|  |
| --- |
| Spearman’s Rank Correlation Coefficient and bike theft |

**Specification links**

AQA

A Level 3.4.2.4 Statistical skills *Inferential and relational statistical techniques to include Spearman’s rank correlation and Chi-square test and the application of significance tests.*

Edexcel

A Level Appendix 1: Geographical skills. *This specification requires students to collect, analyse and interpret such information, and demonstrate the ability to understand and apply suitable analytical approaches for the different information types including, qualitative approaches such as coding and sampling and quantitative approaches such as measures of dispersion, measures of correlation and association from the following statistical tests: t-tests, Spearman’s rank, Chi-squared, Gini Co-efficient, Lorenz curve.*

OCR

A Level Geographical Skills 4.4 Quantitative skills *b) tests of association and significance tests, such as Chi-squared, Spearman’s rank, Mann-Whitney U test and T-test.*

Eduqas

A Level Appendix A Geographical Skills. *2. measures of correlation, including a scatter plot, lines of best fit and Spearman Rank.*

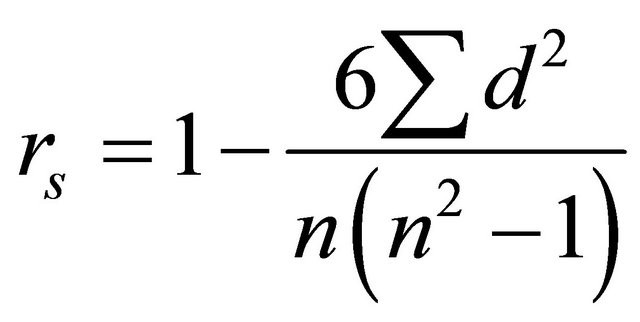
**What is Spearman’s Rank Correlation Coefficient?**

Spearman’s Rank is a statistical method used to test the strength of the relationship between two variables. It uses ranked data to test the relationship and calculates a fixed figure (between –1 and +1) to show the strength of that relationship:

* +1 indicates a perfect positive correlation
* -1 indicates a perfect negative correlation
* 0 shows there is no correlation at all between the two variables

Once the fixed figure has been calculated and a +1, -1 or 0 result is identified the numerical value must be tested statistically to see how significant the result is.

The test can be used for any two sets of data (variables) so long as it is raw data (or percentages or indices) which can be ranked. Below is the equation for Spearman’s Rank Correlation Coefficient (with each part explained on the following page):



*r***s** Spearman’s Rank

∑ the sum of

*d²* difference

*n* number of values

**Test 1 a worked example. Spearman’s Rank Correlation Coefficient: are areas with high levels of cycling more susceptible to bike theft?**

This is a worked example to investigate whether cycling hotspots, such as the city of Oxford, have higher levels of bike theft. This statistical equation will calculate if there is a relationship (correlation) between the number of cyclists and the level of bike theft in an area.

Evidence at the national level shows that cycle theft is often higher in counties where more people cycle. This suggests that because there are greater opportunities for bike theft individuals suffer more from bike crime.

This resource uses bike theft data from [www.data.police.uk](http://www.data.police.uk), the government [Walking and cycling statistics](https://www.gov.uk/government/statistics/walking-and-cycling-statistics-england-2019) PDF 2019 report for England and specifically the [CW0302: Proportion of adults that cycle, by frequency, purpose and local authority: England](https://www.gov.uk/government/statistical-data-sets/walking-and-cycling-statistics-cw) dataset from 2018-2019, based on the National Travel Survey and Active Lives Survey. Population data for each Police Force area was extracted from the 2020 ONS dataset [Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland](https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland).

In order to identify individuals who either commute to work or cycle regularly for leisure the *Total % of adults who cycle for at least 5 x p/w* has been selected for column 4 in Table 1 below.

|  |  |  |  |
| --- | --- | --- | --- |
| January 2011 to December 2020 inclusive - Residential Population | | | |
|  | | Variable 1 | Variable 2 |
| Police Constabulary | Bike theft over a 10-year period | Bike theft per annum | The number of people who cycle x 5 times a week per annum |
| Metropolitan Police Service | 149929 | 14993 | 394327 |
| Greater Manchester Police | 27065 | 2707 | 82234 |
| West Yorkshire Police | 17672 | 1767 | 48981 |
| West Midlands Police | 23385 | 2339 | 124614 |
| Lancashire Constabulary | 13577 | 1358 | 26835 |
| South Yorkshire Police | 9506 | 951 | 25362 |
| Kent Police | 12727 | 1273 | 37957 |
| Hampshire Constabulary | 27003 | 2700 | 42858 |
| Thames Valley Police | 41965 | 4197 | 85089 |
| Northumbria Police | 15012 | 1501 | 34166 |
| Essex Police | 17277 | 1728 | 31272 |
| Avon and Somerset Constabulary | 24313 | 2431 | 59254 |
| Merseyside Police | 13991 | 1399 | 35747 |
| Sussex Police | 18290 | 1829 | 41634 |
| Police Service of Northern Ireland | 6388 | 639 | No data |
| South Wales Police | 15344 | 1534 | No data |
| Devon & Cornwall Police | 6858 | 686 | 50271 |
| Nottinghamshire Police | 14033 | 1403 | 24846 |
| West Mercia Police | 8212 | 821 | 34756 |
| Staffordshire Police | 7630 | 763 | 12313 |
| Derbyshire Constabulary | 7577 | 758 | 20211 |
| Hertfordshire Constabulary | 11080 | 1108 | 26169 |
| Cheshire Constabulary | 9963 | 996 | 26494 |
| Humberside Police | 15709 | 1571 | 26666 |
| Surrey Police | 9370 | 937 | 49045 |
| Cleveland Police | 7275 | 937 | 3428 |
| Leicestershire Police | 14581 | 728 | 21184 |
| Northamptonshire Police | 6761 | 1458 | 16572 |
| Cambridgeshire Constabulary | 29233 | 676 | 62086 |
| Norfolk Constabulary | 9657 | 2923 | 38125 |
| Dorset Police | 9663 | 966 | 11733 |
| Durham Constabulary | 4043 | 966 | 7951 |
| North Yorkshire Police | 9563 | 404 | 14215 |
| North Wales Police | 3833 | 956 | No data |
| Bedfordshire Police | 7620 | 383 | 11481 |
| Lincolnshire Police | 8702 | 762 | 24359 |
| Gwent Police | 2999 | 870 | No data |
| Suffolk Constabulary | 8317 | 300 | 23601 |
| Gloucestershire Constabulary | 7177 | 832 | 2493318 |
| Wiltshire Police | 6089 | 718 | 10500 |
| Warwickshire Police | 5139 | 609 | 15026 |
| Cumbria Constabulary | 2380 | 514 | 14000 |
| Dyfed-Powys Police | 1275 | 238 |  |
| City of London Police | 2994 | 128 | 944978 |

Table 1

This is a correlation exercise using the two variables. These two variables are two separate data sources: Bike theft (variable 1) and % of adults who cycle for at least 5 x p/w (variable 2). Ultimately you are investigating the relationship between these two variables i.e. is there a relationship between high levels of bike theft and areas with a high level of cycling. Spearman’s Rank Correlation Coefficient statistically proveswhether there is or is not a link (and how strong it is).

You will finish with a *r***s** (Spearman’s Rank) value which will be used to identify whether correlation is:

* Nearer to 0, indicating the correlation is weak (either weak positive or weak negative).
* Close to ± 1, showing there is a strong positive correlation.
* Close to -1, revealing there is a strong negative correlation.

Finally, it is important to verify the result as meaningful. This proves the result was not calculated by chance and is revealed by a confidence level and a significance table.

**Step 1**

Before this statistical test is applied you formulate a **null hypothesis**. This is a theory which says there is no statistical relationship or significance between variables. This could be:

“There will be **no** significant relationship between bike theft and the number of people who cycle”.

**Step 2**

It is first important to rearrange the data into a new format, to allow for ease of analysis (see Table 2 below).

The Police Service of Northern Ireland and the 4 Police Forces of Wales have been removed because the *Participation in walking and cycling* survey did not cover these regions.

Rank the data by filling out the (R¹) and (R²) columns. Rank both sets of data from lowest to highest i.e., the **lowest** value gets rank **1**, the 2nd lowest gets rank 2 and so on.

To help you get started ranks 1, 2 and 3 for both variables have been filled in.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Number | Police Constabulary | Bike theft p/a | Rank (R¹) | The number of people who cycle 5 times a week p/a | Rank (R²) | *d*  R¹ - R² | *d²* |
| 1 | Metropolitan Police Service | 14993 |  | 394327 |  |  |  |
| 2 | Greater Manchester Police | 2707 |  | 82234 |  |  |  |
| 3 | West Yorkshire Police | 1767 |  | 48981 |  |  |  |
| 4 | West Midlands Police | 2339 |  | 124614 |  |  |  |
| 5 | Lancashire Constabulary | 1358 |  | 26835 |  |  |  |
| 6 | South Yorkshire Police | 951 |  | 25362 |  |  |  |
| 7 | Kent Police | 1273 |  | 37957 |  |  |  |
| 8 | Hampshire Constabulary | 2700 |  | 42858 |  |  |  |
| 9 | Thames Valley Police | 4197 |  | 85089 |  |  |  |
| 10 | Northumbria Police | 1501 |  | 34166 |  |  |  |
| 11 | Essex Police | 1728 |  | 31272 |  |  |  |
| 12 | Avon and Somerset Constabulary | 2431 |  | 59254 |  |  |  |
| 13 | Merseyside Police | 1399 |  | 35747 |  |  |  |
| 14 | Sussex Police | 1829 |  | 41634 |  |  |  |
| 15 | Devon & Cornwall Police | 686 |  | 50271 |  |  |  |
| 16 | Nottinghamshire Police | 1403 |  | 24846 |  |  |  |
| 17 | West Mercia Police | 821 |  | 34756 |  |  |  |
| 18 | Staffordshire Police | 763 |  | 12313 |  |  |  |
| 19 | Derbyshire Constabulary | 758 |  | 20211 |  |  |  |
| 20 | Hertfordshire Constabulary | 1108 |  | 26169 |  |  |  |
| 21 | Cheshire Constabulary | 996 |  | 26494 |  |  |  |
| 22 | Humberside Police | 1571 |  | 26666 |  |  |  |
| 23 | Surrey Police | 937 |  | 49045 |  |  |  |
| 24 | Cleveland Police | 937 |  | 3428 | 1 |  |  |
| 25 | Leicestershire Police | 728 |  | 21184 |  |  |  |
| 26 | Northamptonshire Police | 1458 |  | 16572 |  |  |  |
| 27 | Cambridgeshire Constabulary | 676 |  | 62086 |  |  |  |
| 28 | Norfolk Constabulary | 2923 |  | 38125 |  |  |  |
| 29 | Dorset Police | 966 |  | 11733 |  |  |  |
| 30 | Durham Constabulary | 966 |  | 7951 | 2 |  |  |
| 31 | North Yorkshire Police | 404 |  | 14215 |  |  |  |
| 32 | Bedfordshire Police | 383 | 3 | 11481 |  |  |  |
| 33 | Lincolnshire Police | 762 |  | 24359 |  |  |  |
| 34 | Suffolk Constabulary | 300 | 2 | 23601 |  |  |  |
| 35 | Gloucestershire Constabulary | 832 |  | 2493318 |  |  |  |
| 36 | Wiltshire Police | 718 |  | 10500 | 3 |  |  |
| 37 | Warwickshire Police | 609 |  | 15026 |  |  |  |
| 38 | Cumbria Constabulary | 514 |  | 14000 |  |  |  |
| 39 | City of London Police | 128 | 1 | 944978 |  |  |  |
| ∑ | | | | | | |  |

Table 2

If you have two numbers that are the same you cannot give them the same whole number, neither can you rank one over the other — so you simply find the average of the two ranks. Be careful to “jump” to the next rank correctly. Mrs Spicer gives an excellent explanation [in this video](https://www.youtube.com/watch?v=5PnyKjxWXDA) (from 3 minutes in) on how to do this.

**Step 1**

Fill out column 7 of Table 2 by working out R¹ - R² (subtracting each piece of R¹ data from R²) to calculate *d* i.e., the difference.

**Step 2**

For column 8 square the data from step 1.

**Step 3**

Sum up the final column and fill in the ∑ value in the bottom right cell. Go back to the equation. This figure is your ∑*d²* part of the equation. The n value is the number of values, in this case, there are 39.

At this stage it is advisable to rewrite out the Spearman’s Rank Correlation Coefficient equation and populate it these figures. When you rewrite the final equation (or type it into a calculator) remember to start with 1- before the fraction.

**Step 4**

Is your *r***s** your value a positive or negative result? How strong is the correlation: is it close to ± 1 or -1? On page 4 in [A Guide to Spearman’s Rank](https://www.rgs.org/CMSPages/GetFile.aspx?nodeguid=882169d2-8f96-4c55-84f5-fbb7614870e9&lang=en-GB) by the Royal Geographical Society (with IBG) there is a useful spectrum to visualise the continuum from ± 1 to -1.

**Step 5**

How likely is it that this result was calculated by chance? To work this final bit out you need to consult a significance table (Table 3 below). Normally, you look for the column shown as 95% confidence (or sometimes listed in reverse as 0.5 level or 5% (i.e., there is a 5% likelihood that this result was calculated by chance).

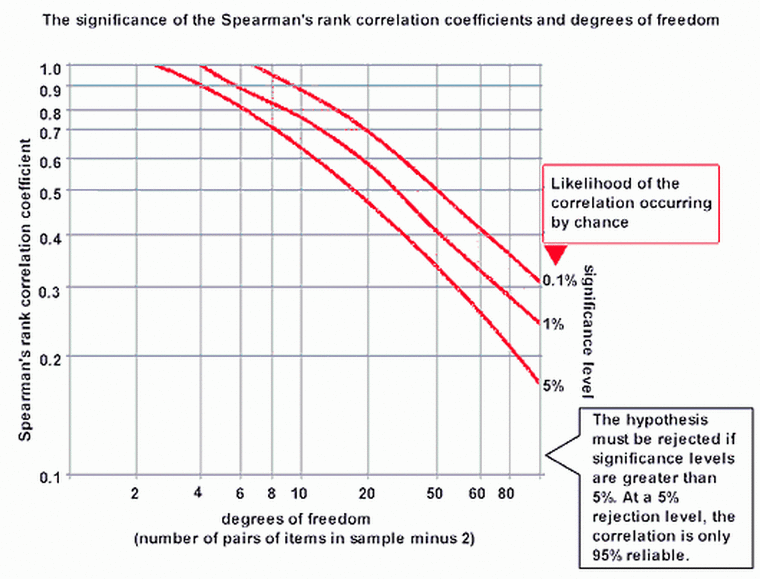
If your result comes out as higher than the number shown for the number of data sets used (39) then you can be confident the result is not down to random chance. 95% is a high rate of confidence, and it is generally considered good enough.

|  |  |  |  |
| --- | --- | --- | --- |
| Number of pairs of data (n) | 0.10  10% chance 90% confident | 0.05  5% chance 95% confident | 0.01  1% chance 99% confident |
| 5 | 0.9 | 1 | 1 |
| 6 | 0.829 | 0.886 | 1 |
| 7 | 0.714 | 0.786 | 0.929 |
| 8 | 0.643 | 0.738 | 0.881 |
| 9 | 0.6 | 0.700 | 0.833 |
| 10 | 0.564 | 0.648 | 0.794 |
| 12 | 0.503 | 0.587 | 0.727 |
| 14 | 0.464 | 0.538 | 0.679 |
| 16 | 0.429 | 0.503 | 0.635 |
| 18 | 0.401 | 0.472 | 0.600 |
| 20 | 0.380 | 0.447 | 0.570 |
| 22 | 0.361 | 0.425 | 0.544 |
| 24 | 0.344 | 0.406 | 0.521 |
| 26 | 0.331 | 0.390 | 0.501 |
| 28 | 0.317 | 0.375 | 0.483 |
| 30 | 0.306 | 0.362 | 0.467 |
| 31 | 0.301 | 0.356 | 0.459 |
| 32 | 0.296 | 0.350 | 0.452 |
| 33 | 0.291 | 0.345 | 0.446 |
| 34 | 0.287 | 0.340 | 0.439 |
| 35 | 0.283 | 0.335 | 0.433 |
| 36 | 0.279 | 0.330 | 0.427 |
| 37 | 0.275 | 0.325 | 0.421 |
| 38 | 0.271 | 0.321 | 0.415 |
| 39 | 0.267 | 0.317 | 0.410 |

Table 3 <http://webspace.ship.edu/pgmarr/geo441/tables/spearman%20ranked%20correlation%20table.pdf>

**Step 6**

Sometimes you may be presented with a graph to calculate the significance of a result. It is important that you can work this out from both a table (such as Table 3) *and* by using a degrees of freedom graph, like the one below (Graph 1).



Graph 1 A degrees of freedom graph

**Test 2 Spearman’s Rank: is there correlation between high rates of bike theft and areas with high levels of cycling for leisure?**

The CSEW offer a breakdown of the data set *Total % of adults who cycle for at least 5 x p/w* (used in Test 1) into those that cycle for work, and those that cycle for leisure.

This is a second Spearman’s Rank Correlation Coefficient task. Return to the [CW0302: Proportion of adults that cycle, by frequency, purpose and local authority: England](https://www.gov.uk/government/statistical-data-sets/walking-and-cycling-statistics-cw) dataset from 2018-2019 this time to complete a Spearman’s Rank Correlation Coefficient statistical test using the data column *Cycling for leisure five times per week*.

Is correlation between high rates of bike theft and areas with high levels of cycling for leisure? Use Table 4 on the next page to calculate the correlation. Repeat and work through the same steps 1-6 as outlined in Test 1.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Number | Police Constabulary | Bike theft p/a | Rank (R¹) | The number of people who cycle **for** **Leisure** 5 times a week p/a | Rank (R²) | *d*  R¹ - R² | *d²* |
| 1 | Metropolitan Police Service | 14993 |  | 71695 |  |  |  |
| 2 | Greater Manchester Police | 2707 |  | 34028 |  |  |  |
| 3 | West Yorkshire Police | 1767 |  | 23324 |  |  |  |
| 4 | West Midlands Police | 2339 |  | 53406 |  |  |  |
| 5 | Lancashire Constabulary | 1358 |  | 19516 |  |  |  |
| 6 | South Yorkshire Police | 951 |  | 7045 |  |  |  |
| 7 | Kent Police | 1273 |  | 11070 |  |  |  |
| 8 | Hampshire Constabulary | 2700 |  | 15207 |  |  |  |
| 9 | Thames Valley Police | 4197 |  | 24655 |  |  |  |
| 10 | Northumbria Police | 1501 |  | 9725 |  |  |  |
| 11 | Essex Police | 1728 |  | 11913 |  |  |  |
| 12 | Avon and Somerset Constabulary | 2431 |  | 11954 |  |  |  |
| 13 | Merseyside Police | 1399 |  | 8579 |  |  |  |
| 14 | Sussex Police | 1829 |  | 13961 |  |  |  |
| 15 | Devon & Cornwall Police | 686 |  | 25343 |  |  |  |
| 16 | Nottinghamshire Police | 1403 |  | 9110 |  |  |  |
| 17 | West Mercia Police | 821 |  | 15427 |  |  |  |
| 18 | Staffordshire Police | 763 |  | 4397 |  |  |  |
| 19 | Derbyshire Constabulary | 758 |  | 7432 |  |  |  |
| 20 | Hertfordshire Constabulary | 1108 |  | 10705 |  |  |  |
| 21 | Cheshire Constabulary | 996 |  | 1592 |  |  |  |
| 22 | Humberside Police | 1571 |  | 8892 |  |  |  |
| 23 | Surrey Police | 937 |  | 16747 |  |  |  |
| 24 | Cleveland Police | 937 |  | 1234 |  |  |  |
| 25 | Leicestershire Police | 728 |  | 6355 |  |  |  |
| 26 | Northamptonshire Police | 1458 |  | 6779 |  |  |  |
| 27 | Cambridgeshire Constabulary | 676 |  | 9149 |  |  |  |
| 28 | Norfolk Constabulary | 2923 |  | 15431 |  |  |  |
| 29 | Dorset Police | 966 |  | 4920 |  |  |  |
| 30 | Durham Constabulary | 966 |  | 3710 |  |  |  |
| 31 | North Yorkshire Police | 404 |  | 6798 |  |  |  |
| 32 | Bedfordshire Police | 383 |  | 2908 |  |  |  |
| 33 | Lincolnshire Police | 762 |  | 8373 |  |  |  |
| 34 | Suffolk Constabulary | 300 |  | 8374 |  |  |  |
| 35 | Gloucestershire Constabulary | 832 |  | 6768 |  |  |  |
| 36 | Wiltshire Police | 718 |  | 7500 |  |  |  |
| 37 | Warwickshire Police | 609 |  | 4623 |  |  |  |
| 38 | Cumbria Constabulary | 514 |  | 7500 |  |  |  |
| 39 | City of London Police | 128 |  | 9 |  |  |  |
| ∑ | | | | | | |  |

Table 7

**Further work**

* Mrs Spicer on Spearman’s Rank Correlation Coefficient, measuring correlation between earthquakes and the number of deaths <https://www.youtube.com/watch?v=5PnyKjxWXDA>
* A guide to Spearman’s Rank from the Royal Geographical Society (with IBG) <https://www.rgs.org/CMSPages/GetFile.aspx?nodeguid=882169d2-8f96-4c55-84f5-fbb7614870e9&lang=en-GB>
* Best Song Ever <https://www.stem.org.uk/resources/elibrary/resource/36017/best-song-ever>
* How safe is your bike? [https://www.met.police.uk/cp/crime-prevention/theft-of-a-bike/how-safe-is-your-bike/](https://www.met.police.uk/cp/crime-prevention/theft-of-a-bicycle/how-safe-is-your-bike/)

**Answers**

1. Below is the complete calculation for Spearman’s Rank Correlation Coefficient: are areas with high levels of cycling more susceptible to bike theft?

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Number | Police Constabulary | Bike theft p/a | Rank (R¹) | The number of people who cycle 5 times a week p/a | Rank (R²) | *d*  R¹ - R² | *d²* |
| 1 | Metropolitan Police Service | 14993 | 39 | 394327 | 37 | 2 | 4 |
| 2 | Greater Manchester Police | 2707 | 36 | 82234 | 34 | 2 | 4 |
| 3 | West Yorkshire Police | 1767 | 31 | 48981 | 29 | 2 | 4 |
| 4 | West Midlands Police | 2339 | 33 | 124614 | 36 | -3 | 9 |
| 5 | Lancashire Constabulary | 1358 | 24 | 26835 | 20 | 4 | 16 |
| 6 | South Yorkshire Police | 951 | 18 | 25362 | 16 | 2 | 4 |
| 7 | Kent Police | 1273 | 23 | 37957 | 25 | -2 | 4 |
| 8 | Hampshire Constabulary | 2700 | 35 | 42858 | 28 | 7 | 49 |
| 9 | Thames Valley Police | 4197 | 38 | 85089 | 35 | 3 | 9 |
| 10 | Northumbria Police | 1501 | 28 | 34166 | 22 | 6 | 36 |
| 11 | Essex Police | 1728 | 30 | 31272 | 21 | 9 | 81 |
| 12 | Avon and Somerset Constabulary | 2431 | 34 | 59254 | 32 | 2 | 4 |
| 13 | Merseyside Police | 1399 | 25 | 35747 | 24 | 1 | 1 |
| 14 | Sussex Police | 1829 | 32 | 41634 | 27 | 5 | 25 |
| 15 | Devon & Cornwall Police | 686 | 8 | 50271 | 31 | -23 | 529 |
| 16 | Nottinghamshire Police | 1403 | 26 | 24846 | 15 | 11 | 121 |
| 17 | West Mercia Police | 821 | 14 | 34756 | 23 | -9 | 81 |
| 18 | Staffordshire Police | 763 | 13 | 12313 | 6 | 7 | 49 |
| 19 | Derbyshire Constabulary | 758 | 11 | 20211 | 11 | 0 | 0 |
| 20 | Hertfordshire Constabulary | 1108 | 22 | 26169 | 17 | 5 | 25 |
| 21 | Cheshire Constabulary | 996 | 21 | 26494 | 18 | 3 | 9 |
| 22 | Humberside Police | 1571 | 29 | 26666 | 19 | 10 | 100 |
| 23 | Surrey Police | 937 | 16.5 | 49045 | 30 | -13.5 | 182.25 |
| 24 | Cleveland Police | 937 | 16.5 | 3428 | 1 | 15.5 | 240.25 |
| 25 | Leicestershire Police | 728 | 10 | 21184 | 12 | -2 | 4 |
| 26 | Northamptonshire Police | 1458 | 27 | 16572 | 10 | 17 | 289 |
| 27 | Cambridgeshire Constabulary | 676 | 7 | 62086 | 33 | -26 | 676 |
| 28 | Norfolk Constabulary | 2923 | 37 | 38125 | 26 | 11 | 121 |
| 29 | Dorset Police | 966 | 19.5 | 11733 | 5 | 14.5 | 210.25 |
| 30 | Durham Constabulary | 966 | 19.5 | 7951 | 2 | 17.5 | 306.25 |
| 31 | North Yorkshire Police | 404 | 4 | 14215 | 8 | -4 | 16 |
| 32 | Bedfordshire Police | 383 | 3 | 11481 | 4 | -1 | 1 |
| 33 | Lincolnshire Police | 762 | 12 | 24359 | 14 | -2 | 4 |
| 34 | Suffolk Constabulary | 300 | 2 | 23601 | 13 | -11 | 121 |
| 35 | Gloucestershire Constabulary | 832 | 15 | 2493318 | 39 | -24 | 576 |
| 36 | Wiltshire Police | 718 | 9 | 10500 | 3 | 6 | 36 |
| 37 | Warwickshire Police | 609 | 6 | 15026 | 9 | -3 | 9 |
| 38 | Cumbria Constabulary | 514 | 5 | 14000 | 7 | -2 | 4 |
| 39 | City of London Police | 128 | 1 | 944978 | 38 | -37 | 1369 |
| ∑ | | | | | | | 5329 |

*r***s** (Spearman’s Rank Correlation Coefficient) final equation is 31974 ÷ 57798 = 0.5532025329596. This is a positive result showing that there *is* correlation between the amount people cycle and how often their bikes are stolen.

Using the significance table, we can say this result did not occur by chance with 99% confidence.

1. Below is the second complete calculation for Spearman’s Rank Correlation Coefficient: is there correlation between high rates of bike theft and areas with high levels of cycling for leisure?

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Number | Police Constabulary | Bike theft p/a | Rank (R¹) | The number of people who cycle **for** **Leisure** 5 times a week p/a | Rank (R²) | *d*  R¹ - R² | *d²* |
| 1 | Metropolitan Police Service | 14993 | 39 | 71695 | 39 | 0 | 0 |
| 2 | Greater Manchester Police | 2707 | 36 | 34028 | 37 | -1 | 1 |
| 3 | West Yorkshire Police | 1767 | 31 | 23324 | 34 | -3 | 9 |
| 4 | West Midlands Police | 2339 | 33 | 53406 | 38 | -5 | 25 |
| 5 | Lancashire Constabulary | 1358 | 24 | 19516 | 33 | -9 | 81 |
| 6 | South Yorkshire Police | 951 | 18 | 7045 | 13 | 5 | 25 |
| 7 | Kent Police | 1273 | 23 | 11070 | 25 | -2 | 4 |
| 8 | Hampshire Constabulary | 2700 | 35 | 15207 | 29 | 6 | 36 |
| 9 | Thames Valley Police | 4197 | 38 | 24655 | 35 | 3 | 9 |
| 10 | Northumbria Police | 1501 | 28 | 9725 | 23 | 5 | 25 |
| 11 | Essex Police | 1728 | 30 | 11913 | 26 | 4 | 16 |
| 12 | Avon and Somerset Constabulary | 2431 | 34 | 11954 | 27 | 7 | 49 |
| 13 | Merseyside Police | 1399 | 25 | 8579 | 19 | 6 | 36 |
| 14 | Sussex Police | 1829 | 32 | 13961 | 28 | 4 | 16 |
| 15 | Devon & Cornwall Police | 686 | 8 | 25343 | 36 | -28 | 784 |
| 16 | Nottinghamshire Police | 1403 | 26 | 9110 | 21 | 5 | 25 |
| 17 | West Mercia Police | 821 | 14 | 15427 | 30 | -16 | 256 |
| 18 | Staffordshire Police | 763 | 13 | 4397 | 6 | 7 | 49 |
| 19 | Derbyshire Constabulary | 758 | 11 | 7432 | 14 | -3 | 9 |
| 20 | Hertfordshire Constabulary | 1108 | 22 | 10705 | 24 | -2 | 4 |
| 21 | Cheshire Constabulary | 996 | 21 | 1592 | 3 | 18 | 324 |
| 22 | Humberside Police | 1571 | 29 | 8892 | 20 | 9 | 81 |
| 23 | Surrey Police | 937 | 16.5 | 16747 | 32 | -15.5 | 240.25 |
| 24 | Cleveland Police | 937 | 16.5 | 1234 | 2 | 14.5 | 210.25 |
| 25 | Leicestershire Police | 728 | 10 | 6355 | 9 | 1 | 1 |
| 26 | Northamptonshire Police | 1458 | 27 | 6779 | 11 | 16 | 256 |
| 27 | Cambridgeshire Constabulary | 676 | 7 | 9149 | 22 | -15 | 225 |
| 28 | Norfolk Constabulary | 2923 | 37 | 15431 | 31 | 6 | 36 |
| 29 | Dorset Police | 966 | 19.5 | 4920 | 8 | 11.5 | 132.25 |
| 30 | Durham Constabulary | 966 | 19.5 | 3710 | 5 | 14.5 | 210.25 |
| 31 | North Yorkshire Police | 404 | 4 | 6798 | 12 | -8 | 64 |
| 32 | Bedfordshire Police | 383 | 3 | 2908 | 4 | -1 | 1 |
| 33 | Lincolnshire Police | 762 | 12 | 8373 | 17 | -5 | 25 |
| 34 | Suffolk Constabulary | 300 | 2 | 8374 | 18 | -16 | 256 |
| 35 | Gloucestershire Constabulary | 832 | 15 | 6768 | 10 | 5 | 25 |
| 36 | Wiltshire Police | 718 | 9 | 7500 | 15.5 | -6.5 | 42.25 |
| 37 | Warwickshire Police | 609 | 6 | 4623 | 7 | -1 | 1 |
| 38 | Cumbria Constabulary | 514 | 5 | 7500 | 15.5 | -10.5 | 110.25 |
| 39 | City of London Police | 128 | 1 | 9 | 1 | 0 | 0 |
| ∑ | | | | | | | 3699.5 |

*r***s** (Spearman’s Rank Correlation Coefficient) final equation for this second test is 3699.5 ÷ 57798 = 0.0640074051005. This proves there *is* positive correlation between areas that have high levels of cycling **for leisure** and high level of bike theft — but it is not a strong relationship between the two data sets.

This result is interesting because we can deduce that the result from test 1 (for all cycling, 5 times a week) might have a stronger positive correlation due to cyclists who are commuters (i.e., not cycling for leisure). In order to know for certain you will need to do a third Spearman’s Rank Correlation Coefficient test on *Cycling for travel 5 times a week*, again using the [CW0302: Proportion of adults that cycle, by frequency, purpose and local authority: England](https://www.gov.uk/government/statistical-data-sets/walking-and-cycling-statistics-cw) dataset from 2018-2019.

However, after using the significance table, we cannot be confident in this result as the value fails to meet the 90% confidence threshold (0.267).

