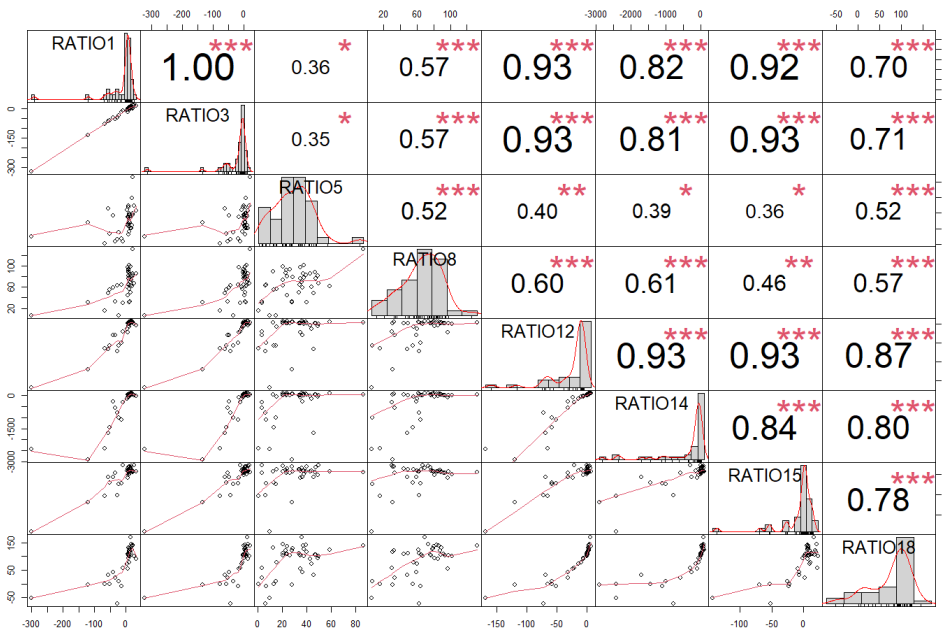


1.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
RATIO1	41	-300,89	33,11	-8,8146	55,46487
RATIO3	41	-321,06	20,29	-17,5937	57,42040
RATIO5	41	,42	86,71	29,9263	17,19171
RATIO8	41	5,87	132,20	65,5649	27,30802
RATIO12	41	-169,23	6,98	-17,9854	37,67320
RATIO14	41	-2913,60	106,99	-345,9383	769,39431
RATIO15	41	-142,82	22,33	-2,7978	30,14706
RATIO18	41	-73,58	170,41	83,8934	57,27756
Valid N (listwise)	41				

Tablo 1.1



Grafik 1.1

Değişkenler arasındaki korelasyonların büyük bir çoğunluğu anlamlıdır.

Aralarındaki korelasyon 0.80'in üzerinde olan 11 adet değişken çifti vardır. Bunlardan 7 tanesi 0.90'ın üzerinde korelasyona sahiptir.

RATIO1 ve RATIO3 arasında tam korelasyon vardır.

RATIO5 ve RATIO8'in diğer değişkenlerle arasında çok yüksek korelasyonlar yoktur.

Coefficients^a

		Collinearity Statistics	
Model		Tolerance	VIF
1	RATIO1	,008	124,847
	RATIO3	,007	144,814
	RATIO5	,688	1,454
	RATIO8	,403	2,479
	RATIO12	,041	24,366
	RATIO14	,114	8,792
	RATIO15	,073	13,664

a. Dependent Variable: RATIO18

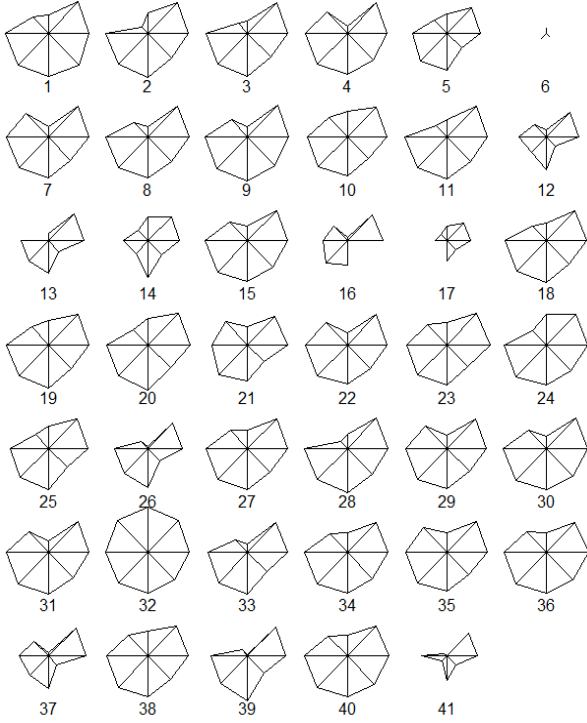
Tablo 1.2

- VIF = 1: Değişkenler arasında korelasyon yok.
- 1 < VIF < 5: Değişkenler arasında orta derecede korelasyon var.
- VIF > 5: Değişkenler arasında yüksek derecede korelasyon var.

RATIO5 ve RATIO8'in diğer değişkenlerle arasında orta derece korelasyon vardır. Geriye kalan değişkenlerin hepsi yüksek korelasyonludur. RATIO1 ve RATIO3 çok yüksek korelasyonlara sahiptir.

Veri setinde çoklu doğrusal bağımlılık vardır.

2.

Yıldız İkon Grafiği

Diğer gözlemlere kıyasla 6, 16, 17 ve 41'inci gözlemlerin aldığı değerlerde gözle görülür farklılıklar vardır.

6. gözlem tüm değişkenler için çok küçük değerler almıştır.

Genel anlamda bozulmalar gözlemin çok küçük değerler aldığı durumlarda ortaya çıkıyor.

Grafik 2.1

3.

Univariate Statistics

	N	Mean	Std. Deviation	Missing		No. of Extremes ^a	
				Count	Percent	Low	High
RATIO1	32	-1,2894	30,70034	9	22,0	3	0
RATIO3	33	-15,7755	59,28758	8	19,5	5	0
RATIO5	31	31,9194	17,49459	10	24,4	0	1
RATIO8	36	65,3361	23,99640	5	12,2	0	0
RATIO12	33	-19,8248	40,71291	8	19,5	2	0
RATIO14	34	-336,8588	753,08836	7	17,1	8	0
RATIO15	33	-,7364	31,41288	8	19,5	3	0
RATIO18	30	90,4433	52,74659	11	26,8	1	0

Gözlemlerin yaklaşık olarak %20'si eksiktir.

En çok eksik verisi olan değişken %26,8 ile RATIO18'dir. RATIO14'te 8 tane, değeri çok düşük aşırı gözlem vardır.

a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

Tablo 3.1

Tabulated Patterns								
Number of Cases	Missing Patterns ^a							
	RATIO8	RATIO3	RATIO12	RATIO18	RATIO15	RATIO14	RATIO5	RATIO1
8								
1						X		
2		X				X		
1		X						
1		X						X
4								X
2	X							
1	X				X			
2				X	X			
1				X	X	X		
1				X	X	X		X
3				X			X	X
2							X	
1							X	
1			X					
2			X	X				
1		X	X	X				
1		X	X	X				
1			X		X			
1					X	X	X	
1						X	X	
1	X						X	
1	X	X						
1		X	X				X	

a. Variables are sorted on missing patterns.
b. Number of complete cases if variables missing in that pattern (marked with X) are not used.

Tablo 3.2

Separate Variance t Tests ^a									
	RATIO1	RATIO3	RATIO5	RATIO8	RATIO12	RATIO14	RATIO15	RATIO18	
RATIO1	t	.	1,1	1,3	,7	,8	,1	1,0	,4
	df	.	7,1	7,8	9,8	9,7	10,1	7,4	4,5
	P(2-tail)	.	,307	,228	,523	,435	,899	,342	,716
	# Present	32	25	25	27	24	26	25	25
	# Missing	0	8	6	9	9	8	8	5
	Mean(Present)	-1,2894	-4,9356	33,8676	67,3433	-15,1646	-326,4254	4,2108	92,9868
	Mean(Missing)	.	-49,6500	23,8017	59,3144	-32,2522	-370,7675	-16,1963	77,7260
RATIO3	t	1,3	.	-1,6	1,5	,4	,9	,3	,2
	df	6,4	.	7,3	16,9	4,8	5,5	10,1	7,5
	P(2-tail)	,245	.	,154	,156	,698	,412	,797	,860
	# Present	25	33	24	29	28	26	24	24
	# Missing	7	0	7	7	5	6	7	6
	Mean(Present)	4,4024	-15,7755	28,5200	67,4048	-18,2618	-256,4332	-.0081	91,3554
	Mean(Missing)	-21,6171	.	43,5743	56,7657	-28,5780	-712,1783	-3,4414	86,7950
RATIO5	t	-,7	,9	.	,5	,4	,1	,6	,3
	df	16,0	8,3	.	10,2	8,5	10,4	9,0	13,6
	P(2-tail)	,465	,389	.	,599	,712	,921	,560	,734
	# Present	25	24	31	27	25	26	24	23
	# Missing	7	9	0	9	8	8	9	7
	Mean(Present)	-2,9536	-6,8058	31,9194	66,9263	-17,7848	-328,7612	2,1654	92,0265
	Mean(Missing)	4,6543	-39,6944	.	60,5656	-26,2000	-363,1762	-8,4744	85,2414
RATIO8	t	-1,1	-,7	-,5	.	-,9	-,8	-1,7	-,8
	df	9,5	15,3	3,2	.	7,9	8,6	28,7	5,5
	P(2-tail)	,296	,465	,624	.	,390	,427	,095	,469
	# Present	27	29	27	36	28	29	29	25
	# Missing	5	4	4	0	5	5	4	5
	Mean(Present)	-3,0678	-17,1410	30,7841	65,3361	-21,8646	-368,3779	-2,1634	86,9372
	Mean(Missing)	8,3140	-5,8750	39,5825	.	-8,4020	-154,0480	9,6100	107,9740

For each quantitative variable, pairs of groups are formed by indicator variables (present, missing).
a. Indicator variables with less than 5% missing are not displayed.

Tablo 3.3

Separate Variance t Tests ^a									
	RATIO1	RATIO3	RATIO5	RATIO8	RATIO12	RATIO14	RATIO15	RATIO18	
RATIO12	t	-,5	-1,1	,5	-1,3	.	-,8	-1,3	-,4
	df	17,7	29,9	10,1	15,5	.	19,0	29,6	5,2
	P(2-tail)	,592	,270	,651	,210	.	,413	,211	,739
	# Present	24	28	25	28	33	26	27	26
	# Missing	8	5	6	8	0	8	6	4
	Mean(Present)	-2,7304	-17,9114	32,5184	63,0096	-19,8248	-384,4273	-2,4989	89,4231
	Mean(Missing)	3,0338	-3,8140	29,4233	73,4787	.	-182,2612	7,1950	97,0750
RATIO14	t	-,4	-,3	,3	1,1	-,5	.	-1,3	-,4
	df	9,4	8,4	4,9	8,8	16,3	.	10,5	5,9
	P(2-tail)	,673	,779	,751	,291	,644	.	,215	,711
	# Present	26	28	26	29	26	34	29	25
	# Missing	6	5	5	7	7	0	4	5
	Mean(Present)	-2,2635	-16,6964	32,5004	67,6117	-21,1408	-336,8588	-2,1386	88,7816
	Mean(Missing)	2,9317	-10,6180	28,8980	55,9086	-14,9371	.	9,4300	98,7520
RATIO15	t	,8	,2	2,1	,6	1,1	1,3	.	,6
	df	12,4	23,1	9,0	8,1	10,7	5,9	.	2,3
	P(2-tail)	,416	,823	,060	,569	,285	,240	.	,612
	# Present	25	26	24	29	27	29	33	27
	# Missing	7	7	7	7	6	5	0	3
	Mean(Present)	,7968	-14,9635	35,5492	66,6631	-16,8911	-273,3466	-,7364	92,7833
	Mean(Missing)	-8,7400	-18,7914	19,4743	59,8386	-33,0267	-705,2300	.	69,3833
RATIO18	t	,3	1,1	1,5	-,1	1,1	,8	,6	.
	df	13,8	8,2	11,2	14,6	6,8	11,6	5,2	.
	P(2-tail)	,771	,311	,170	,883	,329	,415	,547	.
	# Present	25	24	23	25	26	25	27	30
	# Missing	7	9	8	11	7	9	6	0
	Mean(Present)	-,5824	-5,1629	34,7304	64,8756	-14,3062	-262,6008	2,3174	90,4433
	Mean(Missing)	-3,8143	-44,0756	23,8375	66,3827	-40,3229	-543,1311	-14,4783	.

For each quantitative variable, pairs of groups are formed by indicator variables (present, missing).
a. Indicator variables with less than 5% missing are not displayed.

Tablo 3.4

RATIO5'teki 31 gözlemin, RATIO15'te 24 tam ve 7 eksik gözlem olarak bölümlenmesi $p = 0,60$ değeri ile anlamlılığa en yakın olan değerdir. Dolayısıyla anlamlı olan bir durum yoktur ve eksik veriler rastgele dağılmıştır.

Aynı anda pek çok değişkeni etkileyen bir değişken göze çarpmamaktadır.

EM Means^{a,b}

RATIO1	RATIO3	RATIO5	RATIO8	RATIO12	RATIO14	RATIO15	RATIO18
-4,4199	-20,8485	32,9528	66,3206	-18,1138	-328,9428	-2,7425	80,3709

a. Little's MCAR test: Chi-Square = 128,861, DF = 135, Sig. = ,633

b. The EM algorithm failed to converge in 25 iterations.

H_0 : Eksik veriler rastgele dağılmıştır.

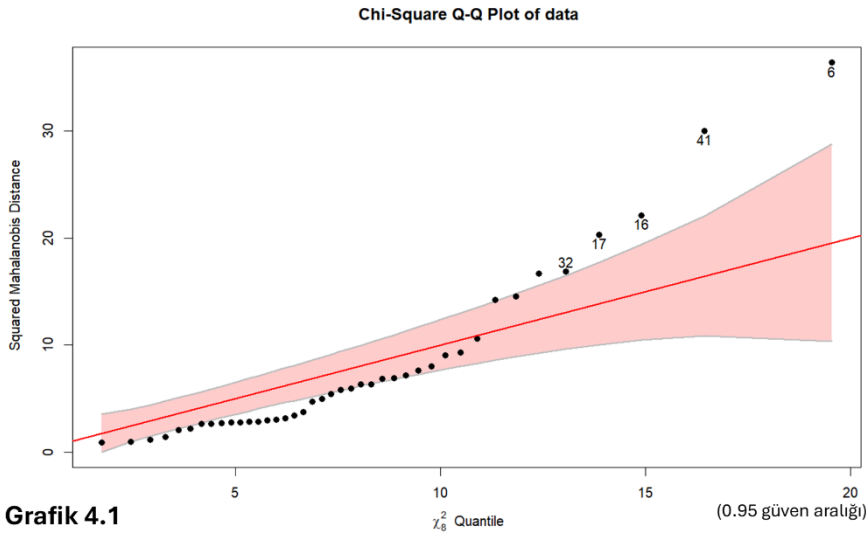
H_1 : Eksik veriler rastgele dağılmamıştır.

Tablo 3.5

Sig. = 0,633 > $\alpha(0,05)$ olduğundan H_0 kabul edilir. Eksik veriler rastgele dağılmıştır.

- Gözlem boyutunu %20 azaltmak istemedğimden eksik verileri çıkarma yanlısı değilim. Çoklu doğrusal bağımlılık sorunu olduğu için regresyon ataması yapmak doğru gelmiyor. Çoklu atama veya EM atama yöntemini kullanmayı öneririm.

4. a)



Kare Mahalanobis uzaklığı hesaplanarak oluşturulmuş ki-kare saçılım grafiğine bakıldığında; 6, 41, 16, 17 ve 32'nci gözlemlerin aşırı değerli olma ihtimali olduğu görülüyor.

Grafik 4.1

b)

Tablo 4.1'de, R programında 'outlier' fonksiyonu ile elde edilen m_i^2 değerleri görülüyor.

Mahalanobis Uzaklıkları

Göz. No	m_i^2	Göz. No	m_i^2
1	7,611333	21	9,28104
2	6,293968	22	2,617582
3	3,402171	23	5,776063
4	6,332286	24	7,156944
5	6,875722	25	2,786335
6	36,40557	26	14,51235
7	3,725154	27	0,897722
8	2,173261	28	4,711255
9	5,411295	29	2,773235
10	2,856083	30	2,69555
11	2,020793	31	2,935777
12	6,833471	32	16,8975
13	14,19177	33	3,118469
14	16,67237	34	1,155338
15	1,405953	35	4,985348
16	22,11744	36	7,974936
17	20,31239	37	9,060917
18	0,941649	38	2,65461
19	2,847932	39	10,59609
20	2,986314	40	5,95453
		41	30,04148

Tablo 4.1

$$X_{(8; \alpha=0,005)}^2 = 21.95495$$

$\alpha=0,005$ anlamlılık düzeyi için serbestlik derecesi 8 olan ki-kare değeri 21,95'tir.

Gözlemlerin kare Mahalanobis değerlerine bakıldığında 21,95'ten büyük olan 3 adet gözlem olduğu görülüyor.

6, 16 ve 41'inci gözlemler çok değişkenli aşırı gözlemlerdir.

İki yöntem de aynı gözlemleri aşırı değer olarak belirlemiştir. b yöntemi daha kesin bir sonuç vermiştir. 6, 16 ve 41'inci gözlemleri çıkarmayı öneririm.

Mahalanobis Uzaklıkları

Göz. No	m_i^2	Göz. No	m_i^2
1	9,277321	19	16,921669
2	6,205371	20	2,749951
3	3,398358	21	9,040737
4	6,236268	22	7,310716
5	22,348835	23	2,955049
6	3,681897	24	15,362698
7	2,758835	25	1,062477
8	6,559522	26	4,979126
9	3,385847	27	2,972147
10	2,274079	28	2,658471
11	10,872321	29	2,788448
12	19,359027	30	16,606837
13	18,409186	31	3,208891
14	1,525891	32	1,290937
15	27,538447	33	5,137509
16	1,229471	34	7,573366
17	2,888116	35	17,985533
18	6,5708	36	3,033653
		37	12,30284

Tablo 4.2

6, 16 ve 41'inci gözlemleri çıkardığımızda aşırı değerler için Mahalanobis değerlerinin azaldığını görüyoruz. Yeni aşırı değerler 5. ve 15. gözlemler olarak belirlenmiştir.

6, 16, 41'inci gözlemleri çıkarmanın büyük bir katkısı olduğunu söyleyemiyorum.

5.

Tablo 4.1'e göre 41 gözlemde 3'ü (%7.317073), $X^2_{(8; \alpha=0,005)} = 21.95495$ değerinden büyüktür. Gözlemler %50'lik konturun içine düşer. Dolayısıyla dağılım normaldir.

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RATIO1	,316	41	,000	,532	41	,000
RATIO3	,283	41	,000	,525	41	,000
RATIO5	,081	41	,200*	,956	41	,115
RATIO8	,094	41	,200*	,970	41	,334
RATIO12	,294	41	,000	,655	41	,000
RATIO14	,337	41	,000	,608	41	,000
RATIO15	,274	41	,000	,643	41	,000
RATIO18	,233	41	,000	,883	41	,001

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tablo 5.1**Tests of Normality**

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RATIO1	,324	38	,000	,691	38	,000
RATIO3	,280	38	,000	,681	38	,000
RATIO5	,084	38	,200*	,950	38	,092
RATIO8	,093	38	,200*	,976	38	,581
RATIO12	,321	38	,000	,645	38	,000
RATIO14	,347	38	,000	,543	38	,000
RATIO15	,234	38	,000	,750	38	,000
RATIO18	,228	38	,000	,893	38	,002

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tablo 5.2

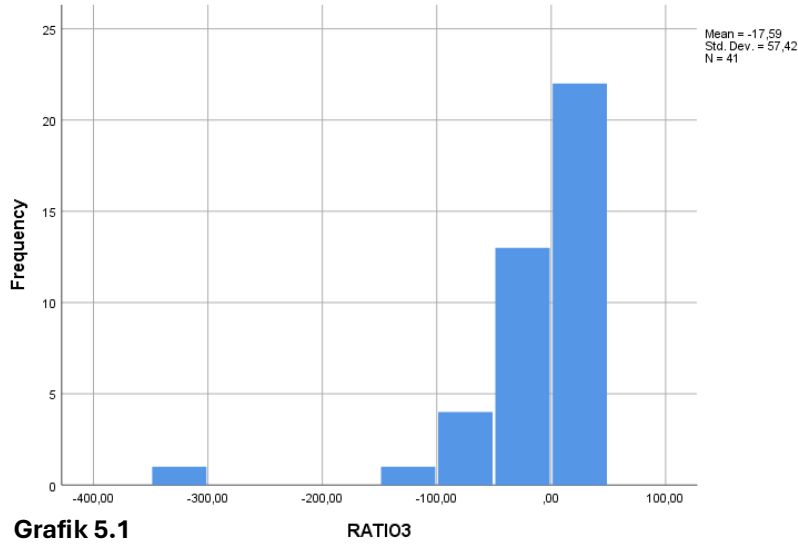
H_0 : Veri seti normal dağılıma sahiptir.

H_1 : Veri seti normal dağılıma sahip değildir.

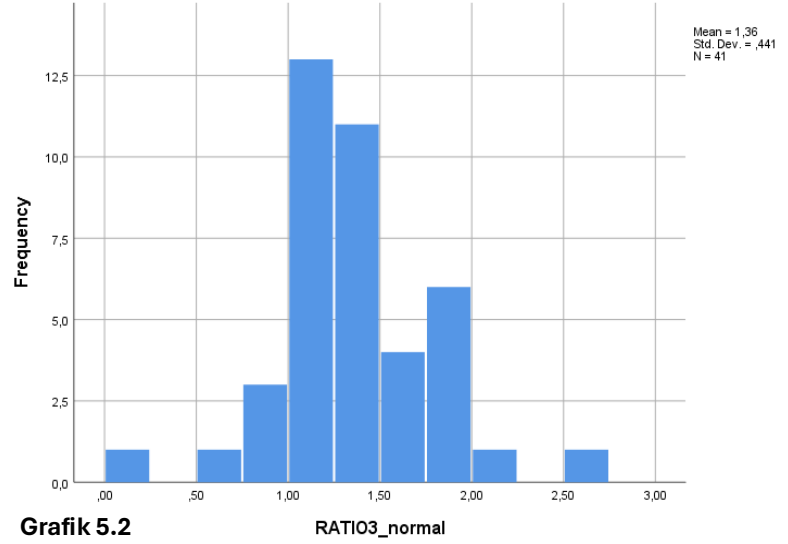
($\alpha = 0,05$)

Sadece RATIO5 ve RATIO8 değişkenleri normal dağılıma sahiptir.

6, 16 ve 41'inci gözlemler çıkarıldığında RATIO8 değişkeninin normalliği artarken RATIO5 değişkeninin normalliği azalmıştır. Bu değişkenleri çıkarmak normallik açısından anlamlı gözüküyor.



RATIO3 değişkeni şiddetli negatif çarpıktır.



RATIO3'e normallik dönüşümü yaptığımızda, normal dağılıma yakın bir dağılım gösterir.

Normallik dönüşümü yapılmış RATIO3 ile R programında elde ettiğim m_i^2 değerleri Tablo 4.1 ile aynı çıktı fakat SPSS üzerinden, bağımlı değişkeni RATIO 18 seçerek m_i^2 değerlerini elde ettiğimde ortaya iki farklı sonuç çıktı.

Tablo 5.2, normallik dönüşümü yapılmamış RATIO3 ve Tablo 5.3, normallik dönüşümü yapılmış RATIO3 ile oluşturulmuştur. RATIO3'e yapılan normallik dönüşümü 32. gözlemi de aşırı olarak belirlemiştir. RATIO3'e normallik dönüşümü yapmak çok değişkenli normallik varsayımına çok bir katkı sağlamamıştır.

Mahalanobis Uzaklıkları			
Göz. No	m_i^2	Göz. No	m_i^2
1	1,61442	21	9,27695
2	6,29393	22	2,61489
3	3,40211	23	5,71547
4	4,74922	24	7,06583
5	3,75634	25	1,98587
6	36,21907	26	10,46690
7	3,71987	27	,56902
8	2,04149	28	4,52266
9	2,71758	29	2,58318
10	2,17510	30	1,68489
11	2,01105	31	2,72980
12	5,06752	32	16,79294
13	12,80589	33	3,11265
14	16,57204	34	1,09735
15	1,14742	35	4,51852
16	21,35968	36	6,97085
17	14,61648	37	8,92125
18	,63889	38	1,54659
19	2,06046	39	10,59178
20	2,83850	40	4,46810
		41	26,95743

Tablo 5.2

Mahalanobis Uzaklıkları			
Göz. No	m_i^2	Göz. No	m_i^2
1	1,47042	21	7,03332
2	5,94491	22	2,40915
3	2,42146	23	4,06370
4	4,48160	24	7,59605
5	6,21436	25	2,84813
6	36,98973	26	5,60672
7	3,61790	27	1,03498
8	2,21607	28	4,65519
9	5,52564	29	3,00331
10	3,55443	30	2,15321
11	2,00625	31	2,28531
12	5,08116	32	22,32832
13	12,25407	33	1,87463
14	17,56984	34	1,45106
15	1,07809	35	2,34643
16	20,72300	36	1,75467
17	14,41097	37	6,54853
18	1,00932	38	3,52101
19	2,12577	39	10,83663
20	9,58246	40	1,45474
		41	26,91747

Tablo 5.3

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
RATIO3_normal	,139	41	,045	,940	41	,033

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tablo 5.4

Normallik dönüşümü yapılmamış RATIO3 için 0 olan Shapiro-Wilk anlamlılık değeri(Tablo 5.1), normallik dönüşümü yapılmış RATIO3 için 0,033 olmuştur. Anlamlılık artmış olsa da $\alpha(0,05)$ değerinin altında kalmıştır. Dolayısıyla normallik dönüşümü yapılmış RATIO3'ün normal dağıldığını söyleyemeyiz.

RATIO3 değişkenine, normallik dönüşümü yapılması çok değişkenli normallik varsayımının sağlanmasında anlamlı bir değişiklik meydana getirmemiştir.