



Varsayımlar



Veri inceleme




Dr. Kübra Atalay Kabasakal
Bahar 2023

İlk olarak veriyi okuma

```
library(haven)
screen <- read_sav("SCREEN.sav")
head(screen)
```

```
## # A tibble: 6 × 7
##   subno timedrs attdrug atthouse income mstatus  race
##   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl> <dbl>
## 1     1     1     8     27     5     2     1
## 2     2     3     7     20     6     2     1
## 3     3     0     8     23     3     2     1
## 4     4    13     9     28     8     2     1
## 5     5    15     7     24     1     2     1
## 6     6     3     8     25     4     2     1
```

- Bu sunumda  SCREEN.sav adlı veri seti kullanılmıştır. Bu veri setinde 6 değişken, 20-59 yaşları arasında 465 kadın vardır. Değişkenlerden timedrs, attdrug, atthouse ve income değişkenleri sürekli, mstatus ve race değişkenleri ise iki kategorili değişkenlerdir.

Veri Setini inceleme

- Daha sonra ise veri setindeki maksimum ve minimum değerleri belirleme

```
summary(screen)
```

```
##      subno      timedrs      attdrug      atthouse
## Min.      : 1.0      Min.      : 0.000      Min.      : 5.000      Min.      : 2.00
## 1st Qu.:137.0      1st Qu.: 2.000      1st Qu.: 7.000      1st Qu.:21.00
## Median :314.0      Median : 4.000      Median : 8.000      Median :24.00
## Mean      :317.4      Mean      : 7.901      Mean      : 7.686      Mean      :23.54
## 3rd Qu.:483.0      3rd Qu.:10.000      3rd Qu.: 9.000      3rd Qu.:27.00
## Max.      :758.0      Max.      :81.000      Max.      :10.000      Max.      :35.00
##                                     NA's      :1
##      income      mstatus      race
## Min.      : 1.00      Min.      :1.000      Min.      :1.000
## 1st Qu.: 2.50      1st Qu.:2.000      1st Qu.:1.000
## Median : 4.00      Median :2.000      Median :1.000
## Mean      : 4.21      Mean      :1.778      Mean      :1.088
## 3rd Qu.: 6.00      3rd Qu.:2.000      3rd Qu.:1.000
## Max.      :10.00      Max.      :2.000      Max.      :2.000
## NA's      :26
```

- Elde edilen değerlerin makul olduğu söylenebilir. Ancak bunu elde etmek için başka yollar da bulunmaktadır.

Veri Setini inceleme

- **psych** paketi ile inceleme daha ayrıntılı yapılabilir.

```
library(psych)
describe(screen[,-1])
```

```
##          vars    n  mean    sd median trimmed  mad min max range  skew
## timedrs      1 465  7.90 10.95      4    5.61 4.45    0  81   81  3.23
## attdrug      2 465  7.69  1.16      8    7.71 1.48    5  10    5 -0.12
## atthouse     3 464 23.54  4.48     24   23.62 4.45    2  35   33 -0.45
## income       4 439  4.21  2.42      4    4.01 2.97    1  10    9  0.58
## mstatus      5 465  1.78  0.42      2    1.85 0.00    1   2    1 -1.34
## race         6 465  1.09  0.28      1    1.00 0.00    1   2    1  2.90
##
##          kurtosis    se
## timedrs      12.88 0.51
## attdrug      -0.47 0.05
## atthouse      1.51 0.21
## income       -0.38 0.12
## mstatus      -0.21 0.02
## race         6.40 0.01
```

-  personality-project sayfasını daha fazla örnek için inceleyebilirsiniz

Veri Setini inceleme

- **gtsummary** paketi ile inceleme

```
library(gtsummary)
screen %>% select(2:6) %>%
tbl_summary(
  statistic = all_continuous() ~ c("{min}", "{max}"),
  missing = "always")
```

-  Presentation-Ready Summary Tables with gtsummary

Characteristic	N = 465 ¹
Visits to health professionals	0, 81
Unknown	0
Attitudes toward medication	
5	13 (2.8%)
6	60 (13%)
7	126 (27%)
8	149 (32%)
9	95 (20%)
10	22 (4.7%)
Unknown	0
Attitudes toward housework	2.0, 35.0
Unknown	1
Income	1.00, 10.00

Veri Setini inceleme

- **vtable** paketi ile inceleme

```
library(vtable)
sumtable(screen, summ=c('notNA(x)', 'min(x)', 'max(x)'))
```

Summary Statistics			
Variable	NotNA	Min	Max
subno	465	1	758
timedrs	465	0	81
attdrug	465	5	10
atthouse	464	2	35
income	439	1	10
mstatus	465	1	2
race	465	1	2

-  vtable paketi için örnekler

Veri Setini inceleme

- **vtable** paketi ile inceleme

```
st(screen, summ = c('notNA(x)', 'min(x)', 'max(x)'),  
summ.names = c('Frekans', 'Minimum', 'Maximum'))
```

Summary Statistics			
Variable	Frekans	Minimum	Maximum
subno	465	1	758
timedrs	465	0	81
attdrug	465	5	10
atthouse	464	2	35
income	439	1	10
mstatus	465	1	2
race	465	1	2

Veri Setini inceleme

- **kable** paketi ile **psych** paketi çıktılarını düzenleme

```
ozet <- describe(screen[,-1])  
kable(ozet,format='markdown',caption="Betimsel İstatistikler",digits=2)
```

Table: Betimsel İstatistikler

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
timedrs	1	465	7.90	10.95	4	5.61	4.45	0	81	81	3.23	12.88	0.51
attdrug	2	465	7.69	1.16	8	7.71	1.48	5	10	5	-0.12	-0.47	0.05
atthouse	3	464	23.54	4.48	24	23.62	4.45	2	35	33	-0.45	1.51	0.21
income	4	439	4.21	2.42	4	4.01	2.97	1	10	9	0.58	-0.38	0.12
mstatus	5	465	1.78	0.42	2	1.85	0.00	1	2	1	-1.34	-0.21	0.02
race	6	465	1.09	0.28	1	1.00	0.00	1	2	1	2.90	6.40	0.01

-  rmarkdown-cookbook

Ödev

- Önce datacampa üye olunuz.. Ödeviniz atanacaktır.

teşekkürler

