

Applied Data Analytics

Statistics — Dispersion & concentration

The Gini coefficient

Hans-Martin von Gaudecker and Aapo Stenhammar

Example

Ordered data

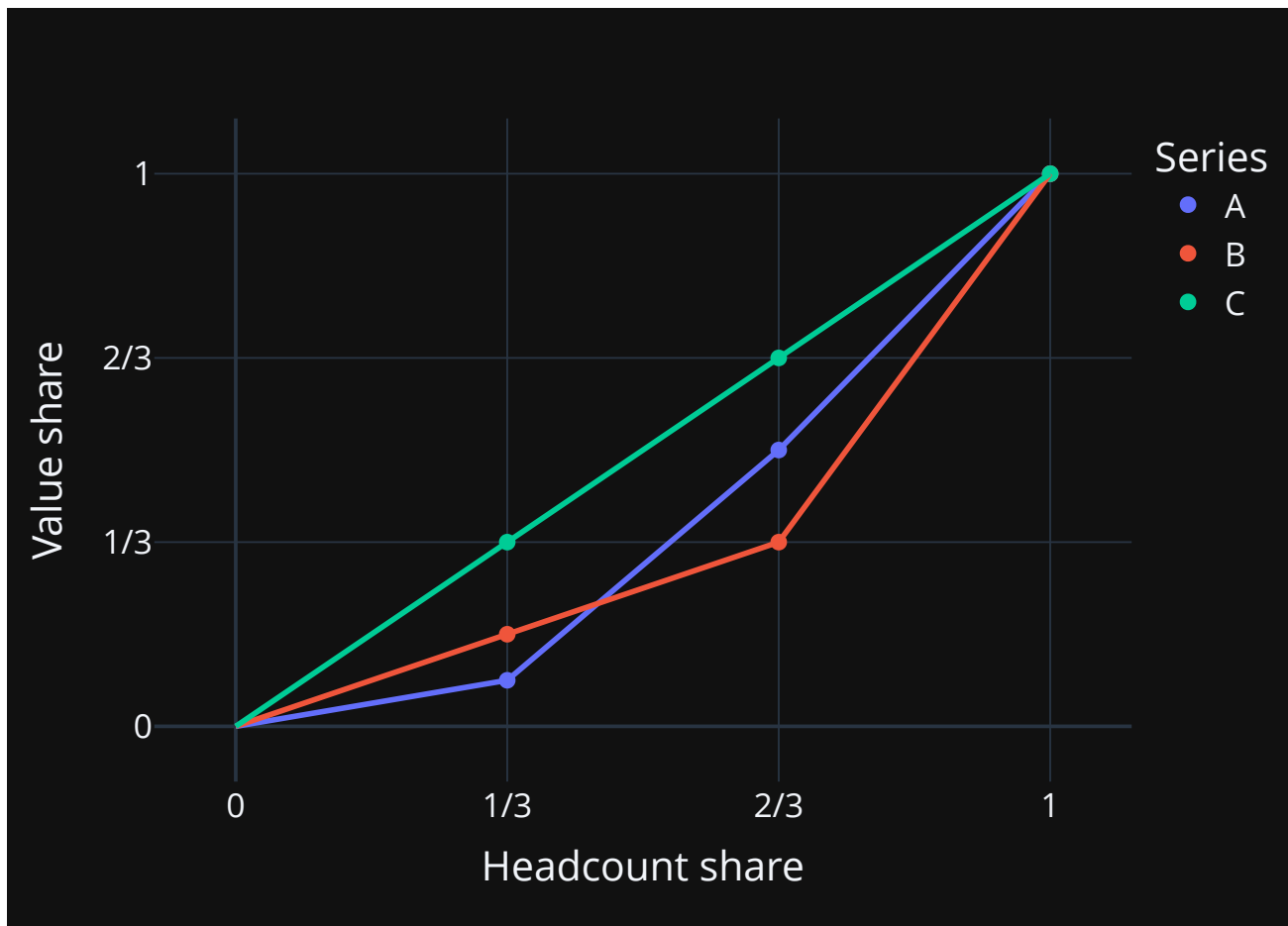
A	B	C
1	2	4
5	2	4
6	8	4

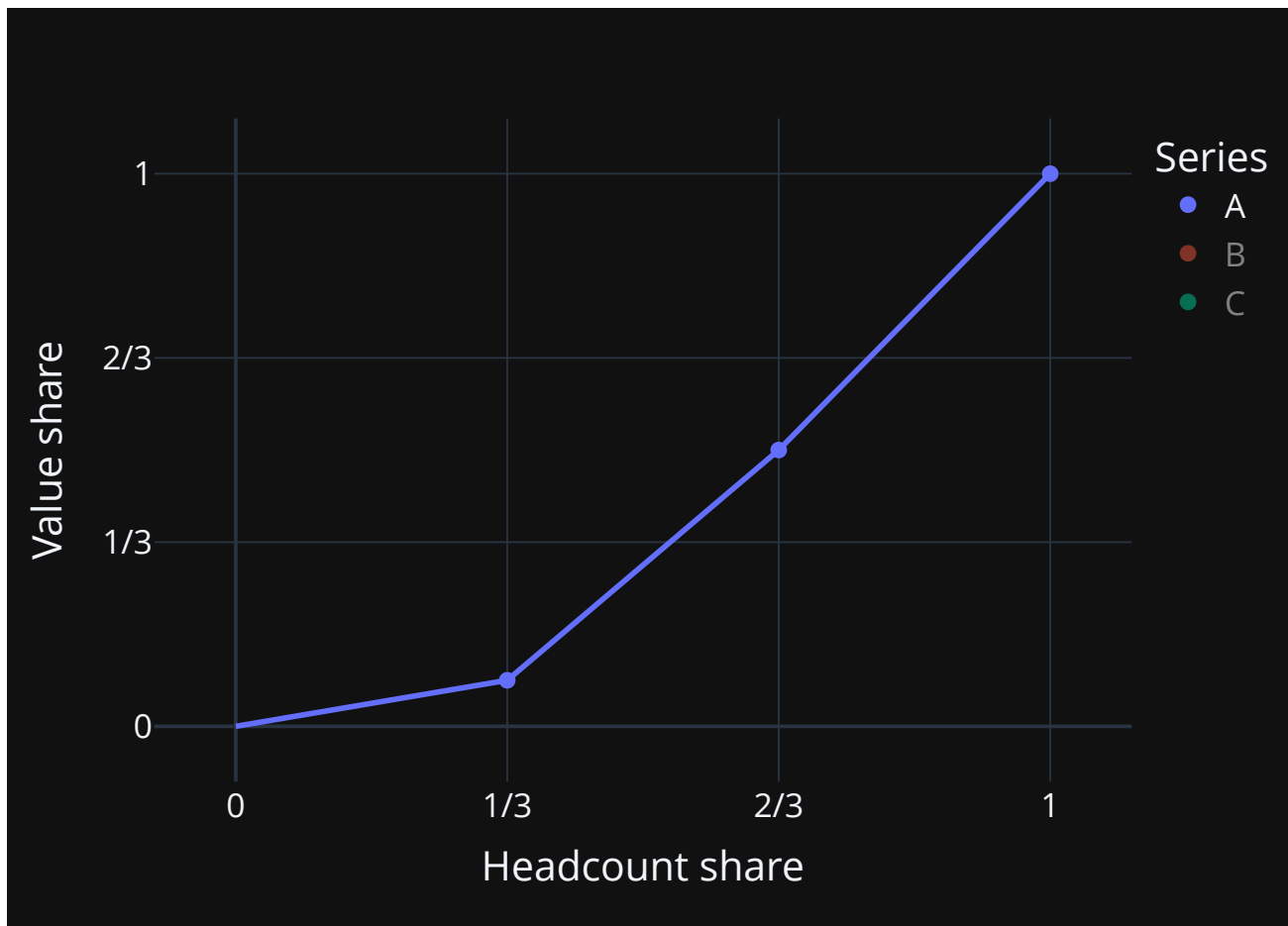
Cum. sum

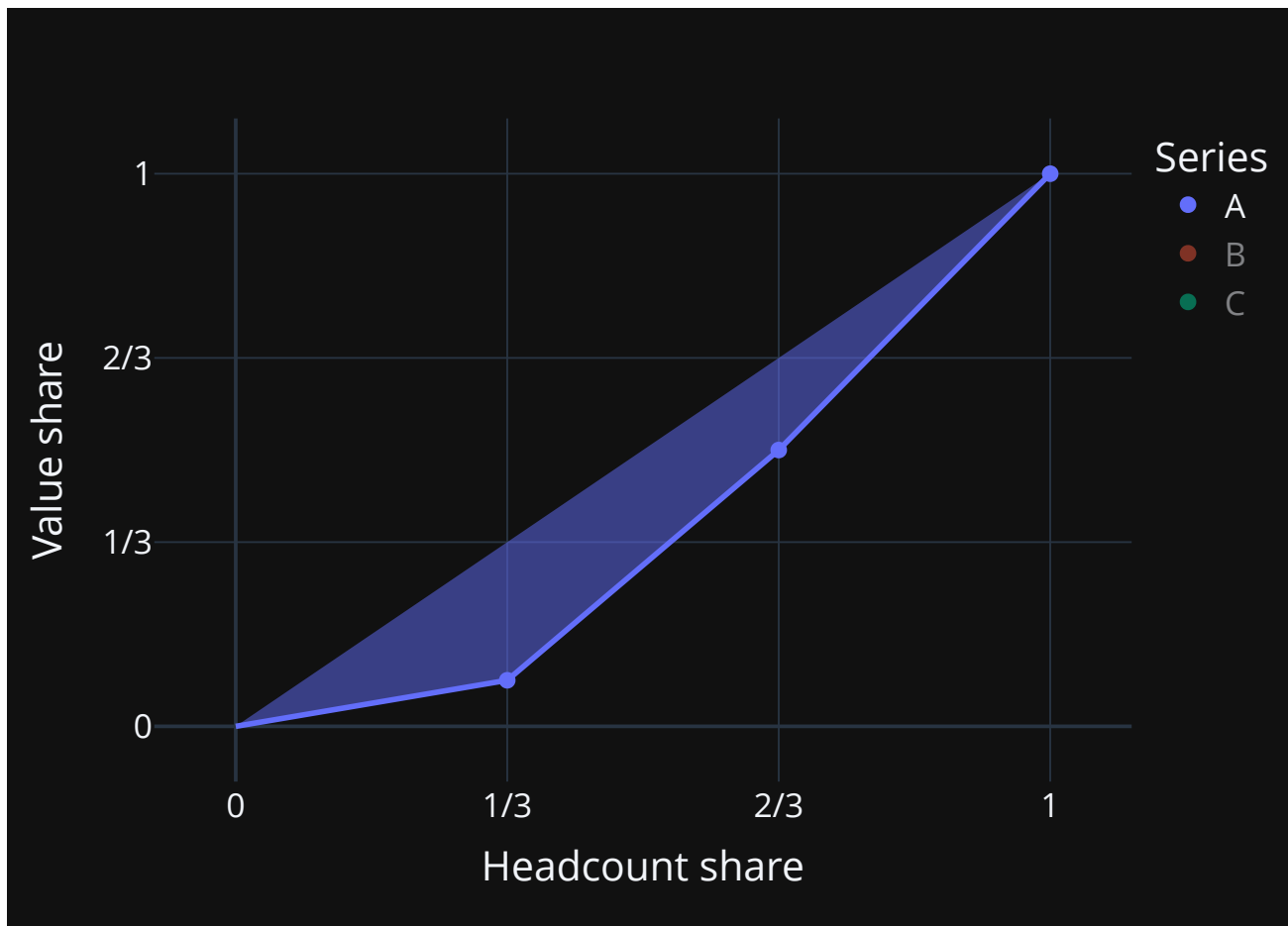
A	B	C
1	2	4
6	4	8
12	12	12

Cum. dist.

A	B	C
0.08	0.17	0.33
0.5	0.33	0.67
1	1	1







Calculation: Formula

$$p_i = \frac{x_i}{\sum_{i=1}^k x_i}$$

$$G = \frac{2 \cdot \sum_{i=1}^n i \cdot p_i - (n + 1)}{n}$$

Calculation: Example

Ordered data

A	B	C
1	2	4
5	2	4
6	8	4

Normal. data

A	B	C
0.08	0.17	0.33
0.42	0.17	0.33
0.5	0.67	0.33

Gini

A	B	C
0.28	0.33	0

