

Complete Solutions

Covariance and Correlation

Ratiomera Statistics

These complete solutions use the same identifiers and order as the Exercise Sheet. Intermediate values are retained until the stated rounding step, so small differences caused by earlier rounding are acceptable where noted. All settings, values, data, and software outputs are constructed teaching material; they are not empirical findings.

Part I: Theory

A02: Third Variables and the Boundary of Causal Claims

T04-A02-V01: Library visits and course marks

Identify the issue

A plausible diagram places study motivation and course workload before both focal variables, with one arrow toward weekly library visits and another toward course marks.

The mechanism is that motivated learners may both visit more often and study more effectively.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would measure relevant study habits prospectively and avoid calling the adjusted association a causal effect.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V02: Park access and neighborhood trust

Identify the issue

A plausible diagram places neighborhood income and public investment before both focal variables, with one arrow toward nearby park space and another toward average trust.

The mechanism is that better-resourced areas may receive both parks and services that support trust.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would compare changes around planned park openings with comparable neighborhoods and resident-level measurements.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V03: Remote work and volunteer hours

Identify the issue

A plausible diagram places schedule flexibility and occupation before both focal variables, with one arrow toward remote-work frequency and another toward volunteer hours.

The mechanism is that flexible occupations can permit both remote work and volunteering.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would measure the time order and use a credible policy comparison or randomized access when feasible.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V04: Museum membership and cultural participation

Identify the issue

A plausible diagram places prior interest in cultural activities before both focal variables, with one arrow toward museum membership and another toward cultural-event attendance.

The mechanism is that existing interest can lead to both membership and attendance.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would measure baseline interest and distinguish prediction from a membership effect.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V05: Cycling lanes and retail visits

Identify the issue

A plausible diagram places street centrality and store mix before both focal variables, with one arrow toward protected cycling infrastructure and another toward weekend shop visits.

The mechanism is that central streets can attract infrastructure and visitors for separate reasons.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would use repeated before-and-after counts on treated and comparable untreated streets.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V06: Caption availability and course completion

Identify the issue

A plausible diagram places course production quality and subject area before both focal variables, with one arrow toward caption availability and another toward course completion.

The mechanism is that better-funded courses may provide both captions and stronger learner support.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would compare otherwise similar courses or randomize a phased caption rollout.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V07: Garden plots and wellbeing

Identify the issue

A plausible diagram places baseline health and social connection before both focal variables, with one arrow toward community-garden participation and another toward wellbeing.

The mechanism is that healthier or more connected residents may be more likely to obtain and use a plot.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would measure wellbeing before allocation and exploit randomized plot offers if demand exceeds supply.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V08: Transit use and daily activity

Identify the issue

A plausible diagram places urban density and car access before both focal variables, with one arrow toward public-transit use and another toward daily step count.

The mechanism is that dense neighborhoods may encourage both transit use and walking.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would collect neighborhood and mobility information and use longitudinal or policy-change evidence.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V09: Music lessons and memory

Identify the issue

A plausible diagram places family resources and educational support before both focal variables, with one arrow toward years of music lessons and another toward memory score.

The mechanism is that resources can affect access to lessons and cognitive opportunities.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would measure baseline scores and relevant family conditions, or evaluate randomized lesson access.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

T04-A02-V10: Online discussion and exam performance

Identify the issue

A plausible diagram places prior understanding and engagement before both focal variables, with one arrow toward discussion-board posts and another toward exam score.

The mechanism is that engaged students may both post more and prepare more thoroughly.

Reason through the evidence

The reported coefficient describes the sample's linear association between the two measured variables.

It does not identify the direction of influence, rule out common causes, or establish a causal effect.

State the conclusion and its limits

A stronger approach would measure prior achievement and use a designed encouragement or other credible comparison.

Even after measured third variables are adjusted for, unmeasured confounding and design limitations remain, so the conclusion must stay model-based rather than causal unless the design supplies causal identification.

A07: Diagnosing Association and Causal Claims

T04-A07-V01: Puzzle apps and navigation

Identify the issue

The observational units and measurement level are individual survey respondents, with both variables recorded at the individual level.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include self-selection, self-reported measurement, uncertain time order, and differences in age or prior spatial interest.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would measure navigation objectively at baseline and follow-up and, when feasible, randomly assign access to comparable activities.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V02: Parks and individual trust

Identify the issue

The observational units and measurement level are districts, with both measurements aggregated to the district level before the claim jumps to individual residents.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include the ecological jump from district averages to individuals, income, density, safety, and public investment.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would collect representative resident-level repeated measurements around park changes and use comparable districts.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V03: Blue notebooks and memory

Identify the issue

The observational units and measurement level are individual online-poll respondents, with notebook choice and reported grade recorded at the individual level.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include volunteer sampling, unverified outcomes, study habits, school differences, and reverse timing.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would randomly assign otherwise equivalent notebook colors and assess a common objective outcome.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V04: Remote work and volunteering**Identify the issue**

The observational units and measurement level are individual employees, with work arrangement and volunteer hours recorded at the individual level.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include occupation, schedule flexibility, commuting time, income, caregiving, and employee preference.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would measure outcomes before and after a credible policy change with comparable employees.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V05: Morning markets and diet**Identify the issue**

The observational units and measurement level are towns, with dietary variety aggregated across households before the claim shifts to every household.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include town-level aggregation, wealth, agriculture, transport, tourism, and unmeasured market attendance.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would sample households directly and compare changes in matched towns before and after openings.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V06: Digital reminders and attendance**Identify the issue**

The observational units and measurement level are scheduled appointments or patients summarized within two clinic time periods, making this a before-and-after aggregate comparison.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include patient mix, staffing, appointment types, season, scheduling, and changes in the recorded denominator.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would use randomized reminder assignment or a concurrent comparable clinic with consistent outcome definitions.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V07: Poetry clubs and empathy

Identify the issue

The observational units and measurement level are individual club members and separate online respondents, with individual-level scores drawn from incomparable samples.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include self-selection, incomparable samples, self-report measurement, and pre-existing literary interest.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would measure baseline empathy in comparable groups and use randomized invitations if participation cannot be assigned.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V08: Cycle lanes and retail traffic

Identify the issue

The observational units and measurement level are streets, with visitor counts aggregated at the street level.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include street centrality, shop mix, events, parking, pedestrian access, and pre-existing visitor levels.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would collect repeated counts before and after installation on treated and matched untreated streets.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V09: Garden plots and wellbeing

Identify the issue

The observational units and measurement level are individual plot renters and waiting-list members, with participation and wellbeing recorded at the individual level.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include health, mobility, motivation, social ties, selective allocation, and lack of baseline measurement.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would measure validated outcomes before allocation and use a lottery when scarce plots permit one.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

T04-A07-V10: Captioned courses and completion

Identify the issue

The observational units and measurement level are courses, with completion aggregated across learners at the course level.

The reported comparison supports, at most, an association at that recorded level.

Reason through the evidence

Key threats include course topic, budget, instructional quality, support, learner composition, and course-level aggregation.

A defensible statement would say that the measured variables were associated in this particular sample or aggregate dataset, without saying that one produced the other or that every individual follows the group pattern.

State the conclusion and its limits

A stronger evaluation would compare matched courses or a randomized phased caption introduction using learner-level outcomes.

The revised design must also report its sampling frame, missingness, measurement quality, time order, and any remaining limits.

No correlation coefficient alone can convert an observational comparison into a causal effect.

Part II: Calculator Practice

A01: Pearson Correlation and Hidden Nonlinearity

T04-A01-V01: Background sound and proofreading errors

Reason before calculating

The means are $\bar{x} = 4.000$ and $\bar{y} = 6.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(84.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of background-sound setting in coded units and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V02: Room temperature and concentration loss

Reason before calculating

The means are $\bar{x} = 22.000$ and $\bar{y} = 10.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(336.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of room temperature in degrees Celsius and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V03: Daily practice and fatigue at both extremes

Reason before calculating

The means are $\bar{x} = 4.000$ and $\bar{y} = 6.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(84.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of daily practice time in hours and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V04: Museum crowding and visit discomfort

Reason before calculating

The means are $\bar{x} = 4.000$ and $\bar{y} = 14.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(756.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of museum-crowding score and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V05: Text size and reading difficulty

Reason before calculating

The means are $\bar{x} = 12.000$ and $\bar{y} = 10.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(336.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of text size in typographic points and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V06: Notification frequency and task disruption

Reason before calculating

The means are $\bar{x} = 3.000$ and $\bar{y} = 6.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(84.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of notifications per hour and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V07: Route complexity and navigation strain

Reason before calculating

The means are $\bar{x} = 5.000$ and $\bar{y} = 14.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(756.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of route-complexity score and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V08: Session length and restlessness

Reason before calculating

The means are $\bar{x} = 6.000$ and $\bar{y} = 6.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(84.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of session length in ten-minute units and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V09: Archive humidity and preservation risk

Reason before calculating

The means are $\bar{x} = 5.000$ and $\bar{y} = 10.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(336.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of relative humidity in ten-percentage-point units and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

T04-A01-V10: Lighting intensity and visual discomfort

Reason before calculating

The means are $\bar{x} = 4.000$ and $\bar{y} = 10.000$.

Symmetric low and high X values have the same Y values, so the positive and negative cross-products cancel:
 $\sum (x_i - \bar{x})(y_i - \bar{y}) = 0.000$.

Work through the calculation

The denominator is $\sqrt{28.000(336.000)}$, giving $r = 0.0000$.

A scatterplot forms a U shape: the outcome is lowest near the middle of lighting intensity in coded units and rises toward either extreme.

Interpret and check the result

Pearson's r summarizes a straight-line pattern only.

Here it is near zero because there is no overall upward or downward line, not because the visible nonlinear relationship is absent.

A03: Testing a Directional Population Correlation

T04-A03-V01: Practice time and reasoning score

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = 4.1000/[2.00(5.00)] = 0.4100$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = 0.4100\sqrt{(20-2)/(1-(0.4100)^2)} = 1.9071$ with $df = 18$.

It crosses the directional boundary 1.7341, so we reject H_0 at 5%.

The sample association is positive: larger values of weekly practice hours tend to accompany larger values of reasoning score in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.41$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V02: Search time and archive accuracy

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = -4.3200/[4.00(3.00)] = -0.3600$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = -0.3600\sqrt{(25-2)/(1-(-0.3600)^2)} = -1.8506$ with $df = 23$.

It crosses the directional boundary -1.7139 , so we reject H_0 at 5%.

The sample association is negative: larger values of search time tend to accompany smaller values of accuracy score in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.36$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V03: Reading time and comprehension

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = 5.6000/[2.50(8.00)] = 0.2800$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = 0.2800\sqrt{(30-2)/(1-(0.2800)^2)} = 1.5434$ with $df = 28$.

It does not cross the directional boundary 1.7011 , so we fail to reject H_0 at 5%.

The sample association is positive: larger values of reading time tend to accompany larger values of comprehension score in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.28$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V04: Navigation errors and route confidence

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = -5.9400/[3.00(6.00)] = -0.3300$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = -0.3300\sqrt{(35-2)/(1-(-0.3300)^2)} = -2.0082$ with $df = 33$.

It crosses the directional boundary -1.6924 , so we reject H_0 at 5%.

The sample association is negative: larger values of navigation-error count tend to accompany smaller values of route-confidence score in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.33$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V05: Workshop attendance and concept score

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = 3.6000/[1.50(10.00)] = 0.2400$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = 0.2400\sqrt{(40-2)/(1-(0.2400)^2)} = 1.5240$ with $df = 38$.

It does not cross the directional boundary 1.6860, so we fail to reject H_0 at 5%.

The sample association is positive: larger values of sessions attended tend to accompany larger values of concept score in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.24$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V06: Notification count and focus rating

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = -2.2000/[5.00(2.00)] = -0.2200$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = -0.2200\sqrt{(45-2)/(1-(-0.2200)^2)} = -1.4789$ with $df = 43$.

It does not cross the directional boundary -1.6811, so we fail to reject H_0 at 5%.

The sample association is negative: larger values of daily notifications tend to accompany smaller values of focus rating in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.22$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V07: Museum visits and historical knowledge

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = 2.8000/[2.00(7.00)] = 0.2000$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = 0.2000\sqrt{(50-2)/(1-(0.2000)^2)} = 1.4142$ with $df = 48$.

It does not cross the directional boundary 1.6772, so we fail to reject H_0 at 5%.

The sample association is positive: larger values of museum visits tend to accompany larger values of knowledge score in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.20$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V08: Response delay and satisfaction

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = -2.8800/[4.00(4.00)] = -0.1800$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = -0.1800\sqrt{(60-2)/(1-(-0.1800)^2)} = -1.3936$ with $df = 58$.

It does not cross the directional boundary -1.6716 , so we fail to reject H_0 at 5%.

The sample association is negative: larger values of response delay tend to accompany smaller values of satisfaction score in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.18$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V09: Practice sets and confidence

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = 2.5500/[3.00(5.00)] = 0.1700$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = 0.1700\sqrt{(80-2)/(1-(0.1700)^2)} = 1.5236$ with $df = 78$.

It does not cross the directional boundary 1.6646 , so we fail to reject H_0 at 5%.

The sample association is positive: larger values of practice sets completed tend to accompany larger values of confidence score in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.17$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

T04-A03-V10: Route familiarity and travel time

Reason before calculating

First, $r = s_{xy}/(s_x s_y) = -2.1000/[6.00(2.50)] = -0.1400$.

The null is $H_0 : \rho = 0$.

Work through the calculation

The test statistic is $t = -0.1400\sqrt{(100-2)/(1-(-0.1400)^2)} = -1.3997$ with $df = 98$.

It does not cross the directional boundary -1.6606 , so we fail to reject H_0 at 5%.

The sample association is negative: larger values of route-familiarity score tend to accompany smaller values of travel time in a linear sense.

Interpret and check the result

Its magnitude is $|r| = 0.14$, which describes the observed linear association.

Statistical evidence depends on both r and n , while practical importance depends on the variables, consequences, precision, and design.

A significant coefficient would still not establish causation.

A04: Spearman Correlation from Separate Ranks

T04-A04-V01: Reading frequency and vocabulary rank

Set up the calculation

The rank differences are $d = (0, -1, 1, -1, 1, 0)$, so $\sum d_i^2 = 4$.

Work through the calculation

With $n = 6$, $r_s = 1 - 6(4)/[6(6^2 - 1)] = 0.8857$.

The positive sign means that cases with higher ranks on one variable tend to have higher ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.8857$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V02: Search efficiency and accuracy rank

Set up the calculation

The rank differences are $d = (-1, 1, -1, 1, -1, 1, 0)$, so $\sum d_i^2 = 6$.

Work through the calculation

With $n = 7$, $r_s = 1 - 6(6)/[7(7^2 - 1)] = 0.8929$.

The positive sign means that cases with higher ranks on one variable tend to have higher ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.8929$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V03: Practice rank and reasoning rank

Set up the calculation

The rank differences are $d = (0, 0, -1, 1, -1, 1)$, so $\sum d_i^2 = 4$.

Work through the calculation

With $n = 6$, $r_s = 1 - 6(4)/[6(6^2 - 1)] = 0.8857$.

The positive sign means that cases with higher ranks on one variable tend to have higher ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.8857$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V04: Attendance rank and confidence rank

Set up the calculation

The rank differences are $d = (0, -1, 1, 0, -1, -1, 2)$, so $\sum d_i^2 = 8$.

Work through the calculation

With $n = 7$, $r_s = 1 - 6(8)/[7(7^2 - 1)] = 0.8571$.

The positive sign means that cases with higher ranks on one variable tend to have higher ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.8571$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V05: Navigation skill and error rank

Set up the calculation

The rank differences are $d = (-5, -3, 0, 0, 3, 5)$, so $\sum d_i^2 = 68$.

Work through the calculation

With $n = 6$, $r_s = 1 - 6(68)/[6(6^2 - 1)] = -0.9429$.

The negative sign means that cases with higher ranks on one variable tend to have lower ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.9429$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V06: Response-delay and satisfaction rank

Set up the calculation

The rank differences are $d = (-6, -3, -3, 0, 2, 5, 5)$, so $\sum d_i^2 = 108$.

Work through the calculation

With $n = 7$, $r_s = 1 - 6(108)/[7(7^2 - 1)] = -0.9286$.

The negative sign means that cases with higher ranks on one variable tend to have lower ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.9286$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V07: Museum visits and knowledge rank

Set up the calculation

The rank differences are $d = (-1, 1, 0, -1, 1, 0)$, so $\sum d_i^2 = 4$.

Work through the calculation

With $n = 6$, $r_s = 1 - 6(4)/[6(6^2 - 1)] = 0.8857$.

The positive sign means that cases with higher ranks on one variable tend to have higher ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.8857$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V08: Practice regularity and retention rank

Set up the calculation

The rank differences are $d = (0, 0, -1, 1, 0, -1, 1)$, so $\sum d_i^2 = 4$.

Work through the calculation

With $n = 7$, $r_s = 1 - 6(4)/[7(7^2 - 1)] = 0.9286$.

The positive sign means that cases with higher ranks on one variable tend to have higher ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.9286$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V09: Task switching and focus rank

Set up the calculation

The rank differences are $d = (-4, -4, -1, 1, 4, 4)$, so $\sum d_i^2 = 66$.

Work through the calculation

With $n = 6$, $r_s = 1 - 6(66)/[6(6^2 - 1)] = -0.8857$.

The negative sign means that cases with higher ranks on one variable tend to have lower ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.8857$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

T04-A04-V10: Archive experience and search rank

Set up the calculation

The rank differences are $d = (0, -1, 1, -1, 1, -1, 1)$, so $\sum d_i^2 = 6$.

Work through the calculation

With $n = 7$, $r_s = 1 - 6(6)/[7(7^2 - 1)] = 0.8929$.

The positive sign means that cases with higher ranks on one variable tend to have higher ranks on the other.

Interpret and check the result

The magnitude $|r_s| = 0.8929$ describes how consistently the ordering follows one monotonic direction.

Because ranks replace original values, the result does not describe an original-unit slope.

A05: Pearson and Spearman under Monotonic Curvature

T04-A05-V01: Practice hours and fluent-recall score

Reason before calculating

Plotting the coordinates (1, 2), (2, 3), (3, 5), (4, 8), (5, 12), (6, 17), (7, 23) gives an increasing curve.

Every increase in X accompanies an increase in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 10.0000$, $S_{xy} = 98.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 364.0000$.

Thus $r = S_{xy}/\sqrt{S_{xx}S_{yy}} = 98.0000/\sqrt{28.0000(364.0000)} = 0.9707$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (1, 2, 3, 4, 5, 6, 7)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is 28.0000, and their squared-deviation sums are 28.0000 and 28.0000, so $r_s = 28.0000/\sqrt{28.0000(28.0000)} = 1.0000$.

Spearman reaches 1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V02: Archive experience and retrieval speed

Reason before calculating

Plotting the coordinates (1, 70), (2, 53), (3, 41), (4, 32), (5, 26), (6, 22), (7, 19) gives a decreasing curve.

Every increase in X accompanies a decrease in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 37.5714$, $S_{xy} = -230.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 2053.7143$.

Thus $r = S_{xy}/\sqrt{S_{xx}S_{yy}} = -230.0000/\sqrt{28.0000(2053.7143)} = -0.9591$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (7, 6, 5, 4, 3, 2, 1)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is -28.0000, and their squared-deviation sums are 28.0000 and 28.0000, so $r_s = -28.0000/\sqrt{28.0000(28.0000)} = -1.0000$.

Spearman reaches -1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V03: Museum visits and knowledge score

Reason before calculating

Plotting the coordinates (1, 10), (2, 13), (3, 17), (4, 22), (5, 28), (6, 35), (7, 43) gives an increasing curve.

Every increase in X accompanies an increase in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 24.0000$, $S_{xy} = 154.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 868.0000$.

Thus $r = S_{xy}/\sqrt{S_{xx}S_{yy}} = 154.0000/\sqrt{28.0000(868.0000)} = 0.9878$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (1, 2, 3, 4, 5, 6, 7)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is 28.0000, and their squared-deviation sums are 28.0000 and 28.0000, so $r_s = 28.0000/\sqrt{28.0000(28.0000)} = 1.0000$.

Spearman reaches 1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V04: Route familiarity and navigation time

Reason before calculating

Plotting the coordinates (1, 65), (2, 48), (3, 37), (4, 29), (5, 24), (6, 20), (7, 17) gives a decreasing curve.

Every increase in X accompanies a decrease in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 34.2857$, $S_{xy} = -213.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 1775.4286$.

Thus $r = S_{xy} / \sqrt{S_{xx}S_{yy}} = -213.0000 / \sqrt{28.0000(1775.4286)} = -0.9553$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (7, 6, 5, 4, 3, 2, 1)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is -28.0000 , and their squared-deviation sums are 28.0000 and 28.0000 , so $r_s = -28.0000 / \sqrt{28.0000(28.0000)} = -1.0000$.

Spearman reaches -1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V05: Practice sets and confidence score

Reason before calculating

Plotting the coordinates $(1, 20)$, $(2, 22)$, $(3, 25)$, $(4, 29)$, $(5, 34)$, $(6, 40)$, $(7, 47)$ gives an increasing curve.

Every increase in X accompanies an increase in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 31.0000$, $S_{xy} = 126.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 588.0000$.

Thus $r = S_{xy} / \sqrt{S_{xx}S_{yy}} = 126.0000 / \sqrt{28.0000(588.0000)} = 0.9820$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (1, 2, 3, 4, 5, 6, 7)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is 28.0000 , and their squared-deviation sums are 28.0000 and 28.0000 , so $r_s = 28.0000 / \sqrt{28.0000(28.0000)} = 1.0000$.

Spearman reaches 1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V06: Notification load and focus score

Reason before calculating

Plotting the coordinates $(1, 82)$, $(2, 69)$, $(3, 59)$, $(4, 51)$, $(5, 45)$, $(6, 40)$, $(7, 36)$ gives a decreasing curve.

Every increase in X accompanies a decrease in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 54.5714$, $S_{xy} = -210.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 1641.7143$.

Thus $r = S_{xy} / \sqrt{S_{xx}S_{yy}} = -210.0000 / \sqrt{28.0000(1641.7143)} = -0.9795$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (7, 6, 5, 4, 3, 2, 1)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is -28.0000 , and their squared-deviation sums are 28.0000 and 28.0000 , so $r_s = -28.0000 / \sqrt{28.0000(28.0000)} = -1.0000$.

Spearman reaches -1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V07: Reading sessions and vocabulary score

Reason before calculating

Plotting the coordinates (1, 30), (2, 31), (3, 34), (4, 39), (5, 46), (6, 55), (7, 66) gives an increasing curve.

Every increase in X accompanies an increase in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 43.0000$, $S_{xy} = 168.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 1092.0000$.

Thus $r = S_{xy} / \sqrt{S_{xx}S_{yy}} = 168.0000 / \sqrt{28.0000(1092.0000)} = 0.9608$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (1, 2, 3, 4, 5, 6, 7)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is 28.0000, and their squared-deviation sums are 28.0000 and 28.0000, so $r_s = 28.0000 / \sqrt{28.0000(28.0000)} = 1.0000$.

Spearman reaches 1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V08: Search attempts and remaining errors

Reason before calculating

Plotting the coordinates (1, 24), (2, 18), (3, 14), (4, 11), (5, 9), (6, 8), (7, 7) gives a decreasing curve.

Every increase in X accompanies a decrease in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 13.0000$, $S_{xy} = -76.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 228.0000$.

Thus $r = S_{xy} / \sqrt{S_{xx}S_{yy}} = -76.0000 / \sqrt{28.0000(228.0000)} = -0.9512$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (7, 6, 5, 4, 3, 2, 1)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is -28.0000, and their squared-deviation sums are 28.0000 and 28.0000, so $r_s = -28.0000 / \sqrt{28.0000(28.0000)} = -1.0000$.

Spearman reaches -1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V09: Workshop sessions and reasoning score

Reason before calculating

Plotting the coordinates (1, 40), (2, 43), (3, 47), (4, 52), (5, 58), (6, 65), (7, 73) gives an increasing curve.

Every increase in X accompanies an increase in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 54.0000$, $S_{xy} = 154.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 868.0000$.

Thus $r = S_{xy} / \sqrt{S_{xx}S_{yy}} = 154.0000 / \sqrt{28.0000(868.0000)} = 0.9878$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (1, 2, 3, 4, 5, 6, 7)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is 28.0000, and their squared-deviation sums are 28.0000 and 28.0000, so $r_s = 28.0000/\sqrt{28.0000(28.0000)} = 1.0000$.

Spearman reaches 1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

T04-A05-V10: Route complexity and completion rate

Reason before calculating

Plotting the coordinates (1, 95), (2, 83), (3, 72), (4, 62), (5, 53), (6, 45), (7, 38) gives a decreasing curve.

Every increase in X accompanies a decrease in Y , but the amount of change is not constant.

Work through the calculation

For the original values, $\bar{x} = 4.0000$, $\bar{y} = 64.0000$, $S_{xy} = -266.0000$, $S_{xx} = 28.0000$, and $S_{yy} = 2548.0000$.

Thus $r = S_{xy}/\sqrt{S_{xx}S_{yy}} = -266.0000/\sqrt{28.0000(2548.0000)} = -0.9959$.

The rank vectors are $R_X = (1, 2, 3, 4, 5, 6, 7)$ and $R_Y = (7, 6, 5, 4, 3, 2, 1)$, both with mean rank 4.0.

Interpret and check the result

Their cross-product sum is -28.0000, and their squared-deviation sums are 28.0000 and 28.0000, so $r_s = -28.0000/\sqrt{28.0000(28.0000)} = -1.0000$.

Spearman reaches -1 because the ordering is perfectly monotonic.

Pearson is smaller in magnitude because it asks how closely the points follow a straight line and therefore responds to the changing slope.

A06: Spearman Correlation with Ties and Positive Transformations

T04-A06-V01: Practice days and confidence

Reason before calculating

Average ranks give $R_X = (1.5, 1.5, 3, 4.5, 4.5, 6, 7.5, 7.5)$ and $R_Y = (1, 2.5, 2.5, 4, 5.5, 5.5, 7, 8)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is 39.0000, while the two squared-deviation sums are 40.5000 and 41.0000.

Therefore $r_s = 39.0000/\sqrt{40.5000(41.0000)} = 0.9571$.

The transformation $Y^* = 5 + 2Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (1, 2.5, 2.5, 4, 5.5, 5.5, 7, 8)$ and $r_s^* = 0.9571$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V02: Museum visits and knowledge

Reason before calculating

Average ranks give $R_X = (1, 2.5, 2.5, 4, 5.5, 5.5, 7.5, 7.5)$ and $R_Y = (1, 2, 3.5, 3.5, 5, 6.5, 6.5, 8)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is 38.7500, while the two squared-deviation sums are 40.5000 and 41.0000.

Therefore $r_s = 38.7500 / \sqrt{40.5000(41.0000)} = 0.9509$.

The transformation $Y^* = 7 + 3Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (1, 2, 3.5, 3.5, 5, 6.5, 6.5, 8)$ and $r_s^* = 0.9509$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V03: Archive shifts and retrieval skill

Reason before calculating

Average ranks give $R_X = (1, 3, 3, 3, 5, 6.5, 6.5, 8)$ and $R_Y = (1, 2.5, 2.5, 4, 5.5, 5.5, 7, 8)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is 38.7500, while the two squared-deviation sums are 39.5000 and 41.0000.

Therefore $r_s = 38.7500 / \sqrt{39.5000(41.0000)} = 0.9629$.

The transformation $Y^* = 4 + 2Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (1, 2.5, 2.5, 4, 5.5, 5.5, 7, 8)$ and $r_s^* = 0.9629$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V04: Reading sessions and recall

Reason before calculating

Average ranks give $R_X = (1.5, 1.5, 3, 4.5, 4.5, 6, 7.5, 7.5)$ and $R_Y = (1.5, 1.5, 3, 4, 5.5, 5.5, 7, 8)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is 39.7500, while the two squared-deviation sums are 40.5000 and 41.0000.

Therefore $r_s = 39.7500 / \sqrt{40.5000(41.0000)} = 0.9755$.

The transformation $Y^* = 3 + 4Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (1.5, 1.5, 3, 4, 5.5, 5.5, 7, 8)$ and $r_s^* = 0.9755$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V05: Route attempts and accuracy

Reason before calculating

Average ranks give $R_X = (1.5, 1.5, 3.5, 3.5, 5, 6, 7.5, 7.5)$ and $R_Y = (1, 2.5, 2.5, 4, 5.5, 5.5, 7, 8)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is 39.0000, while the two squared-deviation sums are 40.5000 and 41.0000.

Therefore $r_s = 39.0000 / \sqrt{40.5000(41.0000)} = 0.9571$.

The transformation $Y^* = 6 + 2Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (1, 2.5, 2.5, 4, 5.5, 5.5, 7, 8)$ and $r_s^* = 0.9571$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V06: Workshop attendance and reasoning

Reason before calculating

Average ranks give $R_X = (1, 2.5, 2.5, 4.5, 4.5, 6, 7.5, 7.5)$ and $R_Y = (1, 2, 3.5, 3.5, 5, 6.5, 6.5, 8)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is 38.7500, while the two squared-deviation sums are 40.5000 and 41.0000.

Therefore $r_s = 38.7500 / \sqrt{40.5000(41.0000)} = 0.9509$.

The transformation $Y^* = 8 + 3Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (1, 2, 3.5, 3.5, 5, 6.5, 6.5, 8)$ and $r_s^* = 0.9509$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V07: Practice sets and fluency

Reason before calculating

Average ranks give $R_X = (1, 2.5, 2.5, 4, 5.5, 5.5, 7.5, 7.5)$ and $R_Y = (1, 2.5, 2.5, 4, 5, 6.5, 6.5, 8)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is 39.5000, while the two squared-deviation sums are 40.5000 and 41.0000.

Therefore $r_s = 39.5000 / \sqrt{40.5000(41.0000)} = 0.9693$.

The transformation $Y^* = 2 + 5Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (1, 2.5, 2.5, 4, 5, 6.5, 6.5, 8)$ and $r_s^* = 0.9693$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V08: Search experience and speed

Reason before calculating

Average ranks give $R_X = (1.5, 1.5, 3, 4.5, 4.5, 6.5, 6.5, 8)$ and $R_Y = (8, 6.5, 6.5, 5, 3.5, 3.5, 2, 1)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is -38.7500 , while the two squared-deviation sums are 40.5000 and 41.0000 .

Therefore $r_s = -38.7500 / \sqrt{40.5000(41.0000)} = -0.9509$.

The transformation $Y^* = 100 + 2Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (8, 6.5, 6.5, 5, 3.5, 3.5, 2, 1)$ and $r_s^* = -0.9509$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V09: Discussion posts and confidence

Reason before calculating

Average ranks give $R_X = (1, 2.5, 2.5, 4, 5.5, 5.5, 7, 8)$ and $R_Y = (1, 2, 3.5, 3.5, 5, 6.5, 6.5, 8)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is 39.5000 , while the two squared-deviation sums are 41.0000 and 41.0000 .

Therefore $r_s = 39.5000 / \sqrt{41.0000(41.0000)} = 0.9634$.

The transformation $Y^* = 9 + 2Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (1, 2, 3.5, 3.5, 5, 6.5, 6.5, 8)$ and $r_s^* = 0.9634$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.

T04-A06-V10: Navigation practice and errors

Reason before calculating

Average ranks give $R_X = (1, 2.5, 2.5, 4.5, 4.5, 6, 7.5, 7.5)$ and $R_Y = (8, 6.5, 6.5, 5, 3.5, 3.5, 2, 1)$.

Both rank columns have mean 4.5.

Work through the calculation

Their cross-product sum is -39.7500 , while the two squared-deviation sums are 40.5000 and 41.0000 .

Therefore $r_s = -39.7500 / \sqrt{40.5000(41.0000)} = -0.9755$.

The transformation $Y^* = 20 + 3Y$ has a positive multiplier, so it preserves every ordering and every tie.

Interpret and check the result

Therefore $R_{Y^*} = (8, 6.5, 6.5, 5, 3.5, 3.5, 2, 1)$ and $r_s^* = -0.9755$, exactly the same coefficient.

Adding a constant changes location and multiplying by a positive constant changes scale, but neither changes ranks.

A negative multiplier would reverse the ordering and therefore reverse the sign of Spearman's coefficient.