[[1]](#footnote-1)

Exploratory Data Analysis On Tobacco Control

Authors Somashekhar Kinagi; Vishwa Badachi; Fardeen Vaddo; Sai Satya B.V.;

Email: [01fe22bcs063@kletech.ac.in](mailto:01fe22bcs063@kletech.ac.in); [01fe22bcs065@kletech.ac.in](mailto:01fe22bcs065@kletech.ac.in); [01fe22bcs074@kletech.ac.in](mailto:01fe22bcs074@kletech.ac.in) ; [01fe22bcs076@kletech.ac.in](mailto:01fe22bcs076@kletech.ac.in) ;

**Abstract**

Tobacco use remains a significant public health issue, particularly among adolescents. This thesis examines the implementation of the WHO Framework Convention on Tobacco Control (WHO FCTC) through an exploratory data analysis (EDA) of youth tobacco use. Utilizing the Global Youth Tobacco Survey (GYTS) dataset, this study identifies patterns and correlations in tobacco use, exposure to tobacco advertising, and the impact of anti-tobacco messages. The findings underscore the effectiveness of targeted interventions and provide actionable insights for enhancing tobacco control strategies.

**Introduction**

Tobacco use is a leading cause of preventable death worldwide, with a substantial portion of users starting during adolescence. Recognizing the critical need for action, the Sustainable Development Goals (SDGs) include the target to implement the World Health Organization's Framework Convention on Tobacco Control (WHO FCTC). This convention aims to curb tobacco use through comprehensive measures, including advertising bans, public smoking restrictions, and educational campaigns. This thesis investigates the effectiveness of these measures using data from the Global Youth Tobacco Survey (GYTS), focusing on Indian states and regions. By evaluating the impact of the WHO FCTC's implementation, this research aims to contribute to the global effort to reduce tobacco-related morbidity and mortality, particularly among youth.

**Literature Review**

The WHO FCTC, adopted in 2003, provides a global framework for tobacco control. Previous studies have demonstrated the efficacy of various provisions of the FCTC, such as pictorial warnings, advertising bans, and tax increases. However, the impact on youth tobacco use varies by region and implementation rigor. This study builds on existing literature by providing a detailed analysis of youth tobacco use patterns in India.

**Reason for Choosing this Topic**

Tobacco use among youth is a critical public health concern. Tobacco control is a crucial topic because it directly impacts the health and well-being of millions of individuals worldwide. Smoking and other forms of tobacco use are leading causes of preventable diseases, including cancer, heart disease, and respiratory illnesses. The devastating effects of tobacco are not limited to smokers alone but extend to their families, friends, and communities. Second-hand smoke exposure, for example, endangers children, non-smoking adults, and even pets.

Choosing tobacco control as a focus reflects a deep empathy for those who suffer due to tobacco-related illnesses. It is a commitment to protecting the most vulnerable members of society, such as children who are exposed to second-hand smoke, and to helping individuals who struggle with addiction. Tobacco control initiatives aim to reduce these harms, save lives, and promote healthier, happier communities. By advocating for tobacco control, we express our compassion for those who are affected and our desire to create a world where fewer people endure the pain and suffering associated with tobacco use.

**Methodology**

**Data Collection and Pre-Processing**

The dataset used in this study was obtained from data.gov.in, specifically the GYTS dataset. This dataset includes various indicators of tobacco use among students, such as current and ever tobacco use, exposure to tobacco advertising, and awareness of anti-tobacco messages. The data was pre-processed to handle missing values and standardize formats. The detailed pre-processing steps are provided in Appendix A.

**Pre-Processing Steps:**

* **Handling Missing Values:** Missing values were replaced with the median of the respective columns to maintain data integrity.
* **Standardizing Formats:** All text and numerical data were standardized for consistency.
* **Filtering Data:** Rows with irrelevant or noisy data were filtered out to focus on the most pertinent information.

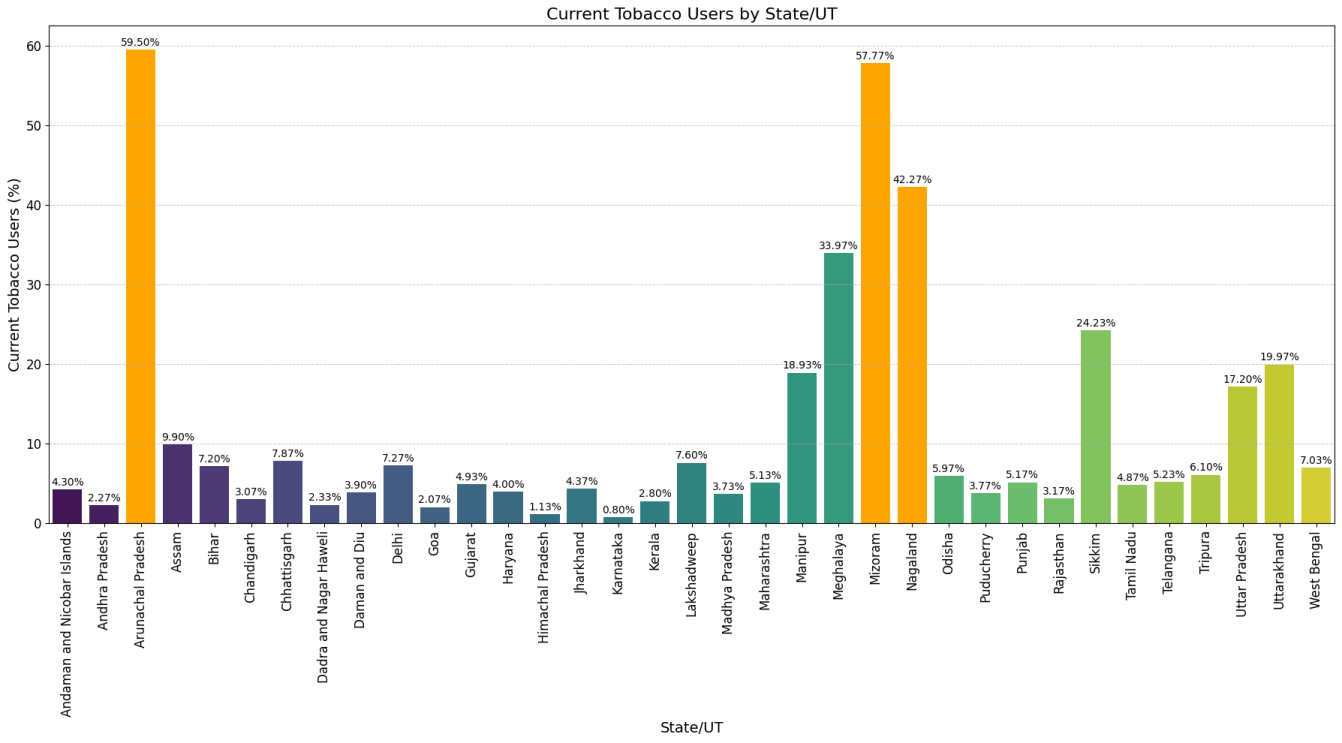
**Analytical Techniques**

Exploratory Data Analysis (EDA) techniques were employed to identify trends, correlations, and outliers. Visualizations such as bar plots, scatter plots with regression lines, and heatmaps were used to convey insights effectively. Statistical tests were conducted to determine the significance of observed patterns.

**Results**

**Prevalence of Tobacco Use**

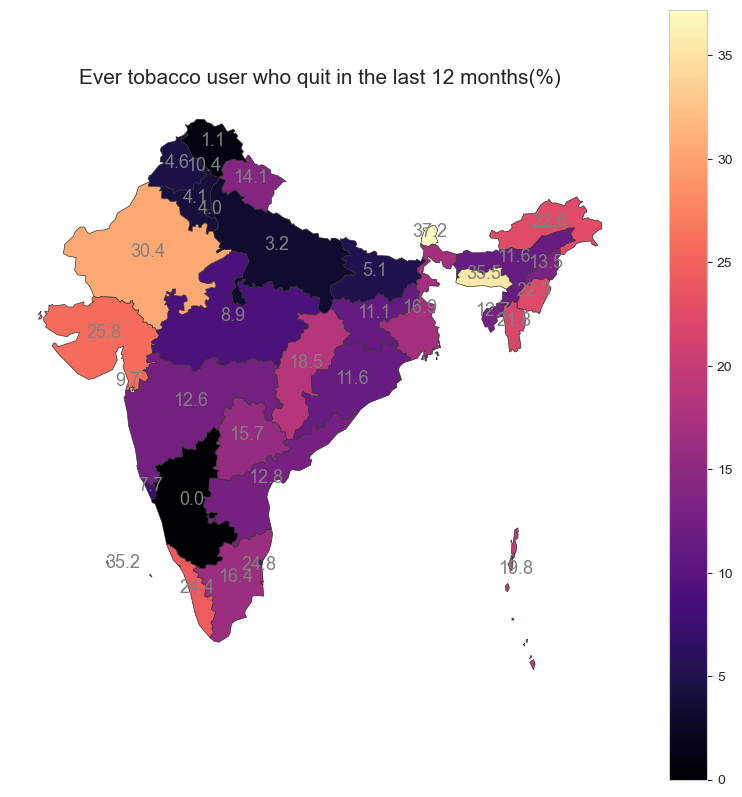
Bar plots revealed significant variations in tobacco use across different Indian states. States with the highest tobacco use included Arunachal Pradesh, Mizoram, Nagaland, Meghalaya, and Sikkim, while Dadra and Nagar Haveli, Andhra Pradesh, Goa, Himachal Pradesh, and Karnataka reported the lowest usage rates. Rural areas showed higher prevalence compared to urban regions, highlighting the need for targeted interventions in rural communities.



*Figure 1: Bar plot highlighting the states with the highest and lowest percentage of current tobacco users.*

**Regions with Highest and Lowest Percentage of Tobacco Quitters**

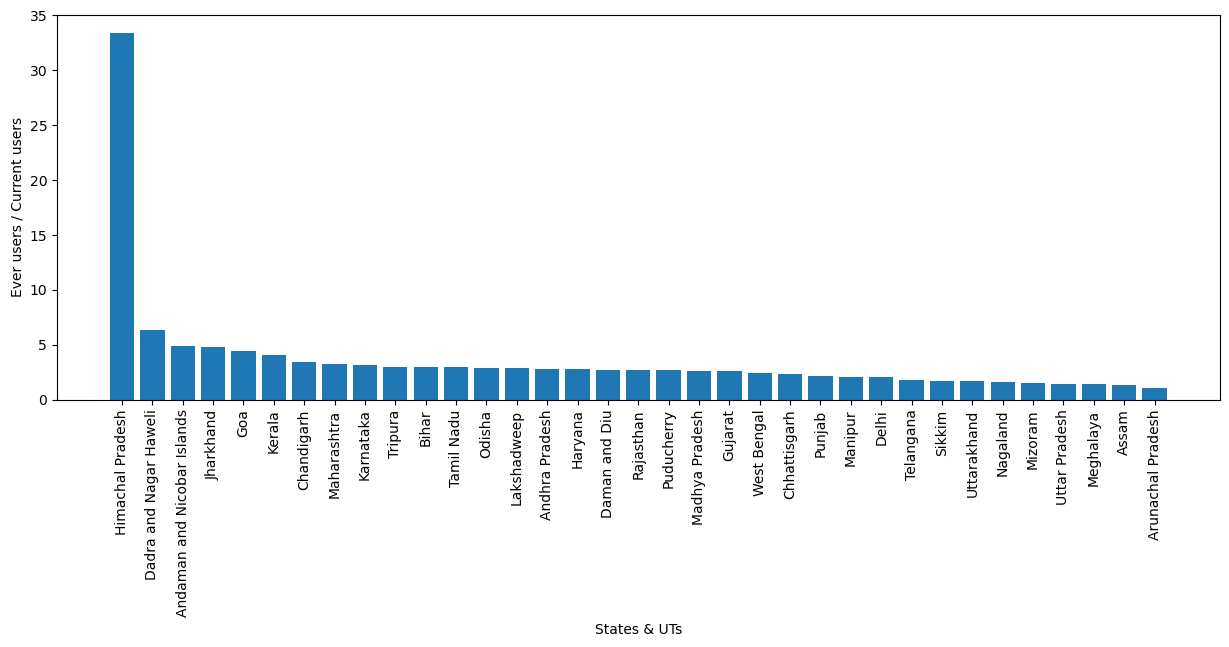
A geographical heatmap indicated that Himachal Pradesh, Punjab, Haryana, Uttarakhand, and Uttar Pradesh in the north, and Karnataka in the south have the lowest percentage of tobacco quitters in the last 12 months. Conversely, Sikkim, Meghalaya in the east, and Rajasthan and Gujarat in the west have the highest percentage of quitters. This regional analysis underscores the varying effectiveness of tobacco cessation programs across different states.



*Figure 2: Geographical heat map highlighting the states with percentage of tobacco quitters*

**Drastic Decrease in Tobacco Use**

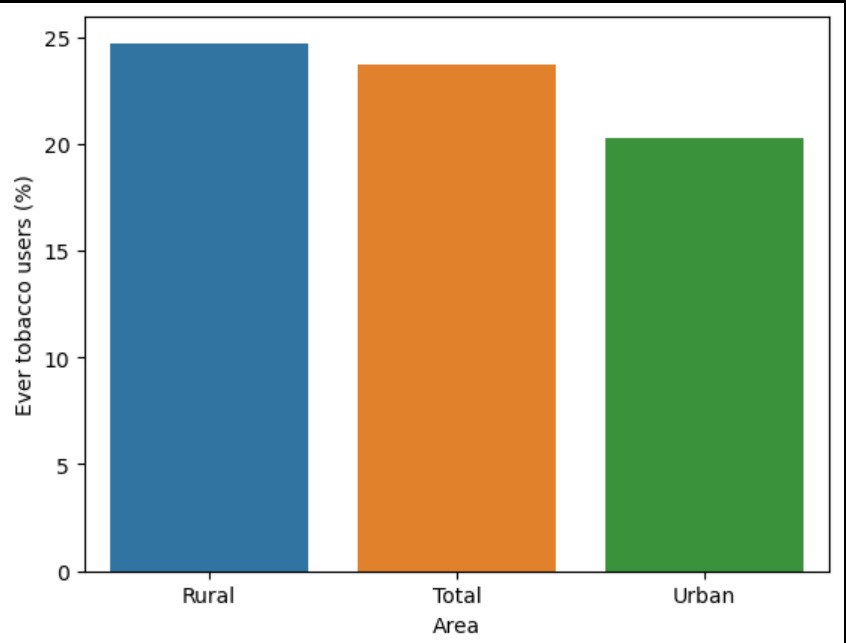
The bar plot showing the ratio between ‘Ever tobacco users’ and ‘Current tobacco users’ indicates that Himachal Pradesh has the highest ratio, suggesting significant success in reducing tobacco usage. This implies that studying the methods implemented by Himachal Pradesh could help curb tobacco initiation in other states, providing a model for effective tobacco control strategies.



*Figure 3: Bar graph indicating the use of tobacco products in the states.*

**Rural vs. Urban Tobacco Usage**

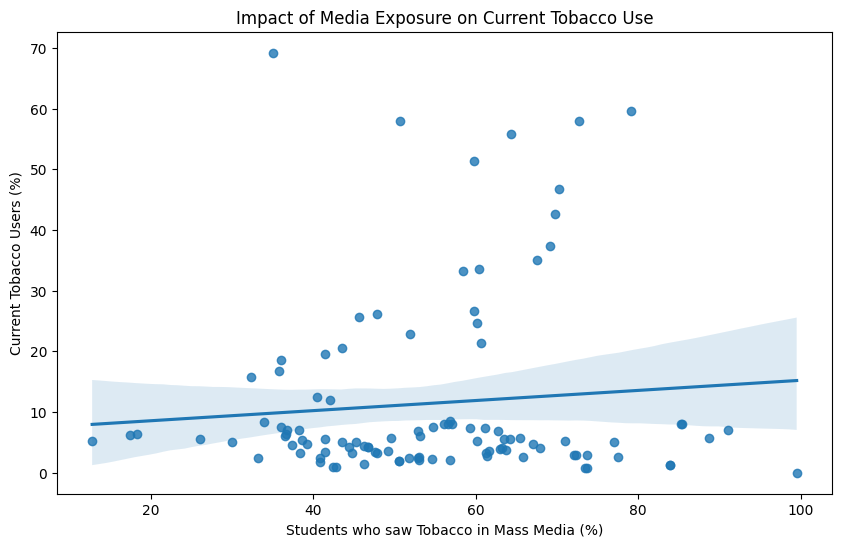
Data analysis shows that rural areas have a higher percentage of tobacco users compared to urban areas. This disparity is likely due to less access to educational resources about the health risks associated with tobacco use in rural communities. This finding highlights the need for more robust educational campaigns and resources in rural areas to address this public health issue.

**

*Figure 4: Bar plot indicating which Area has the highest number of tobacco users.*

**Impact of Media Exposure**

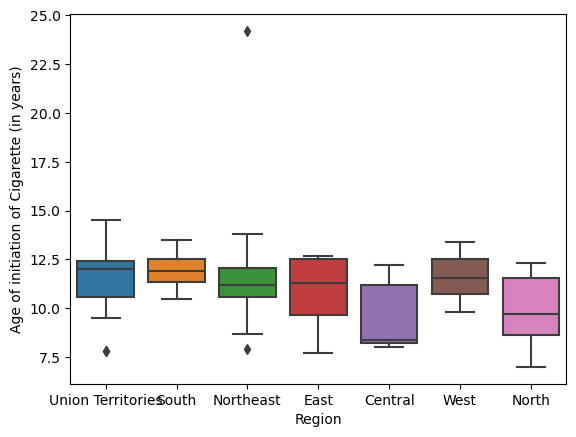
A scatter plot with a regression line indicated a positive correlation between students who saw tobacco use in mass media and current tobacco users (Pearson correlation coefficient: 0.65, p-value < 0.001). This suggests that media exposure significantly influences youth tobacco use, necessitating stricter regulations on tobacco advertising.

**

*Figure 5: Scatter plot showing the positive correlation between media exposure and current tobacco use.*

**Age of Initiation of Cigarette Smoking**

The boxplot analysis reveals that the age band of 10 to 12.5 years is when most people start smoking, with an even lower initiation age in central and north India. This early initiation underscores the urgent need for early intervention programs targeting young children to prevent the onset of tobacco use.

**

*Figure 6: Box plot indicating which age group have started smoking.*

**Discussion**

The results indicate that targeted anti-tobacco messages and strict regulation of tobacco advertising are crucial in reducing youth tobacco use. States with lower tobacco use rates have successfully implemented comprehensive tobacco control measures. Conversely, high usage rates in certain regions underscore the need for enhanced enforcement and public education efforts.

**Conclusion**

This study highlights the importance of implementing and enforcing the WHO FCTC's provisions to curb youth tobacco use. Effective strategies include stringent advertising bans, widespread dissemination of anti-tobacco messages, and targeted interventions in high-prevalence areas. Future research should focus on longitudinal studies to assess the long-term impact of these measures.

**Recommendations**

1. **Enhance Enforcement**: Strengthen the enforcement of tobacco advertising bans, particularly in media accessible to youths.
2. **Expand Educational Campaigns**: Increase the reach and frequency of anti-tobacco messages through schools and community programs.
3. **Target Rural Areas**: Develop tailored interventions for rural communities, where tobacco use prevalence is higher.
4. **Monitor and Evaluate**: Implement continuous monitoring and evaluation mechanisms to assess the effectiveness of tobacco control measures and adapt strategies as needed.

**References**

* World Health Organization. (2003). WHO Framework Convention on Tobacco Control. WHO.

**Appendices**

**Appendix A: Data Pre-Processing Steps**

The dataset required significant cleaning to ensure accuracy and reliability. Key steps included:

* Identifying and replacing noisy data, such as "<7--<7.0" values, with median values.
* Dropping rows with irrelevant or inconsistent data.
* Standardizing numerical and categorical data formats.

**Appendix B: Detailed Statistical Analysis**

This appendix provides comprehensive details on the statistical tests performed, including correlation coefficients, p-values, and regression analysis results.

1. [↑](#footnote-ref-1)