

Final Projesi: Kadın Okuryazarlığı ile Doğurganlık Oranı İlişkisinin İncelenmesi



Bu projede dünya genelindeki kadın okuryazarlığı ile doğurganlık oranı arasındaki ilişkiyi inceleyeceğiz. Bununla ilgili olarak "female_literacy_fertility.csv" dataseti verilmiştir. Bu veri setindeki değişkenler aşağıda verilmiştir.

Country: ülkeler

Continent: kıta

literacy: kadın okuryazarlığı

fertility: doğurganlık yani kadın başına doğan ortalama çocuk sayısı olarak tanımlanır.

population: ülkelerin nüfusu

- "female_literacy_fertility.csv" verisinin ilk 5 satırını okutunuz.
- df deki değişkenlerin yani özelliklerin ne olduğunu gösteriniz.
- değişkenlerin nitel mi nicel mi olduğunu kayıp gözlem olup olmadığını kontrol ediniz ve yorumlayınız.
- veride kaç kıta ve her bir kıtada kaç ülke olduğunu bulunuz.
- okuryazar ve doğum oranı değişkenleriyle ilgili istatistikleri elde ediniz ve yorumlayınız. bu değişkenlerin min ve max değerleri uygun aralıkta mı, kontrol ediniz.
- minimum ve maksimum doğum oranına sahip ilk 5 ülkeyi bulunuz.

- Bu datada Türkiye olup olmadığını kontrol ediniz ve Türkiye ile ilgili değerleri inceleyiniz.
- okuyazar ve doğum oranı arasında ilişki var mıdır? bunun için hem scatter plot(saçılım grafiği) çizin ve hem de korelasyonu bulunuz ve yorumlayınız.
- Elde edilen scatter plot da noktaların renklendirmesini kıtalara göre yapınız ve sonucu yorumlayınız:
- bu ilişki grafiğinde,noktaların boyutunu popülasyona göre de ayarlayıp yorumlayınız.
- her bir kıtanın kadın doğum ve okuyazar oranı değerleri için ortalama, min ve max ve medyan değerlerini bulunuz ve yorumlayınız
- kıtalara göre kadınların okuyazar oranını gösteren çubuk grafiğini çizin ve yorumlayınız.
- kıtalara göre, kadınların doğum oranını gösteren çubuk grafiğini çizin ve yorumlayınız.
- kıtalara göre doğum oranı bakımından uç değer olup olmadığını kontrol ediniz. bunun için aşağıdaki 2 yöntemi kullanınız ve sonuçları yorumlayınız. kutu grafiği çizerek ve uç değer kontrolü yapan bir fonksiyon oluşturarak.
- Popülasyonu 80 milyonun üzerinde olan ülkeleri "büyük ülke" olarak sınıflandırın, aksi takdirde "küçük ülke" olarak sınıflandırın ve bunu df ye yeni bir değişken olarak ekleyiniz.
- büyük ve küçük ülke olarak sınıflandırılan ülkelerin sayısını bulunuz
- büyük ve küçük ülkelerin doğum oranı bakımından min, max ve ortalama değerlerini bulunuz ve sonucu yorumlayınız

""df_year.csv"" verisini okutunuz. bu veri seti 1960-2022 yılları arasında ülkelerin kadın doğum oranlarını vermektedir. burda amacımız Türkiyenin 1960-2021 arasında doğum oranı değerlerinin nasıl değiştiğini belirlemek.

- verinin ilk 5 satırını yazdırınız.
- Türkiye'nin 1960-2021 arası doğum oranının nasıl değiştiğini grafikte çizerek yorumlayınız.

```
import pandas as pd
df = pd.read_csv("female_literacy_fertility.csv", sep=";")
df.head()
```

	Country	Continent	female literacy	fertility	population
0	Chine	ASI	90.5	1.769	1,324,655,000
1	Inde	ASI	50.8	2.682	1,139,964,932
2	USA	NAM	99.0	2.077	304,060,000
3	Indonésie	ASI	88.8	2.132	227,345,082
4	Brésil	LAT	90.2	1.827	191,971,506

```
df.columns

Index(['Country ', 'Continent', 'female literacy', 'fertility',
      'population'], dtype='object')

df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 162 entries, 0 to 161
Data columns (total 5 columns):
#   Column                Non-Null Count  Dtype
#   ...
#   Continent              162 non-null    object
#   female literacy        162 non-null    float64
#   fertility              162 non-null    float64
#   population             162 non-null    int64
```

```

---
0 Country 162 non-null object
1 Continent 162 non-null object
2 female literacy 162 non-null float64
3 fertility 162 non-null float64
4 population 162 non-null object

```

dtypes: float64(2), object(3)

memory usage: 6.5+ KB

#RangeIndex: 162 ,bize 162 satır olduğunu söyler.

#Female literacy ve fertility değişkenlerinin sayısal olduğunu söylerken diğer değişkenlerin metin olduğunu söylüyor.

#Ancak population değeri girilirken , kullanıldığı için onuda metin algılıyor.Yani bu değişkeni kullanacağımız zaman temizlememiz gerekir.

#Ayrıca non-null değerlerlerine bakıldığında boş veri yoktur.

Kaç farklı kıta var?

```
print("Farklı kıta sayısı:", df["Continent"].nunique())
```

Her kıtada kaç ülke var?

```
print("\nHer kıtadaki ülke sayısı:")
```

```
print(df["Continent"].value_counts())
```

Farklı kıta sayısı: 6

Her kıtadaki ülke sayısı:

Continent

AF 49

ASI 47

EUR 36

LAT 24

OCE 4

NAM 2

Name: count, dtype: int64

Temel istatistikler

```
print("Kadın Okuryazarlık Oranı (female literacy)
```

```
İstatistikleri:",df["female literacy"].describe())
```

```
print("Doğurganlık Oranı (fertility)
```

```
İstatistikleri:",df["fertility"].describe())
```

Kadın Okuryazarlık Oranı (female literacy) İstatistikleri: count

162.000000

mean 80.107407

std 23.052415

min 12.600000

25% 66.425000

50% 90.000000

75% 98.500000

max 100.000000

Name: female literacy, dtype: float64

```
Doğurganlık Oranı (fertility) İstatistikleri: count    162.000000
mean          8.835673
std           75.682339
min           1.212000
25%           1.829250
50%           2.388500
75%           3.884500
max           966.000000
Name: fertility, dtype: float64
```

*#Ortalama okuryazarlık oranı %80 civarında, bu da dünya genelinde kadınların büyük çoğunluğunun okuryazar olduğunu gösteriyor.
#Ancak minimum değer %12.6 gibi oldukça düşük bir seviyede; bu da bazı ülkelerde okuryazarlığın hâlâ çok ciddi bir sorun olduğunu ortaya koyuyor.*

##25'lik alt çeyrek değeri %66 civarında, yani ülkelerin en az dörtte birinde kadın okuryazarlığı oldukça düşük.

Min-max kontrolü

```
min_lit = df["female literacy"].min()
```

```
max_lit = df["female literacy"].max()
```

```
min_fert = df["fertility"].min()
```

```
max_fert = df["fertility"].max()
```

```
print(f"Okuryazarlık oranı minimum: {min_lit}, maksimum: {max_lit}")
```

```
print(f"Doğurganlık oranı minimum: {min_fert}, maksimum: {max_fert}")
```

```
Okuryazarlık oranı minimum: 12.6, maksimum: 100.0
```

```
Doğurganlık oranı minimum: 1.212, maksimum: 966.0
```

#Ortalama 8.84 görünse de bu gerçek ortalamayı yansıtmıyor çünkü 966 gibi açıkça veri hatası olan bir aykırı değer ortalamayı yukarı çekmiş.

#Medyan 2.39, yani ülkelerin yarısında doğurganlık oranı 2.4 civarında. Bu değer gelişmiş ülkelerdeki ortalama doğurganlık seviyelerine yakındır.

#Standart sapma çok yüksek (75.68), bu da dağılımda ciddi sapmalar olduğunu gösteriyor.

#966 değeri, doğurganlık oranı için mantıksızdır. Gerçek dünyada bir kadının doğurganlık oranı 7-8'i bile zor bulur.

#Bu değer uç değerdir ve değiştirilmesi gerekir. Fertility oranı genelde 0-10 arası olmalıdır.

Fertility oranı genelde 0-10 arası olmalı; aykırı değeri çıkar

Hatalı olan 966 değerini ortalama ile değiştir

```
df.loc[df['fertility'] > 10, 'fertility'] = 8.835673 # Çünkü 966 gibi aykırı değerler >10 olacak şekilde fitrelenir.
```

Sonuçları kontrol et

```
print(df[df['fertility'] == 8.835673])
```

	Country	Continent	female literacy	fertility	population
149	Macao, Chine	ASI	90.7	8.835673	526,178

```
df.columns = df.columns.str.strip() # baştaki/sondaki boşlukları temizle
```

```
# En düşük doğurganlık oranına sahip 5 ülke
```

```
min_five= df.nsmallest(5, "fertility")["Country"].tolist()
```

```
# En yüksek doğurganlık oranına sahip 5 ülke
```

```
max_five= df.nlargest(5, "fertility")["Country"].tolist()
```

```
print("En düşük doğurganlığa sahip 5 ülke:", min_five)
```

```
print("En yüksek doğurganlığa sahip 5 ülke:", max_five)
```

```
En düşük doğurganlığa sahip 5 ülke: ['Bosnie-Herzégovine', 'South Korea', 'Malte', 'Japan', 'Singapour']
```

```
En yüksek doğurganlığa sahip 5 ülke: ['Macao, Chine', 'Niger', 'Afghanistan', 'Ouganda', 'Tchad']
```

```
print("Türkiye veri setinde var mı?:", "Turquie" in df["Country"].values )
```

```
Türkiye veri setinde var mı?: True
```

```
#Türkiye değerleri
```

```
df[df["Country"] == "Turquie"]
```

	Country	Continent	female literacy	fertility	population
16	Turquie	ASI	81.3	2.1	73,914,260

```
#Türkiye’de kadınların yaklaşık %81’inin okuryazar olduğunu
```

```
gösteriyor. Yani, kadın nüfusun büyük çoğunluğu okuma yazma biliyor.
```

```
#Ancak tam okuryazarlık seviyesine ulaşmamış. Bu, eğitim alanında hala geliştirilecek alanlar olduğunu düşündürebilir.
```

```
#Türkiye’de ortalama doğurganlık oranı 2.1 çocuk. Bu, nüfusun kendini yenileme seviyesi olarak kabul edilen yaklaşık 2.1’in hemen üzerinde.
```

```
#Yani, Türkiye’de nüfus dengeli bir şekilde artıyor diyebiliriz.
```

```
import matplotlib.pyplot as plt
```

```
plt.figure(figsize=(10, 6))
```

```
plt.scatter(df["female literacy"], df["fertility"], s=10, alpha=0.7)
```

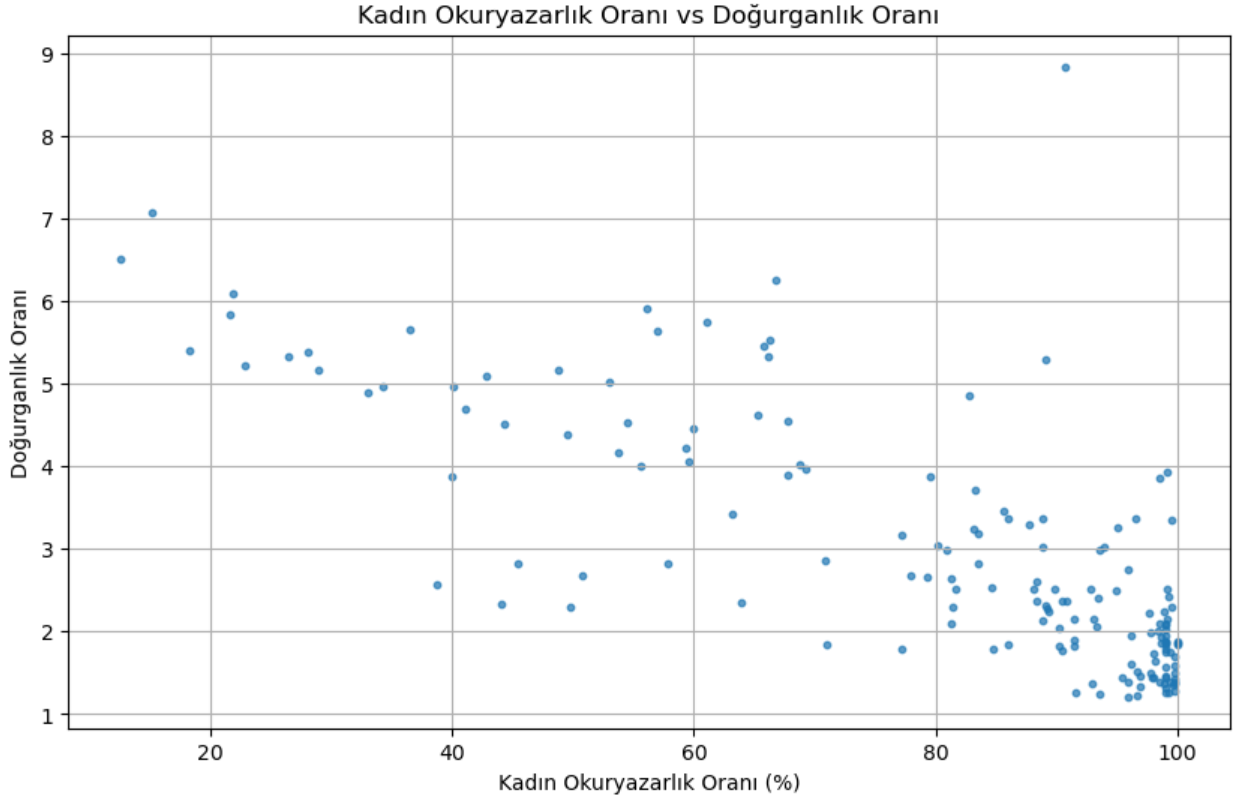
```
plt.title("Kadın Okuryazarlık Oranı vs Doğurganlık Oranı")
```

```
plt.xlabel("Kadın Okuryazarlık Oranı (%)")
```

```
plt.ylabel("Doğurganlık Oranı")
```

```
plt.grid(True)
```

```
plt.show()
```



```
# Korelasyon hesaplama
```

```
correlation = df['female literacy'].corr(df['fertility'])  
print(f"Korelasyon katsayısı: {correlation:.3f}")
```

Korelasyon katsayısı: -0.753

*#Korelasyon katsayısı negatif olduğundan kadın okuryazarlık ile doğurganlık oranı arasında ters bir ilişki vardır.Yani
#biri artarken diğeri azalır ve mutlak değeri 0.7 den büyük buda oldukça güçlü bir bağ demektir.*

```
import matplotlib.pyplot as plt
```

```
# Kıtalar ve onlara atanacak renkler
```

```
continents = df['Continent'].unique()  
colors = ['red', 'blue', 'green', 'orange', 'purple', 'brown', 'cyan']  
color_map = dict(zip(continents, colors))
```

```
# Grafik boyutu
```

```
plt.figure(figsize=(12, 8))
```

```
# Her kıta için ayrı scatter noktaları
```

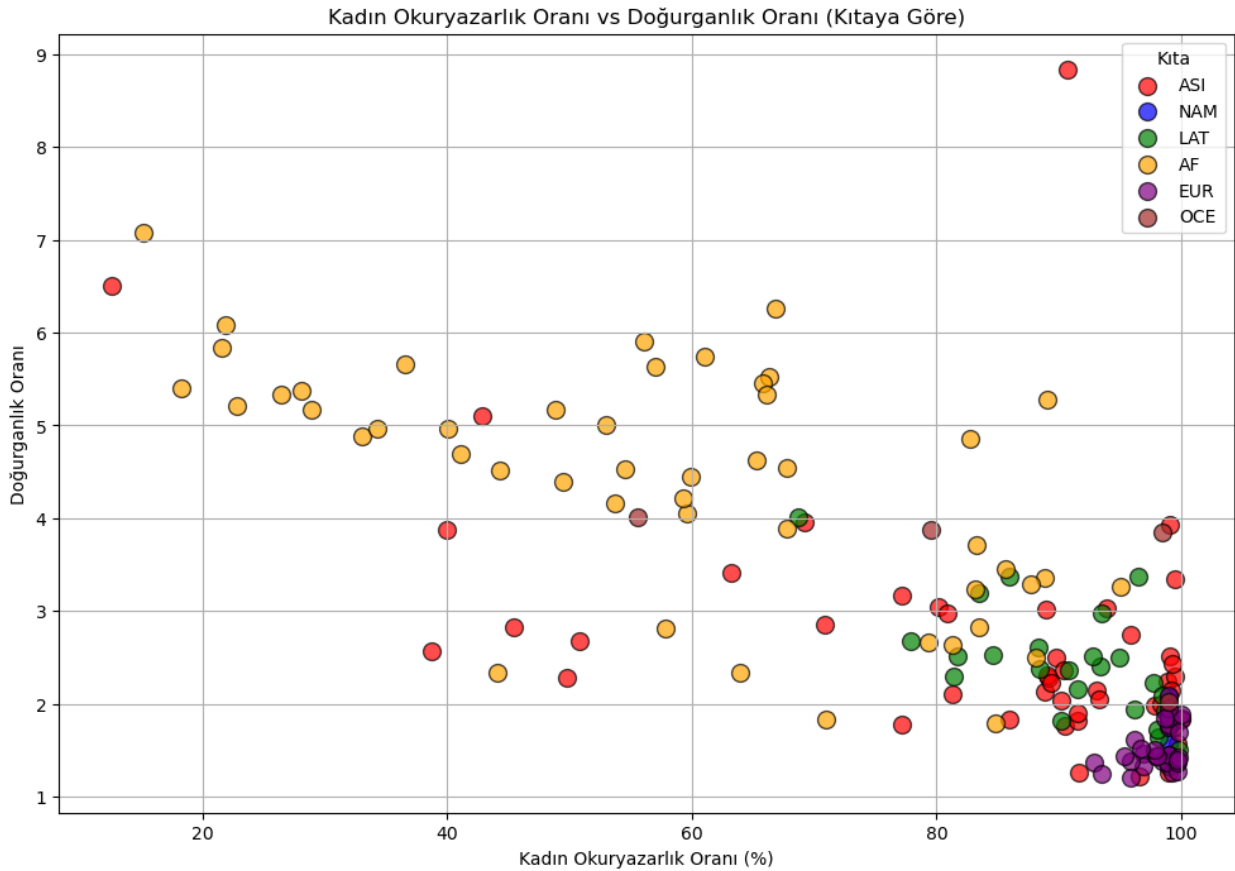
```
for continent in continents:  
    subset = df[df['Continent'] == continent]  
    plt.scatter(  
        subset['female literacy'],
```

```

subset['fertility'],
label=continent,
color=color_map[continent],
s=100,
alpha=0.7,
edgecolor='black'
)

plt.title('Kadın Okuryazarlık Oranı vs Doğurganlık Oranı (Kıtaya Göre)')
plt.xlabel('Kadın Okuryazarlık Oranı (%)')
plt.ylabel('Doğurganlık Oranı')
plt.legend(title='Kıta')
plt.grid(True)
plt.show()

```



#Afrika kıtasının okuryazarlık oranının(%20-%60) genelde düşük olduğunu söyleyebiliriz.
 #Ayrıca doğurganlık oranının yüksek olduğunda belirtebiliriz.(4-7)
 #Avrupa ve Kuzey Amerika kıtasınında kadın okuryazarlık oranının oldukça yüksek olduğunu görüyoruz.(%90-%100)
 #Doğurganlık oranı düşük (1-2 çocuk arası).Gelişmiş ülkelerde kadın

eğitimi arttıkça doğurganlık oranı genelde azalıyor.
#Çok geniş dağılım var. Bazı ülkeler (örneğin Japonya, Güney Kore)
yüksek okuryazarlık ve düşük doğurganlık gösterirken
#diğerleri (örneğin Pakistan, Afganistan) düşük okuryazarlık ve yüksek
doğurganlık oranına sahip.
#Yani Asya'da sosyo-ekonomik gelişmişliğe göre büyük farklılıklar var.

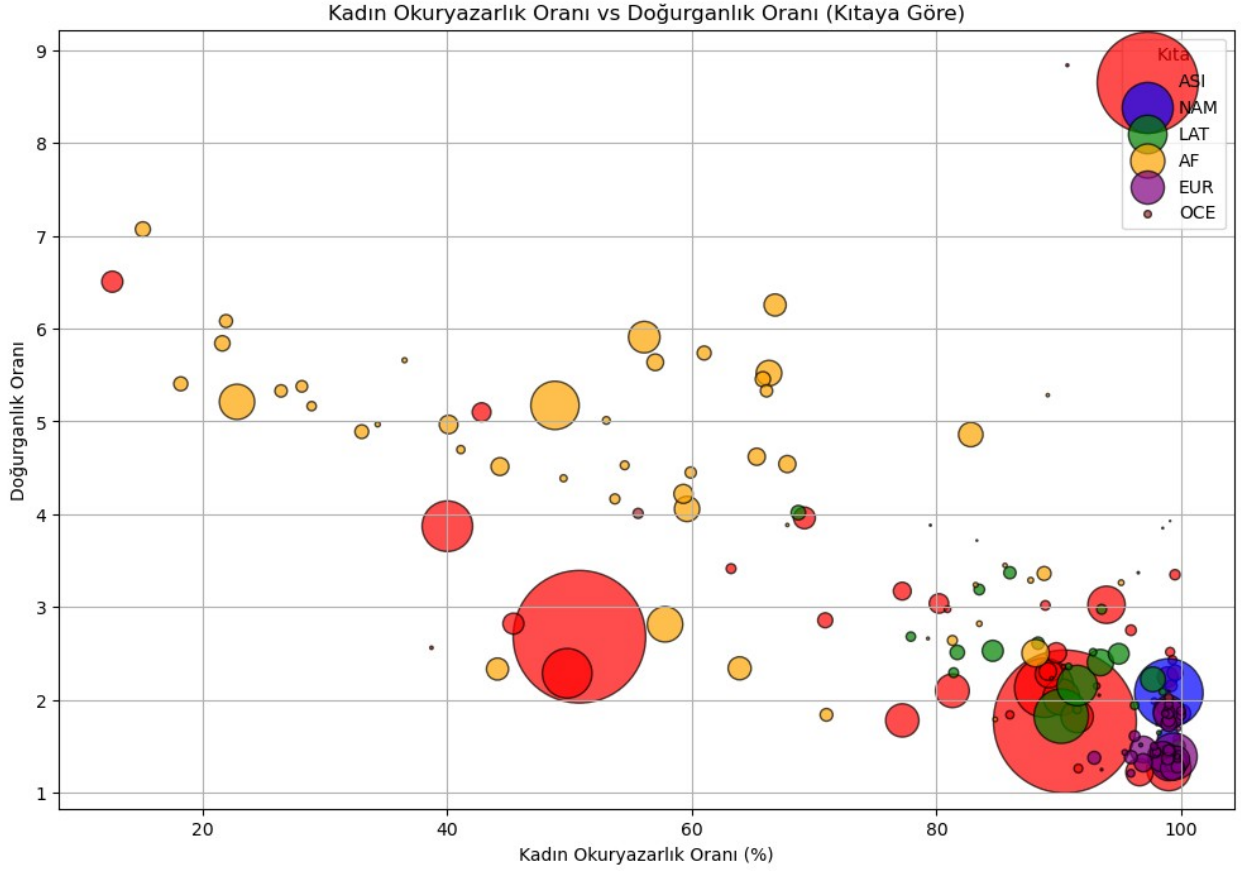
```
import matplotlib.pyplot as plt
import pandas as pd
# Virgül kaldırma ve sayıya çevirme
df['population'] = df['population'].str.replace(',', '').astype(float)

# Kıtalar ve onlara atanacak renkler
continents = df['Continent'].unique()
colors = ['red', 'blue', 'green', 'orange', 'purple', 'brown', 'cyan']
color_map = dict(zip(continents, colors))

# Grafik boyutu
plt.figure(figsize=(12, 8))

# Her kıta için ayrı scatter noktaları
for continent in continents:
    subset = df[df['Continent'] == continent]
    plt.scatter(
        subset['female literacy'],
        subset['fertility'],
        label=continent,
        color=color_map[continent],
        s=subset['population'] / 1e6 * 5, # Nokta büyüklüğü nüfusa
        alpha=0.7,
        edgecolor='black'
    )

plt.title('Kadın Okuryazarlık Oranı vs Doğurganlık Oranı (Kıtaya Göre)')
plt.xlabel('Kadın Okuryazarlık Oranı (%)')
plt.ylabel('Doğurganlık Oranı')
plt.legend(title='Kıta')
plt.grid(True)
plt.show()
```

#Eğitim seviyesi ve doğurganlık arasında negatif bir ilişki bulunmaktadır.
 #Büyük noktalar, nüfusun yüksek olduğu ülkeleri gösteriyor.
 #Bu ülkelerdeki kadın okuryazarlık ve doğurganlık oranları, kıta ortalamasını daha fazla etkileyebilir.

```
fl_stats = df.groupby('Continent')['female literacy'].agg(['mean', 'min', 'max', 'median'])
fert_stats = df.groupby('Continent')['fertility'].agg(['mean', 'min', 'max', 'median'])

stats = pd.concat([fl_stats, fert_stats], axis=1)
print(stats)
```

	mean	min	max	median	mean	min	max
Continent							
AF	57.959184	15.1	95.1	59.60	4.413878	1.792	7.069000
ASI	82.968085	12.6	100.0	90.40	2.680950	1.224	8.835673

2.2940							
EUR	98.225000	92.9	100.0	99.00	1.548417	1.212	2.093000
1.4460							
LAT	90.466667	68.7	99.8	92.15	2.448375	1.505	4.018000
2.3885							
NAM	99.000000	99.0	99.0	99.00	1.828000	1.579	2.077000
1.8280							
OCE	83.150000	55.6	99.0	89.00	3.442500	2.025	4.010000
3.8675							

*#Kadın okuryazarlık oranı en düşük Afrika'da (%58 civarında), en yüksek Kuzey Amerika ve Avrupa (%98-99 civarında).
 #Doğurganlık oranı en yüksek Afrika'da (~4.4 çocuk), en düşük Avrupa ve Kuzey Amerika (~1.5-1.8 çocuk).
 #Genel olarak kadın okuryazarlık oranı yüksek olan kıtalarda doğurganlık oranı düşük, okuryazarlık oranı düşük olan kıtalarda doğurganlık oranı yüksek.
 #Asya, Latin Amerika ve Okyanusya orta seviyede kadın okuryazarlığı ve doğurganlık oranları gösteriyor.
 #En yüksek doğurganlık Afrika'da, en düşük Avrupa ve Kuzey Amerika'da.
 #Kadın eğitimi ile doğurganlık arasında negatif bir ilişki olduğunu destekliyor: Kadınların eğitimi arttıkça doğurganlık oranı düşme eğiliminde.*

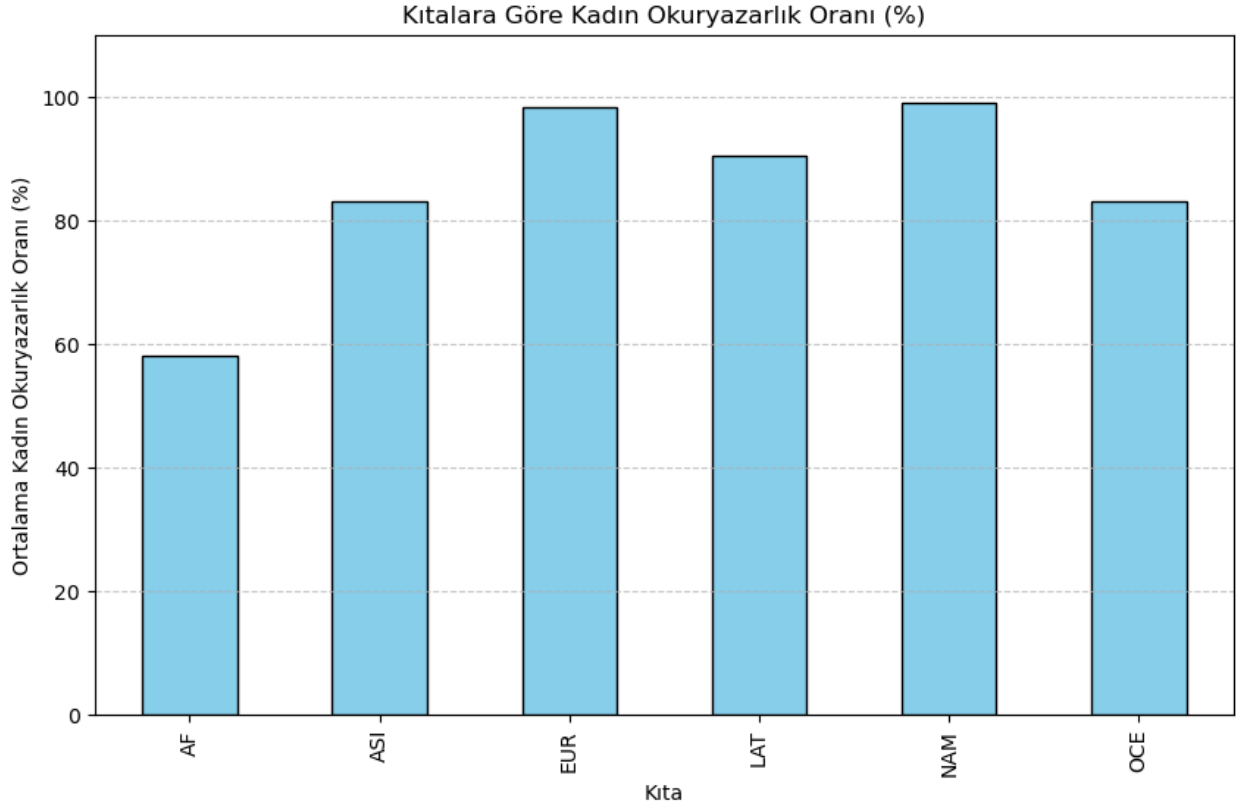
```
import matplotlib.pyplot as plt

# Kıtalarla göre kadın okuryazarlık oranlarının ortalaması
fl_mean = df.groupby('Continent')['female literacy'].mean()

# Çubuk grafik çizimi
plt.figure(figsize=(10, 6))
fl_mean.plot(kind='bar', color='skyblue', edgecolor='black')

plt.title('Kıtalarla Göre Kadın Okuryazarlık Oranı (%)')
plt.xlabel('Kıta')
plt.ylabel('Ortalama Kadın Okuryazarlık Oranı (%)')
plt.ylim(0, 110) # Y eksenini %0-110 aralığına ayarladım
plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.show()
```



#Grafik sonucuna göre Afrika kıtası, diğer kıtalara göre en düşük kadın okuryazarlık oranına sahip.
#Avrupa ve Kuzey Amerika kıtaları kadın okuryazarlık oranının en yüksek olduğu bölgeler.
#Asya, Latin Amerika ve Okyanusya kıtaları ise orta seviyede kadın okuryazarlık oranları sergiliyor.

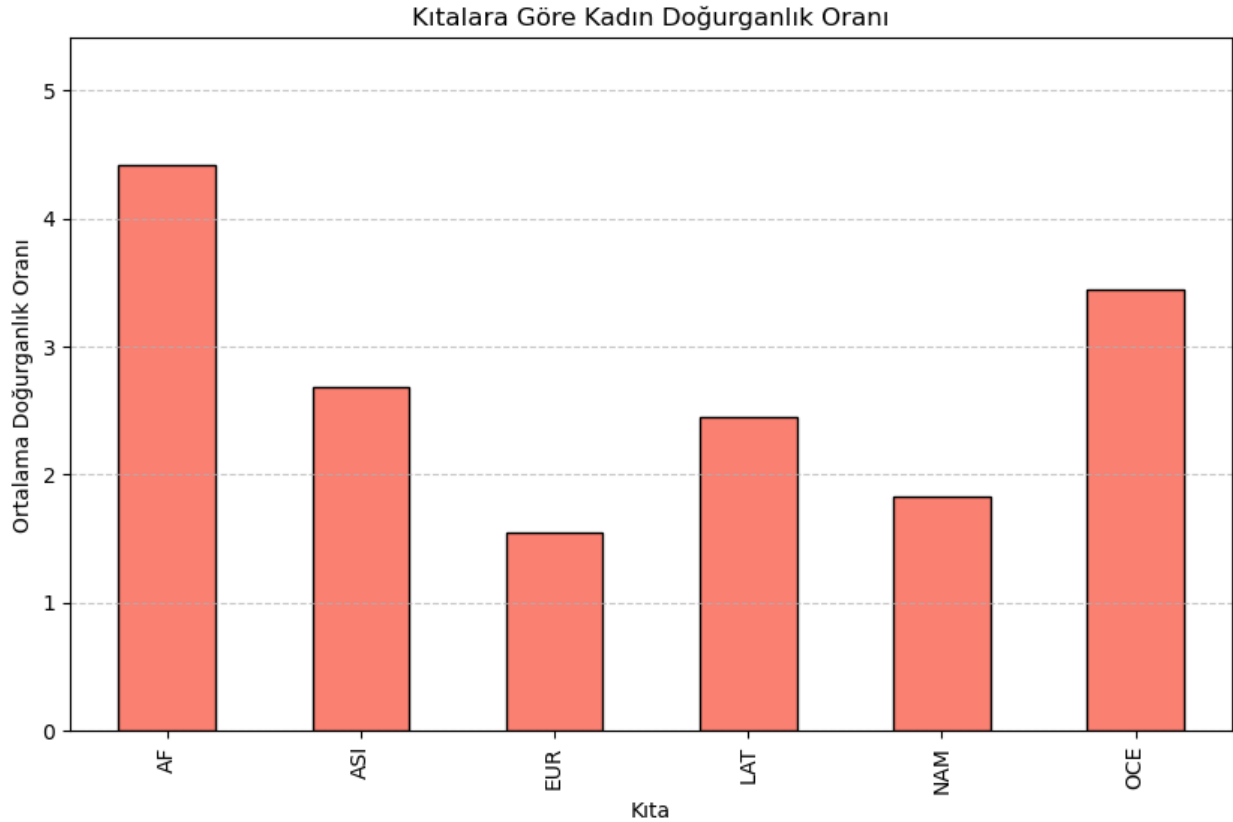
```
import matplotlib.pyplot as plt

# Kıtalara göre doğurganlık oranının ortalaması
fert_mean = df.groupby('Continent')['fertility'].mean()

# Çubuk grafik çizimi
plt.figure(figsize=(10, 6))
fert_mean.plot(kind='bar', color='salmon', edgecolor='black')

plt.title('Kıtalara Göre Kadın Doğurganlık Oranı')
plt.xlabel('Kıta')
plt.ylabel('Ortalama Doğurganlık Oranı')
plt.ylim(0, fert_mean.max() + 1) # Y eksenini üst sınırı biraz boşluk bırakacak şekilde ayarladım.
plt.grid(axis='y', linestyle='--', alpha=0.7)

plt.show()
```



#Grafikte Afrika kıtası en yüksek kadın doğurganlık oranına sahip; yani kadın başına düşen ortalama doğum sayısı burada diğer kıtalara göre belirgin şekilde daha fazla.

#Avrupa ve Kuzey Amerika ise en düşük doğurganlık oranlarını gösteriyor, kadın başına düşen doğum sayısı 2'nin altında.

#Asya, Latin Amerika ve Okyanusya kıtaları ise orta seviyede doğurganlık oranlarına sahip.

```
import matplotlib.pyplot as plt
```

```
plt.figure(figsize=(12, 8))
```

```
df.boxplot(column='fertility', by='Continent', grid=False)
```

```
plt.title('Kıtalara Göre Doğurganlık Oranı Dağılımı ve Uç Değerler')
```

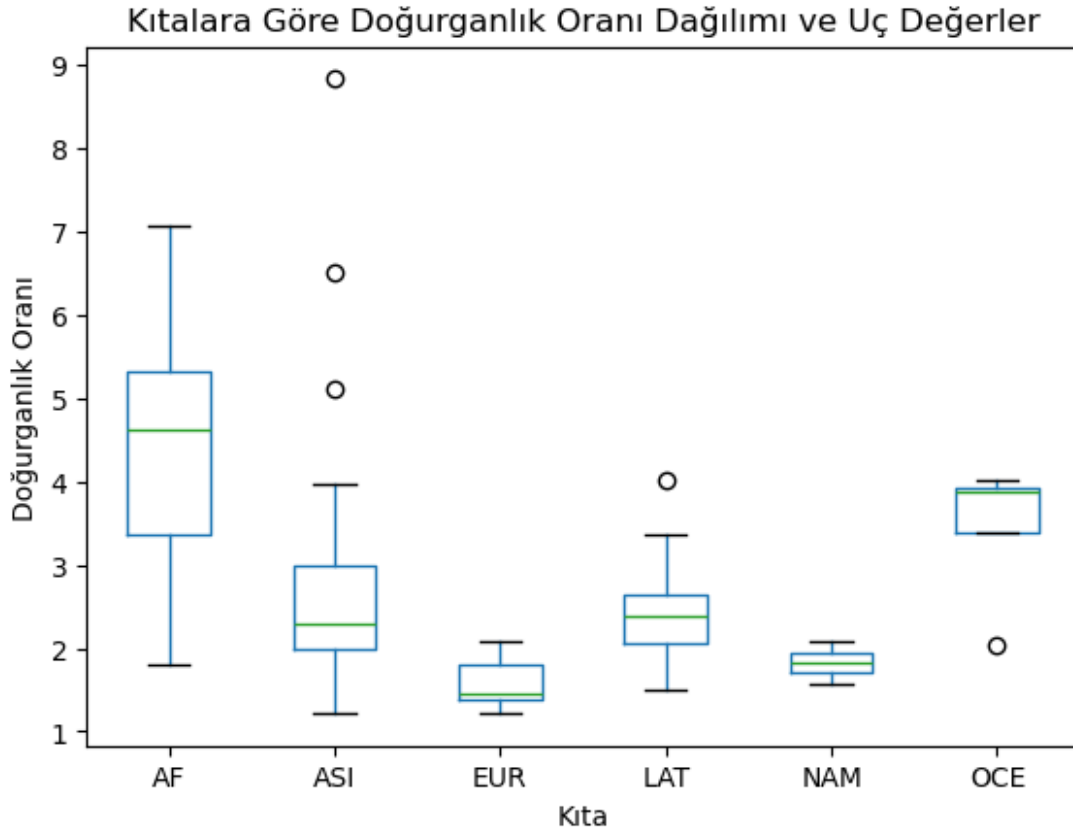
```
plt.suptitle('') # otomatik başlığı kaldırmak için
```

```
plt.xlabel('Kıta')
```

```
plt.ylabel('Doğurganlık Oranı')
```

```
plt.show()
```

<Figure size 1200x800 with 0 Axes>



```
def find_outliers(data):
    Q1 = data.quantile(0.25)
    Q3 = data.quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5 * IQR
    upper_bound = Q3 + 1.5 * IQR
    outliers = data[(data < lower_bound) | (data > upper_bound)]
    return outliers

# Kıtalara göre uç değerlerin bulunması
outliers_by_continent = {}

for continent in df['Continent'].unique():
    fertility_data = df[df['Continent'] == continent]['fertility']
    outliers = find_outliers(fertility_data)
    outliers_by_continent[continent] = outliers

# Sonuçların yazdırılması
for continent, outliers in outliers_by_continent.items():
    print(f"{continent} kıtasında bulunan uç değerler ({len(outliers)} adet):")
    print(outliers.values)
    print('----')
```

```
ASI kıtasında bulunan uç değerler (3 adet):  
[6.505      5.1      8.835673]
```

```
---
```

```
NAM kıtasında bulunan uç değerler (0 adet):  
[]
```

```
---
```

```
LAT kıtasında bulunan uç değerler (1 adet):  
[4.018]
```

```
---
```

```
AF kıtasında bulunan uç değerler (0 adet):  
[]
```

```
---
```

```
EUR kıtasında bulunan uç değerler (0 adet):  
[]
```

```
---
```

```
OCE kıtasında bulunan uç değerler (1 adet):  
[2.025]
```

```
---
```

*#İki yöntemede baktığımızda uç değerler sayıları aynı çıkmıştır.
#Afrika (AF) ve Asya (ASI) kıtalarında doğurganlık oranı dağılımı oldukça geniş ve uç değer sayısı fazladır.
#Avrupa (EUR) ve Kuzey Amerika (NAM) gibi kıtalarda ise uç değerler daha az, dağılım daha sıkışık ve ortalama daha düşüktür.
#Bu grafik hızlıca hangi kıtalarda uç değerlerin daha yaygın olduğunu anlamamıza olanak sağlar
#Afrika için 5 uç değer tespit edilmişse, bu doğurganlık oranı çok yüksek veya çok düşük olan ülkelerin değerleridir.*

```
import pandas as pd
```

```
# Nüfusu sayıya çevir
```

```
df['population'] = df['population'].replace(',', '',  
regex=True).astype(float)
```

```
# Yeni 'boyut' sütununu ekle
```

```
df['boyut'] = df['population'].apply(lambda x: 'büyük ülke' if x >  
80000000 else 'küçük ülke')
```

```
df
```

	Country	Continent	female literacy	fertility
population \				
0	Chine	ASI	90.5	1.769
1.324655e+09				
1	Inde	ASI	50.8	2.682
1.139965e+09				
2	USA	NAM	99.0	2.077

3.040600e+08				
3	Indonésie	ASI	88.8	2.132
2.273451e+08				
4	Brésil	LAT	90.2	1.827
1.919715e+08				
..
...				
157	Vanuatu	OCE	79.5	3.883
2.338660e+05				
158	Samoa	OCE	98.5	3.852
1.788690e+05				
159	Sao Tomé-et-Principe	AF	83.3	3.718
1.601740e+05				
160	Aruba	LAT	98.0	1.732
1.054550e+05				
161	Tonga	ASI	99.1	3.928
1.035660e+05				

	boyut
0	büyük ülke
1	büyük ülke
2	büyük ülke
3	büyük ülke
4	büyük ülke
..	...
157	küçük ülke
158	küçük ülke
159	küçük ülke
160	küçük ülke
161	küçük ülke

[162 rows x 6 columns]

```
df['boyut'].value_counts()
```

```
boyut
küçük ülke    146
büyük ülke     16
Name: count, dtype: int64
```

```
df.groupby('boyut')['fertility'].agg(['min', 'max', 'mean'])
```

	min	max	mean
boyut			
büyük ülke	1.262	5.211000	2.565062
küçük ülke	1.212	8.835673	2.966943

#Ortalama için;

*#Küçük ülkelerin ortalama doğurganlık oranı 2.97, büyük
ülkelerinkinden daha yüksek (2.56).*

#Bu durum, küçük ülkelerin genellikle gelişmekte olan ülkeler

olmasıyla ilişkilendirilebilir.

#Bu ülkelerde sağlık hizmetleri sınırlı olabilir, doğum kontrolüne erişim düşüktür ve nüfus artış hızları yüksektir.

#Büyük ülkeler, genelde daha sanayileşmiş ve şehirleşmiş yapıya sahiptir. Bu da daha düşük doğurganlık oranlarına neden olabilir.

#Min değer için;

#Minimum değerler birbirine çok yakın.Bu, düşük doğurganlık oranlarının her iki grupta da gözlemlenebileceğini gösteriyor.

#Özellikle küçük ama gelişmiş ülkeler bu düşük değerlere sahip olabilir.

#Max değer için;

#Küçük ülkelerdeki maksimum değer 8.83 gibi oldukça yüksek. Bu, bazı küçük ülkelerde hâlâ çok yüksek doğurganlık oranlarının bulunduğunu gösteriyor.

#Büyük ülkelerde maksimum değer 5.21 ile sınırlı kalmış. Yani büyük ülkeler grubunda çok yüksek doğurganlık değerleri pek görülmemekte.

#Küçük ülkeler, doğurganlık oranı açısından daha geniş bir dağılıma sahip. Hem çok düşük hem de çok yüksek doğum oranlarını içerebiliyorlar.

#Büyük ülkelerin doğurganlık oranları ise genellikle daha ortalama seviyelerde yoğunlaşmış ve uç değerlerden uzak.

```
import pandas as pd
df = pd.read_csv("df_year.csv")
df.head()
```

	Country Name	Country Code	\
0	Aruba	ABW	
1	Africa Eastern and Southern	AFE	
2	Afghanistan	AFG	
3	Africa Western and Central	AFW	
4	Angola	AGO	

	Indicator Name	Indicator Code	1960
\			
0	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	4.820000
1	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	6.724125
2	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	7.282000
3	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	6.458448
4	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	6.708000

	1961	1962	1963	1964	1965	...	2014
2015 \							
0	4.655000	4.471000	4.271000	4.059000	3.842000	...	2.148000
1	6.742752	6.762930	6.778712	6.788420	6.800322	...	4.739863
2	7.284000	7.292000	7.302000	7.304000	7.305000	...	5.560000
3	6.471518	6.491826	6.506088	6.525355	6.541102	...	5.437493
4	6.790000	6.872000	6.954000	7.036000	7.116000	...	5.864000

	2016	2017	2018	2019	2020	2021	2022 \
0	1.953000	1.839000	1.587000	1.486000	1.325000	1.180000	NaN
1	4.615671	4.570410	4.527707	4.482899	4.416902	4.354710	NaN
2	5.262000	5.129000	5.002000	4.870000	4.750000	4.643000	NaN
3	5.328709	5.255345	5.186319	5.118932	5.049329	4.978662	NaN
4	5.686000	5.600000	5.519000	5.442000	5.371000	5.304000	NaN

```

Unnamed: 67
0      NaN
1      NaN
2      NaN
3      NaN
4      NaN

```

[5 rows x 68 columns]

```
df.head(5)
```

	Country Name	Country Code \
0	Aruba	ABW
1	Africa Eastern and Southern	AFE
2	Afghanistan	AFG
3	Africa Western and Central	AFW
4	Angola	AGO

	Indicator Name	Indicator Code	1960
0	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	4.820000
1	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	6.724125
2	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	7.282000
3	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	6.458448
4	Fertility rate, total (births per woman)	SP.DYN.TFRT.IN	6.708000

	1961	1962	1963	1964	1965	...	2014
2015 \							
0	4.655000	4.471000	4.271000	4.059000	3.842000	...	2.148000
1	6.742752	6.762930	6.778712	6.788420	6.800322	...	4.739863
2	7.284000	7.292000	7.302000	7.304000	7.305000	...	5.560000
3	6.471518	6.491826	6.506088	6.525355	6.541102	...	5.437493
4	6.790000	6.872000	6.954000	7.036000	7.116000	...	5.864000

	2016	2017	2018	2019	2020	2021	2022 \
0	1.953000	1.839000	1.587000	1.486000	1.325000	1.180000	NaN
1	4.615671	4.570410	4.527707	4.482899	4.416902	4.354710	NaN
2	5.262000	5.129000	5.002000	4.870000	4.750000	4.643000	NaN
3	5.328709	5.255345	5.186319	5.118932	5.049329	4.978662	NaN
4	5.686000	5.600000	5.519000	5.442000	5.371000	5.304000	NaN

	Unnamed: 67
0	NaN
1	NaN
2	NaN
3	NaN
4	NaN

[5 rows x 68 columns]

```
import pandas as pd
import matplotlib.pyplot as plt

# df zaten yüklü varsayalım

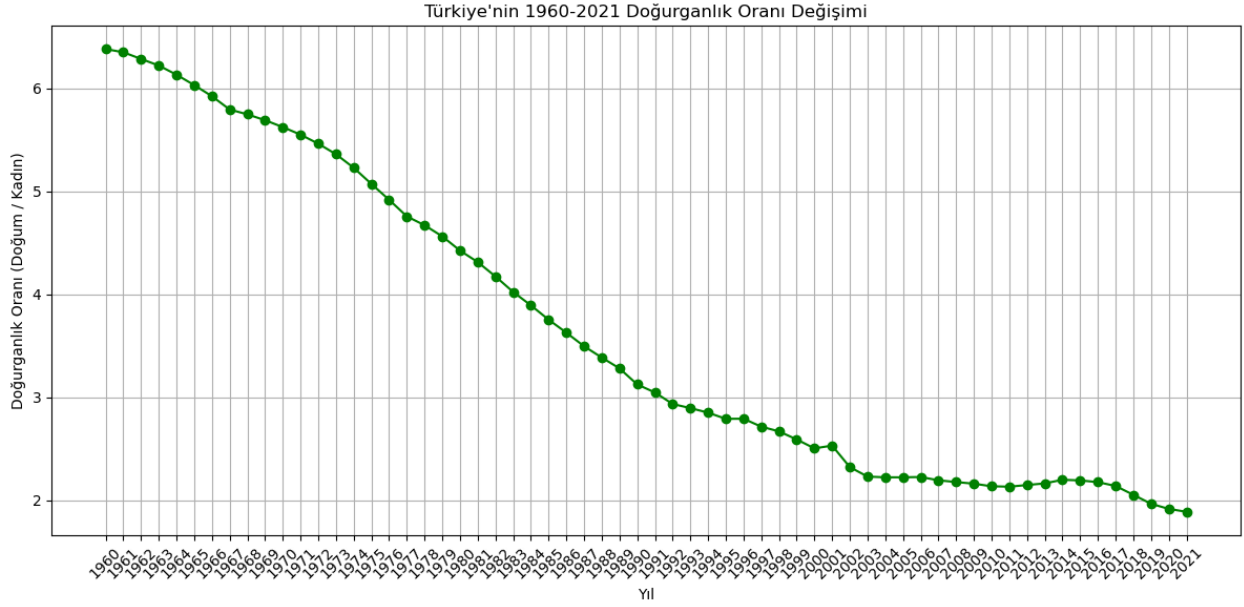
# "Türkiye" satırını seç
turkiye = df[df['Country Name'] == 'Türkiye']

# 1960-2021 yılları sütun isimleri
years = list(map(str, range(1960, 2022)))

# Doğurganlık değerlerini al
fertility_values = turkiye[years].values.flatten()

# Grafik çizimi
plt.figure(figsize=(12, 6))
plt.plot(years, fertility_values, marker='o', linestyle='-',
color='green')
plt.xticks(rotation=45)
plt.title("Türkiye'nin 1960-2021 Doğurganlık Oranı Değişimi")
plt.xlabel("Yıl")
```

```
plt.ylabel("Doğurganlık Oranı (Doğum / Kadın)")  
plt.grid(True)  
plt.tight_layout()  
plt.show()
```



#Grafiğe göre 2021 yılına doğru ilerledikçe doğurganlık oranının azaldığını göre biliriz. Bir önceki veri çıktılarınıda
#göz önüne alırsak kadın okuryazarlığında yıllara göre arttığını söyleyebiliriz.