

Data Analysis and Visualisation

ACCB4002

What is Data Science?

- Better decisions
- Predictive analysis
- Pattern discoveries



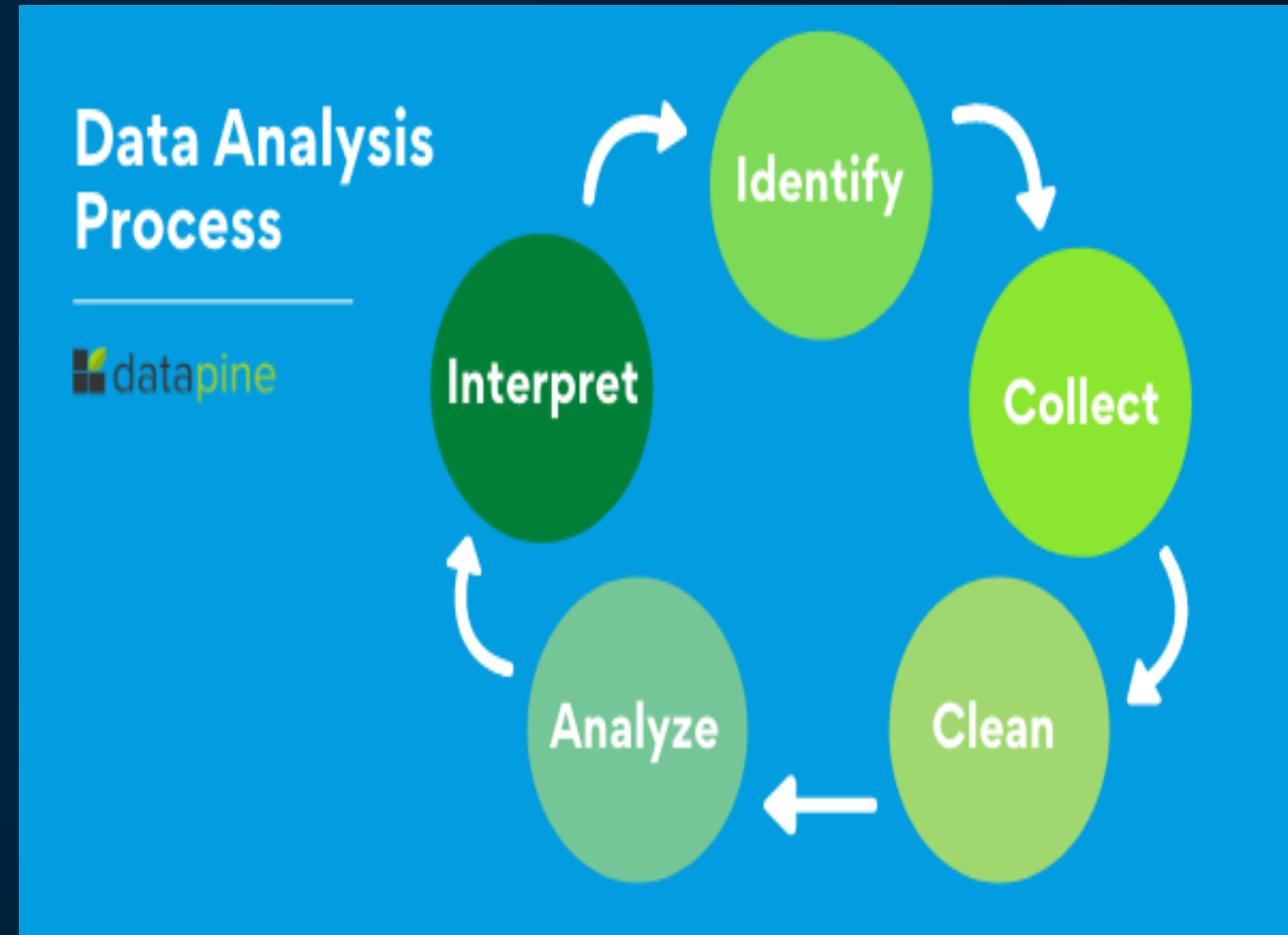
Where is Data Science Needed?

Where is Data Science Needed?

- For route planning
- To foresee delays for flight/ship/train etc.
- To create promotional offers
- To find the best suited time to deliver goods
- To forecast the next years revenue for a company
- To analyze health benefit of training
- To predict who will win elections

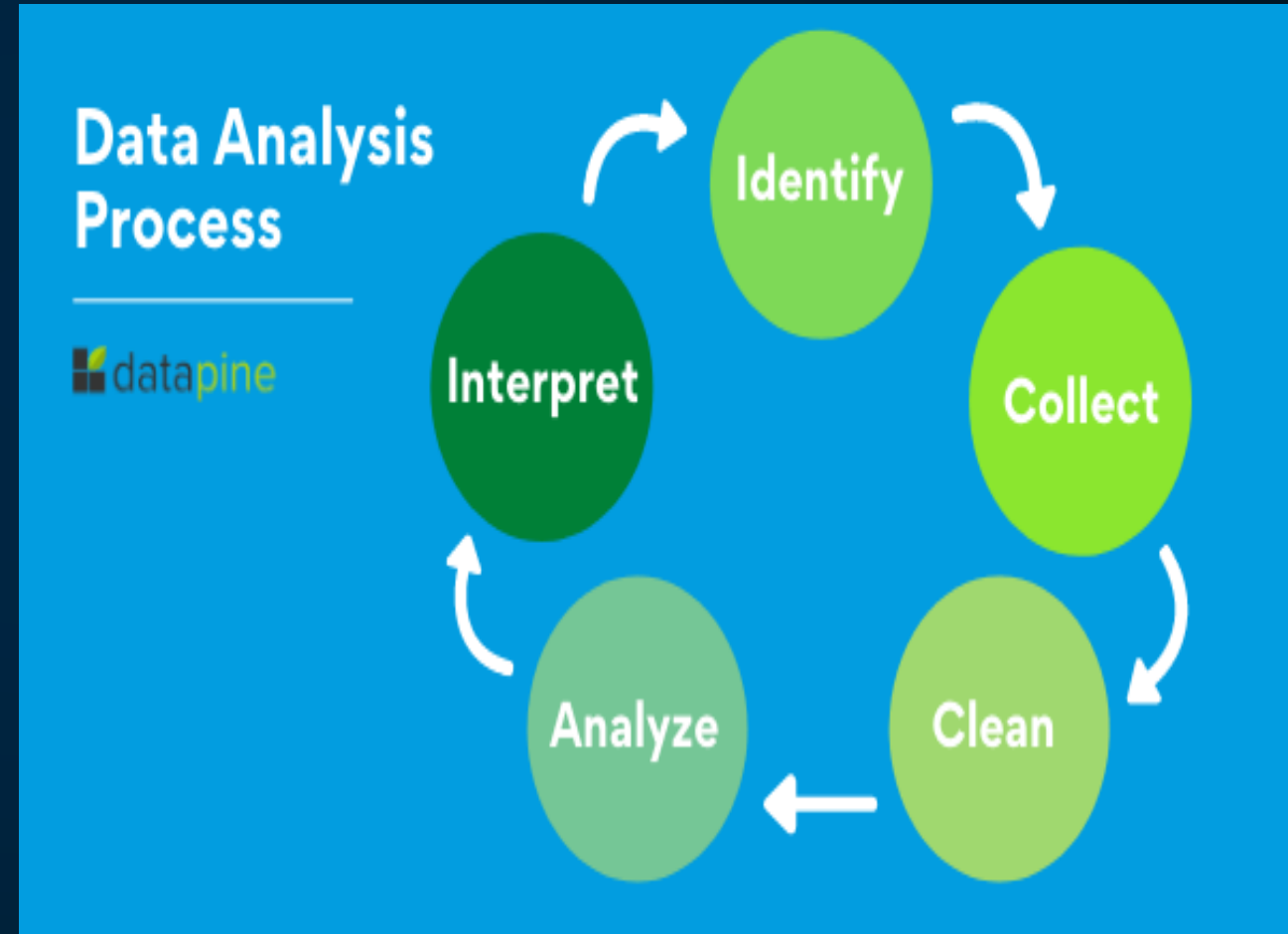
Here is how a Data Analyst works:

- **Ask the right questions**
- **Explore and collect data**
- **Extract the data**
- **Clean the data**
- **Find and replace missing values**
- **Normalize data**



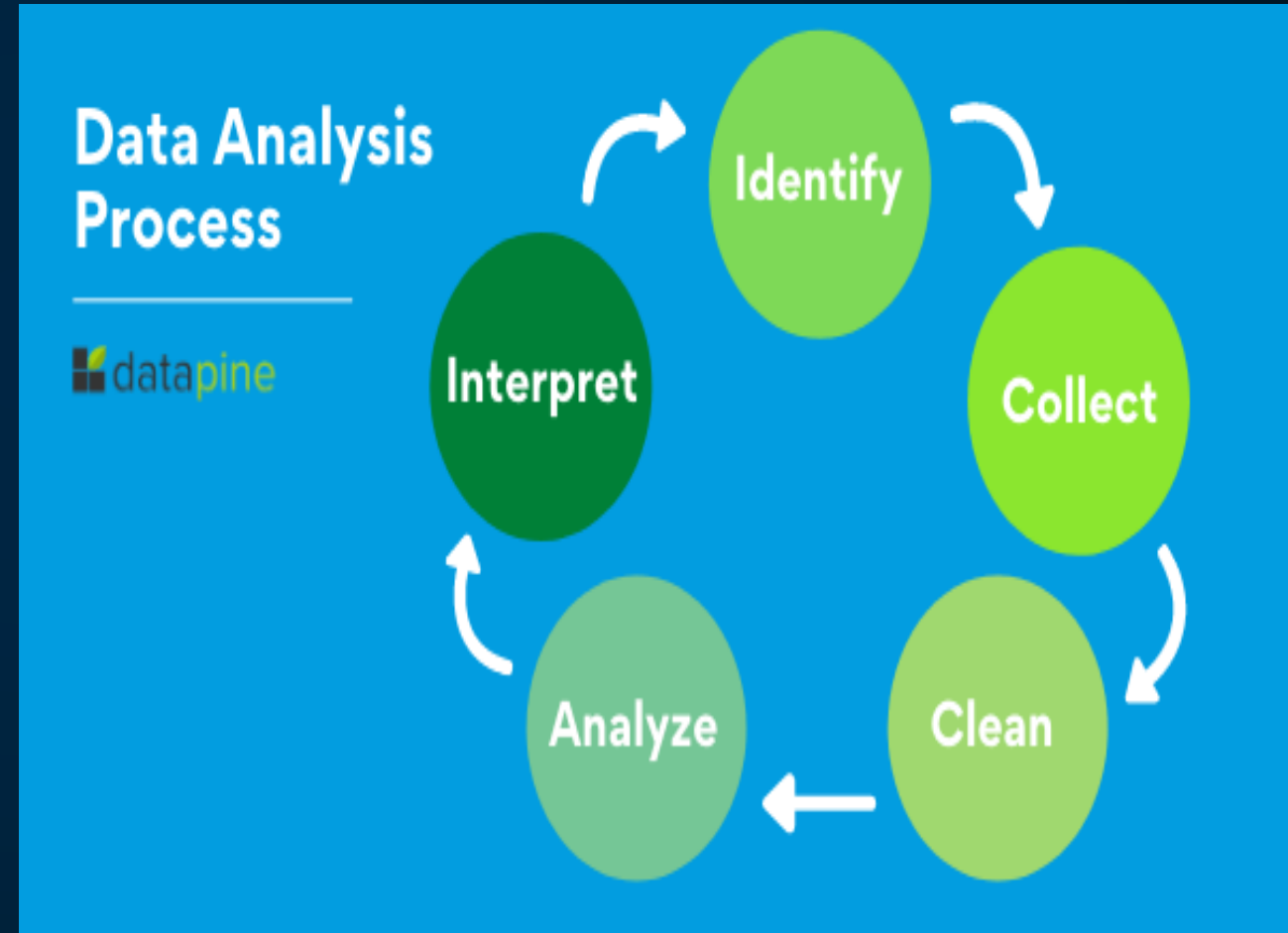
Here is how a Data Analyst works:

- Ask the right questions
- Explore and collect data
- **Extract the data**
- **Clean the data**
- Find and replace missing values
- Normalize data



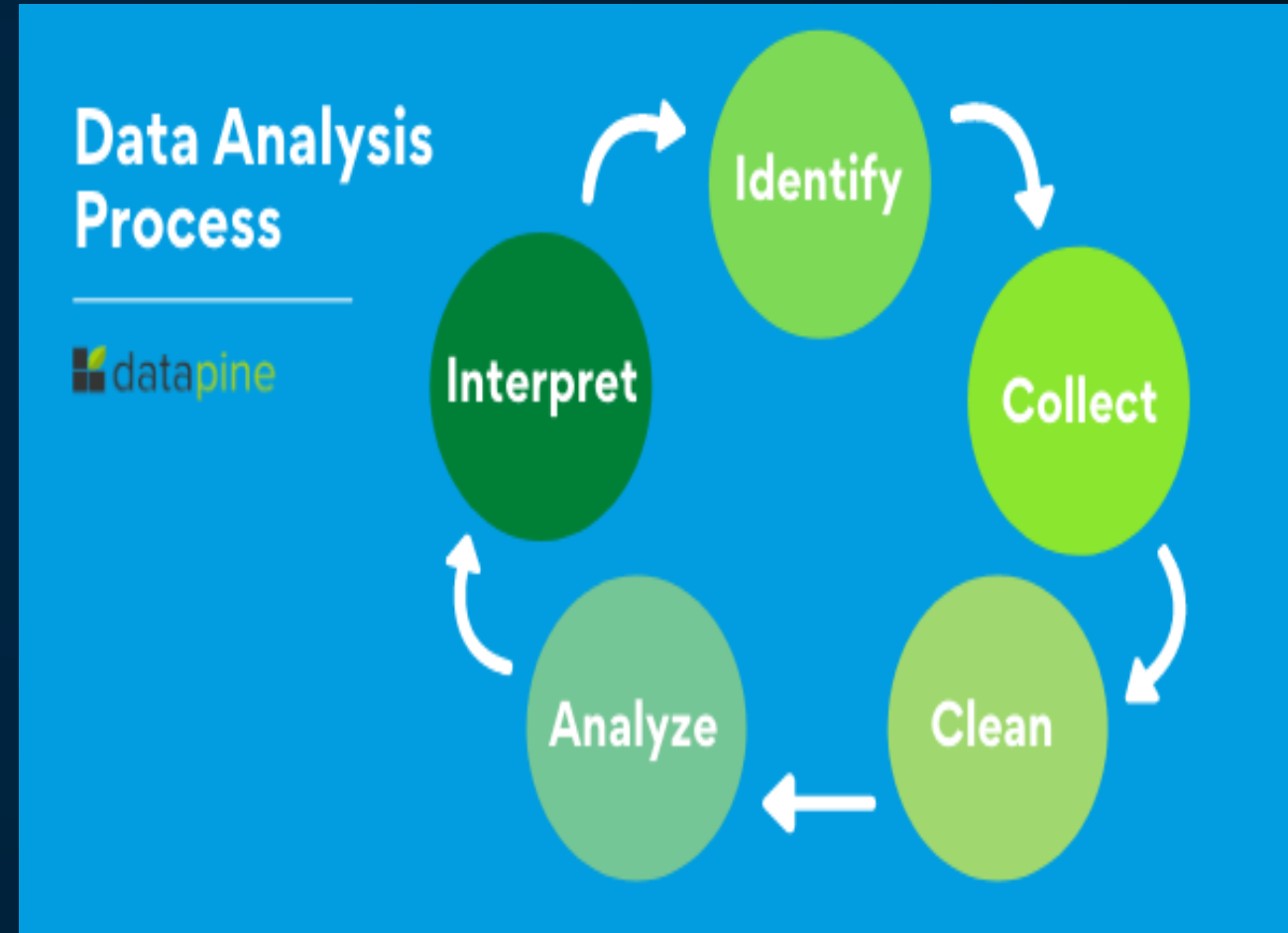
Here is how a Data Analyst works:

- Extract the data
- Clean the data
- Find and replace missing values
- Normalize data



Here is how a Data Analyst works:

- Find and replace missing values
- Normalize data
- Analyze data, find patterns and make future predictions.
- Represent the result

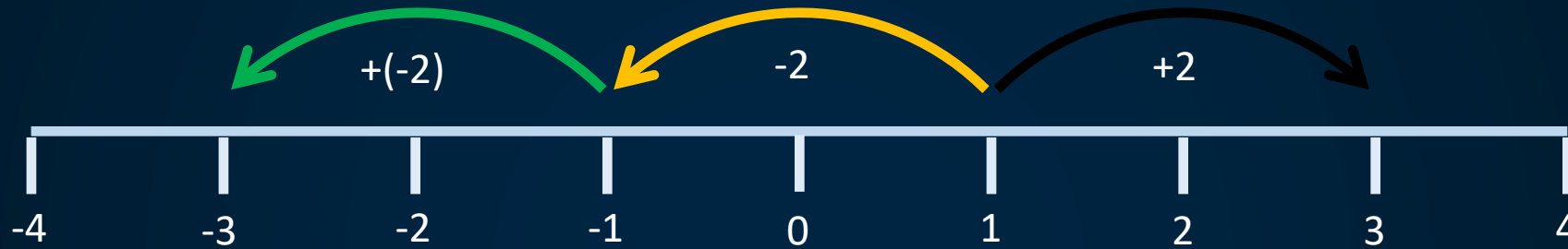


Basic Math review

Basic math review

- We'll start with a gentle review of the necessary maths
 - Positive & Negative Numbers
 - Fractions
 - Decimals & Percentages
 - Exponents and Roots
 - Order of Computation
 - Summation
 - Equations

Positive & Negative numbers



- **Rule 1:** If something is written without a sign, it's positive

$$+2 = 2 \quad +x = x$$

- **Rule 2:** When adding numbers of the same sign, add them up and prefix them with the same sign that the individual numbers had

$$+3 + +5 = +8 \quad 2 + 7 = 9 \quad -4 + -2 + -1 = -7$$

$$-x + -5x = -6x$$

Positive & Negative numbers (Contd..)

- **Rule 3:** When summing a group of numbers with mixed signs follow these three steps

1. Add the positive numbers; add the negative numbers as if they were positive
2. Subtract the smaller sum from the larger sum
3. Prefix your answer with the sign of the larger sum

$$5 - 3 + 2 - 1 \rightarrow (+)(5 + 2) \text{ and } (-)(3 + 1) = (+)(7) \text{ and } (-)(4)$$

$$7 \text{ (larger)} - 4 \text{ (smaller)} = 3$$

Take the sign of the larger number (+) therefore +3

Positive & Negative numbers (contd...)

What about $-2 + 6 - 14 + 3 - 4$?

Positive & Negative numbers (contd...)

What about $-2 + 6 - 14 + 3 - 4$?

$$-2 + 6 - 14 + 3 - 4 \rightarrow (+)(6 + 3) \text{ and } (-)(2 + 14 + 4)$$

$$(+)(9) \text{ and } (-)(20)$$

$$20 \text{ (larger)} - 9 \text{ (smaller)} = 11$$

Take the sign of the larger number (-) therefore -11

Positive & Negative numbers(Contd..)

- **Rule 4:** Adding a negative number is the same as subtracting a positive number; subtracting a negative number is the same as adding a positive number
 - Two positive sign make a plus
 - Two negative signs make a plus
 - $5 - (-4) = 5 + 4 = 9$
- A positive and a negative make a minus
- A negative and a positive make a minus
- $5 - (+3) = 5 + (-3) = 5 - 3 = 2$
- $7 + (-12) = 7 - (+12) = 7 - 12 = -5$

Positive & Negative numbers(Contd..)

Rule 5: When multiplying or dividing two things of the *same* sign the result is always *positive*

$$3 \bullet 7 = 21 \quad -8 \times -3 = 24 \quad 12(8) = 96$$

$$15 \div 3 = 5 \quad (-9) \div (-1) = 9 \quad -6 / -4 = 1.5$$

Rule 6: When multiplying or dividing two things of *different* signs the result is always *negative*

$$3 \bullet -7 = -21 \quad -8 \times 3 = -24 \quad 12(-8) = -96$$

$$-15 \div 3 = -5 \quad (9) \div (-1) = -9 \quad -6 / 4 = -1.5$$

Positive & Negative numbers(Contd...)

Rule 7: With more than two numbers to multiply or divide, take them pairwise in order and follow rules 5 and 6

If there's an odd number of negatives the result will be negative

If there's an even number of negatives the result will be positive

$$3 \times -2 \times -4 \times 3 = -6 \times -4 \times 3 = 24 \times 3 = 72$$

$$40 \div 2 \div 2 \div -2 = 20 \div 2 \div -2 = 10 \div -2 = -5$$

$$6 \div 3 \times -1 \div 5 = 2 \times -1 \div 5 = -2 \div 5 = -4$$

Exercise

- Do the positive and negative numbers multiple choice quiz on Moodle

Fractions

Rule 1: Fractions symbolise division – divide the top expression (the numerator) by the bottom expression (the *denominator*)

$$\frac{-9}{3} = -3 \qquad \frac{6}{4-1} = 2 \qquad \frac{(13-3)(7+3)}{-3+2} = \frac{(10)(10)}{-1} = -100$$

Rule 2: Dividing by 0 is impossible. If you're asked to divide by 0 at any point you've probably made a mistake somewhere

Rule 3: You can turn anything into a fraction by dividing it by 1

$$3 = \frac{3}{1} \qquad -6.2 = \frac{-6.2}{1} \qquad 3x + 4 = \frac{3x+4}{1}$$

Fractions (Contd.,)

Rule 4: To multiply fractions, multiply the numerators together and multiply the denominators together

$$\frac{2}{3} \cdot \frac{1}{2} = \frac{2}{6} = 0.33$$

$$\frac{1}{5} \cdot 10 = \frac{1}{5} \cdot \frac{10}{1} = \frac{10}{5} = 2$$

$$3xy \cdot \frac{3}{4} = \frac{3xy}{1} \cdot \frac{3}{4} = \frac{3(3xy)}{4} = \frac{9xy}{4}$$

Rule 5: Multiplying both the numerator and denominator by the same number doesn't change its value

$$\frac{1}{2} = \frac{2 \cdot 1}{2 \cdot 2} = \frac{2}{4} = 0.5$$

Fractions (Contd.,)

- **Rule 6:** To divide by a fraction, invert and multiply. Switch around the numerator and the denominator of the fraction you want to divide by, then multiply it by the thing you're dividing

$$\begin{aligned} 21 \div \frac{3}{5} &= \frac{21}{1} \div \frac{3}{5} \\ &= \frac{21}{1} \cdot \frac{5}{3} = \frac{105}{3} = 35 \end{aligned}$$

Fractions (Contd.,)

- Rule 7: To add or subtract fractions, they must have the a common denominator, i.e. their denominators must be the same, then the nominators can be added or subtracted as usual

$$\bullet \frac{2}{3} + \frac{1}{2} = \frac{2 \cdot 2}{2 \cdot 3} + \frac{3 \cdot 1}{3 \cdot 2} = \frac{4}{6} + \frac{3}{6} = \frac{7}{6} = 1 \frac{1}{6}$$

$$\bullet \frac{5}{8} - \frac{1}{2} = \frac{1 \cdot 5}{1 \cdot 8} - \frac{4 \cdot 1}{4 \cdot 2} = \frac{5}{8} - \frac{4}{8} = \frac{1}{8} = \frac{1}{8}$$

$$\bullet \frac{1}{3} + \frac{1}{2} + \frac{3}{4} - \frac{1}{12} = \frac{4 \cdot 1}{4 \cdot 3} + \frac{6 \cdot 1}{6 \cdot 2} + \frac{3 \cdot 3}{3 \cdot 4} - \frac{1}{12} = \frac{4}{12} + \frac{6}{12} + \frac{9}{12} - \frac{1}{12} = \frac{18}{12} = 1 \frac{1}{2}$$

Exercise

- Do the fractions multiple choice quiz on Moodle

Decimals and Percentages

Rule 1: Decimals indicate that the part of the number following the decimal point is a fraction with 10, 100, 1000 and so on as the denominator

$$3.2 = 3\frac{2}{10} \quad 3.25 = 3\frac{25}{100} \quad 3.257 = 3\frac{257}{1000}$$

Rule 2: Percentages are fractions of 100 (two-place decimals)

$$45\% = \frac{45}{100} = 0.45 \quad 3.2\% = \frac{3.2}{100} = 0.032$$

$$110\% = \frac{110}{100} = 1.1$$

Exercise

- Do the decimals and percentages multiple choice quiz on Moodle

Exponents and roots

Rule 1: An exponent is a small number placed slightly higher than and to the right of a number. 3^3 has an exponent of 3. $(x-y)^2$ has an exponent of 2.

The exponent tells how many times a number or expression is to be multiplied by itself

$$2^4 = 2 \cdot 2 \cdot 2 \cdot 2 \quad Y^2 = Y(Y) \quad 10^3 = 10 \cdot 10 \cdot 10$$

Rule 2: Roots are the opposite of exponents. The square root is the opposite of an exponent of 2, the cube root is the opposite of the exponent of 3. Statistics generally uses only square roots.

$$\sqrt{9} = 3 \quad \sqrt{36 + 13} = \sqrt{49} = 7$$

$$\sqrt{36} + 13 = 6 + 13 = 19$$

Exercise

- Do the exponents and roots multiple choice quiz on Moodle

Order of Computation

Rule 1: Any expressions within brackets has higher priority with nested brackets work from the inside out.

$$4 (7 - 2) = 4 (5) = 20$$

Treat the numerator and the denominator of a complicated fraction as if each were contained in brackets

$$\frac{3 + 2}{5(7 - 3)} = \frac{5}{5(4)} = \frac{5}{20} = \frac{1}{4}$$

Rule 2: Roots and exponents are handled first

$$(2 - 3^2)(4 + 2^2) = (2 - 9)(4 + 4) = (-7)(8) = -56$$

Order of computation

Rule 3: The order of operation is **MULTIPLY, DIVIDE, ADD, SUBTRACT**

$$\begin{aligned} & 5 - (3 \times 4) \times (8 - 2^2)(-3 + 1) / 3 && \text{(exponent)} \\ & = 5 - (3 \times 4) \times (8 - 4)(-3 + 1) / 3 && \text{(things in} \\ & \text{brackets)} \\ & = 5 - (12) \times (4)(-2) / 3 && \text{(multiply)} \\ & = 5 - 48 \times -2 / 3 && \text{(multiply again)} \\ & = 5 - -96 / 3 && \text{(divide)} \\ & = 5 - (-32) && \text{(subtract)} \\ & = 37 \end{aligned}$$

Exercise

- Do the order of computation multiple choice quiz on Moodle

Summation

These are shown using a Σ symbol, which means 'add up all of whatever comes next'.

$$\sum x = \sum_{i=1}^N x \quad \text{but we normally use the short version}$$

Suppose 5 people take a quiz and get 10, 10, 8, 12 and 10

$$x_1 = 10, x_2 = 10, x_3 = 8, x_4 = 12, x_5 = 10$$

$$\sum x = 10 + 10 + 8 + 12 + 10 = 50$$

$$\sum x^2 = 100 + 100 + 64 + 144 + 100 = 508$$

$$(\sum x)^2 = 50^2 = 2500$$

Summation

Suppose they take a second quiz and get 4, 5, 6, 5 and 4

$$y_1 = 4, y_2 = 5, y_3 = 6, y_4 = 5, y_5 = 4$$

X values : 10, 10 , 8, 12 and 10

$$\sum y$$

$$\sum y^2$$

$$\sum x + \sum y$$

$$\sum x^2 + \sum y^2$$

$$(\sum(x + y))^2$$

Summation

Suppose they take a second quiz and get 4, 5, 6, 5 and 4

$$y_1 = 4, y_2 = 5, y_3 = 6, y_4 = 5, y_5 = 4$$

$$\sum y = 4 + 5 + 6 + 5 + 4 = 24$$

$$\begin{aligned}\sum y^2 &= (4)^2 + (5)^2 + (6)^2 + (5)^2 + (4)^2 \\ &= 16 + 25 + 36 + 25 + 16 = 118\end{aligned}$$

$$\sum x + \sum y = 50 + 24 = 74$$

$$\sum x^2 + \sum y^2 = 2500 + 118 = 2618$$

$$\begin{aligned}(\sum (x + y))^2 &= ((10+4) + (10+5) + (8+6) + (12+5) + (10+4))^2 \\ &= (14 + 15 + 14 + 17 + 14)^2 = 74^2 = 5476\end{aligned}$$

$$\begin{aligned}\sum (x + y)^2 &= (10+4)^2 + (10+5)^2 + (8+6)^2 + (12+5)^2 + (10+4)^2 \\ &= (14)^2 + (15)^2 + (14)^2 + (17)^2 + (14)^2 \\ &= 196 + 225 + 196 + 289 + 196 = 1102\end{aligned}$$

Exercise

- Do the summation multiple choice quiz on Moodle

Equations

Rule 1: Adding or subtracting the same number from each side of an equation is acceptable and the two sides will still be equivalent

$$8 - 2 = 3 + 3$$

$$2 + 8 - 2 = 2 + 3 + 3 \quad 8 = 8$$

$$2x + 7 = 35$$

$$2x + 7 - 7 = 35 - 7 \quad 2x = 28$$

Rule 3: If you multiply or divide one side of an equation by a number then you must do the same to the other side

$$3 + 2 - 1 = 7 - 5 + 2$$

$$6(3 + 2 - 1) = 6(7 - 5 + 2) \quad 6(4) = 6(4) \quad 24 = 24$$

Exercise

- Do the equations multiple choice quiz on Moodle

Summary

- Topics covered
 - Positive & Negative Numbers
 - Fractions
 - Decimals & Percentages
 - Exponents and Roots
 - Order of Computation
 - Summation
 - Equations

References

- Chapter 1 -Kranzler, J.H. (2010) *Statistics for the Terrified*, Pearson