# Chapter 5
# Data : The Foundation of your Data Story

The data story that you seek to tell are as strong as the data foundation on which they're built.

- **RELEVANCE :**

Is the data even relevant to the problem you are trying to solve?

The Data must be applicable to the situation or problem being attempted to analyze and understand.

- More than often, we will be working with less than ideal datasets

- **Trustworthiness :**

Are the insights coming from data that is accurate & trustworthy?

- Trust worthy data is correct or valid, free from significant defect & gaps

DATE / 20
PAGE No

- Trustworthiness of your data begins with proper collection, processing and maintenance of data.

⇒ Every Data Story needs a central insight.

- When analyzing data, it can often feel like an endless maze with multiple potential paths to follow.
- The right question can provide direction and purpose.

Key points to be considered for finding a central insight :

- Problem (Current state)
- Outcome (Future state)
- Actions (Approach)
- Measure (progress)

⇒ Do you have an actionable insight?

- Once we have found an insight, next thing to ask is whether its actionable or not?
- It should pass 'So what' test

- We should have the answers of the following question regarding our 'insight' :-
1. Why should your audience care?
2. What should they do about it?
3. What's the potential business impact?

Breaking down the above questions →

1. Why should your audience care?
It should be :-
   i) Valuable
   ii) Relevant

2. What should they do about it?
Action should be :-
   i) Practical
   ii) Specific

3. What's the potential business impact?
The potential impact should be -
   i) Concrete
   ii) Contextualized.

## SUMMARY SO FAR :-

- We started out with data, the foundations, which should be relevant to the problem and should also be trustable.

- Once we have the data, insights, we follow the 4-D rule, so that we don't get lost in the maze.

- Once we have insights, we go for the 'so what' test, and check the 6 conditions to make sure that our insight is actionable.

⇒ The Analysis process : Exploration to Explanation.

| | Exploratory | Explanatory |
|---|---|---|
| Goal | Understand | Communicate |
| Audience | You | Other people |
| Data familiarity | very familiar (you) | Less familiar (others) |
| Visualization focus | Flexibility & speed | Simplicity, clarity & cohesion |
| Narrative | Unknown | Known |
| Outcome | Insight | Action |

⇒ The three kinds of Data Forging

| | Data | Narrative | Visuals. |
|---|---|---|---|
| Data Cut | ✓ | X | X |
| Data Cameo | X | ✓ | X |
| Data decoration | X | X | ✓ |

⇒ Analyzing and communicating Data Demands discipline :-

⇒ How cognitive bias can distort your data story ? (by system - 1)

i) Confirmationd bias

- To search for and accept evidence that support your existing beliefs or views ; and to ignore information against your existing opinions.

- If you succumb to your own confirmation bias, you find what you expected to see in numbers - and miss what's really happening.

ii) Survivorship Bias :

To focus only on what succeeded or survived while ignoring what failed or did not survive.

iii) The Curse of Knowledge :

To assume other people have the necessary knowledge to follow what you're communicating

Knowing your audience, being self aware of your own proficiencies and using data stories can help to contract this.

⇒ How fallacies can erode data foundation → (by a system 2)

i) Correlation Fallacy :

Correlation : Different variables are fluctuating together in a similar or inverse manner.

Causation : Change in one variation causes change in other.

Correlation does not imply causation.

ii) Texas Sharpshooter Fallacy:

When someone assigns significance to random set of coincidences

iii) Hasty Generalization Fallacy:

When someone makes a broad claim that isn't justified by sufficient or unbiased data.

⇒ When Too much of a good thing is Bad.

Even if we had done all the previously mentioned skills correctly, there is still a chance that we do one mistake. one more thing to be careful of - information overload

- Don't overwhelm your audience.

To Avoid Information Overload:
- Don't use irrelevant data or chart.
- Avoid dense text in slides.
- Layout your content that is easy to follow