

Correlation Analysis 101

A Farm to Table Overview

Presenter...



Esther
LaVielle

Participants who attend the full webinar and complete a quiz will receive a Correlation Analysis Certificate from QuestionPro





About QuestionPro

Overview

Founded: 2005

Employees: > 100

Customer Renewal Rate: +90%

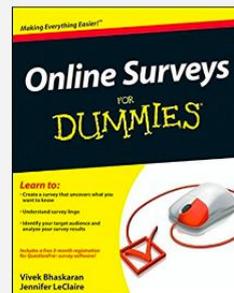
Transition from Non-QuestionPro platform: 70%

Implementation Success Rate: 100%



United States, Mexico, Germany, India, Dubai, Singapore

Publications



careerbuilder®

Adecco Group

Ivy League Magazine Network

groupm

CDC

StubHub!

Poibelly SANDWICH SHOP

FERGUSON
Nobody expects more from us than we do™

zynga

THALES

ROKU

Agenda

1

What is Correlation Analysis?

What's the Difference Correlation & Regression Analysis

2

Whys and Pitfalls of Correlation Analysis

3

When to Use Correlation Analysis

4

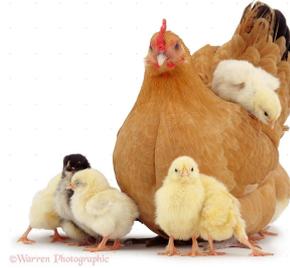
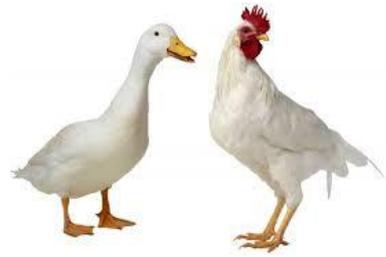
Analytics & Best Practices

5

Setting Up Correlation Analysis in QuestionPro + Demo



What is Correlation Analysis?



- Correlation analysis studies the closeness of the relationship between two or more variables
- It is used to describe the linear relationship between two continuous variables
- For market research, it is a bivariate analysis that measures the strength of association between two variables and the direction of the relationship

Which farm animals are more correlly related than others?

What is Correlation Analysis?



- A positive correlation (+) means that as one variable increases, the other does too.
- A negative correlation (–) means that when one variable increases, the other one decreases.

Example: Farmers who utilize more farming technology (var1) can farm larger field sizes (var2)

The Difference between Correlation Analysis and Regression Analysis

Comparison	Correlation Analysis	Regression Analysis
Definition	Determines co-relationship or association or absence of a relationship of two or more variables	Predicts the value of the dependent variable based on the known value of the independent variable, assuming that average mathematical relationship between two or more variables
Use Case	To represent linear relationship between two variables	To fit a best line and estimate one variable on the basis of another variable.
Indictates	The extent to which two variables move together strength-wise	Regression indicates the impact of a unit change in the known variable (x) on the estimated variable (y)
Objective	To find a numerical value expressing the relationship between variables -1, 0, +1	To estimate values of random variable on the basis of the values of fixed variable

Correlation Analysis vs Regression Analysis

Example for Correlation Analysis:

Farmers who utilize more farming technology can farm larger field sizes

Example for Regression Analysis:

Based on what farmers made in years past, they can estimate their total earnings for the following year



Why Correlation Analysis?



1

Points market researchers towards better questions to ask

2

Allows us to identify strong to weaker relationships between variables

3

Can be used for drill down analysis of specific segments of respondents

4

Strong correlations may be as close as market researchers can get to establishing that critical variables underly other variables in question



Pitfalls of Correlation Analysis

1

Correlation **does not** imply causation

2

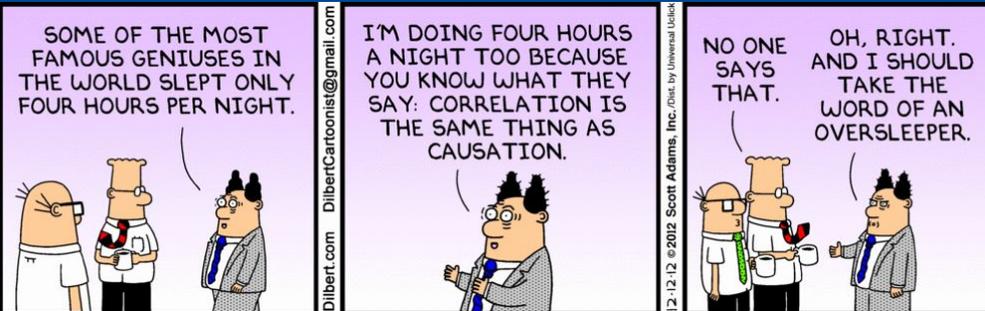
It does not prove that one variable can affect another variable

3

It's easy to get carried away - Correlation analysis is the most abused statistical measure

4

Correlation only measures the strength of linear relationships and does not necessarily imply a relationship between the variables



When to Use Correlation Analysis

When you want to determine the possible association between variables

Q: How much of a relationship is there and what is the strength of the relationship?

Measure **Correlation Coefficient**: A number between **-1 and +1** calculated so as to represent the linear dependence of two variables or sets of data



VS.



When to Use Correlation Analysis

Check for meaningful or a fluke relationships

Correlation analysis will help with testing the significance of the relationship

Significance testing of correlation analysis should be paired with:

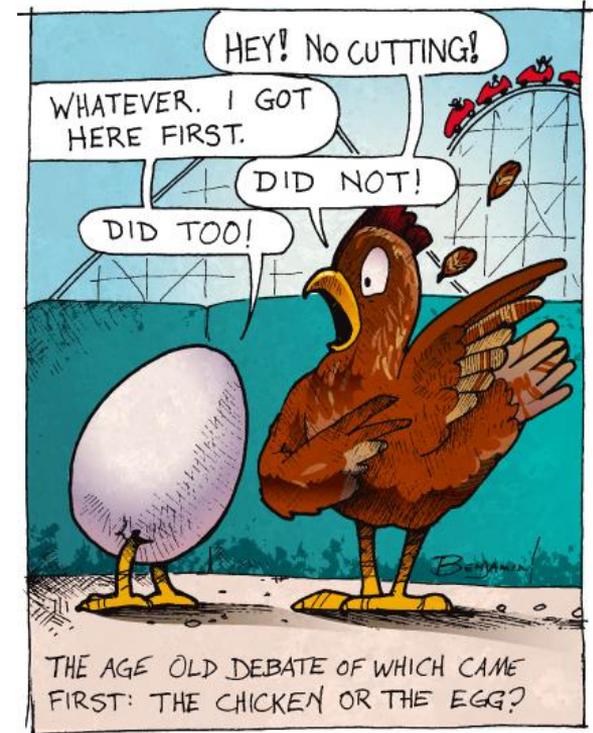
- **Sample Size** - the larger the better
- **Weighted or balanced data** - based on profile and other responses collected
- **Number of questions & responses** reviewed at the same time



When to Use Correlation Analysis

Helps to analyze the cause-and-effect relation...if there is one

- But correlation *does not* imply causation! Beware!
- We are focused on strength of relationships only
- A correlation between variables does not automatically mean that the change in one variable is the cause of the change in the values of the other variables
- A causal relation between two variables exists if the occurrence of the first variables causes the other variable to change - know the difference!

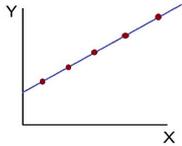


Correlation Analysis

Analytics Overview



Analyzing Correlation Results: 4 Methods



Scatter Diagram Method

Based on the study of graphs, the values for each pair of a variable is plotted on a graph and observed for degree of correlation denoted as "r"

- Simplest method to study the correlation - observational
- The more points scattered over the chart, the lower the correlation
- The more the points plotted are closer to the line, the higher is the degree of correlation

$$r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2} \sqrt{\sum(Y-\bar{Y})^2}}$$

Where, \bar{X} = mean of X variable
 \bar{Y} = mean of Y variable

Pearson Correlation Coefficient

Mathematical method used to calculate the degree and direction of the relationship between linear related variables

- The most extensively used quantitative method to measure correlation
- Denoted as "r"
- Correlation always lies between -1, 0 +1
- Variables and linear & independent of each other

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

Spearman Rho

Is a non-parametric test used to measure the strength of association between two variables

r = +1 means a perfect positive correlation and the value r = -1 means a perfect negative correlation.

- Also known as Spearman rank correlation, the test does not carry any assumptions about the distribution of the data
- Substitute all values with their ranks - based on # of responses
- Spearman's rho test uses ordinal data so show preference

$$r = \sqrt{b_{xy} \times b_{yx}}$$

Method of Least Squares

Mathematical method that tells the degree of correlation between the variables by using the square root of the product of two regression coefficient

- Square root of x on y and y on x
- AKA Line of Best Fit
- Calculate the mean of the x values and the mean of the y values

Best Practices

Setting up to Correlation Analysis



1

Use close-ended question types

All QuestionPro close ended questions can be used for correlation analysis and calculated using Spearman Rho

2

Keep track of a logic skip patterns used within the survey

3

Keep questions simple to reference for analytics

4

Keep answer options short - if possible

5

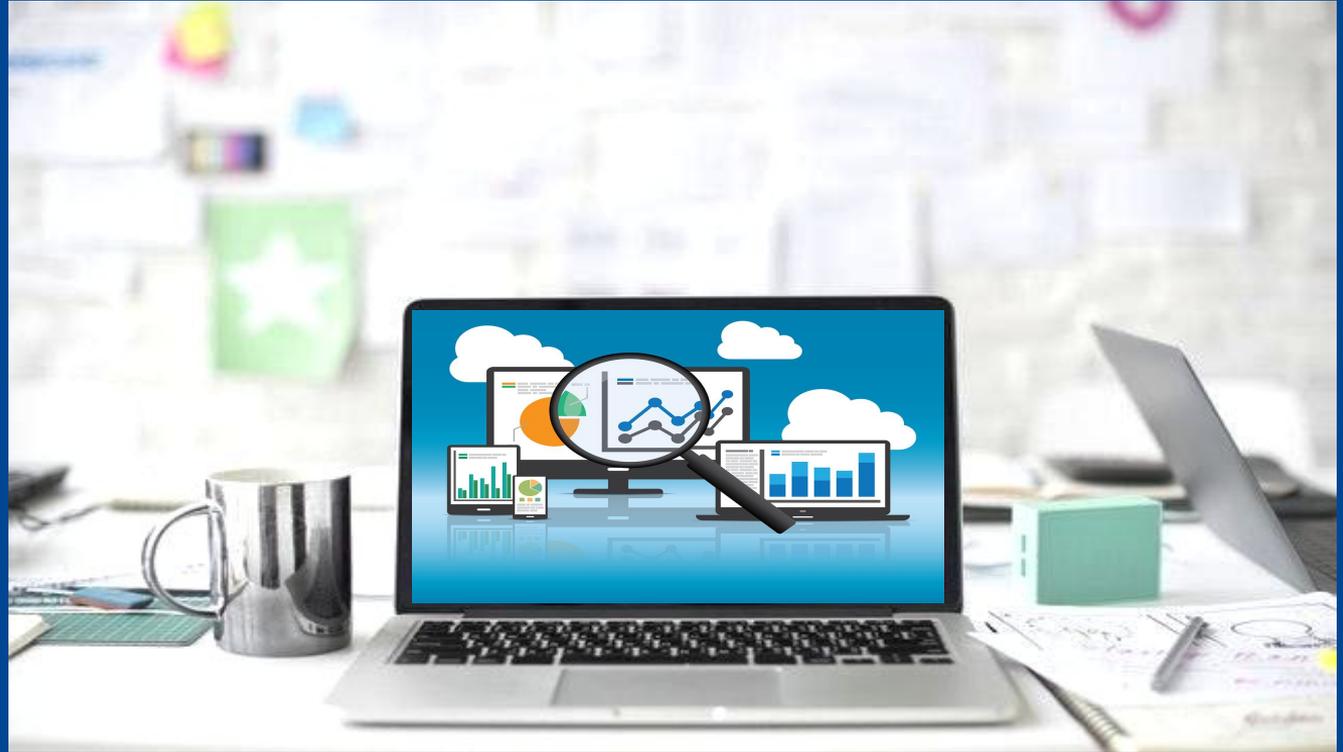
Is this an ongoing or short-term project?

Track changes of correlation results over time for ongoing projects



Demo

Correlation Analysis



Example Correlation Analysis



2018 Market Research Trend Survey by QuestionPro

- Wanted to know people's thoughts on buzzwords and trends that impact the market research industry
- What's most important for their business in 2018
- Had 598 views, 171 started, 110 completed = 64% completion rate



Correlation Analysis Steps



Create survey in QuestionPro with best practices noted in slide 14, field survey to desired respondent population



2018 Trends in MR

83%

[Back](#) Questions marked with a * are required [Exit Survey](#)

How important are the following for your business in 2018?

	1	2	3	4	5	6	7	8	9	10
* Building customer connections	<input type="radio"/>									
* Increase engagement with prospects and existing customers	<input type="radio"/>									
* Validation research to confirm qualitative and quantitative feedback	<input type="radio"/>									

[Next](#)

Online Survey Tool Powered by QuestionPro

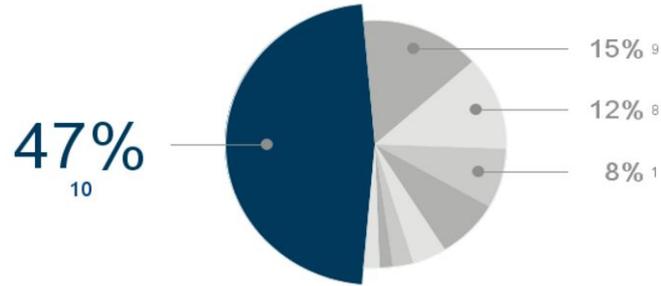
Correlation Analysis Steps



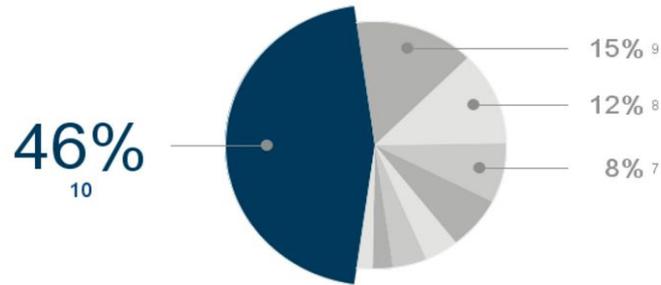
Analyze given data using QuestionPro Tools or other Analytic tools such as SPSS, etc.



Building customer connections



Increase engagement with prospects and existing customers



Validation research to confirm qualitative and quantitative feedback



Correlation Analysis Steps



The screenshot shows the SurveyMonkey interface for a survey titled "Webinar > 2018 Trends in MR". The navigation bar includes "Edit", "Send", "Reports", "Manage Data", "Mobile", and "Integration". The "Reports" menu is open, showing options: "Basic", "Correlation Analysis", "Choice Modelling", "Text / Media", and "Custom". The "Correlation Analysis" option is highlighted with a blue box. Below the menu, the "Correlation Analysis" section is visible, showing two questions selected for correlation: "First Question" and "Second Question", both set to "2. [Q6] How interested are you in watc...". A "Correlate All Questions" toggle is turned off. A blue button labeled "Generate Spearman Coefficient" is at the bottom right.

- Select Reports > Advanced > Correlation Analysis
- Select First and Second Question to be used for correlation analysis
- OR - select Correlate All Questions
- Click “Generate Spearman Coefficient”



Analyzing Results: Correlation Analysis

Correlation Analysis (Spearman rho) [?](#)

First Question 2. [Q6] How interested are you in watc...

Second Question 3. [Q6] Here is a list of buzzwords th...

Correlate All Questions

Generate Spearman Coefficient

Correlation Result [?](#)

#	2. [Q6] How interested are you in watc...	3. [Q6] Here is a list of buzzwords th...
2. [Q6] How interested are you in watc...	1.0	0.15
3. [Q6] Here is a list of buzzwords th...	0.15	1.0

Results show a **positive correlation** (+0.15) between market researchers who are interested in watching market research trends (51% very interested) and their feelings about market research buzzwords (50% says buzzword trends haven't arrived, but still on its way).



Analyzing Results: Correlation Analysis

Correlation Analysis (Spearman rho) ?

First Question 2. [Q6] How interested are you in watc... ▾

Second Question 3. [Q6] Here is a list of buzzwords th... ▾

Correlate All Questions

Generate Spearman Coefficient

Correlation Result ?

#	2. [Q6] How interested are you in watc...	3. [Q6] Here is a list of buzzwords th...	5. [Q6] How important a... Building custom...	5. [Q6] How important a... Increase engage...	5. [Q6] How important a... Validation rese...
2. [Q6] How interested are you in watc...	1	0.15	-0.12	-0.26	-0.18
3. [Q6] Here is a list of buzzwords th...	0.15	1	-0.07	-0.16	-0.14
5. [Q6] How important a... Building custom...	-0.12	-0.07	1	0.72	0.55
5. [Q6] How important a... Increase engage...	-0.26	-0.16	0.72	1	0.59
5. [Q6] How important a... Validation rese...	-0.18	-0.14	0.55	0.59	1

Highest positive correlation between market researchers who wishes to improve building relationships and increase engagement with their customers (+0.72)

Lowest negative correlation is between the researchers who watch market research trends and their wish to improve their level of engagement with their customers (-0.26).



Recap & Next Steps



1

Correlation Analysis measures the strength of association and its direction

2

Correlation **does not** imply causation!

But it can point market researchers towards better questions to ask

3

QuestionPro uses **Spearman Rho** to calculate correlation coefficient, but data can be downloaded to analyze using other correlation methodologies

4

Get started with a QuestionPro license - Available on Corporate license and higher



Thank You!

Q & A Session

esther.lavielle@questionpro.com