

- 1 New technologies have vastly improved data collection, generating a deluge of information across different disciplines.
- 2 - This large amount of data provides new opportunities to address unanswered scientific questions, provided we have efficient tools capable of identifying multiple types of underlying patterns.
- 3 - Correlation analysis is an essential statistical technique for discovering relationships between variables [ @pmid:21310971 ].
- 4 - Correlation coefficients are often used in exploratory data mining techniques, such as clustering or community detection algorithms, to compute a similarity value between a pair of objects of interest such as genes [ @pmid:27479844 ] or disease-relevant lifestyle factors [ @doi:10.1073/pnas.1217269109 ].
- 5 - Correlation methods are also used in supervised tasks, for example, for feature selection to improve prediction accuracy [ @pmid:27006077; @pmid:33729976 ].
- 6 - The Pearson correlation coefficient is ubiquitously deployed across application domains and diverse scientific areas.
- 7 - Thus, even minor and significant improvements in these techniques could have enormous consequences in industry and research.

- 1 + The increasing availability of data has opened up new possibilities for scientific exploration.
- 2 + To take advantage of this, we need efficient tools to identify multiple types of relationships between variables.
- 3 + Correlation analysis is a useful statistical technique to uncover such relationships [ @pmid:21310971 ].
- 4 + Correlation coefficients are often used in data mining techniques, such as clustering or community detection, to calculate the similarity between two objects, like genes [ @pmid:27479844 ] or lifestyle factors related to diseases [ @doi:10.1073/pnas.1217269109 ].
- 5 + They are also used in supervised tasks, like feature selection, to boost prediction accuracy [ @pmid:27006077; @pmid:33729976 ].
- 6 + The Pearson correlation coefficient is widely used across many application domains and scientific disciplines.
- 7 + Therefore, even small improvements in this technique can have a huge impact on industry and research.