

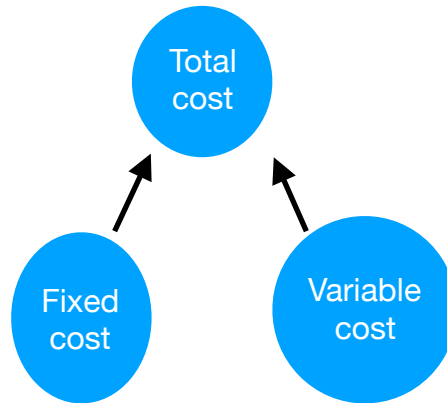
Business Analysis Notes

- CHAPTER 1

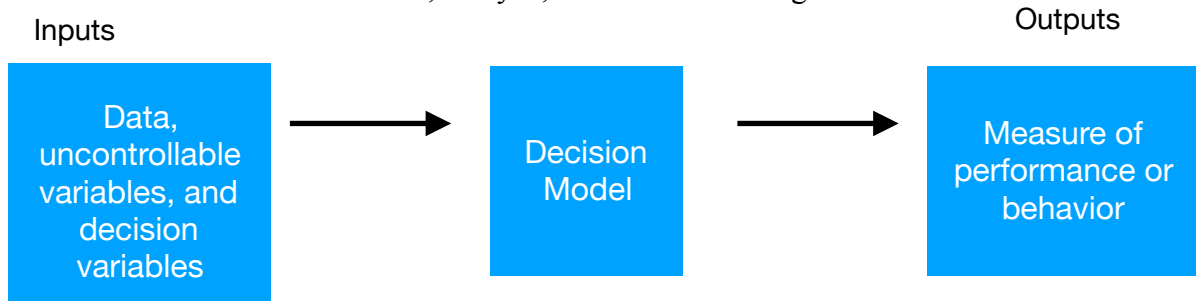
- **Business Analytics:** the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their business operations and make better decisions
 - **Descriptive analytics:** the use of data to understand past and current business performance and make informed decisions.
 - Categorize, characterize, consolidate, and classify data to convert it into useful information for the purposes of understanding and analyzing business performance.
 - **Predictive analytics:** seeks to predict the future by examining historical data, detecting patterns, and then extrapolating these relationships forward in time
 - Predict risk, detect hidden patterns, detect trends
 - **Prescriptive analytics:** uses optimization to identify the best alternatives to minimize or maximize some objective
- **Data:** numerical facts and figures that are collected through some type of measurement process
- **Information:** extracting meaning from data to support evaluation and decision making
- **Data set:** a collection of data
- **Database:** collection of related files containing records on people, places, and things
- **Entities:** people, places, or things for which we store or maintain information
- **Big data:** massive amount of business data from a wide variety of sources
- **Veracity:** the level of reliability associated with data
- **Metric:** a unit of measurement that provides a way to objectively quantify performance
 - **Discrete:** derived from counting something
 - **Continuous:** a continuous scale of measurement
- **Data:**
 - **Categorical data:** sorted into categories by specified characteristics
 - EX: classified by region
 - **Ordinal data:** ordered according to some relationship to one another
 - EX: basketball rankings
 - **Interval data:** ordinal, but have constant differences between observations and have arbitrary zero points
 - EX: time and temperature
 - **Ratio data:** continuous and have natural zero
 - Strongest form of data
- **Reliability:** data are accurate and consistent
- **Validity:** data correctly measure what they are supposed to measure

Business Analysis Notes

- **Model:** an abstraction or representation of a real system, idea, or object
- **Influence diagram:** a simple descriptive model
 - **Nodes:** circular symbols
 - **Branches:** arrows



- **Decision model:** a logical representation of a problem situation that can be used to understand, analyze, or facilitate making a decision



- **Price elasticity:** the ratio of the percentage change in demand to the percentage change in price
- **Uncertainty:** imperfect knowledge of what will happen
- **Risk:** associated with the consequences and likelihood of what might happen
- **Optimization:** process of finding a set of values for decision variables that minimize or maximize same quantity of interest
 - **Objective function:** quantity of interest
 - **Optimal solution:** any set of decision variables that optimizes the objective function
- **Algorithm:** a systematic procedure that finds a solution to a problem
- **Deterministic model:** all model input information is either known or assumed to be known with certainty
- **Stochastic model:** some of the model input information is uncertain
- **Problem solving:** the activity associated with defining, analyzing, and solving a problem.
 - **Recognizing a problem:** problems exist when there is a gap between what is happening and what we think should be happening

Business Analysis Notes

- **Defining the problem:** develop meaningful objectives, characterizing the possible decisions, and identifying any constraints or restrictions
- **Analyzing the problem:** evaluating different scenarios, analyzing risks associated with various decision alternatives, finding a solution that meets certain goals, or determining an optimal solution
- **Interpreting results and making a decision**
- **Implementing the solution:** making it work in the organization or translating the results of the model back to the real world

Business Analysis Notes

- CHAPTER 2

- **Relative address:** uses just the row and column label in the cell reference
- **Absolute address:** uses a dollar sign before either the row or column label or both
- **Net present value:** measures the worth of a stream of cash flows, taking into account the time value of money
- **Discount rate:** reflects the opportunity costs of spending funds now versus achieving a return through another investment
- **MIN():** finds the smallest value in a range of cells
- **MAX():** finds the largest value in a range of cells
- **SUM():** finds the sum of values
- **AVERAGE():** finds the average
- **COUNT():** finds the number of cells within a range that meet a specified criterion
- **COUNTIF():** finds the number of cells within a range that meet a specified criterion
- **IF():** a logical function that returns one value if the condition is true and another if the condition is false
- **AND():** a logical function that returns TRUE if all conditions are true and FALSE if not
- **OR():** a logical function that returns TRUE if any condition is true and FALSE if not
- **VLOOKUP():** looks up a value in the leftmost column
- **HLOOKUP():** look up a value in the top row of a table and returns a value in the same column from a row you specify
- **INDEX():** returns a value or reference of the cell at the intersection of a particular row and column in a given range
- **MATCH():** returns the relative position of an item in an array that matches a specified value in a specified order

Business Analysis Notes

- CHAPTER 3

- Data visualization: the process of displaying data in a meaningful fashion to provide insights that will support better decisions
- Dashboard: a visual representation of a set of key business measures
- A clustered column chart: compares values from across categories using vertical rectangles
- Line charts: provide a useful means for displaying data over time
- Pie chart: partitions a circle into pie-shaped areas showing the relative proportion
- Area chart: combines the features of a pie chart with those of line charts
- Scatter charts: relationship between two variables
- Bubble chart: a type of chart in which the size of the data marker corresponds to the value of a third variable
- Sparklines: graphics that summarize a row or column of data in a single cell
- Pareto analysis: sorting data and calculating the cumulative percentage of the characteristic of interest
- Descriptive statistics: methods of describing and summarizing data using tubular, visual, and quantitative techniques
- Frequency distribution: a table that shows the number of observations in each of several non overlapping groups
- Relative frequency: $\text{FREQUENCY OF CATEGORY}/n$
- Cumulative frequency: list the fraction of data that is below or above a particular value
 - Cumulative distribution: Insert-Recommended charts-Pareto
- Histogram: a graphical depiction of a frequency distribution for numerical data in the form of a column chart
 - Excel: Insert-Chart-Histogram
- Group width: $(\text{UPPER LIMIT} - \text{LOWER LIMIT})/\# \text{ OF GROUPS}$
- kth percentile: a value at or below which at least k percent of the observations lie
 - $=\text{PERCENTILE.INC}(\text{array}, k)$
- Quartiles: break the data into four parts
 - $=\text{QUARTIL.INC}(\text{array}, \text{quart})$
- Tufte's Principles:
 - Clear, detailed through labeling and appropriate scales
 - Size of graphic effect should be directly proportional to numerical quantities
 - Maximize data-ink ratio
 - Avoid chart junk

Business Analysis Notes

- CHAPTER 4

- Arithmetic mean: average; sum of the observations divided by the number of observations
- Outliers: observations that are radically different from the rest
- Median: the measure of location that specifies the middle value when the data are arranged from least to greatest
- Mode: observation that occurs most frequently
 - Excel: =MODE.SNGL(data range)
- Dispersion: the degree of variation in the data
 - Range
 - Interquartile range
 - Variance
 - Population variance: average of the squared deviations from the mean
 - Excel: VAR.P(data range)
 - Excel: VAR.S(data range)
 - High variance: many observations away from mean
 - Low variance: few observations away from mean
 - Standard deviation: equals the square root of variance
 - Excel of population: STDEV.P(data range)
 - Excel for sample: STDEV.S(data range)
 - Small deviation: few observations far from mean
 - Big deviation: many observations far from mean
 - Coefficient of variance: STANDARD DEVIATION/MEAN
 - Covariance: a descriptive measure of linear relationship
 - A positive value indicates an increasing relationship
 - A negative value indicates a decreasing relationship
 - Excel: Data-Data Analysis-Covariance
 - Correlation coefficient: does not depend on units but gives us a measure of linear association
 - Negative correlation: data in quadrants 2 and 4
 - Positive correlation: data in quadrants 1 and 3
 - Excel: CORREL(array 1, array 2)

Business Analysis Notes

- CHAPTER 5

- Probability: likelihood that an outcome occurs; values between 0 and 1
- Experiment: a process that results in an outcome
- Outcome: a result that we observe
- Sample space: collection of all possible outcomes from a sample space
- Rule 1: The probability of any event is the sum of the probabilities of the outcomes that comprise that event
- Rule 2: The probability of the complement of any event A is $P(A^c) = 1 - P(A)$
-