Absolute frequency

- Absolute frequency is a statistical term describing the number of time a particular piece of data or a particular value appears during a trial or set of trials. Essentially, absolute frequency is a simple count of the number of times a value is observed. Absolute frequency is usually expressed as a whole number and is considered a very basic level of statistical analysis

Frequency : 1. Absolute frequency

- number of times that a value appears. It is represented as fi where the subscript represents each of the values.
- The sum of th data, represent as N $f_1 + f_2 + f_3 + \dots + f_n = N$ total number o

$$\sum_{i=1}^{n} f_2 = N$$

~~Exam~~

Example — 6-Sided cube dice

⚀	HH	5
⚁	IIII	4
⚂	IIII	5
⚃	IIII III	8
⚄	III	3

- The absolute frequency of the number 1 is 2
- " " " " " 2 is 4
- " " " " " 3 is 5
- " " " " " 4 is 8
- We add up the no. of throws: 5 + 4 + 5 + 8 + 3 + 5 = 30

1111 5 So we have N = 50

Code

```
import pandas as pd
df = pd.DataFrame(data = ['apple', 'apple',
        'banana', 'orange', 'apple', 'apple', 'banana',
        'banana', 'orange', 'banana', 'apple'],
        columns = ['Fruits'])
absolute_frequencies = df['Fruits'].value_counts
print(absolute_frequencies)
```

Output

Fruit	
apple	5
banana	4
orange	2

count

Name: Fruits, dtype: int64

Relative Frequency

- Relative frequency, for example, is derived from absolu absolute frequency. when the absdute fo of values is tracked over the entire trial, the absolute frequency for a particular value can then be divided by the total number of values for that variable throughout the trial to get the relative frequency.
- The relative frequency is what we most often reference, whether it is the winning percentage of our favorite sports team or the percentage of fund mangers that beat the market. unlike absolute frequency, relative frequency is usually expressed as a percentage or fraction rather than a whole number.
- Consider that p is the population proportion and is the sample proportion

Hence, the formula for the standard error of proportion is defined as follows

SE of a Proportion

The standard error for a sample proportion can be calculated by

$$SE = \sqrt{\frac{P(1-P)}{n}}$$

Note

Notice the sample size in the denominator ... as the sample size increases, the standard error decreases

(2) Relative frequency:

- The result of dividing the absolute frequency of a certain value by the total number of data.
- It is represented as n_i
- The sum of the relative frequencies is equal to 1.

$$n_i = \frac{f_a}{N}$$

Example - 6 - sided cube dice

Number	Absolute	Relative
1	5	5/30 = 0.1667
2	4	4/30 = 0.1333
3	5	5/30 = 0.1667
4	8	8/30 = 0.2667
5	3	3/30 = 0.1
6	5	5/30 = 0.1667

```
df = pd.DataFrame(data = ['apple', 'apple', 'banana',
                          'orange', 'apple', 'apple', 'banana', 'banana',
                          'orange', 'banana', 'apple'], columns = ['Fruit'])
relative_frequencies = df['Fruit'].value_counts(normalize=
                                                 True)
print(relative_frequencies)
```

Output

Fruit	
apple	0.454545
banana	0.363636
orange	0.181818

Name: ~~Fruit~~ proportion, dtype: float64