

World Happiness Report Data Visualization - Global AI Hub

Tuğçe Elmacı

VERİ GÖRSELLEŞTİRME WORLD HAPPINESS REPORT DATA

Dünya Mutluluk Raporu, küresel mutluluğun durumuna ilişkin önemli bir araştırmadır. Ekonomi, psikoloji, anket analizi, ulusal istatistik, sağlık, kamu politikası ve daha fazlası gibi alanlarda önde gelen uzmanlar, ulusların ilerlemesini değerlendirmek için refah ölçümlerinin nasıl etkili bir şekilde kullanılabileceğini anlatıyor. Raporlar, günümüz dünyasında mutluluğun durumunu gözden geçirmekte ve yeni mutluluk biliminin mutluluktaki kişisel ve ulusal farklılıkları nasıl açıkladığını göstermektedir.

Mutluluk puanları ve sıralamaları, Gallup Dünya Anketinden alınan verileri kullanır. Mutluluk puanını izleyen sütunlar, ekonomik üretim, sosyal destek, yaşam beklentisi, özgürlük, yolsuzluğun olmaması ve cömertlik olmak üzere altı faktörün her birinin, her ülkede yaşam değerlendirmelerini Distopya'dakinden daha yüksek hale getirmeye ne ölçüde katkıda bulunduğunu tahmin ediyor.

BÖLÜM 1) VERİ OKUTMA VE ÖN İŞLEME

#KÜTÜPHANELER

```
library(corrplot)
library(stringr)
library(ggplot2)
library(dplyr)
library(arm)
library(ggrepel)
library(RColorBrewer)
library(GGally)
library(tidyr)
library(egg)
library(ggExtra)
library(Hmisc)
```

Değişken Adı	Açıklama
Country name	Ülke Adı
Year	Yıl
Life Ladder	Yaşam Merdiveni
Log GDP per capita	Kişi başına düşen GSYİH
Social support	Sosyal Destek
Healthy life expectancy at birth	Doğumda Sağlıklı Yaşam Beklentisi

Freedom to make life choices	Yaşam Seçimlerini Yapma Özgürlüğü
Generosity	Cömertlik
Perceptions of corruption	Yolsuzluk Algıları
Positive affect	Olumlu Etki
Negative affect	Olumsuz Etki

- **Gözlem sayısı:** 1949
- **Değişken sayısı:** 11

Kaynak; World Happiness Report- Kaggle

1.1 VERİ SETİNİN ÖZETİ

summary(data)

```
## Country_name          year      Life_Ladder  Log_GDP_per_capita
## Length:1949          2017   : 147   Min.      :2.375   Min.      : 6.635
## Class :character     2011   : 146   1st Qu.:4.640   1st Qu.: 8.464
## Mode  :character     2014   : 145   Median :5.386   Median : 9.460
##                               2019   : 144   Mean    :5.467   Mean    : 9.368
##                               2015   : 143   3rd Qu.:6.283   3rd Qu.:10.353
##                               2012   : 142   Max.    :8.019   Max.    :11.648
##                               (Other):1082   NA's    :36
## Social_support        Healthy_life_expectancy_at_birth Freedom_to_make_life_ch
oices
## Min.      :0.2900    Min.      :32.30          Min.      :0.2580
## 1st Qu.:0.7498    1st Qu.:58.69          1st Qu.:0.6470
## Median :0.8355    Median :65.20          Median :0.7630
## Mean    :0.8126    Mean    :63.36          Mean    :0.7426
## 3rd Qu.:0.9050    3rd Qu.:68.59          3rd Qu.:0.8560
## Max.    :0.9870    Max.    :77.10          Max.    :0.9850
## NA's    :13        NA's     :55            NA's     :32
## Generosity          Perceptions_of_corruption Positive_affect  Negative_affe
ct
## Min.      :-0.3350    Min.      :0.0350          Min.      :0.3220    Min.      :0.08
30
## 1st Qu.: -0.1130    1st Qu.:0.6900          1st Qu.:0.6255    1st Qu.:0.20
60
## Median : -0.0255    Median :0.8020          Median :0.7220    Median :0.25
80
## Mean    : 0.0001    Mean    :0.7471          Mean    :0.7100    Mean    :0.26
85
## 3rd Qu.: 0.0910    3rd Qu.:0.8720          3rd Qu.:0.7990    3rd Qu.:0.32
00
## Max.    : 0.6980    Max.    :0.9830          Max.    :0.9440    Max.    :0.70
50
## NA's     :89        NA's     :110            NA's     :22        NA's     :16
```

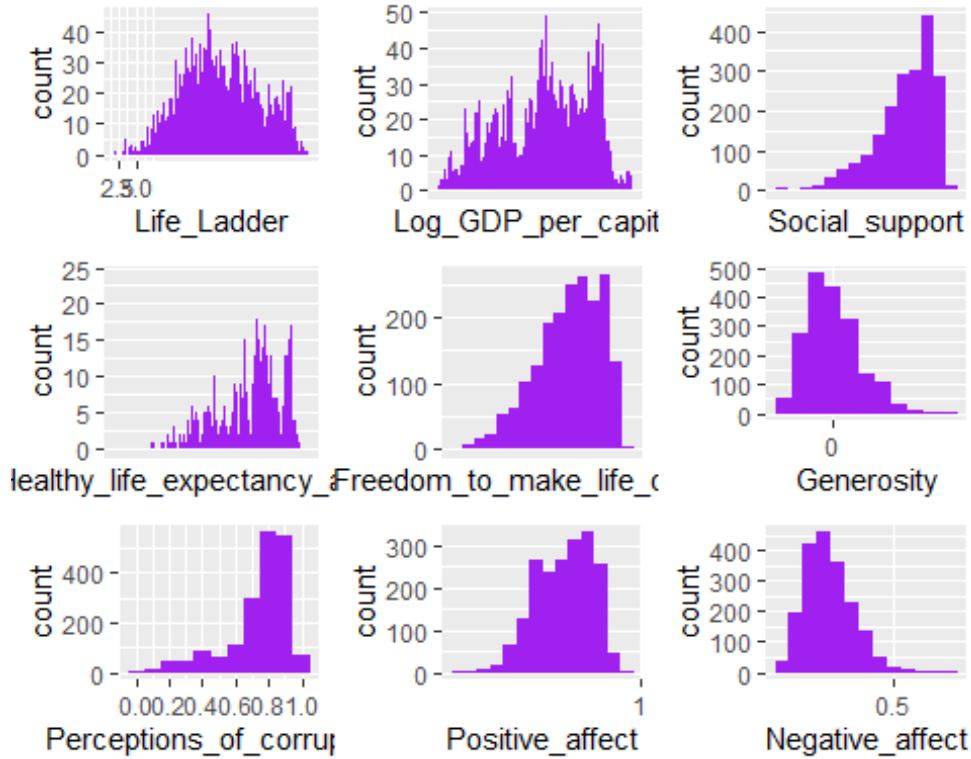
Özeti inceleyelim. Veri setimiz 11 değişken ve 1949 gözleminden oluşmaktadır. Counrty_name değişkeni karakter, year değişkeni de faktör olarak tanımlanmıştır, onun

dışında kalan tüm değişkenler sayısalıdır. 8 değişkende toplam 373 eksik gözlemler bulunmaktadır.

BÖLÜM 2) VERİ GÖRSELLEŞTİRME

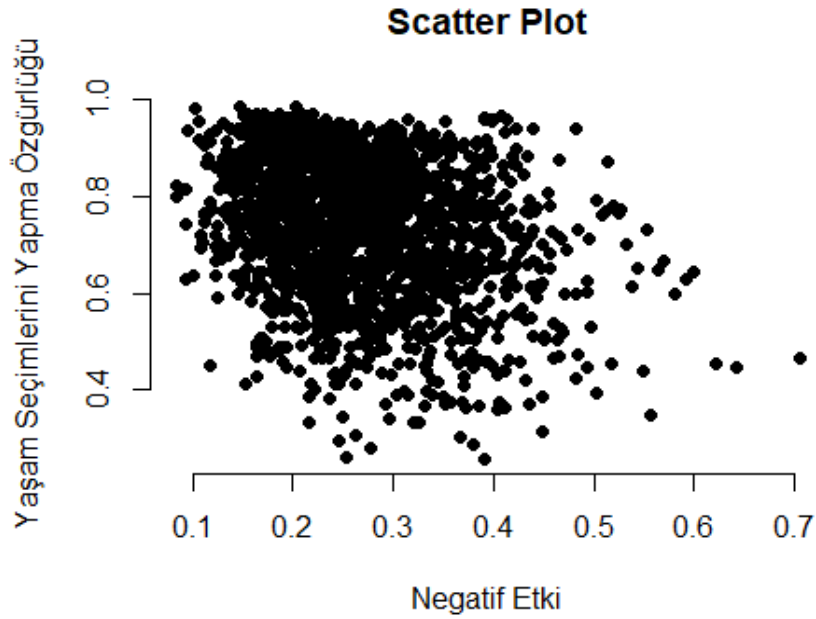
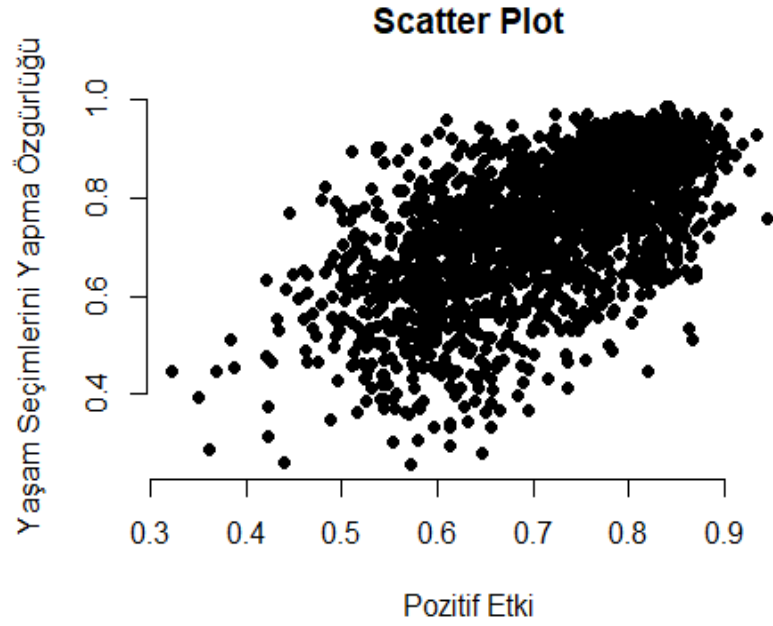
2.1 HISTOGRAM GRAFİĞİ

Aşağıdaki toplu grafikte veri setimizdeki tüm sayısal değişkenlerin histogram grafikleri vardır. Histogramlar, değerlerin nerede yoğunlaştığını, uç noktaların neler olduğunu ve veri kümesinde herhangi bir boşluk veya olağandışı değerler olup olmadığı konusunda bir tahmin verir. Bu grafiklere bakarak normal dağılıma sahip olup olmadıkları hakkında bilgi sahibi olabiliriz.



2.2 SAÇILIM GRAFİĞİ (SCATTER PLOT)

Bir saçılım grafiği, iki sayısal değişken arasındaki ilişkiyi dağınık noktalar şeklinde çizmek için kullanılır. Grafikteki her nokta tek bir gözlemi temsil edecek şekilde gösterilir.

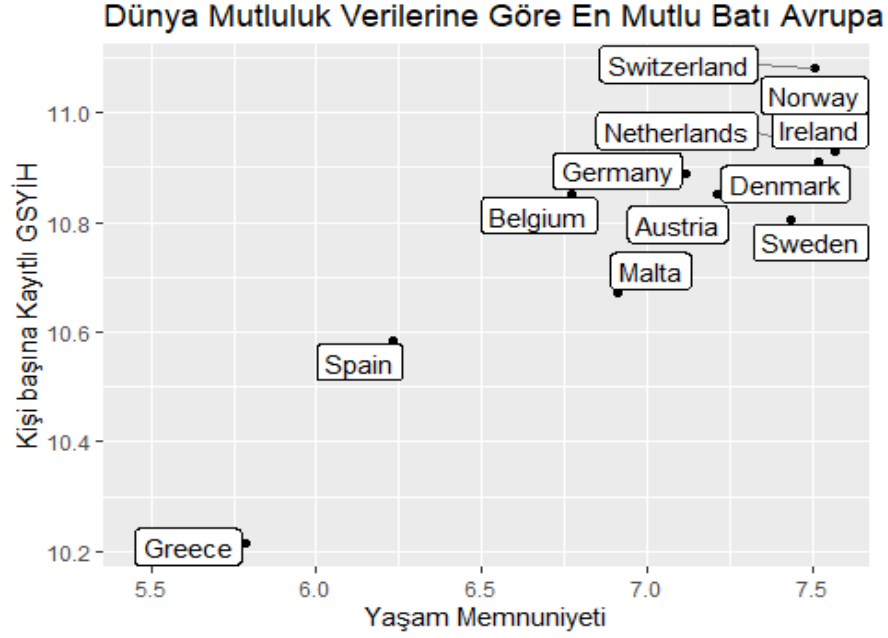


İki farklı scatter plot çizdirdik. Bu grafiklere bakarak Yaşam Seçimlerini Yapma Özgürlüğünün (Freedom_to_make_life_choices) Pozitif veya Negatif bir etki oluşturup oluşturmadığını inceleyebiliyoruz. Yaşam Seçimlerini Yapma Özgürlüğü arttıkça pozitif etki artmıştır, negatif etki azalmıştır. Bu da insanların mutluluklarında yaşamdaki seçimlerini özgürce yapmanın büyük bir etkisi olduğunu gösterir.

Saçılım grafiğini farklı bir alanda kullanarak En Mutlu Ülkelere bakalım. Bunun için Yaşam memnuniyeti (Life_Ladder) değişkeninin ortalama değerini kullanarak **“Life_Ladder > 5”**

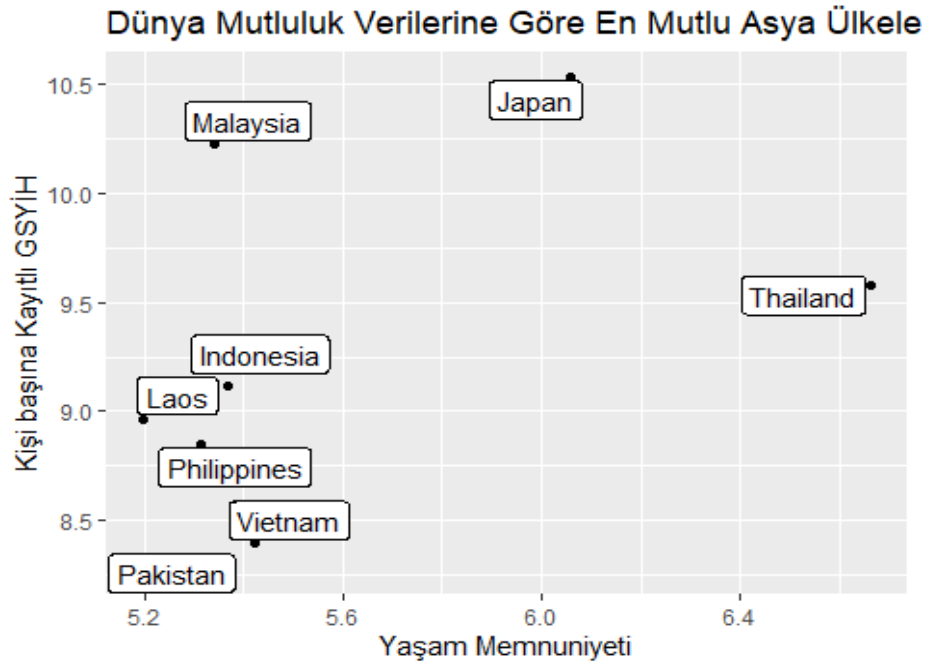
olacak şekilde filtreliyoruz. Bu şekilde yaşam memnuniyeti 5 puanının üzerinde çıkan ülkeler mutlu ülkeler olacaktır.

a) En Mutlu Batı Avrupa Ülkeleri



Baktığımızda 21 tane Batı Avrupa Ülkesinin 12 tanesi yani %57'si mutlu çıkmıştır. Yunanistan en az mutlu, İrlanda ise en mutlu ülke çıkmıştır.

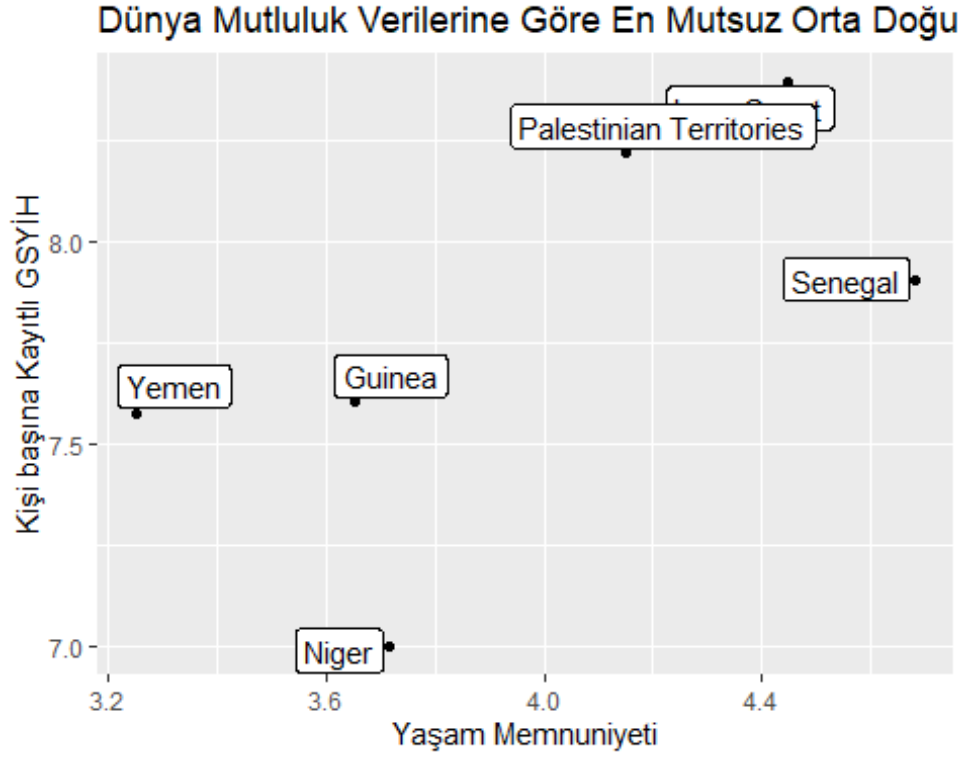
b) En Mutlu Batı Asya Ülkeleri



Baktığımızda 22 tane Asya Ülkesinin 8 tanesi yani %36'si mutlu çıkmıştır. Laos en az mutlu, Tayland ise en mutlu ülke çıkmıştır.

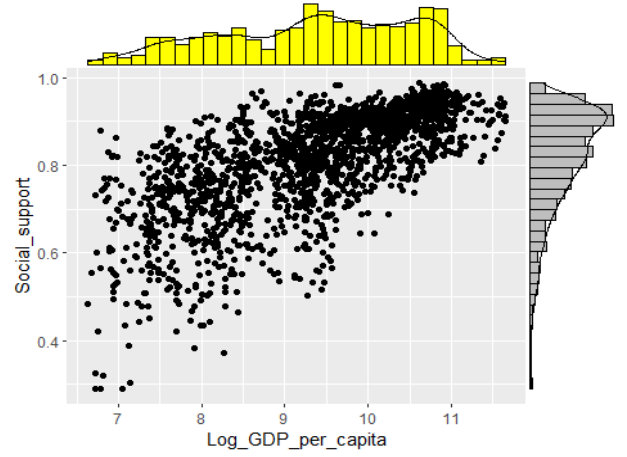
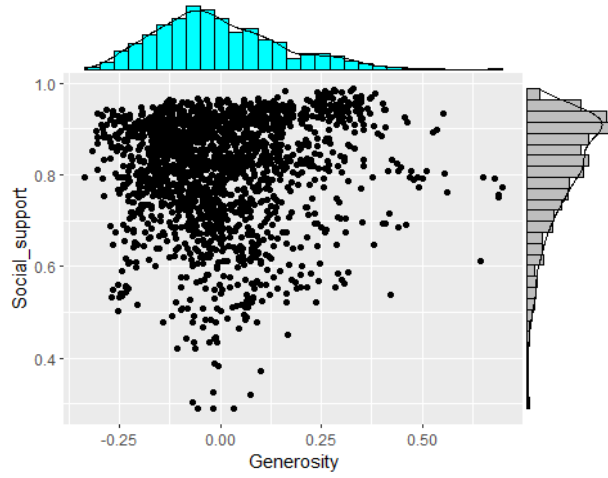
Bir de En Mutsuz Ülkelere bakalım. Bu sefer Yaşam memnuniyeti(Life_Ladder) değişkeninin ortalama değerini kullanarak "**Life_Ladder < 5**" olacak şekilde filtreliyoruz. Bu şekilde yaşam memnuniyeti 5 puanın altında çıkan ülkeler mutsuz ülkeler olacaktır.

c) En Mutsuz Batı Orta Doğu ve Afrika Ülkeleri



Baktığımızda 27 tane Asya Ülkesinin 6 tanesi yani %22'si mutsuz çıkmıştır. Yemen en mutsuz, Senegal ise Yemen'e göre daha az mutsuz bir ülke çıkmıştır.

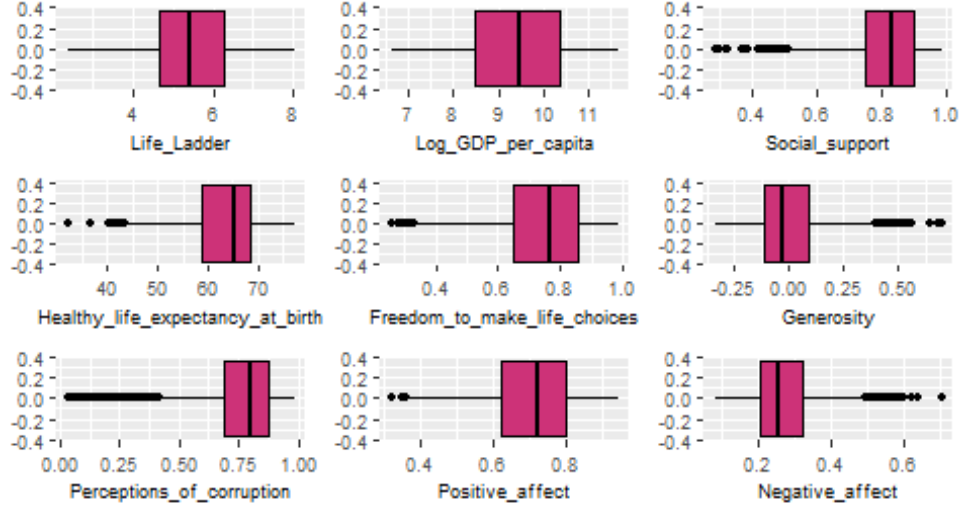
Hem histogram hemde saçılım grafiğini öğrendiğimize göre bu ikisini birleştirerek daha farklı bir grafiksel gösterim elde edebiliriz.



2.3 KUTU GRAFIĞİ (BOX-PLOT)

Bir kutu grafiği, veri çeyreklerini, yüzdelikleri, ortalamaları görüntüleyerek sayısal gözlemlerin görsel olarak dağılımını göstermek için kullanılır. Aynı zamanda aykırı değerlerin tespiti için de kullanılan bir grafiksel yöntemdir.

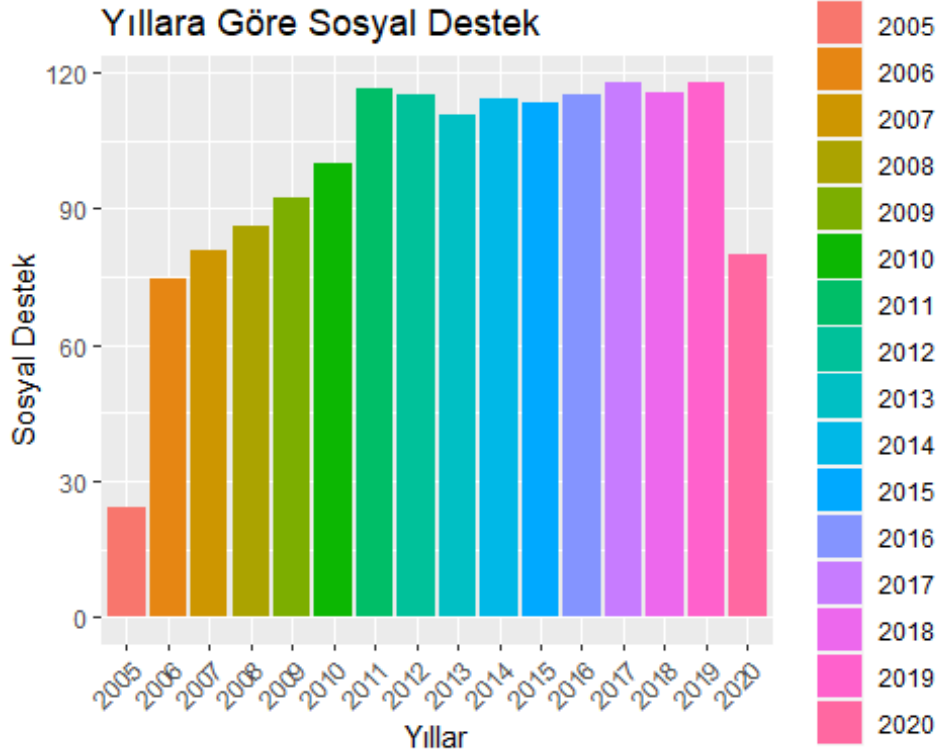
Kutu grafikleri, bir veri kümesinin beş özelliğini gösterir; * Minimum değer * İlk çeyreklik * Medyan * Üçüncü çeyreklik * Maksimum değer



Çizdirdiğimiz kutu grafiklerini incelediğiniz zaman yukarıda bahsettiğimiz gibi hem çeyreklikler, hem medyan hem de aykırı gözlemler hakkında fikir sahibi olabiliyoruz.

2.4 ÇUBUK GRAFIĞİ (BAR-PLOT)

Çubuk Grafiği, yatay veya düşey sütunları kullanarak kategoriler arası ayırım yapmaya ve sayısal karşılaştırmaları göstermeye yarar. Grafiğin bir ekseni, karşılaştırılan kategorileri, diğer ekseni ise bir değer ölçeğini gösterir.



Grafikte yıllara göre sosyal desteğin değişimini görüyoruz. 2011 yılına kadar bir yükseliş varken daha sonra 2019 yılına kadar bir duraklama yaşanmış. Daha sonra 2020 yılında net bir düşüş görünmektedir. Bu sosyal destek azalmasının sebebi 2020 yılında başlayan covid-19 pandemisi olabilir.