



JMP Instructional data set 017

Parents' Opinions of a New School Food Program

JMP platforms and statistical methods

Graph Builder:	Mosaic plot
Distribution:	Bar charts; confidence intervals for proportions
Categorical:	Contingency table chi-square analysis; multiple comparisons; confidence intervals for proportions
Formula Editor:	Creating a new variable

Problem statement

A public school district has proposed a new food program aimed at improving student nutrition by providing more meals and healthy snacks throughout the school day. The program would include free breakfast, lunch, and an afternoon snack for all students, regardless of their parents' economic level. The district believes these changes could improve focus, reduce hunger, and support better academic outcomes. However, funding the food program would require cuts in other school programs, such as after-school activities and clubs.



To assess support for such a program, a survey was distributed to parents of children at three elementary schools in the district. One of the primary survey questions asked if parents are in favor of the program with only two possible answers (In Favor and Opposed). The response rate was about 60%.

The analyses conducted will focus on summarizing the opinions of the responders, conducting statistical tests to assess differences across the three schools, determining if parents' opinions depend on how many children they have in school.

Data

IDS_017.jmp

School	The school that responders' children are attending (Hawthorne, Eastwood, North Canyon)
Children	Number of their children attending the school (One Child, Two Children, 3+ Children)
Responders	Number of completed surveys received for each category of school and children
Opinion	Two possible answers ('In Favor' , 'Opposed')
Frequency	Number of responders for each of the two possible answers

Note: The data is provided in summarized format and will require that to be specified as such when running analyses.

Exercises

1. Create a graphical and numerical summary of the opinions of the responders overall, across schools and numbers of children. Write a brief sentence summarizing the results.

Instructions:

*Analyze > Distribution. Use 'Opinion' in the Y, Columns role.
Use 'Frequency' for the Freq role. Click OK.*

2. Add 95% confidence intervals for each of the two possible answers. Write a brief sentence interpreting these confidence intervals. Is there enough statistical evidence to conclude that more than half of parents would be in favor of the program?

Instructions:

Select Confidence Interval > 0.95 from the Red Triangle next to the 'Opinion' variable name.

3. Create a graphical and numerical summary of the opinions for each school. Add 95% confidence intervals for each of the two answers for each school. Write a few brief sentences interpreting these interval estimates. Compare the results across the three schools. Does there appear to be differences between the schools?

Instructions:

*Analyze > Distribution. Use 'Opinion' in the Y, Columns role.
Use 'Frequency' for the Freq role. Use 'School' for the By role. Click OK.
Select Confidence Interval > 0.95 from the Red Triangle next to the 'Opinion' variable name for each of the three schools.*

4. Conduct a formal test to determine if there is statistically significant evidence indicating a difference in parents' opinions between the schools. Is this result consistent with what you had concluded in Exercise 3?

Instructions:

Analyze > Fit Y by X. Use 'Opinion' in the Y, Response role, 'School' in the X, Factor role, and 'Frequency' for the Freq role. Click OK.

5. Create a graphical and numerical summary of the opinions for the different number of children. Add 95% Confidence Intervals for each of the two possible answers. Write a few brief sentences interpreting these interval estimates. Compare the results across the three children groups. Does there appear to be differences in how responders feel between these groups?

Instructions:

Analyze > Distribution. Use 'Opinion' in the Y, Columns role, 'Frequency' for the Freq role, and 'Children' for the By role. Click OK.

Select Confidence Interval > 0.95 from the Red Triangle next to the 'Opinion' variable name for each of the three schools.

6. Conduct a formal statistical test to determine if there is statistically significant evidence indicating a difference in parents' opinions between the three children groups. Is this result consistent with what you had concluded in Exercise 5?

Instructions:

Analyze > Fit Y by X. Use 'Opinion' in the Y, Response role, 'Children' in the X, Factor role, and 'Frequency' for the Freq role. Use. Click OK.

7. Create a new variable that is the proportion of responders that have each opinion for each category of school and number of children.

Instructions:

Cols > New Column. Name the column 'Proportion'. Click on Column Properties drop down and select Formula. Then select 'Edit Formula'. Within the Formula Editor, create the formula: 'Frequency' / 'Responders'. Click OK. For the format, choose Fixed Dec and specify to have the data display to 3 decimal places.

8. Create a bar chart that displays the proportion of responders that have each opinion for each category of school and number of children. Identify the groups where there is the largest and smallest proportion of responders In Favor of the program. Identify where there appears to be large differences and where there appears to be little or no differences between the groups.

Instructions:

Graph > Graph Builder. Place 'Proportion' on the Y axis. Place 'Children' on the X axis. Place 'School' in the Group X role. Place 'Opinion' in the Overlay role. Then select the Bar Chart icon. 

Add the data values by selecting Label by Value in the Label drop down controls on the left. Choose the Label Format to display 2 decimal places.

It can often be easier to illustrate specific differences when the variables are used in different roles in a visualization. Recreate the following alternative versions of the graph:

- 2) 'School' on the X axis, 'Children' as the Group X, 'Opinion' as the Overlay
- 3) 'Opinion' on the X axis, 'School' as the Group X, and 'Children' as the Overlay
- 4) 'School' on the X axis, 'Opinion' as the Group X, and 'Children' as the Overlay
- 5) 'Children' on the X axis, 'Opinion' as the Group X, and 'School' as the Overlay
- 6) 'Opinion' on the X axis, 'Children' as the Group X, and 'School' as the Overlay

Note: All the bars are identical in each graph. They are simply grouped differently. Select just one you think best illustrates the key results described in your summary.

9. Perform a set of statistical tests comparing the proportions of those In Favor and those Opposed across the nine different responder groups defined by school and number of children. Identify where there is statistically significance evidence of a difference between specific pairs of groups and where there is not. Are these results consistent with what you had concluded in Exercise 8?

Instructions:

Launch Analyze > Consumer Research > Categorical. Use 'Opinion' for the Response role in the Simple tab. Choose 'School' and 'Children' for the X, Grouping Category role, and 'Frequency' for the Freq role. Click OK.

In the Categorical report window created, select Compare Each Sample under the top red triangle.

Note: Each of the nine categories are labeled in the contingency table (A, B, ..., I). At the bottom of the report window, two tables of p-values are added (Likelihood ratio Chi-Square and Pearsons Chi-Square). As you'll see, the results are almost identical. Examine just the results from Pearsons. With nine categories, there are 36 pairwise comparisons (A vs. B, A vs. C, ..., A vs. I, B vs. C, ... H vs. I).

10. Add 95% Confidence Intervals for each of the two opinions across the nine groups. Identify the groups where it is estimated that at least 50% of the parents are In Favor of the program. Which groups have the largest percentage of parents In Favor of the program? Are there any groups where it is estimated that more than 50% of parents would be Opposed to the program?

Instructions:

Choose More Cell Options > Share Confidence Intervals under the top red triangle.

11. Recall that the response rate was about 60%. That is, we don't have the opinions of approximately 40% of the parents. Based on this, what are some potential issues that could cause the results of our analyses to not accurately generalize to the full population of parents?
12. What additional information could be gathered from surveys that would be valuable to analyze to better understand how parents feel about this and other school programs?