

# [Statistics: correlation between retail online sales and total retail sales](https://assignbuster.com/statistics-correlation-between-retail-online-sales-and-total-retail-sales/)

Statistics: correlation between retail online sales and total retail sales To calculate the correlation coefficient for quarterly retail online sales and corresponding total retail sales for the recent few years, we use the formula:   
Let Xi be total retail sales, and Yi - retail online sales for period i. N is total number of pairs of observations and is equal to total number of periods available, thus N = 19.   
#   
x   
y   
x\*y   
x\*x   
y\*y   
1   
787, 212   
5, 335   
4, 199, 776, 020   
619, 702, 732, 944   
28, 462, 225   
2   
714, 561   
5, 663   
4, 046, 558, 943   
510, 597, 422, 721   
32, 069, 569   
3   
774, 677   
6, 185   
4, 791, 377, 245   
600, 124, 454, 329   
38, 254, 225   
4   
768, 139   
7, 009   
5, 383, 886, 251   
590, 037, 523, 321   
49, 126, 081   
5   
812, 809   
9, 143   
7, 431, 512, 687   
660, 658, 470, 481   
83, 594, 449   
6   
724, 731   
7, 893   
5, 720, 301, 783   
525, 235, 022, 361   
62, 299, 449   
7   
802, 662   
7, 794   
6, 255, 947, 628   
644, 266, 286, 244   
60, 746, 436   
8   
779, 096   
7, 821   
6, 093, 309, 816   
606, 990, 577, 216   
61, 168, 041   
9   
850, 265   
10, 755   
9, 144, 600, 075   
722, 950, 570, 225   
115, 670, 025   
10   
738, 185   
9, 549   
7, 048, 928, 565   
544, 917, 094, 225   
91, 183, 401   
11   
814, 626   
10, 005   
8, 150, 333, 130   
663, 615, 519, 876   
100, 100, 025   
12   
818, 061   
10, 734   
8, 781, 066, 774   
669, 223, 799, 721   
115, 218, 756   
13   
859, 250   
13, 999   
12, 028, 640, 750   
738, 310, 562, 500   
195, 972, 001   
14   
767, 433   
12, 115   
9, 297, 450, 795   
588, 953, 409, 489   
146, 773, 225   
15   
852, 760   
12, 718   
10, 845, 401, 680   
727, 199, 617, 600   
161, 747, 524   
16   
867, 242   
13, 651   
11, 838, 720, 542   
752, 108, 686, 564   
186, 349, 801   
17   
912, 109   
17, 512   
15, 972, 852, 808   
831, 942, 827, 881   
306, 670, 144   
18   
834, 716   
15, 515   
12, 950, 618, 740   
696, 750, 800, 656   
240, 715, 225   
19   
919, 041   
15, 654   
14, 386, 667, 814   
844, 636, 359, 681   
245, 047, 716   
Total   
15, 397, 575   
199, 050   
164, 367, 952, 046   
12, 538, 221, 738, 035   
2, 321, 168, 318   
First, we calculate x\*y, x\*x and y\*y for each pair (xi, yi) and fill out the three last column of the table about. Second, the total for each column is found. Now we have all the numbers necessary to determine the correlation coefficient:   
R = (19\*164, 367, 952, 046 - 15, 397, 575\*199, 050)/ (19\*12, 538, 221, 738, 035 - 15, 397, 575\*15, 397, 575)\*(19\*2, 321, 168, 318 - 199, 050\*199, 050) =   
= 58, 103, 785, 124/ 71, 503, 127, 739 = 0. 813   
The correlation coefficient of 0. 813 indicates a strong positive between linear relationship retail online sales and total retail sales variables (X and Y). Therefore, we would expect for the year characterized by the higher level of total retail sales to have also higher level of online retail sales and vice versa. The relationship between two variables can be described using the linear equation.   
Correlation is widely used in the everyday business situations when it is necessary to determine the presence and strength of the relationship between two variables. For example, correlation between the advertising expenditures or promotional activities in certain area and the total sales in the corresponding period in that area is of interest for the managers. They want to determine how effective is the advertising/promotion and forecast the sales given increase or decrease in advertising.   
However, the correlation values should be treated carefully. It characterizes the strength of the linear relationship; therefore, if the relationship between two given variables is quadratic, logarithmic or exponential, then the correlation coefficient will not adequately reflect how strong the relationship is. Moreover, even the relationship between variables is linear, high positive value of the coefficient does not indicate which of the variables causes the change in the value of another. The situation when both of them are influenced by the third variable and change simultaneously is also possible.   
In a business context, a lot of attention should be paid to the economic relationship between variables. For example if the correlation coefficient between sales figures and the amount of calls received by HR department indicates strong positive relationship, it can be a pure coincidence and should not be automatically considered as means to increase sales.