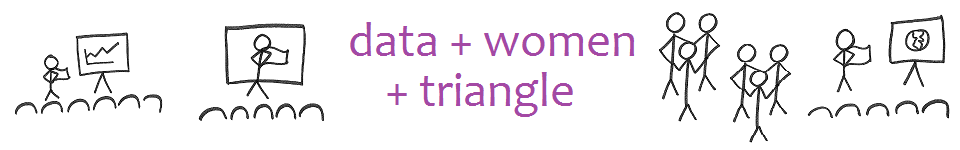
****

CAP Chat Rooms

## STAR Story and Thumbbook

# S Is for Situation

The Certified Analytics Program “defines analytics as the scientific process of transforming data into insight for making better decisions.” The program describes an “end-to-end process beginning with identifying the business problem to evaluating and drawing conclusions about the solution arrived at through the use of analytics.” Further, followers of the program are expected to understand some 160+ vocabulary terms.

# T Is for Task

Meanwhile, the mission of Data+Women+Triangle is to build relationships among co-located women pursuing careers in data science, a.k.a *data scientistas*, while also practicing mutual supportiveness. Our task is to expose these women to the CAP vocabulary to build their careers. If this exposure is in a live-meeting format, it will also build relationships.

Thus, our plan is to configure a series of in-person meetings to cover the CAP vocabulary. A CAP Chat Room is a local, in-person meetup where data scientistas can exchange their experiences relevant to a cluster of CAP vocabulary terms in complete sentences. This way, novices (white belts) to any of the topics can come to a general understanding of the vocabulary, and, the moderately experienced can develop a map towards proficiency, all while practicing the key career skill of the STAR story and practicing the Certified Analytics Program framing.

# A is for Action

Our action plan is to

1. Develop clusters of vocabulary terms
2. Find moderators and managers for each of the cluster chats, a.k.a *real chat rooms*.
3. Develop a *real chat room* meeting framework.

## Clusters of Vocabulary Terms

### The Clusters

Below is the first version of a list of clusters. This list of 9 suggests that the meeting series will have 9 meetings.

|  |  |  |  |
| --- | --- | --- | --- |
| Cluster | Cluster Description | Best Friendly Resource | Moderator & Email |
| A | The Certified Analytics Professional Program | [Certified Analytics Professional Study Guide](https://www.certifiedanalytics.org/sitemap.php) by INFORMS (139 pages) |  |
| B | Business, Accounting, Marketing | [Tableau Analytics in the Accounting Classroom](https://www.barnesandnoble.com/w/tableau-analytics-in-the-accounting-classroom-charles-e-hooper/1129067908?ean=9781544936154) by Hooper (84 pages) |  |
| C | Quality, Lean, and Six Sigma | [What Is Lean Six Sigma?](https://www.barnesandnoble.com/w/what-is-lean-six-sigma-michael-george/1122979790?ean=9780071426688) By George, Rolands, and Kastle (96 pages) |  |
| D | Statistics | [Cartoon Guide to Statistics](https://www.barnesandnoble.com/w/cartoon-guide-to-statistics-larry-gonick/1100632890?ean=9780062731029#/) by Larry Golnick and Woolcott Smith (240 pages, but it’s cartoons) |  |
| E | Analytics | Search on [Data Analytics for Beginners](https://www.barnesandnoble.com/s/data+analytics+for+beginners?_requestid=16814859) |  |
| F | Decision Science | [Decision Science and Queueing Theory](https://www.barnesandnoble.com/w/decision-science-and-queueing-theory-david-a-brown/1103908361?ean=9781257623969) by David A. Brown (152 pages) |  |
| G | Optimization | [A Gentle Introduction to Optimization](https://www.barnesandnoble.com/w/a-gentle-introduction-to-optimization-b-guenin/1124330562?ean=9781107658790) by Guenin, Koneman, and Tuncel (269 pages) |  |
| H | Data Engineering & Visualization | [Why Visual Analytics?](https://www.tableau.com/learn/whitepapers/why-visual-analytics) By Tableau |  |
| J | Security, Privacy, and Threats |  |  |

### Best Friendly Resource

For each cluster, we will honor one, maybe two, best friendly resources. Each best friendly resource:

* Best friendly resources are oriented to novices.
* Best friendly resources are short in length (<150 pages) and low in cost (<$25)
* The participants of the CAP Chat Room are responsible for building consensus to make the designation.

## Chat Room Moderators

Each CAP Chat room has one moderator. The responsibilities of the moderator are as follows:

* Recruit and confirm the participation of 2 or 3 specialists for the CAP Chat Room
* Run the CAP Chat Room, using the meeting format. Retain the sign-in sheet.
* Advocate for the usage of STAR stories and exercise other group facilitation skills.
* Cultivate the group consensus on the Best Friendly Resource of that CAP Chat Room
* Cultivate the group consensus about which terms belong in which cluster
* Find a Chat Room Manager, who will arrange the date, time, place and security of the physical meeting, while conforming to the agreed cadence (e.g., the 2nd Monday of every month from 6:30 to 8:30)

## Real Chat Room Meeting Format

### Sign In

CAP Chat room participants will need to sign-in on a physical sheet and create a nametag. The sign-in area includes Best Friendly Resources on display. Copies of this thumbbook should be available.

### The Update Sheet

Chatters who participate in the meeting will begin with the standard Lean In update sheet. This worksheet proscribes that each participant briefly give the best/worst events of her life professionally/personally during last/next month.

### The STAR Story

Moderators and other chatters alike will practice their use of the STAR story. These stories rehearse answers to interview questions or résumé line items. [Colgate University](http://www.colgate.edu/portaldata/imagegallerywww/0c5ec2b8-5500-4c50-b897-50041d78ee46/imagegallery/telling%20your%20story%20using%20the%20star.pdf) describes the 4 components of a STAR story:

* **Situation** – Describe vividly the environmental condition; ie. department, organization, timeframe.
* **Task** or **Obstacle**– Describe what was expected or asked of you in terms of responsibilities OR the challenges or obstacles you had to overcome.
* **Action** – Detail the specific action steps that you took to handle the tasks or address the obstacle.
* **Result** – What impact did you have on the situation? Identify concrete outcomes of your actions, providing quantitative or qualitative results when possible.

# R is for Results

The following is conceptually a worksheet to capture the results of each CAP Chat Room: For each term and definition, note where the term is discussed in the Best Friendly Resource, and re-tell the STAR story about that vocabulary term

|  |  |  |  |
| --- | --- | --- | --- |
| Term | Definition | Citation | STAR |
| The Certified Analytics Professional Program (A) | | | |
|  | (An introduction and overview of the program) |  |  |
|  |  |  |  |
| Business, Accounting, Marketing (B) | | | |
| Activity-based costing | method of assigning costs to products or services on the resources that they consume (http://www.economist. com/node/13933812) |  |  |
| Amortization | allocation of cost of an item or items over a time period such that the actual cost is recovered; often used to account for capital expenditures |  |  |
| Business analytics (BA) | refers to the skills, technologies, applications, and practices for continuous iterative exploration and investigation of past business performance to gain insight and drive business planning; can be descriptive, prescriptive, or predictive; focuses on ... |  |  |
| Business case | reasoning underlying and supporting the estimates of business consequences of an action |  |  |
| Business intelligence (BI) | a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information (http://en.wikipedia.org/wiki/ Business\_intelligence) |  |  |
| Chief Analytics Officer (CAO) | possible title of one overseeing analytics for a company; may include mobilizing data, people, and systems for successful deployment, working with others to inject analytics into company strategy and decisions, supervising activities of analytical peopl... |  |  |
| Cost of capital | the cost of funds used for financing a business. Cost of capital depends on the mode of financing used—it refers to the cost of equity if the business is financed solely through equity, or to the cost of debt if it is financed solely through debt (www.i... |  |  |
| Enterprise resource planning (ERP) | a cross-functional enterprise system driven by an integrated suite of software modules that supports the basic internal business processes of a company (http:// en.wikipedia.org/wiki/Enterprise\_resource\_planning) |  |  |
| Fixed cost | a cost that is some value, say C, regardless of the level as long as the level is positive; otherwise the fixed charge is zero. This is represented by Cv, where v is a binary variable. When v = 0, the fixed charge is 0; when v = 1, the fixed charge is C... |  |  |
| Internal rate of return (IRR) | the rate of growth that a project or investment is expected to create, expressed as a percentage, over a specified term. IRR is, in essence, the theoretical interest rate earned by the project (http://www.askjim.biz/ answers/internal-rate-of-return-irr-... |  |  |
| Mean time between failures (MTBF) | a measure of how reliable a hardware product or component is. For most components, the measure is typically in thousands or even tens of thousands of hours between failures (http://whatis.techtarget.com/ definition/MTBF-mean-time-between-failures) |  |  |
| Net present value | value in today’s currency of an item or service (Davenport, Enterprise Analytics, p. 22) |  |  |
| Opportunity cost | the cost of an alternative that must be forgone to pursue a certain action (http://www.investopedia.com/terms/o/ opportunitycost.asp) |  |  |
| Payback | the length of time required to recover the cost of an investment (http://www.investopedia.com/terms/p/ paybackperiod.asp) |  |  |
| Proprietary data | data that no other organization possesses; produced by a company to enhance its competitive posture (Davenport, Enterprise Analytics, p. 37) |  |  |
| Return on investment (ROI) | calculations that provide a basis for comparison with other investment opportunities; typically calculated using ROI = ((Total value/benefits) – (total investment costs))/Total investment costs (Davenport; Enterprise Analytics, p. 20) |  |  |
| Revenue management | the science and art of enhancing revenues while selling essentially the same amount of product (http://www.ivey. uwo.ca/faculty/Peter\_Bell/RM%20Ahmedabad%202005. pdf) |  |  |
| RFM | data related to customer relationship management; refers to recency, frequency, and monetary value of purchases (Davenport, Enterprise Analytics, p. 49) |  |  |
| Risk | the potential of loss (an undesirable outcome, however not necessarily so) resulting from a given action, activity, and/or inaction (http://en.wikipedia.org/wiki/Risk) |  |  |
| Shadow price | an economic term to denote the rate at which the optimal value changes with respect to a change in some right-hand side that represents a resource supply or demand requirement (A. Holder, editor. Mathematical Programming Glossary. INFORMS Computing Soci... |  |  |
| Spreadsheet analysis | the analysis of data using special computer software to anticipate marketing performance under a given set of circumstances (http://www.marketing-dictionary.com/s. php) |  |  |
| Variable cost | a periodic cost that varies in step with the output or the sales revenue of a company. Variable costs include raw material, energy usage, labor, distribution costs, etc. (http://www.businessdictionary.com/definition/variablecost.html) |  |  |
| Quality, Lean, and Six Sigma (D) | | | |
| 5 Whys | iterative process of discovery through repetitively asking ‘why’; used to explore cause and effect relationships underlying and/or leading to problem (http:// en.wikipedia.org/wiki/5\_Whys) |  |  |
| 5S | workplace organization method promoting efficiency and effectiveness; five terms based on Japanese words for: sorting, set in order, systematic cleaning, standardizing, and sustaining (http://en.wikipedia.org/ wiki/5S\_(methodology) |  |  |
| 80/20 | Rule AKA the Pareto principle: roughly 80% of results come from 20% of effort |  |  |
| Assemble-to-Order (ATO) | manufacturing process where products are assembled as they are ordered; characterized by rapid production and customization (http://www.investopedia.com/terms/a/ assemble-to-order.asp) |  |  |
| Automation | use of mechanical means to perform work previously done by human effort |  |  |
| Batch production | method of production where components are produced in groups rather than a continual stream of production; see also, Continuous production |  |  |
| Batch production | method of production where components are produced in groups rather than a continual stream of production; see also, Continuous production |  |  |
| Benchmark problems | comparison of different algorithms using a large test set (http://www.cs.cmu.edu/afs/cs/project/jair/pub/ volume24/ortizboyer05a-html/node6.html) |  |  |
| Benchmarking | act of comparison against a standard or the behavior of another in attempt to determine degree of conformity to standard or behavior |  |  |
| Business Process Modeling or Mapping (BPM) | act of representing processes of an enterprise so that the current process may be analyzed and improved; typically action performed by business analysis and managers seeking improved efficiency and quality (http:// en.wikipedia.org/wiki/Business\_process... |  |  |
| Continuous production | method of production where components are produced in a continuous stream; see also, Batch production |  |  |
| Efficiency | the comparison of what is actually produced or performed with what can be achieved with the same consumption of resources (money, time, labor, etc.). It is an important factor in determination of productivity (www.businessdictionary.com) |  |  |
| Experimental design | in quality management, a written plan that describes the specifics for conducting an experiment, such as which conditions, factors, responses, tools, and treatments are to be included or used; see also, Design of experiments (www.businessdictionary.com) |  |  |
| Failure Mode and Effects Analysis (FMEA) | a systematic, proactive method for evaluating a process to identify where and how it might fail, and to assess the relative impact of different failures to identify the parts of the process that are most in need of change (http://intranet.uchicago.edu/q... |  |  |
| Lead time | time between the initial phase of a process and the emergence of results, as between the planning and completed manufacture of a product (http://www. thefreedictionary.com/lead+time) |  |  |
| Lean production | a Japanese approach to management that focuses on cutting out waste while ensuring quality. This approach can be applied to all aspects of a business – from design through production to distribution (http://www. tutor2u.net/business/production/introduct... |  |  |
| Nominal group technique (NGT) | a structured method for group brainstorming that encourages contributions from everyone (http://asq. org/learn-about-quality/idea-creation-tools/overview/ nominal-group.html) |  |  |
| Pareto concept | See, 80/20 rule |  |  |
| Six Sigma | a set of strategies, techniques, and tools for process improvement. It seeks to improve the quality of process outputs by identifying and removing the causes of defects (errors) and minimizing variability in manufacturing and business processes (http://... |  |  |
| Variation reduction | reference to process variation where reduction leads to stable and predication process results (http://www. businessdictionary.com) |  |  |
| Yield | percentage of ‘good’ product in a batch; has three main components: functional (defect driven), parametric (performance driven), and production efficiency/ equipment utilization (http://www-inst.eecs.berkeley. edu/~ee290h/fa05/Lectures/PDF/lecture%201%2... |  |  |
| Statistics | | | |
| Accuracy | quality or state of being correct or precise, or the degree to which the result of a measurement, calculation, or specification conforms to the correct value or standard (https://www.google.com/#q=accuracy) |  |  |
| ANCOVA | acronym for analysis of covariance |  |  |
| ANOVA | acronym for analysis of variance |  |  |
| Average | sum of a range of values divided by the number of values to arrive at a value characteristic of the midpoint of the range; see also, Mean |  |  |
| Bias | a tendency for or against a thing, person, or group in a way as to appear unfair; in statistics, data calculated so that it is systematically different from the population parameter of interest (http://en.wikipedia.org/wiki/ Bias\_(statistics) |  |  |
| Confidence interval | a type of interval estimate of a population parameter used to indicate the reliability of an estimate. It is an observed interval (i.e., it is calculated from the observations), in principle different from sample to sample, that frequently includes the ... |  |  |
| Confidence level | if confidence intervals are constructed across many separate data analyses of repeated (and possibly different) experiments, the proportion of such intervals that contain the true value of the parameter will match the confidence level (http://www.usable... |  |  |
| Correlation | a broad class of statistical relationships involving dependence (http://en.wikipedia.org/wiki/Correlation\_ and\_dependence) |  |  |
| Cumulative density function | probability that a real-valued random variable X with a given probability distribution will be found at a value less than or equal to x; used to specify the distribution of multivariate random variables (http://en.wikipedia.org/ wiki/Cumulative\_distribu... |  |  |
| Factor analysis | a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors. Factor analysis searches for such joint variations in response to unobserved latent var... |  |  |
| Goodness of fit | degree of assurance or confidence to which the results of a sample survey or test can be relied upon for making dependable projections. Described as the degree of linear correlation of variables, it is computed with the statistical methods such as chi-s... |  |  |
| Hypothesis testing | the theory, methods, and practice of testing a hypothesis by comparing it with the null hypothesis. The null hypothesis is only rejected if its probability falls below a predetermined significance level, in which case the hypothesis being tested is said... |  |  |
| Logistic regression | a type of probabilistic classification model [1] used for predicting the outcome of a categorical dependent variable (i.e., a class label) based on one or more predictor variables (features). Logistic regression can be binomial or multinomial. Binomial or binary logistic regression deals with situations in which the observed outcome for a dependent variable can have only two possible types (for example, “dead” versus “alive”). Multinomial logistic regression de... |  |  |
| MANOVA | acronym for multivariate analysis of variance for use with multiple independent variables |  |  |
| Mean | the arithmetic average of a set of values or distribution; however, for skewed distributions, the mean is not necessarily the same as the middle value (median), or the most likely (mode); see also, Average (http:// en.wikipedia.org/wiki/Mean) |  |  |
| Mean squared error (MSE) | the unbiased estimator of population variance. MSE divides by the error degrees of freedom, e.g., if only the mean is estimated, MSE divides by N-1, if four parameters are estimated, MSE divides by N-4, and so on (http://en.wikipedia.org/wiki/Mean\_squar... |  |  |
| Median | the value such that the number of terms having values greater than or equal to it is the same as the number of terms having values less than or equal to it (http:// searchdatacenter.techtarget.com/definition/statisticalmean-median-mode-and-range) |  |  |
| Mode | value of the term that occurs the most often (http:// searchdatacenter.techtarget.com/definition/statisticalmean-median-mode-and-range) |  |  |
| Precision | the degree to which repeated measurements under unchanged conditions show the same results (http:// en.wikipedia.org/wiki/Accuracy\_and\_precision) |  |  |
| Pricing | a tactic in the simplex method, by which each variable is evaluated for its potential to improve the value of the objective function. Let p = c\_B[B^-1], where B is a basis, and c\_B is a vector of costs associated with the basic variables. The vector p i... |  |  |
| Probability density function | the equation used to describe a continuous probability distribution (http://stattrek.com/statistics/dictionary. aspx?definition=Continuous\_probability\_distribution) |  |  |
| Random | of or characterizing a process of selection in which each item of a set has an equal probability of being chosen (http://dictionary.reference.com/browse/random) |  |  |
| Range | the difference between the maximum and minimum observations providing an estimate of the spread of the data (http://explorable.com/range-in-statistics) |  |  |
| Regression | a statistical measure that attempts to determine the strength of the relationship between one dependent variable (usually denoted by Y) and a series of other changing variables (known as independent variables) (http://www.investopedia.com/terms/r/regres... |  |  |
| Regression analysis | statistical approach to forecasting change in a dependent variable (e.g., sales revenue) on the basis of change in one or more independent variables (e.g., population and income); AKA curve fitting or line fitting (www.businessdictionary.com) |  |  |
| Response surface methodology (RSM) | a surface in (n+1) dimensions that represents the variations in the expected value of a response variable (see, regression) as the values of n explanatory variables are varied. Usually the interest is in finding the combination that gives a global maxim... |  |  |
| Standard deviation | measure of the unpredictability of a random variable, expressed as the average deviation of a set of data from its arithmetic mean and computed as the positive square root of the variance. Customarily represented by the lower-case Greek letter sigma (λ)... |  |  |
| Statistical significance | probability of obtaining a test result that occurs by chance and not by systematic manipulation of data (www.businessdictionary.com) |  |  |
| Statistics | branch of mathematics concerned with collection, classification, analysis, and interpretation of numerical facts, for drawing inferences on the basis of their quantifiable likelihood (probability). Statistics can interpret aggregates of data too large t... |  |  |
| Variability | describes how spread out or closely clustered a set of data is (http://en.wikipedia.org/wiki/Variability) |  |  |
| Variance | a parameter in a distribution that describes how far the values are spread apart. Variance is a characteristic of some probability distribution, which distinguishes the concept of variance from the ways to estimate it from sample data(http://en.wikipedi... |  |  |
| Analytics (E) | | | |
| Analytics | scientific process of transforming data into insight for making better decisions (INFORMS) |  |  |
| Artificial Intelligence (AI) | branch of computer science that studies and develops intelligent machines and software (http://en.wikipedia. org/wiki/Artificial\_intelligence) |  |  |
| Artificial Neural Networks | computer-based models inspired by animal central nervous systems (https://www.google.com/#q=artificial+ neural+networks) |  |  |
| Chi-squared Automated Interaction Detection (CHAID) | a technique for performing decision tree analysis developed by Gordon V. Kass. CHAID is one of several commonly used techniques for decision trees and is based upon hypothesis testing using Bonferroni correction. |  |  |
| Classification | assortment of items or entities into predetermined categories |  |  |
| Clustering | grouping of a set of objects in such a way that objects in the same group (cluster) are more similar to each other than to those in other groups or clusters (http:// en.wikipedia.org/wiki/Cluster\_analysis) |  |  |
| Data mining | relatively young and interdisciplinary field of computer science; the process of discovering new patterns from large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems; see also... |  |  |
| Descriptive analytics | prepares and analyzes historical data to identify patterns for reporting trends (http://www.informs.org/ Community/Analytics/About-Us) |  |  |
| Effective domain | the domain of a function for which its value is finite (A. Holder, editor. Mathematical Programming Glossary. INFORMS Computing Society, http://glossary. computing.society.informs.org/, 2006-08. Originally authored by Harvey J. Greenberg, 1999-2006.) |  |  |
| Forecasting | the use of historic data to determine the direction of future trends (http://www.investopedia.com/terms/f/ forecasting.asp) |  |  |
| Lift/lift curve | a measure of the effectiveness of a predictive model calculated as the ratio between the results obtained with and without the predictive model; lift charts consisting of lift curve and a baseline are visuals aids for measuring model performance (http:/... |  |  |
| Machine learning | an artificial intelligence (AI) discipline geared toward the technological development of human knowledge. Machine learning allows computers to handle new situations via analysis, self-training, observation, and experience (http://www.techopedia.com/def... |  |  |
| Next best offer (NBO) | a targeted offer or proposed action for customers based on analyses of past history and behavior, other customer preferences, purchasing context, attributes of the produces, or services from which they can choose (Davenport, Enterprise Analytics, p. 83) |  |  |
| Pattern recognition | in machine learning, pattern recognition is the assignment of a label to a given input value (http:// en.wikipedia.org/wiki/Pattern\_recognition) |  |  |
| Principal Component Analysis (PCA) | a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set (ftp://statgen. ncsu.edu/pub/thorne/molevoclass/AtchleyOct19.pdf) |  |  |
| Problem assessment/ framing | initial step in the analytics process; involves buy in from all parties involved on what the problem is before a solution can be found |  |  |
| Scenario analysis | a process of analyzing possible future events by considering alternative possible outcomes (scenarios). The analysis is designed to allow improved decision making by allowing more complete consideration of outcomes and their implications (http://www. in... |  |  |
| Stepwise regression | a semi-automated process of building a model by successively adding or removing variables based solely on the t-statistics of their estimated coefficients (http:// people.duke.edu/~rnau/regstep.htm) |  |  |
| Validation (of a model) | determining how well the model depicts the real-world situation it is describing (http://www.easterbrook.ca/ steve/2010/11/the-difference-between-verification-andvalidation/) |  |  |
| Verification (of a model) | includes all the activities associated with the producing high quality software: testing, inspection, design analysis, specification analysis (http://www.easterbrook. ca/steve/2010/11/the-difference-between-verificationand-validation/) |  |  |
| Decision Science (F) | | | |
| Agent-based modeling | a class of computation models for simulating actions and interactions of autonomous agents with a view to assessing their effects on the system as a whole (http:// en.wikipedia.org/wiki/Agent-based\_model) |  |  |
| Analytics professional | person capable of making actionable decisions through the analytic process; also a person holding the Certified Analytics Professional (CAP®) credential |  |  |
| Conjoint analysis | allows calculation of relative importance of varying features and attributes to customers |  |  |
| Constraint | a condition that a solution to an optimization problem is required by the problem itself to satisfy. There are several types of constraints—primarily equality constraints, inequality constraints, and integer constraints (http://en.wikipedia.org/wiki/Con... |  |  |
| Constraint programming | a programming paradigm wherein relations between variables are stated in the form of constraints (http:// en.wikipedia.org/wiki/Constraint\_programming) |  |  |
| Decision variables | a decision variable represents a problem entity for which a choice must be made. For instance, a decision variable might represent the position of a queen on a chessboard, for which there are 100 different possibilities (choices) on a 10x10 chessboard o... |  |  |
| Design of experiments | design of any information gathering exercise where variation is present, whether under the control of the experimenter or not; see also, Experimental design (http://en.wikipedia.org/wiki/Design\_of\_experiments) |  |  |
| Discrete event simulation | models the operation of a system as a discrete sequence of events in time; between events, no change in the system is assumed thus a simulation can move in time from one event to the next (http://en.wikipedia.org/wiki/ Discrete\_event\_simulation) |  |  |
| Dynamic programming | based on the Principle of Optimality, this was originally concerned with optimal decisions over time. For continuous time, it addresses problems in variational calculus. For discrete time, each period is sometimes called a stage, and the DP is called a ... |  |  |
| Fuzzy logic | a form of mathematical logic in which truth can assume a continuum of values between 0 and 1 (http:// wordnetweb.princeton.edu/perl/webwn?s=fuzzy logic) |  |  |
| Game Theory | in general, a (mathematical) game can be played by one player, such as a puzzle, but its main connection with mathematical programming is when there are at least two players, and they are in conflict. Each player chooses a strategy that maximizes his pa... |  |  |
| INFORMS | the largest professional society in the world for professionals in the field of operations research (OR), management science, and analytics (www.informs.org/ About) |  |  |
| Innovative Applications in Analytics Award | award administered by the Analytics Section of INFORMS to recognize creative and unique developments, applications, or combinations of analytical techniques. The prize promotes the awareness of the value of analytics techniques in unusual applications, ... |  |  |
| Integer program | the variables are required to be integer-valued. Historically, this term implied the mathematical program was otherwise linear, so one often qualifies a nonlinear integer program versus a linear IP (A. Holder, editor. Mathematical Programming Glossary. ... |  |  |
| Linear program | opt{cx: Ax = b, x >= 0}. (Other forms of the constraints are possible, such as Ax <= b.) The standard form assumes A has full row rank. Computer systems ensure this by having a logical variable (y) augmented, so the form appears as Opt{cx: Ax + y = b, L... |  |  |
| Little’s Law | queuing theory where numerator and denominator are halved so queues are roughly equivalent no matter how many are in line; the long-term average number of customers in a stable system L is equal to the long-term average effective arrival rate, λ, multip... |  |  |
| Monte Carlo simulation | a computerized mathematical technique that allows people to account for risk in quantitative analysis and decision making. The technique is used by professionals in such widely disparate fields as finance, project management, energy, manufacturing, engi... |  |  |
| Operations management | deals with the design and management of products, processes, services, and supply chains. It considers the acquisition, development, and utilization of resources that firms need to deliver the goods and services their clients want (http://mitsloan.mit.e... |  |  |
| Operations Research | a discipline that deals with the application of advanced analytical methods to help make better decisions (http:// en.wikipedia.org/wiki/Operations\_research) |  |  |
| Project management | the application of knowledge, skills, and techniques to execute projects effectively and efficiently. A strategic competency for organizations, enabling them to tie project results to business goals (http://www.pmi.org/ About-Us/About-Us-What-is-Project... |  |  |
| Queuing theory | mathematical study of waiting in lines; results are used when making business decisions about the resources needed to provide service; research begun by A. K. Erlang (http://en.wikipedia.org/wiki/Queuing\_theory on 2/20/13) |  |  |
| Scheduling | a schedule for a sequence of jobs, say j1,...,jn, is a specification of start times, say t1,...,tn, such that certain constraints are met. A schedule is sought that minimizes cost and/or some measure of time, like the overall project completion time (wh... |  |  |
| Sensitivity analysis | the concern with how the solution changes if some changes are made in either the data or in some of the solution values (by fixing their value). Marginal analysis is concerned with the effects of small perturbations, maybe measurable by derivatives. Par... |  |  |
| Supply chain management | the active management of supply chain activities to maximize customer value and achieve a sustainable competitive advantage (http://scm.ncsu.edu/scmarticles/article/what-is-supply-chain-management) |  |  |
| Tolerance | an approach to sensitivity analysis in linear programming that expresses the common range that parameters can change while preserving the character of the solution (A. Holder, editor. Mathematical Programming Glossary. INFORMS Computing Society, http://... |  |  |
| Uncertainty | the estimated amount or percentage by which an observed or calculated value may differ from the true value (http://www.thefreedictionary.com/uncertainty) |  |  |
| Optimization & Specialty Fields (G) | | | |
| Assignment problem | one of the fundamental combinatorial optimization problems in the branch of optimization or operations research in mathematics; consists of finding a maximumweight matching in a weighted bipartite graph (http:// en.wikipedia.org/wiki/Assignment\_problem) |  |  |
| Branch-and-Bound | a general algorithm for finding optimal solutions of various optimization problems; consists of a system enumeration of all candidate solutions where large subsets of fruitless candidates are discarded en masse using upper and lower estimated bounds of ... |  |  |
| Combinatorial optimization | a topic that consists of finding an optimal object from a finite series of objects; used in applied mathematics and theoretical computer science (http://en.wikipedia.org/ wiki/Combinatorial\_optimization) |  |  |
| Cutting stock problem | optimization or integer linear programming problem arising from applications in industry where high production problems exist (http://en.wikipedia.org/wiki/ Cutting\_stock\_problem) |  |  |
| Engagement | an estimate of the depth of visitor interaction against a clearly defined set of goals; may be measured through analytical models (Davenport, Enterprise Analytics, p. 73-74) |  |  |
| Expert systems | a computer program that simulates the judgment and behavior of a human or an organization that has expert knowledge and experience in a particular field. Typically, such a system contains a knowledge base containing accumulated experience and a set of r... |  |  |
| Genetic algorithms | a class of algorithms inspired by the mechanisms of genetics, which has been applied to global optimization (especially for combinatorial programs). It requires the specification of three operations (each is typically probabilistic) on objects, called “... |  |  |
| Global optimal | refers to mathematical programming without convexity assumptions, which are NP-hard. In general, there could be a local optimum that is not a global optimum. Some authors use this term to imply the stronger condition there are multiple local optima. Som... |  |  |
| Greedy heuristics | an algorithm that follows the problem-solving heuristic of making the locally-optimal choice at each stage with the hope of finding a global optimum (http:// en.wikipedia.org/wiki/Greedy\_heuristic) |  |  |
| Heuristic | in mathematical programming, this usually means a procedure that seeks an optimal solution but does not guarantee it will find one, even if one exists. It is often used in contrast to an algorithm, so branch and bound would not be considered a heuristic... |  |  |
| Integrity | the measure of the trust that can be placed in the correctness of the information supplied by a navigation system (http://www.navipedia.net/index.php/Integrity; http://www.genengnews.com/gen-articles/preservingthe-integrity-of-statistics/3081/) |  |  |
| Knapsack problem | an integer program of the form, Max{cx: x in Zn+ and ax <= b}, where a > 0. The original problem models the maximum value of a knapsack that is limited by volume or weight (b), where x\_j = number of items of type j put into the knapsack at unit return c... |  |  |
| Local optimal | a solution that is optimal (either maximal or minimal) within a neighboring set of candidate solutions (http:// en.wikipedia.org/wiki/Local\_optimum) |  |  |
| Metaheuristics | a general framework for heuristics in solving hard problems. The idea of ``meta’’ is that of level. An analogy is the use of a metalanguage to explain a language. For computer languages, we use symbols, like brackets, in the metalanguage to denote prope... |  |  |
| Network optimization | the process of striking the best possible balance between network performance and network costs, in consideration of grade of service requirements (www. yourdictionary.com) |  |  |
| Objective function | the (real-valued) function to be optimized. In a mathematical program in standard form, this is denoted f (A. Holder, editor. Mathematical Programming Glossary. INFORMS Computing Society, http://glossary. computing.society.informs.org/, 2006-08. Origina... |  |  |
| Optimization | procedure or procedures used to make a system or design as effective or functional as possible, especially the mathematical techniques involved (http://www. thefreedictionary.com/optimization) |  |  |
| Robust optimization | a term given to an approach to deal with uncertainty, similar to the recourse model of stochastic programming, except that feasibility for all possible realizations (called scenarios) is replaced by a penalty function in the objective. As such, the appr... |  |  |
| Simulated annealing | an algorithm for solving hard problems, notably combinatorial programs, based on the metaphor of how annealing works: reach a minimum energy state upon cooling a substance, but not too quickly in order to avoid reaching an undesirable final state. As a ... |  |  |
| System dynamics | a computer-aided approach to policy analysis and design. It applies to dynamic problems arising in complex social, managerial, economic, or ecological systems (http://www.systemdynamics.org/what\_is\_ system\_dynamics.html) |  |  |
| Traveling salesman problem (TSP) | given n points and a cost matrix [cij], a tour is a permutation of the n points. The points can be cities, and the permutation the visitation of each city exactly once, then returning to the first city (called home). (A. Holder, editor. Mathematical Pro... |  |  |
| Vehicle routing problem (VRP) | finding optimal delivery routes from one or more depots to a set of geographically scattered points (e.g., population centers). A simple case is finding a route for snow removal, garbage collection, or street sweeping (without complications, this is aki... |  |  |
| Web analytics | ability to use data generated through Internet-based activities; typically used to assess customer behaviors; see also, RFM (Davenport, Enterprise Analytics, p. 49-51) |  |  |
| Data Engineering & Visualization (H) | | | |
| Algorithm | set of specific steps to solve a problem |  |  |
| Big data | data sets too voluminous or too unstructured to be analyzed by traditional means |  |  |
| Cleansing | AKA cleaning or scrubbing: the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database; may also involve harmonization of data, and standardization of data (http://en.wikipedia.org/wiki/Data\_... |  |  |
| Cube | see OLAP cube |  |  |
| Data | (plural form of datum) values of qualitative or quantitative variables, belonging to a set of items; represented in a structure, often tabular (represented by rows and columns), a tree (a set of nodes with parent-children relationship), or a graph struc... |  |  |
| Data warehouse | a central repository of data that is created by integrating data from one or more disparate sources; used for reporting and data analysis (http://en.wikipedia.org/wiki/ Data\_warehouse) |  |  |
| Database | an organized collection of data organized to model relevant aspects of reality to support processes requiring this information (http://en.wikipedia.org/wiki/Database) |  |  |
| ETL (extract, transform, load) | refers to three separate functions combined into a single programming tool. First, the extract function reads data from a specified source database and extracts a desired subset of data. Next, the transform function works with the acquired data—using ru... |  |  |
| Graphical User Interface (GUI) | a human–computer interface (i.e., a way for humans to interact with computers) that uses windows, icons, and menus, and that can be manipulated by a mouse (and often to a limited extent by a keyboard as well) (http:// www.linfo.org/gui.html) |  |  |
| KDD | acronym for knowledge discovery in databases process; see also, Data mining (Davenport, Enterprise Analytics, p. 14) |  |  |
| Normalization | splits up data to avoid redundancy (duplication) by moving commonly repeating groups of data into new tables. Normalization therefore tends to increase the number of tables that need to be joined to perform a given query, but reduces the space required ... |  |  |
| OLAP | an abbreviation for “Online Analysis and Processing”; a type of database technology that has long been used by the business community to analyze and interactively explore large financial data sets. The basic idea is that data sets are viewed as cubes wi... |  |  |
| OLAP cub | an array of data understood in terms of its zero or more dimensions; each cell of the cube holds a number that represents some measure of the business, such as sales, profits, expenses, budget, and forecast (http:// en.wikipedia.org/wiki/OLAP\_cube) |  |  |
| Predictive analytics | any approach to data mining with four attributes: an emphasis on prediction (rather than description, classification, or clustering), rapid analysis measured in hours or days (rather than the stereotypical months of traditional data mining), an emphasis... |  |  |
| Prescriptive analytics | evaluates and determines new ways of operating targeting business objective and balancing all constraints (http://www.informs.org/Community/ Analytics/About-Us) |  |  |
| Box-and-whisker plot | a simple way of representing statistical data on a plot in which a rectangle is drawn to represent the second and third quartiles, usually with a vertical line inside to indicate the median value. The lower and upper quartiles are shown as horizontal li... |  |  |
| Decision tree | graphic illustration of how data leads to decision when branches of the tree are followed to their conclusion; different branches may lead to different decisions |  |  |
| Histogram | graphic depiction of data using columns to represent relative size/importance of data grouping |  |  |
| Influence diagram | depicts structure of decision process and notes the data needed to make the decision |  |  |
| Pie chart | graphic depiction of data using a pie with different ‘slices’ to represent the relative size of different groupings of data points to the size of the whole |  |  |
| Scatter plot | graphic depiction of data, used to show/identify relationship between independent variables |  |  |
| Security, Privacy, and Threats (H) | | | |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |