General Mathematics UDF Guide

Version 1.0

LazyMath

Acknowledgements

A huge thank you to everyone who helped test the UDFs 😊

Contents

Contact
lazymath2024@gmail.com

Data Analysis

Summary Statistics

Determines the quartiles, fences, mean, and standard deviation of the input data.

Syntax

sum_stats(Data)

Where, *Data* represents a list containing the data to be analysed.

Example

The number of points a pro gamer scores on Flappy Bird over 10 games is shown in the table below.

Game	1	2	3	4	5	6	7	8	9	10
Score	12	47	58	73	20	31	10	22	17	250

Determine the quartiles, fences, and outliers (if any).

score_data $\{12.,47.,58.,73.,20.,31.,10.,22.,17.,250.\}$

summary_stats(score_data)

Total = 10

▶Data Summary:

"Minimum"	10.
"Q1"	17.
"Q2"	26.5
"Q3"	58.
"Maximum"	250.
"IQR"	41.
"Lower Fence"	-44.5
"Upper Fence"	119.5
"Range"	240.
"Mean"	54.
"Standard Dev"	71.972
"Skew"	"Positive"

Warning: Skew may be inaccurate

▶Possible outliers:

$\{250.\}$

Sorted data saved as data.summary_stats

Done

Warning: Skew may be inaccurate

Contact
lazymath2024@gmail.com

Dot Plot

Determines the summary statistics of an input dot plot.

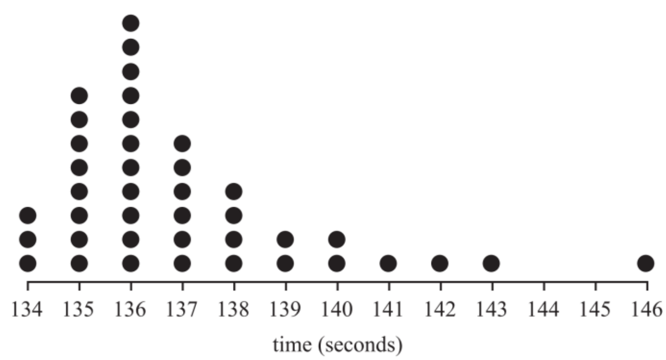
Syntax

dot_plot(Data)

Where, *Data* represents a matrix with the x-values in the top row, and the y-values in the bottom row.

Example

The dot plot shows the times, in seconds, of 40 runners in the qualifying heats of their 800 m club championship.



Source: *VCAA 2023 General Mathematics Examination 1 Question 1*

Determine the median and skew of the data.

$$dot\_plot\left(\begin{bmatrix} 134 & 135 & 136 & 137 & 138 & 139 & 140 & 141 & 142 & 143 & 146 \\ 3 & 8 & 11 & 6 & 4 & 2 & 2 & 1 & 1 & 1 & 1 \end{bmatrix}\right)$$

Total = 40

▸Data Summary:

"Minimum"	134.
"Q1"	135.
"Q2"	136.
"Q3"	138.
"Maximum"	146.
"IQR"	3.
"Lower Fence"	130.5
"Upper Fence"	142.5
"Range"	12.
"Mean"	137.05
"Standard Dev"	2.5715
"Skew"	"Positive"

▸Possible outliers:

{143.,146.}

Sorted data saved as data.dot_plot

Done

Warning: Skew may be inaccurate.

Contact
lazymath2024@gmail.com

Histogram

Determines the summary statistics of an input histogram.

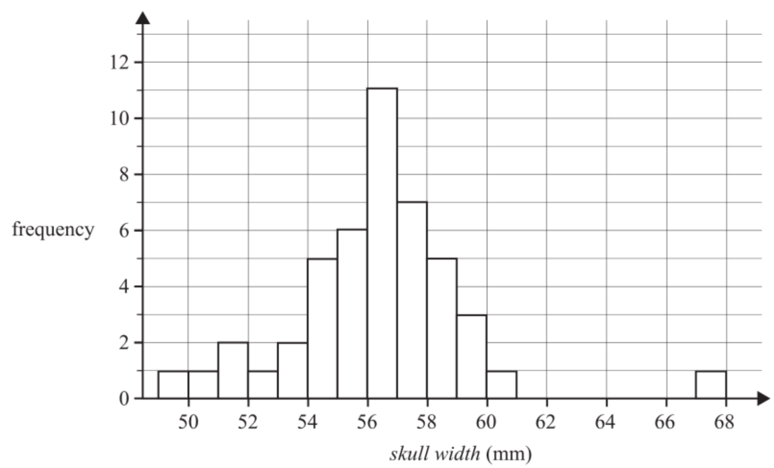
Syntax

histogram(Data)

Where, *Data* represents a matrix with the x-values in the top row, and the y-values in the bottom row.

Example

The histogram below displays the distribution of *skull width*, in millimeters, for 46 female possums.



Source: VCAA 2022 Further Mathematics Written Examination 1 Question 1

$histogram\left(\begin{bmatrix} 49 & 50 & 51 & 52 & 53 & 54 & 55 & 56 & 57 & 58 & 59 & 60 & 67 \\ 1 & 1 & 2 & 1 & 2 & 5 & 6 & 11 & 7 & 5 & 3 & 1 & 1 \end{bmatrix}, 1\right)$

Total = 46

▸Data Summary:

"Minimum"	"49–50"
"Q1"	"54–55"
"Q2"	"56–57"
"Q3"	"57–58"
"Maximum"	"67–68"
"IQR"	"2–4"
"Lower Fence"	"48–52"
"Upper Fence"	"60–64"
"Range"	"17–19"

▸Approximate values:

"Mean"	56.326
"Standard Dev"	2.9235
"Skew"	"Negative"

Warning: Skew may be inaccurate

▸Possible outliers:

{"49–50","67–68"}

Sample data saved as data.histogram

Done

Warning: Skew may be inaccurate.

Contact
lazymath2024@gmail.com

Frequency Table

Determines the frequency table of the input data list.

Syntax

freq_table(Data, *Minimum*, *Bin Size*)

Where,

Data represents a list containing the data

Minimum represents the starting point of the frequency table

Bin Size represents the size of each bin in the frequency table

Example

Determine the frequency table of the following data.

{35, 48, 45, 43, 38.2, 50, 39.8, 40.7, 40, 50, 35.4, 38.8, 40.2, 45, 45, 40, 43.3, 53.1, 35.6, 44.1, 34.8}

Start your table from 30 and use a bin size of 5.

freq_table({35,48,45,43,38.2,50,39.8,40.7,40,50,35.4,38.8,40.2,45,45,40,43.3,53.1,35.6,44.1,34.8},30,5)

▸Frequency Table:

"Interval"	"Frequency"	"Percentage"
"30−<35"	1.	4.7619
"35−<40"	6.	28.571
"40−<45"	7.	33.333
"45−<50"	4.	19.048
"50−<55"	3.	14.286
"Total"	21.	100.

Done

Contact
lazymath2024@gmail.com

Inverse Normal

Uses the 68-95-99.7% rule alongside the given mean and standard deviation to determine the values for which $Pr(X > x) = \%p$ and $Pr(X < x) = \%p$.

Syntax

norm_inverse(Mean, Standard Deviation, Percentage Probability)

Where,

Mean represents the mean of the normal distribution

Standard Deviation represents the standard deviation of the normal distribution

Percentage Probability represents the percentage probability of being less than or greater than a value

Example

The weight of dogs is normally distributed with a mean of 30 kg with a standard deviation of 3.4 kg.

Using the 68-95-99.7% rule, determine the weight which 16% of dogs are less than.

norm_inverse(30,3.4,16)

▸Given:

$\bar{x}$ = 30 and sx = 3.4

▸Answer:

16% of values are less than 26.6

16% of values are greater than 33.4

Done

Contact
lazymath2024@gmail.com

Normal Bound

Uses the 68-95-99.7% rule to determine the cumulative percentage probability between two bounds, that is, $Pr(x_1 < X < x_2)$ %.

Syntax

norm_bound(Mean, Standard Deviation, Lower Bound, Upper Bound)

Where,

Mean represents the mean of the normal distribution

Standard Deviation represents the standard deviation of the normal distribution

Lower Bound represents the lower bound in the probability expression

Upper Bound represents the upper bound in the probability expression

Example

The weight of dogs is normally distributed with a mean of 30 kg with a standard deviation of 3.4 kg.

Using the 68-95-99.7% rule, determine the percentage of dogs which weigh between 26.6 kg and 36.8 kg.

norm_bound(30,3.4,26.6,36.8)

▸Given:

$\bar{x} = 30$ and $sx = 3.4$

▸Answer:

81.5% of the values are between 26.6 and 36.8

Done

Contact
lazymath2024@gmail.com

Normal Solve

Uses the 68-95-99.7% rule to determine the mean and standard deviation of a normal distribution, given two probabilities, $Pr(X < x_1) = p_1\%$ and $Pr(X > x_2) = p_2\%$.

Syntax

normsolve(Lower, % Pr(Lower), Upper, %Pr(Upper))

Where,

Lower represents the value, x_1

% Pr(Lower) represents the percentage probability of $X < x_1$, in other words, $p_1\%$

Upper represents the value, x_2

% Pr(Upper) represents the percentage probability of $X > x_2$, in other words, $p_2\%$

Example

The mean and standard deviation for the average weight of dogs is unknown.

After conducting some measurements, scientists determined that:

- 2.5% of dogs weigh more than 36.8 kg
- 16% of dogs weigh less than 26.6 kg

Use the 68-95-99.7% rule to determine, in kilograms, the mean and standard deviation.

norm_solve(26.6,16,36.8,2.5)

▸Given:
16% of values are less than 26.6
2.5% of values are greater than 36.8
▸Determine the number of sx from x̄ using 68−95−99.7% rule:
26.6 is ⁻1 sx from x̄
36.8 is 2 sx from x̄
▸Determine the equations:
26.6 = x̄ − sx
36.8 = x̄ + 2sx
▸Solve equations simultaneously for x̄ and sx:
sx=3.4 and x̄=30.

Done

Contact
lazymath2024@gmail.com

Line Solve

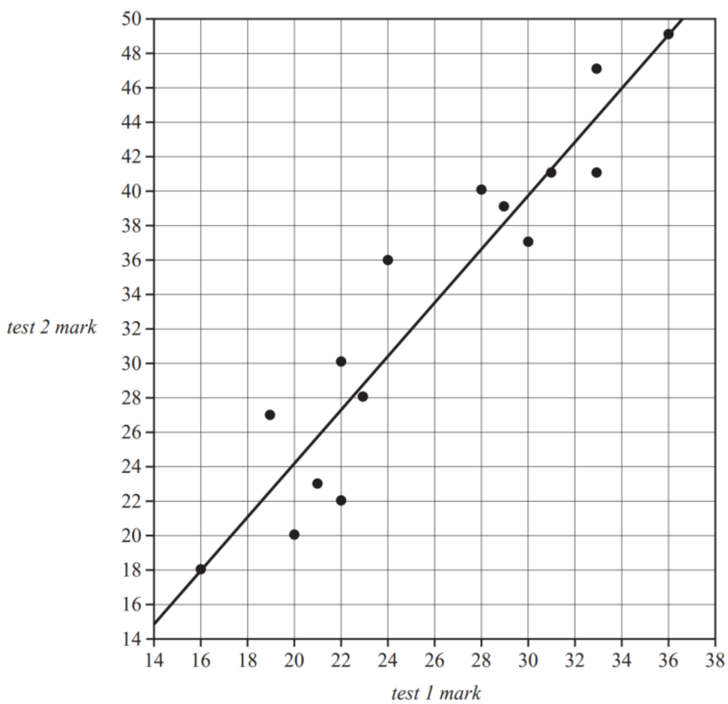
Determines the equation of the line passing through two input points.

Syntax

lin_solve(x1, y1, x2, y2)

Where, *x1*, *y1*, *x2*, *y2* represent the *x* and *y* coordinates of the two points respectively

Example



Source: VCAA 2023 General Mathematics Examination 2 Q7 & Q8

Determine the equation for the least squares line.

$$lin\_solve(16,18,34,46) \quad y=1.5556\cdot x-6.8889$$

Contact
lazymath2024@gmail.com

Linear Regression

Determines the least squares line, R, R^2, and association between the explanatory variable and the response variable.

Syntax

lin_reg(EV, RV)

Where,

EV represents a list containing the values of the explanatory variable

RV represents a list containing the values of the response variable

Example

The amount of money a student earns from their stocks each year is shown in the table below.

Year	1	2	3	4	5	6	7	8
Amount ($)	2.50	6.70	8.90	10.50	11.70	16.20	17.50	19.20

Determine the equation for the line of best fit of the data.

ev $\{1.,2.,3.,4.,5.,6.,7.,8.\}$

rv $\{2.5,6.7,8.9,10.5,11.7,16.2,17.5,19.2\}$

lin_reg(ev,rv)

Length = 8

▶Linear Regression:

$$\begin{bmatrix} \text{"Equation"} & y=2.3095\cdot x+1.2571 \\ \text{"R"} & 0.98899 \\ \text{"R}^2\text{"} & 0.9781 \\ \text{"Association"} & \text{"strong positive"} \\ \text{"Interpolation"} & \text{"}1\le x\le 8\text{"} \end{bmatrix}$$

Done

Linear Transformations

Determines the least squares line and R^2 of various transformations of the explanatory and response variables. These include squaring, reciprocal, and log10.

Syntax

lin_trans(EV, RV)

Where,

EV represents a list containing the values of the explanatory variable

RV represents a list containing the values of the response variable

Example

The amount of money a student earns from their stocks each year is shown in the table below.

Year	1	2	3	4	5	6	7	8
Amount ($)	2.50	6.70	8.90	10.50	11.70	16.20	17.50	19.20

Determine the least squares line with $log_{10}(amount)$ as the explanatory variable.

ev {1.,2.,3.,4.,5.,6.,7.,8.}

rv {2.5,6.7,8.9,10.5,11.7,16.2,17.5,19.2}

lin_trans(ev,rv)

▸Transforms:

"Trans"	"R²"	"Equation"
"x"	"0.9781"	"y=2.3095x+1.2571"
"x^2"	"0.9104"	"$y=0.24168x^2+5.4871$"
"y^2"	"0.9556"	"$y^2=52.617x-72.426$"
"log(x)"	"0.9375"	"y=18.13log(x)+1.2126"
"log(y)"	"0.8464"	"log(y)=0.10845x+0.51181"
"x^{-1}"	"0.7640"	"$y=^{-}17.02x^{-1}+17.432$"
"y^{-1}"	"0.6002"	"$y^{-1}=^{-}0.0364x+0.29046$"

Done

Contact
lazymath2024@gmail.com

Residuals

Determines the least squares line fit and the differences between the true values and predicted values.

Syntax

residual(EV, RV)

Where,

EV represents a list containing the values of the explanatory variable

RV represents a list containing the values of the response variable

Example

The amount of money a student earns from their stocks each year is shown in the table below.

Year	1	2	3	4	5	6	7	8
Amount ($)	2.50	6.70	8.90	10.50	11.70	16.20	17.50	19.20

Determine the residual value for each year.

ev $\{1.,2.,3.,4.,5.,6.,7.,8.\}$

rv $\{2.5,6.7,8.9,10.5,11.7,16.2,17.5,19.2\}$

residual(ev,rv)

$$\begin{bmatrix} \text{"y"} & \text{"}\hat{\text{y}}\text{"} & \text{"Residual"} \\ 2.5 & 3.5667 & -1.0667 \\ 6.7 & 5.8762 & 0.82381 \\ 8.9 & 8.1857 & 0.71429 \\ 10.5 & 10.495 & 0.00476 \\ 11.7 & 12.805 & -1.1048 \\ 16.2 & 15.114 & 1.0857 \\ 17.5 & 17.424 & 0.07619 \\ 19.2 & 19.733 & -0.53333 \end{bmatrix}$$

Done

Contact
lazymath2024@gmail.com

Mean Smoothing

Performs mean smoothing on the provided dataset and returns the result. Points which are marked with a blank string indicate they are invalid.

Syntax

mean_smooth(Data, Size)

Where,

Data represents a list containing the data to be mean smoothed

Size represents the group size which is used in smoothing

Example

The number of sales made by a company for the first eight months of 2025 is shown in the table below.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Sales	200	250	100	350	450	500	890	320

Determine the four-mean smoothed data, with centering.

data $\{200.,250.,100.,350.,450.,500.,890.,320.\}$

$mean\_smooth(data,4)$

$$\begin{bmatrix} \text{"x"} & \text{"y"} \\ 1. & \text{"}\square\text{"} \\ 2. & \text{"}\square\text{"} \\ 3. & 256.25 \\ 4. & 318.75 \\ 5. & 448.75 \\ 6. & 543.75 \\ 7. & \text{"}\square\text{"} \\ 8. & \text{"}\square\text{"} \end{bmatrix}$$

Done

Contact
lazymath2024@gmail.com

Median Smoothing

Performs median smoothing on the provided dataset and returns the result. Points which are marked with a blank string indicate they are invalid.

Syntax

med_smooth(Data, Size)

Where,

Data represents a list containing the data to be median smoothed

Size represents the group size which is used in smoothing

Example

The number of sales made by a company for the first eight months of 2025 is shown in the table below.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Sales	200	250	100	350	450	500	890	320

Determine the three-median smoothed data.

$$data \qquad \{200.,250.,100.,350.,450.,500.,890.,320.\}$$

$$med\_smooth(data,3)$$

$$\begin{bmatrix} \text{"x"} & \text{"y"} \\ 1. & \text{"}\square\text{"} \\ 2. & 200. \\ 3. & 250. \\ 4. & 350. \\ 5. & 450. \\ 6. & 500. \\ 7. & 500. \\ 8. & \text{"}\square\text{"} \end{bmatrix}$$

Done

Contact
lazymath2024@gmail.com

Seasonal Data

Determines the seasonal averages, seasonal indices, deseasonalised data, and the least square line fit of the deasonalised data. Rounding for each calculation step can be specified using the appropriate syntax.

Syntax

Case 1: Exact values

season(Data)

Where, *Data* represents the matrix containing the data, with each row representing one cycle and each column representing one period.

Note

In SACs and exams, you will have to round your answers at each stage. This case would be useful for checking your answers rather than obtaining the answers.

Case 2: Rounded values

season({"Data", Round_1, Round_2, Round_3, Round_4})

Where,

"Data" represents a string containing the name of the variable used to store the data

Round_1 represents the number of decimal places to round the average of each cycle to

Round_2 represents the number of decimal places to round the seasonal indices to

Round_3 represents the number of decimal places to round the average of the seasonal indices to

Round_4 represents the number of decimal places to round the deseasonilised data to

Note

All of the above must be inputted in sequence as a list

Example

The sales data for a clothing store was tracked quarterly for three years.

Year	2025				2026				2027			
Quarter	1	2	3	4	1	2	3	4	1	2	3	4
Sales	82	57	42	43	88	59	48	50	97	65	52	55

a. Calculate the sales average for each quarter. Give your answer correct to two decimal places.

b. Calculate the seasonal indices for each sale. Give your answer correct to three decimal places.

c. Calculate the average of the seasonal indices for each sale. Give your answer correct to two decimal places.

d. Deseasonalise the data. Give your answer correct to the nearest whole number.

e. Determine the least squares line fit for the deseasonalised data.

Case 1

data $\begin{bmatrix} 82. & 57. & 42. & 43. \\ 88. & 59. & 48. & 50. \\ 97. & 65. & 52. & 55. \end{bmatrix}$

season(data)

▸Season UDF:

▸Find averages for each cycle:

$$\begin{bmatrix} \text{"Season"} & 1. & 2. & 3. & 4. & \text{"Avg"} \\ \text{"Cycle 1."} & 82. & 57. & 42. & 43. & 56. \\ \text{"Cycle 2."} & 88. & 59. & 48. & 50. & 61.25 \\ \text{"Cycle 3."} & 97. & 65. & 52. & 55. & 67.25 \end{bmatrix}$$

▸Find indicies and take their average:

$$\begin{bmatrix} \text{"Season"} & 1. & 2. & 3. & 4. \\ \text{"Cycle 1."} & 1.4643 & 1.0179 & 0.75 & 0.76786 \\ \text{"Cycle 2."} & 1.4367 & 0.96327 & 0.78367 & 0.81633 \\ \text{"Cycle 3."} & 1.4424 & 0.96654 & 0.77323 & 0.81784 \\ \text{"Avg"} & 1.4478 & 0.98256 & 0.76897 & 0.80068 \end{bmatrix}$$

▸Deseasonalise the data:

$$\begin{bmatrix} \text{"Season"} & 1. & 2. & 3. & 4. \\ \text{"Cycle 1."} & 56.638 & 58.012 & 54.619 & 53.705 \\ \text{"Cycle 2."} & 60.782 & 60.048 & 62.421 & 62.447 \\ \text{"Cycle 3."} & 66.998 & 66.154 & 67.623 & 68.692 \end{bmatrix}$$

▸Find LSR fit of deseasonalised data:

y=1.3066x+53.019

Note: This may **not** provide the answers the marker will be looking for since it uses the exact value at each stage rather than the rounded values.

Contact
lazymath2024@gmail.com

Case 2

data

$$\begin{bmatrix} 82. & 57. & 42. & 43. \\ 88. & 59. & 48. & 50. \\ 97. & 65. & 52. & 55. \end{bmatrix}$$

season({"data",2,3,2,0})

▶Season UDF:

▶Find averages for each cycle:

"Season"	1.	2.	3.	4.	"Avg"
"Cycle 1."	82.	57.	42.	43.	56.
"Cycle 2."	88.	59.	48.	50.	61.25
"Cycle 3."	97.	65.	52.	55.	67.25

▶Find indicies and take their average:

"Season"	1.	2.	3.	4.
"Cycle 1."	1.464	1.018	0.75	0.768
"Cycle 2."	1.437	0.963	0.784	0.816
"Cycle 3."	1.442	0.967	0.773	0.818
"Avg"	1.45	0.98	0.77	0.8

▶Deseasonalise the data:

"Season"	1.	2.	3.	4.
"Cycle 1."	57.	58.	55.	54.
"Cycle 2."	61.	60.	62.	63.
"Cycle 3."	67.	66.	68.	69.

▶Find LSR fit of deseasonalised data:

y=1.3007x+53.212

Significant Figures

Rounds an input number to a specific number of significant figures.

Syntax

sig_fig(Number, *SF*)

Where,

Number represents the number to round

SF represents the number of significant figures to round the number to

Example

Round the number 14.520010 to five significant figures.

$$sig\_fig(14.52001,5) \qquad \text{"14.520"}$$

Contact
lazymath2024@gmail.com