JMP Academic Case Study 004

Film on the Rocks

Bar Charts, Cross-Tabulations and Mosaic Plots

Produced by

Marlene Smith, University of Colorado Denver Business School

jmp STATISTICAL DISCOVERY

Film on the Rocks[1]

Bar Charts, Cross-Tabulations and Mosaic Plots

Background

Film on the Rocks is a summer movie series held at the world-renowned Red Rocks Amphitheatre, which is situated on a hillside of the Rocky Mountains in Morrison, Colorado fifteen miles west of Denver. The film series features classic films, and pre-show entertainment including bands and comedians. Among the features that have made Red Rocks an internationally famous concert stage is sweeping views of Denver, outstanding acoustics, and cool, dry Colorado summers.

The series is jointly promoted by the Denver Film Society (DFS) and the City and County of Denver's Division of Theatres and Arenas (DT&A). It is marketed through various outlets including newspapers, radio, and the Red Rocks and Denver Film Society websites. Film on the Rocks patrons also benefit from corporate sponsorship. In return for on-site posters and banners at Red Rocks, and recognition in pre-show marketing materials, corporations donate funds that keep ticket prices low.

Although the Red Rocks Amphitheatre provides a cinematic experience unlike any other venue, there *are* tradeoffs. Red Rocks is a farther commute for most people than the local movie theater or movie rental store. Given the uphill walk to the amphitheatre from the parking lot, getting there can be challenging. And, as an outdoor venue, the viewing experience is dependent on the weather.

Patron satisfaction with Red Rocks as the venue for the film series is critical to its success. But, the series promoters would also like to increase attendance at the film series, and are unsure how to do this. Promoters recognize that they need a better understanding of the customer base, and of the current level of satisfaction. Knowing the demographics of those who attend the film series will help attract and expand corporate sponsorship. In addition, knowing which media outlets are most effective will provide information about how best to target future marketing campaigns.

To this end, the promoters conducted surveys during a recent Film on the Rocks season. Questionnaires were handed out at the entrance. Volunteers walked through the crowd to remind people about the free soft drink given to those who returned the survey.

The Task

Use the survey results to address the following questions:

- What is the overall level of customer satisfaction?
- What factors are linked to satisfaction?
- What is the demographic profile of Film on the Rocks patrons?
- In what media outlet(s) should the film series be advertised?

[1]Data provided by the Denver Film Society

The Data films.jmp

The data set contains 330 surveys collected during three Film on the Rocks movies: Ferris Bueller's Day Off, Old School, and Willy Wonka and the Chocolate Factory. The variables are:

Gender	The patron's gender: 1 = male; 2 = female
Marital Status	The patron's marital status: 1 = married; 2 = single
Age	The patron's age in years: 1= 1-12; 2 = 13-30; 3 = 31-60; 4 = 60+
Income	The patron's annual household income: 1 = Less than $50,000; 2 = $50-$100,000; 3 = $100,000+
Hear About	The patron's response to this question: "How did you hear about Film on the Rocks?" Respondents could check any of the following that applied: 1 = television; 2 = newspaper; 3 = radio; 4 = website; 5 = word of mouth

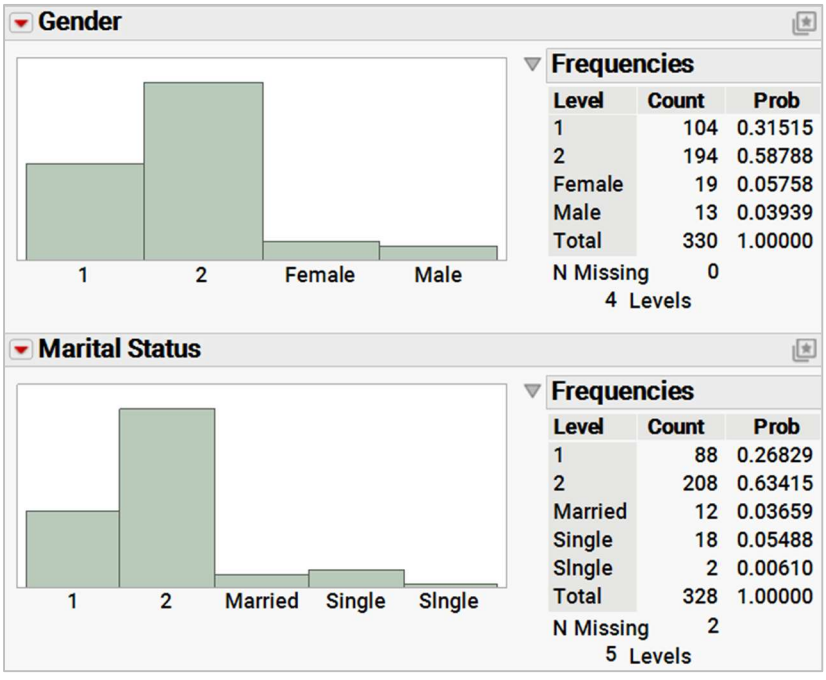
The survey also contained four Likert-scaled questions about satisfaction; each of the following questions is coded: 1 = Excellent; 2 = Good; 3 = Average/Fair; 4 = Poor; 5 = Very Poor.

Signage	"How was the signage directing you to Red Rocks?"
Parking	"How was the venue's parking?"
Clean	"How was the cleanliness of the venue?"
Overall	"How was your overall customer service experience?"

Analysis

With any data that is manually entered into a data table, it's especially important to begin with a check of data quality. This can be accomplished by examining distributions of each variable in the data set. Of special interest are invalid entries and typos. In Exhibit 1, we see coding issues for Gender and Marital Status. Note, for example, the typo "SIngle" in the Marital Status variable.

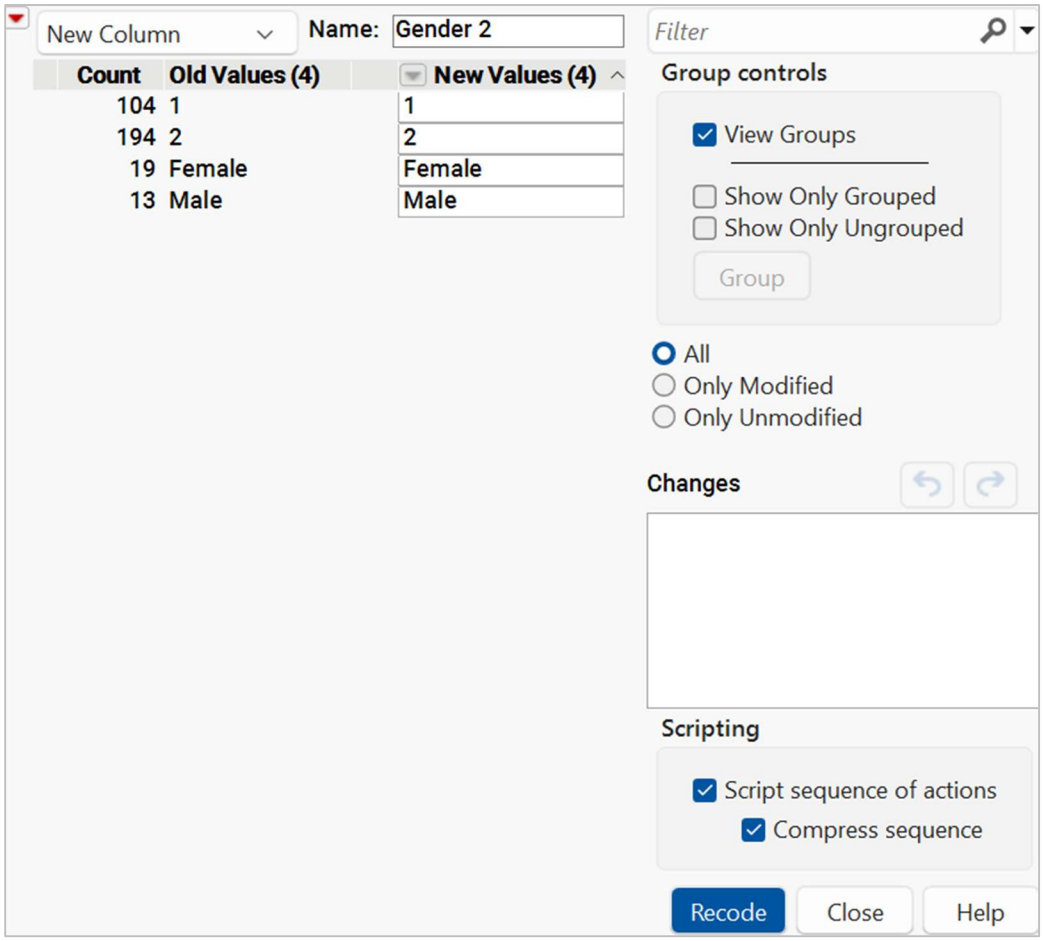
Exhibit 1 Distributions of Gender and Marital Status



(Analyze > Distribution; for a horizontal layout select Stack under the top red triangle)

Coding and other data quality issues should be corrected before beginning any analysis. A simple approach is to use Recode (Exhibit 2).

Exhibit 2 Recoding Variables

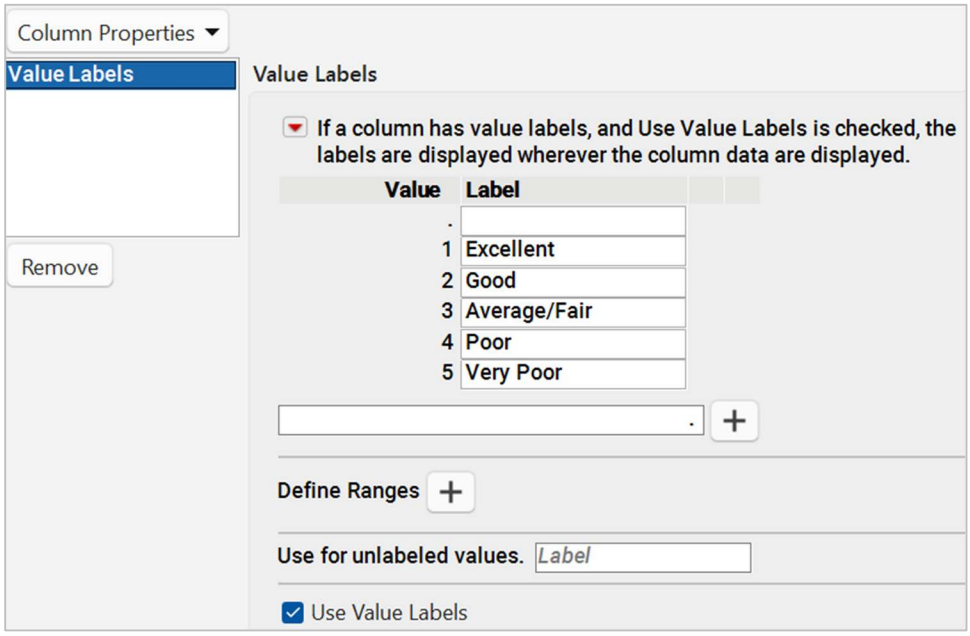


(From the data table, select the column(s) to recode, then select Recode from Col. Under New Value enter the corrected values then select Recode. Choose from the options to create a new column with recoded values or create a formula column or In place)

For ease of presentation and interpretation, all other variables can be recoded to replace numbers with text. **Hear About**, which can contain multiple responses, can be recoded to display each response. For example, observation 61, which contains the value "1,5", could be recoded to display "TV & Word of Mouth".

For the other columns, which contain ordinal data, we've used another approach; a column property has been set to display value labels instead of values. This allows the ordering of the values to be respected in analyses and graphs. In Exhibit 3 we show how to set value labels for the satisfaction variables.

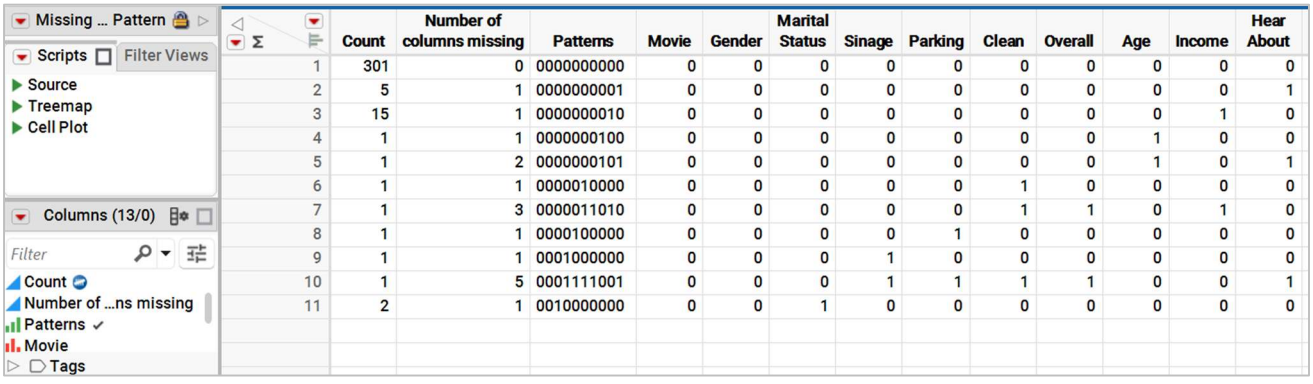
Exhibit 3 Setting Value Labels



(To set value labels for one variable, go to Column Info for the variable, and under Column Properties select Value Labels. Enter the value and the label you'd like to display, and click Add. Repeat for all values, then click OK. Note: To change multiple columns at the same time, select the columns. Then, from the Cols menu select Standardize Attributes and follow the steps above.)

The next step in checking data quality is to look for missing values. Exhibit 4 shows that for 301 of the 330 returned surveys all questions were answered (all columns have zeros in the first row). Five respondents answered all questions except **Hear About** (only **Hear About** has a "1" in the second row). The most frequently skipped question concerned household income. In general, though, the respondents did an unusually good job of completing the questionnaire. Because the number of missing values is small, we'll ignore them in the subsequent analyses.

Exhibit 4 Missing Data Patterns

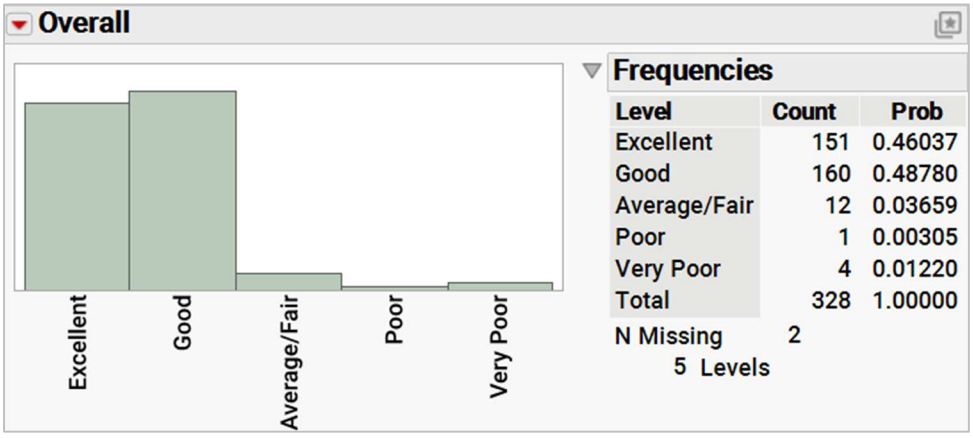


	Count	Number of columns missing	Patterns	Movie	Gender	Marital Status	Sinage	Parking	Clean	Overall	Age	Income	Hear About
1	301	0	0000000000	0	0	0	0	0	0	0	0	0	0
2	5	1	0000000001	0	0	0	0	0	0	0	0	0	1
3	15	1	0000000010	0	0	0	0	0	0	0	0	1	0
4	1	1	0000000100	0	0	0	0	0	0	0	1	0	0
5	1	2	0000000101	0	0	0	0	0	0	0	1	0	1
6	1	1	0000010000	0	0	0	0	0	1	0	0	0	0
7	1	3	0000011010	0	0	0	0	0	1	1	0	1	0
8	1	1	0000100000	0	0	0	0	1	0	0	0	0	0
9	1	1	0001000000	0	0	0	1	0	0	0	0	0	0
10	1	5	0001111001	0	0	0	1	1	1	1	0	0	1
11	2	1	0010000000	0	0	1	0	0	0	0	0	0	0

(Tables > Missing Data Pattern; Select all columns > Add Columns > OK)

Now that we've prepared the data, it's time for analysis. We start with the key variable of interest, **Overall** (satisfaction). Overall customer satisfaction is very high, with 94.8% of respondents reporting either good or excellent customer service experience (Exhibit 5).

Exhibit 5 Distribution of Overall Satisfaction



(Analyze > Distribution; Select Overall = Y columns > OK. Right click on the Y -Axis > Axis Setting > Label Orientation = Vertical or Overall red triangle > Histogram Options > uncheck Vertical*)*

These high levels of satisfaction are wonderful news for the promoters, sponsors, and customers of Film on the Rocks. Only five customers reported Poor or Very Poor satisfaction!

Ironically, this limits our ability, statistically, to learn more about drivers of low satisfaction. Consider, for example, an analysis of overall satisfaction and satisfaction with the venue's parking. The mosaic plot, cross-tabulation and chi-squared tests are shown in Exhibit 6. Certainly, higher levels of overall satisfaction are associated with higher levels of satisfaction with parking, as is evidenced by the mosaic plot. Yet the large number of zero or small counts in the Poor and Very Poor cells should give pause, as JMP warns at the bottom of Exhibit 6. Specifically, although the p-values (Prob > ChiSq) for the Likelihood Ratio and Pearson tests are very small (< .0001), suggesting a strong association between these two satisfaction measures, we should be wary of those results because of limited information about dissatisfaction.

Exhibit 6 Overall Satisfaction versus Satisfaction with Parking

Contingency Analysis of Overall By Parking

Mosaic Plot

Contingency Table

Parking	Count / Total % / Col % / Row %	Overall: Excellent	Good	Average/ Fair	Poor	Very Poor	Total
Excellent		105	26	2	0	1	134
		32.11	7.95	0.61	0.00	0.31	40.98
		69.54	16.35	16.67	0.00	25.00	
		78.36	19.40	1.49	0.00	0.75	
Good		40	109	3	0	1	153
		12.23	33.33	0.92	0.00	0.31	46.79
		26.49	68.55	25.00	0.00	25.00	
		26.14	71.24	1.96	0.00	0.65	
Average/Fair		4	20	5	0	0	29
		1.22	6.12	1.53	0.00	0.00	8.87
		2.65	12.58	41.67	0.00	0.00	
		13.79	68.97	17.24	0.00	0.00	
Poor		2	3	1	0	1	7
		0.61	0.92	0.31	0.00	0.31	2.14
		1.32	1.89	8.33	0.00	25.00	
		28.57	42.86	14.29	0.00	14.29	
Very Poor		0	1	1	1	1	4
		0.00	0.31	0.31	0.31	0.31	1.22
		0.00	0.63	8.33	100.00	25.00	
		0.00	0.25	25.00	25.00	25.00	
Total		151	159	12	1	4	327
		46.18	48.62	3.67	0.31	1.22	

Tests

N	DF	-LogLike	RSquare (U)
327	16	66.761951	0.2268

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	133.524	<.0001*
Pearson	229.479	<.0001*

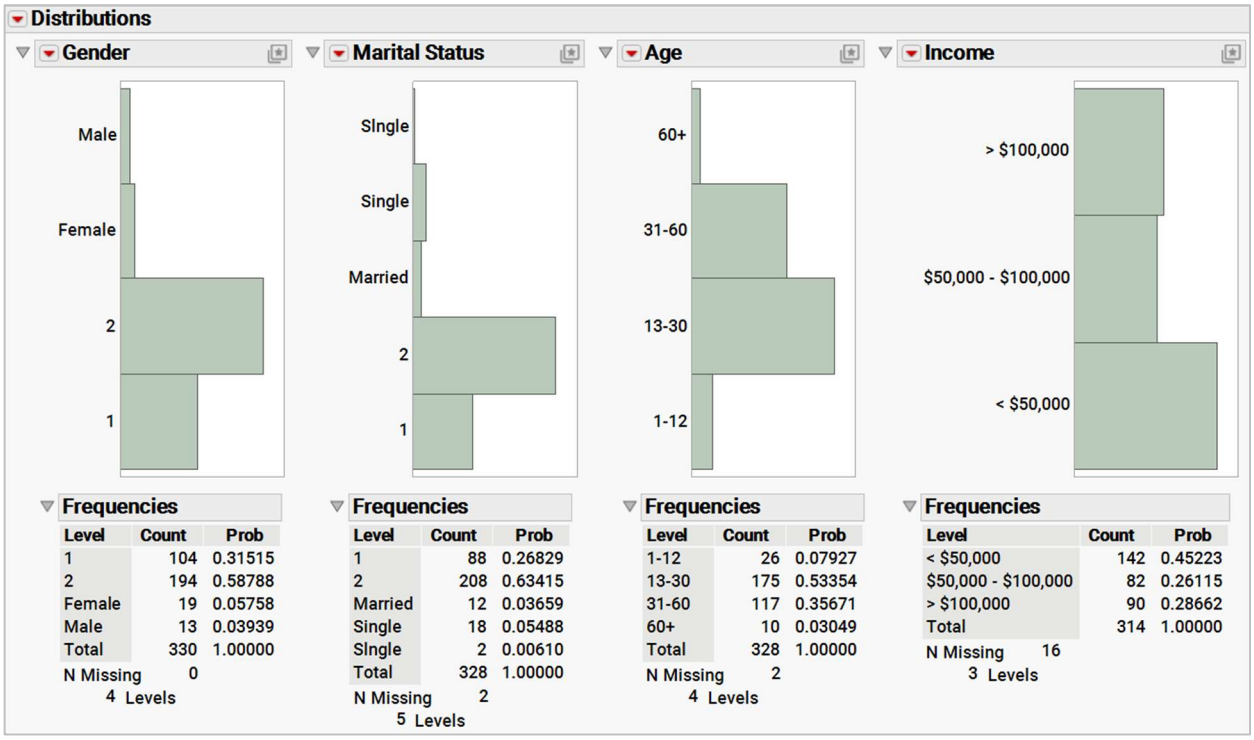
Warning: 20% of cells have expected count less than 5, ChiSquare suspect.

(Analyze > Fit Y by X; use Overall as Y, Response and Parking as X, Factor in the Contingency Table)

We can say more about the demographic make-up of these respondents. Exhibit 7 indicates that:

- 64.5% are female,
- almost 70% are single,
- the largest age group (53%) is 13 to 30 years old, although just over one-third are aged 31 to 60 years; and
- 45% of the respondents have household incomes of less than $50,000.

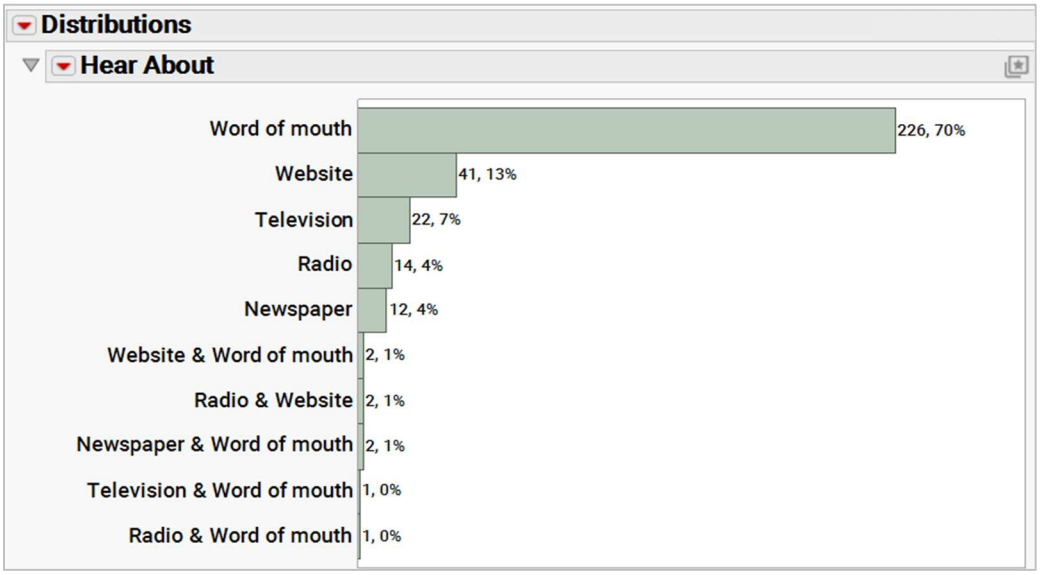
Exhibit 7 Summary of Demographic Information



(Analyze > Distribution. Select Gender, Marital Status, Age and Income as Y columns > OK)

Surprising (to the promoters) is that the vast majority (226 responses, or 70%, of those answering the question) said they heard about the film series solely through word of mouth (Exhibit 8). The traditional media outlets (newspapers, radio, and television) were among the least effective with regard to getting the word out among these respondents.

Exhibit 8 Summary of “How Did You Hear About the Series?”



(Analyze > Distribution > Select Hear About as Y columns > OK. Under the second red triangle, select Histogram Options, Show Counts and Show Percents. To sort in descending, select Order By > Count Ascending under the lower red triangle.)

Summary

Statistical Insights

Although great news for the film series sponsors, it should be noted that the distributions for the satisfaction variables are heavily weighted toward excellent or good reports of satisfaction. In some circumstances, small cell counts can be handled by combining categories; that won’t help here because even combining average/fair with poor and very poor leaves small cell counts. We can’t know much about the dissatisfied customers because there aren’t many of them in the sample. This outcome might have been expected, given that the survey was distributed at the film series. Knowing about who doesn’t come to Film on the Rocks, and why, would only be possible using a different survey design and collection method.

Was the financial cost of giving away soft drinks to respondents worth it? Although virtually all of the people who returned the surveys answered all of the items on the questionnaire, response rates (calculated as the number of surveys returned divided by the number of tickets sold for each show) were low, ranging from 0.4% for Old School to 5.2% for Willy Wonka.

The extent to which the gift and surveyed population might have induced non-representativeness is unclear, as is the extent of a potential voluntary response bias. These are nontrivial matters that should be made clear when reporting the results of the survey.

Managerial Implications

Given the low response rates, not much can be ascertained about overall satisfaction. If it can be assumed that the question, “How did you hear about Film on the Rocks?” is not influenced by voluntary response bias, decision-makers should weigh the costs and benefits of advertising via newspaper, radio, and television. A reasonable percentage of respondents learned about the film series from the websites. The efficacy of the websites in communicating upcoming events should be evaluated.

For this particular market, word of mouth seems to be an important outlet; determining how to take advantage of this vehicle for communication should be the next step in getting the word out about the film series. Forms of social media, such as Twitter and Facebook, might prove to be important vehicles for word of mouth communication.

JMP Features and Hints

This case demonstrates JMP capabilities with univariate and bivariate categorical variable analyses: Distribution was used to produce bar charts and frequency distributions, and Fit Y by X for mosaic plots and cross-tabulations (contingency tables).

We also used some tools for data preparation, including Recode, Missing Data Pattern, and the column property Value Labels. JMP will ignore rows of data that contain missing values. However, if there are a large number of missing values, recoding these as "missing" allows those observations to be included in analyses. Also, you may not want to override the original data. To create new columns with the corrected data when recoding, click on the down arrow next to In Place and change to New Column or Formula Column.

Exercises

1. Do the other satisfaction variables, **Signage** and **Clean**, provide any useful information?

2. Are there any other useful insights that can be gleaned from these data? (Explore the variables using Distribution and dynamic plot-linking, Fit Y by X, and the Graph Builder.)

3. Consider the survey that was administered. If you were to design a second survey (or other type of study) for the promoters of the series, what changes would you consider? Recall that the study should allow the promoters to answer the following questions:

    - What is the overall level of customer satisfaction?
    - What factors are linked to satisfaction?
    - What is the demographic profile of Film on the Rocks patrons?
    - In what media outlet(s) should the film series be advertised?

JMP and all other JMP Statistical Discovery LLC product or service names are registered trademarks or trademarks of JMP Statistical Discovery LLC® indicates USA registration. Other brand and product names are trademarks of their respective companies. in the USA and other countries. Copyright © 2025 JMP Statistical Discovery LLC. All rights reserved.