JMP® ENHANCED DATA SET

SOFT DRINK TRIANGLE TASTE TEST

RELEVANT JMP PLATFORMS AND STATISTICAL TECHNIQUES

Graph Builder : Heat Map

Distribution : Bar Charts and Frequency Table ; Confidence Intervals and Tests for Proportions

PROBLEM STATEMENT

A manufacturer of soft drinks has been experimenting with a new ingredient formulation for one of its drinks popular with teenagers. Though this new formulation would be slightly more expensive to produce, it uses a healthier substitute for the current sweetener that the company believes could help increase sales by appealing to parents. If consumers of the drink are unable to tell the difference in taste, it could be more advantageous to use this new ingredient formulation.

The Triangle test is a common method used in sensory studies to determine if consumers can discriminate between two formulations. In this type of test, assessors (taste testers) are given three samples to taste. Two of them are the same and one is different. Each assessor tries all three and is asked to identify which one is different. In this particular study, the assessors are not asked to rate which one they think tastes better or what characteristics they perceive as different, but merely to identify which one they believe is different.

Each assessor gets up to one minute to taste the three samples, as often as they need and in any order. Each must make a decision on which one they believe is different even if they feel they can't tell the difference and simply guess. There are two possible test sets one can receive – either two of the samples being the current formulation and one new (Current, Current, New) or two being the new formulation and one the current (New, New, Current).

120 teenagers (60 males and 60 females) ranging in age from 13-17 (24 for each age) were recruited for the study with all being consumers of the soft drink's current formulation at least once a week. A balanced experimental design for age and gender and test set was created resulting in 6 assessors in each of the 5 x 2 x 2 = 20 groups. Below is a table of the experimental design for these factors.

		Test Set	
Age	Gender	C C N	N N C
13	F	6	6
	M	6	6
14	F	6	6
	M	6	6
15	F	6	6
	M	6	6
16	F	6	6
	M	6	6
17	F	6	6
	M	6	6

The frequency with which the assessors consume the product is another important factor to examine, but it was not part of the experimental design. Each assessor was asked this question, and responses were coded as one of three categories: 1-2 per week , 3-5 per week , 6+ per week. The table below shows the experimental design with this additional factor. As you can see, this factor is not incorporated in a balanced manner, resulting in some groups not having any assessors or only 1 while other groups have 4 or 5.

		Test Set					
		C C N			N N C		
		Frequency			Frequency		
Age	Gender	1-2/week	3-5/week	6+/week	1-2/week	3-5/week	6+/week
13	F	1	3	2	3	1	2
	M	2	2	2	3	3	0
14	F	0	1	5	2	2	2
	M	1	1	4	1	2	3
15	F	1	3	2	1	3	2
	M	1	4	1	3	1	2
16	F	2	3	1	3	0	3
	M	0	1	5	3	1	2
17	F	0	3	3	2	1	3
	M	4	0	2	1	3	2

DATA SET

Soft_Drink_Triangle_Taste_Test.jmp

Assessor	ID of the assessor coded as *"Age-Gender-Number in Age/Gender group-Taste Set"*
Age	Age of assessor (13, 14, 15, 16, 17)
Gender	Gender of assessor (M or F)
Test Set	Test set received (Current, Current, New) or (New, New, Current)
Frequency	Frequency assessor consumes the drink (1-2, 3-5, or 6+ per week)
Correct Answer	Yes/No indicating if assessor correctly identified the one that was different

EXERCISES

The exercises consist of conducting a variety of statistical tests to determine if the subjects can discriminate between the two formulations (overall and based upon their Age, Gender, and Frequency of using the product). You'll also create a visualization that summarizes and communicates the results of the study in one graph.

Note: Each Assessor must make a choice on which sample they believe is different even if they feel they can't distinguish and simply must guess. Since 1 of the 3 samples are different and 2 the same, there is a 1/3 chance that an Assessor will correctly guess the one that is different by randomness alone. As a result, the statistical tests you'll perform will be to determine if the statistical evidence exists indicating that the percentage of consumers in a population who can identify a difference is greater than 33%.

1. Create a bar chart of the overall proportion of times the Assessors correctly identified the sample that is different. Provide a 95% Confidence Interval for the population proportions and conduct a statistical test (at 0.05 significance level) to determine if a population can correctly identify the different sample greater than 33% of the time. What conclusion can you reach based on these results? Provide an interpretation of the Confidence Interval. Why is it incorrect to conclude that between (LCL to UCL) of a population of teenage consumers would be able to notice that the soft drink has changed if the new formulation was released in the market?

> ***Instructions:*** *Choose Analyze > Distribution. Place the variable 'Correct Answer' in the Y role. Click OK.*
> *For the Confidence Intervals: Choose Confidence Intervals > 0.95 under the red triangle menu next to the variable name 'Correct Answer'.*
> *For the Statistical Test: Choose Test Probabilities under the red triangle menu. Type in '0.33' in the 'Hypoth Prob / Yes' field. Select 'probability greater than hypothesized value" for the alternative hypothesis.*

2. Create a bar chart of each Gender showing the proportion of times the Assessors correctly identified the sample that is different. Provide 95% Confidence Intervals and conduct statistical tests for each Gender. Examine the p-values for the hypothesis tests. What conclusions can you reach based on these results? Provide an interpretation of the Confidence Intervals.

Instructions: *Choose Analyze > Distribution. Place the variable 'Correct Answer' in the Y role. Place 'Gender' in the By Variable role. Click OK.*
For Confidence Intervals: For each Gender, choose Confidence Intervals > 0.95 under the red triangle menu next to the variable name 'Correct Answer'
For Statistical Test: For each Gender, choose Test Probabilities under the red triangle menu. Type in '0.33' in the 'Hypoth Prob / Yes' field and select 'probability greater than hypothesized value" for the alternative hypothesis.

3. Create a bar chart for each Age showing the proportion of times the Assessors correctly identified the sample that is different for each Age. Provide 95% Confidence Intervals and conduct statistical tests for each Age. Examine the p-values for the hypothesis tests. What conclusions can you reach based on these results? Provide an interpretation of the Confidence Intervals.

Instructions: *Choose Analyze > Distribution. Place the variable 'Correct Answer' in the Y role. Place 'Age' in the By Variable role. Click OK.*
For Confidence Intervals: For each 'Age', choose Confidence Intervals > 0.95 under the red triangle menu next to the variable name 'Correct Answer'
For Statistical Test: For each 'Age', choose Test Probabilities under the red triangle menu. Type in '0.33' in the 'Hypoth Prob / Yes' field. Select 'probability greater than hypothesized value" for the alternative hypothesis.

4. Create a bar chart for each Frequency group showing the proportion of times the Assessors correctly identified the sample that is different. Provide 95% Confidence Intervals and conduct statistical tests for each group. Examine the p-values for the hypothesis tests. What conclusions can you reach based on these results? Provide an interpretation of the Confidence Intervals.

Instructions: *Choose Analyze > Distribution. Place the variable 'Correct Answer' in the Y role. Place 'Frequency' in the By Variable role. Click OK.*
For Confidence Intervals: For each 'Frequency', choose Confidence Intervals > 0.95 under the red triangle menu next to the variable name 'Correct Answer'
For Statistical Test: For each 'Frequency', choose Test Probabilities under the red triangle menu. Type in '0.33' in the 'Hypoth Prob / Yes' field. Select 'probability greater than hypothesized value" for the alternative hypothesis.

5. Create a bar chart of the proportion of times the Assessors correctly identified the sample that is different for each Test Set. Provide 95% Confidence Intervals and conduct statistical tests for each. Examine the p-values for the hypothesis tests. What conclusions can you reach based on these results? Provide an interpretation of the Confidence Intervals.

> ***Instructions:*** *Choose Analyze > Distribution. Place the variable 'Correct Answer' in the Y role. Place 'Test Set' in the By Variable role. Click OK.*
> *For Confidence Intervals: For each 'Test Set', choose Confidence Intervals > 0.95 under the red triangle menu next to the variable name 'Correct Answer'*
> *For Statistical Test: For each 'Test Set', choose Test Probabilities under the red triangle menu. Type in '0.33' in the 'Hypoth Prob / Yes' field. Select 'probability greater than hypothesized value" for the alternative hypothesis.*

6. To summarize the results of the study in one graph, create a heat map that displays each Assessors' answer with the Frequency group, Age and Gender identified clearly on the graph. Does this graph do an adequate job at visually communicating the conclusions you reached in the statistical tests in Exercises 2-4?

> ***Instructions:*** *Use Graph Builder. Place 'Assessor' on the X axis. Place 'Gender' on the X axis underneath 'Assessor'. Place 'Age' on the X axis underneath 'Gender'. Place 'Frequency' in the Group Y role. Place 'Correct Answer' in the Color role. Select the heat map icon.* 
> *Note: Each of the 120 tasters can only be in one 'Frequency' group. Thus there's only one colored cell for each column in the heat map. The color indicates whether or not that particular taster correctly identified the different sample.*

7. Provide ideas you have for other data that could have been collected in this study (e.g., 'other characteristics of the assessors' , 'response variables' , etc.)? What other studies could be done to better understand consumers' ability to distinguish between the current and new formulation and if switching to this new formulation would be a good idea. Describe a study that could be done that would allow you to create an estimate of the proportion of a population of teenage consumers who would notice that the soft drink has changed if the new formulation was released in the market?