

BACS2042 Research Methods

Data Analysis

Recap

Types of Validity

Internal Validity

Construct Validity

External Validity

Statistical Validity

Threats to Internal Validity (11)

Threats to External Validity (3)

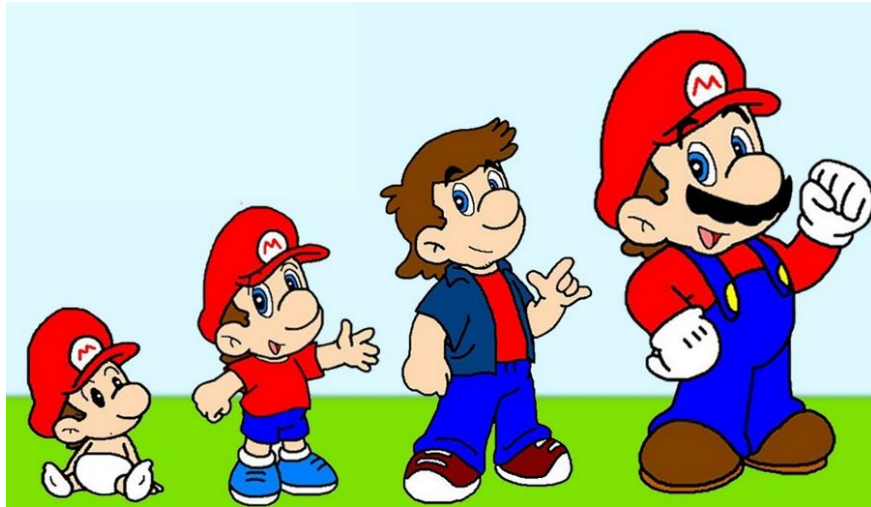
Confident that **changes in DV** are caused by the **IV**

How well the **RESULT** support the **theory/principle**

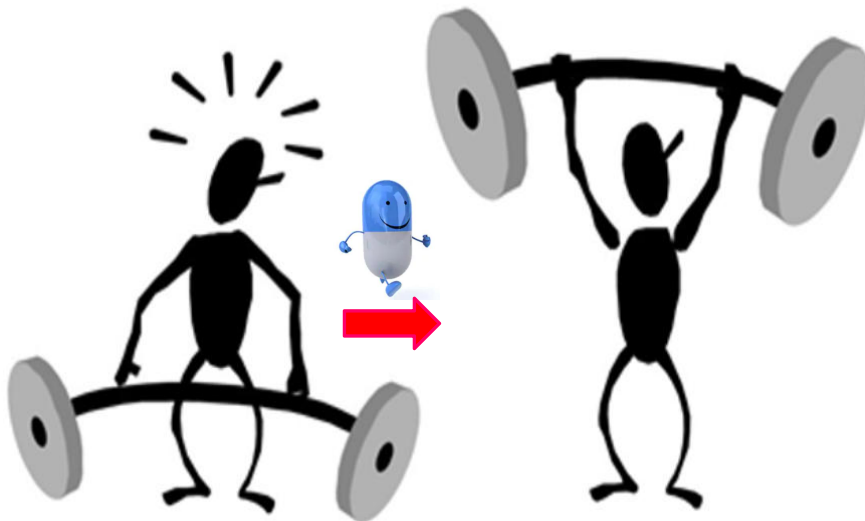
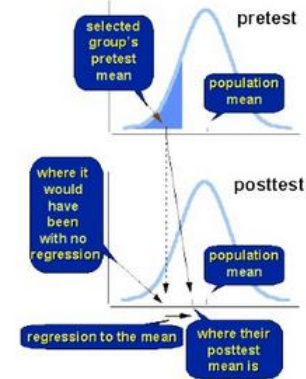
extent to which **data are shown to be the result of cause-effect relationships** rather than accident

To what extent would the results **be transferable or generalizable** to the actual world

Threats to Internal Validity



Sequencing effect



Threats to External Validity

Threats (External)	
1. Other subjects	we cannot assume that animal can be substitute for any other (human) in all situations.
2. Other times	would the same experiment conducted at another time (e.g. after 20 years) produce the same results?
3. Other settings	how the phenomenon observed in one laboratory can be related to a similar phenomenon observed in another laboratory or in the real world?



Will technology gadgets bring negative effects to children?



Data Analysis

Investigating
Differences

Investigating
Relationship

Non-parametric

T-Test

ANOVA

Correlation

Regression

Chi-square

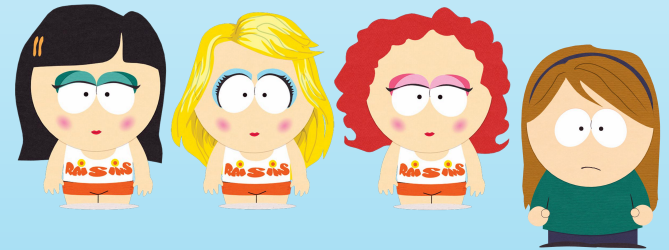
1-way
ANOVA

2-way
ANOVA

Investigating differences: T-Test

Independent Variable	Dependent Variable
1 (2 levels)	1
Gender (M & F)	Hours of sleep

“Is there a significant difference of sleeping hours between boys and girls in this class?”



Male	Hours of sleep
Student A	12
Student B	10
Student C	11
Student D	9
μ_m	10.5

Female	Hours of sleep
Student W	6
Student X	7
Student Y	8
Student Z	9
μ_f	7.5

Investigating differences: T-Test

$$H_0: \mu_m = \mu_f$$

Group Statistics					
	gender	N	Mean	Std. Deviation	Std. Error Mean
hour	M	4	10.50	1.291	.645
	F	4	7.50	1.291	.645

Homogeneity of variance

Independent Samples Test

Levene's Test for Equality of Variances

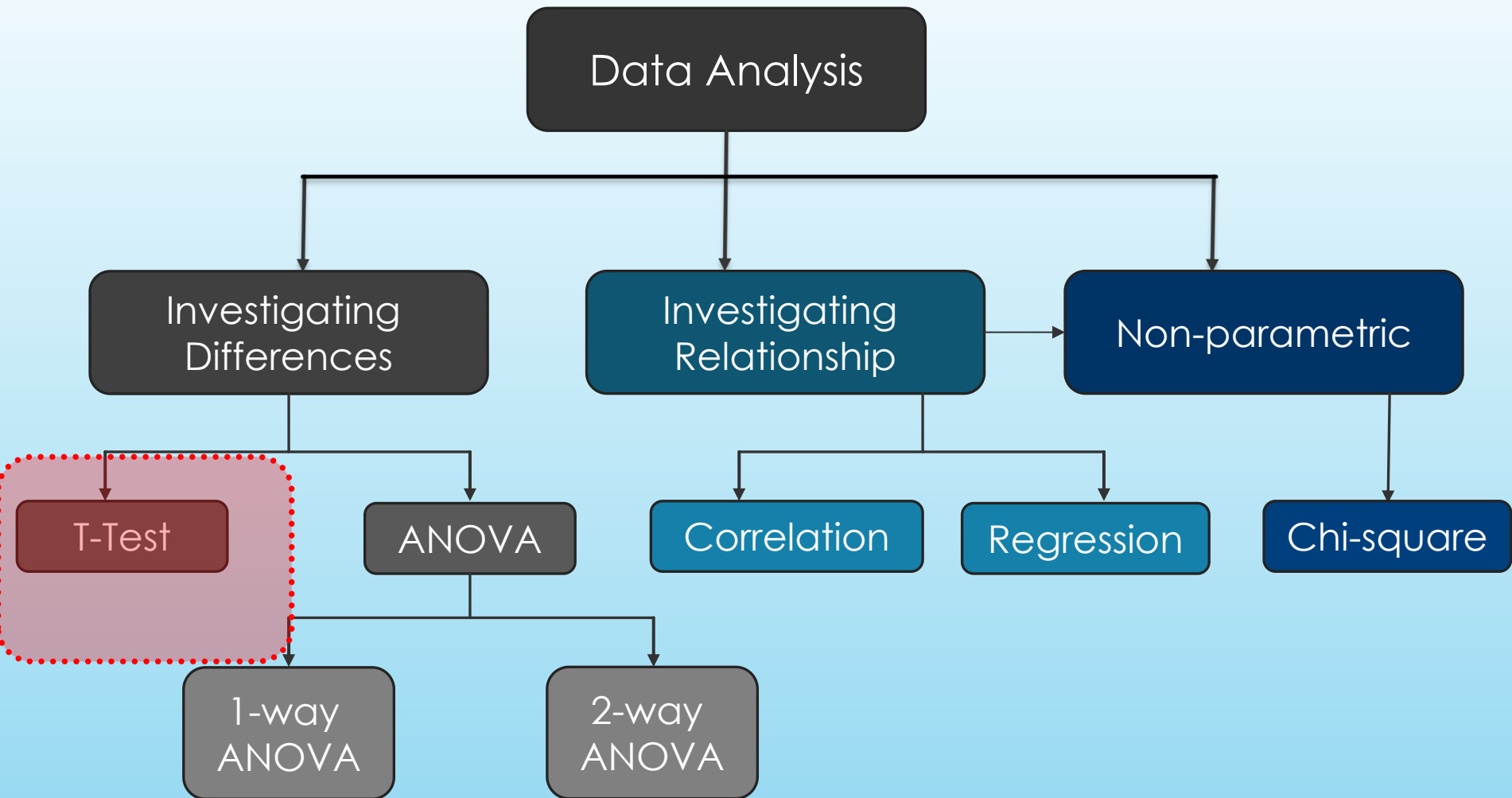
t-test for Equality of Means

95% Confidence Interval of the Difference

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
hour	Equal variances assumed	.000	1.000	3.286	6	.017	3.000	.913	.766	5.234
	Equal variances not assumed			3.286	6.000	.017	3.000	.913	.766	5.234

Reject null hypothesis

There is a significant difference of sleeping hours between the boys and girls in this class



Investigating differences: ANOVA

Independent Variable	Dependent Variable
1 (>2 levels) (cat)	1 (continuous data)
Program (RSF, RST, REI)	Hours of sleep

“Is there a significant difference in sleeping hours among the students of different prog?”



RSF



RST



REI

RSF	Hours of sleep
Student A	7
Student B	8
Student C	9
Student D	6

RST	Hours of sleep
Student a	7
Student b	8
Student c	9
Student d	7

REI	Hours of sleep
Student W	6
Student X	7
Student Y	8
Student Z	9

Investigating differences: ANOVA

$$H_0: \mu_{RSF} = \mu_{RST} = \mu_{REI}$$

Descriptives								
hour								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
REI	4	7.50	1.291	.645	5.45	9.55	6	9
RSF	4	7.75	.957	.479	6.23	9.27	7	9
RST	4	7.50	1.291	.645	5.45	9.55	6	9
Total	12	7.58	1.084	.313	6.89	8.27	6	9

Test of Homogeneity of Variances				
hour				
	Levene Statistic	df1	df2	Sig.
	.300	2	9	.748

ANOVA					
hour					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.167	2	.083	.059	.943
Within Groups	12.750	9	1.417		
Total	12.917	11			

Accept null hypothesis

Data Analysis

Investigating
Differences

Investigating
Relationship

Non-parametric

T-Test

ANOVA

Correlation

Regression

Chi-square

1-way
ANOVA

2-way
ANOVA

Investigating differences: 2-wayANOVA

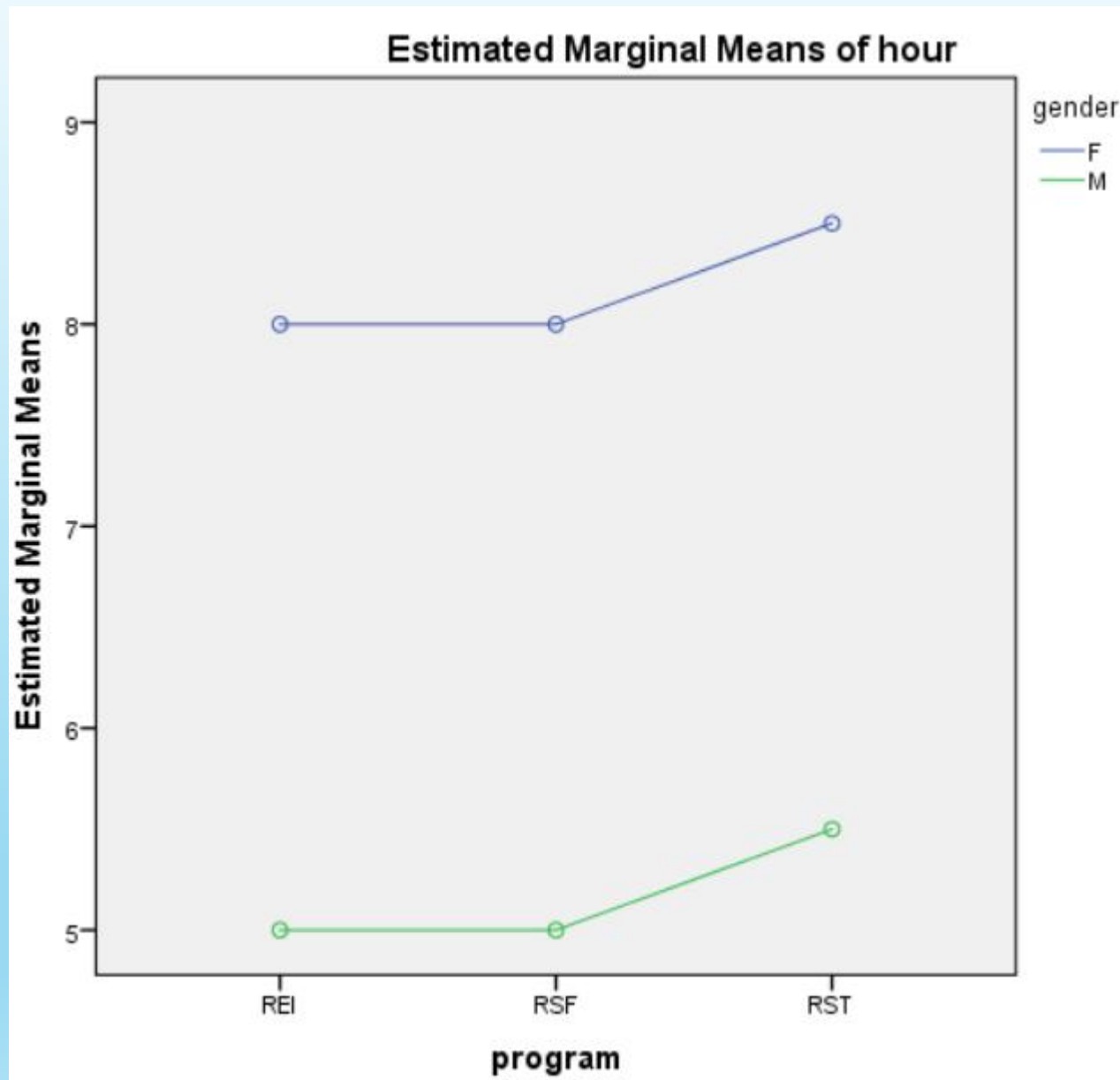
Independent Variable	Dependent Variable
2 (cat.)	1 (continuous data)
Gender (M & F)	Hours of sleep
Program (RSF, RST, REI)	

$$H_0: \mu_m = \mu_f$$

$$H_0: \mu_{RSF} = \mu_{RST} = \mu_{REI}$$

H_0 : *There's no interaction between the factors*

Tests of Between-Subjects Effects					
Dependent Variable: hour					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	27.667 ^a	5	5.533	33.200	.000
Intercept	533.333	1	533.333	3200.000	.000
Program	.667	2	.333	2.000	.216
Gender	27.000	1	27.000	162.000	.000
Program * Gender	.000	2	.000	.000	1.000
Error	1.000	6	.167		
Total	562.000	12			
Corrected Total	28.667	11			



Data Analysis

Investigating Differences

Investigating Relationship

Non-parametric

T-Test

ANOVA

Correlation

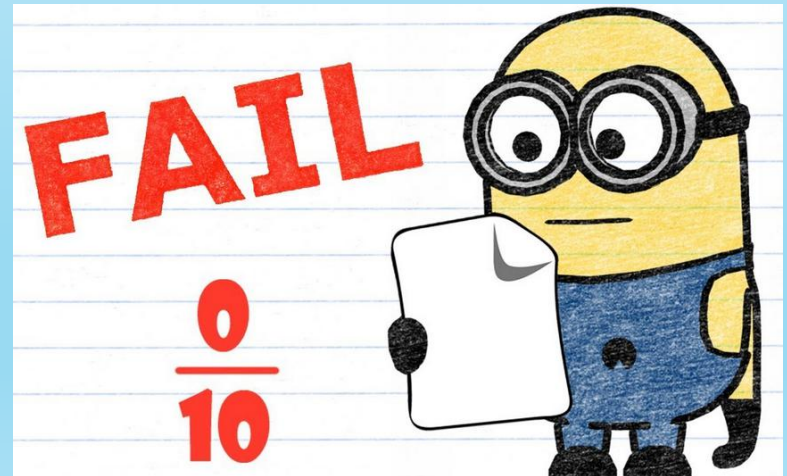
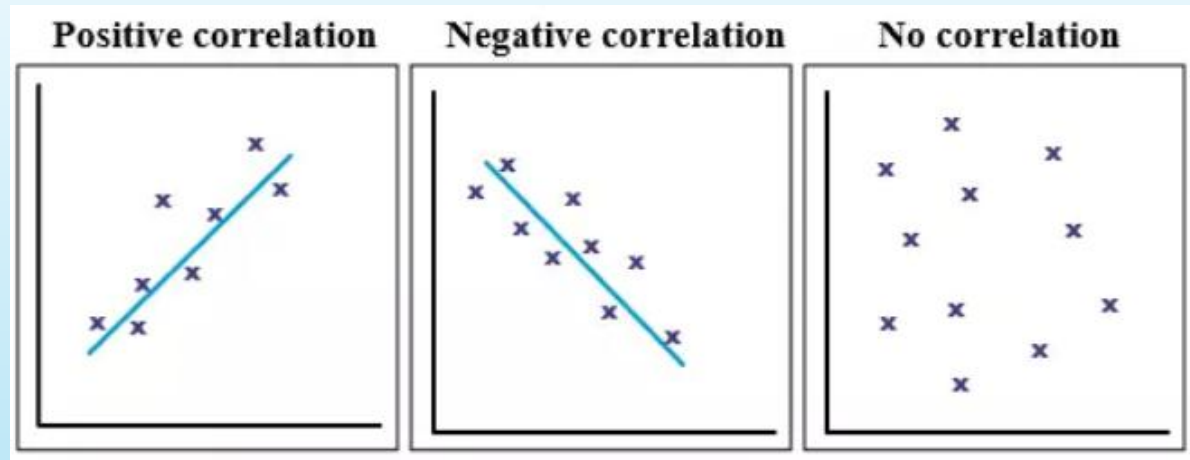
Regression

Chi-square

1-way ANOVA

2-way ANOVA

Investigating relationship: **Correlation**



Investigating relationship: **Correlation**

Independent Variable	Dependent Variable
Hours of sleep	Final Score

Correlations		score	hour
score	Pearson Correlation	1	-.893**
	Sig. (2-tailed)		.000
	N	12	12
hour	Pearson Correlation	-.893**	1
	Sig. (2-tailed)	.000	
	N	12	12



Data Analysis

Investigating
Differences

Investigating
Relationship

Non-parametric

T-Test

ANOVA

Correlation

Regression

Chi-square

1-way
ANOVA

2-way
ANOVA

Investigating relationship: **Regression**

- “What is the individual and combined power of diploma CGPA and MUET result in predicting CGPA in bachelor?”

Diploma

CGPA
3.75 - 4.0
3.5 - 3.74
3.0 - 3.49
2.5 - 2.99
2.0 - 2.49
< 2.0

MUET

Test Component	Maximum Score	Obtained Score
LISTENING	45	14
SPEAKING	45	28
READING	120	51
WRITING	90	55
AGGREGATED SCORE	300	148
BAND ACHIEVED		3



predict

Bachelor

CGPA
3.75 - 4.0
3.5 - 3.74
3.0 - 3.49
2.5 - 2.99
2.0 - 2.49
< 2.0

Investigating relationship: Regression

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.997 ^a	.993	.992	.06053	2.492

a. Predictors: (Constant), diploma cgpa, muet band

b. Dependent Variable: bachelor cgpa

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.116	.133		.874	.403					
	muet band	.011	.028	.029	.412	.689	.930	.129	.011	.136	7.375
	diploma cgpa	.965	.070	.970	13.755	.000	.997	.975	.357	.136	7.375

$$\text{Bachelor CGPA} = 0.116 + 0.11(\text{MUET}) + 0.965(\text{Dip. CGPA})$$

Data Analysis

Investigating Differences

Investigating Relationship

Non-parametric

T-Test

ANOVA

Correlation

Regression

Chi-square

1-way ANOVA

2-way ANOVA

Chi-Square

- Investigate if 2 categorical variables are related

RSF RST
RIT REI

Program



R/ship status

H_0 : Program is independent of r/ship status

program * status Crosstabulation					
Count		status			Total
		single	in relationship	complicated	
program	RSF	2	1	1	4
	RST	1	2	1	4
	REI	1	2	1	4
	RSD	2	0	2	4
	RIT	1	2	1	4
	RIS	2	1	1	4
Total		9	8	7	24

Chi-Square Tests			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	4.214 ^a	10	.937
Likelihood Ratio	5.349	10	.867
Linear-by-Linear Association	.000	1	1.000
N of Valid Cases	24		

Accept null hypothesis

No association was found between program and r/ship status.

Data Analysis

Investigating
Differences

Investigating
Relationship

Non-parametric

T-Test

ANOVA

Correlation

Regression

Chi-square

1-way
ANOVA

2-way
ANOVA