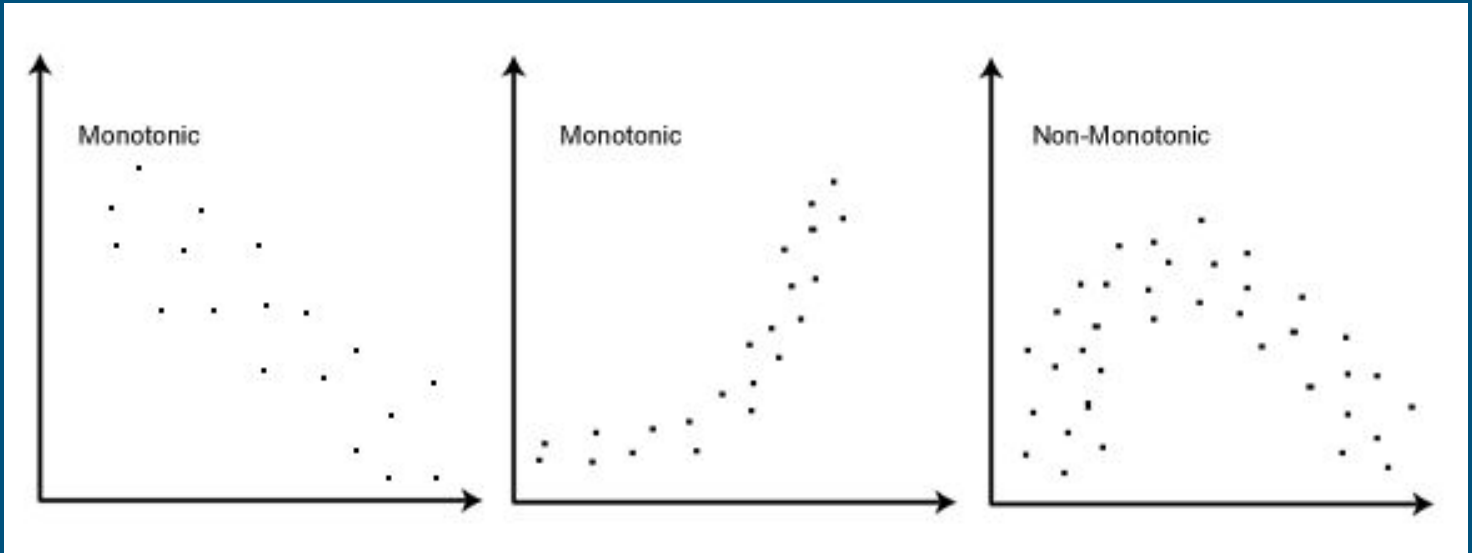
Spearman's Rank Correlation

Katy, Daiwik, Rachel

What is Spearman's Rank Correlation?

Used to show correlation of two variables with a monotonic relationship



Variables

Ordinal - a categorical variable that can be ranked in ascending/descending order

Example: low income, middle income, high income

Interval - numerical data that are on a continuum

Example: -4°F, 55°F, 28°F

Ratio - a type of interval data; at quantity 0, there is nothing of that variable

Example: 9 mL, 0 mL, 13 mL

Why Test is Needed

- Find the strength of the correlation between two variables
- Finds whether the relationship is positive or negative

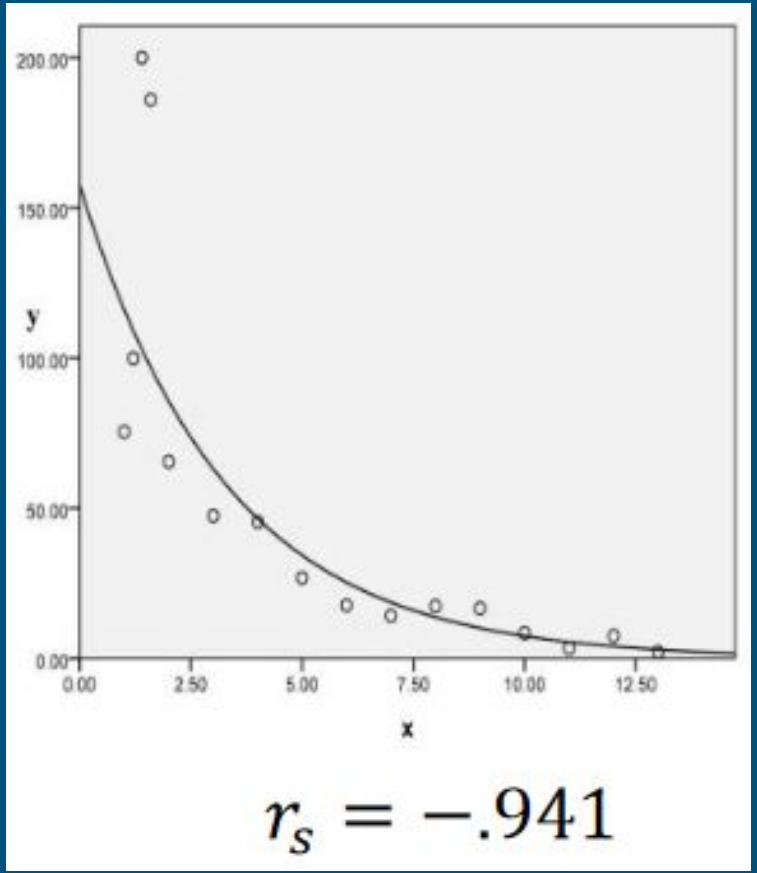
Spearman's ρ (r_s) vs. Pearson's r

- Can increase at changing rates (monotonic, doesn't have to be linear)
- Not required to have a normal distribution
- Uses ranked values

- Increases at the same rate (linear)
- Normal distribution
- Doesn't use ranked values

Correlation Coefficient

- Correlation coefficient is a value between -1 and +1
  - 0.00 - ±0.19 “very weak”
  - ±0.20 - ±0.39 “weak”
  - ±0.40 - ±0.59 “moderate”
  - ±0.60 - ±0.79 “strong”
  - ±0.80 - ±1.00 “very strong”
- Correlation is based on ranked values rather than raw data



Statistically Significant P-Value

- Calculations with ρ to determine statistical significance
- P-value < 0.05 - statistically significant

- Student’s t-test
- df = n - 2

$$t = \rho \sqrt{\frac{n-2}{1-\rho^2}}$$

Ranking Data

English (mark)	Maths (mark)	Rank (English)	Rank (maths)
56	66	9	4
75	70	3	2
45	40	10	10
71	60	4	7
61	65	6.5	5
64	56	5	9
58	59	8	8
80	77	1	1
76	67	2	3
61	63	6.5	6

- Give the highest score of rank 1 and lowest score last rank
  - Last rank is the number of terms. In this case, it is 10.
- If there are two identical values, then average their ranks
  - In this case, English Grade 61 is repeated.
  - Their ranks are averaged (7 + 6)/2

Calculations of r_s or ρ (rho)

For ranked data without ties:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

For ranked data with ties:

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

Example Data & Calculations

Does a student's Math grade correlate to their English grade?

English (mark)	Maths (mark)	Rank (English)	Rank (maths)
56	66	9	4
75	70	3	2
45	40	10	10
71	60	4	7
62	65	6	5
64	56	5	9
58	59	8	8
80	77	1	1
76	67	2	3
61	63	7	6

Null Hypothesis: There is no monotonic relationship between Math and English grades.

Alternative Hypothesis: There is a monotonic relationship between Math and English grades.

Calculation for ρ

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

1. Calculate Σd^2

$$\Sigma d^2 = (9-4)^2 + (3-2)^2 + (10-10)^2 + (4-7)^2 + (6-5)^2 + (5-9)^2 + (8-8)^2 + (1-1)^2 + (2-3)^2 + (7-6)^2$$

$$\Sigma d^2 = 54$$

2. Substitute Σd^2 and n = 10 (sample size) into the equation to find ρ

$$\rho = 1 - \frac{6\Sigma d^2}{n\left(n^2-1\right)} \qquad \rho = 1 - \frac{324}{990}$$

$$\rho = 1 - \frac{6 \times 54}{10\left(10^2-1\right)} \qquad \rho = 0.67$$

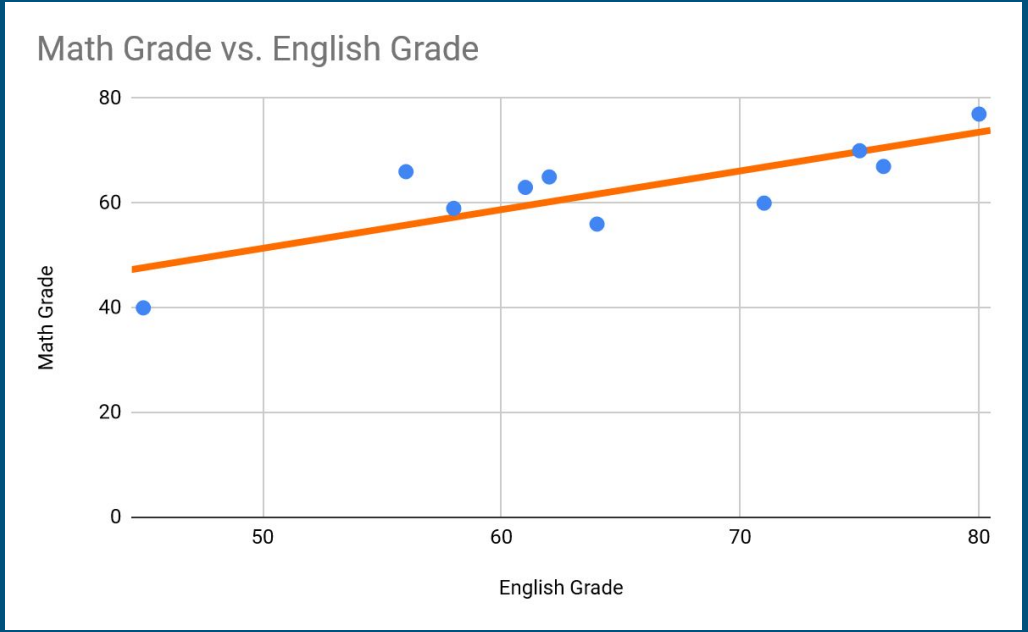
Analysis and Significance Test

Spearman Correlation: +0.67

P-value: 0.033

What does this mean?

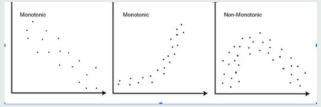
- Indicates strong positive relationship
- Statistically significant
- Rejects null hypothesis



Infographic

SPEARMAN'S RANK CORRELATION

This statistical test shows the correlation strength between two monotonic variables.



RANKING DATA

English (mark)	Maths (mark)	Rank (English)	Rank (maths)
56	66	9	4
75	70	3	2
45	40	10	10
71	60	4	7
62	65	6	5
64	56	5	9
58	59	8	8
80	77	1	1

Give the highest score the rank 1
Give the lowest score the highest rank (# of terms)
If there are two identical values, then average their ranks

EQUATION FOR RANKED DATA WITHOUT TIES

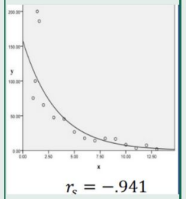
$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

d = difference in paired ranks
n = number of cases

EQUATION FOR RANKED DATA WITH TIES

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

i = paired score (tied scores)



CORRELATION RATIO

- Outputs a value between -1 and +1:
  - **0.00 - ±0.19** "very weak"
  - **±0.20 - ±0.39** "weak"
  - **±0.40 - ±0.59** "moderate"
  - **±0.60 - ±0.79** "strong"
  - **±0.80 - ±1.00** "very strong"
- Tested on ranked data rather than raw values.
- Positive ratio = positive relationship
- Negative ratio = negative relationship

INFORMATION SOURCES

Clearly explained: Pearson V/S Spearman Correlation Coefficient | by Juhi Ramzai | Towards Data Science. (n.d.). Retrieved February 18, 2021, from https://towardsdatascience.com/clearly-explained-pearson-v-s-spearman-correlation-coefficient-ada2f473b8

Spearman's Rank-Order Correlation—A guide to when to use it, what it does and what the assumptions are. (n.d.). Retrieved February 18, 2021, from https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-statistical-guide.php

Spearmans.pdf. (n.d.). Retrieved February 18, 2021, from https://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf

References

A comparison of the Pearson and Spearman correlation methods. (n.d.). [Mtbconcept]. Retrieved February 18, 2021, from https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/modeling-statistics/regression/supporting-topics/basics/a-comparison-of-the-pearson-and-spearman-correlation-methods/

Clearly explained: Pearson V/S Spearman Correlation Coefficient | by Juhi Ramzai | Towards Data Science. (n.d.). Retrieved February 18, 2021, from https://towardsdatascience.com/clearly-explained-pearson-v-s-spearman-correlation-coefficient-ada2f473b8

Foundation, C.-12. (n.d.). *Scatter Plots*. Retrieved February 18, 2021, from https://www.ck12.org/statistics/scatter-plots/lesson/Scatter-Plots-MSM7/

Spearman Rank Correlations—The Ultimate Guide. (n.d.). Retrieved February 18, 2021, from https://www.spss-tutorials.com/spearman-rank-correlation/

Spearman's Rank Calculator: Rs, p-value, Scatter Graph and Conclusion. (n.d.). Retrieved February 18, 2021, from https://geographyfieldwork.com/SpearmansRankCalculator.html#:~:text=In%20geography%20we%20work%20generally,they%20show%20a%20true%20relationship

Spearman's Rank Correlation Coefficient. (n.d.). Retrieved February 18, 2021, from https://geographyfieldwork.com/SpearmansRank.htm#:~:text=The%20Spearman%27s%20Rank%20Correlation%20Coefficient,Museum%20in%20El%20Raval%2C%20Barcelona

Spearman's Rank-Order Correlation—A guide to when to use it, what it does and what the assumptions are. (n.d.). Retrieved February 18, 2021, from https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-statistical-guide.php

Spearmans.pdf. (n.d.). Retrieved February 18, 2021, from https://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf

Any Questions?