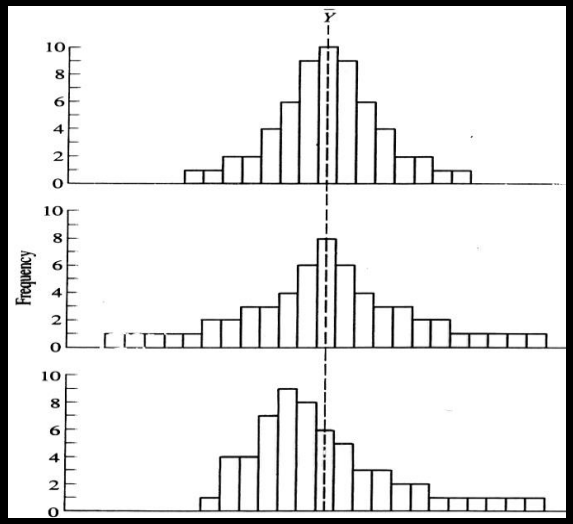
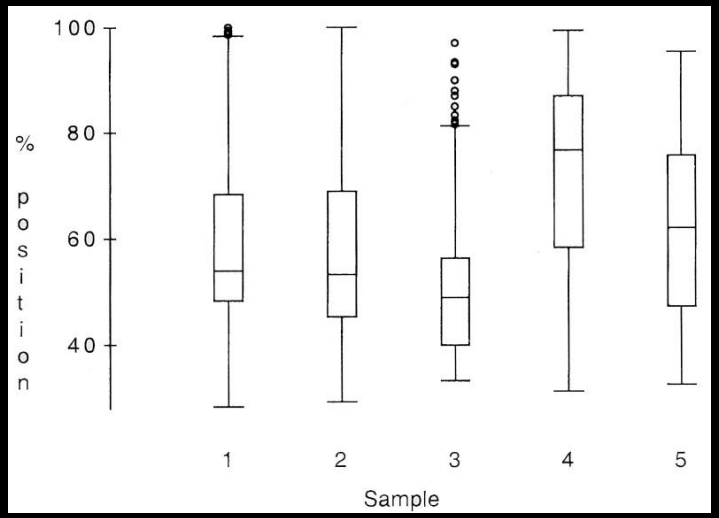
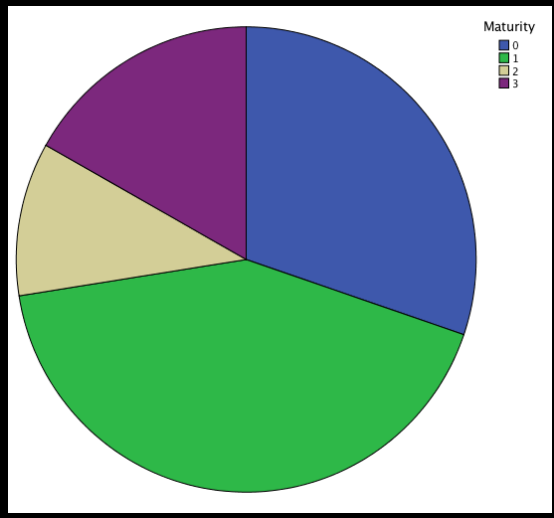
Data Handling :

A practical approach







Lecture 6 ANOVA II

Dr Yu Mo, Zoology

moyu@tcd.ie | https://github.com/github-moyu/Teaching

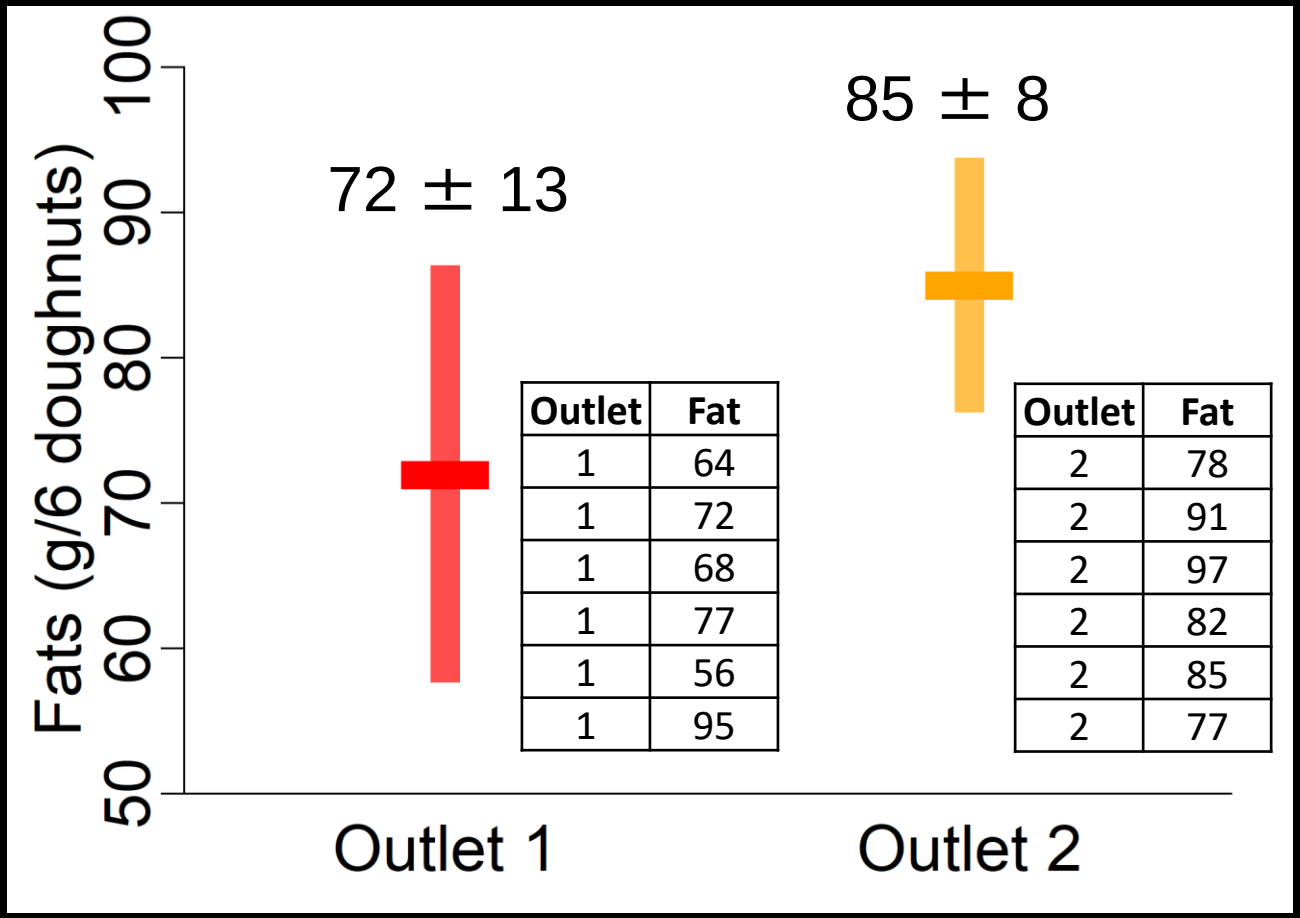
Fats in doughnuts

Independent t-test

t = -2.0624
df = 11
p = 0.06612

If alpha 0.05

Fail to reject H_0



Outlet	Fat
1	64
1	72
1	68
1	77
1	56
1	95

Outlet	Fat
2	78
2	91
2	97
2	82
2	85
2	77

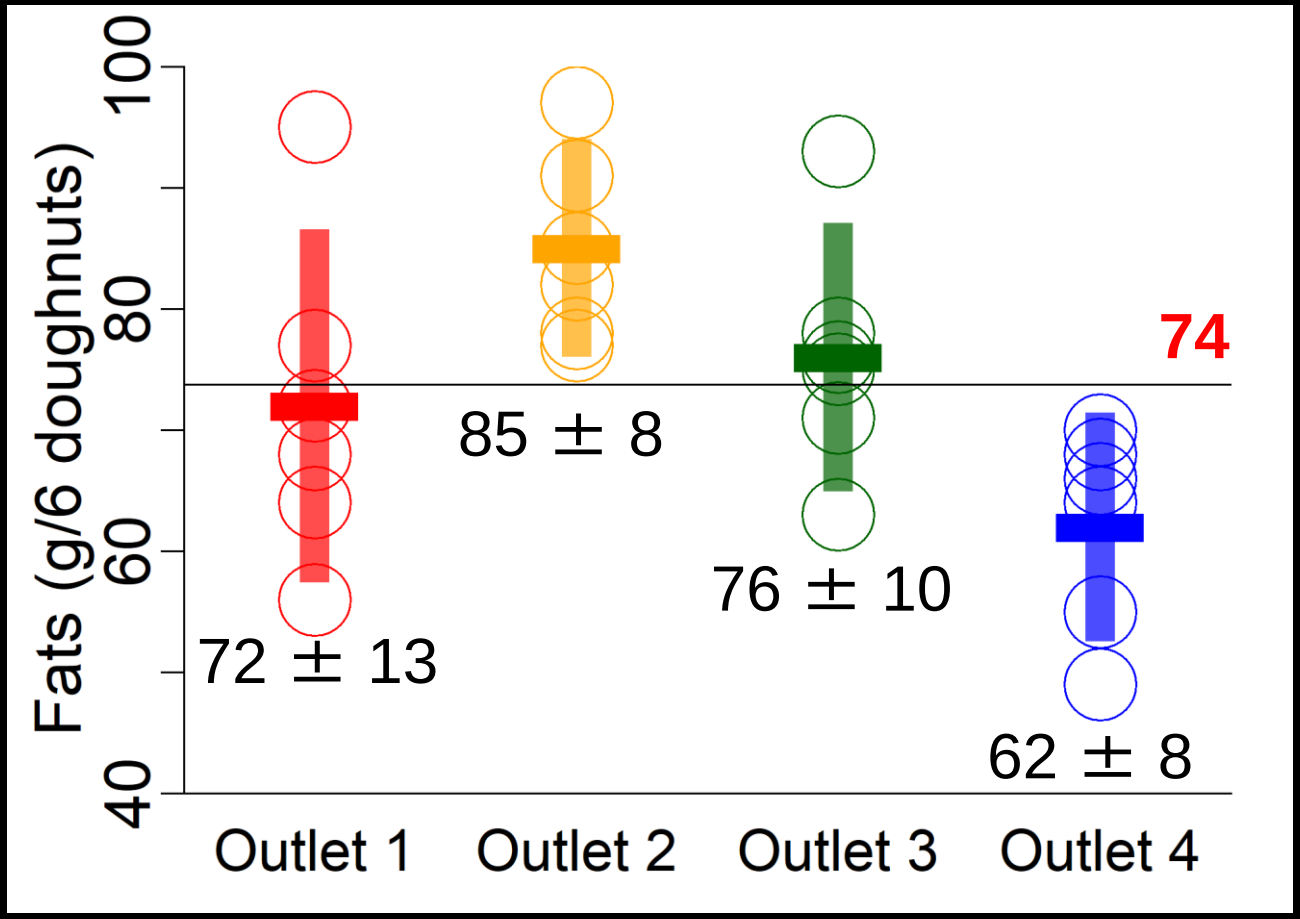
Summary of lecture 5

- Comparison of multiple groups using ANOVA
- Between group variability, Within group variability
- F value, df of denominator, df of numerator, and p value

Fats in doughnuts

$$SS_w = \sum_{j=1}^{k} \sum_{i=1}^{n} \left(X_{i,j} - \bar{X}_{,j}\right)^2$$

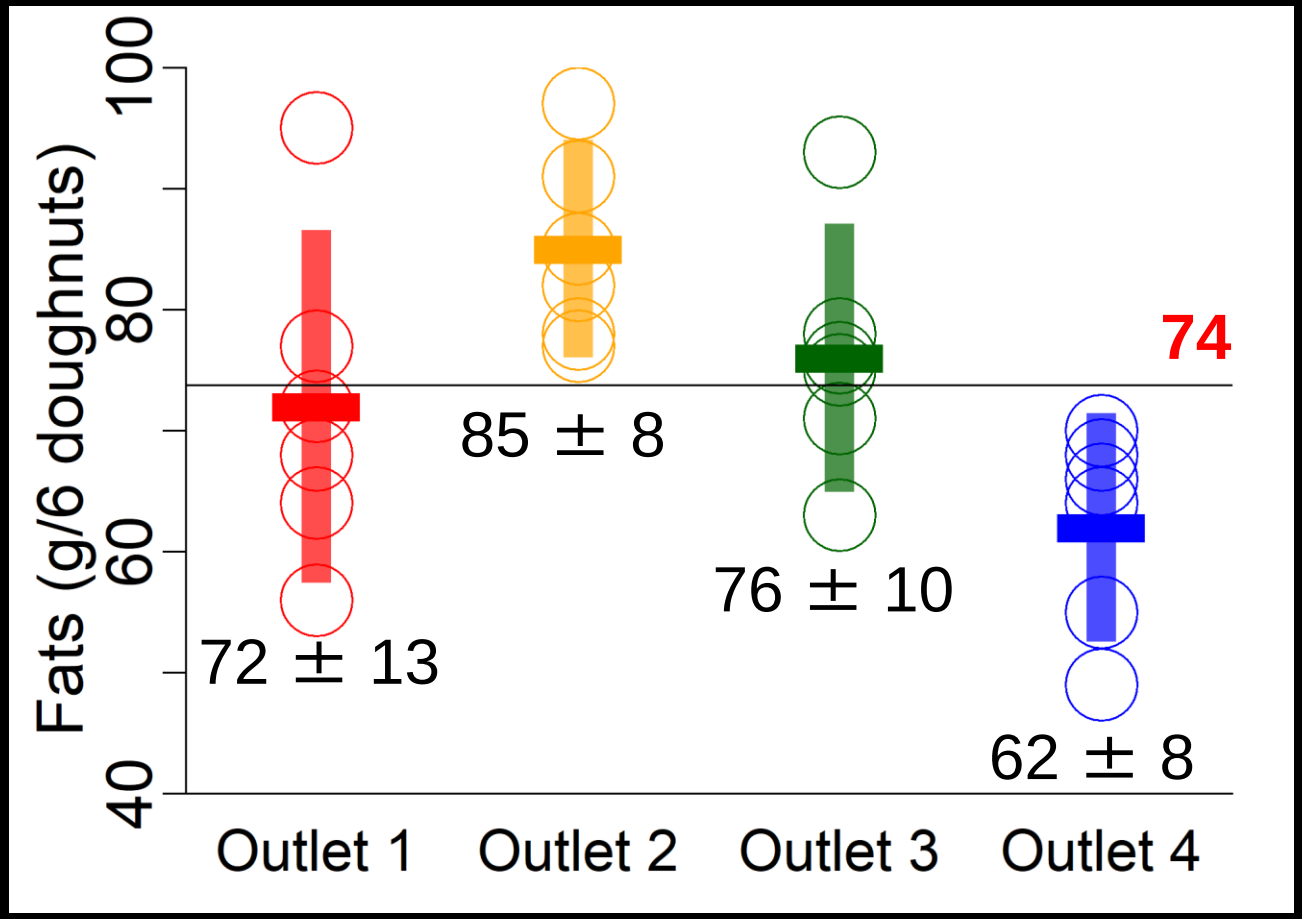
$$SS_B = \sum n_j (\bar{X}_{,j} - \bar{\bar{X}})^2$$



Fats in doughnuts

$F = MS_B/MS_W$

j-1 = numerator (j number of groups)
n-j = denominator (n number of obs.)



df Denominator n-j: 24-4 = 20

	Alpha	1	2	3	4	5	V1
V2							
16							
17							
18							
19							
	0.75						
	0.5						
	0.25						
	0.1						
20	0.05	4.35	3.49	3.1	2.87	2.71	
	0.025			3.86			
	0.01			4.94			
	0.005			5.82			
	0.001			8.10			

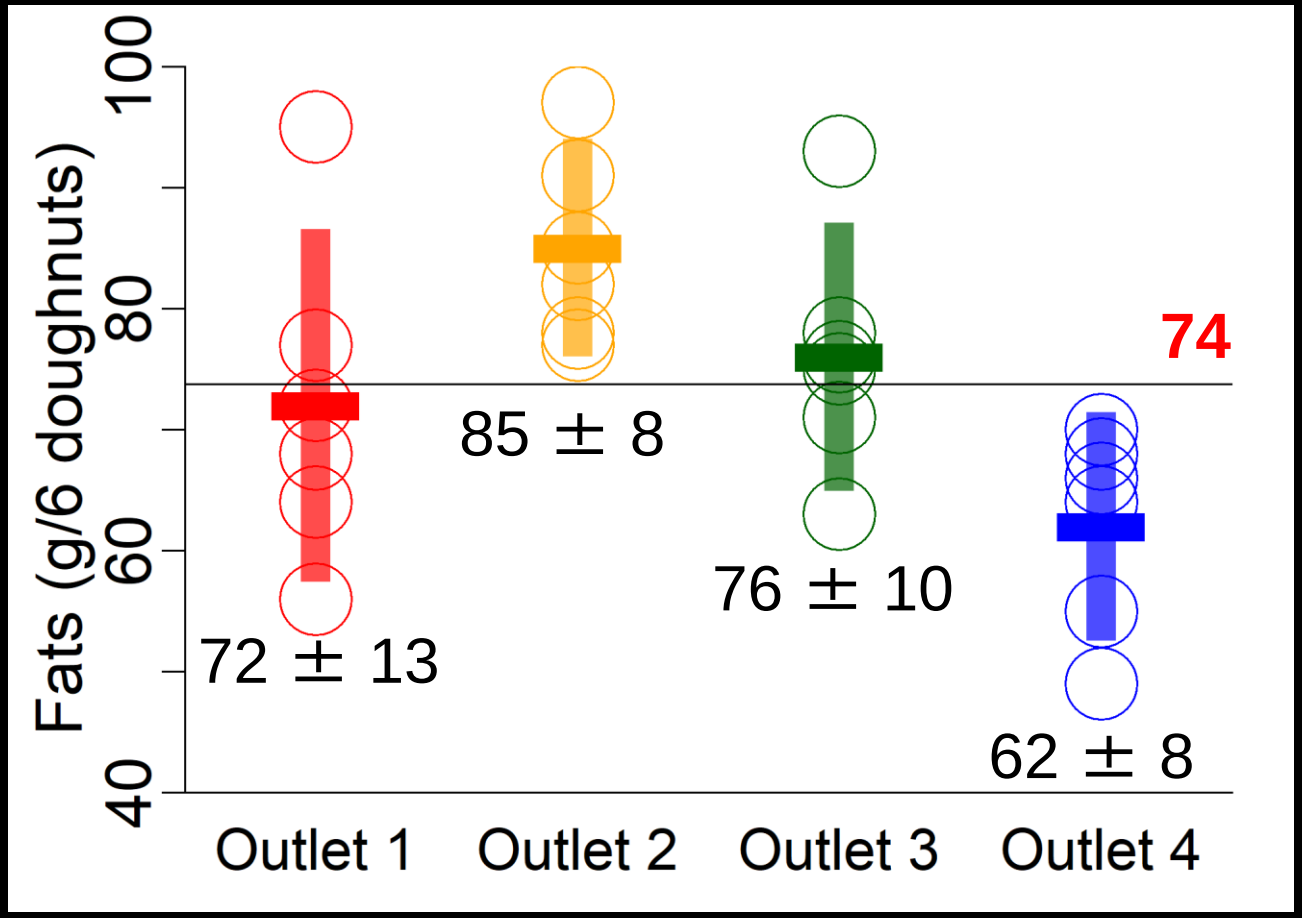
df Numerator
j-1
4-1 = 3

Critical values of the F distribution

```
>       #ANOVA
>       oneway.test(Fat~Outlet, data=data,var.equal = TRUE )

        One-way analysis of means

data:  Fat and Outlet
F = 5.4063, num df = 3, denom df = 20, p-value = 0.006876
```

Fats in doughnuts

Which ones?



Post Hoc Tests with Tukey's Method

- Control familywise error (VS series of t-test)

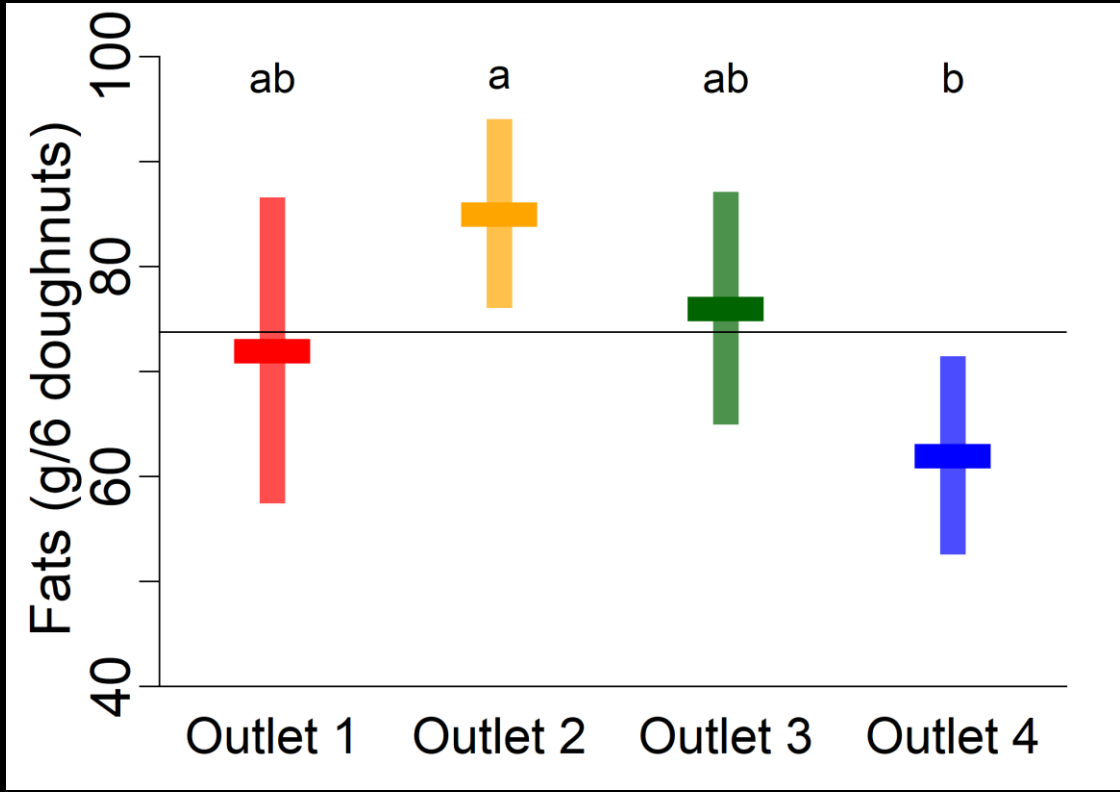
	t-test	ANOVA
Ho:	$\mu_1 = \mu_2$ $\mu_1 = \mu_3$ $\mu_1 = \mu_4$ $\mu_2 = \mu_3$ $\mu_2 = \mu_4$ $\mu_3 = \mu_4$	$\mu_1 = \mu_2 = \mu_3 = \mu_4$
Type I error:	$1-(.95)^6 = .26$	0.05

26% chance of suggesting an effect when there isn't one!

Post Hoc Tests with Tukey's Method

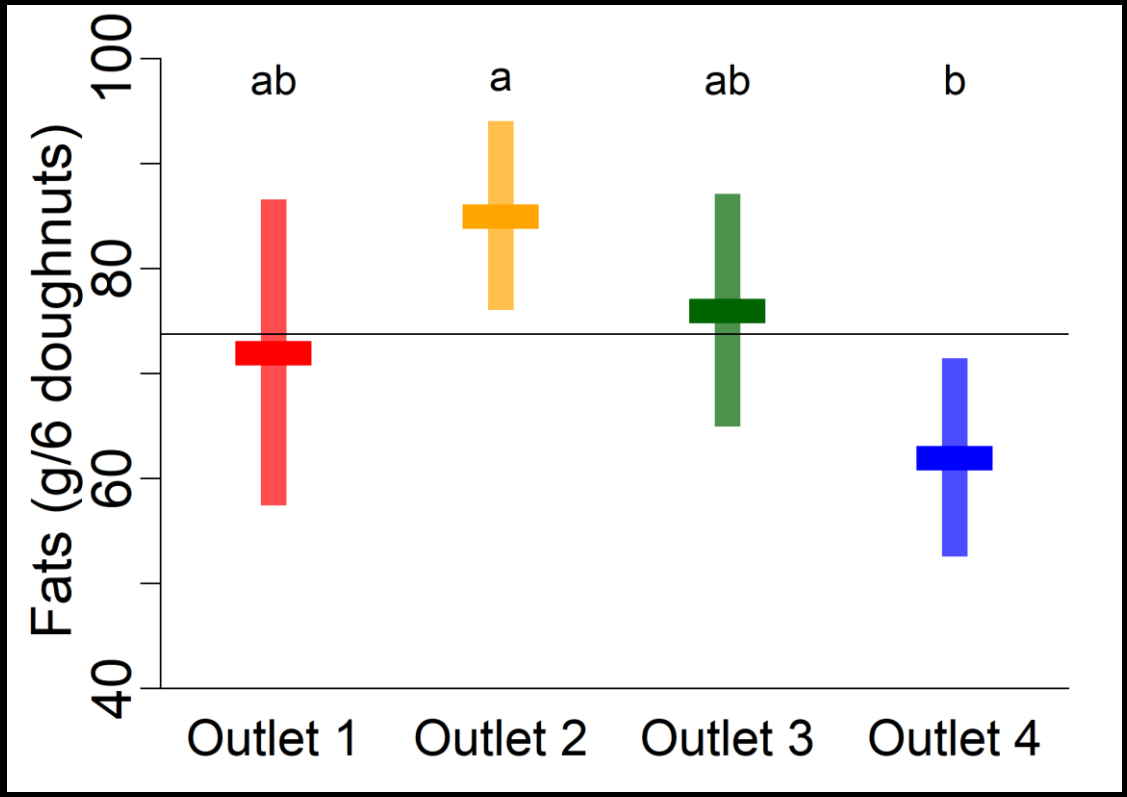
```
>   TukeyHSD(aov)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Fat ~ Outlet2, data = data)

$Outlet2
     diff         lwr       upr     p adj
2-1    13  -3.232221 29.232221 0.1461929
3-1     4 -12.232221 20.232221 0.8998057
4-1   -10 -26.232221  6.232221 0.3378150
3-2    -9 -25.232221  7.232221 0.4270717
4-2   -23 -39.232221 -6.767779 0.0039064
4-3   -14 -30.232221  2.232221 0.1065573
```

Post Hoc Tests with Tukey's Method



	diff	lwr	upr	p adj
2-1	13	-3.232221	29.232221	0.1461929
3-1	4	-12.232221	20.232221	0.8998057
4-1	-10	-26.232221	6.232221	0.3378150
3-2	-9	-25.232221	7.232221	0.4270717
4-2	-23	-39.232221	-6.767779	0.0039064
4-3	-14	-30.232221	2.232221	0.1065573

Post Hoc Tests with Tukey's Method

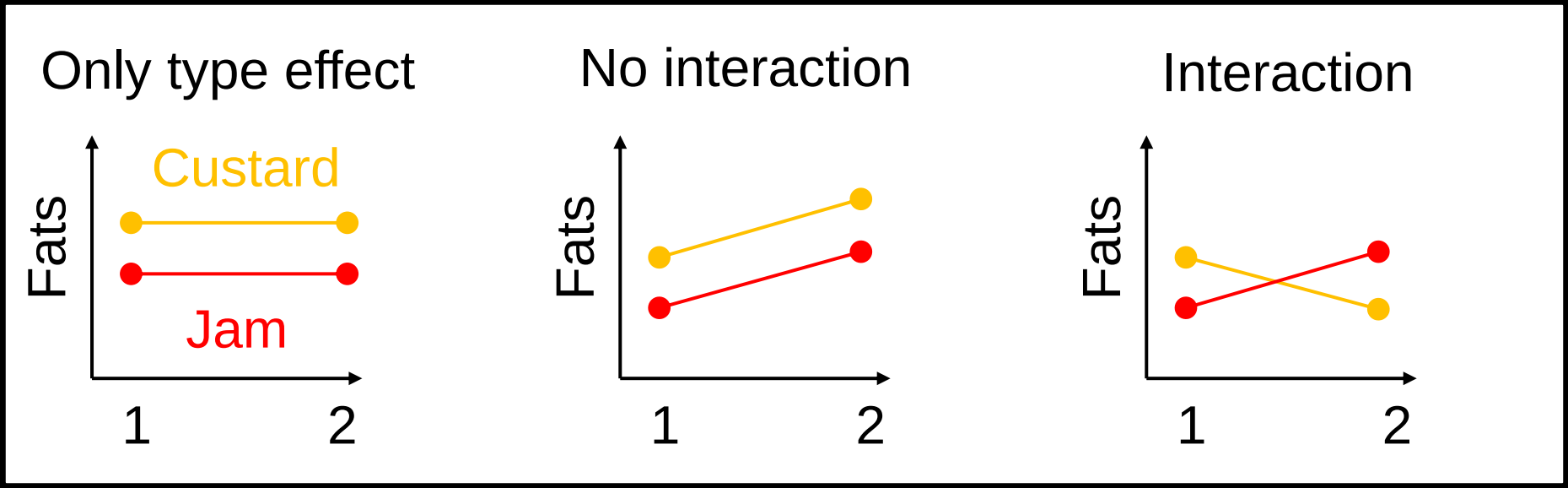


What if?				
	b	**a**	**b**	**b**
	a	**a**	**a**	**b**
	b	**a**	**b**	**c**

Fats in doughnuts

Two-way ANOVA

- Effect of factor 1
- Effect of factor 2
- Interaction: the effect of one factor depends on the other factor



Two-way ANOVA

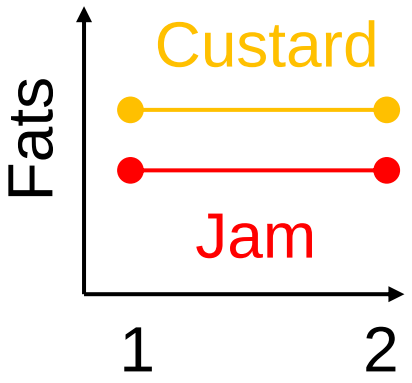
H_0:

1 No effect of factor 1
2 No effect of factor 2
3 No interaction

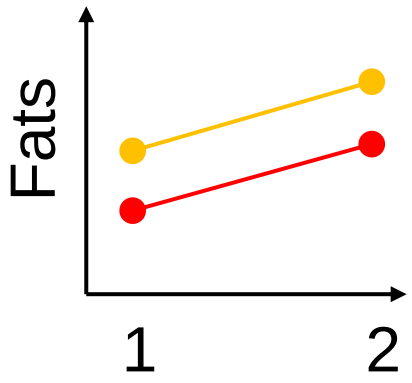
H_1:

Factor 1 | factor 2 | interaction

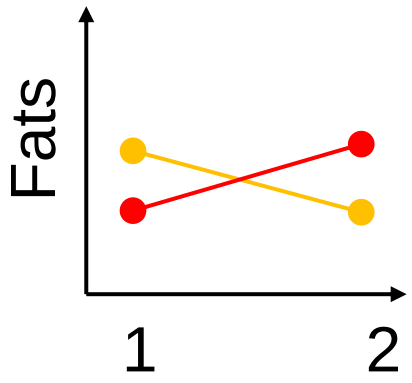
Only type effect



No interaction



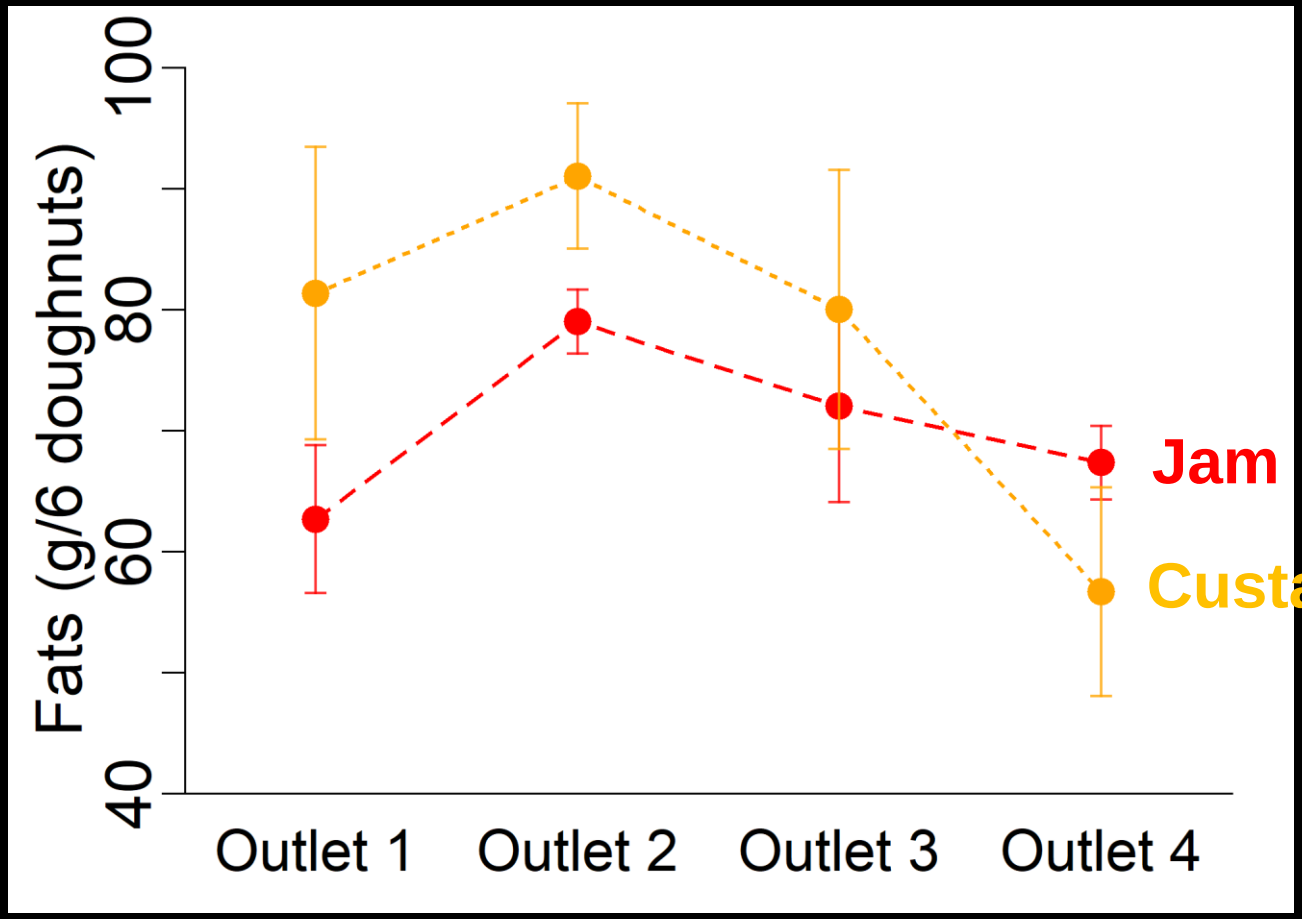
Interaction



Fats in doughnuts

Obs	Type	Outlet 1	Outlet 2	Outlet 3	Outlet 4
1	Jam	64	78	75	64
2	Jam	56	82	63	70
3	Jam	68	77	78	68
4	Custard	77	91	71	55
5	Custard	72	85	93	66
6	Custard	95	97	76	49

Fats in doughnuts



Analysis : two way ANOVA

- Fasts in doughnuts
- Shop, level = 4
- Type, level = 2
- Interaction between outlet and type

Analysis : two way ANOVA

```
> aov2way <- aov(Fat ~ Outlet2 * Type, data = data)
> summary(aov2way)
             Df Sum Sq Mean Sq F value  Pr(>F)
Outlet2       3 1636.5   545.5   8.619 0.00124 **
Type          1  294.0   294.0   4.645 0.04672 *
Outlet2:Type  3  711.3   237.1   3.746 0.03265 *
Residuals    16 1012.7    63.3
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Analysis : two way ANOVA

```
> aov2way <- aov(Fat ~ Outlet2 * Type, data = data)
> summary(aov2way)
             Df Sum Sq Mean Sq F value  Pr(>F)
Outlet2       3 1636.5   545.5   8.619 0.00124 **
Type          1  294.0   294.0   4.645 0.04672 *
Outlet2:Type  3  711.3   237.1   3.746 0.03265 *
Residuals    16 1012.7    63.3
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> aov2way <- aov(Fat ~ Outlet2 + Type, data = data)
> summary(aov2way)
             Df Sum Sq Mean Sq F value  Pr(>F)
Outlet2       3   1636   545.5   6.012 0.00466 **
Type          1    294   294.0   3.240 0.08775 .
Residuals    19   1724    90.7
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Post Hoc test for ANOVA

```
> aov2way <- aov(Fat ~ Outlet2 * Type, data = data)
> TukeyHSD(aov2way)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Fat ~ Outlet2 * Type, data = data)

$Outlet2
    diff         lwr        upr     p adj
2-1   13  -0.141154 26.141154 0.0530280
3-1    4  -9.141154 17.141154 0.8196711
4-1  -10 -23.141154  3.141154 0.1718718
3-2   -9 -22.141154  4.141154 0.2437391
4-2  -23 -36.141154 -9.858846 0.0006717
4-3  -14 -27.141154 -0.858846 0.0347989

$Type
             diff       lwr         upr     p adj
jam-custard    -7 -13.88516 -0.1148392 0.0467162
```

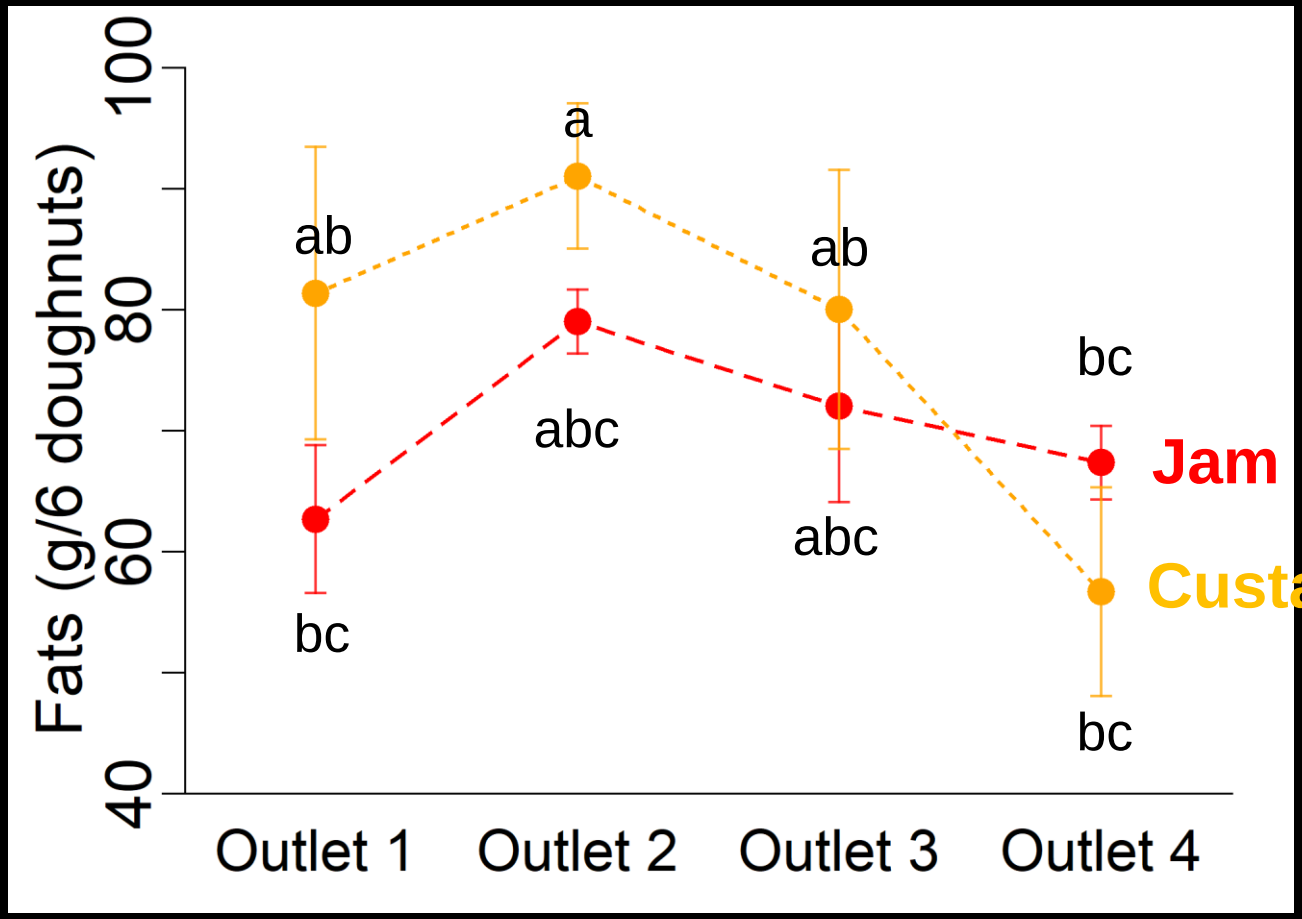
$`Outlet2:Type`

	diff	lwr	upr	p adj
2:custard-1:custard	9.666667	-12.8225125	32.1558458	0.8029236
3:custard-1:custard	-1.333333	-23.8225125	21.1558458	0.9999988
4:custard-1:custard	-24.666667	-47.1558458	-2.1774875	0.0264504
1:jam-1:custard	-18.666667	-41.1558458	3.8225125	0.1438285
2:jam-1:custard	-2.333333	-24.8225125	20.1558458	0.9999433
3:jam-1:custard	-9.333333	-31.8225125	13.1558458	0.8277981
4:jam-1:custard	-14.000000	-36.4891791	8.4891791	0.4239003
3:custard-2:custard	-11.000000	-33.4891791	11.4891791	0.6913308
4:custard-2:custard	-34.333333	-56.8225125	-11.8441542	0.0014775
1:jam-2:custard	-28.333333	-50.8225125	-5.8441542	0.0088415
2:jam-2:custard	-12.000000	-34.4891791	10.4891791	0.6006090
3:jam-2:custard	-19.000000	-41.4891791	3.4891791	0.1317609
4:jam-2:custard	-23.666667	-46.1558458	-1.1774875	0.0355076
4:custard-3:custard	-23.333333	-45.8225125	-0.8441542	0.0391418
1:jam-3:custard	-17.333333	-39.8225125	5.1558458	0.2018598
2:jam-3:custard	-1.000000	-23.4891791	21.4891791	0.9999998
3:jam-3:custard	-8.000000	-30.4891791	14.4891791	0.9104520
4:jam-3:custard	-12.666667	-35.1558458	9.8225125	0.5398511
1:jam-4:custard	6.000000	-16.4891791	28.4891791	0.9789576
2:jam-4:custard	22.333333	-0.1558458	44.8225125	0.0522941
3:jam-4:custard	15.333333	-7.1558458	37.8225125	0.3217512
4:jam-4:custard	10.666667	-11.8225125	33.1558458	0.7206509
2:jam-1:jam	16.333333	-6.1558458	38.8225125	0.2566627
3:jam-1:jam	9.333333	-13.1558458	31.8225125	0.8277981
4:jam-1:jam	4.666667	-17.8225125	27.1558458	0.9950507
3:jam-2:jam	-7.000000	-29.4891791	15.4891791	0.9527458
4:jam-2:jam	-11.666667	-34.1558458	10.8225125	0.6311015
4:jam-3:jam	-4.666667	-27.1558458	17.8225125	0.9950507

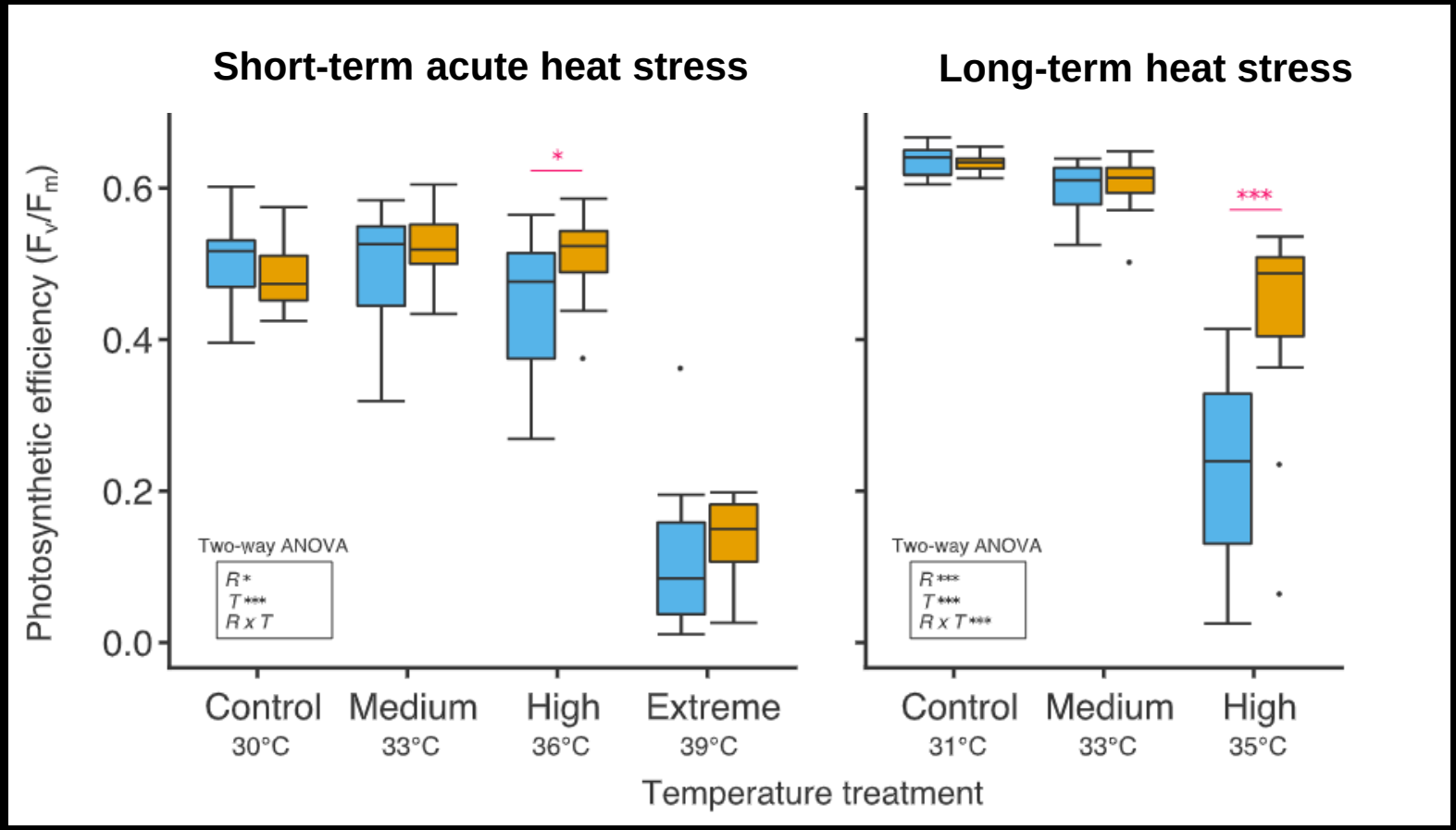
Post Hoc test for ANOVA

```
> sigLetter <- multcompView ::multcompLetters4(aov2way, tuk)
> print(sigLetter)
$Outlet2
   2    3    1    4
 "a"  "a" "ab"  "b"

$Type
custard     jam
    "a"     "b"

$`Outlet2:Type`
2:custard 1:custard 3:custard     2:jam     3:jam     4:jam     1:jam 4:custard
      "a"      "ab"      "ab"     "abc"     "abc"      "bc"      "bc"       "c"
```

Post Hoc test for ANOVA

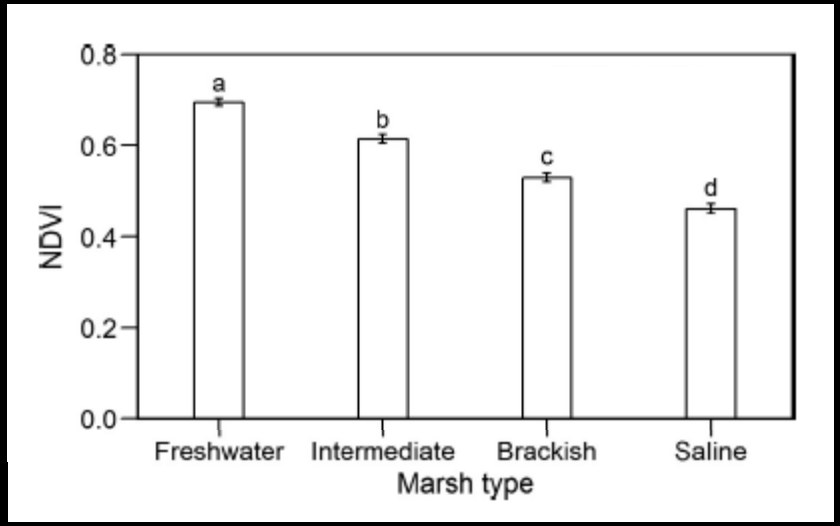


Heat stress on coral photosynthesis



Impact of drought on marsh phenology

Peak biomass



Length of growing season