Data Analytics

Mrinal Das

Scenario 1

Often we have processes that transforms some inputs into outputs

We want to optimize the process

- to minimize resource consumption
- to minimize time
- to maximize output quality

We need to investigate the process:

interaction among inputs themselves, and with outputs to make it better

Scenario 2

Often we have processes that transforms some inputs into outputs

The process might be having some issues, output is not as desired

We want to fix the process.

We need to investigate the process:

interaction among inputs themselves, and with outputs to make it better

Scenario 3

Often we have to implement a new processes

The process is unknown involving new tools

We want to design the process:

- choose tools
- choose materials

We need to do pilot runs and understand

interaction among inputs themselves, and with outputs to design

Design of Experiments

Experiments are any activity, physical, chemical, computational

that given some input produces some output

Design of Experiments is systematic approach to plan experiments

given the goal of the experiment

What is the Goal

We need to **identify** the root cause of a problem

We need to **improve** product performance

We need to **optimize** a design / Process

We need to **improve** yield

We need to **reduce** cost

We need to **improve** OEE

Goal

What is DoE

A **designed experiment** or **DOE** is a statistical method that allows you to **quantify** the **relationship** between the **inputs** and **outputs** of a **process**.

The outputs are also called Responses or Dependent Variables or Y.

The inputs are also called Factors or Independent Variables, or Controllable Variable or X

Example

A car manufacturer wants to reduce the **paint defect rate**.

Potential factors:

- A: Paint temperature
- B: Spray pressure
- C: Drying time

Response:

- Y: Percentage of defective painted parts

DOE helps determine the settings of A, B, and C that minimize Y.

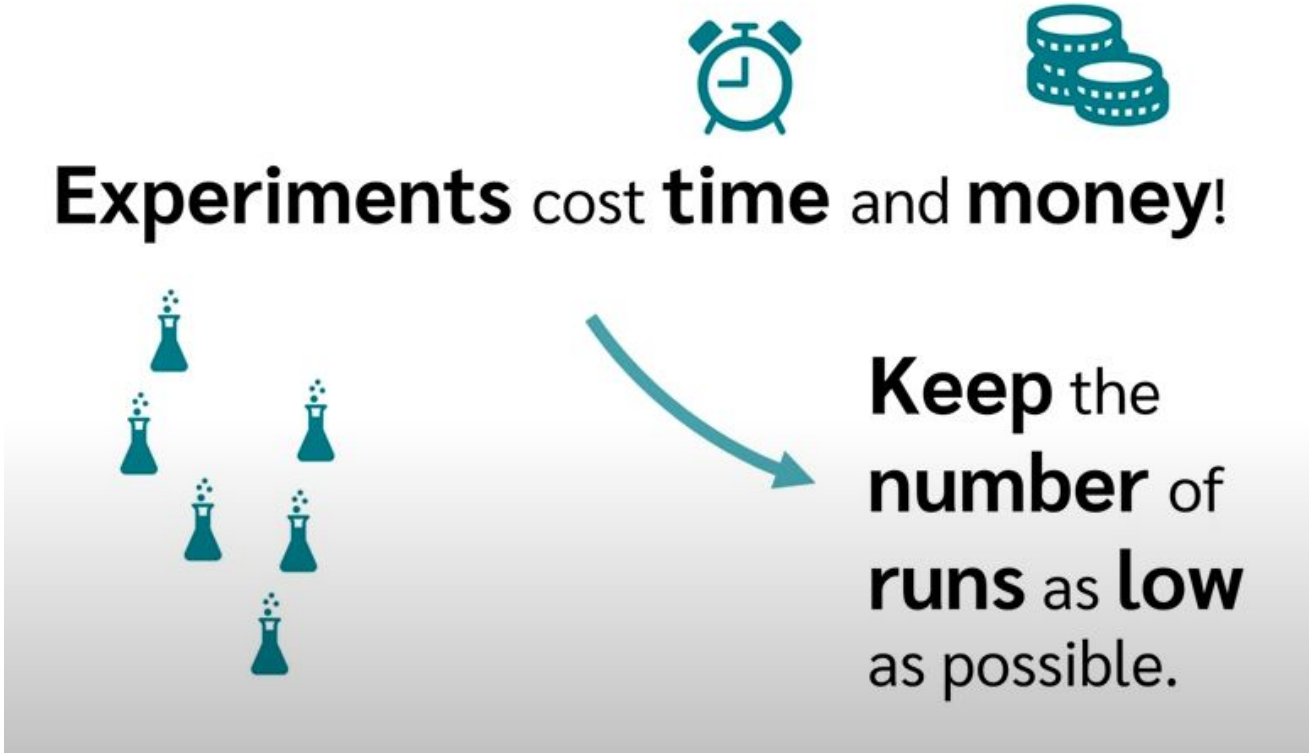
Systematic Planning

Factor	Low	High
Temperature	60°C	80°C
Pressure	2 bar	4 bar
Drying time	10 min	20 min

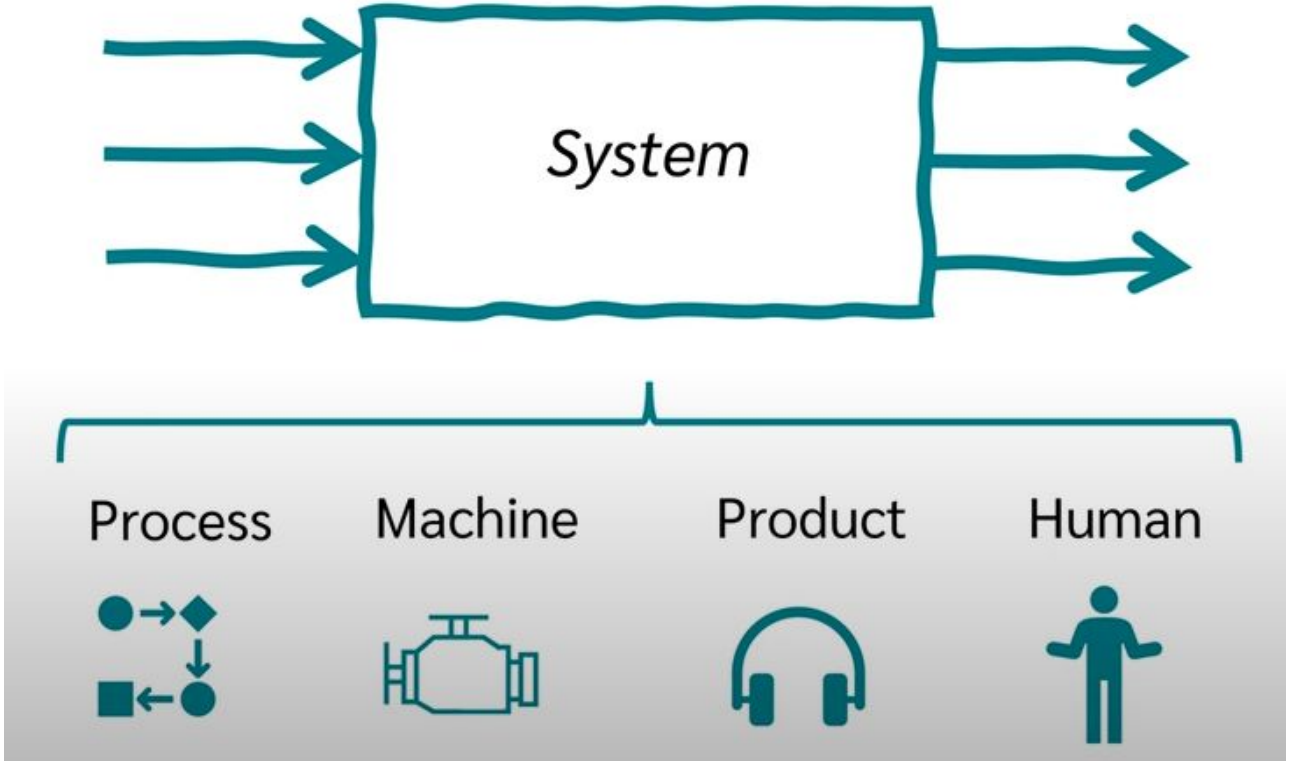
Many combinations possible and many values are possible.
Too many experiments and too much time to be spent.

How can we make it efficient ?

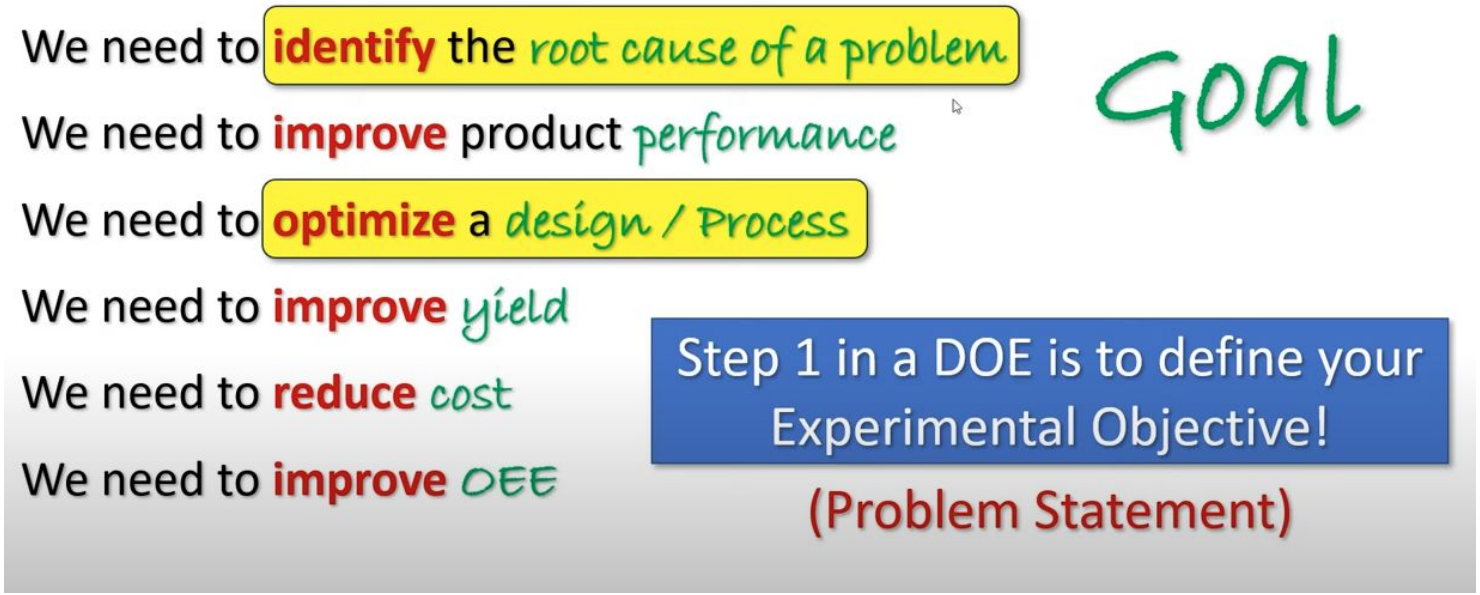
Advantage of DoE



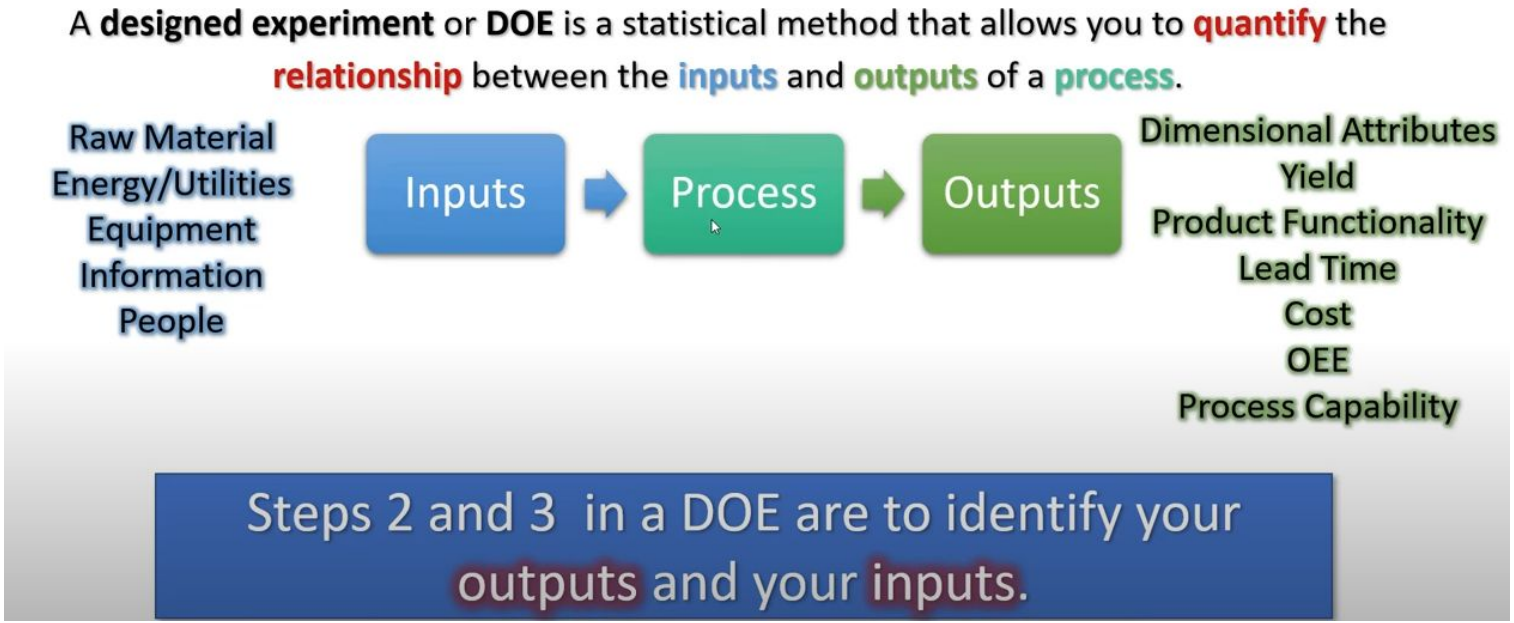
Where can we apply DOE



Step-1: Identify the Goal



Inputs and Outputs



Measurements are Important

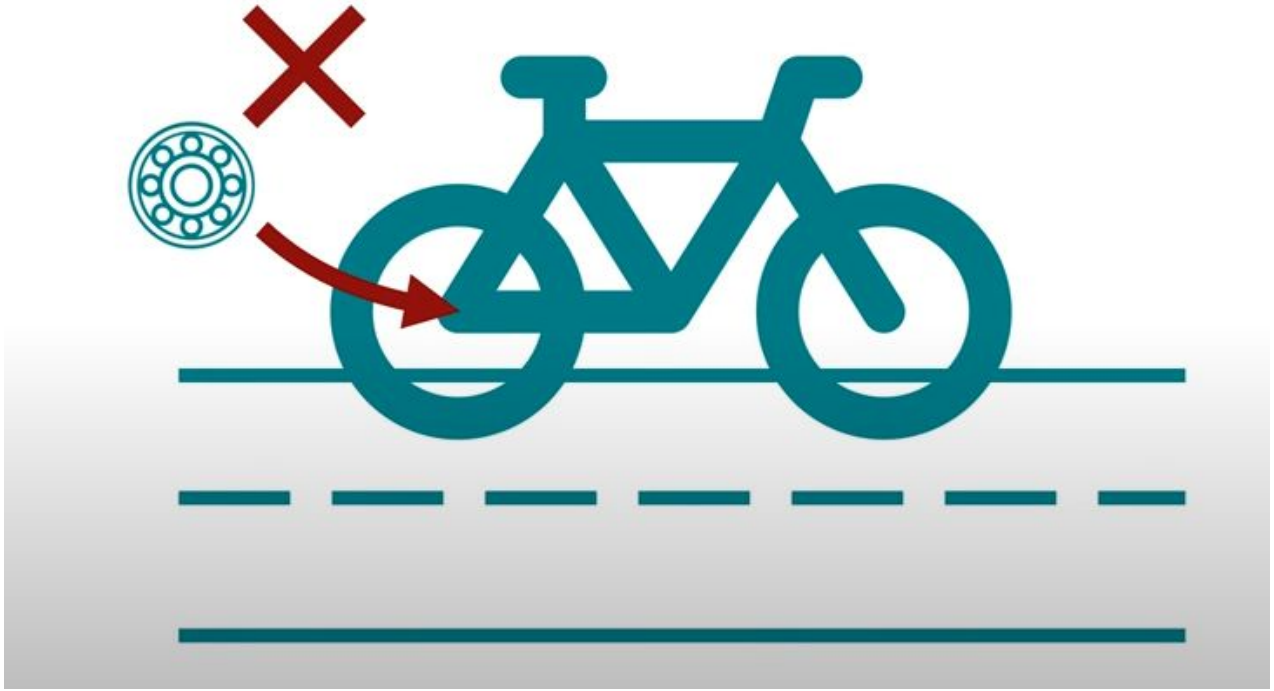
A **designed experiment** or **DOE** is a statistical method that allows you to **quantify** the **relationship** between the **inputs** and **outputs** of a **process**.

Inputs → Process → Outputs

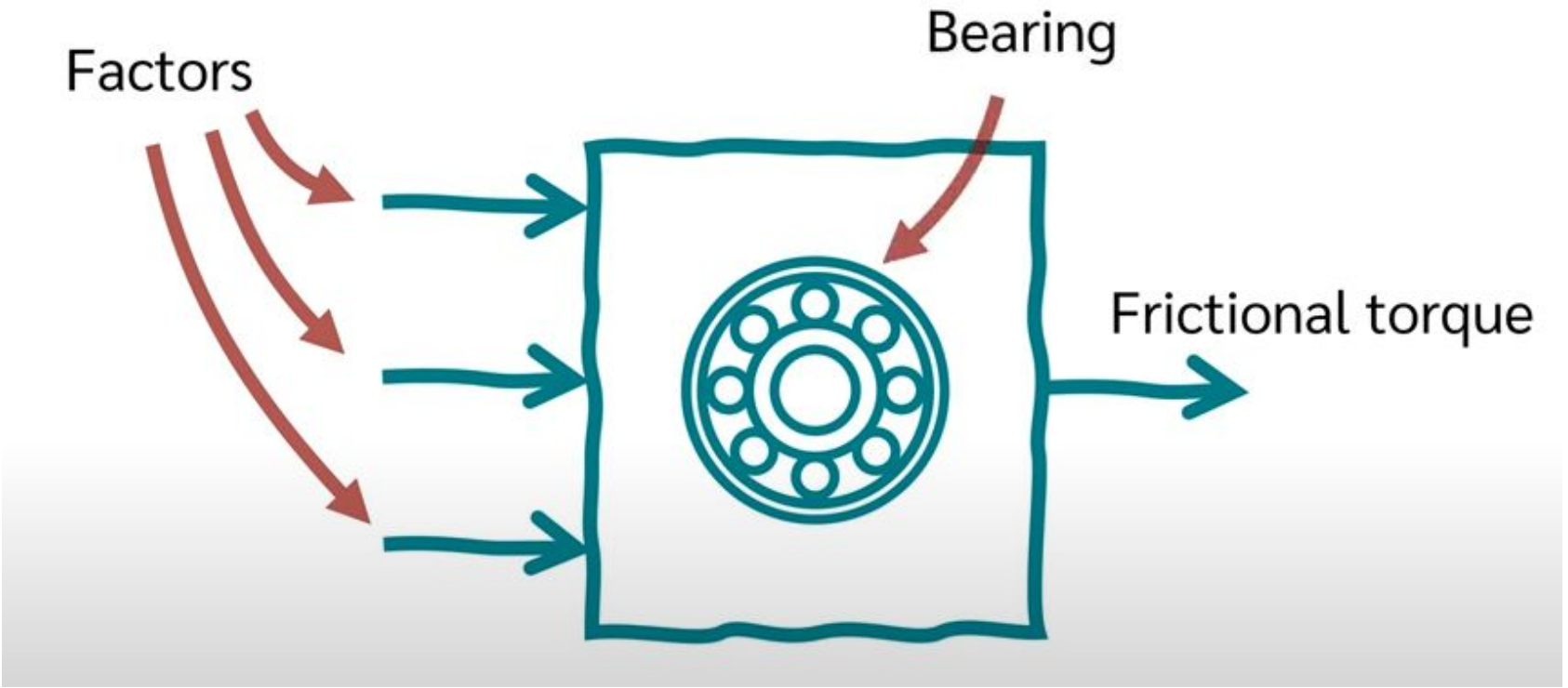
Dimensional Attributes
Product Functionality
Process Capability




Problem: Fix the issue with bearing



Input, Process, Output



In Nutshell

Systematic approach

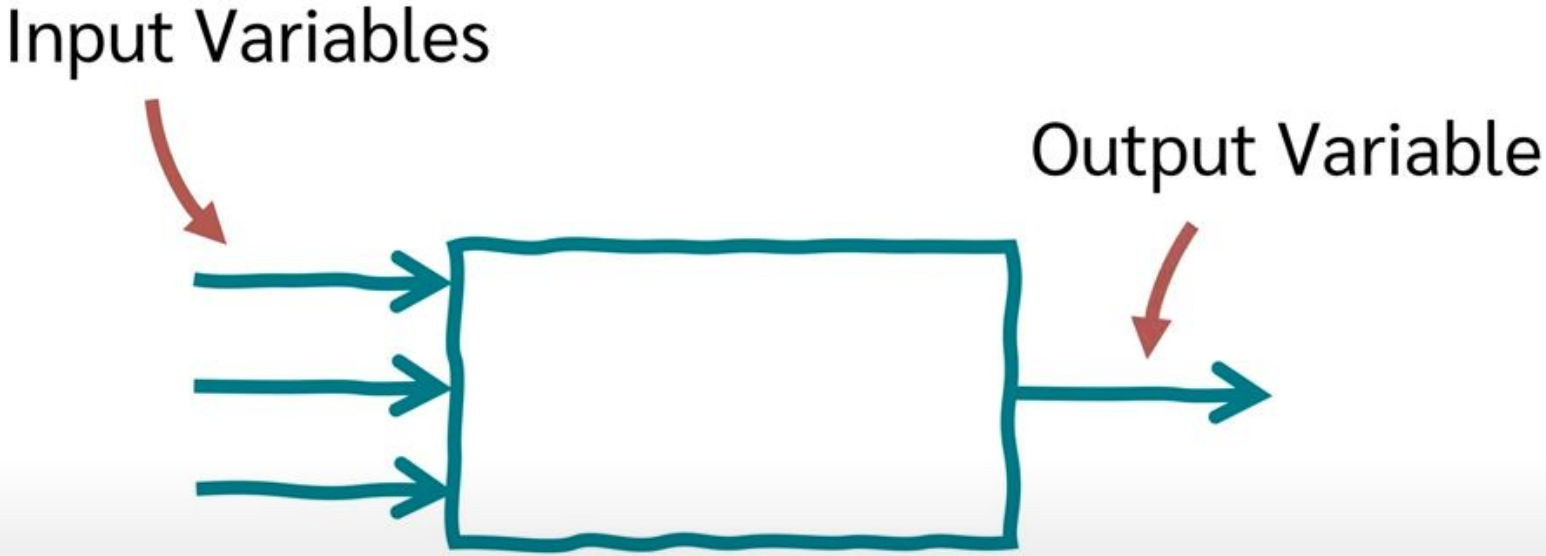
Planning

Conducting

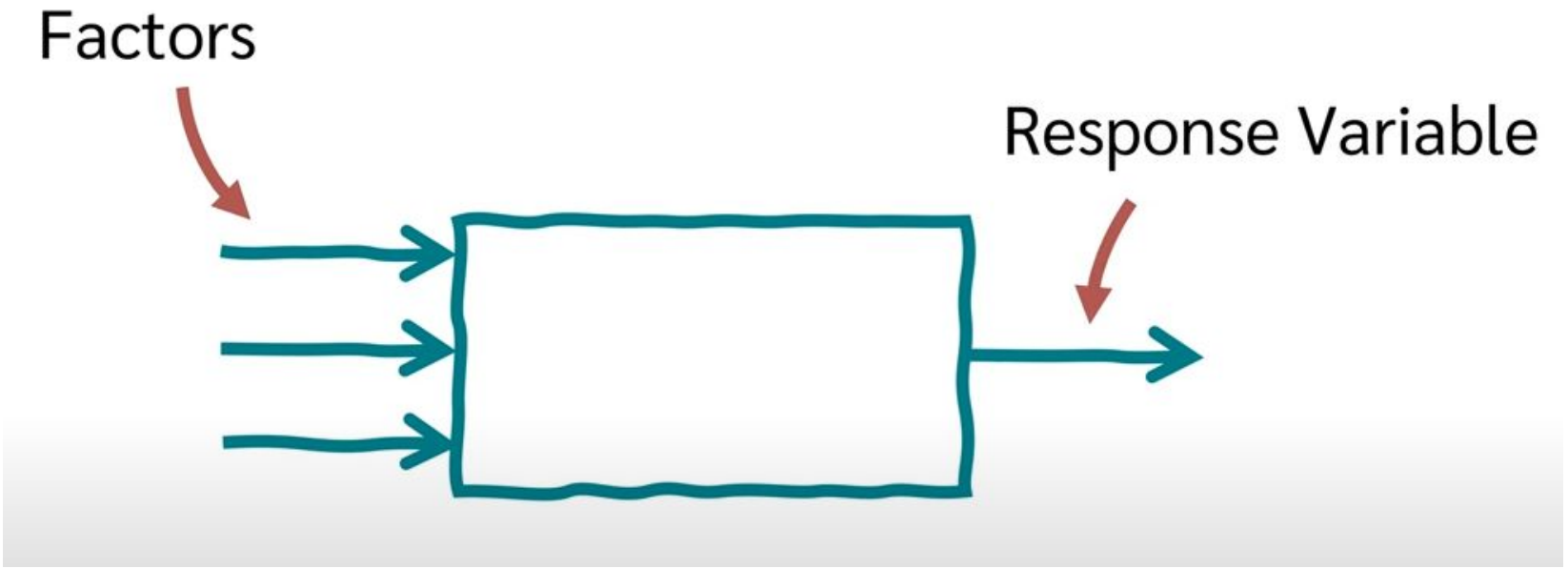
Analyzing

Aim of DoE

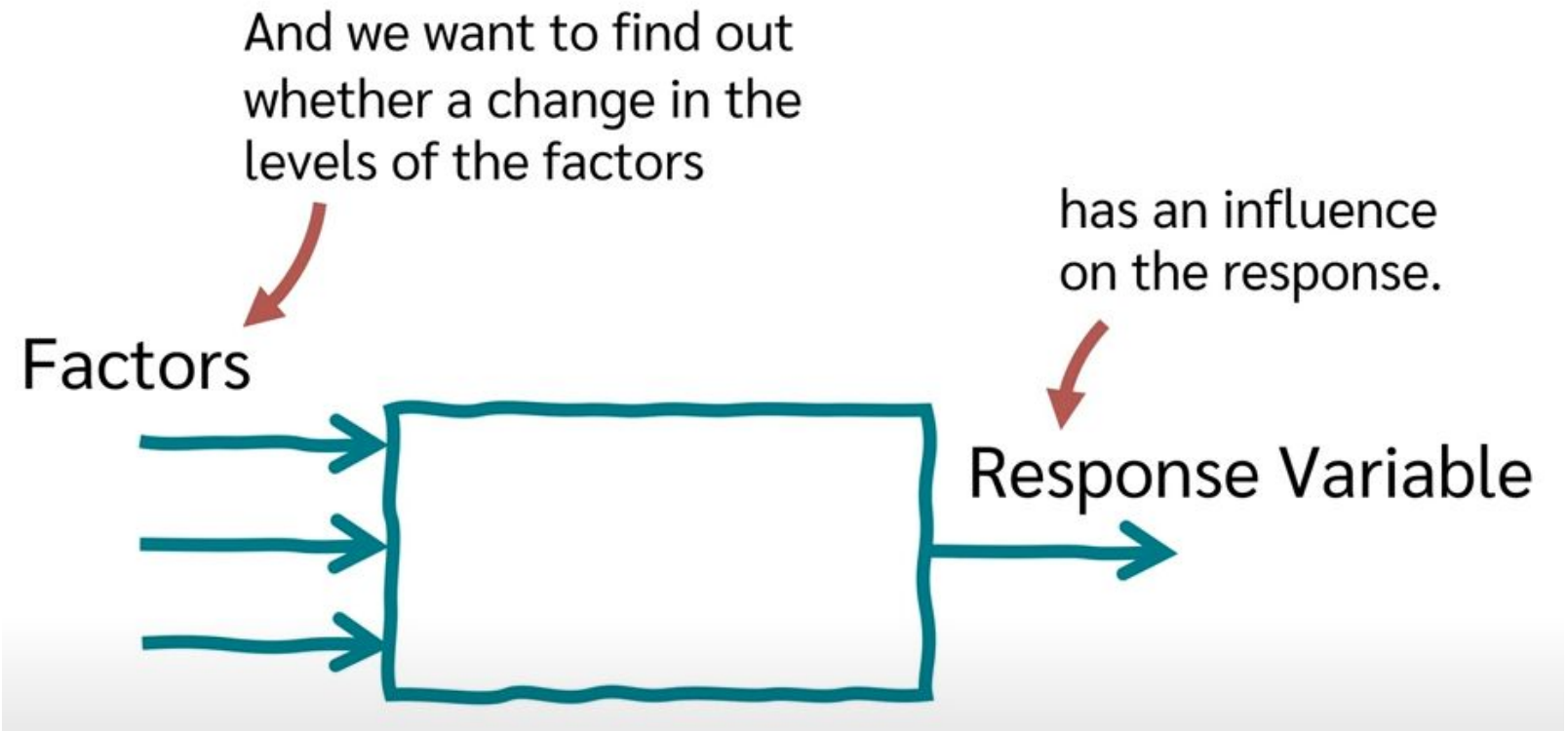
Analyse Relationship between Input and Output Variables



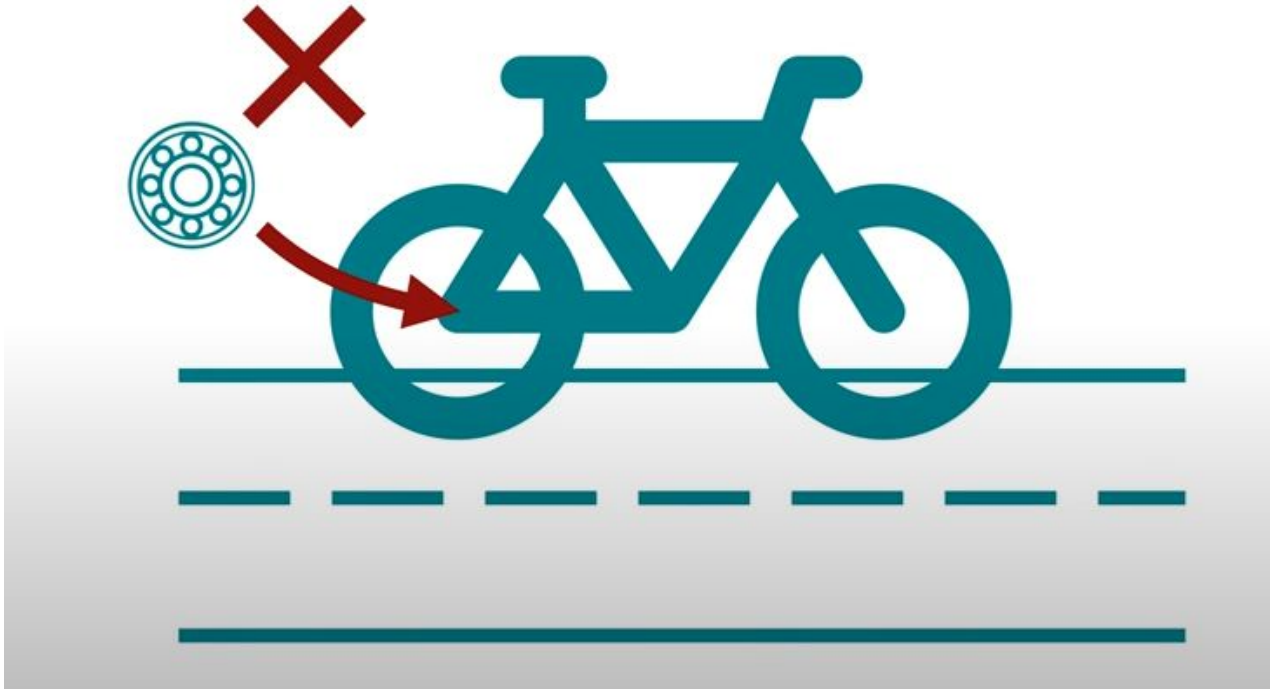
Factors and Response



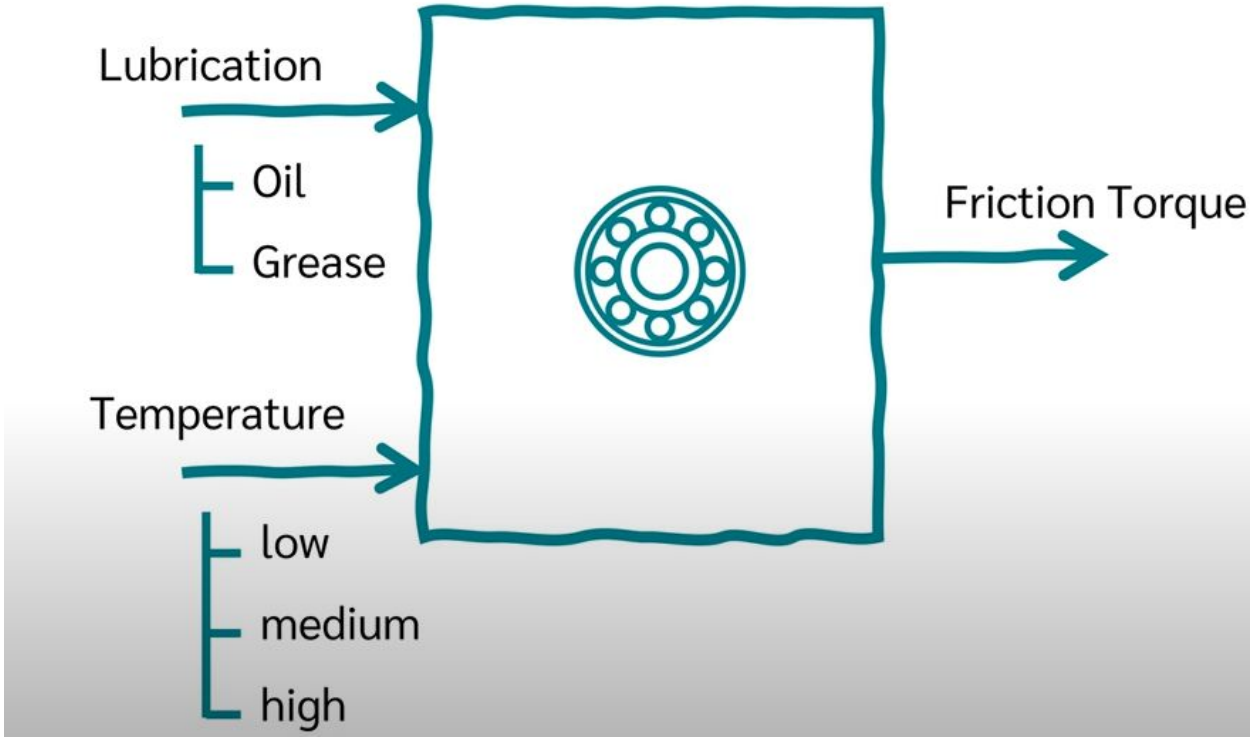
Objective


And we want to find out
whether a change in the
levels of the factors
has an influence
on the response.
Factors
Response Variable

Analysing the Friction in Bearing

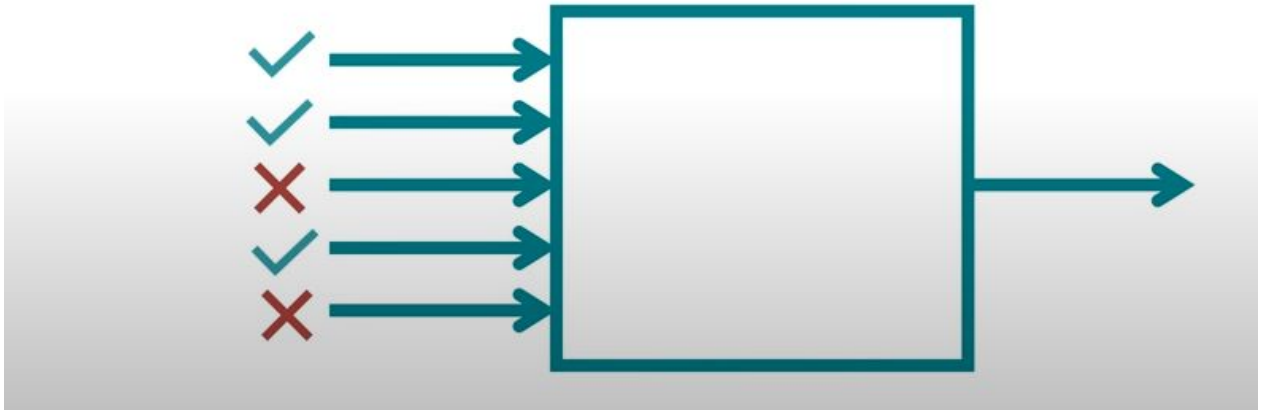


Factors and Levels



Step 2

1 Identify relevant factors



Step 3

1 Identify relevant factors

2 Determine optimal input variables

4 Phases in DOE

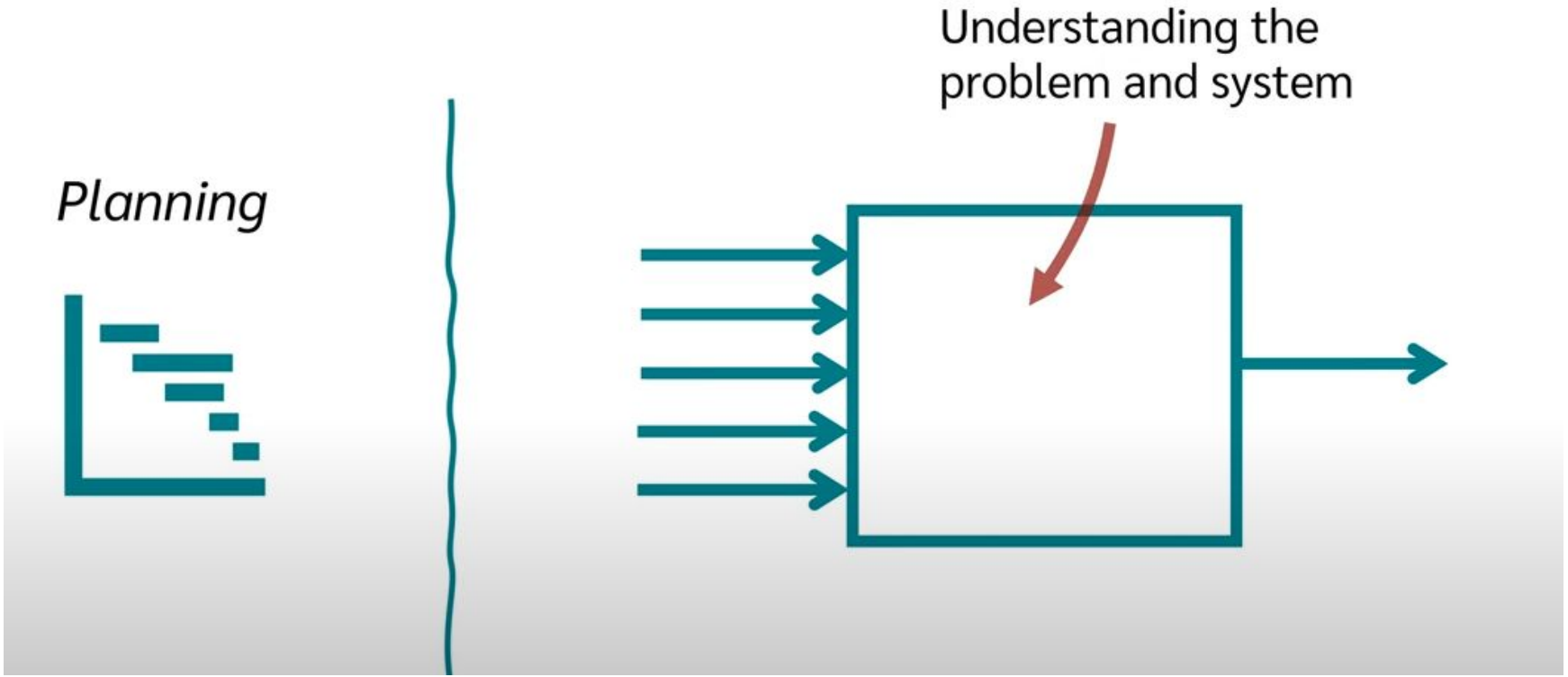
Planning

Screening

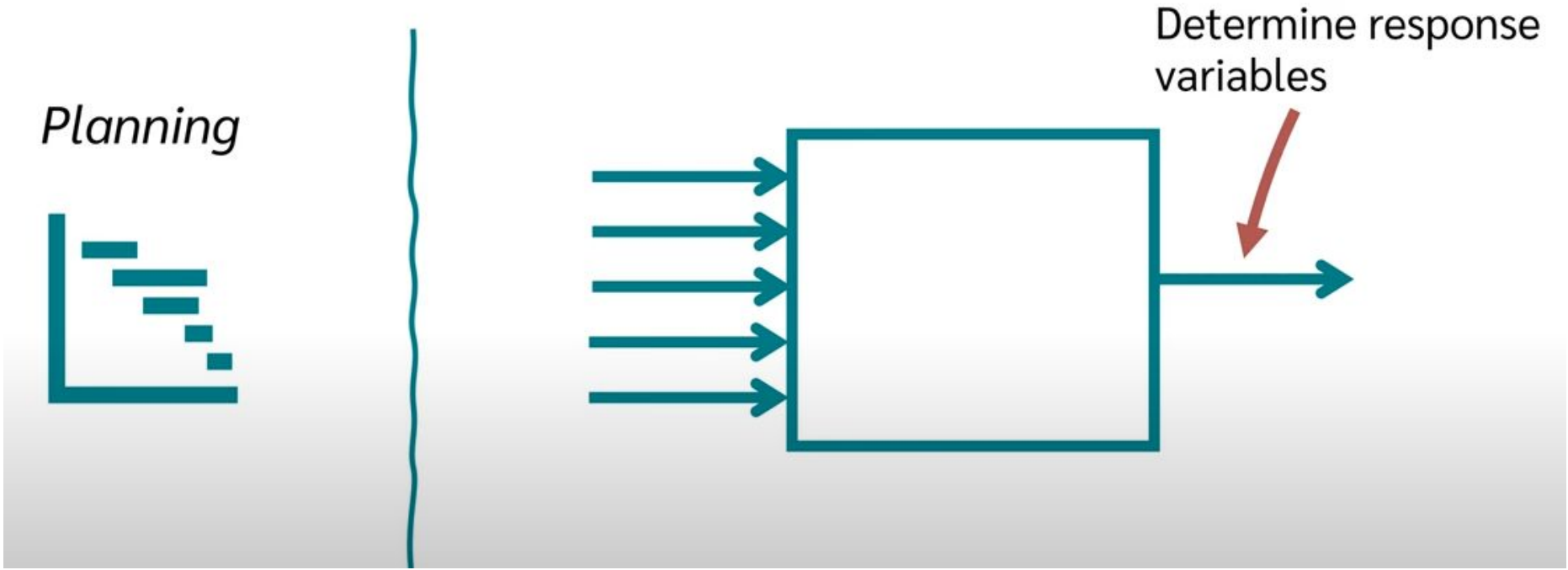
Optimization

Verification

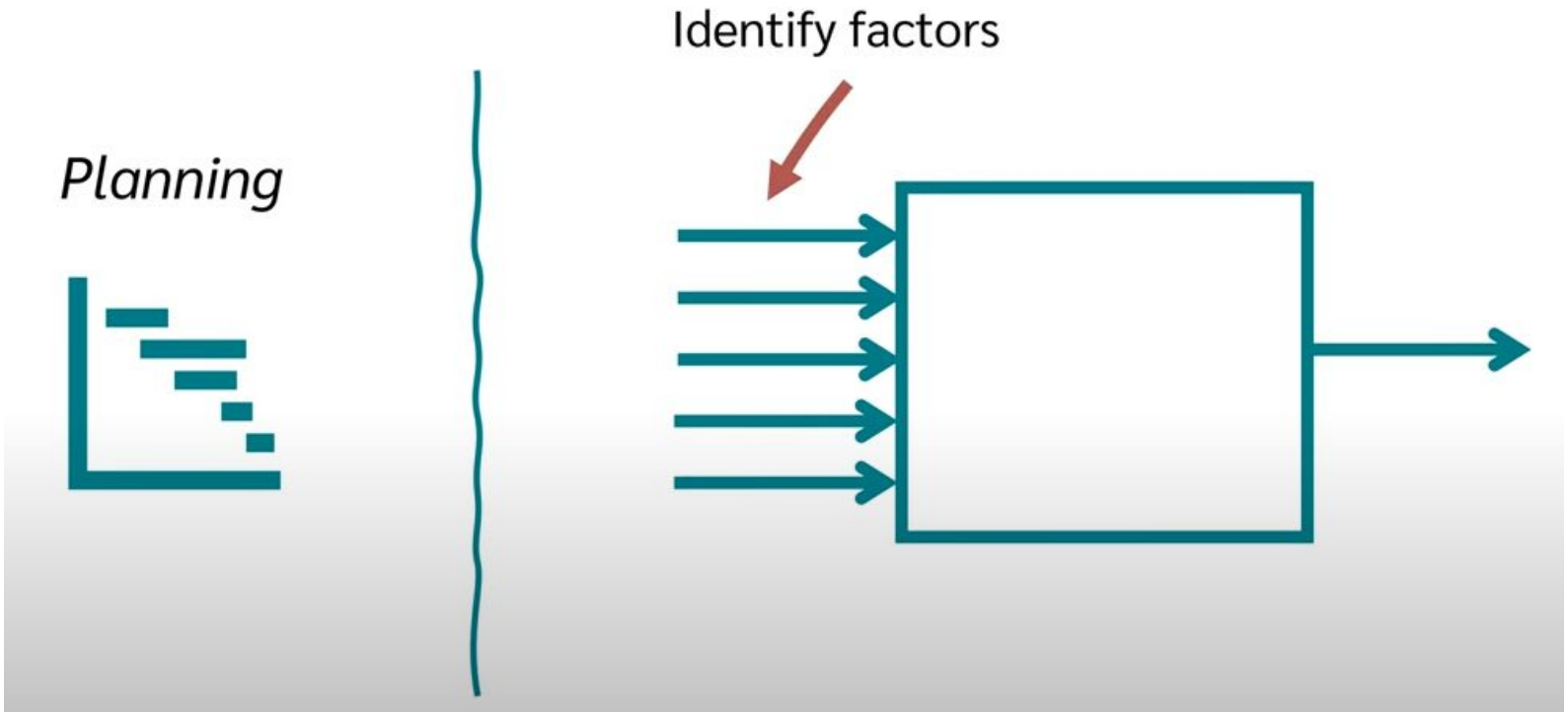
Phase 1



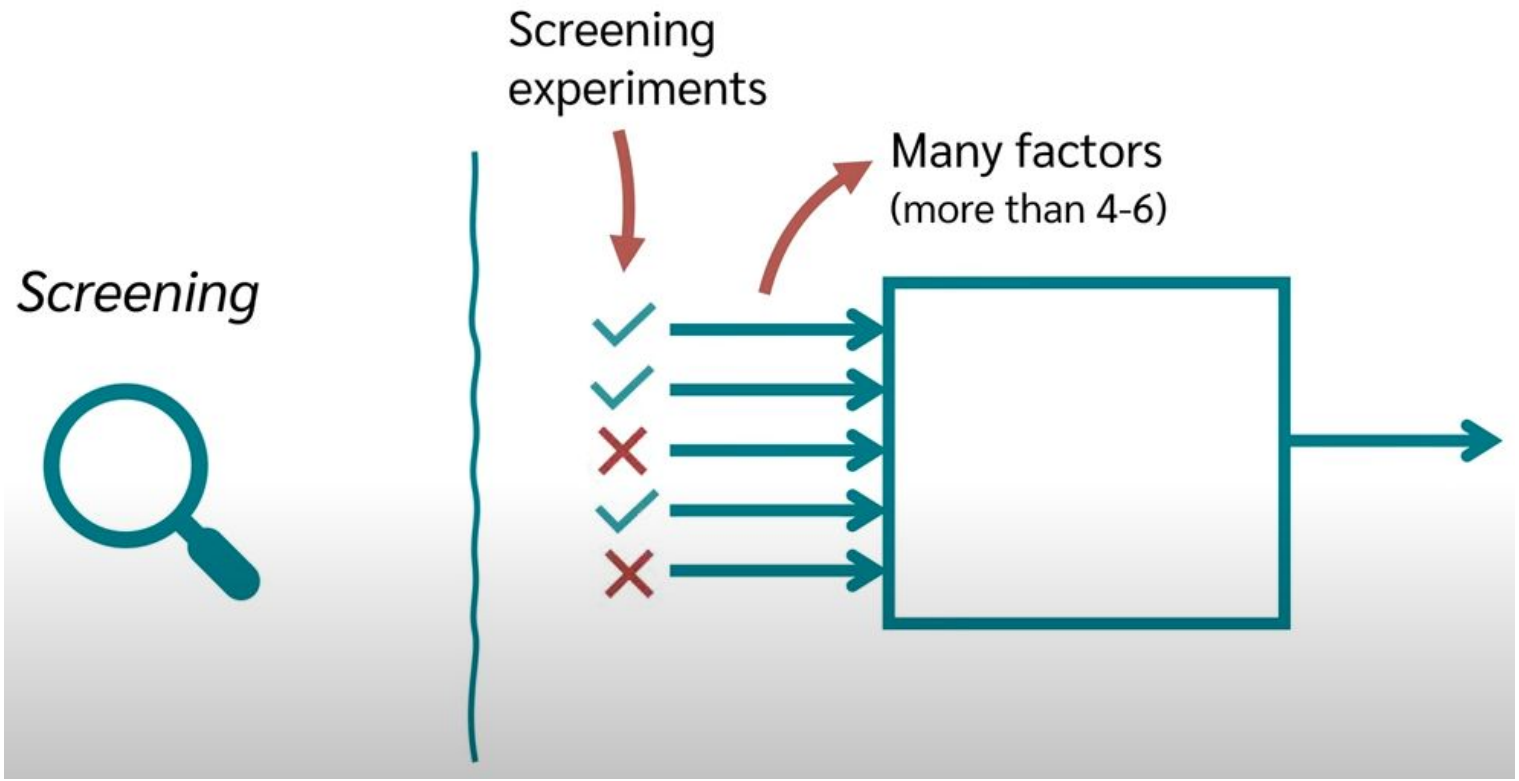
Phase 1



Phase 1



Phase 2


Screening
experiments
Many factors
(more than 4-6)
Screening

Phase 3

Full factorial design:

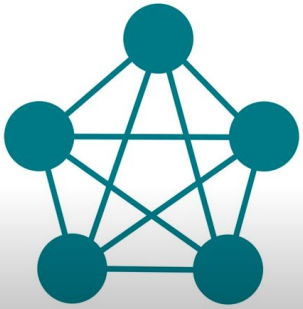
$$N = 2^k$$

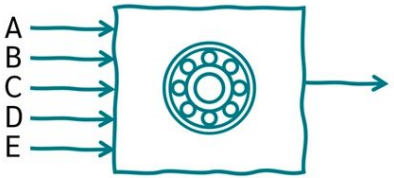
Number of runs

Number of factors

Full Factorial Design

Full factorial designs take **all interactions** into account.





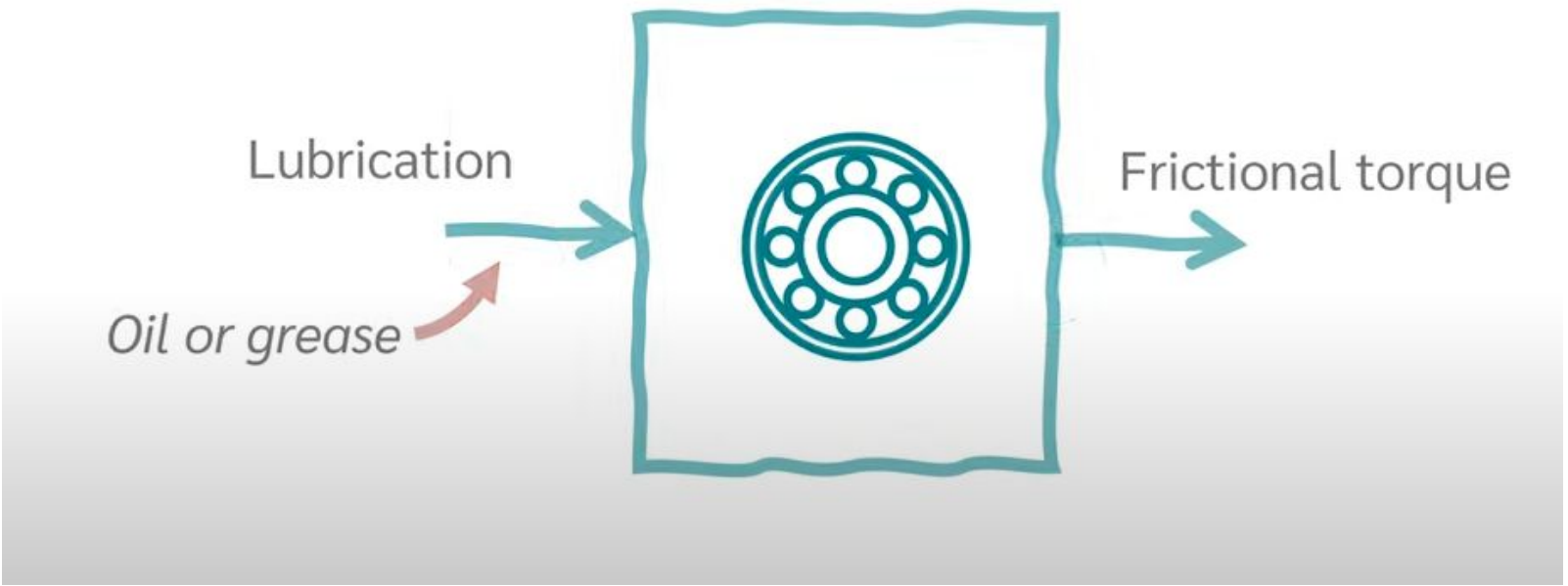
AB, AC, AD, AE, BC, BD, CD, ...

ABC, ABD, BCD, ABE, BCE ...

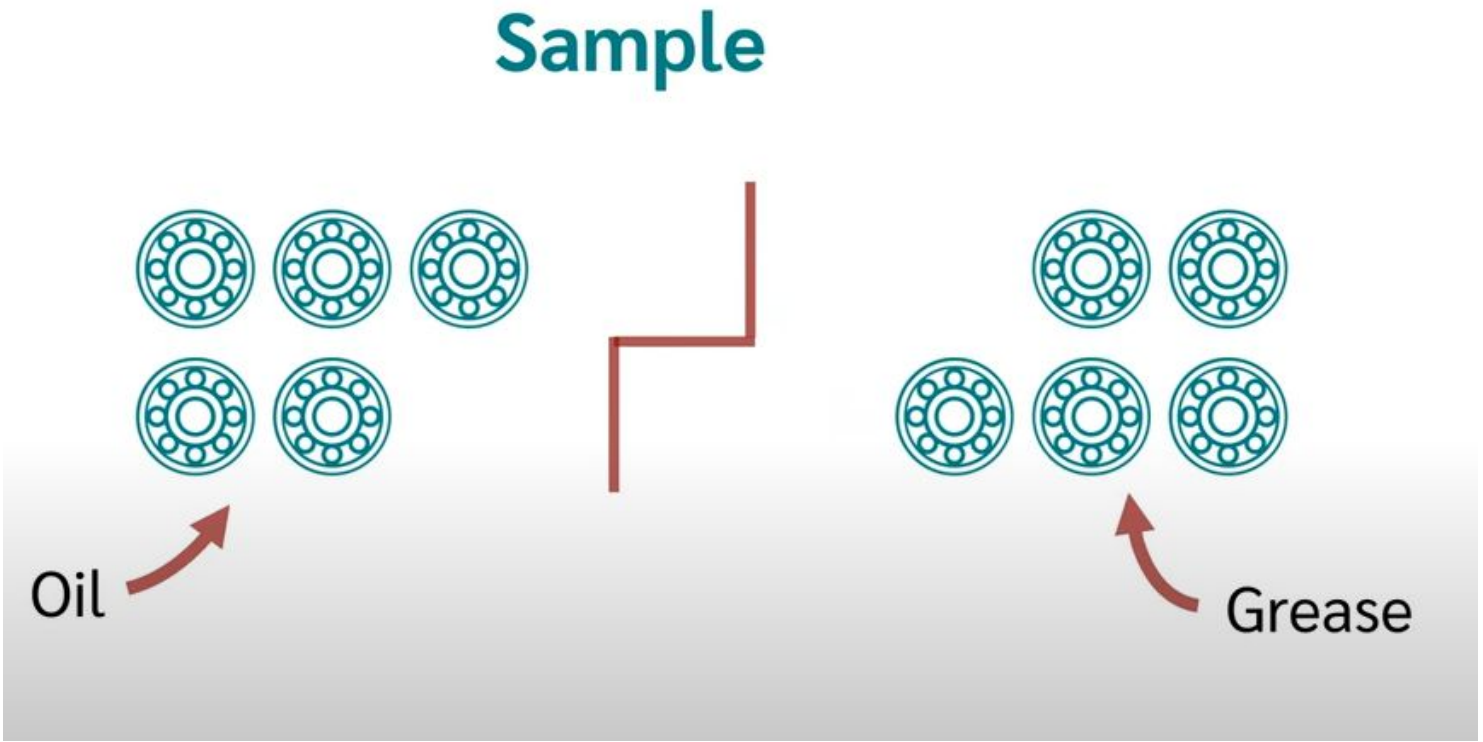
ABCD, ABCE, ACDE, ...

ABCDE

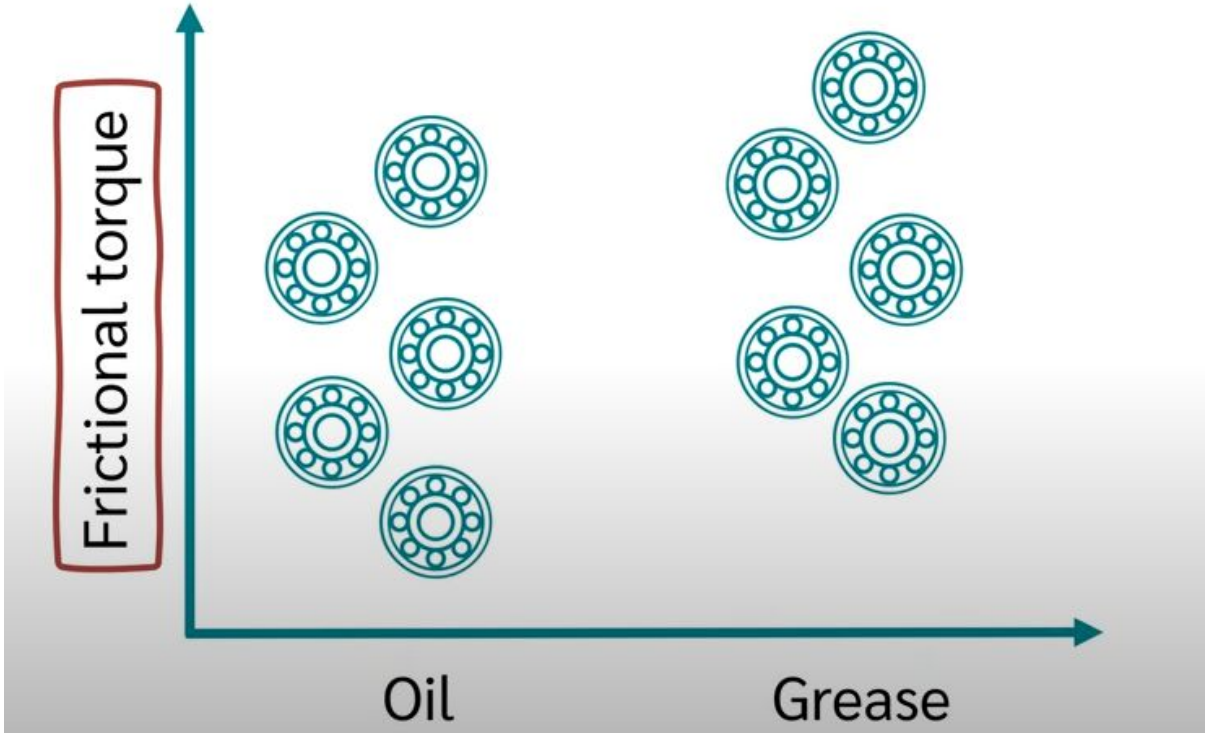
Example of Bearing



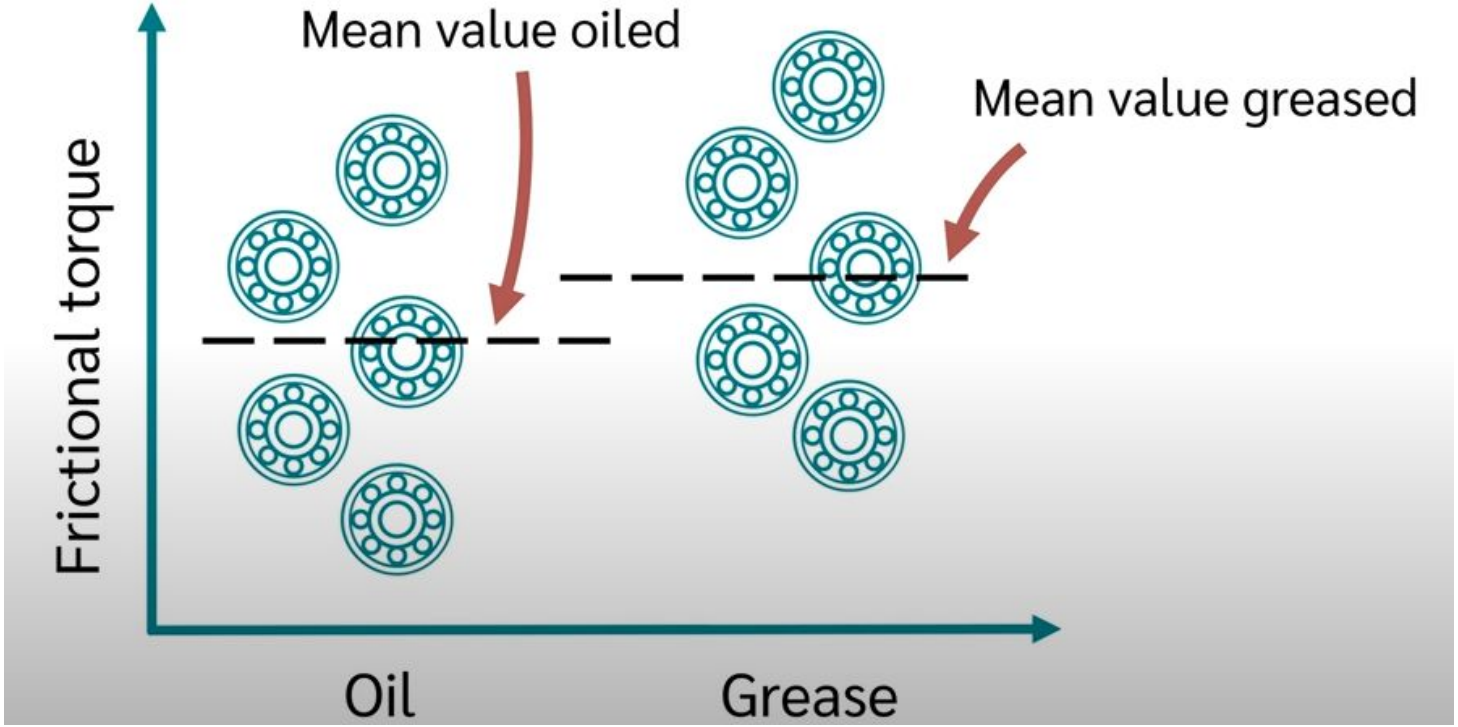
2 Sets of Experiments for 2 Levels



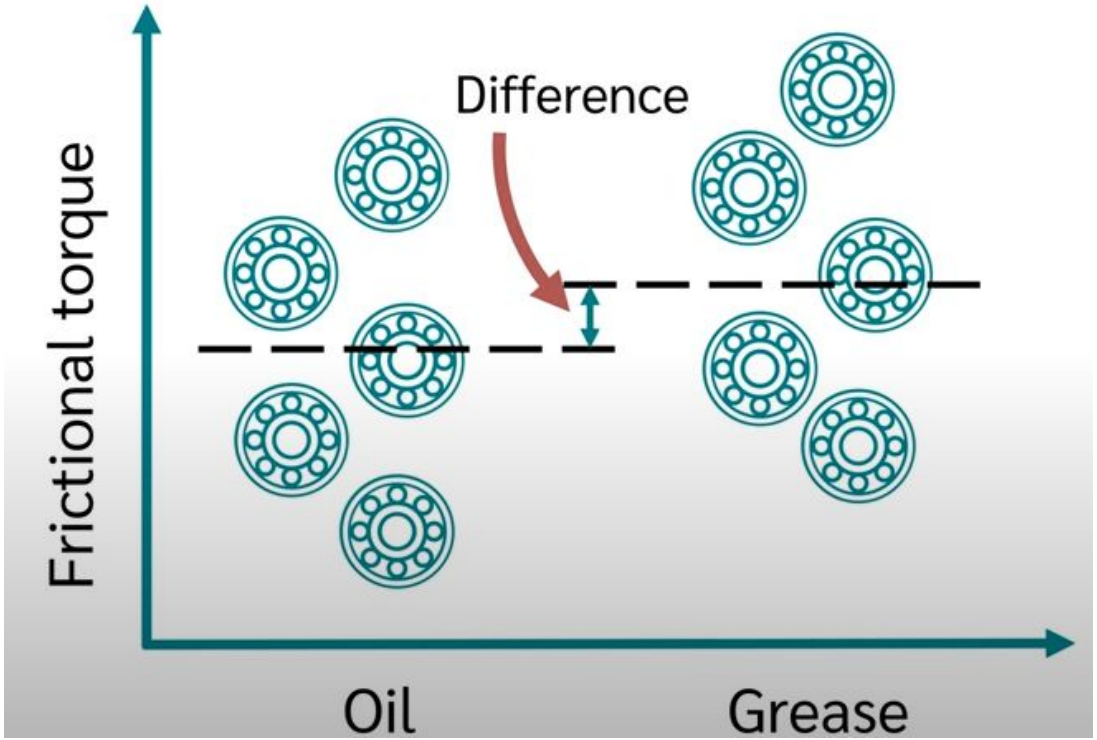
Pilot Runs



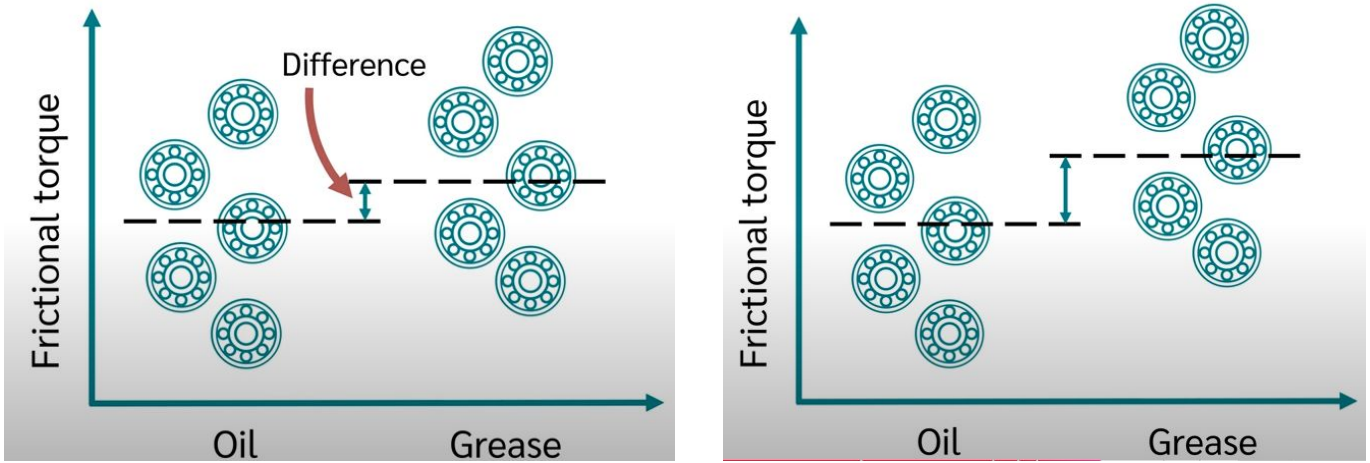
Collect Measurements



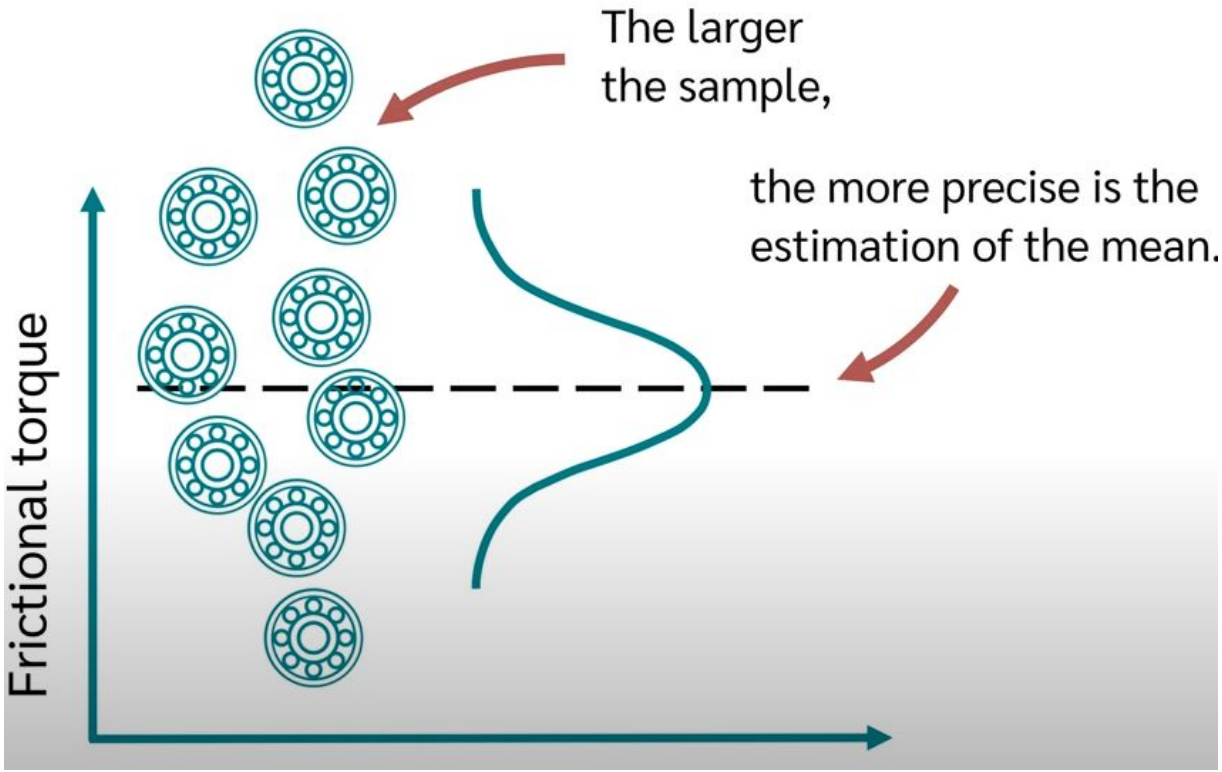
Difference in Effect



Different Effect for Two Different Samples



Need More Experiments



Full Pipeline

Inputs, Outputs and The Process

Suppliers	Inputs	Process	Outputs	Customers
Providers of Resources	*Resources Required for Process*	*How we Transform inputs into outputs*	*Deliverables of Process*	*Receivers of Process Outputs*

Example: Survey Reveals that Cake is not Tasty

Suppliers	Inputs	Process	Outputs	Customers
Providers of Resources	*Resources Required for Process*	*How we Transform inputs into outputs*	*Deliverables of Process*	***Receivers** of Process Outputs*
				Bakery Customers Party Guests

List the Outputs

Suppliers	Inputs	Process	Outputs	Customers
Providers of Resources	*Resources Required for Process*	*How we Transform inputs into outputs*	***Deliverables** of Process*	*Receivers of Process Outputs*
			Cake **flavor** Cake **texture** Cake **color** Cake **size** Cake **shape** **Yield,** **Lead Time,** **Product Cost**	Bakery Customers Party Guests

List the Process Steps

Suppliers	Inputs	Process	Outputs	Customers
Providers of Resources	*Resources Required for Process*	*How we **Transform** inputs into outputs*	*Deliverables of Process*	*Receivers of Process Outputs*
		1. Preheat Oven 2. **Mix Dry Ingredients** 3. **Mix Wet Ingredients** 4. Combine Wet and Dry 5. Pour into Pan 6. **Bake the cake** 7. Remove and cool 8. Decorate 9. Package and Ship	Cake **flavor** Cake **texture** Cake **color** Cake **size** Cake **shape** **Yield,** **Lead Time,** **Product Cost**	Bakery Customers Party Guests

List the Inputs

Suppliers	Inputs	Process	Outputs	Customers
Providers of Resources	***Resources Required** for Process*	*How we Transform inputs into outputs*	*Deliverables of Process*	*Receivers of Process Outputs*
	Cake Ingredients **Mixing Equipment** **Mixing Bowls** **Cake Pan** **The Oven** **Baking Temperature** **Baking Time** **Measuring Cup** **The Bakers**	1. Preheat Oven 2. **Mix Dry Ingredients** 3. **Mix Wet Ingredients** 4. Combine Wet and Dry 5. Pour into Pan 6. **Bake the cake** 7. Remove and cool 8. Decorate 9. Package and Ship	Cake **flavor** Cake **texture** Cake **color** Cake **size** Cake **shape** **Yield,** **Lead Time,** **Product Cost**	Bakery Customers Party Guests

Additional Variables

Suppliers	Inputs	Process	Outputs	Customers
***Providers** of Resources*	*Resources Required for Process*	*How we Transform inputs into outputs*	*Deliverables of Process*	*Receivers of Process Outputs*
Ingredient Vendors Grocery Store Home Goods Store (Equipment)	**Cake Ingredients** **Mixing Equipment** **Mixing Bowls** **Cake Pan** **The Oven** **Baking Temperature** **Baking Time** **Measuring Cup** **The Bakers**	1. Preheat Oven 2. **Mix Dry Ingredients** 3. **Mix Wet Ingredients** 4. Combine Wet and Dry 5. Pour into Pan 6. **Bake the cake** 7. Remove and cool 8. Decorate 9. Package and Ship	Cake **flavor** Cake **texture** Cake **color** Cake **size** Cake **shape** **Yield,** **Lead Time,** **Product Cost**	Bakery Customers Party Guests

Narrow the Study

Suppliers	Inputs	Process	Outputs	Customers
Providers of Resources	*Resources Required for Process*	*How we Transform inputs into outputs*	*Deliverables of Process*	*Receivers of Process Outputs*
Ingredient Vendors Grocery Store Home Goods Store (Equipment)	**Cake Ingredients** **Mixing Equipment** **Mixing Bowls** **Cake Pan** **The Oven** **Baking Temperature** **Baking Time** **Measuring Cup** **The Bakers**	1. Preheat Oven 2. **Mix Dry Ingredients** 3. **Mix Wet Ingredients** 4. Combine Wet and Dry 5. Pour into Pan 6. **Bake the cake** 7. Remove and cool 8. Decorate 9. Package and Ship	Cake **flavor** Cake **texture** Cake **color** Cake **size** Cake **shape** **Yield,** **Lead Time,** **Product Cost**	Bakery Customers Party Guests

Factors, Controllable Variables, Independent Variables

Response Variables, Dependent Variables

Select some Variables

What combination of **input factors** (**Temperature, Time** and **Sugar**) produces the **output (best tasting cake)**?

Consider the Levels

What combination of **input factors** (**Temperature, Time** and **Sugar**) produces the **output (best tasting cake)**?

Factor	High Level (+)	Low Level (-)
Temperature		
Time		
Sugar		

A level refers to a specific setting (Set point) of a **factor (Input)**.

Design the Experiment

What combination of **input factors** (**Temperature, Time** and **Sugar**) produces the **output (best tasting cake)**?

Factor	High Level (+)	Low Level (-)
Temperature	400° F	350° F
Time	20 Min.	15 Min.
Sugar	3 Cups	2.5 Cups

A level refers to a specific setting (Set point) of a **factor (Input)**.

Treatments

What combination of **input factors** (**Temperature, Time** and **Sugar**) produces the **output (best tasting cake)**?

Factor	High Level (+)	Low Level (-)
Temperature	400° F	350° F
Time	20 Min.	15 Min.
Sugar	3 Cups	2.5 Cups

	Factors		
	Temperature	Time	Sugar
Treatment 1	+	+	+
Treatment 2	+	+	-

A treatment is a unique combination of factors and levels within an experiment.

List all Treatments

(Treatments)

A treatment is a unique combination of factors and levels within an experiment.

Factor	High Level (+)	Low Level (-)
Temperature	400° F	350° F
Time	20 Min.	15 Min.
Sugar	3 Cups	2.5 Cups

We're looking for the **best tasting cake** across all these unique combinations!

	Factors		
	Temperature	Time	Sugar
Treatment 1	+	+	+
Treatment 2	+	+	-
Treatment 3	+	-	+
Treatment 4	+	-	-
Treatment 5	-	+	+
Treatment 6	-	+	-
Treatment 7	-	-	+
Treatment 8	-	-	-

Block and Randomize

Systematic Error is the **variation** in your **response variable** that has **a special cause** outside of the process being studied.

Random error is the **variation** caused by controllable or uncontrollable factors (noise) or random variation in the response variable.

Block what you can.

randomize what you can't.

Inputs
Cake Ingredients
Mixer
Mixing Bowls
Cake Pan
The Oven
Baking Temperature
Baking Time
Measuring Cup

Eliminate by Blocking

Blocking is a method to **minimize** the **variation (**random error**)** in a **DOE**, where you execute your experiment at a **single setting (level) of that factor (Block)**.

Blocking means that we're going to:

- Use the same person to bake the cakes

By "Blocking" this factor out, we are eliminating any person-to-person variation!

We're reducing the error in our experiment!

Block and Select

Blocking is a method to **minimize** the **variation** (*random error*) in a **DOE**, where you execute your experiment at a **single setting (level) of that factor (Block)**.

Suppliers	Inputs	Process	Outputs	Customers
Providers of Resources	*Resources Required for Process*	*How we Transform inputs into outputs*	*Deliverables of Process*	*Receivers of Process Outputs*
Ingredient Vendors Grocery Store Home Goods Store (Equipment)	**Cake Ingredients** **Mixing Equipment** **Mixing Bowls** **Cake Pan** **The Oven** **Baking Temperature** **Baking Time** **Measuring Cup** **The Bakers**	1. Preheat Oven 2. **Mix Dry Ingredients** 3. **Mix Wet Ingredients** 4. Combine Wet and Dry 5. Pour into Pan 6. **Bake the cake** 7. Remove and cool 8. Decorate	Cake **flavor** Cake **texture** Cake **color** Cake **size** Cake **shape** **Yield,** **Lead Time,** **Product Cost**	Bakery Customers Party Guests

Replications

Replication is the act of performing an **experiment all over again.**

Replicates

Replication provides **degrees of freedom** for **ANOVA Analysis.**

Replications!

87 Octane	89 Octane	91 Octane	93 Octane
168.8	178.0	179.9	185.3
171.5	178.0	183.4	189.7
172.6	173.6	180.1	187.1
168.1	173.7	181.8	187.7
167.1	172.8	179.4	186.2
170.9	176.4	183.9	188.3
168.3	172.9	184.2	186.6
169.2	178.5	184.0	191.2
168.4	176.8	183.3	187.4
171.8	172.3	182.5	185.4
169.7	175.3	182.2	187.5

Variation **WITHIN** a Treatment (MS_E)

Entire Process

7 Steps to Plan and Organize a DOE

1. Determining Experimental Objectives (Problem Statement)
2. Identifying Responses (Outputs) and the Measurement Systems
3. Identifying Input Factors to Study along with Levels (and those to Block)
4. Choosing the Right Design
   - Replication, Sample size, Randomization
   - Full Factorial or Fractional Factorial
5. Running the Experiment
6. Analyzing the Results - ANOVA, Main Effects, Interaction Effects
7. Conclusions, and Iterations (Rinse & Repeat) – Iterative Approach

Quiz 9

1. Write 2 things that you liked/understood today.
2. Write 2 things that you did not like/understand today.

3. Describe the Design of Experiment for your case study, or take any simple case study and show the steps.