K2 Analytics
Building Skills, Building Individuals

Analytics Basics

- Rajesh Jakhotia

16-Feb-2017

Agenda

- Analytics - What it is? and Why Analytics?
- Supervised and Unsupervised Learning
- Target
- Dependent and Independent Variable
- Hypothesis Creation
- Modeling Base (Dev, Validation and Hold-out)
- Window Definition
- Test & Control

Analytics – What it is?

- Analytics – dictionary definition

  "the science of logical analysis"

- Analytics – technical definition

  "the application of computer technology, operational research, and statistics to solve problems in business and industry. Analytics is carried out within an information system"

- Analytics – business definition

  "*simplify data to amplify its meaning*"

What is Predictive Analytics?

- ***Predictive Analytics*** *is an area of statistical analysis that deals with* ***extracting information (insights)*** *from data and using it to* ***predict future*** *trends and behavior patterns*

Why Analytics Matters?

Validates your Gut Feeling

"Without data you're just another person with an opinion"

- W. Edwards Deming

Quantifies the Impact

"If you **can't measure** it, you **can't** improve it"

- Peter Drucker

Empowers you take decision

K2 Analytics
Building Skills, Building Individuals

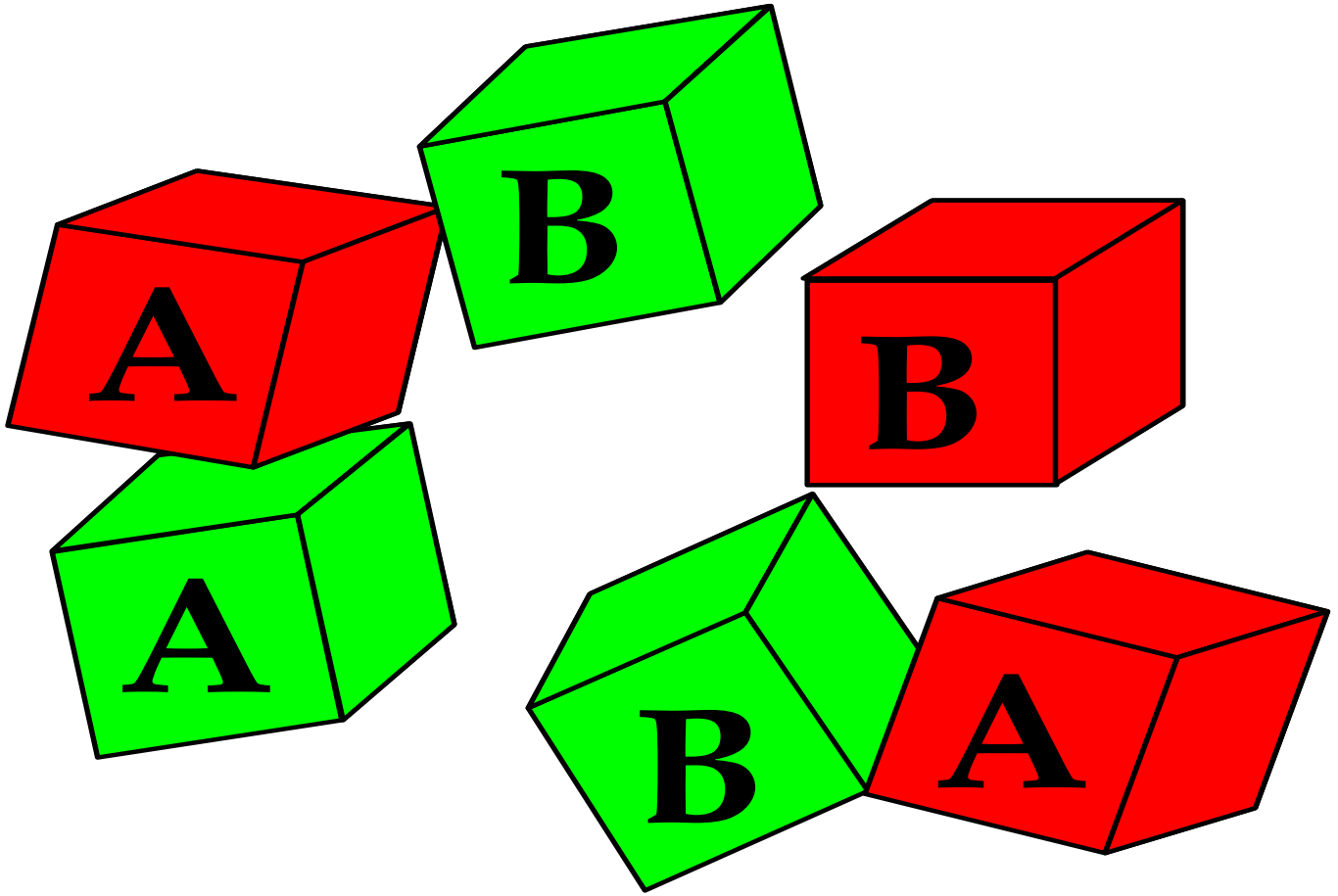
Supervised vs Unsupervised Learning

Types of Data Mining Techniques

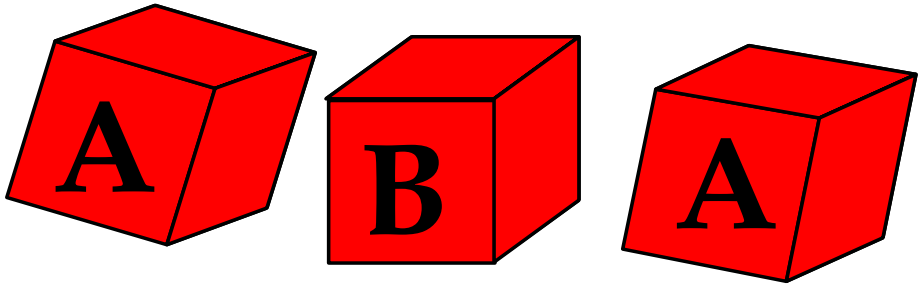
- Supervised learning: The target output expected is clearly defined

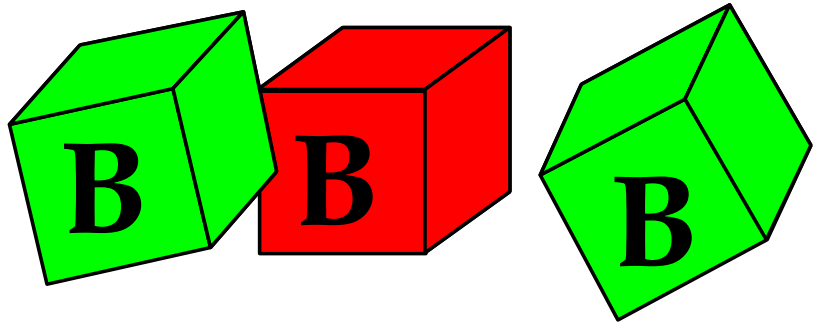
- Unsupervised learning: The data have no target attribute.
  - We want to explore the data to find some intrinsic structures in them.

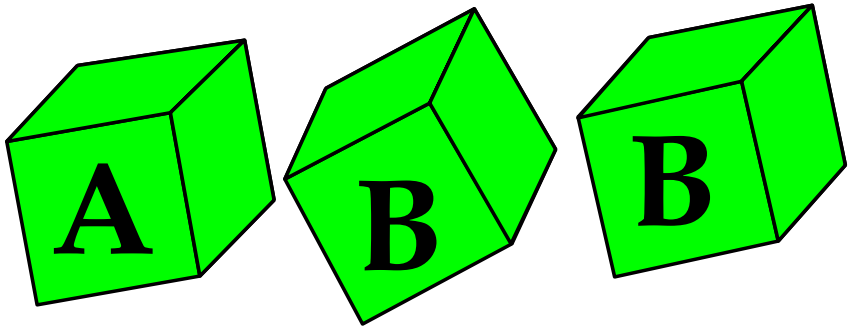
Understanding Supervised vs. Unsupervised Learning

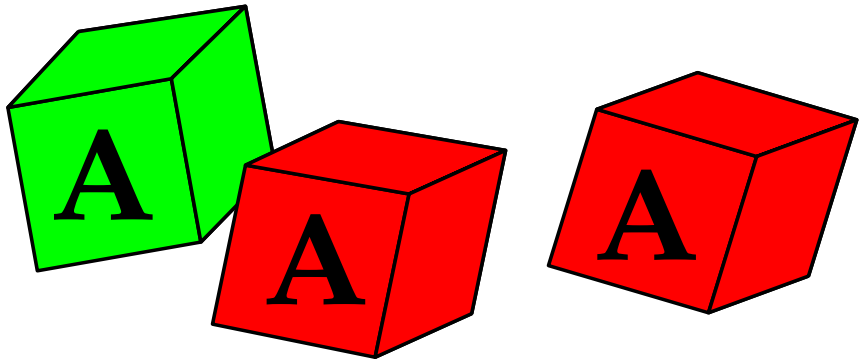


Two possible solutions



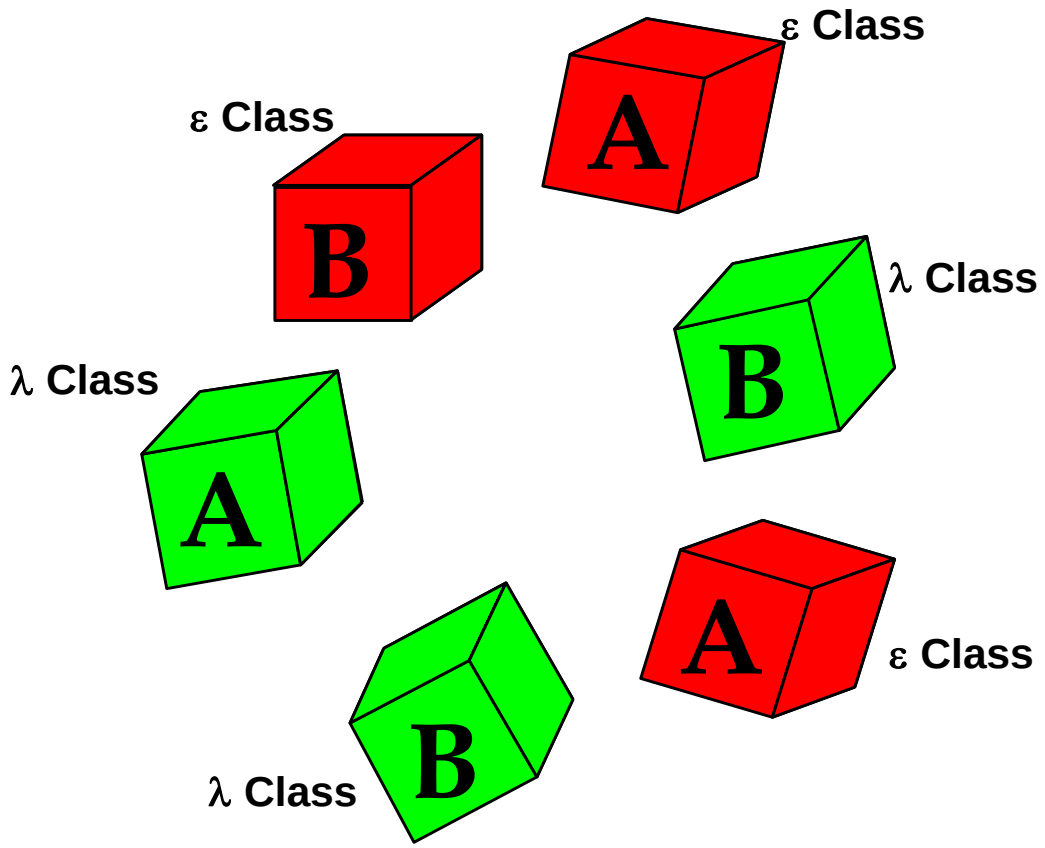






Supervised Learning

- It is based on a labeled training set.
- The class of each piece of data in training set is known.
- Class labels are pre-determined and provided in the training phase.



Examples of Supervised Learning

Industry / Vertical	Business Objective	Modeling Objective	Predicted Event Terminology
HR	To arrest employee attrition	To build a model that will help predict employee attrition	Attrition Rate
Telecom	To reduce customer churn	To build a model that will help predict the likelihood of customer churn	Churn Rate
Retail / Ecommerce	To win back churned customers	To build a model that will help predict and identify the customers who are likely to transact again	Win Back Rate
Banking	To cross-sell banking product / service	To build a model that will help assign the probability to a customer to take a product / service	Response Rate
Insurance	To reduce policy lapsation	To build a model to assess the likelihood of customer not renewing his / her policy	Lapsation Rate

K2 Analytics
Building Skills, Building Individuals

Modeling Terminologies

Business Objective vs. Modeling Objective

- **Business Objective**: The goal that a Business User wish to achieve
    - Typical Business Objective in Predictive Modeling is related to Customer Lifecycle Management

  Acquire → Activate → Upsell / Cross-sell → Retain → **(Grow)**

- **Modeling Objective**: The Business Objective converted into a Mathematical form for analysis is Modeling Objective
    - Acquisition Rate, Activation Rate, Response Rate, Retention Rate, Churn Rate, Win back Rate

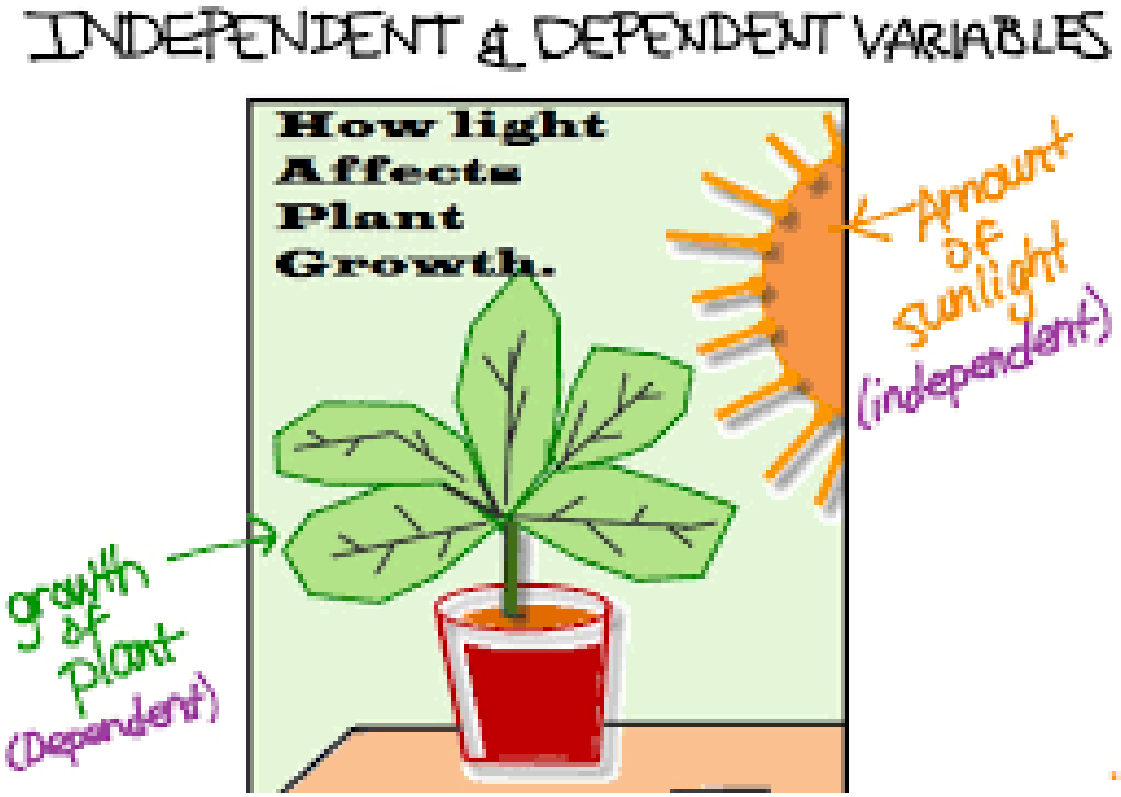
Target

Target: The phenomenon or the event being analysed is called Target

Target is often referred as Dependent Variable in statistics

- Possible outcomes of the Target
  - **Positive**: The occurrence of the event; typically it is represented as 1
  - **Negative**: The non-occurrence of the event; typically it is represented as 0
  - **Indeterminate**: Grey zone where we cannot clearly demarcate occurrence / non-occurrence of the event; typically the indeterminate cases are removed from analysis purview

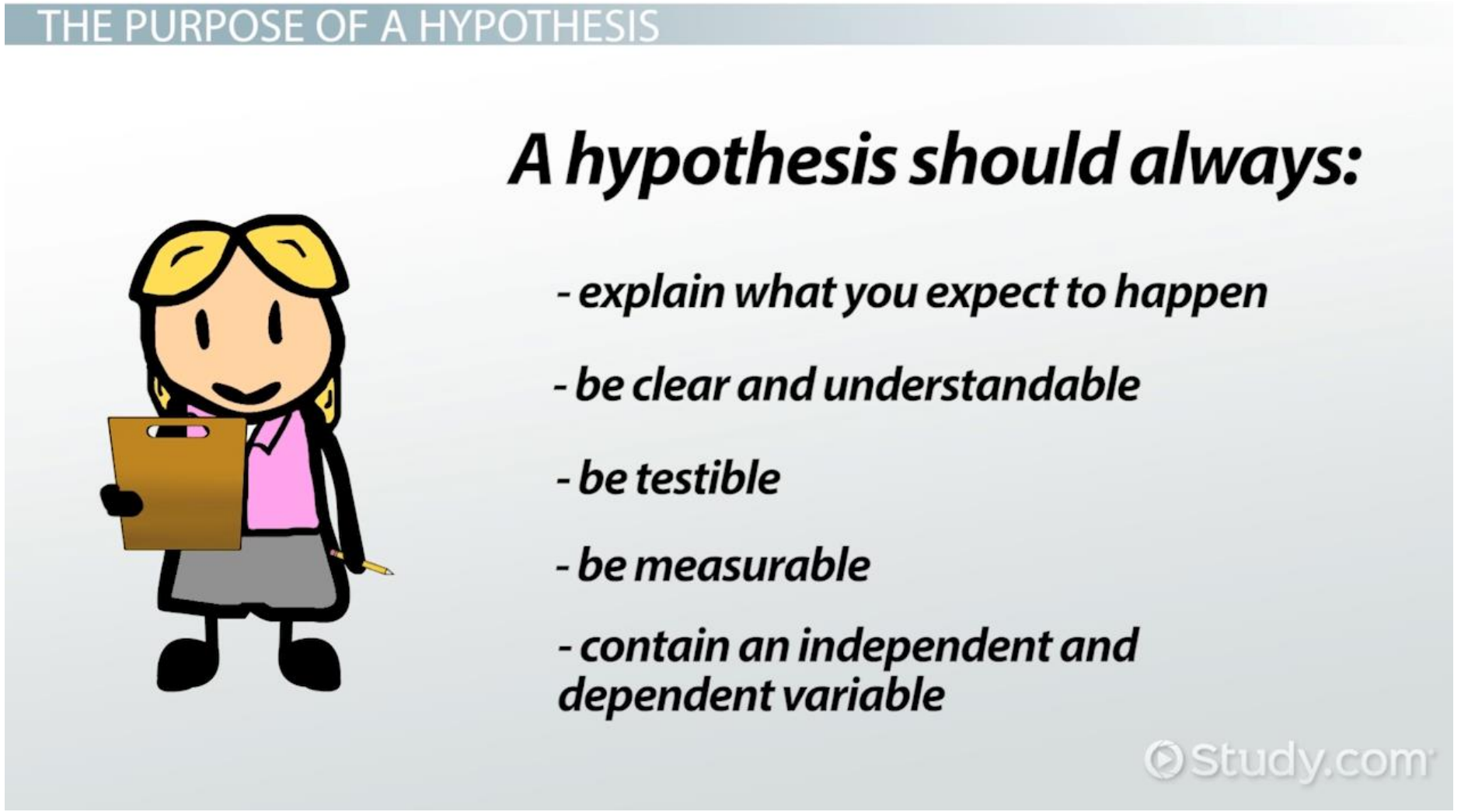
Independent Variables



- Independent Variable : The variable which is **experimented** in order to observe its effect on the Dependent Variable
- It is also often referred as Predictor Variable

http://www.showme.com/sh/?h=9WsGXQG

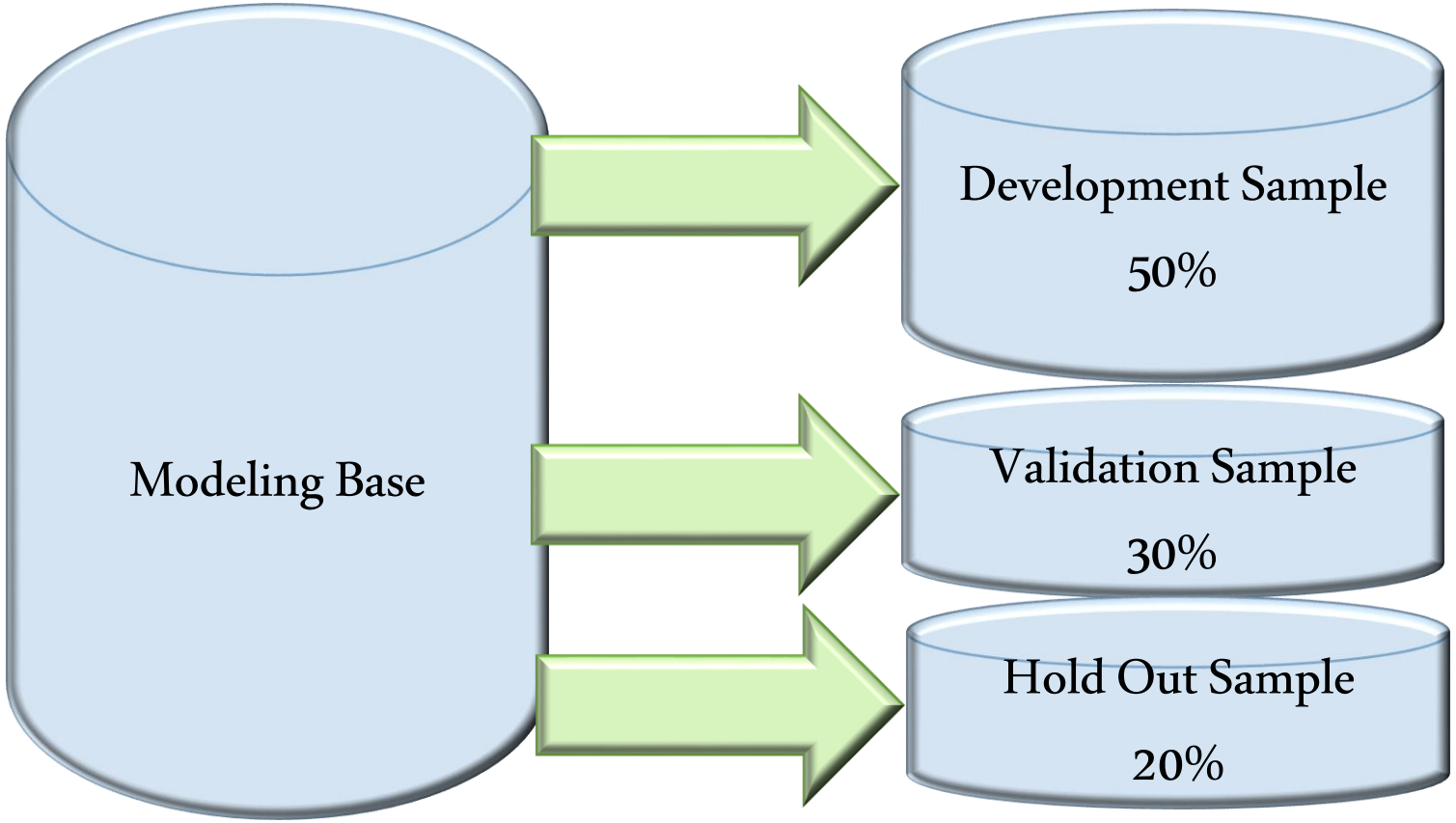
Hypothesis Creation



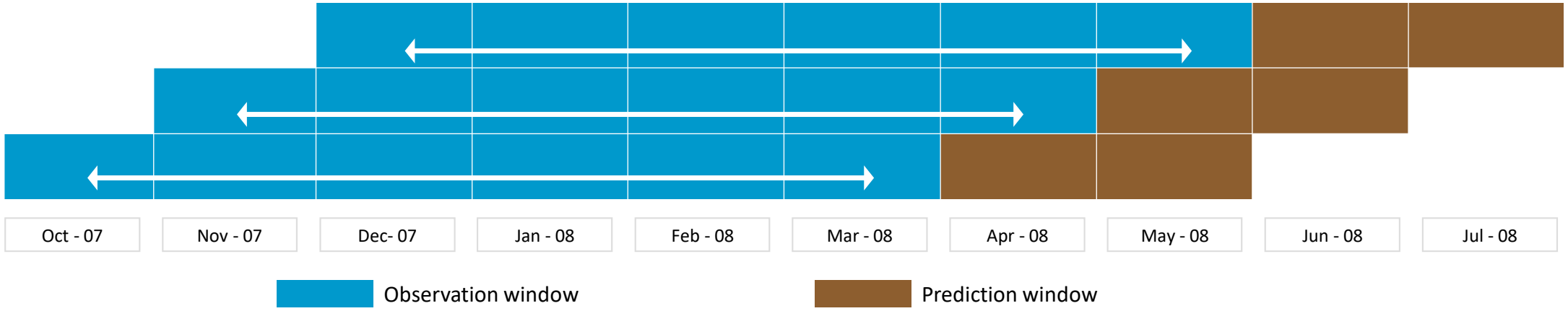
http://study.com/academy/lesson/what-is-a-hypothesis-definition-lesson-quiz.html

Modeling Base

- Modeling Base: The base on which you build the model
- Modeling base comprises of Dependent and Independent Variables



Window Definition



- Observation /Analysis Window - The window in which the behavior of the customer is analyzed for modeling
- Prediction window - The period for which the customers behavior is predicted for creation of target variable

Test & Control

Test vs. Control

What is Target Base?

- The base eligible to be targeted for campaign is called Target Base.
- The target base is typically generated after applying various filters, de-duplication, model based prioritization (if any)

Experimental vs. Control Groups

Experimental (Test) Group	Control Group
• 1 variable (thing) changes or is tested	•Comparison group •No changes •"Normal" conditions

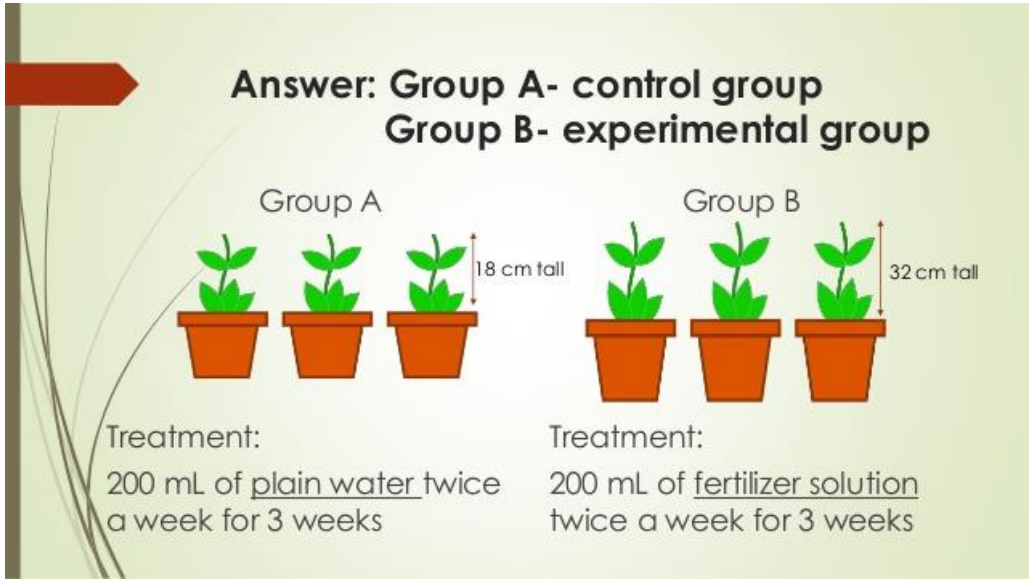
What is Control Group?

- A control group is a subset of the customers eligible to receive your campaign but you hold back this subset and do not send the campaign

What is Test Group?

- The set of customers eligible to receive your campaign and to whom you send the campaign

Why use Control Group?



- Control Group provides you a baseline
- Control Group helps you do "**what-if**" analysis if you would not have executed the campaign

- In absence of Control Group, you will not have a reference point and as such you will not be able to measure the efficacy of your model

AB Testing

- AB Testing is about comparing two different versions of Offer (or design) to see which one works better. The one which works better wins!!!
- AB Testing is also sometimes called Split Testing
- Also sometimes called as Champion-Challenger
- Quite often A/B Testing is used to test various creative / web page design
- E.g. You wish to increase sales of Laptop of particular brand. You wish to understand customer preference for the below two offers:
  - 10% discount on Laptop
  - One year additional warranty
- **Limitations** : A/B testing can’t isolate which variables drive customers to act, and the rate of organizational learning is relatively slow.

K2 Analytics
Building Skills, Building Individuals

Questions

Contact us:
ar.jakhotia@k2analytics.co.in