

T.C.
KIRKLARELİ ÜNİVERSİTESİ
PINARHİSAR HÜSEYİN KEÇECİ YERLEŞKESİ
BİLGİSAYAR PROGRAMCILIĞI

PYTHON PROGRAMLAMA
GÜZ DÖNEMİ VERİ ANALİZİ PROJESİ

HAZIRLAYAN
BARAN ALYAR

SORUM ÖĞRETİM GÖREVLİSİ
DR. MURAT ASLANYÜREK

OCAK – 2024
KIRKLARELİ

Merhaba Sevgili Hocam ve Değerli Öğrenci Arkadaşlarım,

Bu projemde, sizlere aktarmak istediğim bir yazılım dili olan Python aracılığıyla veri setleri kullanımı, gösterimi ve görselleştirmesi üzerine odaklanan bir çalışma sunuyorum. Projemi kısaca anlatmam gerekirse, "Cost Of Living" yani Yaşam Maliyeti'nin ülkelere göre değerlerini analiz ediyorum. Toplamda 139 farklı ülkenin sırasıyla (Yaşam Maliyeti Endeksi, Kiralama Endeksi, Yaşam Maliyeti ve Kira Endeksi, Market Endeksi, Restoran Ödeme Endeksi, Yerel Satın Alma Endeksi) verilerini CSV formatında (tablo verilerini düz metin olarak depolayan, özel bir biçimde kaydedilmiş bir metin dosya formatı) türünde sunuyorum.

Python yazılım dilini kullanarak geliştirdiğim kodlar ile bu verileri işleyerek çeşitli sorulara yanıtlar buluyor ve görselleştirmelerle sunuyorum. Toplamda 13 farklı soru hazırladım, her biri farklı ve görselleştirmeleriyle desteklenmiş. Bu şekilde çıktılarım hem bilgilendirici hem de estetik bir görünüme sahip olacak şekilde tasarlandı.

Projem, yaşam maliyeti üzerine geniş kapsamlı bir analizi içermekte olup, her bir soru özelinde detaylı bilgiler sunmayı amaçlamaktadır. Sırasıyla 13 sorunun kodları ve kodlara dair açıklamalar aşağıya sıralanmıştır. **Ve en aşağıda paketlerin ve bazı kodların açıklaması vardır.**

Soru1 : En düşük yaşam maliyetine sahip 10 ülkeyi Donut Chart ile görselleştirilmiştir.

Soru2 : En yüksek yaşam maliyetine sahip 10 ülkeyi Donut Chart ile görselleştirilmiştir.

Soru3 : Yerel satın alma gücü endeksi 40'tan yüksek olan ülkeleri barplot ile görselleştirilmiştir.

Soru4 : Kiralama endeksine göre ülkelerin yaşam maliyeti endeksine olan dağılımını scatterplot ile görselleştirilmiştir.

Soru5 : İtalya'nın alt kategorilerdeki endekslerini pie chart ile görselleştirilmiştir.

Soru6 : Ülkelerin restoran fiyat endeksine göre sıralamasını barplot ile görselleştirilmiştir.

Soru7 : Ülkelerin ortalama kiralama endeksi ve yaşam maliyeti endeksi ilişkisini scatterplot ile görselleştirilmiştir.

Soru8 : En düşük kira endeksine sahip 10 ülkeyi barplot ile görselleştirilmiştir.

Soru9 : En Yüksek Kiralama endeksine 10 ülkeyi barplot ile görselleştirilmiştir.

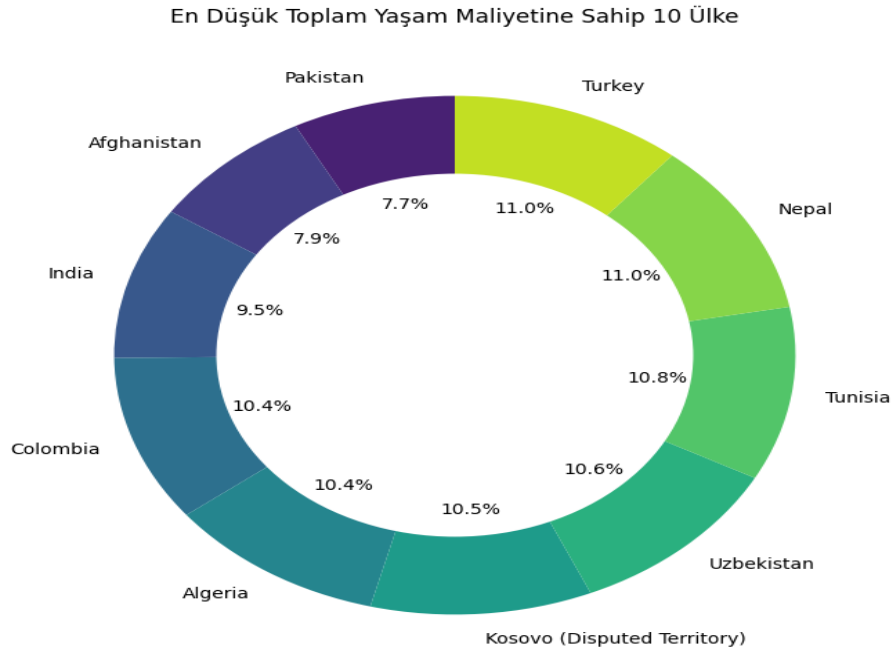
Soru10 :Yüksek yerel satın alma gücüne sahip 10 ülkeyi barplot ile görselleştirilmiştir.

Soru11 : Düşük yerel satın alma gücüne sahip 14 ülkeyi barplot ile görselleştirilmiştir.

Soru12 : Ülkelerin genel maliyet endeksine göre renkli haritasını heatmap ile görselleştirilmiştir

Soru13 : Kira endeksi 6-8 arasındaki ülkelerin değişimini lineplot ile görselleştirilmiştir.

Soru 1



Bu fonksiyon, veri setindeki ülkeleri yaşam maliyeti endeksine göre sıralar ve en düşük yaşam maliyetine sahip ilk 10 ülkeyi seçerek oluşturulmuştur.

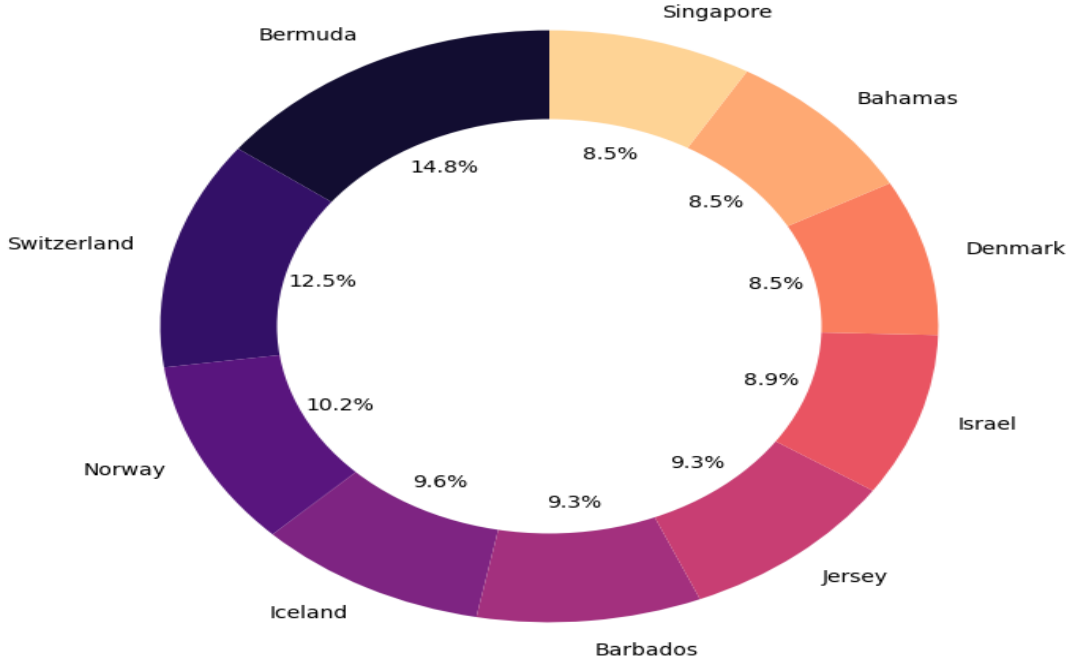
Seçilen bu ülkelerin yaşam maliyeti endeksini kullanarak bir Pasta Grafiği oluşturur. Her dilim, bir ülkeyi temsil eder ve dilimlerin büyüklükleri yaşam maliyeti endeksine göre orantılıdır.

Grafik üzerinde, her bir dilimin yanında ülke adı ve yaşam maliyeti endeksi yüzdesi bulunmaktadır.

Bu görselleştirme, en düşük yaşam maliyetine sahip olan ülkelerin oranlarını anlamak için kullanılır.

Soru 2

En Yüksek Toplam Yaşam Maliyetine Sahip 10 Ülke



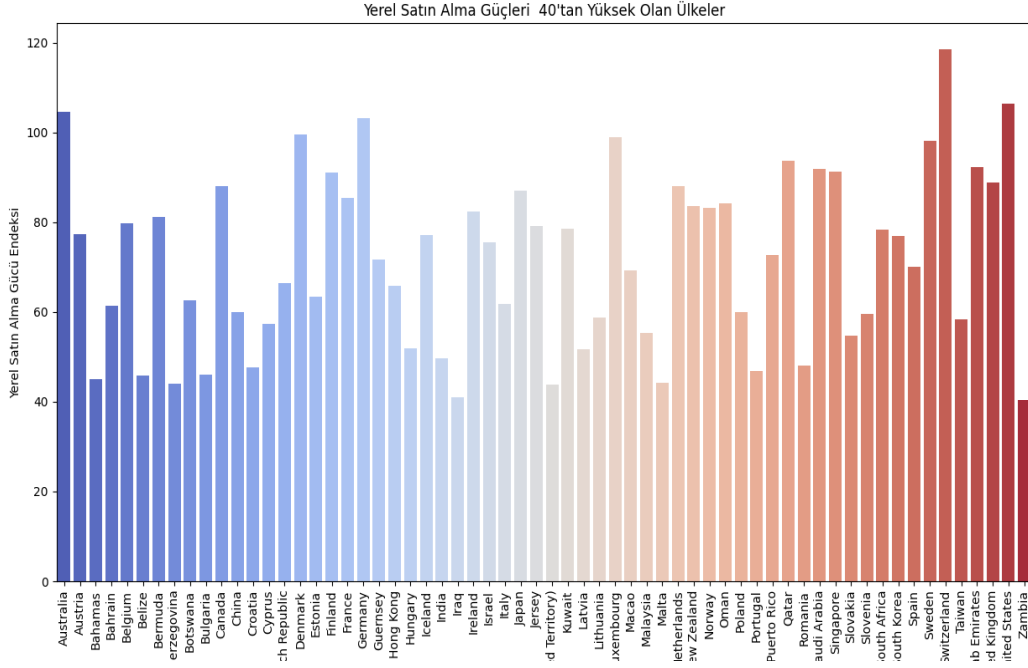
Fonksiyon, veri集中的i ülkeleri yaşam maliyeti endeksine göre sıralar ve en yüksek yaşam maliyetine sahip ilk 10 ülkeyi seçerek oluşturulmuştur.

Seçilen bu ülkelerin yaşam maliyeti endeksini kullanarak bir Pasta Grafiğı oluşturur. Her dilim, bir ülkeyi temsil eder ve dilimlerin büyüklükleri yaşam maliyeti endeksine göre orantılıdır.

Grafik üzerinde, her bir dilimin yanında ülke adı ve yaşam maliyeti endeksi yüzdesi bulunmaktadır.

Bu görselleştirme, en yüksek yaşam maliyetine sahip olan ülkelerin oranlarını anlamak için kullanılır.

Soru 3

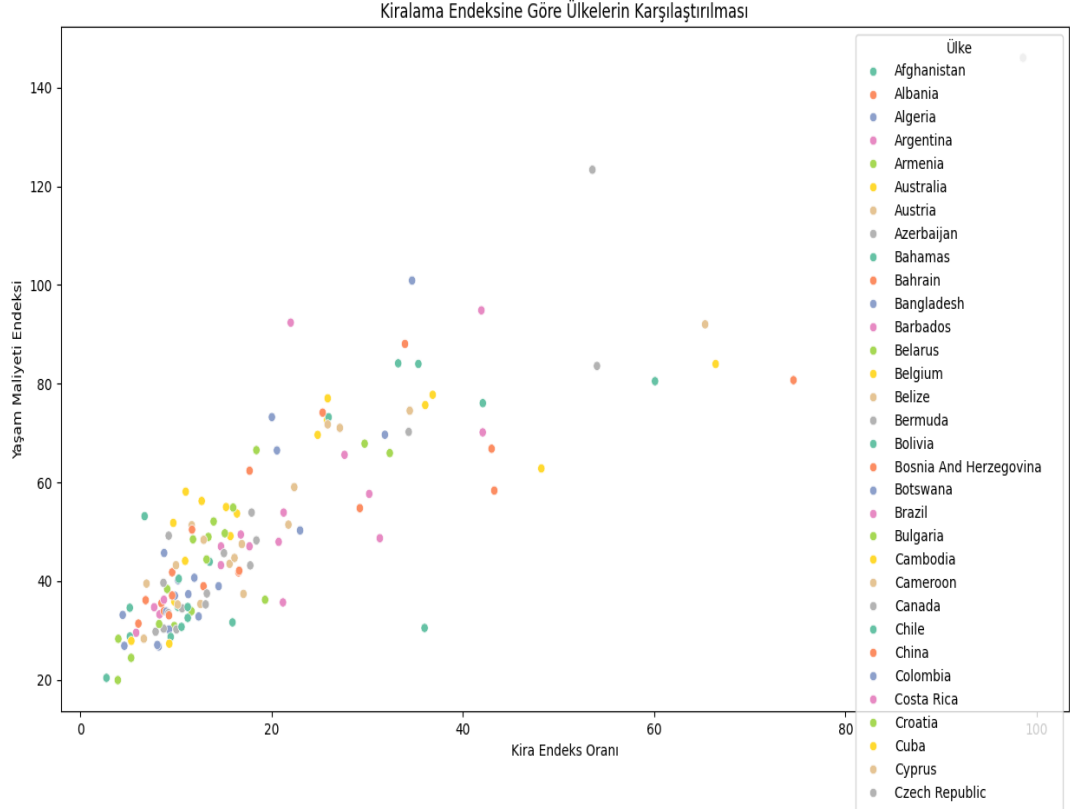


Fonksiyon, veri setindeki ülkeleri yerel satın alma gücü endeksine göre filtreler ve bu endeksi 40'tan yüksek olan ülkeyi seçerek oluşturulmuştur. Seçilen bu ülkelerin yerel satın alma gücü endeksini bir çubuk grafiği üzerinde görselleştirir.

Çubuk grafik, her bir ülkeyi temsil eder ve yüksek yerel satın alma gücüne sahip olan ülkelerin karşılaştırılmasını sağlar.

Bu görselleştirme, yerel satın alma gücü endeksi 40'tan yüksek olan ülkelerin sıralamasını ve karşılaştırmasını sunar.

Soru 4

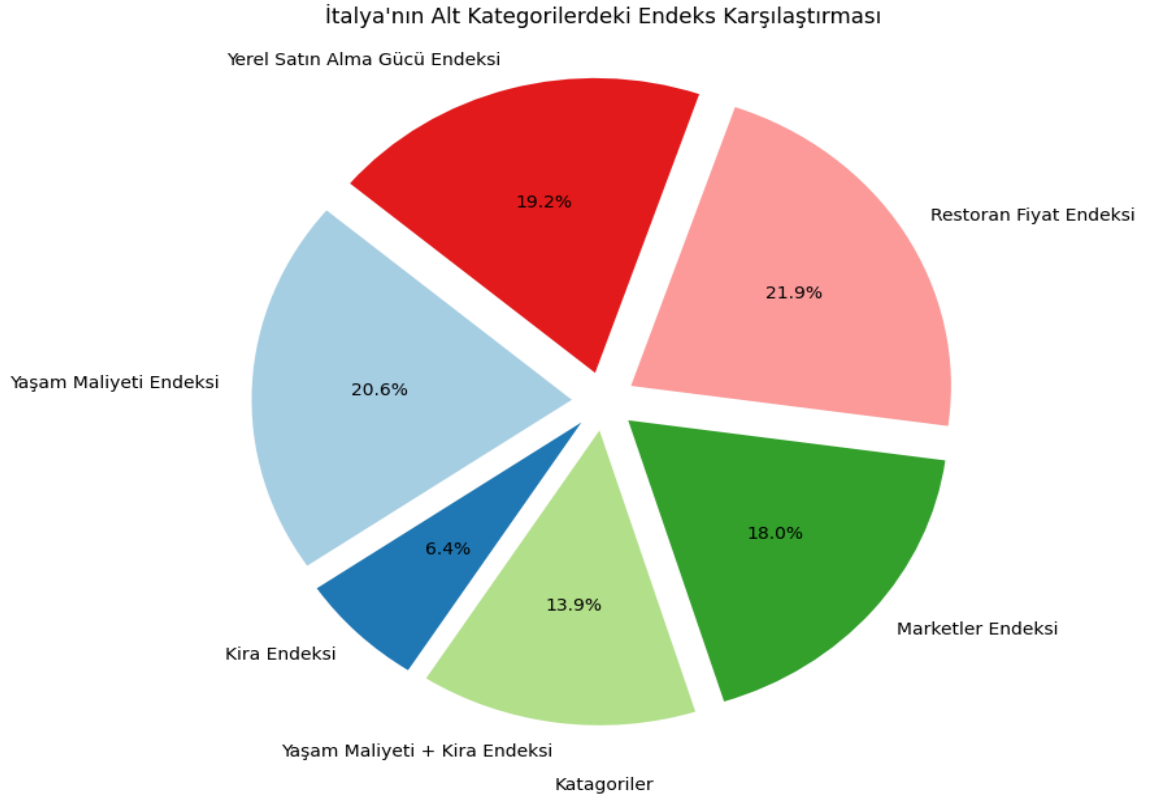


Bu fonksiyon, veri setindeki ülkeleri kira endeksi ve yaşam maliyeti endeksi üzerinden bir saçılım grafiği ile karşılaştırır. X eksenini kira endeksini, Y eksenini ise yaşam maliyeti endeksini temsil eder.

Her bir nokta bir ülkeyi temsil eder ve noktaların renkleri ülkeleri ayırt etmek için kullanılır.

Bu görselleştirme, kira endeksi ile yaşam maliyeti endeksi arasındaki ilişkiyi anlamak ve karşılaştırmak için kullanılır.

Soru 5



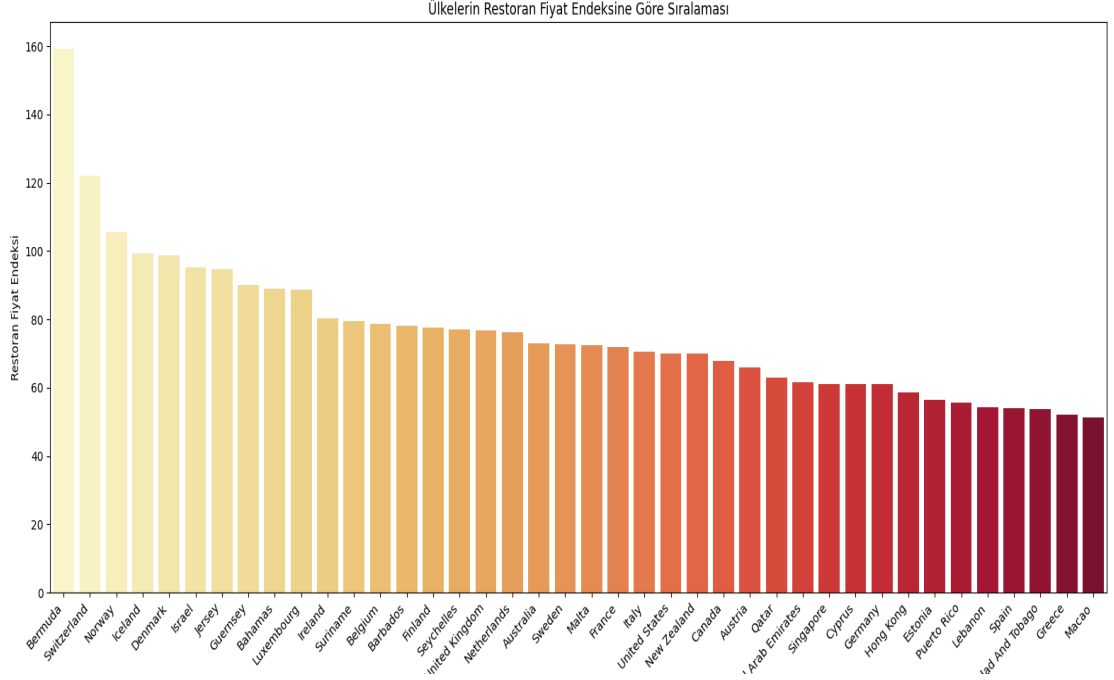
Fonksiyon, İtalya'nın alt kategorilerdeki yaşam maliyeti endeksini bir pasta grafiği ile görselleştirir.

İtalya'nın yaşam maliyeti endeksi alt kategorileri, pasta dilimleri olarak temsil edilir.

Pasta dilimlerinin büyüklükleri, İtalya'nın her bir alt kategorideki endeksine göre orantılıdır.

Bu görselleştirme, İtalya'nın yaşam maliyeti endeksinin alt kategorileri arasındaki dağılımı hakkında bilgi sağlar.

Soru 6



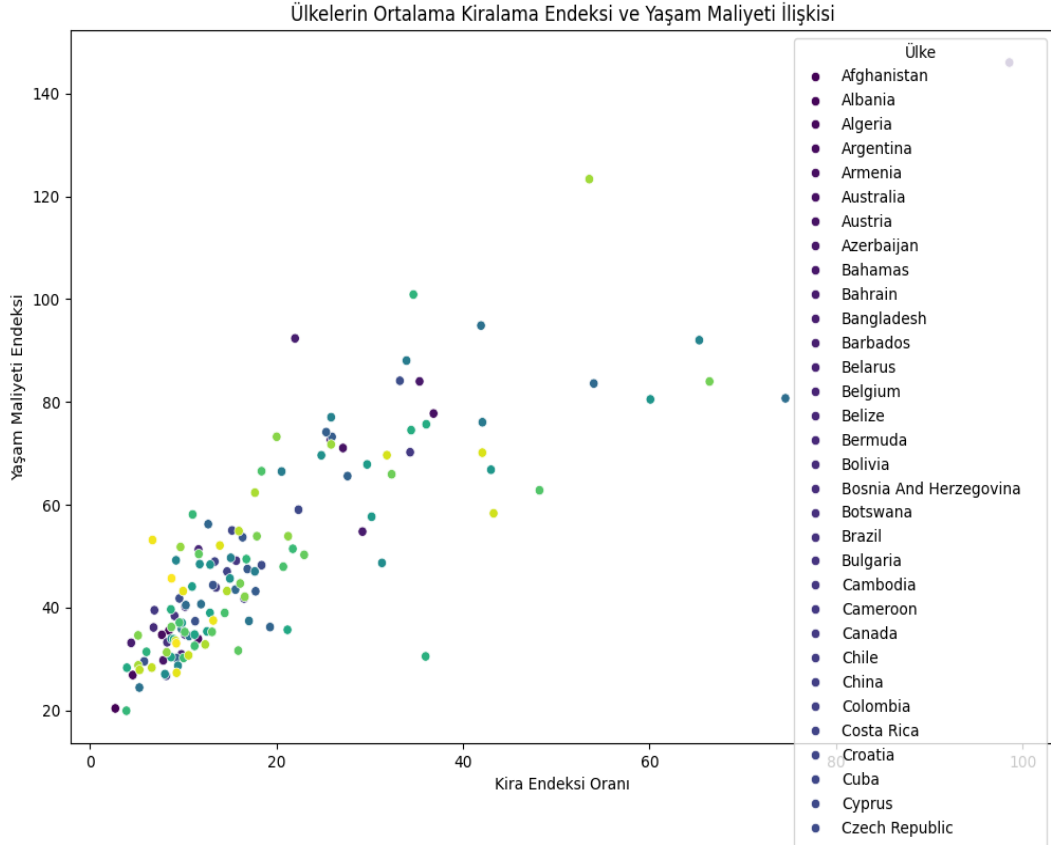
Fonksiyon, ülkeleri restoran fiyat endeksine göre sıralar ve en yüksek restoran fiyat endeksine sahip ilk 40 ülkeyi seçer.

Seçilen bu ülkelerin restoran fiyat endeksini bir çubuk grafiği ile görselleştirir.

Çubuk grafik, her bir ülkeyi temsil eder ve yüksek restoran fiyat endeksine sahip olan ülkelerin sıralamasını sunar.

Bu görselleştirme, restoran fiyat endeksine göre ülkeleri karşılaştırmak için kullanılır.

Soru 7



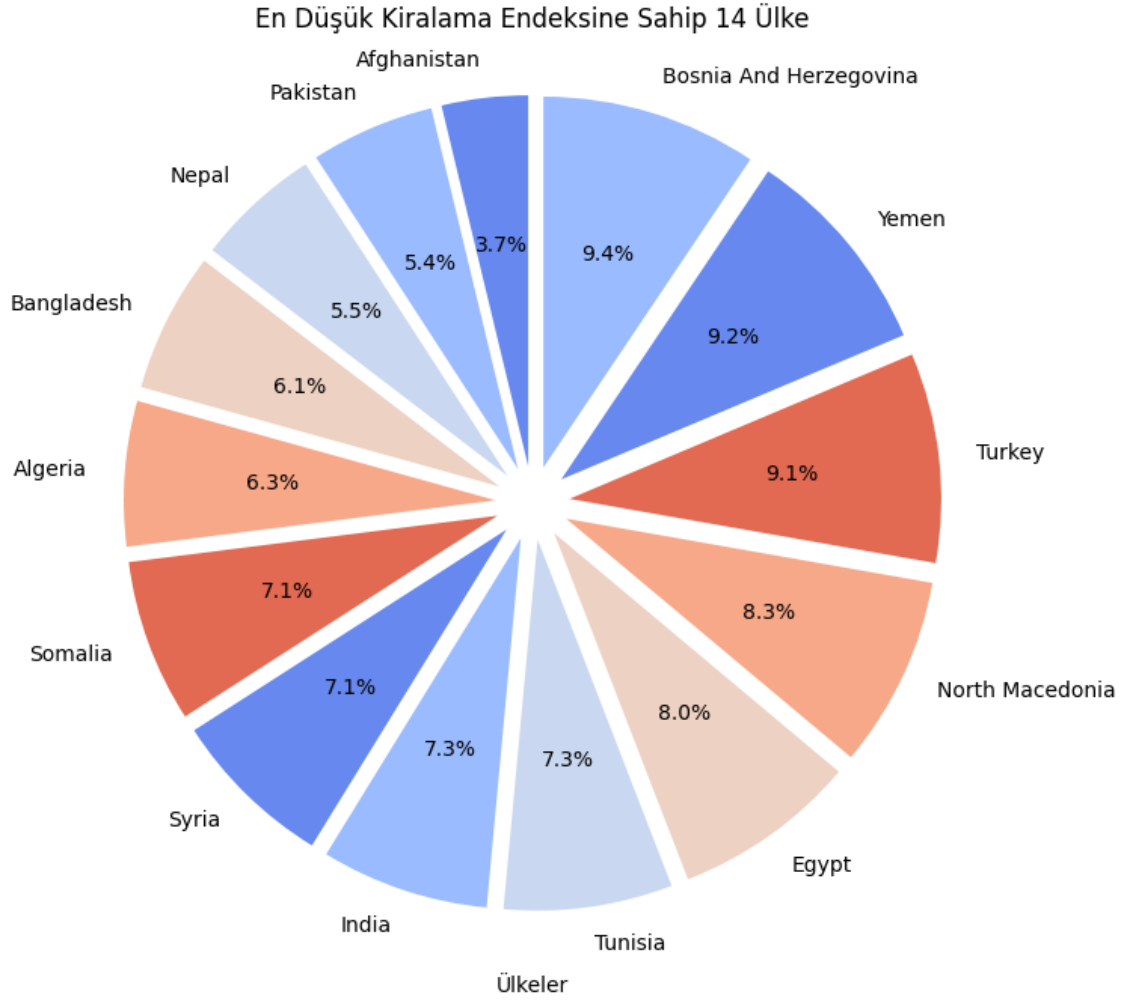
Bu fonksiyon, ülkeleri ortalama kira endeksi ve yaşam maliyeti endeksi üzerinden bir saçılım grafiği ile görselleştirir.

X eksenini ortalama kira endeksini, Y eksenini ise yaşam maliyeti endeksini temsil eder.

Her bir nokta bir ülkeyi temsil eder ve noktaların renkleri ülkeleri ayırt etmek için kullanılır.

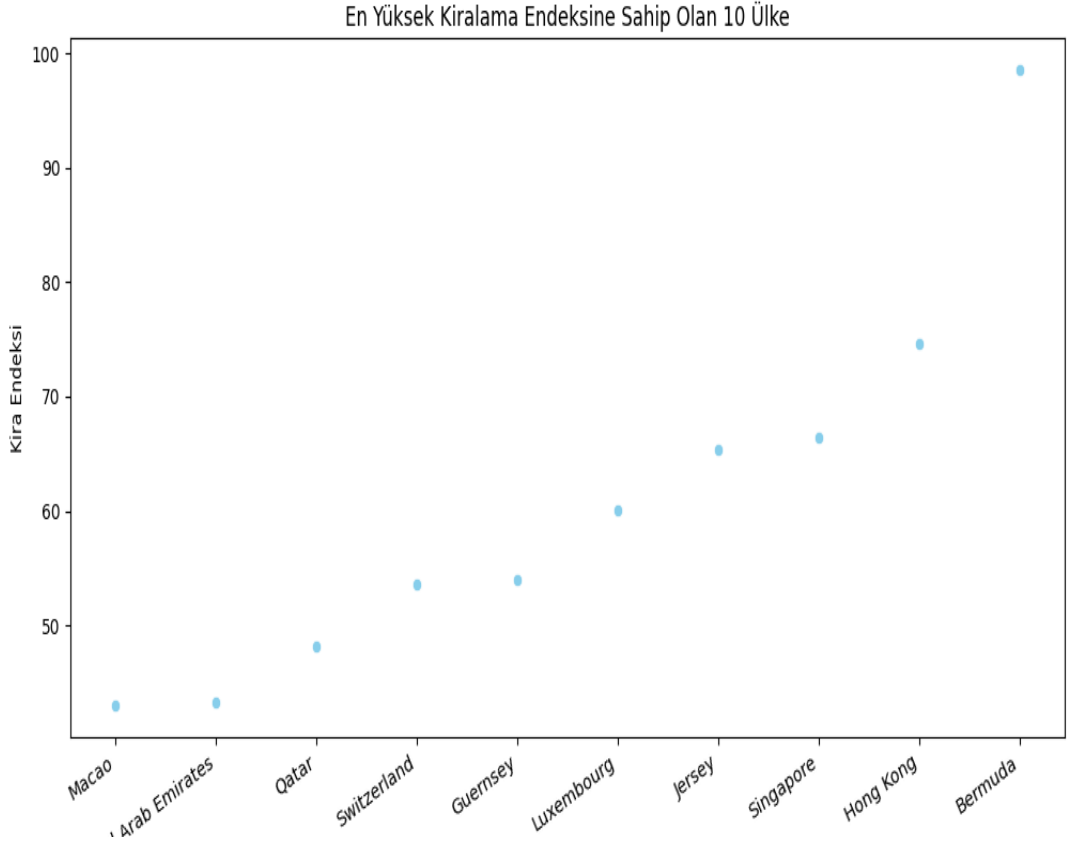
Bu görselleştirme, ortalama kira endeksi ile yaşam maliyeti endeksi arasındaki genel ilişkiyi anlamak için kullanılır.

Soru 8



Bu fonksiyon, ülkeleri kira endeksine göre sıralar ve en düşük kira endeksine sahip ilk 14 ülkeyi seçer. Seçilen bu ülkelerin kira endeksini bir pasta grafiği ile görselleştirir. Pasta dilimleri, her bir ülkeyi temsil eder ve dilimlerin büyüklükleri kira endeksine göre orantılıdır. Bu görselleştirme, en düşük kira endeksine sahip olan ülkelerin oranlarını anlamak için kullanılır.

Soru 9



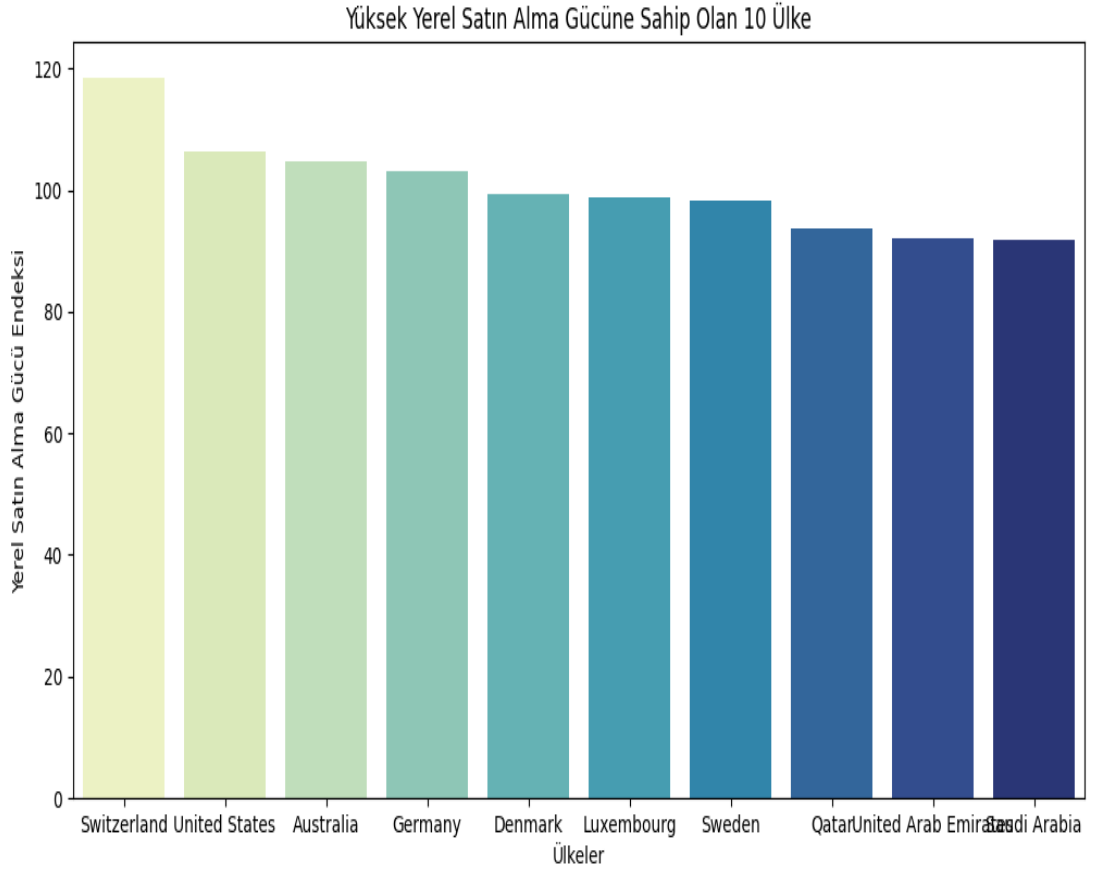
Bu fonksiyon, ülkeleri kira endeksine göre sıralar ve en yüksek kira endeksine sahip ilk 10 ülkeyi seçer.

Seçilen bu ülkelerin kira endeksini bir nokta grafiği ile görselleştirir.

Noktalar, her bir ülkeyi temsil eder ve yüksek kira endeksine sahip olan ülkelerin sıralamasını sunar.

Bu görselleştirme, kira endeksine göre ülkeleri karşılaştırmak için kullanılır.

Soru 10



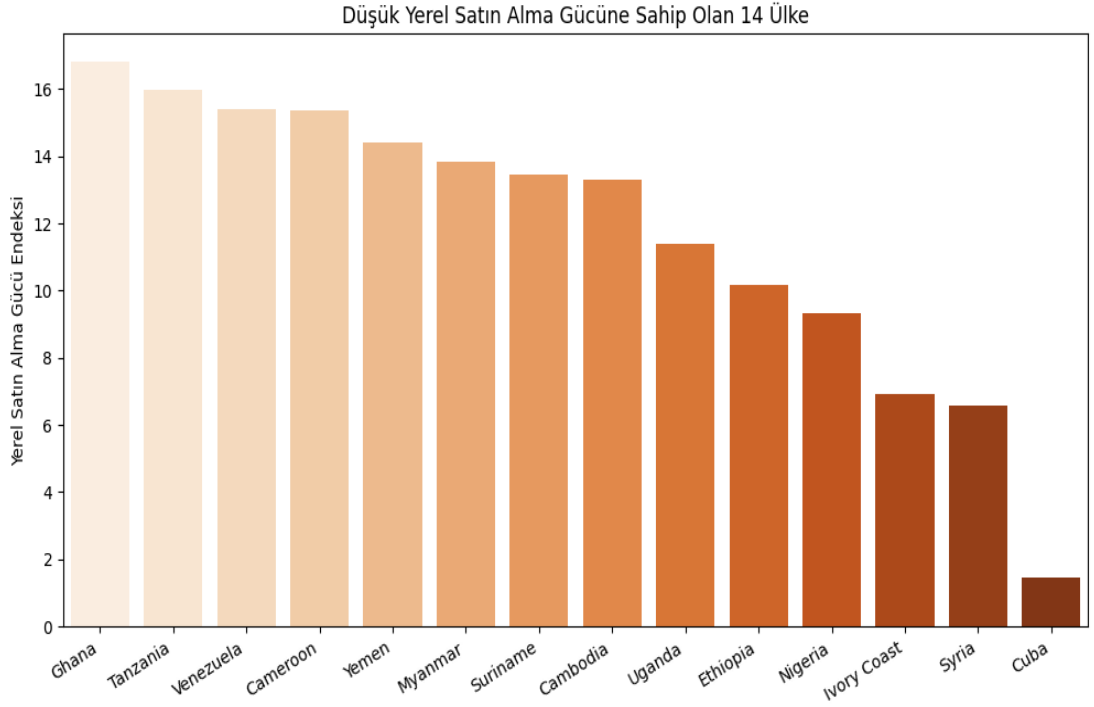
Bu fonksiyon, veri setindeki ülkeleri yerel satın alma gücü endeksine göre sıralar ve bu endeksi 40'tan yüksek olan ülkeleri seçer.

Seçilen bu ülkelerin yerel satın alma gücü endeksini bir çubuk grafiği ile görselleştirir.

Çubuk grafik, her bir ülkeyi temsil eder ve yüksek yerel satın alma gücüne sahip olan ülkelerin sıralamasını sunar.

Bu görselleştirme, yerel satın alma gücü endeksi 40'tan yüksek olan ülkelerin sıralamasını ve karşılaştırmasını sunar.

Soru 11



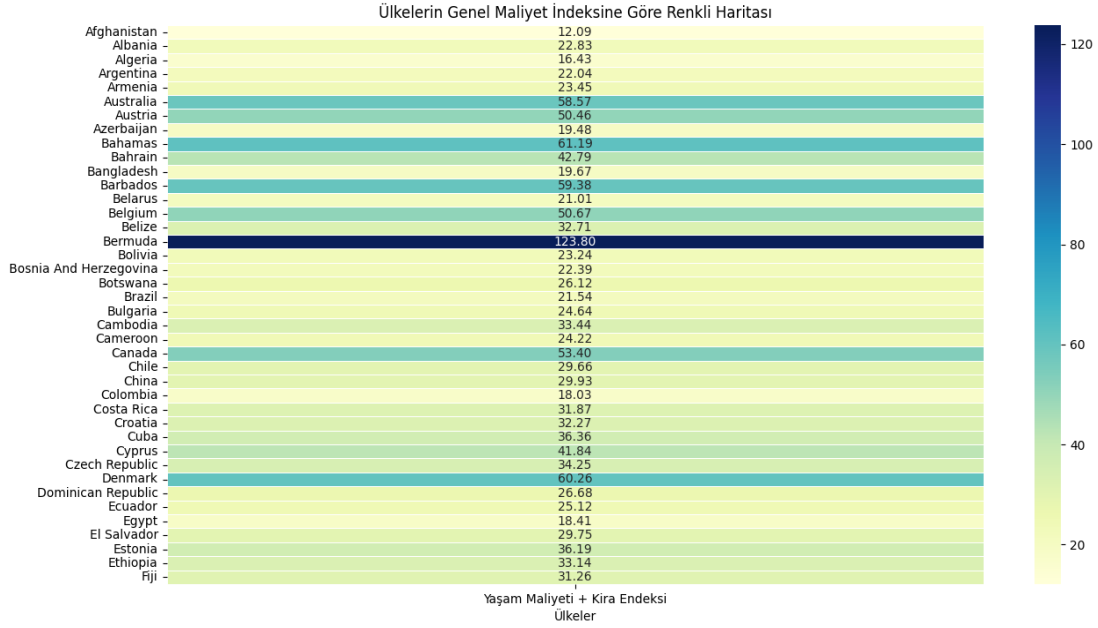
Bu fonksiyon, veri setindeki ülkeleri yerel satın alma gücü endeksine göre sıralar ve bu endeksi 40'tan küçük olan 14 ülkeyi seçer.

Seçilen bu ülkelerin yerel satın alma gücü endeksini bir çubuk grafiği ile görselleştirir.

Çubuk grafik, her bir ülkeyi temsil eder ve düşük yerel satın alma gücüne sahip olan ülkelerin sıralamasını sunar.

Bu görselleştirme, yerel satın alma gücü endeksi 40'tan küçük olan ülkelerin sıralamasını ve karşılaştırmasını sunar.

Soru 12

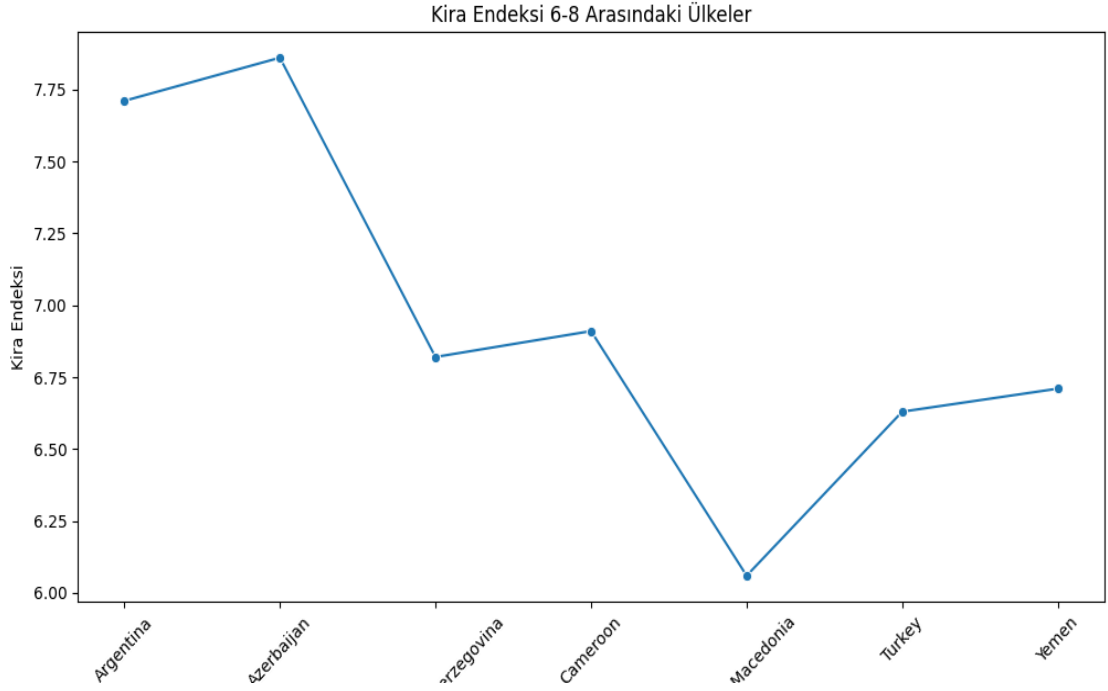


Bu fonksiyon, ülkelerin genel maliyet endeksine göre renkli bir harita oluşturur.

Haritadaki renk tonları, genel maliyet endeksini temsil eder ve her ülkeyi farklı renklerle ayırt eder.

Bu görselleştirme, genel maliyet endeksine göre dünya genelindeki ülkelerin dağılımını anlamak için kullanılır.

Soru 13



Bu fonksiyon, ülkeleri kira endeksine göre filtreler ve kira endeksi 6 ile 8 arasında olan ülkeleri bir çizgi grafiği ile görselleştirir.

Her çizgi, bir ülkeyi temsil eder ve kira endeksi 6 ile 8 arasındaki değişimleri gösterir.

Bu görselleştirme, belirli bir kira endeksi aralığındaki ülkelerin değişimlerini ve karşılaştırmalarını sunar.

Son olarak kodlarım ve bazı gerekli açıklamalar aşağıdadır.

pandas (pd)

Pandas, veri manipölasyonu ve analizi için kullanılan güçlü bir kütüphanedir. **pd.read_csv** fonksiyonu, CSV dosyalarını okumak ve veri çerçeveleri oluşturmak için kullanılır.

Veri çerçeveleri, tabloları ve veri manipölasyonunu kolaylaştıran bir veri yapısı sağlar.

numpy (np)

NumPy, bilimsel hesaplamalar ve çok boyutlu dizilerle çalışma için kullanılan temel bir kütüphanedir.

NumPy dizileri, veri manipölasyonu ve matematiksel işlemler için optimize edilmiş hızlı bir veri yapısı sunar.

matplotlib.pyplot (plt)

Matplotlib, veri görselleştirmesi için kullanılan bir kütüphanedir.

plt.figure, **plt.pie**, **plt.scatter**, **plt.barplot**, **plt.xticks**, **plt.title**, vb. gibi fonksiyonlar, çeşitli grafik türlerini oluşturmak ve özelleştirmek için kullanılır.

seaborn (sns)

Seaborn, Matplotlib üzerine inşa edilmiş bir veri görselleştirme kütüphanesidir.

sns.color_palette, **sns.barplot**, **sns.scatterplot** gibi fonksiyonlar, çeşitli istatistiksel grafiklerin oluşturulmasını kolaylaştırır.

Kodların Açıklanması

```
colors = sns.color_palette("viridis", len(top10))
```

Renk paletini oluşturuyoruz, burada 'viridis' rengini kullanıyoruz.

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

Şekilimizin boyutunu belirliyoruz.

```
plt.pie(colors=colors, autopct='%1.1f%%', startangle=90)
```

Pizza şeklinde dairesel dilim görselleştirmesi yapıyoruz.

```
categories = ["Yaşam Maliyeti Endeksi", "Kira Endeksi", "Yaşam Maliyeti + Kira Endeksi", "Marketler Endeksi", "Restoran Fiyat Endeksi", "Yerel Satın Alma Gücü Endeksi"]
```

Categories adını verdiğimiz değerin içinde hangi sütunlar olacak onları belirliyoruz.

```
values = italy_data[categories].values.flatten()
```

İtalya'ya ait veri çerçevesinde belirtilen kategorilere ait değerleri alıyoruz.

```
sorted_data = data_frame.sort_values(ascending=False)
```

Verimizde belirlediğimiz değerlere göre düzenleme yapıyoruz.

```
sns.barplot(x="x", y="y", data=sorted_data)
```

Çubuk grafik görünümü için kullanıyoruz.

```
plt.xticks(rotation=90)
```

x eksenindeki verileri 90 derece döndürme komudu.

```
explode = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1)
```

Pie Chart gibi görselleştirmelerde veri kümelerini birbirinden ayırmak için kullanıyoruz.

```
plt.xlabel("X Eksen Etiket")
```

X eksenindeki başlığı değiştirmek için kullanıyoruz.

```
plt.ylabel("Y Eksen Etiket")
```

y eksenindeki başlığı değiştirmek için kullanıyoruz.

```
plt.gca().add_artist(plt.Circle((0,0),0.70,fc='white'))
```

İç içe geçmiş bir beyaz daire ekliyoruz.

```
plt.title("Dairesel Grafik Başlığı")
```

Başlık eklemek için kullanıyoruz.

```
plt.show()
```

Grafiği göstermek için kullanıyoruz.

Kodlar

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

veri = pd.read_csv("cost_of_living.csv", names=["Sıralama", "Ülke", "Yaşam Maliyeti Endeksi", "Kira Endeksi", "Yaşam Maliyeti + Kira Endeksi", "Marketler Endeksi", "Restoran Fiyat Endeksi", "Yerel Satın Alma Gücü Endeksi"], header=0)

print(veri.columns)

def soru1(data_frame):
    sorted_data = data_frame.sort_values("Yaşam Maliyeti Endeksi", ascending=True)
    top10 = sorted_data.head(10)
    colors = sns.color_palette("viridis", len(top10))
    plt.figure(figsize=(8, 8))
    plt.pie(top10["Yaşam Maliyeti Endeksi"], labels=top10["Ülke"], colors=colors, autopct='%1.1f%%', startangle=90)
    plt.gca().add_artist(plt.Circle((0,0),0.70,fc='white'))
    plt.title("En Düşük Toplam Yaşam Maliyetine Sahip 10 Ülke")
    plt.show()

def soru2(data_frame):
    sorted_data = data_frame.sort_values("Yaşam Maliyeti Endeksi", ascending=False)
    bottom10 = sorted_data.head(10)
    colors = sns.color_palette("magma", len(bottom10))
    plt.figure(figsize=(8, 8))
    plt.pie(bottom10["Yaşam Maliyeti Endeksi"], labels=bottom10["Ülke"], colors=colors, autopct='%1.1f%%', startangle=90)
    plt.gca().add_artist(plt.Circle((0,0),0.70,fc='white'))
    plt.title("En Yüksek Toplam Yaşam Maliyetine Sahip 10 Ülke")
    plt.show()
```

```
def soru3(data_frame):  
    komut = data_frame[data_frame["Yerel Satın Alma Gücü Endeksi"] > 40]  
    plt.figure(figsize=(15, 8))  
    sns.barplot(x="Ülke", y="Yerel Satın Alma Gücü Endeksi", data=komut,  
palette="coolwarm")  
    plt.title("Yerel Satın Alma Güçleri 40'tan Yüksek Olan Ülkeler")  
    plt.xticks(rotation=90)  
    plt.ylabel("Yerel Satın Alma Gücü Endeksi")  
    plt.xlabel("Ülkeler")  
    plt.show()
```

```
def soru4(data_frame):  
    plt.figure(figsize=(15, 8))  
    sns.scatterplot(x="Kira Endeksi", y="Yaşam Maliyeti Endeksi", data=data_frame,  
hue="Ülke", palette="Set2")  
    plt.title("Kiralama Endeksine Göre Ülkelerin Karşılaştırılması")  
    plt.ylabel("Yaşam Maliyeti Endeksi")  
    plt.xlabel("Ülkeler")  
    plt.xlabel("Kira Endeks Oranı")  
    plt.show()
```

```
def soru5(data_frame):  
    italy_data = data_frame[data_frame["Ülke"] == "Italy"]  
    categories = ["Yaşam Maliyeti Endeksi", "Kira Endeksi", "Yaşam Maliyeti + Kira Endeksi",  
"Marketler Endeksi", "Restoran Fiyat Endeksi", "Yerel Satın Alma Gücü Endeksi"]  
    values = italy_data[categories].values.flatten()  
    explode = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1)  
    plt.figure(figsize=(10, 8))  
    plt.pie(values, labels=categories, autopct='%1.1f%%', startangle=140,  
colors=plt.cm.Paired.colors, explode=explode)  
    plt.title("İtalya'nın Alt Kategorilerdeki Endeks Karşılaştırması")  
    plt.xlabel("Katagoriler")  
    plt.show()
```

#Sığmadığından dolayı sadece ilk csv dosyasındaki ilk 40 veriyi baz alarak yapıldı

```
def soru6(data_frame):
```

```
    plt.figure(figsize=(25, 8))

    sns.barplot(x="Ülke", y="Restoran Fiyat Endeksi",
data=data_frame.sort_values("Restoran Fiyat Endeksi", ascending=False).head(40),
palette="YlOrRd")

    plt.title("Ülkelerin Restoran Fiyat Endeksine Göre Sıralaması")

    plt.xlabel("Ülkeler")

    plt.xticks(rotation=45, ha="right")

    plt.ylabel("Restoran Fiyat Endeksi")

    plt.show()
```

```
def soru7(data_frame):
```

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

    sns.scatterplot(x="Kira Endeksi", y="Yaşam Maliyeti Endeksi", data=data_frame,
hue="Ülke", palette="viridis")

    plt.title("Ülkelerin Ortalama Kiralama Endeksi ve Yaşam Maliyeti İlişkisi")

    plt.xlabel("Ülkeler")

    plt.ylabel("Yaşam Maliyeti Endeksi")

    plt.xlabel("Kira Endeksi Oranı")

    plt.show()
```

```
def soru8(data_frame):
```

```
    kira_endeksi_sıralı = data_frame.sort_values("Kira Endeksi", ascending=True)

    en_düşük_10_kira = kira_endeksi_sıralı.head(14)

    explode = [0.1] * len(en_düşük_10_kira)

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

    plt.pie(en_düşük_10_kira["Kira Endeksi"], labels=en_düşük_10_kira["Ülke"],
autopct='%1.1f%%', startangle=90, colors=sns.color_palette("coolwarm"), explode=explode)

    plt.xlabel("Ülkeler")

    plt.xticks(rotation=30, ha="right")

    plt.title("En Düşük Kiralama Endeksine Sahip 14 Ülke")

    plt.show()
```

```
def soru9(data_frame):  
    kira_endeksi_siralı = data_frame.sort_values("Kira Endeksi", ascending=True)  
    en_yüksek_10_kira = kira_endeksi_siralı.tail(10)  
    plt.figure(figsize=(12, 6))  
    sns.scatterplot(x="Ülke", y="Kira Endeksi", data=en_yüksek_10_kira, color="skyblue")  
    plt.title("En Yüksek Kiralama Endeksine Sahip Olan 10 Ülke")  
    plt.xlabel("Ülkeler")  
    plt.ylabel("Kira Endeksi")  
    plt.xticks(rotation=30, ha="right")  
    plt.show()
```

```
def soru10(data_frame):  
    yerel_alma_gücü_siralı = data_frame.sort_values("Yerel Satın Alma Gücü Endeksi",  
    ascending=False)  
    en_yüksek_10_alma_gücü = yerel_alma_gücü_siralı.head(10)  
    plt.figure(figsize=(12, 6))  
    sns.barplot(x="Ülke", y="Yerel Satın Alma Gücü Endeksi",  
    data=en_yüksek_10_alma_gücü, palette="YlGnBu")  
    plt.title("Yüksek Yerel Satın Alma Gücüne Sahip Olan 10 Ülke")  
    plt.xlabel("Ülkeler")  
    plt.ylabel("Yerel Satın Alma Gücü Endeksi")  
    plt.xlabel("Ülkeler")  
    plt.show()
```

```
def soru11(data_frame):  
    yerel_alma_gücü_sıralı = data_frame.sort_values("Yerel Satın Alma Gücü Endeksi",  
ascending=False)  
    en_düşük_5_alma_gücü = yerel_alma_gücü_sıralı.tail(14)  
    plt.figure(figsize=(12, 6))  
    sns.barplot(x="Ülke", y="Yerel Satın Alma Gücü Endeksi", data=en_düşük_5_alma_gücü,  
palette="Oranges")  
    plt.title("Düşük Yerel Satın Alma Gücüne Sahip Olan 14 Ülke")  
    plt.xlabel("Ülkeler")  
    plt.ylabel("Yerel Satın Alma Gücü Endeksi")  
    plt.xticks(rotation=30, ha="right")  
    plt.show()
```

#Sığmadığından dolayı sadece ilk csv dosyasındaki ilk 40 veriyi baz alarak yapıldı

```
def soru12(data_frame):  
    plt.figure(figsize=(15, 8))  
    pivot_table = data_frame.pivot_table(index="Ülke", values="Yaşam Maliyeti + Kira  
Endeksi")  
    sns.heatmap(pivot_table.head(40), cmap="YlGnBu", annot=True, fmt=".2f", linewidths=.5)  
    plt.title("Ülkelerin Genel Maliyet İndeksine Göre Renkli Haritası")  
    plt.xlabel("Ülkeler")  
    plt.show()
```

```
def soru13(data_frame):  
    kira_endeksi_filtrelenmiş_güncel = data_frame[(data_frame["Kira Endeksi"] >= 6) &  
(data_frame["Kira Endeksi"] <= 8)]  
    plt.figure(figsize=(12, 6))  
    sns.lineplot(x="Ülke", y="Kira Endeksi", data=kira_endeksi_filtrelenmiş_güncel,  
marker="o")  
    plt.title("Kira Endeksi 6-8 Arasındaki Ülkeler")  
    plt.xlabel("Ülkeler")  
    plt.xticks(rotation=45)  
    plt.ylabel("Kira Endeksi")  
    plt.show()
```

