veri inceleme ve bazı görselleştirmeler

ilke

2022-07-27

```r
library(tidyverse)
```

```
## — Attaching packages ————————————————————————— tidyverse 1.3.2 —
## ✔ ggplot2 3.3.6     ✔ purrr   0.3.4
## ✔ tibble  3.1.7     ✔ dplyr   1.0.9
## ✔ tidyr   1.2.0     ✔ stringr 1.4.0
## ✔ readr   2.1.2     ✔ forcats 0.5.1
## — Conflicts ——————————————————————————— tidyverse_conflicts() —
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

```r
car<- mpg

summary(car)
```

```
##  manufacturer         model              displ          year
##  Length:234         Length:234         Min.  :1.600   Min.   :1999
##  Class :character   Class :character   1st Qu.:2.400  1st Qu.:1999
##  Mode  :character   Mode  :character   Median :3.300  Median :2004
##                                        Mean  :3.472   Mean   :2004
##                                        3rd Qu.:4.600  3rd Qu.:2008
##                                        Max.  :7.000   Max.   :2008
##       cyl           trans               drv                cty
##  Min.  :4.000   Length:234         Length:234         Min.  : 9.00
##  1st Qu.:4.000  Class :character   Class :character   1st Qu.:14.00
##  Median :6.000  Mode  :character   Mode  :character   Median :17.00
##  Mean  :5.889                                         Mean   :16.86
##  3rd Qu.:8.000                                        3rd Qu.:19.00
##  Max.  :8.000                                         Max.   :35.00
##       hwy            fl               class
##  Min.  :12.00   Length:234         Length:234
##  1st Qu.:18.00  Class :character   Class :character
##  Median :24.00  Mode  :character   Mode  :character
##  Mean  :23.44
##  3rd Qu.:27.00
##  Max.  :44.00
```

```r
colnames(car)
```

```
## [1] "manufacturer" "model"        "displ"        "year"         "cyl"
## [6] "trans"        "drv"          "cty"          "hwy"          "fl"
## [11] "class"
```

```r
rownames(car)
```

```
##   [1] "1"   "2"   "3"   "4"   "5"   "6"   "7"   "8"   "9"   "10"  "11"  "12"
##  [13] "13"  "14"  "15"  "16"  "17"  "18"  "19"  "20"  "21"  "22"  "23"  "24"
##  [25] "25"  "26"  "27"  "28"  "29"  "30"  "31"  "32"  "33"  "34"  "35"  "36"
##  [37] "37"  "38"  "39"  "40"  "41"  "42"  "43"  "44"  "45"  "46"  "47"  "48"
##  [49] "49"  "50"  "51"  "52"  "53"  "54"  "55"  "56"  "57"  "58"  "59"  "60"
##  [61] "61"  "62"  "63"  "64"  "65"  "66"  "67"  "68"  "69"  "70"  "71"  "72"
##  [73] "73"  "74"  "75"  "76"  "77"  "78"  "79"  "80"  "81"  "82"  "83"  "84"
##  [85] "85"  "86"  "87"  "88"  "89"  "90"  "91"  "92"  "93"  "94"  "95"  "96"
##  [97] "97"  "98"  "99"  "100" "101" "102" "103" "104" "105" "106" "107" "108"
## [109] "109" "110" "111" "112" "113" "114" "115" "116" "117" "118" "119" "120"
## [121] "121" "122" "123" "124" "125" "126" "127" "128" "129" "130" "131" "132"
## [133] "133" "134" "135" "136" "137" "138" "139" "140" "141" "142" "143" "144"
## [145] "145" "146" "147" "148" "149" "150" "151" "152" "153" "154" "155" "156"
## [157] "157" "158" "159" "160" "161" "162" "163" "164" "165" "166" "167" "168"
## [169] "169" "170" "171" "172" "173" "174" "175" "176" "177" "178" "179" "180"
## [181] "181" "182" "183" "184" "185" "186" "187" "188" "189" "190" "191" "192"
## [193] "193" "194" "195" "196" "197" "198" "199" "200" "201" "202" "203" "204"
## [205] "205" "206" "207" "208" "209" "210" "211" "212" "213" "214" "215" "216"
## [217] "217" "218" "219" "220" "221" "222" "223" "224" "225" "226" "227" "228"
## [229] "229" "230" "231" "232" "233" "234"
```

```
nrow(car)
```

```
## [1] 234
```

```
ncol(car)
```

```
## [1] 11
```

```
head(car)
```

```
## # A tibble: 6 × 11
##   manufacturer model displ  year   cyl trans      drv      cty   hwy fl    class
##   <chr>        <chr> <dbl> <int> <int> <chr>      <chr> <int> <int> <chr> <chr>
## 1 audi         a4      1.8  1999     4 auto(l5)   f        18    29 p     compa…
## 2 audi         a4      1.8  1999     4 manual(m5) f        21    29 p     compa…
## 3 audi         a4      2    2008     4 manual(m6) f        20    31 p     compa…
## 4 audi         a4      2    2008     4 auto(av)   f        21    30 p     compa…
## 5 audi         a4      2.8  1999     6 auto(l5)   f        16    26 p     compa…
## 6 audi         a4      2.8  1999     6 manual(m5) f        18    26 p     compa…
```

```
glimpse(car)
```

```
## Rows: 234
## Columns: 11
## $ manufacturer <chr> "audi", "audi", "audi", "audi", "audi", "audi", "audi", "…
## $ model        <chr> "a4", "a4", "a4", "a4", "a4", "a4", "a4", "a4 quattro", "…
## $ displ        <dbl> 1.8, 1.8, 2.0, 2.0, 2.8, 2.8, 3.1, 1.8, 1.8, 2.0, 2.0, 2.…
## $ year         <int> 1999, 1999, 2008, 2008, 1999, 1999, 2008, 1999, 1999, 200…
## $ cyl          <int> 4, 4, 4, 4, 6, 6, 6, 4, 4, 4, 4, 6, 6, 6, 6, 6, 6, 8, 8, …
## $ trans        <chr> "auto(l5)", "manual(m5)", "manual(m6)", "auto(av)", "auto…
## $ drv          <chr> "f", "f", "f", "f", "f", "f", "f", "4", "4", "4", "4", "4…
## $ cty          <int> 18, 21, 20, 21, 16, 18, 18, 18, 16, 20, 19, 15, 17, 17, 1…
## $ hwy          <int> 29, 29, 31, 30, 26, 26, 27, 26, 25, 28, 27, 25, 25, 25, 2…
## $ fl           <chr> "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p…
## $ class        <chr> "compact", "compact", "compact", "compact", "compact", "c…
```

```
fivenum(car$hwy) #diskteki ilk 5 farklıyı gösterir
```
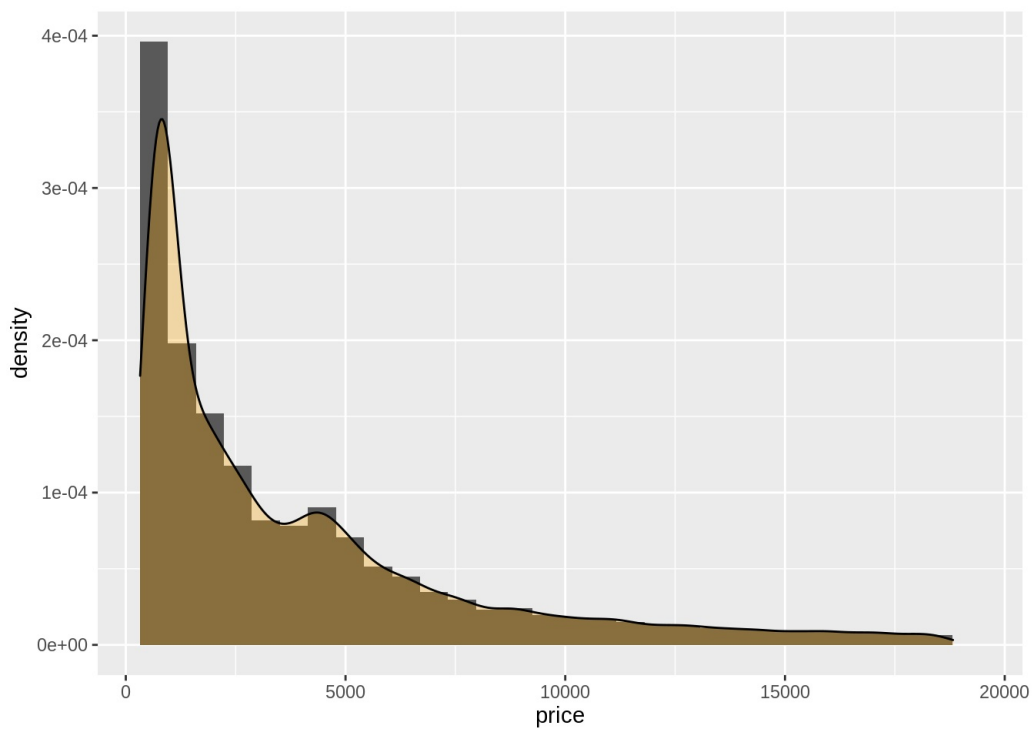
```
## [1] 12 18 24 27 44
```

```
d<- diamonds

ggplot(d, aes(x = cut, y = price, fill = cut)) +
        geom_violin(alpha = 0.6)                    #boxplottan farklı olarak frekansa duyarlı
```
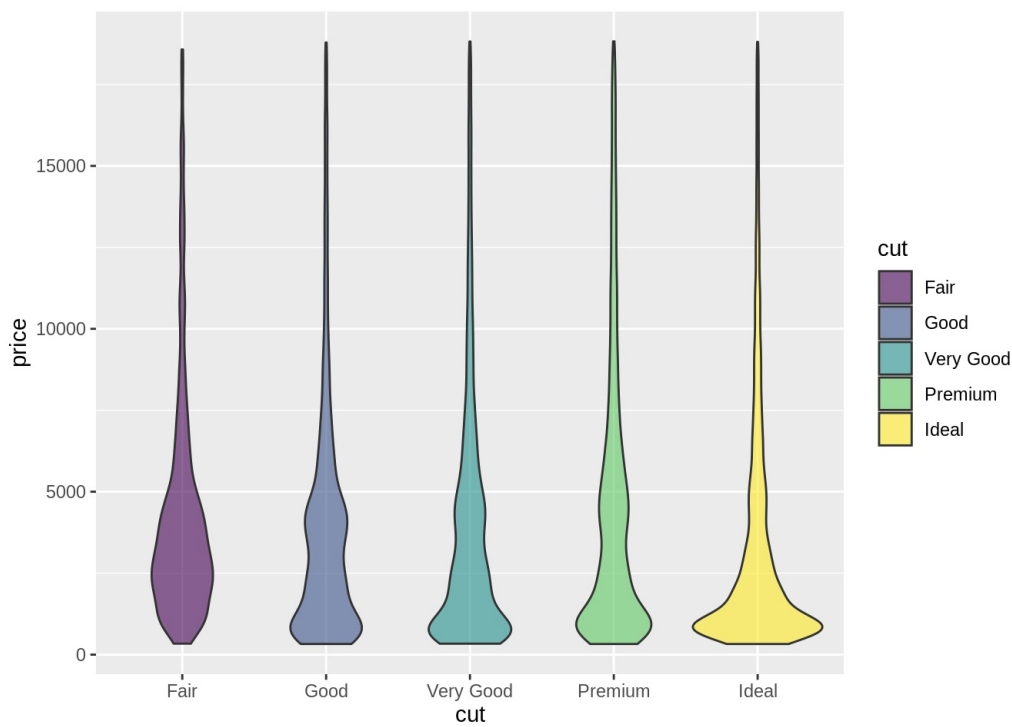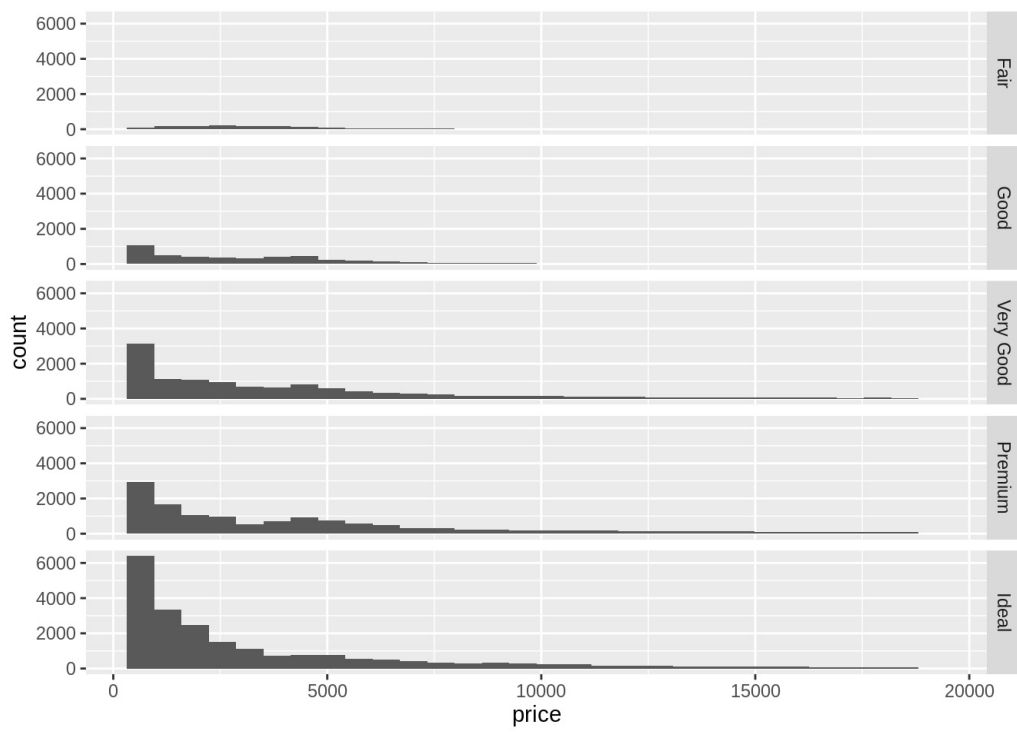
```
ggplot(d, aes(price)) +
  geom_histogram(aes(y = ..density..)) +
  geom_density(alpha = .3, fill = "orange")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
ggplot(d, aes(price)) +
  geom_histogram() +
  facet_grid(cut ~. )
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

dağılım-çoklu frekans

```
ggplot(d, aes(price, y = ..density..)) +
  geom_density(aes(colour = cut), binwidth = 500)
```

```
## Warning: Ignoring unknown parameters: binwidth
```