PYTHON BLOK PROJESI

Nobel Ödülü belki de dünyanın en tanınmış bilimsel ödülüdür. Her yıl kimya, edebiyat, fizik, tıp, ekonomi ve barış alanlarında bilim insanlarına ve akademisyenlere verilmektedir. Bu projede, Nobel Ödülü kazananları inceleyeceğiz.



Proje Sorulari

- 1. Nobel Ödüllerini en çok kazanan ilk on ülkeyi bulunuz.
- 2. Nobel Ödüllerini kazanan ilk kadınları listeleyiniz.
- 3. Nobel Ödüllerini kazanan ilk erkekleri listeleyiniz.
- 4. Nobel ödülünü en çok kazanan ülkenin hangi yıldan itibaren hakimiyet sağladığını görselleştirip bu hakimiyette rol oynayan Şeyler nelerdir? İçgörülerinizi paylaşır mısınız?
- 5. Nobel Ödülü kazananların cinsiyetlerini, yaşlarını, ödül kategorisi ve yılları kullanarak görselleştiriniz.(Her bir ödül kategorisi için ayrı grafik gösteriniz) Çıkan sonuçlara göre görseli yorumlayınız.
- 6. 1938-1945 yılı arasında Nobel Ödülü kazananların kategorilerini ve ülkelerini görselleştirip yorumlayınız.
- 7. 1947–1991 yılları arasında Nobel Ödülü kazananların kategorilerini ve ülkelerini görselleŞtirip yorumlayınız.(Her kategori için ayrı bir grafik olması istenmektedir)
- 8. Kimya, Edebiyat, Barış, Fizik ve Tıp kategorilerindeki 2000 sonrasındaki kişilerin ülkelerini, yaşlarını görselleştirin.(Her bir Kategori için ayrı görselleştirme yapılması istenmektedir) Veriyi yorumlayınız.

Proje Soruları Öncesi .csv uzantılı Veriye Erişim

```
import pandas as pd
import numpy as np

df=pd.read_csv("nobel.csv")

df.head(10)
```

ye	ar cat	tegory	prize	motivation	prize_share	laureate_id	laureate_type	full_name	birth_date	birth_city	birth_country	sex	organization_name	organization_city	organization_country	death_date	death_city	death_country
0 190	1 Che	emistry	The Nobel Prize in Chemistry 1901	"in recognition of the extraordinary services	1/1	160	Individual	Jacobus Henricus van 't Hoff	1852-08- 30	Rotterdam	Netherlands	Male	Berlin University	Berlin	Germany	1911-03-01	Berlin	Germany
1 190	1 Lite	erature	The Nobel Prize in Literature 1901	"in special recognition of his poetic composit	1/1	569	Individual	Sully Prudhomme	1839-03- 16	Paris	France	Male	NaN	NaN	NaN	1907-09-07	Châtenay	France
2 190	1 Me	edicine	The Nobel Prize in Physiology or Medicine 1901	"for his work on serum therapy, especially its	1/1	293	Individual	Emil Adolf von Behring	1854-03- 15	Hansdorf (Lawice)	Prussia (Poland)	Male	Marburg University	Marburg	Germany	1917-03-31	Marburg	Germany
3 190	1	Peace	The Nobel Peace Prize 1901	NaN	1/2	462	Individual	Jean Henry Dunant	1828-05- 08	Geneva	Switzerland	Male	NaN	NaN	NaN	1910-10-30	Heiden	Switzerland
4 190	1	Peace	The Nobel Peace Prize 1901	NaN	1/2	463	Individual	Frédéric Passy	1822-05- 20	Paris	France	Male	NaN	NaN	NaN	1912-06-12	Paris	France
5 190	1 6	Physics	The Nobel Prize in Physics 1901	"in recognition of the extraordinary services	1/1	1	Individual	Wilhelm Conrad Röntgen	1845-03- 27	Lennep (Remscheid)	Prussia (Germany)	Male	Munich University	Munich	Germany	1923-02-10	Munich	Germany
6 190	2 Che	emistry	The Nobel Prize in Chemistry 1902	"in recognition of the extraordinary services	1/1	161	Individual	Hermann Emil Fischer	1852-10-	Euskirchen	Prussia (Germany)	Male	Berlin University	Berlin	Germany	1919-07-15	Berlin	Germany
7 190	2 Lite	erature	The Nobel Prize in Literature 1902	"the greatest living master of the art of hist	1/1	571	Individual	Christian Matthias Theodor Mommsen	1817-11- 30	Garding	Schleswig (Germany)	Male	NaN	NaN	NaN	1903-11-01	Charlottenburg	Germany
8 190	12 Me	edicine	The Nobel Prize in Physiology or Medicine 1902	"for his work on malaria, by which he has show	1/1	294	Individual	Ronald Ross	1857-05- 13	Almora	India	Male	University College	Liverpool	United Kingdom	1932-09-16	Putney Heath	United Kingdom
			The Nahel Dance Drive 1003	NeN	1/2	161	la divisional	Élia Durannous	1833-02-	Ganna	Curiosadand	Male	NeN	NaN	NeW	1004 12 07	0	Curionariana

Verileri incelemeden önce pandas ve numpy kütüphanelerini import ederek, "nobel.csv" diye isimlendirdiğim data setini read_csv metodu ile okuyup dataframe isimlendirmesini "df" olarak kullandım ve verileri görmek adına ilk 10 satırı döndürdüm.

Veri içeriğinde bulunan sütun adları şunlardır;

```
year,
category,
prize,
motivation,
prize_share,
laureate_id,
laureate_type,
full_name,
birth_date,
birth_city,
birth_country,
sex,
organization_name,
organization_city,
organization_country,
death_date,
death_city,
death_country
```

1. Nobel Ödüllerini en çok kazanan ilk on ülkeyi bulunuz.

```
top_ten_winning_countries=df['birth_country'].value_counts().head(10)
print(top_ten_winning_countries)
United States of America
United Kingdom
                              85
Germany
                              61
France
                              51
Sweden
                              29
Japan
                              24
Canada
                              18
Netherlands
                              18
                              17
Italy
Russia
Name: birth_country, dtype: int64
```

Nobel ödüllerini en çok kazanan ülkeleri,yukarıda import ettiğim pandas kütüphanesinden value counts() fonksiyonunu kullanarak listeledim.

Listenin en başında **Amerika** açık ara farkla en çok ödül alan ülke.Onları takip eden diğer ülkeler **Birleşik Krallık** (İngiltere, İskoçya, Galler, Kuzey İrlanda) ve **Almanya**'dır.

2. Nobel Ödüllerini kazanan ilk kadınları listeleyiniz.

```
import pandas as pd
pd.set option('display.max columns', None)
pd.set_option('display.width', 1000)
female winners = df[df["sex"] == "Female"]
female winners = female_winners.dropna(subset=["birth country"])
female_winners = female_winners.sort_values("year")
first_female_winners = female_winners.drop_duplicates(subset=["category"], keep="first")
print(first_female_winners[["full_name", "birth_country", "category","year"]])
                              full_name birth_country category year
Marie Curie, née Sklodowska Russian Empire (Poland) Physics 1903
Austrian Empire (Czech Republic) Peace 1905
19
29
     Baroness Bertha Sophie Felicita von Suttner, n... Austrian Empire (Czech Republic)
                           Selma Ottilia Lovisa Lagerlöf
Marie Curie, née Sklodowska
51
                                                                                              Sweden Literature 1909
                                                                          Russian Empire (Poland) Chemistry 1911
62
                          Gerty Theresa Cori, née Radnitz Austria-Hungary (Czech Republic) Medicine 1947
Elinor Ostrom United States of America Economics 2009
235
```

Nobel ödüllerini kazanan ilk kadınları kategori, yıllar ve ülkelerine göre listeledim.İlk önce sorguda istediğim verileri yanyana yazdırması için set_option() metodunu kullandım.İstenilen liste için öncelikle sadece "Female" olan kadınları listeledim."birth_country" de varsa nan değerlerini ortadan kaldırmak için dropna() metodunu kullandım.Yıllara göre sıraladıktan sonra,"drop_duplicates" metodu ile tekrarlanan kategorileri silip kategorilerde bulunan ilk kadın isimlerini,ülkelerini,kategorilerini ve yıllarını listeledim.

3. Nobel Ödüllerini kazanan ilk erkekleri listeleyiniz.

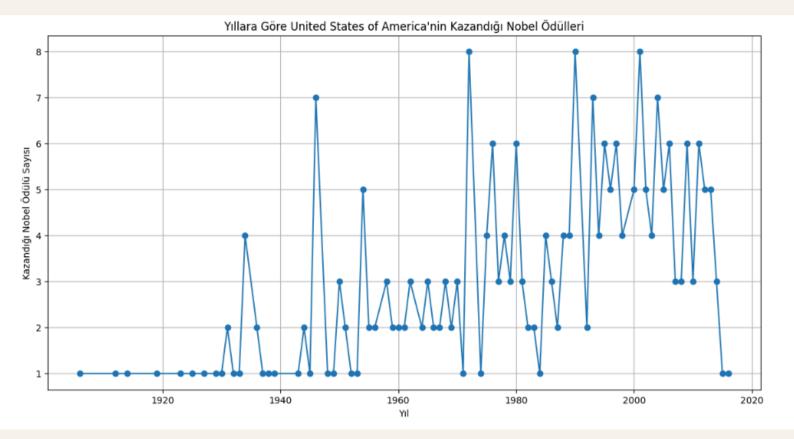
```
import pandas as pd
pd.set option('display.max columns', None)
pd.set_option('display.width', 1000)
male_winners = df[df["sex"] == "Male"]
male_winners = male_winners.dropna(subset=["birth_country"])
male winners = male winners.sort values("year")
first_male_winners = male_winners.drop_duplicates(subset=["category"], keep="first")
print(first_male_winners[["full_name", "birth_country", "category","year"]])
                       full name
                                      birth_country
                                                      category year
0
    Jacobus Henricus van 't Hoff
                                       Netherlands
                                                     Chemistry 1901
1
                 Sully Prudhomme
                                            France Literature 1901
2
          Emil Adolf von Behring Prussia (Poland) Medicine 1901
3
               Jean Henry Dunant
                                        Switzerland
                                                         Peace 1901
                                                      Physics 1901
          Wilhelm Conrad Röntgen Prussia (Germany)
5
393
                   Ragnar Frisch
                                            Norway
                                                     Economics 1969
```

Nobel ödüllerini kazanan ilk erkekleri kategori, yıllar ve ülkelerine göre listeledim.İlk önce sorguda istediğim verileri yanyana yazdırması için set_option() metodunu kullandım. İstenilen liste için öncelikle sadece "Male" olan erkekleri listeledim."birth_country" de varsa nan değerlerini ortadan kaldırmak için dropna() metodunu kullandım.Yıllara göre sıraladıktan sonra,"drop_duplicates" metodu ile tekrarlanan kategorileri silip kategorilerde bulunan ilk erkek isimlerini,ülkelerini,kategorilerini ve yıllarını listeledim.

4. Nobel ödülünü en çok kazanan ülkenin hangi yıldan itibaren hakimiyet sağladığını görselleştirip bu hakimiyette rol oynayan şeyler nelerdir? İçgörülerinizi paylaşır mısınız?

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

birth_country = "United States of America"
country_winners = df[df['birth_country'] == birth_country]
country_winners_count = country_winners.groupby('year').size().reset_index(name='count')
plt.figure(figsize=(14, 7))
plt.plot(country_winners_count['year'], country_winners_count['count'], marker='o', linestyle='-')
plt.xlabel('Yıl')
plt.ylabel('Kazandığı Nobel Ödülü Sayısı')
plt.title(f'Yıllara Göre {birth_country}\'nin Kazandığı Nobel Ödülleri')
plt.grid(True)
plt.show()
```



Öncelikle görselleştirme için gerekli olan matplotlib kütüphanesini import ettim.Daha önce cevabını bulduğumuz en çok nobel kazanan ülke Amerika olduğu için "birth_country" direk ülke adını yazıp filtreleyerek sorguya başladım.Yıllara göre u ülkenin kazandığı nobel ödül sayısını hesaplayıp bu sayıları görselleştirdim.

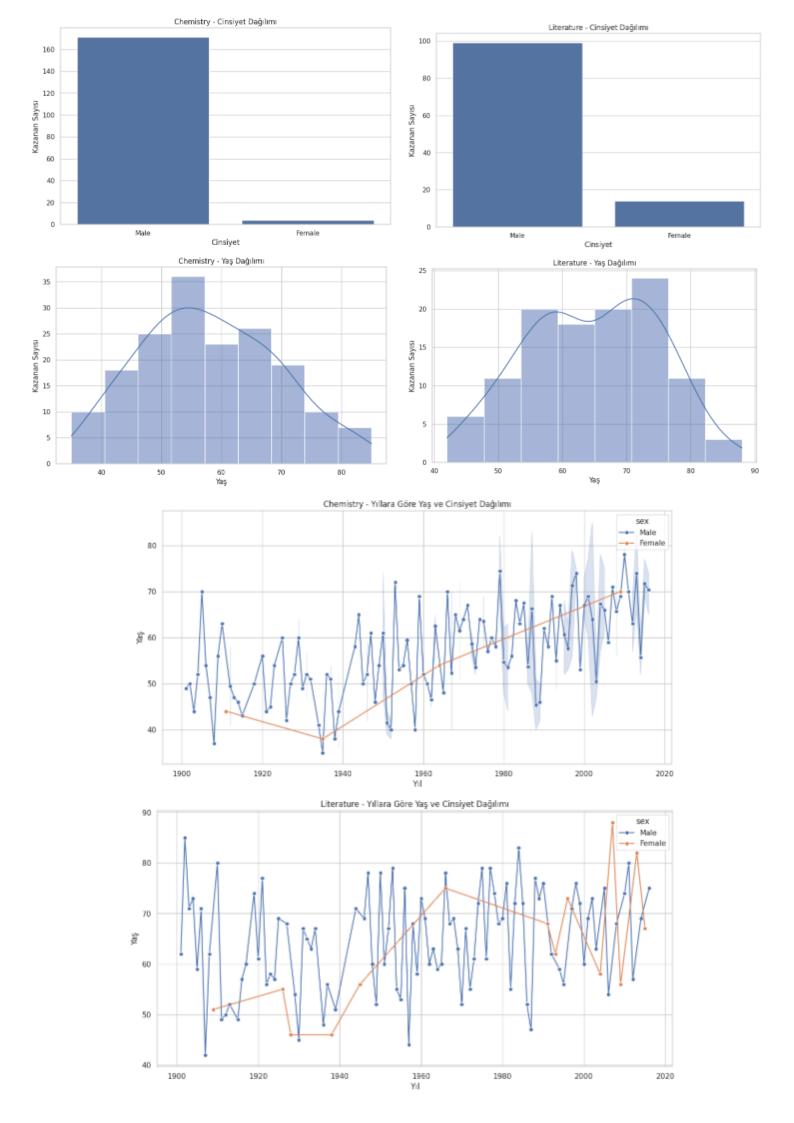
Amerika'daki nobel ödül sayısının çok olmasını istikrarlı bir şekilde bilimsel araştırmalara ,teknolojiye veya eğitime bütçe ayırması,üstüne düşmesine bağlayabiliriz. İkinci Dünya Savaşı'ndan sonra bu özellikle kazanılan ilk yüksek ödül sayılarıyla gözle görülmektedir.

Bu artışlarla ilgili diğer nedenlerde İkinci Dünya Savaşı sonrası Avrupadaki bilim insanları Amerika'ya göç ederek Amerika'nın bilimsel başarılarına katkı sağlamış olmaları durumu da söz konusu olabilir.

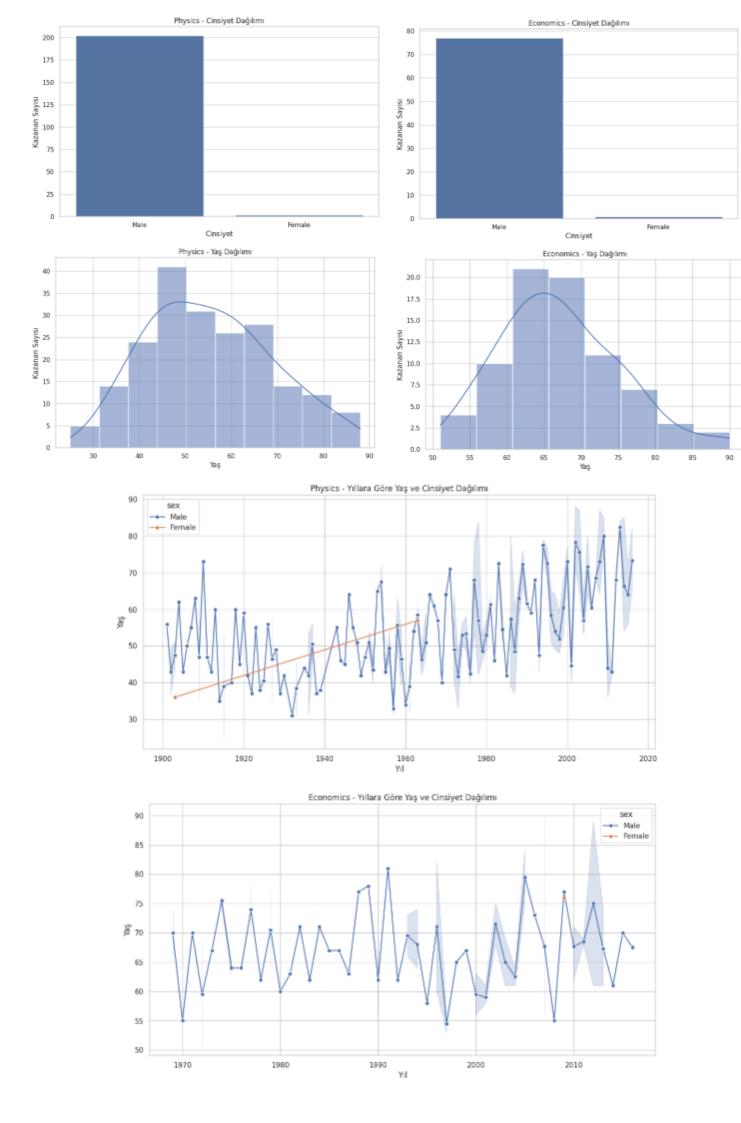
Sonuç olarak belirli dönem aralıklarında Amerika'da Nobel ödülü kazanımı istikrarlı şekilde devam etmiştir.Bu da uzun yıllar boyunca geliştirme 'araştırma ve bütçelemeye devam ettiklerini göstermektedir.

5. Nobel Ödülü kazananların cinsiyetlerini, yaşlarını, ödül kategorisi ve yılları kullanarak görselleştiriniz.(Her bir ödül kategorisi için ayrı grafik gösteriniz) Çıkan sonuçlara göre görseli yorumlayınız.

```
%pip install seaborn
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv("nobel.csv")
df = df[['year', 'category', 'full name', 'birth date', 'sex']]
df['birth date'] = pd.to datetime(df['birth date'], errors='coerce')
df['age'] = df['year'] - df['birth_date'].dt.year
categories = df['category'].unique()
sns.set(style="whitegrid")
for category in categories:
    category data = df[df['category'] == category]
    plt.figure(figsize=(10, 6))
    sns.countplot(data=category_data, x='sex')
    plt.title(f'{category} - Cinsiyet Dağılımı')
    plt.xlabel('Cinsiyet')
    plt.ylabel('Kazanan Sayısı')
    plt.show()
    plt.figure(figsize=(10, 6))
    sns.histplot(data=category_data, x='age', kde=True)
    plt.title(f'{category} - Yaş Dağılımı')
    plt.xlabel('Yas')
    plt.ylabel('Kazanan Sayısı')
    plt.show()
    plt.figure(figsize=(14, 7))
    sns.lineplot(data=category_data, x='year', y='age', hue='sex', marker='o')
    plt.title(f'{category} - Yıllara Göre Yaş ve Cinsiyet Dağılımı')
    plt.xlabel('Y11')
    plt.ylabel('Yas')
    plt.show()
```







Kimya Kategorisi: Cinsiyet dağılımında, erkek kazananların sayısı çok daha fazladır.Yaş dağılımında, en çok ödül alanların 50-60 yaş aralığında olduğunu söyleyebiliriz. Yıllara göre yaş ve cinsiyet dağılımında, yıllara göre kazananların yaşı değişiklik gösterdiğini söyleyebilirim.

Edebiyat Kategorisi: Cinsiyet dağılımında, erkek kazananların sayısı çok daha fazladır.Ancak kadınların diğer kategorilere oranla bu kategoride 0-20 aralığı daha yüksektir. Yaş dağılımında, en çok ödül alanların 50 yaş üstü olduğunu söyleyebiliriz. Yıllara göre yaş ve cinsiyet dağılımında, bu kategoride daha yaşlı ödül alanların fazla olduğunu söyleyebiliriz.

Tıp Kategorisi: Cinsiyet dağılımında, erkek kazananların sayısı çok daha fazladır. Yaş dağılımında, en çok ödül alanların 50-60 yaş aralığında olduğunu söyleyebiliriz. Yıllara göre yaş ve cinsiyet dağılımında, bu kategoride son yıllarda daha yaşlı adayların ödül aldığını söyleyebiliriz.

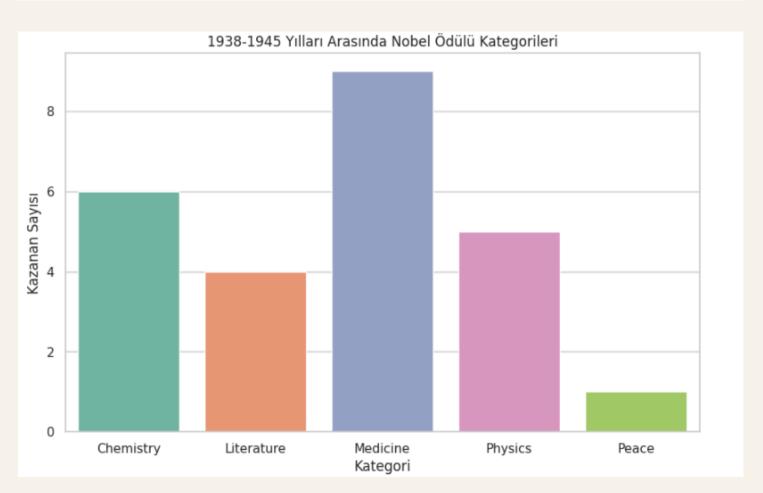
Barış Kategorisi: Cinsiyet dağılımında, erkek kazananların sayısı çok daha fazladır.Ancak kadınlar diğer kategorilere göre bu kategoride en yüksek sayıya sahiptir.Yaş dağılımında, en çok ödül alanların 50-80 yaş aralığında olduğunu söyleyebiliriz. Yıllara göre yaş ve cinsiyet dağılımında, 2014 yılında 17 yaşındaki Malala Yusufzay'a nobel ödülü verilmesiyle grafik genç yaşlar skalasında bulunmaktadır.Aynı şekilde bu kategorinin erkek cinsiyetinde diğer kategorilere göre daha genç yaştakilerin ödül aldığını söyleyebiliriz.

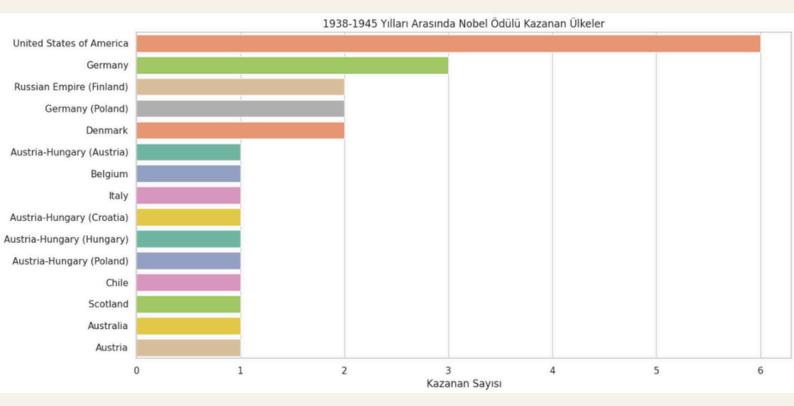
Fizik Kategorisi: Cinsiyet dağılımında, erkek kazananların sayısı çok daha fazladır. Yaş dağılımında, en çok ödül alanların 40-70 yaş aralığında olduğunu söyleyebiliriz. Yıllara göre yaş ve cinsiyet dağılımında, bu kategoride son yıllarda daha yaşlı adayların ödül aldığını söyleyebiliriz.

Ekonomi Kategorisi: Cinsiyet dağılımında, erkek kazananların sayısı çok daha fazladır. Yaş dağılımında, en çok ödül alanların 60-70 yaş aralığında olduğunu söyleyebiliriz. Yıllara göre yaş ve cinsiyet dağılımında, bu kategoride son yıllarda daha genç adayların ödül aldığını söyleyebiliriz.

6. 1938–1945 yılı arasında Nobel Ödülü kazananların kategorilerini ve ülkelerini görselleştirip yorumlayınız.

```
%pip install seaborn
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv('nobel.csv')
df_filtered = df[(df['year'] >= 1938) & (df['year'] <= 1945)]</pre>
df_filtered = df_filtered[['year', 'category', 'birth_country']].dropna()
sns.set(style="whitegrid")
plt.figure(figsize=(10, 6))
sns.countplot(data=df_filtered, x='category',hue='category', palette='Set2',dodge=False, legend=False)
plt.title('1938-1945 Yılları Arasında Nobel Ödülü Kategorileri')
plt.xlabel('Kategori')
plt.ylabel('Kazanan Sayısı')
plt.show()
plt.figure(figsize=(14, 7))
sns.countplot(data=df_filtered, y='birth_country',hue='birth_country', palette='Set2', dodge=False,
legend=False, order=df_filtered['birth_country'].value_counts().index)
plt.title('1938-1945 Yılları Arasında Nobel Ödülü Kazanan Ülkeler')
plt.xlabel('Kazanan Sayısı')
plt.ylabel('Ülke')
plt.show()
```



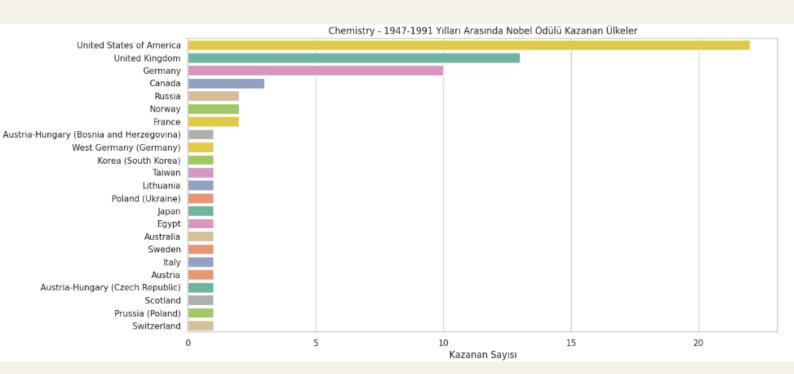


1938-1945 yılları arasında en fazlaTıp kategorisinde ödül verilmiştir. Ancak genel olarak verilen ödül sayılarında 7 yılda az bir ödül dağıtımı söz konusudur. Bu durum 2.Dünya Savaşı zamanlarında olunmasından ve savaşın etkilerinden kaynaklı olduğunu söylemek mümkün. Ayrıca o yıllarda ekonomi kategorisinde ödül verilmediğinden görselde bulunmamaktadır.

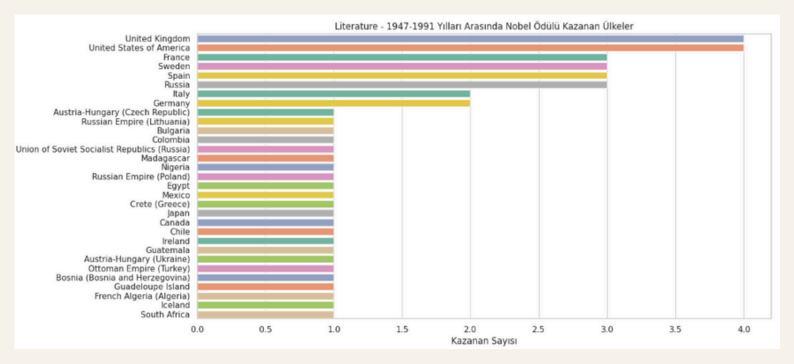
En çok Nobel ödülü kazanan ülke Amerika'dır.Avrupa'nın belli ülkelerinde araştırma ve geliştirme Amerika ölçüsünde önemsense de Amerika'nın nobel ödülü sayısının 2 kattan bile daha fazla çıkması durumunu göç almasına bağlayabiliriz.

7. 1947–1991 yılları arasında Nobel Ödülü kazananların kategorilerini ve ülkelerini görselleştirip yorumlayınız.(Her kategori için ayrı bir grafik olması istenmektedir)

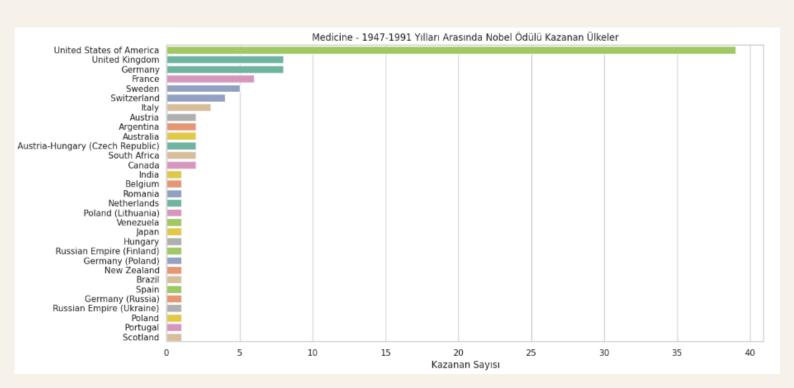
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv('nobel.csv')
df_filtered = df[(df['year'] >= 1947) & (df['year'] <= 1991)]</pre>
df_filtered = df_filtered[['year', 'category', 'birth_country']].dropna()
categories = df_filtered['category'].unique()
sns.set(style="whitegrid")
for category in categories:
    category_data = df_filtered[df_filtered['category'] == category]
    plt.figure(figsize=(14, 7))
    sns.countplot(data=category_data, y='birth_country', hue='birth_country', palette='Set2', dodge=False,
                  legend=False, order=category_data['birth_country'].value_counts().index)
    plt.title(f'{category} - 1947-1991 Yılları Arasında Nobel Ödülü Kazanan Ülkeler')
    plt.xlabel('Kazanan Sayısı')
    plt.ylabel('Ülke')
    plt.show()
```



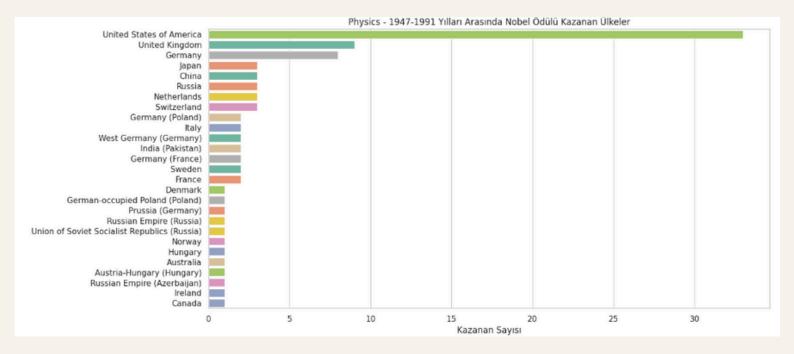
1947-1991 yılları arasında 1. sırada Amerika, bu ülkeyi takiben Birleşik Krallık ve Almanya Kimya ödülü kategorisinde en çok ödüle sahip ülkelerdir.



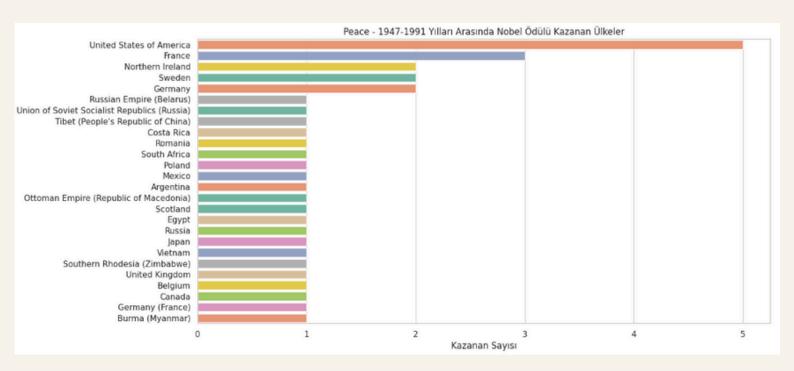
1947-1991 yılları arasında Edebiyat kategorisinde Birleşik Krallık ve Amerika en çok ödül alan ülkelerdir. Bu kategoride diğer kategorilere göre ödül alan diğer ülkelerin ödül sayısı daha fazladır.



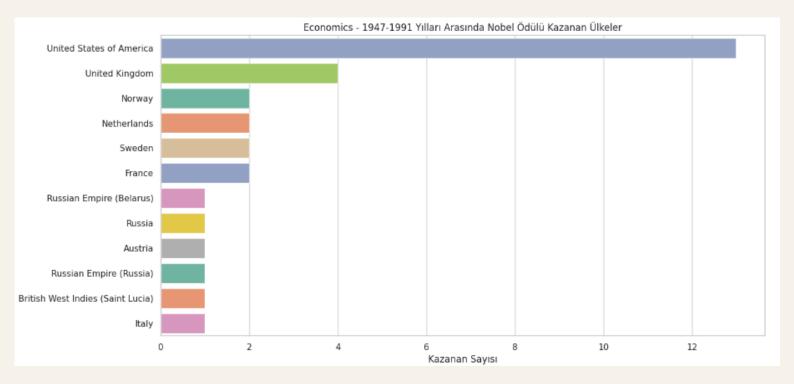
1947-1991 yılları arasında tıp kategorisinde açık ara farkla 1. sırada Amerika, bu ülkeyi takiben Birleşik Krallık ve Almanya en çok ödüle sahip ülkelerdir.



1947-1991 yılları arasında fizik kategorisinde açık ara farkla 1. sırada Amerika, bu ülkeyi takiben Birleşik Krallık ve Almanya en çok ödüle sahip ülkelerdir.



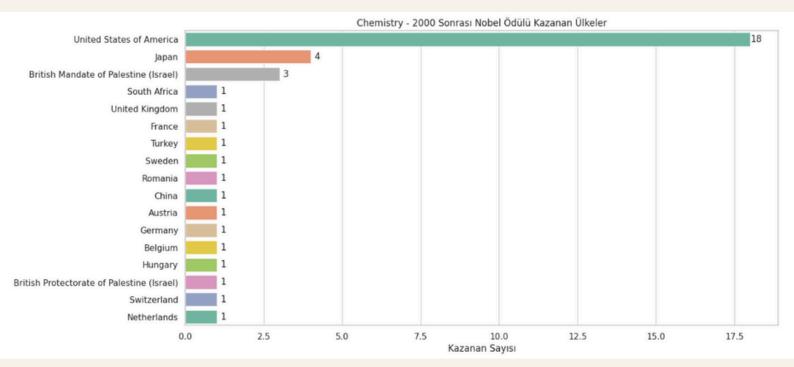
1947-1991 yılları arasında barış kategorisinde 1. sırada Amerika, bu ülkeyi takiben Fransa en çok ödüle sahip ülkelerdir.

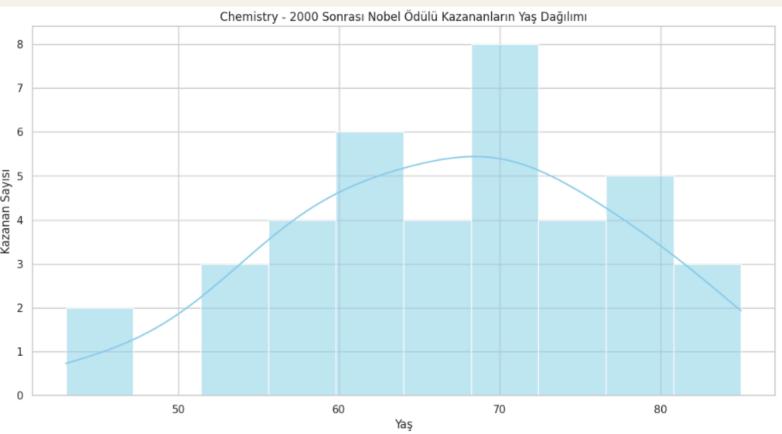


1947-1991 yılları arasında ekonomi kategorisinde 1. sırada Amerika, bu ülkeyi takiben Birleişk Krallık en çok ödüle sahip ülkelerdir.

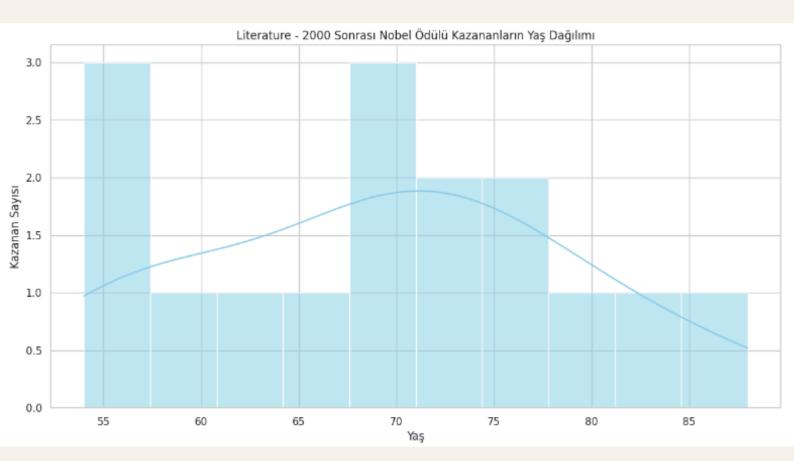
8. Kimya, Edebiyat, Barış, Fizik ve Tıp kategorilerindeki 2000 sonrasındaki kişilerin ülkelerini, yaşlarını görselleştirin. (Her bir Kategori için ayrı görselleştirme yapılması istenmektedir) Veriyi yorumlayınız.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv('nobel.csv')
df_filtered = df[df['year'] > 2000]
df_filtered = df_filtered[['year', 'category', 'birth_country', 'birth_date']].dropna()
df_filtered['birth_date'] = pd.to_datetime(df_filtered['birth_date'], errors='coerce')
df_filtered['age'] = df_filtered['year'] - df_filtered['birth_date'].dt.year
categories = ['Chemistry', 'Literature', 'Peace', 'Physics', 'Medicine']
sns.set(style="whitegrid")
for category in categories:
   category_data = df_filtered[df_filtered['category'] == category]
   plt.figure(figsize=(14, 7))
   ax1 = sns.countplot(data=category_data, y='birth_country', hue='birth_country', palette='Set2',
                        dodge=False, legend=False, order=category_data['birth_country'].value_counts().index)
   plt.title(f'{category} - 2000 Sonrası Nobel Ödülü Kazanan Ülkeler')
   plt.xlabel('Kazanan Sayısı')
   plt.ylabel('')
   for p in ax1.patches:
       width = p.get_width()
        plt.text(width + 0.2,
                p.get_y() + p.get_height() / 2,
                int(width),
                ha='center',
                va='center')
   plt.show()
   plt.figure(figsize=(14, 7))
   ax2 = sns.histplot(data=category_data, x='age', kde=True, bins=10, color='skyblue')
   plt.title(f'{category} - 2000 Sonrası Nobel Ödülü Kazananların Yaş Dağılımı')
   plt.xlabel('Yaş')
   plt.ylabel('Kazanan Sayısı')
   plt.show()
```

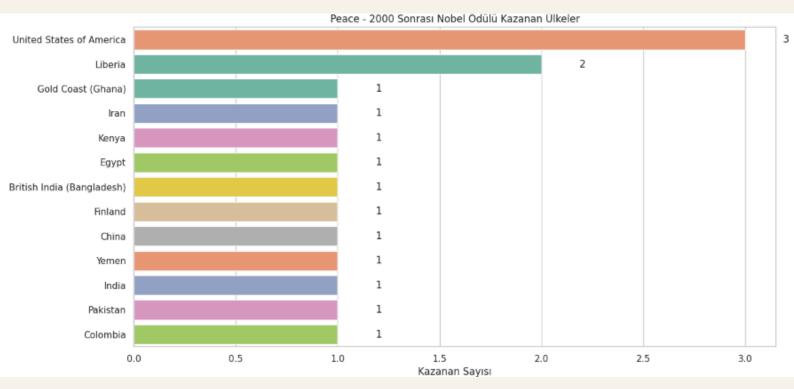


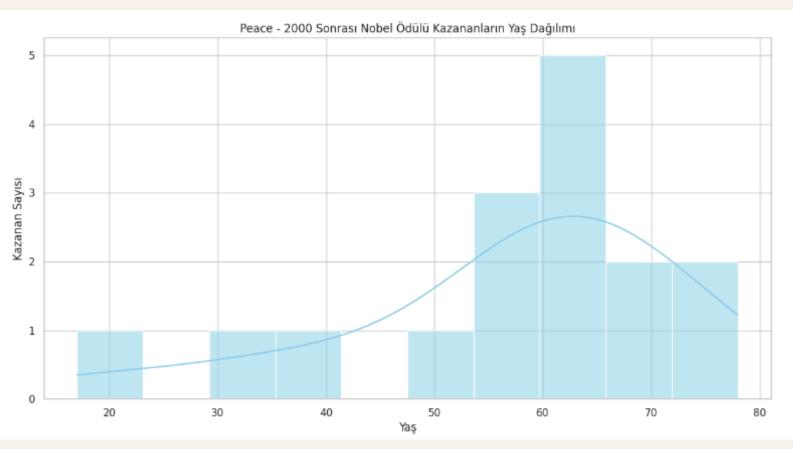


2000 yılı sonrası kimya kategorisinde sırayla Amerika, Japonya ve Israil öne çıkmaktadır.Nobel kazananların yaş dağılımına baktığımızda 50-70 yaş aralığında yoğunlaştığını söyleyebilirim.

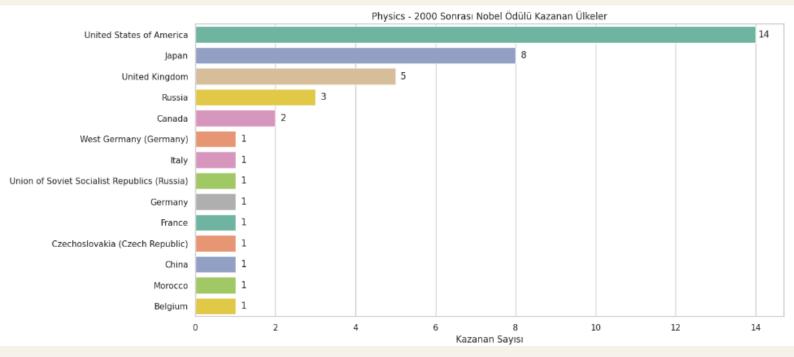


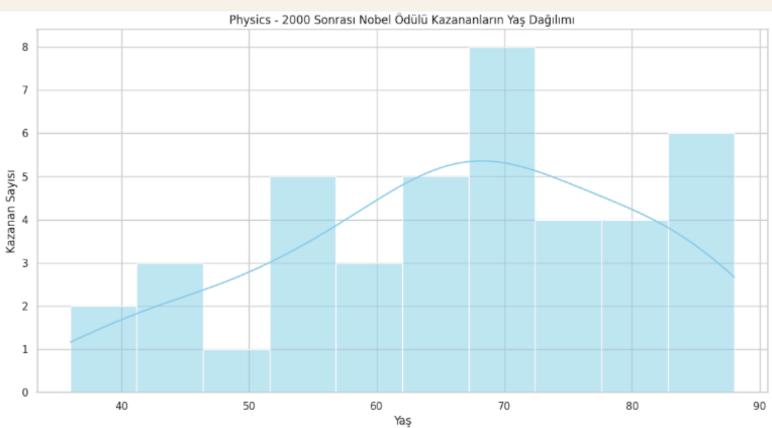
2000 yılı sonrası edebiyat kategorisinde 2 ödülle Fransa, diğer ülkelerde eşit şekilde 1 ödülle öne çıkmaktadır. Nobel kazananların yaş dağılımına baktığımızda grafikte 50 yaş diliminde yükseliş ve 70 yaş aralığında yükseliş olduğunu söyleyebilirim.



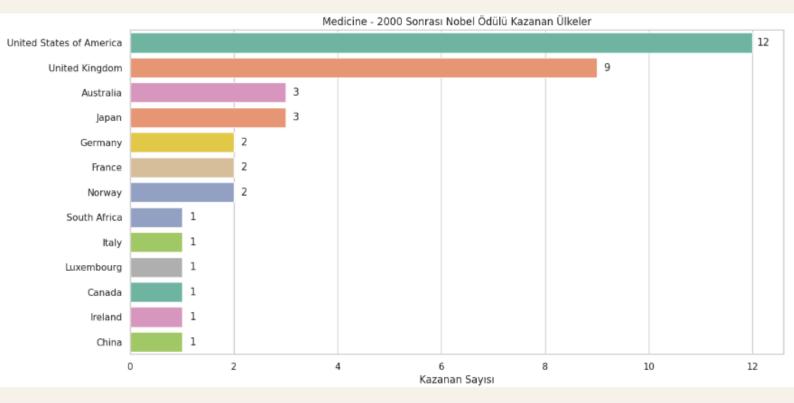


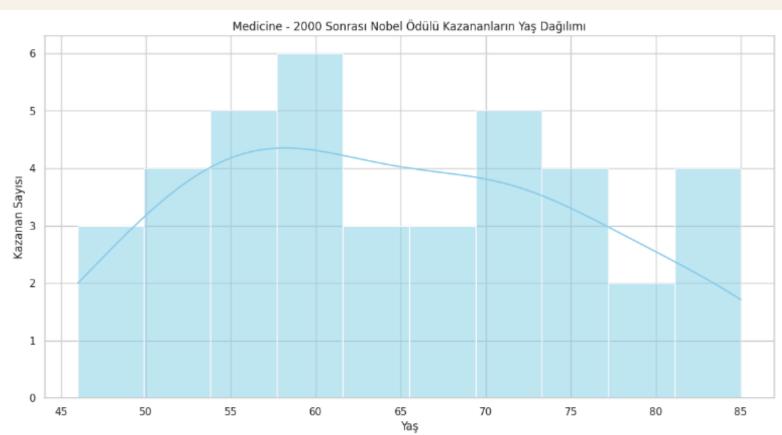
2000 yılı sonrası barış kategorisinde 3 ödülle Amerika,2 ödülle Liberya, diğer ülkelerde eşit şekilde 1 ödülle öne çıkmaktadır. Nobel kazananların yaş dağılımına baktığımızda grafikte 1 ödül alan ülkelerde ödül alanlar daha genç yaşlara sahipler. Ayrıca 60-70 yaş aralığında ödül alanlarda sayıca fazla olduğunu söyleyebilirim.





2000 yılı sonrası fizik kategorisinde sırayla Amerika 14, Japonya 8, Birleşik Krallık 5 ödülle öne çıkmaktadır. Nobel kazananların yaş dağılımına baktığımızda en çok ödül alanların 70 yaşlarında sayıca fazla olduğunu söyleyebilirim.





2000 yılı sonrası tıp kategorisinde sırayla Amerika 12, Birleşik Krallık 9, Avustralya ve Japonya 3 ödülle öne çıkmaktadır. Nobel kazananların yaş dağılımına baktığımızda 50-65 yaş ve 70-85 yaş aralığında en çok ödül alanların sayıca fazla olduğunu söyleyebilirim.