Conditional Probability

Find your dataset. Put just the data you want to analyse into a spreadsheet tab.

	A	B	C
1	Food Costs		
2		Oct-22	Sep-22
3		12-mon % change	12-mon % change
4	Meat	5.5	7.6
5	Dairy products	10.6	9.7
6	Eggs	13.8	12.1
7	Bakery products	15.2	14.8
8	Rice	14.7	13.6
9	Fresh fruit	8.9	12.7
10	Fresh vegetables	11	11.8
11	Sugar and syrup	12.5	24.3
12	Coffee	13.6	16.5

Add these columns for analysis:

	A	B	C	D	E	F	G
1	Food Costs						
2		Oct-22	Sep-22				
3		12-mon % change	12-mon % change		A	B	A and B
4	Meat	5.5	7.6				
5	Dairy products	10.6	9.7				
6	Eggs	13.8	12.1				
7	Bakery products	15.2	14.8				
8	Rice	14.7	13.6				
9	Fresh fruit	8.9	12.7				
10	Fresh vegetables	11	11.8				
11	Sugar and syrup	12.5	24.3				
12	Coffee	13.6	16.5				
13		count		Prob			
14							
15						P(A|B)	

ORANGE CELLS

- First, count the number of things in the dataset.
- In C13: =COUNT(C4:C12)

BLUE CELLS

- Then, pick your A Event and B event. For me, A was Oct's values over 11 and B was September's values were over 11.
- I want to find the probability that Oct was over 11%, given that Sept was over 11%.

	A	B	C	D	E	F	G
1	Food Costs				Oct 22 Over	Sept 22 over	
2		Oct-22	Sep-22		11	11	
3		12-mon % change	12-mon % change		A	B	A and B

GREEN CELLS

- Then, build your formulas for your events:
- In E4: =B4>E2 (note: E2 was the value I picked for Oct)
- In F4: =C4>F2 (note: F2 was the value I picked for Sept)
- In G4: =AND(E4:F4)
- Fill down.

	A	B	C	D	E	F	G
1	Food Costs				Oct 22 Over	Sept 22 over	
2		Oct-22	Sep-22		11	11	
3		12-mon % change	12-mon % change		A	B	A and B
4	Meat	5.5	7.6		FALSE	FALSE	FALSE
5	Dairy products	10.6	9.7		FALSE	FALSE	FALSE

YELLOW CELLS

- Count up the columns and determine the probability of each event.
- In E13: =ROUND(COUNTIF(E4:E12,TRUE)/C13,3)
- For F13 and G13, fill right.

	A	B	C	D	E	F	G
1	Food Costs				Oct 22 Over	Sept 22 over	
2		Oct-22	Sep-22		11	11	
3		12-mon % change	12-mon % change		A	B	A and B
4	Meat	5.5	7.6		FALSE	FALSE	FALSE
5	Dairy products	10.6	9.7		FALSE	FALSE	FALSE
6	Eggs	13.8	12.1		TRUE	TRUE	TRUE
7	Bakery products	15.2	14.8		TRUE	TRUE	TRUE
8	Rice	14.7	13.6		TRUE	TRUE	TRUE
9	Fresh fruit	8.9	12.7		FALSE	TRUE	FALSE
10	Fresh vegetables	11	11.8		FALSE	TRUE	FALSE
11	Sugar and syrup	12.5	24.3		TRUE	TRUE	TRUE
12	Coffee	13.6	16.5		TRUE	TRUE	TRUE
13		count	9	Prob	0.556	0.778	0.556
14							
15						P(A|B)	0.714653
16							

PINK CELLS

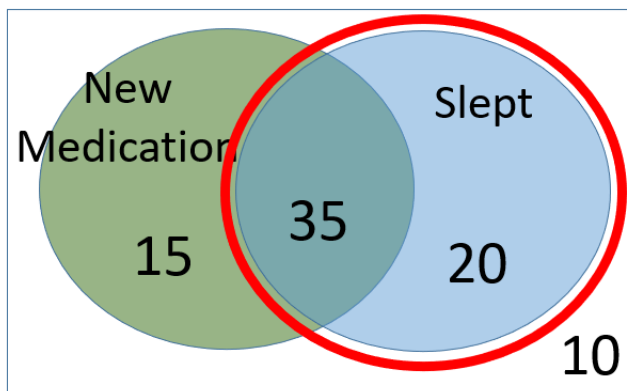The conditional probability formula was:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

So, we need to grab the cells with (A and B) and divide it by the cell with (B).

In G15: =ROUND(G13/F13,3)

This is the probability that the food will have a 12-month percentage increase over 10% in October, given that it was over 10% in September.

A useful slide from the Conditional Probability Lecture:

Given that they slept well, what was the probability they had medication?

$$= \frac{35}{55} \quad = 63.6\%$$

Given that they slept well, what was the probability they had a placebo?

$$= \frac{20}{55} \quad = 36.3\%$$

Because the people with the medication were more likely to sleep well, we can tell that the medication works better than a placebo.