Tutorial: Mean Difference Test

T-test, ANOVA, Chi-sq

Number Analytics LLC

March 2017

Coca cola vs Pepsi, taste better?

How to test blind taste test results

Is there significant taste difference?

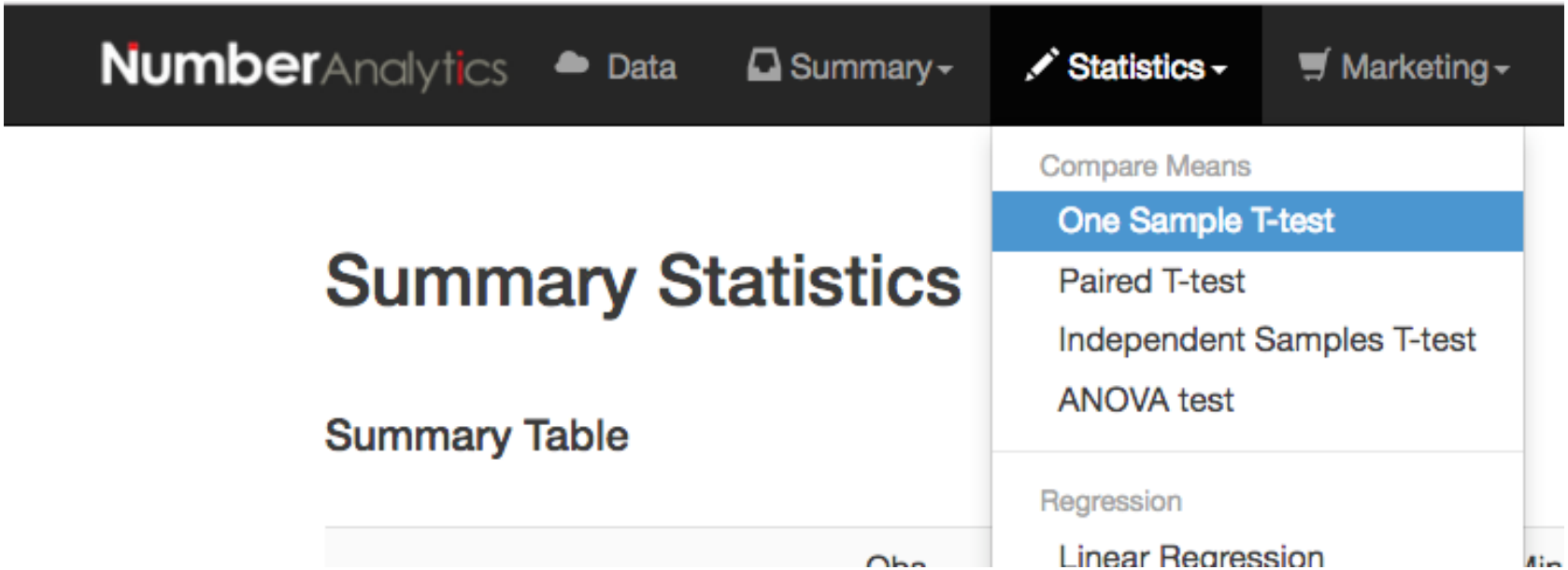
NumberAnalytics

Statistical Tests for Differences

Choose the correct statistic test

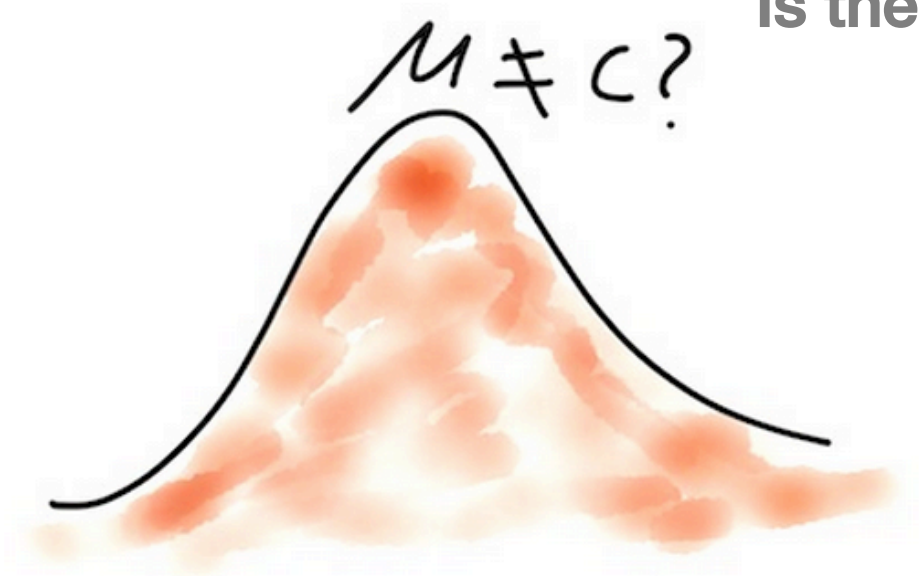
Area of Application	Level Scaling	Subgroups	Test	Example
Hypotheses about frequency	Nominal	All	Chi-square	Do customer industry types differ by company size ?
Hypotheses about means	Metric (Interval or ratio)	One	One Sample T-test	Is the purchase frequency different from 1.5?
		Two	Independent Samples T-test	Is the purchase frequency greater for email promotion responders than that for non-responders?
			Paired Sample T-test	Is the average brand score different after exposure to a new TV ad? (before and after)
		Three or more	One-way ANOVA	Is the purchase frequency different by company size?

One Sample T-test



One Sample T-test

One Sample T-test



Please select the data file below

- (•) Sample
- ( ) Your own file

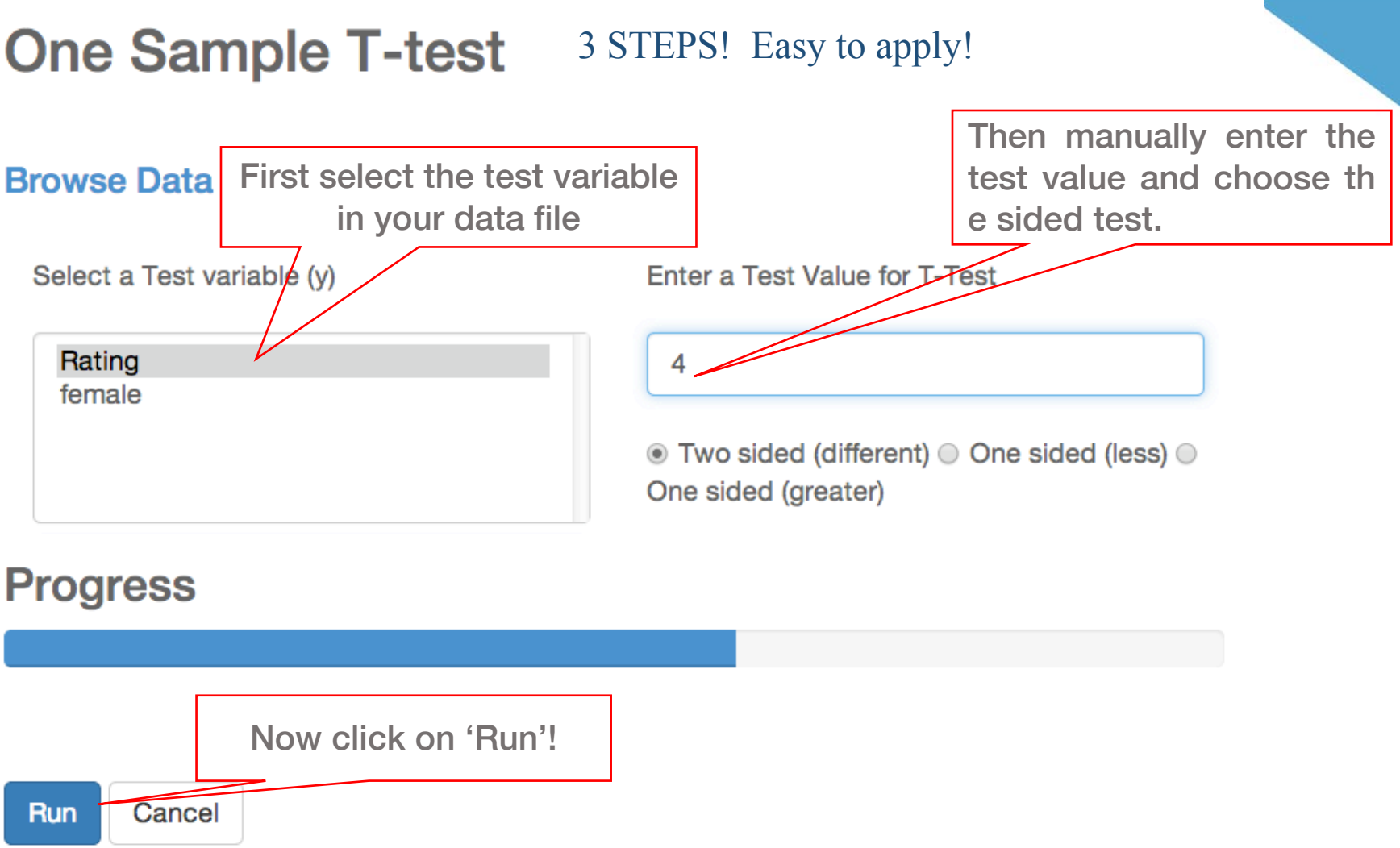
You can choose your own file by uploading it to the cloud.

Is the overall rating significantly different than 4?

data_ttest

	A	B
1	Rating	female
2	9	1
3	4	1
4	3	1
5	6	0
6	9	0
7	7	1
8	5	0
9	6	0

One Sample T-test

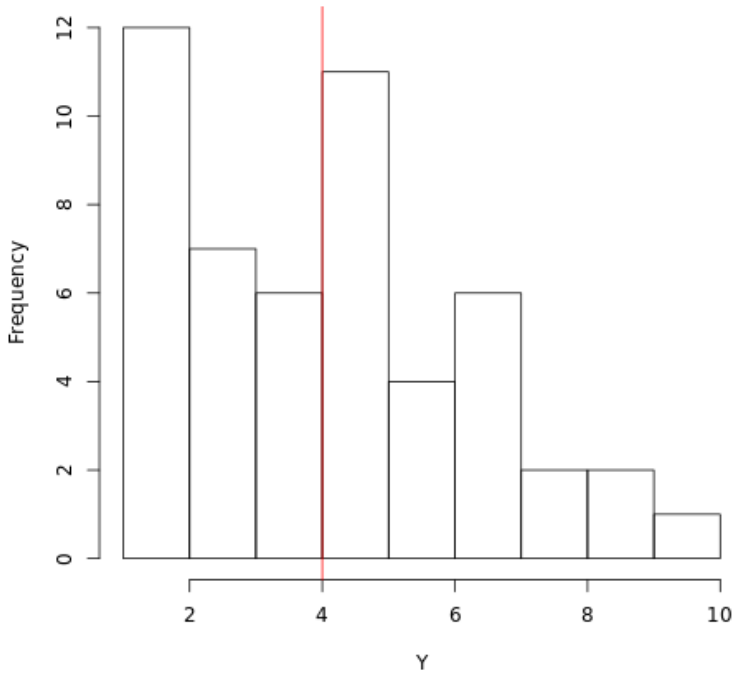


One Sample t-test

One Samples T-test

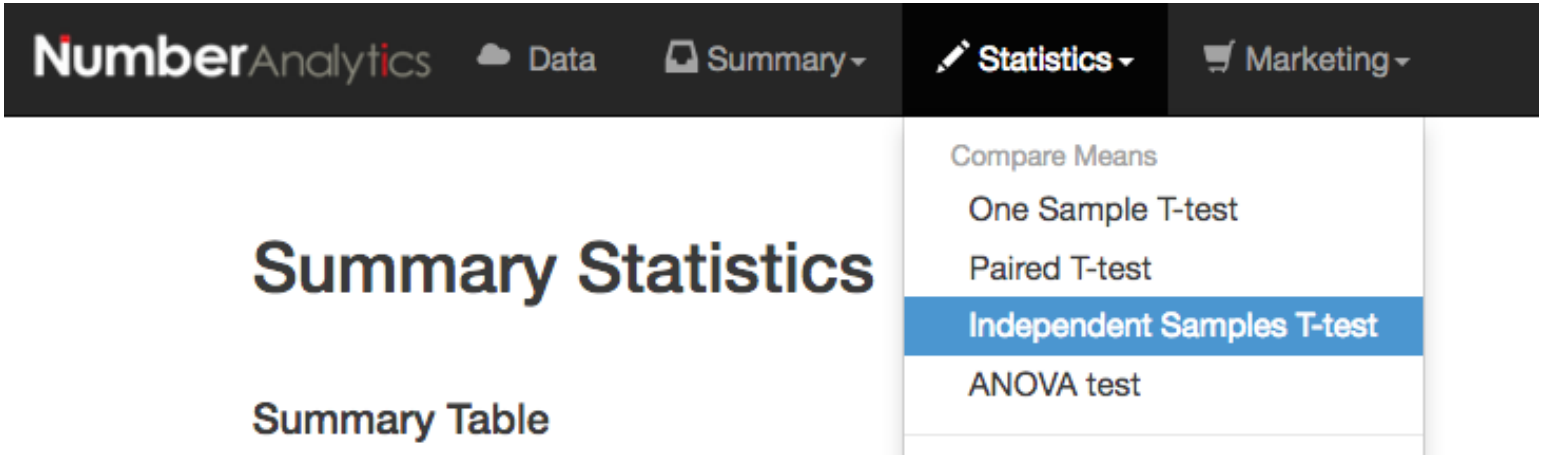
t_value	pvalue	
1.439	0.16	Rating is indifferent from 4

Conclusion: Rating is not different from 4 on average.



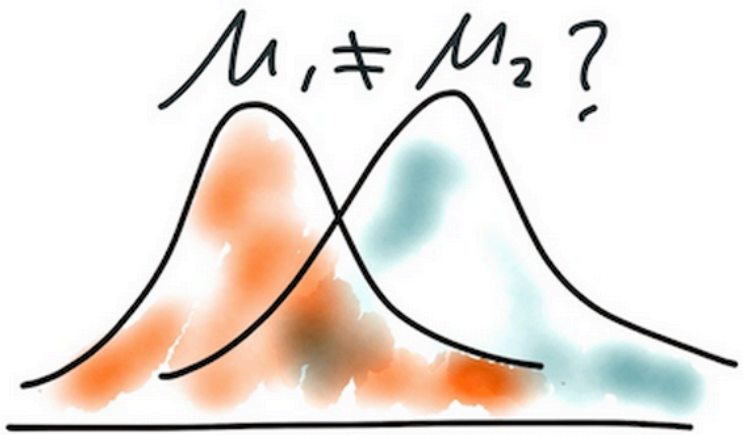
N (# of observation)	Average of Y	S.D.	MAX	MIN
51.00	4.47	2.34	10.00	1.00

Independent Samples T-test



Independent Samples T-test

Independent Samples T-test

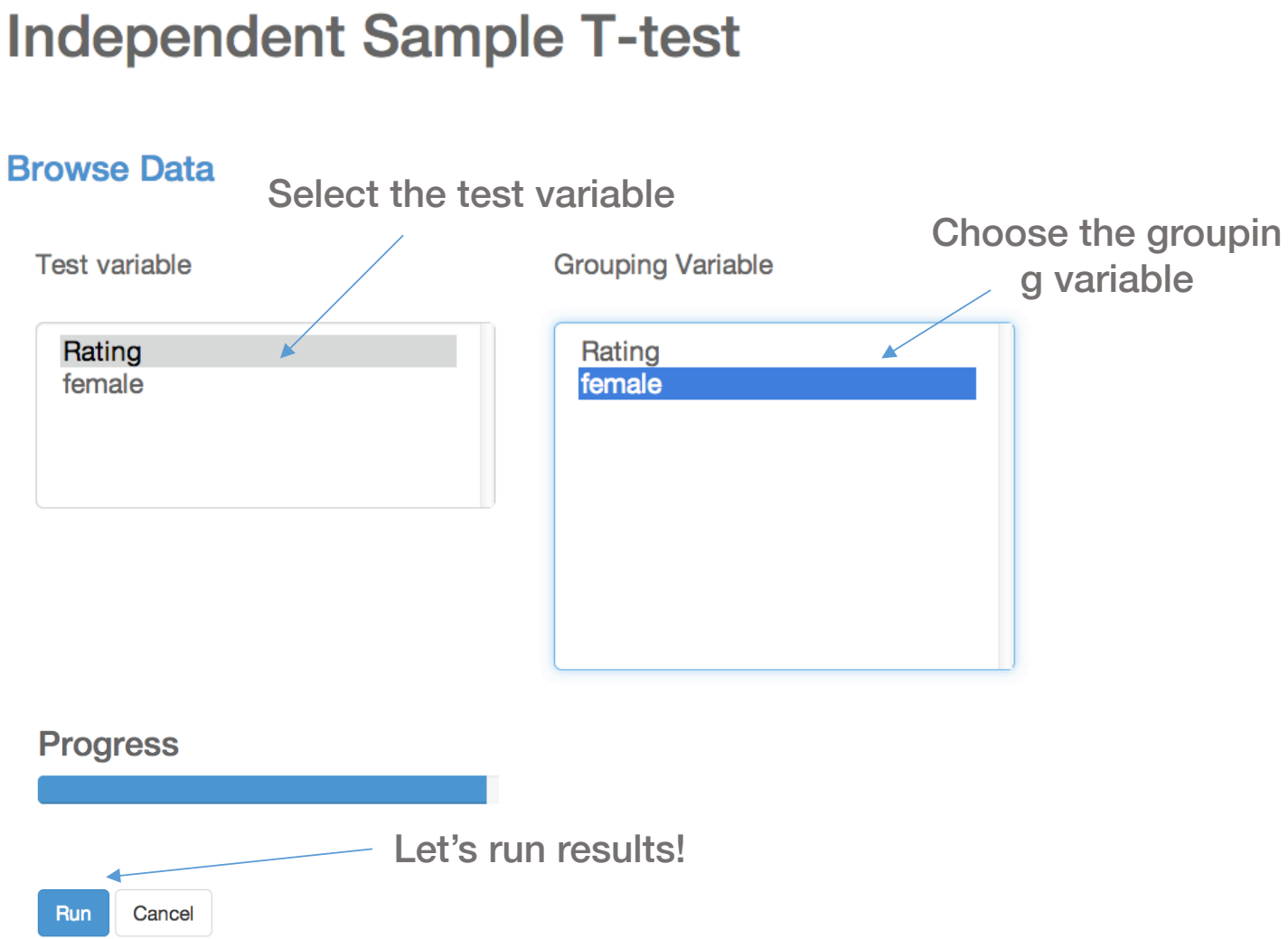


Please upload the Data file below

- (•) **Sample**
- ( ) **Your own file**

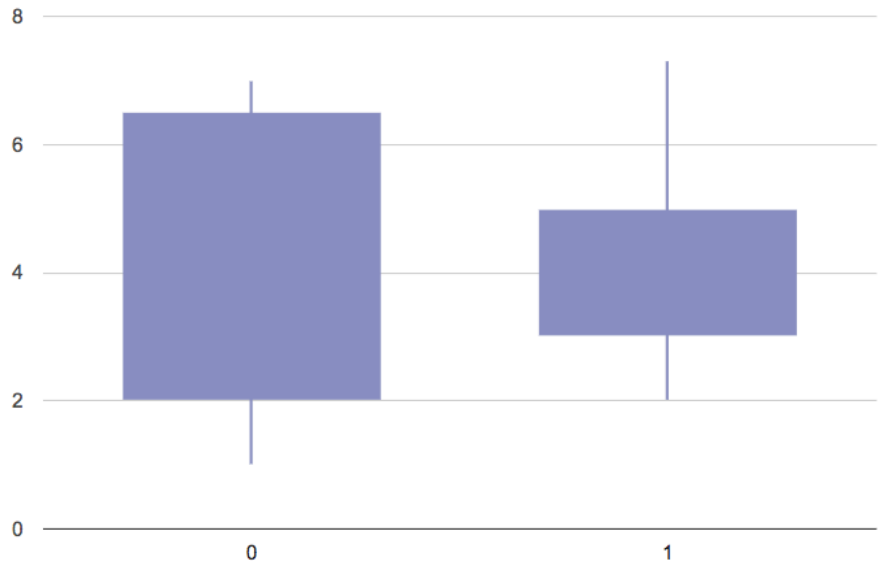
Now we're interested to know whether the rating for female group is significantly different from rating for male group.

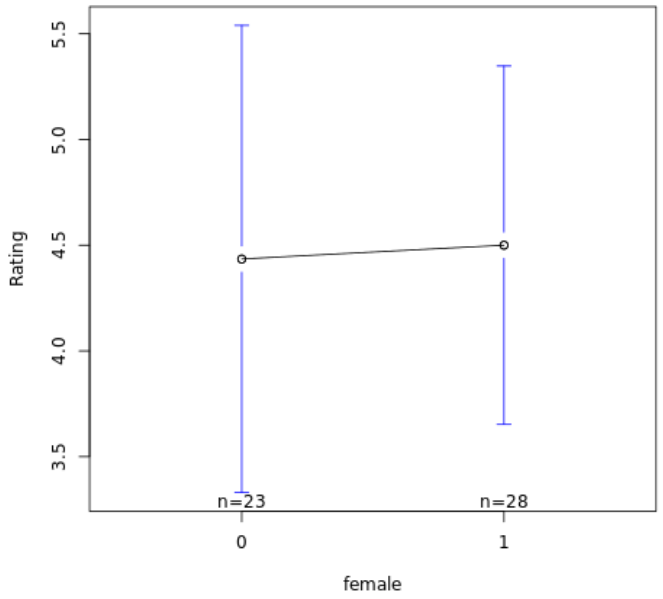
Independent Samples T-test



Independent Samples T-test

t value	p value	Equal variance test	Interpretation
0.09825	0.922	Equal variance supported	There is no significant difference by female on the means of Rating



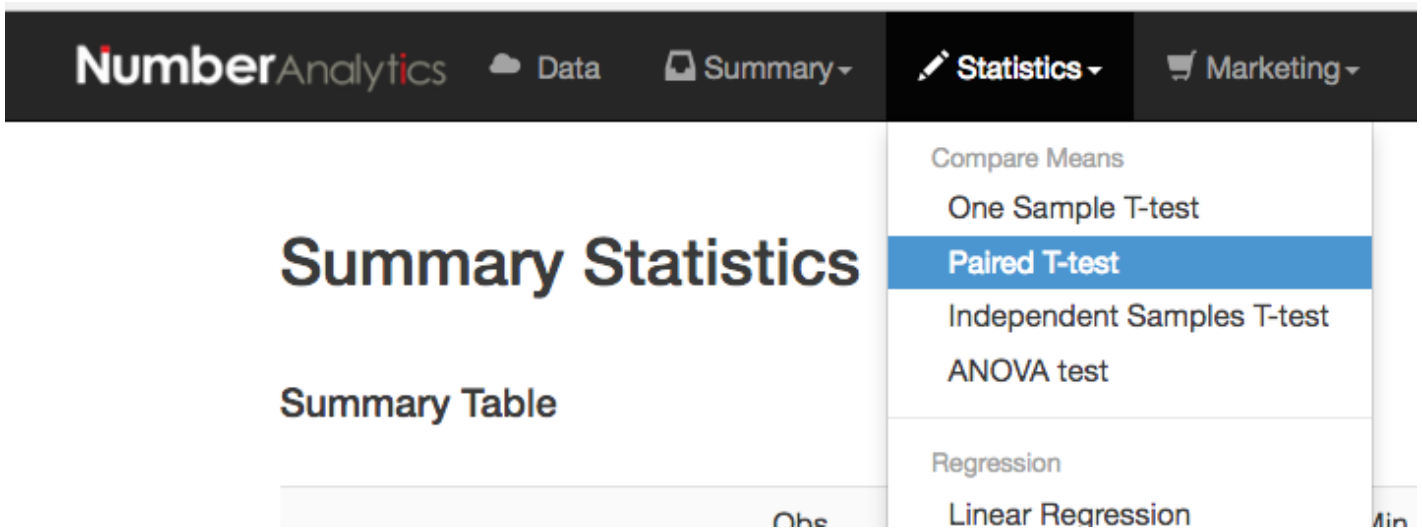


SPSS style output

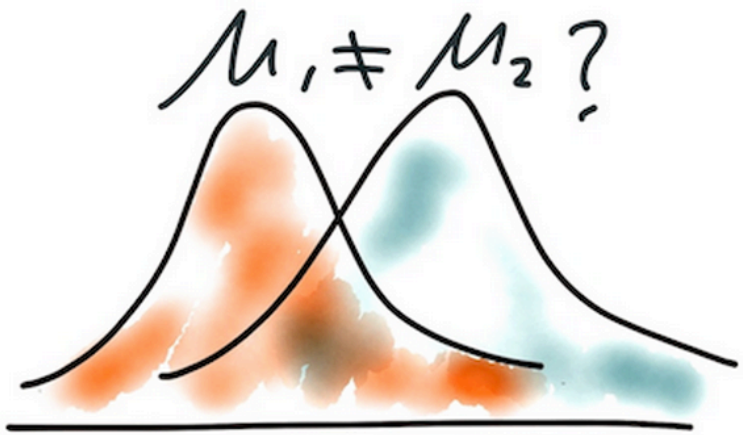
Independent Samples Test

		Levene's Test for Equal Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig (2-tailed)	Mean difference	Std. Error difference	Lower 95% Conf. Int.	Higher 95% Conf. Int.
Rating	Equal variances assumed	2.73	0.105	0.09825	49	0.922	0.0652174	0.663817	-1.26877	1.39921
	Equal variances not assumed			0.09673	43.569	0.923	0.0652174	0.67419	-1.2939	1.42434

Paired sample t-test



Paired Sample T-test



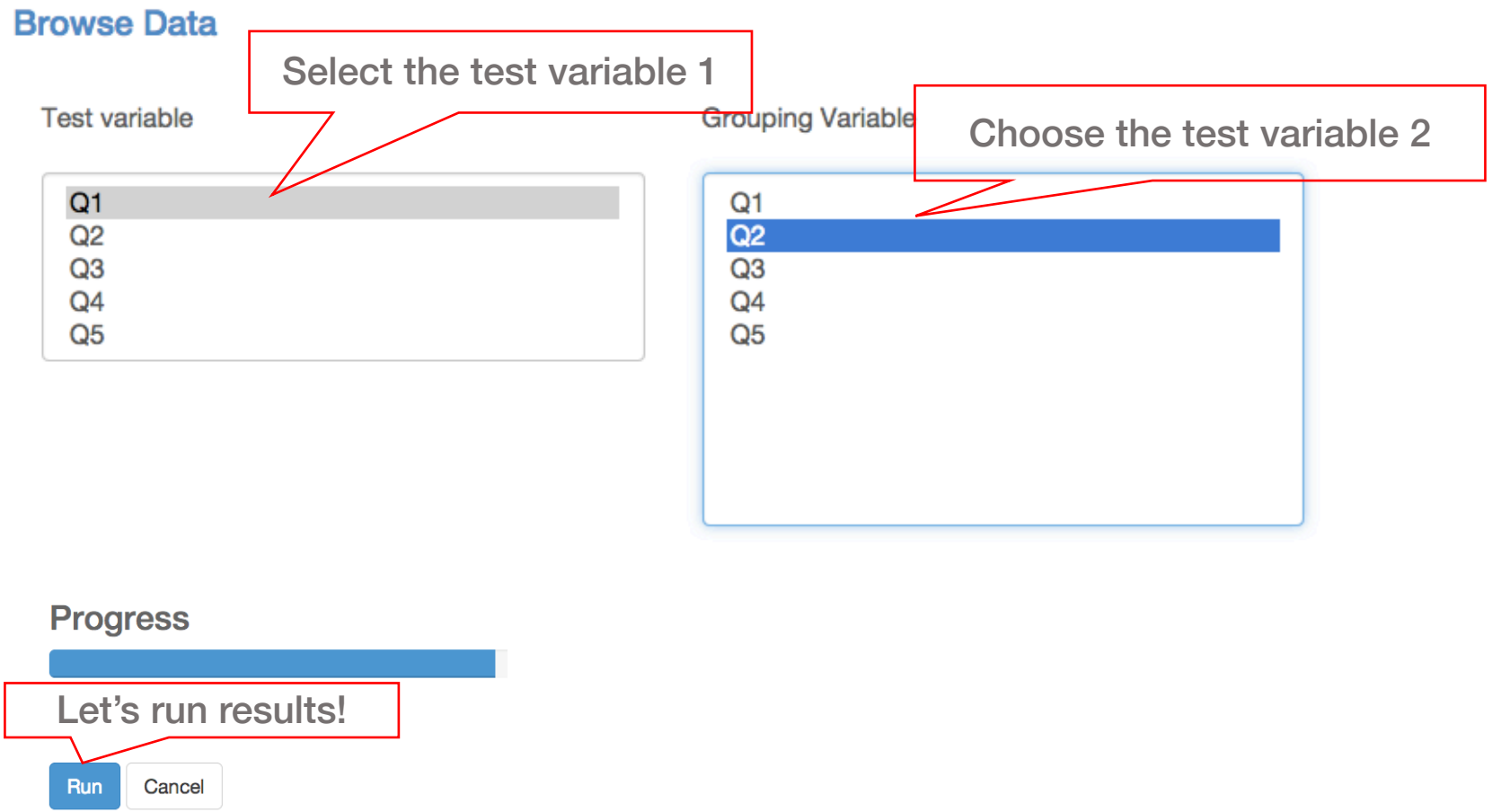
Please upload the Data file below

- (•) **Sample**
- ( ) **Your own file**

Unlike the independent samples T-test, paired sample T-test is applied to the same group comparing two outcomes like music preference rating scores across different songs

Independent Samples T-test

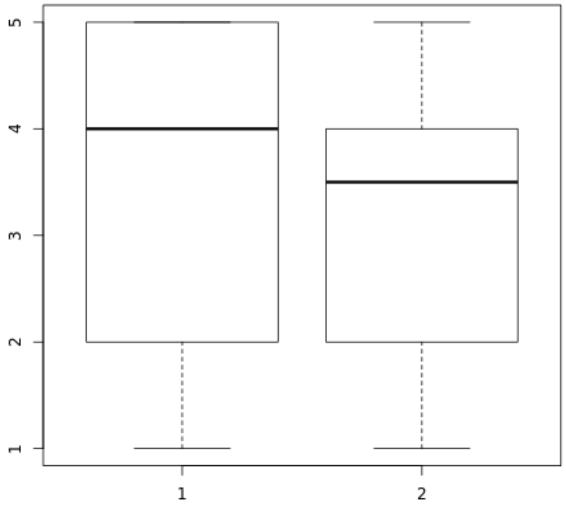
Paired Sample T-test



Independent Samples T-test

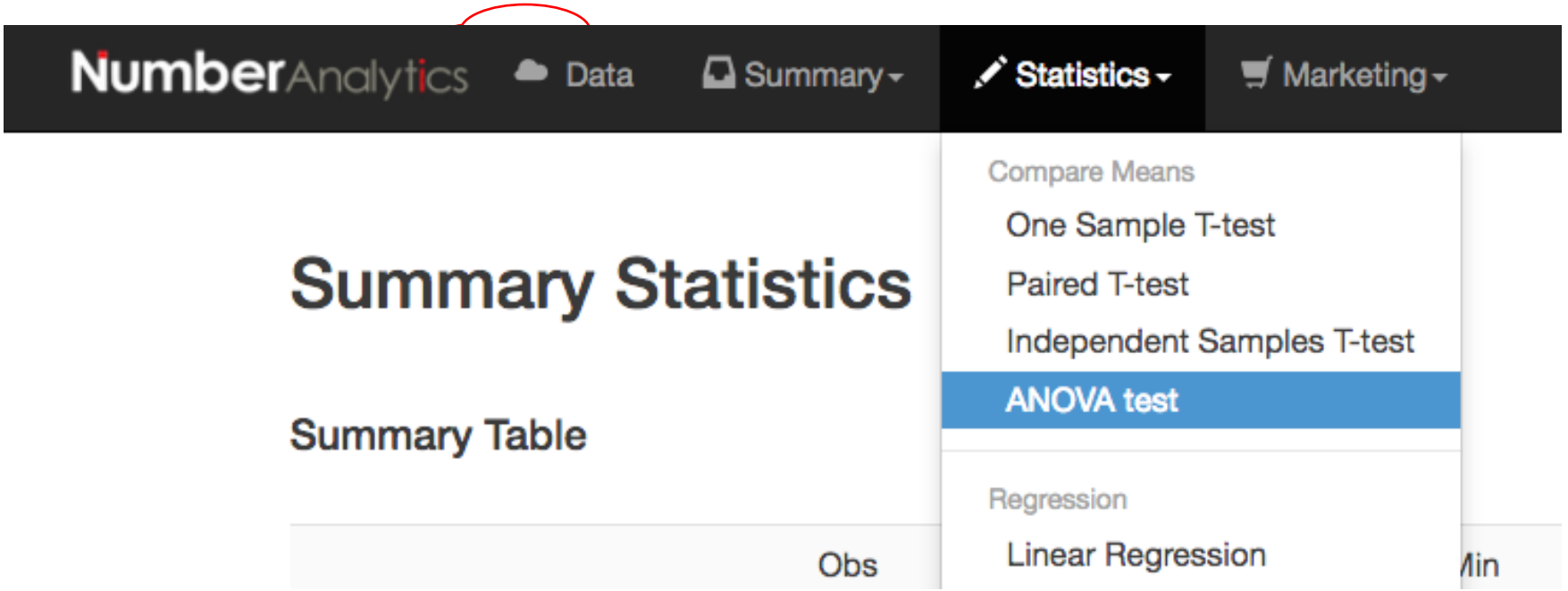
Paired Samples T-test

t_value	pvalue	
3.261	0.0013	There is a very significant difference on the means of Q1 and Q2 with confidence level 99%



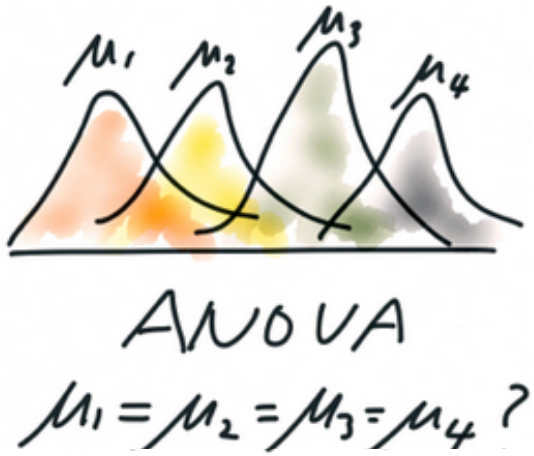
	Q1	Q2
Mean	3.467	3.234
S.D.	0.1039	0.1009

ANOVA


NumberAnalytics
Data
Summary
Statistics
Marketing
Compare Means
One Sample T-test
Paired T-test
Independent Samples T-test
ANOVA test
Regression
Linear Regression
Summary Statistics
Summary Table
Obs
Min

ANOVA

ANOVA (Analysis of Variance) Test



Please select data below

- **Sample**
- **Yours** (change)

Rating	Ethnicity
9	Japanese
4	Hispanic
3	Chinese

download

How about the rating among different ethnical groups?

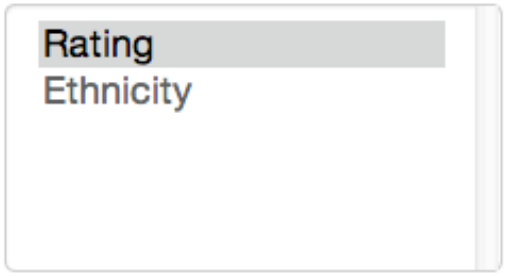
data_anova

	A	B
1	Rating	Ethnicity
2	9	Japanese
3	4	Hispanic
4	3	Chinese
5	6	Chinese
6	9	White
7	7	Korean
8	5	White
9	6	Hispanic
10	1	White
11	3	Other Asian

ANOVA difference test

Browse Data

Select a dependent variable



Select a grouping variable

Progress

Run Cancel

ANOVA

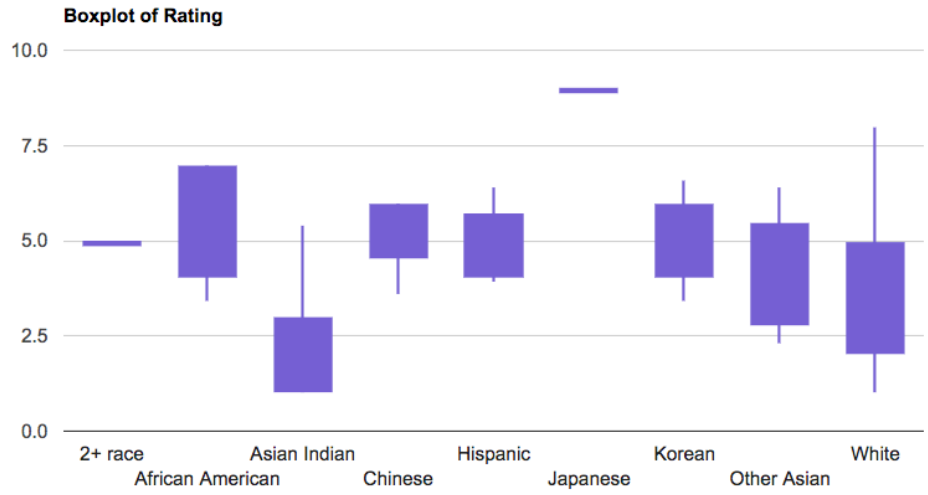
ANOVA Result

P value 0.32 (>0.05)

Conclusion: Rating is indifferent across ethnicity.

Fvalue	pvalue	
1.207	0.32	Rating is indifferent acrosss Ethnicity

P Value: Exact Probability of getting a computed test statistic that is due
to chance. The smaller the p value,
the smaller the probability that the
observed result occurred by chance

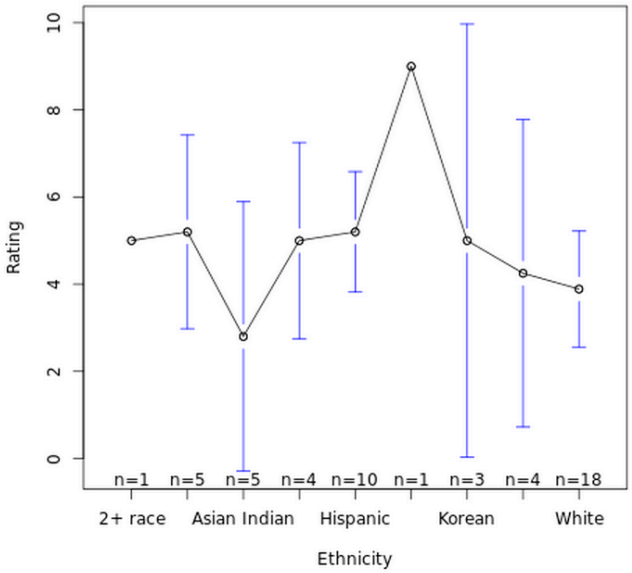


SPSS style output

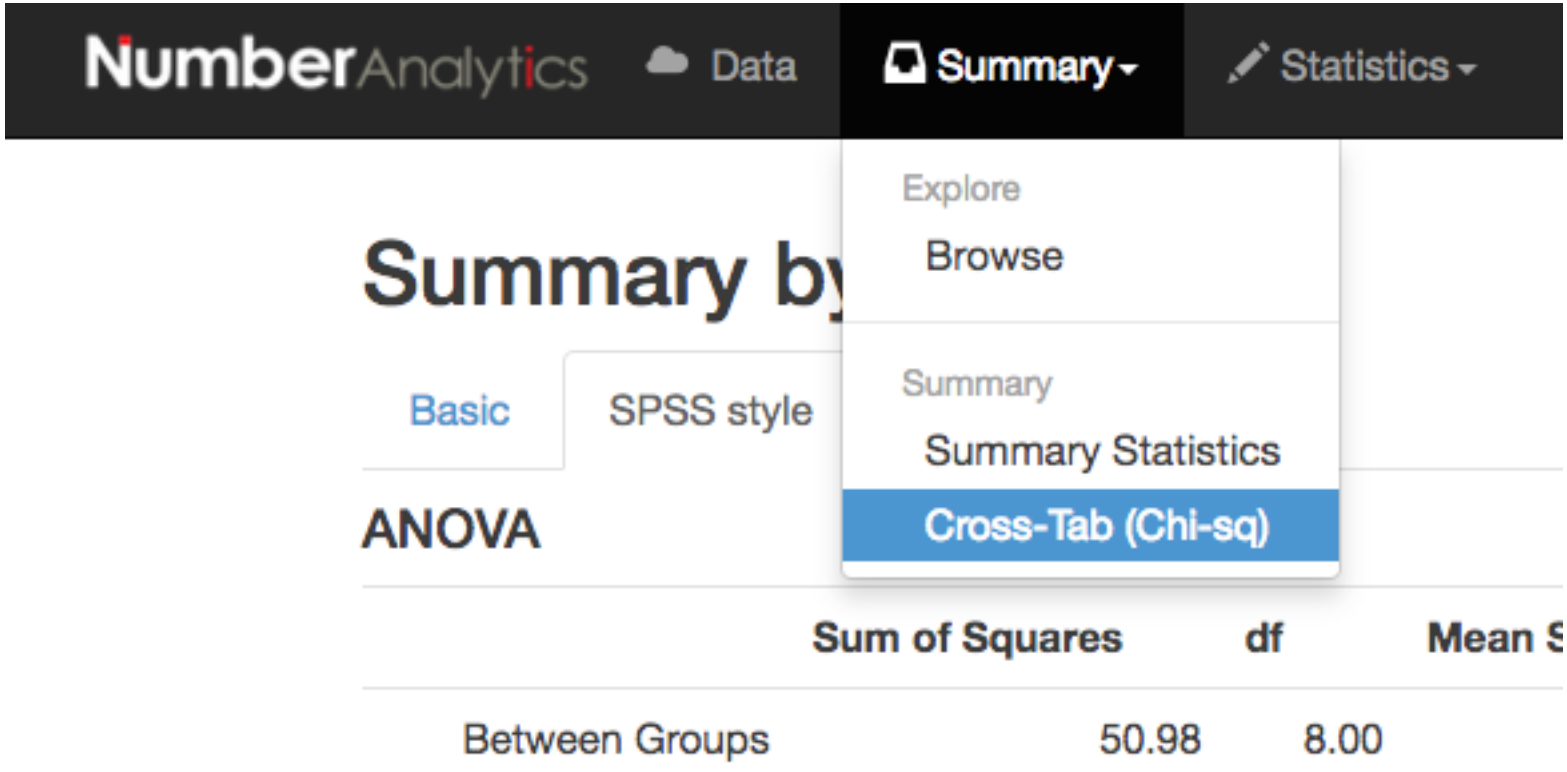
Basic | SPSS style | Warning

ANOVA

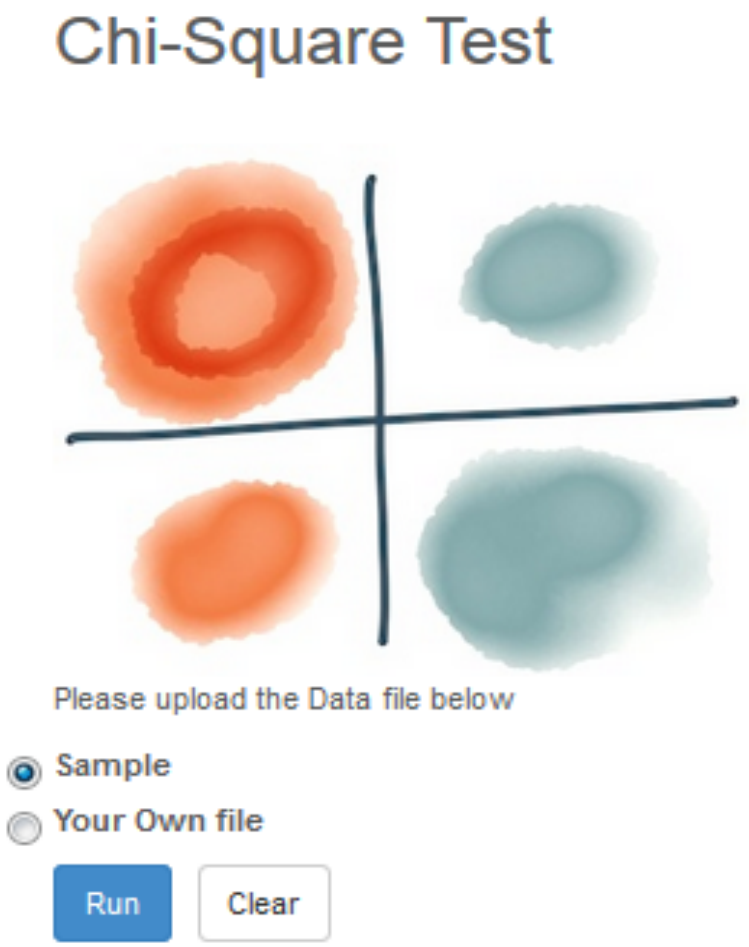
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	50.98	8.00	6.37	1.21	0.32
Within Groups	221.73	42.00	5.28		
Total	272.71	50.00			



Cross-tab (Chi-square test)



Cross-tab (Chi-square test)



- Cross-tab is a frequency table of two or three variables
- Used to examine association between two or 3 variables (usually 2)
- H_0: there is a relation between variable X and variable Y
- Variables take a limited number of values, for example:
  Consumers: gender, ethnicity
  Business: industry, company size

Cross-tab (Chi-square test)

Chi-Square Test

Browse Data

Select a Dependent Variable (y)

Ethnicity
female

Select Independent Variables (X)

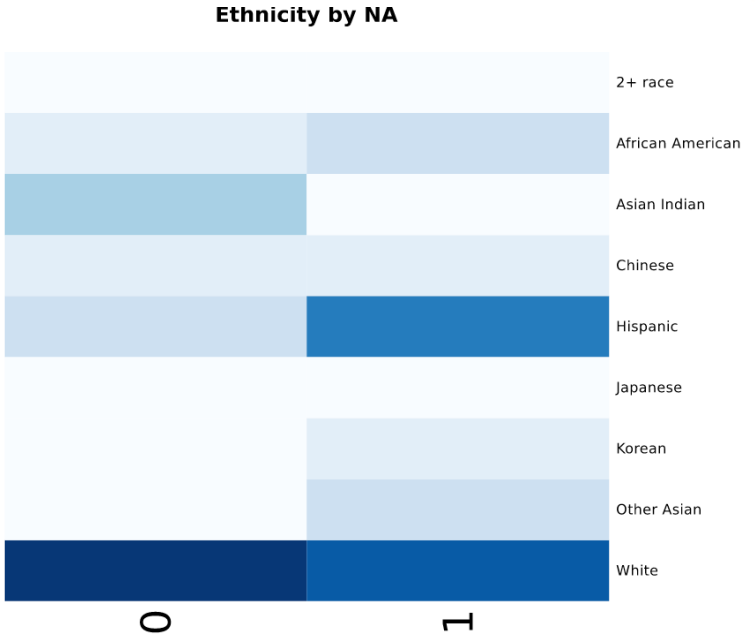
Ethnicity
female

Cross-tab (Chi-square test)

Chi-square test

Chisq_value	pvalue	
6.73	0.57	Ethnicity is indifferent acrosss female

P value greater than 0.05, reject H0.
Conclusion: There is no relation between gender and ethnicity



Crosstab table

FACTOR	0	1
2+ race	0	1
African American	2	3
Asian Indian	4	1
Chinese	2	2
Hispanic	3	7
Japanese	0	1
Korean	1	2
Other Asian	1	3
White	10	8