ComMetrics
Benchmark smarter. Perform better.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Just the facts

Title	**Raise your hand if you hate research** **=> ComMetrics Intel eBook series #2011-01**
Author	Urs E. Gattiker – benchmark your blog(s) at My.ComMetrics.com – powered by CyTRAP Labs
Short description	The web provides an array of resources to help with strategic planning and business policy so this report takes you to research methods school. It expands your toolbox by providing you with new insights, tricks and tips on how to use intel offered by consultants and research institutes to improve your risk analysis and social media marketing.
Version number	2011-01-01 ComMetrics Intel eBook series #2011-01

Citation: Gattiker, U. E. (January 4, 2011). Raise your hand if you hate research. **ComMetrics eBook series #2011-01**, Zurich: CyTRAP Labs GmbH.

Get a FREE login to download: http://University.ComMetrics.com/download/21/ (250 KB), Last access: 2011-01-01.

Or download from Facebook (click on Like to access our EXCLUSIVE page): http://www.facebook.com/pages/ComMetrics/188946538373?v=app_6009294086, Last access: 2011-01-01.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

ComMetrics
Benchmark smarter. Perform better.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

ComMetrics eBook series

The ComMetrics eBook series helps readers benchmark smarter to improve performance. Topics and issues may range from risk management to social media monitoring to blog benchmarking.

An important focus is to provide insights that empower readers to develop approaches customized to their own needs. In turn, these insights can be applied in for-profit as well as not-for-profit organizations. It provides the framework to make any necessary adjustments in the face of important organizational context issues that affect you (e.g., industry, resources and other key parameters).

This book begins a series on **business strategies for information value** that will continue throughout 2011, so watch for the next book coming shortly.

It is not the strongest of the species that survive, nor the most intelligent, but the one most responsive to change.
– Charles Darwin

*You can **get these books for FREE** by registering yourself at My.ComMetrics.com - **benchmark your blog** => improve performance and then visit university.commetrics.com/?page_id=188 to download these e-books*

These days, even the most successful organizations face new and unexpected challenges. Traditional approaches to marketing and customer relations that worked in the past are no longer sufficient to meet the demands of a changing marketplace.

The ComMetrics eBook series provides the latest insights, strategies, and tools needed to navigate uncertain times and position your company for the next evolution of social media and the Internet.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

What is the issue?

This ComMetrics Intel eBook on **research methodology** explains why worrying about things like getting your theory, methods and procedures right is even necessary.[1]

The book guides you through the maze of research terminology, data collection methods and research tools. In turn, you can decide better which ones to use, and how to use them most effectively to conduct research or evaluate the quality of a consultant's research report.

With an objective approach and clear, straightforward language, Urs E. Gattiker, aka 'the social media marketing and metrics scientist', shows you how to plan and implement your research venture intelligently, to provide information that assists your decision-making.

The challenge is to make sure that the research you use to evaluate your company's social media marketing strategy passes the quality test.

There is nothing worse than making a case for more effort and resources to be spent on Facebook marketing or the corporate blog when you base your strategic review on consulting reports and research papers based on questionable information.

This book will help you choose the best – and avoid the worst – of the opportunities offered by research, thereby helping you take better advantage of social media for your business or organization.

1. Understand the features, functionality, and protocols that are used to collect data from consumers, customers and voters.

2. Choose the tools and research tactics most relevant to your goals.

3. Learn how to determine sample size and confidence interval and why margin of error does make a difference.

4. Get clear-cut explanations of the methods you need to assess whether a research report can help you make a better decision.

[1] You can receive the ComMetrics eBook series within 48 hours by registering yourself at My.ComMetrics.com or University.ComMetrics.com and download e-books: university.commetrics.com/?page_id=188

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Chapter 1: Types of business research

Something strange is happening. Your advertising does not work anymore, at least not as well as it used to. But not everything is just doom and gloom: you don't have to outspend the biggest companies if you know your target audience better than the next guy.

For this reason, organizations either purchase research from others (e.g., Nielsen marketing research) or conduct their own. Larger companies also offer a lot of free business research, in the hope that it might convince you to purchase their product.

We use the term business research because all research techniques (e.g., case analysis, observations and surveys) can be used in business settings.

In order to better find our way through this maze, we explain the three prevalent types of research that you should know about, and address qualitative and quantitative matters regarding your methodology.

Exploratory studies

This type of research is conducted to clarify and define the nature of a problem that requires research in order to gain a better understanding of the dimensions of the issue(s). An exploratory approach is often chosen because knowledge about the phenomenon is limited and research that addresses similar issues is not available.

To some degree, social research is always exploratory and generally requires extensive fieldwork of various kinds, such as observing pupils in an education setting or workers commuting to and from work.

Researchers explore the group, process, activity or situation they want to examine when they have little or no scientific knowledge about it. Nevertheless, they may have reason to believe it contains elements worth discovering.

Conducting such research effectively requires open-mindedness about where to find relevant data and flexibility in retrieval methods.

Descriptive research

While exploratory research investigates the issues you have no clear insight of but may have a hunch about, descriptive studies are usually undertaken when you have some idea about what you will find. For instance, a pollster might do quarterly polls about consumer attitudes regarding current economic conditions.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Descriptive research usually illustrates the characteristics of a population, such as finding out how clients focus on the most critical risks regarding information security and data management.

Most countries have a national statistical agency such as Statistics Canada. Such an agency will collect quarterly or more frequently data to answer questions of who, what, when, where, and how regarding the country's labor markets. This yields information such as the country's unemployment rate, new jobs created in the last quarter and so forth.

Descriptive research is based on some previous understanding of the nature of the research problem. Hence, we know the way companies felt about the security risks of social media use in 2009, but may conduct research to learn if or how those feelings have changed in 2010.

For instance, the fact that increasing numbers of people have smartphones with mobile Internet may change the types of threats companies must concern themselves with, since workers may use their private mobile Internet connection to visit Facebook during breaks. This situation could necessitate a change in a company's social media risk management strategy for 2011 and a descriptive survey could be helpful.[2]

Causal research

Generally, exploratory and descriptive researches precede cause-and-effect relationship studies. For example, in causal studies we may want to know if a prediction about the influence of price on sales is true or not.

In order to determine causality, it is important to ensure that the variable assumed to cause change in others remains constant before measuring any changes.

This is sometimes very complex, especially if several variables are involved. Therefore, in order to predict a relationship between gender, price and perceived quality of a product for example, causal studies often create statistical controls to establish 'contrast groups', such as comparing men and women.

[2] For instance, Ernst & Young's Annual Global Information Security Survey for 2010: www.ey.com/GL/en/Services/Advisory/IT-Risk-and-Assurance/13th-Global-Information-Security-Survey-2010---Survey-parameters

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Quantitative or qualitative research?

Quantitative research tends to be objective, deductive, based on numbers and can be generalized. By contrast, qualitative research tends to be subjective, inductive, uses words and cannot be generalized.

Both quantitative and qualitative researches explain phenomena, such as why people prefer one smartphone over another.

In quantitative research, one collects numerical data that is analyzed using mathematical methods, specifically statistics (see Chapter 3 and 4).

Quantitative research has some disadvantages, such as requiring a large sample group in order to make applying inferences about that sample to the entire population reasonable (e.g., all consumers in a region). Margin of error and confidence interval specifically address this issue (see Chapter 5).

Because exploratory research is used for problems that have not yet been clearly defined, it makes better sense to use a qualitative approach than a quantitative one, since there may not be anything to quantify.

Qualitative and quantitative research may use descriptive statistics to gain a better understanding of the data set by describing its basic features, and providing simple summaries about the sample and the measures that were used. An example might be the percentage of people who answer 'Yes' to the question, "Do you use the Internet outside work?" versus the percentage who answer 'No'.

In a qualitative study, descriptive statistics are used to summarize information about the sample. In a quantitative study, summarizing information about the sample is used to infer information about the population (see further below).

Business research and you

Your customers and your competition are already using business research to help them better serve their clients. The following chapters will teach you much of what you need to know to pick the right tools and get started.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Chapter 2: Data collection

Once we conclude whether to conduct some kind of study we need to decide how to collect the necessary data and describe what phenomenon we hope to find answers for before we start our research. For instance, why do people spend an average of 30 seconds on the landing page of our website instead of the two-minute objective we have set? Our research question will guide how we collect what kind of data.

Data collection is both an art and a science. While much of the practice is based on well-researched factors, many decisions about data collection are guided by common sense and practicality.[3] Therefore, you need to know which method costs less and, most importantly, how using one method over another may affect the validity and reliability of your data. There is nothing worse than basing decisions on incorrect data.

Why collect data?

Primary data collection entails gathering your own using different methods as described below. It is important to remember that this makes the data unique to you and your research or work. Accordingly, until it is published no one else has access to it.

Secondary data have been collected and recorded by someone else prior to and for purposes other than your current needs. For instance, a hotel's guest history system may have information about clients such as gender and postal address. One reason to systematically collect information in an organization could be to evaluate some activity or program, including its outcomes, personnel, products and so forth. As mentioned earlier, a national statistical agency collects information about consumer inflation and economic growth.[4]

Other examples include wanting a better understanding of how your social media marketing strategy supports customer engagement, how efforts to raise awareness help combat AIDS, or how a literacy program improves reading scores.

Before collecting data to evaluate a program or test hypotheses, we must spell out what we intend to evaluate, and how.[5] The easiest way to do this is by formulating some research

[3] For instance, various approaches can be used to measure poverty but each one has advantages and disadvantages, such as being more or less resource-intensive: go.worldbank.org/6Y1HRHQCN0

[4] More information about secondary data: brent.tvu.ac.uk/dissguide/hm1u3/hm1u3text2.htm

[5] How to set up an evaluation program (no date or author): ohioline.osu.edu/b868/pdf/b868.pdf

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

questions we would like answered, which will also help reduce the risk of collecting incorrect or inaccurate data that will yield invalid results.

How to collect data

In the social sciences (e.g., market research, political science, anthropology), the **dominant data collection method is interviews and questionnaires**. Unfortunately, no method of data collection is without bias.

For instance, one could be surveying households about how much money is spent each week on digital media downloads via the Internet. Respondents might give inaccurate estimates, either unintentionally or to place themselves in a better light with the researcher. The latter is likely to happen if the issue is sensitive and personal. For instance, when asking information about income, sexual promiscuity, alcohol consumption and so forth, people may not want to tell a pollster the truth. The same can happen in an exit poll during an election, when people might try to give the 'right' answer.

In both cases, the subject must cooperate. But there are ways and means to supplement and cross-validate survey or interview data with measures that do not require participant cooperation.

One way to confirm how much money people spend on digital media downloads might be to check their credit card records or download statistics from their Internet service provider (ISP). Besides checking digital traces (of course, only with the participant's prior consent), one can also observe people for a period of time to ascertain whether they really exercise as much as they claim to.

Data collection methods

Other forms of data collection include focus group interviews, observation, case studies, diaries, critical incidents and portfolios.

For instance, focus group interviews may include anywhere from four to twelve people who are not familiar with one another but all share certain characteristics relevant to the study's questions.[6]

The type of person invited to participate in focus groups (e.g., busy professionals) may inherently reduce your sample size. Hence, non-response bias must be addressed (see

[6] Learn more about the advantages and disadvantages to focus groups: nnlm.gov/evaluation/workshops/measuring_your_impact/DataCollectionHandout.pdf

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

below). Whatever data collection method works best for your project may depend on how many subjects – sample size – are needed and the resources available.

Sample size

Choosing a new corporate brand image or logo requires a more in-depth study with many more subjects than ensuring your friends like the menu you have planned for a dinner party (e.g., who is allergic to garlic or nuts).

Sample size is also important when calculating margin of error, and ensuring that it is large enough to allow looking at groups.

For instance, if we want to look at women with back pain who work part-time, we must compare women and men (two groups), those with or without back pain (two groups), and people working full-time, part-time or not at all (three groups). This already requires a sample size of somewhere around 300 people in the hope that we end up with a sub-sample of at least twenty women with back pain who work part-time.

Accordingly, it is necessary to address whether it is important to collect data about sub-groups such as different types of customers beforehand. The more heterogeneous a population, the larger the sample size required to achieve a certain level of precision (formulas can be applied to calculate this).[7]

[7] More about strategies for determining sample size: edis.ifas.ufl.edu/pd006
Sample size calculator: www.surveysystem.com/sscalc.htm

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Chapter 3: What kind of data do we need?

The researcher must first decide what kind of data collection method will be used (i.e. survey, focus groups, case studies or archives). Then, sample size is addressed, but this is impossible without determining the type of data required to investigate the chosen phenomena.

Why do we care about data types?

If our data does not meet our needs or suit our method(s) of analysis, we are lost. Each type of data builds on the previous category (e.g., ordinal data is also nominal, etc.).

There are four types of data in social research: nominal, ordinal, interval and ratio.

- **Nominal data** (also called categorical) is unordered and usually expressed through labels, such as a respondent's sex (male or female) or a type of rock (igneous, sedimentary or metamorphic). Such variables allow information to be grouped, but no one grouping has any more value than another (see also correlation and multiple regression below).

- **Ordinal data** requires ranking, but the differences or ratios between values are not meaningful. For example, patients might be asked to rank their pain on a scale of 1 to 7. A score of 5 means more pain than a score of 3, and that is more than a score of 2. But the difference between 5 and 4 may not be the same as the difference between 4 and 3. The values simply express an order.

- **Interval data** can be ranked, *and* exhibits meaningful differences between values. For instance, the difference between 25 and 30 degrees Celsius is the same as the difference between 5 and 10 degrees Celsius.

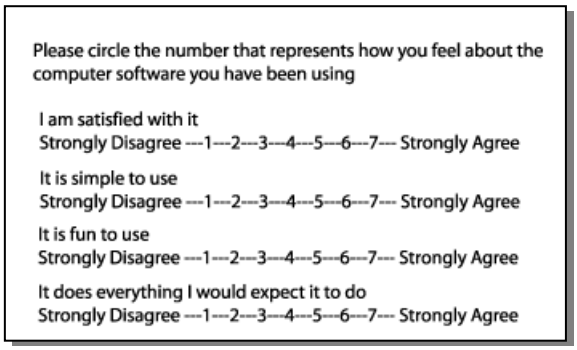
- **Ratio data** has an absolute zero. Variables like height, weight, and enzyme activity are ratio variables, but not temperature, because 0 degrees Celsius is still a temperature measurement.

Most business research, which collects data through interviews or surveys, includes several variable types, resulting in different data types. We may ask or note a subject's gender, a categorical variable, and also ask or note their height, a ratio variable.

We may also have ordinal and interval data, but the differences are not as clear-cut as they might appear.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

For instance, Likert-type scales are often used in organizational behavior or consumer research for respondents to indicate their level of agreement with a given statement, such as where 1 means total disagreement and 7 means complete agreement. Since there is no quantifiable difference between ranks on this kind of scale, it would be considered ordinal. However, this type of data is often treated as interval and subjected to certain statistical techniques (e.g., Factor Analysis).

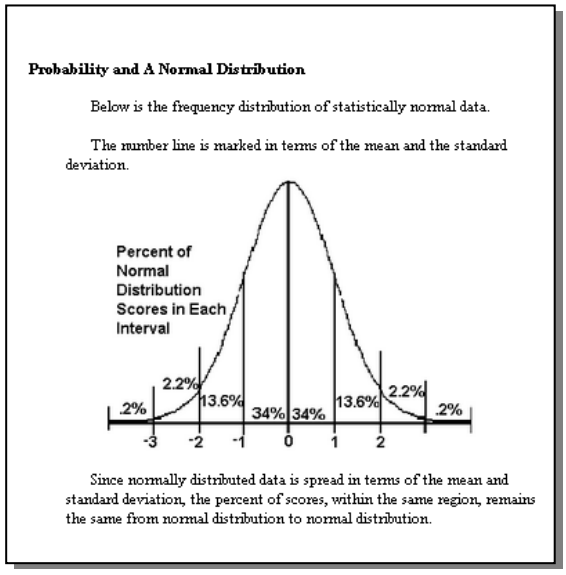


What kind of statistics can we use?

As outlined above, choosing the type of data to collect is important if we hope to draw any conclusions through the use of statistics.

Interval and ratio data are considered **parametric**, and used in conjunction with statistical tools in which distributions are predictable and often **normal**. Parametric statistics assume a normal underlying bell-shaped distribution, which is often forced through means of samples.

They are considered more powerful than **non-parametric** data (nominal or ordinal only), which do not assume any particular distribution and are used with non-parametric tools, such as the histogram or the run-test.[8]



Parametric statistics use a **Pearson Correlation Coefficient**, while non-parametric statistics use the **Spearman Rank Correlation Coefficient**.[9]

The Pearson Correlation Coefficient is the proportion of variation that can be explained, with a value of 1 denoting perfect correlation. It

[8] More information about the differences between parametric and non-parametric data: www.csse.monash.edu.au/~smarkham/resources/param.htm

[9] More information about correlation coefficients: changingminds.org/explanations/research/analysis/correlation.htm

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

is only used with parametric data, but the Spearman Rank Correlation Coefficient, a variation of Pearson, is used with non-parametric data after its conversion to rankings.

Once the researcher decides what method(s) will be used to gather data (e.g., survey, focus groups, case studies or archives), it is necessary to determine what type of data will be collected, which in turn will determine what kind of quantitative analysis can be done. For simplicity's sake and future reference, we have assembled Table 1 below.

Table 1 – What statistics can be used with which data

	Nominal	Ordinal	Interval	Ratio
Possible types of statistical analysis	non-parametric		parametric	
Frequency distribution, mode	Yes	Yes	Yes	Yes
Median, percentiles	No	Yes	Yes	Yes
Add or subtract	No	No	Yes	Yes
Mean, standard deviation, standard error of the mean, correlation, regression, analysis of variance	No	No	Yes	Yes
Coefficient of variation, logarithms	No	No	No	Yes

Note. Naming different kinds of variables or data helps prevent mistakes.

In other words, research may accumulate data about customers' satisfaction with their accommodations using a Likert-type scale and then use ordinal variables to group responses (e.g., women versus men or frequency of visits within the last 12 months).

In the case of a self-administered survey, asking the participant to mark the appropriate response on a scale is not the best way to go. Instead, it is advisable to record the absolute value, so the client might be asked, "How often have you stayed at this hotel in the last 12 months while on business?" The same could be asked about how often a purchase was made from a specific retailer in the last 12 months, but a company's loyalty program may also provide this information. Comparing survey responses with the loyalty program data allows the researcher to determine whether the sample of respondents is similar to the population.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Chapter 4: What do correlations tell us?

Understanding correlations and what they are used for is essential and seems particularly sensible since parametric statistics such as multiple regression, Anova (analysis of variance), and Manova (multiple Anova) all use the Pearson Correlation Coefficient.

A correlation is used to estimate the relationship between at least two variables. If one of the variables is ordinal and the other an interval or ratio variable, then we use the Spearman Rank Correlation Coefficient, also known as rho or rs.

Interval and ratio data necessitate using Pearson, which can fall anywhere between 0.00 (no correlation) and 1.00 (perfect correlation).[10] A correlation of -1 indicates a perfect negative linear relationship between variables, while a correlation of +1 means the opposite. Put differently, the larger one's feet, the higher the shoe size number required for the wearer to be comfortable.

Other factors such as group or sample size will determine if the correlation is significant. Generally, correlations above 0.80 are considered pretty high. In most social sciences context correlations between .10 and .50 are prevalent.

Correlations – assumptions

When examining survey or poll findings regarding risk, consumer attitudes and so forth we must verify whether it is based on correlations. This is done using statistical methods such as analysis of variance and multiple regression, in all of which correlations play a central role.

Correlational studies mean that when an association between two variables (A and B) is observed, six possible explanations are conceivable.

1. A influences B.
2. B influences A.
3. The relationship between A and B is reciprocal (combination of 1 and 2).
4. A and B are both influenced by a third variable, C.

[10] More about Pearson: davidmlane.com/hyperstat/A62891.html

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

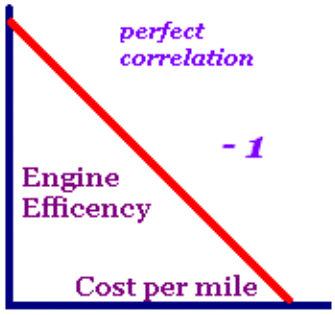
5. A combination of the previous four explanations (e.g., A influences B and B influences A and C influences A and B).
6. The correlation between A and B is partly or entirely spurious (the result of errors in definitions, measurement, or data aggregation).

In addition, further variables can influence only A or B, but may have no effect on the relationship between A and B.

It is also important to remember that historical processes can change relationships – due to changes in the types and content of cultural and social practices relevant for education, competence development, marketing and innovation management (e.g., research and development), current influences may differ widely from those noted in the past.

Based on the above it is clear that correlation is nothing more than a statistical term that indicates whether two variables move together. When it snows, it tends to be cold; therefore these two factors are positively correlated.[11]

Sunshine and fog, meanwhile, are negatively correlated (-1.0 = perfect correlation as shown to the right), since they do not occur simultaneously.



This seems easy as long as only two variables are studied, but with 20 or more, things get a bit more complicated.

Regression analysis

We use regression analysis to sort out accumulated data and further illustrate the use of correlations by **artificially holding every variable constant except the two being focused on**, and demonstrating how they co-vary.

> **Question 1**: Does having many books in the home lead to increased academic achievement in your child?

Regression analysis cannot answer question 1, however, it can answer a subtly different one:

> **Question 2**: Does a child who grows up in a home with many books achieve more academically than a child who grows up in a home with few or no books?

[11] See www.currency-forex-trading-online.com/currency-correlation.htm

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

The difference between the above questions is the difference between **causality** (question 1) and **correlation** (question 2).

A regression analysis can demonstrate correlation but not cause, so this method alone cannot tell us whether it snows because it is cold, whether it is cold because it snows, or if the two just happen to go together.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Chapter 5: What is margin of error?

At this stage we have discussed what kind of research, using what method to collect data can be used, including choosing between qualitative and quantitative research. Depending on the type of data we collect, we can also choose between parametric or non-parametric statistics, but correlations play an important role regardless, as outlined in Chapter 4.

Now we focus on margin of error and level of confidence, whereby **level of confidence is a measure of how confident we are in a given margin of error** (people often use 95 percent). The margin of error and level of confidence depend on the sample size, NOT population size.

Sample selection

Sample selection is critical to figuring its margin of error based on confidence levels. Surveying customers (e.g., shoppers, industrial buyers) has been likened to taste-testing soup – a few spoonfuls will tell you what the whole pot tastes like. In turn, surveying a sample of your customers should give you a pretty good indication of how they rate such things as customer service, product quality and so forth.

The key to the validity of any survey is randomness. To illustrate, soup must be stirred in order for a ladleful to represent the whole pot. Similarly, when sampling a population, one must ensure that the group is mixed before respondents are chosen, so random selection is required.

Student samples

One of the most popular sample types is asking students to volunteer to participate in a laboratory exercise or fill out a survey in the professor's class.

The validity of using student participants has been debated for quite some time. Researchers increasingly seem to agree that if results from such a sample can be generalized for the real-life situations in question, using students as research subjects is perfectly acceptable.

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Asking students about their use of social media such as Facebook makes sense when focusing on this type of user group. Some exploratory research with students in the US has reported that using Facebook seems to negatively affect their grades.[12]

However, using a student sample limits the applicability of those findings for other Facebook user groups.[13]

Stratified random sampling

Besides finding the right sample (e.g., students or managers) for investigating a particular phenomenon, many studies also ensure they are stratified.

Often, certain factors divide the population into sub-populations (groups / strata) in which we may expect the measurement of interest to vary, such as retirees and teenagers. This must be accounted for when we select a sample from the population to ensure that it is representative, which is achieved by stratification.

Sampling carried out at random from each stratum of a stratified population is critical to enable making inferences from our sample to the general population.[14] When we sample a population with several strata (e.g., people with back pain between 30 and 40 years of age), we generally require that the proportion of each stratum in the sample should be the same as in the population.

In business research, stratified sampling is often used to reduce costs, get estimates of the population parameters for each sub-population of the target market and increase accuracy.

---

[12] This exploratory study reports that students who use Facebook for several hours each week seem to have lower grades than those who use it sparingly or not all (see Saturday): ComMetrics.com/?p=8983

[13] Bello, D., Leung, K., Radebaugh, L., Tung, R. L., van Witteloostuijn, A. (2009). EDITORIAL - From the Editors: Student samples international business research. **Journal of International Business Studies**, **40**, 361–364. Last access 2010-12-04: www.palgrave-journals.com/jibs/journal/v40/n3/pdf/jibs2008101a.pdf

[14] More information about sampling: www.stat.yale.edu/Courses/1997-98/101/sample.htm

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

If a company wanted to find out how well-known its drinks brand is among people under the age of 30 and single parents with children under the age of 14, the company could select a proportion of each stratum in the sample that is the same as the general population. To further learn about the difference between apple juice users and energy drink aficionados or those who love cola, those would also have to be included in the sample.

Margin of error

How well the survey sample represents the population is gauged by two important statistics.

1. Margin of error
2. Confidence level

These tell us how well the spoonfuls from our soup represent the entire pot.

A survey may have a margin of error of plus or minus 3 percent at a 95 percent level of confidence, which simply means that if the survey were conducted 100 times, data would fall within a certain number of percentage points above or below the percentage reported in 95 of the 100 surveys.

In theory, in 95 cases out of 100, the results based on such samples will differ by no more than three percentage points from what would have been obtained by seeking out all clients (the whole population).

To illustrate, let us say that a university surveys its students a year after they graduate, and finds that 50 percent of the respondents consider the university's student career services 'very good'.

A **confidence level** of 95 percent, plus or minus 3 percent means that if this student or client **survey is conducted 100 times**, the percentage who consider career services 'very good' will **range between 47 and 53 percent most of the time** (i.e. 95 percent).[15]

Margin of error – the plus or minus 3 percent in the above example – decreases as the sample size increases, but because of diminished returns, this applies only up to a certain point.

A very small sample, such as 50 respondents, has a margin of error of about 14 percent. Conversely, a sample of 1,000 has a margin of error of 3 percent. Assuming it is at least

[15] Calculators for calculating required sample size to achieve a certain confidence level with a specified margin of error: www.raosoft.com/samplesize.html
www.comres.co.uk/star-calculator.html#
www.marketingcounsel.com/tools/margin_of_error_calculator2.htm

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

larger than the sample, the **size of the population** (the group being surveyed) **does not matter**. However, doubling the sample to 2,000, only decreases the margin of error from plus or minus 3 percent to plus or minus 2 percent.

Also, a 90 percent level of confidence may suffice in some instances, even though a 95 percent level is an industry standard. A 90 percent level can be obtained with a smaller sample, which usually translates into a less expensive survey. Obtaining a 3 percent margin of error at a 90 percent level of confidence requires a sample size of about 750, but for a 95 percent level of confidence, the sample required would jump to about 1,000.

This explains why most pre-election or national referendum polls report including more than 1,000 respondents.[16] Nevertheless, people may lie to pollsters, and for that matter, one should never pay too much attention to just one poll or survey.[17] Finally, sometimes research reports are rather unclear about how things work – Nielsen and others.[18]

Non-response bias

When participants fail to return a survey or fill it out completely, the results can affect the size and characteristic of the sample, which in turn can compromise the external validity of your research project.

Of course, in cases of self-administered surveys the researcher can send a second one to follow up with those who have not responded within a certain timeframe, which usually increases the response rate.

---

[16] How to calculate margin of error for any study, including examples:
www.wordiq.com/definition/Margin_of_error
stattrek.com/AP-Statistics-4/Margin-of-Error.aspx?Tutorial=Stat

[17] Poll position – How to read the statistics:
www.columbia.edu/cu/news/record/pollposition.html

[18] "*The survey polled more than 27,000 consumers across 55 countries throughout Asia Pacific, Europe, Latin America, the Middle East and North America. The sample has quotas based on age and sex for each country based on their Internet users and is weighted to be representative of Internet consumers with a maximum margin of error of ±0.6 %.*"
From "How People Watch: A Global Nielsen Consumer Report August 2010", p. 14.
Though it uses a stratified sample, the report is unclear about how people were selected while ensuring random sampling. Nor does it explain how the issue of non-response bias was dealt with, so one cannot re-calculate the margin of error.
Also, calculating across 55 countries with vastly different economies and Internet penetration cannot be considered 'best practice', for lack of a better term, and nowhere does the study address cross-national or cross-cultural differences that could affect interview responses as well. So, generalizing from this study is actually pretty much impossible – oops!

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Similarly, if someone is not home when an interviewer calls, they might try to reach the subject again on another day and/or at another time.

One can also compare data of respondents and non-respondents to see if there are differences regarding such data as gender, age, education. Differences indicate response bias and a need for caution when making inferences from the sample to apply to the wider population.[19]

Other approaches include interviewing a sub-sample of those that did not fill out the self-administered survey and those that did reply, using a sub-set of the overall questionnaire by telephone. Responses can then be compared to ensure there is no bias.[20]

Reliability and validity

Understanding bias (or lack of bias = validity) in one's research data is important for reaching conclusions. This means the questions must measure what they are supposed to and unless we can be certain of that, conclusions from data are incorrect.[21]

Measurement is likely still reliable on an incorrectly set bathroom scale, because if you stand on it twice, it will show you weighing pretty much the same. However, it cannot be trusted, since it is not valid if the zero is set wrong and may therefore show you weighing a bit less than you actually do.

There are also different types of variables that play an important role when performing statistics such as multiple regression: moderating, control or mediating.[22]

Here is another great online resource about research methodology, the process and the issues addressed above:

> Joppé, M. (not dated). The research process. Last access 2010-12-06: http://www.uoguelph.ca/htm/MJResearch/rp.htm

[19] See also: edis.ifas.ufl.edu/pd008
[20] Detailed explanation: papers.ssrn.com/sol3/papers.cfm?abstract_id=949815
[21] See also: ComMetrics.com/?p=5653
[22] For more information, see:
info.cytrap.eu/articles/haagen-dazs-or-magnum-ice-cream-methodology-artefacts-and-bias
blog.CyTRAP.eu/2007/09/12/3-enisa-awareness-raising-study-better-prevention-thanks-to-data-crunching/

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Chapter 5: Conclusion – why do we care?

This eBook tried to provide information regarding four important issues that must be addressed before using information from a research report to make strategic decisions:

1. Understand the features, functionality, and protocols that are used to collect data from consumers, customers and voters.
2. Choose the tools and research tactics most relevant to your goals.
3. Learn how to determine sample size and confidence interval and why margin of error does make a difference.
4. Get clear-cut explanations of the methods you need to assess whether a research report can help you make a better decision.

A short checklist that further helps assess the merits of any research study is found below. These help get a first feel for the quality of the research and whether to trust its findings.

Takeaway Tips

A. Whenever you read about a study, report or research that piques your curiosity, try to get your hands on the original.

B. Check the methodology section to see what type of sample was used (e.g., students, age group, 1 or 2 countries, consumers or industrial buyers).

C. Don't forget to look for information about what kind of data were collected.

D. Find out what kind of analysis was done.

E. Ask yourself if the study allows for generalizations based on its findings.

We come across studies and rankings providing information that we might use to make decisions in our own lives or work on a daily basis, such as the Financial Times' annual ranking of Executive Master of Business Administration programs and others that continue to be popular with readers.[23]

[23] See rankings.ft.com/businessschoolrankings/

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

But if we need to choose an MBA program that best suits our individual needs, looking at the methodology used to determine rankings is a good start and wise risk management.[24]

For this and other reasons, next time we hear a finding reported in the popular press, we should check the original report that led journalists draw their conclusions before trusting them. For various reasons, journalists may not have gone beyond the press release before writing their piece.

As the material above indicates, familiarizing oneself with the methodology researchers have used before making decisions based on such findings that require substantial investments of resources is just good strategy. Similarly, before we trust the next university, business education or mobile subscription contract rankings published with great fanfare, let us check the facts, and start by going over the methodology with a fine-tooth comb. This will help ensure that we spend our money on the option that best suits our needs.

Good luck!

[24] St. Gallen (Switzerland) had one program participate and supported this decision with some smart strategic moves to further improve its rankings. Business Education (December 6, 2010) ranks 35 European business schools, with several moving in the rankings by 15 positions between 2009 and 2010.
Since it is unlikely that faculty research/publications changed so much within 12 months, these data should probably be supplemented and cross-validated with measures that do not require the respondents' cooperation (e.g., alumni), or contaminate the response (e.g., how to deal with missing data).
The biggest flaw in FT's MBA rankings is that such a system implicitly suggests that one institution is better for all MBA students than any other, which is a patent falsehood (see ComMetrics.com/?p=19). Risk assessment and risk management are critical when interpreting data that will be used as a basis for strategic decisions: info.cytrap.eu/?p=118

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

Checklist: Can I trust this research?

Here is a list of issues to keep in mind and questions to ponder when checking out a research report or consultant briefing in order to get a handle on how useful the data may be to your organization.

1. **Research methodology**: Look for a description of the methodology used to collect the data, get a first feel about the study.

2. **Type of research**: Was exploratory, descriptive or causal research conducted and why was that type chosen?

3. **Data collection method**: Were primary or secondary data used? What methods were used to collect data (i.e. interviews, surveys, observations, laboratory work).

4. **Sample selection**: How and from which population was the sample selected?

5. **Survey or interview**: Were the questions all designed with a mid-point (e.g., undecided) and personal questions asked at the end (e.g., subject's gender)? Study the full instrument (list of questions) provided in the appendix.

6. **Variables**: What kind of variables were used to collect data; nominal, ordinal, interval and/or ratio?

7. **Non-response bias or missing data, validity and reliability**: How did the authors deal with these issues?

8. **Multi-method**: Did the authors use one method to collect data (e.g., interviews, archive data) and another to offset its weaknesses (e.g., surveys)?

9. **Statistics**: What kind of descriptive statistics were used and did the authors make inferences from their sample about the population. Was the use of statistics applied appropriately (see Table 1)?

10. **Context:** In what context was this study conducted (e.g., during the financial crisis, war, high unemployment, what country, type of organization)? How could this have influenced the findings (i.e. can we infer that the data applies to our setting as well)?

REMEMBER that no study will have answers for the above questions to fully satisfy the inquiring mind, but the answers you jot down will give you a feel for how much you can trust the findings. This will hopefully enable you to base strategic decisions on good research or go look for answers elsewhere.

For regular updates please see: ComMetrics.com/?p=12344

Roentgenstrasse 49 Street
CH-8005 Zuerich Zip Code
Switzerland Country
+41(0)44 272 1876 Voice
+41(0)76 200 7778 Cell
www.CyTRAP.eu URL
info@CyTRAP.eu E-Mail

About ComMetrics

ComMetrics, a division of CyTRAP Labs GmbH, is an industry leader in the analysis of corporate blogs in Europe, creating web-based software used by social media experts, in-house professionals and advertising agencies helping companies improve social media performance in the blogosphere.

The social media channel presents new and exciting opportunities for brands to better understand, manage and grow online communities.
As a leader in social brand analytics, strategy and programs, ComMetrics helps clients engage directly with the people who matter most – their customers.

As an independent consulting body, we have been monitoring best practice since its inception at the turn of the 21st century and advise corporations, non-profits and governmental organizations on their social media presence. For these tasks we have developed a range of web-based software tools, some of which are offered for public use. These are available at My.ComMetrics.com.

One of the popular services we offer is called "Social Readiness Roadmap" which helps brands get ready internally for social business.

Professor Urs E. Gattiker, Ph.D.[25]

An analytics and social media industry leader, widely-read blogger and author of several books, Urs provides training seminars on best practice for developing and implementing social media strategies that deliver actionable metrics and return on investment (ROI).

There is a need for better quality information to be shared in online channels. Merchants and brands must make it easy for customers to reach out when they have questions, concerns or are trying to decide what product/service to buy.

- Urs E. Gattiker

As an analyst, Urs draws on a wealth of experience managing and executing social media projects for clients, enabling them to optimize their online presence using data. In his position at ComMetrics, he monitors trends and advises and supports clients in leveraging their talents to gain competitive advantages in the marketplace.

One of the popular offerings he has been spearheading is called “Social Risk Monitor” that helps brands managing governance risk and compliance for better customer engagement.

As Chief Technology Officer, Urs is responsible for further developing My.ComMetrics.com, a web-based benchmarking software for corporate blogs.

[25] Read more about Urs E. Gattiker: info.cytrap.eu/?page_id=114