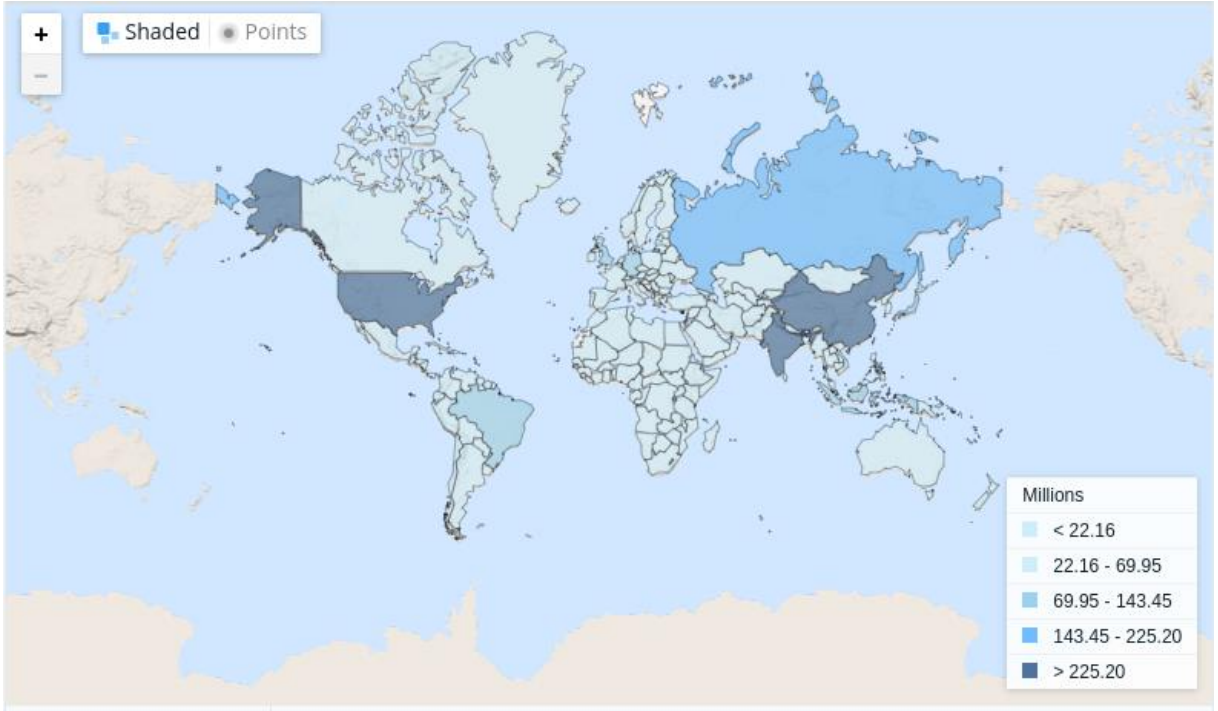


DÜNYA NÜFUS VERİLERİNE DAYANARAK ÜLKELERİN GELECEK NÜFUSLARININ TAHMİNLENMESİ

14 Kasım 2022 yi 15 kasım 2022 ye bağlayan gece dünya tarihinde önemli bir tarih oldu. Bu tarihte dünya nüfusu 8 milyara ulaştı ve dünya literatürüne bugün 8 milyar günü olarak geçti. 1925 yılında 2 milyar olan nüfusumuz 1974 yılında bunu katlayarak 4 milyar seviyelerine geldi. Dünya nüfusunun 2 milyardan 4 milyara çıkması 49 yıl alırken, bu nüfusu tekrar ikiye katlaması 47 yıl sürdü.

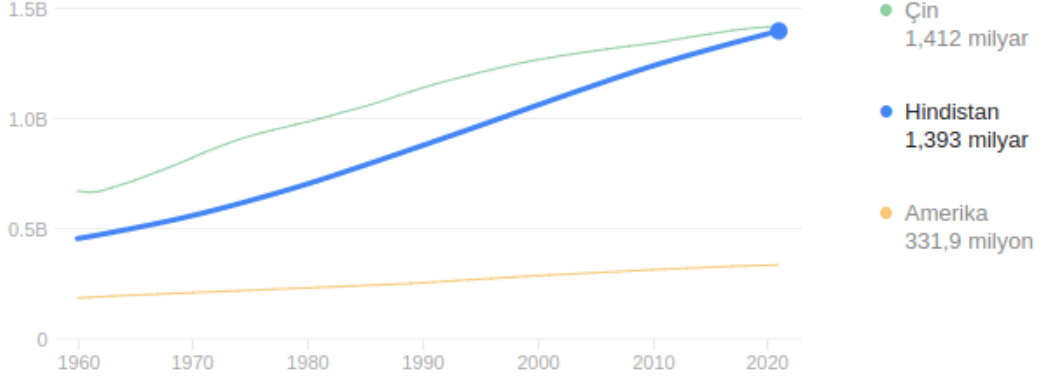
Bugün dünya nüfusu artışa devam etmekle birlikte gitgide nüfus artış hızının düştüğü gözlenmektedir. Nüfusumuzun 9 milyar seviyesini aştıktan sonra 10 milyara ulaşmadan nüfus artış hızının eksiye döneceği ve nüfusun tekrar 8 milyara altına düşeceğine dair hesaplamalar bulunmaktadır.

İstatistiksel olarak bugüne kadar yaşanmış en iyi durumdayız. İnsanlar daha uzun yaşıyor, geçen nesillere göre yiyeceği, temiz suya ve sağlık hizmetlerine erişim daha kolay, çok daha az insan fakirlik sınırının altında yaşıyor.



Her ne kadar dünyada haberler iyi olsa da ülkelerin nüfus ve olanaklarına göre bu durumun tersine yaşandığı ülkelerde mevcuttur. Günümüzde en kalabalık ülkenin Çin Halk Cumhuriyeti olduğu bilinmektedir ancak nüfus artış hızının yavaşlamasıyla birlikte Çin yakın bir gelecekte dünyanın en kalabalık ikinci ülkesi konumuna düşecektir. Nüfus artış hızı yükselmeye devam eden ve şu anda en kalabalık ikinci ülke konumunda bulunan Hindistan yakın bir gelecekte birinci sırayı alacaktır.

1,393 milyar (2021)



Kaynaklar şunları içerir: Dünya Bankası, Amerika Birleşik Devletleri Nüfus Sayım Bürosu

Geri Bildirim

Bu veriler ışığında yapılan çalışmamızda ülkelerin dünya nüfusundaki yeri ve yıllara göre nüfusunun alacağı hali görmek adına dünya bankası nüfus verilerine dayalı olarak lineer regresyon kullanarak hesaplama yapıldı.

Bu işlem için yine dünya bankasının sitesi üzerinde data bölümünde bulunan veriler ile çalıştık. İlgili sayfa da verileri üç farklı şekilde alabiliyorduk. Bu şekiller csv, xml ve xlsx veri tipleriydi. Öncelikle csv dosyası olarak aldık ancak dosya içerisinde Python tarafından tanınmayacak şekilde veriler olduğu görüldü. Bu nedenle xlsx tipindeki verileri bilgisayara indirerek dosyada bulunan data sayfası csv dosyası olarak dışa aktarıldı.

Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964
Aruba	ABW	Population, total	SP.POP.TOTL	54208	55434	56234	56699	57181
Africa Eastern and Southern	AFE	Population, total	SP.POP.TOTL	130836765	134159786	137614644	141202036	144789118
Afghanistan	AFG	Population, total	SP.POP.TOTL	8996967	9169406	9351442	9543200	9734958
Africa Western and Central	AFW	Population, total	SP.POP.TOTL	96396419	98407221	100506960	102691339	104885718
Angola	AGO	Population, total	SP.POP.TOTL	5454938	5531451	5608499	5679409	5750319
Albania	ALB	Population, total	SP.POP.TOTL	1608800	1659800	1711319	1762621	1813923
Andorra	AND	Population, total	SP.POP.TOTL	13410	14378	15379	16407	17435
Arab World	ARB	Population, total	SP.POP.TOTL	92197715	94724540	97334438	100034191	102738944
United Arab Emirates	ARE	Population, total	SP.POP.TOTL	92417	100801	112112	125130	138148
Argentina	ARG	Population, total	SP.POP.TOTL	20481781	20817270	21153042	21488916	21824790
Armenia	ARM	Population, total	SP.POP.TOTL	1874119	1941498	2009524	2077584	2145644
American Samoa	ASM	Population, total	SP.POP.TOTL	20127	20605	21246	22029	22812
Antigua and Barbuda	ATG	Population, total	SP.POP.TOTL	54132	55005	55849	56701	57553
Australia	AUS	Population, total	SP.POP.TOTL	10276477	10483000	10742000	10950000	11158000
Austria	AUT	Population, total	SP.POP.TOTL	7047539	7086299	7129864	7175811	7221768
Azerbaijan	AZE	Population, total	SP.POP.TOTL	3895398	4030325	4171428	4315127	4458826
Burundi	BDI	Population, total	SP.POP.TOTL	2797925	2852438	2907320	2964416	3021512
Belgium	BEL	Population, total	SP.POP.TOTL	9153489	9183948	9220578	9289770	9358962
Benin	BEN	Population, total	SP.POP.TOTL	2431617	2465865	2502897	2542864	2582831
Burkina Faso	BFA	Population, total	SP.POP.TOTL	4829289	4894580	4960328	5027811	5095294
Bangladesh	BGD	Population, total	SP.POP.TOTL	48013505	49362834	50752150	52202008	53692466
Bulgaria	BGR	Population, total	SP.POP.TOTL	7867374	7943118	8012946	8078145	8143344
Bahrain	BHR	Population, total	SP.POP.TOTL	162429	167899	173140	178142	183144
Bahamas, The	BHS	Population, total	SP.POP.TOTL	109532	115119	121092	127340	133588
Bosnia and Herzegovina	BIH	Population, total	SP.POP.TOTL	3225664	3288604	3353228	3417573	3482113

Elde ettiğimiz veri seti 1960 yılından 2021 yılına kadar olan 266 ülkeye ait nüfus verileridir. Veri seti içerisinde ülke Country Name, Country Code, Indicator Name ve Indicator Code sütunları bulunmaktadır.

Çalışmaya başlarken ilk olarak ilgili kütüphaneler sisteme eklendi. Bu kütüphaneler öncelikli olarak pandas, matplotlib ve sklearn kütüphanesidir. Ayrıca arka planda oluşacak bazı hataların atlanabilmesi ve ekrana yansıtılmaması amacıyla warning kütüphanesi ile ignore tipindeki hata mesajları filtrelenmiştir.

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
import warnings
warnings.filterwarnings("ignore")
```

Sonrasında edinilen veri seti programımıza import edilerek dataframe olarak kaydedilmiştir.

```
In [19]: data = pd.read_csv('pop2.csv')
```

Sonrasında elde bulunan veri setiyle ilgili elde bulunan verileri gözlemek amacıyla bazı denemeler yapıldı.

Veri setinin ilk 5 satırı.

```
In [6]: data.head()
```

	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	...	2012	2013
0	Aruba	ABW	Population, total	SP.POP.TOTL	54208.0	55434.0	56234.0	56699.0	57029.0	57357.0	...	102565.0	103165.0
1	Africa Eastern and Southern	AFE	Population, total	SP.POP.TOTL	130836765.0	134159786.0	137614644.0	141202036.0	144920186.0	148769974.0	...	547482863.0	562601578.0
2	Afghanistan	AFG	Population, total	SP.POP.TOTL	8996967.0	9169406.0	9351442.0	9543200.0	9744772.0	9956318.0	...	31161378.0	32269592.0
3	Africa Western and Central	AFW	Population, total	SP.POP.TOTL	96396419.0	98407221.0	100506960.0	102691339.0	104953470.0	107289875.0	...	370243017.0	380437896.0
4	Angola	AGO	Population, total	SP.POP.TOTL	5454938.0	5531451.0	5608499.0	5679409.0	5734995.0	5770573.0	...	25107925.0	26015786.0

5 rows × 66 columns

Veri setinin son 5 satırı.

```
In [7]: data.tail()
```

	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965	...	2012	2013
261	Kosovo	XXK	Population, total	SP.POP.TOTL	947000.0	966000.0	994000.0	1022000.0	1050000.0	1078000.0	...	1807106.0	1818117.0
262	Yemen, Rep.	YEM	Population, total	SP.POP.TOTL	5315351.0	5393034.0	5473671.0	5556767.0	5641598.0	5727745.0	...	24473176.0	25147112.0
263	South Africa	ZAF	Population, total	SP.POP.TOTL	17099836.0	17524533.0	17965733.0	18423157.0	18896303.0	19384838.0	...	52832659.0	53687125.0
264	Zambia	ZMB	Population, total	SP.POP.TOTL	3070780.0	3164330.0	3260645.0	3360099.0	3463211.0	3570466.0	...	14465148.0	14926551.0
265	Zimbabwe	ZWE	Population, total	SP.POP.TOTL	3776679.0	3905038.0	4039209.0	4178726.0	4322854.0	4471178.0	...	13115149.0	13350378.0

5 rows × 66 columns

Veri setiyle ilgili bazı bilgiler.

```
In [22]: data.describe()
```

Out[22]:

	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	...
count	2.640000e+02	2.640000e+02	2.640000e+02	2.640000e+02	2.640000e+02	2.640000e+02	2.640000e+02	2.640000e+02	2.640000e+02	2.640000e+02	...
mean	1.172174e+08	1.187633e+08	1.208717e+08	1.234910e+08	1.261315e+08	1.288372e+08	1.316853e+08	1.345256e+08	1.374350e+08	1.404490e+08	...
std	3.695745e+08	3.739180e+08	3.804316e+08	3.889142e+08	3.974401e+08	4.062000e+08	4.155171e+08	4.247722e+08	4.342805e+08	4.441772e+08	...
min	2.833000e+03	3.077000e+03	3.367000e+03	3.703000e+03	4.063000e+03	4.460000e+03	4.675000e+03	4.922000e+03	5.194000e+03	5.461000e+03	...
25%	5.022802e+05	5.109642e+05	5.206540e+05	5.311622e+05	5.421252e+05	5.533362e+05	5.647475e+05	5.823645e+05	5.981078e+05	6.100030e+05	...
50%	3.718330e+06	3.826398e+06	3.929109e+06	4.015834e+06	4.124521e+06	4.242788e+06	4.326013e+06	4.387887e+06	4.474171e+06	4.550402e+06	...
75%	2.636053e+07	2.721235e+07	2.808607e+07	2.890669e+07	2.972333e+07	3.055227e+07	3.134845e+07	3.200449e+07	3.244145e+07	3.277149e+07	...
max	3.032156e+09	3.071596e+09	3.124561e+09	3.189656e+09	3.255146e+09	3.322047e+09	3.392098e+09	3.461620e+09	3.532783e+09	3.606554e+09	...

8 rows x 62 columns

Bu aşamadan sonra veriseti üzerinden model eğitilmesi ve regresyon analizi yaparak nüfusun belirlenmesi üzerine çalışıldı. Bu aşamada 4 adet fonksiyon geliştirildi. İlk fonksiyon modelin eğitilmesi amacıyla oluşturulan `populationreg` fonksiyonudur.

```
In [2]: def populationreg(x_train, y_train):
```

```
    reg = LinearRegression()
    reg.fit(x_train, y_train)
    return reg
```

İkinci fonksiyon model eğitilmesi için yapılan düzenlemeler, sonucun gösterilmesi ve bu verilerin matplotlib lib kütüphanesi kullanılarak analiz edilen verilerin gösterilmesini içermektedir.

```
In [15]: def regReturn(country, year, country_list, data):
```

```
    if country in country_list:
        data = data.loc[data['Country Name'] == country]
        data.drop(['Country Name', 'Country Code', 'Indicator Name', 'Indicator Code'], axis=1, inplace=True)
        data = data.T
        data.dropna(inplace=True)
        data = data.reset_index()
        x1 = data.iloc[:, 0]
        y1 = data.iloc[:, 1]
        x = x1.to_numpy().reshape(-1, 1).astype(int)
        y = y1.to_numpy().reshape(-1, 1).astype(int)

        x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.1)

        reg = populationreg(x_train, y_train)
        predict_pop = int(reg.coef_[0][0] * year + reg.intercept_[0])
        print("%s ülkesinin %d yılında hesaplanan nüfusu %d " % (country, year, predict_pop))
        plt.scatter(x_train, y_train, color="red")
        plt.scatter(x_test, y_test, color="blue")

        plt.plot(x_train, reg.predict(x_train), color="green")
        plt.title("Nüfus verisi için oluşturulan regresyon doğrusu")
        plt.xlabel("Yıl")
        plt.ylabel("Nüfus")
        plt.show()
    else:
        print("Verisetinde ülke adı bulunamadı")
```

Burada öncelikli olarak gereksiz sütunlar silinmiş ve veriler içinde sadece verilen ülkeye ait nüfus verileri ile yıl verilerinin kalması sağlanmıştır. Sonrasında dataset transpoze edilerek yıl ve nüfus verileri satır halinden sütun haline çevrilmiştir. Daha önce incelediğimiz

verilerde null veri bulunmamasına rağmen dropna metodu kullanılarak bundan emin olunmuştur. Indexlerin düzenlenmesinden sonra veriler integer haline getirilip veriler test ve train veriler olarak yapılandırılmıştır.

Elde edilen veriler ilk fonksiyona gönderilerek model eğitimi tamamlanmış ve sonrasında gelen regresyon doğrusu yıl ile çarpılıp bir düzeltme fonksiyonundan geçirilip nüfus verisi ile grafik çizdirilmiştir.

Aynı işlem tüm dünya nüfusu için 3. Fonksiyonda da yapılmıştır.

```
In [4]: def regAllCountry(year,data):
        data = data.T
        data.dropna(inplace=True)
        data = data.reset_index()
        x1 = data.iloc[:, 0]
        y1 = data.iloc[:, 1]
        x = x1.to_numpy().reshape(-1, 1).astype(int)
        y = y1.to_numpy().reshape(-1, 1).astype(int)
        from sklearn.model_selection import train_test_split
        x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.1)
        reg = populationreg(x_train, y_train)
        predict_pop = int(reg.coef_[0][0] * year + reg.intercept_[0])
        print("Dünyanın %d yili için hesaplanan nüfusu : %d " % (year, predict_pop))
        plt.scatter(x_train, y_train, color="red")
        plt.scatter(x_test, y_test, color="blue")

        plt.plot(x_train, reg.predict(x_train),color="green")
        plt.title("Nüfus verisi için oluşturulan regresyon doğrusu")
        plt.xlabel("Yıl")
        plt.ylabel("Nüfus")
        plt.show()
```

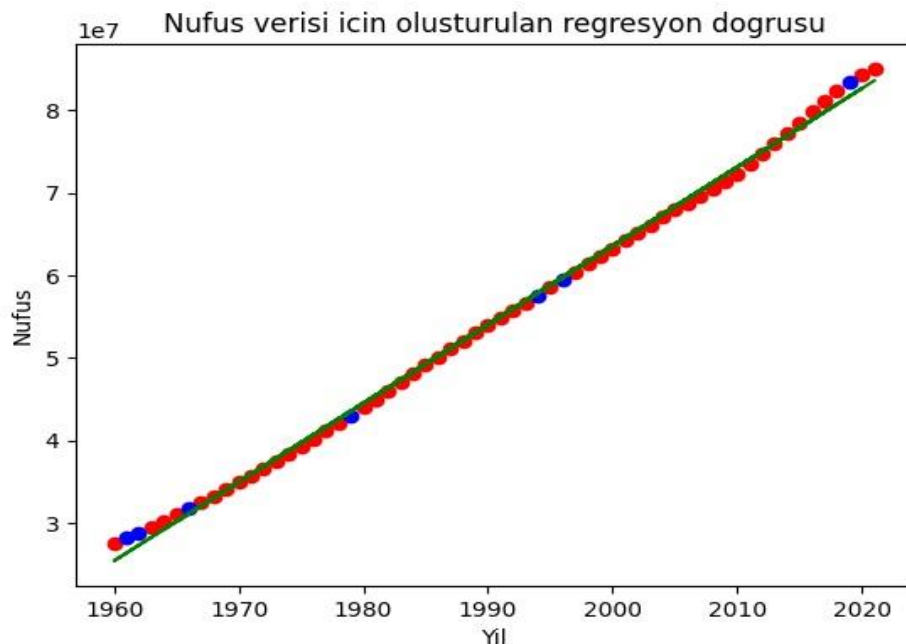
Burada elde edilen fonksiyon aslında 4. fonksiyondan alınan verileri ülke kısıtlaması olmadan modeli eğitmektedir. Son fonksiyonumuzda gereksiz satırlar temizlenmiş ve sütun toplamları alınarak dünya nüfusu verileri hesaplanmış ve eğitime gönderilmiştir.

```
In [5]: def allCountryPop(data, year):
        data.drop(['Country Name', 'Country Code', 'Indicator Name', 'Indicator Code'], axis=1, inplace=True)
        dataAll = data.max()
        regAllCountry(year,dataAll)
```

Sonrasında main fonksiyonunda ülke ve yıl adları alınıp fonksiyonlara gönderilmiştir.

```
In [13]: regReturn(country, year, country_list, data)
```

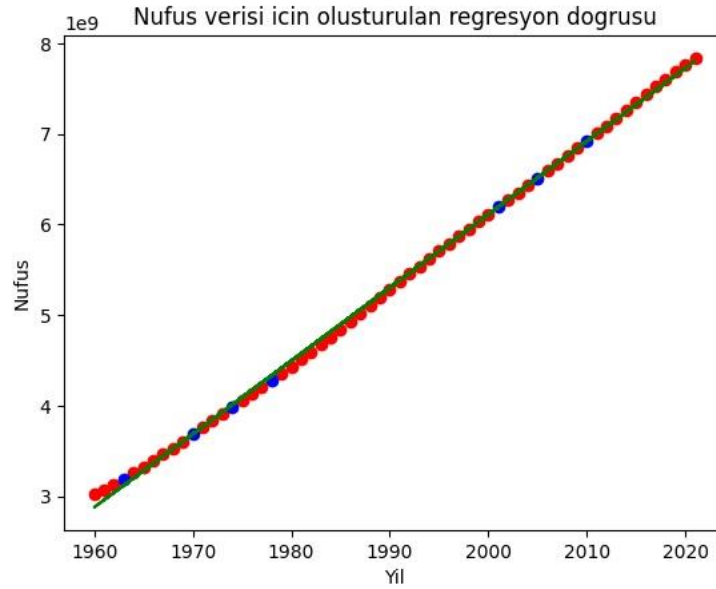
Türkiye ülkesinin 2023 yılında hesaplanan nüfusu 85519614



Regresyon verileri içerisinde train verileri kırmızı işaretlenirken test verileri mavi gösterilmiştir. Ülkemiz için yapılan hesaplama göre 2023 yılı nüfusumuzun 85 519 614, 2030 yılında 96 076 004 olacağı öngörülmüştür.

```
In [14]: allCountryPop(data, year)
```

Dünyanın 2023 yılı için hesaplanan nüfusu : 7968401256



Dünya nüfusu için yapılan hesaplama da ise 2030 yılında 8 528 648 226 olacağı 2040 yılında ise 9 338 074 482 olacağı öngörülmüştür.