

Across platforms for data manipulation

Sample dataset: mtcars (it already exists in the R installation)

```
mtcars <- data(mtcars)
```

Goal: Arrange the dataset in descending order of horse power, select the observations with more than 4 cylinders (> 4, not >= 4), exclude the “vs” column, create a new column where you convert miles per gallon to km per litre (multiply miles per gallon by 0.354) and group observations by transmission type (automatic = 0, manual = 1)

\$ + [] + which()

```
cars.ord <- mtcars[order(-mtcars$hp),]  
cars.4 <- cars.ord[which(cars.ord$cyl > 4),]  
cars.lc <- subset(cars.4, select = -vs)  
cars.mc <- cbind(cars.lc, km1 = cars.lc$mpg*0.354)  
mean(cars.mc$mpg[which(cars.mc$am == 0)])  
sd(cars.mc$mpg[which(cars.mc$am == 0)])  
mean(cars.mc$mpg[which(cars.mc$am == 1)])  
sd(cars.mc$mpg[which(cars.mc$am == 1)])  
[1] 16.06875  
[1] 3.08193  
[1] 18.5  
[1] 2.893095
```

with() + [] + subset()

```
cars.ord <- with(mtcars, mtcars[order(-hp),])  
cars.4 <- subset(cars.ord, cyl > 4)  
cars.lc <- subset(cars.4, select = -vs)  
cars.mc <- with(cars.lc, cbind(cars.lc, km1 = mpg*0.354))  
with(cars.mc, aggregate(mpg, by = list(am), FUN = mean))  
with(cars.mc, aggregate(mpg, by = list(am), FUN = sd))  
  
Group.1      x  
1          0 16.06875  
2          1 18.50000  
  
Group.1      x  
1          0 3.081930  
2          1 2.893095
```

%>% + {dplyr}

```
mtcars %>%  
  arrange(desc(hp)) %>%  
  filter(cyl > 4) %>%  
  select(-vs) %>%  
  mutate(km1 = mpg*0.354) %>%  
  group_by(am) %>%  
  summarise(mpg.m = mean(mpg), mpg.sd = sd(mpg))  
# A tibble: 2 x 3  
  am mpg.m mpg.sd  
<dbl> <dbl> <dbl>  
1  0     16.1  3.08  
2  1     18.5  2.89
```