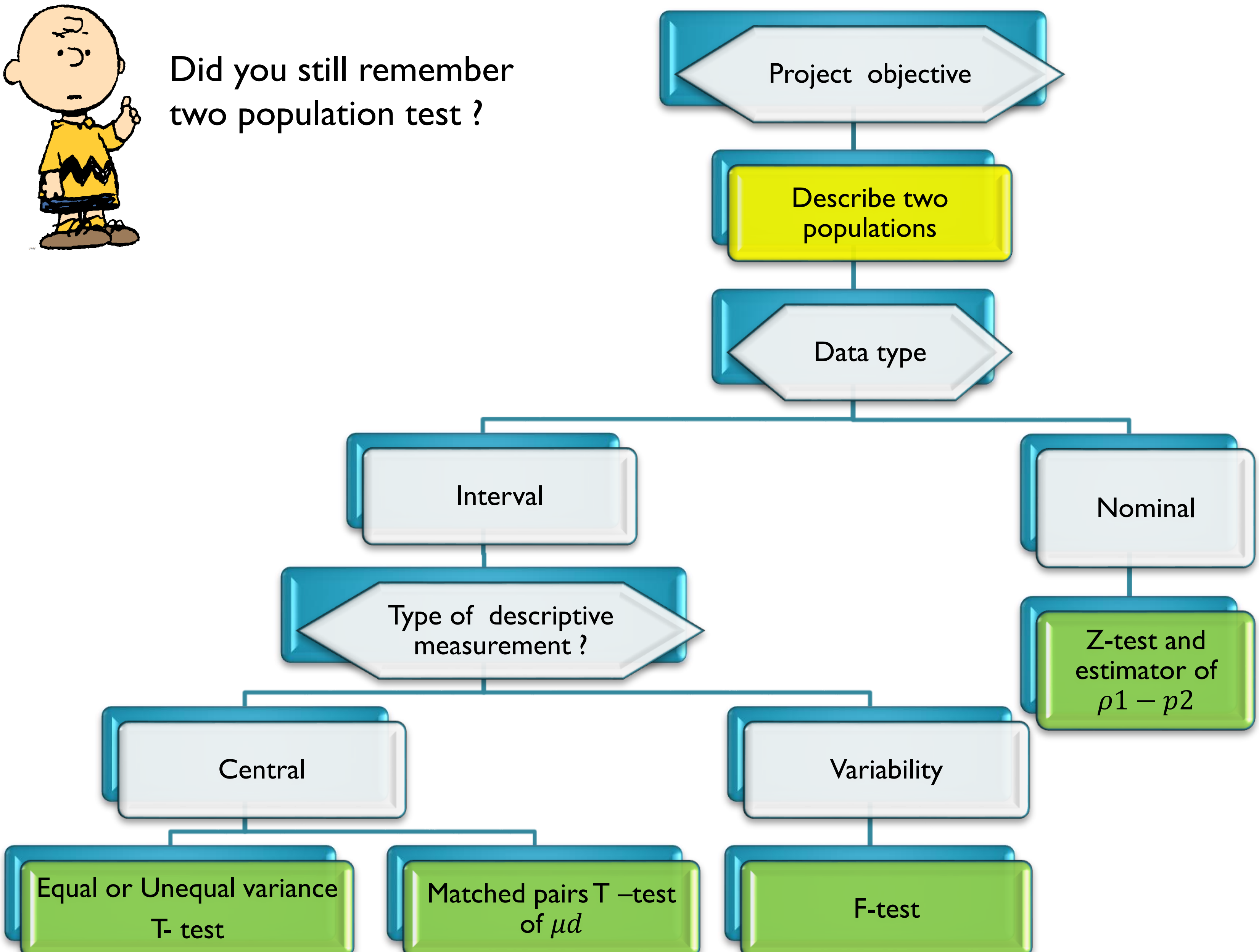
Welcome

Did you still remember two population test ?

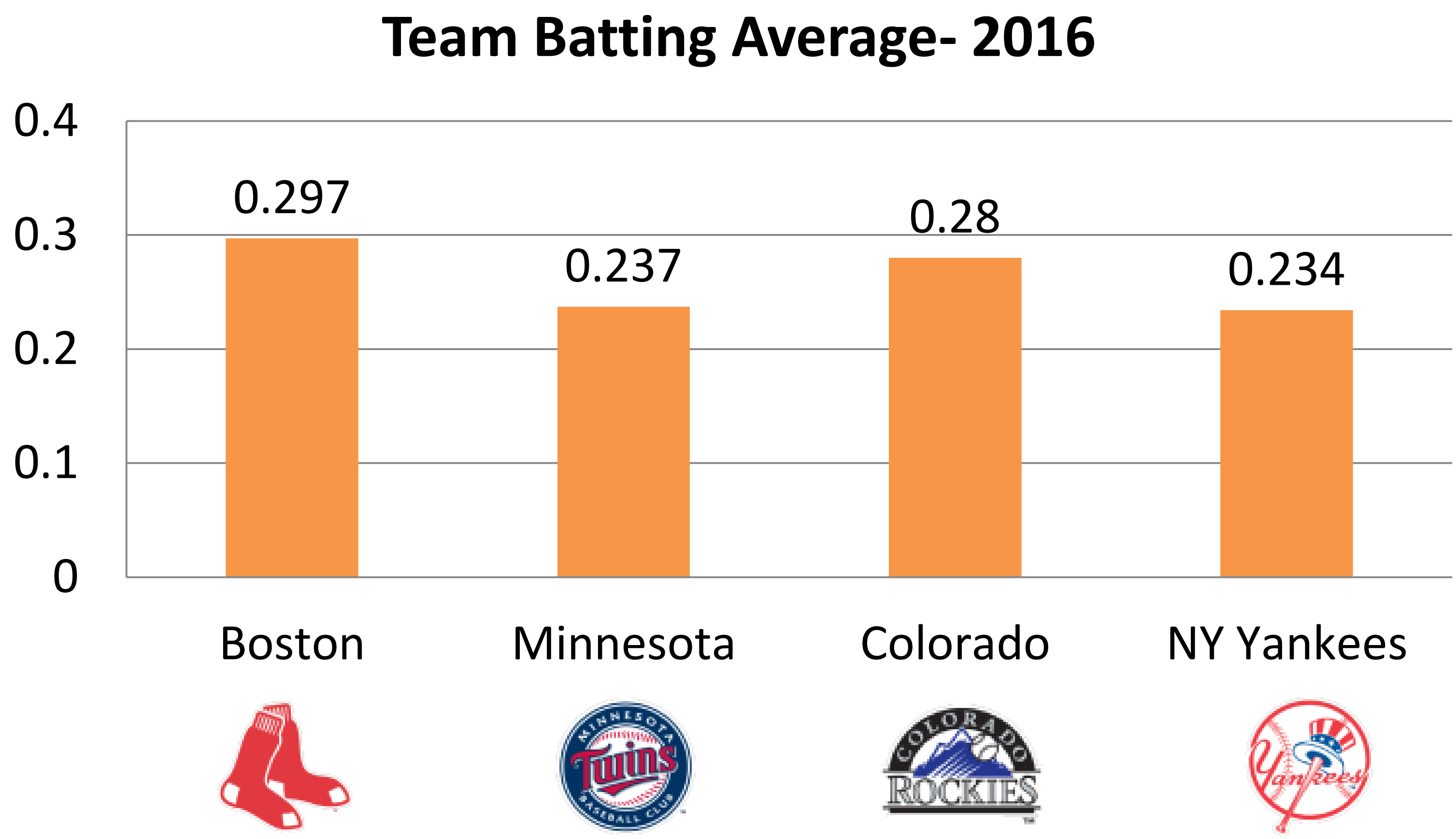


Topics

- One-Way ANOVA without repeated measures
  - One-way ANOVA
- One-Way ANOVA with repeated measures
  - Randomized complete block ANOVA

Batting average

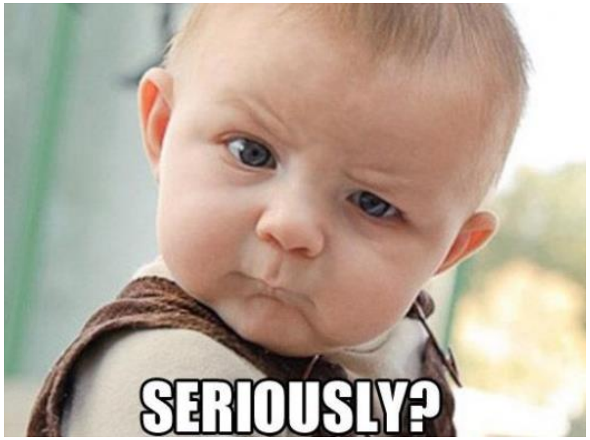
Batting average



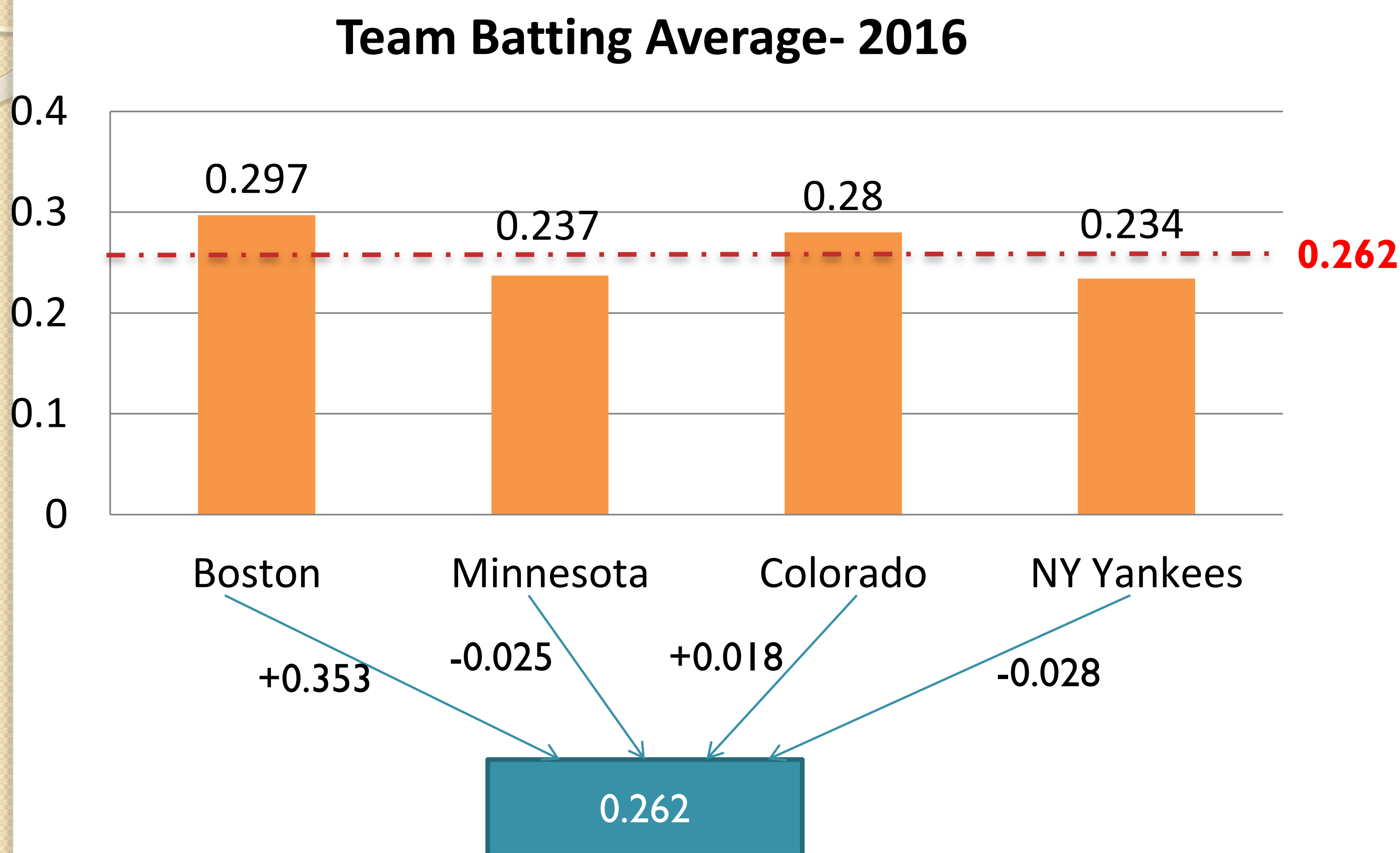
If you want to infer that the population means differ

What if

- We want to compare more than two populations
- And, a mix of interval and nominal variables



Batting average problem- Eyeball test?

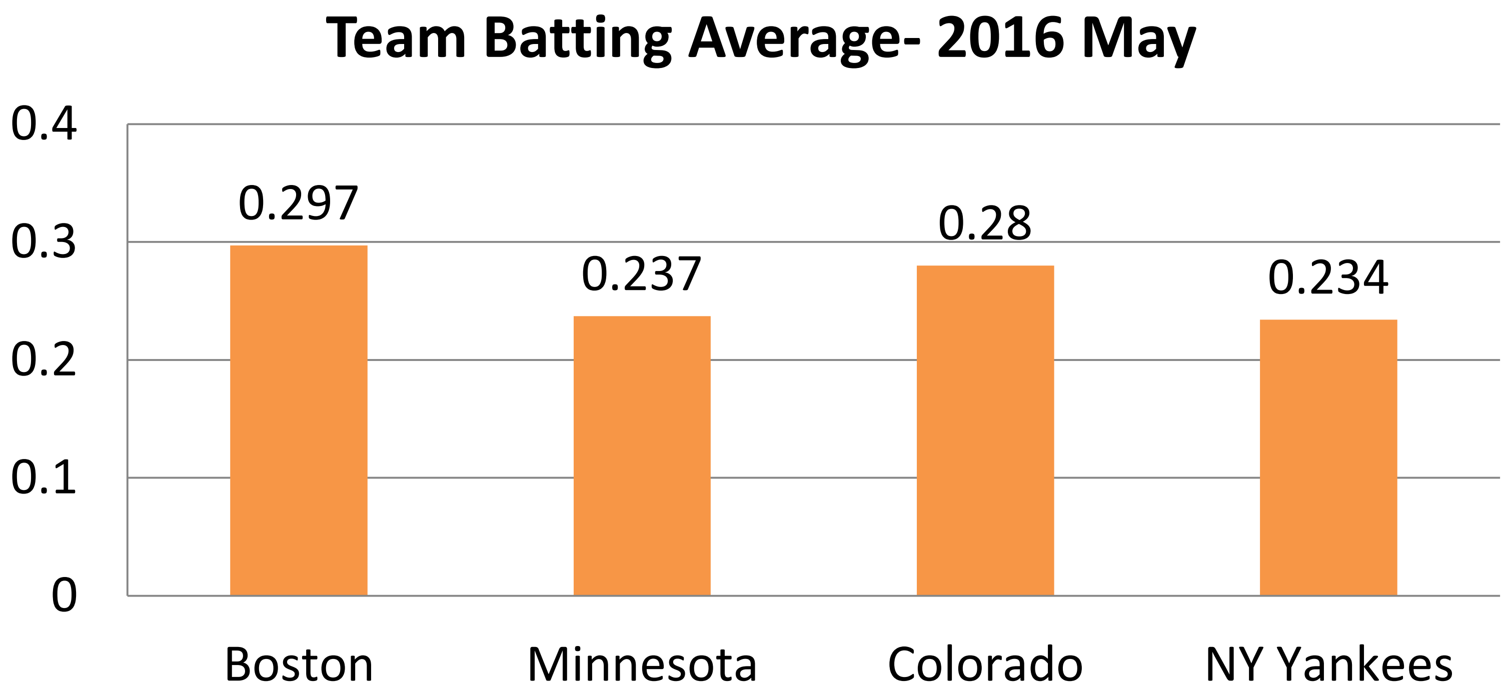


The grand mean of these four teams is 0.262

One-way ANOVA

- A mix of interval and nominal variables
- To test hypotheses about the mean on one variable for three or more groups
- The hypotheses:
  - Null hypothesis (H0): all teams have the same mean
  - Research hypothesis (H1): at least one team has a different mean

Batting average



If you want to infer that the population mean differ at 5 % significance level

H0: μ1 = μ2 = μ3 = μ4

H1: At least two means differ

The real data looks like

x_{ij} refers to the i^{th} observation in the j^{th} sample
Ex : $X_{11} = 0.284$; $X_{12} = 0.221$

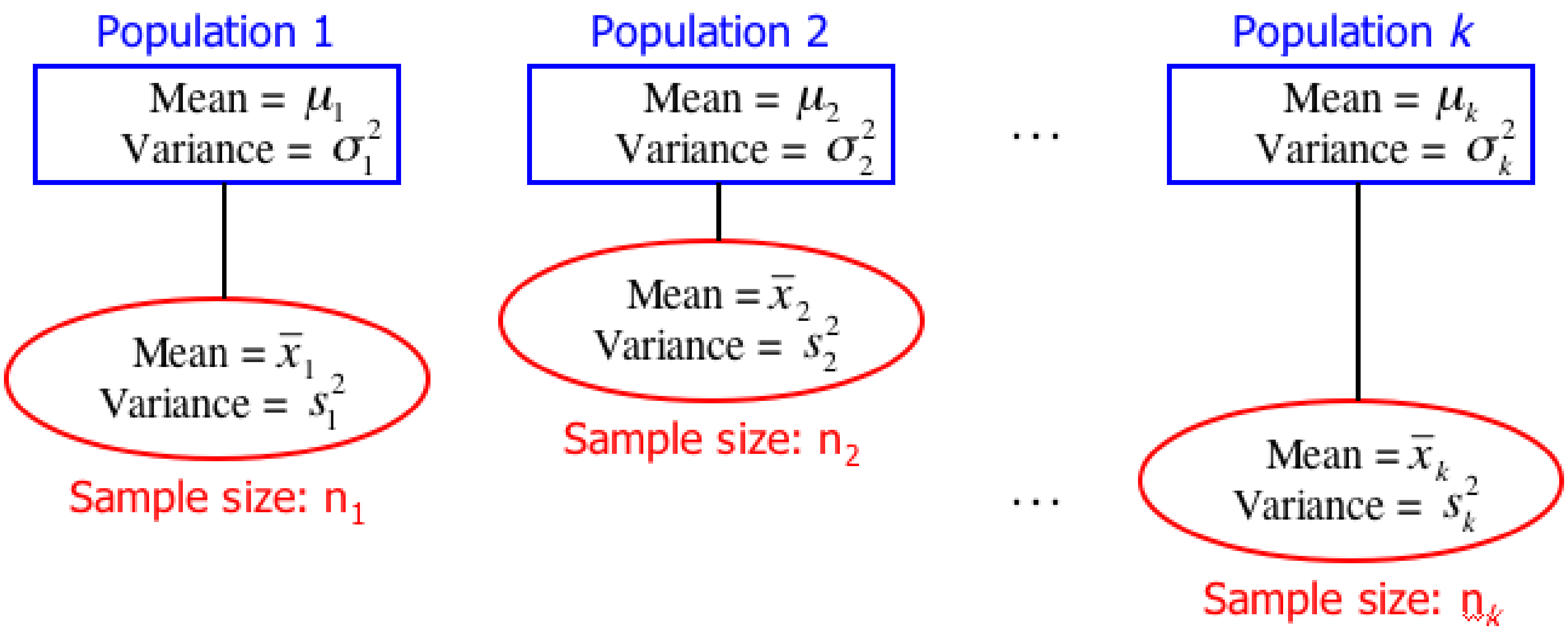
These populations are referred to as treatments

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

x is the response variable

Analysis of variance

- Analysis of variance (ANOVA) is a technique that allows us to compare more than two populations (4 baseball teams) of interval data (batting average).



ANOVA – Between treatments variation

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

The ***grand mean***, $\bar{\bar{x}}$, is the mean of all the observations, i.e.:

$$\bar{\bar{x}} = \frac{\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}}{n}$$

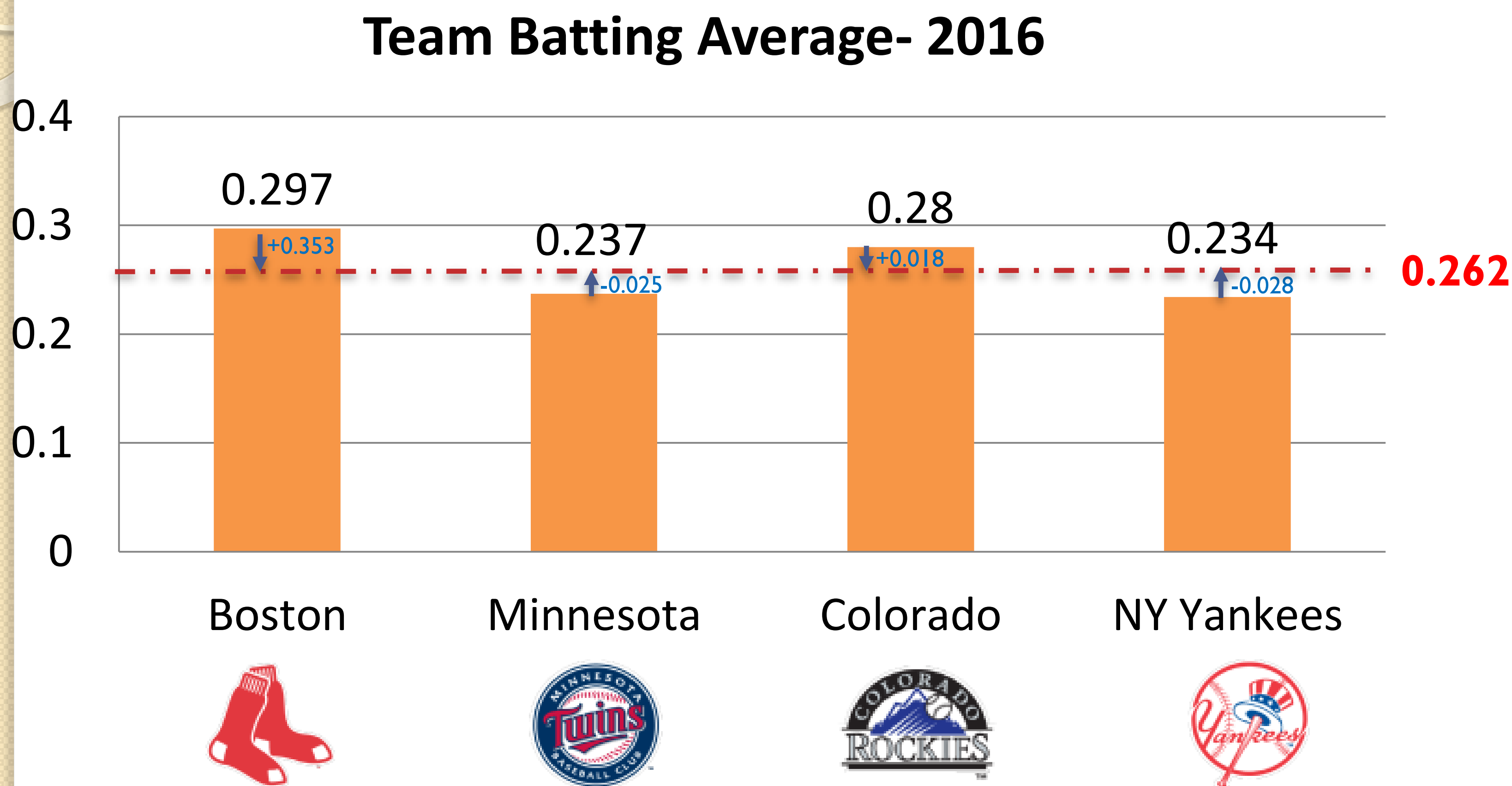
(n = $n_1 + n_2 + \ldots + n_k$)

Larger or small ?

0.262

What do you think if $\mu 1 = \mu 2 = \mu 3 = \mu 4$

Batting average problem- Eyeball test?



ANOVA – Between treatments

- Between-treatments variation (Sum of squares for treatments)

$$SST = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{\bar{x}})^2$$

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

0.262

ANOVA – Between treatments variation

- Between-treatments variation (Sum of squares for treatments)

$$SST = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{\bar{x}})^2$$

$$SST = 3(0.297 - 0.262)^2 + 3\,(0.237 - 0.262)^2 + 3(0.280 - 0262)^2 + 3(0.234 - 0.262)^2 = 0.0089$$

- The mean square for treatments (MST)

$$MST = \frac{SST}{k-1}$$

$$MST = \frac{0.0089}{3} = 0.00297$$

What do you think if

$$\mu 1 = \mu 2 = \mu 3 = \mu 4$$

ANOVA – Between treatments

What do you think if $\mu1 = \mu2 = \mu3 = \mu4$

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.281	0.280	0.287	0.289
Player 2	0.280	0.278	0.277	0.280
Player 3	0.279	0.282	0.276	0.271
BA	0.280	0.280	0.280	0.280

The ***grand mean***, $\bar{\bar{x}}$ is the mean of all the observations, i.e.:

$$\bar{\bar{x}} = \frac{\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}}{n}$$

$(n = n_1 + n_2 + \ldots + n_k)$

Larger or small ?

0.280

ANOVA – Between treatments

- Between-treatments variation (Sum of squares for treatments)

$$SST = \sum_{j=1}^{k} n_j (\bar{x}_j - \bar{\bar{x}})^2$$

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.281	0.280	0.287	0.289
Player 2	0.280	0.278	0.277	0.280
Player 3	0.279	0.282	0.276	0.271
BA	0.280	0.280	0.280	0.280

0.280

ANOVA – Within treatments variation

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

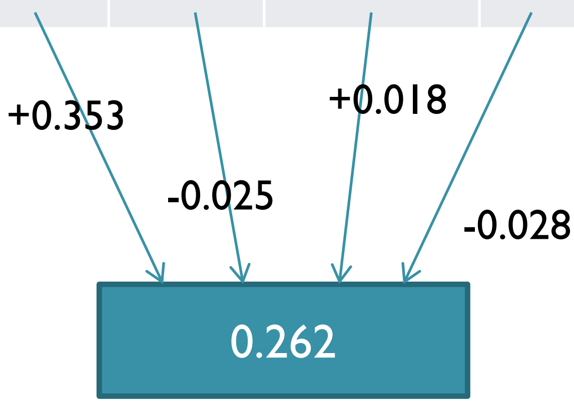
SST tells us the between-treatments variation

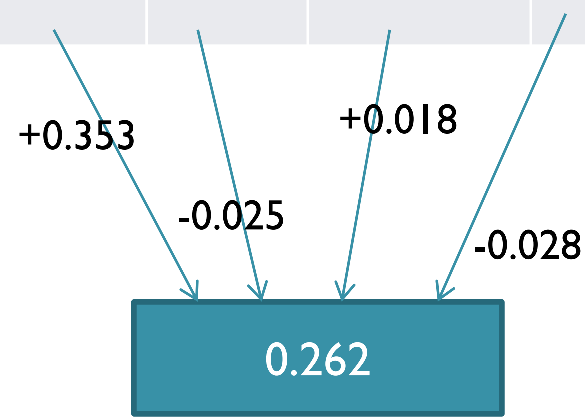
We need to know the within-treatments variation that is not caused by the treatments

ANOVA – Within treatments

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.335	0.221	0.287	0.112
Player 2	0.189	0.159	0.199	0.288
Player 3	0.367	0.331	0.354	0.302
BA	**0.297**	**0.237**	**0.280**	**0.234**





In which case, you will have more confidence to conclude at least two means differ?

ANOVA – Within treatments

- Within-treatments variation (Sum of Squares for Error)

$$SSE = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

ANOVA – Within treatments variation

- Within-treatments variation (Sum of Squares for Error)

$$SSE = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

$$\begin{aligned} SSE &= (0.284 - 0.297)^2 + (0.331 - 0.297)^2 + 3(0.276 - 0.297)^2 + (0.221 - 0.237)^2 \\ &+ (0.234 - 0.237)^2 + (0.256 - 0.237)^2 + (0.287 - 0.280)^2 + (0.277 - 0.280)^2 \\ &+ (0.276 - 0.280)^2 + (0.221 - 0.234)^2 + (0.256 - 0.234)^2 + (0.225 - 0.234)^2 \\ &= 0.0032 \end{aligned}$$

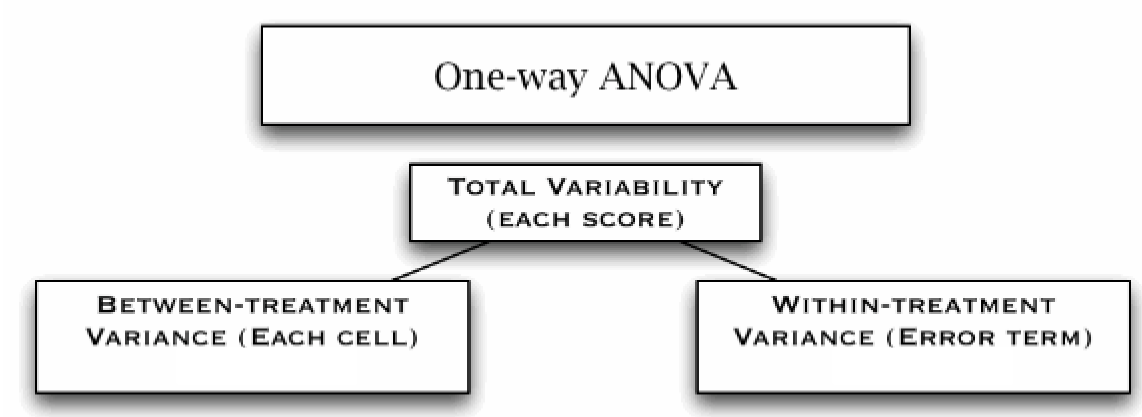
- The mean square for errors (MSE)

$$MSE = \frac{SSE}{n-k}$$

- $MSE = \frac{0.0032}{12-4} = 0.0004$

ANOVA

- We compare MST and MSE using an F-test and conclusions are drawn using the value of F.



$$F = \frac{MST}{MSE}$$

F-distributed with k–1 and n–k degrees of freedom.

Remember?

- When SST will equal to 0 ?
- If SST = 0
- Our decision : non-reject H0
  - H0: $\mu 1 = \mu 2 = \mu 3 = \mu 4$
  - H1: At least two means differ

$$SST = \sum_{j=1}^{k} n_j (\bar{x}_j - \bar{\bar{x}})^2$$

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.281	0.280	0.287	0.289
Player 2	0.280	0.278	0.277	0.280
Player 3	0.279	0.282	0.276	0.271
BA	0.280	0.280	0.280	0.280

Hypothesis testing

- Since:

$$SST = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{\bar{x}})^2$$

- If SST > 0
- Our decision : may or may not reject H0
  - H0: μ1 = μ2 = μ3 = μ4
  - H1: At least two means differ

- A large value of SST rejects the H0.
  - But, how large is "large enough" to reject the H0?

ANOVA table

Source of Variation	Sum of Squares	degrees of freedom	Mean Square	F
Between	SST	$k-1$	MST=SST/($k-1$)	MST/MSE
Within	SSE	$n-k$	MSE=SSE/($n-k$)	
Total	SS(Total)	$n-1$		

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Boston	3	0.891	0.297	0.000883
Minnesota	3	0.711	0.237	0.000313
Colorado	3	0.84	0.28	0.00004
NY Yankees	3	0.702	0.234	0.000367

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.008874	3	0.002958	7.395	0.010774	4.066181
Within Groups	0.0032	8	0.0004			
Total	0.012074	11				

R codes

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

- b<-c(0.284,0.331,0.276)
- m<-c(0.221,0.234,0.256)
- c<-c(0.287,0.277,0.276)
- ny<-c(0.221,0.256,0.225)
- mydata<-data.frame(b,m,c,ny)
- mydata2 <- stack(mydata)
- mydata2
- aov1<- aov(values ~ ind, data=mydata2)
- summary(aov1)

R codes

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

```
> mydata2
   values ind
1   0.284   b
2   0.331   b
3   0.276   b
4   0.221   m
5   0.234   m
6   0.256   m
7   0.287   c
8   0.277   c
9   0.276   c
10  0.221  ny
11  0.256  ny
12  0.225  ny
```

H0: μ1 = μ2 = μ3 = μ4

H1: At least two means differ

```
> aov1<- aov(values ~ ind, data=mydata2)
> summary(aov1)
            Df   Sum Sq  Mean Sq F value Pr(>F)
ind          3 0.008874 0.002958   7.395 0.0108 *
Residuals    8 0.003200 0.000400
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.335	0.221	0.287	0.112
Player 2	0.189	0.159	0.199	0.288
Player 3	0.367	0.331	0.354	0.302
BA	**0.297**	**0.237**	**0.280**	**0.234**

Anova: Single Factor

Consistent variance group

SUMMARY

Groups	Count	Sum	Average	Variance
Boston	3	0.891	0.297	0.000883
Minnesota	3	0.711	0.237	0.000313
Colorado	3	0.84	0.28	0.00004
NY Yankees	3	0.702	0.234	0.000367

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.008874	3	0.002958	7.395	0.010774	4.066181
Within Groups	0.0032	8	0.0004			
Total	0.012074	11				

Anova: Single Factor

Inconsistent variance group

SUMMARY

Groups	Count	Sum	Average	Variance
Boston	3	0.891	0.297	0.009004
Minnesota	3	0.711	0.237	0.007588
Colorado	3	0.84	0.28	0.006043
NY Yankees	3	0.702	0.234	0.011212

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.008874	3	0.002958	0.349573	0.790752	4.066181
Within Groups	0.067694	8	0.008462			
Total	0.076568	11				

Post hoc test – Multiple comparison

- ANOVA tells you whether you have an overall difference between your treatments (four teams), but it does not tell you which specific treatments differed .

- We have multiple comparison method

Post hoc test – Multiple comparison

- b<-c(0.284,0.331,0.276)
- m<-c(0.221,0.234,0.256)
- c<-c(0.287,0.277,0.276)
- ny<-c(0.221,0.256,0.225)
- mydata<-data.frame(b,m,c,ny)
- mydata2 <- stack(mydata)
- mydata2
- aov1<- aov(values ~ ind, data=mydata2)
- summary(aov1)
- pairwise.t.test(mydata2$values,mydata2$ind)

Post hoc test – Multiple comparison

- We are looking for a critical number to compare the absolute differences of the sample means against.

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

Post hoc test – Multiple comparison

	Boston	Minnesota	Colorado	NY Yankees
Player 1	0.284	0.221	0.287	0.221
Player 2	0.331	0.234	0.277	0.256
Player 3	0.276	0.256	0.276	0.225
BA	**0.297**	**0.237**	**0.280**	**0.234**

```
> pairwise.t.test(mydata2$values,mydata2$ind)

        Pairwise comparisons using t tests with pooled SD

data:  mydata2$values and mydata2$ind

   b     m     c
m  0.031 -     -
c  0.657 0.090 -
ny 0.029 0.859 0.090

P value adjustment method: holm
```

Problem

- How does an MBA major affect the number of job offers received? An MBA student randomly sampled four recent graduates, one each in finance, marketing, and management, and asked them to report the number of job offers.

- Can we conclude at the 5% significance level that there are differences in the number of job offers between the three MBA majors

Finance	Marketing	Management
3	1	8
1	5	5
4	3	4
1	4	6

R codes

- fin<-c(3,1,4,1)
- market<-c(1,5,3,4)
- manage<-c(8,5,4,6)
- mydata<-data.frame(fin,market,manage)
- mydata2 <- stack(mydata)
- mydata2
- aov1<- aov(values ~ ind, data=mydata2)
- summary(aov1)
- pairwise.t.test(mydata2$values,mydata2$ind)


Happy Hour

Exercise 1

- A research study was conducted to examine the clinical efficacy of a new antidepressant. Depressed patients were randomly assigned to one of three groups: a placebo group, a group that received a low dose of the drug, and a group that received a moderate dose of the drug.

- After four weeks of treatment, the patients completed the Depression Inventory Test. The higher the score, the **more** depressed the patient. The data are presented below. Compute the appropriate test by using 5% sig. level. What is your finding?

Placebo	Low Dose	Moderate Dose
38	22	14
47	19	26
39	8	11
25	23	18
42	31	5

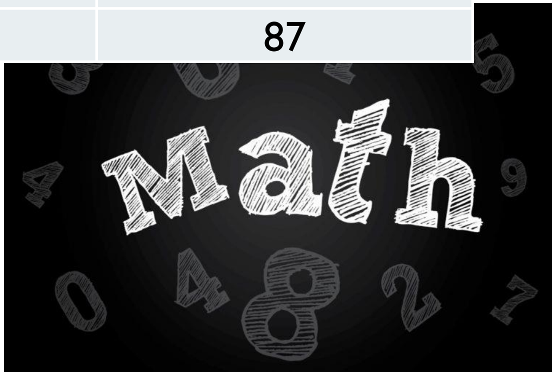
Exercise 2

- A researcher is concerned about the level of knowledge possessed by university students regarding math.

- Students completed a high school senior level standardized math exam. Major for students was also recorded. Data in terms of percent correct is recorded below for 12 students.

- Compute the appropriate test by using 5% sig. level. What is your finding?

Education	Business	Social Science	Fine Arts
62	72	42	80
81	49	52	57
75	63	31	87



Exercise 3

- Neuroscience researchers examined the impact of environment on rat development. Rats were randomly assigned to be raised in one of the four following test conditions: loser (wire mesh cage - housed alone), standard (cage with other rats), enriched (cage with other rats and toys), super enriched (cage with rats and toys changes on a periodic basis).

- After two months, the rats were tested on a variety of learning a maze. Their time to complete the maze task is below. Compute the appropriate test by using 5% sig. level. What is your finding?

Loser	Standard	Enriched	Super Enriched
22	17	12	8
19	21	14	7
15	15	11	10
24	12	9	9
18	19	15	12

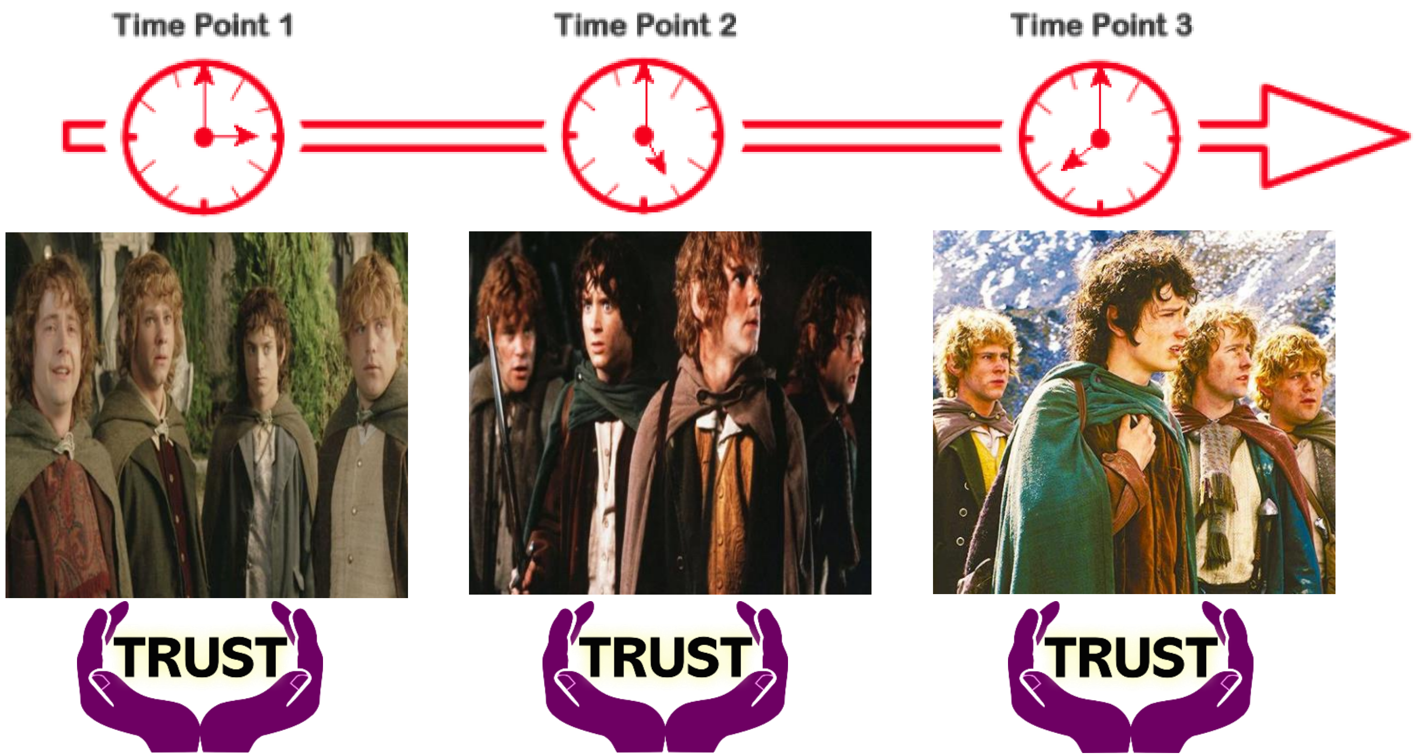
Topics

- One-Way ANOVA without repeated measures
  - One-way ANOVA
- One-Way ANOVA with repeated measures
  - Randomized complete block ANOVA

Repeated measures ANOVA

- Repeated measures ANOVA is the equivalent of the one-way ANOVA, but for related, not independent groups
- The nature of the repeated measures ANOVA, that of a test to detect any overall differences between related means
- Studies that investigate either
    - changes in mean scores over three or more time points
    - differences in mean scores under three or more different conditions

Repeated measures ANOVA



In repeated measures ANOVA, the independent variable has categories called **related groups**.

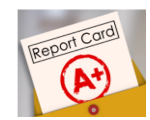
Where measurements are repeated over time in this case, such as when measuring changes in team trust due to an adventure journey, the independent variable is **time**.

Repeated measures ANOVA

Condition1

Condition2

Condition3



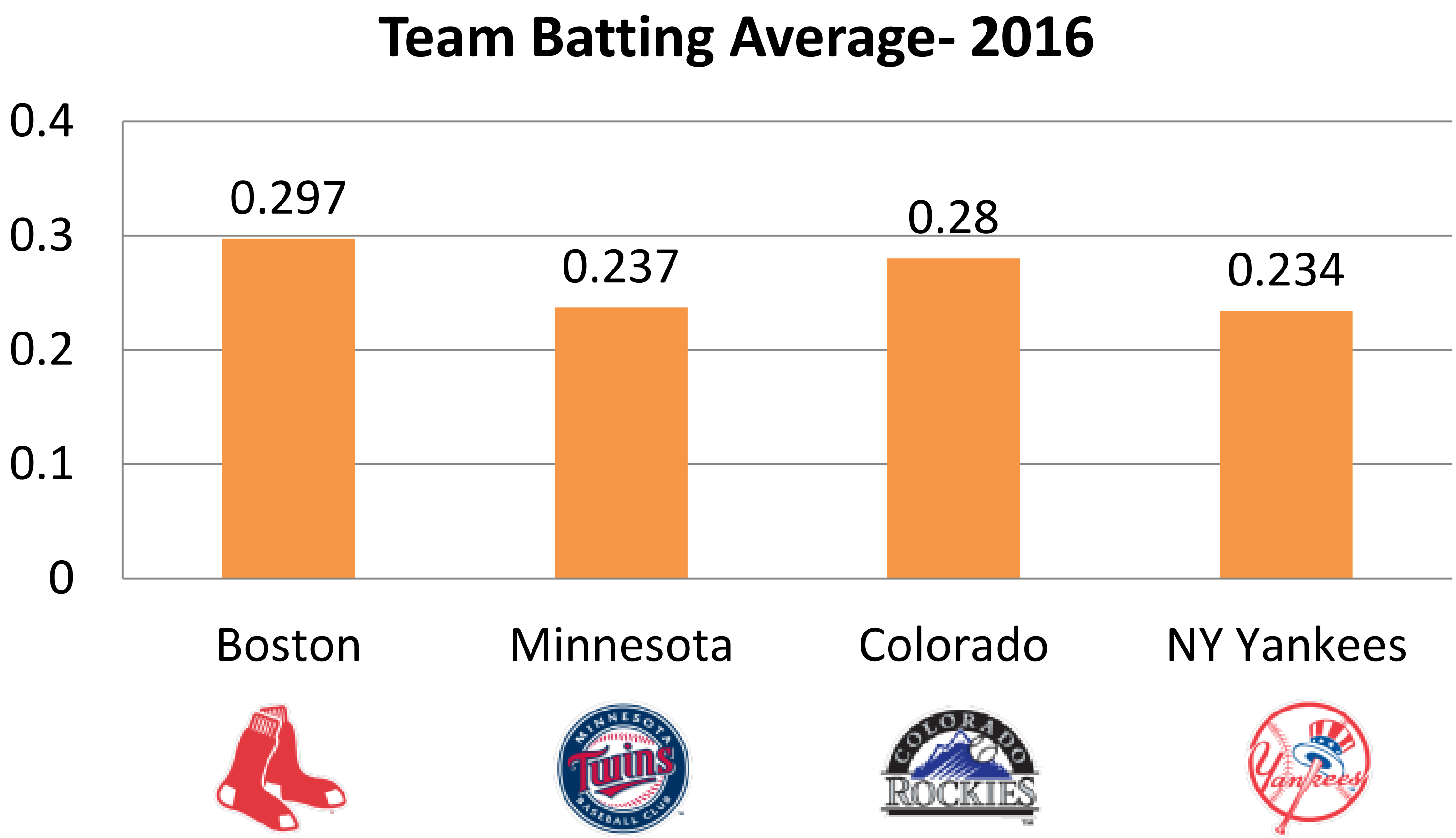
In repeated measures ANOVA, the independent variable has categories called **related groups**.

Where measurements are repeated under different conditions, such as when measuring changes in performance due to class room, the independent variable is **conditions**.

Repeated measures ANOVA

Batting average

Batting average



If you want to infer that the population means differ

This time, the real data looks like

	Boston	Minnesota	Colorado	NY Yankees
Player A	Player A 0.284	Player A 0.221	Player A 0.287	Player A 0.221
Player B	Player B 0.331	Player B 0.234	Player B 0.277	Player B 0.256
Player C	Player C 0.276	Player C 0.256	Player C 0.276	Player C 0.225

Repeated measures ANOVA

	Boston	Minnesota	Colorado	NY Yankees	Subjects
Player A	Player A 0.284	Player A 0.221	Player A 0.287	Player A 0.221	
Player B	Player B 0.331	Player B 0.234	Player B 0.277	Player B 0.256	
Player C	Player C 0.276	Player C 0.256	Player C 0.276	Player C 0.225	
BA	**0.297**	**0.237**	**0.280**	**0.234**	0.262

$$SST = 3(0.297-0.262)^2 + 3\,(0.237-0.262)^2 + 3(0.280-0262)^2 + 3(0.234-0.262)^2 = 0.0089$$

$$SST = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{\bar{x}})^2$$

Repeated measures ANOVA

	Boston	Minnesota	Colorado	NY Yankees	Subjects
Player A	Player A 0.284	Player A 0.221	Player A 0.287	Player A 0.221	
Player B	Player B 0.331	Player B 0.234	Player B 0.277	Player B 0.256	
Player C	Player C 0.276	Player C 0.256	Player C 0.276	Player C 0.225	
BA	**0.297**	**0.237**	**0.280**	**0.234**	

$$SSE = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

$$\begin{aligned} SSE &= (0.284-0.297)^2 + (0.331-0.297)^2 + (0.276-0.297)^2 \\ &+ (0.221-0.237)^2 + (0.234-0.237)^2 + (0.256-0.237)^2 \\ &+ (0.287-0.280)^2 + (0.277-0.280)^2 + (0.276-0.280)^2 \\ &+ (0.221-0.234)^2 + (0.256-0.234)^2 + (0.225-0.234)^2 = 0.0032 \end{aligned}$$

Repeated measures ANOVA

	Boston	Minnesota	Colorado	NY Yankees	Subjects
Player A	Player A 0.284	Player A 0.221	Player A 0.287	Player A 0.221	**0.253**
Player B	Player B 0.331	Player B 0.234	Player B 0.277	Player B 0.256	**0.275**
Player C	Player C 0.276	Player C 0.256	Player C 0.276	Player C 0.225	**0.258**
BA	**0.297**	**0.237**	**0.280**	**0.234**	**0.262**

Player B is excellent

$$SS_{subjects} = k \cdot \sum (\bar{x}_i - \bar{x})^2$$

$$SSB = 4(0.253 - 0.262)^2 + 4(0.275 - 0.262)^2 + 4(0.258 - 0.262)^2 = 0.001$$

	Boston	Minnesota	Colorado	NY Yankees	Subjects
Player A	Player A 0.284	Player A 0.221	Player A 0.287	Player A 0.221	
Player B	Player B 0.331	Player B 0.234	Player B 0.277	Player B 0.256	
Player C	Player C 0.276	Player C 0.256	Player C 0.276	Player C 0.225	
BA	0.297	0.237	0.280	0.234	

$$SSE = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

Real within-treatments variation

= SSE – $SS_{subject}$ = 0.0032-0.001 = 0.0022

	Boston	Minnesota	Colorado	NY Yankees	Subjects
Player A	Player A 0.284	Player A 0.221	Player A 0.287	Player A 0.221	0.253
Player B	Player B 0.331	Player B 0.234	Player B 0.277	Player B 0.256	0.275
Player C	Player C 0.276	Player C 0.256	Player C 0.276	Player C 0.225	0.258
BA	0.297	0.237	0.280	0.234	0.262

$$SS_{subjects} = k \cdot \sum (\bar{x}_i - \bar{x})^2$$

R codes

- b<-c(0.284,0.331,0.276)
- m<-c(0.221,0.234,0.256)
- c<-c(0.287,0.277,0.276)
- ny<-c(0.221,0.256,0.225)
- mydata<-data.frame(b,m,c,ny)
- mydata2<- stack(mydata)
- id<-factor(rep(1:3, times=4))
- mydata2<-data.frame(mydata2,id)
- mydata2
- aov2<- aov(values ~ ind + Error(id/ind), data=mydata2)
- summary(aov2)
- pairwise.t.test(mydata2$values,mydata2$ind)

R codes

```
> b<-c(0.284,0.331,0.276)
> m<-c(0.221,0.234,0.256)
> c<-c(0.287,0.277,0.276)
> ny<-c(0.221,0.256,0.225)
> mydata<-data.frame(b,m,c,ny)
> mydata2<- stack(mydata)
> id<-factor(rep(1:3, times=4))
> mydata2<-data.frame(mydata2,id)
> mydata2
   values ind id
1   0.284   b  1
2   0.331   b  2
3   0.276   b  3
4   0.221   m  1
5   0.234   m  2
6   0.256   m  3
7   0.287   c  1
8   0.277   c  2
9   0.276   c  3
10  0.221  ny  1
11  0.256  ny  2
12  0.225  ny  3
```

```
> aov2<- aov(values ~ ind + Error(id/ind), data=mydata2)
> summary(aov2)

Error: id
          Df    Sum Sq   Mean Sq F value Pr(>F)
Residuals  2 0.0009875 0.0004937

Error: id:ind
          Df   Sum Sq   Mean Sq F value Pr(>F)
ind        3 0.008874 0.0029580   8.022  0.016 *
Residuals  6 0.002212 0.0003687
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R codes

```
> pairwise.t.test(mydata2$values,mydata2$ind)

        Pairwise comparisons using t tests with pooled SD

data:  mydata2$values and mydata2$ind

   b     m     c
m  0.031 -     -
c  0.657 0.090 -
ny 0.029 0.859 0.090
```

What is the difference ?

One-way ANOVA without Repeated measures

```
> aov1<- aov(values ~ ind, data=mydata2)
> summary(aov1)
            Df   Sum Sq  Mean Sq F value Pr(>F)
ind          3 0.008874 0.002958   7.395 0.0108 *
Residuals    8 0.003200 0.000400
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

One-way ANOVA with Repeated measures

```
> aov2<- aov(values ~ ind + Error(id/ind), data=mydata2)
> summary(aov2)

Error: id
          Df    Sum Sq   Mean Sq F value Pr(>F)
Residuals  2 0.0009875 0.0004937

Error: id:ind
          Df   Sum Sq   Mean Sq F value Pr(>F)
ind        3 0.008874 0.0029580   8.022  0.016 *
Residuals  6 0.002212 0.0003687
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Where are we going ?

Populations and Samples

Continuous probability

Business decisions with sampling

Decisions about variables

- ANOVA
- Regression