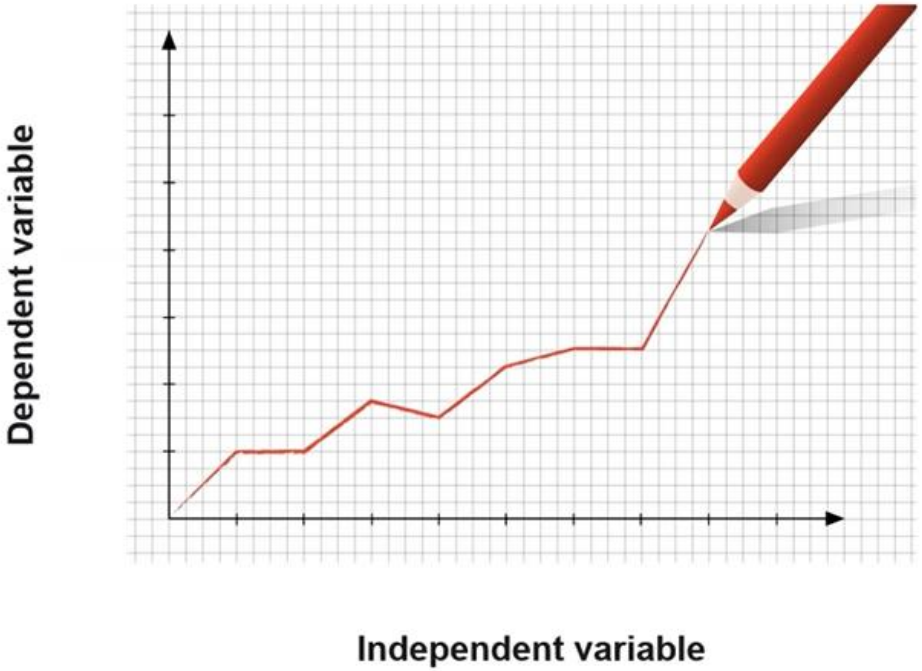
Data Collection From Surveys and Experimentation Design

Department of Computer Science

Variables



- **Dependent variables** receive this name because, in an experiment, their values are studied under the supposition or demand that they depend, by some law or rule (e.g., by a mathematical function), on the values of other variables.
- **Independent variables**, in turn, are not seen as depending on any other variable in the scope of the experiment in question.

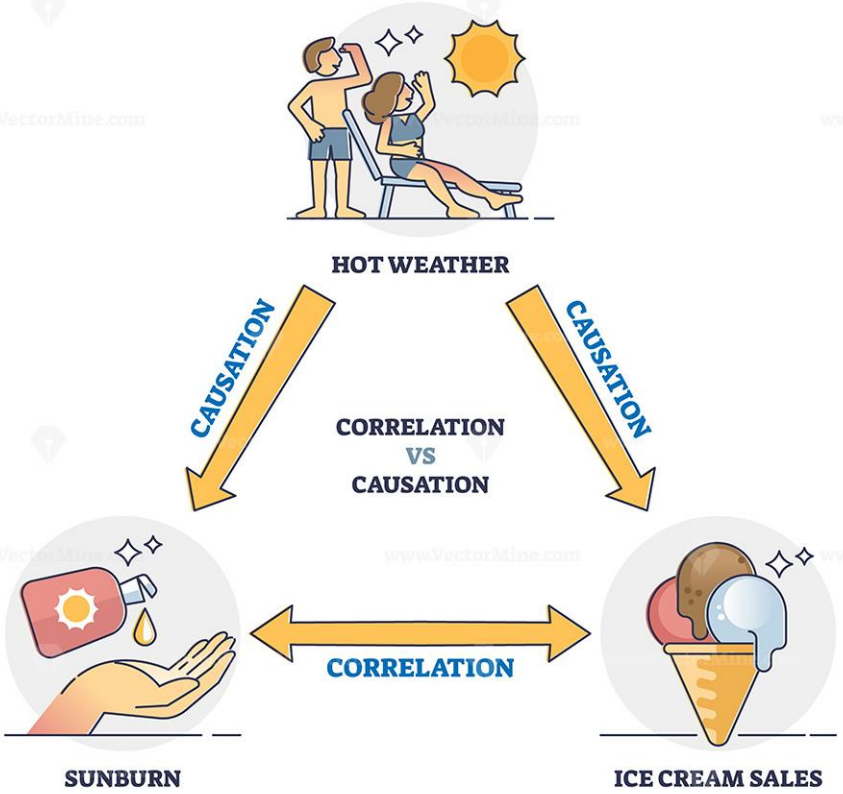
Use-Case

$\propto$

- The presence or absence of amount caffeine is **independent variable** where as how much hungry you are is **dependent variable**
- **Caffeine may reduce feelings of hunger and your desire to eat for a brief time.**

Caffeine effects appetite

Concept of Causality



- Causality (also referred to as **causation**, or **cause and effect**) is influence by which one event, process, state, or object (a cause) contributes to the production of another event, process, state, or object (an effect) **where the cause is partly responsible for the effect, and the effect is partly dependent on the cause**.

Concept of Causality

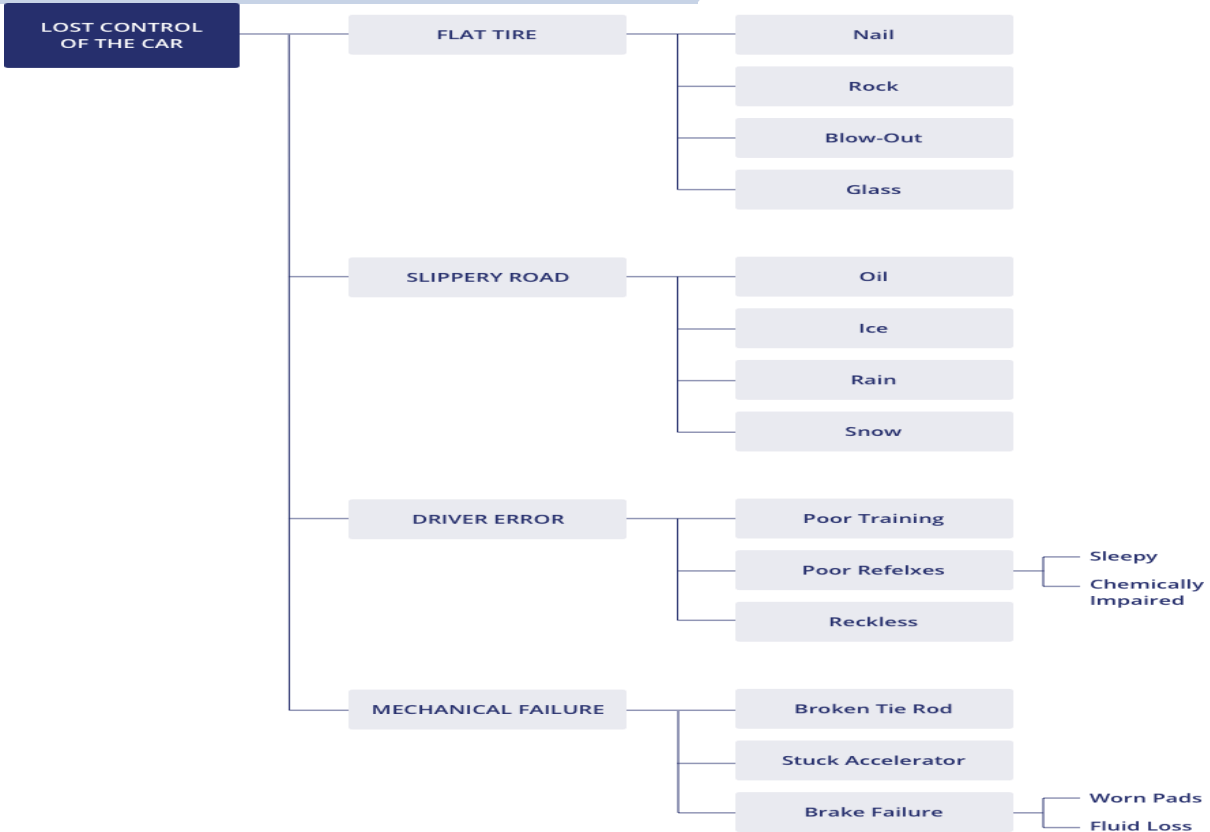
Ordinary Meaning	Scientific Meaning
X is the only cause of *Y*.	*X* is only one of a number of possible causes of *Y*.
X must always lead to *Y* (*X* is a deterministic cause of *Y*).	The occurrence of *X* makes the occurrence of *Y* more probable (*X* is a probabilistic cause of *Y*).
It is possible to prove that *X* is a cause of *Y*.	We can never prove that *X* is a cause of *Y*. At best, we can infer that *X* is a cause of *Y*.

In **deterministic models**, the output of the model is fully determined by the **parameter values** and the initial values, whereas **probabilistic (or stochastic) models** incorporate **randomness in their approach**.

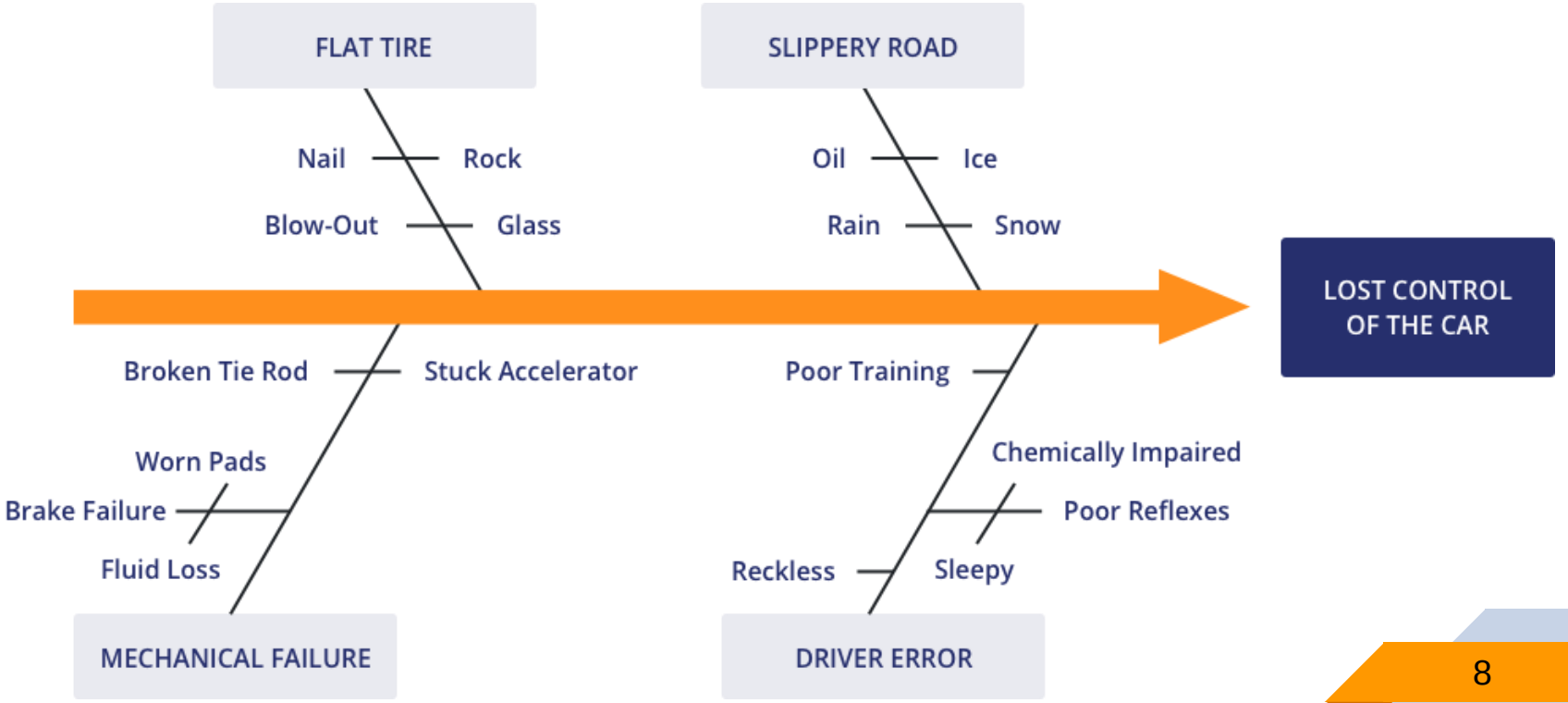
Condition for Causality

- ***Concomitant variation*** is the extent to which a cause, X and an effect Y occur together or vary together in the way predicted by the hypothesis under consideration. (**e.g: if a brand's advertising expenditures have been cut in half and the brands sales fell, we may suspect that the reduced advertising support caused sales to fall.)**
- **The time order of occurrence** condition states that the causing event must occur either before or simultaneously with the effect it cannot occur afterward
- The **absence of other possible causal factors** means that the factor or variable being investigated should be the only possible causal explanation**.**

Cause and Effect Diagram Examples



Cause and Effect Diagram Examples



Ingredients of an Experiment

- **Independent variables** are variables that are manipulated and whose effects are measured and compared, e.g., price levels.
- **Test units** are individuals, organizations, or other entities whose response to the independent variable or treatment is being examined, e.g., consumers or stores.
- **Dependent variables** are the variables which measure the effect of the independent variables on the test units, e.g., sales, profits, and market shares.
- **Extraneous variables** are all variables other than the independent variables that affect the response of the test units, e.g., store size, store location, and competitive effort.
- **Mediating variables** provide the causal link between other variables.
- **Moderating variables** influence the relation between two other variables and provide an 'interaction effect'.
- **Experimental group** is exposed to the treatment variable.
- **Control group** is not exposed to the treatment variable.

Validity in Experiment

- When conducting an experiment, A researcher has two goals;
  - Draw valid conclusions about the effects of independent variables on the study group
  - Make valid generalizations to a larger population of interest.
- The first goal concerns internal validity and the second goal concerns external validity
- Internal and external validity are concepts that reflect whether or not the results of a study are trustworthy and meaningful.
- While internal validity relates to how well a study is conducted (its structure), external validity relates to how applicable the findings are to the real world

Internal Validity

- A **measure of accuracy** of the experiment.
- It measure **whether the manipulation of the independent variables or treatments, actually caused the effects** on the dependent variables.
- Control of extraneous variables is a necessary condition for establishing internal validity

External Validity

- Refers to whether cause and effect relationship found in the experiment can be generalized.

Example;

Can the effects of customer preferences be generalized to fast food chains.

Internal and external validity comparison

Internal Validity

- Conclusions are warranted
- Controls extraneous variables
- Eliminates alternative explanations
- Focus on accuracy and strong research methods

External Validity

- Findings can be generalized
- Outcomes apply to practical situations
- Results apply to the world at large
- Results can be translated into another context

Classification of Experimental design

- Experimental design may be classified as;
  - Pre-experimental design
  - True-experimental design
  - Quasi-experimental design
  - A statistical design

Pre-experimental design

- It does not employ randomization procedures to control for extraneous factors.

  Example;
  One shot case study
  One group pre test – post test design
  Static-group Comparison

- A group, or various groups, are kept under observation after implementing factors of cause and effect.
- Pre-experimental design being the simplest form of **exploratory design**, include **one or more than one experimental group to be observed against certain treatment**.

True-experimental design

- The researcher can randomly assign test units and treatments to experimental groups

  Example;

  Pre test – post test control group design

- True experimental design is a **statistical approach** of establishing a cause and effect relationship between different variables. This is one of the most accurate forms of research designs which provides a substantial backing to support the existence of relationships.
- Randomization means sudden selection without any plan.
- Example: A lottery system. The lottery numbers are announced at random so everyone who buys a lottery has an equal chance. Hence, it means you select a sample without any plan and everyone has an equal chance of getting into any one of the experimental groups.

Quasi-experimental design

- When the researcher is unable to achieve full manipulation of scheduling or allocation of treatments to test unit but can still apply part of the apparatus of true experimentation.

- Like a true experiment, a quasi-experimental design aims to establish a cause-and-effect relationship between an independent and dependent variable.

- However, unlike a true experiment, a quasi-experiment does not rely on random assignment. Instead, subjects are assigned to groups based on non-random criteria

Quasi-experimental design

- A chocolate and a crying child.
- Here, the Independent variable is: the type of chocolate
- And dependent variable is: the child is crying for a chocolate
- So manipulation means the effect of an independent variable that is chocolate, on the dependent variable that is the crying child.
- In short, **you are using an outside source on the dependent variable**.
- This proves that after getting the chocolate (independent variable), the child stops crying (dependent variable).

Comparison of 3 Experiments

	Pre-Experimental Design	True Experimental Design	Quasi- Experimental Design
Presence of a control group?	In some cases, Always but usually not	Always	Often
Random selection of subjects from a population?	No	Yes	No
Random assignment of subjects to groups?	No	Yes	No
Random assignment of treatments to groups?	No	Yes	No
Degree of control over extraneous variables?	None	Yes	Some