Analysis of the ARENA Simulation Data from ECS526 Take Home (Test #2) Fall 2018

Jorge Luis Romeu, Ph.D, Instructor

Airport Security Operation

Statement of the Problem

The objective of the present exercise is to show how to build, using statistical analysis, a set of system management tools to help its owners: (1) measure the performance of its variables; (2) compare their results, (3) improve and optimize the system, characterized by such variables.

The system here is the Airport Security Operation. We have collected data from some variables that characterize the status of said system (e.g. arrival and service rates, security levels) and some others that characterize its performance (e.g. time and number of customers in Queue, utilization of the servers).

Some of these variables, that characterize a scenario and are obtained first, become the factors in ANOVA or regressors in regression analyses. The others, that depend on them are system performance measures.

In the present analysis, the data collection has been implemented via a simulation program in ARENA. The parameters characterizing the different scenarios were given to the simulation program. The Performance Measures were collected from running said programs.

The overall objective is to find equations that allow the manager to predict the performance of the system under some scenario, characterized by several input variables or regressors.

As usual in a multivariate analysis situation, we start by implementing a multiple correlation analysis, to see which variables are impacted by which others, and which are not. Then, from those having high correlations, we select some useful variables to work with.

In our case, we select as response variable of interest the Utilization (of the servers). Other performance measures such as size of the waiting queue, times in queue and system, customers lost, etc. are highly correlated and well understood, having the value of Utilization.

Scenarios are characterized by arrival rates (e.g. the arrival of several large airplanes, such as a Boeing 707, loaded with passengers, as opposed to arrival of a few smaller, such as Embrier. Service rates may imply having all security inspection stations open, versus having half of them opened. Security levels imply being in a normal status versus a high security alert, one.

We will attempt to identify which of the characterizing variables are statistically significant.

The correlation analysis, our first step, follows.

Correlations: Arr. Mean, Triang. Mode, Sec. Prob, Wait in Sys, Wait in Q, ...

Cell Contents: Pearson correlation
P-Value

	Arr. Mean	Triang. Mode	Sec. Prob	Wait in Sys	
Triang. Mode	0.245 0.000				
Sec. Prob	-0.187 0.001	-0.272 0.000			
Wait in Sys	-0.570 0.000	0.428 0.000	0.169 0.003		
Wait in Q	-0.620 0.000	0.356 0.000	0.205 0.000	0.995 0.000	
No. in Q	-0.722 0.000	0.132 0.022	0.322 0.000	0.931 0.000	
Utilization	-0.869 0.000	0.146 0.011	0.243 0.000	0.863 0.000	

	Wait in Q	No. in Q
No. in Q	0.959 0.000	
Utilization	0.888 0.000	0.920 0.000

Thence, we regress Utilization over the three variables that characterize a scenario.

Regression Analysis: Utilization versus Arr. Mean, Triang. Mode, ...

The regression equation is:

Utilization = - 1.67 - 0.188 Arr. Mean + 0.219 Triang. Mode + 2.62 Sec. Prob

300 cases used, 5 cases contain missing values

Predictor	Coef	SE Coef	T	P
Constant	-1.6723	0.2361	-7.08	0.000
Arr. Mean	-0.187517	0.003331	-56.30	0.000
Triang. Mode	0.218975	0.008747	25.03	0.000
Sec. Prob	2.6214	0.2411	10.87	0.000

S = 0.0438356 R-Sq = 92.4% R-Sq(adj) = 92.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	6.8949	2.2983	1196.06	0.000
Residual Error	296	0.5688	0.0019		
Lack of Fit	6	0.4577	0.0763	199.21	0.000
Pure Error	290	0.1111	0.0004		
Total	299	7.4637			

Source	DF	Seq SS
Arr. Mean	1	5.6413
Triang. Mode	1	1.0264
Sec. Prob	1	0.2271

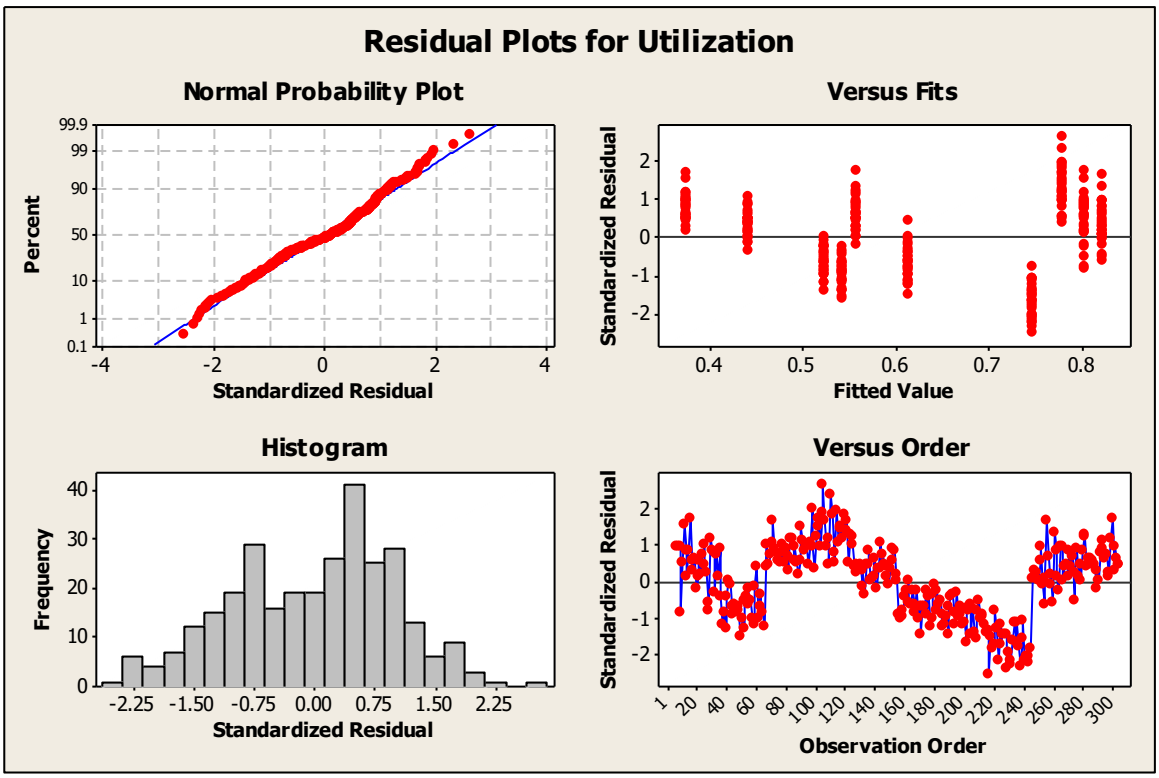
Unusual Observations

Obs	Arr. Mean	Utilization	Fit	SE Fit	Residual	St Resid
105	2.00	0.89580	0.78008	0.00520	0.11572	2.66R
110	2.00	0.88280	0.78008	0.00520	0.10272	2.36R
217	2.50	0.63700	0.74770	0.00331	-0.11070	-2.53R
223	2.50	0.65540	0.74770	0.00331	-0.09230	-2.11R
229	2.50	0.64520	0.74770	0.00331	-0.10250	-2.34R
231	2.50	0.64900	0.74770	0.00331	-0.09870	-2.26R
232	2.50	0.65410	0.74770	0.00331	-0.09360	-2.14R
238	2.50	0.64780	0.74770	0.00331	-0.09990	-2.29R
241	2.50	0.65460	0.74770	0.00331	-0.09310	-2.13R
242	2.50	0.65890	0.74770	0.00331	-0.08880	-2.03R
243	2.50	0.65810	0.74770	0.00331	-0.08960	-2.05R
244	2.50	0.65160	0.74770	0.00331	-0.09610	-2.20R

R denotes an observation with a large standardized residual.

All three variables are statistically significant. The graphs below provide information about the regression analysis assumptions (Normality, variance etc)

Residual Plots for Utilization



Assumptions are not totally compliant. Results have to be taken with care.

We now regress in two variables, as the Response Surface graph (see further ahead) requires only two regressors. The Arrival and Service rates (or means, as used here) are implemented.

Regression Analysis: Utilization versus Arr. Mean, Triang. Mode

The regression equation is:

Utilization = 0.888 - 0.192 Arr. Mean + 0.196 Triang. Mode

Predictor	Coef	SE Coef	T	P
Constant	0.88834	0.01825	48.67	0.000
Arr. Mean	-0.192201	0.003900	-49.28	0.000
Triang. Mode	0.19638	0.01003	19.57	0.000

S = 0.0517662 R-Sq = 89.3% R-Sq(adj) = 89.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	6.6678	3.3339	1244.11	0.000
Residual Error	297	0.7959	0.0027		
Lack of Fit	7	0.6848	0.0978	255.47	0.000
Pure Error	290	0.1111	0.0004		
Total	299	7.4637			

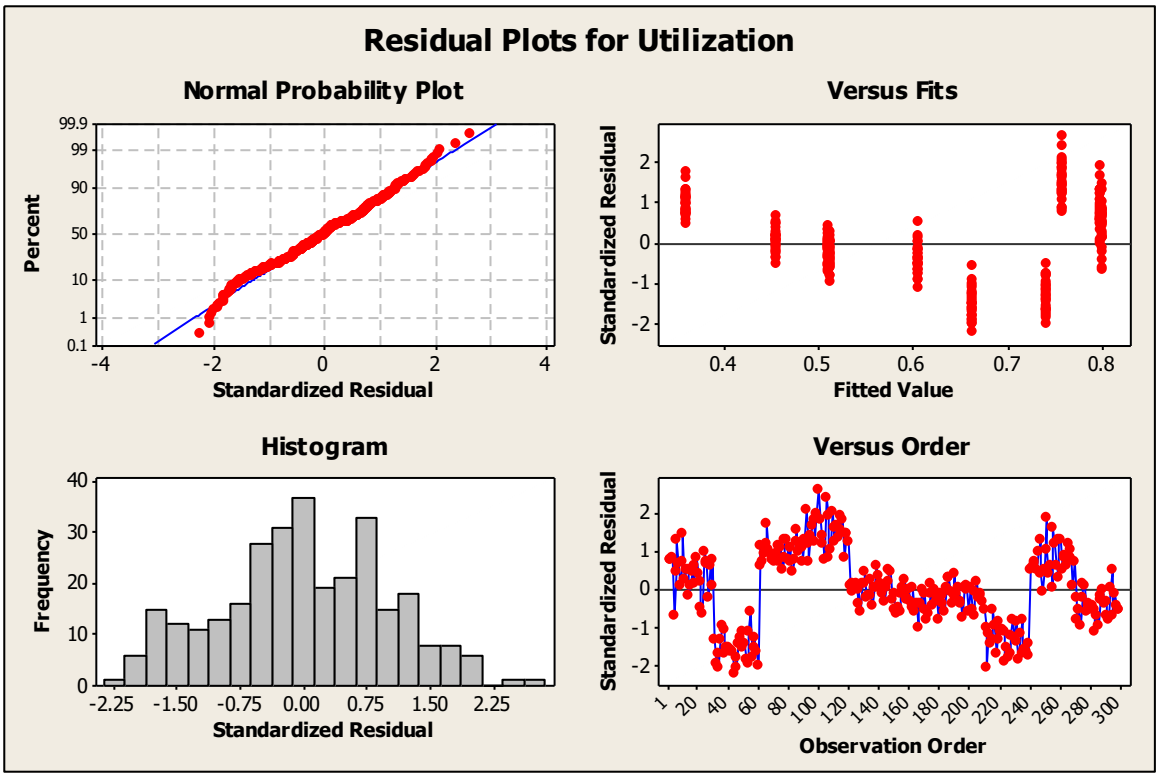
Source	DF	Seq SS
Arr. Mean	1	5.6413
Triang. Mode	1	1.0264

Unusual Observations

Obs	Arr. Mean	Utilization	Fit	SE Fit	Residual	St Resid	
1	3.00	0.84320	0.80270	0.00903	0.04050	0.79	X
2	3.00	0.84390	0.80270	0.00903	0.04120	0.81	X
3	3.00	0.84460	0.80270	0.00903	0.04190	0.82	X
4	3.00	0.76670	0.80270	0.00903	-0.03600	-0.71	X
5	3.00	0.82650	0.80270	0.00903	0.02380	0.47	X
6	3.00	0.87010	0.80270	0.00903	0.06740	1.32	X
7	3.00	0.83850	0.80270	0.00903	0.03580	0.70	X
8	3.00	0.80870	0.80270	0.00903	0.00600	0.12	X
9	3.00	0.83910	0.80270	0.00903	0.03640	0.71	X
10	3.00	0.87770	0.80270	0.00903	0.07500	1.47	X
11	3.00	0.81700	0.80270	0.00903	0.01430	0.28	X
12	3.00	0.82870	0.80270	0.00903	0.02600	0.51	X
13	3.00	0.82930	0.80270	0.00903	0.02660	0.52	X
14	3.00	0.79500	0.80270	0.00903	-0.00770	-0.15	X
15	3.00	0.80820	0.80270	0.00903	0.00550	0.11	X
16	3.00	0.82690	0.80270	0.00903	0.02420	0.47	X
17	3.00	0.81130	0.80270	0.00903	0.00860	0.17	X
18	3.00	0.83580	0.80270	0.00903	0.03310	0.65	X
19	3.00	0.84650	0.80270	0.00903	0.04380	0.86	X
20	3.00	0.82400	0.80270	0.00903	0.02130	0.42	X
21	3.00	0.81410	0.80270	0.00903	0.01140	0.22	X
22	3.00	0.77950	0.80270	0.00903	-0.02320	-0.46	X
23	3.00	0.77000	0.80270	0.00903	-0.03270	-0.64	X
24	3.00	0.85310	0.80270	0.00903	0.05040	0.99	X
25	3.00	0.83940	0.80270	0.00903	0.03670	0.72	X
26	3.00	0.83680	0.80270	0.00903	0.03410	0.67	X
27	3.00	0.79140	0.80270	0.00903	-0.01130	-0.22	X
28	3.00	0.83470	0.80270	0.00903	0.03200	0.63	X
29	3.00	0.80920	0.80270	0.00903	0.00650	0.13	X

```
 30   3.00       0.84220   0.80270   0.00903    0.03950       0.77 X
 34   2.70       0.55830   0.66398   0.00355   -0.10568      -2.05R
 44   2.70       0.54870   0.66398   0.00355   -0.11528      -2.23R
 46   2.70       0.55810   0.66398   0.00355   -0.10588      -2.05R
 92   2.00       0.86660   0.75924   0.00570    0.10736       2.09R
 99   2.00       0.86220   0.75924   0.00570    0.10296       2.00R
100   2.00       0.89580   0.75924   0.00570    0.13656       2.65R
105   2.00       0.88280   0.75924   0.00570    0.12356       2.40R
109   2.00       0.86520   0.75924   0.00570    0.10596       2.06R
212   2.50       0.63700   0.74169   0.00385   -0.10469      -2.03R

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
```

Residual Plots for Utilization

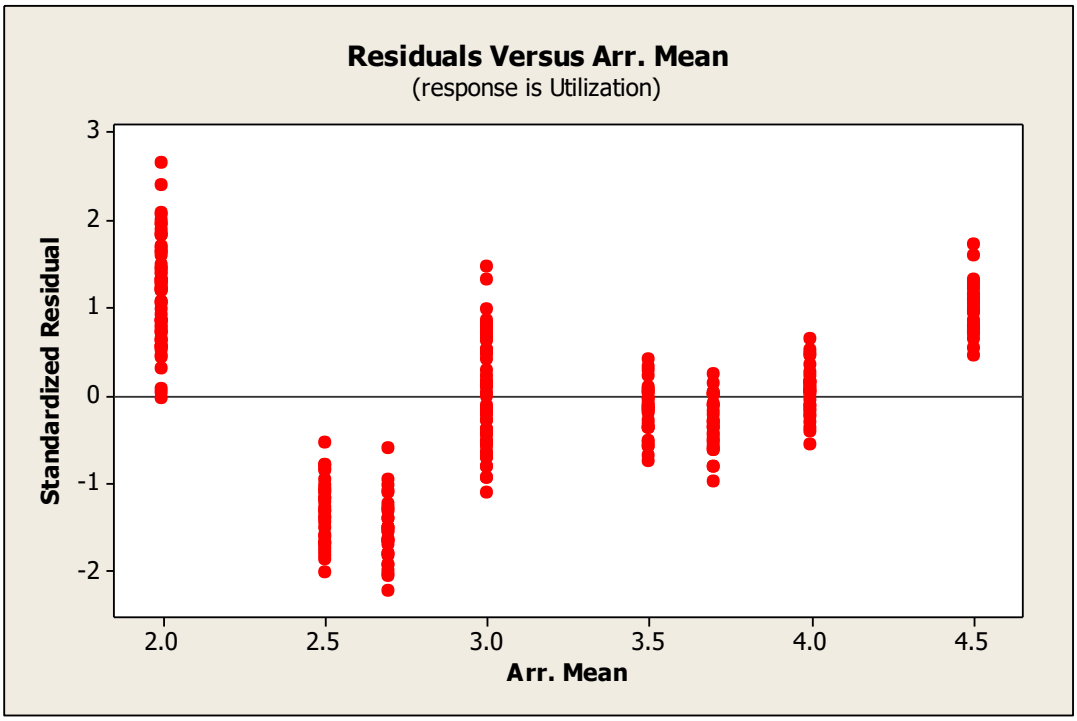


Comparison of Full (with three regressors) and Reduced (w/o Sec. Probability) Models:

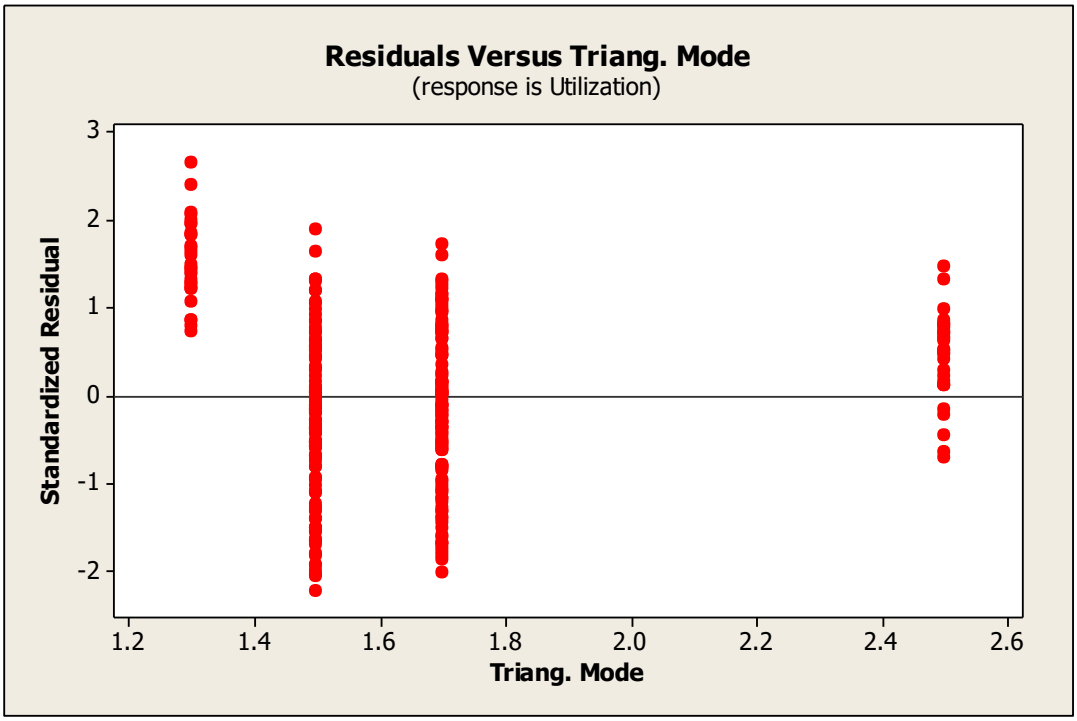
$$F = \frac{\left[\dfrac{SSR_{RM} - SSR_{FM}}{DF_{RM} - DF_{FM}}\right]}{\left[\dfrac{SSR_{FM}}{DF_{FM}}\right]} = \frac{\left[\dfrac{R^2_{FM} - R^2_{RM}}{\Delta_{DF}}\right]}{\dfrac{1 - R^2_{FM}}{DF_{FM}}}$$

Equations above are used to assess whether a Full Model (with all variables) or the Reduced Model (with fewer ones) is adequate (as per information provided).

Residuals from Utilization vs Arr. Mean



Residuals from Utilization vs Triang. Mode



We can do similar analyses using other PM such as time in queue, No. in queue, etc.

Regression Analysis: Wait in Q versus Arr. Mean, Triang. Mode, Sec. Prob

The regression equation is:

Wait in Q = - 40.9 - 1.93 Arr. Mean + 4.10 Triang. Mode + 44.4 Sec. Prob

Predictor	Coef	SE Coef	T	P
Constant	-40.929	6.192	-6.61	0.000
Arr. Mean	-1.93163	0.08737	-22.11	0.000
Triang. Mode	4.0993	0.2295	17.86	0.000
Sec. Prob	44.421	6.326	7.02	0.000

S = 1.14995 R-Sq = 70.8% R-Sq(adj) = 70.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	947.26	315.75	238.77	0.000
Residual Error	296	391.43	1.32		
Lack of Fit	6	211.71	35.29	56.94	0.000
Pure Error	290	179.72	0.62		
Total	299	1338.69			

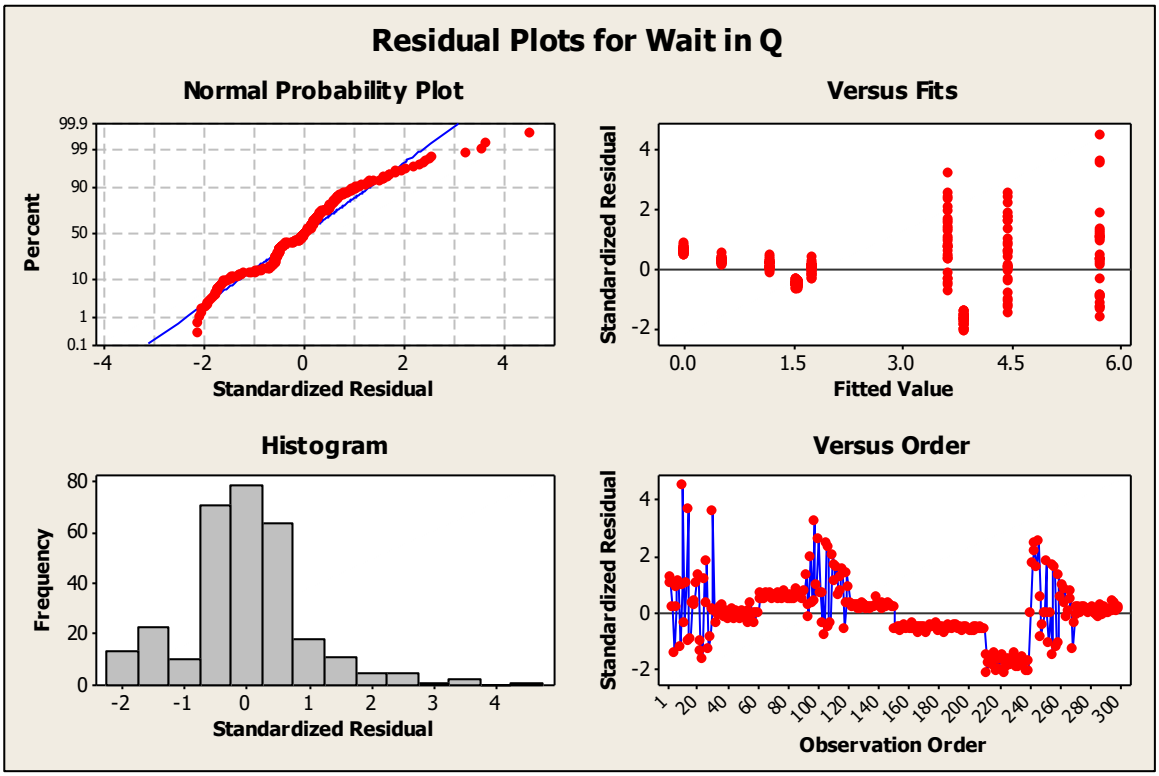
Source	DF	Seq SS
Arr. Mean	1	514.43
Triang. Mode	1	367.62
Sec. Prob	1	65.21

Unusual Observations

Obs	Arr. Mean	Wait in Q	Fit	SE Fit	Residual	St Resid
10	3.00	10.8243	5.7246	0.2006	5.0997	4.50R
13	3.00	9.8448	5.7246	0.2006	4.1202	3.64R
30	3.00	9.7440	5.7246	0.2006	4.0194	3.55R
97	2.00	7.3303	3.6254	0.1363	3.7049	3.24R
100	2.00	6.5444	3.6254	0.1363	2.9190	2.56R
105	2.00	6.4028	3.6254	0.1363	2.7774	2.43R
106	2.00	6.2680	3.6254	0.1363	2.6426	2.31R
109	2.00	5.9581	3.6254	0.1363	2.3327	2.04R
212	2.50	1.4146	3.8552	0.0868	-2.4406	-2.13R
218	2.50	1.5304	3.8552	0.0868	-2.3248	-2.03R
224	2.50	1.4426	3.8552	0.0868	-2.4126	-2.10R
238	2.50	1.4905	3.8552	0.0868	-2.3647	-2.06R
239	2.50	1.5237	3.8552	0.0868	-2.3315	-2.03R
243	2.00	7.2104	4.4453	0.1292	2.7651	2.42R
244	2.00	6.9427	4.4453	0.1292	2.4974	2.19R
246	2.00	7.3223	4.4453	0.1292	2.8770	2.52R

R denotes an observation with a large standardized residual.

Residual Plots for Wait in Q



Regression Analysis: No. in Q versus Arr. Mean, Triang. Mode, Sec. Prob

The regression equation is:

No. in Q = - 21.6 - 0.905 Arr. Mean + 1.21 Triang. Mode + 24.4 Sec. Prob

Predictor	Coef	SE Coef	T	P
Constant	-21.615	2.783	-7.77	0.000
Arr. Mean	-0.90523	0.03927	-23.05	0.000
Triang. Mode	1.2118	0.1031	11.75	0.000
Sec. Prob	24.396	2.843	8.58	0.000

S = 0.516832 R-Sq = 69.8% R-Sq(adj) = 69.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	182.864	60.955	228.20	0.000
Residual Error	296	79.066	0.267		
Lack of Fit	6	47.207	7.868	71.62	0.000
Pure Error	290	31.859	0.110		
Total	299	261.930			

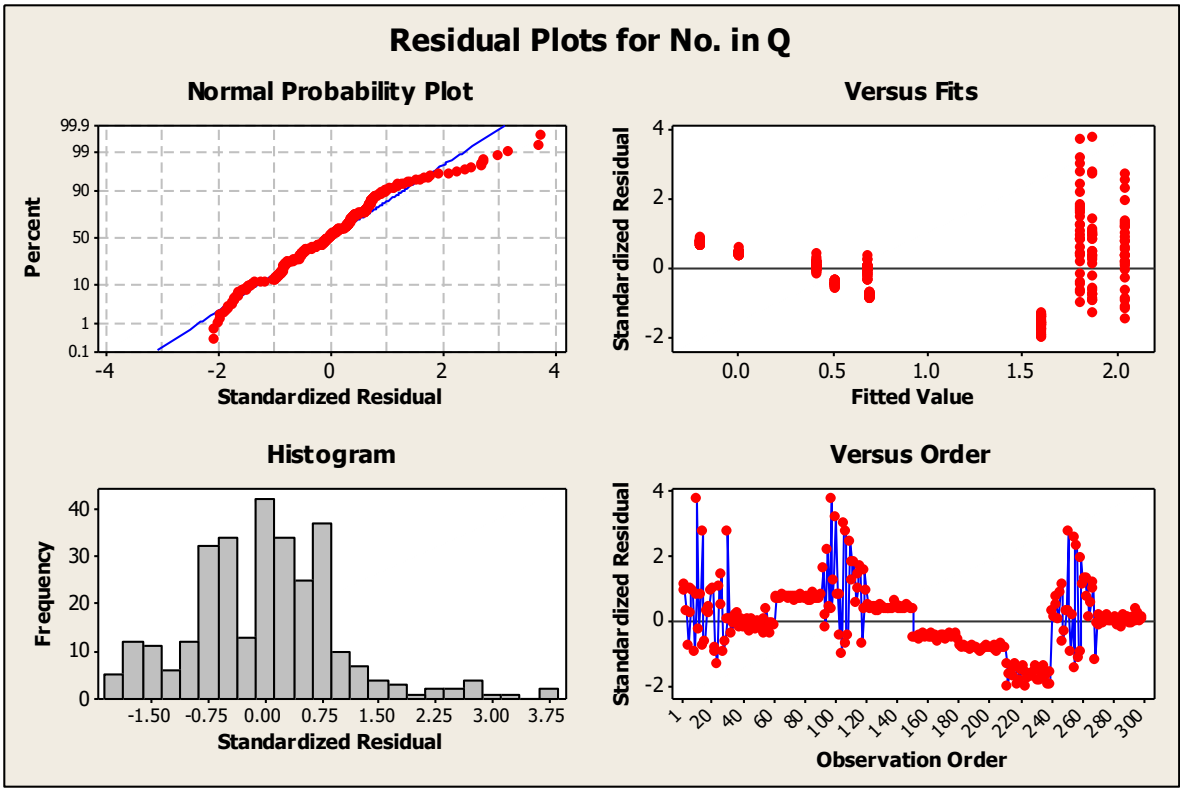
Source	DF	Seq SS
Arr. Mean	1	136.496
Triang. Mode	1	26.700
Sec. Prob	1	19.668

Unusual Observations

Obs	Arr. Mean	No. in Q	Fit	SE Fit	Residual	St Resid
10	3.00	3.7858	1.8746	0.0902	1.9112	3.76R
13	3.00	3.2808	1.8746	0.0902	1.4062	2.76R
30	3.00	3.2594	1.8746	0.0902	1.3848	2.72R
95	2.00	2.9142	1.8135	0.0613	1.1007	2.14R
97	2.00	3.7279	1.8135	0.0613	1.9144	3.73R
100	2.00	3.4513	1.8135	0.0613	1.6378	3.19R
105	2.00	3.3514	1.8135	0.0613	1.5379	3.00R
106	2.00	3.2228	1.8135	0.0613	1.4093	2.75R
109	2.00	3.0487	1.8135	0.0613	1.2352	2.41R
212	2.50	0.5432	1.6017	0.0390	-1.0585	-2.05R
224	2.50	0.5445	1.6017	0.0390	-1.0572	-2.05R
251	2.00	3.4513	2.0559	0.0581	1.3954	2.72R
256	2.00	3.3514	2.0559	0.0581	1.2955	2.52R
257	2.00	3.2228	2.0559	0.0581	1.1669	2.27R

R denotes an observation with a large standardized residual.

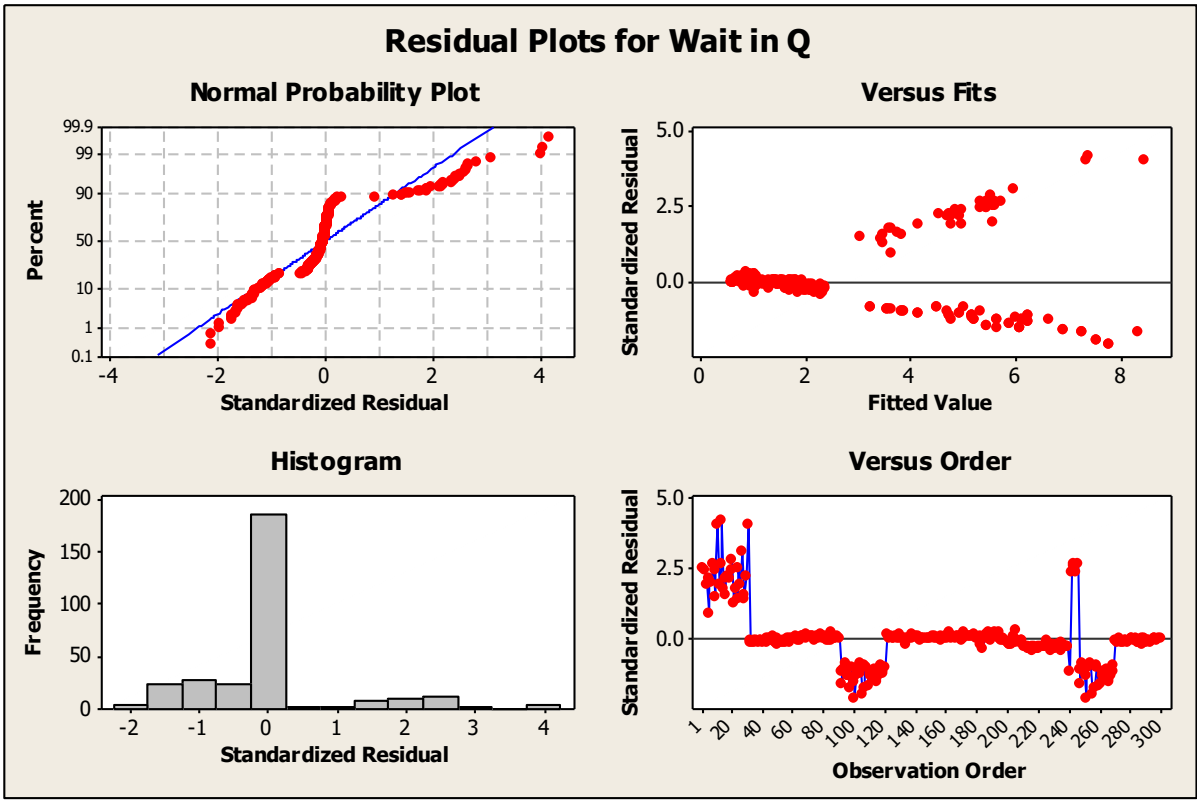
Residual Plots for No. in Q



Regression Analysis: Wait in Q versus No. in Q, Utilization

```
The regression equation is:

Wait in Q = 0.109 + 2.08 No. in Q + 0.549 Utilization

Predictor       Coef  SE Coef      T      P
Constant      0.1086   0.2713   0.40  0.689
No. in Q     2.08195  0.09508  21.90  0.000
Utilization   0.5485   0.5632   0.97  0.331

S = 0.603596   R-Sq = 91.9%   R-Sq(adj) = 91.9%

Analysis of Variance

Source           DF       SS      MS        F      P
Regression        2  1230.48  615.24  1688.69  0.000
Residual Error  297   108.21    0.36
Total           299  1338.69

Sum of squares for pure error is (nearly) zero.
Cannot do pure error test.


Source       DF   Seq SS
No. in Q      1  1230.13
Utilization   1     0.35
```

Residual Plots for Wait in Q



We have Ten Groups of Data, corresponding to Ten different Scenarios. Each scenario could be characterized by some input variables. We will attempt to identify them via an ANOVA on Util.

One-way ANOVA: Util versus Grp No

```
Source  DF        SS        MS       F      P
Grp No   8  1.191740  0.148968  353.37  0.000
Error   36  0.015176  0.000422
Total   44  1.206916

S = 0.02053   R-Sq = 98.74%   R-Sq(adj) = 98.46%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level  N     Mean    StDev  --------+---------+---------+---------+-
1      5  0.82498  0.03344                                    (-*)
2      5  0.40534  0.01282  (-*)
3      5  0.82906  0.02508                                     (*-)
4      5  0.45712  0.00440       (*-)
5      5  0.48730  0.01020         (-*)
6      5  0.50102  0.00848          (-*)
7      5  0.66980  0.02003                      (-*)
8      5  0.83206  0.02674                                     (*-)
9      5  0.59108  0.02389                 (*-)
                          --------+---------+---------+---------+-
                                0.48      0.60      0.72      0.84

Pooled StDev = 0.02053
```

ANOVA is highly Significant: different scenarios actually impact Utilization.

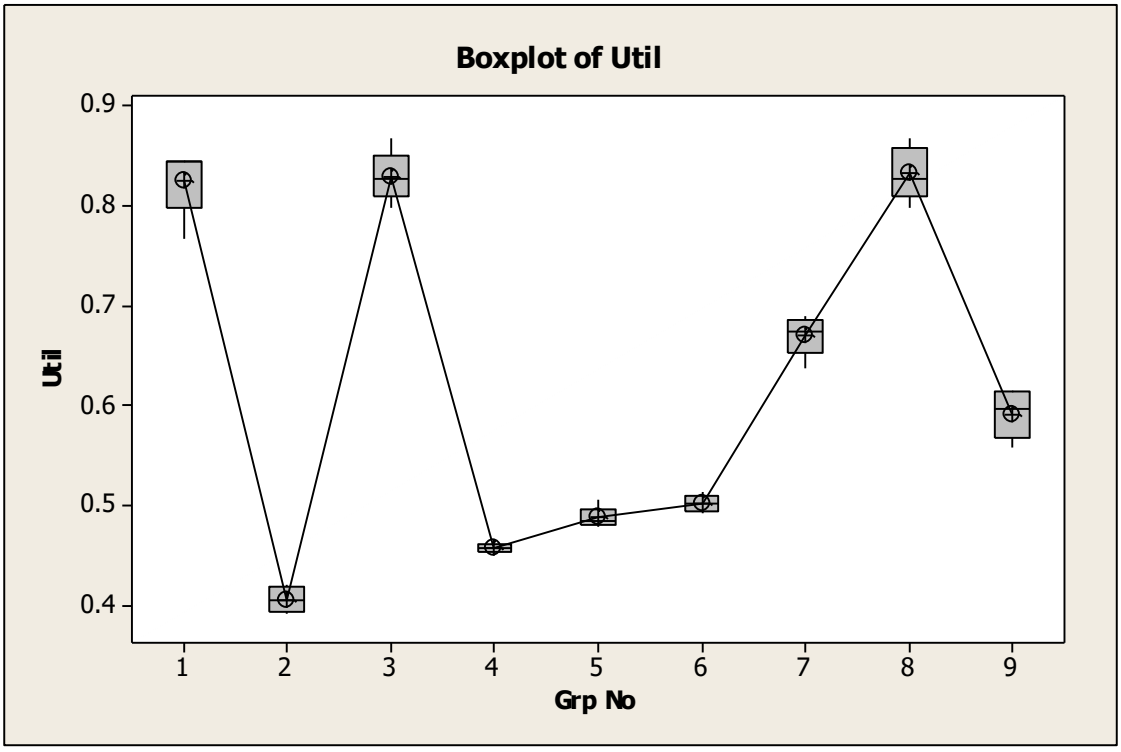
Tukey 95% Simultaneous Confidence Intervals

```
All Pairwise Comparisons among Levels of Grp No

Individual confidence level = 99.78%

Grp No = 1 subtracted from:
No      Lower    Center     Upper  ---------+---------+---------+---------+
2    -0.46243  -0.41964  -0.37685   (*-)
3    -0.03871   0.00408   0.04687                  (-*-)
4    -0.41065  -0.36786  -0.32507     (*-)
5    -0.38047  -0.33768  -0.29489      (*-)
6    -0.36675  -0.32396  -0.28117      (-*-)
7    -0.19797  -0.15518  -0.11239             (-*-)
8    -0.03571   0.00708   0.04987                   (*-)
9    -0.27669  -0.23390  -0.19111          (-*)
                         ---------+---------+---------+---------+
                               -0.25      0.00      0.25      0.50


Grp No = 2 subtracted from:
No      Lower    Center     Upper  ---------+---------+---------+---------+
3    0.38093  0.42372  0.46651                                    (-*-)
4    0.00899  0.05178  0.09457                      (-*-)
5    0.03917  0.08196  0.12475                        (*-)
6    0.05289  0.09568  0.13847                        (-*-)
7    0.22167  0.26446  0.30725                               (-*)
8    0.38393  0.42672  0.46951                                    (-*-)
9    0.14295  0.18574  0.22853                          (*-)
                         ---------+---------+---------+---------+
                               -0.25      0.00      0.25      0.50
```

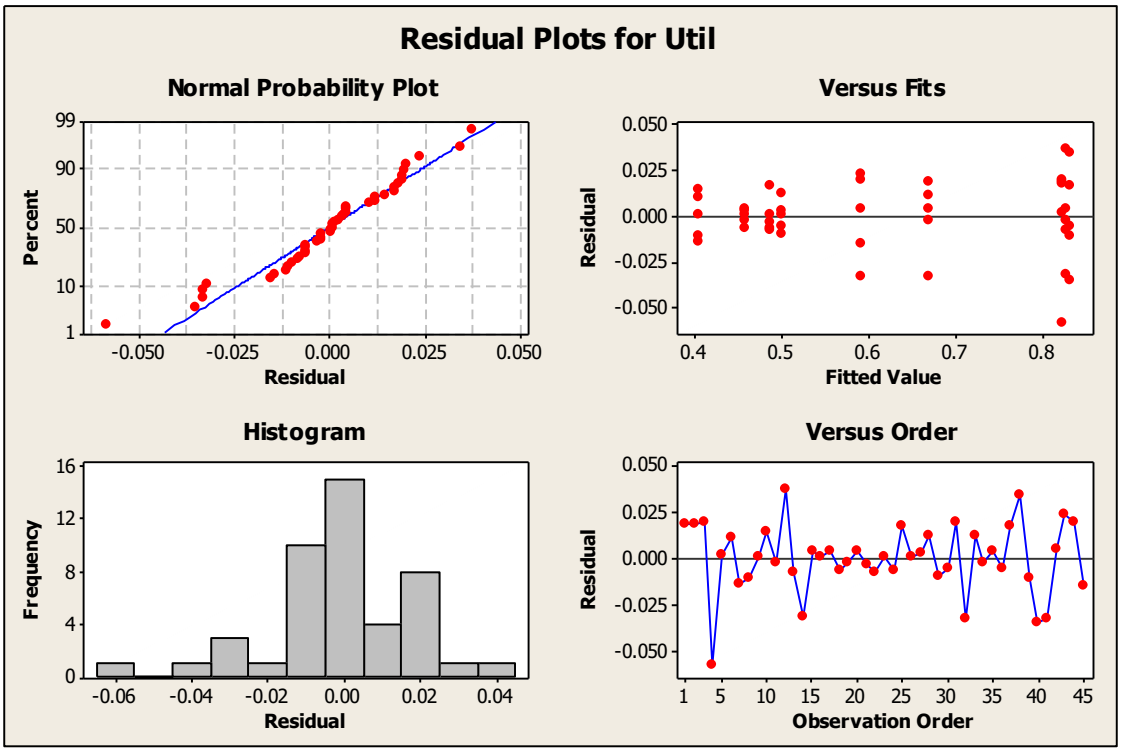
```
Grp No = 3 subtracted from:
No      Lower    Center     Upper  ---------+---------+---------+---------+
4    -0.41473  -0.37194  -0.32915    (-*-)
5    -0.38455  -0.34176  -0.29897      (*-)
6    -0.37083  -0.32804  -0.28525      (-*-)
7    -0.20205  -0.15926  -0.11647             (-*)
8    -0.03979   0.00300   0.04579                  (-*-)
9    -0.28077  -0.23798  -0.19519          (*-)
                                  ---------+---------+---------+---------+
                                        -0.25      0.00      0.25      0.50


Grp No = 4 subtracted from:
No      Lower   Center    Upper  ---------+---------+---------+---------+
5    -0.01261  0.03018  0.07297                   (-*-)
6     0.00111  0.04390  0.08669                    (-*)
7     0.16989  0.21268  0.25547                          (-*)
8     0.33215  0.37494  0.41773                                (-*-)
9     0.09117  0.13396  0.17675                       (*-)
                                  ---------+---------+---------+---------+
                                        -0.25      0.00      0.25      0.50


Grp No = 5 subtracted from:
No      Lower   Center    Upper  ---------+---------+---------+---------+
6    -0.02907  0.01372  0.05651                   (-*)
7     0.13971  0.18250  0.22529                          (*-)
8     0.30197  0.34476  0.38755                               (-*-)
9     0.06099  0.10378  0.14657                     (-*-)
                                  ---------+---------+---------+---------+
                                        -0.25      0.00      0.25      0.50

Grp No = 6 subtracted from:
No     Lower   Center    Upper  ---------+---------+---------+---------+
7    0.12599  0.16878  0.21157                         (-*)
8    0.28825  0.33104  0.37383                               (*-)
9    0.04727  0.09006  0.13285                     (-*)
                                 ---------+---------+---------+---------+
                                       -0.25      0.00      0.25      0.50

Grp No = 7 subtracted from:
Grp
No      Lower    Center     Upper  ---------+---------+---------+---------+
8     0.11947   0.16226   0.20505                          (*-)
9    -0.12151  -0.07872  -0.03593               (-*-)
                                  ---------+---------+---------+---------+
                                        -0.25      0.00      0.25      0.50

Grp No = 8 subtracted from:
Grp
No      Lower    Center     Upper  ---------+---------+---------+---------+
9    -0.28377  -0.24098  -0.19819          (*-)
                                  ---------+---------+---------+---------+
                                        -0.25      0.00      0.25      0.50
```

From Tukey's comparisons we can see which scenarios differ from which other.

The Graph of the Boxplot of Util shows how there are three groups of scenarios: one of very high Utilization, one of very low utilization and one in between. If we can establish an equation with variables that characterizes them, we will have a very useful tool.



Residual Plots for Util show how ANOVA assumptions fit (or not) the model.



We will implement a Discriminant Analysis between Three groups: the high Utilizations versus the low Utilizations. The scenarios with intermediate Utilizations serve as a buffer in this work:

Discriminant Analysis: DisGrp versus Mean, Mode, W-Q, Util, Sec, InQ

Linear Method for Response: DisGrp

Predictors: Mean, Mode, W-Q, Util, Sec, InQ

Group	1	2	3
Count	15	10	20

Summary of classification

	True Group		
Put into Group	1	2	3
1	15	0	0
2	0	10	0
3	0	0	20
Total N	15	10	20
N correct	15	10	20
Proportion	1.000	1.000	1.000

N = 45 N Correct = 45 Proportion Correct = 1.000

Squared Distance Between the Three Utilization Groups

	1	2	3
1	0.000	53.791	152.471
2	53.791	0.000	63.815
3	152.471	63.815	0.000

Linear Discriminant Function for the Three Groups

	1	2	3
Constant	-15129	-15092	-16116
Mean	322	312	330
Mode	207	214	214
W-Q	22	20	23
Util	-1214	-1345	-1524
Sec	31311	31416	32522
InQ	-61	-79	-78

		Means for Group		
Variable	Pooled Mean	1	2	3
Mean	3.2222	2.3333	2.7500	4.1250
Mode	1.6778	1.7667	1.6000	1.6500
W-Q	2.4772	5.2751	1.5588	0.8379
Util	0.62197	0.82870	0.63044	0.46269
Sec	0.95889	0.96333	0.95000	0.96000
InQ	0.95715	2.14929	0.58680	0.24823

		StDev for Group		
Variable	Pooled StDev	1	2	3
Mean	0.4196	0.4880	0.2635	0.4253
Mode	0.3231	0.5434	0.1054	0.0889
W-Q	0.7862	1.3232	0.3454	0.1405
Util	0.03709	0.02669	0.04640	0.03867
Sec	0.008909	0.009759	0.010541	0.007255
InQ	0.2699	0.4066	0.1820	0.1533

Pooled Covariance Matrix

	Mean	Mode	W-Q	Util	Sec	InQ
Mean	0.176091					
Mode	0.096230	0.104365				
W-Q	0.059310	0.109369	0.618088			
Util	-0.007016	-0.000427	0.011930	0.001375		
Sec	-0.002778	-0.001746	-0.001003	0.000155	0.000079	
InQ	-0.047773	-0.015797	0.068601	0.004364	0.000927	0.072841

Covariance matrix for Group 1

	Mean	Mode	W-Q	Util	Sec	InQ
Mean	0.23810					
Mode	0.26190	0.29524				
W-Q	0.26770	0.31783	1.75080			
Util	-0.00133	-0.00135	0.01931	0.00071		
Sec	-0.00476	-0.00524	-0.00535	0.00003	0.00010	
InQ	-0.03376	-0.03993	0.14952	0.00308	0.00068	0.16532

Covariance matrix for Group 2

	Mean	Mode	W-Q	Util	Sec	InQ
Mean	0.06944					
Mode	-0.02778	0.01111				
W-Q	-0.07365	0.02946	0.11927			
Util	-0.01093	0.00437	0.01504	0.00215		
Sec	-0.00278	0.00111	0.00295	0.00044	0.00011	
InQ	-0.04119	0.01648	0.06253	0.00812	0.00165	0.03311

Covariance matrix for Group 3

	Mean	Mode	W-Q	Util	Sec	InQ
Mean	0.18092					
Mode	0.03289	0.00789				
W-Q	-0.03126	-0.00638	0.01974			
Util	-0.00935	-0.00202	0.00502	0.00150		
Sec	-0.00132	-0.00053	0.00033	0.00012	0.00005	
InQ	-0.06122	-0.01330	0.01185	0.00353	0.00077	0.02352

From the above analysis we obtain: (1) three equations for each of the three Groups. A new scenario, characterized with these variables, will be implemented in them. It will be classified into the Group with the Shortest Distance and highest probability.

Summary of Classified Observations

Observation	True Group	Pred Group	Group	Squared Distance	Probability
1	1	1	1	8.881	1.000
			2	87.386	0.000
			3	188.154	0.000
2	1	1	1	7.591	1.000
			2	82.785	0.000
			3	185.023	0.000
3	1	1	1	6.018	1.000
			2	66.278	0.000
			3	174.261	0.000
4	1	1	1	15.10	0.999
			2	29.25	0.001

			3	119.12	0.000
5	1	1	1	5.246	1.000
			2	59.702	0.000
			3	161.100	0.000
6	3	3	1	214.230	0.000
			2	108.065	0.000
			3	6.576	1.000
7	3	3	1	211.858	0.000
			2	106.215	0.000
			3	5.997	1.000
8	3	3	1	202.796	0.000
			2	100.637	0.000
			3	4.418	1.000
9	3	3	1	191.766	0.000
			2	95.110	0.000
			3	3.174	1.000
10	3	3	1	194.170	0.000
			2	97.168	0.000
			3	3.612	1.000
11	1	1	1	7.514	1.000
			2	74.829	0.000
			3	171.708	0.000
12	1	1	1	10.78	1.000
			2	57.37	0.000
			3	168.58	0.000
13	1	1	1	6.541	1.000
			2	42.221	0.000
			3	144.103	0.000
14	1	1	1	18.98	1.000
			2	86.65	0.000
			3	174.32	0.000
15	1	1	1	7.063	1.000
			2	58.914	0.000
			3	153.772	0.000
16	3	3	1	134.250	0.000
			2	43.390	0.000
			3	4.165	1.000
17	3	3	1	133.904	0.000
			2	41.927	0.000
			3	4.225	1.000
18	3	3	1	140.384	0.000
			2	45.025	0.000
			3	4.218	1.000
19	3	3	1	137.038	0.000
			2	43.967	0.000
			3	4.164	1.000
20	3	3	1	133.370	0.000
			2	41.418	0.000
			3	4.271	1.000
21	3	3	1	154.520	0.000
			2	73.797	0.000
			3	4.163	1.000
22	3	3	1	157.197	0.000
			2	74.947	0.000
			3	3.668	1.000
23	3	3	1	152.534	0.000
			2	73.090	0.000
			3	4.531	1.000
24	3	3	1	154.765	0.000
			2	74.404	0.000
			3	3.895	1.000
25	3	3	1	144.249	0.000
			2	69.683	0.000

			3	7.025	1.000
26	3	3	1	134.517	0.000
			2	53.816	0.000
			3	2.984	1.000
27	3	3	1	133.425	0.000
			2	53.491	0.000
			3	2.996	1.000
28	3	3	1	127.333	0.000
			2	49.846	0.000
			3	3.207	1.000
29	3	3	1	141.611	0.000
			2	57.540	0.000
			3	3.161	1.000
30	3	3	1	139.025	0.000
			2	56.275	0.000
			3	3.069	1.000
31	2	2	1	37.282	0.000
			2	3.696	1.000
			3	71.621	0.000
32	2	2	1	67.146	0.000
			2	4.555	1.000
			3	58.560	0.000
33	2	2	1	44.480	0.000
			2	3.210	1.000
			3	70.420	0.000
34	2	2	1	48.069	0.000
			2	2.900	1.000
			3	64.989	0.000
35	2	2	1	48.922	0.000
			2	3.047	1.000
			3	68.187	0.000
36	1	1	1	8.196	1.000
			2	58.108	0.000
			3	148.567	0.000
37	1	1	1	13.16	1.000
			2	87.66	0.000
			3	180.28	0.000
38	1	1	1	5.259	1.000
			2	82.114	0.000
			3	190.558	0.000
39	1	1	1	8.092	1.000
			2	34.146	0.000
			3	136.854	0.000
40	1	1	1	4.705	1.000
			2	32.592	0.000
			3	123.798	0.000
41	2	2	1	78.085	0.000
			2	4.704	1.000
			3	58.195	0.000
42	2	2	1	64.324	0.000
			2	2.908	1.000
			3	69.165	0.000
43	2	2	1	57.221	0.000
			2	3.575	1.000
			3	75.833	0.000
44	2	2	1	57.168	0.000
			2	3.366	1.000
			3	73.939	0.000
45	2	2	1	70.562	0.000
			2	3.383	1.000
			3	62.579	0.000

We see how all the data have been re-classified in their correct groups using these equations

The previous analysis was implemented using the Minitab Discriminant procedure. The discriminant analysis below can be implemented with any regression package. It is limited to TWO groups.

Regression Analysis: DisGrp_1 versus Mean_1, Mode_1, ...

The regression equation is:

```
DisGrp_1 = 12.3 - 0.272 Mean_1 + 0.084 Mode_1 + 0.0511 W-Q_1 + 1.36 Util_1
           - 12.8 Sec_1 - 0.167 InQ_1
```

Predictor	Coef	SE Coef	T	P
Constant	12.294	9.511	1.29	0.207
Mean_1	-0.2723	0.1826	-1.49	0.147
Mode_1	0.0841	0.2758	0.31	0.763
W-Q_1	0.05111	0.06706	0.76	0.452
Util_1	1.358	1.264	1.07	0.292
Sec_1	-12.820	9.580	-1.34	0.192
InQ_1	-0.1674	0.1711	-0.98	0.336

S = 0.248506 R-Sq = 79.8% R-Sq(adj) = 75.5%

This equation has too many regressors (six). We will find ways to reduce their number.

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	6	6.8423	1.1404	18.47	0.000
Residual Error	28	1.7291	0.0618		
Total	34	8.5714			

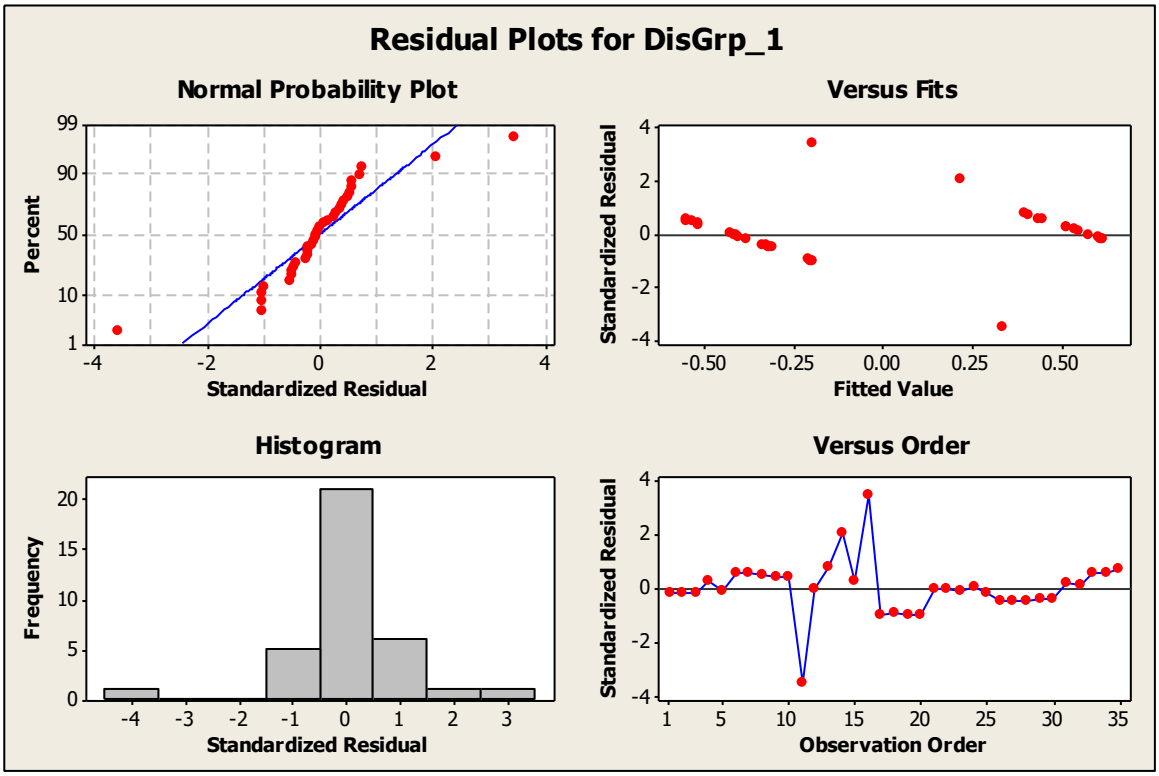
Source	DF	Seq SS
Mean_1	1	5.2037
Mode_1	1	1.3362
W-Q_1	1	0.0919
Util_1	1	0.0319
Sec_1	1	0.1195
InQ_1	1	0.0591

Obs	Mean_1	DisGrp_1	Fit	SE Fit	Residual	St Resid
1	3.00	0.5720	0.6114	0.1272	-0.0394	-0.18
2	3.00	0.5720	0.6142	0.1203	-0.0422	-0.19
3	3.00	0.5720	0.6195	0.1141	-0.0475	-0.22
4	3.00	0.5720	0.5146	0.1464	0.0574	0.29
5	3.00	0.5720	0.6032	0.1102	-0.0312	-0.14
6	4.50	-0.4280	-0.5561	0.0825	0.1281	0.55
7	4.50	-0.4280	-0.5527	0.0797	0.1247	0.53
8	4.50	-0.4280	-0.5373	0.0747	0.1093	0.46
9	4.50	-0.4280	-0.5170	0.0735	0.0890	0.37
10	4.50	-0.4280	-0.5211	0.0758	0.0931	0.39
11	2.00	-0.4280	0.3336	0.1264	-0.7616	-3.56R
12	2.00	0.5720	0.5750	0.1435	-0.0030	-0.01
13	2.00	0.5720	0.3997	0.1140	0.1723	0.78
14	2.00	0.5720	0.2195	0.1812	0.3525	2.07R
15	2.00	0.5720	0.5154	0.1221	0.0566	0.26
16	4.00	0.5720	-0.2016	0.1082	0.7736	3.46R
17	4.00	-0.4280	-0.1974	0.1077	-0.2306	-1.03
18	4.00	-0.4280	-0.2122	0.1102	-0.2158	-0.97
19	4.00	-0.4280	-0.2057	0.1090	-0.2223	-1.00
20	4.00	-0.4280	-0.1977	0.1076	-0.2303	-1.03
21	4.50	-0.4280	-0.4129	0.0969	-0.0151	-0.07
22	4.50	-0.4280	-0.4198	0.0932	-0.0082	-0.04
23	4.50	-0.4280	-0.4066	0.0994	-0.0214	-0.09

24	4.50	-0.4280	-0.4288	0.0949	0.0008	0.00
25	4.50	-0.4280	-0.3841	0.1141	-0.0439	-0.20
26	3.50	-0.4280	-0.3216	0.0954	-0.1064	-0.46
27	3.50	-0.4280	-0.3160	0.0944	-0.1120	-0.49
28	3.50	-0.4280	-0.3101	0.0908	-0.1179	-0.51
29	3.50	-0.4280	-0.3371	0.1015	-0.0909	-0.40
30	3.50	-0.4280	-0.3301	0.0991	-0.0979	-0.43
31	2.00	0.5720	0.5370	0.1249	0.0350	0.16
32	2.00	0.5720	0.5520	0.1533	0.0200	0.10
33	2.00	0.5720	0.4369	0.1012	0.1351	0.60
34	2.00	0.5720	0.4471	0.1226	0.1249	0.58
35	2.00	0.5720	0.4072	0.0909	0.1648	0.71

R denotes an observation with a large standardized residual.

Residual Plots for DisGrp_1



Stepwise Regression: DisGrp_1 versus Mean_1, Mode_1, etc is a procedure to find a smaller set of regressors that will accomplish a similar task:

Forward selection. Alpha-to-Enter: 0.25

Response is DisGrp_1 on 6 predictors, with N = 35

Step	1	2	3
Constant	-1.445	7.666	14.673
Util_1	2.33	2.44	1.63
T-Value	10.04	10.56	2.45
P-Value	0.000	0.000	0.020

Sec_1		-9.5	-15.7
T-Value		-1.87	-2.25
P-Value		0.071	0.031
Mean_1			-0.18
T-Value			-1.28
P-Value			0.209
S	0.253	0.244	0.242
R-Sq	75.35	77.77	78.90
R-Sq(adj)	74.60	76.39	76.85
Mallows Cp	3.2	1.8	2.3

Regression Analysis: DisGrp_1 versus Mean_1, Mode_1, Sec_1

The resulting smaller regression equation is:

DisGrp_1 = 12.0 - 0.465 Mean_1 + 0.372 Mode_1 - 11.6 Sec_1

Predictor	Coef	SE Coef	T	P	
Constant	12.036	9.553	1.26	0.217	
Mean_1	-0.46486	0.05202	-8.94	0.000	
Mode_1	0.3723	0.1992	1.87	0.071	
Sec_1	-11.553	9.556	-1.21	0.236	(Non significant; can be removed)

S = 0.250164 R-Sq = 77.4% R-Sq(adj) = 75.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	6.6314	2.2105	35.32	0.000
Residual Error	31	1.9400	0.0626		
Lack of Fit	2	0.3400	0.1700	3.08	0.061
Pure Error	29	1.6000	0.0552		
Total	34	8.5714			

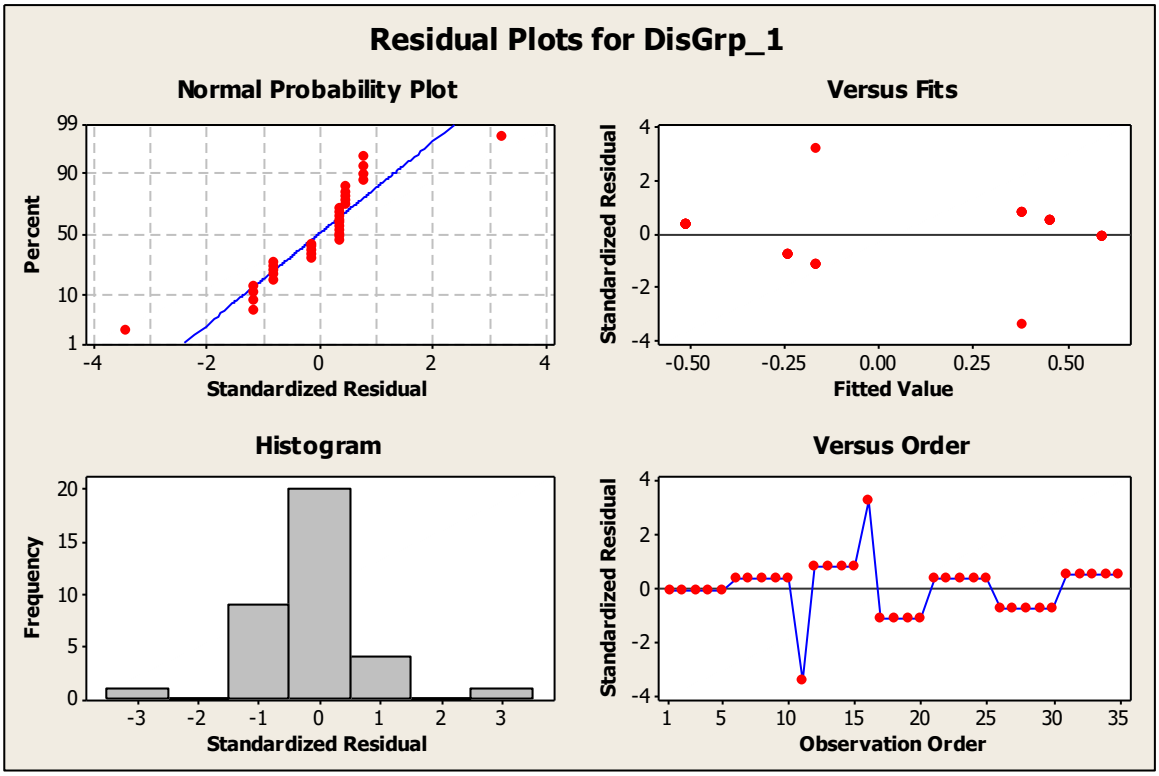
Source	DF	Seq SS
Mean_1	1	5.2037
Mode_1	1	1.3362
Sec_1	1	0.0915

Obs	Mean_1	DisGrp_1	Fit	SE Fit	Residual	St Resid
1	3.00	0.5720	0.5964	0.1104	-0.0244	-0.11
2	3.00	0.5720	0.5964	0.1104	-0.0244	-0.11
3	3.00	0.5720	0.5964	0.1104	-0.0244	-0.11
4	3.00	0.5720	0.5964	0.1104	-0.0244	-0.11
5	3.00	0.5720	0.5964	0.1104	-0.0244	-0.11
6	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
7	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
8	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
9	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
10	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
11	2.00	-0.4280	0.3835	0.0829	-0.8115	-3.44R
12	2.00	0.5720	0.3835	0.0829	0.1885	0.80
13	2.00	0.5720	0.3835	0.0829	0.1885	0.80
14	2.00	0.5720	0.3835	0.0829	0.1885	0.80
15	2.00	0.5720	0.3835	0.0829	0.1885	0.80
16	4.00	0.5720	-0.1662	0.1042	0.7382	3.25R
17	4.00	-0.4280	-0.1662	0.1042	-0.2618	-1.15
18	4.00	-0.4280	-0.1662	0.1042	-0.2618	-1.15
19	4.00	-0.4280	-0.1662	0.1042	-0.2618	-1.15
20	4.00	-0.4280	-0.1662	0.1042	-0.2618	-1.15

21	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
22	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
23	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
24	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
25	4.50	-0.4280	-0.5142	0.0680	0.0862	0.36
26	3.50	-0.4280	-0.2393	0.0738	-0.1887	-0.79
27	3.50	-0.4280	-0.2393	0.0738	-0.1887	-0.79
28	3.50	-0.4280	-0.2393	0.0738	-0.1887	-0.79
29	3.50	-0.4280	-0.2393	0.0738	-0.1887	-0.79
30	3.50	-0.4280	-0.2393	0.0738	-0.1887	-0.79
31	2.00	0.5720	0.4580	0.0738	0.1140	0.48
32	2.00	0.5720	0.4580	0.0738	0.1140	0.48
33	2.00	0.5720	0.4580	0.0738	0.1140	0.48
34	2.00	0.5720	0.4580	0.0738	0.1140	0.48
35	2.00	0.5720	0.4580	0.0738	0.1140	0.48

R denotes an observation with a large standardized residual.

Residual Plots for DisGrp_1



Regression Analysis: DisGrp_1 versus Mean_1, Mode_1

The resulting Two-Variable regression equation is:

DisGrp_1 = 0.489 - 0.431 Mean_1 + 0.563 Mode_1

Predictor	Coef	SE Coef	T	P
Constant	0.4894	0.2336	2.10	0.044
Mean_1	**-0.43062**	**0.04395**	**-9.80**	**0.000**
Mode_1	**0.5629**	**0.1227**	**4.59**	**0.000**

S = 0.251962 R-Sq = 76.3% R-Sq(adj) = 74.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	**2**	**6.5399**	**3.2700**	**51.51**	**0.000**
Residual Error	32	2.0315	0.0635		
Lack of Fit	3	0.4315	0.1438	2.61	0.071
Pure Error	29	1.6000	0.0552		
Total	34	8.5714			

Source	DF	Seq SS
Mean_1	1	5.2037
Mode_1	1	1.3362

Obs	Mean 1	DisGrp 1	Fit	SE Fit	Residual	St Resid
1	3.00	0.5720	0.6047	0.1110	-0.0327	-0.14
2	3.00	0.5720	0.6047	0.1110	-0.0327	-0.14
3	3.00	0.5720	0.6047	0.1110	-0.0327	-0.14
4	3.00	0.5720	0.6047	0.1110	-0.0327	-0.14
5	3.00	0.5720	0.6047	0.1110	-0.0327	-0.14
6	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
7	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
8	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
9	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
10	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
11	**2.00**	**-0.4280**	**0.3598**	**0.0812**	**-0.7878**	**-3.30R**
12	2.00	0.5720	0.3598	0.0812	0.2122	0.89
13	2.00	0.5720	0.3598	0.0812	0.2122	0.89
14	2.00	0.5720	0.3598	0.0812	0.2122	0.89
15	2.00	0.5720	0.3598	0.0812	0.2122	0.89
16	**4.00**	**0.5720**	**-0.2763**	**0.0511**	**0.8483**	**3.44R**
17	4.00	-0.4280	-0.2763	0.0511	-0.1517	-0.62
18	4.00	-0.4280	-0.2763	0.0511	-0.1517	-0.62
19	4.00	-0.4280	-0.2763	0.0511	-0.1517	-0.62
20	4.00	-0.4280	-0.2763	0.0511	-0.1517	-0.62
21	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
22	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
23	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
24	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
25	4.50	-0.4280	-0.4916	0.0659	0.0636	0.26
26	3.50	-0.4280	-0.1735	0.0502	-0.2545	-1.03
27	3.50	-0.4280	-0.1735	0.0502	-0.2545	-1.03
28	3.50	-0.4280	-0.1735	0.0502	-0.2545	-1.03
29	3.50	-0.4280	-0.1735	0.0502	-0.2545	-1.03
30	3.50	-0.4280	-0.1735	0.0502	-0.2545	-1.03
31	2.00	0.5720	0.4724	0.0733	0.0996	0.41
32	2.00	0.5720	0.4724	0.0733	0.0996	0.41
33	2.00	0.5720	0.4724	0.0733	0.0996	0.41
34	2.00	0.5720	0.4724	0.0733	0.0996	0.41
35	2.00	0.5720	0.4724	0.0733	0.0996	0.41

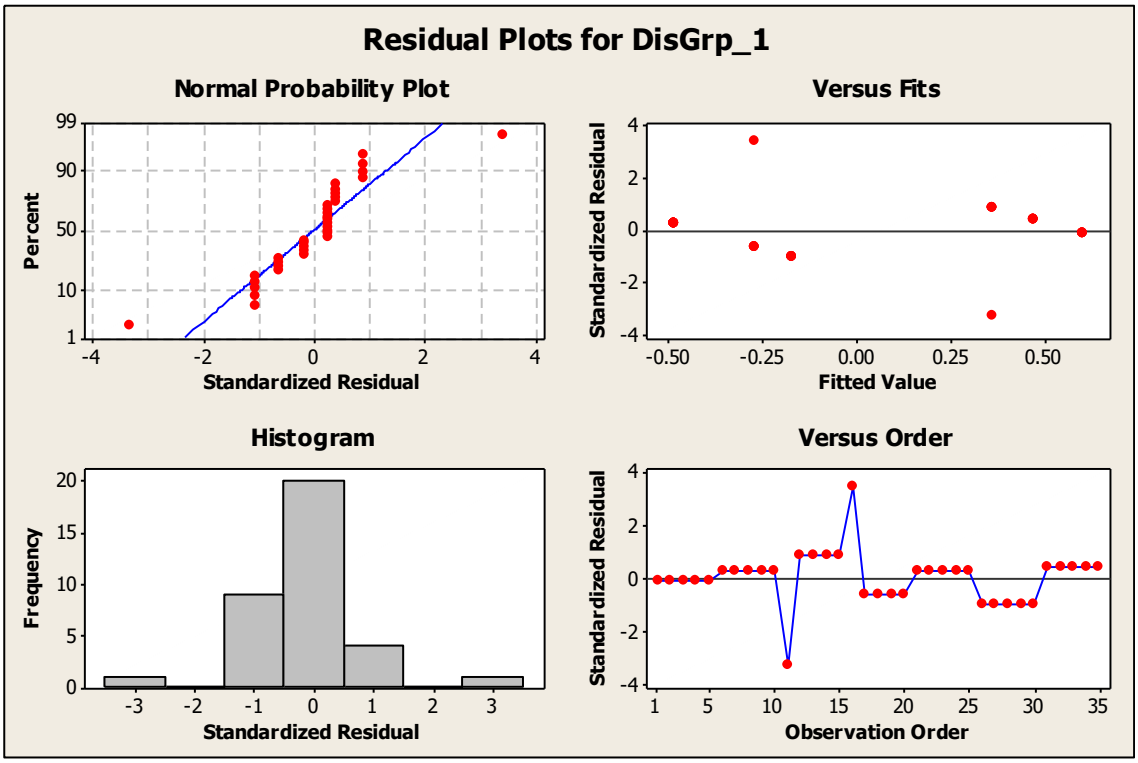
R denotes an observation with a large standardized residual.

We will use this resulting two-variable equation in our Discriminant Analysis.

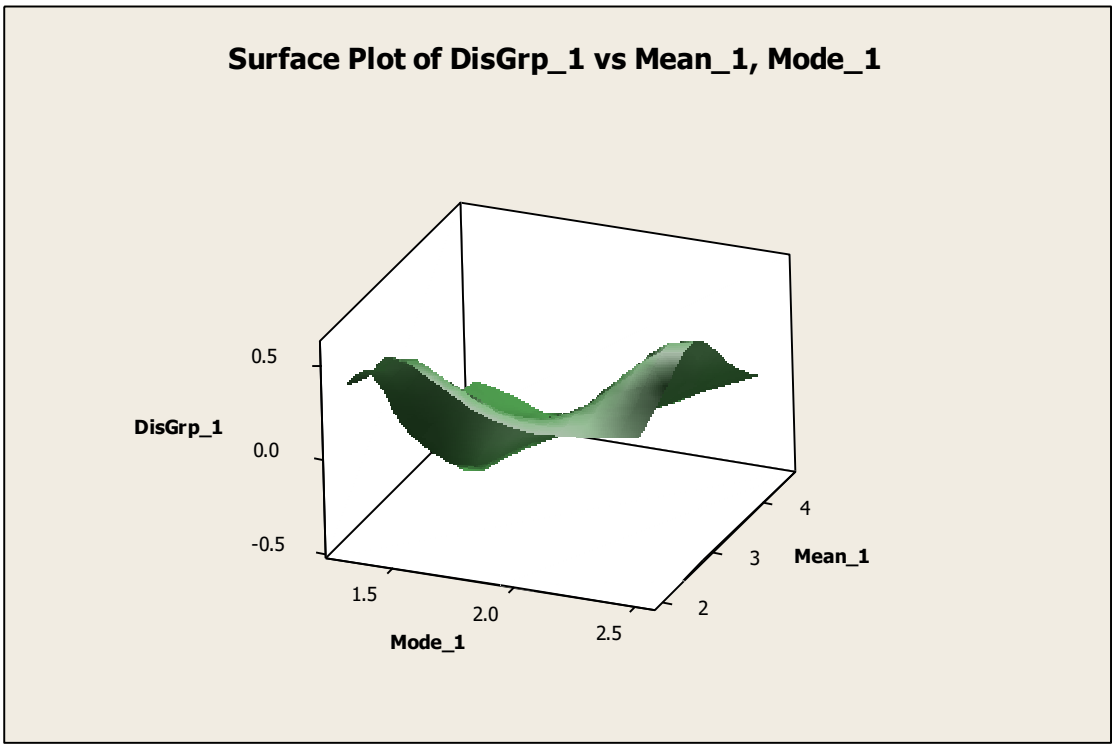
Notice how classification, based upon these two input variables (arrival and service rates) is quite good: only two (out of 35) misclassifications (about 5%).

Both regressors are highly significant (p value is zero) and it is possible to come up with a response surface.

Residual Plots for Assessing the DisGrp_1 regression:



The Surface Plot allows us to easily assess which scenarios, determined by the input and service rates/means, will yield high, intermediate or low utilizations.



END of the ANALISIS.