

JMP Instructional data set 026

Alcohol Taste Perception Study

JMP platforms and statistical methods

Graph Builder:	Comparative line graphs and bar charts
Formula Editor:	Creating new variables
Fit Y by X:	Analysis of variance (ANOVA); simple linear regression
Fit Model:	Multi-factor analysis of variance (ANOVA)

Problem statement

A food and beverage sensory testing laboratory has been engaged by a beverage producer to help better understand how one's perception of the quality of an alcoholic beverage changes throughout the time of drinking it.

Three different studies were done for three different alcoholic beverages (wine, beer, bourbon).



For each beverage, a panel of 10 different tasters were selected from a pool of consumers who consider themselves to be regular drinkers of that beverage, with an average level of knowledge/ability in distinguishing quality. The tasters were told they would be drinking 16 different unknown brands across four different tasting sessions and asked to rate the quality of each on a 1-10 scale, with the scoring of 1-2: OK; 3-4: Decent; 5-6: Good; 7-8: Very Good; 9-10: Excellent. Unknown to them, however, is that they would in fact be tasting the same four brands in each session.

Experimental Design

To examine how the tasters' ratings might change across a typical drinking experience, the experiment was conducted as follows:

Four different tasting sessions were scheduled (each one a week apart).

Each session consisted of the tasters drinking the four brands in a half-size serving (6 oz. for the wine and beer and 0.5 oz for the bourbon), resulting in the taster essentially having two drinks. For the wine, the tasters drank the four brands while eating a meal of pasta. For the beer, the tasters ate a beef burger and fries. For the bourbon, they ate a small serving of mixed nuts. It was designed to simulate a typical drinking experience for the beverages.

The order of tasting for each session was created such that each taster experienced each of the 4 brands in the 4 possible tasting positions. For example, one such arrangement was:

Session 1: A, B, C, D
Session 2: B, A, D, C
Session 3: C, D, A, B
Session 4: D, C, B, A

Data

IDS_026_Wine.jmp

IDS_026_Beer.jmp

IDS_026_Bourbon.jmp

Tasting Session	Four tasting sessions (1, 2, 3, 4)
Taster	Name of the taster
Order_Cat	Position of the tasting as a categorical variable (1, 2, 3, 4)
Order_Num	Position of the tasting as a numerical variable (1, 2, 3, 4)
Brand	Name of each of the 4 brands
Score	Quality rating provided by the taster

Exercises

Perform the following analyses for each beverage.

1. Create a comparative line graph that compares the scores across tasting order for each taster and brand. Does there appear to be any differences or trend in the scores across the tasting order? If so, describe those differences.

Instructions:

Graph > Graph Builder. Place 'Score' on the Y axis, 'Order_Cat' on the X axis, 'Brand' in the overlay role, and 'Taster' in the wrap role.

Select to display both the data points and connect lines.  

2. Create 2 new variables. 1) Difference in the score for a given brand and tasting session from the average of all scores from that taster across all brands ; 2) Difference in the score for a given brand and tasting session from the average of all scores for a given brand across all tasters

Instructions:

1) Choose Cols > New Columns. Name the column 'Diff from Avg of All Scores from Taster'. Select Column Properties > Formula. In the formula editor, write the following formula:

$$\text{Score} - \text{Col Mean} (\text{Score}, \text{Taster})$$

Note: The 'Col Mean' function is under the Statistical Functions.

2) Choose Cols > New Columns. Name the column 'Diff from Avg of All Scores from Brand'. Select Column Properties > Formula. In the formula editor, write the following formula:

$$\text{Score} - \text{Col Mean} (\text{Score}, \text{Brand})$$

3. Create a comparative bar charts that compares these new variables across tasting order for each taster and brand. Are there any differences in scores between brands and tasters? In other words, do certain tasters tend to score higher/lower regardless of the brand, and do certain brands tend to have higher/lower scores regardless of the taster? Explain why it is important when evaluating if there are differences or trends in the scores across tasting order to account for the different tasters and brands?

Instructions:

Graph > Graph Builder. Place 'Diff from Avg of All Scores from Taster' on the Y axis, 'Order_Cat' on the X axis, 'Brand' in the Group Y role, and 'Taster' in the Group X role. Select the bar chart icon from the graph palette. 

Choose to display the data values by selecting Label by Value from the Label drop down menu in the controls on the left.

Instead of creating a separate graph for the 'Diff from Avg of All Scores from Brand' variable, utilize the Column Switcher tool so either variable can be chosen to display. To do so, select Redo > Column Switcher under the red triangle. Select the current variable used for the Y as the Column to Switch, and choose that along with the other variable to use as the Replacement Columns.

4. Conduct a One-factor Analysis of Variance of the Scores between Order.
Is there statistical evidence that the average scores are not equal between order of tasting?
Perform multiple comparisons across all possible pairs. Identify all significant differences.

Instructions:

Analyze > Fit Y by X. Place 'Score' in the Y, Response role. Place 'Order_Cat' in the X, Factor role. Click OK. Select Means/ANOVA under the red triangle.
Select Compare Means > Each Pair, Student's t.

5. Conduct a simple linear regression analysis of the Scores by Order.
Is there statistical evidence that there is a linear trend between the scores across order of tasting?
How is this analysis different than that in Exercise 4? Provide and interpretation of the R-Squared value. Interpret the slope coefficient from the fitted regression equation. Which type of model (One-factor ANOVA from Exercise 4 or the linear regression done in this exercise) do you think is better at describing the change in perception of the quality of the beverage across amount consumed?

Instructions:

Analyze > Fit Y by X. Place 'Score' in the Y, Response role. Place 'Order_Num' in the X, Factor role. Click OK. Select Fit Line under the red triangle.

6. What are the statistical models created in Exercise 4 and 5 not accounting for?
7. Build a multiple linear regression model for the Scores by the three factors ('Order', 'Brand', and 'Taster'). Start with a model that includes all 3 possible two-way interactions. Evaluate statistical tests for each of those interactions. Remove any that are not significant. Then evaluate statistical tests for each of the three main effects. Remove any that are not significant. The final model should only contain significant terms. Provide an interpretation of R-Squared. Why is the value different than the one from the simple linear regression done in Exercise 5?

Instructions:

Analyze > Fit Model. Place 'Score' in the Y, Response role. Highlight 'Taster', 'Order_Num', and 'Brand' in the columns list. Then select Factorial to Degree in the Macros drop down menu. Change Emphasis to Minimal Report. Click Run.
In the Effects Summary Table, examine the p-values for the three two-way interactions. Highlight any that are not statistically significant and select Remove.
Next examine the p-values for the three main effects. Remove any that are not statistically significant.

8. Display the profiler for the final model created in Exercise 7. What does this profiler tool allow you to do that you were simply visually evaluating in the graphs created in Exercises 2 and 3? Do you think it's reasonable to use the regression equation to predict what the scores would be beyond 4 (e.g. 6+)? If not, describe why

Instructions:

Select Factor Profiling > Profiler under the red triangle.

9. Is the slope coefficient for 'Order_Num' in this multi-factor regression model the same or different from the simple linear regression model. Explain why.
10. Compare the estimates of the slope coefficient for Order between the multiple linear regression model to the simple linear regression model. Are these estimates the same or different? Why is this so? Examine the Confidence Intervals for these slope coefficients for each model. Are they the same or different? Why?

Instructions:

Right-click on the Parameter Estimates table in the report for each of the models, and select Columns > Lower 95% and Upper 95%.

11. What are some ideas for further experimentation that could help in better understanding how people perceive the quality of an alcoholic beverage over time?