SE

Introduction to summation notation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_{i,j} = ?$$

Summation notation

Summation notation is used in math and physics classes:

$$\sum_{i=1}^{6} i = 1+2+3+4+5+6 = 21$$

$$\sum_{i=2}^{5} i^2 = 2^2+3^2+4^2+5^2 = 4+9+16+25 = 54$$

$$4\left(\sum_{i=0}^{\infty}\frac{(-1)^i}{1+2i}\right) = \frac{4}{1}-\frac{4}{3}+\frac{4}{5}-\frac{4}{7}+\frac{4}{9}\ldots = \pi$$

Data tables

The most common way to organize data is in tables.

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4

Usually each column is a data set and the rows represent individual values taken from each set.

Data tables

The most common way to organize data is in tables.

ID#	West	North	East	South
1	5	8	9	4
2	3	8	7	9
3	2	4	6	8
4	1	2	3	4
5	9	8	7	4

Note that the left set of values are just labels, not data, so we won't be using those values in calculations.

Data tables

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4

$$x_{3,2} = 8$$

$$x_{4,1} = 1$$

$$x_{1,4} = 4$$

We can identify any value in the data table by specifying the row and column. Most common is row, then column.

Data tables

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4

$$x_{row,column}$$

$$x_{2,3} = 7$$

We can identify any value in the data table by specifying the row and column. Most common is row, then column.

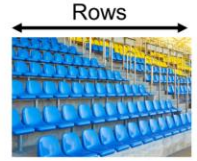
Data tables

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4

COLUMNS: We refer to the values in the same vertical line as being in the same column. Like the columns on a building.

Data tables

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4



ROWS: We refer to the values in the same horizontal line as being in the same row. Like rows of seats in a movie theater or stadium.

Data tables

Once we have a bunch of data and a way to specify values, we typically want to compute other values from them.

- A huge data set is impossible to understand, but one or two summary statistics are.

Two symbols used to describe calculations across rows and columns.

$\sum$ Sum and $\prod$ Product

Describe calculations across rows and columns.
The "i" and "j" values are dummy variables to change in the range specified below and above the symbols.

Sum $\sum_{i=1}^{3} x_{i,1} = x_{1,1}+x_{2,1}+x_{3,1}$

Product $\prod_{j=2}^{4} x_{1,j} = x_{1,2}\times x_{1,3}\times x_{1,4}$

Best approach until you're comfortable is to carefully rewrite using parentheses.

Be careful, the order can matter.

$$\sum_{i=1}^{3}\sum_{j=2}^{4} x_{i,j} = \sum_{i=2}^{4}\sum_{j=1}^{3} x_{j,i} \qquad \sum_{i=1}^{3}\sum_{j=2}^{4} x_{i,j} \neq \sum_{i=2}^{4}\sum_{j=1}^{3} x_{i,j}$$

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4

In truth, Σ is used much more than Π in statistics, so we won't worry about Π any more in this video.

$$\sum_{i=1}^{3} x_{i,1} = x_{1,1}+x_{2,1}+x_{3,1} = 5+3+1 = 9$$

$$\prod_{j=2}^{4} x_{1,j} = x_{1,2}\times x_{1,3}\times x_{1,4} = 8\times 9\times 4 = 288$$

Data tables

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4

$$\sum_{i=2}^{4} x_{2,i} = ?$$

$$\sum_{i=2}^{4} x_{2,i} = x_{2,2}+x_{2,3}+x_{2,4} = 5+7+9 = 21$$

We can specify calculations using this notation.

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4

$$\bar{x} = \frac{\sum_{i=1}^{2}\sum_{j=1}^{4} x_{i,j}}{8}$$

e.g., what is the mean of the top two rows?

$$\bar{x} = \frac{\sum_{i=1}^{2}\sum_{j=1}^{4} x_{i,j}}{8}$$

ID#	West	North	East	South
1	5	8	9	4
2	3	5	7	9
3	2	8	6	8
4	1	2	3	4
5	9	8	7	4

$$\sum_{i=1}^{3}\sum_{j=2}^{4} x_{i,j} = ?$$

$$\sum_{i=1}^{3}\sum_{j=2}^{4} x_{i,j} = \left(\sum_{j=2}^{4} x_{1,j}\right)+\left(\sum_{j=2}^{4} x_{2,j}\right)+\left(\sum_{j=2}^{4} x_{3,j}\right) = (x_{1,2}+x_{1,3}+x_{1,4})+(x_{2,2}+x_{2,3}+x_{2,4})+(x_{3,2}+x_{3,3}+x_{3,4})$$

$$= (8+9+4)+(5+7+9)+(8+6+8) = 21+21+22 = 64$$

If we want to perform calculations using more than one row or column, we use double summation notation.
Best approach until you're comfortable is to carefully rewrite using parentheses.

$$\sum_{i=1}^{3}\sum_{j=2}^{4} x_{i,j} = \sum_{i=1}^{3}\left(\sum_{j=2}^{4} x_{i,j}\right) = \left(\sum_{j=2}^{4} x_{1,j}\right)+\left(\sum_{j=2}^{4} x_{2,j}\right)+\left(\sum_{j=2}^{4} x_{3,j}\right)$$