



20 Analytical Tools That Every Market Researcher Should Know

Here's a look at some key statistical analysis tools that all researchers should have within reach, ranging from the basic foundational pieces ("that rings a bell from seventh grade math") to some more sophisticated tools ("graduate-level stat geek").

- 1 Mean
- 2 Median
- 3 Frequency Distribution
- 4 Histogram
- 5 Variance
- 6 Interquartile Range
- 7 Crosstab
- 8 Chi-Square
- 9 T-Test
- 10 Correlation
- 11 Regression
- 12 Multiple Regression
- 13 Supervised Bayesian Network
- 14 Factor Analysis
- 15 Bayesian Network Community
- 16 Perceptual Map
- 17 Cluster Analysis
- 18 Conjoint Analysis
- 19 MaxDiff
- 20 TURF

1 MEAN

A simple mathematical average. It's a quick and easy way to determine the central tendency of your data set. Use it to get a benchmark of "typical" behaviors in your population.

2 MEDIAN

The value where half the data lies above and half the data falls below. Like the mean, it can give you an idea of central tendency, except a median won't be distorted by outliers. It can help you identify typical benchmarks, without as much influence from extreme behaviors.

3 FREQUENCY DISTRIBUTION

Measures how dispersed the data is around the mean. It can help you better understand exactly how statistically significant your data is.

4 HISTOGRAM

A bar chart that tallies how many responses take on each value (the frequency distribution in summary chart form). When the number of distinct response values is large, it's a great way to visualize "normal" behaviors and tendencies in your data.

5 VARIANCE

Measures how dispersed the data is around the mean. It's useful in statistical significance testing, which can help you tell whether results from two groups or waves of data are genuinely different or just appear different due to sample sampling error.

6 INTERQUARTILE RANGE

Like when Olympics judges throw out the highest and lowest scores, this identifies where the middle half of the data falls. It tells you something similar to the variance, but without getting bogged down in complex statistical concepts.

7 CROSTAB

A table that crosses responses to two separate questions. It shows you how the two responses relate to each other. Use it to compare sub-groups or to discover overlaps between target demographics or behaviors.

8 CHI-SQUARE

A statistic applied to a crosstab that tests whether the distribution of responses to one question varies depending upon responses to another. Use it to determine whether the results from one subgroup or wave of data truly follow a different pattern from another subgroup's results or simply appear to as a result of random variation.

9 T-TEST

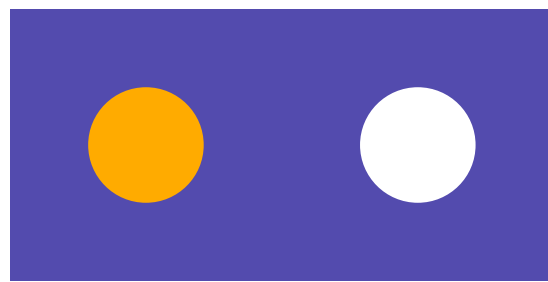
Determines whether your sample results (means or proportions) from two sub-groups represent a true difference between the populations they were drawn from. The larger the "t-score," the greater the likelihood that the groups are truly different.

10 CORRELATION

A statistical analysis that indicates the strength of relationship between measures (how consistently does one move together with the other), or whether a (linear) relationship exists at all. Coupled with knowledge about the category or situation, it can be used to make assumptions about causality.

11 REGRESSION A.K.A. BIVARIATE ORDINARY LEAST SQUARES (OLS)

A descriptive technique quantifying how much a dependent variable tends to change given an observed change in a single independent variable. Use it to get a relative sense of impact a marketing or product decision may have on customer engagement.



12 MULTIPLE REGRESSION

A predictive model that quantifies the effect on one dependent variable based on changes to two or more independent variables. A multiple regression analysis is useful to isolate impact uniquely attributed to a given variable in the context of others or measuring how multiple decisions can impact customer engagement when executed simultaneously (more advanced techniques are required when independent variables are badly inter-correlated, as is frequently the case with survey data).

13 SUPERVISED BAYESIAN NETWORK

A predictive model that quantifies the effect on the dependent variable while taking into account that the independent variables impact each other. It's helpful to understand how a single marketing or product decision can set off a chain reaction of outcomes for your business.

14 FACTOR ANALYSIS

Sometimes you can be swimming in too much data. Factor Analysis is a data reduction process where you can boil down your data to a smaller number of underlying dimensions. Use it to synthesize your results and simplify the story.

15 BAYESIAN NETWORK COMMUNITY

A group of closely inter-related perceptions in a Bayesian Network. It helps you understand the path by which changes in one perception can ripple through other perceptions to have a cumulative impact on your business, and inform which combination of levers you should try to pull.

16 PERCEPTUAL MAP

A visual display of how a set of brands are perceived in the market. It can be used to more clearly identify a brand's strengths and weaknesses relative to competitors, as well as "white space" in the market.

17 CLUSTER ANALYSIS

The process of grouping respondents into “clusters” that they are more similar to than they are to members of other “clusters.” It’s a great way to arrange potential customers into distinct market groups, and select your target(s).

18 CONJOINT ANALYSIS

A technique used to determine the importance of product features and preferred levels of those attributes. It’s very useful for understanding how shoppers choose products and what trade-offs they’re willing to make in order to inform product and pricing decisions.

19 MAXDIFF

A question format used to understand preference or importance scores for multiple items without overwhelming the respondent with a long list of features. This question format asks a few questions about “most” and “least” preferred or important items within systematically chosen subsets of your list. Use it when you need to prioritize features, benefits or claims and want to avoid “everything’s important” or cross-cultural scale biases that can distort the results.

20 TURF (TOTAL UNDUPLICATED REACH & FREQUENCY)

A technique used to determine the importance of product features and preferred levels of those attributes. It’s very useful for understanding how shoppers choose products and what trade-offs they’re willing to make in order to inform product and pricing decisions.