

Classification of the pre-exam anxiety levels in physically inactive students from 48 countries: application of the K-means clustering machine learning algorithm.

Authors:

Samuel Encarnação | Universidad Autónoma de Madrid (UAM) | Instituto Politécnico de Bragança (IPB) | samuel01.encarnacao@gmail.com

António Miguel Monteiro | Instituto Politécnico de Bragança (IPB) | mmonteiro@ipb.pt

Tatiana Sampaio | Instituto Politécnico de Bragança (IPB) | Centro de Investigação em Desporto, Saúde e Desenvolvimento (CIDESD) | tatiana_sampaio30@hotmail.com

Paula Vaz | Centro de Investigação em Educação Básica (CIEB) | Instituto Politécnico de Bragança (IPB) | paulavaz@ipb.pt

Pedro Miguel Forte | Instituto Superior de Ciências Educativas do Douro | Centro de Investigação em Desporto, Saúde e Desenvolvimento (CIDESD) | pedromiguelforte@gmail.com

Abstract

Physical inactivity is a behavioral factor that is strongly associated with anxiety symptoms in children and adolescents. Anxiety is a disabling condition that reduces academic performance (Kandola et al., 2020).

Keywords

Mental health; K-means clustering; Physical inactivity; Artificial intelligence; Academic performance.

Research question | Objective

Identifying the countries that have the higher incidence of physical inactivity and anxiety in students should be determinant in targeting vulnerable public and building goals to make youths more physically active and prevent anxiety damage to mental health and academic development. The objective of this study was to identify and cluster the percentual levels of pre-exam anxiety of physically inactive students from 48 countries.

Methods

Data from percentual pre-exam anxiety levels from 600,000 scholar students, average age 15-year-old, from 48 countries was recorded. The dataset was obtained from the PISA study 2018 (OECD, 2018). A K-means clustering machine learning algorithm was applied in Python (TM), programming language, to identify the clusters of countries based on their percentual pre-exam anxiety levels (Ikotun et al., 2023, Unpingco, 2016). Four clusters were determined based on the percentual pre-exam anxiety levels of physically inactive students.

Results

The Figure 1 shows the word map of the countries by the pre-exam levels of anxiety. The size of the countries name is adjusted according with the percentagem level [higher percentages = bigger word size].



Figure 1. Word map from pre-exam anxiety levels.

The Table 1 below indicate the percentual rates of pre-exam anxiety in high school students. The K-means algorithm split the data into four clusters 1 to 4 [highest to lowest rates].

Table 1 – Percentual rates of pre-exam anxiety in high school students

Cluster	n	Mean_%	Min_%	Max_%
1	8	77.6	74	82
2	14	65.6	61	70
3	16	54.8	50	60
4	8	41.3	35	40

Note. n: number of countries on each cluster.

Conclusion

The levels of pre-exam anxiety in the physically inactive students from the 48 countries ranged from 35 to 82%. These findings strengthen the importance of promoting a physically active lifestyle in the most affected countries and, thus, prevent anxiety and its complications for the youths' mental health and academic performance. This conclusion is of paramount importance in the sense that we can think that promoting physical activity in students of this age could, in the long term, reduce their pre-test anxiety. In assessment situations and, consequently, promote their success at school.

References

- Kandola, A., Lewis, G., Osborn, D. P. J., Stubbs, B., & Hayes, J. F. (2020). Depressive symptoms and objectively measured physical activity and sedentary behaviour throughout adolescence: A prospective cohort study. *The Lancet Psychiatry*, *7*(3), 262–271. [https://doi.org/10.1016/S2215-0366\(20\)30034-1](https://doi.org/10.1016/S2215-0366(20)30034-1)
- Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhajja, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, *622*, 178–210. <https://doi.org/10.1016/j.ins.2022.11.139>
- OECD. (2019). PISA 2018 Results (Volume I): What Students Know and Can Do. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/education/pisa-2018-results-volume-i_5f07c754-en
- Unpingco, J. (2016). *Python for probability, statistics, and machine learning* (Vol. 1). Springer.