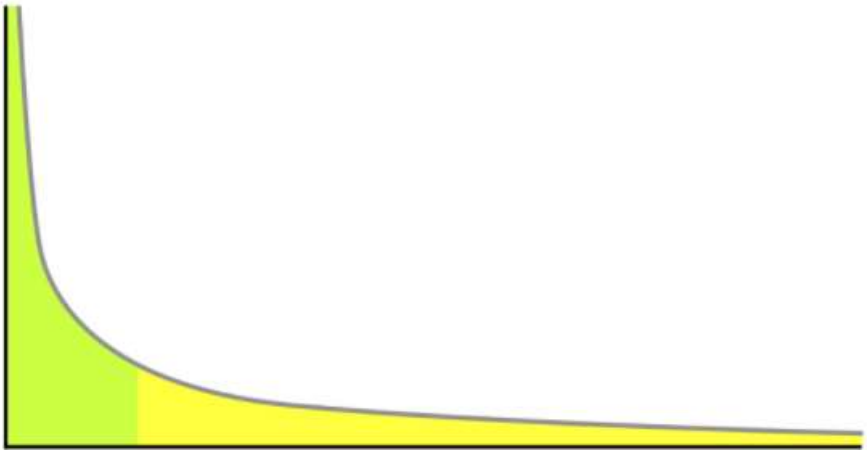
Power-Law Distribution

Power-Law Distribution

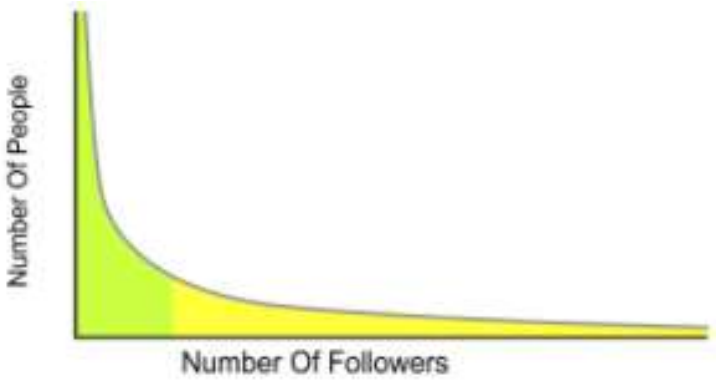
"Power-Law Distribution" is a term used when there are only a small number of observations which have a very high value of a certain characteristic, and a large number of observations which have a small value for that characteristic. Many unequally-distributed phenomena in the real world follow such a distribution.

The graph of a power-law distribution looks like the following image:



Let's consider some examples:

Example 1: The number of followers for each user on Instagram.

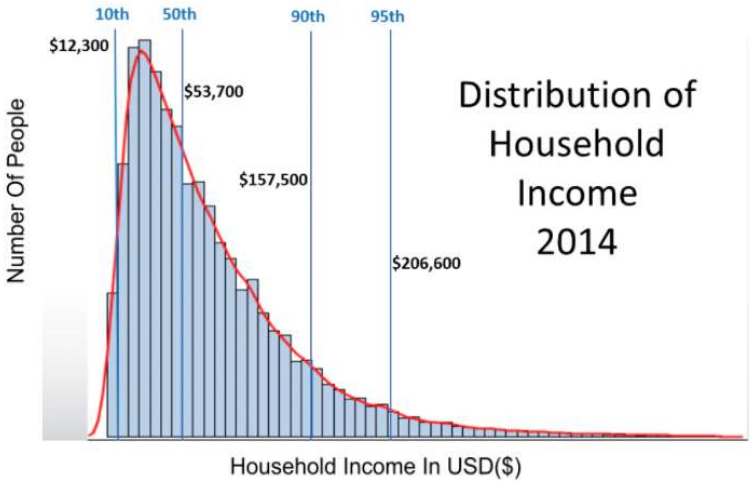


On Instagram, **there are a lot of people** (regular profiles) **with a small number of followers each**, and **only a few people percentage-wise** (verified handles, celebrities etc.) **who have an extremely high number of followers**.

Example 2: Distribution of US household income

As per a 2014 survey of US household incomes, the 90th percentile of income was $157,500. Hence 90% of US household incomes are less than $157,500. As we proceed further in the x-axis in the figure below, we observe that there are very few people who have extremely high household incomes.

The graph below depicts a frequency distribution of household incomes, where the x-axis represents "Household Income in USD($) in 2014" and the y-axis represents the "Number of people".



Mathematically, a power-law distribution is written as $P_k = c\left(k^{-\alpha}\right)$

Alternatively, we can rewrite this by taking the logarithm on both sides:

$$\log P_k = -\alpha\left(\log k\right) + c$$

Where:

- **c** is a constant
- **P**, **k** are the variables of interest - For example: In the Instagram power-law distribution, **k** would be the number of followers and **P** is the number of Instagram users.
- **α (alpha)** is the law's exponent.

In the graphs below, we see that as we increase the magnitude of the exponent term (alpha), the steepness/slope of the curve also increases.