**HACETTEPE ÜNİVERSİTESİ**

taslak, çizim, kırpıntı çizim, tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu

**IST347 – İSTATİSTİKSEL ANALİZ UYGULAMALARI**

**FİNAL ÖDEVİ**

**Hazırlayanlar:**

**Hande Nur BANUŞ**

**Ders Sorumlusu:** **Dr.OnurToka**

**8 ARALIK 2024**

**2. Gerekli paketler**

**Kullandığımız paketler:**

**Readr paketi**, özellikle metin tabanlı veri dosyalarını (CSV, TSV gibi) okuma, düşük bellek kullanımı ve performans optimizasyonu gibi konularda geliştirilmiştir. readr paketi, veri analizi projelerinde veri içe aktarma aşamasını daha etkili ve hızlı hale getirmek amacıyla tasarlanmıştır.

**Ggplot2 paketi,** veri görselleştirmeiçin kullanılır.ggiraph paketi, veri görselleştirme, makine öğrenmesi, sinyal işleme, yapay zeka, veri madenciliği ve daha birçok alanda kullanılabilir.

**Corrplot paket**i, korelasyon matrislerini görselleştirmek için kullanılır. Corrplot paketi, bu matrisleri renkli grafiklere ve görsel özelliklerle görselleştirmeyi sağlar. Bu sayede veri setindeki değişkenler arasındaki ilişkileri daha iyi anlayabiliriz.

**Dplyr paketi**, R programlama dilinde veri manipülasyonu ve veri analizi için kullanılan bir pakettir. Veri çerçeveleri üzerinde veri manipülasyonu ve filtreleme işlemleri için tasarlanmıştır. Bu paket, filter (), select(), mutate(), group\_by() gibi fonksiyonlar aracılığıyla veri setlerini hızlı ve anlaşılır bir şekilde işlemenizi sağlar.

**Tidyr paketi**, veri setlerini düzenleme ve temizleme işlemleri için kullanılır. Bu paket, özellikle veri çerçevelerindeki geniş formatlı veriyi daha uzun formatlı bir hale getirme veya tam tersini yapma (veri normalleştirme) gibi işlemleri gerçekleştirmek için tasarlanmıştır.

**Xts paketi**, zaman serileri için veri yapılarını ve işlevlerini sağlayan bir pakettir. Bu paket zaman serisi verilerini depolamak, alt kümelere ayırmak, birleştirmek, dönüştürmek ve analiz etmek için çeşitli işlevler sunar.

summarise() fonksiyonu ile veri setlerini özetleyebilir ve arrange() ile sıralama işlemleri gerçekleştirebilirsiniz.

join() fonksiyonları, farklı veri çerçevelerini birleştirme işlemleri için kullanılır.

gather() ve spread() gibi fonksiyonlar, geniş formatlı veriyi uzun formatlıya ve uzun formatlı veriyi geniş formatlıya dönüştürmek için sıklıkla kullanılır.

Veri setindeki eksik veya boş değerleri düzenleme ve temizleme işlemlerinde drop\_na() gibi fonksiyonlar kullanılabilir.

**3.1. Veriyi açıklayınız. Değişken sayısı, gözlem sayısı, değişkenlerin yapısı hakkında bilgiler veriniz.**

**Değişken Sayısı**: Veri setinizde toplam 35 değişken bulunmaktadır.

**Gözlem Sayısı**: Billionaires veri setimizden rasgele 1000 veri çekiyoruz. Yani gözlem sayımız 1000’dir. Oluşturduğumuz yeni veri setine veri\_f ismini verdik.

**Değişkenlerin Yapısı**: Veri setimizde nicel, kategorik ve mantıksal (logical) türlerde değişkenler bulunuyor.

**3.2. Veri yükleme ve düzenlemesi konusunda bilgi veriniz.**

Öncelikle Billioners veri setimizi import dataset kısmından çektik. Daha sonra rasgele 1000 verimizi çektik ve bu verimizi set.seed(777) olarak sabitledik. Bu sayede kodu her çalıştırdığımızda aynı veri setini verecek.

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

**3.3. Verideki tüm değişkenleri özetleyip, yorumlayınız.**

***NİCEL VERİLER:***

**rank**: Milyarderin servet açısından sıralaması

**finalWorth**: Milyarder serveti

**age**: Yaş

**birthYear**: Doğum yılı

**birthMonth**: Doğum ayı

**birthDay**: Doğum günü

**cpi\_country**: TÜFE

**cpi\_change\_country**: Milyarderin ülkesi için TÜFE değişimi

**gdp\_country**: Milyarderin ülkesi için Gayri Safi Yurtiçi Hasıla (GSYİH)

**gross\_tertiary\_education\_enrollment**: Milyarderin ülkesindeki yükseköğretime kayıt

**gross\_primary\_education\_enrollment\_country**: Milyarderin ülkesinde ilköğretime kayıt

**life\_expectancy\_country**: Milyarderin ülkesindeki yaşam beklentisi

**tax\_revenue\_country\_country**: Milyarderin ülkesindeki vergi geliri

**total\_tax\_rate\_country**: Milyarderin ülkesindeki toplam vergi oranı

**population\_country**: Milyarderin ülkesinin nüfusu

**latitude\_country**: Milyarderin ülkesinin enlem koordinatı

**longitude\_country**: Milyarderin ülkesinin boylam koordinatı

***KATEGORİK VERİLER:***

**category**: Milyarderin işinin faaliyet gösterdiği kategori veya sektör

**personName:** Milyarderin tam adı

**country:** Milyarderin ikamet ettiği ülke

**city:** Milyarderin ikamet ettiği şehir

**source:** Milyarderin servetinin kaynağı

**industries:** Milyarderin ticari çıkarlarıyla ilişkili endüstriler

**countryOfCitizenship:** Milyarderin vatandaşı olduğu ülke

**organization:** Milyarderle ilişkili kuruluş veya şirketin adı

**status:** "D" kendi kendini yetiştirmiş milyarderleri (Kurucular/Girişimciler) temsil eder ve "U" miras alınan veya devralınan kişileri belirtir

**gender:** Cinsiyet

**birthDate:** Doğum tarihi

**lastName:** Milyarderin soyadı

**firstName:** Milyarderin ilk adı

**title:** Milyarderin unvanı

**date:** Veri toplanma tarihi

**state:** Milyarderin ikamet ettiği eyalet

**residenceStateRegion:** Milyarderin ikamet ettiği bölge veya eyalet

***MANTIKSAL (LOGİCAL) VERİ:***

**selfMade**: Milyarderin kendi kendini yetiştirmiş olup olmadığını belirtir (Doğru/Yanlış).

**4.1.Servet dağılım analizi: Milyarderlerin servetinin farklı sektörlere, ülkelere ve bölgelere dağılımını keşfedin.**

Veri görselleştirmesi için R studioda ggplot2 paketi kullanılır. Paketi indirdikten sonra library(ggplot2) ile paketi kullanıma hazır hale getiririz.

Geom\_col ile sütun grafiği çizdirdik. Show.legend=FALSE ile lejantları göstermesini istemedik.

Coord\_flip() ile eksenlerin yerini değiştirdik. Böylece grafik daha okunaklı oldu.

Theme\_minimal() ile arka plan temasını seçtik.

Sektör:

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, ekran görüntüsü, diyagram, renklilik içeren bir resim

Açıklama otomatik olarak oluşturuldu

Çıktıya göre milyarder servetinin en fazla olduğu sektör finance & ınvestments (finans ve yatırım). Yani en çok bu sektörde iş yapan kişiler daha zengin olmuş denebilir. En az ise gambling and casinos (kumar ve kumarhaneler) olduğunu görüyoruz.

**Ülke:**

Billionaries veri setinde “county” sütunun adını 4.5 ‘deki soruda sorun yaşamamız nedeniyle “region” olarak değiştirdik.

**metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**metin, ekran görüntüsü, sayı, numara, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu**

Çıktıya göre milyarderlerin servetinin en fazla olduğu United States. Yani Amerika Birleşik Devletleri’nde yaşayan milyarderlerin kazandığı servet daha fazla.

**Bölgelere göre:**

**metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**metin, ekran görüntüsü, ekran, görüntüleme, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu**

Çıktıya göre güney, batı ve kuzeydoğu bölgesinde yaşayan milyarderlerin serveti orta batıda yaşayan milyarderlerin servetinden daha fazla.

**4.2 Demografik analiz: Milyarderlerin yaşını, cinsiyetini ve doğum yeri demografisini araştırın.**

Bu soru için öncelikle ggplot2 paketimizi yüklüyoruz.

metin, yazı tipi, el yazısı, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

a) Yaş histogramını çizelim:



metin, ekran görüntüsü, diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu

* Bu histogram grafiği, milyarderlerin yaş dağılımını gösteriyor. Bu histogramda, 50-70 yaş arası milyarderlerin diğer yaş gruplarına göre daha fazla olduğu görülüyor.
* Bu histogramda, verilerin ortalama değeri yaklaşık 60 yaş civarındadır. Medyan değer ise, verilerin ortasında kalan değerdir. Bu histogramda, medyan değeri 60-70 yaş aralığındadır. Mod değeri ise, verilerin en çok görüldüğü değeridir. Bu histogramda, mod değeri 50-60 yaş aralığındadır.
* Yayılımı ölçmek için standart sapma, varyans veya aralık gibi ölçüler kullanılabilir. Bu histogramda, verilerin aralığı, en büyük değerden en küçük değeri çıkartarak bulunur. Aralık = 100-20 = 80’dir. Bu, verilerin 80 yaşlık bir aralıkta dağıldığını gösterir. Standart sapma ve varyans ise, verilerin ortalama değerden ne kadar uzaklaştığını gösterir. Bu histogramda, standart sapma ve varyans değerleri hesaplanabilir, ancak görsel olarak da verilerin ortalama değerden ne kadar yayıldığını görmek mümkündür. Verilerin ortalama değerden uzaklaştıkça, sütunların yüksekliği azalır. Bu, verilerin ortalama değere yakın olduğunu, yani düşük bir standart sapma ve varyans değerine sahip olduğunu gösterir.
* Bu, verilerin sağa çarpık olduğunu gösterir. Sağa çarpık dağılımlarda, ortalama değer medyan değerden büyüktür. Bu histogramda, ortalama değer yaklaşık 60, medyan değer ise 50-60 yaş aralığındadır. Bu, ortalama değerin medyan değerden büyük olduğunu doğrular.

b) Cinsiyet dağılımını çizelim:

metin, yazılım, multimedya yazılımı, ekran görüntüsü içeren bir resim

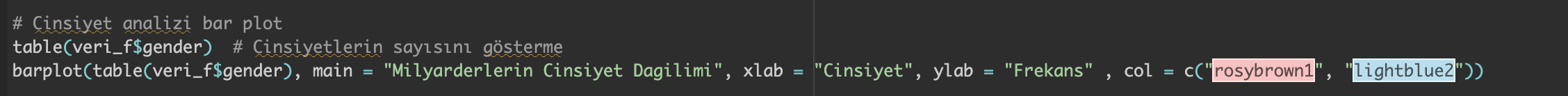
Açıklama otomatik olarak oluşturuldu

metin, diyagram, ekran görüntüsü, daire içeren bir resim

Açıklama otomatik olarak oluşturuldu

Bu pasta grafiği, milyarderlerin cinsiyet dağılımını gösteriyor. Bu pasta grafiğinde, milyarderlerin büyük çoğunluğunun erkek olduğu görüyoruz. Erkeklerin oranı %87.3 , kadınların oranı ise %12.7’dir.

Cinsiyet dağılımını bar grafiği ile de çizebiliriz:



metin, yazı tipi, ekran görüntüsü, siyah içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, ekran görüntüsü, diyagram, dikdörtgen içeren bir resim

Açıklama otomatik olarak oluşturuldu

Buna göre milyarderlerin büyük çoğunluğu erkeklerden oluşuyor. Erkeklerin sayısı 873, kadınların sayısı ise 127’dir.

b) Doğum yeri dağılımını çizelim:

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

ekran görüntüsü, metin, diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu

Bu grafiğe göre en fazla milyarderin bulunduğu ülkenin United States olduğunu söyleyebiliriz. Burada çok fazla ülke görülmekte, aşağıdaki kod ile 66 ülkenin her birindeki milyarder sayısına bakalım:

Burada düzenleme için dplyr paketini kullandık.

metin, yazı tipi, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, ekran görüntüsü, menü içeren bir resim

Açıklama otomatik olarak oluşturuldu metin, menü, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

Burada çok fazla ülke bulunmakta ve bazı ülkelerdeki milyarder sayısı 1, 2, 3 gibi düşük değerlerdir. Bu ülkelerden nüfusu fazla olan ilk 15 ülkeyi alalım ve buna top\_15\_country ismini vereli, kalan ülkelere “Diger Ulkeler” diyelim. Bu şekilde grafiğimiz daha sade olacaktır.

metin, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, yazı tipi, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, ekran görüntüsü, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, diyagram içeren bir resim

Açıklama otomatik olarak oluşturuldu

Buna göre; en fazla milyarderin bulunduğu ülke Amerika’dır (United States). En fazla milyarder bulunan 2. Ülke Çin (China), en fazla milyarder bulunduran 3. Ülke ise Hindistan (India) olarak yorum yapabiliriz.

**4.3.Kendi kendine yapılan ve miras alınan zenginlik: Kendi kendine milyarder olanların ve onların servetini miras alan kişilerin oranını analiz edin.**

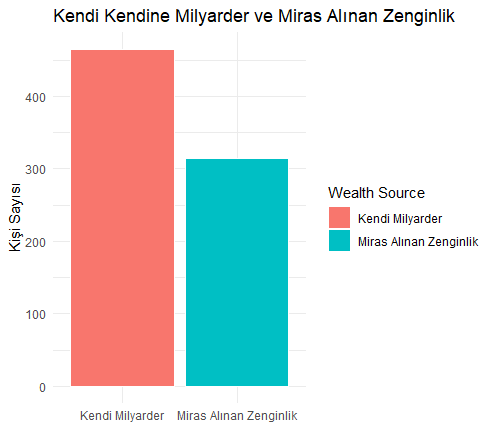
veri3\_soru adında bir veri çerçevesi oluşturduk. Bu çerçeve, "category" adında bir sınıf ve "count" adında bir sayım içeriyor. "category" sütunu "Kendi Milyarder" ve "Miras Alınan Zenginlik" olarak iki kategori içerirken, "count" sütunu bu kategorilere ait kişi sayılarını içerir. veri\_f verisinden "status" sütunundaki değerlere göre kişi sayılarını hesaplayarak oluşturuluyor. Status sütunuda D kendi kendine edinilen zenginlik, U miras alınan zenginliği ifade ediyor.

geom\_bar() fonksiyonu, çubuk grafiğini oluşturmak için kullanılır. "stat = "identity"" parametresi, veri çerçevesindeki "count" sütununun direkt olarak çubuk yüksekliklerine karşılık gelmesini sağlar. Veri3\_soru değişkeninden category ve count değişkenlerini aes içine atadık.

**metin, yazılım, yazı tipi, multimedya yazılımı içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**Bar grafiği:**

****

Kendi kendine milyarder olan kişi sayısı daha fazladır.

**Pasta grafiği:**

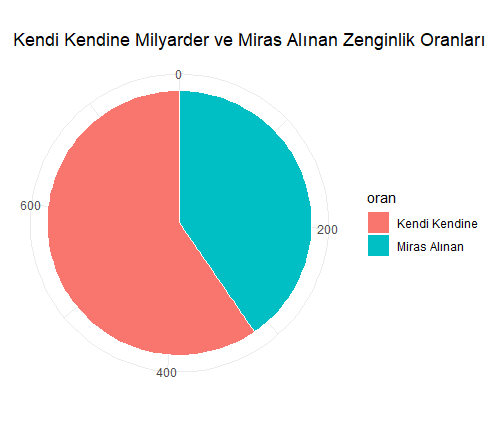
**metin, ekran görüntüsü, yazılım, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu**

İlk olarak, "Kendi kendine Milyarder(D)" ve "Miras Alınan Zenginlik(U)" kategorilerine ait kişi sayılarını toplanıp , **veri\_yeni** adında yeni bir veri çerçevesi oluşturduk.

**coord\_polar("y")** fonksiyonu, polar koordinatlarda bir grafik oluşturulmasını sağlar, bu da pasta grafiğini oluşturur. Yani bar grafiğinden pasta grafiğine dönüşüyor.

**theme(legend.position = "right")** ise grafikteki lejantın sağ üst köşeye yerleştirilmesini sağlar.

****

**metin, yazı tipi, ekran görüntüsü, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu**

Kendi kendine milyarder olan kişi oranı 45.6 iken miras ile milyarder olan kişi oranı 34.3’tür.

**4.4. Ekonomik göstergeler: Milyarder serveti ile GSYİH, TÜFE ve vergi oranları gibi ekonomik göstergeler arasındaki korelasyonları inceleyin.**



metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Burada korelasyon matrisinde NA değerleri çıkıyor. Bunun için NA değerlerini çıkartmalı ve gdp\_country sütunundaki dolar işareti ve virgülü kaldırmalıyız.

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

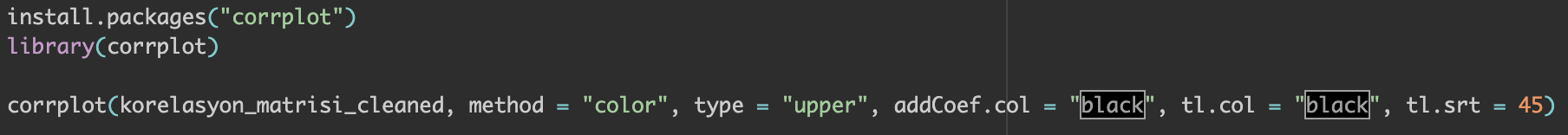


Korelasyon\_matrisi\_cleaned çıktısı şu şekildedir:

metin, yazı tipi, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

Artık matrisimizde NA değerleri bulunmuyor. Şimdi bu korelasyon matrisimizi görselleştirelim. Bunun için öncelikle corrplot paketini yüklememiz gerekiyor. Bu paketi korelasyon matrisimizi görselleştirmek için kullanacağız. Daha sonra kodumuzu yazıyoruz.



metin, ekran görüntüsü, diyagram, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Bu çıktı, finalWorth (milyarder serveti), gdp\_country (GSYİH), cpi\_country (TÜFE) ve total\_tax\_rate\_country (vergi oranı) arasındaki korelasyonları gösteriyor. Korelasyon katsayısı, -1 ile 1 arasında değişen bir değerdir.

* 1’e yaklaşan pozitif bir korelasyon, iki değişken arasında olan güçlü bir pozitif ilişki olduğunu gösterir. Pozitif korelasyonda iki değişken arasında aynı yönlü ilişki vardır.
* -1’e yaklaşan negatif bir korelasyon, iki değişken arasında olan güçlü bir negatif ilişki olduğunu gösterir. Negatif korelasyonda iki değişken arasında ters yönlü ilişki vardır.
* 0’a yaklaşan bir değer, iki değişken arasında ilişki olmadığını gösterir.

Burada her bir değişkenin kendisiyle olan korelasyon katsayısı 1’dir.

***Milyarder serveti(finalWorth)*** ile ***GSYİH (gdp\_country)*** arasındaki korelasyon -0.02’dir. Yani, iki değişken arasında çok zayıf bir negatif ilişki vardır.

***Milyarder serveti (finalWorth)*** ile ***TÜFE (cpi\_country)*** arasındaki korelasyon -0.02’dir. Yani, iki değişken arasında çok zayıf bir negatif ilişki vardır.

***Milyarder serveti (finalWorth)*** ile ***vergi oranı (total\_tax\_rate\_country)*** arasındaki korelasyon -0.08’dir. Yani, iki değişken arasında çok zayıf bir negatif ilişki vardır.

***GSYİH (gdp\_country)*** ile ***TÜFE (cpi\_country)*** arasındaki korelasyon 0’dır. Yani, iki değişken arasındaki ilişki çok zayıftır, yoktur diyebiliriz.

***GSYİH (gdp\_country)*** ile ***vergi oranı (total\_tax\_rate\_country)*** arasındaki korelasyon 0.11’dir. Yani, iki değişken arasında çok zayıf bir pozitif ilişki olduğunu söyleyebiliriz.

***TÜFE (cpi\_country)*** ile ***vergi oranı (total\_tax\_rate\_country)*** arasındaki korelasyon 0.26’dır. Yani, iki değişken arasında orta güçlükte pozitif ilişki olduğunu söyleyebiliriz.

**4.5.Mekansal analiz: Milyarderlerin coğrafi dağılımını ve servetlerini bir harita üzerinde görselleştirin.**

**1.yol:**

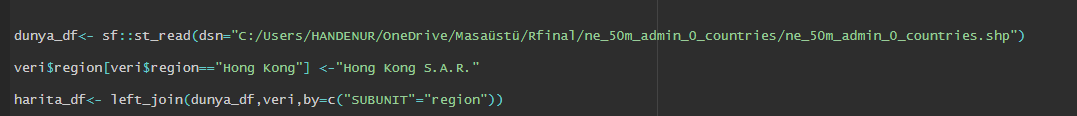
Gerekli kütüphaneleri yükledik. Veri\_f veri setinin sadece iki sütunuyla (finalworth ve region) ilgilendiğimiz için o ikisini seçip “veri” adlı değişkene atadık.

**1)**

**metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**2)**

****

**Rvest** paketi, web sitesinden bilgiyi almak için yani web sitesinden bilgiyi “scrape(kazımak)” ediyoruz.

**Magrittr** paketi, veri manipülasyonunu daha okunabilir ve zincirleme tarzında (chaining) yapmayı sağlayan bir pakettir. Bu paket kodun daha anlaşılır hale gelmesini ve işlemlerin sıralı bir şekilde gerçekleştirilmesini sağlar. Burada gerekli düzenleme yapmak için kullandık.

**Tidyverse** paketini bilgileri filtreleme yapmak için kullandık.

Sf paketi yani shape file dediğimiz haritalarda şekil dosyalarını oluşturması için işlem gördüğümüz kütüphanedir.

**tmap** paketi, tematik haritalar oluşturmak ve görselleştirmek için kullanılan bir pakettir. Bu paket, harita oluşturma sürecini kolaylaştırmak ve çeşitli özelleştirme seçenekleri sunmak amacıyla geliştirilmiştir. Tmap paketi, tematik haritalar oluşturmak için **tmap** ve **tm** adlı iki ana sınıf içermektedir.

**dunya\_df<sf::st\_read(dsn="C:/Users/HANDENUR/OneDrive/Masaüstü/Rfinal/ne\_50m\_admin\_0\_countries/ne\_50m\_admin\_0\_countries.shp"):** sf paketinden st\_read isimli fonksiyonu ile coğrafi verileri çeşitli formatlardan okumak için çağırdık. Dsn, veri kaynağının adresini belirtir. Oluşturulan bu veriyi dunya\_df değişkene atadık.

Bu dunya\_df verisinde ülkelerin isimleri, ölçüleri, konumları, sınırları gibi ülkelerle alakalı coğrafik verileri gösterir.

**veri$region[veri$region=="Hong Kong"] <-"Hong Kong S.A.R." :** burada “veri” isimli verimiz ile “dunya\_df “ verisindeki ülke isimlerini (subunit sütunu) ile karşılaştırdım ve aynı olmayanları düzenleyip aynı isim ile gösterilmesini sağladık. Yani “veri “deki Hong Kong ismi yerine Hong Kong S.A.R. yazdırdık.

**harita\_df<- left\_join(dunya\_df,veri,by=c("SUBUNIT"="region")):**  **dplyr** paketinin **left\_join** fonksiyonunu kullanarak iki veri çerçevesini birleştirme işlemi gerçekleştirir. İki veri çerçevesi arasında birleştirme işlemi, **SUBUNIT** sütunundaki değerlerle **region** sütunundaki değerlerin eşleştirilmesine dayanır. Yani dünya veri çerçevesinden (dunya\_df) hepsini al, “veri” den ise eşleşenleri al ve birleştir. Ama SUBUNIT=region’ a göre eşleştir.

**3)**

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

**tmap\_style("white")**: Bu kod satırı, harita stili için kullanılacak temayı belirler. Haritanın arka plan rengini beyaz olarak ayarlar.

Veri haritamızı oluşturmak istediğimizde, tm\_shape fonksiyonu ile daha önce tanımladığımız harita\_df’yi aldık. Böylece tm kütüphanesi bana bir veri haritası oluştuşturacak. Enlem ve boylamlarını yerleştirmesi gerekiyor.

Tm\_fill , içini “finalworth” ile dolduracak. Stil, renk paleti,başlık gibi farklı parametreler kullandık. **style = "quantile**.” değerlerin yüzdelik dilimlerine göre renklendirme yapar. **legend.is.portrait = TRUE**: Bu, harita üzerindeki renk skalasının dikey (portrait) olarak görüntülenmesini sağlar.

**4)**

metin, ekran görüntüsü, yazılım, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

* **main.title**: Haritanın başlığını belirtir.
* **main.title.position**: Başlığın pozisyonunu belirtir. "center" değeri, başlığın haritanın ortasında yer almasını sağlar.
* **main.title.size**: Başlık font boyutunu belirtir.
* **legend.height** ve **legend.width**: Bu değerler, renk skalasının boyutunu kontrol eder.
* **legend.outside**: Renk skalasının harita dışında olup olmamasını belirtir. **FALSE** değeri, renk skalasının harita içinde yer alacağını gösterir.
* **legend.position**: Renk skalasının konumunu belirtir. Bu durumda, sol alt köşede (**"left", "bottom"**) yer alacaktır.
* **legend.frame**:Renk skalasının etrafındaki çerçevenin kullanılmasını istediğimiz için **TRUE** yazdık.
* **legend.bg.color** ve **legend.bg.alpha**: Renk skalası arka planının rengi ve şeffaflığı belirtilir.
* **bg.color**: Haritanın genel arka plan rengini belirtir. Bu durumda, "skyblue" olarak ayarlanmıştır.
* **inner.margins**: Harita içerisindeki kenar boşluklarını belirtir. Bu değerler, sırasıyla sol, üst, sağ ve alt kenarlardaki boşlukları temsil eder.

Bütün bu değerler tek tek denenerek özelleştirilmiş ve son hali paylaşılmıştır.

**5)**

metin, yazı tipi, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

**tm\_borders** fonksiyonu, harita üzerindeki ülke sınırlarını çizmek için kullanılır. **alpha** parametresi, sınırların şeffaflığını belirler. Burada **alpha = 0.5** seçerek sınırların yarı şeffaf bir şekilde görüntülenmesini sağladık.

**tm\_text("ISO\_A3", size = "AREA")**: Bu satır, harita üzerine ülke etiketlerini ekler. **tm\_text** fonksiyonu, "ISO\_A3" sütunundaki değerleri harita üzerine yazdırmak için kullanılır. **size = "AREA"** parametresi, etiketlerin boyutunu, harita üzerindeki alanları temsil eden bir özellik olan "AREA" sütununa göre belirler. Yani, büyük ülkelerin etiketleri daha büyük, küçük ülkelerin etiketleri daha küçük olacaktır.

**6)**



**tm\_grid(alpha = 0.2)**: harita üzerine ızgara eklemek için kullanılır. **alpha = 0.2** seçerek, ızgaranın şeffaflığını belirledik.Burada ızgara, haritanın altında hafif şeffaflıkla görünür olacaktır.

Tm\_compass ile de hairtaya istediğimiz şekil ve formatta pusula ekleriz.

* **type = "8star"**: Pusula tipini belirtir. Burada "8star", sekiz yıldızlı bir pusula kullanılacağını gösterir.
* **position = c(.65, .15)**: Pusulanın konumunu belirtir. İki elemanlı bir vektör kullanılarak pusulanın x ve y koordinatları belirlenir. Burada, pusula sağ üst köşede (**x = 0.65, y = 0.15**) görünecektir.
* **size = 3**: Pusula boyutunu belirtir.
* **color.light = "grey90"**: Pusula üzerindeki ışıkların rengini belirtir. Burada gri tonları kullanılır.

**7)**

metin, yazı tipi, ekran görüntüsü, tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu

Haritayı plot (sabit bir çizim) şeklinde ya da interaktif şekilde göstermemizi sağlayan kod satırlarıdır.

Plot modunu çalıştıralım:



metin, harita, atlas, diyagram içeren bir resim

Açıklama otomatik olarak oluşturuldu

İnteraktif hale getirelim:



metin, harita, diyagram, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

Dilediğimiz gibi yakınlaşabilir, istediğimiz ülkeyi inceleyebiliriz.

metin, ekran görüntüsü, harita, diyagram içeren bir resim

Açıklama otomatik olarak oluşturuldu

**4.6. Zaman içindeki eğilimler: Milyarder demografisindeki ve zenginlikteki yıllar içindeki değişiklikleri izleyin.**

Burada öncelikle xts paketini yüklüyoruz. Bu paket zaman serileri için veri yapılarını ve işlevlerini sağlayan bir pakettir.

Milyarderlerin zenginliğinin zaman içindeki değişimi:

metin, yazı tipi, ekran görüntüsü, grafik içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, ekran görüntüsü, yazılım, multimedya yazılımı içeren bir resim

Açıklama otomatik olarak oluşturuldu

veri\_f veri çerçevesinde bulunan "birthYear" sütunundaki eksik değerleri kontrol ettik ve bunları missing\_years değişkenine atadık.

İf-else kısmında eğer eksik değer varsa eksik olmayan değerler arasındaki min ve max yılları belirliyoruz. Eğer eksik değer yoksa tüm yıllar arasındaki min ve max yılları belirliyoruz.

ts() fonksiyonu veri\_f deki “finalWorth” sütununu zaman serisine dönüştürür.

Start ve end parametreleri, zaman serisinin başlangıç ve bitiş tarihlerini belirler.

frequency parametresi ise zaman serisinin frekansını belirler. (Bir yılda 12 ay bulunduğu için 12 olarak ayarladık)

Son olarak plot() fonksiyonu ile zaman serisi grafiğimizi çizdiriyoruz. Col parametresi ile de grafiğin rengini belirliyoruz.

metin, ekran görüntüsü, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Bu grafiğe göre; yaklaşık 1950 yılında bir artış görülmektedir. Daha sonra 2000 yılına kadar geçen sürede zenginlik seviyesi neredeyse stabil şekilde ilerlemiştir. Fakat 1990-2000 yılları arasında zenginlikte büyük bir artış vardır. Bu yıllar arasında zenginliği etkileyen bir faktör olduğu düşünülebilir, örneğin milyarder sayısındaki artış zenginliğin artışını etkilemiştir. Bu yıllardaki artışın ardından zenginlikte büyük bir düşüş görülmektedir.

Milyarderlerin cinsiyet dağılımının zaman içindeki değişimi:

metin, ekran görüntüsü, yazı tipi içeren bir resim

Açıklama otomatik olarak oluşturuldu

ekran görüntüsü, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, diyagram, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

Milyarderlerin yıllara göre cinsiyet dağılımı grafiği verilmiştir. Mavi barlar erkek milyarderleri, pembe barlar ise kadın milyarderleri göstermektedir. Bu grafiğe göre erkek milyarder sayısının kadın milyarder sayısından fazla olduğunu söyleyebiliriz. 1962 yılında erkek ve kadın milyarder sayısında büyük bir artış görüyoruz. 1930-1960 yılları arasında erkek milyarder sayısı önemli ölçüde artmıştır. Kadın milyarder sayısı tüm yıllarda erkek milyarder sayısından azdır.

**Kaynakça:**

* <https://campus.datacamp.com/courses/>
* <https://www.data-to-viz.com/>
* <https://bookdown.org/ugurdar/rileverigorsellestirme/ggplot2-k%C3%BCt%C3%BCphanesi.html>