

Untitled

R Markdown

```
library(tidyverse)
```

```
## -- Attaching packages -----  
## v ggplot2 3.1.0      v purrr  0.3.1  
## v tibble  2.0.1      v dplyr  0.8.0.1  
## v tidyr   0.8.3      v stringr 1.4.0  
## v readr   1.3.1      v forcats 0.4.0
```

```
## -- Conflicts -----  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()     masks stats::lag()
```

```
library(rvest)
```

```
## Loading required package: xml2
```

```
##
```

```
## Attaching package: 'rvest'
```

```
## The following object is masked from 'package:purrr':
```

```
##
```

```
##   pluck
```

```
## The following object is masked from 'package:readr':
```

```
##
```

```
##   guess_encoding
```

```
raw_html <- read_html("https://bkm.com.tr/secilen-aya-ait-sektorel-gelisim/?filter_year=2018&filter_mon  
raw_html
```

```
## {xml_document}
```

```
## <html lang="tr-TR">
```

```
## [1] <head>\n<meta http-equiv="Content-Type" content="text/html; charset= ...
```

```
## [2] <body>\n    <div id="anaKapsul">\n                                <table width="735" ...
```

```
html_tables <- raw_html %>% html_table(fill=TRUE)
```

```
raw_df <- html_tables[[4]] %>% tbl_df()
```

```
raw_df
```

```
## # A tibble: 29 x 5
```

```
##   X1          X2          X3          X4          X5  
##   <chr>      <chr>      <chr>      <chr>      <chr>  
## 1 İşyeri Grubu İşlem Adedi İşlem Ade~ İşlem Tutarı (M~ İşlem Tutarı (Mi~  
## 2 İşyeri Grubu İşlem Aded~ İşlem Ade~ "İşlem Tutarı \~ "İşlem Tutarı \n~  
## 3 ARABA KİRALAMA 292.290    57.920    158,36      12,48  
## 4 ARAÇ KİRALAMA~ 2.725.497    540.511    1.920,41     100,50  
## 5 BENZİN VE YAK~ 25.260.026    8.000.259    4.641,49     567,45  
## 6 BİREYSEL EMEK~ 2.582.916     527        667,90       0,19  
## 7 ÇEŞİTLİ GIDA   24.041.009    12.196.630    3.483,63     477,84  
## 8 DOĞRUDAN PAZA~ 973.180       24.216       719,08        5,66  
## 9 EĞİTİM / KIRT~ 5.368.531     2.203.468    1.444,43     153,74  
## 10 ELEKTRİK-ELEK~ 7.016.143     1.933.429    2.984,89     298,30
```

```
## # ... with 19 more rows
raw_df %>% slice(-(1:2)) %>% filter(X1 != "TOPLAM") %>% mutate(X2_new = gsub("\\.", "", X2))

## # A tibble: 26 x 6
##   X1                X2      X3      X4      X5      X2_new
##   <chr>            <chr>    <chr>    <chr>    <chr> <chr>
## 1 ARABA KİRALAMA    292.290  57.920   158,36  12,48 292290
## 2 ARAÇ KİRALAMA-SATIŞ/SERVİS/YE~ 2.725.497 540.511   1.920,~ 100,~ 2725497
## 3 BENZİN VE YAKIT İSTASYONLARI 25.260.0~ 8.000.259 4.641,~ 567,~ 252600~
## 4 BİREYSEL EMEKLİLİK 2.582.916 527      667,90  0,19 2582916
## 5 ÇEŞİTLİ GIDA      24.041.0~ 12.196.6~ 3.483,~ 477,~ 240410~
## 6 DOĞRUDAN PAZARLAMA 973.180   24.216   719,08  5,66 973180
## 7 EĞİTİM / KIRTASIYE / OFİS MAL~ 5.368.531 2.203.468 1.444,~ 153,~ 5368531
## 8 ELEKTRİK-ELEKTRONİK EŞYA, BİL~ 7.016.143 1.933.429 2.984,~ 298,~ 7016143
## 9 GİYİM VE AKSESUAR 25.171.0~ 8.781.408 3.893,~ 870,~ 251710~
## 10 HAVAYOLLARI      1.974.862 347.591   1.013,~ 267,~ 1974862
## # ... with 16 more rows
raw_df %>% slice(-(1:2)) %>% filter(X1 != "TOPLAM") %>% mutate_at(vars(X2:X5),funs(gsub("\\.", "", .)))

## Warning: funs() is soft deprecated as of dplyr 0.8.0
## please use list() instead
##
## # Before:
## funs(name = f(.))
##
## # After:
## list(name = ~f(.))
## This warning is displayed once per session.

## # A tibble: 26 x 5
##   X1                X2      X3      X4      X5
##   <chr>            <chr>    <chr>    <chr>    <chr>
## 1 ARABA KİRALAMA    292290  57920   158,36  12,48
## 2 ARAÇ KİRALAMA-SATIŞ/SERVİS/YEDEK PARÇA 2725497 540511   1920,41 100,50
## 3 BENZİN VE YAKIT İSTASYONLARI 25260026 8000259 4641,49 567,45
## 4 BİREYSEL EMEKLİLİK 2582916 527      667,90  0,19
## 5 ÇEŞİTLİ GIDA      24041009 12196630 3483,63 477,84
## 6 DOĞRUDAN PAZARLAMA 973180   24216   719,08  5,66
## 7 EĞİTİM / KIRTASIYE / OFİS MALZEMELERİ 5368531 2203468 1444,43 153,74
## 8 ELEKTRİK-ELEKTRONİK EŞYA, BİLGİSAYAR 7016143 1933429 2984,89 298,30
## 9 GİYİM VE AKSESUAR 25171042 8781408 3893,09 870,49
## 10 HAVAYOLLARI      1974862 347591   1013,06 267,60
## # ... with 16 more rows
raw_df %>% slice(-(1:2)) %>% filter(X1 != "TOPLAM") %>% mutate_at(vars(X2:X5),funs(gsub("\\.", "", .))) %>%

## # A tibble: 26 x 5
##   X1                X2      X3      X4      X5
##   <chr>            <dbl>    <dbl> <dbl>    <dbl>
## 1 ARABA KİRALAMA    292290  57920  158.   12.5
## 2 ARAÇ KİRALAMA-SATIŞ/SERVİS/YEDEK PARÇA 2725497 540511 1920.  100.
## 3 BENZİN VE YAKIT İSTASYONLARI 25260026 8000259 4641.  567.
## 4 BİREYSEL EMEKLİLİK 2582916 527     668.   0.19
## 5 ÇEŞİTLİ GIDA      24041009 12196630 3484.  478.
## 6 DOĞRUDAN PAZARLAMA 973180   24216  719.   5.66
```

```
## 7 EĞİTİM / KIRTASIYE / OFİS MALZEMELERİ 5368531 2203468 1444. 154.
## 8 ELEKTRİK-ELEKTRONİK EŞYA, BİLGİSAYAR 7016143 1933429 2985. 298.
## 9 GİYİM VE AKSESUAR 25171042 8781408 3893. 870.
## 10 HAVAYOLLARI 1974862 347591 1013. 268.
## # ... with 16 more rows
```

```
cards_df <- raw_df %>% slice(-(1:2)) %>% filter(X1 != "TOPLAM") %>% mutate_at(vars(X2:X5), funs(gsub("\\",
colnames(cards_df) <- c("tur", "kk_ia", "bk_ia", "kk_it", "bk_it")
cards_df
```

```
## # A tibble: 26 x 5
##   tur          kk_ia    bk_ia kk_it  bk_it
##   <chr>      <dbl>    <dbl> <dbl> <dbl>
## 1 ARABA KİRALAMA 292290    57920 158. 12.5
## 2 ARAÇ KİRALAMA-SATIŞ/SERVİS/YEDEK PARÇA 2725497    540511 1920. 100.
## 3 BENZİN VE YAKIT İSTASYONLARI 25260026 8000259 4641. 567.
## 4 BİREYSEL EMEKLİLİK 2582916    527 668. 0.19
## 5 ÇEŞİTLİ GIDA 24041009 12196630 3484. 478.
## 6 DOĞRUDAN PAZARLAMA 973180    24216 719. 5.66
## 7 EĞİTİM / KIRTASIYE / OFİS MALZEMELERİ 5368531 2203468 1444. 154.
## 8 ELEKTRİK-ELEKTRONİK EŞYA, BİLGİSAYAR 7016143 1933429 2985. 298.
## 9 GİYİM VE AKSESUAR 25171042 8781408 3893. 870.
## 10 HAVAYOLLARI 1974862 347591 1013. 268.
## # ... with 16 more rows
```

```
cards_df %>% arrange(desc(kk_ia)) %>% slice(1:5) %>% select(tur, kk_ia)
```

```
## # A tibble: 5 x 2
##   tur          kk_ia
##   <chr>      <dbl>
## 1 MARKET VE ALIŞVERİŞ MERKEZLERİ 87585116
## 2 YEMEK 34477734
## 3 BENZİN VE YAKIT İSTASYONLARI 25260026
## 4 GİYİM VE AKSESUAR 25171042
## 5 ÇEŞİTLİ GIDA 24041009
```

```
cards_df %>% mutate(kk_br = kk_it*10^6/kk_ia) %>% arrange(desc(kk_br)) %>% slice(1:5) %>% select(tur, kk_br)
```

```
## # A tibble: 5 x 2
##   tur          kk_br
##   <chr>      <dbl>
## 1 MÜTEAHHİT İŞLERİ 1068.
## 2 KUYUMCULAR 833.
## 3 DOĞRUDAN PAZARLAMA 739.
## 4 ARAÇ KİRALAMA-SATIŞ/SERVİS/YEDEK PARÇA 705.
## 5 YAPI MALZEMELERİ, HIRDAVAT, NALBURİYE 685.
```