Descriptive vs Diagnostic Analytics & Correlation vs Causation

☑️ Descriptive vs Diagnostic Analytics

📊 Descriptive Analytics — "What happened?"

Descriptive analytics summarizes historical data to show what has already happened in the business. It focuses on reporting and aggregation, not explanation.

It answers questions like:

- What were total sales last month?
- How many customers purchased?
- Which product sold the most?
- What was the average order value?

Typical methods include:

- Reports
- Dashboards
- Totals and averages
- Trend charts
- Grouped summaries

Real Business Example

An e-commerce company dashboard shows:

- March revenue = ₹48 lakh
- Orders = 12,400
- Top category = Electronics

This is descriptive — it tells the result, not the reason.

Deliverable: Clear summary of past performance.

🔍 Diagnostic Analytics — "Why did it happen?"

Diagnostic analytics goes one step deeper. It investigates causes and contributing factors behind the results seen in descriptive analysis.

It answers questions like:

- Why did sales drop in March?
- Why is one region underperforming?
- Why did customer churn increase?
- Why did conversion rate fall?

Typical methods include:

- Drill-down analysis
- Segment comparison
- Correlation checks
- Before–after comparisons
- Root cause analysis

Real Business Example

After seeing March sales dropped, analyst checks:

- Region-wise sales → drop only in North region
- Delivery time → increased by 2 days
- Complaint tickets → increased

Conclusion:Sales dropped because delivery delays increased in that region.

This is diagnostic — it explains the cause.

Deliverable: Root-cause explanation supported by data.

🎯 Key Difference [Practical View]

Descriptive = reports the outcomeDiagnostic = investigates the reason

Descriptive is monitoring.Diagnostic is investigation.

Both are required in real companies — descriptive first, diagnostic next.

☑ Correlation vs Causation

This is one of the most misunderstood concepts in data analysis.

🔗 Correlation — "Two variables move together"

Correlation means two variables show a statistical relationship — when one changes, the other also tends to change.
It does NOT prove that one causes the other.

Example
Data shows:Ice cream sales ↑ Cold drink sales ↑

They are correlated — both rise in summer.
But:Ice cream sales do NOT cause cold drink sales.

A third factor (temperature) drives both.
Correlation = relationship patternNot proof of cause.

⚠ Causation — "One variable directly causes change in another"

Causation means change in one variable directly produces change in another — proven through controlled testing or strong evidence.

Example
Company runs A/B test:

Group A → shows discount bannerGroup B → no banner

Group A conversion = 5%Group B conversion = 3%

Because experiment controlled other factors, we can say:Discount banner caused higher conversion.
That is causation.

🖊 How Businesses Prove Causation

Real companies use:
- A/B testing
- Controlled experiments
- Time-based intervention analysis
- Randomized trials

Without controlled testing, analysts should say:"Associated with" — not "caused by."

◎ Practical Warning for Analysts

Wrong statement:"Higher ad spend caused higher sales."

Correct statement:"Higher ad spend is correlated with higher sales. Further testing is needed to confirm causation."

This distinction is important in professional reporting.