# 5.2 LESSON NOTES

## AGGREGATORS

- Essential when you're summarising or clustering data
- Available aggregators are:
  - **SUM** - Sums up all values to give a total
  - **AVG** - the SUM of values divided by the COUNT of values
  - **COUNT** - counts the number of values
  - **COUNT DISTICT** - counts the number of **different** values
  - **MIN** - returns the minimum value
  - **MAX** - returns the maximum value
  - **MED** - returns the middle value in a string of numbers. Useful when you'd like to calculate an average but you have outliers in your data that skew the average value

### MEDIAN

String a)
1 , 2 , 3 , 4 , 5

Average = 3 (15 / 5)
Median = 3

String b)
~~1~~ , ~~2~~ , 3 , ~~4~~ , ~~500~~

Average = 102 (510 / 5)
Median = 3

**vitamin** BI
BUSINESS INTELLIGENCE

- **STANDARD DEVIATION** - measures the amount of variation or dispersion of a set of values from the mean (or average) value
- **VARIANCE** - measures the spread between numbers in a dataset. Or how far each number in the set is from the mean

- The same metric can be added to a query multiple times so that you can apply different aggregators to it
- **AUT** (Automatic) - some metrics will not allow you to change the aggregator because it has already been specified in the definition or formula of that metric
- Dimensions can be added as metrics with the COUNT or COUNT DISTINCT aggregator