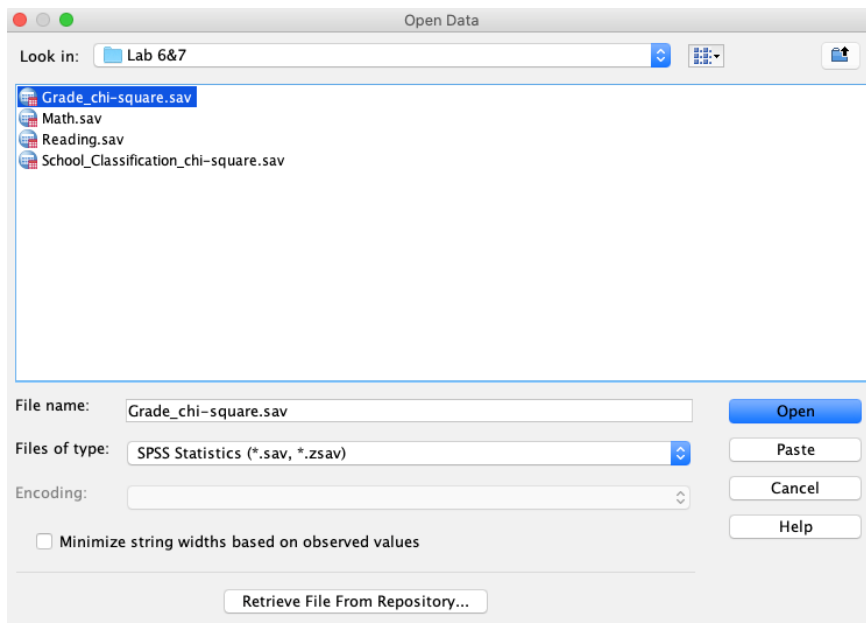
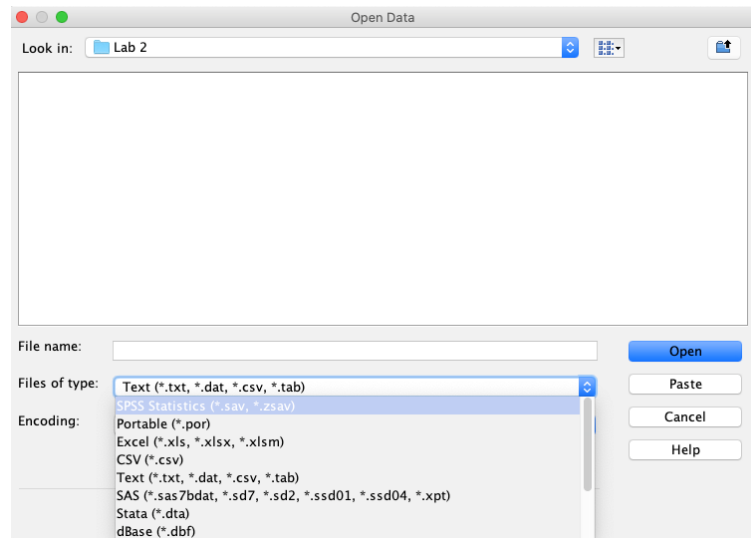
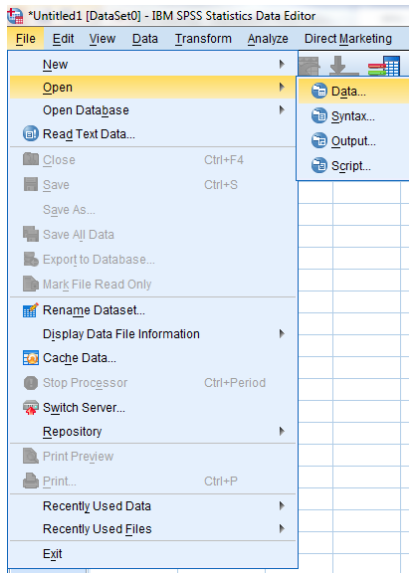


Chi-Square Test and Correlation in SPSS

Dataset: Grade_Chi-square.sav

1. Open/Import/Read Dataset: File -> Open -> Data. Select **All Files** from Files of type to show the file you're looking for. Select the correct file and click **Open**.



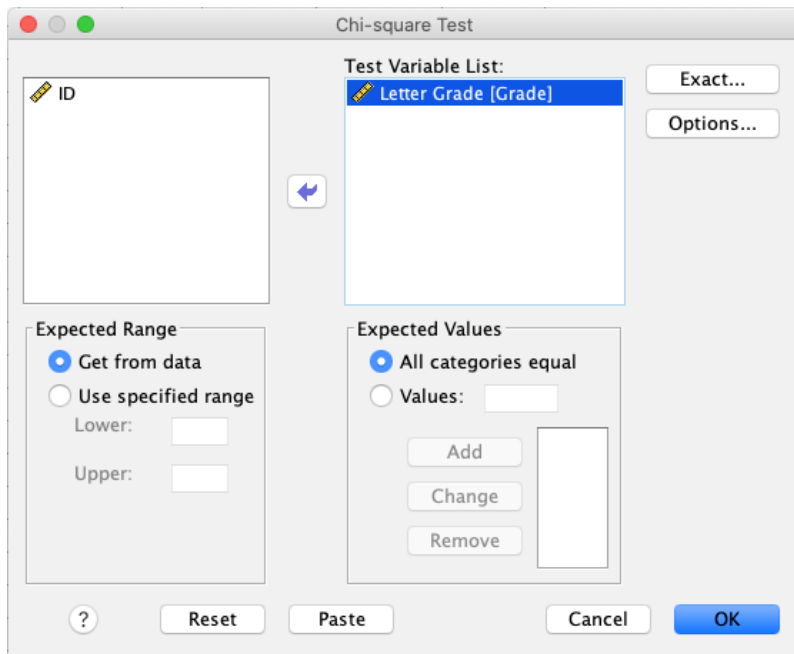
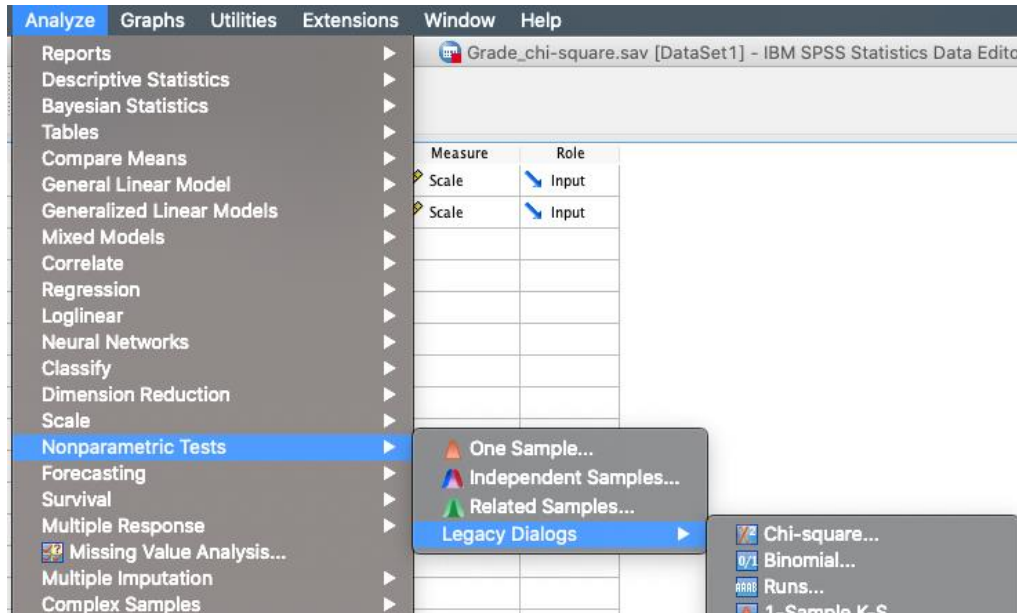
2. Goodness of Fit. Test whether students' GPA were equally distributed across five letter grades.

State the null and alternative hypotheses


H₀: Students' GPA were equally distributed across five letter grades.

H₁: Students' GPA were not equally distributed across five letter grades.

Analyze -> Nonparametric Tests -> Legacy Dialogs -> Chi-square. Move variable "Letter Grade" to Test Variable List.



Take a screenshot of your syntax and paste it here:


NPAR TESTS
 /CHISQUARE=Grade
 /EXPECTED=EQUAL
 /MISSING ANALYSIS.

Run the test (highlight the syntax and click the green triangle button). Take a screenshot of the output and paste it below:

Chi-Square Test

Frequencies

Letter Grade			
	Observed N	Expected N	Residual
A	18	22.0	-4.0
B	24	22.0	2.0
C	40	22.0	18.0
D	19	22.0	-3.0
F	9	22.0	-13.0
Total	110		

Test Statistics	
	Letter Grade
Chi-Square	23.727 ^a
df	4
Asymp. Sig.	.000

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 22.0.

What do your results indicate?

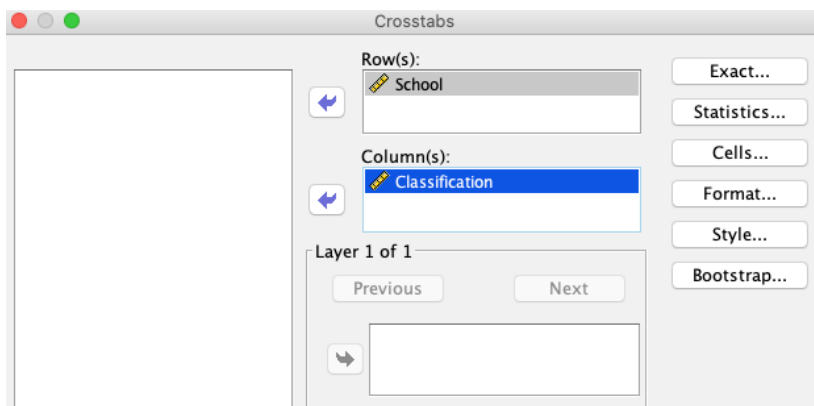
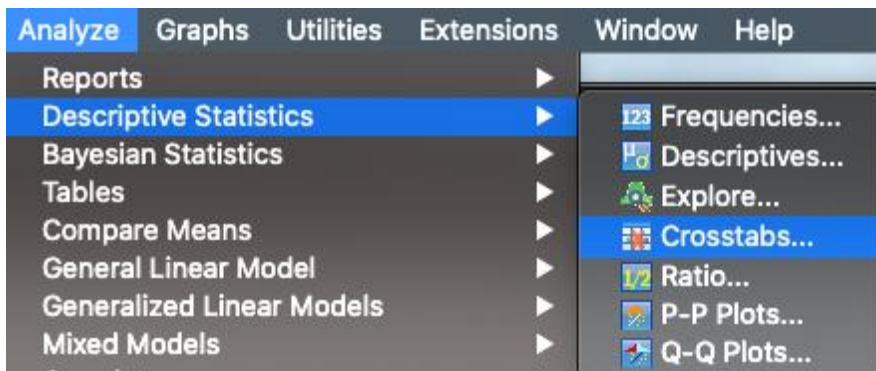
The results indicate that the students' GPA are not equally distributed across the five letter grades. The p value was <.000.

3. Test for Independence Open the dataset **School_Type** and determine if there is a significant relationship between school types and classifications.

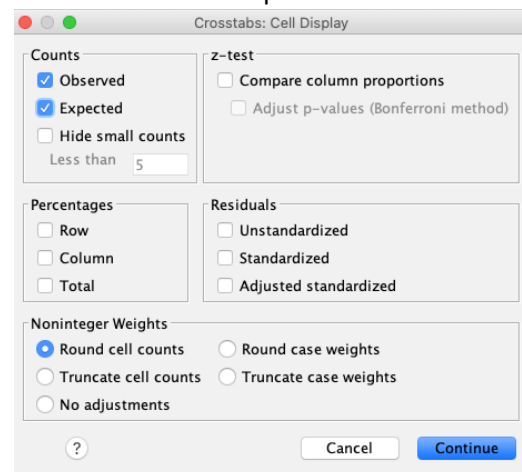
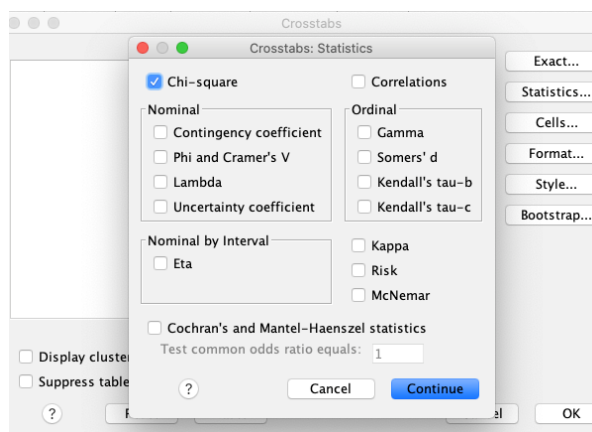
School type: 1 = school 1, 2 = school 2, 3 = school 3

Classification: 1 = Remedial, 2 = Passable, 3 = Proficient, 4 = Outstanding

Analyze -> Descriptive Statistics -> Crosstabs. Move a variable to Row and the other variable to Column.



Click statistics... => Check Chi-square -> Click continue. Click Check Expected -> Continue -> Paste.



Screenshot your Syntax and results below:

CROSSTABS

```
/TABLES=School BY Classification
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ
/CELLS=COUNT EXPECTED
/COUNT ROUND CELL.
```

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
School * Classification	243	100.0%	0	0.0%	243	100.0%

School * Classification Crosstabulation

			Classification				Total
			1	2	3	4	
School	1	Count	7	23	24	11	65
		Expected Count	5.1	20.3	22.2	17.4	65.0
	2	Count	7	19	18	14	58
		Expected Count	4.5	18.1	19.8	15.5	58.0
	3	Count	5	34	41	40	120
		Expected Count	9.4	37.5	41.0	32.1	120.0
Total	Count	19	76	83	65	243	
	Expected Count	19.0	76.0	83.0	65.0	243.0	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	9.585 ^a	6	.143
Likelihood Ratio	9.954	6	.127
Linear-by-Linear Association	7.415	1	.006
N of Valid Cases	243		

a. 1 cells (8.3%) have expected count less than 5. The minimum expected count is 4.53.

What do your results mean?

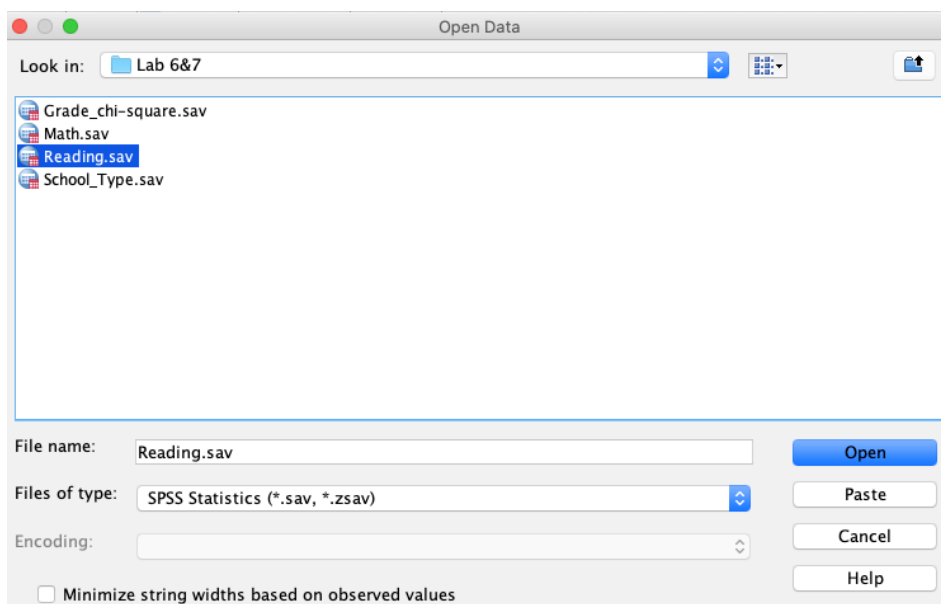
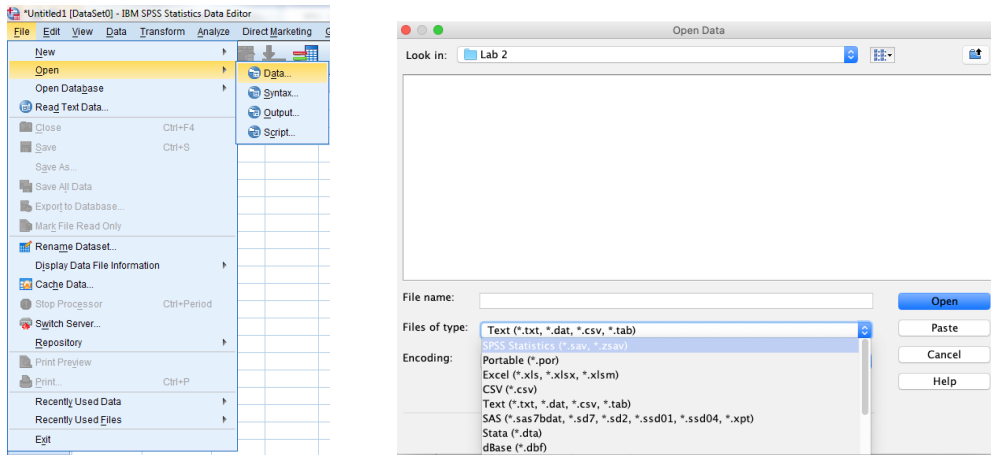
The p-value of Pearson Chi-square being .143 which is greater than alpha indicates that there is no significant relationship between school type and classifications.

How does the test for independence relate to correlation?

Both tests are used to determine whether there is a relationship between two variables. The test for independence is used for categorical variables while correlation is used for continuous variables.

Dataset: Reading.sav

4. Open/Import/Read Dataset: File -> Open -> Data. Select **All Files** from Files of type to show the file you're looking for. Select the correct file and click **Open**.



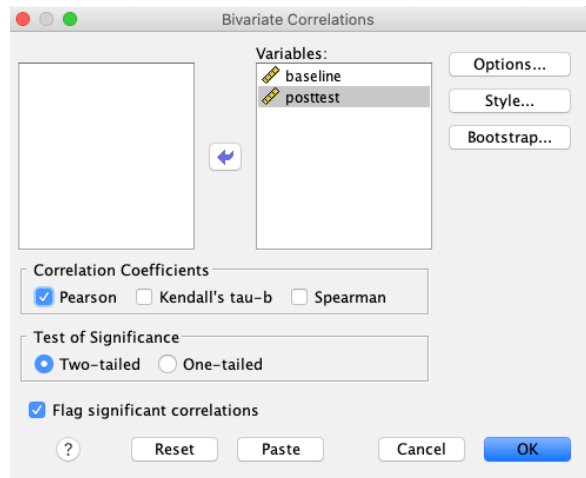
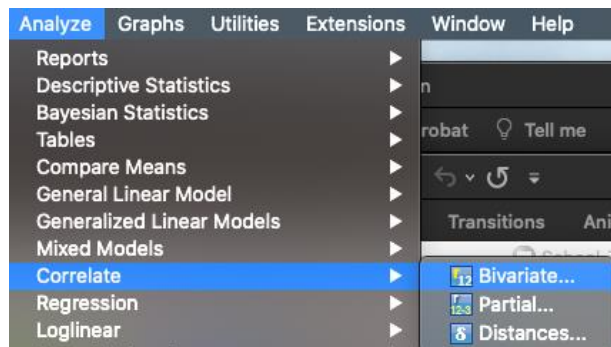
5. Correlation. Test whether there is a statistically significant correlation between students' baseline test score and their posttest score.

State the null and alternative hypotheses

H₀: There is no significant correlation between students' baseline test score and their posttest score

H₁: There is significant correlation between students' baseline test score and their posttest score

Analyze -> Correlate -> Bivariate -> Move variables to Variables box -> Check Pearson -> Paste.



Take a screenshot of your syntax and paste it here:

```

CORRELATIONS
/VARIABLES=baseline posttest
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
  
```

Run the test (highlight the syntax and click the green triangle button). Take a screenshot of the output and paste it below:

Correlations			
		baseline	posttest
baseline	Pearson Correlation	1	.074
	Sig. (2-tailed)		.461
	N	100	100
posttest	Pearson Correlation	.074	1
	Sig. (2-tailed)	.461	
	N	100	100

What do your results indicate (explain in your own words)?

From the results, there is no significant correlation between students' baseline test score and their posttest score.

APA write up: To assess the linear relationship between students' baseline score and their posttest score, a Pearson correlation coefficient was computed. Based on the Pearson correlation coefficient, there is not a significant correlation between baseline score and posttest score ($r(98) = .074, p = .461$).

Practice:

Dataset: Math.sav

6. Do a Pearson correlation test to evaluate whether there is a statistically significant correlation between time spent on math homework and math achievement score.

State the null and alternative hypotheses

H₀: There is no statistically significant correlation between time spent on math homework and math achievement score.

H₁: There is statistically significant correlation between time spent on math homework and math achievement score.

Screenshot your Syntax and results below:

```
CORRELATIONS
/VARIABLES=mathhome mathach
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

Correlations

		Time Spent on Math Homework per Week	Math Achievement Test Score
Time Spent on Math Homework per Week	Pearson Correlation	1	.320**
	Sig. (2-tailed)		.001
	N	100	100
Math Achievement Test Score	Pearson Correlation	.320**	1
	Sig. (2-tailed)	.001	
	N	100	100

** . Correlation is significant at the 0.01 level (2-tailed).

What is the effect size? (hint: you'll need to calculate it by hand; r^2)

$$r^2 = 0.1024$$

Using APA format, write up your results. Be sure to include a statement regarding the direction of the relationship (if there is one), your effect size, and an interpretation of the effect size.

From the results of Pearson correlation test above, there is a significant relationship between the time spent on Math Homework per week and math achievement test score ($r(98) = .320$, $p = 0.001$). The relationship is positive given the r is positive and with the effect size of 0.1024 the relationship is small.