



# conrad

School of Entrepreneurship and Business

# Agenda

01

Recap

02

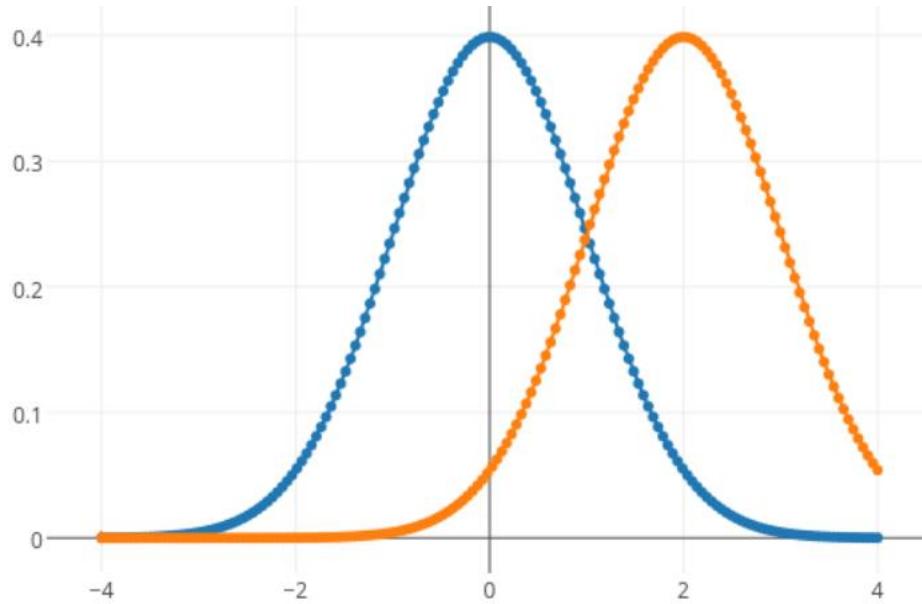
Comparing  
more than 2  
populations

03

Correlation  
Analysis

# Statistical Inference About Difference in Two Populations

# Statistical Inference About Means of Two Populations



**Inferences About the Difference Between  
Two Population Means:  $\sigma_1$  and  $\sigma_2$  Known**

**Inferences About the Difference Between  
Two Population Means:  $\sigma_1$  and  $\sigma_2$   
Unknown**

# Hypothesis Tests About $\mu_1 - \mu_2$ : $\sigma_1$ and $\sigma_2$ Known

## Hypotheses

$$H_0: \mu_1 - \mu_2 \geq D_0$$

$$H_a: \mu_1 - \mu_2 < D_0$$

Left-tailed

$$H_0: \mu_1 - \mu_2 \leq D_0$$

$$H_a: \mu_1 - \mu_2 > D_0$$

Right-tailed

$$H_0: \mu_1 - \mu_2 = D_0$$

$$H_a: \mu_1 - \mu_2 \neq D_0$$

Two-tailed

## Test Statistic

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

# Hypothesis Tests About $\mu_F - \mu_M$ $\sigma_F$ and $\sigma_M$ Unknown

## Hypotheses

$$H_0: \mu_1 - \mu_2 \geq D_0$$

$$H_a: \mu_1 - \mu_2 < D_0$$

$$H_0: \mu_1 - \mu_2 \leq D_0$$

$$H_a: \mu_1 - \mu_2 > D_0$$

$$H_0: \mu_1 - \mu_2 = D_0$$

$$H_a: \mu_1 - \mu_2 \neq D_0$$

Left-tailed

Right-tailed

Two-tailed

## Test Statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

# Statistical Inference About Proportions of Two Populations

# Hypothesis Tests about $p_1 - p_2$

## Hypotheses

$$H_0: p_1 - p_2 \geq 0$$

$$H_a: p_1 - p_2 < 0$$

Left-tailed

$$H_0: p_1 - p_2 \leq 0$$

$$H_a: p_1 - p_2 > 0$$

Right-tailed

$$H_0: p_1 - p_2 = 0$$

$$H_a: p_1 - p_2 \neq 0$$

Two-tailed

# Hypothesis Tests about $p_1 - p_2$

## Test Statistic

$$z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where:

$$\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2}$$



# Group Work

- Since we have 79 registered students, 20 groups will be created with students being randomly assigned
- Please submit your work no later than 10:00pm EST. There is no drop box, work needs to be submitted via email to [harvir.bansal@uwaterloo.ca](mailto:harvir.bansal@uwaterloo.ca)
- Please ensure that you clearly outline the steps taken in your analysis as well as the results.
- As I have said multiple times, data analysis is not a spectator sport so PLAY and have fun



**7.5/7.5**  
For all those who submitted the  
in-class exercise

# Hands-on Analysis

- All questions below refer to the brand tracking questionnaire and data made available for this class on LEARN
  1. Are there any stores where males and females different in their commitment to the brand? Explain via analysis
  2. Are Kohl's customers equally likely to purchase a product within the next 9 months as compared to TJ Maxx's customers?
  3. As compared to Amazon's NPS score which retailer must work hardest to compete?

# Hands-on Analysis

- All questions below refer to the brand tracking questionnaire and data made available for this class on LEARN
- 1. Are there any stores where males and females different in their commitment to the brand? Explain via analysis

S1 Are you...	GO TO
Male	S2
Female	

**SPSS®**

**SPSS** Analyze | Compare Means | Independent-Samples T Test

Brand Commitment						
Q11 Imagine you had to shop at a retail store, which of these statements best describes how much you would consider shopping at each of these stores? (Please select one answer for each brand)						
1 PROGRAMMER: ALLOW ONLY ONE ANSWER PER BRAND.						
ACCEPT ONLY ONE ANSWER IN FIRST ROW "ONLY STORE WOULD CONSIDER"						
	Brand X	JC Penney	Kohl's	Nordstrom	Amazon	TJ Maxx
Favorite store; only one I consider	<input type="radio"/>					
Store I prefer and consider highly	<input type="radio"/>					
Store I consider equally with others	<input type="radio"/>					
Store I might consider, less so than others	<input type="radio"/>					
Not a store I usually consider	<input type="radio"/>					
Store I would never consider	<input type="radio"/>					

# SPSS® Analyze | Compare Means| Independent-Samples T Test

Independent-Samples T Test

X

Test Variable(s):

- respid
- status
- interview\_start
- interview\_end
- htotaltimeinminutes
- userid
- subsid
- os2
- dregion
- dregion1
- s3
- happ

Grouping Variable:

s1(2 1)

Estimate effect sizes

OK Paste Reset Cancel Help

Options...  
Bootstrap...

## Group Statistics

	Are you...	N	Mean	Std. Deviation	Std. Error Mean
Brand X	Female	3044	3.06	1.146	.021
	Male	1287	2.98	1.140	.032
JC Penney	Female	77	2.99	1.006	.115
	Male	21	2.67	.966	.211
Kohl's	Female	1117	2.97	1.183	.035
	Male	493	2.86	1.065	.048
Nordstrom	Female	1165	3.04	1.192	.035
	Male	425	3.06	1.148	.056
Amazon	Female	1140	2.26	.822	.024
	Male	468	2.11	.789	.036
TJ Maxx	Female	1233	3.15	1.129	.032
	Male	370	3.29	1.083	.056

# Significant Differences

## Independent Samples Test

		Levene's Test for Equality of Variances				t-test for Equality of Means				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Brand X	Equal variances assumed	3.718	.054	2.022	4329	.022	.043	.077	.038	.002	.152
	Equal variances not assumed			2.027	2432.065	.021	.043	.077	.038	.002	.151
JC Penney	Equal variances assumed	.622	.432	1.304	96	.098	.195	.320	.246	-.167	.808
	Equal variances not assumed			1.335	32.835	.096	.191	.320	.240	-.168	.809
Kohl's	Equal variances assumed	3.356	.067	1.742	1608	.041	.082	.108	.062	-.014	.230
	Equal variances not assumed			1.814	1038.285	.035	.070	.108	.060	-.009	.225
Nordstrom	Equal variances assumed	.896	.344	-.215	1588	.415	.829	-.014	.067	-.146	.117
	Equal variances not assumed			-.219	779.528	.413	.827	-.014	.066	-.143	.115
Amazon	Equal variances assumed	9.608	.002	3.203	1606	<.001	.001	.143	.045	.055	.230
	Equal variances not assumed			3.258	901.812	<.001	.001	.143	.044	.057	.229
TJ Maxx	Equal variances assumed	.596	.440	-2.164	1601	.015	.031	-.143	.066	-.273	-.013
	Equal variances not assumed			-2.214	628.995	.014	.027	-.143	.065	-.271	-.016

# Hands-on Analysis

- All questions below refer to the brand tracking questionnaire and data made available for this class on LEARN
  1. Are there any stores where males and females different in their commitment to the brand? Explain via analysis
  2. Are Kohl's customers equally likely to purchase a product within the next 9 months as compared to TJ Maxx's customers?

SPSS®



Likelihood to Purchase						
Q12 For the following retailers, when do you anticipate your next purchase at that store might be? PROGRAMMER: ALLOW ONLY ONE ANSWER PER BRAND.						
	Brand X	JC Penney	Kohl's	Nordstrom	Amazon	TJ Maxx
Within the next month	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Within <u>next</u> 1-3 months	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
With next 3-6 months						
Within next 6-9 months	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Within 9 months – 1 year	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Longer than 1 year	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I'm not sure when I will shop there again						

# SPSS® Transform | Recode Into Different Variables

Recode into Different Variables

Numeric Variable -> Output Variable:

- q12x1 --> BrandX9Months
- q12x2 --> JCP9Months
- q12x11 --> Kohlls9Months
- q12x3 --> Nordstrom9Months
- q12x23 --> Amazon9Months
- q12x15 --> TJM9Months

Output Variable

Name: TJM9Months

Label:

Change

Old and New Values...

If... (optional case selection condition)

OK Paste Reset Cancel Help

Old Value

Value:

System-missing

System- or user-missing

Range:   
through

Range, LOWEST through value:

Range, value through HIGHEST:

All other values

New Value

Value:

System-missing

Copy old value(s)

Old --> New:

1 thru 4 --> 1  
5 thru 7 --> 2

Add Change Remove

Output variables are strings Width: 8

Convert numeric strings to numbers ('5'->5)

Continue Cancel Help

# SPSS® Analyze | Descriptive Statistics | Frequencies



		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Purchase Within 9 Months	1241	28.7	77.1	77.1
Missing	System	2721	62.8		
Total		4331	100.0		



		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Purchase Within 9 Months	1141	26.3	71.2	71.2
Missing	System	2728	63.0		
Total		4331	100.0		

## Z Score Calculator for 2 Population Proportions

Success!

You'll find the values for  $z$  and  $p$  below. Blue means your result is significant, red means it's not.

Sample 1 Proportion (or total number)

Sample 1 Size ( $N_1$ )

Sample 2 Proportion (or total number)

Sample 2 Size ( $N_2$ )

Significance Level:

0.01

0.05

0.10

One-tailed or two-tailed hypothesis?:

One-tailed

Two-tailed

The value of  $z$  is 3.8198. The value of  $p$  is .00014. The result is significant at  $p < .05$ .

[Calculate Z](#)

[Reset](#)

Social Science Statistics

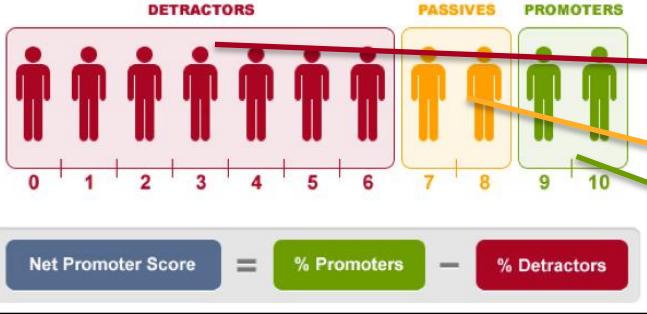
# Hands-on Analysis

- All questions below refer to the brand tracking questionnaire and data made available for this class on LEARN

1. Are there any stores where males and females different in their commitment to the brand? Explain via analysis
2. Are Kohl's customers equally likely to purchase a product within the next 9 months as compared to TJ Maxx's customers?
3. As compared to Amazon's NPS score which retailer must work hardest to compete?

**SPSS®**

**SPSS® Transform | Recode Into Different Variables**



Likelihood Recommend						
Q13 How likely are you to recommend the following retailer to friends and family members, on a scale from 0-10, where 0=Not at all likely to recommend, and 10= Will definitely recommend.						
PROGRAMMER: ALLOW ONLY ONE ANSWER PER BRAND. SHOW EACH RETAILER ON SEPARATE SCREENS.						
	Brand X	JC Penney	Kohl's	Nordstrom	Amazon	TJ Maxx
0	<input type="radio"/>					
1	<input type="radio"/>					
2	<input type="radio"/>					
3	<input type="radio"/>					
4	<input type="radio"/>					
5	<input type="radio"/>					
6	<input type="radio"/>					
7	<input type="radio"/>					
8	<input type="radio"/>					
9	<input type="radio"/>					
10	<input type="radio"/>					

Numeric Variable -> Output Variable:

- q12x11
- q12x3
- q12x23
- q12x15
- q19a\_24
- q19anew\_1
- q19anew\_2
- q19anew\_3
- q19anew\_4
- q19anew\_5
- q19anew\_6
- q19anew\_7
- q19anew\_8
- q19anew\_9

**Old and New Values...**

**If... (optional case selection condition)**

OK | Paste | Reset | Cancel | Help

Output Variable

Name: TJMaxxNPS

Label:

Change

Recode into Different Variables: Old and New Values

Old Value

Value:

System-missing

System- or user-missing

Range:  
  
 through

Range, LOWEST through value:

Range, value through HIGHEST:

All other values

New Value

Value:

System-missing

Copy old value(s)

Old --> New:

- 0 thru 6 --> 1
- 7 thru 8 --> 2
- 9 thru 10 --> 3

Add | Change | Remove

Output variables are strings | Width: 8

Convert numeric strings to numbers ('5'->5)

Continue | Cancel | Help

Variable(s):

- FashionForwardAmazon
- NPSAmazon
- BrandX9Months
- JCP9Months
- Kohls9Months
- Nordstrom9Months
- Amazon9Months
- TJM9Months
- BrandXNPSP
- JCPNPS
- KohlsNPS
- NordstromNPS
- AmazonNPS
- TJMaxxNPS

Statistics... | Charts... | Format... | Style... | Bootstrap...

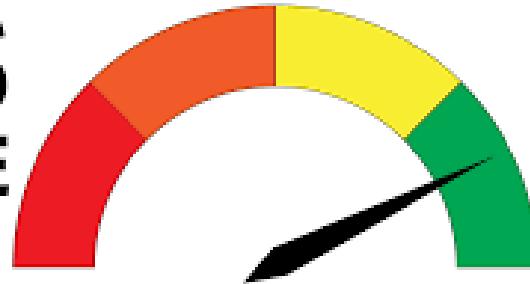
Display frequency tables  Create APA style tables

OK | Paste | Reset | Cancel | Help





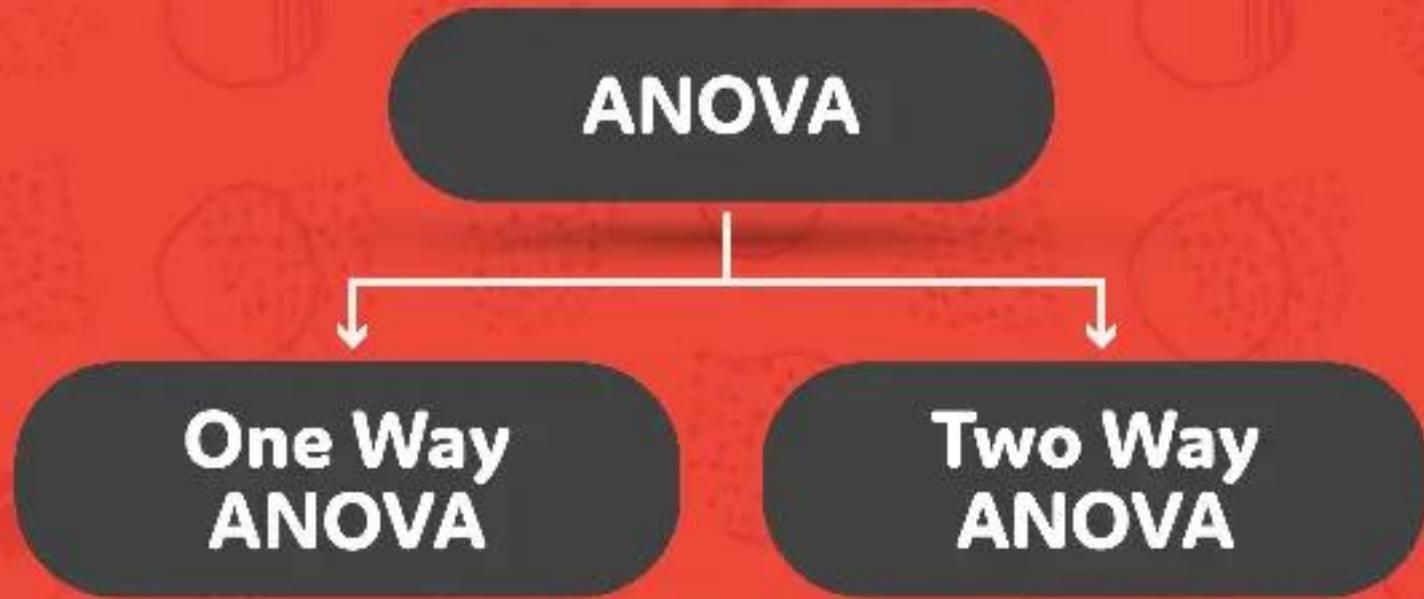
# NPS SCORE



Brand	Promoters	Detractors	NPS
amazon.com®	65.7	8.8	56.9
JCPenney	29.6	36.7	-7.1
KOHL'S®	32.1	32.9	-0.8
NORDSTROM	31.9	34	-2.1
T.J.maxx®	29.8	38.6	-8.8
BrandX	27.7	37.8	-10.1



# Statistical Inference About Difference in Means of More than Two Populations



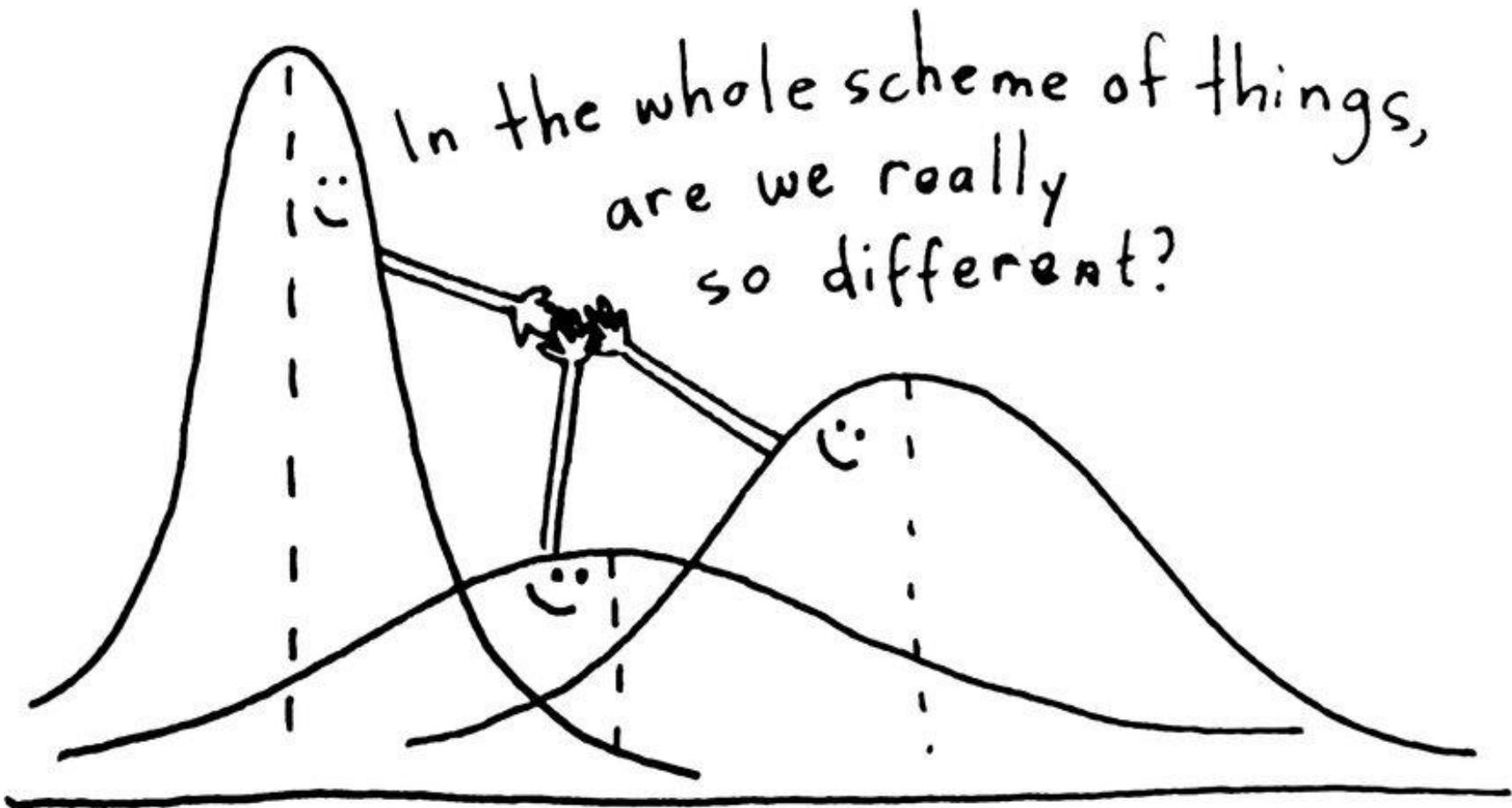
# Analysis of Variance

- Analysis of Variance (ANOVA) can be used to test for the equality of three or more population means using data obtained from observational or experimental studies.
  - We want to use the sample results to test the following hypotheses.
- 
- If  $H_0$  is rejected, we cannot conclude that all population means are different.
  - Rejecting  $H_0$  means that at least two population means have different values.

# ANOVA

- The ***response/dependent*** variable is the variable you're comparing
- The ***factor/treatment/independent*** variable is the categorical variable being used to define the groups
- **One-way ANOVA** - to test hypotheses about the mean on one dependent variable created by one independent variable. E.g. Income – dependent and Gender (Male/Female) - independent
- **Two-way ANOVA** - Tests hypotheses about the mean on one dependent variable for groups created by two independent variables (or factors). E.g. Income – dependent and Gender (Male/Female), Level of Education (HS/College/University) - independents

# ANOVA



# Test for the Equality of $k$ Population Means

## Hypotheses

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$H_a$ : Not all population means are equal

## Test Statistic

$$F = \text{MSTR}/\text{MSE}$$

# One-way ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F
Treatment	SSTR	$k - 1$	MSTR	MSTR/MSE
Error	SSE	$n_T - k$	MSE	
Total	SST	$n_T - 1$		

SST divided by its degrees of freedom  $n_T - 1$  is simply the overall sample variance that would be obtained if we treated the entire  $n_T$  observations as one data set.

# One-way ANOVA

- **Post hoc multiple comparison tests**
- Tukey, Tukey's b, and Bonferroni most common tests
  - One-way ANOVA can only tell you if  $F$  ratio is significant but not which groups are significantly different from one another
  - Post hoc tests can identify pairs of groups that significantly differ
  - If  $F$  ratio for model not significant, post hoc tests are not needed

# One-Way ANOVA Research Question

- Let's revert to the brand tracking questionnaire and data made available for this class on LEARN
  1. The brand who paid for the research (Brand X) is interested in knowing whether the satisfaction with their reward program differs by the marital status?

# One-way ANOVA Research Question

S7 Which of the following statements best describes your current marital status? <i>(Please select one answer)</i>	GO TO	1
Single (not living with a partner) Married/living with a partner Widowed/divorced/separated	S8	<b>Independent</b>

Q21. For the following programs that you are enrolled in, how satisfied are you with the value of the rewards? (only show responses of "enrolled" from previous question)

SAME ORDER AS Q20 and ONLY SHOW BRANDS ENROLLED IN	Very Satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Very dissatisfied
Plenti Rewards					
Brand X Star Rewards					
Amazon Prime					
TJX Rewards Access					
Nordstrom Fashion Rewards					
Bloomingdales Loyallist					
JC Penney JCP Rewards					
Kohl's Rewards					
Banana Republic/Gap/Old Navy					
Sephora Beauty Insider					
Nordstrom Fashion Rewards					
Ulta Ultamate Rewards					
DSW Rewards					
Victoria's Secret Angel Card					



**Dependent**

Q22. Which of the following life events have taken place for you in the past year? (select all that apply)

# Research Question

*The brand who paid for the research (Brand X) is interested in knowing whether the satisfaction with their reward program differs by the marital status?*

$$H_0: \mu_{\text{Single}} = \mu_{\text{Married}} = \mu_{\text{Widowed}}$$

$H_a$ : Not all population means are equal

# Analyze | Compare Means | One-Way ANOVA

(Use Post-Hoc to request Multiple Comparison Test)

One-Way ANOVA X

Dependent List:

- Brand X Star Rewards (For the follo...

Contrasts...  
Post Hoc...  
Options...  
Bootstrap...

Factor:

- Which of the following statements b...

Estimate effect size for overall tests

OK   Paste   Reset   Cancel   Help

## Statistics

- Descriptive
- Fixed and random effects
- Homogeneity of variance test
- Brown-Forsythe test
- Welch test

- Means plot

## Missing Values

- Exclude cases analysis by analysis
- Exclude cases listwise

## Confidence Intervals

Level(%):

**Continue****Cancel****Help**

# Requesting Sample Means and Post Hoc Multiple Comparisons

## Equal Variances Assumed

- |                                     |   |  |
|-------------------------------------|---|--|
| <input type="checkbox"/> LSD        | <input type="checkbox"/> S-N-K            | <input type="checkbox"/> Waller-Duncan   |
| <input type="checkbox"/> Bonferroni | <input checked="" type="checkbox"/> Tukey | Type I/Type II Error Ratio: <input type="text" value="100"/>   |
| <input type="checkbox"/> Sidak      | <input type="checkbox"/> Tukey's-b        | <input type="checkbox"/> Dunnett   |
| <input type="checkbox"/> Scheffe    | <input type="checkbox"/> Duncan           | Control Category : <input type="text" value="Last"/>   |
| <input type="checkbox"/> R-E-G-W F  | <input type="checkbox"/> Hochberg's GT2   | Test   |
| <input type="checkbox"/> R-E-G-W Q  | <input type="checkbox"/> Gabriel          | <input checked="" type="radio"/> 2-sided <input type="radio"/> < Control <input type="radio"/> > Control |

## Equal Variances Not Assumed

- Tamhane's T2
- Dunnett's T3
- Games-Howell
- Dunnett's C

## Null Hypothesis test

- Use the same significance level [alpha] as the setting in Options

- Specify the significance level [alpha] for the post hoc test

Level:

**Continue****Cancel****Help**

# Sample Means for Each Marital Status

## Descriptives

Brand X Star Rewards (For the following programs that you are enrolled in, how satisfied are you with the value of the rewards?)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Single (not living with a partner)	283	1.86	.874	.052	1.76	1.96	1	5
Married/living with a partner	768	1.99	.988	.036	1.92	2.06	1	5
Widowed/divorced/separated	79	2.15	1.220	.137	1.88	2.43	1	5
Total	1130	1.97	.981	.029	1.91	2.03	1	5

# Marital Status and Rewards Satisfaction

What is the probability of getting the sample results if the null hypothesis is true?

## ANOVA

Brand X Star Rewards (For the following programs that you are enrolled in, how satisfied are

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.281	2	3.140	3.278	.038
Within Groups	1079.756	1127	.958		
Total	1086.036	1129			

### Conventions\*

- $P > 0.10 \Rightarrow$  non-significant evidence against  $H_0$
- $0.05 < P \leq 0.10 \Rightarrow$  marginally significant evidence
- $0.01 < P \leq 0.05 \Rightarrow$  significant evidence against  $H_0$
- $P \leq 0.01 \Rightarrow$  highly significant evidence against  $H_0$

Reject or do not reject null hypothesis.

$p = .038$ , **Reject** null hypothesis, there is a significant evidence against null. Which groups are different?

See post-hoc multiple test (next slide)

# Marital Status and Rewards Satisfaction

## Multiple Comparisons

Dependent Variable: Brand X Star Rewards (For the following programs that you are enrolled in, how satisfied are you with the value of the rewards?)

Tukey HSD

(I) Which of the following statements best describes your current marital status?

(J) Which of the following statements best describes your current marital status?

Mean Difference (I-J)

Single (not living with a partner)

Married/living with a partner

-.130

.068

.136

-.29

.03

Married/living with a partner

Single (not living with a partner)

.130

.068

.136

duplicate

.29

Widowed/divorced/separated

Single (not living with a partner)

.290

.125

.053

duplicate

.58

Married/living with a partner

Married/living with a partner

.160

.116

.351

duplicate

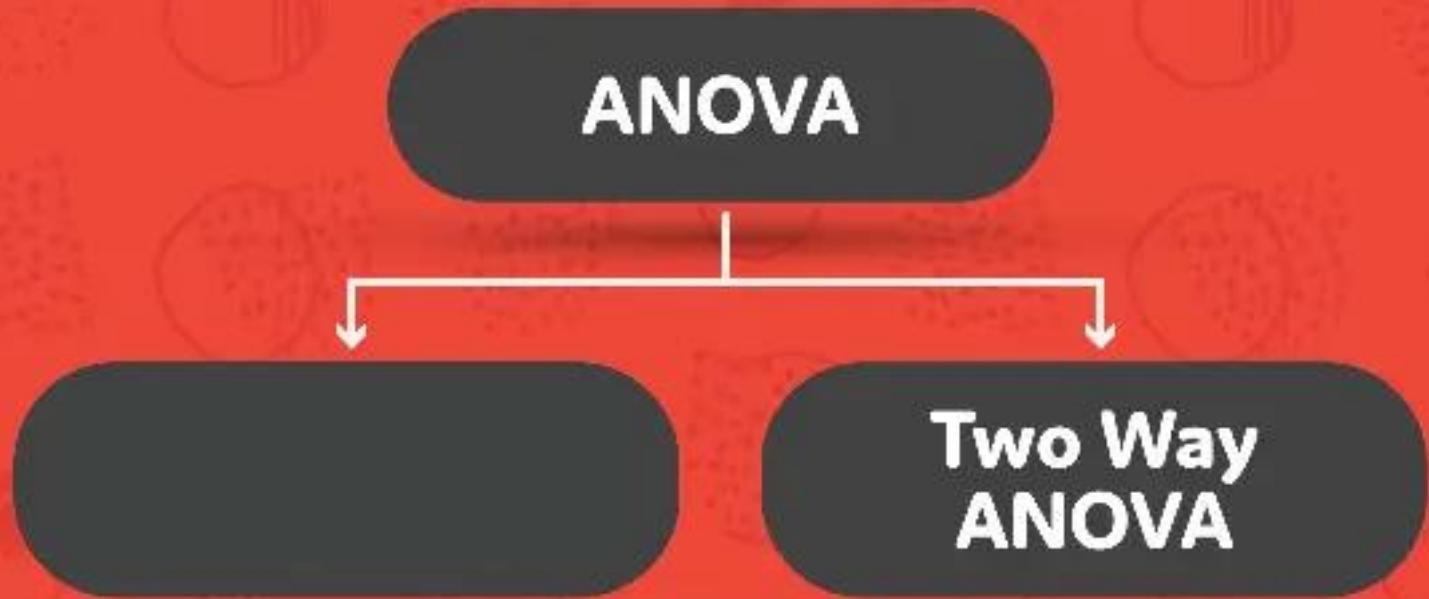
.43

Descriptives  
Brand X Star Rewards (For the following programs that you are enrolled in, how satisfied are you with the value of the rewards?)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean			
					Lower Bound	Upper Bound	Minimum	Maximum
Single (not living with a partner)	283	1.86	.874	.052	1.76	1.96	1	5
Married/living with a partner	768	1.99	.988	.036	1.92	2.06	1	5
Widowed/divorced/separated	79	2.15	1.220	.137	1.88	2.43	1	5
Total	1130	1.97	.981	.029	1.91	2.03	1	5

*The brand who paid for the research (Brand X) is interested in knowing whether the satisfaction with their reward program differs by the marital status?*

*Single people are significantly more satisfied than widowed/divorced/seperated people with Brand X's rewards programs.*



*What if more than two  
factors/independent variables  
are involved?*

# Two-way ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F
Factor A	SSA	$a - 1$	$MSA = \frac{SSA}{a-1}$	$\frac{MSA}{MSE}$
Factor B	SSB	$b - 1$	$MSB = \frac{SSB}{b-1}$	$\frac{MSB}{MSE}$
Interaction	SSAB	$(a - 1)(b - 1)$	$MSAB = \frac{SSAB}{(a-1)(b-1)}$	$\frac{MSAB}{MSE}$
Error	SSE	$ab(n_T - 1)$	$MSE = \frac{SSE}{ab(n_T - 1)}$	
Total	SST	$n_T - 1$		

# Two-way ANOVA Research Question

- Let's revert to the brand tracking questionnaire and data made available for this class on LEARN
  1. The brand who paid for the research (Brand X) is interested in knowing whether the satisfaction with their reward program differs by the marital status and Gender?

# Two-way ANOVA Research Question

S1 Are you...	GO TO
	Male S2
	Female

1

Independent

S7 Which of the following statements best describes your current marital status? <i>(Please select one answer)</i>	GO TO
Single (not living with a partner)	S8
Married/living with a partner	
Widowed/divorced/separated	

2

Independent

Q21. For the following programs that you are enrolled in, how satisfied are you with the value of the rewards? (only show responses of "enrolled" from previous question)

SAME ORDER AS Q20 and ONLY SHOW BRANDS ENROLLED IN	Very Satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Very dissatisfied
Plenti Rewards					
Brand X Star Rewards					
Amazon Prime					
TJX Rewards Access					
Nordstrom Fashion Rewards					
Bloomingdales Loyallist					
JC Penney JCP Rewards					
Kohl's Rewards					
Banana Republic/Gap/Old Navy					
Sephora Beauty Insider					
Nordstrom Fashion Rewards					
Ulta Ultamate Rewards					
DSW Rewards					
Victoria's Secret Angel Card					

Dependent

Q22. Which of the following life events have taken place for you in the past year? (select all that

## Analyze | General Linear Model| Univariate

(Use **Plots** and **Post-Hoc** to request Subgroup Means and a Plot of Subgroup Means)

Univariate

Dependent Variable: Brand X Star Rewards (For the follow...  
Fixed Factor(s): Are you... [s1]  
Which of the following statements b...  
Random Factor(s):  
Covariate(s):  
WLS Weight:  
Model...  
Contrasts...  
Plots...  
Post Hoc...  
EM Means...  
Save...  
Options...  
Bootstrap...

Yes, Cuban (Are you of Hispanic, ...  
Yes, Central / South American (A...  
Yes, other Hispanic, Latino or Sp...  
White / Caucasian (Which of the ...  
Black / African American (Which ...  
Asian or Pacific Islander (Which o...  
Alaskan Native / American Indian ...  
Other (Specify) (Which of the follo...  
Other (Specify) (Which of the follo...  
Which of the following best descri...  
An advertising agency/advertising ...  
A company or store that makes o...  
A company or store that makes o...  
A marketing or marketing researc...  
Banking (Are you or any member...  
OK Paste Reset Cancel Help

# Requesting Subgroup Means and a Plot of Subgroup Means

**Display**

Descriptive statistics       Homogeneity tests  
 Estimates of effect size       Spread-vs.-level plots  
 Observed power       Residual plots  
 Parameter estimates       Lack-of-fit test  
 Contrast coefficient matrix       General estimable function(s)

**Heteroskedasticity Tests**

Modified Breusch-Pagan test       F test  
 Model...       Model...  
 Breusch-Pagan test       White's test  
 Model...

Parameter estimates with robust standard errors

HC0       HC1  
 HC2       HC3  
 HC4

Significance level

**Univariate: Estimated Marginal Means**

**Estimated Marginal Means**

**Factor(s) and Factor Interactions:** (OVERALL)  
 s1  
 s7  
 s1\*s7

**Display Means for:**

Compare main effects

**Confidence interval adjustment:** LSD(none)

**Continue****Cancel****Help**

**Factors:** s1, s7

**Horizontal Axis:** s7

**Separate Lines:** s1

**Separate Plots:**

**Plots:**

s1\*s7

**Chart Type:**  Line Chart       Bar Chart

**Error Bars**

Include Error bars  
 Confidence Interval (95.0%)  
 Standard Error      Multiplier: 2

Include reference line for grand mean  
 Y axis starts at 0

**Continue** **Cancel** **Help**

# Significance Tests

1. Overall Model: **significant**
2. Gender Effect: **irrelevant**
3. Interaction Effect: **significant**

## Tests of Between-Subjects Effects

Dependent Variable: Brand X Star Rewards (For the following programs that you are enrolled in)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	12.881 <sup>a</sup>	5	2.576	2.698	.020 1
Intercept	1178.558	1	1178.558	1234.396	.000
s1	5.792	1	5.792	6.067	.014 2
s7	2.322	2	1.161	1.216	.297
s1 * s7	4.599	2	2.300	2.409	.090 3
Error	1073.156	1124	.955		
Total	5475.000	1130			
Corrected Total	1086.036	1129			

a. R Squared = .012 (Adjusted R Squared = .007)

## Estimated Marginal Means

Which of the following statements best describes your current marital status? \* Are you...

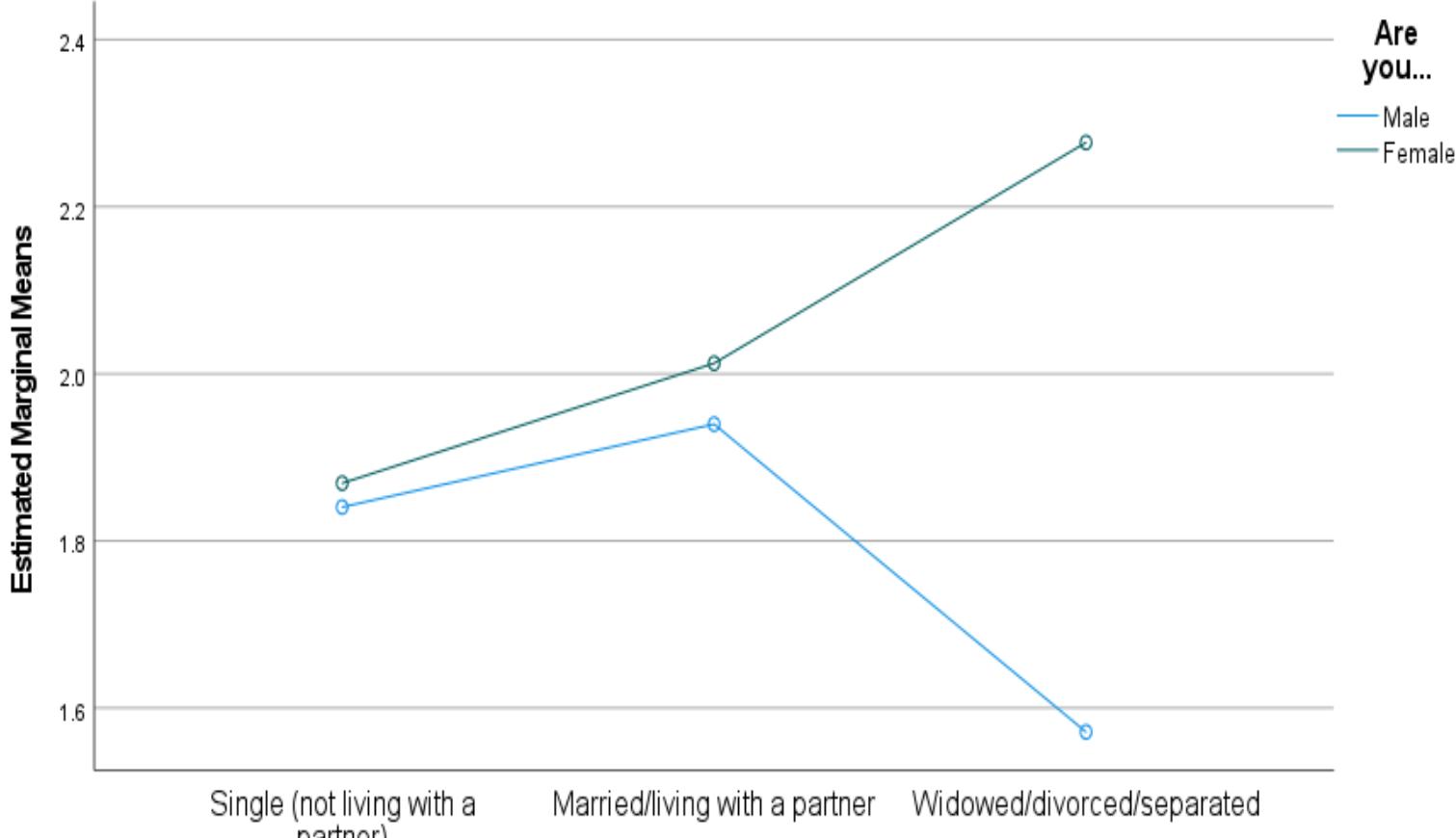
Dependent Variable: Brand X Star Rewards (For the following programs that you are enrolled in, how satisfied are you with the value of the rewards?)

Which of the following statements best describes your current marital status?	Are you...	95% Confidence Interval			
		Mean	Std. Error	Lower Bound	Upper Bound
Single (not living with a partner)	Male	1.841	.118	1.610	2.071
	Female	1.869	.067	1.738	2.000
Married/living with a partner	Male	1.940	.066	1.809	2.070
	Female	2.013	.042	1.931	2.094
Widowed/divorced/separated	Male	1.571	.261	1.059	2.084
	Female	2.277	.121	2.039	2.515

# Subgroup Means

# Subgroup Plot

Estimated Marginal Means of Brand X Star Rewards (For the following programs that you are enrolled in, how satisfied are you with the value of the rewards?)



Which of the following statements best describes your current marital status?

# Two-Way ANOVA

- If interaction is significant, then interpret it along with means and plot.
  - This indicates that the IV's are not acting separately from one another in their effect on the DV. Main effect becomes irrelevant.
- If interaction is not significant, interpret main effects.
  - This indicates that IV effects on DV are independent of one another and that there is no significant interaction of the two IV's in the population.

# Examining Relationships Among Variables

# The Question

- Are two variables related?
  - Does one increase as the other increases?
    - e. g. skills and income
  - Does one decrease as the other increases?
    - e. g. nutrition and health problems
- How can we get a numerical measure of the degree of relationship?

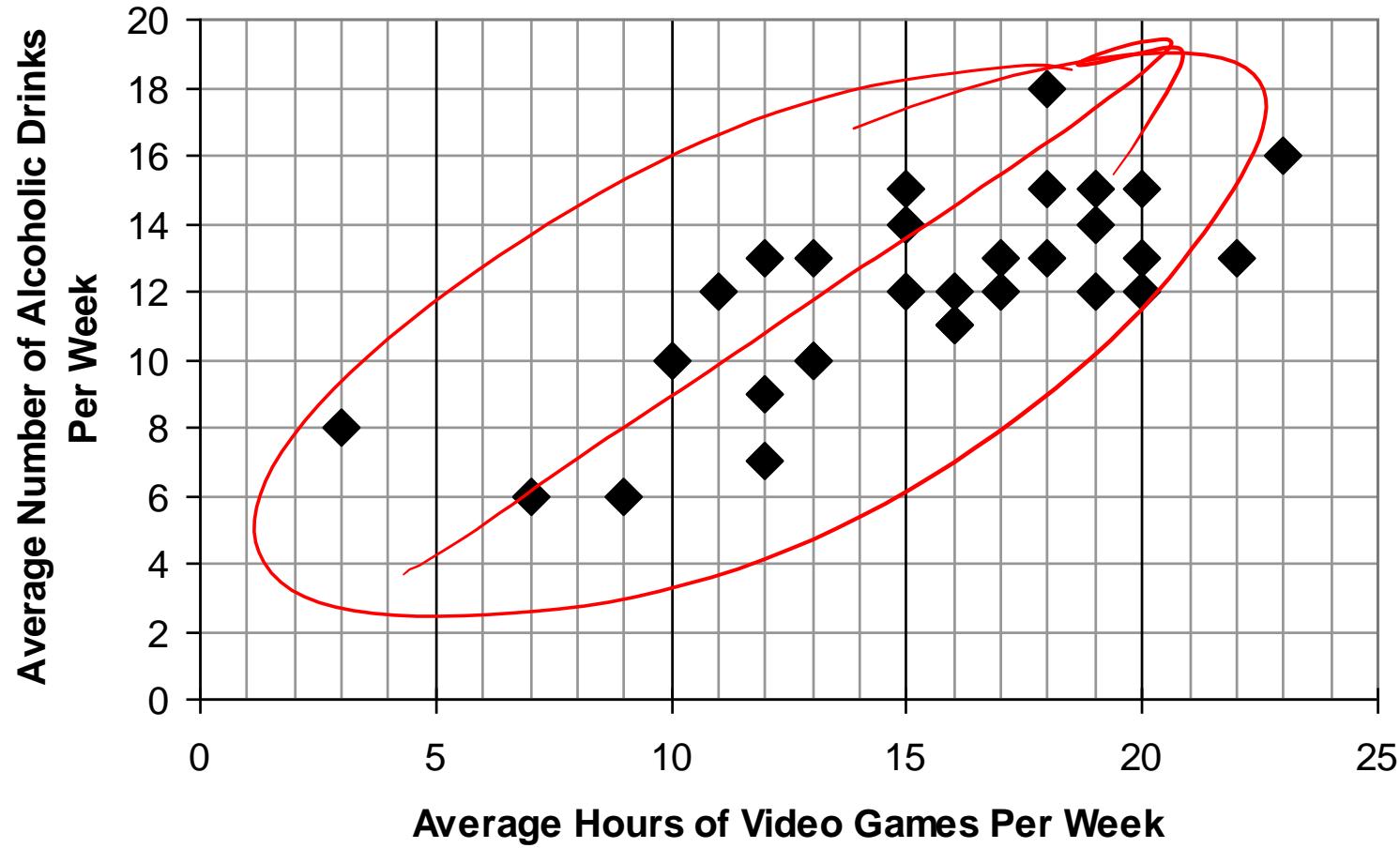


## Scatter diagram or scattergram

- Graphically depicts the relationship between two variables in two-dimensional space.

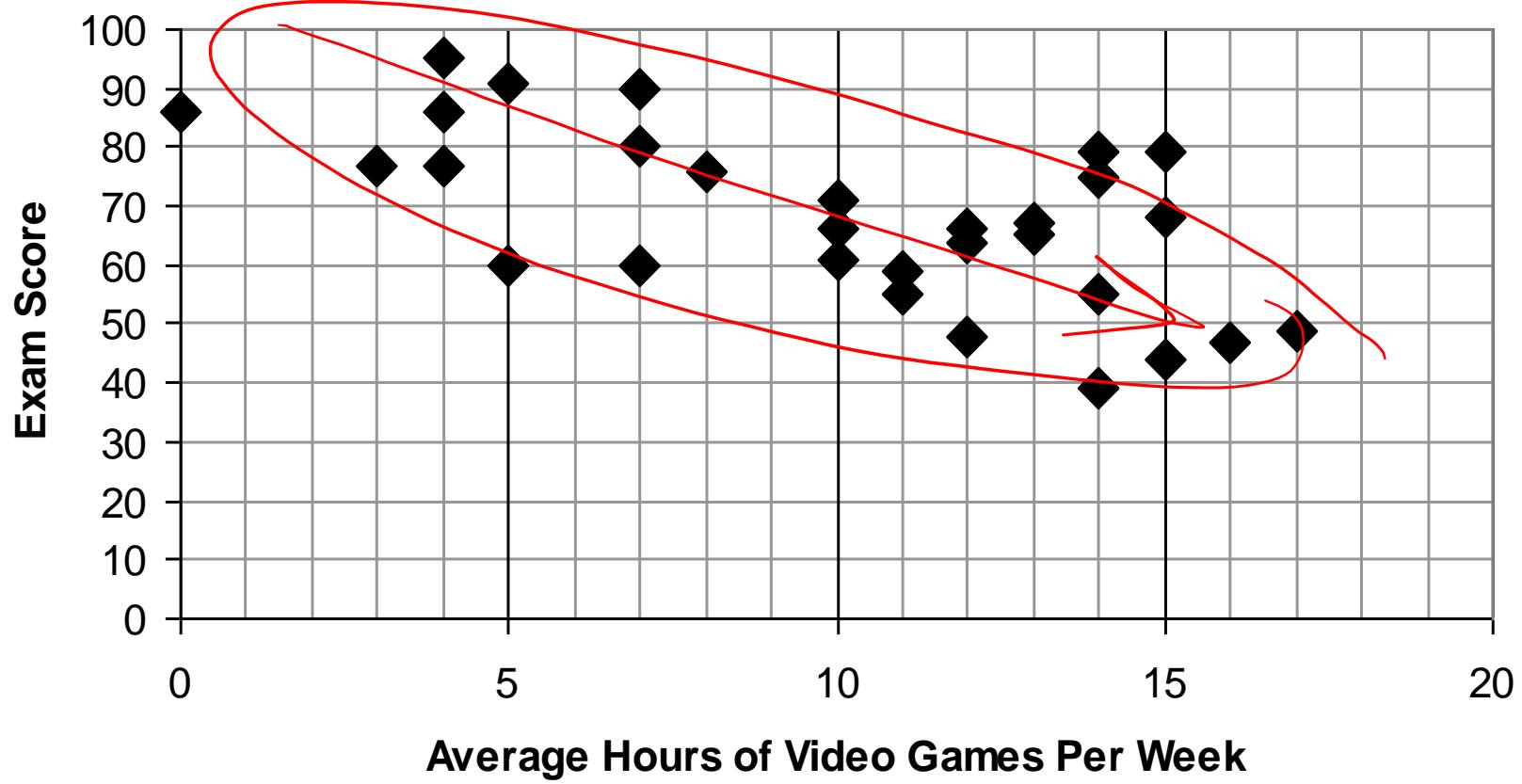
# Direct Relationship

Scatterplot: Video Games and Alcohol Consumption



# Inverse Relationship

Scatterplot: Video Games and Test Score



# Numerical Measure of a Simple Linear Relationship

# Correlation

- The relationship between two variables
- Measured with a correlation coefficient
- Most popularly seen **correlation coefficient: Pearson Product-Moment Correlation**
- Symbolized by  $r$  (*sample*) and  $\rho$  (*population*)
- A measure of degree of a **linear relationship**
- **Varies between 1 and -1**, where the sign refers to relational direction.
- **Based on covariance**
  - Measure of degree to which large scores on X go with large scores on Y, and small scores on X go with small scores on Y
- **Does not imply causation**

# Types of Correlation

## ■ Positive correlation

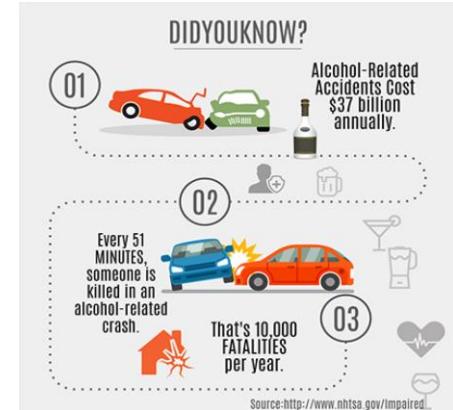
- High values of X tend to be associated with high values of Y.
- As X increases, Y increases

## ■ Negative correlation

- High values of X tend to be associated with low values of Y.
- As X increases, Y decreases

## ■ No correlation

- No consistent tendency for values on Y to increase or decrease as X increases



SL  
p  
MORE  
STRESS  
LESS

## Can Plant Protein Get You Jacked?

Or are you better off sticking with an animal-based protein like whey?

**OPENING ARGUMENTS**

**DEFENSE**

- Protein is protein, no matter where it comes from. Your body (and muscles) don't care about the source.
- A study in the Journal of the American College of Nutrition in 2013 found that when vegans ate plant-based protein supplements after resistance training, they increased muscle mass and strength more than those who ended up with almost double the protein from animal sources.
- A recent study from Penn International University looked at the effects of whey protein supplemented with either whey or plant-based protein after resistance training. They found no significant difference in muscle mass or body composition after six weeks.

**PROSECUTION**

- Protein is a simple molecule that helps build up tissue, muscles, and bone—but not all proteins are created equal. The amino acids that make up protein are important, and not all protein sources contain the same amino acids.
- The essential amino acids our bodies need to grow and repair tissue are called "essential amino acids." Animal sources usually contain all the essential amino acids, whereas plant-based protein powders have to have them added, which can be less effective.

**EVIDENCE**

- In 2012, a study published in Nutrition Journal had 24 guys lift weights and supplement with either whey or rice protein. At a high dose (by weight), the whey protein group found no difference in muscle mass, strength, or choice fat loss, strength, and power.
- Some claim that soy protein can cause breast enlargement in men (and spur breast growths), but that's not true. A study from Catherine University in Minneapolis found that soy protein isolate had no effect.

**VERDICT**

There are differences between protein sourced from animals and protein derived from plants, but there's not enough evidence to recommend one over the other.

**SENTENCING**

Some guys say that whey protein is the best way to go if that's the case for you, go with it. If you want to get the most protein per pound of body weight, whey protein can cover the bases.

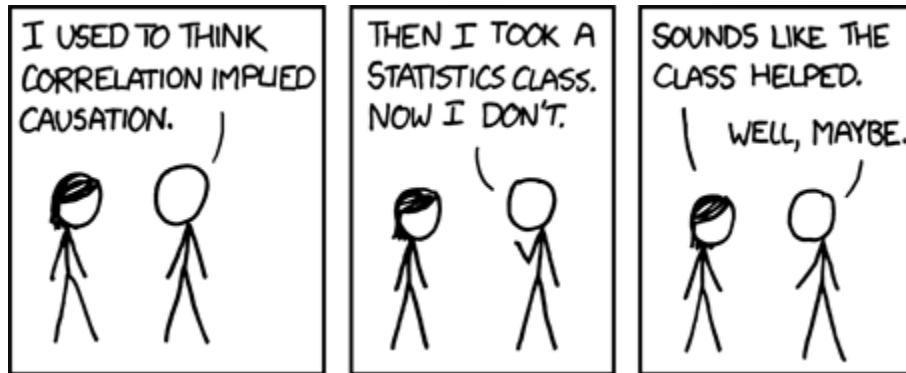
Illustration: Michael Hickey

PHOTOGRAPH BY JEFFREY M. CAMPION FOR PLATE

# Correlation Size

Size of Correlation	Interpretation
$\pm 1$	Perfect Positive/Negative Correlation
$\pm .90$ to $\pm .99$	Very High Positive/Negative Correlation
$\pm .70$ to $\pm .90$	High Positive/Negative Correlation
$\pm .50$ to $\pm .70$	Moderate Positive/Negative Correlation
$\pm .30$ to $\pm .50$	Low Positive/Negative Correlation
$\pm .10$ to $\pm .30$	Very Low Positive/Negative Correlation
$\pm .00$ to $\pm .10$	Markedly Low and Negligible Positive/ Negative Correlation

# What do you mean no causality if there is a correlation?

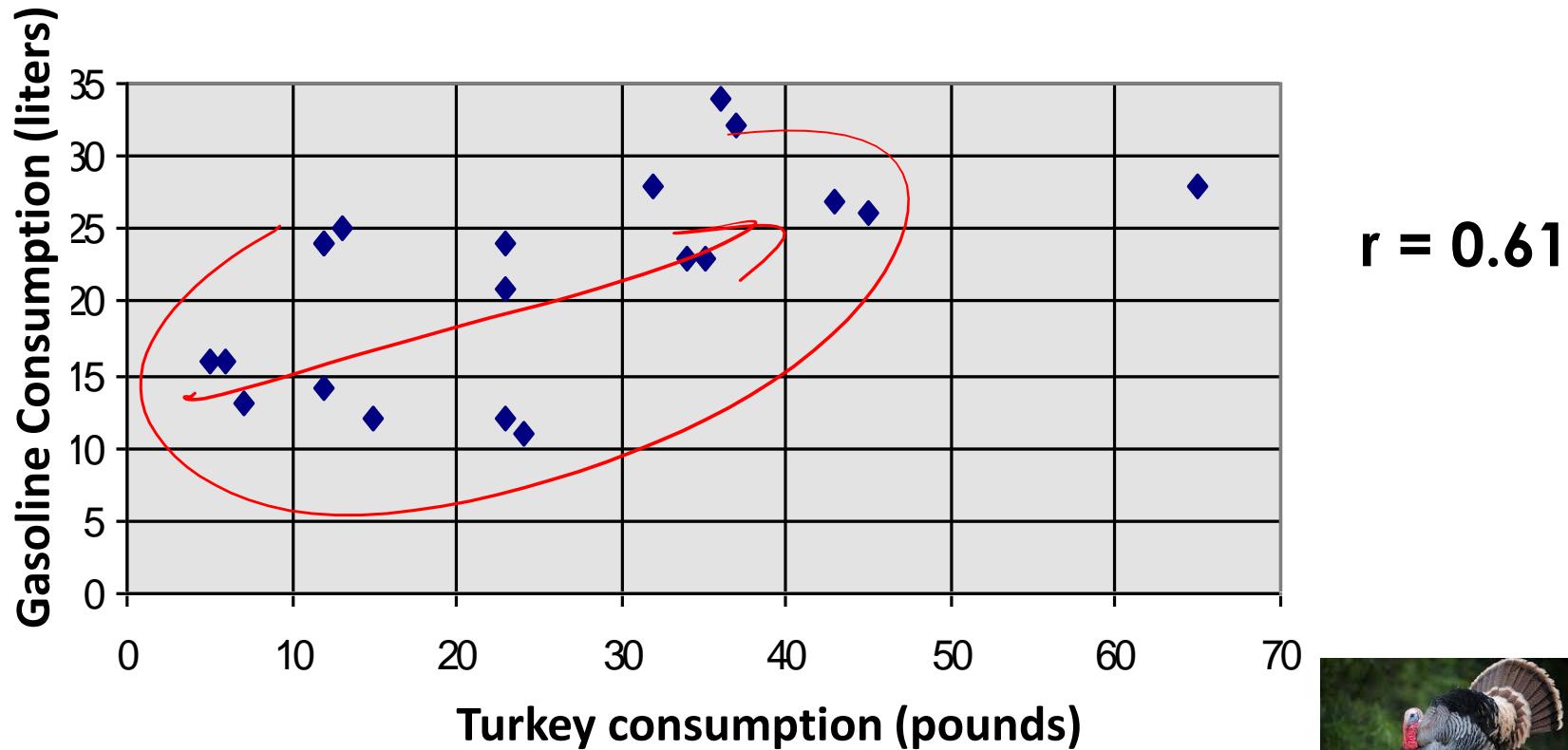




# Correlation and Causation

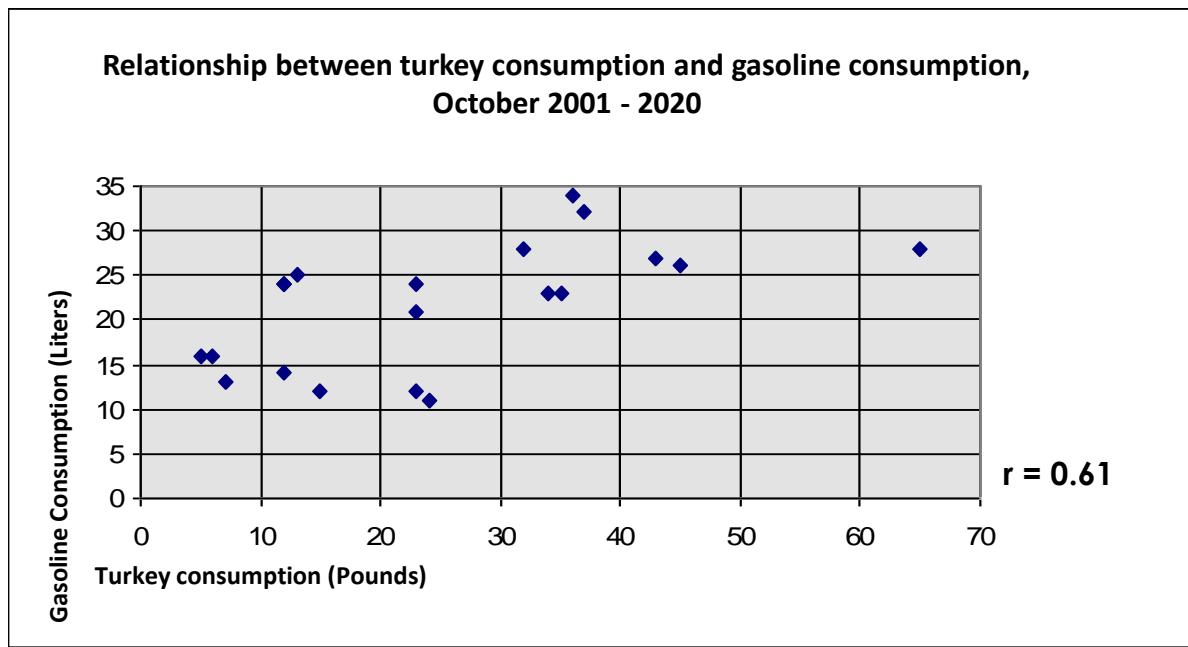


Relationship between turkey consumption and gasoline consumption, October, 2001-2020



# Correlation and Causation

- Does it mean that one variable causes the other ?
- Turkey consumption and gasoline consumption are **strongly related** ( $r = 0.61$ ) but that does not indicate that consuming turkeys explains increased gasoline consumption.
  - Maybe a third variable????



TILL  
NEXT  
TIME

