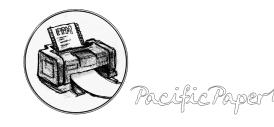
Chi-Squared Tests

These are simple tests that compare categorical variables... to other categorical variables. This is different from a t-test. There are no continuous variables here. Your independents and dependents are both categorical. For example, who is more likely to own a pet snake: men or women? Let's say you ask 200 men and 200 women. Within your 400-subject sample, the frequency of snake ownership among men is 2.5% and the ownership among women is 0.9%. That's a pretty big difference in your *sample*, but is it big enough to indicate a difference in the *population*? The chi-squared test is an inferential statistic. It is used to make inferences about the larger population.



Examples of Questions for Chi-Square Tests:

Do football players and baseball players differ in divorce rate?

Are cigarette smokers more likely to get lung cancer?

What is the difference in STD frequency between professional and amateur hockey players?

Do republicans and democrats differ in astrological sign?



	Total	Men	Women	Sig.
Ν	45	23	22	
Age (years)	20.2 ± 0.7	21.0 ± 0.5	22.3 ± 0.9	p=0.412
BMI (kg/m ²)	26.5 ± 2.8	27.4 ± 0.8	26.0 ± 0.6	p=0.518
GPA	2.9 ± 0.7	2.6 ± 1.0	3.2 ± 0.5	p = 0.048
Nightly sleep (hours)	7.5 ± 1.4	7.1 ± 1.2	7.6 ± 1.9	p=0.525
Employed (%)	22.5%			
Weekly work (hours)	4.0 ± 5.2	8.5 ± 4.1	2.3 ± 6.6	p=0.126
Academic Scholarship (%)	12.9%			
Athletic Scholarship (%)	11.5%			



	Total	Men	Women	Sig.
Ν	45	23	22	
Age (years)	20.2 ± 0.7	21.0 ± 0.5	22.3 ± 0.9	p = 0.412
BMI (kg/m ²)	26.5 ± 2.8	27.4 ± 0.8	26.0 ± 0.6	p = 0.518
GPA	2.9 ± 0.7	2.6 ± 1.0	3.2 ± 0.5	p = 0.048
Nightly sleep (hours)	7.5 ± 1.4	7.1 ± 1.2	7.6 ± 1.9	p = 0.525
Employed (%)	22.5%	\bigcirc	()	
Weekly work (hours)	4.0 ± 5.2	4.1	2.3 ± 6.6	p=0.126
Academic Scholarship (%)	12.9%	\bigcirc	\bigcirc	
Athletic Scholarship (%)	11.5%	\bigcirc	\bigcirc	



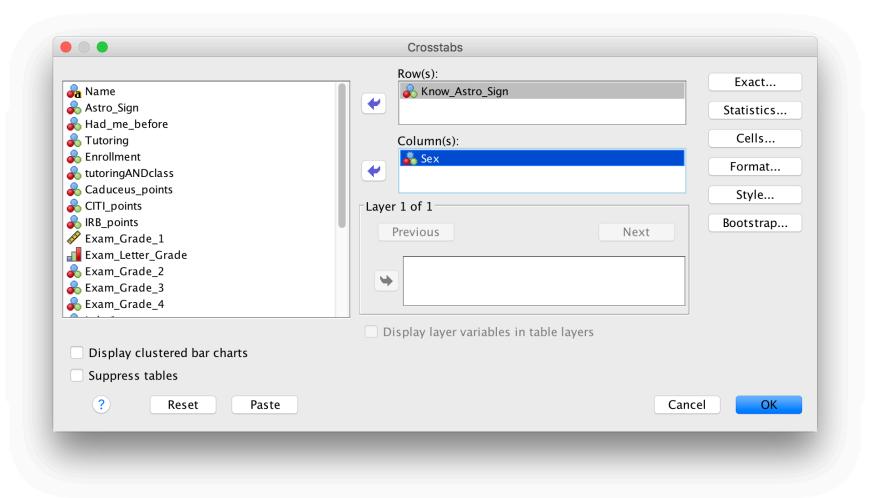


SPSS Statistics	File	Edit	View	Data	Transform	Analyze	Graphs	Utilities	Extensions	Window Help	
						Reports	5		►		
						Descrip	tive Statis	tics	•	123 Frequencies	
						Compar	re Means		•	Ha Descriptives	
						General	l Linear Mo	odel	•	4 Explore	
						General	lized Linea	r Models	•	🐺 Crosstabs	
						Mixed N	/lodels		•	+ TURF Analysis	
						Correlat				🔽 Ratio	
						Regress			•	P-P Plots	
						Loglinea			•	🛃 Q-Q Plots	
						Classify					
							ion Reduc	tion			
						Scale					
						-	ametric Te	ests			
						Forecas	-				
						Surviva					
							Response	5			
							ulation				
							Control				
							Curve	a wal b fa alal			
						Spatial	and temp	oral Model	ing 🕨		



	Row(s):	-
a Name		Exact
Know_Astro_Sign		Statistics
Astro_Sign Had_me_before	Column(s):	Cells
Sex		
b Tutoring		Format
Enrollment		Style
tutoringANDclass	Layer 1 of 1	Bootstrap
CITI_points	Previous	
RB_points		
Exam_Grade_1 Exam_Letter_Grade		
Exam_Grade_2		
	Display layer variables in table layers	1
Display clustered bar charts		
Suppress tables		
suppress lables		
? Reset Paste	Car	ncel OK







	CI	osstabs		
	🔴 🔘 🔴 Crosstabs: Sta	tistics		Exact
 Name Astro_Sign Had_me_before Tutoring Enrollment tutoringANDclass Caduceus_points CITI_points IRB_points Exam_Grade_1 Exam_Grade_2 Exam_Grade_3 Exam_Grade_4 	 Chi-square Nominal Contingency coefficient Phi and Cramer's V Lambda Uncertainty coefficient Nominal by Interval Eta 	Correlations Ordinal Gamma Somers' d Kendall's tau-b Kendall's tau-c Kappa Risk McNemar	Next	Exact Statistics Cells Format Style Bootstrap
 Display clustered bar cha Suppress tables ? Reset 	Cochran's and Mantel-Haer Test common odds ratio equ ?Canc	ials: 1	ayers Canc	el OK



	Cr	osstabs		
	🔴 🔵 🌒 Crosstabs: Sta	tistics		Evect
 Name Astro_Sign Had_me_before Tutoring Enrollment tutoringANDclass Caduceus_points CITI_points IRB_points Exam_Grade_1 Exam_Grade_2 Exam_Grade_3 Exam_Grade_4 	 Chi-square Nominal Contingency coefficient Phi and Cramer's V Lambda Uncertainty coefficient Nominal by Interval Eta 	Correlations Ordinal Gamma Somers' d Kendall's tau-b Kendall's tau-c Kappa Risk McNemar	Next	Exact Statistics Cells Format Style Bootstrap
 Display clustered bar cha Suppress tables Reset 	Cochran's and Mantel-Haer Test common odds ratio equ ? Can	ials: 1	ayers	el OK



••		Crosstabs: Cell Display	
a Name	Counts	z-test	Exact
Astro_Sign	🗹 Observed	Compare column proportions	Statistics
Had_me_before	Expected	Adjust p-values (Bonferroni method)	
Tutoring	Hide small counts		Cells
Enrollment			Format
tutoringANDclass	Less than 5		Tormat
Caduceus_points	Demonstration	Desiduala	Style
CITI_points IRB_points	Percentages	Residuals	Destatues
Exam_Grade_1	Row	Unstandardized	Bootstrap
Exam_Letter_Grade	Column	Standardized	
Exam_Grade_2	🗌 Total	Adjusted standardized	
Exam_Grade_3			
Exam_Grade_4	Noninteger Weights		
	Round cell counts	O Round case weights	
Display clustered bar charts	Truncate cell counts	s 🔵 Truncate case weights	
	🔵 No adjustments		
Suppress tables			
? Reset Pa	?	Cancel	cel OK
	\odot		



••		Crosstabs: Cell Display	
Name	Counts	z-test	Exact
Astro_Sign	Observed	Compare column proportions	Statistics
Had_me_before	Expected	Adjust p-values (Bonferroni method)	Statistics
Tutoring		August p values (bonnehonn method)	Cells
Enrollment	Hide small counts		
tutoringANDclass	Less than 5		Format
Caduceus_points			Style
CITI_points	Percentages	Residuals	Stylem
RB_points	Row	Unstandardized	Bootstrap
Exam_Grade_1	Column	Standardized	
Exam_Letter_Grade			
Exam_Grade_2	🗌 Total	Adjusted standardized	
Exam_Grade_3			
Exam_Grade_4	Noninteger Weights		
	Round cell counts	Round case weights	
Disular, directored have showed	Truncate cell counts	5 🔘 Truncate case weights	
Display clustered bar charts	No adjustments		
Suppress tables	ino aujustiments		
? Reset Pa	?	Cancel	cel OK
	\cdot	cuncer	



Crosstabs

Know_Astro_Sign * Sex Crosstabulation

			Sex		
			Male	Female	Total
Know_Astro_Sign	Does not know	Count	3	5	8
		% within Sex	33.3%	20.8%	24.2%
	Knows	Count	6	19	25
		% within Sex	66.7%	79.2%	75.8%
Total		Count	9	24	33
		% within Sex	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.557 ^a	1	.456		
Continuity Correction ^b	.084	1	.772		
Likelihood Ratio	.534	1	.465		
Fisher's Exact Test				.651	.374
Linear-by-Linear Association	.540	1	.462		
N of Valid Cases	33				



Pacific Paper Mill.com

	Total	Men	Women	Sig.
Ν	45	23	22	
Age (years)	20.2 ± 0.7	21.0 ± 0.5	22.3 ± 0.9	p = 0.412
BMI (kg/m ²)	26.5 ± 2.8	27.4 ± 0.8	26.0 ± 0.6	p=0.518
GPA	2.9 ± 0.7	2.6 ± 1.0	3.2 ± 0.5	p = 0.048
Nightly sleep (hours)	7.5 ± 1.4	7.1 ± 1.2	7.6 ± 1.9	p=0.525
Employed (%)	22.5%	24.0%	19.7%	p = 0.099
Weekly work (hours)	4.0 ± 5.2	8.5 ± 4.1	2.3 ± 6.6	p=0.126
Academic Scholarship (%)	12.9%	13.2%	12.6%	p = 0.884
Athletic Scholarship (%)	11.5%	11.9%	11.3%	p=0.768



Chi-Squared Value. Two binary measurements. What's the distribution between these? Is it the same percentage in each group (expected) or is there a difference between groups (observed)?

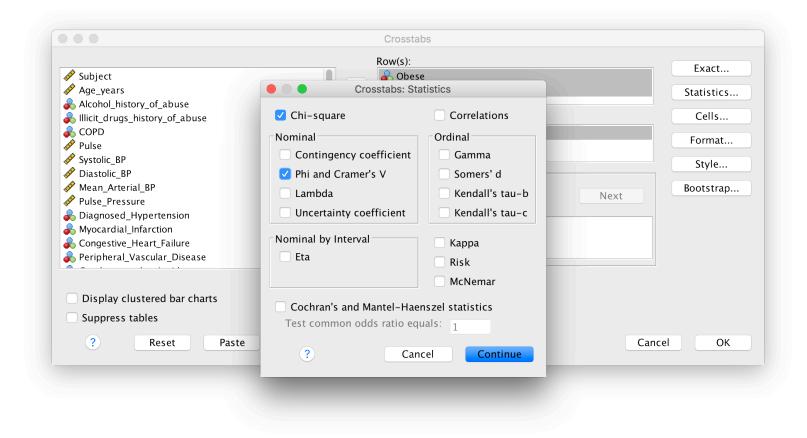


Subject	Row(s):	Exact
Sex		Statistics
Age_years 0 Obese	Calumn(a)	Cells
Alcohol_history_of_abuse	Column(s):	Cents
lllicit_drugs_history_of_abuse		Format
Smoker		Style
COPD	Layer 1 of 1	
Pulse	Previous	Bootstrap
Systolic_BP Diastolic_BP		
Mean_Arterial_BP		
Pulse_Pressure		
Diagnosed_Hypertension		
··· ··· · · ·	Display layer variables in table layers	
Display clustered bar charts		
Suppress tables		
? Reset Paste	Can	cel OK



A	Row(s):	Exact
Subject Age_years	Chese Smoker	
Age_years		Statistics
lllicit_drugs_history_of_abuse	Column(s):	Cells
COPD	Sex .	
🖻 Pulse		Format
Systolic_BP		Style
Diastolic_BP	Layer 1 of 1	
Mean_Arterial_BP	Previous	Bootstrap
Pulse_Pressure Diagnosed_Hypertension		
Myocardial_Infarction		
Congestive_Heart_Failure		
Peripheral_Vascular_Disease		
	Display layer variables in table layers	
Display clustered bar charts		
Suppress tables		
? Reset Paste	Cano	el OK
? Reset Paste	Canc	UK UK

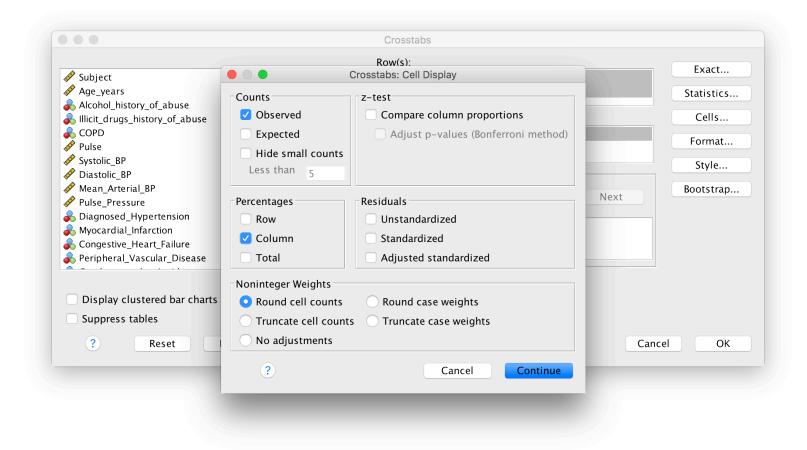






🔗 Subject	Row(s):	Exact
Age_years	Smoker	Statistics
Alcohol_history_of_abuse		
lllicit_drugs_history_of_abuse	Column(s):	Cells
COPD	Sex	Format
Pulse		
Systolic_BP Diastolic_BP	Layer 1 of 1	Style
Mean_Arterial_BP		Bootstrap
Pulse_Pressure	Previous	Dootstrupm
biagnosed_Hypertension		
b Myocardial_Infarction		
Congestive_Heart_Failure		
Peripheral_Vascular_Disease		
	Display layer variables in table layers	
Display clustered bar charts		
Suppress tables		
Suppress tables		
? Reset Paste	Cano	cel OK







Subject	Row(s):	Exact
Age_years	Smoker	Statistics
Alcohol_history_of_abuse		Statistics
lllicit_drugs_history_of_abuse	Column(s):	Cells
COPD	A Sex	
Pulse		Format
Systolic_BP		Style
Diastolic_BP	Layer 1 of 1	
Mean_Arterial_BP	Previous	Bootstrap
Pulse_Pressure		
Diagnosed_Hypertension Myocardial_Infarction		
Congestive_Heart_Failure	↔	
Peripheral_Vascular_Disease		
	Display layer variables in table layers	
Display clustered bar charts		
Suppress tables		
? Reset Paste	Cano	cel OK



Obese * Sex

Crosstab

			Sex				
			Male	Female	Total		
Obese	0	Count	1080	849	1929		
		% within Sex	83.7%	83.6%	83.7%		
	1	Count	211	166	377		
		% within Sex	16.3%	16.4%	16.3%		
Total		Count	1291	1015	2306		
		% within Sex	100.0%	100.0%	100.0%		

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	.000	.994
	Cramer's V	.000	.994
N of Valid Cases		2306	

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.000 ^a	1	.994		
Continuity Correction ^b	.000	1	1.000		
Likelihood Ratio	.000	1	.994		
Fisher's Exact Test				1.000	.519
Linear-by-Linear Association	.000	1	.994		
N of Valid Cases	2306				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 165.94.

b. Computed only for a 2x2 table



Pacific Paper Mill.com

Smoker * Sex

Crosstab

			Se		
			Male	Female	Total
Smoker	0	Count	881	831	1712
		% within Sex	68.2%	81.9%	74.2%
	1	Count	410	184	594
		% within Sex	31.8%	18.1%	25.8%
Total		Count	1291	1015	2306
		% within Sex	100.0%	100.0%	100.0%

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	155	.000
	Cramer's V	.155	.000
N of Valid Cases		2306	

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	55.204 ^a	1	.000		
Continuity Correction ^b	54.493	1	.000		
Likelihood Ratio	56.539	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	55.180	1	.000		
N of Valid Cases	2306				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 261.45.

b. Computed only for a 2x2 table



Pacific Paper Mill.com

When examining the association between two dichotomous variables, the correlation is called Phi.

When examining the association between two categorical variables that have more than two categories, Cramer's V is the appropriate statistic.

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	155	.000
	Cramer's V	.155	.000
N of Valid Cases		2306	

If you want the correlation of one continuous variable to another continuous variable, use the Pearson correlation coefficient (bivariate correlation table); if you want the correlation of a continuous variable and a dichotomous variable, that's a pointbiserial correlation (also bivariate table).



Remember: Chi-squared tests aren't controlling
for anything. What is the difference in lung cancer
frequency among people who do and do not keep
matches in their pockets? Let's say the chi-square
is significant: there's a five-fold increase in the
frequency of lung cancer in the matches-in-pocket
group (p=0.002). Does that mean matches cause
lung cancer (at a 99.8% confidence level)? Or is
causality a different question altogether?NN

	Total	Men	Women	Sig.
N	45	23	22	
Age (years)	20.2 ± 0.7	21.0 ± 0.5	22.3 ± 0.9	p = 0.412
BMI (kg/m ²)	26.5 ± 2.8	27.4 ± 0.8	26.0 ± 0.6	p = 0.518
GPA	2.9 ± 0.7	2.6 ± 1.0	3.2 ± 0.5	p = 0.048
Nightly sleep (hours)	7.5 ± 1.4	7.1 ± 1.2	7.6 ± 1.9	p = 0.525
Employed (%)	22.5%	24.0%	19.7%	p = 0.099
Weekly work (hours)	4.0 ± 5.2	8.5 ± 4.1	2.3 ± 6.6	p = 0.126
Academic Scholarship (%)	12.9%	13.2%	12.6%	p = 0.884
Athletic Scholarship (%)	11.5%	11.9%	11.3%	p = 0.768

