Determinants of Youth Tobacco Use: A Cross-National Analysis of Policy and Country-Level Factors

A graph of a number of numbers

AI-generated content may be incorrect.*Ethan, Zayed and Layne*

Data characteristics

Analyzing global trends and risk factors associated with youth tobacco use by applying statistical learning techniques to predict smoking prevalence and identify high-risk populations and regions.

Topic

**Dataset**: Global Youth Tobacco Survey (GYTS)  
**Source**: CDC and WHO  
**Years Covered**: 1999–2018  
**Initial Observations**: 30,535 rows  
**Post-Cleaning Observations**: 25,048 rows  
**Variables Used**: 12 selected out of 27 original variables  
**Structure**: Time-series panel data with geographic and demographic granularity.



Conclusion

Our analysis of the Global Youth Tobacco Survey revealed clear patterns in youth smoking behavior worldwide. Tobacco use has steadily declined over time, with significant variation by region, gender, and public health policy. Models like logistic regression and LDA proved effective in identifying high-risk areas, achieving over 79% accuracy and strong AUC performance. These insights support evidence-based strategies for global tobacco control and offer a data-driven approach to prioritizing prevention efforts among vulnerable youth populations.