



Hacettepe Üniversitesi

2022-2023Güz

İst347- İstatistiksel Analiz Uygulamaları

Final Ödevi

Öğrenci : Sedanur Gülten

Öğrenci Numara: 21935965

## İÇİNDEKİLER

Veri seti ve Veri Seti Hakkında Bilgi.....	3
Çalışma Özeti.....	4
“Dplyr “ Paketi ile Veri Manipülasyonu .....	5-11
“DataExplorer “Paketi Veri Analizi .....	11-26
“Esquisse” Paketi ile Veri Görselleştirme.....	26-36
Kaynakça.....	37

## Veri Seti ve Veri Seti Hakkında Bilgi

Veri setimiz müzik terapisinin, yani dinlenen müzik türünün, bireyin genel zihinsel durumuna etkisini araştırmak için kullanılır. Müzik ve mental sağlık veri seti, bir bireyin müzik zevki ile kendisinin bildirdiği ruh sağlığı arasında varsa, hangi ilişkilerin var olduğunu belirlemeyi amaçlar. İdeal olarak, bu bulgular müzik terapisinin daha bilinçli bir uygulamasına katkıda bulunabilir veya sadece zihin hakkında ilginç manzaralar sunabilir.

- Toplamda 33 değişken ve 736 gözlemden oluşmuştur.
- Veri setinde kategorik veri bulunmaktadır.

1-**Timestamp:** Formun gönderildiği tarih ve saat

2-**Age:** Formu yanıtlayanların yaşı

3- **Primary streaming service:** Birincil dinlenme servis yeri : [Spotify](#)-[Apple Music](#)-[Youtube Müzik](#)- servis kullanmıyor

4-**Hours Per Day :** Katılımcıların günlük müzik dinleme saat sayısı (1-24 saat)

5-**While Working:** Katılımcı ders çalışırken/ çalışırken müzik dinliyor mu? Evet/hayır

6- **Instrumentalist:** Katılımcı düzenli olarak enstrüman çalışıyor mu? [Evet](#)/[hayır](#)

7-**Composer:** Katılımcı müzik besteliyor mu? [Evet](#)/[hayır](#)

8-**Fav genre:** Yanıtlayanın en favori müzik türü?

9- **Exploratory:** Katılımcı aktif olarak yeni sanatçıları/türleri keşfediyor mu? [Evet](#)/[hayır](#)

10-**BPM:** Favori türün vuruş sayısı (0-1000m aralığında)

11- **Frequency [Classical]:** Katılımcı ne sıklıkta klasik müzik dinliyor? [Nadiren](#), [ara sıra](#), [hiç](#), [çok sık](#)

12- **Frequency [Country]:** Katılımcı ne sıklıkta klasik country dinliyor?

13 **Frequency [EDM]:** Katılımcı ne sıklıkta EDM dinliyor?

14-**Frequency [Folk]:** Katılımcı ne sıklıkta halk müziği dinliyor?

15-**Frequency [Gospel]:**Katılımcı ne sıklıkla gopsel müziği dinliyor?

16- **Frequency [Hip hop]:** Katılımcı ne sıklıkla hip-hop müziği dinliyor?

17 -**Frequency [Jazz]:** Katılımcı ne sıklıkla caz müziği dinliyor?

18- **Frequency [K pop]:**Katılımcı ne sıklıkla kpop müziği dinliyor?

19-**Frequency [Latin]:** Katılımcı ne sıklıkla latin müziği dinliyor?

20-**Frequency [Lofi]:** Katılımcı ne sıklıkla lofi müziği dinliyor?

21- **Frequency [Metal]:** Katılımcı ne sıklıkla metal müziği dinliyor?

22- **Frequency [Pop]:** Katılımcı ne sıklıkla pop müziği dinliyor?

23 -**Frequency [R&B]:** Katılımcı ne sıklıkla r&b müziği dinliyor?

24-**Frequency [Rap]:** Katılımcı ne sıklıkla rap müziği dinliyor?

25-**Frequency [Rock]:** Katılımcı ne sıklıkla lofi müziği dinliyor?

26-**Frequency [Video game music]:** Katılımcı ne sıklıkla oyun müziği dinliyor?

27-**Anxiety:** 0-10 aralığında katılımcıların kendilerine verilen kaygı puanı

28-**Depression:** 0-10 aralığında katılımcıların kendilerine verilen depresyon puanı

29-**Insomnia:** 0-10 aralığında katılımcıların kendilerine verilen uykusuzluk puanı

30-**OCD:** 0-10 aralığında katılımcıların kendilerine verilen OKB puanı

31-**Music effects:** Müzik, yanıtlayanın ruh sağlığı koşullarını, iyileştirir mi/kötüleştirir mi? geliştirir, daha kötü yapar , etkisi yok

32-**Permissions:** verileri yayınlama izinleri, verilen okudum, [anladım/ anlamadım](#) cevabı.

## Çalışma Özeti

Kullanılacak paketler ve fonksiyonların açıklamaları

Veri manipülasyonu ile veriyi hazır hale getirme (dplyr paketi kullanılarak)

Veriyi analiz etme (DataExplorer)

Veri görselleştirme (esquisse)

## Dplyr Paketi

**dplyr** paketi, gruplar arası farklar, değişken gruplamaları, yeni değişkenler oluşturma ve benzeri işlemleri gerçekleştiremeye yarayan R paketlerinden biridir. “Beş büyükler” olarak nitelendirebileceğimiz fonksiyonları sayesinde dplyr paketi verinin düzgün bir şekilde manipülasyonu olanaklı kılar. Tidyverse evreninin bir parçasıdır.

- **Select():** Bu fonksiyon sütunlar/değişkenler arasından istediklerimizi seçer.
- **Filter():** Bu fonksiyon satırlar/gözlemler arasından istediklerimizi seçer.
- **Arrange():** Değişkenlere göre sütunları organize eder
- **Mutate():** Veri tabanındaki sütun/değişkenleri kullanarak aynı veri tabanına yeni sütun/değişken ekler .
- **Summarise()** :İstenilen kriterlere göre gözlemleri gruplar.

**Group\_by():**Veri çerçevesindeki satırları sütunlara göre gruplandırmak için kullanılır. Veri çerçevesi verilerinin nasıl görüldüğünü değiştirmez ve yalnızca özet yapabileceğimiz gruplandırılmış tbl'yi (tibble tablosu) döndürür.Tüm veri çerçevesini kullanmak yerine artık grupları kullanırız.

**%%:**Bir fonksiyonun çıktısını alır ve bunu argüman olarak başka bir fonksiyona iletir. Bu, bir dizi analiz adımı arasında bağlantı kurmamızı sağlar.

**Subset():** Veri çerçevesinden satırları ve sütunları ( gözlemler ve değişkenler) almak için kullanılan genel bir R işlevidir.

## Veri Manipölasyonum

### Veri yükleme

```
library(readxl)
music_m <- read_excel("C:/Users/sedad/Desktop/music_m.xlsx")
view(music_m)
```

“dplyr” paketini R programına indirip, kullanmak için kütüphaneden çağırıyoruz.

```
install.packages("dplyr")
library(dplyr)
```

**glimpse():** Bir veri çerçevesindeki her sütunu görmeyi mümkün kılar. Mümkün olduğu kadar çok veri göstermeye çalışır.

Veri setimizin, nelerden oluştuğunu, özelliklerini görebilmek için

```
glimpse(music_m)
#> # A tibble: 736 x 33
#>   Timestamp      Age Primarystreamingservice Hoursperday whileworking Instrumentalist Composer Favgenre Exploratory Foreignlanguages BPM
#>   <chr>      <dbl> <chr>                                <dbl> <chr>      <chr>      <chr>      <chr>      <chr>      <dbl>
#> 1 "8/27/2022 19:29:02" 18, 63, 18, 61, 18, 18, 18, 21, 19, 18, 18, 19, NA, 19, 18, 17~
#>   Spotify", "Pandora", "Spotify", "YouTube Music", "Spotify", "~
#>   3.0, 1.5, 4.0, 2.5, 4.0, 5.0, 3.0, 1.0, 6.0, 1.0, 3.0, 8.0, 3.~
#>   "Yes", "Yes", "No", "Yes", "Yes", "Yes", "Yes", "Yes", "Yes", ~
#>   "Yes", "No", "No", "No", "No", "Yes", "Yes", "No", "No", "No", ~
#>   "Yes", "No", "No", "Yes", "No", "Yes", "No", "No", "No", "No", ~
#>   "Latin", "Rock", "Video game music", "Jazz", "R&B", "Jazz", "V~
#>   "Yes", "Yes", "No", "Yes", "Yes", "Yes", "Yes", "Yes", "No", "~
#>   "Yes", "No", "Yes", "Yes", "No", "Yes", "Yes", "Yes", "No", "Y~
#>   156, 119, 132, 84, 107, 86, 66, 95, 94, 155, NA, 125, NA, 88, ~
#>   "Rarely", "Sometimes", "Never", "Sometimes", "Never", "Rarely"~
#>   "Never", "Never", "Never", "Never", "Never", "Sometimes", "Nev~
#>   "Rarely", "Never", "Very frequently", "Never", "Rarely", "Neve~
#>   "Never", "Rarely", "Never", "Rarely", "Never", "Never", "Somet~
#>   "Never", "Sometimes", "Never", "Sometimes", "Rarely", "Never", ~
#>   "Sometimes", "Rarely", "Rarely", "Never", "Very frequently", "~
#>   "Never", "Very frequently", "Rarely", "Very frequently", "Neve~
#>   "Very frequently", "Rarely", "Very frequently", "Sometimes", "~
#>   "Very frequently", "Sometimes", "Never", "Very frequently", "S~
#>   "Rarely", "Rarely", "Sometimes", "Sometimes", "Sometimes", "Ve~
#>   "Never", "Never", "Sometimes", "Never", "Never", "Rarely", "Ra~
#>   "Very frequently", "Sometimes", "Rarely", "Sometimes", "Someti~
#>   "Sometimes", "Sometimes", "Never", "Sometimes", "Very frequent~
#>   "Very frequently", "Rarely", "Rarely", "Never", "Very frequent~
#>   "Never", "Very frequently", "Rarely", "Never", "Never", "Very ~
#>   "Sometimes", "Rarely", "Very frequently", "Never", "Rarely", "~
#>   3, 7, 7, 9, 7, 8, 4, 5, 2, 2, 7, 1, 9, 2, 6, 7, 8, 5, 7, 7, 6, ~
#>   0, 2, 7, 7, 2, 8, 8, 3, 0, 2, 7, 0, 3, 1, 4, 5, 8, 7, 5, 3, 9, ~
#>   1, 2, 10, 3, 5, 7, 6, 5, 0, 5, 4, 0, 2, 2, 7, 4, 4, 10, 0, 0, ~
#>   0, 1, 2, 3, 9, 7, 0, 3, 0, 1, 7, 1, 7, 0, 0, 1, 3, 0, 3, 2, 0, ~
#>   NA, NA, "No effect", "Improve", "Improve", "Improve", "Improve~
#>   "I understand.", "I understand.", "I understand.", "I understa~
```

Glimpse(music\_m) komutunu kullandık ve bize 736 gözlem, 33 değişkenli bir veri setinin, değişkenleri, değişken türlerini, gözlemlerin nelerden oluştuğunu gösterdi.

Yine veri setindeki isimlere bakmak için, names():

```
> names(music_m)
[1] "Timestamp"      "Age"             "Primarystreamingservice"
[4] "Hoursperday"    "whileworking"    "Instrumentalist"
[7] "Composer"       "Favgenre"        "Exploratory"
[10] "Foreignlanguages" "BPM"             "Frequency[Classical]"
[13] "Frequency[Country]" "Frequency[EDM]"  "Frequency[Folk]"
[16] "Frequency[Gospel]" "Frequency[Hip hop]" "Frequency[Jazz]"
[19] "Frequency[K pop]" "Frequency[Latin]" "Frequency[Lofi]"
[22] "Frequency[Metal]" "Frequency[Pop]"   "Frequency[R&B]"
[25] "Frequency[Rap]" "Frequency[Rock]"  "Frequency[Videogamemusic]"
[28] "Anxiety"        "Depression"      "Insomnia"
[31] "OCD"            "Musiceffects"    "Permissions"
```

Verimizde hangi değişkende, kaç tane kayıp gözlem olduğuna bakalım.-->**colsums(is.na())**

```
> colsums(is.na(music_m))
Timestamp      Age      Primarystreamingservice
0              1              1
Hoursperday    whileworking    Instrumentalist
0              3              4
Composer       Favgenre       Exploratory
1              0              0
Foreignlanguages    BPM      Frequency[Classical]
4              107           0
Frequency[Country]    Frequency[EDM]      Frequency[Folk]
0              0              0
Frequency[Gospel]    Frequency[Hip hop]      Frequency[Jazz]
0              0              0
Frequency[K pop]    Frequency[Latin]      Frequency[Lofi]
0              0              0
Frequency[Metal]    Frequency[Pop]      Frequency[R&B]
0              0              0
Frequency[Rap]    Frequency[Rock]    Frequency[Videogamemusic]
0              0              0
Anxiety           Depression      Insomnia
0              0              0
OCD              Musiceffects    Permissions
0              8              0
```

En çok “BPM” değişkeninin, 107 kayıp gözlemi olduğunu görüyoruz. “filter” komutu ile her bir kayıp veri bulunan değişkeni kullanarak, kayıp verilerin silinmesini sağlarım, burada göstermek amacı ile beş değişkeni kullandım. Beş değişken ile kayıp verilerin olduğu satırları silince 624 gözlem kalıyor.

```
colsums(is.na(music_m))
filter(music_m, !is.na(Age), !is.na(Primarystreamingservice), !is.na(whileworking), !is.na(Instrumentalist), !is.na(BPM))
```

Fakat uzun geldiği için, direkt aşağıdaki kod ile tüm kayıp verili satırları silebilirim.

-->**na.omit()**

```
music_mh <- na.omit(music_m)
music_veri<-data.frame(music_mh)
music_veri
```

Kayıp verilerin silindiği yeni music\_veri isimli veri setini oluşturduk.

“dim” ile kaç gözlem ve değişken olduğuna bakalım.

```
> dim(music_veri)
[1] 616  33
```

616 gözlemli ve 33 değişkenli olduğunu anlayabiliriz.

**Summary():** Bize bazı özet istatistikleri gösterir.

```
> summary(music_veri)
```

Timestamp Length:616 Class :character Mode :character	Age Min. :10.00 1st Qu.:18.00 Median :21.00 Mean :24.79 3rd Qu.:27.00 Max. :89.00	Primarystreamingservice Length:616 Class :character Mode :character	Hoursperday Min. : 0.000 1st Qu.: 2.000 Median : 3.000 Mean : 3.702 3rd Qu.: 5.000 Max. :24.000
whileworking Length:616 Class :character Mode :character	Instrumentalist Length:616 Class :character Mode :character	Composer Length:616 Class :character Mode :character	Favgenre Length:616 Class :character Mode :character
Exploratory Length:616 Class :character Mode :character	Foreignlanguages Length:616 Class :character Mode :character	BPM Min. : 0 1st Qu.: 100 Median : 120 Mean : 1623500 3rd Qu.: 144 Max. :999999999	Frequency.Classical. Length:616 Class :character Mode :character
Frequency.Country. Length:616 Class :character Mode :character	Frequency.EDM. Length:616 Class :character Mode :character	Frequency.Folk. Length:616 Class :character Mode :character	Frequency.Gospel. Length:616 Class :character Mode :character
Frequency.Hip.hop. Length:616 Class :character Mode :character	Frequency.Jazz. Length:616 Class :character Mode :character	Frequency.K.pop. Length:616 Class :character Mode :character	Frequency.Latin. Length:616 Class :character Mode :character
Frequency.Hip.hop. Length:616 Class :character Mode :character	Frequency.Jazz. Length:616 Class :character Mode :character	Frequency.K.pop. Length:616 Class :character Mode :character	Frequency.Latin. Length:616 Class :character Mode :character
Frequency.Lofi. Length:616 Class :character Mode :character	Frequency.Metal Length:616 Class :character Mode :character	Frequency.Pop. Length:616 Class :character Mode :character	Frequency.R.B. Length:616 Class :character Mode :character
Frequency.Rap. Length:616 Class :character Mode :character	Frequency.Rock. Length:616 Class :character Mode :character	Frequency.Videogamemusic. Length:616 Class :character Mode :character	Anxiety Min. : 0.000 1st Qu.: 4.000 Median : 6.000 Mean : 5.884 3rd Qu.: 8.000 Max. :10.000
Depression Min. : 0.000 1st Qu.: 2.000 Median : 5.000 Mean : 4.894 3rd Qu.: 7.000 Max. :10.000 Permissions Length:616 Class :character Mode :character	Insomnia Min. : 0.000 1st Qu.: 1.000 Median : 3.000 Mean : 3.801 3rd Qu.: 6.000 Max. :10.000	OCD Min. : 0.000 1st Qu.: 0.000 Median : 2.000 Mean : 2.659 3rd Qu.: 5.000 Max. :10.000	Musiceffects Length:616 Class :character Mode :character

Sürekli değişkenlerdeki min-max değer, çeyreklikler, medyan, ortalamaları görebiliriz.

Hoursperday- günlük müzik dinleme sürelerinde, kullanıcıların %50'si (0-24 saat aralığında) 3 saat veya daha az müzik dinliyor. Simetrik veriler için ortanca ve ortalamalar birbirine yakındır.

Bu özetlere bakarak veride aşağıdaki değişiklikleri yapıyorum. ( $Q_1 - 1,5IQR < \min, \max < Q_3 + 1,5IQR$ ) ile aykırı olan değişkenlere karar verebiliriz .Max. ve Çeyrek Yaş(Age), Günlük dinleme süresi(Hoursperday) ve BPM aralığı, verilerin aykırı değerlere sahip olduğunu gösterir.

## Subset()

```
music_veri2 <- subset(music_veri, Age <= 60 )
music_veri2
music_veri3<- subset(music_veri2, Hoursperday <= 15)
music_veri3
music_veri4 <- subset(music_veri3, BPM <= 250 )
music_veri4
music_yeni_veri= subset(music_veri4, select = -c(Timestamp,Permissions) )
music_yeni_veri
```

Yaşı 60, günlük dinleme süresi 15, vuruş sayısı 250'ye kadar olan verileri seçip tarih (Timestamp) ve verileri okuma izinleri (Permission) değişkenlerini veri seti dışında bıraktık.

```
grouped_music <- group_by(music_yeni_veri, Primarystreamingservice)
summarise(grouped_music, mean(BPM))
```

Müzik servislerini gruplayıp, summarise() fonksiyonu ile BPM (vuruş sayıları) ortalamalarını hesaplattık.

```
> grouped_music <- group_by(music_yeni_veri, Primarystreamingservice)
> summarise(grouped_music, mean(BPM))
# A tibble: 6 x 2
  Primarystreamingservice    `mean(BPM)`
  <chr>                    <dbl>
1 Apple Music              127.
2 I do not use a streaming service. 121.
3 Other streaming service    126.
4 Pandora                  131.
5 Spotify                  124.
6 YouTube Music            119.
```

Aynı şeyi yaş değişkeni için de yapalım.

```
grouped_music2 <- group_by(music_yeni_veri, Primarystreamingservice)
summarise(grouped_music2, mean(Age))

> grouped_music2 <- group_by(music_yeni_veri, Primarystreamingservice)
> summarise(grouped_music2, mean(Age))
# A tibble: 6 x 2
  Primarystreamingservice    `mean(Age)`
  <chr>                    <dbl>
1 Apple Music              22.6
2 I do not use a streaming service. 27.3
3 Other streaming service    27.7
4 Pandora                  37.2
5 Spotify                  22.2
6 YouTube Music            27.0
```

Tablo hali için:

```
sumg_mus2<-summarise(grouped_music2, mean(Age))
data.frame(sumg_mus2)
view(sumg_mus2)
```



	Primarystreamingservice	mean(Age)
1	Apple Music	22.65000
2	I do not use a streaming service.	27.30769
3	Other streaming service	27.65789
4	Pandora	37.20000
5	Spotify	22.24936
6	YouTube Music	26.95775

Pandora servisini kullananların yaşları diğerlerine göre daha yüksek, Spotify ve Apple Music kullananların yaş ortalamaları neredeyse aynı.

Bu müzik servislerini kullananların, hangi platformu kaç kullanıcının kullandığına bakalım.

```
kullanici_platform<- music_yeni_veri%>%
  filter(Primarystreamingservice != "") %>%
  group_by(Primarystreamingservice)%>%
  summarize(kullanıcı = n())%>%
  arrange(desc(kullanıcı))
view(kullanici_platform)
```

	Primarystreamingservice	kullanıcı
1	Spotify	369
2	YouTube Music	71
3	I do not use a streaming service.	52
4	Apple Music	40
5	Other streaming service	38
6	Pandora	5

Kullanıcılar en fazla Spotify servisini kullanıyorlar. Spotify'ı ,Youtube Music takip ediyor.

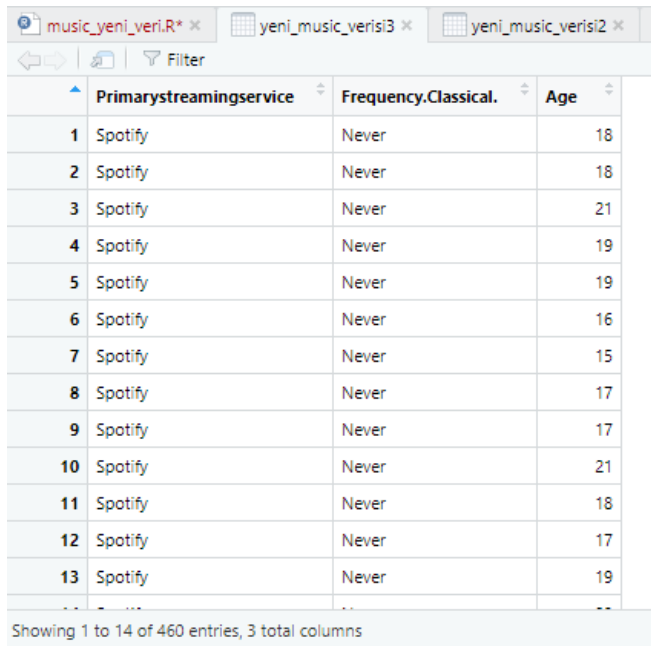
```
yeni_music_verisi2 <- music_yeni_veri %>%
  select(Primarystreamingservice, BPM, Age) %>%
  filter(Primarystreamingservice %in% c("Spotify", "YouTube Music")) %>%
  arrange(Primarystreamingservice)
view(yeni_music_verisi2)
```

	Primarystreamingservice	BPM	Age
1	Spotify	132	18
2	Spotify	107	18
3	Spotify	86	18
4	Spotify	95	21
5	Spotify	94	19
6	Spotify	88	19
7	Spotify	148	18
8	Spotify	103	16
9	Spotify	120	16
10	Spotify	120	15
11	Spotify	125	17
12	Spotify	79	18
13	Spotify	169	18
14	Spotify	136	14
15	Spotify	126	17
16	Spotify	120	19
17	Spotify	143	17

Değişkenler arasından müzik servisleri, yaş ve vuruş sayılarını seçtik. Müzik servisleri değişkeninin içinden Youtube Music ve Spotify'ı seçtik. Müzik servislerini sıralamasını istedik.

```
 yeni_music_verisi3 <- music_yeni_veri %>%  
  select(Primarystreamingservice, Frequency.Classical. , Age) %>%  
  filter(Primarystreamingservice %in% c("Spotify", "YouTube Music")) %>%  
  arrange(Primarystreamingservice,Frequency.Classical.)  
 data.frame(yeni_music_verisi3)  
 view(yeni_music_verisi3)|
```

Yukarıda yapılan işlemi, klasik müzik dinleme sıklığını ekleyerek, BPM değişkeni olmadan yaptık.



	Primarystreamingservice	Frequency.Classical.	Age
1	Spotify	Never	18
2	Spotify	Never	18
3	Spotify	Never	21
4	Spotify	Never	19
5	Spotify	Never	19
6	Spotify	Never	16
7	Spotify	Never	15
8	Spotify	Never	17
9	Spotify	Never	17
10	Spotify	Never	21
11	Spotify	Never	18
12	Spotify	Never	17
13	Spotify	Never	19

Showing 1 to 14 of 460 entries, 3 total columns

```
jazz_music<- music_yeni_veri$Frequency.Jazz.  
Jazz_table<-table(jazz_music)  
view(Jazz_table)
```

Caz müzik dinleyenlerin, dinleme sıklıklarına göre, dinleyen kullanıcı sayısının olduğu bir tablo çıktısı

	jazz_music	Freq
1	Never	204
2	Rarely	208
3	Sometimes	139
4	Very frequently	44

Caz müziğini hiç dinlemeyenler, sık dinleyenlerden daha fazla.

```
jazz_music_bin<- jazz_music[jazz_music == 'never']
jazz_music_bin<- append(jazz_music_bin , jazz_music[jazz_music == 'Very frequently'])
table(jazz_music_bin)
view(table(jazz_music_bin))

binom.test(x=44 , n=248 , p=0.5)
```

Burada değişkenleri, çok sık ve hiç halinde iki kategori haline getirip, binom testini uyguladım.

	jazz_music_bin	Freq
1	Never	204
2	Very frequently	44

$H_0$ : Hiç kullanmayanlar ile sık kullananlar arasındaki oran 0.5'e eşittir.

$H_s$ : Hiç kullanmayanlar ile sık kullananlar arasındaki oran 0.5'e değildir.

```
> binom.test(x=44 , n=248 , p=0.5)

Exact binomial test

data: 44 and 248
number of successes = 44, number of trials = 248, p-value < 2.2e-16
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.1319720 0.2307619
sample estimates:
probability of success
 0.1774194
```

P value:  $0.00 < 0.05$   $H_0$  ret

Caz müziğini hiç dinlemeyenler ile çok sık dinleyenlerin arasındaki oranın 0.5'ten farklı olduğu %95 güven düzeyinde söylenir. Caz müziğini sık dinleyenlerin olasılığı 0.1 ile 0.2 arasında olduğu %95 güven düzeyi ile söylenir. Caz müziği çok sık dinleyenlerin olasılığı %18 olduğu %95 güven düzeyinde söylenir. Yapılan test sonucu genel anlamda caz müziğini insanların dinlemediği söylenebilir.

## DATAEXPLORER: Veri Analizi

```
install.packages("tidyverse")
library(tidyverse)

library(DataExplorer)
create_report(music_yeni_veri)
```

**DataExplorer** paketini indirdikten sonra kütüphaneden çağırıyoruz ve **creat\_report** fonksiyonu ile rapor sayfasına erişiyoruz. Bu fonksiyon bize oldukça hızlı bir şekilde istatistiksel grafikleri çıkartıyor ve bu yönüyle oldukça yararlı bir fonksiyon.

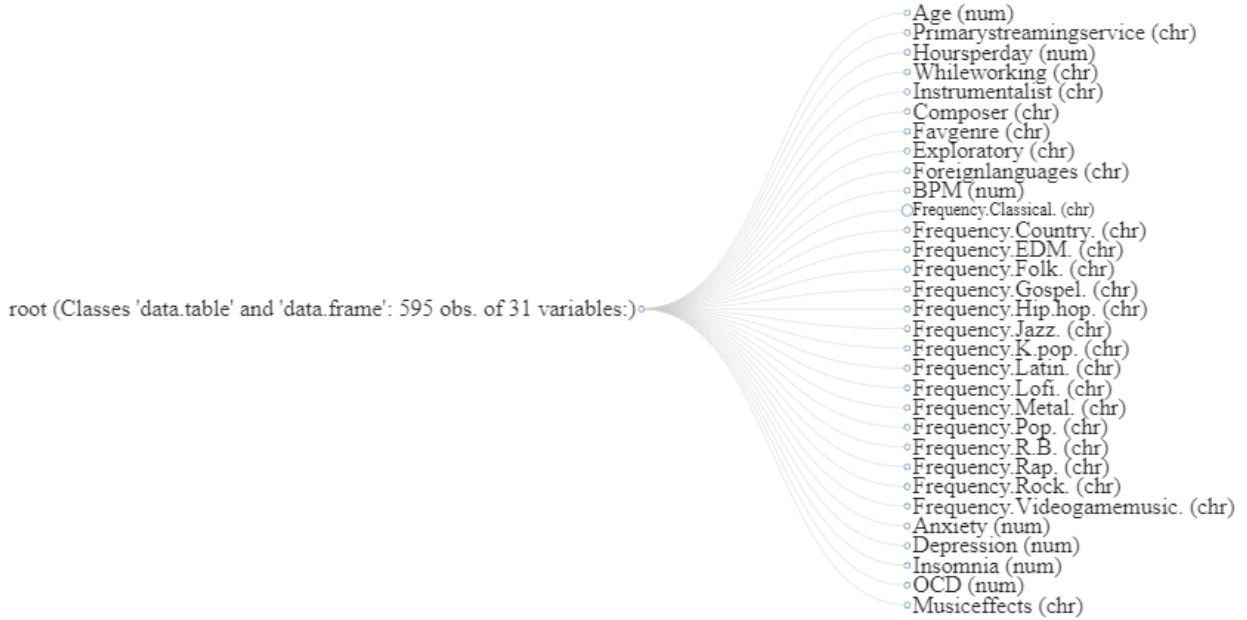
```
> library(DataExplorer)
> create_report(music_yeni_veri)

processing file: report.rmd
|...
2% inline R code fragments
|...
5% label: global_options (with options)
List of 1
 $ include: logi FALSE
|....
7% ordinary text without R code
|.....
100%
```

## Veriyi Yakından Tanıma

### Plot\_str

Aşağıdaki grafikte veri seti hakkında ve değişken türleri hakkında bilgileri alabiliriz. 595 gözlem, yani 595 katılımcı ve 31 değişkenden oluşan, manipüle edilmiş, yeni müzik ve mental sağlık verisinin, değişken türlerinin karakter ve sayısal olduklarını görebiliriz.



### Introduce

```
> introduce(music_yeni_veri)
  rows columns discrete_columns continuous_columns
1  595     31          24              7
all_missing_columns total_missing_values complete_rows
1              0              0              595
total_observations memory_usage
1          18445          160904
```

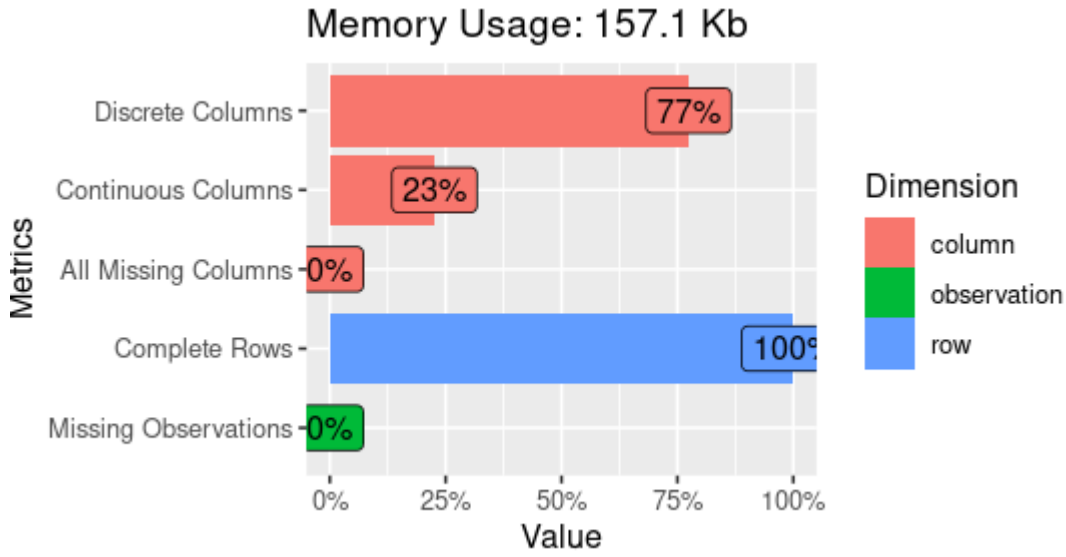
rows	columns	discrete_columns	continuous_columns	all_missing_columns	total_missing_values	complete_rows	total_observations	memory_usage	
1	595	31	24	7	0	0	595	18445	160904

Name	Value
Rows	595
Columns	31
Discrete columns	24
Continuous columns	7
All missing columns	0
Missing observations	0
Complete Rows	595
Total observations	18,445
Memory allocation	157.1 Kb

**Introduce** fonksiyonu ile yine veri seti hakkındaki bilgileri görebiliyoruz. 24 Kategorik değişken, 7 sayısal değişkenin olduğunu toplamda 31 değişkenin olduğunu söyleyebiliriz.

## Plot\_intro

```
> plot_intro(music_yeni_veri)
```



**Plot\_intro** ,hangi değişkenlerin sorunlu olabileceği hakkında,hızlıca fikir sahibi olmamıza yardımcı oluyor.

**Discrete columns** satırında Kategorik veriler, verimizin %77'sini oluşturmaktadır. **Continuous Columns** satırında sürekli verilerin, verilerin %23'ünü oluşturduğunu görebiliriz. Kayıp verilerimizi daha önce veri setinden çıkarmıştık bu yüzden, kayıp satır-sütun satırlarını %0 olarak görüyoruz. Tam olan satırların ise %100 olduğunu söyleyebiliriz.

## Kayıp veri

```
> plot_missing(music_yeni_veri)
```

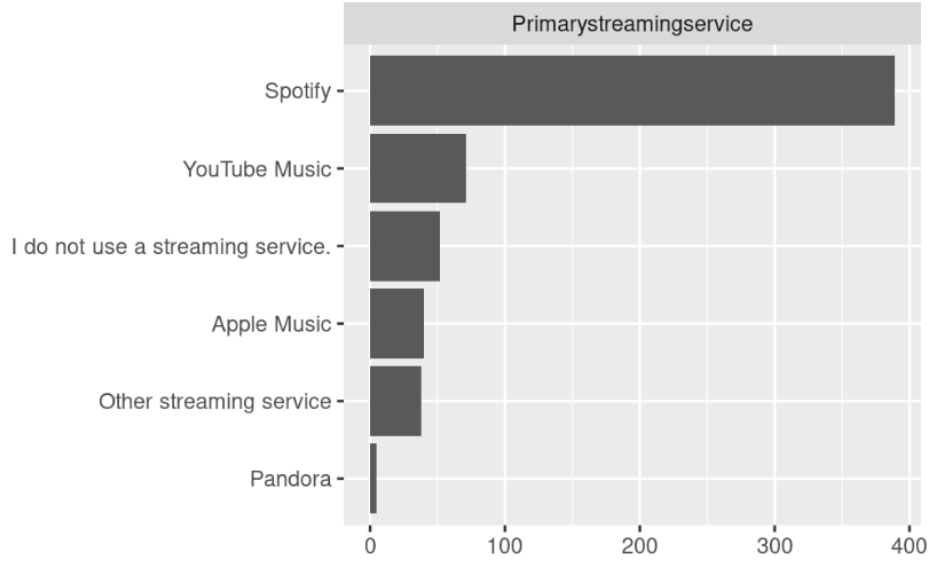


**Plot\_missing**, atılması gerekebilecek özelliklerin hızlı bir şekilde değerlendirilmesine veya değerlerinin atama yoluyla tahmin edilmesine olanak tanır, kayıp verileri bize gösterir. Temizlememiz gereken veri yoktur.

### Plot\_bar

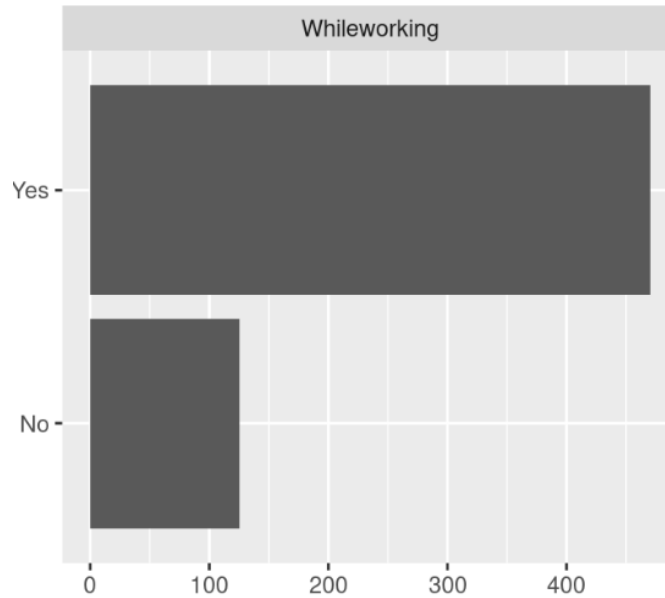
```
> plot_bar(music_yeni_veri)
```

Katılımcıların Müzik Servislerini Kullanma Grafiği



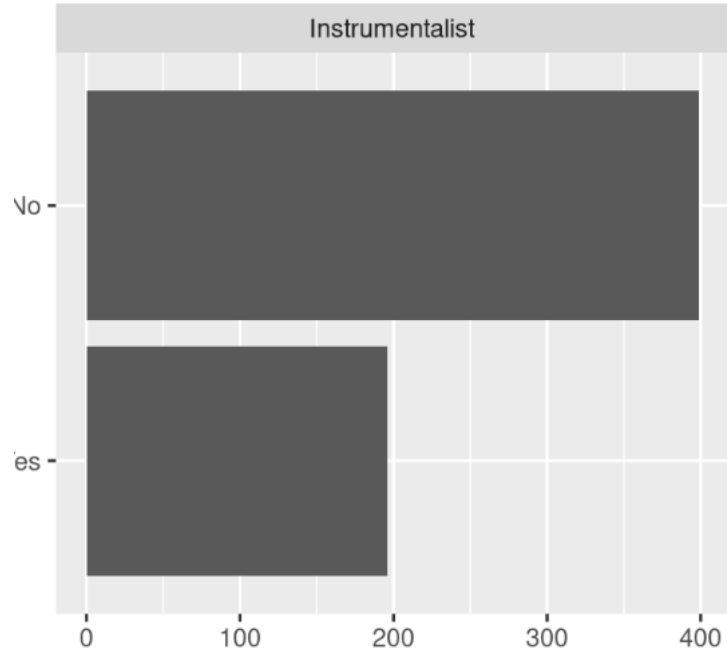
Kullanıcıların en çok Spotify kullandığını, en az da Pandora kullandığını, diğer müzik servislerinin ise birbirlerine yakın düzeyde, bu servisleri kullandıklarını söyleyebiliriz.

Katılımcıların Ders Çalışırken Müzik Dinleme Durumu Grafiği



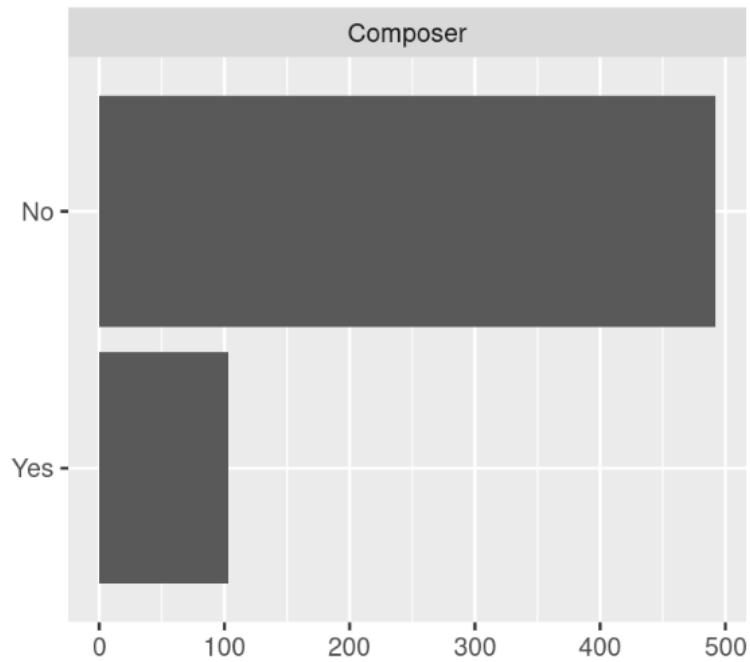
Katılımcıların (595 katılımcı), büyük bir kısmı ders çalışırken müzik dinlemektedir ve 100'e yakın katılımcı, ders çalışırken müzik dinlememektedir.

### Katılımcıların Düzenli Enstrüman Çalma Durumu Grafiği



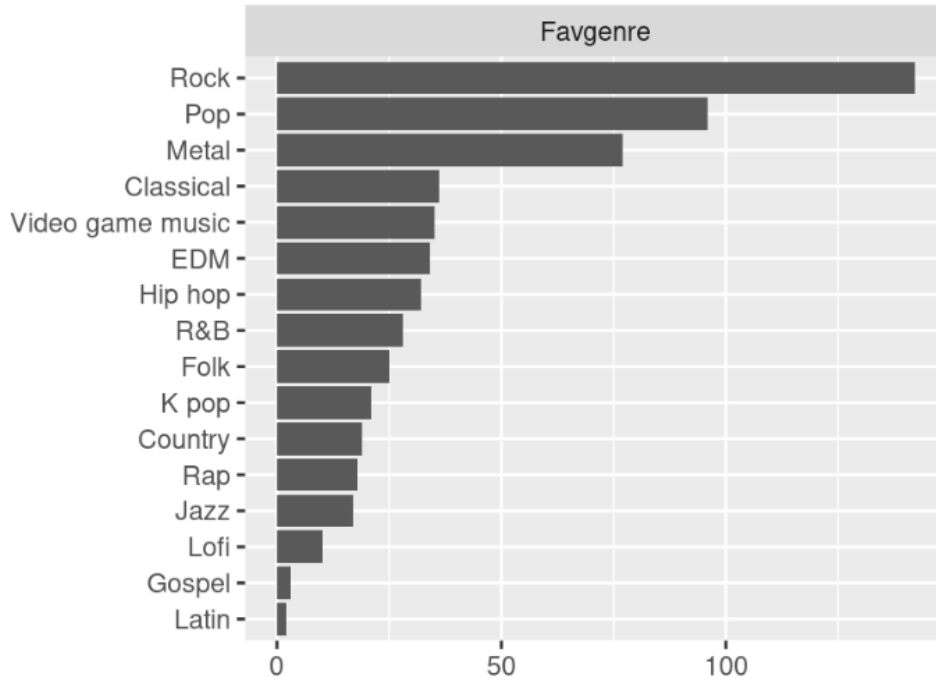
Katılımcıların, 200’e yakını düzenli olarak enstrüman çalıyor. 400’e yakını ise düzenli olarak enstrüman aleti çalmamaktadır.

### Katılımcıların Beste Yapma Durumu Grafiği



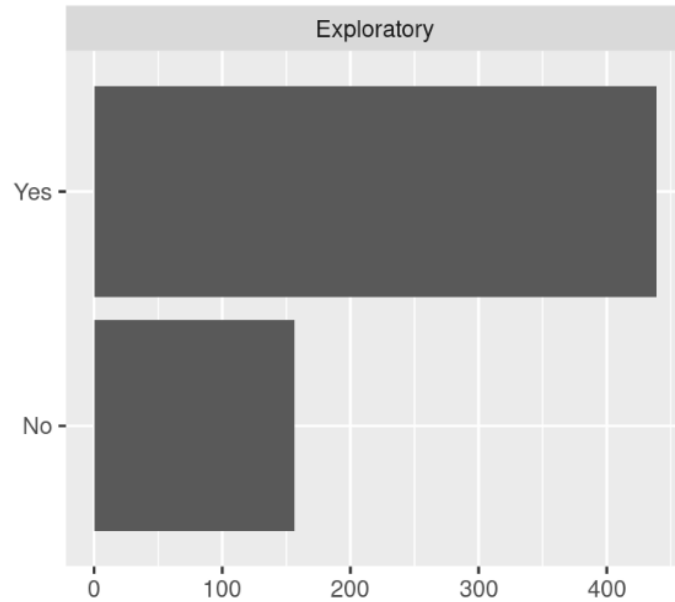
“Katılımcıların müzik besteliyor mu?” sorusuna verdikleri yanıtlarda, 500’e yakın katılımcının “hayır” cevabı vermesiyle, büyük çoğunluğun beste yapmadığı söylenir.

### Katılımcıların Müzik Türlerini Dinleme Grafiği



Katılımcılar, en fazla rock müzik türünü dinlemekte ve en az olarak da latin türü müzikleri dinlemektedirler. Rock müzik türünü, pop müzik ikinci sırada takip etmektedir.

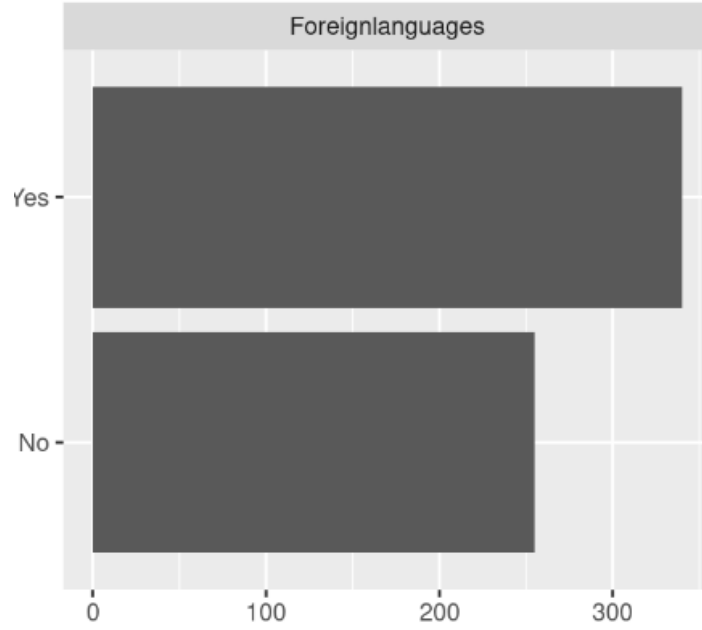
### Katılımcıların Aktif Yeni Sanatçı ve Müzik Türü Keşfetme Durumu Grafiği



Katılımcıların, aktif olarak yeni sanatçı ve türleri keşfetme yanıtları, büyük oranda “evet” yanıtı almıştır.

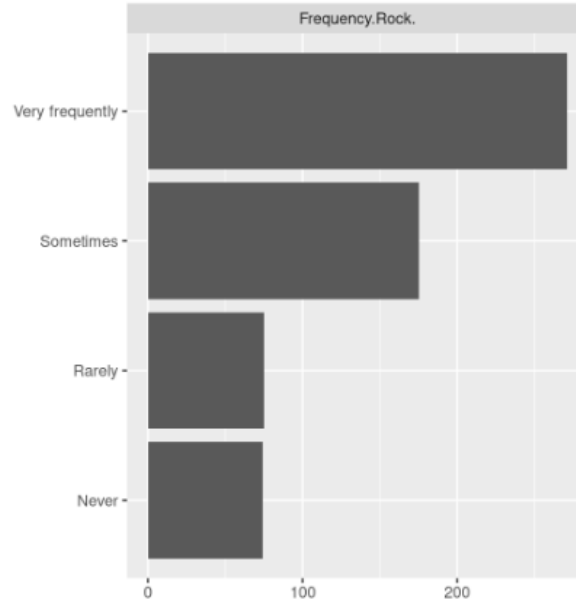


### Katılımcıların Müziği Yabancı Dilde Dinleme Durumu Grafiği



Katılımcıların müzik dinlerken yabancı dil kullanma oranları birbirlerine çok yakın olmasa da yakın sayılabilir. “Müzik dinlerken yabancı dil kullanıyor mu?” sorusuna cevapları çoğunluk olarak “evet” olmuştur.

### Katılımcıların Rock Müzik Türünü Dinleme Sıklığı Grafiği

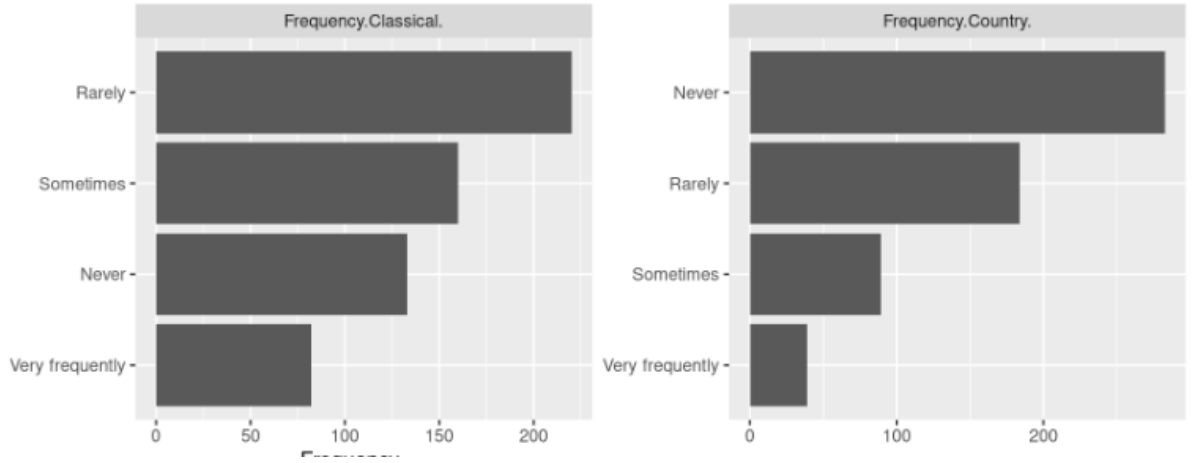


En çok dinlenen müzik türü olan rock müzik türünü ayrı olarak incelemek istedim. Dört kategoriden oluşan bu müzik dinleme sıklıkları;

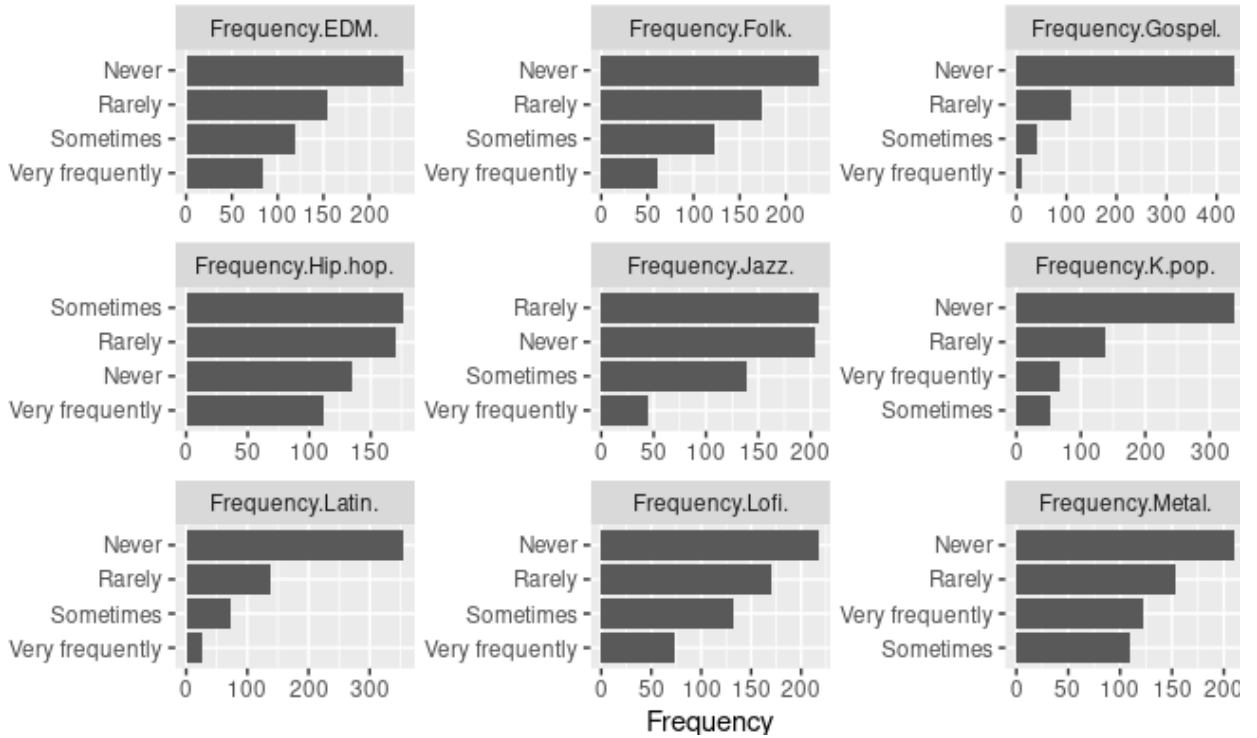
**Very frequently:** oldukça sık, **Somtimes:** Bazen, **Rarely:** Nadiren, **Never:** Hiç olarak sınıflandırılmıştır.

Rock müzik türünü hiç dinlemeyen ve ara sıra dinleyenler birbirlerine yakın giderken oldukça sık dinleyenler bazen dinleyenleri geçmiştir.

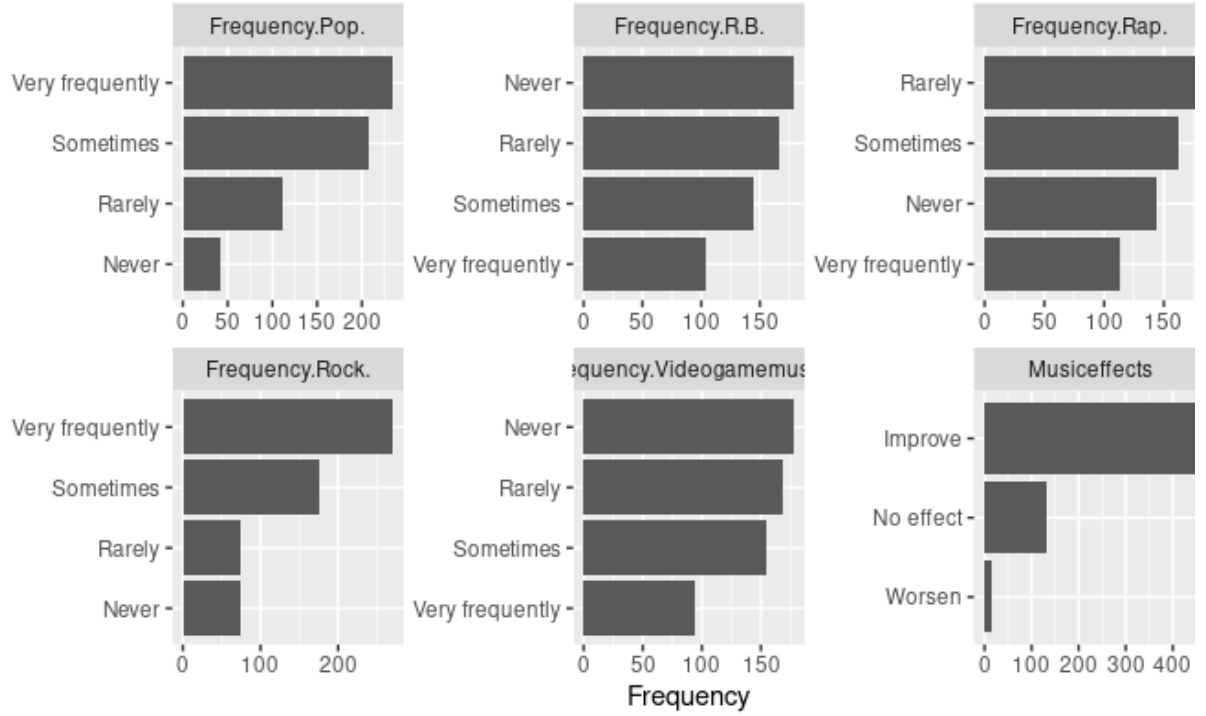
#### Katılımcıların Diğer Müzik Türlerini İçeren Dinleme Sıklıkları Grafikleri



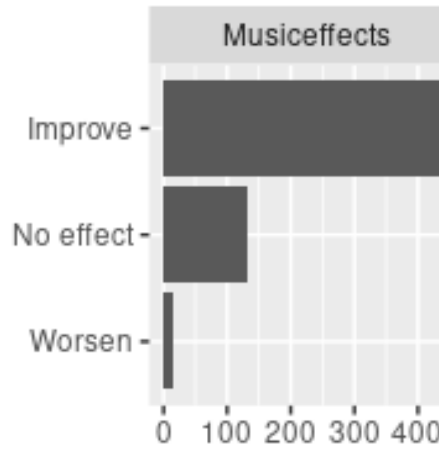
Diğer sıklık durumu grafikleri de aynı şekilde yorumlanabilir. Aşağıda rock müzik türünden farklı olarak EDM, folk ve gospel türlerinin hiç dinlenmeme oranları daha fazla olduğu yorumu da yapılabilir.



İkinci sırada olup, en çok dinlenen müzik türü olan, pop müzik dinlenme sıklığı ise oldukça sıktır.



Katılımcıların Müziğin Ruh Sağlığı Koşullarını Etkileme Durumu Cevap Grafiği



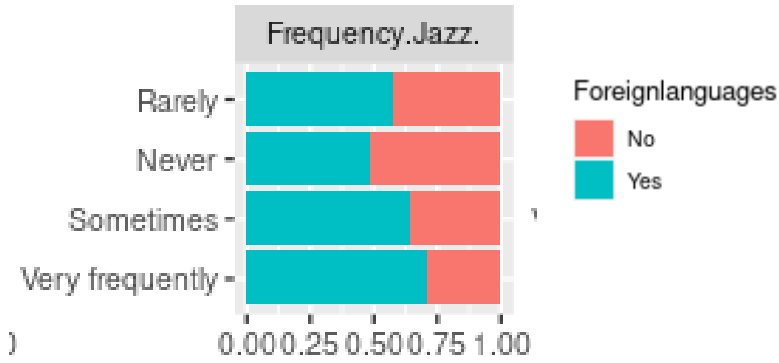
**Improve:** İyileştirir, **No effect:** Etkilemez, **Worsen:** Kötüleştirir

Katılımcılar, müzik dinlemenin ruh sağlığı üstünde iyileştirici olduğuna büyük oranda katılırken, kötü etkileme cevabının çok az verildiğini söylenir. Bir kısmında ise etki yaratmamaktadır.

```
> plot_bar(music_yeni_veri, by = "Foreignlanguages")
```

music\_yeni\_veri setinden değişkenlerin, "Foreignlanguages" yani yabancı dil dinleme durumları değişkenine göre grafik çıktılarını verir.

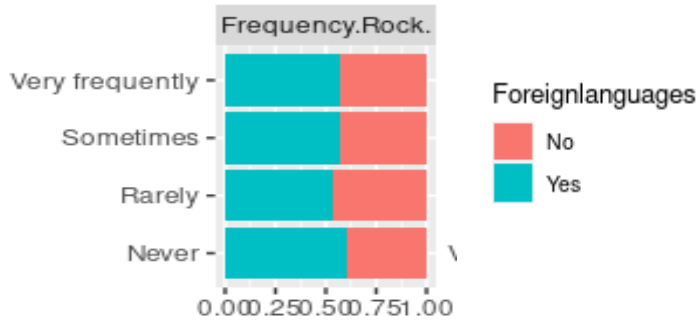
### Katılımcıların Caz Müziğini Yabancı Dil ile Dinleme Durumu Grafiği



**Very frequently:** Oldukça sık, **Somtimes:** Bazen, **Rarely:** Nadiren, **Never:** Hiç

Katılımcıların, Caz müziğini en çok *nadiren* sıklıkta dinlediklerini ve nadiren dinlerken, yabancı dilde daha çok dinledikleri söylenir. Yine caz müziğini oldukça sık dinleyenlerde, yabancı dilde dinlemeler daha fazladır.

### Katılımcıların Rock Müziğini Yabancı Dil ile Dinleme Durumu Grafiği

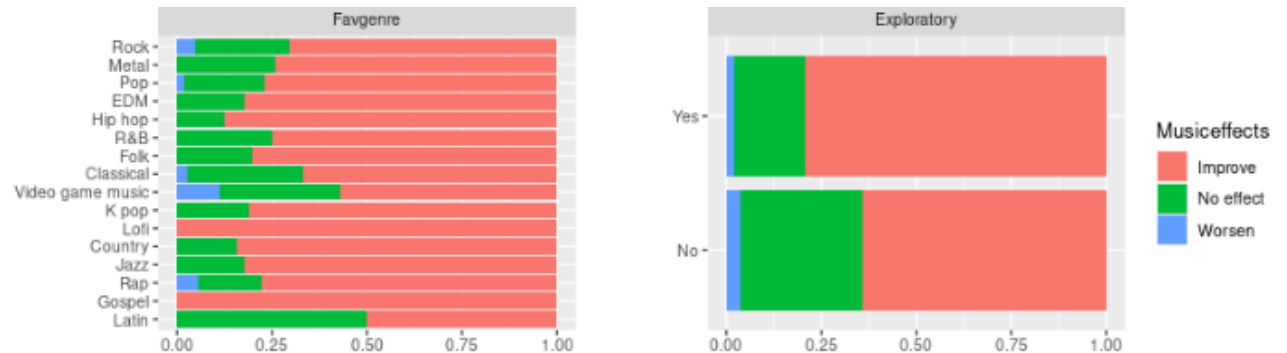


**Very frequently:** Oldukça sık, **Somtimes:** Bazen, **Rarely:** Nadiren, **Never:** Hiç

Katılımcıların, rock müziğini en fazla *oldukça sık* sıklığında dinledikleri, yabancı dilde dinleme ve dinlememe durumlarının yakın olduğu söylenir.

```
> plot_bar(music_yeni_veri, by = "MusicEffects")
```

### Müzik Dinlemenin Ruh Sağlığı Durumuna Etkisinin Diğer Değişkenlere Göre Bazı Grafikler



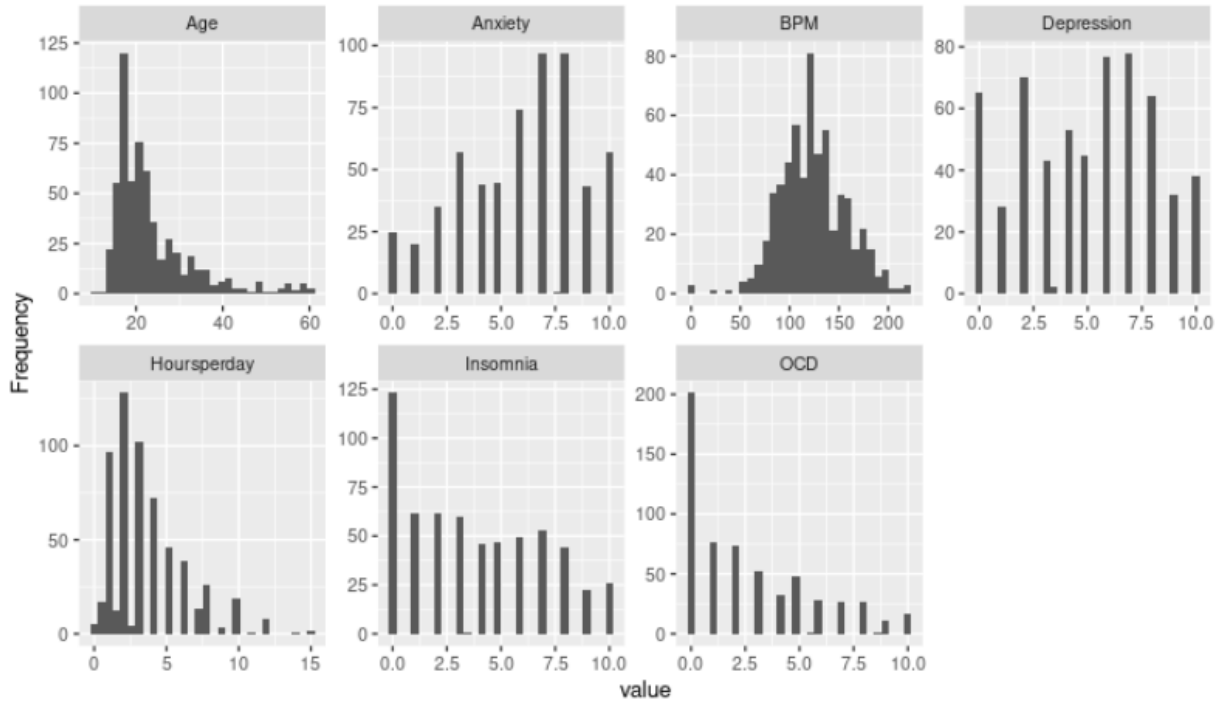
Katılımcıların en çok dinlediği müzik türü olan rock müzik, büyük oranda ruh sağlığına iyi etki ettiği yanıtını alırken, bir kısmı etkisiz, çok azı ise kötü etkiliyor yanıtını vermişlerdir.

Yeni müzik türü/sanatçı keşfetme durumunda ise, keşfetmeye *evet* yanıtını veren katılımcılar, *hayır* diyenlere göre, müzik dinlemenin ruh sağlığı üstünde yararlı etkisi olduğu yanıtını daha çok vermişlerdir.

## Plot\_histogram

```
> plot_histogram(music_yeni_veri)
```

Sürekli veriler için histogram grafiğinde, değişkenlerin dağılımları hakkında bilgi sahibi oluruz.



**Age** Katılımcıların yaş dağılımının en fazla yirmili yaşlarda olduğu görülür

**Anxiety** 0:10 aralığında katılımcıların kendilerine verdikleri kaygı puanlarında, çoğunlukla 7,5 puanına yakın puanlar verilmiştir.

**BPM** 0-1000m aralığında vuruş sayısında, katılımcılar en çok 100-150 bpm aralığını tercih etmişlerdir.

**Depression** 0-10 aralığında katılımcıların kendilerine verdikleri depresyon puanlarında, kendisine 0 puan veren (depresyon yok) ile 2.5 puan (depresyon az olarak nitelendirilebilecek puan) birbirine yakın seçim almıştır, kaygı puanında olduğu gibi depresyon puanında da 7.5 puan en fazla seçim almıştır.

**Hoursperday** Katılımcıların günlük dinleme saat sayısı çoğunlukla 0-5 saat arasındadır.

**Insomnia** 0-10 aralığında uykusuzluk puanına, katılımcıların büyük çoğunluğu 0 seçimini yapmıştır. Katılımcılar genel anlamda uykusuzluk çekmeseler de uykusuzluk çekenler en fazla 2.5 puan

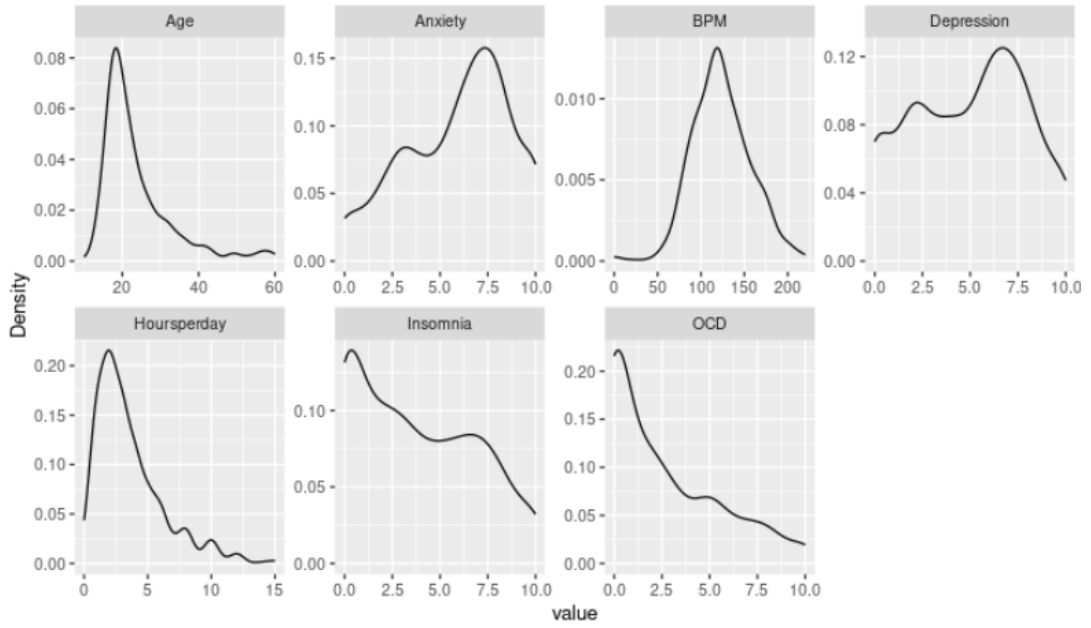
aralığını seçmişler, daha sonra 7.5 puan aralığını seçmişler, kendine çok uykusuz diyenler azınlıktadır.

**OCD** 0-10 aralığında katılımcıların kendilerine verdikleri obsesif kompulsif bozukluk puanlarında, çoğun kendisinde obsesiflik görmemiştir.

### Plot\_density

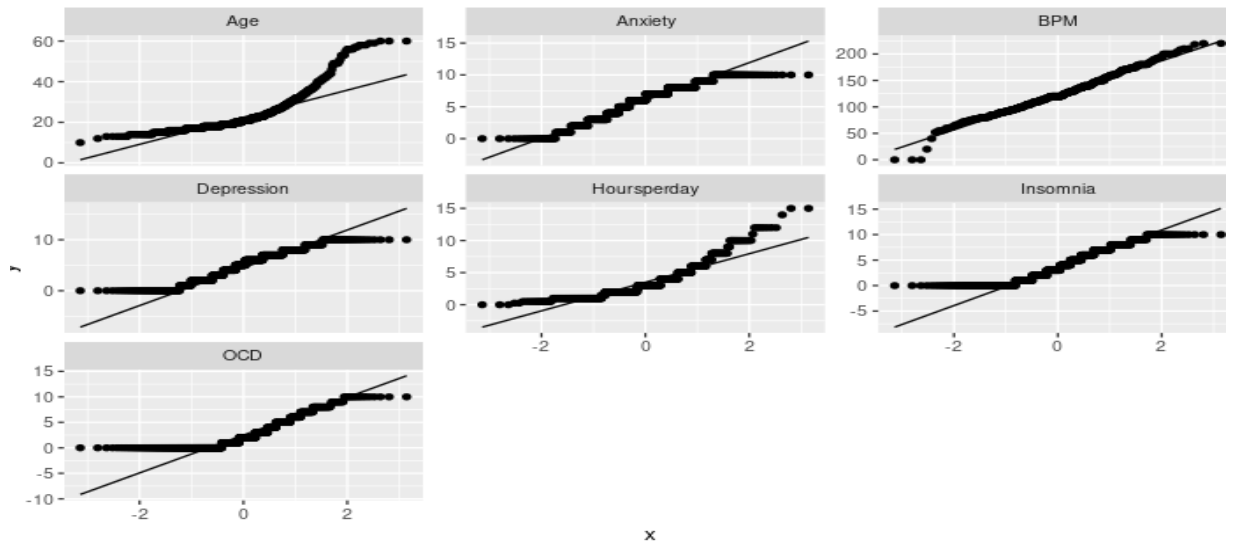
```
> plot_density(music_yeni_veri)
```

Tüm sürekli değişkenlerin, tahmini yoğunluk dağılımını gösteren grafik aşağıdaki gibidir.



```
> plot_qq(music_yeni_veri)
```

Quantile-Quantile grafiği, belirli bir olasılık dağılımından sapmayı görselleştirmenin bir yoludur. Normal dağılım etrafında verilerin ne şekilde dağıldığını görebiliriz.

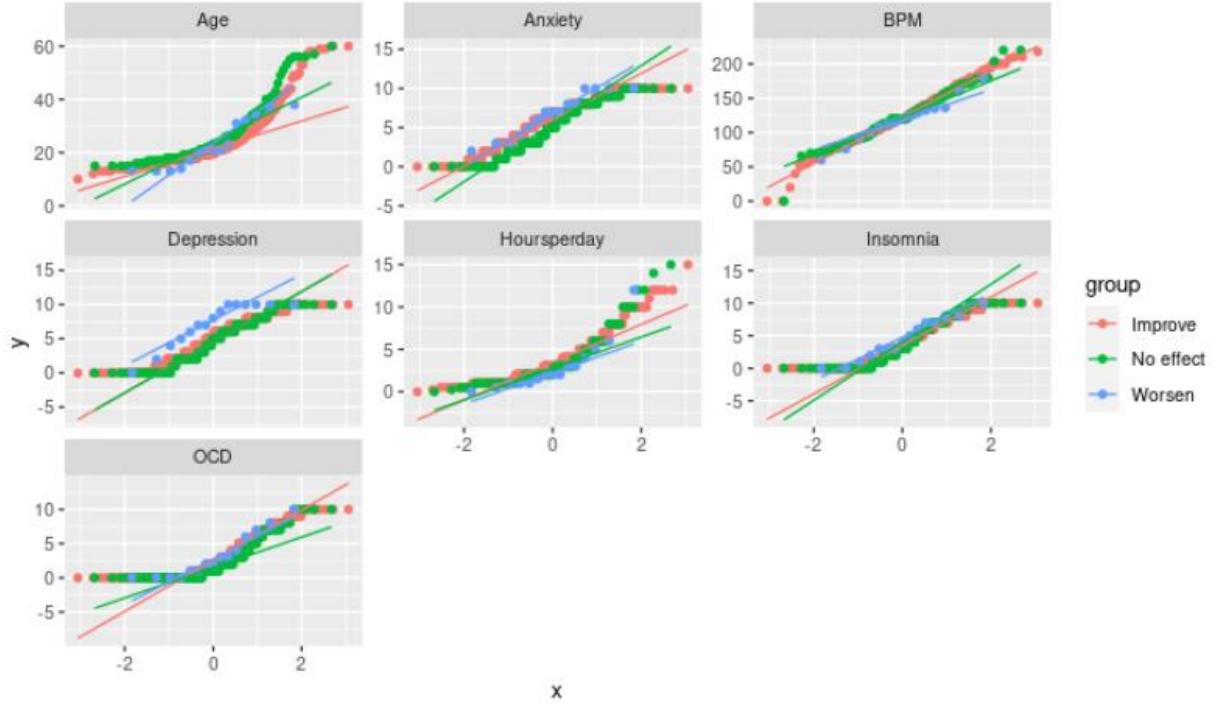


*Age* değişkeni, 40 yaş civarı sonrasında sapmayı arttırmıştır ve *BPM* değişkeni, 50-200 bpm'de daha sağlıklı sonuçlar vermiştir. *Hoursperday*, 0-15 saat aralığında, 6 saatten sonra doğrudan uzaklaşma göstermiştir. Geriye kalan tüm değişkenler 0-10 aralığında puanlanan değişkenlerdir.

## Plot\_qq

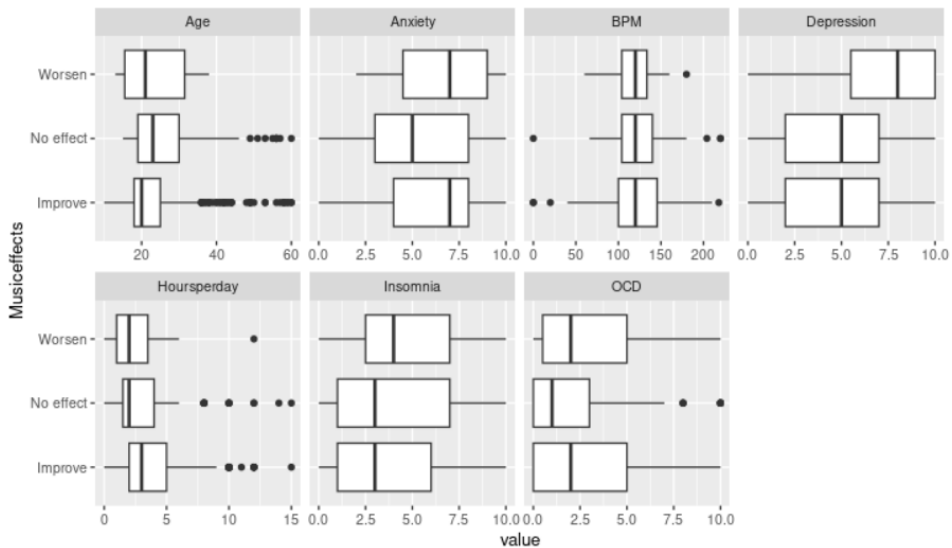
```
> plot_qq(music_yeni_veri, by = "Musiceffects")
```

Müzik Dinlemenin Ruh Sağlığı Üzerindeki Etkisi ile Diğer Sürekli Değişkenler Üzerindeki Sapmaları



## Plot\_boxplot

```
> plot_boxplot(music_yeni_veri, by = "Musiceffects")
```



**Age|** *Müzik dinlemenin ruh sağlığı üzerindeki etkisi, yaşa göre geneli* 20 yaşı civarındadır ve kötü etkiliyor diyenler 40 yaş altındadır. Olumlu etkiliyor diyenlerin ortalama yaşı 20'dir ve 60 yaşına kadar olumlu etki seçimini yapanlar vardır. Etkisiz diyenlerin yaş ortalaması da yaklaşık 23 civarındadır. Noktalar aynı zamanda aykırı gözlemleri de gösterirler.

**Anxiety|** *Müzik dinlemenin ruh sağlığı üzerindeki etkisi, katılımcıların kendilerine verdikleri kaygı puanlarına göre* kötü ve iyi etkiliyor diyenlerin, depresyon puanları ortalamaları 7,5'a yakındır. İyi etkiliyor diyenlerin genişliği kötü etkiliyor diyenlere göre daha küçüktür. Etkisiz olduğunu seçenlerin ortalama depresyon puanı 5'tir ve en fazla genişliğe sahip seçenektir. Müzik dinlemenin ruh sağlığı üstündeki etkisi, iyi de dense kötü de dense, kaygı puanlarında çok büyük bir değişiklik yaratmamıştır. Müzik dinlemek ruh sağlığını etkilemiyor diyenlerin kaygı puanları daha düşüktür.

**BPM|** *Müzik dinlemenin ruh sağlığı üzerindeki etkisi, müzik vuruş sayısına göre* iyi, kötü, etkisiz etkiliyor seçeneklerinin hepsinin ortalama vuruş sayısı 120 civarındadır. İyi etkiliyor diyenler en fazla genişliğe sahiptir ve 100-150 bpm arasındadır, 200 veya 0 bpm yakınlığında da iyi etkiliyor diyenler olmuştur. Sonra genişliği, etkisiz olduğunu söyleyenler, an az genişlik de kötü etkiliyor diyenlerin olmuştur.

**Depression|** *Müzik dinlemenin ruh sağlığı üzerindeki etkisi, katılımcıların kendilerine verdikleri depresyon puanlarına göre* kötü etkiliyor diyenlerin ortalama depresyon puanları 8 civarındadır. Etkisiz ve iyi etkiliyor diyenlerin ortalama puanları ise 5'tir. Katılımcıların kendilerine verdikleri depresyon puanı, müzik dinlemenin ruh sağlığı üzerindeki etkisi kötü diyenlerle orantılı olarak artmıştır.

**Hoursperday|** *Müzik dinlemenin ruh sağlığı üzerindeki etkisi, günlük müzik dinleme sürelerine göre* iyi etkiliyor diyenlerin ortalama dinleme süreleri 3'tür. Etkisiz ve kötü etkiliyor diyenlerin ise ortalama dinleme süreleri 1-2 saat aralığındadır. Genel anlamda günlük 0-5 saat aralığında müzik dinlenildiği söylenebilir.

**Insomnia|** *Müzik dinlemenin ruh sağlığı üzerindeki etkisi, uykusuzluk puanlarına göre* kötü etkiliyor diyenlerin, iyi etkiliyor ve etkilemiyor diyenlere göre daha fazla uykusuzluk yaşadığı görülür. Müzik dinlemek kötü etkiliyor diyenler ortalama 4 uykusuzluk puanı verirken, etkisiz ve iyi etkiliyor diyenler 2,5 civarı puan vermektedir. En fazla genişliğe etkilemiyor diyenler sahiptir.

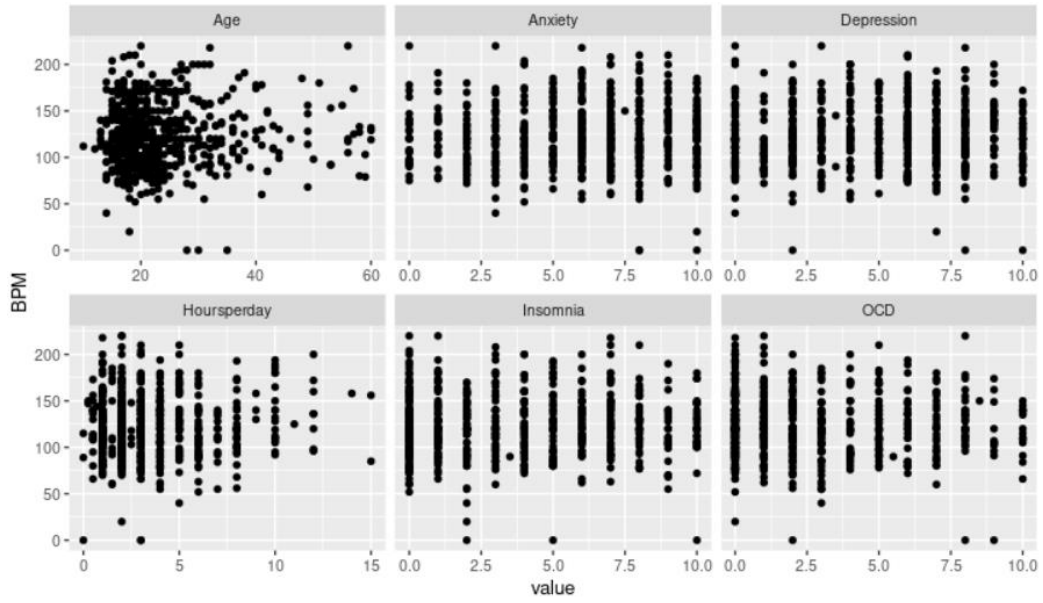
**OCD|** *Müzik dinlemenin ruh sağlığı üzerindeki etkisi, katılımcıların kendilerine verdikleri obsesif kompulsif bozukluk puanlarına göre* ortalama puanları 2,5 altındadır. Müzik dinlemek etkisiz diyenlerin kendilerine çok yüksek obs puanı verdikleri de görülür.

## **Plot\_scatterplot**

```
> plot_scatterplot(split_columns(music_yeni_veri)$continuous, by = "BPM", sampled_rows = 1000L)
```

Sürekli olan değişkenlerin BPM değişkenine göre saçılımalarını görüyoruz.

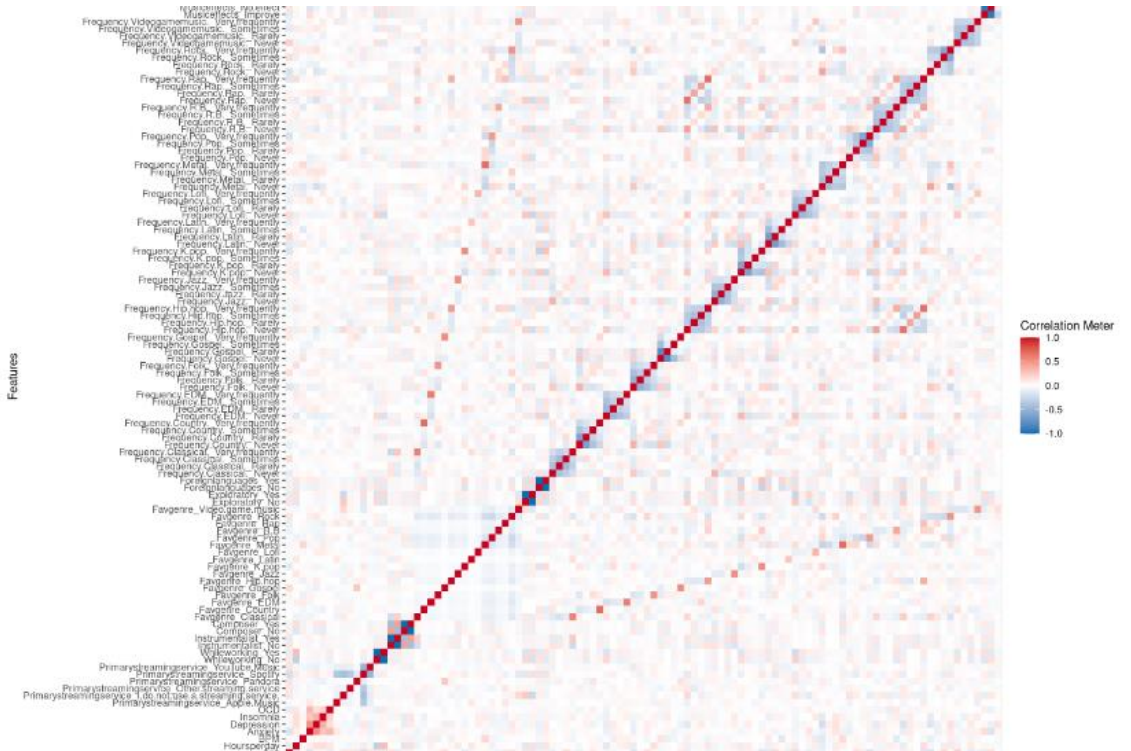




## Plot\_correlation

```
> plot_correlation(music_yeni_veri)
\
```

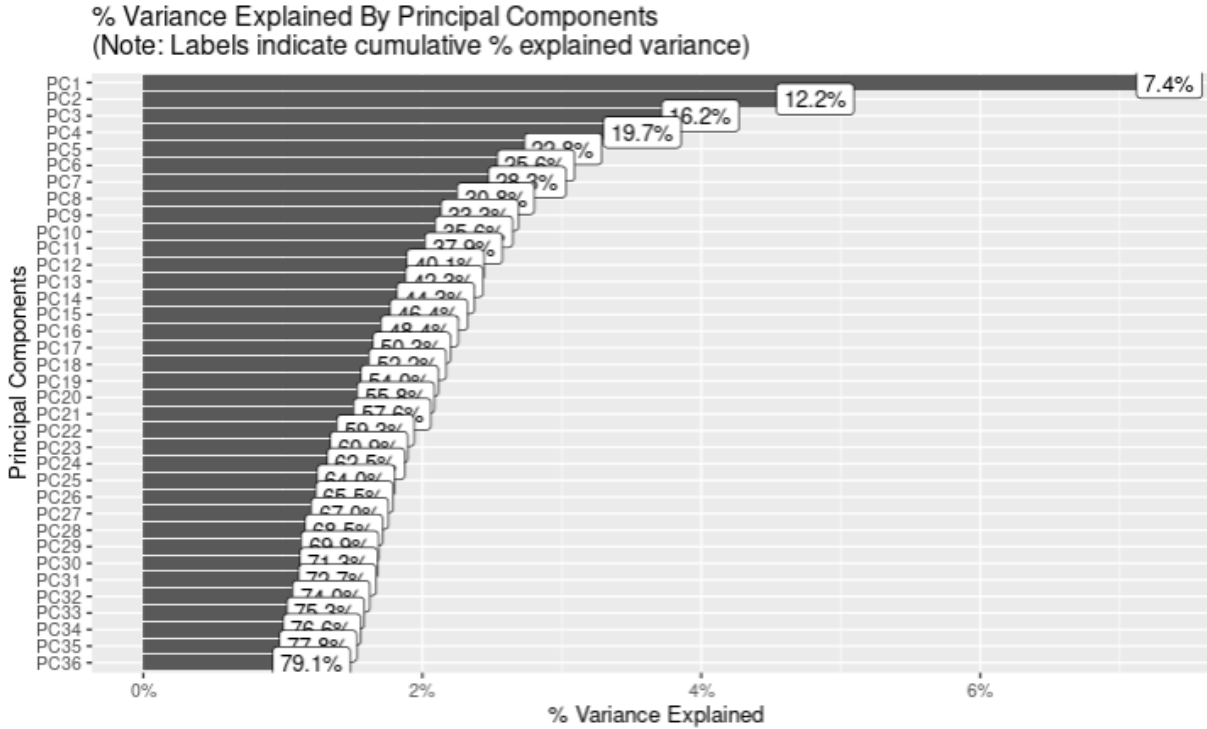
Hangi değişkenlerin güçlü bir şekilde ilişkili olduğu veya çok güçlü şekilde ilişkili olmadığı hakkında hızlıca bilgi sahibi olmamıza yaran bu korelasyon grafiğinde, pozitif ve negatif korelasyonların renklerle gösterildiği grafikdir.



## Temel bileşenler Analizi

```
> plot_prcomp(music_yeni_veri, maxcat = 5L)
2 features with more than 5 categories ignored!
Primarystreamingservice: 6 categories
Favgenre: 16 categories
```

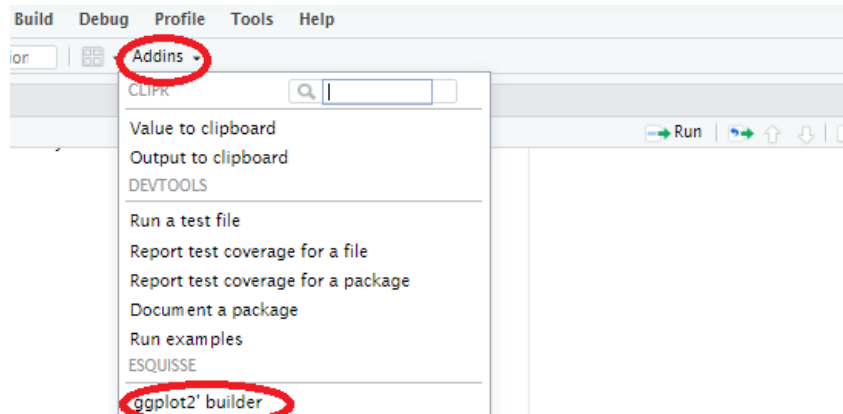
Temiz veriler ile daha iyi çalışır. Değişkenler kullanım açısından yararlı olabilir veya olmayabilir. Çok fazla sayısal değişkenimiz var ise ve daha fazla güçlü şekilde ilişkili öğeler varsa temel bileşenler analizi daha yararlı olur. Varyans açıklama oranını en fazla, en az hangi değişkenleri tutarak sağlayabiliriz bunun cevabını verir.

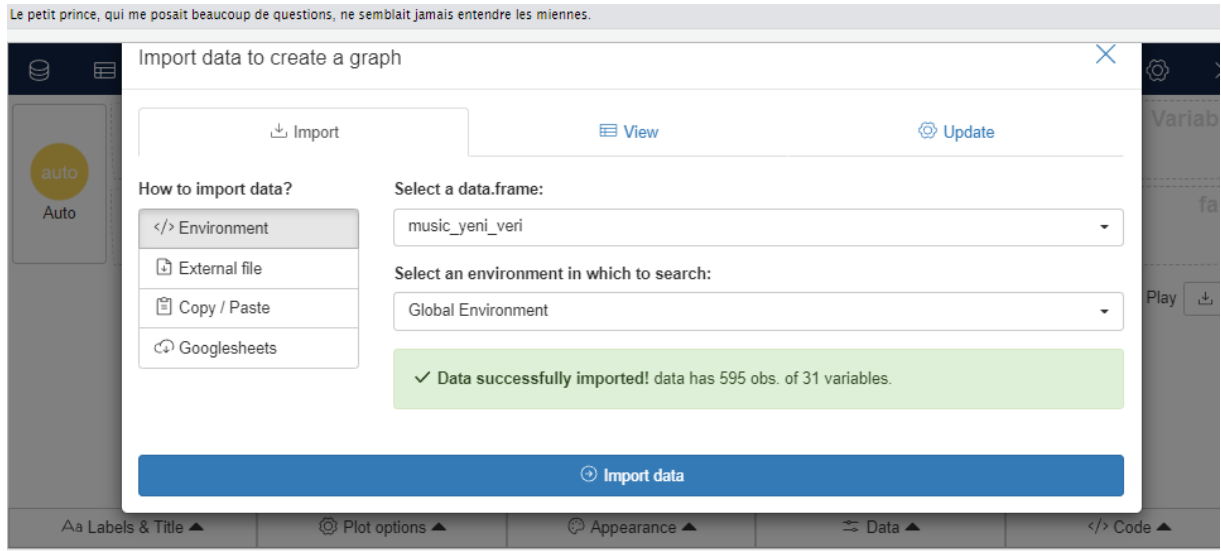


## VERİ GÖRSELLEŞTİRME:

### Esquisse Package

Bu paketin en büyük avantajı, değişkenleri seçip belirli kutucuklara bıraktığımızda bize otomatik olarak grafiği çok hızlı bir şekilde elde edebilmemizi sağlamasıdır.

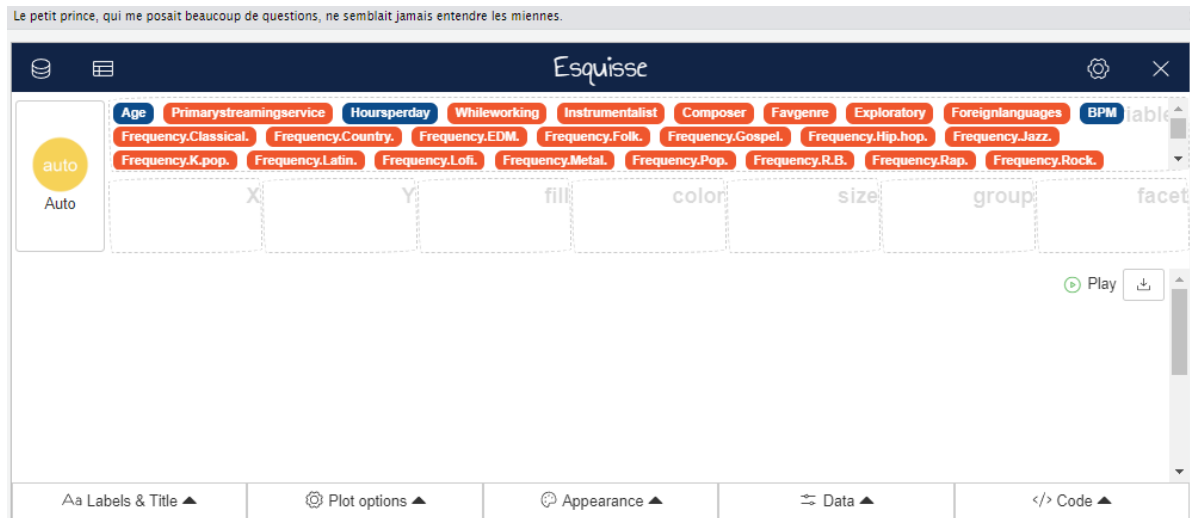




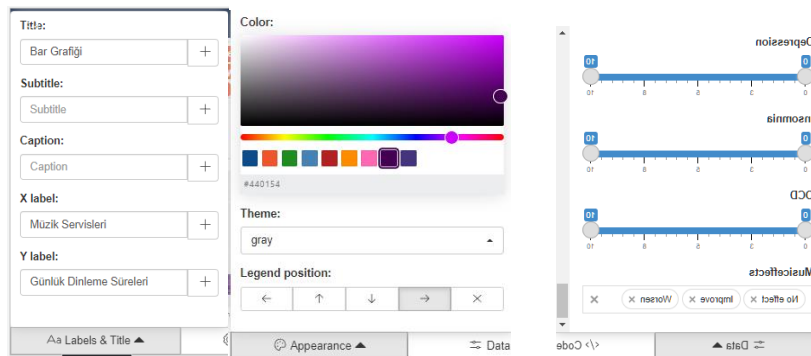
Bu adımları izleyerek veri setini okutabiliriz veya;

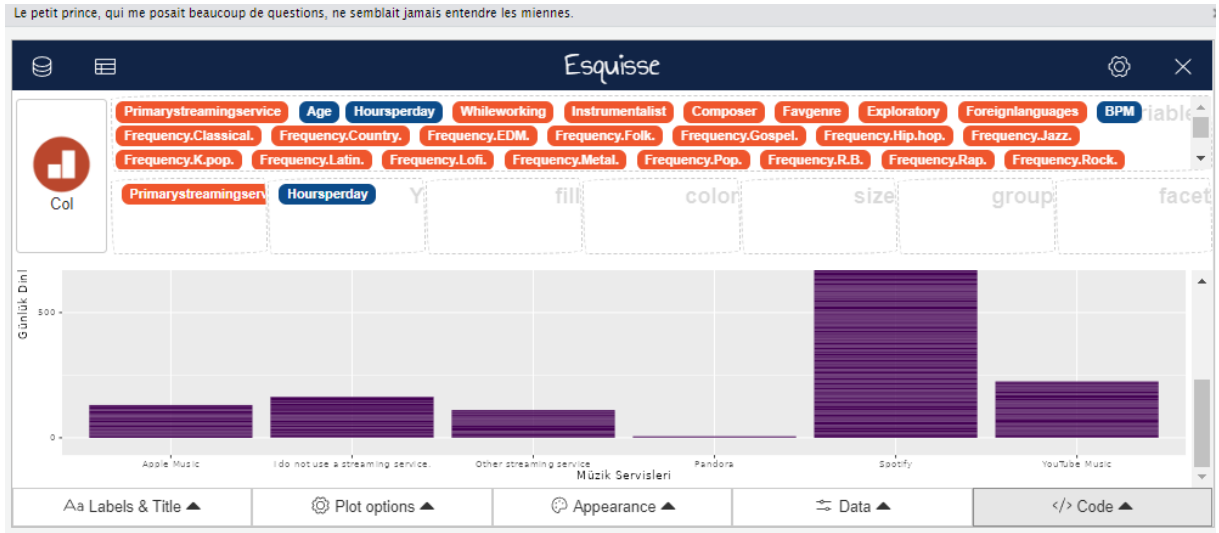
```
> library(esquisse)
> esquisser(music_yeni_veri)
```

Yukarıdaki kod, aşağıdaki gibi üstünde istediğim ayarlamaları yapabileceğim sayfayı veriyor.



Karşımıza çıkan ekranın alt kısımda bulunan, açılır pencereli alanlar, daha detaylı ayarlama yapma olanağı tanır; Başlık verme, renk verme, veri setini ayarlama vb. değişiklikler.

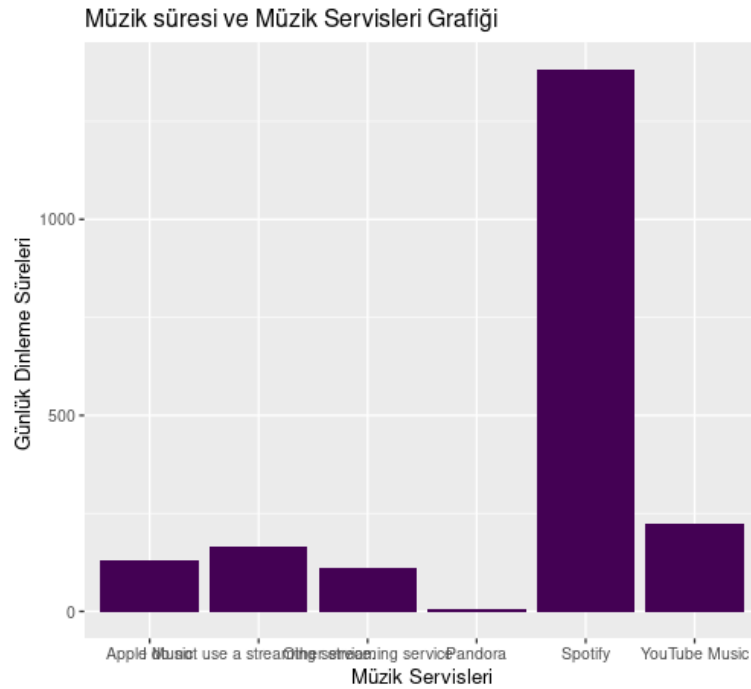




Müzik servisleri ve yaş değişkenlerini sürükleyerek alanlara bıraktım. Bana otomatik grafik oluşturdu.

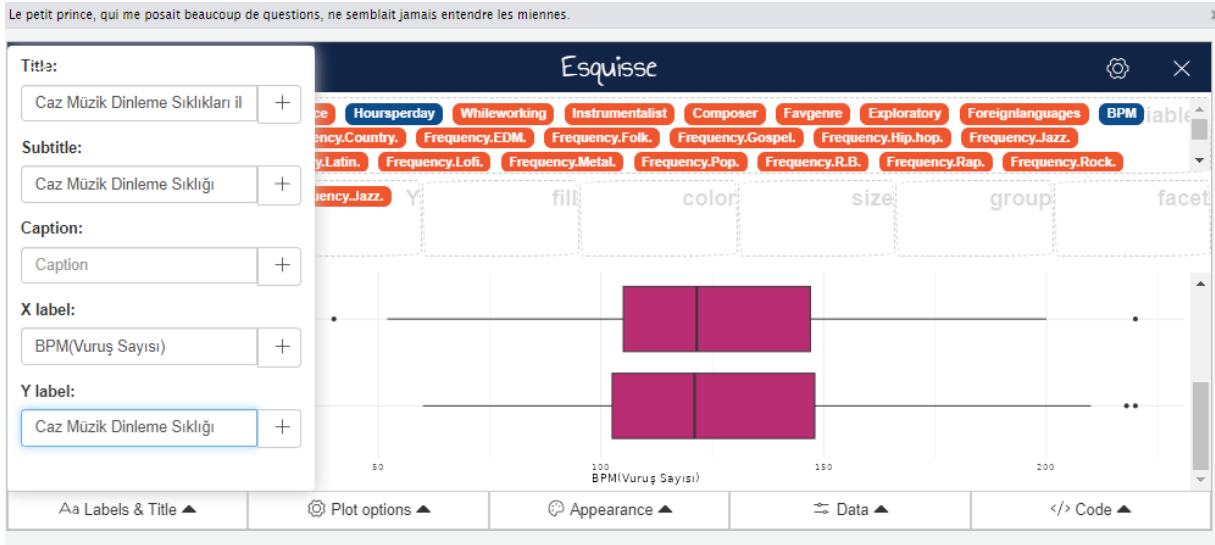
Başlıklar, yazı boyutu, renk gibi ayarlamaların da içinde yer aldığı kod:

```
> music_yeni_veri %>%  
+ filter(!(Favgenre %in% "R&B")) %>%  
+ ggplot() +  
+ aes(x = Primarystreamingservice, y = Hoursperday) +  
+ geom_col(fill = "#440154") +  
+ labs(x = "Müzik Servisleri", y = "Günlük Dinleme Süreleri ", title = "Müzik süresi ve Müzik Servisleri Grafiği") +  
+ theme_gray() +  
+ theme(plot.caption = element_text(size = 25L, face = "bold", hjust = 0.5))
```

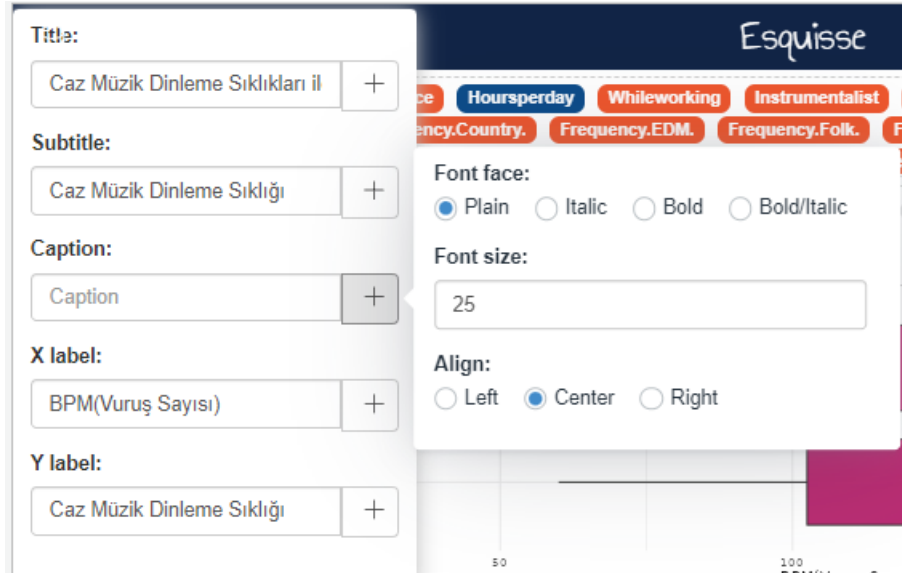


Grafikten en çok Spotify servisinin kullanıldığını görebiliriz.

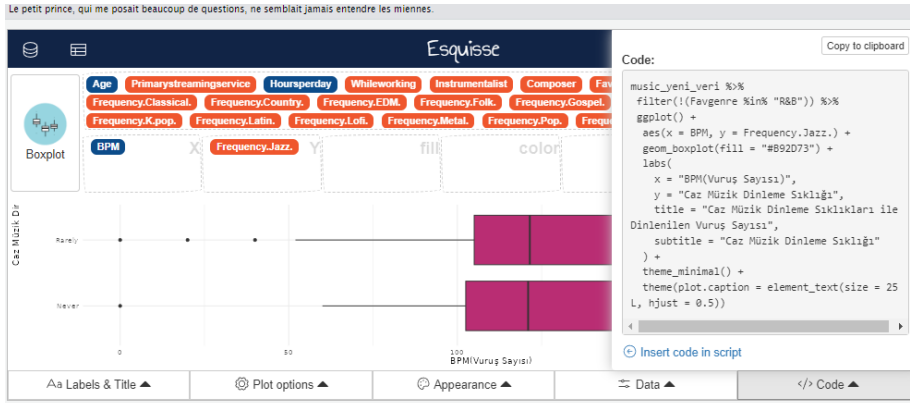
Başka bir grafik ile devam edelim.



Burada x-y eksenlerinin isimlendirmelerini yaptım

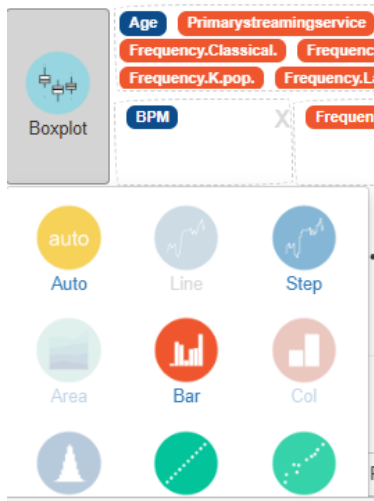


Yazı boyutu, yazı tipi gibi ayarlamaları buradan yapabiliriz.



Burada ise ayarlamaları tamamladıktan sonra kodu, Insert code in script kısmına tıklayarak

program yüzüne yazdırırım ve çalıştırırım.



Burada elimizdeki değişkenlerle elde edebileceğimiz ve elde edemeyeceğimiz grafikleri görebiliriz. Turuncu renkli değişkenler kategorik değişkenler olurken, lacivert değişkenler sayısal değerler içerenlerdir.

**Filter():** Değişkenleri Filtrelememizi sağlar

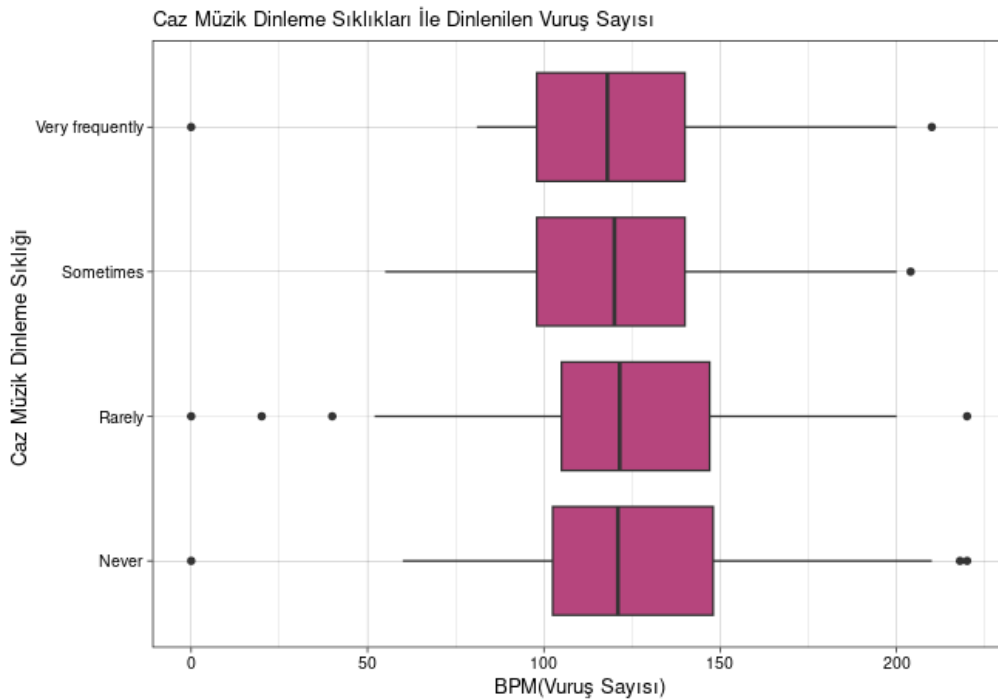
**Aes():** Kullanacağımız değişkenleri seçeriz.

**Geom\_boxplot():** Renklendirme için kullanılır.

**Labs():** İsimlendirme için kullanılır.

**Theme():** Yazı boyutu gibi ayarları yapabiliriz.

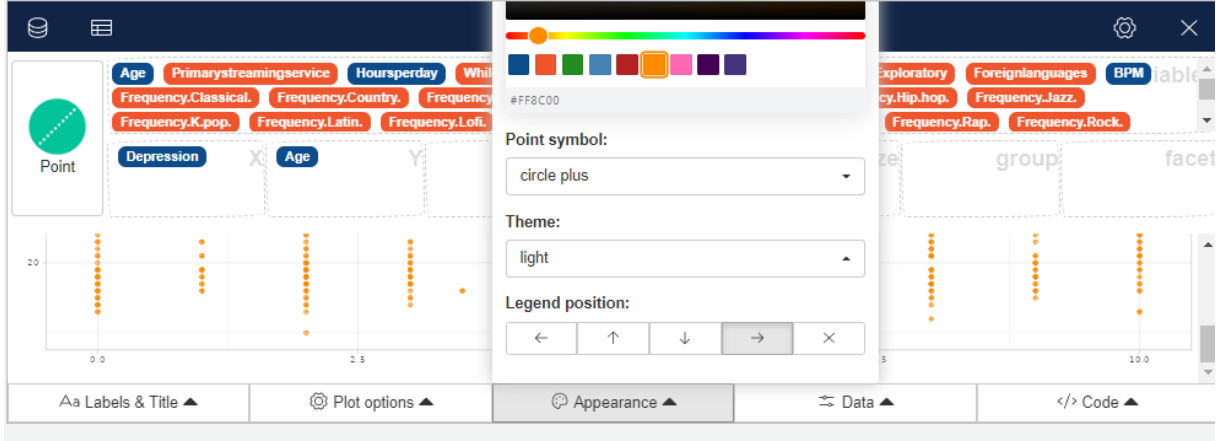
```
> music_yeni_veri %>%
+ filter(!(Favgenre %in% "R&B")) %>%
+ ggplot() +
+ aes(x = BPM, y = Frequency.Jazz.) +
+ geom_boxplot(fill = "#86457D") +
+ labs(x = "BPM(Vuruş Sayısı)",
+ y = "Caz Müzik Dinleme Sıklığı ", subtitle = "Caz Müzik Dinleme Sıklıkları İle Dinlenen Vuruş Sayısı") +
+ theme_linedraw() +
+ theme(plot.caption = element_text(size = 25L, hjust = 0.5))
> |
```



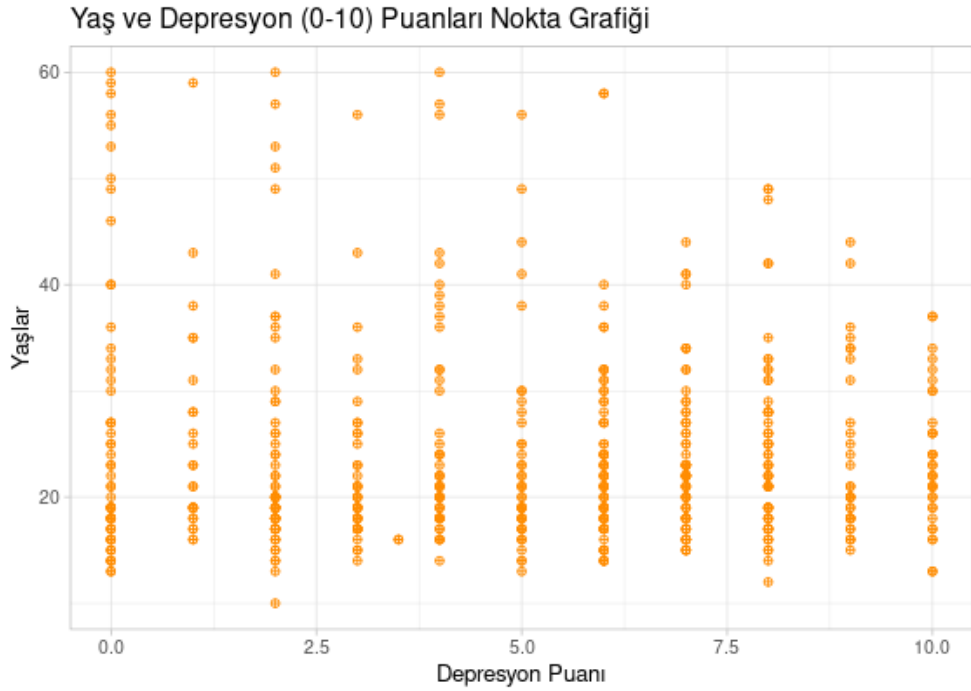
**Very frequently:** çok sık, **Somtimes:** Bazen, **Rarely:** Nadiren, **Never:** Hiç

Caz müziğini dinleyenler, tüm dinleme sıklıklarında ortalama 120 bpm civarında dinlemektedirler.

İki sürekli değişkenin grafiğini çizdirmek istediğimde bana *point* grafiğini sundu. Açılı pencereden de renk ayarlaması yaptım.

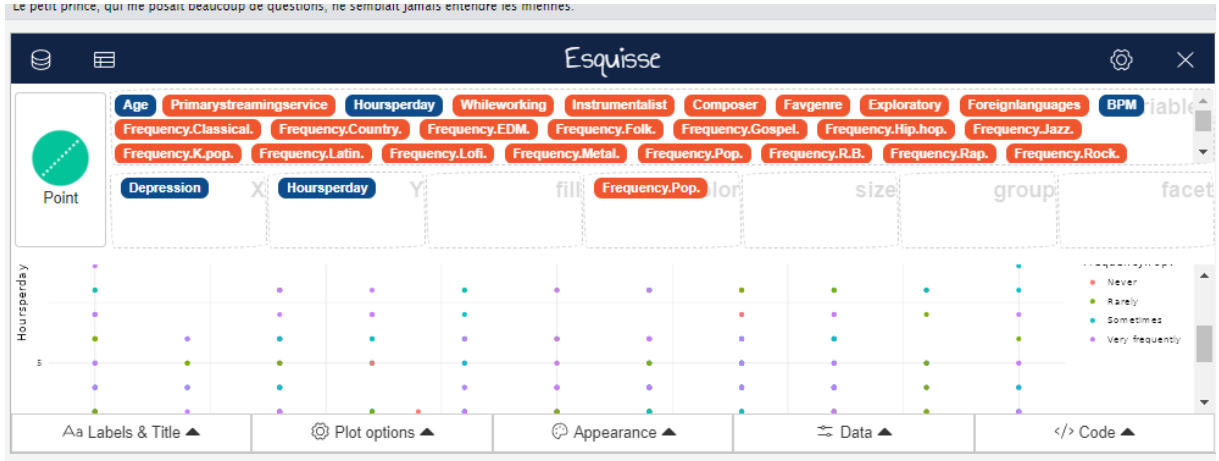


```
> music_yeni_veri %>%
+ filter(!(Favgenre %in% "R&B")) %>%
+ ggplot() +
+ aes(x = Depression, y = Age) +
+ geom_point(shape = "circle plus", size = 1.5,
+ colour = "#FF8C00") +
+ labs(x = "Depresyon Puanı", y = "Yaşlar", title = "Yaş ve Depresyon (0-10) Puanları Nokta Grafiği ") +
+ theme_light() +
+ theme(plot.caption = element_text(size = 25L, face = "italic", hjust = 0.5))
>
```

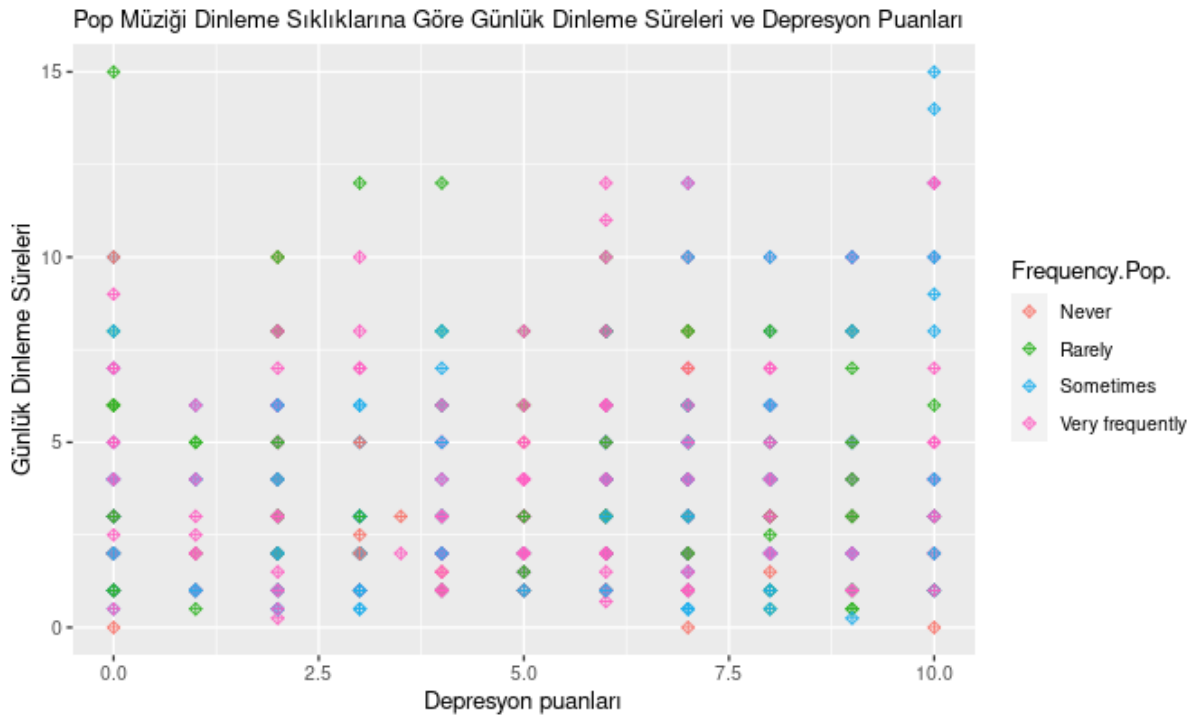


En yüksek genişlik, depresyon puanının en düşük verildiği 0 puan ile 15-60 yaş arasında yayılmaktadır. Katılımcıların kendilerine verdikleri en yüksek depresyon puanı olan 10, 15-38 yaş grubu arasında görülmektedir. Daha genç yaştaki insanlar, daha depresif hissettikleri söylenebilir.

İki sürekli değişkeni, kategorik değişkenle birlikte grafik halinde göreceğiz, kategorileri renklerle ayrıştırıyor.

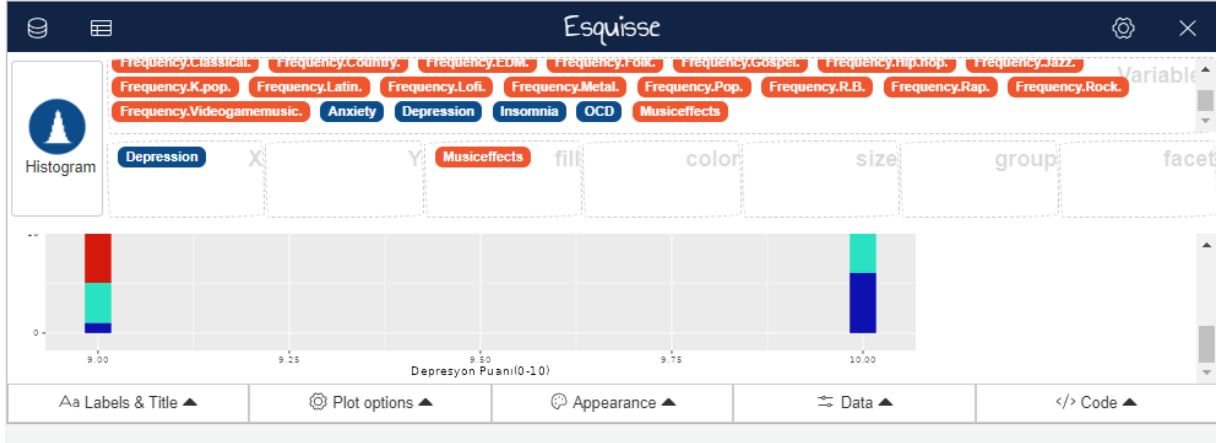


```
> music_yeni_veri %>%
+ filter(!(Favgenre %in% "R&B")) %>%
+ ggplot() +
+ aes(x = Depression, y = Hoursperday, colour = Frequency.Pop.) +
+ geom_point(shape = "diamond plus",
+ size = 1.5) +
+ scale_color_manual(values = c(Never = "#F8766D", Rarely = "#31B425", Sometimes = "#20AFEC",
+ 'Very frequently' = "#FF61C3")) +
+ labs(x = "Depresyon puanları", y = "Günlük Dinleme Süreleri", subtitle = "Pop Müziği Dinleme Sıklıklarına Göre Günlük Dinleme Süreleri ve Depresyon Puanları") +
+ theme_gray()
> |
```

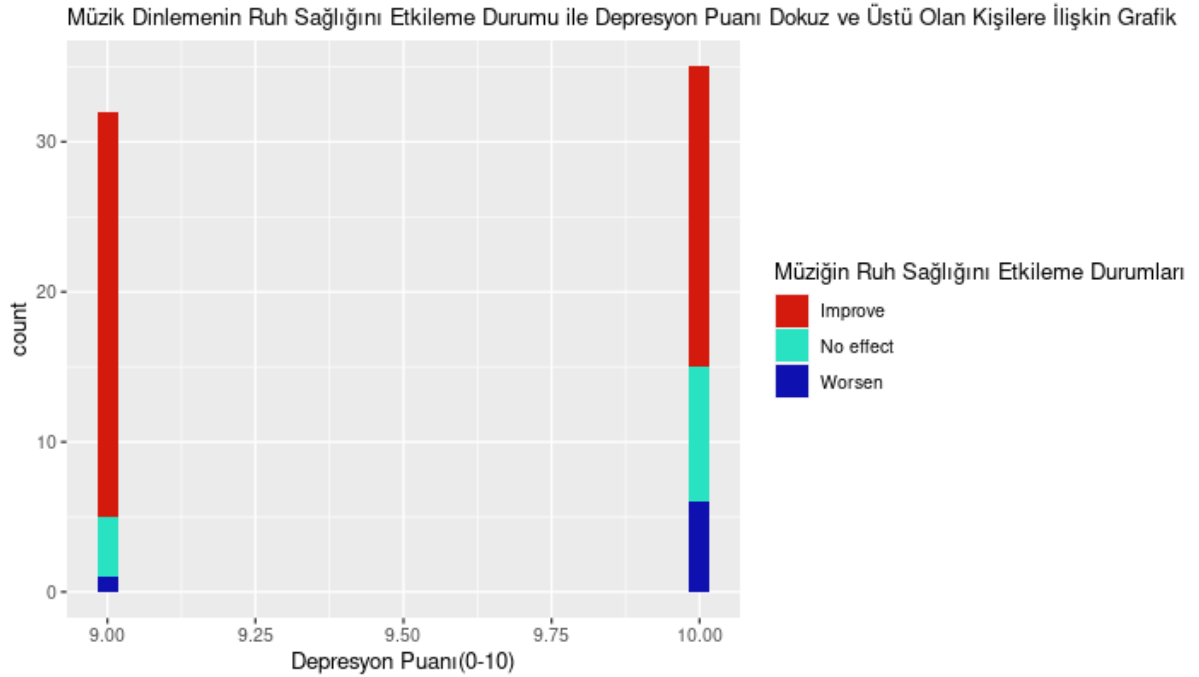




*Depression* değişkeninin 9 ve 10 puanlamada, müzik dinlemenin ruh sağlığını nasıl etkilediğini veren grafik.



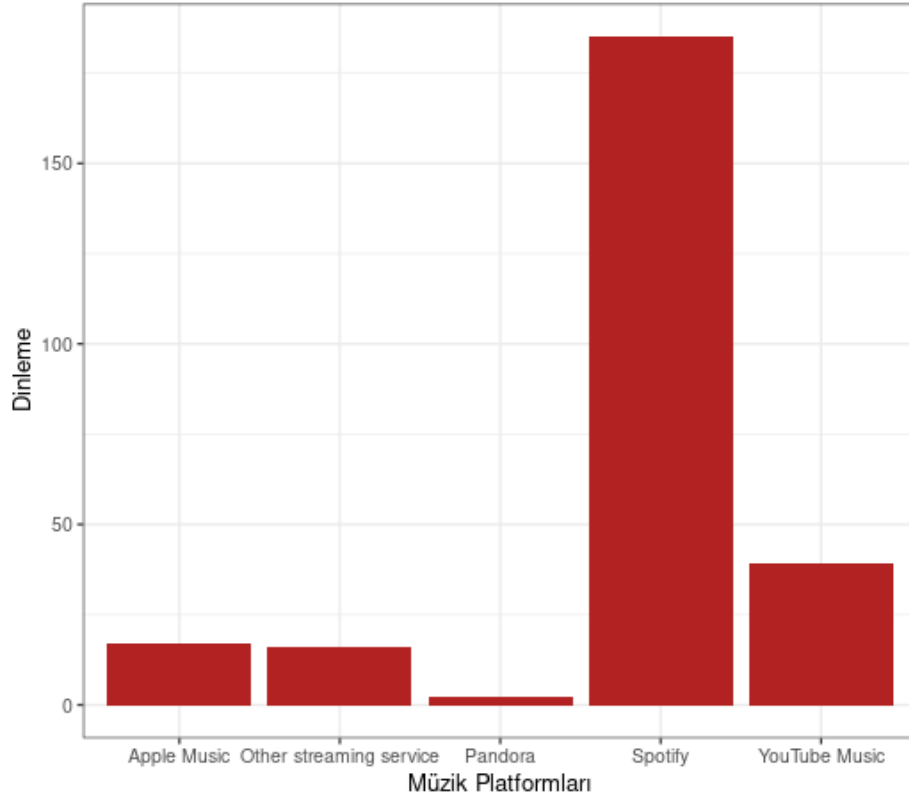
```
> music_yeni_veri %>%
+ filter(!Favgenre %in% "R&B") %>%
+ filter(Depression >= 9L & Depression <= 10L) %>%
+ ggplot() +
+ aes(x = Depression, fill = Musiceffects) +
+ geom_histogram(bins = 30L) +
+ scale_fill_manual(values = c(Improve = "#D41A0D",
+ 'No effect' = "#2B8E2C", Worsen = "#0E11AF")) +
+ labs(x = "Depresyon Puanı(0-10)", subtitle = "Müzik Dinlemenin Ruh Sağlığını Etkileme Durumu ile Depresyon Puanı Dokuz ve Üstü Olan Kişilere İlişkin Grafik",
+ fill = "Müziğin Ruh Sağlığını Etkileme Durumları") +
+ theme_gray()
> |
```



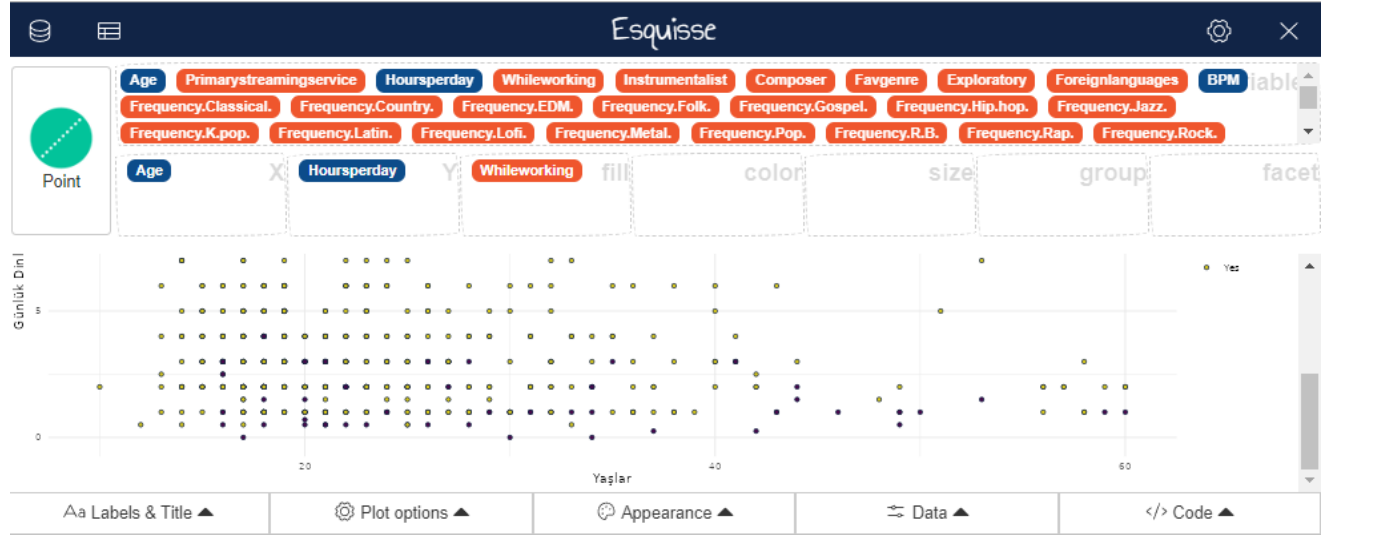
Kendilerine çok yüksek depresyon puanı veren katılımcılar, çoğunlukla müziğin ruh sağlığını iyi etkilediği seçeneği seçmişlerdir

```
> music_yeni_veri %>%
+ filter(!Primarystreamingservice %in% "I do not use a streaming service.") %>%
+ filter(Favgenre %in%
+ c("Jazz", "Rock", "Pop", "Classical")) %>%
+ ggplot() +
+ aes(x = Primarystreamingservice) +
+ geom_bar(fill = "#B22222") +
+ labs(x = "Müzik Platformları",
+ y = "Dinleme") +
+ theme_bw()
> |
```

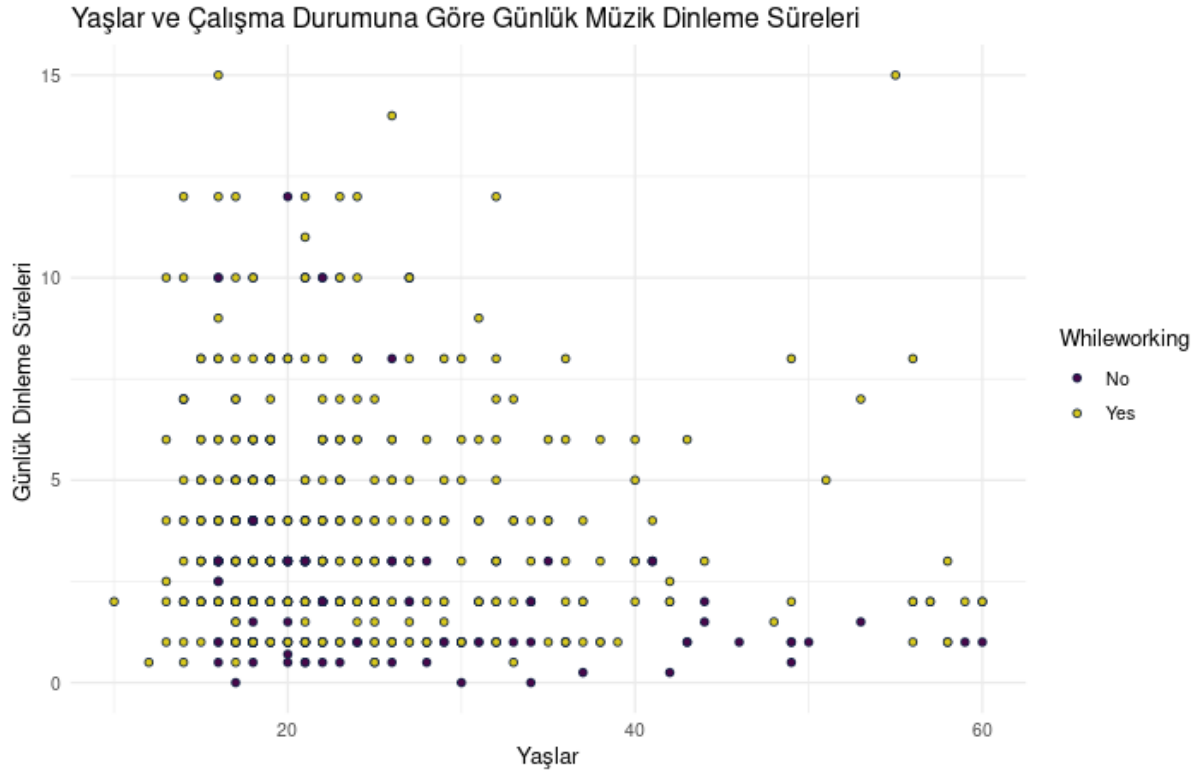
Müzik dinleme servisi kullanmayanları dışarda bıraktığımız ve caz, pop, rock, klasik müzik dinleyenlerin en çok hangi platformları dinledikleri aşağıdaki gibidir.



Bu dört müzik türünü, diğer müzik türlerinde olduğu gibi, Spotify platformundan dinlendiği söylenebilir.



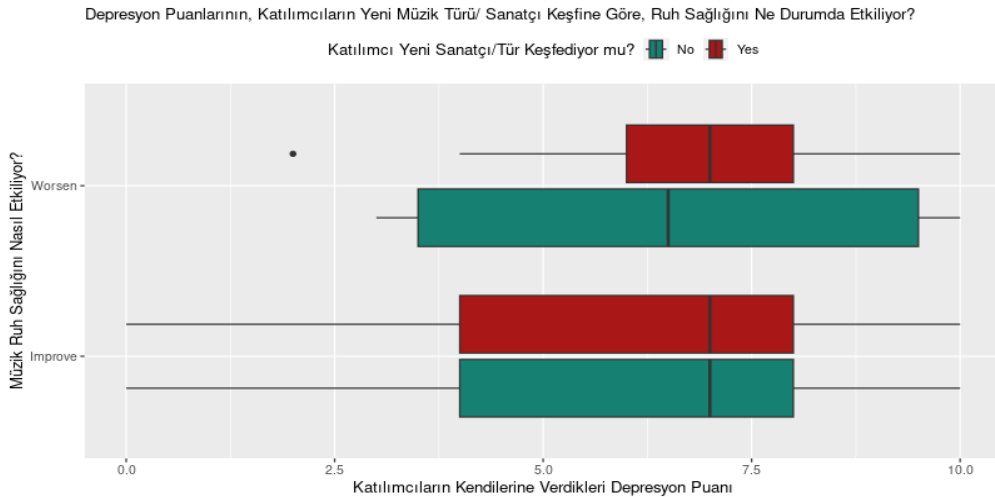
```
> music_yeni_veri %>%
+ filter(!(Favgenre %in% "R&B")) %>%
+ ggplot() +
+ aes(x = Age, y = Hoursperday, fill = Whileworking) +
+ geom_point(shape = "circle filled",
+ size = 1.5, colour = "#112446") +
+ scale_fill_manual(values = c(No = "#4A074A", Yes = "#D8C81C")) +
+ labs(x = "Yaşlar", y = "Günlük Dinleme Süreleri", title = "Yaşlar ve Çalışma Durumuna Göre Günlük Müzik Dinleme Süreleri") +
+ theme_minimal() +
+ theme(plot.caption = element_text(size = 25L, face = "italic", hjust = 0.5))
> |
```



Katılımcıların çoğunun, çalışırken müzik dinlediğini, sarı renklerden anlayabiliyoruz.

Müzik dinlemenin ruh sağlığını etkileme durumunu iki kategori haline getirip(kötü-iyi), depresyon puanlarına göre durumunu incelemek istediğimiz grafik:

```
> music_yeni_veri %>%
+ filter(!(Favgenre %in% "R&B")) %>%
+ filter(!(Musiceffects %in% "No effect")) %>%
+ ggplot() +
+ aes(x = Anxiety, y = Musiceffects, fill = Exploratory) +
+ geom_boxplot() +
+ scale_fill_manual(values = c(No = "#168173",
+ Yes = "#A17171")) +
+ labs(x = "Katılımcıların Kendilerine Verdikleri Depresyon Puanı", y = "Müzik Ruh Sağlığını Nasıl Etkiliyor?",
+ subtitle = "Depresyon Puanlarının, Katılımcıların Yeni Müzik Türü/ Sanatçı Keşfine Göre, Ruh Sağlığını Ne Durumda Etkiliyor?",
+ fill = "Katılımcı Yeni Sanatçı/Tür Keşfediyor mu?") +
+ theme_gray() +
+ theme(legend.position = "top",
+ plot.caption = element_text(size = 25L, face = "italic", hjust = 0.5))
>
```



Katılımcıların müziğin ruh sağlığına *olumlu* etkisi olduğunu söyledikleri durumda, yeni keşiflerin *yapılması* da *yapılmaması* da ortalama olarak 7 puan civarında, kendilerine verilen depresyon puanı ile sonuçlandırılmıştır.

Katılımcıların müziğin ruh sağlığına *kötü* etkisi olduğunu söyledikleri durumda, yeni keşif *yapılmaması* ortalaması 6,5 civarında kendilerine verilen depresyon puanı ile sonuçlanmıştır. Yeni keşif *yapılması* durumunda ise ortalama 7 civarında kendilerine verilen depresyon puanı ile sonuçlanmıştır.

Katılımcıların yeni müzik türü keşfetmesi ile müziğin ruh sağlığının üstündeki durum, katılımcıların kendine verdikleri depresyon puanlarını çok etkilememiştir.

## KAYNAKÇA

[Music & Mental Health EDA | Kaggle](#)

[Music & Mental Health Survey Results | Kaggle](#)

[Music and Mental Health Analysis | Kaggle](#)

<https://bookdown.org/content/2096/dplyr-paketi.html>

<https://www.newslabturkey.org/2018/10/28/r-dplyr-paketi-ile-veri-nasil-islenir-ve-analiz-edilir/>

<https://rpubs.com/sadettind/dplyr>

<https://study.com/academy/lesson/summarize-function-in-r-programming.html>

<https://teknoloji.org/r-ile-veri-manipulasyonu-tidyverse-dplyr/>

<https://cran.r-project.org/web/packages/esquisse/vignettes/get-started.html>

<https://www.exploringdata.org/post/how-to-explore-data-dataexplorer-package/>

<https://cran.r-project.org/web/packages/DataExplorer/DataExplorer.pdf>

[https://www.google.com/search?q=what+means+data+explorer+in+r+programming&sxsrf=ALiCzsZxiBb\\_0ONm7WsbmZxt6Gv3CSoxkw%3A1671805738522&ei](https://www.google.com/search?q=what+means+data+explorer+in+r+programming&sxsrf=ALiCzsZxiBb_0ONm7WsbmZxt6Gv3CSoxkw%3A1671805738522&ei)

<http://boxuancui.github.io/DataExplorer/>

[DataExplorer::dataexplorer-intro](#)

<https://www.youtube.com/watch?v=WzCyY9CyqUY>

<https://www.geeksforgeeks.org/esquisse-package-in-r-programming/>

<https://cran.r-project.org/web/packages/esquisse/vignettes/get-started.html>

<https://www.rdocumentation.org/packages/tibble/versions/3.0.1/topics/glimpse>

<https://sparkbyexamples.com/r-programming/r-group-by-function-from-dplyr/>

[configure\\_report function - RDocumentation](#)

[Esquisse Package in R Programming - GeeksforGeeks](#)

[https://www.youtube.com/watch?v=ONrGJF\\_8onw](https://www.youtube.com/watch?v=ONrGJF_8onw)

<https://www.youtube.com/watch?v=WzCyY9CyqUY>

<https://www.youtube.com/watch?v=VZjysaSfpiU>