation Coefficients, N = 30		
	runtime	rstpulse
runtime	1.00000	0.95805 <.0001
rstpulse	0.95805 <.0001	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	0.9581
p-Value	<.0001

- Since p-value (<0.0001) $<$ alpha (0.05), reject null hypothesis.
- This means Runtime is not independent of Rstpulse.
- Since correlation is 0.95 (between 0.75 and 1) , this means there is a strong positive association.

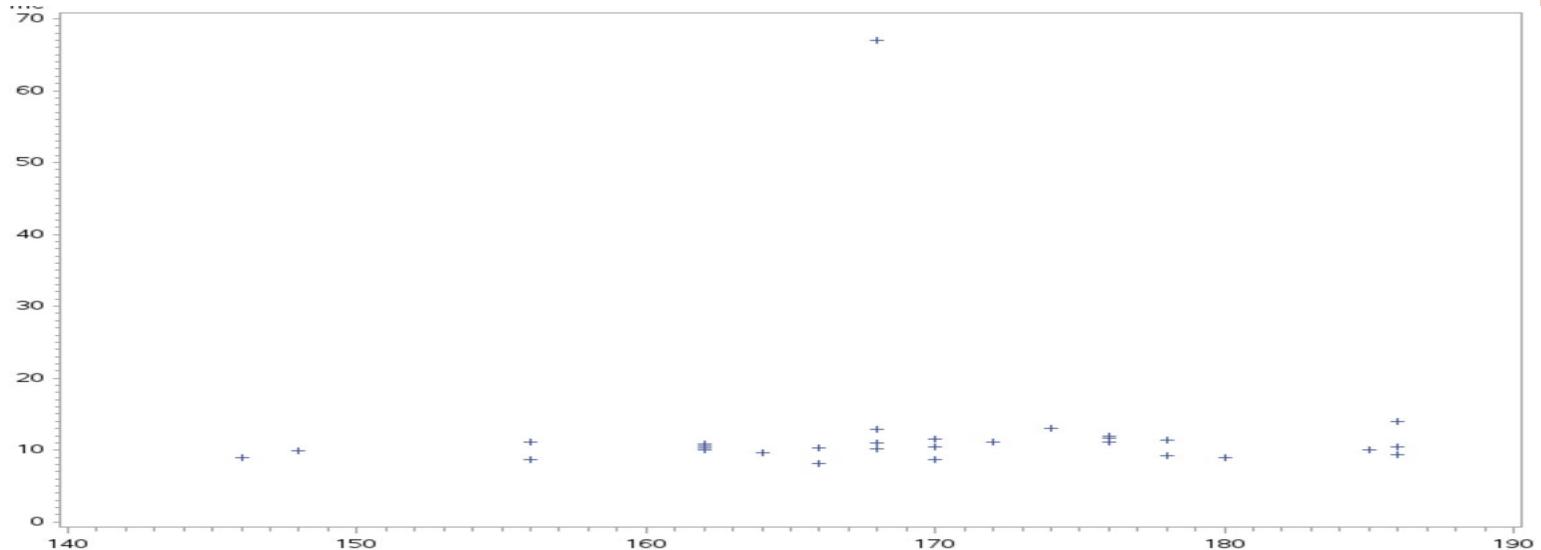
Pearson Test of Independence Hypothesis (Runtime and Runpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var runtime runpulse;  
run;  
proc gplot data=fitness;  
plot runtime * runpulse;  
run;
```

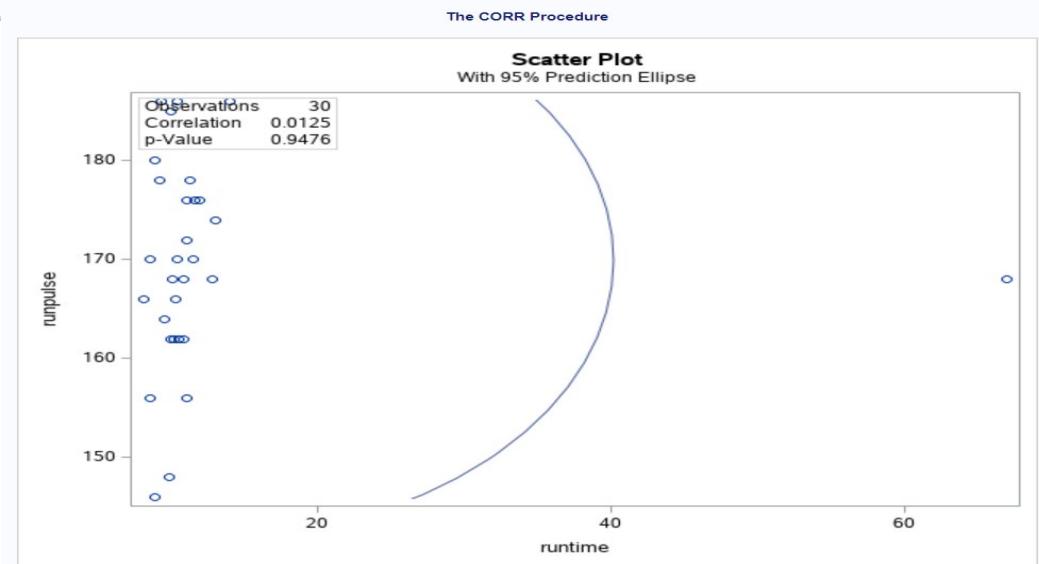
Ho (Null Hypothesis):
Runtime is independent of
Runpulse

H1 (Alternate Hypothesis):
Runtime is not
independent of Runpulse

Pearson Correlation Coefficients, N = 30		
	runtime	runpulse
runtime	1.00000	0.01253 0.9476
runpulse	0.01253 0.9476	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	0.0125
p-Value	0.9476

- Since p-value ($0.9 > 0.05$), we fail to reject null hypothesis.
- This means Runtime is independent of Runpulse.
- Since correlation is 0.01 (between 0 and 0.25), this means there is a negligible positive association.

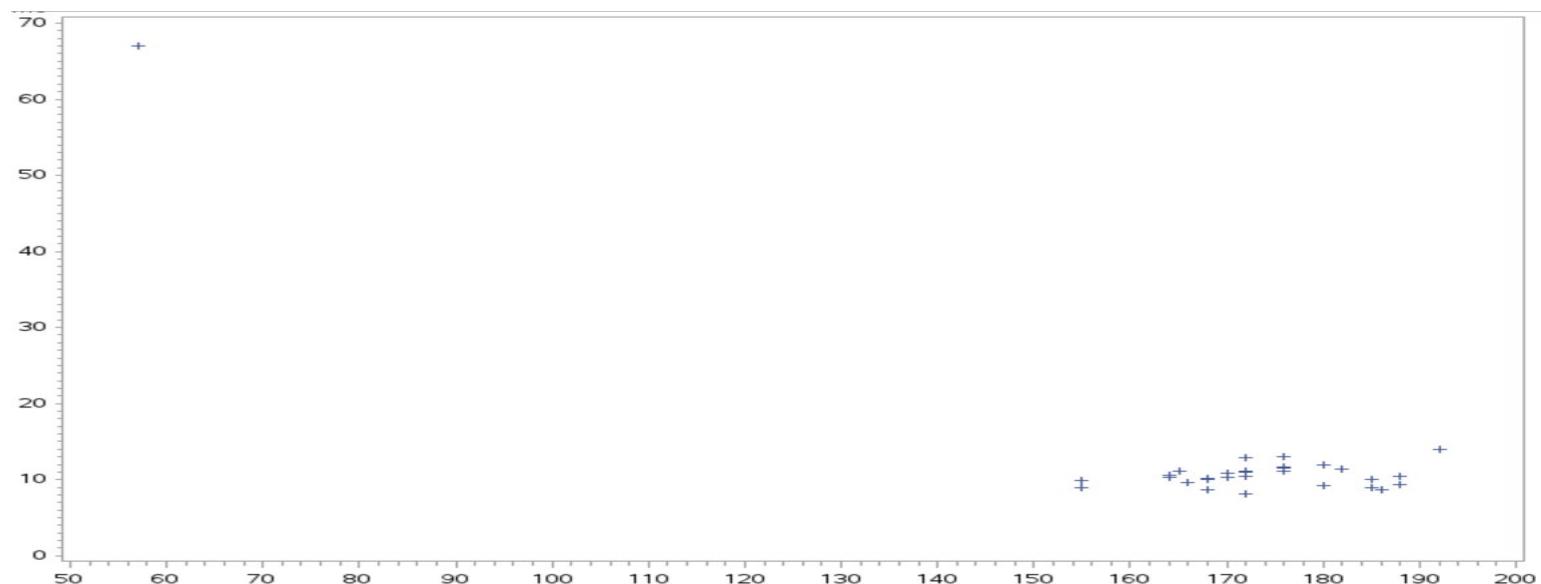
Pearson Test of Independence Hypothesis (Runtime and Maxpulse)

```
proc corr data=fitness nomiss  
  plots=matrix(histogram)  
  plots=scatter(nvar=2 alpha=0.05);  
  var runtime maxpulse;  
run;  
proc gplot data=fitness;  
  plot runtime * maxpulse;  
run;
```

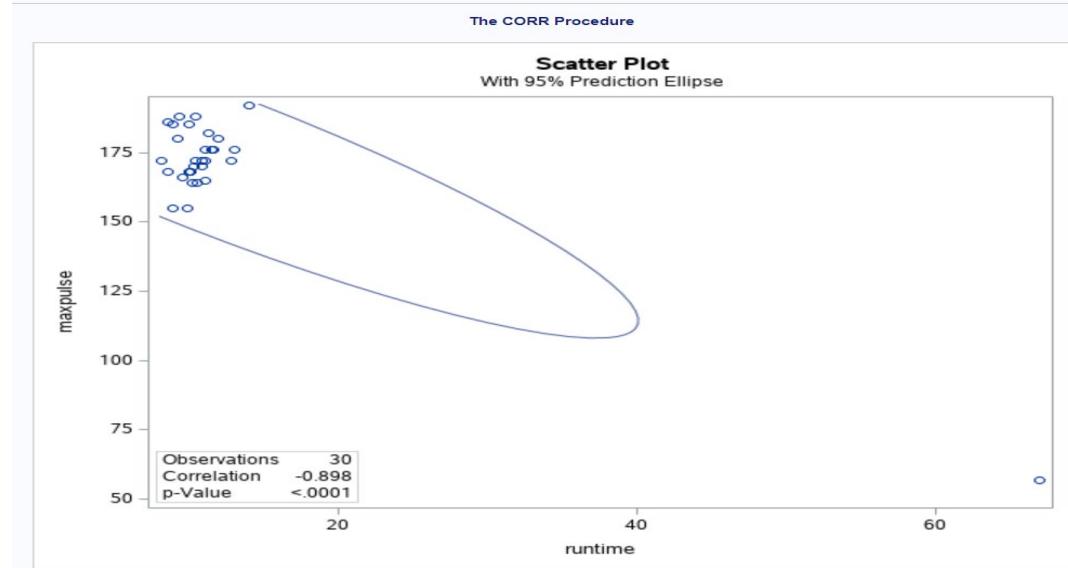
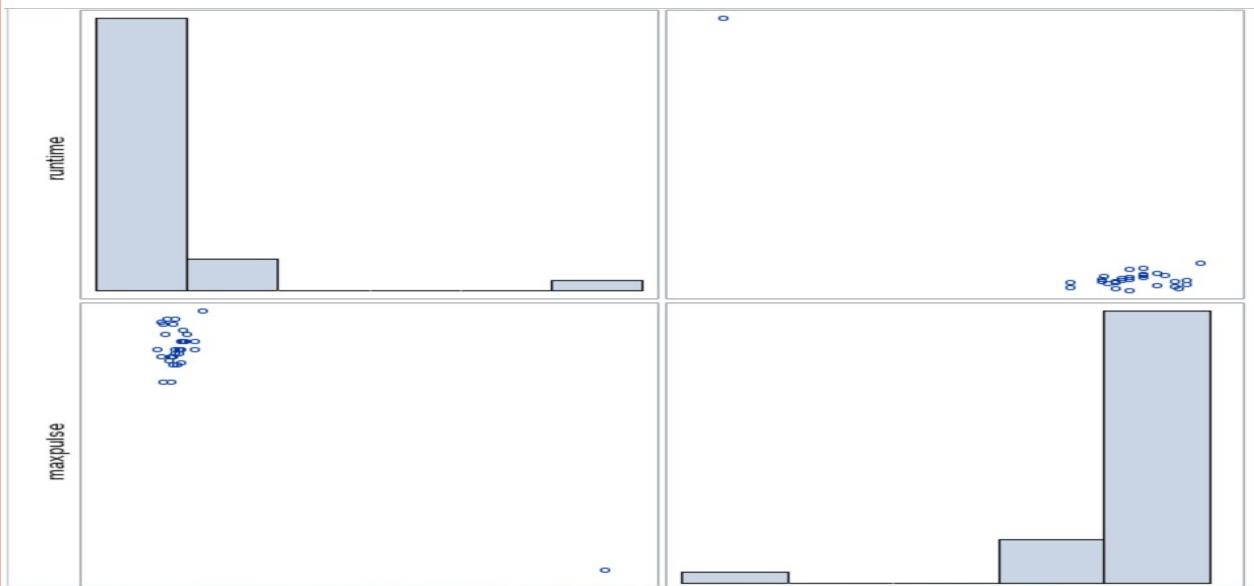
Ho (Null Hypothesis):
Runtime is independent of
Maxpulse

H1 (Alternate Hypothesis):
Runtime is not
independent of Maxpulse

Pearson Correlation Coefficients, N = 30		
	runtime	maxpulse
runtime	1.00000	-0.89847 <.0001
maxpulse	-0.89847 <.0001	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.898
p-Value	<.0001

- Since p-value (<0.0001) $<$ alpha (0.05), reject null hypothesis.
- This means Runtime is not independent of Maxpulse.
- Since correlation is -0.89 (between -0.75 and -1) , this means there is a strong negative association.

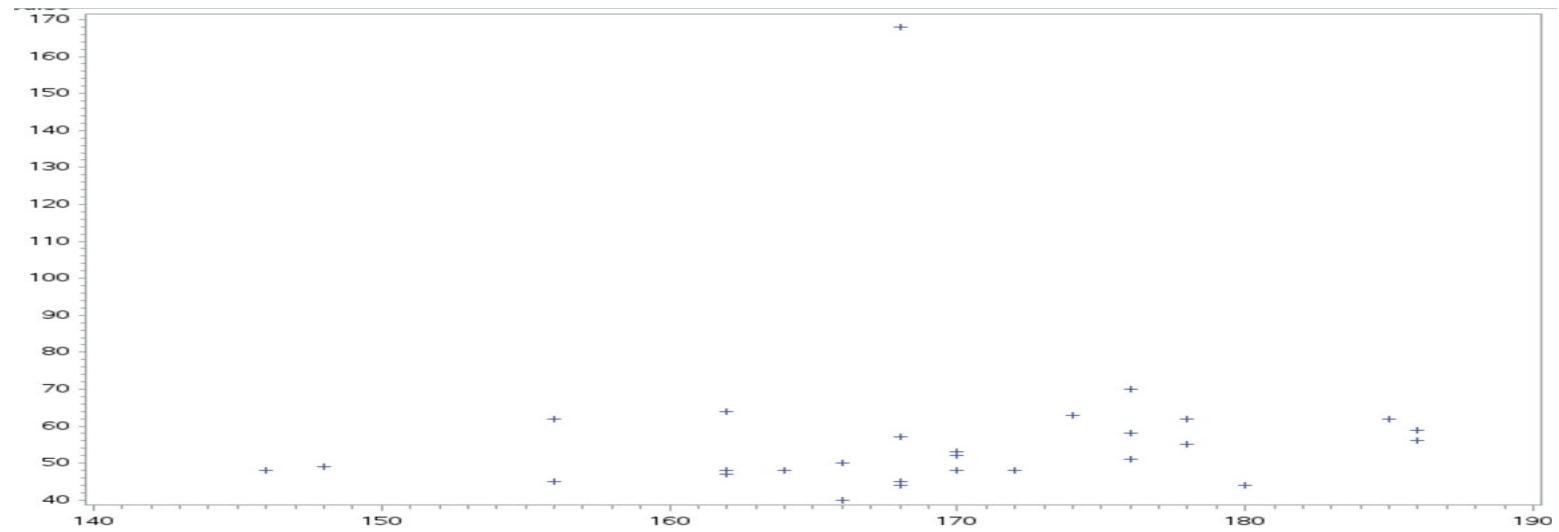
Pearson Test of Independence Hypothesis (Rstpulse and Runpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var rsthrt runpulse;  
run;  
proc gplot data=fitness;  
plot rsthrt * runpulse;  
run;
```

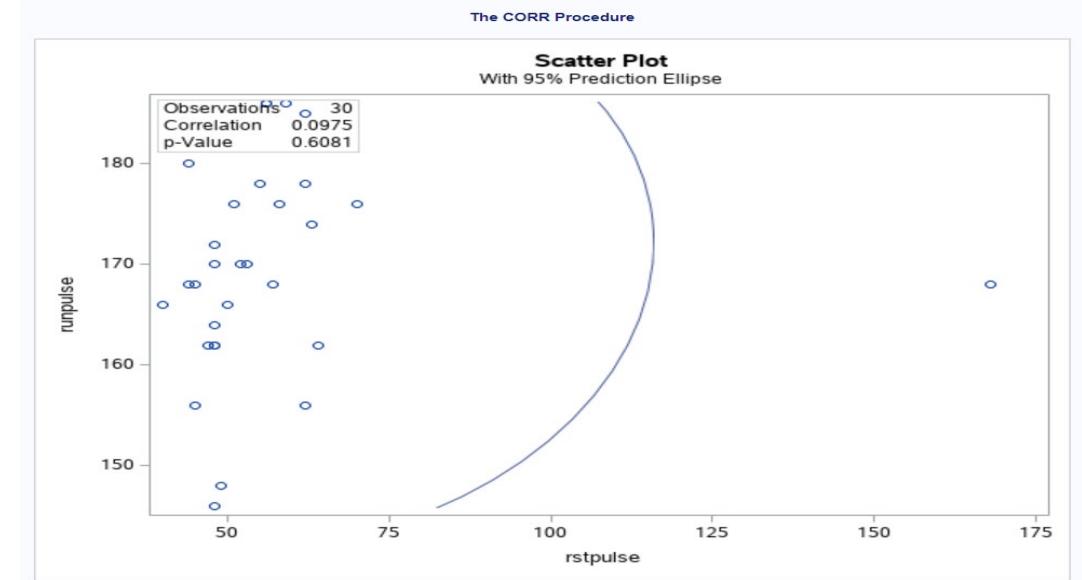
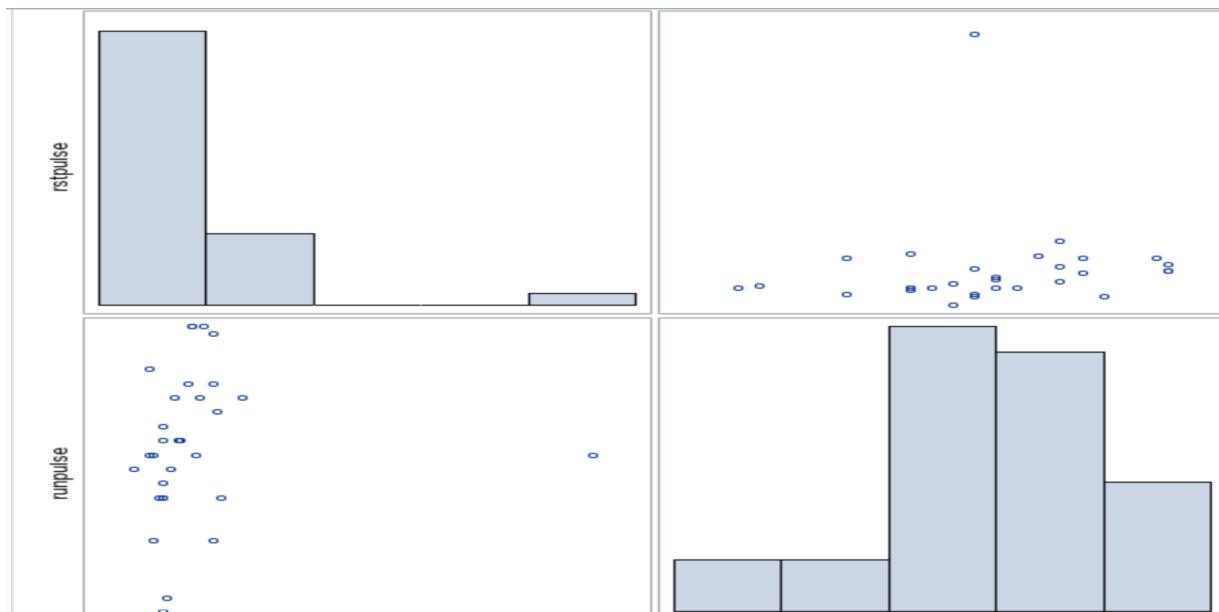
Ho (Null Hypothesis):
Rstpulse is independent of
Runpulse

H1 (Alternate Hypothesis):
Rstpulse is not
independent of Runpulse

Pearson Correlation Coefficients, N = 30		
	rstpulse	runpulse
rstpulse	1.00000	0.09753 0.6081
runpulse	0.09753 0.6081	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	0.0975
p-Value	0.6081

- Since p-value ($0.6 > 0.05$), we fail to reject null hypothesis.
- This means Rstpulse is independent of Runpulse.
- Since correlation is 0.09 (between 0 and 0.25), this means there is a negligible positive association.

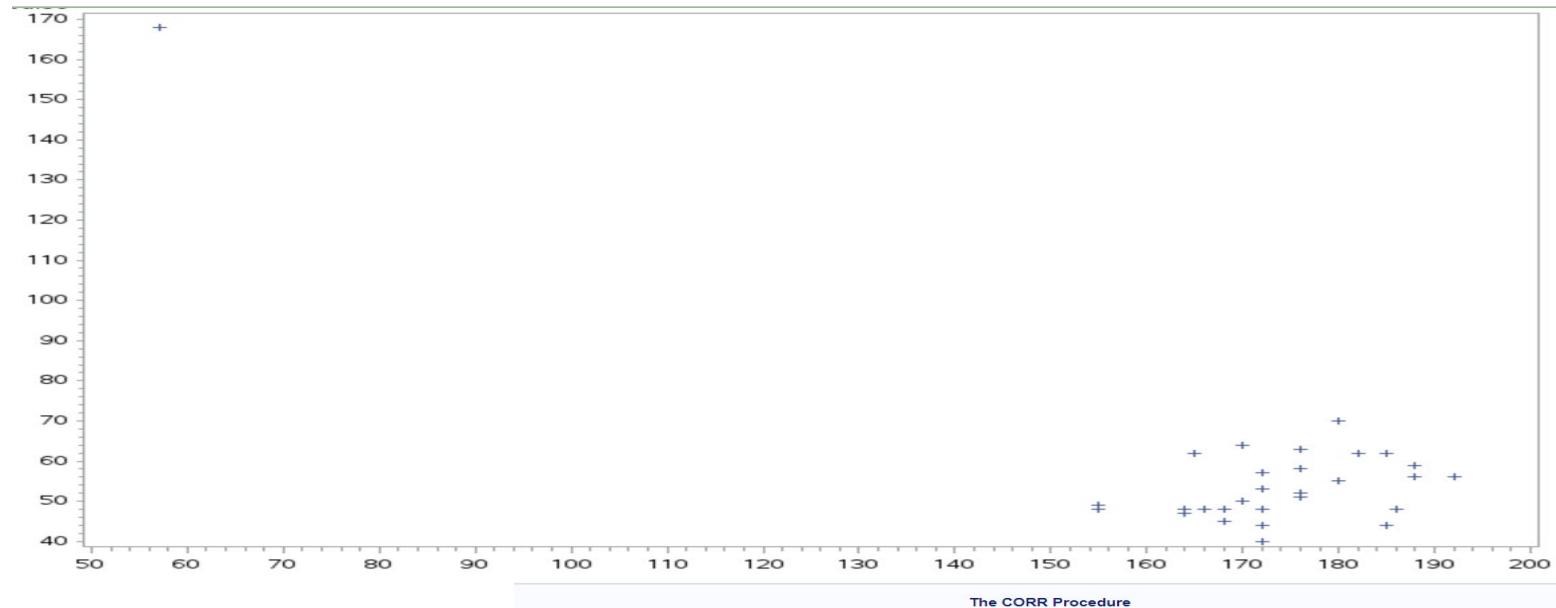
Pearson Test of Independence Hypothesis (Rstpulse and Maxpulse)

```
proc corr data=fitness nomiss  
  plots=matrix(histogram)  
  plots=scatter(nvar=2 alpha=0.05);  
  var rsthpulse maxpulse;  
run;  
proc gplot data=fitness;  
  plot rsthpulse * maxpulse;  
run;
```

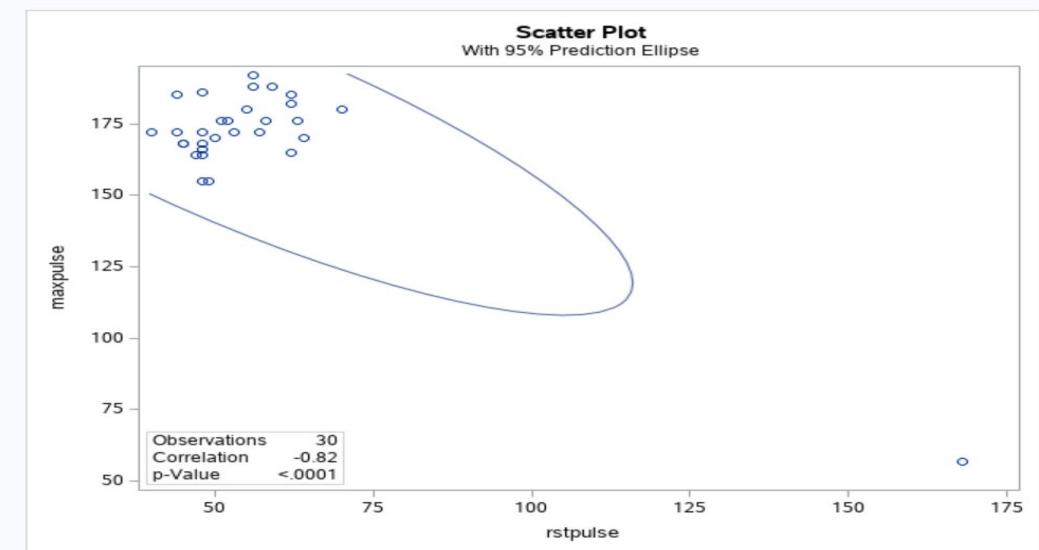
Ho (Null Hypothesis):
Rsthpulse is independent of
Maxpulse

H1 (Alternate Hypothesis):
Rsthpulse is not
independent of Maxpulse

Pearson Correlation Coefficients, N = 30		
	rstpulse	maxpulse
rstpulse	1.00000	-0.82010 <.0001
maxpulse	-0.82010 <.0001	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	-0.82
p-Value	<.0001

- Since p-value (<0.0001) $<$ alpha (0.05), reject null hypothesis.
- This means Rstpulse is not independent of Maxpulse.
- Since correlation is -0.82 (between -0.75 and -1) , this means there is a strong negative association.

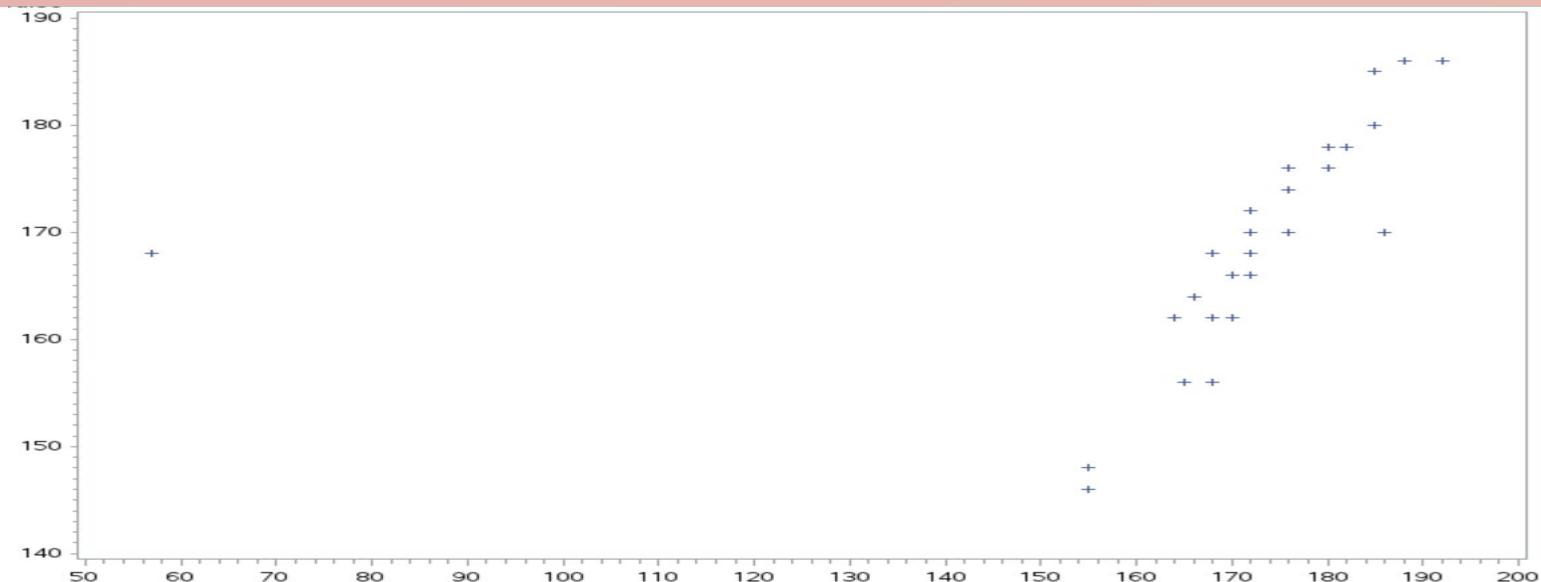
Pearson Test of Independence Hypothesis (Runpulse and Maxpulse)

```
proc corr data=fitness nomiss  
plots=matrix(histogram)  
plots=scatter(nvar=2 alpha=0.05);  
var runpulse maxpulse;  
run;  
proc gplot data=fitness;  
plot runpulse * maxpulse;  
run;
```

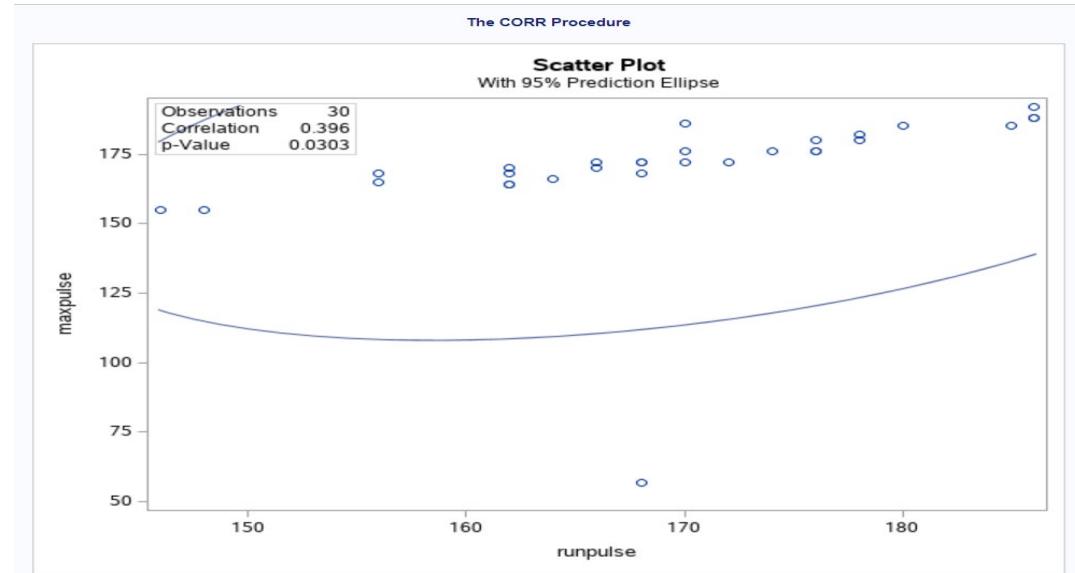
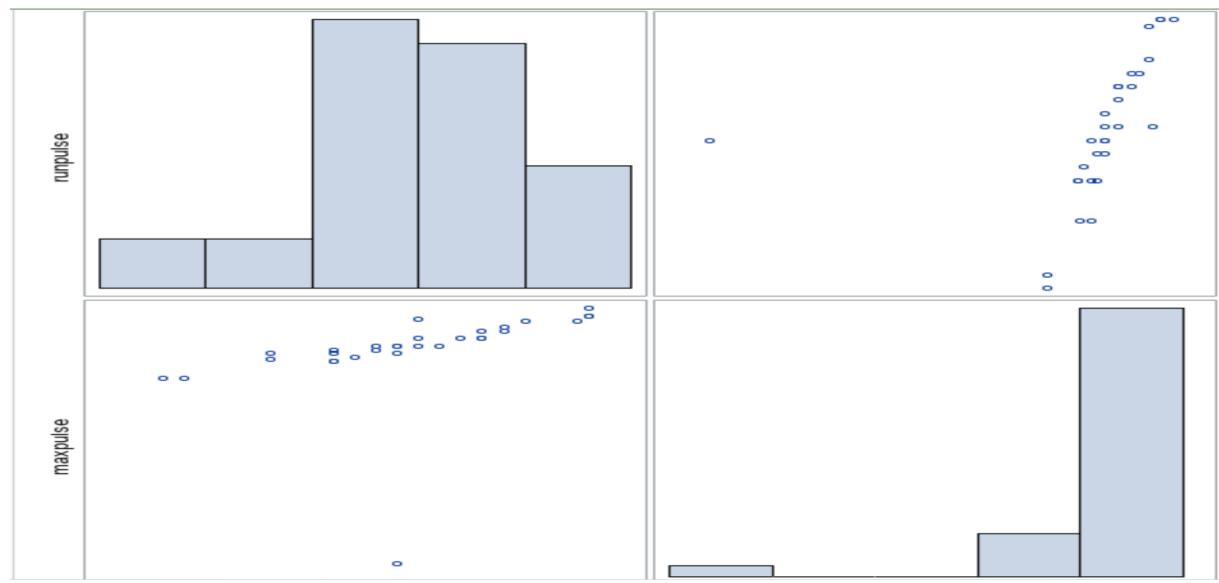
Ho (Null Hypothesis):
Runpulse is independent
of Maxpulse

H1 (Alternate Hypothesis):
Runpulse is not
independent of Maxpulse

Pearson Correlation Coefficients, N = 30		
	runpulse	maxpulse
runpulse	1.00000	0.39605 0.0303
maxpulse	0.39605 0.0303	1.00000



Scatter Plot Matrix



Conclusion

Observations	30
Correlation	0.396
p-Value	0.0303

- Since p-value (0.03) < alpha (0.05), we reject null hypothesis.
- This means Runpulse is not independent of Maxpulse.
- Since correlation is 0.39 (between 0.25 and 0.5), this means there is a weak positive association.

```

proc corr data=fitness nomiss ;
  var age weight oxy runtime rstpulse runpulse maxpulse;
run;

```

Pearson Correlation Procedure Matrix: Relationship between all variables

Pearson Correlation Coefficients, N = 29 Prob > r under H0: Rho=0							
	age	weight	oxy	runtime	rstpulse	runpulse	maxpulse
age	1.00000	-0.21205 0.2695	-0.24222 0.2055	0.11216 0.5624	-0.24999 0.1909	-0.38634 0.0384	-0.47222 0.0097
weight	-0.21205 0.2695	1.00000	-0.18884 0.3265	0.17106 0.3750	0.09199 0.6351	0.18839 0.3277	0.24663 0.1971
oxy	-0.24222 0.2055	-0.18884 0.3265	1.00000	-0.85062 <.0001	-0.44766 0.0149	-0.39208 0.0354	-0.22473 0.2412
runtime	0.11216 0.5624	0.17106 0.3750	-0.85062 <.0001	1.00000	0.48920 0.0071	0.30363 0.1093	0.21728 0.2575
rstpulse	-0.24999 0.1909	0.09199 0.6351	-0.44766 0.0149	0.48920 0.0071	1.00000	0.37821 0.0431	0.36464 0.0518
runpulse	-0.38634 0.0384	0.18839 0.3277	-0.39208 0.0354	0.30363 0.1093	0.37821 0.0431	1.00000	0.93349 <.0001
maxpulse	-0.47222 0.0097	0.24663 0.1971	-0.22473 0.2412	0.21728 0.2575	0.36464 0.0518	0.93349 <.0001	1.00000

```

proc corr data=fitness spearman ;
  var age weight oxy runtime rstpulse runpulse maxpulse;
run;

```

Spearman Correlation Procedure Matrix: Relationship between all variables

		Spearman Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations						
		age	weight	oxy	runtime	rstpulse	runpulse	maxpulse
age		1.00000 0.5512 30	-0.11539 0.4722 29	-0.13645 0.4722 30	0.10965 0.5641 30	-0.15459 0.4147 30	-0.33539 0.0700 30	-0.41655 0.0220 30
weight		-0.11539 0.5512 29	1.00000 29	-0.09117 0.6381 29	0.08870 0.6473 29	0.00087 0.9964 29	0.06714 0.7293 29	0.11886 0.5392 29
oxy		-0.13645 0.4722 30	-0.09117 0.6381 29	1.00000 30	-0.81758 <.0001 30	-0.49084 0.0059 30	-0.38593 0.0352 30	-0.16581 0.3812 30
runtime		0.10965 0.5641 30	0.08870 0.6473 29	-0.81758 <.0001 30	1.00000 30	0.57511 0.0009 30	0.24612 0.1898 30	0.07062 0.7108 30
rstpulse		-0.15459 0.4147 30	0.00087 0.9964 29	-0.49084 0.0059 30	0.57511 0.0009 30	1.00000 30	0.35375 0.0551 30	0.24742 0.1874 30
runpulse		-0.33539 0.0700 30	0.06714 0.7293 29	-0.38593 0.0352 30	0.24612 0.1898 30	0.35375 0.0551 30	1.00000 30	0.90308 <.0001 30
maxpulse		-0.41655 0.0220 30	0.11886 0.5392 29	-0.16581 0.3812 30	0.07062 0.7108 30	0.24742 0.1874 30	0.90308 <.0001 30	1.00000 30

Limitations of correlation

- Most of the correlations are **non-linear** in nature.
- There are **many outliers** in most of the correlations, removing them, can leave huge impact on correlation.