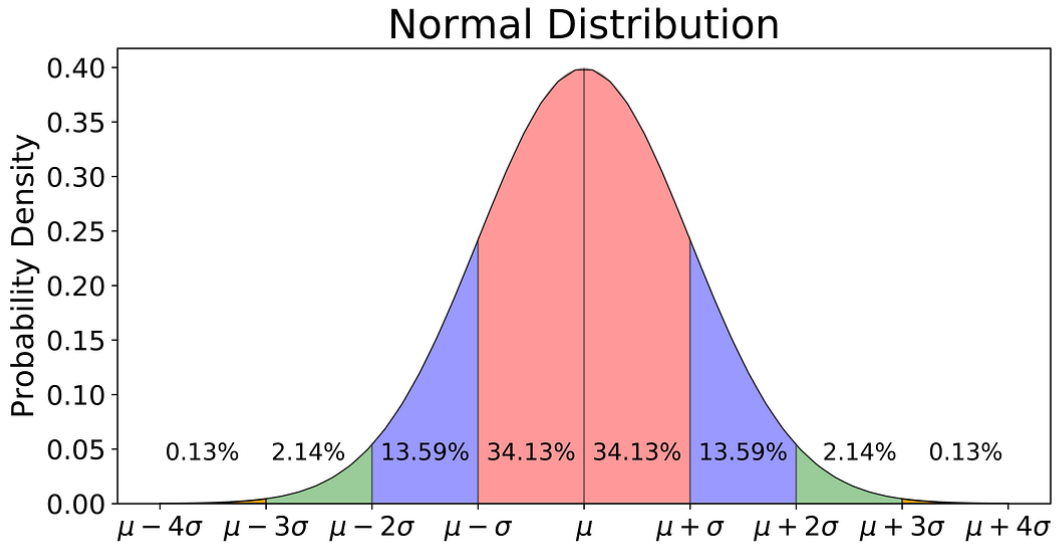
What is Normal Distribution?

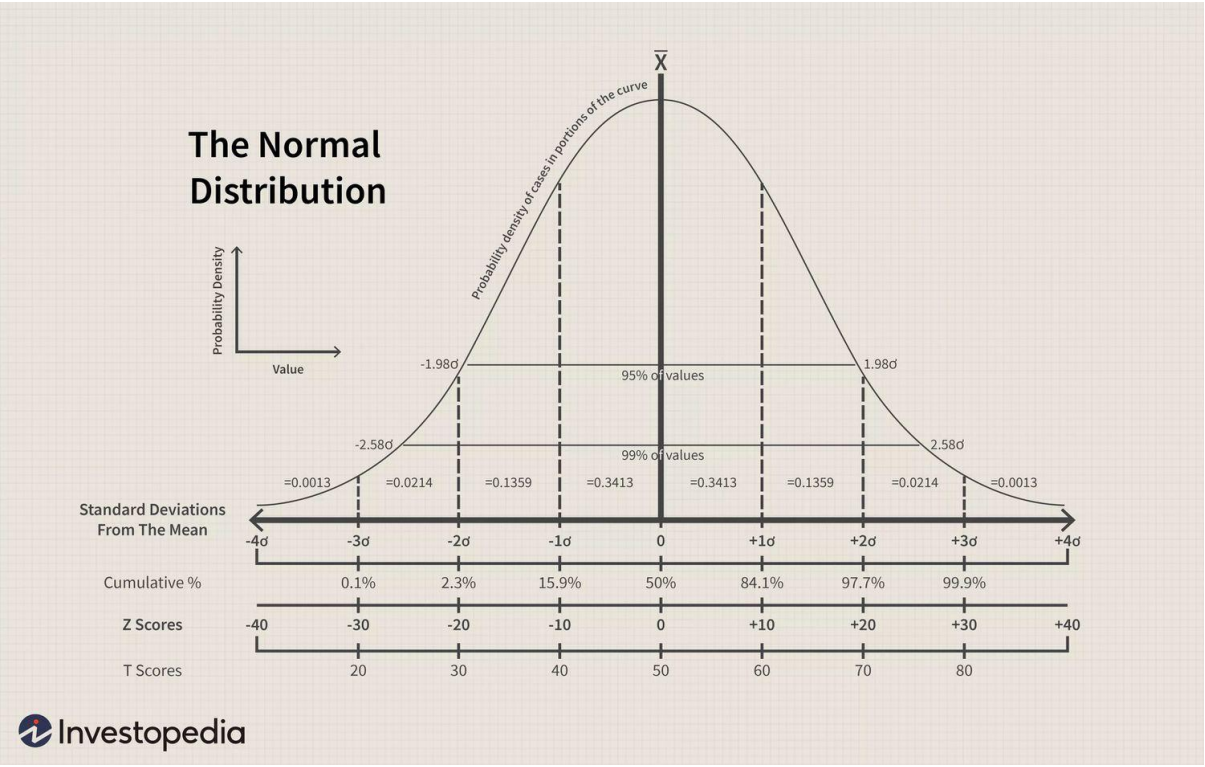
📘 Professional Definition

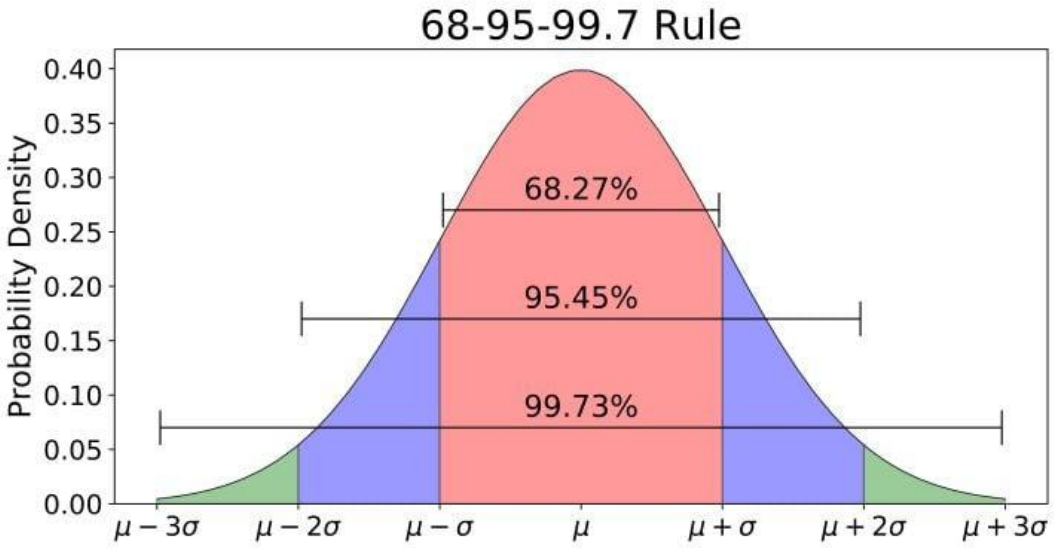
Normal Distribution (also called Gaussian Distribution) is a continuous probability distribution that is symmetric around its mean, where most observations cluster around the central value and probabilities taper off equally toward both extremes.

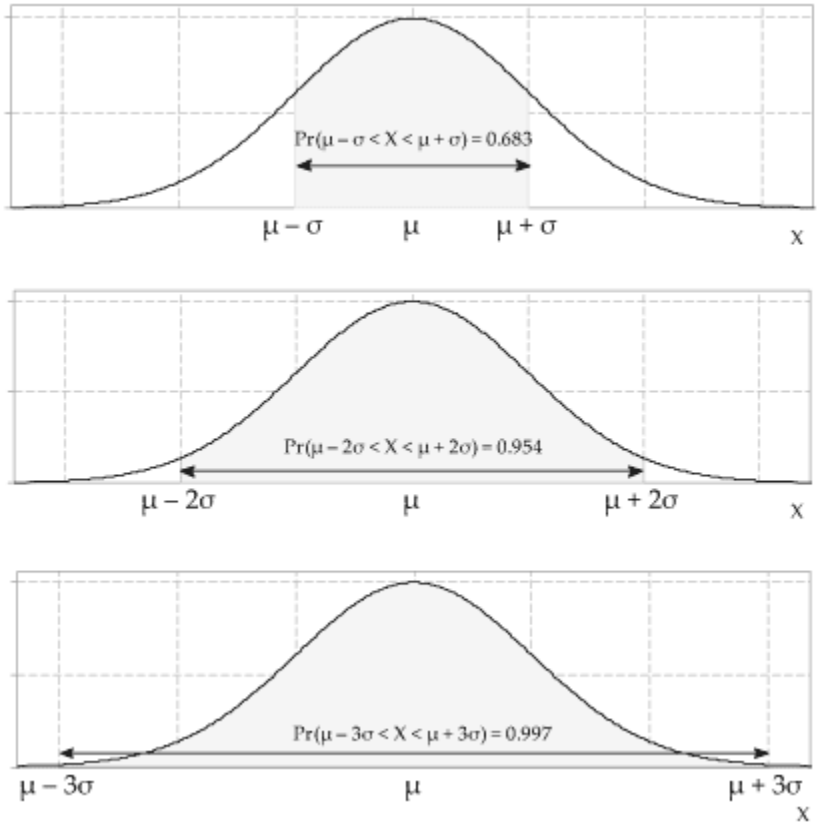
📗 Simple Explanation

It's a **bell-shaped curve** showing that most data points are near the average (mean), and fewer are far from it. It is the most common distribution in statistics used to model real-world phenomena.




The Normal Distribution
Probability Density
Value
X̄
Probability density of cases in portions of the curve
-1.98σ
1.98σ
95% of values
-2.58σ
2.58σ
99% of values
=0.0013
=0.0214
=0.1359
=0.3413
=0.3413
=0.1359
=0.0214
=0.0013
Standard Deviations From The Mean
-4σ
-3σ
-2σ
-1σ
0
+1σ
+2σ
+3σ
+4σ
Cumulative %
0.1%
2.3%
15.9%
50%
84.1%
97.7%
99.9%
Z Scores
-40
-30
-20
-10
0
+10
+20
+30
+40
T Scores
20
30
40
50
60
70
80
Investopedia


68-95-99.7 Rule
Probability Density
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
68.27%
95.45%
99.73%
μ − 3σ
μ − 2σ
μ − σ
μ
μ + σ
μ + 2σ
μ + 3σ



4

📊 2️⃣ Shape of the Curve (Bell Curve)

- 🟢 **Symmetric** around the mean
- 🟢 **Bell-shaped** with one peak
- 🟢 mean = median = mode (all at the center)
- 🟢 Both sides mirror each other
- 🟢 Tails never quite touch the x-axis (they go to infinity) — this means extreme values are possible but rare.

🧮 3️⃣ Formula (Probability Density Function)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Where:
• x= random variable
• μ= mean (center)

• σ= standard deviation (spread)
• e≈ 2.718 (base of natural log)
• π≈ 3.14159

📌 This formula gives the height of the curve at any point x. The full area under the curve = 1 (100% probability).

📐 4️⃣ **Role of Mean and Standard Deviation**

- **Mean (μ)** — determines where the bell curve is located (center).
- **Standard Deviation (σ)** — determines how wide or narrow the curve is:
    - 🔹 Small σ → taller and narrower curve
    - 🔹 Large σ → shorter and wider curve

📏 5️⃣ **Empirical Rule (68 – 95 – 99.7 Rule)**

In a normal distribution:

Range Percentage of Data

μ ± 1σ ~68%

μ ± 2σ ~95%

μ ± 3σ ~99.7%

So, most values are close to the mean, and very few are far away.

🧠 6️⃣ **Why Normal Distribution is Important?**

🎯 **Real-Life Examples**

- 📈 Heights of people
- 🤓 IQ scores
- 🩺 Blood pressure measurements
- 📊 Exam marks in large groups
  All of these often look like a bell-shaped distribution in a histogram.

💡 **Data Science & Machine Learning Uses**

1. **Z-scores & Standardization**
   To compare scores from different scales by converting them into how many standard deviations away from the mean they are.

2. **Feature assumptions in ML models**
   Many algorithms assume normality (e.g., linear regression residuals).
3. **Anomaly detection**
   If a value lies far from the mean (e.g., beyond 3σ), it's considered unusual or an *outlier*.
4. **Statistical testing & confidence intervals**
   Many inferential methods rely on normal distribution properties.

📊 7️⃣ Standard Normal Distribution

When:

- Mean = 0
- Standard deviation = 1

It's called **standard normal distribution**, and tables (Z-tables) are used to find probabilities.

📈 8️⃣ Practical Intuition

Visualize this:

Most students score around average marks → tall center of the curve.
Very few score extremely low or extremely high → short tails on both sides.

This symmetry and predictable structure help in making probability-based decisions.

🧪 9️⃣ Real Data Science Example

Suppose a model's error distribution is roughly normal:

- Most errors are small (near zero).
- Few large mistakes (tails).
- You can quantify *confidence intervals* using ±1σ, ±2σ ranges.

This is why normal distribution is foundational in statistical modeling and ML validation.

🧠 1️⃣0️⃣ Key Properties (Summarized)

✓ Bell-shaped and symmetric
✓ Mean = Median = Mode
✓ Defined by mean (μ) and standard deviation (σ)
✓ Area under curve = 1
✓ Empirical rule (68–95–99.7) holds
✓ Tails are asymptotic (never touch the axis)

CENTRAL LIMIT THEOREM (CLT)

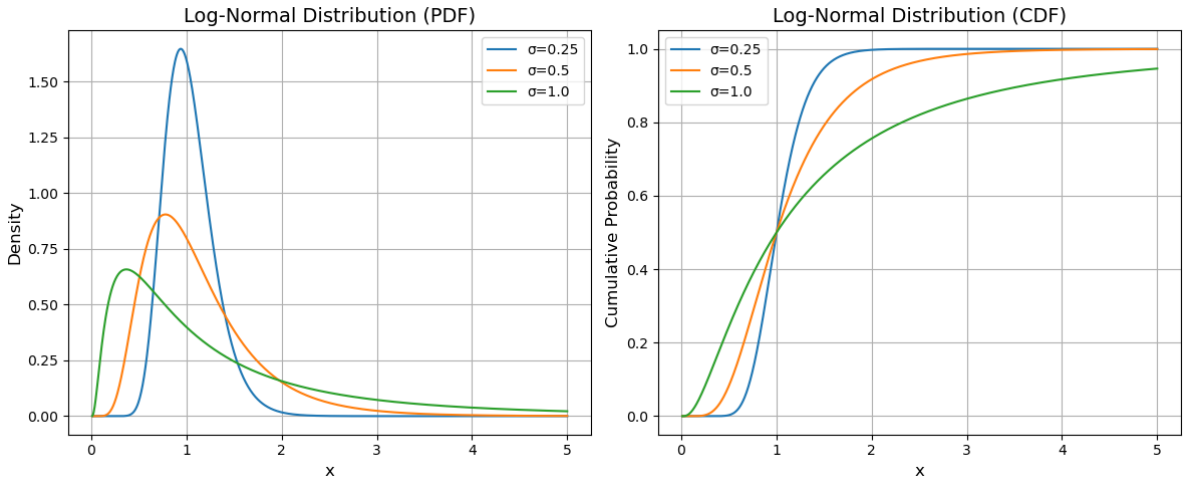
📌 What is CLT?

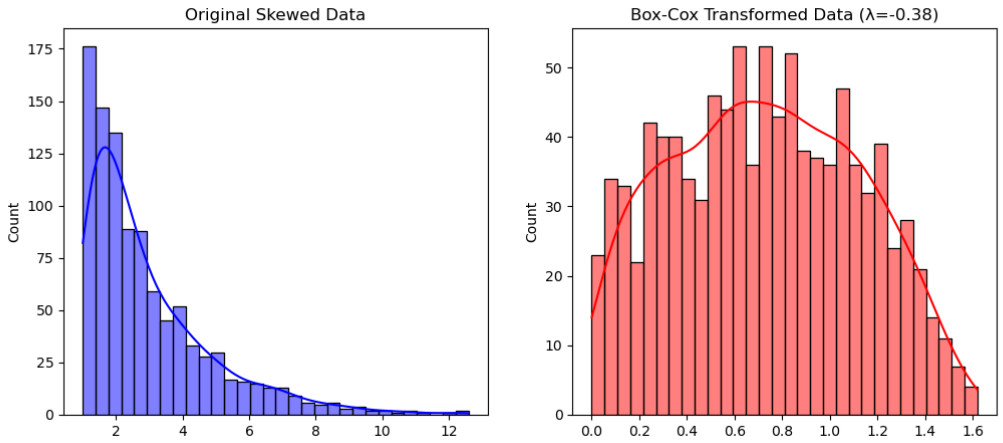
Central Limit Theorem states that if you take large samples from any distribution and calculate their means, the distribution of those means will tend to be Normal (Bell-shaped), even if the original data is not normal.

- Sample size usually ≥ 30.
- This is how Normal distribution becomes powerful in real data analysis.

Practical Example:
Take 50 samples repeatedly of any random data (e.g., uniform, skewed), and the histogram of sample means becomes bell-shaped as sample size increases.

📊 LOG-NORMAL DISTRIBUTION





📌 Professional Definition

A **Log-Normal Distribution** is a distribution where the **logarithm of the variable is normally distributed** — meaning the original data itself is not symmetrical but its log is.

📌 Simple Explanation

If X is log-normal:

$$Y = \ln(X)$$

follows Normal distribution.
Hence X is log-normal.

📌 Shape

✓ Not symmetric
✓ Right-skewed (long tail to the right)
✓ Only positive values

📌 Real-Life Examples

✨ Stock prices (income, prices often log-normal)
✨ Growth processes (multiplicative factors)

👉 In Data Science: used in modeling positive skewed metrics.

📊 POWER LAW DISTRIBUTION

📌 Professional Definition

A **Power Law Distribution** is where one quantity varies as a power of another:

$$P(x) \propto x^{-\alpha}$$

This means **few values are very large and many values are very small.**

📌 **Simple Explanation**

Long tail on one side → **heavy-tail distribution**

📌 **Real-Life Examples**

✓ Frequency of words (few words extremely common)
✓ City population sizes
✓ Income/wealth distribution

Often used in network modelling and social media analysis.

📊 **PARETO DISTRIBUTION**

0image11"]}

📌 **Professional Definition**

Pareto distribution is a type of **power law distribution** that describes situations where **a small number of items contribute disproportionately to an effect.**

📌 **80–20 Rule (Pareto Principle)**

✓ 80% results from 20% causes
✓ 20% customers make 80% revenue

Named after economist Vilfredo Pareto.

📌 **Real-Life Importance**

📌 Business: customer value distribution
📌 Economics: wealth distribution
📌 DS: heavy-tail behavior identification

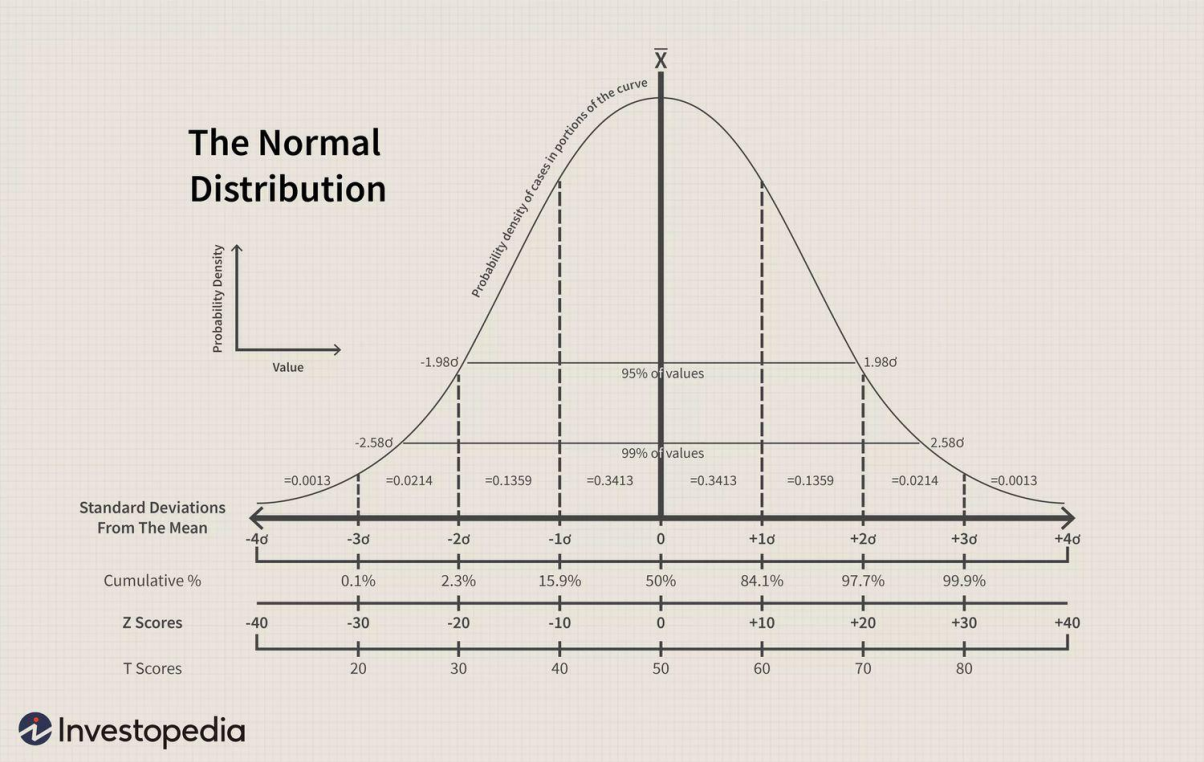
📌 **Relation to Power Law**

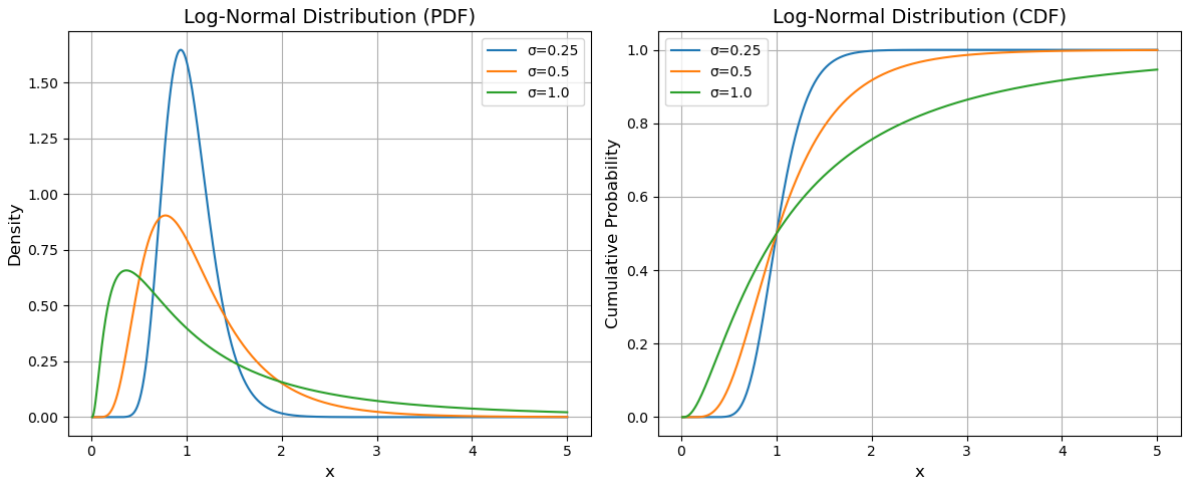
Pareto is basically a power law applied to real world phenomena with minimum value constraint.

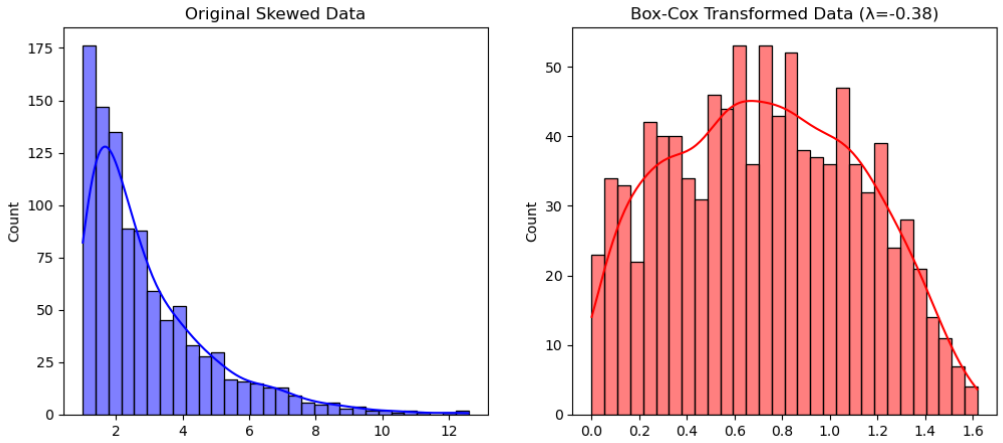
📊 HOW THESE DISTRIBUTIONS RELATE

Distribution	Shape	Use case
Normal	Bell symmetric	Exam scores, heights
Log-Normal	Right skewed	Prices, incomes
Power Law	Heavy tail	City size, social network
Pareto	80–20 type	Business, wealth

📊 VISUAL SUMMARY







4

- First image shows **normal distribution** and its symmetry.
- Second image shows **log-normal curves** with different spreads.
- Third shows skewed data shape like log-normal heavy-tail.
- Fourth shows a **power law heavy-tail** compared to normal.

📊 DATA SCIENCE CONNECTION

💡 Normal Distribution — used in statistical inference and model assumptions.
💡 CLT — justifies many ML models' reliance on approximate normality.
💡 Log-Normal — used for positive skew modeling (ex: stock prices).
💡 Power Law/Pareto — used in anomaly detection, social network metrics & business analytics.