

Module 7 - Market Research on Caffeinated Products

Background

- A start-up company in Boulder has an idea for a new energy product: caffeinated chocolate. Many details about the exact product concept are yet to be decided, however, the company wanted to get some consumer input before it moved ahead with developing the new product. To do that, they created concept statements for seven related product ideas and administered a survey regarding those concepts. You will analyze the data collected from the surveys.
 - The seven concepts are
 1. Caffeinated chocolate bar
 2. Caffeinated cookies
 3. Caffeinated chocolates
 4. Caffeinated potato chips
 5. Caffeinated water
 6. Caffeinated cereal/granola
 7. Caffeinated hamburger meat
 - The concepts numbered 1 and 3 (chocolate bar and chocolates) are closest to what the company has in mind for the new product. Concept number 2 is a similar idea, while concepts 4-6 are quite different categories from their focal idea. Concept 7 (hamburger meat) served as a benchmark: it was designed as an intentionally unattractive idea, and it was used to see if people were really paying attention to the survey or just answering randomly.
 - The survey first asked some background questions (gender, age, employment status, and usage of energy products), then asked four questions for each of the seven concepts. The four questions for each concept were about 1) purchase intent, 2) uniqueness of the product idea, 3) projected usage frequency, and 4) substitution behavior. The content of the module focuses on the analysis of the responses to the intent, uniqueness, and frequency questions.
- The Excel workbook for this module consists of 3 tabs.
 1. The “survey-results” tab contains all the raw data gathered from the 150 survey respondents.
 2. The “concepts” tab contains the exact statements of the concepts.
 3. The “response codes” tab contains the meaning of the response values for the gender, age, intent, uniqueness, and frequency questions.

Part 0: Mean and Weighted Mean

Given a large set of data, it can be desirable to summarize the general tendencies of the data. One of the most common summarizing computations is that of the mean of the data. In some cases, you may want to allocate more/less weight to certain data and compute a weighted mean.

- 1) To begin this Module, view the screencast [Mean and Weighted Mean](#).

Part 1: Purchase Intent

Each respondent was asked a purchase intent question for each concept: “Based on the description, how likely would you be to purchase this product?” This is a standard question in market research for consumer products.

In the data file, the responses to the purchase intent questions are in columns labeled INTENT(1) through INTENT(7), corresponding to Concept 1 through Concept 7 above.

In those columns,

- 1 means the respondent answered “Definitely would not,”
- 2 means “Probably would not,”
- 3 means “Might or might not,”
- 4 means “Probably would,” and
- 5 means “Definitely would.”

1) Average Purchase Intent

- a. For each concept, find the average purchase intent response. Also, find the number of people who answered the purchase intent question for each concept.

#	Concept Name	Average Purchase Intent	Number of Responses
1	Caffeinated chocolate bar	2.49	140
2	Caffeinated cookies		
3	Caffeinated chocolates		140
4	Caffeinated potato chips		
5	Caffeinated water	2.04	139
6	Caffeinated cereal/granola	1.72	
7	Caffeinated hamburger meat		142

- b. Which concept has the highest average score for purchase intent?

- c. Which concept has the lowest average score for purchase intent?

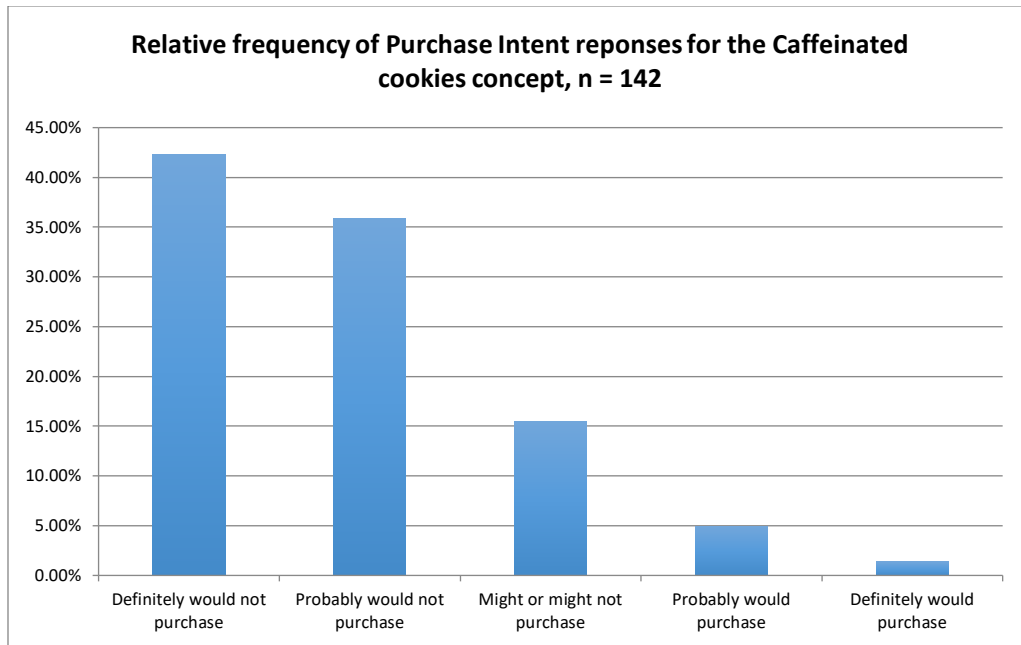
2) Distribution of Responses

For each concept, what percentage of the responses are 1s, 2s, etc.? Give your answers as percentages to two decimal places.

Distribution of Responses for Purchase Intent						
#	Concept	%1s	%2s	%3s	%4s	%5s
2	Caffeinated cookies	42.25%	35.92%	15.49%		
3	Caffeinated chocolates		29.29%	27.14%		

3) Graph of Distribution of Responses

Below is a bar graph showing the distribution of responses for caffeinated cookies (Concept 2).



- a. Create a similar graph in Excel showing the distribution for caffeinated chocolates (Concept 3).

4) Top-Two Boxes

Some companies look at the “top-two boxes” score for concepts. The top-two boxes score is the percentage of respondents who answered “Probably would” plus the percentage of respondents who answered “Definitely would.” For example, if 6% of respondents answered “Definitely would” (a response of 5 in the data) for one of the concepts and 20% of respondents answered “Probably would” (a response of 4 in the data) for that concept, then the top-two boxes score for that concept is 26%.

- a. What is the top-two boxes score for the Caffeinated chocolate bar (Concept 1)?

- b. Which of the seven concepts has the highest top-two boxes score?

Purchase intent responses can be coded as probabilities. One way of translating the response phrases to probabilities is as follows:

- Definitely would not = 0% chance,
- Probably would not = 25% chance,
- Might or might not = 50% chance,
- Probably would = 75% chance, and
- Definitely would = 100% chance.

5) Average Purchase Probabilities

- a. Using the probabilities above, what is the average probability of purchase for each of the seven concepts? Use all the responses, not just the top-two boxes.

Average Purchase Probabilities, Using 0%, 25%, 50%, 75%, 100%		
#	Concept	Average Purchase Probability
1	Caffeinated chocolate bar	37.14%
2	Caffeinated cookies	21.83%
3	Caffeinated chocolates	

4	Caffeinated potato chips	
5	Caffeinated water	
6	Caffeinated cereal/granola	18.04%
7	Caffeinated hamburger meat	

- b. Which concept has the highest average probability of purchase?
- c. Which concept has the lowest average probability of purchase?
- d. If Concept A has a higher Average Purchase Intent (using the 1-5 scale) than Concept B, will Concept A necessarily have a higher Average Purchase Probability (using the 0%, 25%, 50%, 75%, 100% scale) than Concept B?
(Note that you should not answer this question considering only the data presented from this survey, but in general for any survey results.)

People often overstate their intention to purchase. Instead of 0%, 25%, 50%, 75%, and 100%, a more realistic translation from response phrases to probabilities is as follows:

- Definitely would not = 0% chance,
- Probably would not = 2% chance,
- Might or might not = 10% chance,
- Probably would = 25% chance, and
- Definitely would = 50% chance.

- 6) Using these values, what is the average probability of purchase for each of the seven concepts?

Average Purchase Probabilities, Using 0%, 2%, 10%, 25%, 50%		
#	Concept	Average Purchase Probability
1	Caffeinated chocolate bar	9.65%
2	Caffeinated cookies	4.20%
3	Caffeinated chocolates	
4	Caffeinated potato chips	
5	Caffeinated water	

6	Caffeinated cereal/granola	
7	Caffeinated hamburger meat	0.86%

- a. Which concept has the highest average probability of purchase?

 - b. Which concept has the lowest average probability of purchase?

 - c. If Concept A has a higher Average Purchase Intent (using the 1-5 scale) than Concept B, will Concept A necessarily have a higher Average Purchase Probability (using the 0%, 2%, 10%, 25%, 50% scale) than Concept B?
(Note that you should not answer this question considering only the data presented from this survey, but in general for any survey results.)
- 7) Comparing Top-Two Boxes Percentages and Average Purchase Probabilities
TRUE or FALSE: If a concept has the largest top-two boxes score, it will also have the highest average purchase probability. Explain your answer.

Part 2: Frequency

Each respondent was asked, “If you tried this product and it met your expectations for taste and effectiveness, how often could you see yourself using it?” This is a standard question in market research for non-durable consumer products.

In the data file, the responses to the frequency questions are in columns labeled **FREQ(1)** through **FREQ(7)**, corresponding to Concept 1 through Concept 7 as shown at the beginning of the module.

In these columns,

- 1 means the respondent answered “Once a month or less often,”
- 2 means “More than once a month, but less than once a week,”
- 3 means “1 to 3 times a week,”
- 4 means “More than 3 times a week,” and
- 5 means “Don’t know.”

1) Distribution of Responses

For each concept, what percentage of the responses are 1s, 2s, etc.? Give your answers as percentages to two decimal places.

Distribution of Responses for Frequency						
#	Concept	%1s	%2s	%3s	%4s	%5s
2	Caffeinated cookies	61.76%		2.94%		23.53%
3	Caffeinated chocolates	57.46%				11.19%

You can use the frequency responses to estimate annual usage behavior. Similar to our translation of intent responses to probabilities, you can translate the frequency responses to annual usage numbers. Of course, these are just estimates, but they give a reasonable approximation. Use the following values.

- Once a month or less often = 5 times a year
- More than once a month but less than once a week = 15 times a year
- 1 to 3 times a week = 60 times a year
- More than 3 times a week = 100 times a year
- Don’t know = 0 times a year.

Notice that these values are conservative translations, once again assuming that people overstate their intentions.

2) Estimated Annual Usage

Using these values, find the estimated annual usage for each concept.

Estimated Annual Usage, Using 5, 15, 60, 100, 0		
#	Concept	Estimated Annual Usage
1	Caffeinated chocolate bar	11.30
2	Caffeinated cookies	7.87
3	Caffeinated chocolates	
4	Caffeinated potato chips	
5	Caffeinated water	
6	Caffeinated cereal/granola	8.72
7	Caffeinated hamburger meat	

3) TRUE or FALSE: If the numeric values of annual usage were changed from 5, 15, 60, 100, and 0 to the values 6, 20, 60, 110, and 0, the estimated annual usage of every concept would increase or stay the same, no matter what the actual responses to the survey were. Why?

4) Normalizing: Excluding the Don't Knows

Above, you treated "Don't know" as "won't buy." That is, you estimated 0 annual usage for people who responded "Don't know" to the frequency question. That's a somewhat extreme assumption.

- a. Instead, you could just ignore or exclude the "Don't know" responses. To do that, look at the percent of each response out of the total number of responses 1-4. Find those percentages and fill in the chart below.

Distribution of Responses for Frequency, Excluding Don't Know (5)					
#	Concept	%1s	%2s	%3s	%4s
2	Caffeinated cookies		13.46%	3.85%	

6	Caffeinated cereal/granola		16.19%	4.76%	
---	----------------------------	--	--------	-------	--

- b. Using the same estimates of frequency for responses 1-4 from above (e.g., a response of 1, which is “Once a month or less often” means 5 times a year), find the estimated annual usage for each concept.

Estimated Annual Usage, Excluding Don't Know (5)		
#	Concept	Estimated Annual Usage
1	Caffeinated chocolate bar	13.38
2	Caffeinated cookies	10.29
3	Caffeinated chocolates	
4	Caffeinated potato chips	
5	Caffeinated water	
6	Caffeinated cereal/granola	11.05
7	Caffeinated hamburger meat	

- c. The estimated annual usage for each concept is higher when the “Don't know” responses are excluded. Why?

Part 3: By Gender and Age

The column labeled GEN has responses to a question about gender. The responses are coded as follows:

- 1 means Male
- 2 means Female

The column labeled AGE has responses to a question about age. The responses are coded as follows:

- 1 means the respondent answered under 18 years old
- 2 means 18-25 years old
- 3 means 26-35 years old
- 4 means 36-45 years old
- 5 means 46-55 years old
- 6 means 56-65 years old
- 7 means 66-75 years old
- 8 means over 75 years old

- 1) For each concept, find the average purchase intent (INTENT) response by Male respondents.

Purchase Intent by Male Respondents			
#	Concept	Average Purchase Intent	Number of Responses
1	Caffeinated chocolate bar	2.56	39
2	Caffeinated cookies		41
3	Caffeinated chocolates		39
4	Caffeinated potato chips	1.51	
5	Caffeinated water	2.13	
6	Caffeinated cereal/granola	1.56	
7	Caffeinated hamburger meat		40

- 2) For which of the seven concepts do males have higher average purchase intent than females?

- 3) Find the average purchase intent (INTENT) response by respondents 45 and under.

Purchase Intent by Respondents 45 and under			
#	Concept	Average Purchase Intent	Number of Responses
1	Caffeinated chocolate bar	2.58	97
2	Caffeinated cookies	1.94	99
3	Caffeinated chocolates	2.66	
4	Caffeinated potato chips		
5	Caffeinated water		
6	Caffeinated cereal/granola		
7	Caffeinated hamburger meat	1.22	98

- 4) For which of the seven concepts do respondents aged 45 and under have higher average purchase intent than respondents over 45?

- 5) Distribution of Responses

For each concept, determine what percentage of the frequency (FREQ) responses from Male respondents 45 and under are 1s, 2s, etc.? Give your answers as percentages to two decimal places.

#	Concept	% 1s	%2s	%3s	%4s	%5s
2	Caffeinated cookies	54.84%		3.23%		
3	Caffeinated chocolate	60.00%				13.33%

Part 4: Excluding Low Quality Responses

The “caffeinated hamburger meat” concept wasn’t a real concept—it was just a check to see if people were paying attention to the survey. Although it is possible that someone would have a high purchase intent for that concept, the concept was included to help screen out people who were not taking the survey seriously. When people took the survey, the response for “Definitely would” was the closest to the “advance to the next screen” button, so people who were just trying to speed through the survey would be more likely to give that response.

The company also wants to screen out people who didn’t answer all the questions.

- 1) Find the average purchase intent responses, excluding “low quality” responses. By definition, a low quality response is a response in any row in the data set that meets *at least one* of these criteria:

- a response of 3, 4, or 5 for purchase intent on Concept 7 (INTENT(7))
- failure to answer all twenty-one INTENT, UNIQ, and FREQ questions.

There are many ways to exclude the low quality responses. Try to avoid using a manual process like deleting each of the low quality rows. (Imagine if the data set had 150,000 rows instead of 150: you wouldn’t want to go through all of them, one by one.)

Purchase Intent, Excluding Low Quality Responses			
#	Concept	Average Purchase Intent	Number of Responses
1	Caffeinated chocolate bar	2.50	125
2	Caffeinated cookies		
3	Caffeinated chocolates		
4	Caffeinated potato chips		
5	Caffeinated water	2.01	
6	Caffeinated cereal/granola	1.74	
7	Caffeinated hamburger meat		

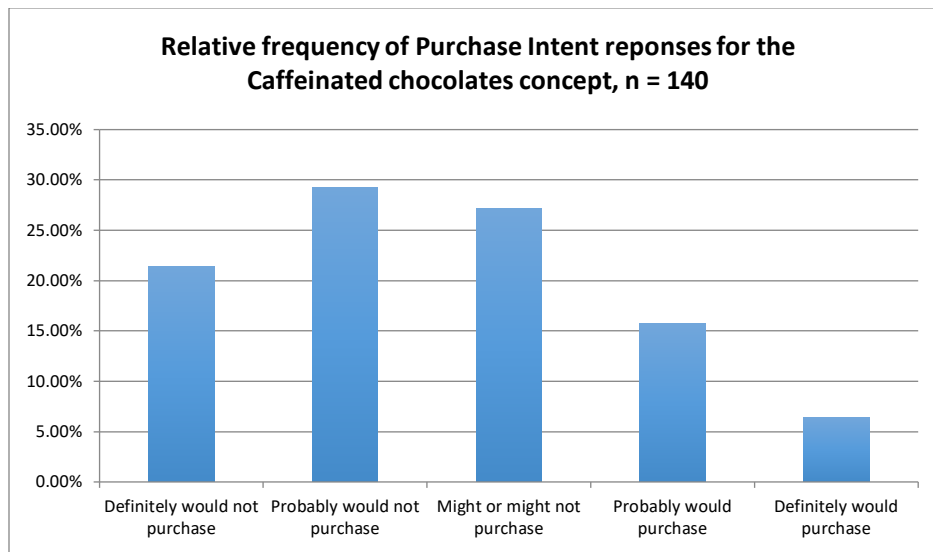
Selected Answers

Part 1

1b) Caffeinated chocolates (Concept 3) has the highest average purchase intent (2.56).

1c) Caffeinated hamburger meat (Concept 7) has the lowest average purchase intent (1.21).

3)



4a) 20.71%

4b) Concept 3 has the highest top-two boxes score (22.14%)

5b) Caffeinated chocolates has the highest average probability of purchase (39.11%)

5c) Caffeinated hamburger meat has the lowest average probability of purchase (5.28%)

5d) The answer is “yes”, independent of the specific survey results.

6a) Caffeinated chocolates has the highest average probability of purchase (10.44%)

6b) Caffeinated hamburger meat has the lowest average probability of purchase (0.86%)

6c) Although the answer is “yes” for the results of this specific survey, in general the answer is “no”.

7) Although it will often be the case that the concept with the largest top-two boxes score is the one with the highest average purchase probability, in general, the statement is false

Part 2

3) True. Because the usage estimate for every category of response increased or stayed the same, the estimated annual usage for each concept will increase or stay the same, no matter what the survey results were.

4c) When included, the “Don’t Know” responses were counted as “0 consumed” which lowered the average.

Part 3

2) Males have higher average purchase intent than females for Concepts, 1,2,4,5,7.

4) Respondents 45 and under have higher average purchase intent than those over 45 for Concepts 1,2,3,4,6,7.

The creation of this work "Mathematical Analysis in Business" was supported by Open CU Boulder 2020-2021, a grant funded by the Colorado Department of Higher Education with additional support from the CU Office of the President, CU Office of Academic Affairs, CU Boulder Office of the Provost, and CU Boulder University Libraries.