

W6 practice

2023-02-16

0. data step

```
library(haven); library(psych); library(dplyr); library(tidyr);
```

```
##  
## Attaching package: 'dplyr'  
  
## The following objects are masked from 'package:stats':  
##  
##   filter, lag  
  
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
library(magrittr); library(ggplot2); library(gridExtra)
```

```
##  
## Attaching package: 'magrittr'  
  
## The following object is masked from 'package:tidyr':  
##  
##   extract  
  
##  
## Attaching package: 'ggplot2'  
  
## The following objects are masked from 'package:psych':  
##  
##   %+%, alpha  
  
##  
## Attaching package: 'gridExtra'  
  
## The following object is masked from 'package:dplyr':  
##  
##   combine
```

```
library(rstatix)
```

```
##
```

```
## Attaching package: 'rstatix'
```

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

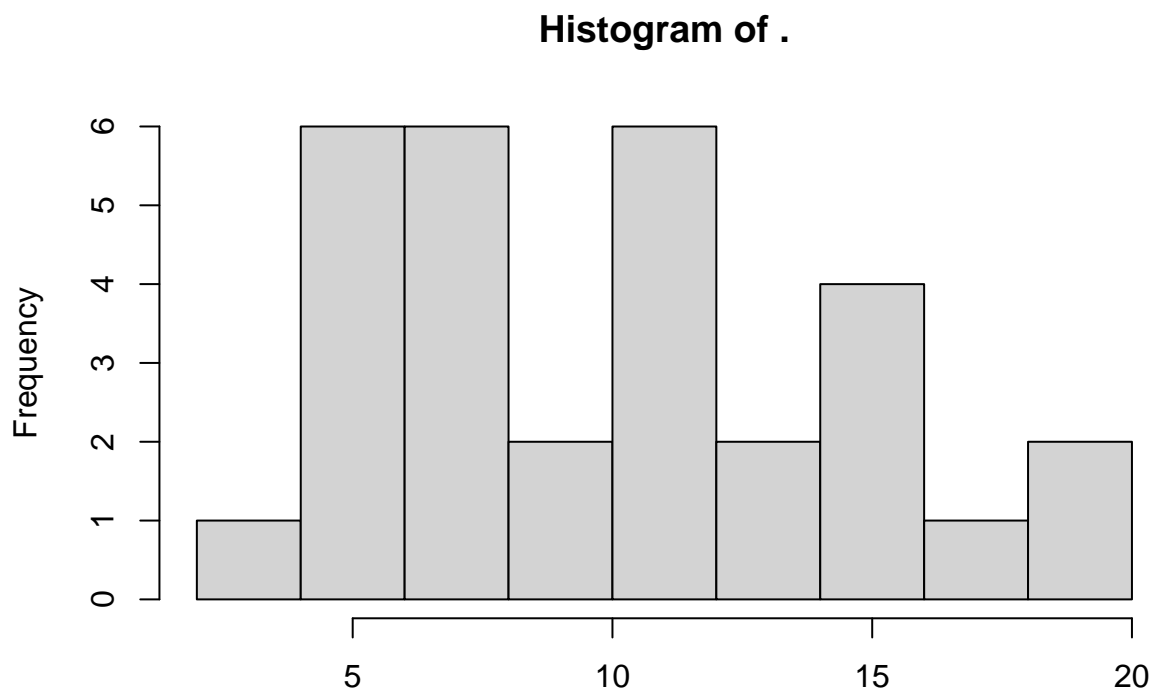
```
##
```

```
## filter
```

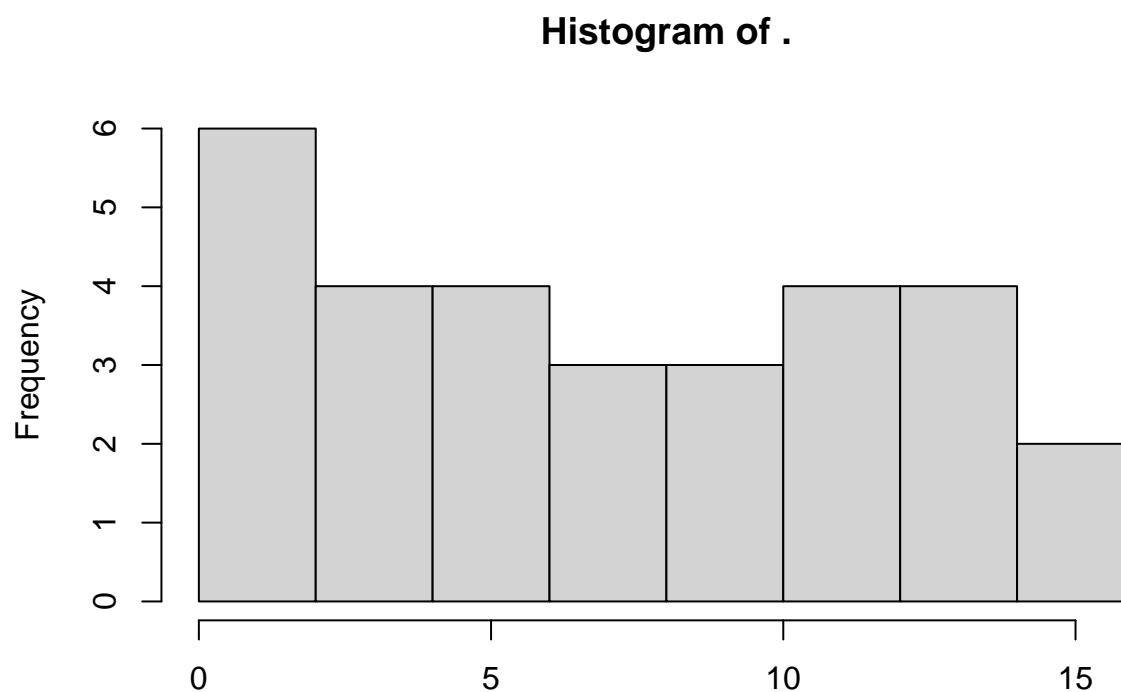
```
drugtest =  
  data.frame(drug = factor(rep(c("A", "D", "F"), each = 2, times = 5),  
              levels = c("F", "A", "D")), # F as reference  
    pretreatment = c(11, 6, 6, 8, 16, 16, 8, 10, 6, 19, 13, 12,  
                     5, 6, 7, 8, 11, 12, 14, 11, 8, 5, 9, 7,  
                     19, 3, 18, 15, 16, 12),  
    posttreatment = c(6, 4, 2, 4, 13, 12, 0, 13, 2, 14, 10, 8,  
                      2, 1, 5, 9, 16, 12, 8, 8, 4, 4, 5, 6,  
                      11, 0, 12, 9, 15, 14))
```

```
# histogram
```

```
drugtest$pretreatment %>% hist
```



```
drugtest$postttreatment %>% hist
```



```
# means & vaiances
drugtest %>% describeBy("drug") # easy way
```

```
##
## Descriptive statistics by group
## drug: F
##      vars  n mean  sd median trimmed  mad min max range  skew
## drug*      1 10  1.0 0.00      1   1.00 0.00   1  1   0  NaN
## pretreatment  2 10 12.4 3.03     12  12.62 2.97   7 16   9 -0.21
## postttreatment 3 10 11.1 3.75     12  11.25 3.71   5 16  11 -0.34
##      kurtosis  se
## drug*      NaN 0.00
## pretreatment -1.22 0.96
## postttreatment -1.47 1.19
## -----
## drug: A
##      vars  n mean  sd median trimmed  mad min max range  skew kurtosis
## drug*      1 10  2.0 0.00      2   2.00 0.00   2  2   0  NaN      NaN
## pretreatment  2 10  9.3 4.76      9   8.88 4.45   3 19  16 0.58    -0.76
## postttreatment 3 10  5.3 4.64      5   5.00 5.19   0 13  13 0.27    -1.56
##      se
## drug*      0.00
## pretreatment 1.51
```

```
## posttreatment 1.47
## -----
## drug: D
##
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis
## drug*	1	10	3.0	0.00	3.0	3.00	0.00	3	3	0	NaN	NaN
## pretreatment	2	10	10.0	5.25	8.0	9.50	2.97	5	19	14	0.73	-1.35
## posttreatment	3	10	6.5	4.22	4.5	6.12	3.71	2	14	12	0.51	-1.40

```
##
```

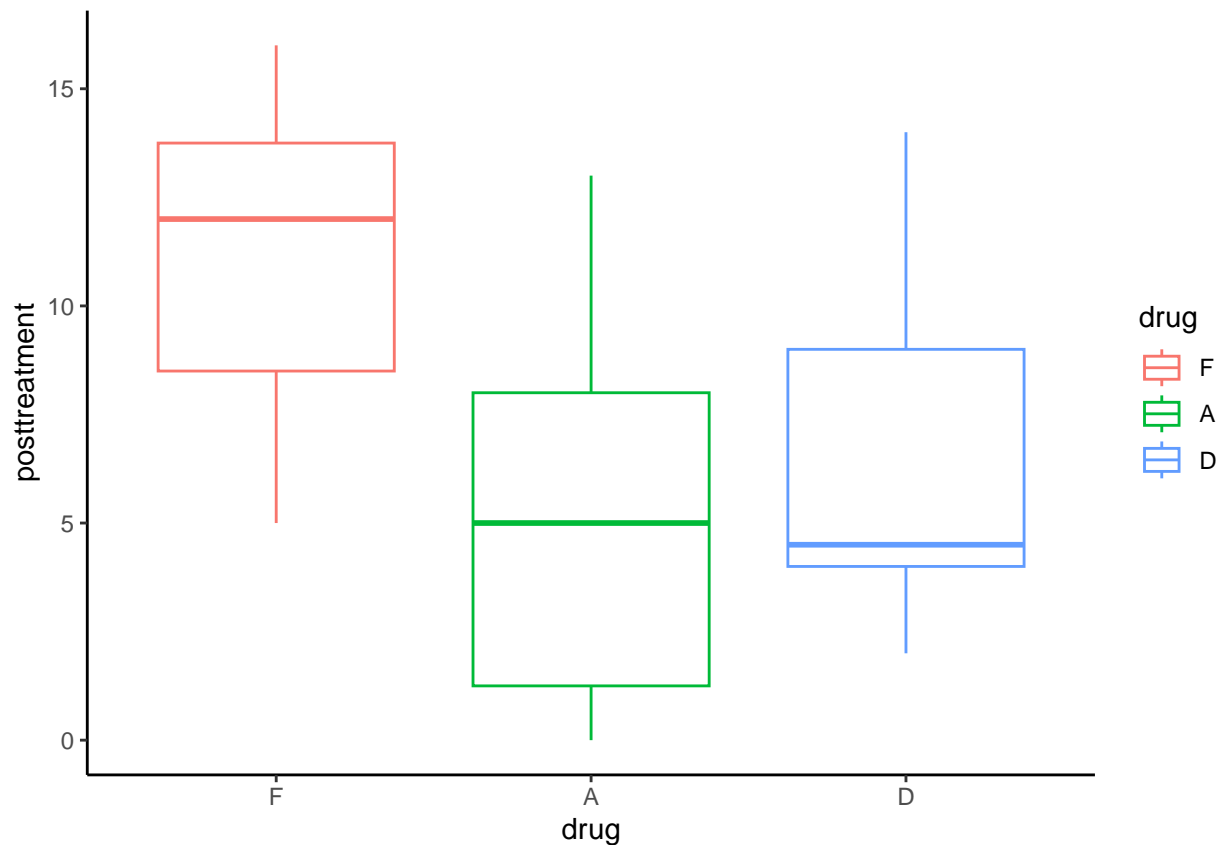
	se
## drug*	0.00
## pretreatment	1.66
## posttreatment	1.34

```
drugtest %>% # another way
  gather(key = treatment, value = value, -drug) %>%
  mutate(treatment = factor(treatment)) %>%
  group_by(drug, treatment) %>%
  summarise(n = n(),
            mean = mean(value),
            median = median(value),
            sd = sd(value),
            se = sd/sqrt(n))
```

'summarise()' has grouped output by 'drug'. You can override using the
'.groups' argument.

```
## # A tibble: 6 x 7
## # Groups:   drug [3]
##   drug treatment      n mean median    sd    se
##   <fct> <fct>      <int> <dbl> <dbl> <dbl> <dbl>
## 1 F     posttreatment    10  11.1    12   3.75  1.19
## 2 F     pretreatment     10  12.4    12   3.03  0.957
## 3 A     posttreatment    10   5.3     5   4.64  1.47
## 4 A     pretreatment     10   9.3     9   4.76  1.51
## 5 D     posttreatment    10   6.5     4.5  4.22  1.34
## 6 D     pretreatment     10  10      8   5.25  1.66
```

```
# box plot
drugtest %>%
  ggplot(aes(x = drug, y = posttreatment, group = drug, color = drug)) +
  geom_boxplot() +
  theme_classic()
```



1. Run ANVOCA

```
# unadjusted
fit1 = lm(posttreatment ~ drug, data = drugtest)
summary(fit1)
```

```
##
## Call:
## lm(formula = posttreatment ~ drug, data = drugtest)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.10  -3.25  -0.20   2.70   7.70
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    11.100     1.335    8.312 6.39e-09 ***
## drugA          -5.800     1.889   -3.071  0.00482 **
## drugD          -4.600     1.889   -2.436  0.02174 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.223 on 27 degrees of freedom
```

```
## Multiple R-squared:  0.2802, Adjusted R-squared:  0.2269
## F-statistic: 5.256 on 2 and 27 DF,  p-value: 0.01181
```

```
confint(fit1)
```

```
##              2.5 %      97.5 %
## (Intercept)  8.359955 13.8400453
## drugA        -9.675009 -1.9249907
## drugD        -8.475009 -0.7249907
```

```
anova(fit1)
```

```
## Analysis of Variance Table
##
## Response: posttreatment
##          Df Sum Sq Mean Sq F value    Pr(>F)
## drug       2 187.47   93.733   5.2561 0.01181 *
## Residuals 27 481.50   17.833
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# adjusted
fit2 = lm(posttreatment ~ drug + pretreatment, data = drugtest)
summary(fit2)
```

```
##
## Call:
## lm(formula = posttreatment ~ drug + pretreatment, data = drugtest)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.330 -1.516 -0.693  1.189  7.178
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.8505     1.6568   1.117  0.27426
## drugA          -3.4876     1.2424  -2.807  0.00935 **
## drugD          -2.8098     1.2217  -2.300  0.02973 *
## pretreatment    0.7459     0.1151   6.481  7.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.661 on 26 degrees of freedom
## Multiple R-squared:  0.7248, Adjusted R-squared:  0.693
## F-statistic: 22.83 on 3 and 26 DF,  p-value: 1.875e-07
```

```
confint(fit2)
```

```
##              2.5 %      97.5 %
## (Intercept) -1.5551316  5.2560483
## drugA        -6.0413071 -0.9339221
## drugD        -5.3209116 -0.2986207
## pretreatment  0.5093459  0.9825157
```

```
anova(fit2)
```

```
## Analysis of Variance Table
##
## Response: posttreatment
##           Df Sum Sq Mean Sq F value    Pr(>F)
## drug         2  187.47   93.733   13.238 0.0001084 ***
## pretreatment 1  297.40  297.403   42.002 7.205e-07 ***
## Residuals    26  184.10    7.081
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# get adjusted means
tmp.adj=
  drugtest %>%
  group_by(pretreatment) %>%
  emmeans_test(posttreatment ~ drug, detailed = T) # multiplied comparison
get_emmeans(tmp.adj) # getting adjusted means
```

```
## # A tibble: 3 x 8
##   pretreatment drug  emmean    se    df conf.low conf.high method
##   <dbl> <fct>    <dbl> <dbl> <dbl>   <dbl>   <dbl> <chr>
## 1     10.6 F      9.66 1.04   24     7.52    11.8 Emmeans test
## 2     10.6 A      6.24 0.909  24     4.37     8.12 Emmeans test
## 3     10.6 D      6.92 0.881  24     5.10     8.73 Emmeans test
```