

```
# Abdurrahman Bulut
import pandas as pd
from scipy.stats import shapiro
from scipy.stats import levene
from scipy.stats import ttest_ind, mannwhitneyu
```

Görev 1: Veriyi Hazırlama ve Analiz Etme

Adım 1: ab_testing_data.xlsx adlı kontrol ve test grubu verilerinden oluşan veri setini okutunuz. Kontrol ve test grubu verilerini ayrı değişkenlere atayınız.

```
df_control = pd.read_excel("ab_testing.xlsx", sheet_name="Control Group")
df_test = pd.read_excel("ab_testing.xlsx", sheet_name="Test Group")

print("Kontrol Grubu Verileri:")
print(df_control.head())
print("\nTest Grubu Verileri:")
print(df_test.head())
```

Kontrol Grubu Verileri:

	Impression	Click	Purchase	Earning
0	82529.459271	6090.077317	665.211255	2311.277143
1	98050.451926	3382.861786	315.084895	1742.806855
2	82696.023549	4167.965750	458.083738	1797.827447
3	109914.400398	4910.882240	487.090773	1696.229178
4	108457.762630	5987.655811	441.034050	1543.720179

Test Grubu Verileri:

	Impression	Click	Purchase	Earning
0	120103.503796	3216.547958	702.160346	1939.611243
1	134775.943363	3635.082422	834.054286	2929.405820
2	107806.620788	3057.143560	422.934258	2526.244877
3	116445.275526	4650.473911	429.033535	2281.428574
4	145082.516838	5201.387724	749.860442	2781.697521

Adım 2: Kontrol ve test grubu verilerini analiz ediniz.

```
# Kontrol grubunun temel istatistiklerini analiz etme
print("Kontrol Grubu Temel İstatistikler:")
print(df_control.describe())

# Test grubunun temel istatistiklerini analiz etme
print("Test Grubu Temel İstatistikler:")
print(df_test.describe())

# Satır ve sütun sayısı
```

```
print(f"Kontrol Grubu Boyut: {df_control.shape}")
print(f"Test Grubu Boyut: {df_test.shape}")
```

Kontrol Grubu Temel İstatistikler:

	Impression	Click	Purchase	Earning
count	40.000000	40.000000	40.000000	40.000000
mean	101711.449068	5100.657373	550.894059	1908.568300
std	20302.157862	1329.985498	134.108201	302.917783
min	45475.942965	2189.753157	267.028943	1253.989525
25%	85726.690349	4124.304129	470.095533	1685.847205
50%	99790.701078	5001.220602	531.206307	1975.160522
75%	115212.816543	5923.803596	637.957088	2119.802784
max	147539.336329	7959.125069	801.795020	2497.295218

Test Grubu Temel İstatistikler:

	Impression	Click	Purchase	Earning
count	40.000000	40.000000	40.000000	40.000000
mean	120512.411758	3967.549761	582.106097	2514.890733
std	18807.448712	923.095073	161.152513	282.730852
min	79033.834921	1836.629861	311.629515	1939.611243
25%	112691.970770	3376.819024	444.626828	2280.537426
50%	119291.300775	3931.359804	551.355732	2544.666107
75%	132050.578933	4660.497911	699.862360	2761.545405
max	158605.920483	6019.695079	889.910460	3171.489708

Kontrol Grubu Boyut: (40, 4)

Test Grubu Boyut: (40, 4)

Adım 3: Analiz işleminden sonra concat metodunu kullanarak kontrol ve test grubu verilerini birleştiriniz.

```
# Grupları belirtmek için yeni bir sütun ekleme
df_control['group'] = 'control'
df_test['group'] = 'test'

# Verileri birleştirme
df = pd.concat([df_control, df_test], ignore_index=True)

print("Birleşmiş Veri:")
print(df.head())
```

Birleşmiş Veri:

	Impression	Click	Purchase	Earning	group
0	82529.459271	6090.077317	665.211255	2311.277143	control
1	98050.451926	3382.861786	315.084895	1742.806855	control
2	82696.023549	4167.965750	458.083738	1797.827447	control
3	109914.400398	4910.882240	487.090773	1696.229178	control
4	108457.762630	5987.655811	441.034050	1543.720179	control

Görev 2: A/B Testinin Hipotezinin Tanımlanması

Adım 1: Hipotezi tanımlayınız.

- $H_0: M_1 = M_2$ (Kontrol grubu ve test grubunun purchase (kazanç) ortalamaları arasında fark yoktur)
- $H_1: M_1 \neq M_2$ (Kontrol grubu ve test grubunun purchase (kazanç) ortalamaları arasında fark vardır)

Adım 2: Kontrol ve test grubu için purchase(kazanç) ortalamalarını analiz ediniz

```
# Kontrol grubunun purchase ortalamasını hesaplama
control_purchase_mean = df_control['Purchase'].mean()
print(f"Kontrol Grubu Purchase Ortalama: {control_purchase_mean}")

# Test grubunun purchase ortalamasını hesaplama
test_purchase_mean = df_test['Purchase'].mean()
print(f"Test Grubu Purchase Ortalama: {test_purchase_mean}")

# Kontrol ve Test gruplarının purchase (kazanç) ortalamalarını
karşılaştırma
print(f"Kontrol Grubu Purchase Ortalama: {control_purchase_mean:.2f}")
print(f"Test Grubu Purchase Ortalama: {test_purchase_mean:.2f}")
```

```
Kontrol Grubu Purchase Ortalama: 550.8940587702316
Test Grubu Purchase Ortalama: 582.1060966484677
Kontrol Grubu Purchase Ortalama: 550.89
Test Grubu Purchase Ortalama: 582.11
```

GÖREV 3: Hipotez Testinin Gerçekleştirilmesi

Adım 1

Normallik Varsayımı:

- H_0 : Normal dağılım varsayımı sağlanmaktadır.
- H_1 : Normal dağılım varsayımı sağlanmamaktadır.
- $p < 0.05$: H_0 RED
- $p > 0.05$: H_0 REDDEDİLEMEZ

```
# Kontrol grubu için normallik testi
stat, p_value_control = shapiro(df_control['Purchase'])
print(f"Kontrol Grubu Normallik Testi p-değeri: {p_value_control}")

# Test grubu için normallik testi
stat, p_value_test = shapiro(df_test['Purchase'])
print(f"Test Grubu Normallik Testi p-değeri: {p_value_test}")
```

```
# Normallik varsayımı sonuçlarının yorumlanması
if p_value_control > 0.05:
    print("Kontrol grubu normal dağılım varsayımını sağlar.")
else:
    print("Kontrol grubu normal dağılım varsayımını sağlamaz.")

if p_value_test > 0.05:
    print("Test grubu normal dağılım varsayımını sağlar.")
else:
    print("Test grubu normal dağılım varsayımını sağlamaz.")
```

Kontrol Grubu Normallik Testi p-değeri: 0.5891125202178955
 Test Grubu Normallik Testi p-değeri: 0.15413342416286469
 Kontrol grubu normal dağılım varsayımını sağlar.
 Test grubu normal dağılım varsayımını sağlar.

Varyans Homojenliği:

- H_0 : Varyanslar homojendir.
- H_1 : Varyanslar homojen değildir.
- $p < 0.05$: H_0 RED
- $p > 0.05$: H_0 REDDEDİLEMEZ

```
# Varyans homojenliği testi
stat, p_value levene = levene(df_control['Purchase'],
df_test['Purchase'])
print(f"Varyans Homojenliği Testi p-değeri: {p_value levene}")

# Varyans homojenliği sonuçlarının yorumlanması
if p_value levene > 0.05:
    print("Varyanslar homojendir.")
else:
    print("Varyanslar homojen değildir.")
```

Varyans Homojenliği Testi p-değeri: 0.108285882718748
 Varyanslar homojendir.

Adım 2

- Eğer normallik varsayımı ve varyans homojenliği sağlanıyorsa bağımsız iki örneklem t testi yapılacaktır.
- Eğer normallik varsayımı sağlanmıyorsa veya varyans homojenliği sağlanmıyorsa Mann-Whitney U testi yapılacaktır.

```
# Normallik ve varyans homojenliği varsayımları
normallik_control = p_value_control > 0.05
normallik_test = p_value_test > 0.05
varyans_homojenligi = p_value levene > 0.05
```

```
# Uygun testin seçilmesi
if normallik_control and normallik_test and varyans_homojenligi:
    # Bağımsız iki örneklem t testi
    stat, p_value_ttest = ttest_ind(df_control['Purchase'],
df_test['Purchase'], equal_var=True)
    test_type = "Bağımsız İki Örneklem T Testi"
else:
    # Mann-Whitney U testi
    stat, p_value_ttest = mannwhitneyu(df_control['Purchase'],
df_test['Purchase'])
    test_type = "Mann-Whitney U Testi"

print(f"{test_type} p-değeri: {p_value_ttest}")
```

Bağımsız İki Örneklem T Testi p-değeri: 0.34932579202108416

Adım 3

```
if p_value_ttest < 0.05:
    print("Kontrol ve test grubu satın alma ortalamaları arasında
istatistiksel olarak anlamlı bir fark vardır.")
else:
    print("Kontrol ve test grubu satın alma ortalamaları arasında
istatistiksel olarak anlamlı bir fark yoktur.")
```

Kontrol ve test grubu satın alma ortalamaları arasında istatistiksel olarak anlamlı bir fark yoktur.

Task 4: Sonuçların Analizi

- Kontrol Grubu Normallik Testi p-değeri: 0.5891 ($p > 0.05$, normal dağılım varsayımını sağlar)
- Test Grubu Normallik Testi p-değeri: 0.1541 ($p > 0.05$, normal dağılım varsayımını sağlar)
- Varyans Homojenliği Testi p-değeri: 0.1083 ($p > 0.05$, varyanslar homojendir)
- Bu varsayımlar göz önüne alındığında, normallik varsayımı ve varyans homojenliği sağlandığı için bağımsız iki örneklem t testi kullanılmıştır.
- Hem kontrol hem de test grubu normal dağılım varsayımını sağladığı için ve varyanslar homojen olduğu için, bağımsız iki örneklem t testi uygun bir test yöntemidir.
- Bağımsız İki Örneklem T Testi p-değeri: 0.3493 ($p > 0.05$, istatistiksel olarak anlamlı bir fark yoktur)
- Kontrol Grubu Purchase Ortalama: 550.89
- Test Grubu Purchase Ortalama: 582.11
- Yapılan bağımsız iki örneklem t testine göre, kontrol ve test grubu satın alma ortalamaları arasında istatistiksel olarak anlamlı bir fark bulunmamaktadır (p-değeri: 0.3493). Bu, maximum bidding ve average bidding yöntemlerinin satın alma üzerinde önemli bir farklılık yaratmadığını göstermektedir.

- Bu veriler göz önüne alındığında, average bidding yönteminin maximum bidding yöntemine kıyasla belirgin bir avantajı yoktur.
- Ancak, test süresi veya örnek büyüklüğü yetersiz olabilir, bu nedenle test süresini uzatmak veya daha fazla veri toplayarak testi tekrarlamak faydalı olabilir.
- Müşteri memnuniyeti, diğer metrikler veya maliyetler göz önüne alındığında average bidding yöntemi tercih edilebilir. Ancak, sadece satın alma oranına odaklanıldığında, mevcut verilerle her iki yöntem arasında anlamlı bir fark bulunmamaktadır.
- average bidding yönteminin diğer potansiyel avantajlarını (örneğin, reklam maliyetleri, kullanıcı etkileşimi vb.) değerlendirmek ve bunları da göz önünde bulundurmak mantıklı olabilir.