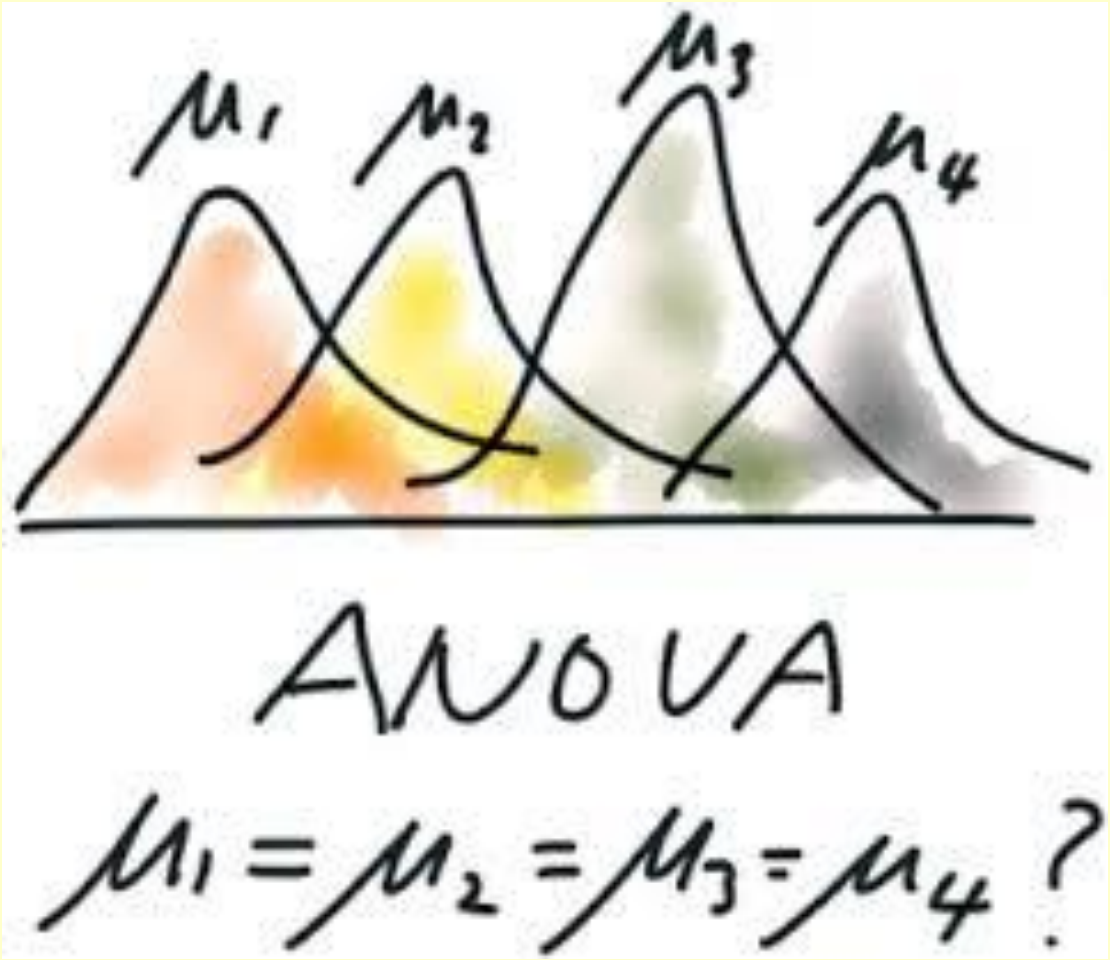
ANOVA - Implementation

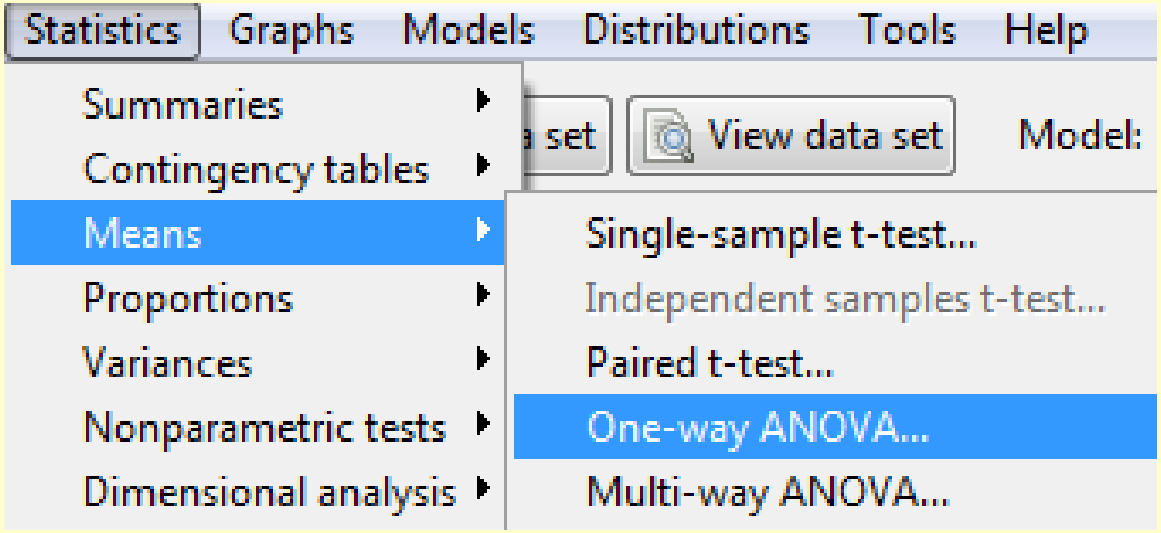


http://www.pelagicos.net/classes_biometry_fa17.htm

Doing an ANOVA – With RCmdr

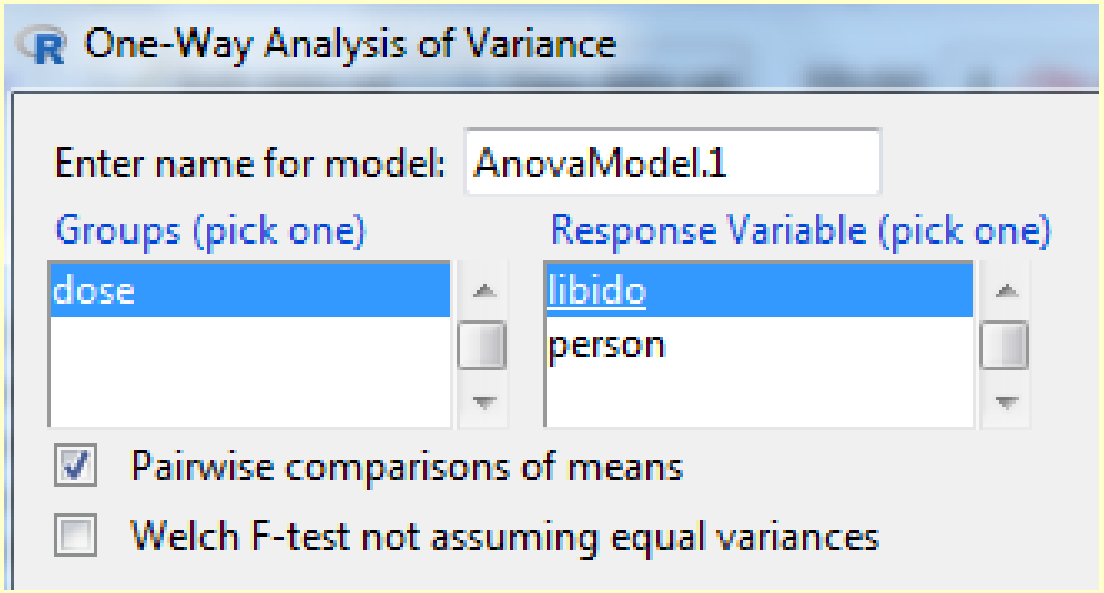
Categorical Variable

	person	dose	libido
1	1	low	3
2	2	low	2
3	3	low	1
4	4	low	1
5	5	low	4
6	6	mid	5
7	7	mid	2
8	8	mid	4
9	9	mid	2
10	10	mid	3
11	11	high	7
12	12	high	4
13	13	high	5
14	14	high	3
15	15	high	6



One-Way ANOVA
Testing a single Factor – "dose"
with 3 treatments (low, mid, high)

Doing an ANOVA – With RCmdr



Assume Equal Variances? YES

What Happens After Significant ANOVA?

Compare means

Doing an ANOVA – Output

```
Rcmdr> AnovaModel.1 <- aov(libido ~ dose, data=viagra)

Rcmdr> summary(AnovaModel.1)

            Df   SumSq   MeanSq  Fvalue  Pr(>F)
dose         2   20.13   10.067  5.119   0.0247 *
Residuals 12   23.60    1.967

--- Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Rcmdr> with(Dataset, numSummary(libido, groups=dose,
statistics=c("mean", "sd")))

                mean       sd    data:n
high    5.0    1.581139      5
low     2.2    1.303840      5
mid     3.2    1.303840      5
```

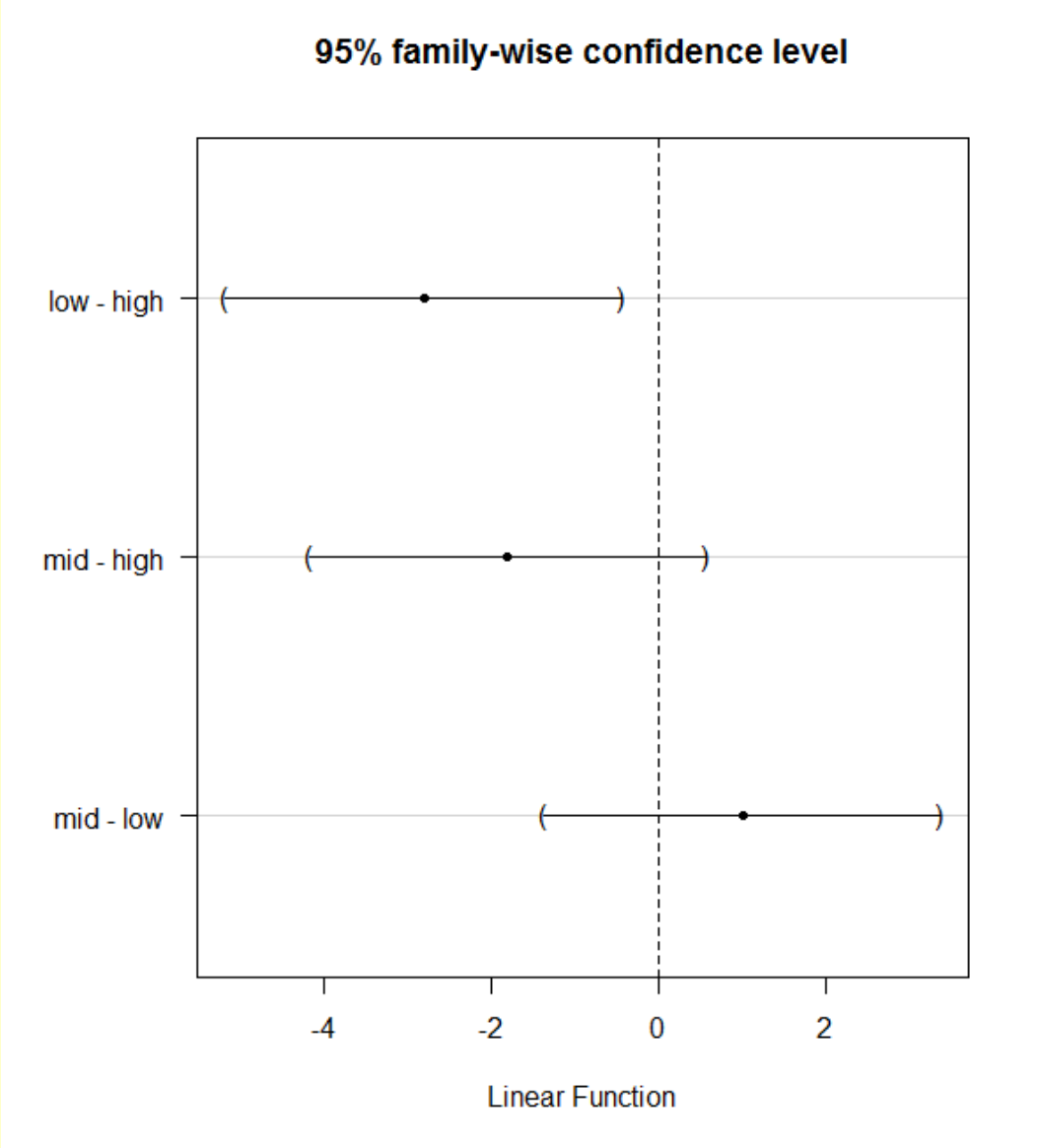
Doing an ANOVA – Output

```
Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means:

Tukey Contrasts Fit:
aov(formula = libido ~ dose, data = Dataset)

Linear Hypotheses:  Estimate Std.Error t value Pr(>|t|)
low - high == 0     -2.8000   0.8869   -3.157   0.021 *
mid - high == 0     -1.8000   0.8869   -2.029   0.148
mid - low  == 0      1.0000   0.8869    1.127   0.516

--- Signif. codes:
 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ‘ 1
(Adjusted p values reported -- single-step method)
```

Doing an ANOVA – Output

```
Simultaneous Confidence Intervals

Multiple Comparisons of Means:

Tukey Contrasts Fit:
aov(formula = libido ~ dose, data = Dataset)

Quantile = 2.667 95% family-wise confidence level

Linear Hypotheses:   Estimate    lwr      upr

low - high == 0      -2.8000   -5.1655 -0.4345
mid - high == 0      -1.8000   -4.1655  0.5655
mid - low  == 0       1.0000   -1.3655  3.3655
```

Doing an ANOVA – Output



Result

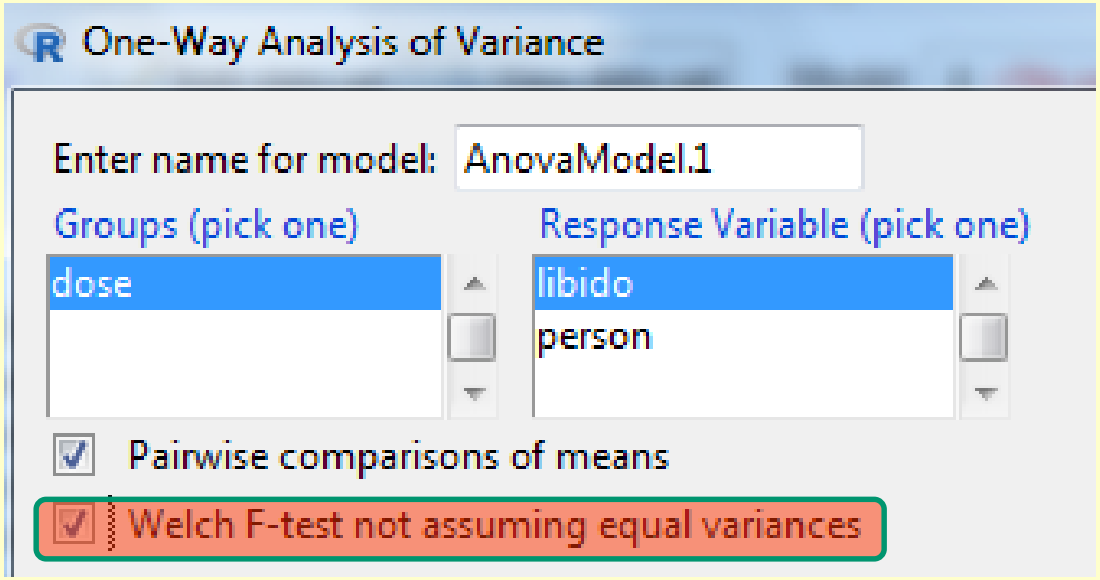
One pair of samples is different

Two pairs of samples are not different

```
high low mid
"b" "a" "ab"
```

Low Mid High

Doing an ANOVA – With RCmdr



Assume Equal Variances? NO

What Happens After Significant ANOVA?

Compare means

PERFORMS AN ANOVA, FOLLOWED BY A Welch F –test, WITHOUT ASSUMPTION OF EQUAL VARIANCES

Doing an ANOVA – Output

SAME OUTPUT, BUT ADDS ANOTHER TEST, WITHOUT THE ASSUMPTION OF EQUAL VARIANCES

```
oneway.test(libido ~ dose, data=viagra)

# Welch test
One-way analysis of means
(not assuming equal variances)

data: libido and dose

F = 4.3205,
num df = 2.0000,
denom df = 7.9434,

p-value = 0.05374
```

NOTE:
DF error decreased from 12 to 7.9434

What is the Result?

Not Significant

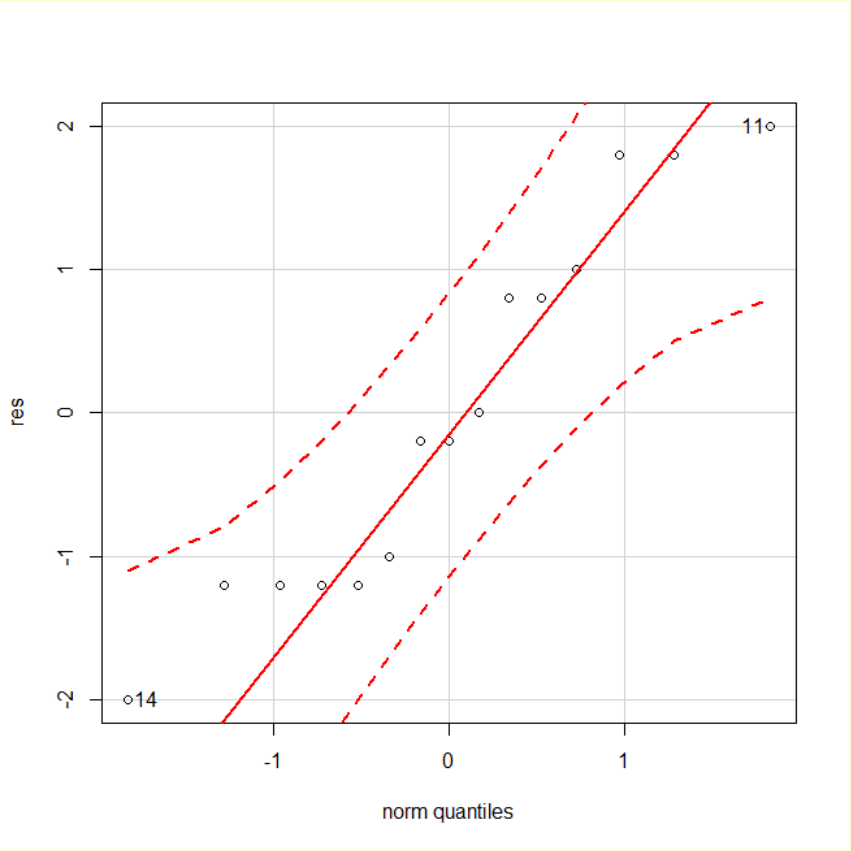
(More Conservative)

After Doing an ANOVA

- Access and plot ANOVA residuals

```
> res <- residuals(AnovaModel.1)
```

```
> qqPlot(res, dist="norm",
id.method="y", id.n=2,
labels=rownames(libido))
```



After Doing an ANOVA

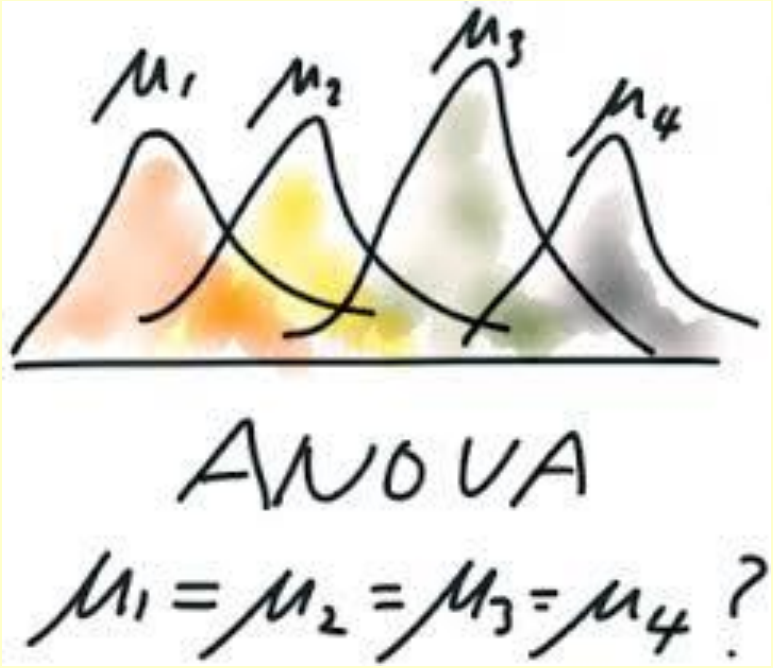
- Analyze ANOVA residuals for normality

```
> normalityTest (~res, test="shapiro.test")
```

```
Shapiro-Wilk normality test data:

res W = 0.91669, p-value = 0.1715
```

Post-Hoc Tests

- After significant ANOVA result, we know that at least two means are different (NOTE: ANOVA is an omnibus test).



- Post-Hoc (After-the-Fact) tests are able to identify which pairs of samples are responsible for the significant ANOVA result

Post-Hoc Tests

- *Which post hoc* test to conduct depends on the assumptions and on how conservative you want to be concerning the chance of a type-I error.
- Tukey's test accurately maintains alpha levels at intended values as long as model assumptions are met (i.e., normality, homogeneous variances).
- NOTE: Tukey is default for ANOVA in Rcmdr.
- In R, Bonferroni and related methods implemented using the *pairwise.t.test()* function.

P-Adjustment Methods

- The adjustment methods include the Bonferroni correction ("bonferroni") in which the p-values are multiplied by the number of comparisons.

- Less conservative corrections included by:
  - Holm (1979) ("holm")
  - Benjamini & Hochberg (1995) ("BH" or "fdr")

Post-Hoc Test Options

- Bonferroni correction is very conservative:

$$Bonferroni\ \alpha = 0.0083 = \frac{0.05}{6}$$

  - More difficult to find a significant result (lower power)
  - Easier to make type-II error

Same comparison for all tests

Bonferroni

	p	$p_{crit} = \frac{\alpha}{k}$	
NT–Super	.0000	.0083	*
Super–Hulk	.0014	.0083	*
Spider–Super	.0127	.0083	
NT–Spider	.0252	.0083	
NT–Hulk	.1704	.0083	
Spider–Hulk	.3431	.0083	

Post-Hoc Test Options

- Holm method is less conservative than the Bonferroni correction:
  - Ranks p values from largest to smallest (using index j)
  - Calculates P crit = alpha / j (where j is rank of each p value)

Different comparison for each test

Holm

p	j	$p_{crit} = \frac{\alpha}{j}$	
.0000	6	.0083	*
.0014	5	.0100	*
.0127	4	.0125	
.0252	3	.0167	
.1704	2	.0250	
.3431	1	.0500	

Post-Hoc Test Options

- Benjamin-Hochberg method estimates type-I error rate using False Discovery Rate (FDR).

FDR = $\frac{\textit{Falsely Rejected Nulls}}{\textit{Total Rejected Nulls}}$

- Ranks p values from smallest to largest (using index j)

- Calculates P crit = alpha * (j / k) (where k = number tests)

Different comparison for each test

Benjamini–Hochberg

p	j	$p_{crit} = \left(\frac{j}{k}\right)\alpha$	
.0000	1	.0083	*
.0014	2	.0167	*
.0127	3	.0250	*
.0252	4	.0333	*
.1704	5	.0417	
.3431	6	.0500	

Post-Hoc Tests

Tukey: Do not consider multiple testing
(each sample pair compared using alpha = 0.05)

Other options consider multiple testing:

Bonferroni: Most Conservative
(Less Likely to Yield a Significant Result)

Holm: Middle-of-the-road Conservative

Benjamin-Hochberg: Least Conservative

Post-Hoc Bonferroni Test

```
> pairwise.t.test(viagra$libido, viagra$dose,
p.adjust.method = "bonferroni")
```

```
Pairwise comparisons using t tests with pooled SD data:

viagra$libido and viagra$dose

     high low
low  0.025   -
mid  0.196 0.845
```

Outcome:

HIGH significantly different from LOW

Post-Hoc Holm Test

```
> pairwise.t.test(viagra$libido, viagra$dose,
p.adjust.method = “holm")
```

```
Pairwise comparisons using t tests with pooled SD data:

viagra$libido and viagra$dose

     high low

low 0.025    -
mid 0.130 0.282
```

Outcome:

HIGH significantly different from LOW

Post-Hoc BH Test

```
> pairwise.t.test(viagra$libido, viagra$dose,
p.adjust.method = "BH")
```

```
Pairwise comparisons using t tests with pooled SD data:

viagra$libido and viagra$dose

     high low
low  0.025    -
mid  0.098 0.282
```

Outcome:

HIGH significantly different from LOW

One-Way ANOVA Summary

- One-way ANOVA allows us to analyze experiments involving only one independent variable (factor) manipulated in multiple ways and only one measured outcome variable. It is an expansion of the T-test.

- ANOVAs that yield significant results need to be followed by post-hoc tests. There are many options, so try several and compare results.

- Use Post-Hoc tests: Perform all three and compare
  - Bonferroni: Most conservative
  - Step-Down: Holm is intermediately conservative
  - Step-Up: Benjamini & Hochberg is least conservative

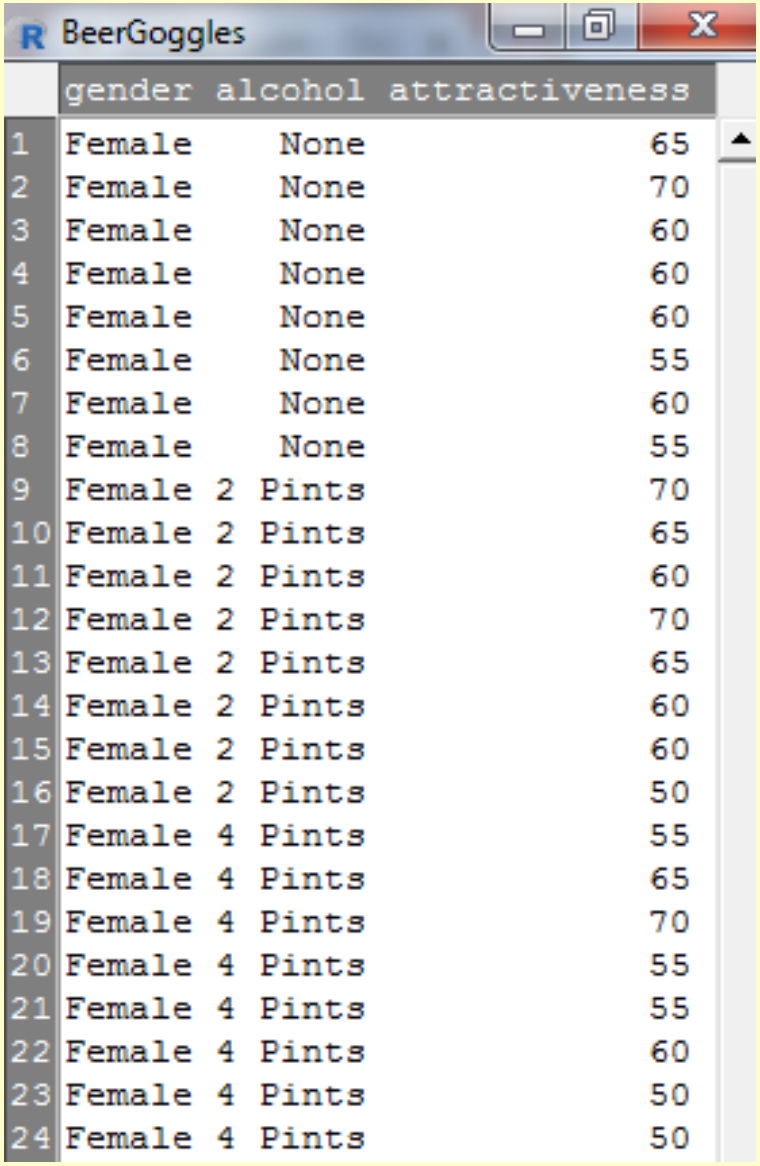
Multi-Way ANOVA in RCmdr

Two-Way ANOVA

(factor1:
Gender – 2 categories)

(factor2:
Alcohol – 3 categories)

Dependent: Attractiveness



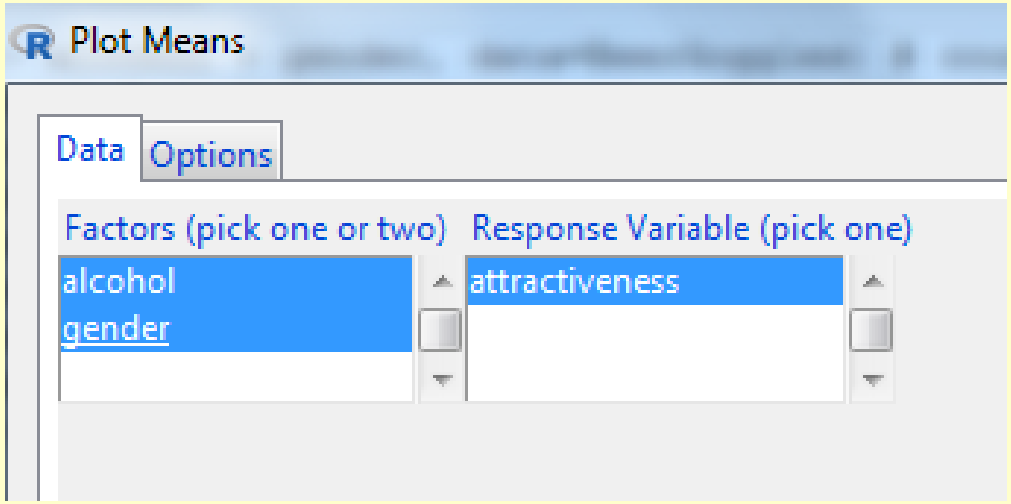
BeerGoggles

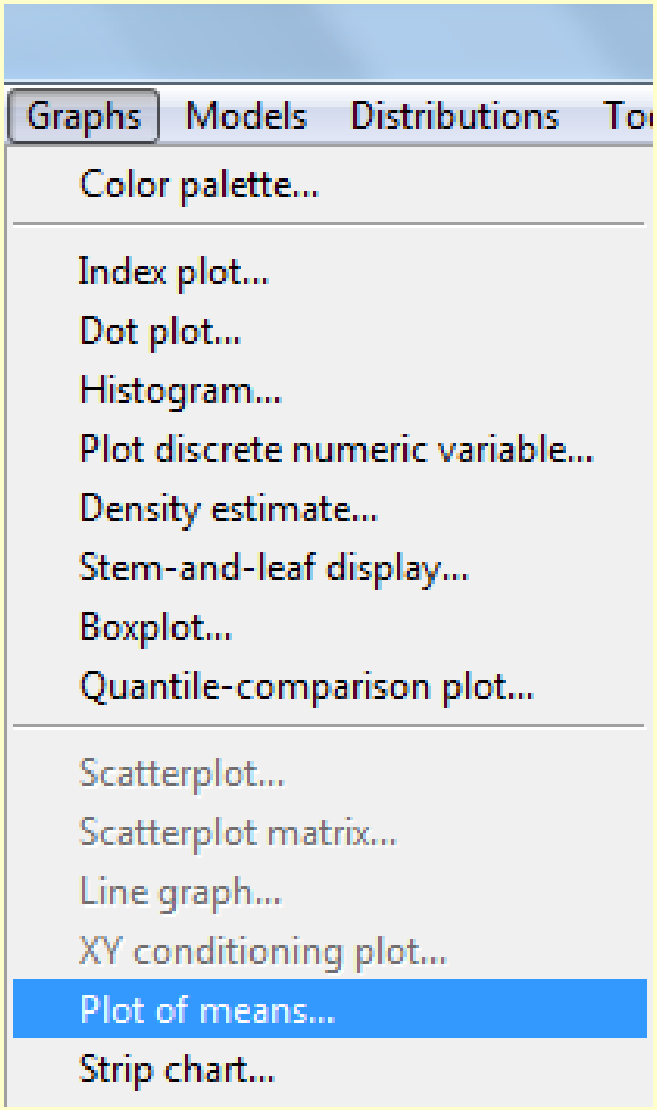
	gender	alcohol	attractiveness
1	Female	None	65
2	Female	None	70
3	Female	None	60
4	Female	None	60
5	Female	None	60
6	Female	None	55
7	Female	None	60
8	Female	None	55
9	Female	2 Pints	70
10	Female	2 Pints	65
11	Female	2 Pints	60
12	Female	2 Pints	70
13	Female	2 Pints	65
14	Female	2 Pints	60
15	Female	2 Pints	60
16	Female	2 Pints	50
17	Female	4 Pints	55
18	Female	4 Pints	65
19	Female	4 Pints	70
20	Female	4 Pints	55
21	Female	4 Pints	55
22	Female	4 Pints	60
23	Female	4 Pints	50
24	Female	4 Pints	50

Multi-Way ANOVA in RCmdr

- Visualizing the data:

  Graphs / Plot of Means





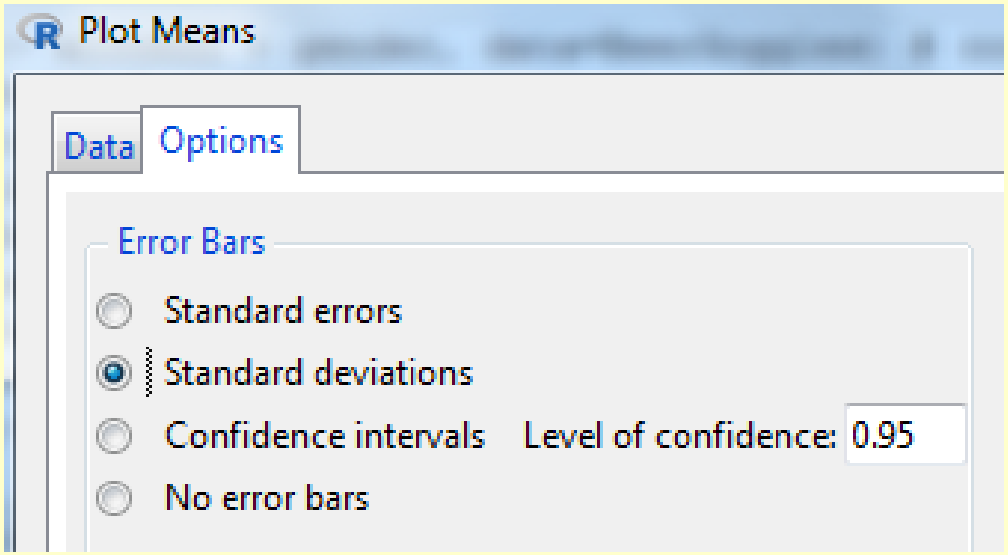
Two-Way ANOVA:

(factor1: Gender)

(factor2: Alcohol)

Multi-Way ANOVA in RCmdr

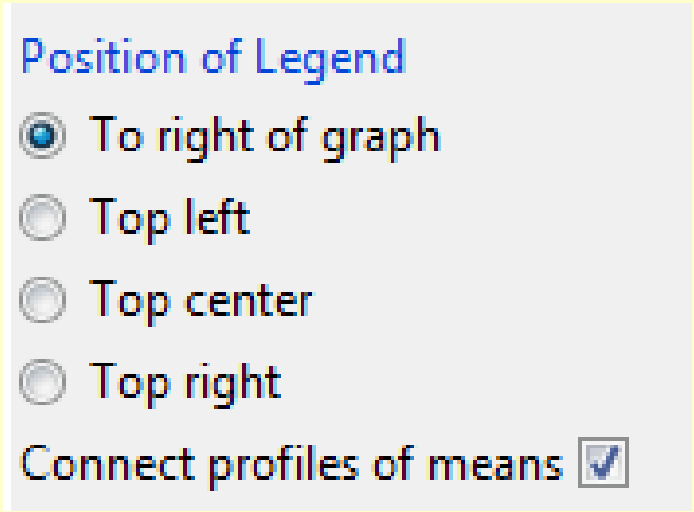
- Visualizing the data:



SE

SD

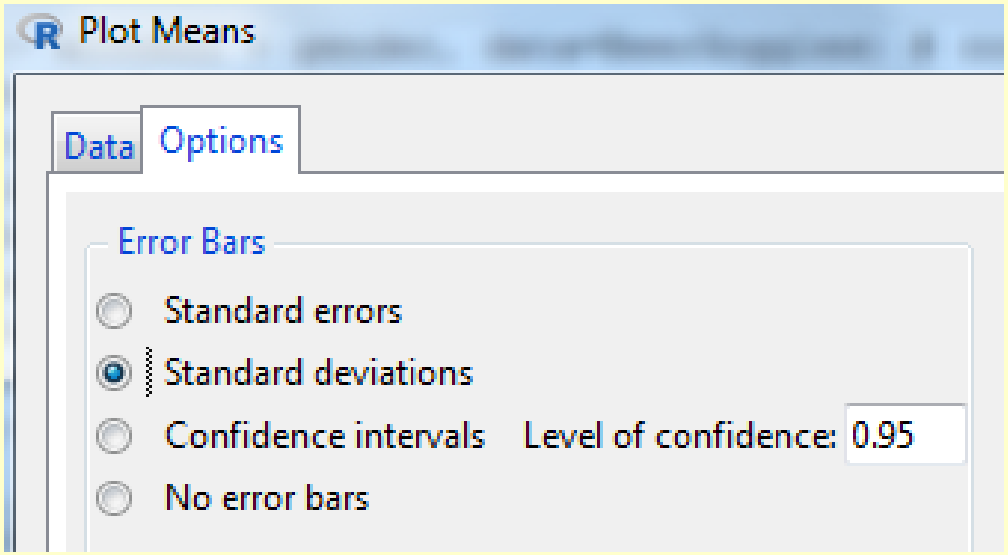
CI



Placement the figure legend

Multi-Way ANOVA in RCmdr

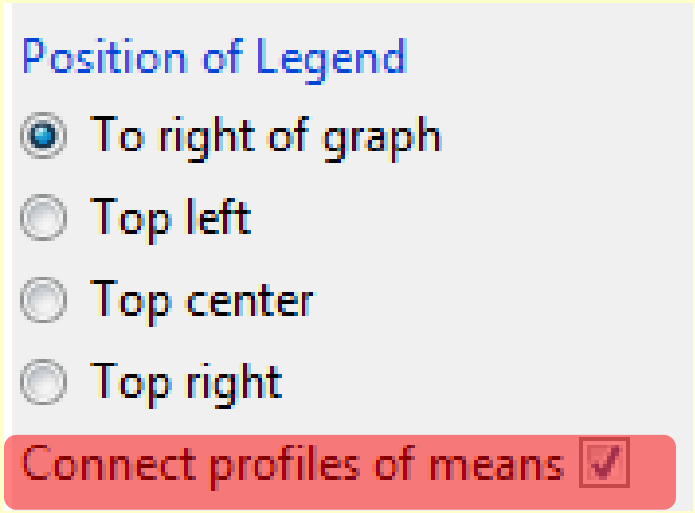
- Visualizing the data:



SE

SD

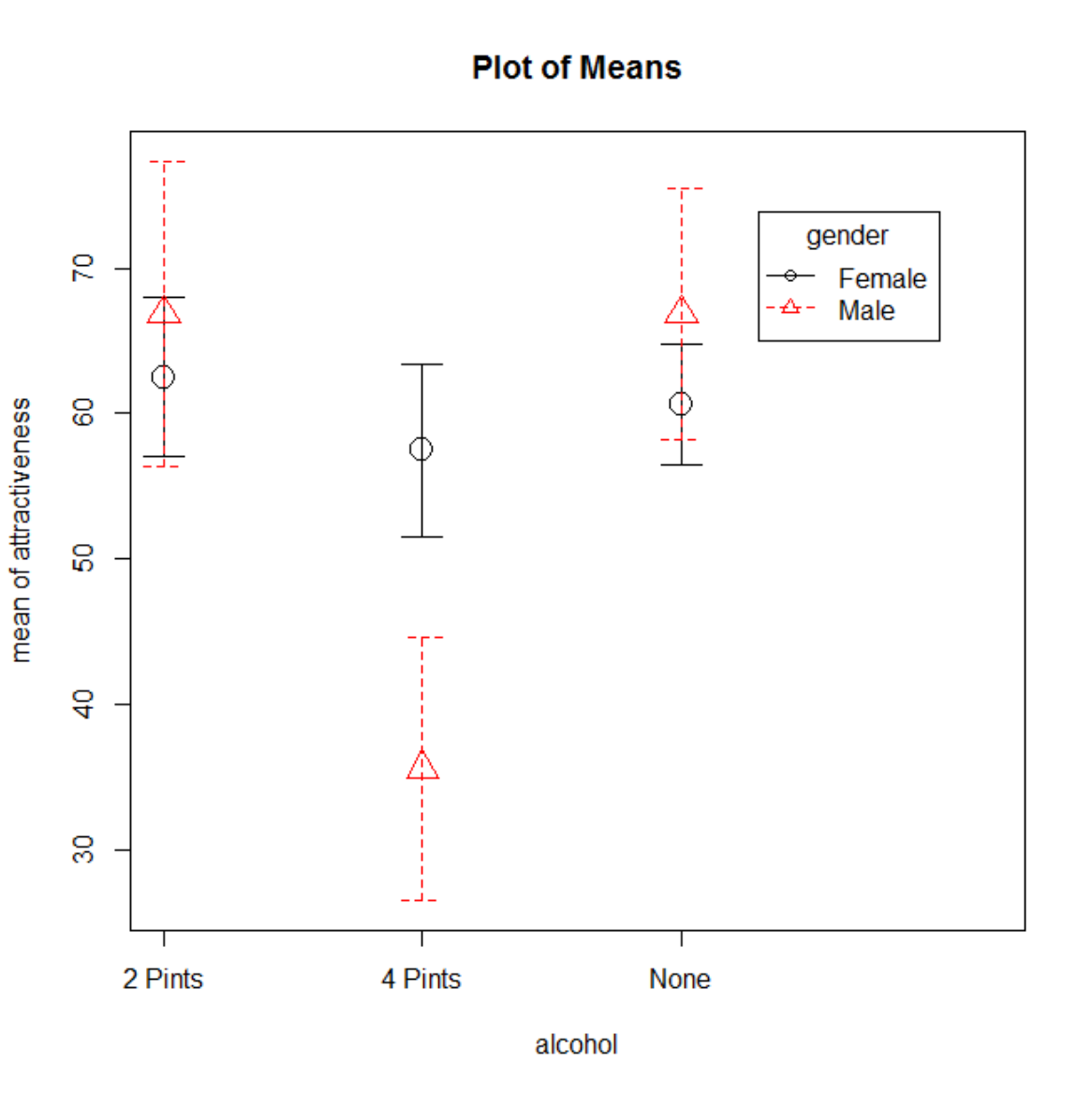
CI



Placement the figure legend

Connect means with lines ?

Multi-Way ANOVA in RCmdr

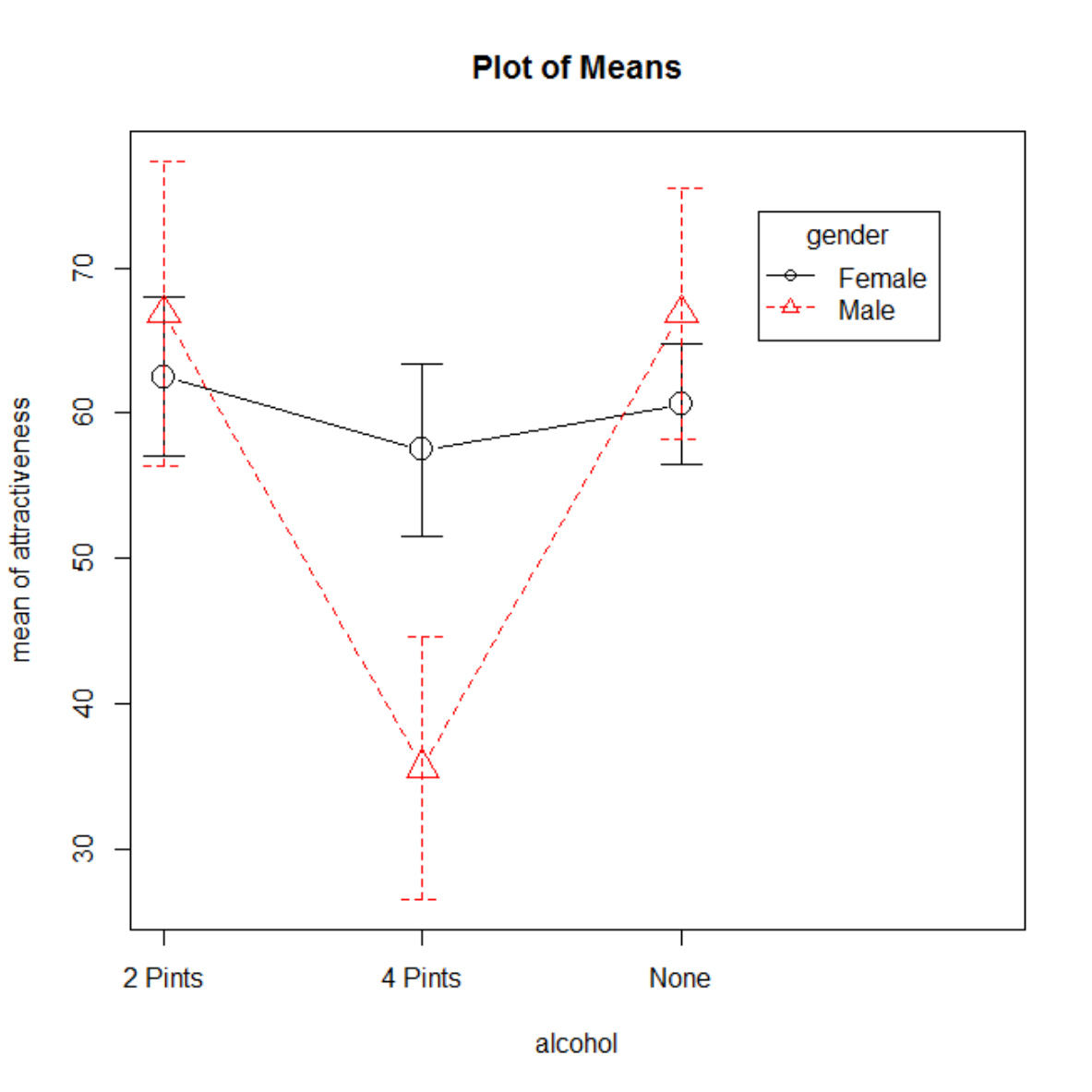


What error bars should you show?

Typically, we use SDs to assess variability about the mean

95% CI easier to interpret

Multi-Way ANOVA in RCmdr

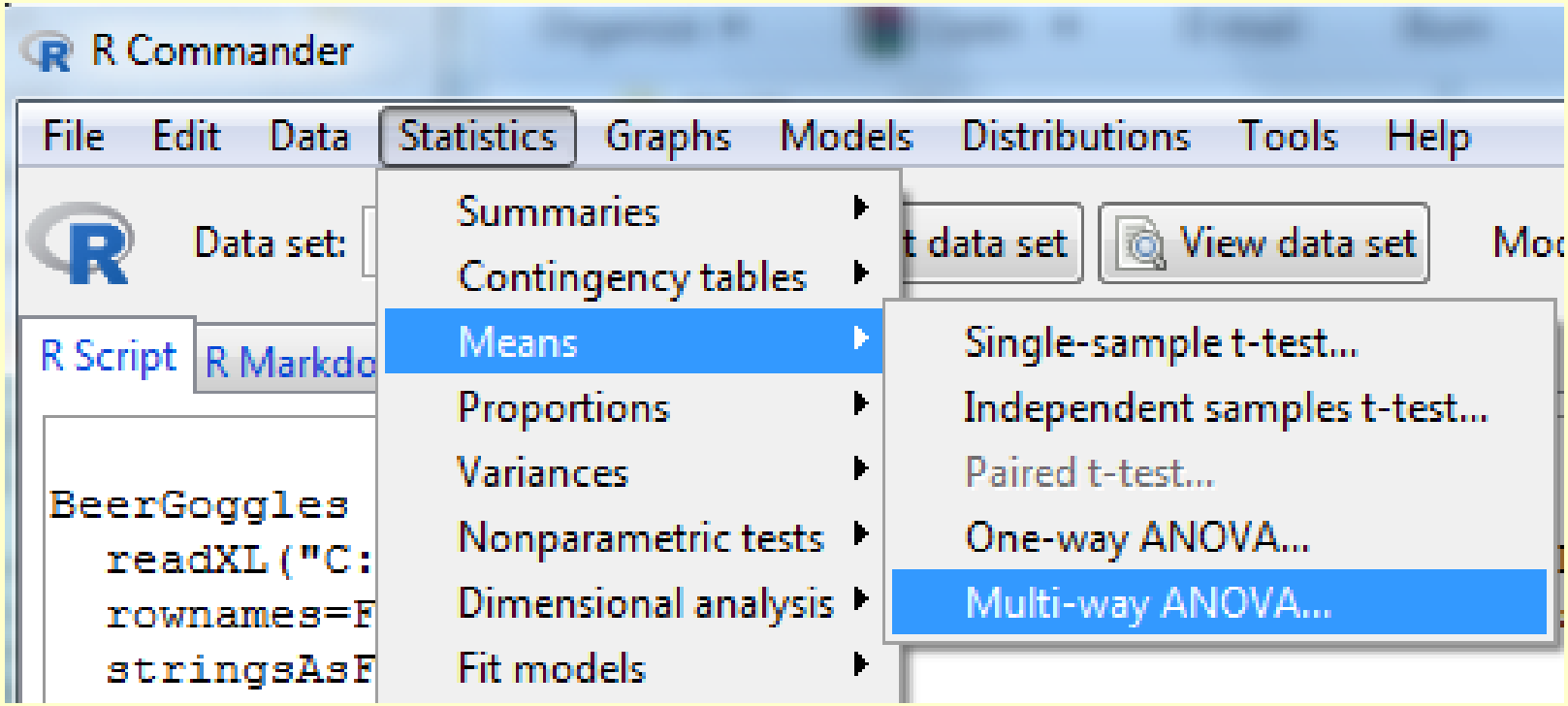


What do the lines show?

The lines connecting group means are critical to assess interactions

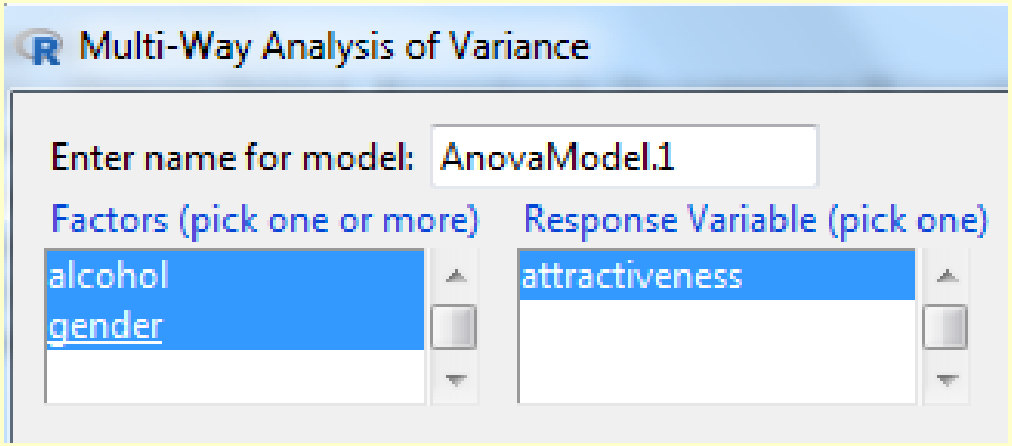
Multi-Way ANOVA in RCmdr

- Multi-way ANOVA
  (one dependent variable,
  two or more independent variables)



Statistics / Means / Multi-way ANOVA

Multi-Way ANOVA in Rcmdr



Select the following:

- Factors

- Response

```
Rcmdr> AnovaModel.1 <- lm(attractiveness ~ alcohol*gender,
data=BeerGoggles,
contrasts=list(alcohol ="contr.Sum", gender ="contr.Sum"))
```

NOTE:

lm = linear model

* = denoted the interaction of two factors

contrasts = compare factor levels and summarize

Multi-Way ANOVA in Rcmdr

Output 1: ANOVA table (SS terms, MS terms, p values)

```
Rcmdr> Anova(AnovaModel.1)
Anova Table (Type II tests)

Response: attractiveness

                 Sum Sq   Df   F value     Pr(>F)
alcohol          3332.3   2    20.0654   0.0000007649 ***
gender            168.8   1     2.0323   0.1614
alcohol:gender   1978.1   2    11.9113   0.0000798660 ***

Residuals         3487.5  42

--- Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Multi-Way ANOVA in Rcmdr

Output 2: Table of means, for all levels in both factors

```
Rcmdr> with(BeerGoggles,

(tapply(attractiveness, list(alcohol, gender), mean,
na.rm=TRUE)))

# means

          Female Male
2 Pints 62.500 66.875
4 Pints 57.500 35.625
None    60.625 66.875
```

tapply function =

Applies one function to each cell of a ragged array, that is to each (non-empty) group of values given by a unique combination of the levels of certain factors.

Multi-Way ANOVA in Rcmdr

Output 3: Table of SDs, for all levels in both factors

```
Rcmdr> with(BeerGoggles,

(tapply(attractiveness, list(alcohol, gender), sd,
na.rm=TRUE)))

# std.deviations

                Female   Male
2 Pints       6.546537 12.51784
4 Pints       7.071068 10.83562
None          4.955156 10.32940
```

Multi-Way ANOVA in Rcmdr

Output 4: Table of sample sizes (n), for all levels in both factors

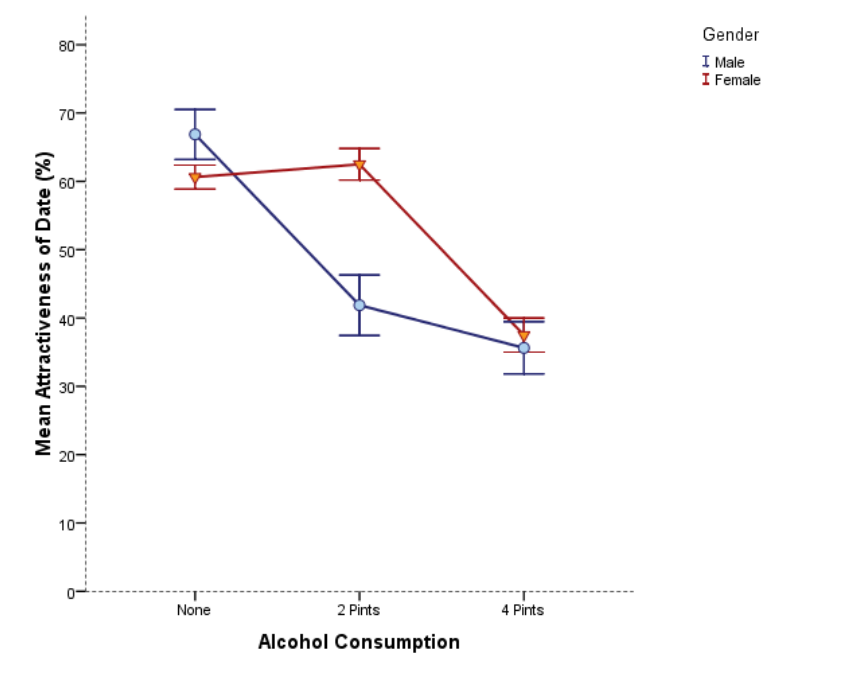
```
Rcmdr> xtabs
(~ alcohol + gender, data=BeerGoggles)

# counts  gender

alcohol
          Female  Male
2 Pints        8     8

4 Pints        8     8

None           8     8
```

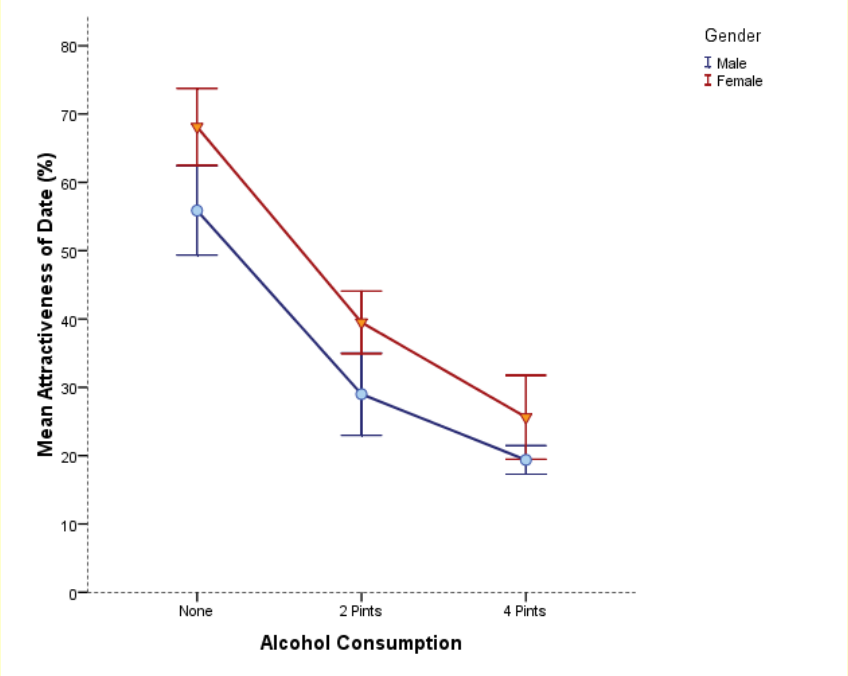
Multi-Way ANOVA – Summary

- Multi-Way ANOVA can deal with multiple independent variables and their interactions.
- However, this design requires a full dataset: every combination of factor treatments must be sampled.
- Another limitation of multi-way ANOVA is that analyses with 3 independent variables (and three-way interactions) are difficult to interpret.
- NOTE: Repeated Measures ANOVA can be used in a One-Way (only that factor) or a Multi-Way design (repeated measure level and one other factors).

Multi-Way ANOVA - Summary

Example: Two different possible results:
Which one shows a significant interaction effect?



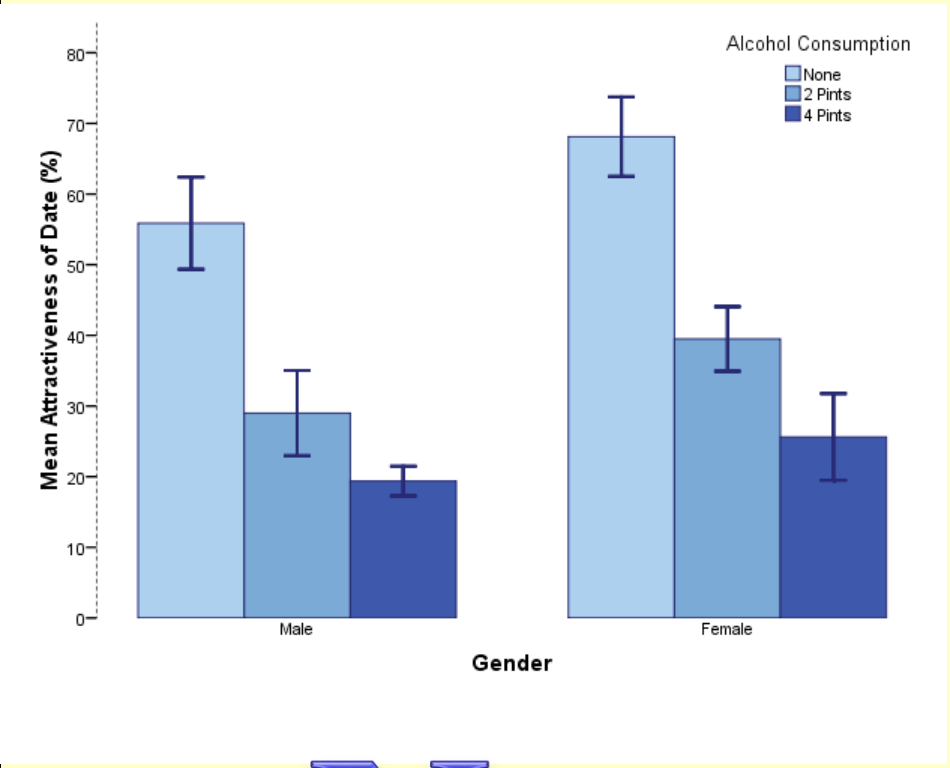


Yes

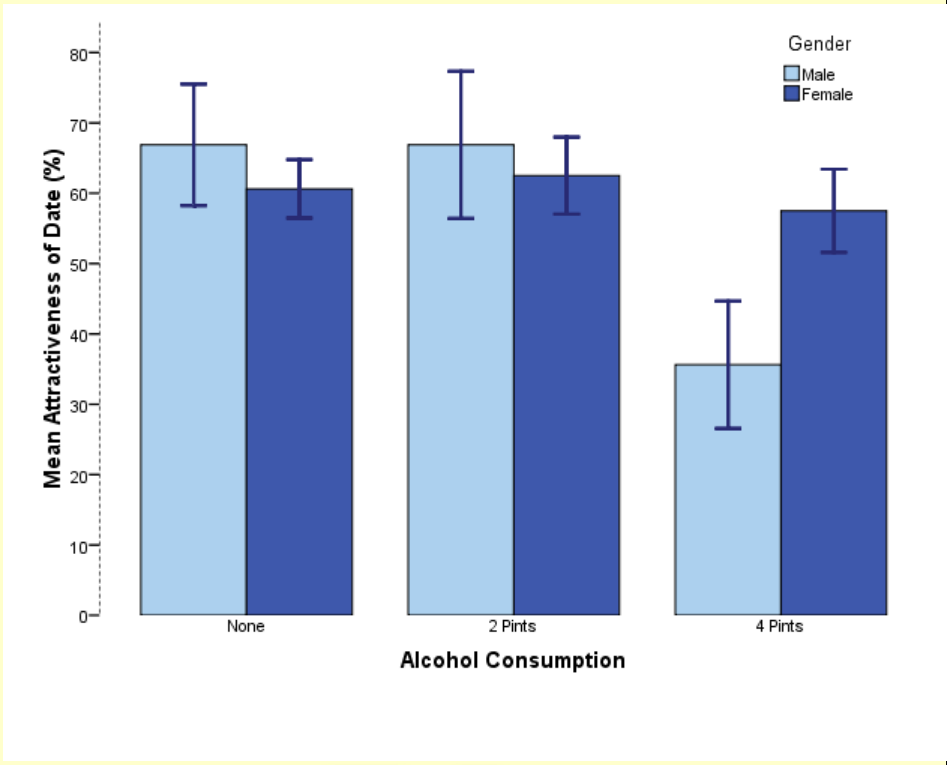
No

Multi-Way ANOVA - Summary

Example: Two different possible results:
Which one shows a significant interaction effect?



No



Yes

References

Benjamini, Y. & Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* **57**, 289–300.

Benjamini, Y. & Yekutieli, D. 2001. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* **29**, 1165–1188.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**, 65–70.

Hommel, G. 1988. A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika* **75**, 383–386.

Hochberg, Y. 1988. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75**, 800–803.

Wright, S. P. 1992. Adjusted P-values for simultaneous inference. *Biometrics* **48**, 1005–1013.