Analysis of Quantitative data
Student’s *t*-test

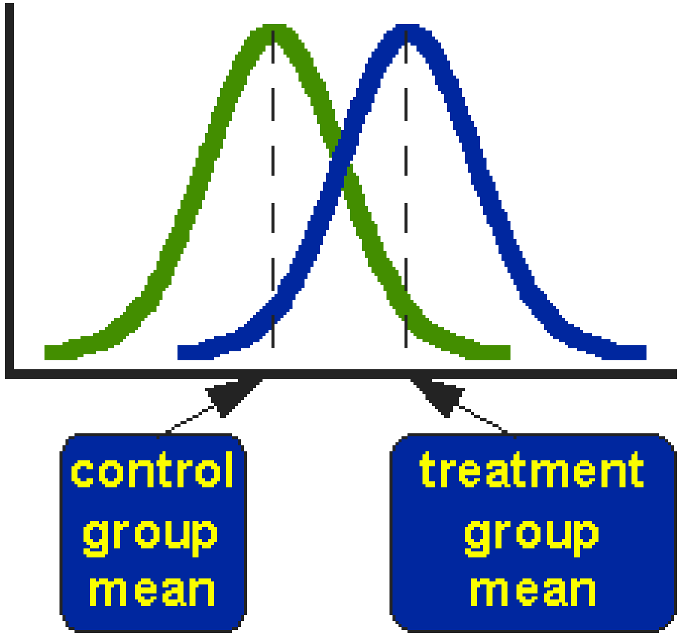
Anne Segonds-Pichon
v2020-12

Comparison between 2 groups

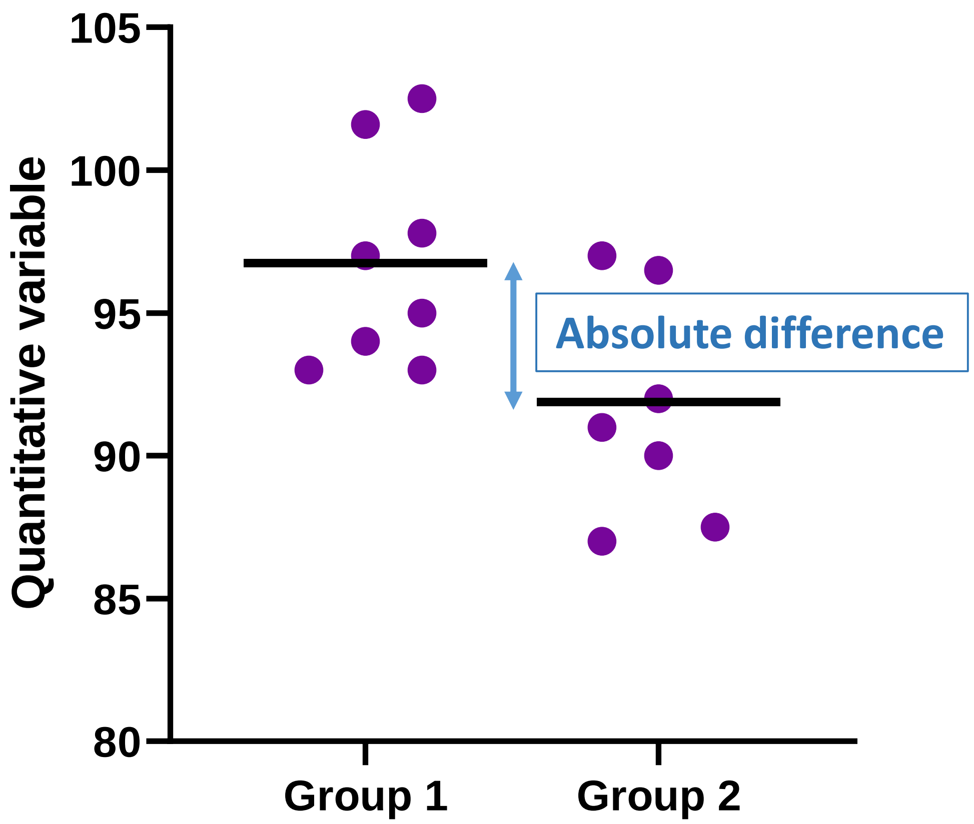
Comparison between 2 groups:

Student's *t*-test

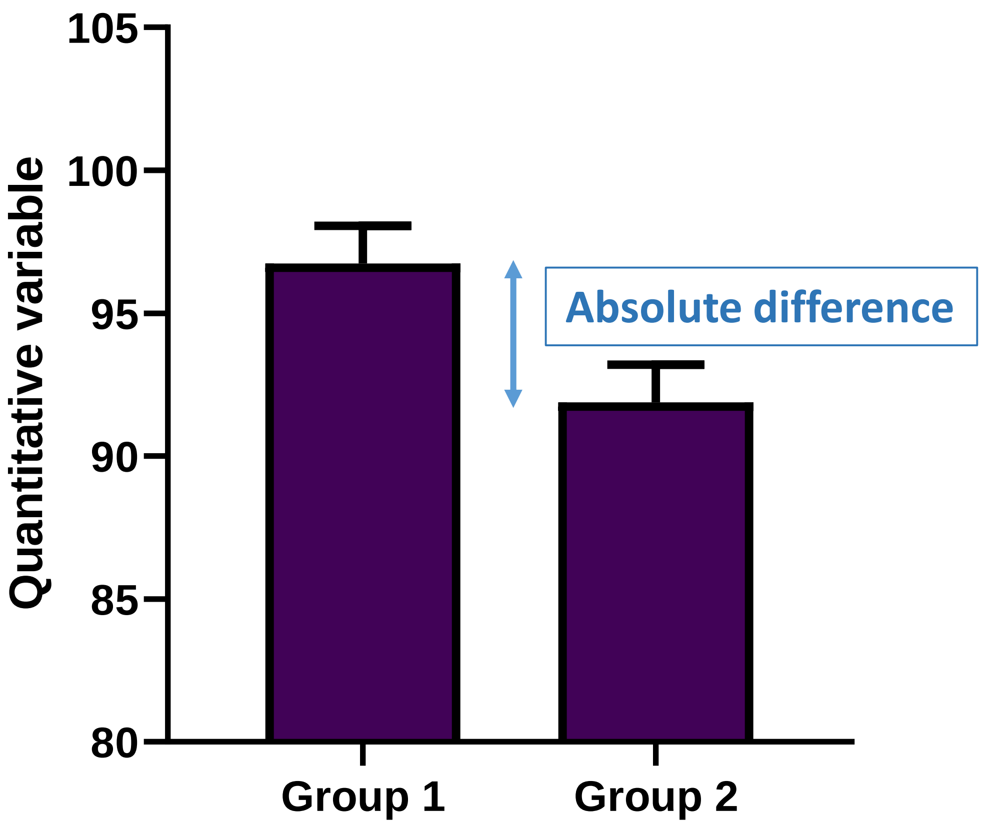
- **Basic idea**:
  - When we are looking at the differences between scores for 2 groups, we have to judge the difference between their means relative to the spread or variability of their scores.
    - Eg: comparison of 2 groups: control and treatment



Variability does matter

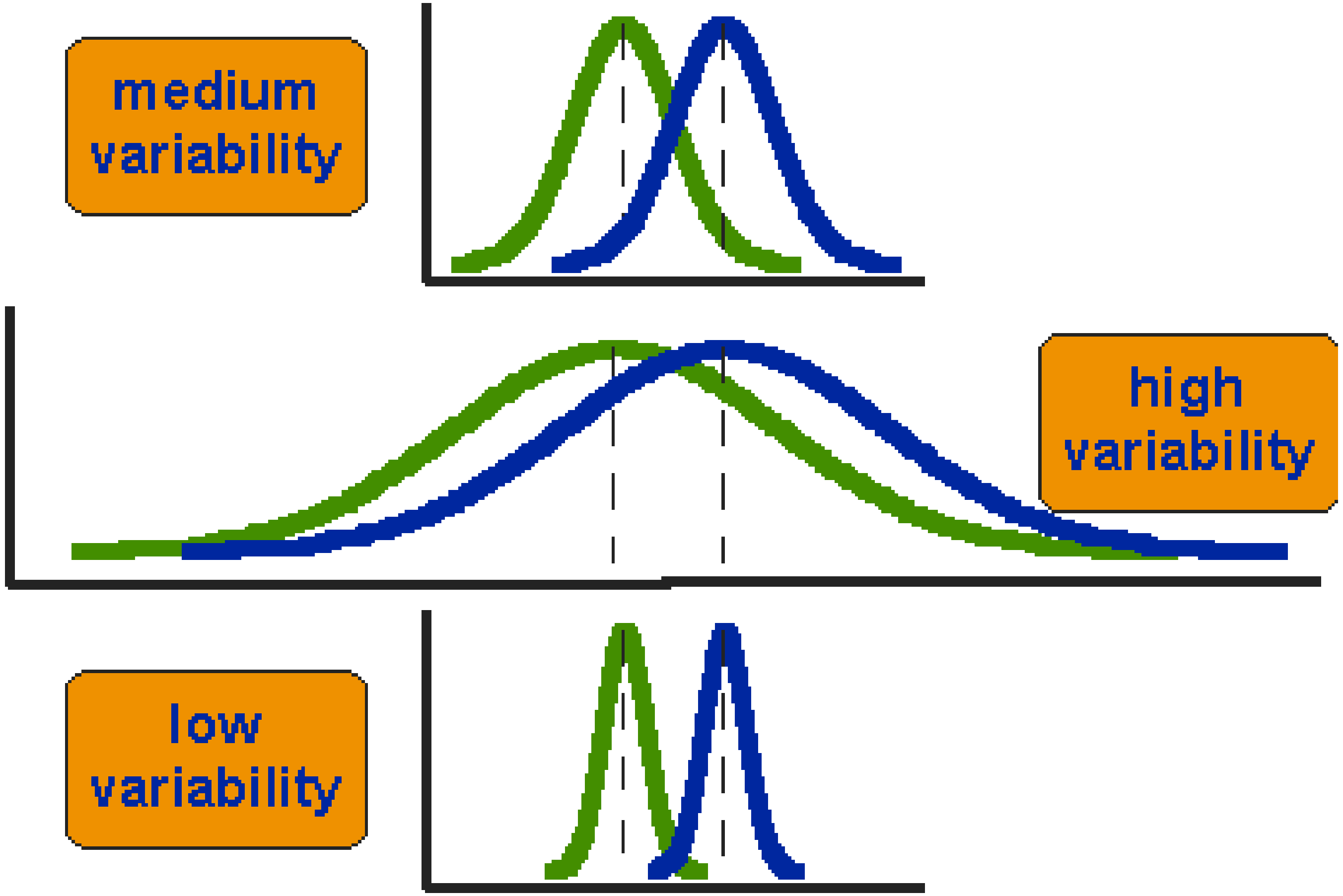


Scatter plot 🙂

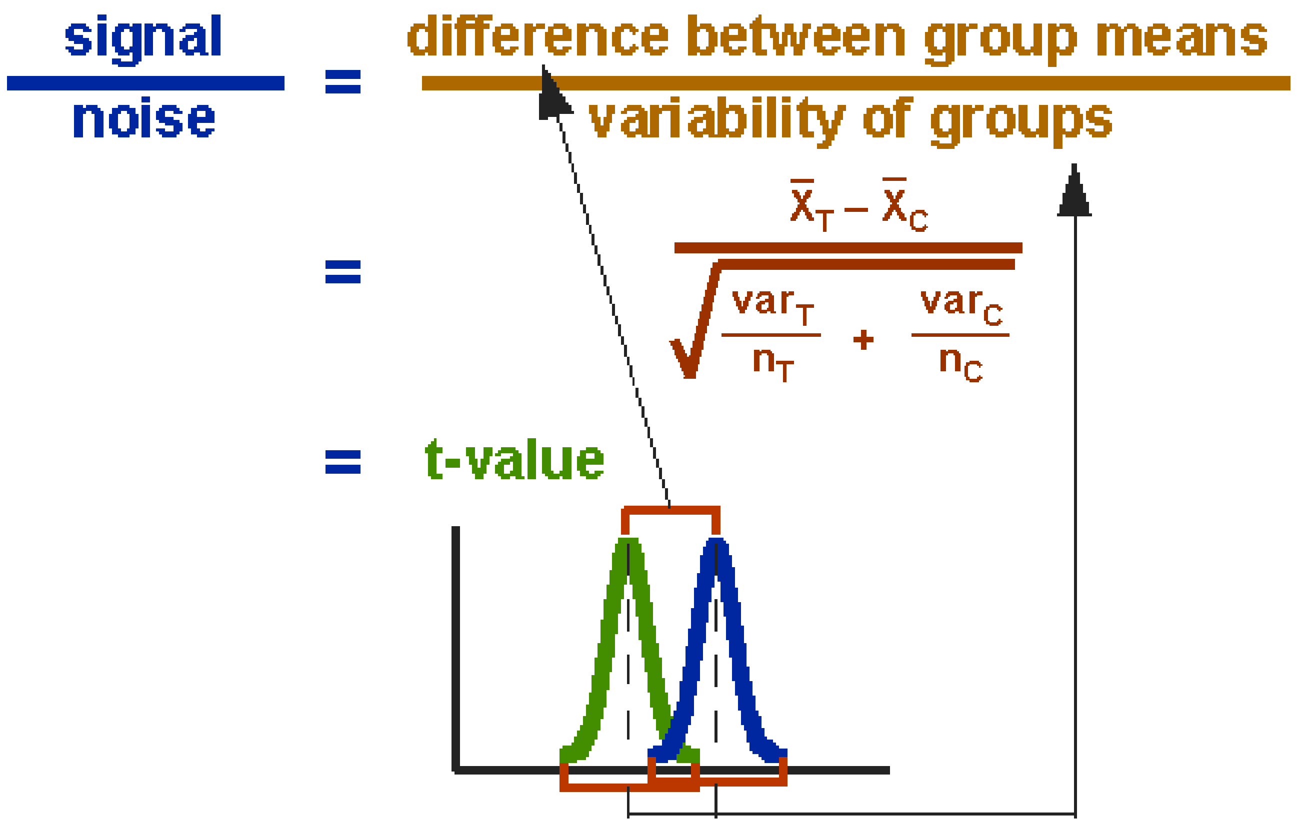


Bar chart ☹

Student's *t*-test



Student’s *t*-test

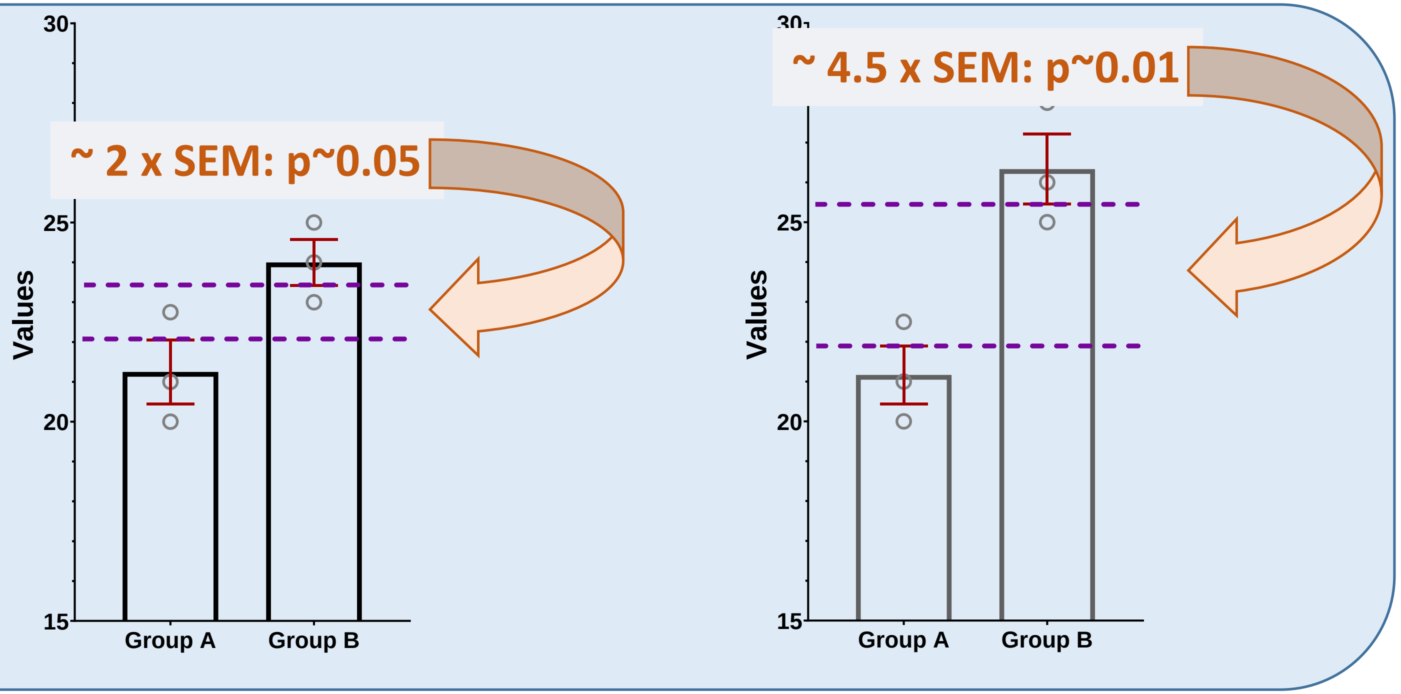


$$\frac{\text{signal}}{\text{noise}} = \frac{\text{difference between group means}}{\text{variability of groups}}$$

$$= \frac{\bar{X}_T - \bar{X}_C}{\sqrt{\frac{var_T}{n_T} + \frac{var_C}{n_C}}}$$

$$= \text{t-value}$$

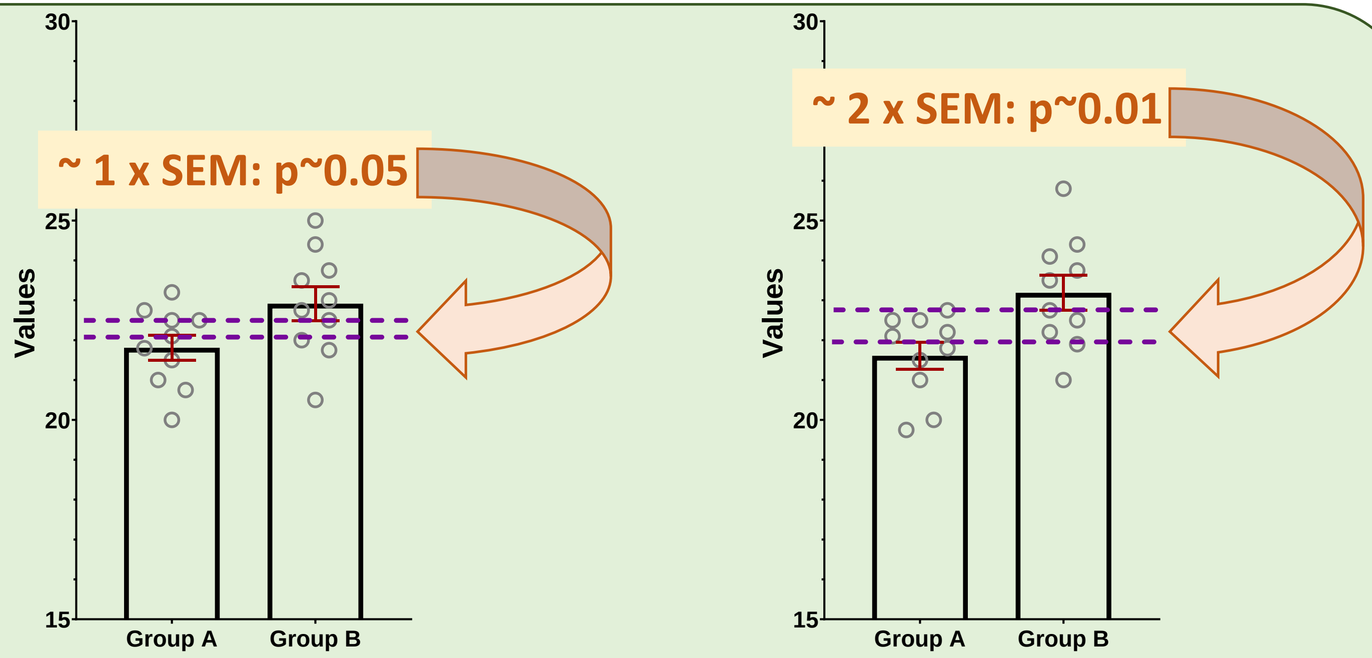
n = 3



~ 2 x SEM: p~0.05

~ 4.5 x SEM: p~0.01

n = 10



~ 1 x SEM: p~0.05

~ 2 x SEM: p~0.01

Student's *t*-test

- **Independent t-test**
  - Difference between 2 means of one variable for <u>two independent groups</u>
    - Example: difference in weight between WT and KO mice

- **Paired t-test**
  - Difference between two measures of one variable for <u>one group</u>:
    - Example: before-after measurements
      - the second 'sample' of values comes from the same subjects (mouse, petri dish …).
    - Importance of experimental design!

- One-Sample t-test
  - Difference between the mean of a single variable and a specified constant.

Example: coyotes

- Question: do male and female coyotes differ in size?
- **Sample size**
- **Data exploration**
- **Check the assumptions for parametric test**
- **Statistical analysis: Independent t-test**

Exercise 3: Power analysis

- Example case:

No data from a pilot study but we have found some information in the literature.

In a study run in similar conditions as in the one we intend to run, **male coyotes** were found to measure: **92cm +/- 7cm (SD)**.

We expect a **5% difference** between genders.

- **smallest biologically meaningful difference**

```
power.t.test(n = NULL, delta = NULL, sd = 1, sig.level = NULL,
power = NULL, type = c("two.sample", "one.sample", "paired"),
alternative = c("two.sided", "one.sided"))
```

Exercise 3: Power analysis - *Answers*

Example case:

We don’t have data from a pilot study but we have found some information in the literature.

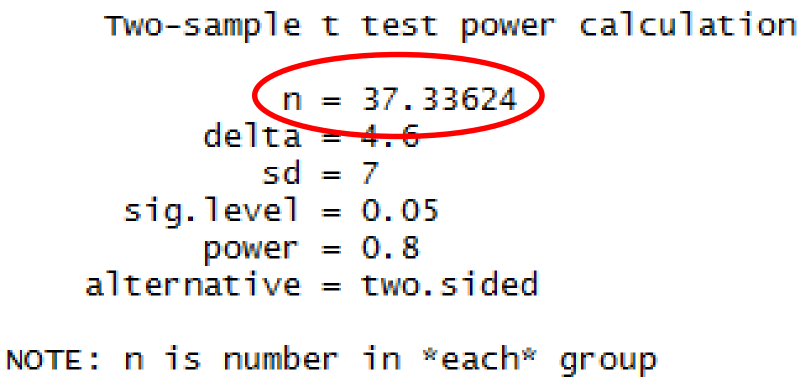
In a study run in similar conditions as in the one we intend to run, **male coyotes** were found to measure: **92cm+/- 7cm (SD**)

We expect a **5% difference** between genders with a similar variability in the female sample.

Mean 1 = 92
Mean 2 = 87.4 (5% less than 92cm)

delta = 92 – 87.4
sd = 7

```
power.t.test(delta=92-87.4, sd=7, sig.level=0.05, power=0.8)
```



```
     Two-sample t test power calculation 

              n = 37.33624
          delta = 4.6
             sd = 7
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

We need a sample size of **n~76 (2*38**)

Data exploration $\neq$ plotting data

Exercise 4: Data exploration

coyote.csv

- The file contains individual body length of male and female coyotes.

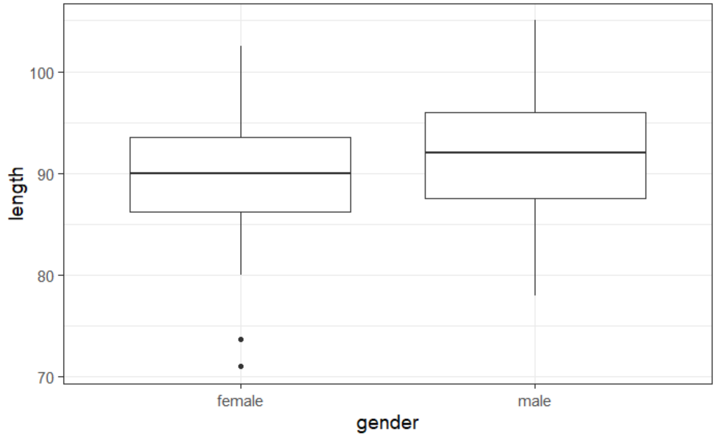
Question: do male and female coyotes differ in size?

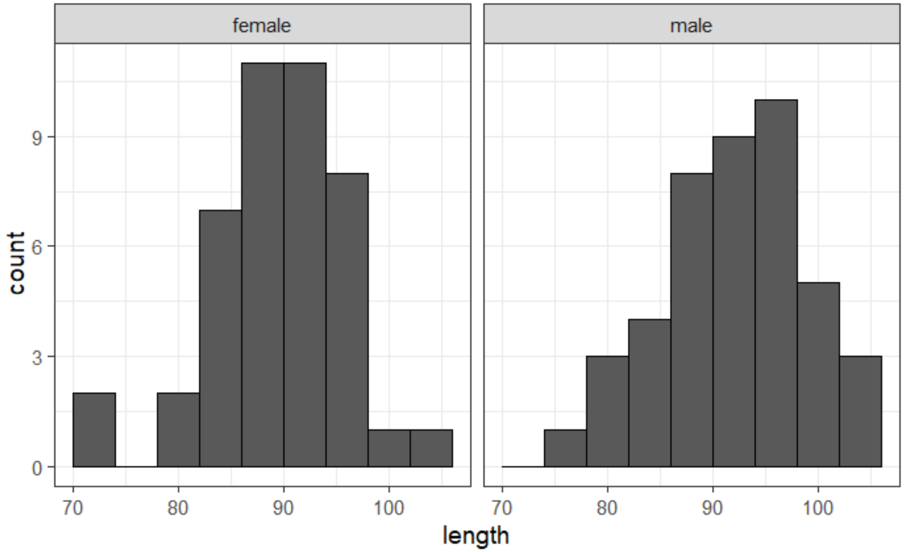
  - Load **coyote.csv**
  - Plot the data as boxplot, histogram, violinplot and stripchart

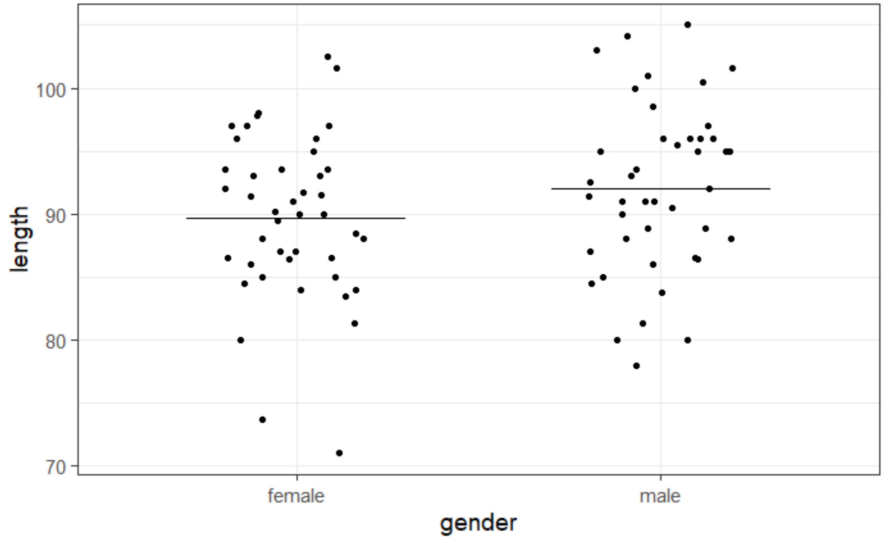
Data exploration ≠ plotting data

Exercise 4: Data exploration

- Explore data using 4 different representations:





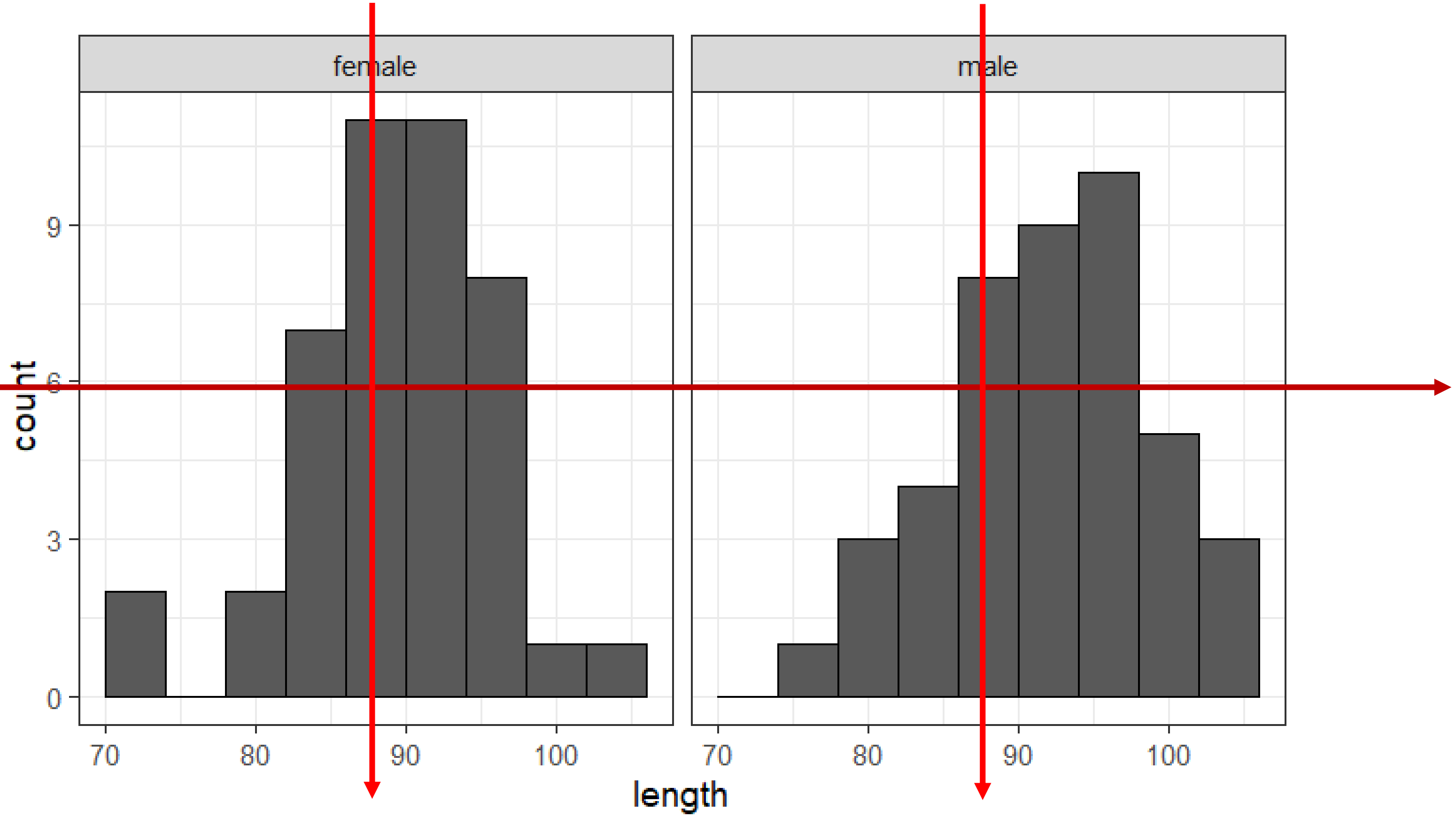


Exercise 4: `facet_grid(rows=vars(row),cols=vars(column))`

`facet_grid(cols=vars(gender))`

2 columns: one per gender

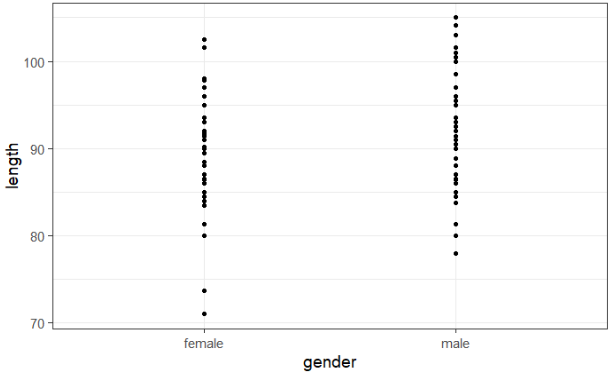
One row



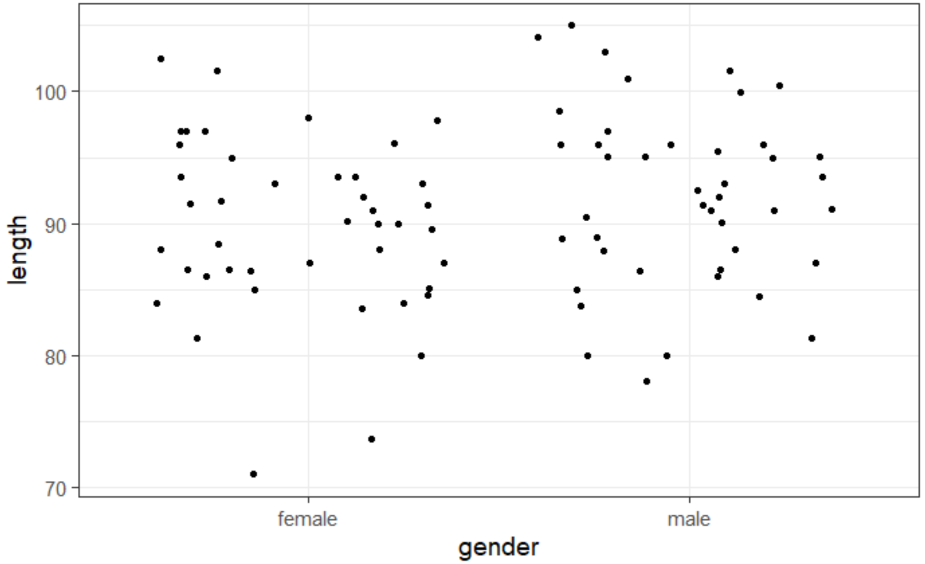
Exercise 4: `geom_jitter()`

- **Stripchart**
  - Variation of `geom_point()`: `geom_jitter()`

```
coyote %>%
  ggplot(aes(x=gender,y=length)) +
  geom_point()
```

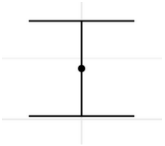


```
coyote %>%
  ggplot(aes(x=gender,y=length)) +
  geom_jitter(height=0, width=0.3)
```



Exercise 4: `stat_summary()`

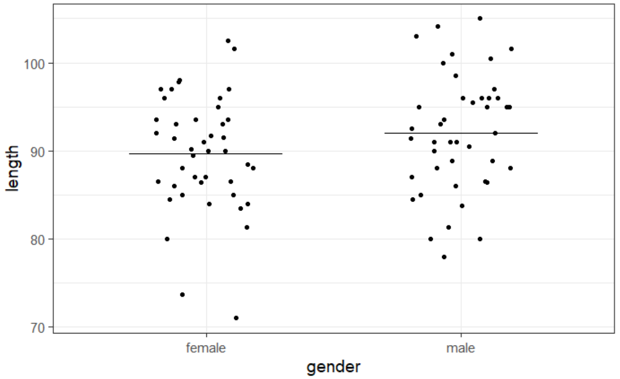
- Stripchart
  - `stat_summary()`
    - What statistical summary: mean: `fun = "mean"`
    - What `geom():` choice of graphical representation: a line: `geom_errorbar()`



```
stat_summary(geom="errorbar", fun="mean",fun.min="mean",fun.max="mean")
```

mean=minimum=max

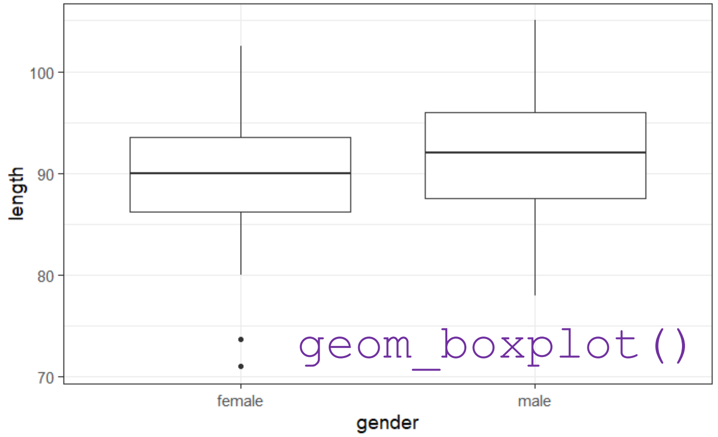
```
coyote %>%
  ggplot(aes(gender,length)) +
    geom_jitter(height=0, width=0.2)+
    stat_summary(geom= "errorbar", fun="mean", fun.min="mean", fun.max="mean")
```



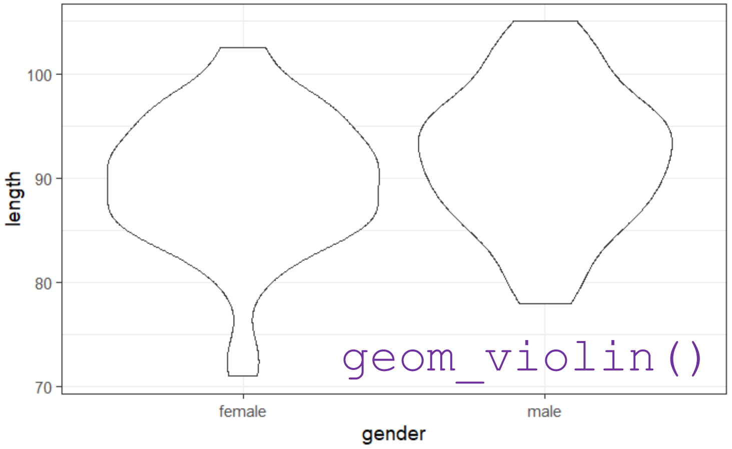
Exercise 4: Data exploration

```
coyote %>%
  ggplot(aes(x=gender, y=length))+
  geom_...()
```

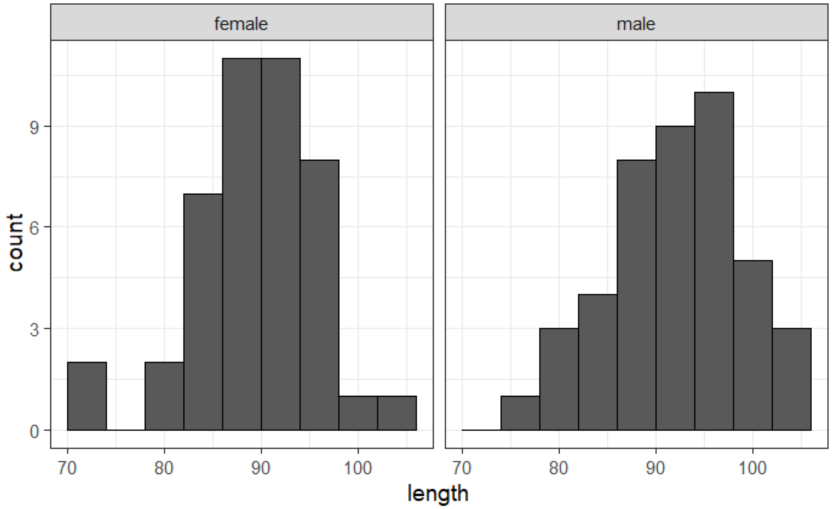
- Explore data using 4 different representations:



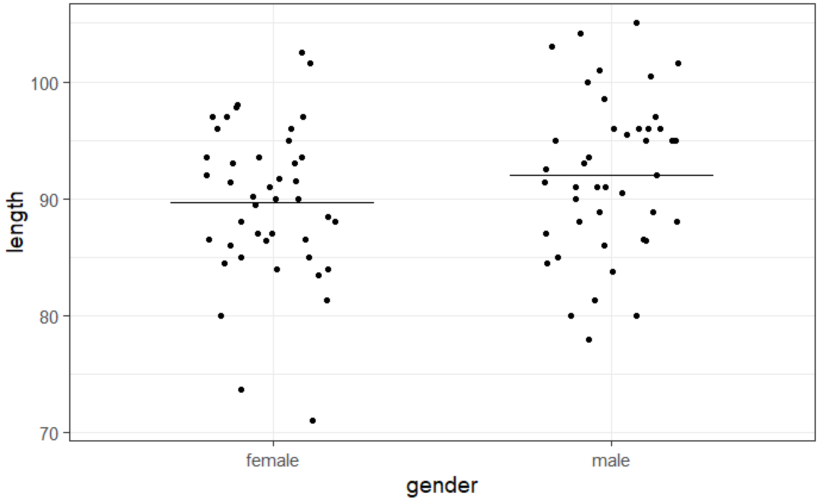
`geom_boxplot()`



`geom_violin()`



```
facet_grid(rows=vars(row),cols=vars(column))
geom_histogram
```

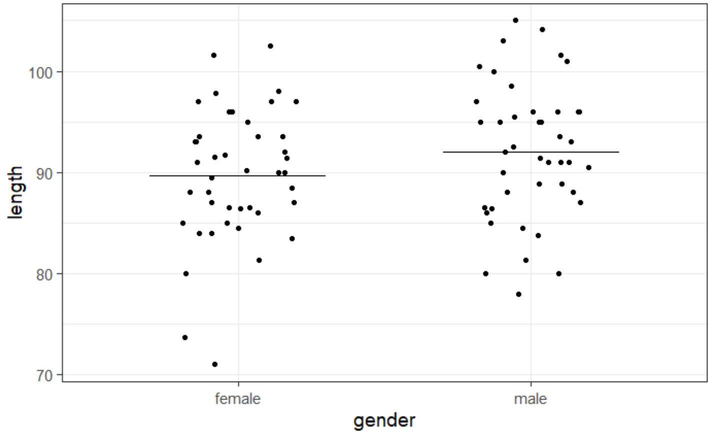


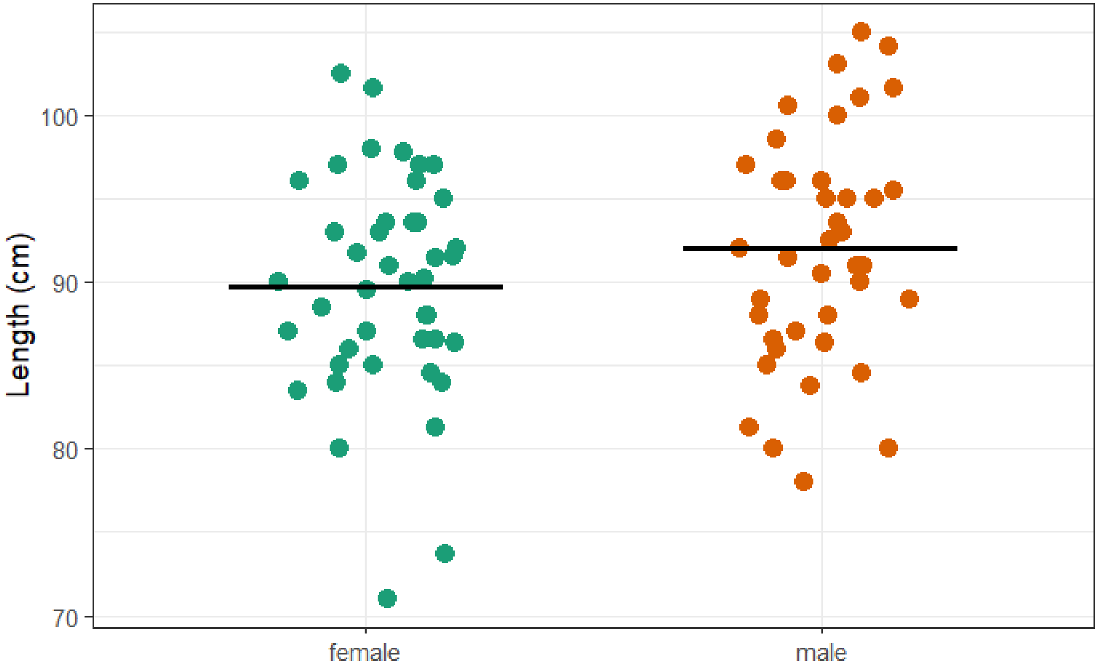
```
geom_jitter()
stat_summary()
```

Have a go!

Exercise 4: Exploring data - Stripchart

```
coyote %>%
  ggplot(aes(gender,length)) +
      geom_jitter(height=0, width=0.2)+
      stat_summary(geom= "errorbar", fun="mean", fun.min="mean", fun.max="mean")
```

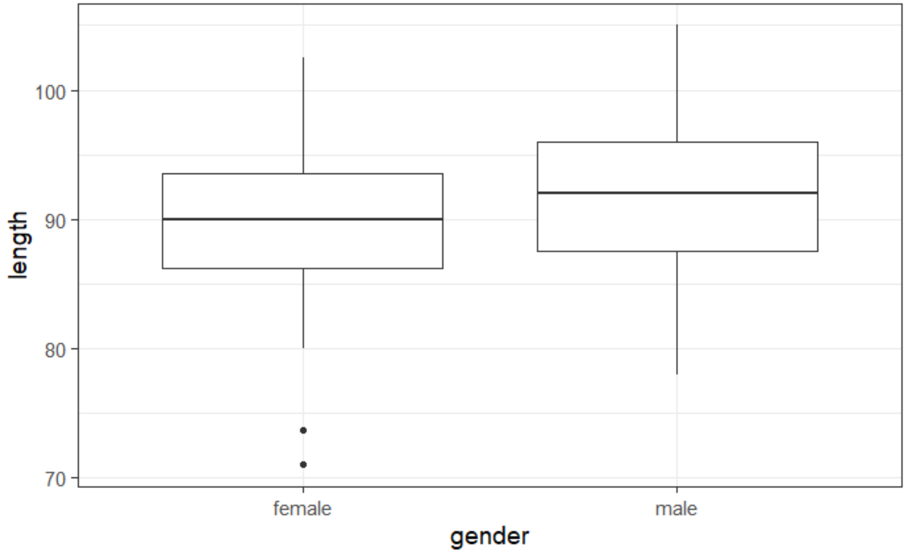


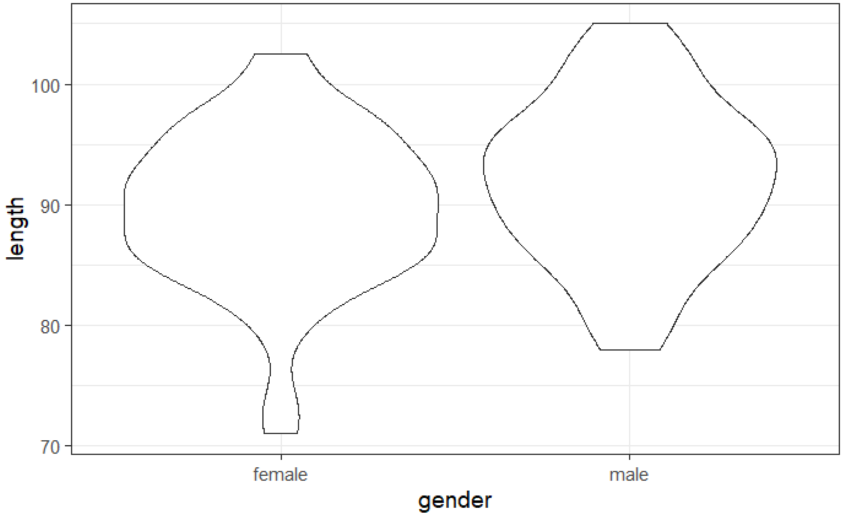


```
coyote %>%
  ggplot(aes(gender,length, colour=gender)) +
    geom_jitter(height=0, size=4, width=0.2, show.legend = FALSE) +
    ylab("Length (cm)")+
    scale_colour_brewer(palette="Dark2")+
    xlab(NULL)+
    stat_summary(geom="errorbar", fun=mean, fun.min=mean, fun.max=mean, colour="black", size=1.2, width=0.6)
```

Exercise 4: Exploring data - Boxplots and beanplots

```
coyote %>%
  ggplot(aes(x=gender, y=length)) +
  geom_boxplot()
```

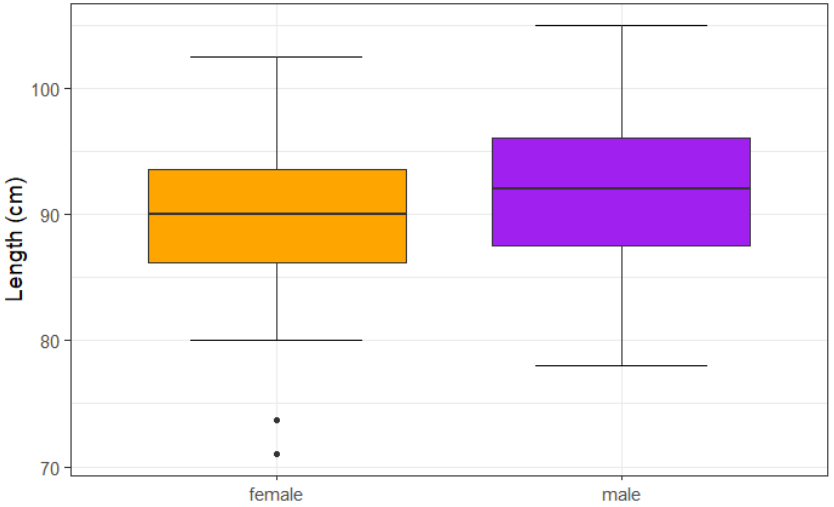




```
coyote %>%
  ggplot(aes(x=gender, y=length)) +
  geom_violin()
```

Exercise 4: Exploring data - Boxplots and beanplots

```
coyote %>%
  ggplot(aes(x=gender, y=length, fill=gender)) +
        stat_boxplot(geom="errorbar",width=0.5) +
        geom_boxplot(show.legend=FALSE)+
        ylab("Length (cm)")+
        xlab(NULL)+
        scale_fill_manual(values = c("orange","purple"))
```

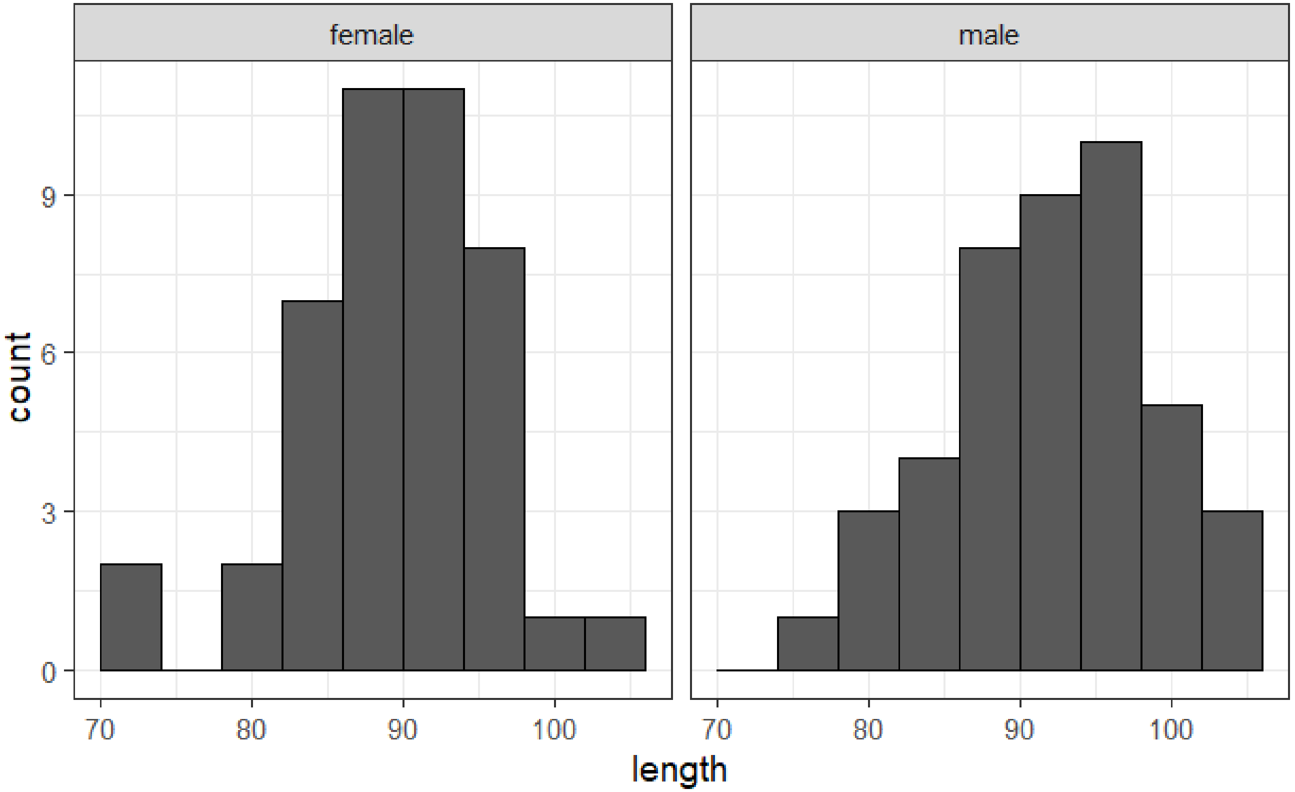


```
coyote %>%
  ggplot(aes(gender, length, fill=gender)) +
    geom_violin(trim=FALSE, size=1, show.legend=FALSE)+
    ylab("Length (cm)")+
    scale_fill_brewer(palette="Dark2")+
    stat_summary(geom = "point", fun = "median",show.legend=FALSE)
```

Exercise 4: Exploring data - Histograms

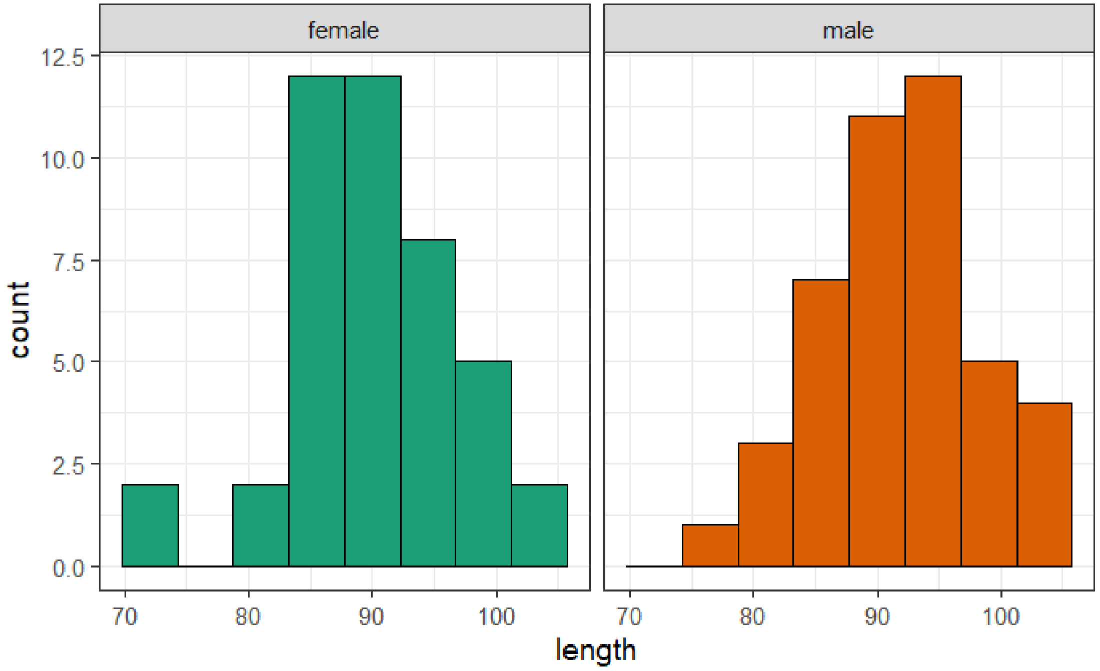
also works
`facet_wrap(vars(gender))`

```
coyote %>%
    ggplot(aes(length))+
      geom_histogram(binwidth = 4, colour="black") +
      facet_grid(cols=vars(gender))
```



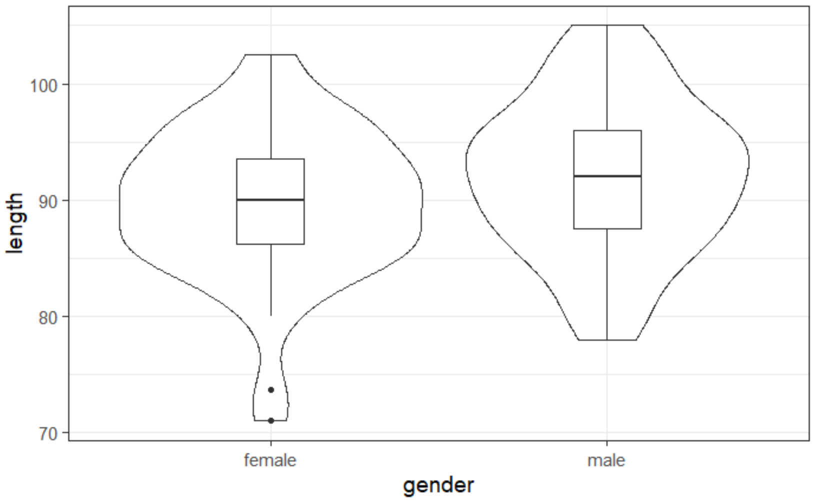
Exercise 4: Exploring data - Histograms

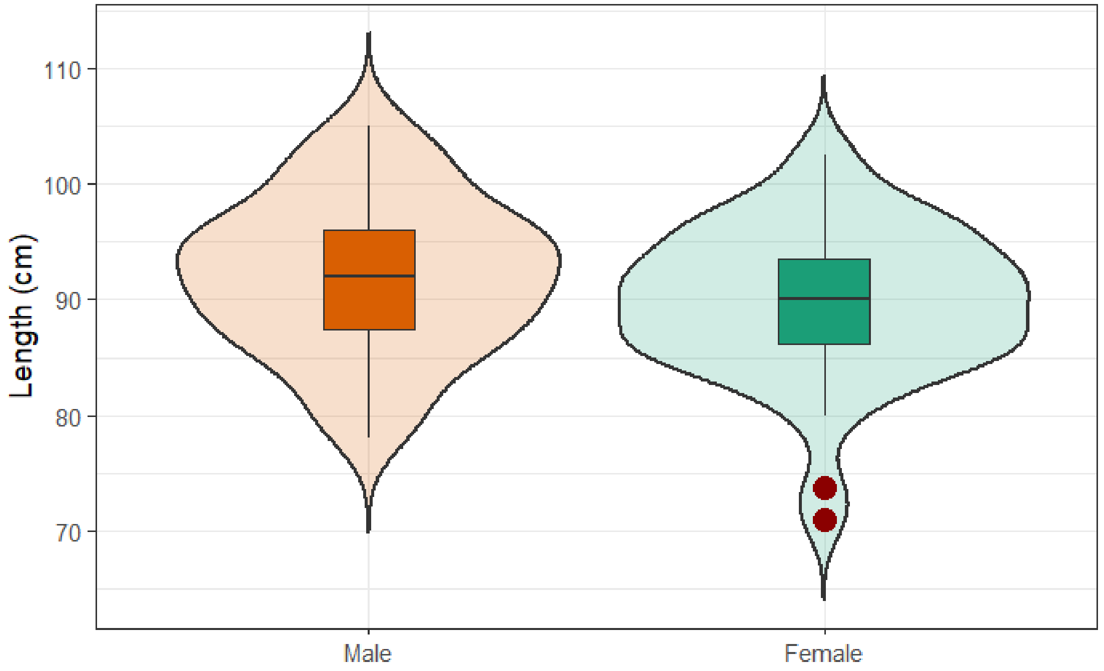
```
coyote %>%
  ggplot(aes(length, fill=gender))+
    geom_histogram(binwidth = 4.5, colour="black", show.legend = FALSE) +
    scale_fill_brewer(palette="Dark2")+
    facet_grid(cols=vars (gender))
```



Exercise 4 extra: Exploring data - Graph combinations

```
coyote %>%
  ggplot(aes(gender, length)) +
       geom_boxplot(width=0.2)+
       geom_violin()
```

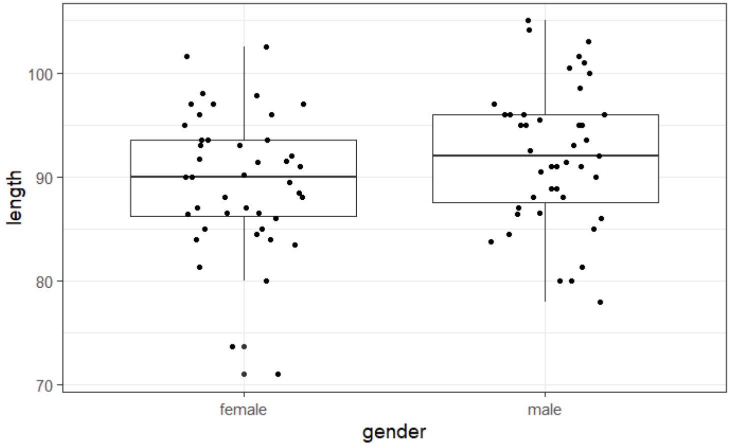


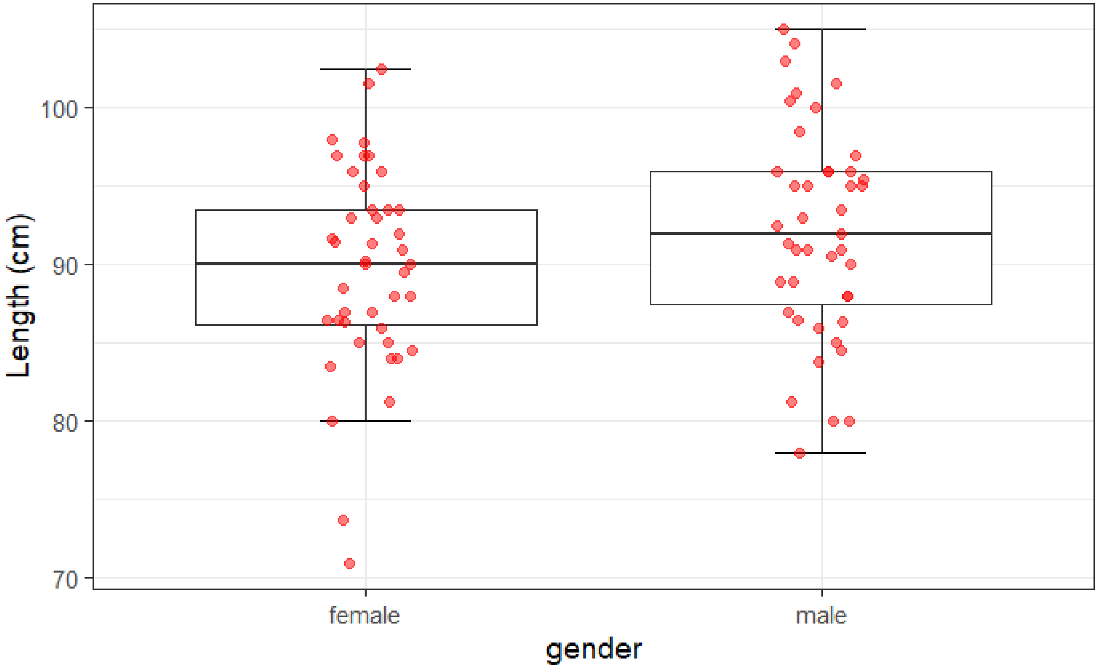


```
coyote %>%
  ggplot(aes(gender,length, fill=gender)) +
       geom_violin(size=1, trim = FALSE, alpha=0.2, show.legend=FALSE) +
       geom_boxplot(width=0.2, outlier.size=5, outlier.colour = "darkred", show.legend=FALSE)+
       scale_fill_brewer(palette="Dark2")+
       ylab("Length (cm)")+
       xlab(NULL)+
       scale_x_discrete(labels=c("female"="Female", "male"="Male"), limits =c("male", "female"))
```

Exercise 4 extra: Exploring data - Graph combinations

```
coyote %>%
  ggplot(aes(gender, length)) +
  geom_boxplot()+
  geom_jitter(height=0, width=0.2)
```



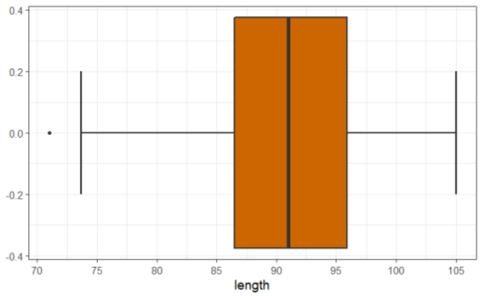


```
coyote %>%
  ggplot(aes(gender, length)) +
       geom_boxplot(outlier.shape=NA)+
       stat_boxplot(geom="errorbar", width=0.2)+
       geom_jitter(height=0, width=0.1, size=2, alpha=0.5, colour="red")+
       ylab("Length (cm)")
```

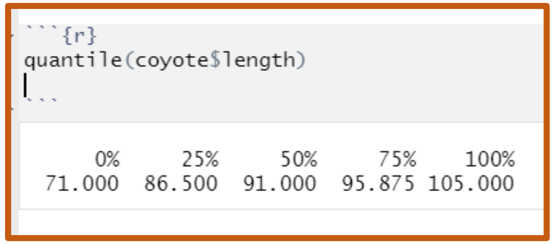
Checking the assumptions

Normality assumption: QQ Plot

QQ plot= Quantile – Quantile plot

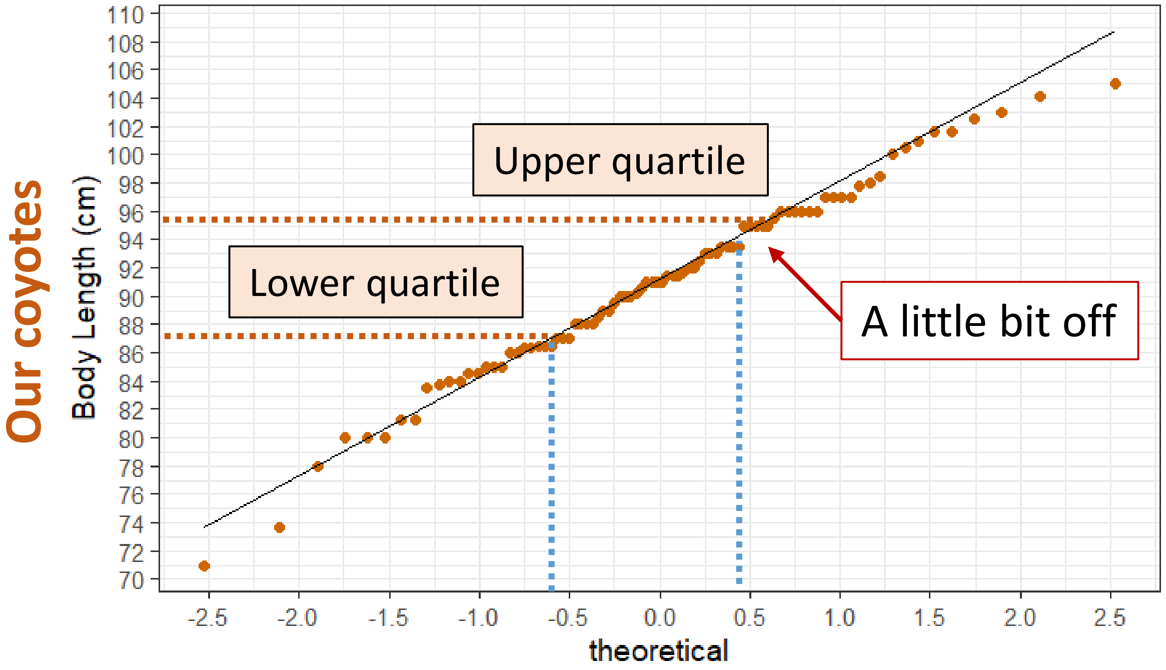


Quantiles:



~~~
```{r}
quantile(coyote$length)
```
~~~

```
     0%     25%     50%     75%    100% 
 71.000  86.500  91.000  95.875 105.000 
```



Upper quartile

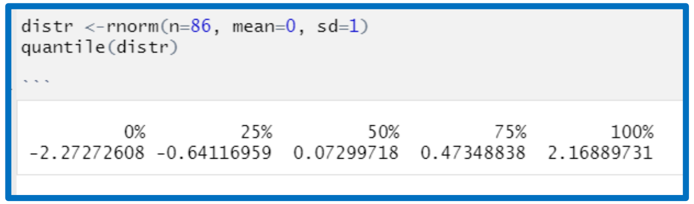
Lower quartile

A little bit off

Normality ☑ (ish)

Mean = 0
SD = 1
Same sample size
Perfectly normal distribution

Quantiles:

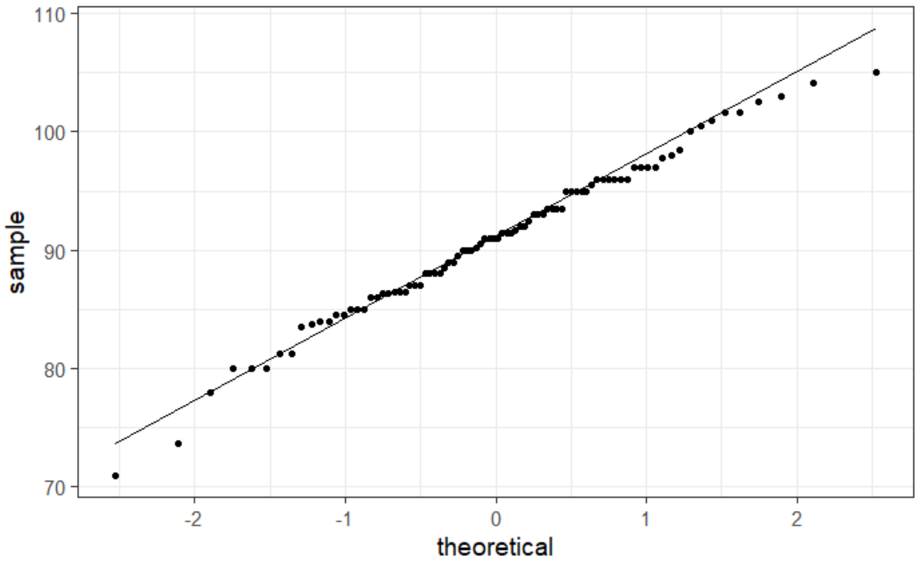


~~~
distr <-rnorm(n=86, mean=0, sd=1)
quantile(distr)
```
~~~

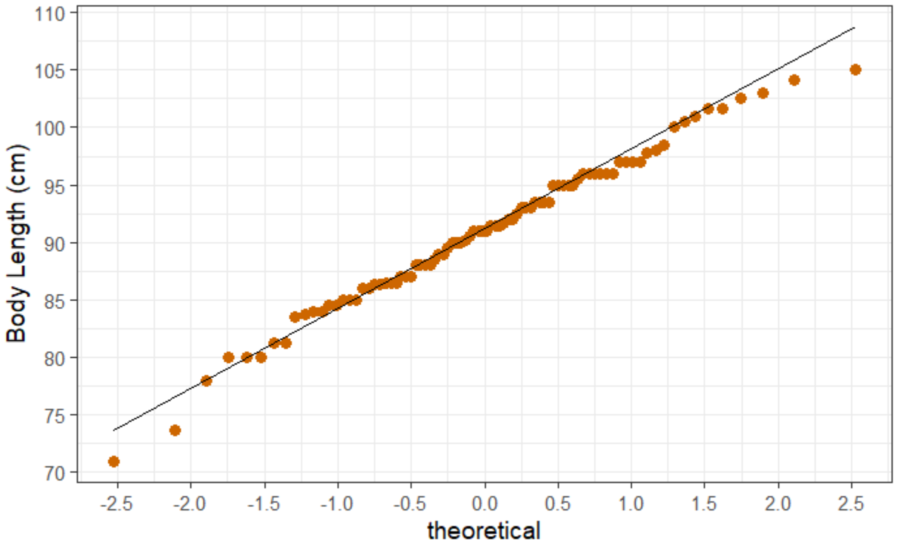
```
         0%         25%         50%         75%        100% 
-2.27272608 -0.64116959  0.07299718  0.47348838  2.16889731 
```

Normality assumption: QQ plot

```
coyote %>%
  ggplot(aes(sample = length)) +
  stat_qq()+
  stat_qq_line()
```



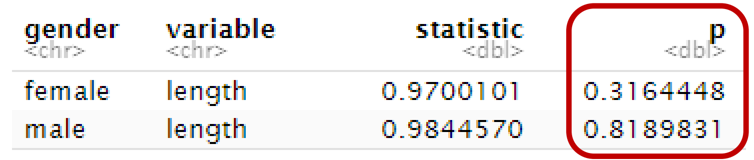
```
coyote %>%
  ggplot(aes(sample = length)) +
  stat_qq(size=2, colour="darkorange3")+
  stat_qq_line()+
  ylab("Body Length (cm)")+
  scale_y_continuous(breaks=seq(from=70, by=5, to=110))+
  scale_x_continuous(breaks=seq(from=-3, by=0.5, to=3))
```



Assumptions of Parametric Data

- First assumption: Normality
  - ❖ Shapiro-Wilk test **`shapiro_test()`** # rstatix package #
    - ❖ It is based on the correlation between the data and the corresponding normal scores.
- Second assumption: Homoscedasticity
  - ❖ Levene test **`levene_test()`**

```
coyote %>%
  group_by(gender) %>%
   shapiro_test(length)%>%
        ungroup()
```



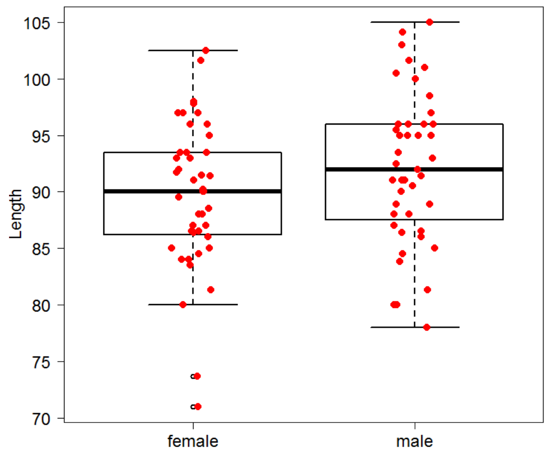
```
gender   variable         statistic          p
<chr>    <chr>               <dbl>       <dbl>
female   length          0.9700101   0.3164448
male     length          0.9844570   0.8189831
```

Normality ☑

```
coyote %>%
  levene_test(length ~ gender)
```

```
df1    df2        statistic             p
<int>  <int>          <dbl>         <dbl>
  1     84         0.167929     0.6830022
```

Homogeneity in variance ☑



Normality

Other classic: D'Agostino-Pearson test
fBasic package
`dagoTest()`

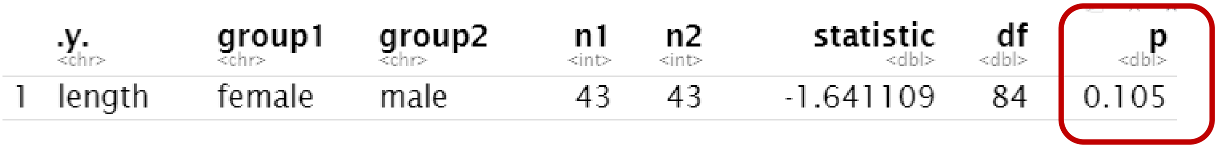
Homoscedasticity

More robust: Brown-Forsythe test
onewaytests package #, `bf()`
Other classic: Bartlett test
`bartlett.test()`

Independent *t*-test: results (tidyverse)

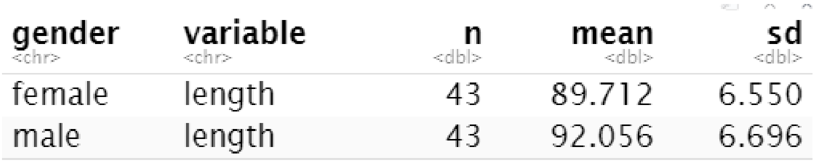
coyote.csv

```
coyote %>%
      t_test(length~gender)
```



```
  .y.     group1  group2   n1   n2   statistic   df     p
  <chr>   <chr>   <chr>  <int> <int>     <dbl> <dbl> <dbl>
1 length  female  male     43   43   -1.641109   84  0.105
```

```
coyote %>%
  group_by(gender) %>%
  get_summary_stats(length, type = "mean_sd") %>%
    ungroup()
```



```
gender  variable      n    mean     sd
<chr>   <chr>     <dbl>   <dbl>  <dbl>
female  length       43  89.712  6.550
male    length       43  92.056  6.696
```

- **<u>Answer</u>: Males tend to be longer than females but not significantly so (p=0.1045).**

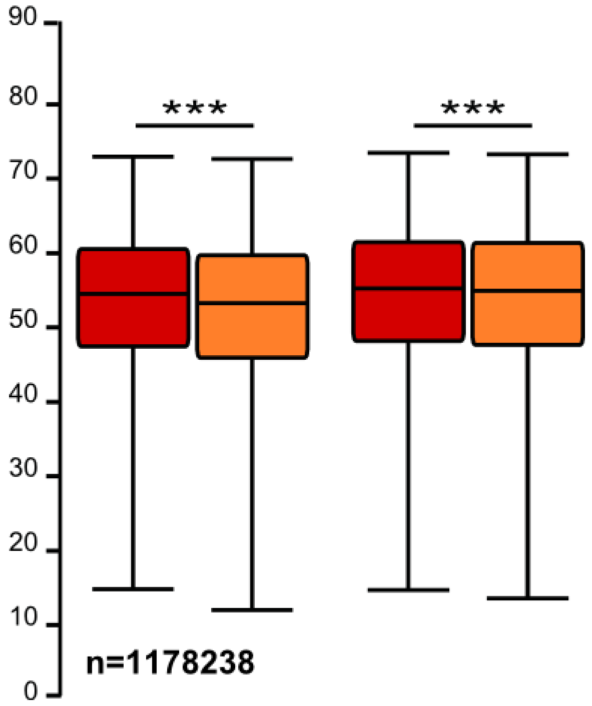
- Power : How many more coyotes to reach significance?
  - Re-run the power analysis with mean=89.7 for females: n~250
    - **But does it make sense?**

Sample size: the bigger the better?

- It takes huge samples to detect tiny differences but tiny samples to detect huge differences.

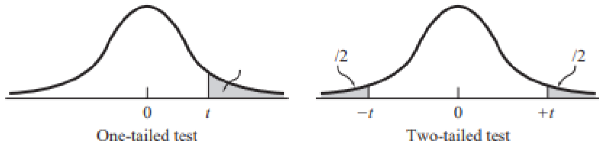
- What if the tiny difference is meaningless?
  - Beware of **overpower**
  - Nothing wrong with the stats: it is all about interpretation of the results of the test.

- Remember the important first step of power analysis
  - **What is the effect size of biological interest?**



Independent *t*-test: results

The old-fashion way



	Level of Significance for One-Tailed Test								
	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.0005
	Level of Significance for Two-Tailed Test								
df	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.620
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.32	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
21	0.686	0.859	1.063	1.32					
22	0.686	0.858	1.061	1.321					
23	0.685	0.858	1.060	1.31					
24	0.685	0.857	1.059	1.31					
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	3.496
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.390
	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

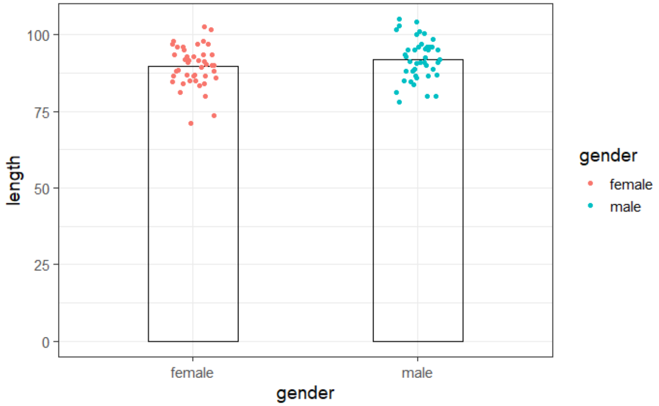
```
   n1    n2   statistic
<int> <int>       <dbl>
   43    43   -1.641109
```

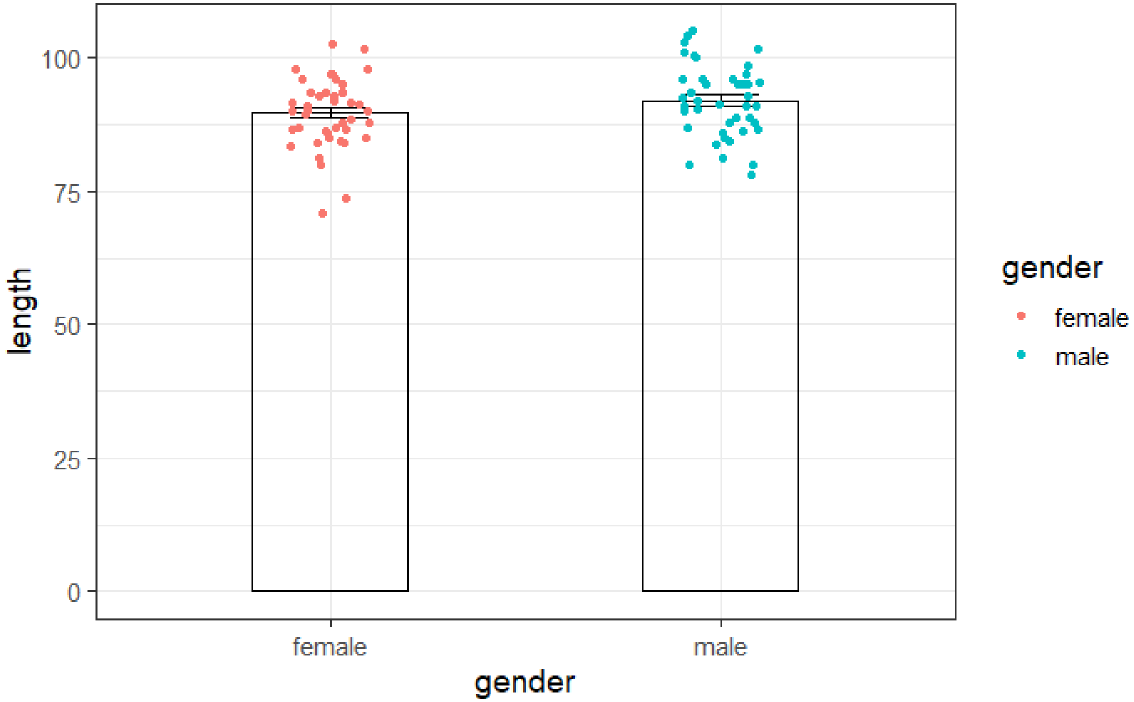
***t* = 1.641 < 1.984: not significant**

Critical value

Plot ‘coyote.csv’ data: Plotting data

```
coyote %>%
  ggplot(aes(gender,length, colour=gender)) +
    geom_bar(stat = "summary", fun="mean", width=0.4, alpha=0, colour="black")+
    geom_jitter(height=0, width=0.1)
```





- Add error bars

```
coyote %>%
  ggplot(aes(gender,length, colour=gender)) +
    geom_bar(stat = "summary", fun="mean", width=0.4, alpha=0, colour="black")+
    geom_jitter(height=0, width=0.1)+
    stat_summary(geom="errorbar", colour="black", width=0.2)
```

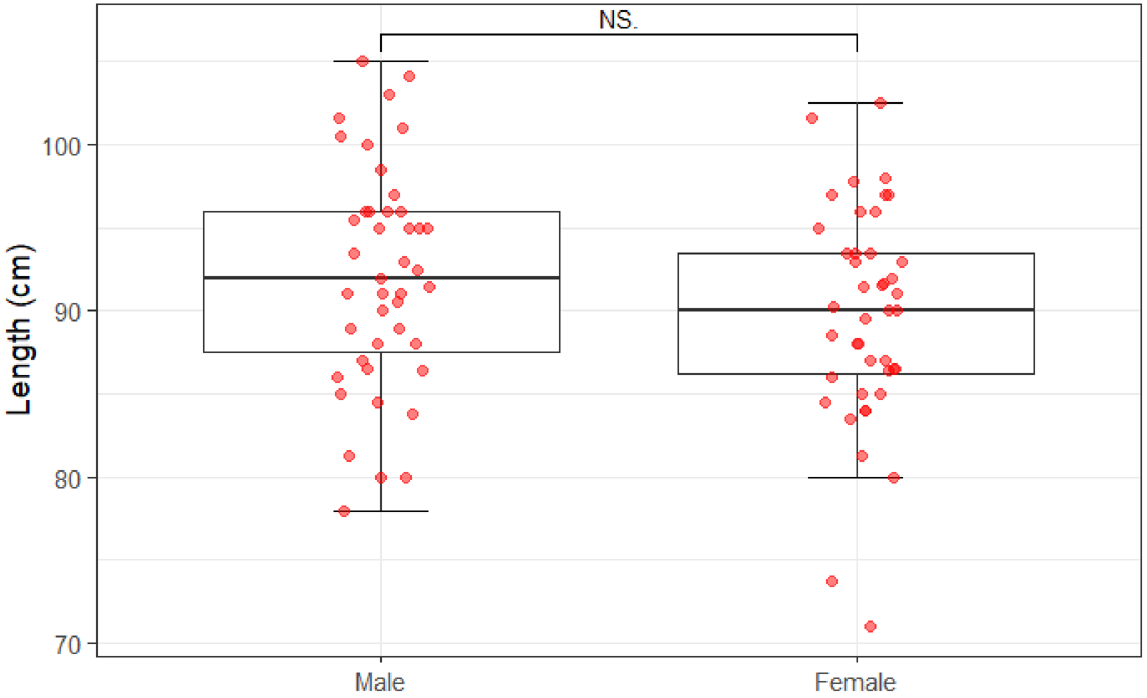
Plot ‘coyote.csv’ data: Plotting data

- Prettier version

```
coyote %>%
  ggplot(aes(gender,length, colour=gender, fill=gender)) +
    geom_bar(stat="summary", fun="mean", width=0.4, alpha=0.2, colour="black", show.legend=FALSE)+
    stat_summary(geom="errorbar", colour="black", width=0.2)+
    geom_jitter(height=0, width=0.1, show.legend=FALSE)+
    scale_colour_brewer(palette="Dark2")+
    scale_fill_brewer(palette="Dark2")+
    theme(legend.position = "none")+
    scale_x_discrete(limits = c("male", "female"), labels = c("male"="Male", "female"="Female"))+
    xlab(NULL)+
    ylab("Length (cm)")
```

Plot ‘coyote.csv’ data: Plotting data

- ***Work in progress*** # ggsignif package #



```
coyote %>%
  ggplot(aes(gender, length)) +
    stat_boxplot(geom="errorbar", width=0.2)+
    geom_boxplot(outlier.shape = NA)+
    geom_jitter(height=0, width=0.1, size = 2, alpha = 0.5, colour="red")+
    scale_x_discrete(limits = c("male", "female"), labels = c("male"="Male", "female"="Female"))+
    ylab("Length (cm)")+
    xlab(NULL)+
    geom_signif(comparisons = list(c("female", "male")), map_signif_level=T, test = "t.test")
```

Exercise 5: Dependent or Paired *t*-test

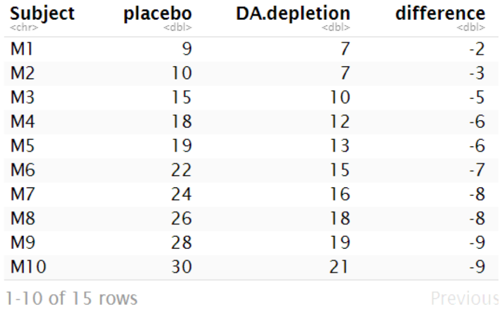
working.memory.csv

- A researcher is studying the effects of dopamine depletion on working memory in rhesus monkeys.
  - A group of rhesus monkeys (n=15) performs a task involving memory after having received a placebo. Their performance is graded on a scale from 0 to 100. They are then asked to perform the same task after having received a dopamine depleting agent.

- **Question**: does dopamine affect working memory in rhesus monkeys?
  - Load **working.memory.csv** and check out the structure of the data.
  - Work out the difference: DA.depletion – placebo and assign the difference to a column: working.memory$difference
  - Plot the difference as a stripchart with a mean
  - Add **confidence intervals as error bars**
    - Clue: `stat_summary(…, fun.data=mean_cl_normal)` # Hmisc package #
  - Run the paired *t*-test. `t_test(var ~ 1, mu=0)`

Exercise 5: Dependent or Paired *t*-test - *Answers*

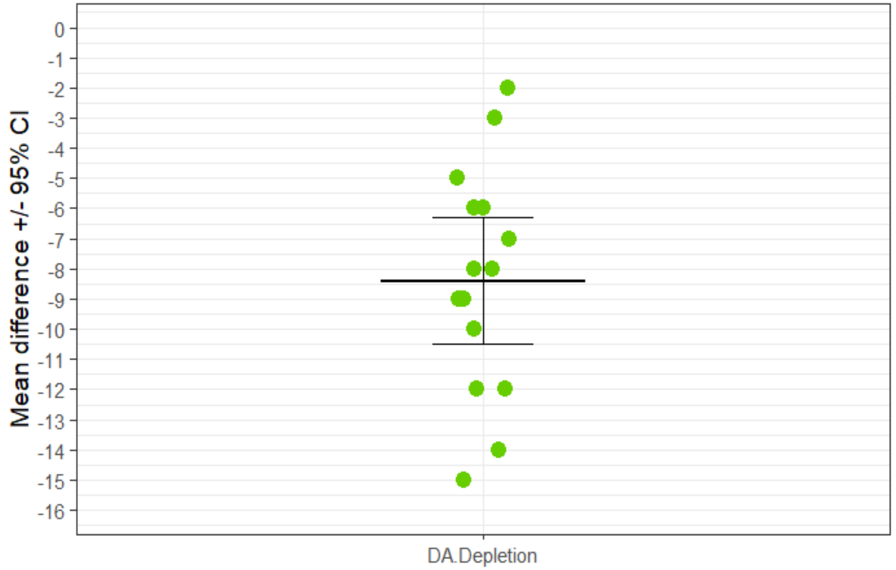
```
working.memory %>%
  mutate(difference = DA.depletion - placebo) -> working.memory
```



Subject <chr>	placebo <dbl>	DA.depletion <dbl>	difference <dbl>
M1	9	7	-2
M2	10	7	-3
M3	15	10	-5
M4	18	12	-6
M5	19	13	-6
M6	22	15	-7
M7	24	16	-8
M8	26	18	-8
M9	28	19	-9
M10	30	21	-9

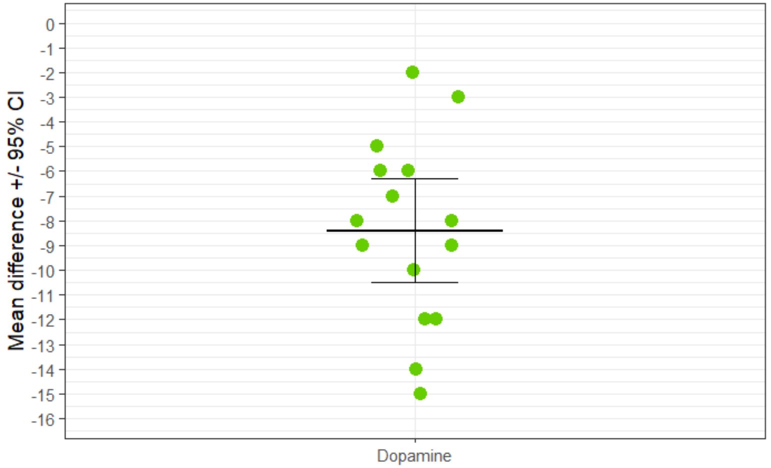
1-10 of 15 rows Previous

```
# Hmisc package #
working.memory %>%
  ggplot(aes("DA.Depletion", difference))+
    geom_jitter(height=0, width=0.05, size=4, colour="chartreuse3")+
    stat_summary(geom="errorbar",fun="mean", fun.min="mean", fun.max="mean",  width=0.3, size=1)+
    stat_summary(geom="errorbar", fun.data=mean_cl_normal, width=0.15)+
    scale_y_continuous(breaks=-16:0, limits=c(-16, 0))+
    xlab(NULL)+
    ylab("Mean difference +/- 95% CI")
```



Exercise 5: Dependent or Paired *t*-test (tidyverse)

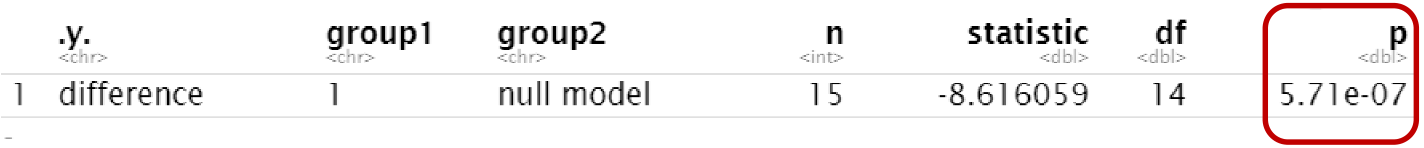
Question: does dopamine affect working memory in rhesus monkeys?



```
working.memory %>%
  shapiro_test(difference)
```

```
variable              statistic           p
<chr>                     <dbl>       <dbl>
difference            0.9772671   0.9474075
```

```
working.memory %>%
  t_test(difference ~ 1, mu=0)
```



```
  .y.          group1   group2        n    statistic   df          p
  <chr>        <chr>    <chr>     <int>        <dbl>  <dbl>      <dbl>
1 difference   1        null model   15    -8.616059   14   5.71e-07
```

Answer: the injection of a dopamine-depleting agent significantly affects working memory in rhesus monkeys (t=-8.62, df=14, p=5.715e-7).

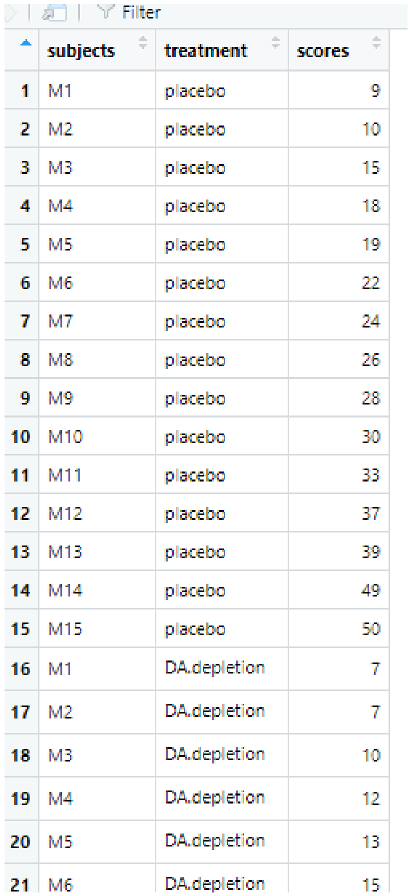
Dependent or Paired *t*-test

- ***Work in progress*** # ggpubr package #

```
working.memory.long %>%
  t_test(scores ~ treatment, paired = TRUE) -> stat.test

working.memory.long %>%
  ggpaired(x = "treatment", y = "scores", color = "treatment",
  palette = "Dark2", line.color = "gray", line.size = 0.4)+
  scale_y_continuous(breaks=seq(from =0, by=5, to=60),
       limits = c(0,60))+
  stat_pvalue_manual(stat.test, label = "p", y.position = 55)
```

working.memory.long



	subjects	treatment	scores
1	M1	placebo	9
2	M2	placebo	10
3	M3	placebo	15
4	M4	placebo	18
5	M5	placebo	19
6	M6	placebo	22
7	M7	placebo	24
8	M8	placebo	26
9	M9	placebo	28
10	M10	placebo	30
11	M11	placebo	33
12	M12	placebo	37
13	M13	placebo	39
14	M14	placebo	49
15	M15	placebo	50
16	M1	DA.depletion	7
17	M2	DA.depletion	7
18	M3	DA.depletion	10
19	M4	DA.depletion	12
20	M5	DA.depletion	13
21	M6	DA.depletion	15