

ONDOKUZ MAYIS ÜNİVERSİTESİ
FEN-EDEBİYAT FAKÜLTESİ
İSTATİSTİK BÖLÜMÜ
2019-2020 Yılı Güz Dönemi
Minitab Ders Notları



Minitab® 17

Minitab'taki Ana Menüler

Minitab programının 11 tane ana menüsü bulunmaktadır. Bu ana menüler altında ise çok sayıda kullanışlı menüler ve fonksiyonlar vardır. İlk ana menü olan File menüsü ile konulara giriş yapabiliriz.

File

New: Yeni bir MINITAB dosyası açma ya da MINITAB projesi açar.

Open Project: Minitab'da daha önce kaydedilmiş MPJ uzantılı proje dosyalarını açar. Hem veri hem de yapılan işlemler ekranda gözükür.

Open Worksheet: MINITAB'da daha önce kaydedilmiş dosyaları açar. Bunlardan MINITAB dosyasını (.MTW), taşınır veri dosyası (.MTP), Excel (.XLS), dBase, .DAT, text gibi dosyaları açmayı sağlar.

Query Database (ODBC): Microsoft Access, Oracle, Sybase ya da SAS programlarında hazırlanmış olan veri tabanı dosyalarını MINITAB veri sayfası olarak açar.

Save Current Worksheet: Veri dosyasını aynı isimle kaydeder.

Save Current Worksheet As: Veri dosyasını farklı isimle kaydeder. Buna ilaveten farklı formatlarda dosya kaydetmeyi sağlar.

Close Worksheet: Veri dosyasını kapatır.

Open Graph: *.MGF uzantılı grafik dosyalarını açar.

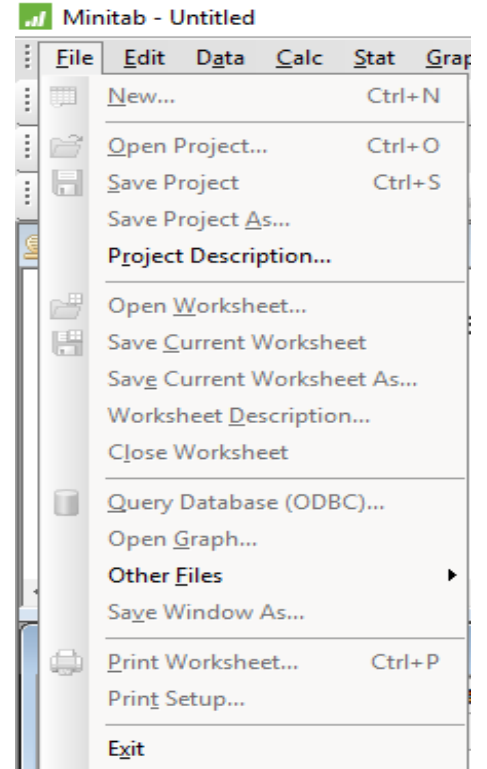
Save Window As: Aktif pencere içeriğini bir dosyaya kayıt eder.

Other Files: import ve export edilebilecek metin dosyaları ile ilgili komutları görüntüler.

Print Session Window: Aktif pencerenin çıktısını alır.

Print Setup: Yapılan işlemlerin çıktısı alınır ve yazıcı hakkında bilgi verir.

Exit: MINITAB programından çıkış yapar.



Edit

Undo Clear: Son yapılan değişikliği geri alır.

Redo: Son yapılan değişikliği ileri alır.

Clear Cells: Seçilmiş hücrenin verisi silinir.

Delete Cells: Seçilen veri gözlerinin satır ya da sütun olarak silinmesini sağlar.

Copy Cells: Seçilen veri ya da metin parçasını kopyalar.

Cut Cells: Seçili alanı temizler ve içeriğini belleğe alır.

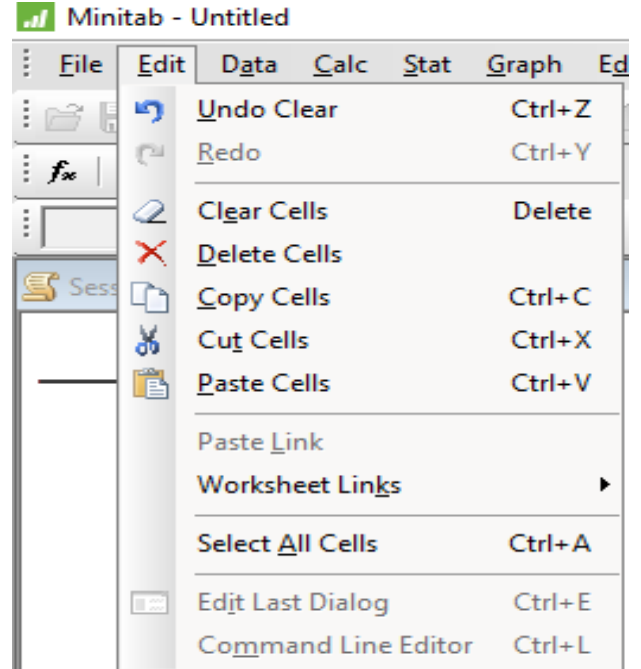
Paste Cells: Kayıtlı olan seçeneği seçilen yere yapıştırır.

Links: Dinamik veri alışveriş bağlantısı kurar, bağlantı kurulmuş başka bir programdan veri alır.

Select All Cells: Aktif pencerede yer alan tüm içeriği seçer.

Edit Last Command Dialog: En son yapılan işlemle ilgili pencereyi ve seçenekleri görüntüler.

Command Line Editor: Daha önce çalıştırılan komutları yeniden çalıştırma izni verir.



Data

Subset Worksheet: Veri sayfasında kayıtlı verilerden belirli özelliklere sahip yeni bir alt set oluşturur.

Split Worksheet: Veri sayfasında kayıtlı verileri belirli özelliklere göre iki parçaya ayırarak, yeni veri dosyası oluşturur.

Merge Worksheet: Farklı iki veri setini yan yana ya da alt alta birleştirir. Farklı dosyalardaki verileri değişkenler farklı ise yan yana, aynı ise alt alta birleştirir.

Copy Columns: Veri sayfasındaki bir sütunu kopyalar.

Unstack Columns: Bir kolonu belirli bir kritere ya da bir başka değişkene göre ayırır.

Stack: Bir sütunu ya da bir satırı çalışılan veri sayfasının belirli bir sütuna yığmayı ve farklı veri dosyasına birleştirerek yazmayı sağlar.

Transpose Columns: Bir sütunu satıra, satırı ise sütuna çevirir.

Short: Bir veya daha fazla sütunda yer alan verileri sıralar. Eğer *Descending* işaretlenirse büyükten küçüğe, işaretlenmezse küçükten büyüğe doğru bir sıralama yapılır.

Rank: Bir sütundaki verileri sıralama puanlarına dönüştürür.

Delete Rows: Verilerde yer alan sütundan bir veriyi siler.

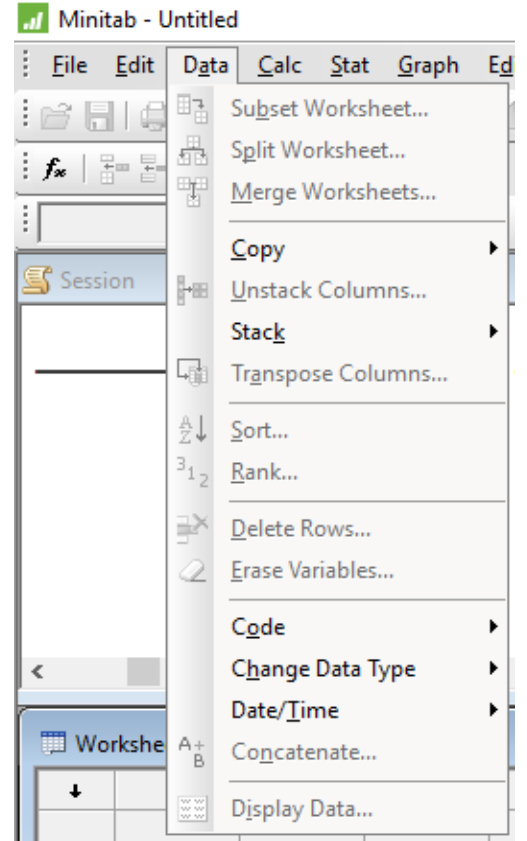
Erase Variables: İşaretlenen değişkenleri siler.

Code: Sütunlarda yer alan verileri kodlayarak yeni değişken oluşturur.

Change Data Type: Karakter verilerini sayısal verilere, sayısal verileri karakter değerlere, tarih/zaman verilerini ise diğer tiplere dönüştürür.

Concatenate: İki veya daha fazla karakter içeren değişkenleri bir sütunda bileştirir.

Display Data: İşlem penceresi aktif iken, aktif veri setinde yer alan değişkenlerin sabitlerin-matrislerin ekranda görüntüler.



Calc

Calculator: MINITAB'da birçok matematiksel ve mantıksal karşılaştırma işlemlerini yapar.

Column Statistics: Sütun (değişkenler) ile ilgili değişik istatistikleri hesaplar. *Sum*-Toplam, *Mean*-ortalama, *Standart deviation*-standart sapma, *Range*-sıra, *Median*-medyan, *N missing*-kayıp gözlem sayısı, *optional*-seçmeli, isteğe bağlı.

Row Statistics: Seçilen değişkenlerin her satırı için temel istatistikleri bulur.

Standardize: Sütunlardaki değişkenlere merkezi bir değere göre standartlaştırma yapar.

Make Patterned Data: Veri setinde değişik sayıda tekrarlanan değerleri kısaca girmek için kullanılır.

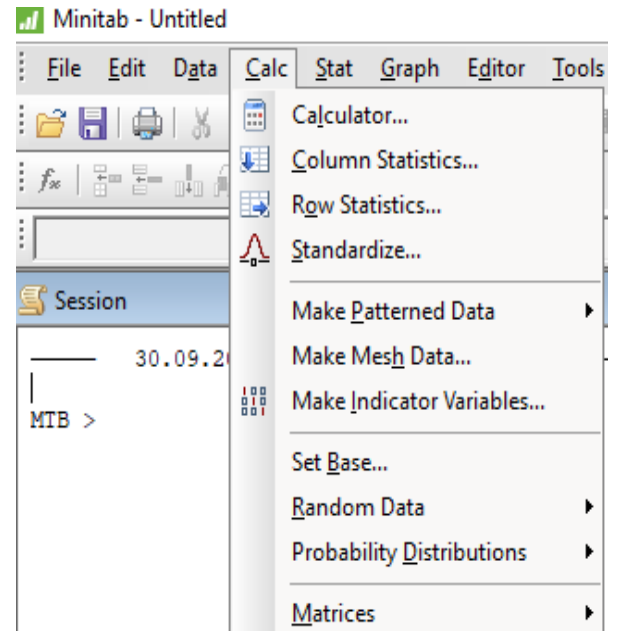
Make Mesh Data: Üç boyutlu yüzey grafiği çizmek için gerekli olan x,y,z değerlerinin veri setinin oluşturur.

Make Indicator Variables: Regresyon analizinde kullanılmak üzere kukla (dummy) değişkenler üretir.

Set Base: Rasgele veri üretimi yapılırken başlangıç sayısı belirler.

Random Data: Bazı teorik dağılımlara uygunluk gösteren örnek veriler üretir.

Probability Distributions: Sürekli ve kesikli dağılımların olasılık yoğunluklarını (*Probability density*), yığılımlı olasılıklarını (*Cumulative probability*) ve ters yığılımlı olasılıklarını (*Inverse Cumulative probability*) hesaplamayı sağlar.



Matrices: Matris işlemlerini yapar.

Read - matris girişi;

Number of rows: satır sayısı,

Number of columns - sütun sayısı,

Read into matrix - Girilen matrise M1, M2.. gibi isimler verilir.

Read from keyboard - işaretlenir ve veriler klavyeden girilir.

Transpose - Girilen matrisin transpozunu alır ve yeni bir M matrisi diye tanımlanır.

Invert - Girilen matrisin tersini alır ve yeni bir M matrisi diye tanımlanır.

Define constant - Sabit bir sayı tanımlanır ve bunun matrisi alınır.

Diagonal - Girilen bir sütundaki verilerden köşegen matris oluşturur.

Copy diagonal - Bir matrisin köşegen elemanlarını kopyalar.

Copy - Bir sütunu veya matrisi kopyalar.

Eigen Analysis - Simetrik olarak girilen bir matrisin özdeğer ve özvektörlerini bulur.

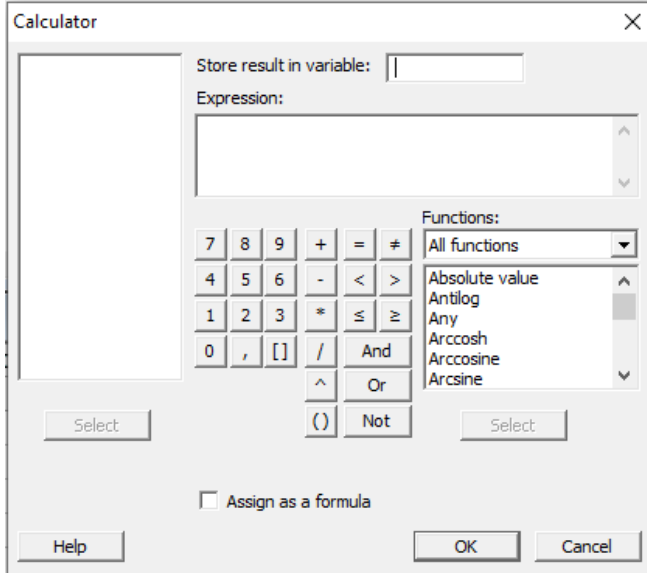
Column of eigenvalues - Matrisin özdeğerlerini bulur. Bu değerleri görmek için C3 gibi bir sütun ismi yazılır.

Matrix of eigenvectors - Özvektörler matrisi bulunur .

Arithmetic - Aritmetik matris işlemleri yapılır.

Add- İki matrisi toplar, *Subtract* - İki matrisi birbirinden çıkarır, *Multiply* - İki matrisi çarpar.

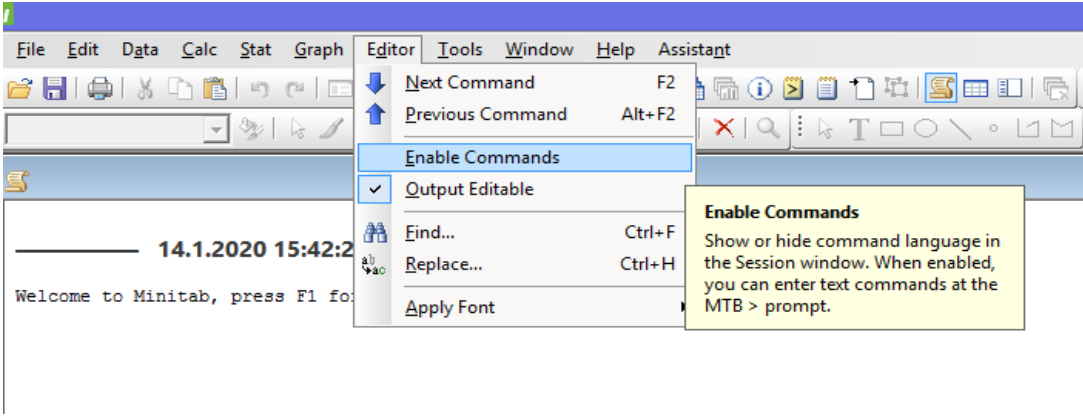
Calc – Calculator: Birçok cebirsel ve matematiksel işlemleri yapmaya, veri dönüştürme işlemleri yapmaya, yeni değişkenler tanımlamaya yönelik çok sayıda fonksiyon içerir.



MINITAB'ta Komutlar

Minitab üzerinde işlem yapılırken menüler aracılığıyla işlem yapılabileceği gibi komutlar üzerinden de bu işlemler yapılabilir. Bu komutlar Session Window üzerinden yazılmaktadır. Bu pencere, program açıldığı zaman aktif değildir. Bu nedenle pencere aktif halde getirilerek gerekli komutların yazılabilmesi için aşağıdaki işlem yapılmalıdır.

Editor - Enable Commands: Oturum penceresinde (session window) işlemlerin komutlarla yapılabilmesi için MTB> iletilisinin aktifleşmesini sağlanmalıdır alttaki resimde görülebileceği gibi komutlar bu şekilde aktifleşir. Bu ileti yoksa oturum penceresinde yalnızca menüler sonucu elde edilen çıktılar yer alır.



Minitab'da kullanılan bazı önemli komutlar için örnekler ve komutların işlevleri aşağıda verilmiştir. Komutlarda C1 değişkenin girildiği sütun numarasını göstermektedir. K1 ise sabit değerleri göstermektedir.

- **NAME (C1)** : C1 sütununa isim verir.
MTB>NAME C1 'Kilogram'
- **READ** : Klavyeden veri girişini sağlar.
MTB>READ C1
DATA>2
DATA>7
DATA>13
DATA>END
- **SET** : Bir değişkenin değerlerinin bir sütuna girilmesini sağlar.
MTB>SET C1
DATA>10
DATA>4
DATA>3
DATA>END

MTB>SET C2
DATA>1:20
DATA>END

MTB>SET C3
DATA>10(1)
DATA>END

MTB>SET C4
DATA>3(1:3)
DATA>END

MTB>SET C5
DATA>(1:3)2
DATA>END

MTB>SET C6
DATA>2(1:3)3
DATA>END
- **ABS (K1)** : Bir sayının mutlak değerini alır.
MTB>ABS -5 K1
- **DESC (C1)** : Bir değişkenin(kolonun) belirtici istatistiklerinin hesaplar.
MTB>DESC C1
- **MEAN (C1)** : Bir değişkenin(kolonun) ortalamasını bulur.
MTB>MEAN C1
- **AVERAGE (C1)**: Bir değişkenin(kolonun) ortalamasını bulur.

MTB> AVERAGE (C1)

- **STDEV (C1)** : Bir deęişkenin(kolonun) standart sapmasını bulur.

MTB>STDEV (C1)

- **EXP (K1)** : Bir sabitin e üzeri x deęerini hesaplar.

MTB>EXP 3 K1

- **SQRT (K1)** : Bir sabitin karekökünü bulur.

MTB>SQRT 4 K1

- **TALLY (C1)** : Bir deęişkenin(kolonun) frekans daęılımını bulur.

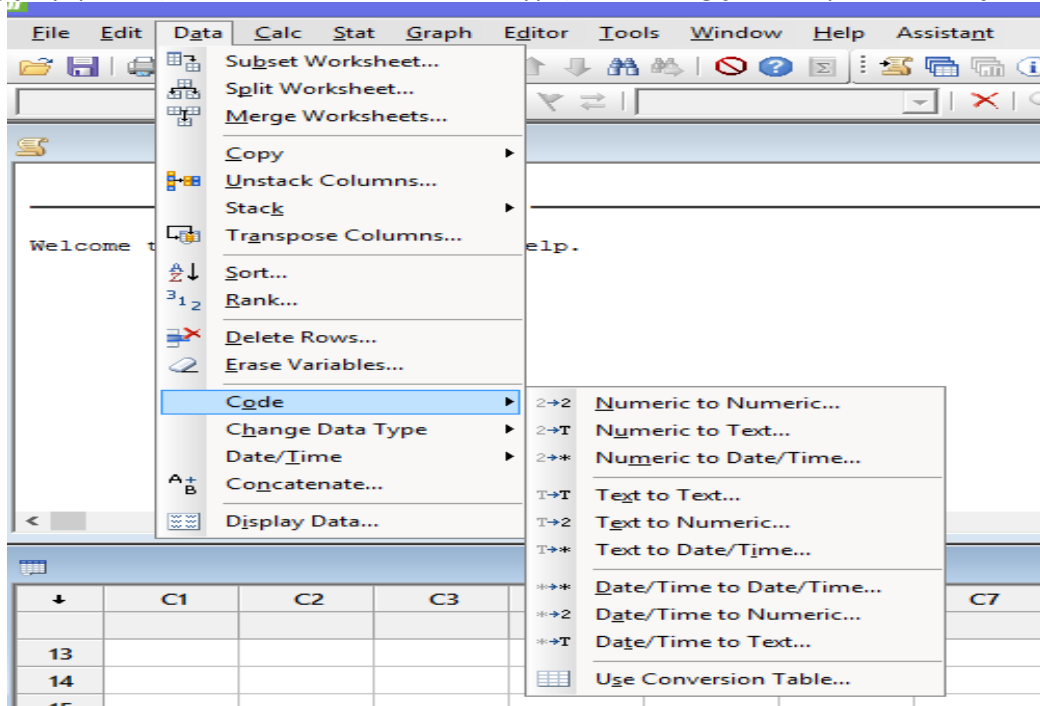
MTB>TALLY C1

- **LET** : Bu komut atama, belirtme ve tanımlama işlemlerini yapar.

LET C2=SQRT(C1) : C1'de yer alan deęerlerin kareköklerini alır ve C2 sütununa yazar.

TEMEL MENULER ÜZERİNDEN YAPILAN İŞLEMLERLE İLGİLİ BAZI ALIŞTIRMALAR

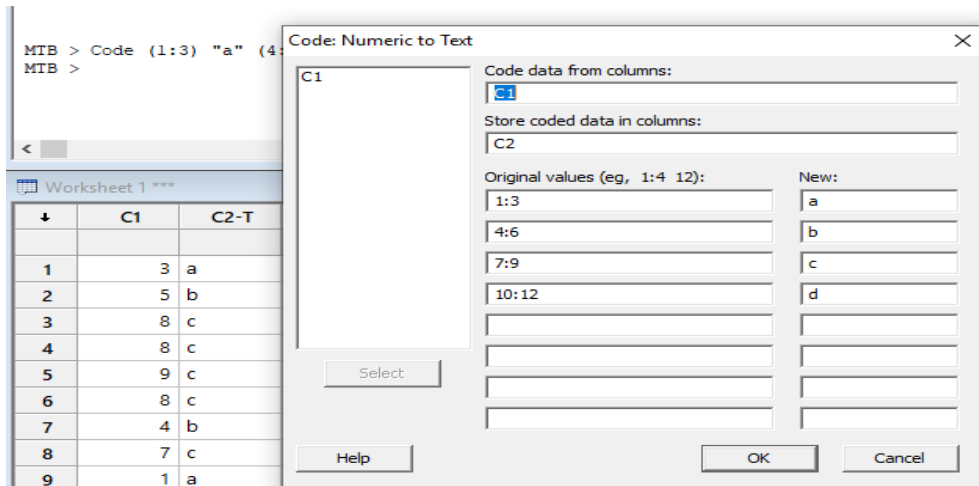
Datadan – Code: Sütunlarda yer alan verileri kodlayarak yeni bir kodlanmış deęişken oluşturur. Bu komut ile veri sayfasındaki deęişik yapıdaki (text, numeric, date and time type) verileri deęişik veri tiplerine dönüştürülebilir.



Örnek:

- **Numeric To Text:**

MTB > Code (1:3) "a" (4:6) "b" (7:9) "c" (10:12) "d" C1 C2



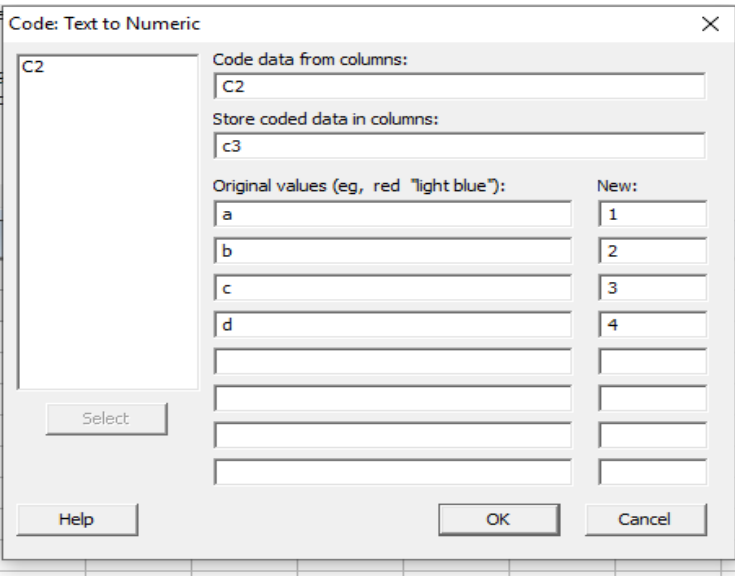
- **Text to Numeric:**

MTB > Code ("a") 1 ("b") 2 ("c") 3 ("d") 4 C2 c3

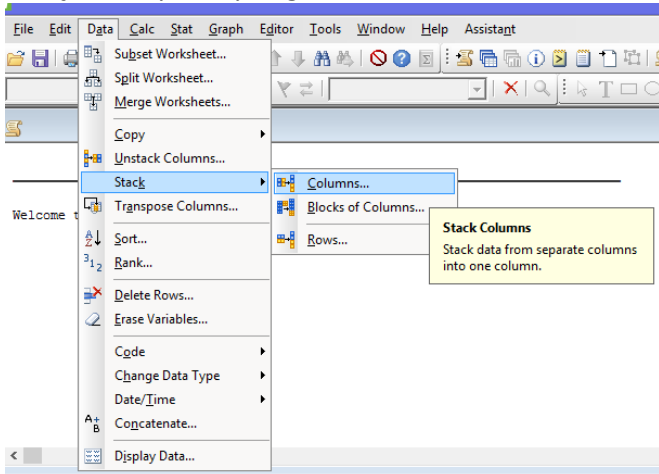
itab, press F1 for help

3) "a" (4:6) "b" (7:9)
a") 1 ("b") 2 ("c") 3 ("d") 4

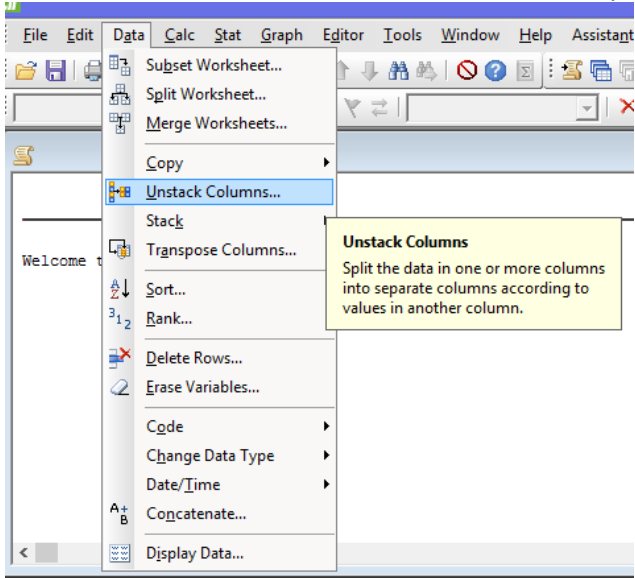
| C2-T | C3 |
|------|----|
| a | 1 |
| b | 2 |
| c | 3 |
| c | 3 |
| c | 3 |
| c | 3 |
| b | 2 |
| c | 3 |



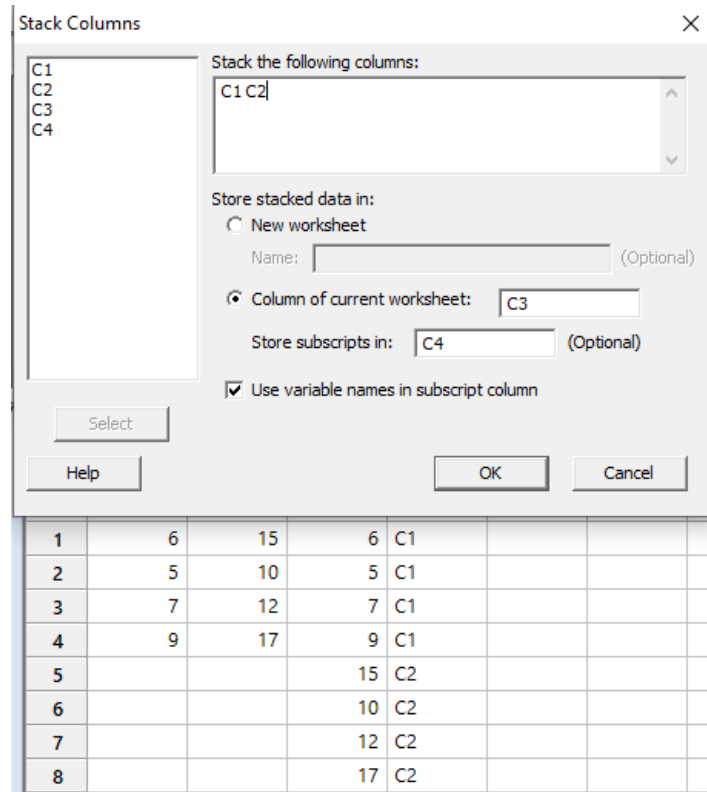
Data – Stack: Bir sütunu ya da bir satırı çalışılan veri sayfasının belirli bir sütuna yığmayı ve farklı veri dosyasına birleştirerek yazmayı sağlar.



Data – Unstack Columns: Bir sütündaki verileri bir ya da daha fazla sütüne bölmeye yarar.



Örnek (Stack Columns): İki gruba ait aşağıdaki veriler elde edilmiştir. Bu iki grup verilerini alt alta birleştirip C3 ünunda alt alta girip, C4 e ise verilerin hangi kolonlardan geldiğini yazdıracak işlemi yapınız.



Minitab'ta Veri Girişi

Minitabta veri setleri elle girileceği zaman 2 farklı metod kullanılır. Bunlar çalışma sayfası ile veri girişi ve komutlar üzerinden veri giriştir.

a) Çalışma sayfası ile veri girişi

| | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | C13 | C14 |
|----|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|
| 1 | | | | | | | | | | | | | | |
| 2 | | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | | | | |
| 4 | | | | | | | | | | | | | | |
| 5 | | | | | | | | | | | | | | |
| 6 | | | | | | | | | | | | | | |
| 7 | | | | | | | | | | | | | | |
| 8 | | | | | | | | | | | | | | |
| 9 | | | | | | | | | | | | | | |
| 10 | | | | | | | | | | | | | | |
| 11 | | | | | | | | | | | | | | |

Şekilde görüldüğü üzere "worksheet" yani çalışma sayfası üzerinden değişkenlere yani kolonlara isim verilip satırlar doldurularak bu işlem yapılabilir.

b) Komutlar ile veri girişi

İki ayrı komut kullanılarak bu işlem yapılabilir. Bunlar "Read" ve "Set" komutlarıdır.

Minitab'ta Normallik Testleri

Parametrik testlerin uygulanabilmesi için verinin belli bir dağılıma sahip olması gerekmektedir. Örnek vermek gerekirse; Anova, Regresyon, t testi gibi testler verilerin normal dağılıma sahip olduğu durumlarda kullanılmaktadır bu yüzden de verilerin dağılımını kontrol etmek, testin belirlenmesinden önce gerçekleştirilmelidir.

Literatüre girmiş birçok normallik varsayımı testi vardır. Bunlar şöyle sıralanabilir:

- Anderson – Darling Testi
- Shapiro – Wilk Testi
- Kolmogorov – Smirnov Tek Örnek Testi
- Pearson Ki Kare Uygunluk Testi
- Cramer- Von Mises Testi
- Shapiro – Francia Testi

Bu testler birbirinin türevi olarak şekilde çoğaltılmıştır. Yani bu listeyi daha da genişletmek mümkündür. Yapılan çalışmalarda ise Shapiro-Wilk testi ve Anderson-Darling testi diğerlerine göre daha güçlü bulunmuştur. Minitab programı bizlere 3 tane normallik testi sunmaktadır. Bunlar Anderson-Darling, Ryan Joiner ve Kolmogorov Smirnov testleridir.

Burada test edilen hipotezler şu şekildedir.

H_0 : Veriler normal dağılmıştır.

H_1 : Veriler normal dağılmamıştır.

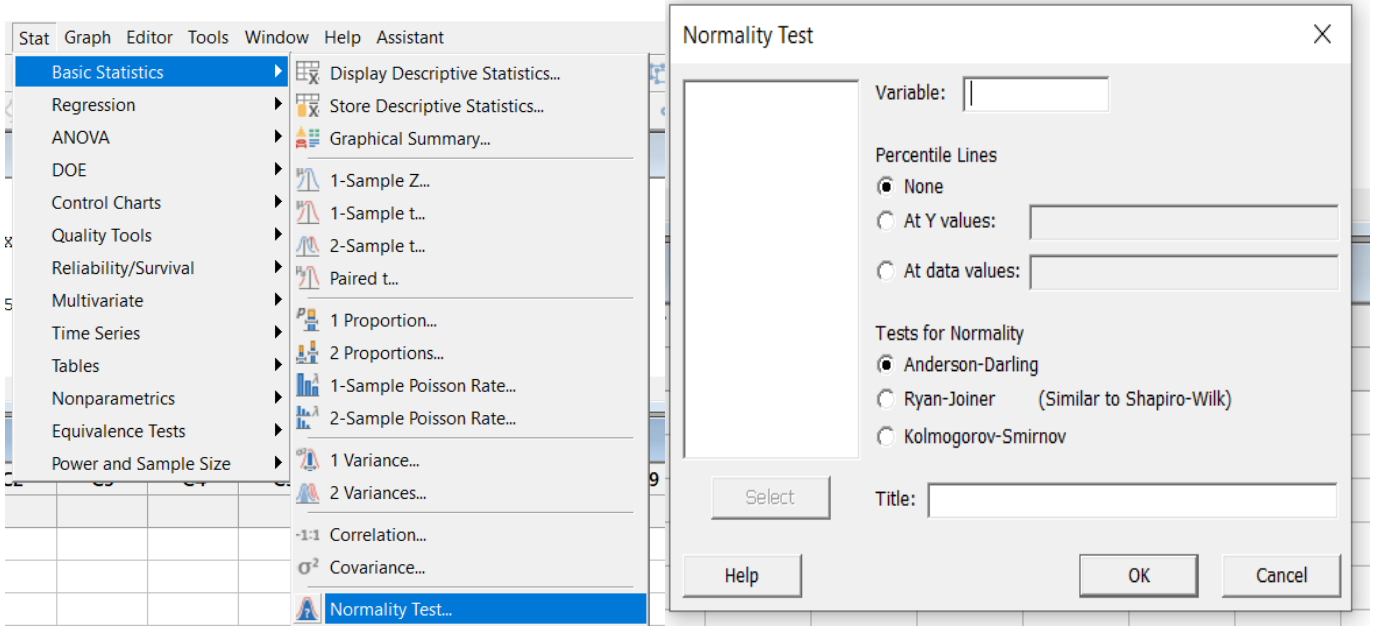
Minitab ta yapılacak testin önem değeri eğer testten önce belirlenen alfa değerinden küçükse H_0 reddedilir, değilse reddedilemez. Bu sonuca göre de karar ve yorum yapılarak normallik testi yapılır.

Eğer normallik testine göre veriler normal dağılımsız değilse verilere log-karekok gibi matematiksel fonksiyonlar uygulanarak normallik sağlanmaya çalışılır. Eğer bu işlem de sonuç vermezse parametrik olmayan testler tercih edilir.

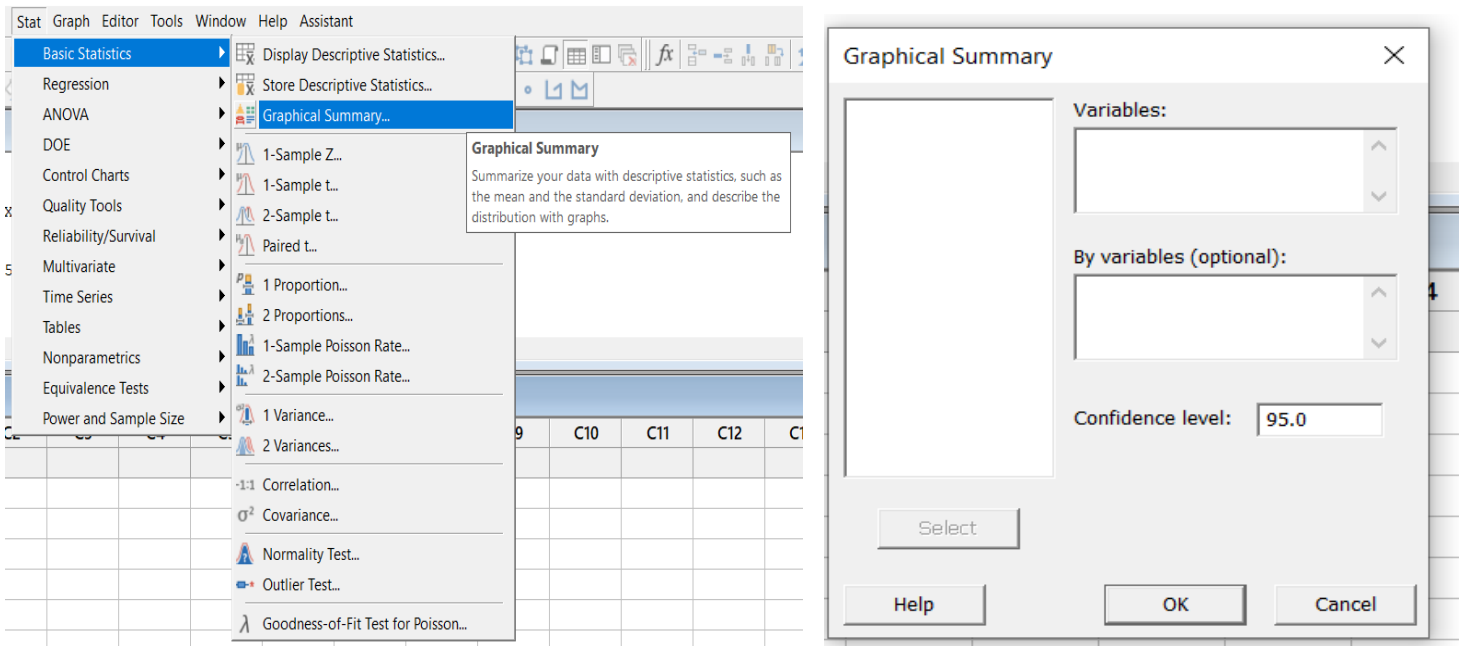
Minitab'ta Anderson Darling Normallik Testi

Anderson Darling testi iki şekilde yapılabilir ilki şu sıralamadır.

[Stat-Basic Statistics-Normality Test-Anderson Darling](#)



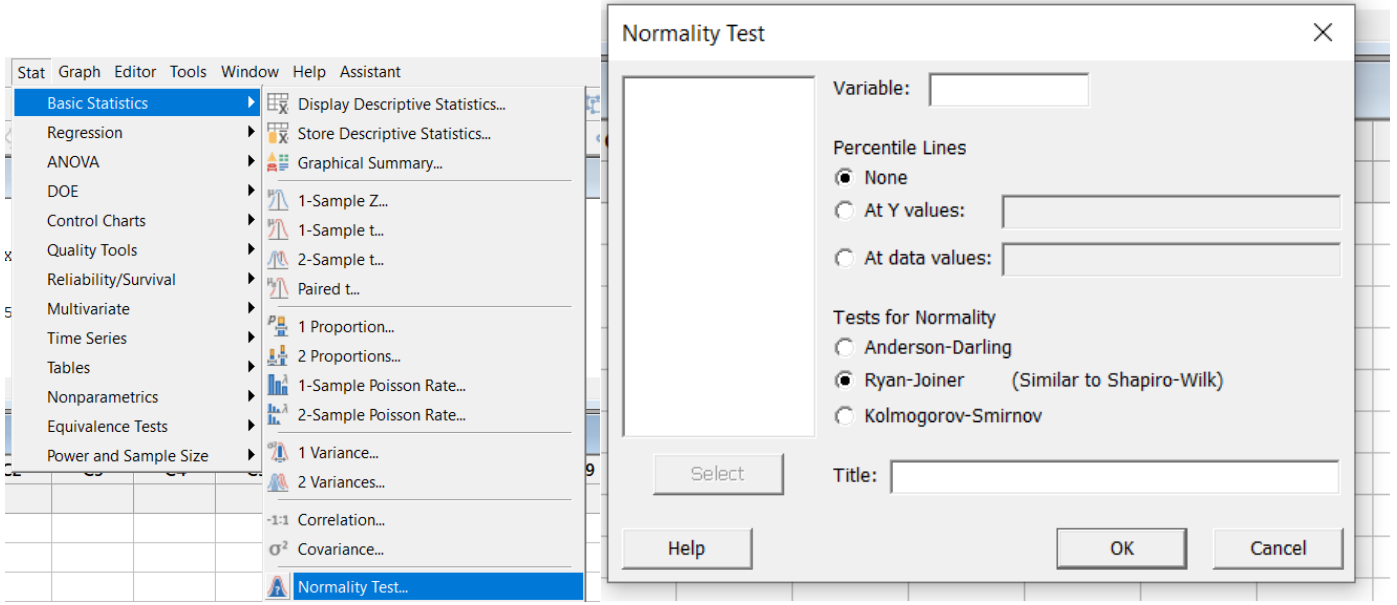
İkinci olarak da **Basic Statistics-Graphical Summary** seçenekleri üzerinden yapılır.



Minitab'ta Ryan Joiner Normallik Testi

Shapiro -Wilk normallik testinin benzeri olan Ryan Joiner testi su sıralamayla yapılmaktadır.

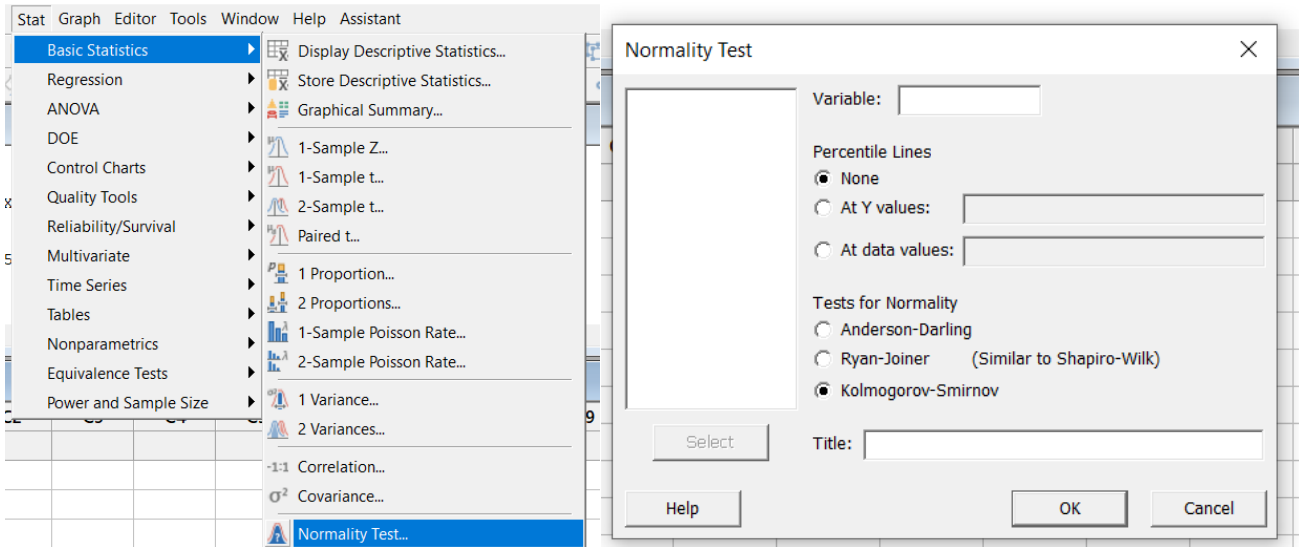
Stat-Basic Statistics-Normality Test – Ryan Joiner



Minitab'ta Kolmogorov Smirnov Normallik Testi

Kolmogorov Smirnov testi su sıralamayla yapılmaktadır.

Stat-Basic Statistics-Normality Test – Kolmogorov Smirnov



Minitab'ta Tek Örnek Ortalama Testleri

n tane gözlemden oluşan nicel verilerin tek örnek düzeyinde toplum parametresine dayalı hipotezlerini, örnek istatistikleri aracılığı ile test etmek için kullanılan bir yöntemdir. Bu testin kullanılabilmesi için verilerin normal dağılıma sahip olması gerekir ve önceden belirlenmiş değerden önemli bir ölçüde farklı olup olmadığının testini yaparken kullanılır.

Parametrik olarak tek örnek ortalama testi 2 farklı şekilde uygulanabilir. Bunlar z ve t testleridir. Bu ikisi arasında karar verirken örnek sayısına ve verilerin geldiği anakütlenin standart sapmasının bilinip bilinmediğine dikkat edilir.

- i) **Anakütle normal dağılışlı ve anakütle varyansı biliniyorsa ve $n \geq 30$ ise, Z test istatistiği kullanılır.**

SORU: ilkokul mezunu öğrencilerin Türkçeyi kullanma basari puanları $X \sim N(65,144)$ parametrelili normal degilim gostermektedir. Rastgele secilen ilkokul mezunu 45 öğrencinin Türkçeyi kullanma basari puanlari TDKO olcegi araciligıyla asagidaki gibi belirlenmistir.

BASARI PUANLARI

23 45 65 47 65 48 56 78 90 87 62 68 79 30 48 65 67 65 45 67 66 88 73 52 44 50 35 76 62 59 83 23 43 54 44 66 74 81 88 32 43 61 74 70 75.

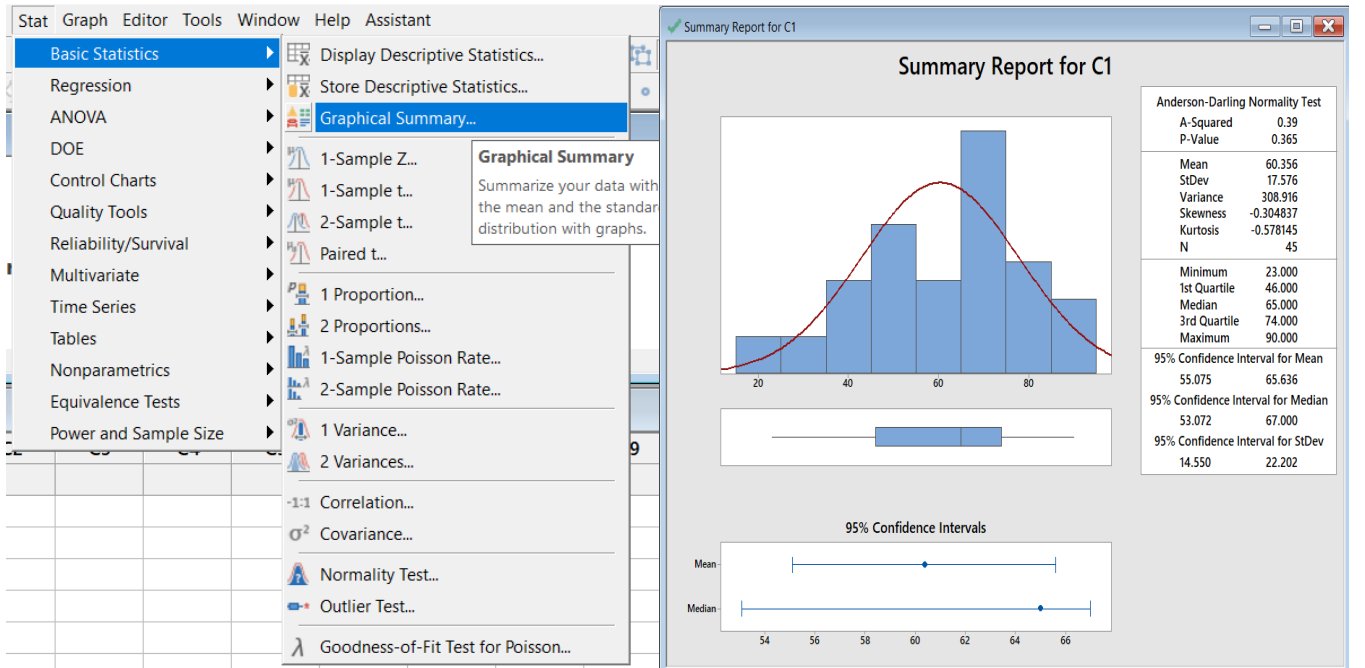
Ogrencilerin basari puanlarinin ortalama deęeri 65'e eřit midir test ediniz.

Çözüm:

Öncelikle rastgele seçilen verilerin normal dağılıma sahip olup olmadığı hipotezi test edilir.

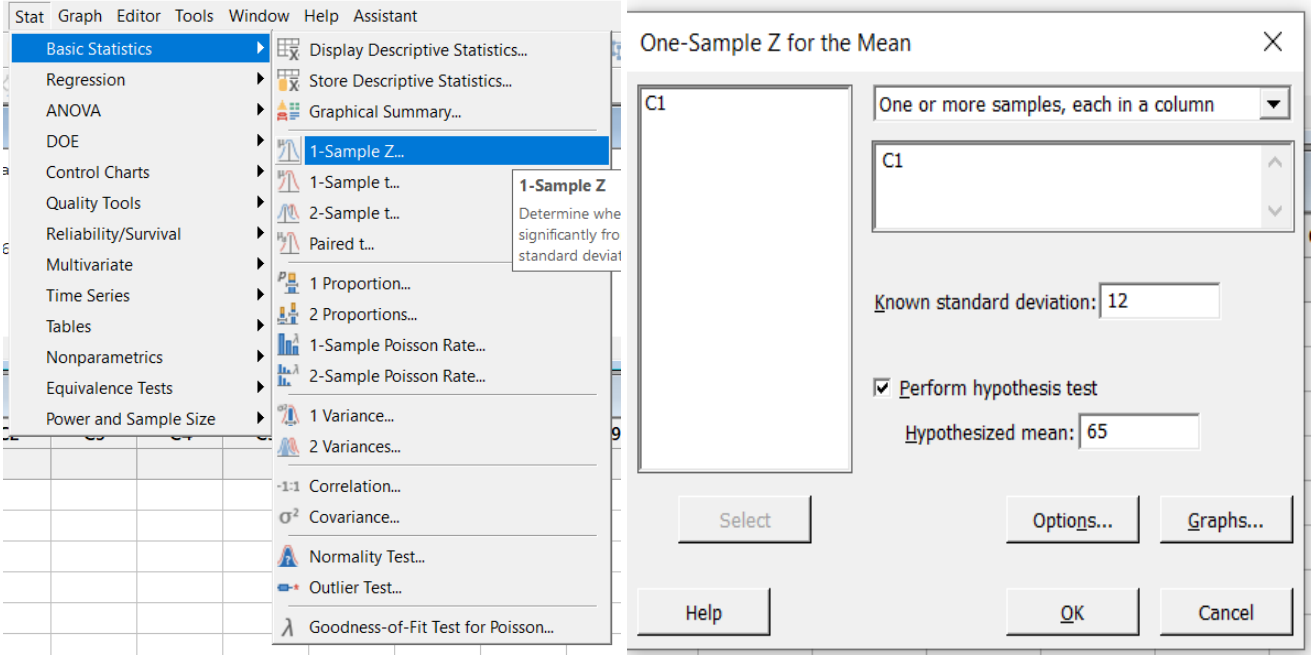
H_0 : Veriler normal dağılmıştır.

H_1 : Veriler normal dağılmamıştır.



Anderson – Darling sonucuna gore p deęeri yüzde 5 önem seviyesinden yüksek olduęu için yokluk hipotezi reddedilemez. Dolayısıyla veriler normal dagilisa sahiptir diyebiliriz.

Elimizde rastgele seçilmiş, normal dagilima sahip, tek grup veriler bulunmaktadır. Ayrıca bu verilerin varyansı bilinmektedir. Bu yazılan nedenler sebebiyle tek grup ortalama testlerinden Z testi seçilir.



One-Sample Z: C1

Test of $\mu = 65$ vs $\neq 65$

The assumed standard deviation = 12

| Variable | N | Mean | StDev | SE Mean | 95% CI | Z | P |
|----------|----|-------|-------|---------|----------------|-------|-------|
| C1 | 45 | 60.36 | 17.58 | 1.79 | (56.85, 63.86) | -2.60 | 0.009 |

Değerler girilip menüden z testi uygulandığında üstteki çıktı elde edilmiştir. Buna göre, p değeri alfa değerinden düşüktür yani yokluk hipotezi reddedilir.

Yani, yüzde 95 güven düzeyinde verilerin ortalaması 65 değildir diyebiliriz.

- ii) **Anakütle standart sapması bilinmiyorsa, örneklem standart sapması kullanılır. $n \leq 30$ olduğunda ve anakütle varyansı bilinmediğinde t dağılımına ait test istatistiği kullanılır.**

SORU: Bir okuldaki 5. Sınıf öğrencilerinin bir sınavdan aldığı skorların ortalaması 65 tir. Rastgele seçilen 10 öğrencinin sınav sonuçları aşağıdaki gibi belirlenmiştir.

(60,55,75,80,90,40,70,100,60,70)

seçilen örnekler tüm okulun bir rastgele örneği midir?

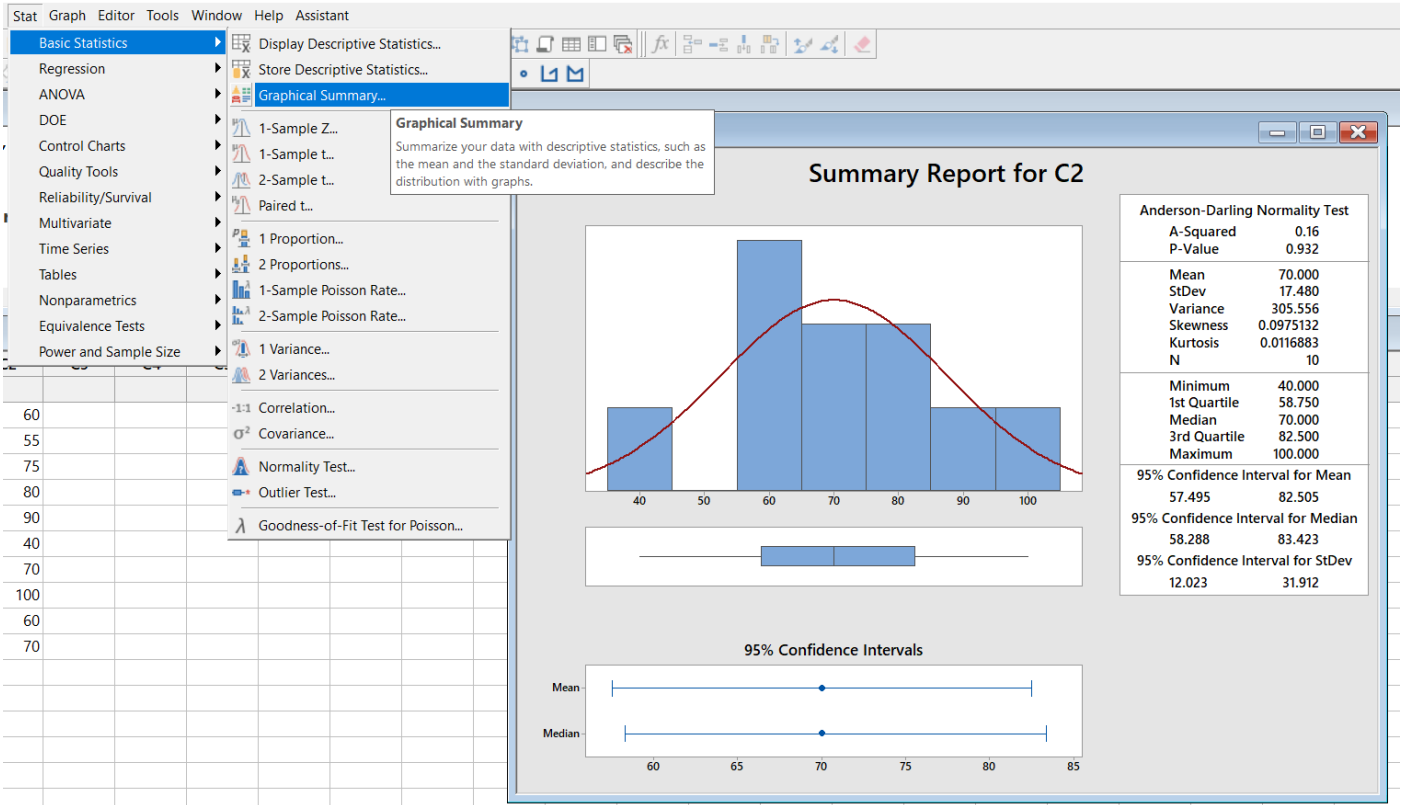
Bu

Cozum:

Tek grup verilerin ortalama testini yapmadan önce verilerin normallik şartını sağlayıp sağlamadığına bakılır.

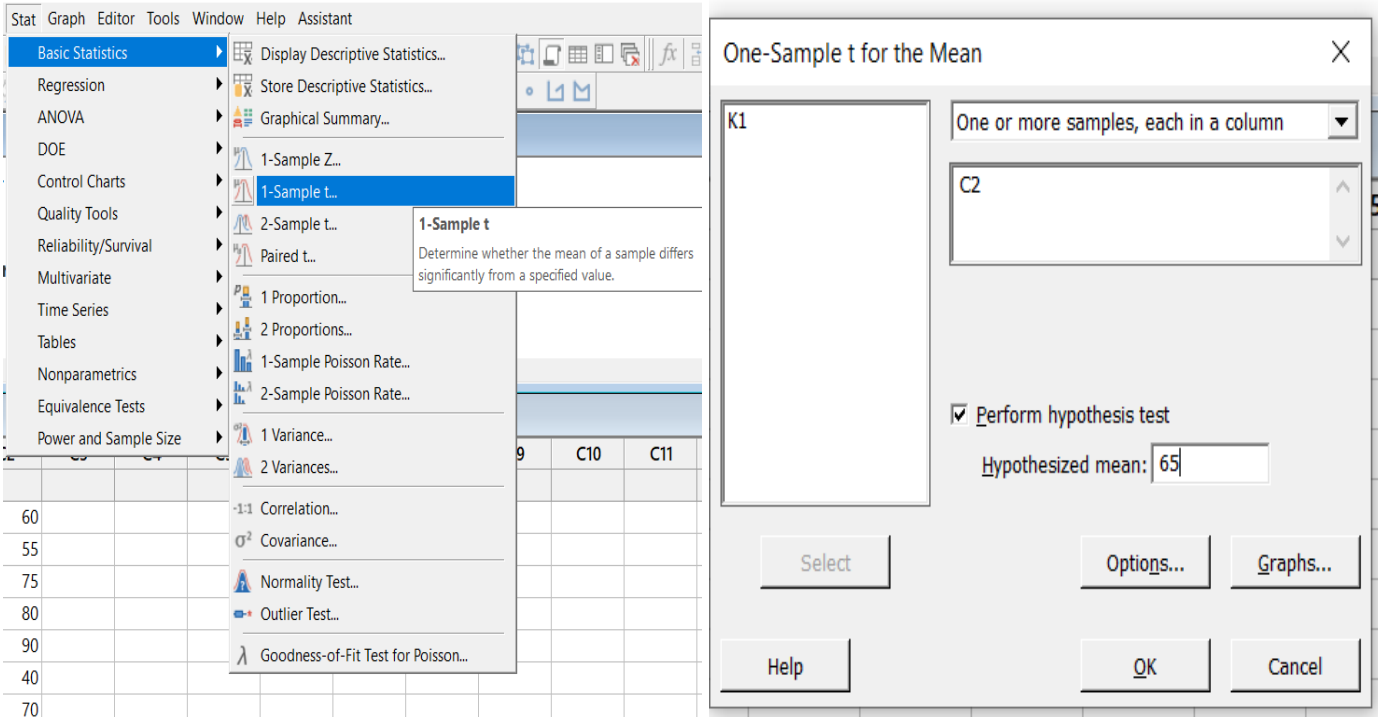
H_0 : Veriler normal dağılmıştır.

H_1 : Veriler normal dağılmamıştır.



Anderson – Darling sonucuna göre p değeri yüzde 5 önem seviyesinden yüksek olduğu için yokluk hipotezi reddedilemez. Dolayısıyla veriler normal dağılıma sahiptir diyebiliriz.

Elimizde rastgele seçilmiş, normal dağılıma sahip, tek grup veriler bulunmaktadır. Veri sayısı 30'dan küçük ve verilerin geldiği örneklem varyansı bilinmemektedir. Buna göre 1 grup için t testi uygulanır.



Bu işlemlerden sonra elde edilen sonuç aşağıdaki gibidir.

One-Sample T: C2

Test of $\mu = 65$ vs $\neq 65$

| Variable | N | Mean | StDev | SE Mean | 95% CI | T | P |
|----------|----|-------|-------|---------|----------------|------|-------|
| C2 | 10 | 70.00 | 17.48 | 5.53 | (57.50, 82.50) | 0.90 | 0.389 |

Buna göre, p değeri alfa değerinden düşüktür yani yokluk hipotezi reddedilir. Yani, yüzde 95 güven düzeyinde verilerin ortalaması 65 değildir diyebiliriz.

Bu iki testi de veriler normal dağılırken uygulayabildik. Fakat veriler normal değilse 1 örneklem için Wilcoxon testi uygulanır.

SORU: Atakum Denizevleri Mahallesiindeki ev fiyat ortalaması Emlakçılar Birliği tarafından 150 birim olarak belirtilmiştir. Yapılan bir araştırmada 29 tane örnek seçilip onların fiyatları aşağıdaki gibi belirtilmiştir. Bu seçilen örneklemin ortalaması 150 birime eşit midir test ediniz.

60 50 56 72 80 80 80 99 101 110 110 110 120 140 144 145 150 180 201 210 220 240 290 309 320 325 400 500 507

Cozum:

Tek grup verilerin ortalama testini yapmadan önce verilerin normallik şartını sağlayıp sağlamadığına bakılır. Eğer veriler normal dağılmıyorsa 1 grup için ortanca testi yapılır.

Öncelikle veriler aşağıdaki komutla 3. Kolona girilir.

MTB > set c3

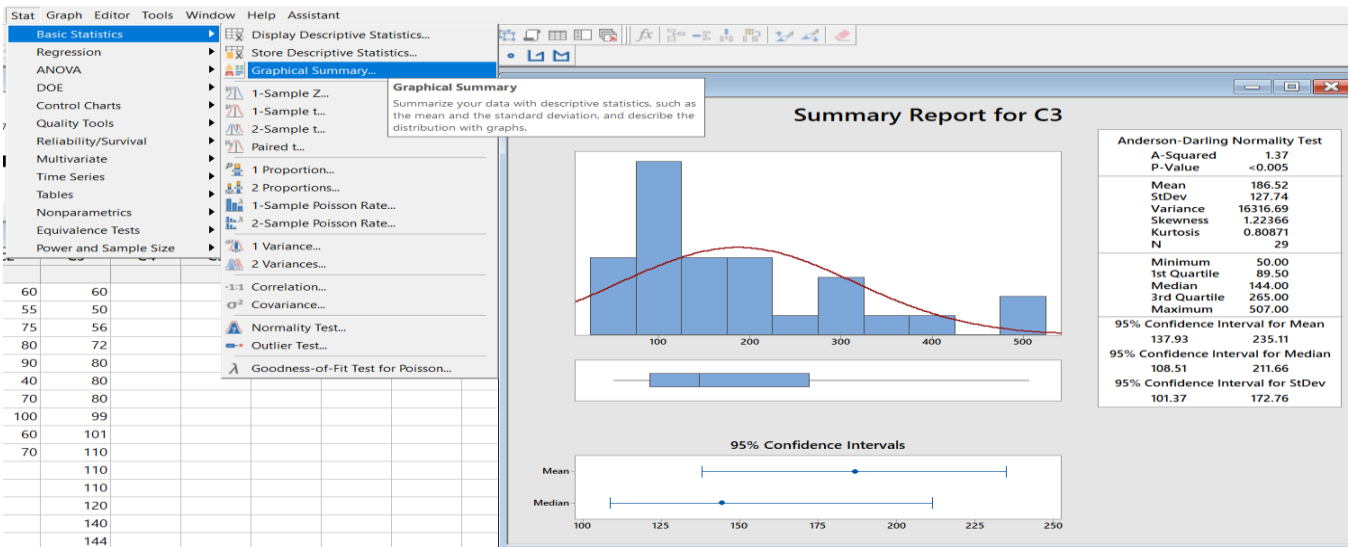
DATA> 60 50 56 72 80 80 80 99 101 110 110 110 120 140 144 145 150 180 201 210 220 240 290 309 320 325 400 500 507

DATA> end

Veriler için normallik varsayımı sağlanıp sağlanmadığı kontrol edilir.

H_0 : Verilerin normal dağılmıştır.

H_1 : Veriler normal dağılmamıştır.



Çıkan sonuca göre p değeri alfa değeri olan yüzde 5 ten küçüktür. Bu da bize yokluk hipotezi reddedilme imkânı sunar. Yani, veriler normal dağılışa sahip değildir, diyebiliriz. Böylece, 1 örneklem için parametrik olmayan test olan Wilcoxon testi uygulanır aşağıdaki hipotezler için uygulanır.

H_0 : Ev fiyatlarının ortancası 150 birimdir.

H1: Ev fiyatlarının ortancası 150 birim değildir.

Stat Graph Editor Tools Window Help Assistant

Basic Statistics
Regression
ANOVA
DOE
Control Charts
Quality Tools
Reliability/Survival
Multivariate
Time Series
Tables

Nonparametrics
Equivalence Tests
Power and Sample Size

150.0
ted
ian
0.0

1-Sample Sign...
1-Sample Wilcoxon...
Mann-Whitney...
Kruskal-Wallis...
Mood's Median Test...
Friedman...
Runs Test...
Pairwise Averages...
Pairwise Differences...
Pairwise Slopes...

1-San
Determ
signific
symm

1-Sample Wilcoxon

Variables:
C3

Confidence interval
Level: 95.0

Test median: 150

Alternative: not equal

Select
Help
OK
Cancel

Bu uygulamalar sonucu asagidaki cikti elde edilir.

Wilcoxon Signed Rank Test: C3

Test of median = 150.0 versus median \neq 150.0

| | N | Test | Wilcoxon Statistic | P | Estimated Median |
|----|----|------|-----------------------|-------|---------------------|
| C3 | 29 | 28 | 235.0 | 0.473 | 170.0 |

P değeri yüzde 5 ten büyüktür bu nedenle H_0 hipotezi reddedilemez. Yani, yüzde 95 güven düzeyinde ev fiyatlarının ortancası 150 birimdir diyebiliriz.

Minitab'ta Bağımsız İki Grup Ortalama Testleri

İki tane bağımsız grup için normal dağılıma bağımlı bir değişkenin ortalamalarının karşılaştırılmak istendiğinde kullanılır. Verilerin rastgeleliği, normallik varsayımı ve varyansların homojenliği varsayımları altında parametrik olanları uygulanır. Varsayımları test etmek için normallik testi ve varyans homojenliği testi uygulanır.

Bu varsayımlar sağlanmadığında ise parametrik olmayan testlerden Mann Whitey testi kullanılır.

Örnek: Psikolog uykunun hatırlama üzerine etkisini araştırıyor. Günde 8 saat uyuyan 12 öğrenci ile günde 5 saat uyuyan diğer 12 öğrenciye hatırlama testi uyguluyor.

Öğrencilerin aldığı puanlar aşağıdaki şekilde ölçülüyor.

8 saat uyuyanlar: 40 45 52 61 65 75

5 saat uyuyanlar: 30 35 48 52 54 60

Bu iki grup arasında önemli bir fark olup olmadığını test ediniz.

Çözüm: İki bağımsız grubun kıyaslamasının uygulandığı bu test için öncelikle parametrik mi parametrik olmayan bir yöntem mi seçileceğine karar verilir.

Hipotezler:

H_0 : İki grup ortalamasi arasında fark yoktur.

H_1 : İki grup ortalamasi arasında fark vardır.

Normallik testi:

H_0 : Veriler normal dağılmıştır.

H_1 : Veriler normal dağılmamıştır.

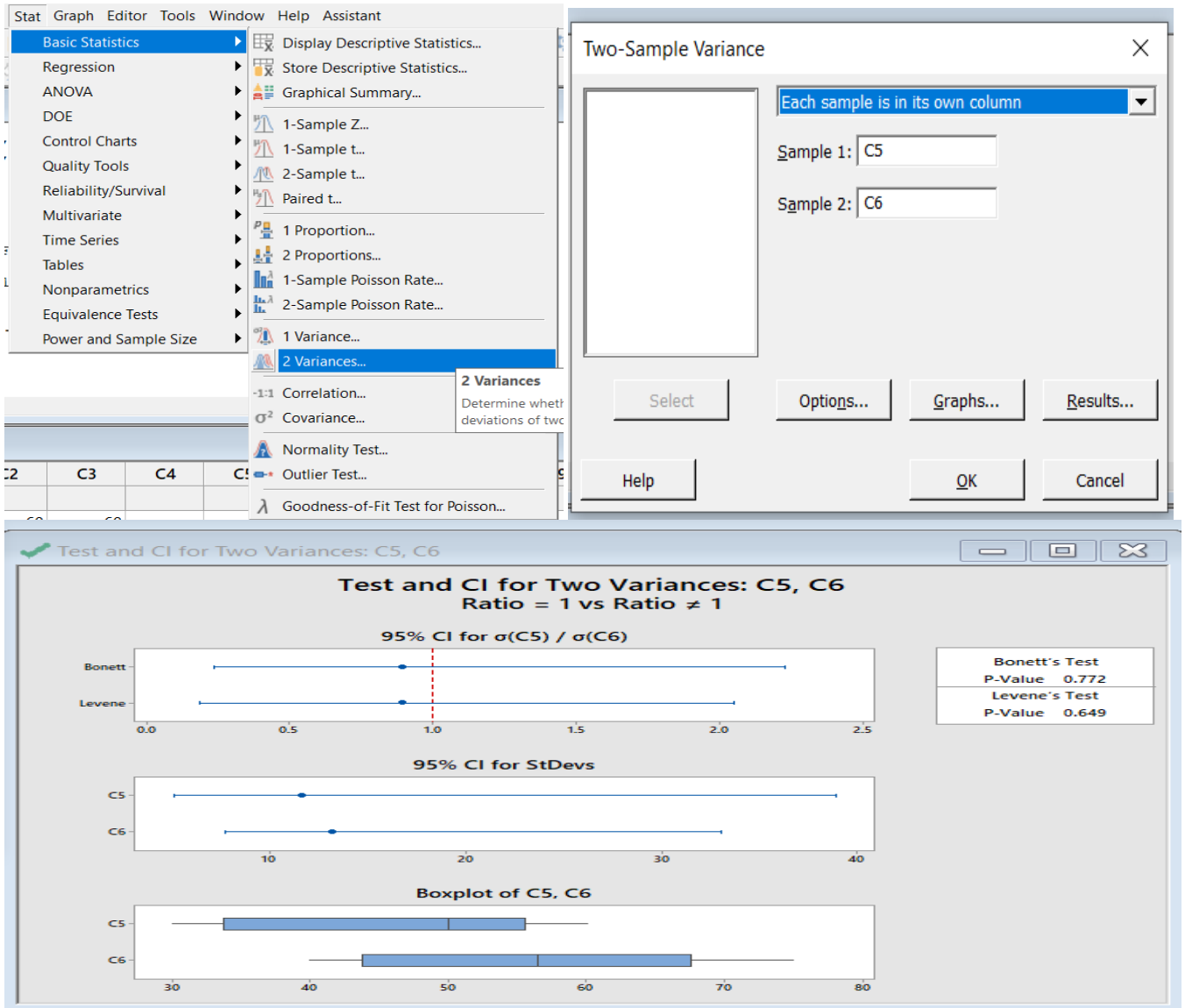
İki grup için p değerleri sırasıyla şöyledir: 0.465 ve 0.897 . Bu sonuçlara göre H_0 hipotezi kabul edilir.Yani veriler normal dağılır diyebiliriz.

Homojenlik testi:

H_0 : Varyanslar homojen.

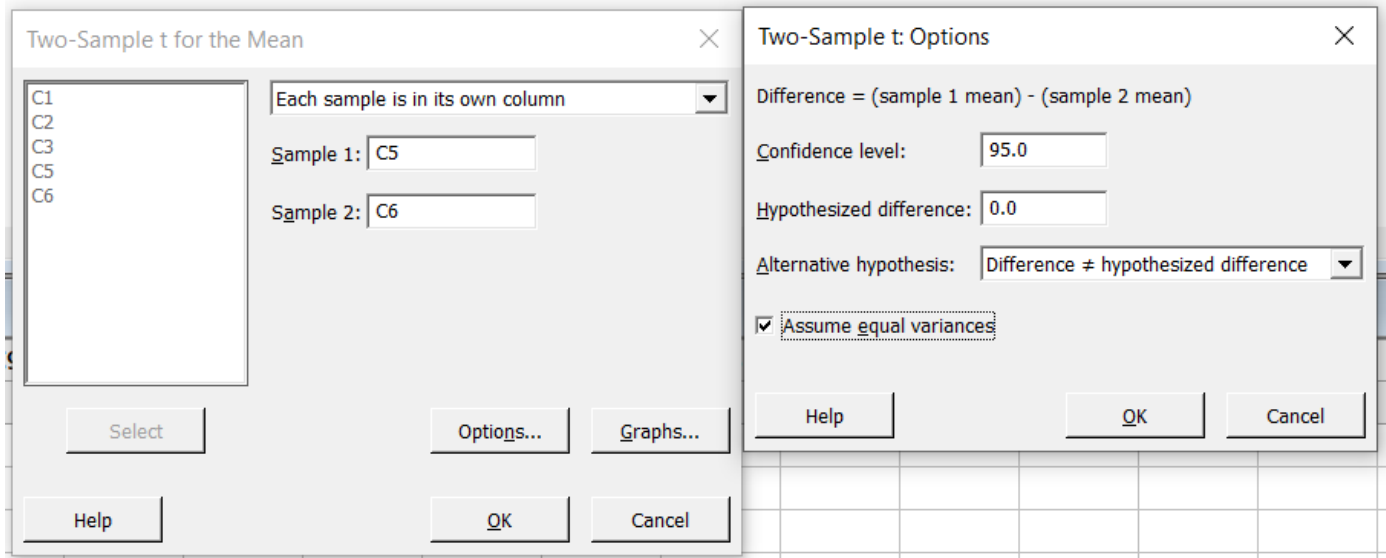
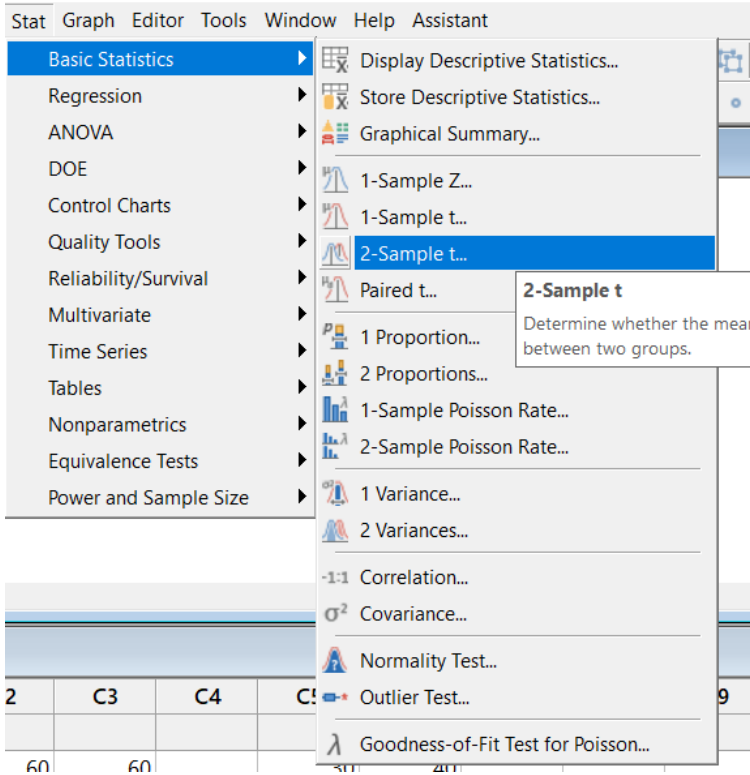
H_1 : Varyanslar heterojendir.

Homojenlik testi Levene's Test aracılığı ile yapılır.



Levene testine göre p değeri yüzde 5 ten büyüktür. Bu nedenle varyanslar homojendir diyebiliriz.

Veriler normal dağılımlı ve gruplararası varyanslar eşit yani homojen olduğuna göre parametrik bir yöntem olan 2 sample t test uygulanır.



Two-sample T for C5 vs C6

| | N | Mean | StDev | SE Mean |
|----|---|------|-------|---------|
| C5 | 6 | 46.5 | 11.6 | 4.7 |
| C6 | 6 | 56.3 | 13.1 | 5.4 |

Difference = μ (C5) - μ (C6)
 Estimate for difference: -9.83
 95% CI for difference: (-25.77, 6.10)
 T-Test of difference = 0 (vs \neq): T-Value = -1.37 P-Value = 0.199 DF = 10
 Both use Pooled StDev = 12.3888

Uygulamadan elde edilen sonuçlara göre; H_0 hipotezi reddedilemez yani iki grup ortalaması arasında fark yoktur diyebiliriz.

Minitab'ta Bağımlı İki Grup Ortalama Testleri

Bağımlı iki örneklem t testi, rastgele seçilen n birimlik gruptan iki farklı zamanda ve iki farklı uygulamaya ilişkin elde edilen iki veri setinin farklarının sıfır ortalamalı toplumun rastgele örnekleri olup olmadığını test etmek için uygulanan bir yöntemdir. Aynı fert üzerinde farklı zamanlarda ölçümler alındığında ve bunların karşılaştırılması söz konusu olduğu durumlarda bağımlı (eşli) grup ortaya çıkar.

Bu testin uygulanabilmesi için verilerin normal dağılıma sahip olması gerekmektedir. Eğer veriler normal dağılıma sahip değil ise parametrik olmayan yöntemlerden Wilcoxon testi uygulanır.

Örnek: Üniversite öğrencilerinin vize ve final notları arasındaki başarı durumunu ölçmek isteyen bir öğretim elemanı 15 kişilik öğrenci grubunun vize ve final notlarında anlamlı bir farklılık olup olmadığını nasıl test edebilir?

| Vize | Final | Vize | Final | Vize | Final |
|------|-------|------|-------|------|-------|
| 45 | 75 | 62 | 73 | 77 | 92 |
| 67 | 73 | 48 | 76 | 81 | 90 |
| 60 | 85 | 63 | 80 | 56 | 70 |
| 55 | 72 | 72 | 95 | 45 | 60 |
| 48 | 56 | 50 | 82 | 68 | 87 |

Çözüm: İki bağımlı grubun kıyaslamasının uygulandığı bu test için öncelikle parametrik mi parametrik olmayan bir yöntem mi seçileceğine karar verilir.

Hipotezler:

H_0 : Vize ve final not ortalaması arasında fark yoktur.

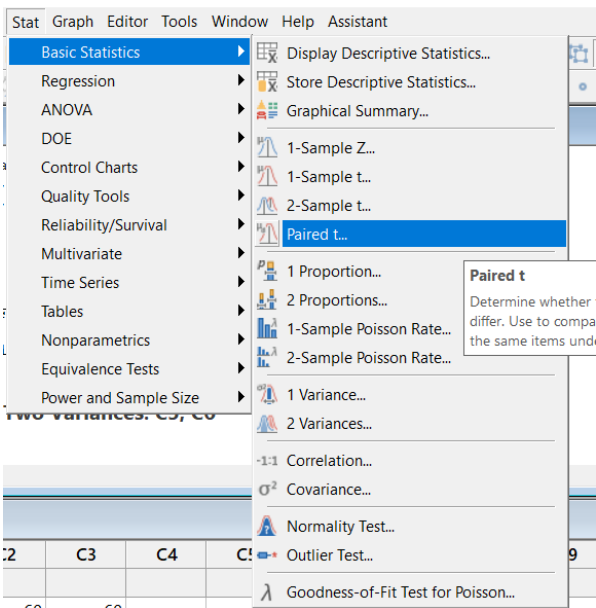
H_1 : Vize ve final not ortalaması arasında fark vardır.

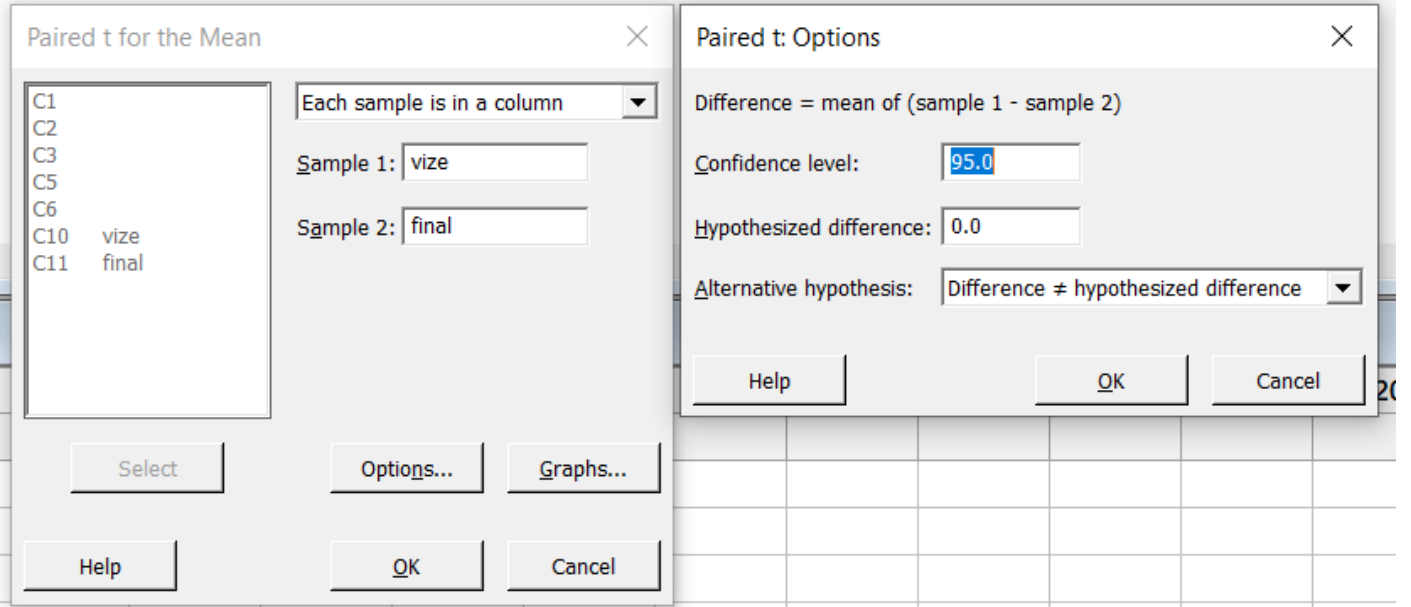
Normallik testi:

H_0 : Veriler normal dağılmıştır.

H_1 : Veriler normal dağılmamıştır.

İki sınav için p değerleri sırasıyla şöyledir: 0.631 ve 0.809 . Bu sonuçlara göre H_0 hipotezi kabul edilir.Yani veriler normal dağılır diyebiliriz. O zaman parametrik bir test olan paired samples t test uygulanır.





Secilen bu teste ve ozellilere gore su sonuc elde edilmektedir.

Paired T-Test and CI: vize, final

Paired T for vize - final

| | N | Mean | StDev | SE Mean |
|------------|----|--------|-------|---------|
| vize | 15 | 59.80 | 11.58 | 2.99 |
| final | 15 | 77.73 | 11.18 | 2.89 |
| Difference | 15 | -17.93 | 8.13 | 2.10 |

95% CI for mean difference: (-22.43, -13.43)

T-Test of mean difference = 0 (vs ≠ 0): T-Value = -8.55 P-Value = 0.000

p değeri alfa için belirlenen yüzde 5 ten az çıktığı için H_0 hipotezi reddedilir. Yani vize ve final sınavları ortalaması arasında farklılık vardır.

Minitab'ta Tek Grup Oran Testi(1-Proportion)

Bir gruptan oluşan verinin oranının önceden yapılmış çalışmanın oranından ya da belirlenmiş bir orandan farklı olup olmadığının araştırmasında bu testten faydalanılır. Burada z testinin istatistiğinden faydalanılır.

Ornek: Bir fabrika ürettiği kaliteli ürün oranının yüzde 65 üzerinde olduğunu iddia etmektedir. Yapılan araştırma sonucu üretilen 2000 urundan 300 unun kalite testinden geçemediği tespit edilmiştir. Buna göre iddia edilen oranı test ediniz.

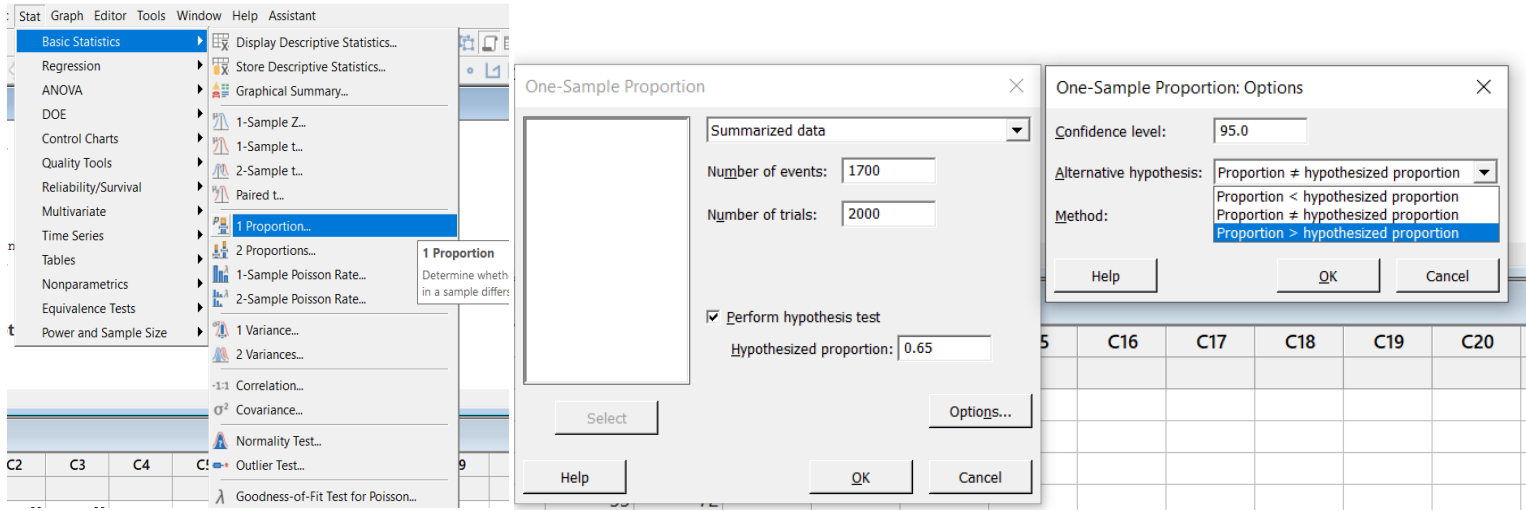
Çözüm: Burada karşılaştırılan değerler oran olduğu için tek grup oran testi uygulanır.

Hipotezler:

H_0 : Kaliteli ürün oranı yüzde 65 tir.

H_1 : Kaliteli ürün oranı yüzde 65 ten fazladır.

Stat- Basic Statistics-1 Proportion adımlarıyla testimize ulaşırız.



Test and CI for One Proportion

Test of $p = 0.65$ vs $p > 0.65$

| Sample | X | N | Sample p | 95% Lower Bound | Exact P-Value |
|--------|------|------|----------|-----------------|---------------|
| 1 | 1700 | 2000 | 0.850000 | 0.836237 | 0.000 |

Yapılan teste göre p değeri 0 çıkmıştır ve alfa dan küçüktür. Bu nedenle H_0 hipotezi reddedilir. Yani kaliteli ürün oranı yüzde 65 ten fazladır diyebiliriz.

Minitab'ta İki Grup Oran Testi(1-Proportion)

İki farklı popülasyona ait oranların karşılaştırılmasında kullanılan bu test için yine z istatistiği kullanılır. Minitab üzerinden ise şu şekilde teste ulaşılır: **Stat- Basic Statistics-2 Proportions**.

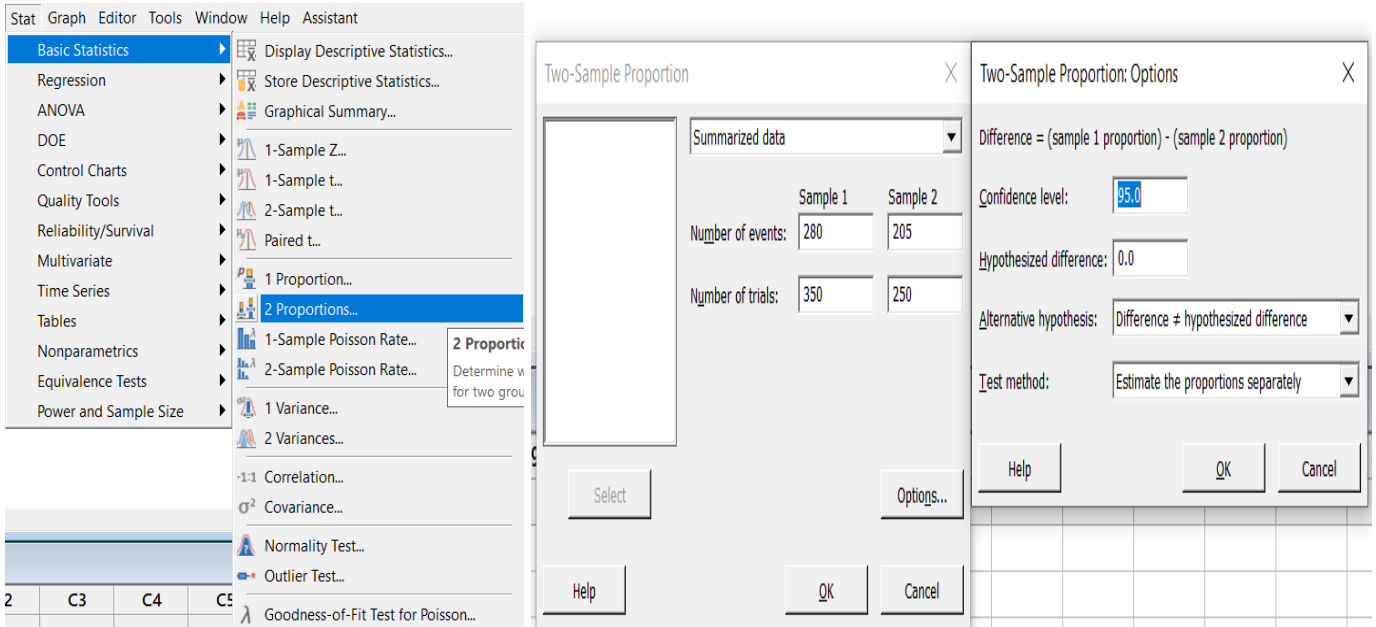
Ornek: Beyaz eşya üretimi yapan bir firma, müşteri memnuniyeti adına ürettiği A ve B ürünlerinin kalite kontrol çalışmaları yürütmeye karar vermiştir. Yapılan çalışmalarda 350 A, 250 B ürünü incelenmiştir. A ürününden 70 tanesi B ürününden 45 tanesi kaliteden geçememiştir. Buna göre A ve B ürünlerini üretim kalitesi açısından kıyaslayınız.

Çözüm: Burada grupların kalitesi için oranlar karşılaştırılmıştır. Bu nedenle iki grup oran testi uygulanılır.

Hipotezler:

H_0 : Ürün kalite oranları arasında farklılık yoktur.

H_1 : Ürün kalite oranları arasında farklılık vardır.



Test and CI for Two Proportions

| Sample | X | N | Sample p |
|--------|-----|-----|----------|
| 1 | 280 | 350 | 0.800000 |
| 2 | 205 | 250 | 0.820000 |

Difference = p (1) - p (2)
 Estimate for difference: -0.02
 95% CI for difference: (-0.0834357, 0.0434357)
 Test for difference = 0 (vs ≠ 0): Z = -0.62 P-Value = 0.537

Fisher's exact test: P-Value = 0.599

Test sonucuna göre p değeri yüzde 5 ten büyük olduğu için H_0 hipotezi reddedilemez. Yani, iki ürün arasında kalite olarak anlamlı bir farklılık yoktur.

Minitab'ta Tek Yönlü Varyans Analizi (ANOVA)

Bu asmaya kadar kadar bir ya da iki gruba sahip verilerin ortalamalarının kıyaslanması üzerinde durulmuştur. Bu bölümde ise bağımsız grup sayısının 3 ve 3 ten fazla olduğu durumların ortalama karşılaştırılmasında kullanılan ANOVA testi yapılacaktır.

ANOVA, normal dağılımlı verilerle oluşturulan birbirinden bağımsız ikiden çok grubun, ilave olarak varyansların homojenliği ve artıkların normal dağılımlı olması varsayımları altında kullanılabilir. ANOVA için varsayımlar sağlanmadığı zaman hangi testler uygulanması gerektiği aşağıdaki gibi açıklanmıştır.

- Eğer veriler normal dağılımlı ve varyanslar homojense: **ANOVA**
- Eğer veriler normal dağılımlı ve varyanslar homojen değilse: **Welch**
- Veriler normal dağılımlı değil ise : **Kruskal Wallis**

ANOVA'nın yorumlanması:

H_0 : Grupların ortalaması arasında fark yoktur.

H_1 : Ortalamalardan en az biri farklıdır.

Analiz sonucunda çıkan p değeri alfa değerinden küçük ise en az iki deney arasında anlamlı fark vardır. Bu durumda ise çoklu karşılaştırma testleri yani Post-hoc testleri kullanılarak ikili farklar elde edilir. Önemli bulunan ikili farklar ve ortalamalar rapor edilir.

Coklu Karsilastirma Testleri

Varyans analizi sonucunda cikan sonu gruplarin birbirinden farkli olduėunu gsterir sekilde ise bu farklılıkları ortaya cikan grup ya da grupların ve ortalamaları farklılık gösteren grupları tespit etmek amacıyla oluşan testlere çoklu karşılaştırma testleri denir.

Bu testler; homojen varyans yaklaşımını kullananlar ve heterojen varyans yaklaşımını kullananlar olarak iki temel gruba ayrılırlar.

Homojen varyans yaklaşımından en çok kullanılan testlere TUKEY ve DUNCAN, heterojen varyans yaklaşımından ise Tamhane ve Games – Howell testlerini sayabiliriz.

Veriler normal dağılıma sahip değilken ise Dunn çoklu karşılaştırma testini kullanabiliriz.

SORU: Dört ayrı grupta ölçülen serum Total Protein değerleri verilmiştir. Total Protein değerinin hastalık türlerine göre farklılık göstermediğini test ediniz.

Veriler

```
MTB > set c1
DATA> 6,4 6,5 6,7 5,9 6,0 8,6 9,7 10,1 9,6 9,4 13,7 13,8 14,9 16,3 15,4 21,8 17,9 21,6 20,5 19,7
DATA> end
```

Çözüm: 4 ayrı bağımsız grup karşılaştırması olduğu için tek yönlü varyans analizi normallik, rastgelelik,bağımsızlık, hataların bağımsızlığıve varyans homojenliği varsayımları altında uygulanır.

Hipotezler:

H_0 : Grup ortalamaları arasında farklılık yoktur.

H_1 : En az bir grup ortalaması diğerlerinden farklıdır.

Normallik testi:

H_0 : Veriler normal dağılmıştır.

H_1 : Veriler normal dağılmamıştır.

Anderson Darling testine göre p değerleri alfa değeri olan yüzde 5 ten büyüktür. Yani veriler normal dağılışıdır.

Homojenlik testi:

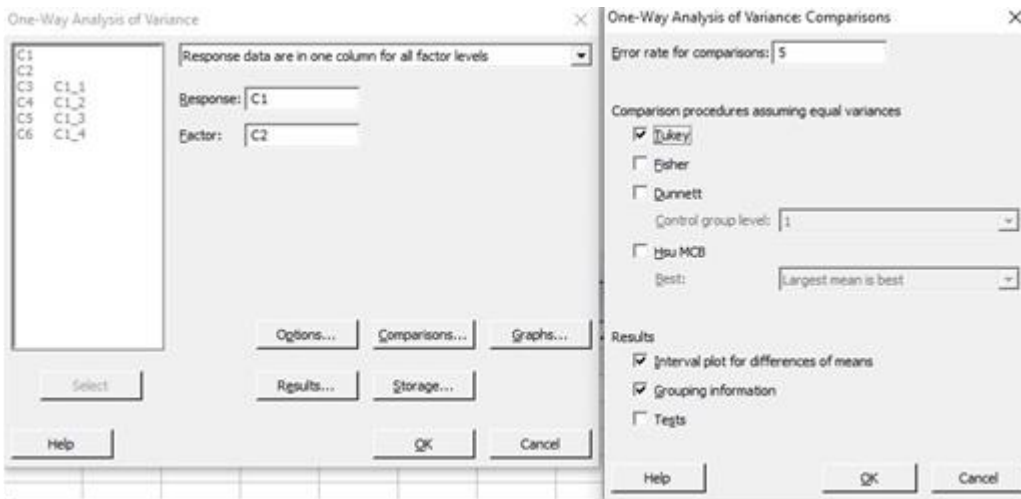
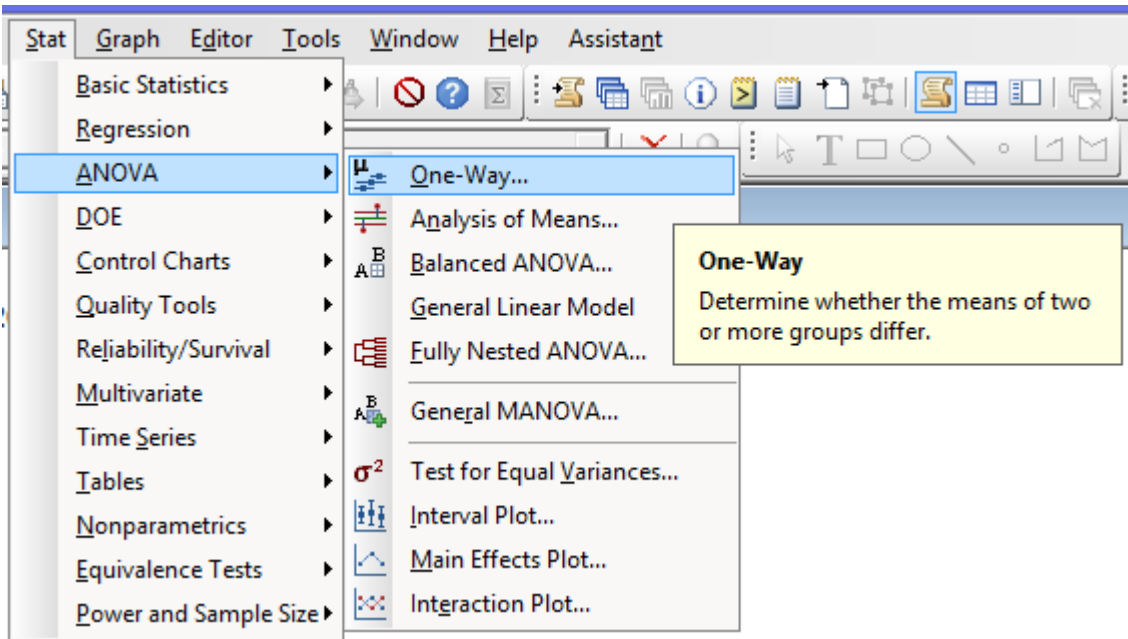
H_0 : Varyanslar homojen.

H_1 : Varyanslar heterojendir.

Homojenlik testi Levene's Test aracılığı ile yapılır.

P değeri 0.099 yani 0.05 ten büyüktür. Bu nedenle H_0 hipotezi reddedilemez ve gruplararası varyanslar homojendir.

Grup sayısı 4, gruplar bağımsız, veriler normal dağılışı,gruplararası varyanslar eş.Bu nedenle Anova testi uygulayabiliriz. Stat > ANOVA > One-Way adımları ile bu teste ulaşırız.



Analysis of Variance

| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|--------|----|--------|---------|---------|---------|
| C2 | 3 | 567,90 | 189,300 | 182,28 | 0,000 |
| Error | 16 | 16,62 | 1,039 | | |
| Total | 19 | 584,52 | | | |

Buna göre p değeri yüzde 5 ten küçüktür ve H_0 hipotezi reddedilir. Yani gruplar arasında farklılık vardır.

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

| C2 | N | Mean | Grouping |
|----|---|--------|----------|
| 4 | 5 | 20,300 | A |
| 3 | 5 | 14,820 | B |
| 2 | 5 | 9,480 | C |
| 1 | 5 | 6,300 | D |

Bu çıktıya göre de tüm grupların birbirinden farklı oldukları görülür.Çünkü hepsi farklı bir harfle adlandırılmıştır.

Ödev:Eğer veriler normal dağılımlı fakat homojen varyanslı olmasaydı(öyle kabul edelim aynı veriler için) hangi test, çoklu karşılaştırma testi uygulanmalıydı Minitab üzerinden çözünüz.

İki Faktörlü Varyans Analizi

NEDİR ?

- Tek yönlü varyans analizinde;bir faktörün(bağımsız değişkenin) bağımlı değişken üzerindeki etkisi araştırılır.
- Bazı araştırmalarda ise deneyin hassaslığını ve araştırmanın başarısını arttırmak adına iki faktörün bağımlı değişken üzerindeki etkisi araştırılır.
- Bu tip varyans analizi iki faktörlü(yönlü) varyans analizi olarak adlandırılır ve parametrik bir yöntemdir.

İki Faktörlü Varyans Analizi

Nicin Kullanilir ?

- Bu analiz yalnızca tek faktörün etkisinin ölçüldüğü ve ayrı ayrı yapılan iki tane tek yönlü varyans analizine tercih edilmektedir.
- Birinci olarak iki kez aynı testin gerçekleştirilmesi 1. tip hatanın büyümesine neden olmaktadır.
- İki yönlü faktör analizinde, iki faktörün yalnızca birbirini etkilemesi sonucu ortaya çıkan bir etkiyi de ölçmesidir.
- İlave olarak da iki kez test yapılması yerine tek seferde varyans analizinin yapılması bu yöntemi cazip ve mantıklı kılmaktadır.

İki Faktörlü Varyans Analizi

Ornek Arastirma Sorusu

- Örneğin, stres düzeyi üzerinde fiziksel aktivite seviyesi ile cinsiyetin etkisi olup olmadığını anlamak için iki yönlü ANOVA kullanılabilir.
- Yani, bağımlı değişken "stres skoru", sürekli bir ölçekte ölçülmüş olup, bağımsız değişkenler ise "düşük", "orta" ve "yüksek" gibi üç farklı grupta belirtilen fiziksel aktivite seviyesi ile "erkekler" ve "kadınlar" olmak üzere iki gruba ayrılan "cinsiyet" olacaktır.

İki Faktörlü Varyans Analizi

- Minitab üzerinde uygulanmasına geçmeden önce bu testin sahip olduğu bazı prosedürlerin yerine getirilmesi gereklidir. Bunları hatırlamak gerekirse;

İki Faktörlü Varyans Analizi

- **Grupların Bağımsızlığı:**

Her gruptaki gözlemler arasında veya grupların kendileri arasında bir ilişki olmadığı anlamına gelir.

Örneğin, hiçbir katılımcının birden fazla grupta olmadığı her grupta farklı katılımcıların yer alması gerekmektedir. Bu sağlanmazsa gruplar birbirine bağlı(esli) olur.

İki Faktörlü Varyans Analizi

- **Verilerin Normalligi** : Parametrik bir yöntem olan ANOVA normallik varsayımı altında uygulanır. Bu varsayım sağlanmazsa Friedman testi uygulanır .
Minitab üzerinden control edilebilir
- **Verilerin Nicelliği**: Bağımlı değişkenin verileri niceldir.Gruplar ise kategoriktir.
- **Örneklemin Rastgeleliği**
- **Hataların Normalligi ve Bağımsızlığı**
- **Varyansların Homojenliği** : Gruplar arası varyansların eşitliği ya da esite yakın olması varsayımına dayanır. Levene's test ile control edilebilir .

İki Faktörlü Varyans Analizi

Uygulama Sorusu

1.Cinsleri T1, T2, T3, T4 olan dört farklı tohum, 3 farklı gübre G1,G2,G3 uygulanmış ve birim başına elde edilen miktarlar aşağıdaki tabloda verilmiştir.

| Gübre Cinsleri | Tohum Cinsleri | | | |
|----------------|----------------|----|----|----|
| | T1 | T2 | T3 | T4 |
| G1 | 3 | 6 | 1 | 2 |
| G2 | 5 | 7 | 4 | 6 |
| G3 | 2 | 3 | 2 | 2 |

İki Faktörlü Varyans Analizi

Uygulama Sorusu 1

Buna göre;

a) Farklı tohum tipleri, verim üzerinde etkili midir?

b) Farklı gübre tipleri, verim üzerinden etkili midir?

% 5 önem seviyesinde test ediniz.

İki Faktörlü Varyans Analizi

Uygulama Sorusu 2

2. Cinsleri T1, T2, T3olan uc farkli tohuma, 3 farkli gubre G1,G2,G3 uygulanmis ve birim basina elde edilen miktarlar asagidaki tabloda verilmistir.

| Verim | 500 | 580 | 460 | 540 | 560 | 600 | 540 | 460 | 560 | 620 | 600 | 580 | 480 | 400 | 420 | 480 | 480 | 410 |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Tohum | T1 | T1 | T1 | T1 | T1 | T1 | T2 | T2 | T2 | T2 | T2 | T2 | T3 | T3 | T3 | T3 | T3 | T3 |
| Gubre | G1 | G1 | G2 | G2 | G3 | G3 | G1 | G1 | G2 | G2 | G3 | G3 | G1 | G1 | G2 | G2 | G3 | G3 |

İki Faktörlü Varyans Analizi

Uygulama Sorusu 2

Buna göre;

a) Farklı tohum tipleri, verim üzerinde etkili midir?

b) Farklı gübre tipleri, verim üzerinden etkili midir?

c) Tohum ve gübre tiplerinin etkileşimi verim üzerinde etkili midir?

% 5 önem seviyesinde test ediniz.

Tek Yonlu MANOVA

NEDİR ?

- Çok degiskenli varyans analizi olarak bilinen Manova icin tek yonlu olaninda bir bagimsiz degiskenin, birden fazla bagimli degiskene etki ettigi durumlarda kullanilir.
- Parametrik bir yontemdir.
- Anovadaki tek degiskenli varsayimlara ilaveten cok degiskenli normallik varsayimi ve varyans - kovaryans matrisi homojenligi varsayimi eklenir.
- Maalesef bu varsayimlarin testi Minitab uzerinden direkt olarak hesaplanamamaktadır.
- Eger bagimsiz degisken sayisi arttirilirse Iki Yonlu Manova halini alır.
- Ayrica Anovanin birden fazla yapilmasiyla tek bir bagimsiz degisken icin hesaplanabilir ama pratik olmamasi, hata oraninin yukselmesi ve yuzde yuz net cevabi alamadigimizdan Manova tercih edilir.

Tek Yonlu MANOVA

Varsayimleri

- En az iki tane bagimli degisken vardir ve bu bagimsiz degiskenler icin very tipi niceldir.
- Iki ve daha fazla gruptan olusan bagimsiz degisken vardir.
- Gozlemler bagimsizdir.
- Cok degiskenli normallik control edilmelidir.(Cok degiskenli normallik spss ve Minitabta yok)
- Ayni sekilde varyans-kovaryans matrisi homojenligi vardir.(Spss Box's M test Minitabta bulunmuyor.)
- Hatalarin normal dagilimi varsayimi da gecerlidir.

Tek Yonlu MANOVA

Uygulama

- Bir muhendis olusturdugu 3 farkli alasimin gucunu ve esneklikligini olcuyor.Bu verilere gore alasim gruplarına göre guc ve esneklik degerleri arasında farklılık olup olmadığını test ediniz.