Youth Substance Abuse Analysis (2020-2024)

Manaswi Kolani

Contents

1 Introduction to the Dataset and Project Objectives

This project examines trends in youth smoking and drug experimentation from 2020 to 2024, utilizing a comprehensive data set comprising 10,000 observations. The data set captures various influencing factors, including age groups, gender, socioeconomic status, family influences, and peer influences. The objective is to inform research and policy development by identifying patterns and trends in youth substance abuse.

Dataset Link: Youth Smoking and Drug Dataset on Kaggle

2 Key Insights from Data Visualizations

The visualizations produced in this analysis uncovered several critical trends, offering valuable insights for policymakers and researchers:

- **Gender Differences in Smoking Rates:** The average smoking rate among females is significantly higher than that of males in the age groups 15-19, 30-39, and 40-49 across the studied years.
- **Peak Smoking Rates:** The highest smoking rates were observed in males aged 20-24 in 2022, and among both males and females aged 60-69 in 2023.
- **Influence of Social Factors:** While multiple factors influence substance use, peer influence and family background play a significant role in the prevalence of smoking and drug experimentation.
- **Impact of Awareness Initiatives:** Interestingly, no instances of high smoking or drug experimentation rates were recorded when comprehensive awareness factors were present, in contrast to 38 such observations when awareness was limited or absent.
- **Family Influence:** An unexpected finding was that 6 out of the 14 instances with the highest smoking and drug experimentation rates were associated with strong family influence.

3 Development of the Interactive Interface

The interactive visualization interface was built using the Shiny package in R, designed to facilitate dynamic data exploration. Data preparation involved extensive cleaning, filtering, pivoting, and selection processes to tailor the data set for effective visualization. The development process was iterative, focusing initially on three primary objectives: enabling comparisons, highlighting influential factors, and allowing user-defined grouping. This approach guided the selection of three key plots and a data table, optimized for user engagement and ease of interpretation. The user interface (UI) was refined to enhance usability and intuitiveness.

4 Reactive Graph Structure

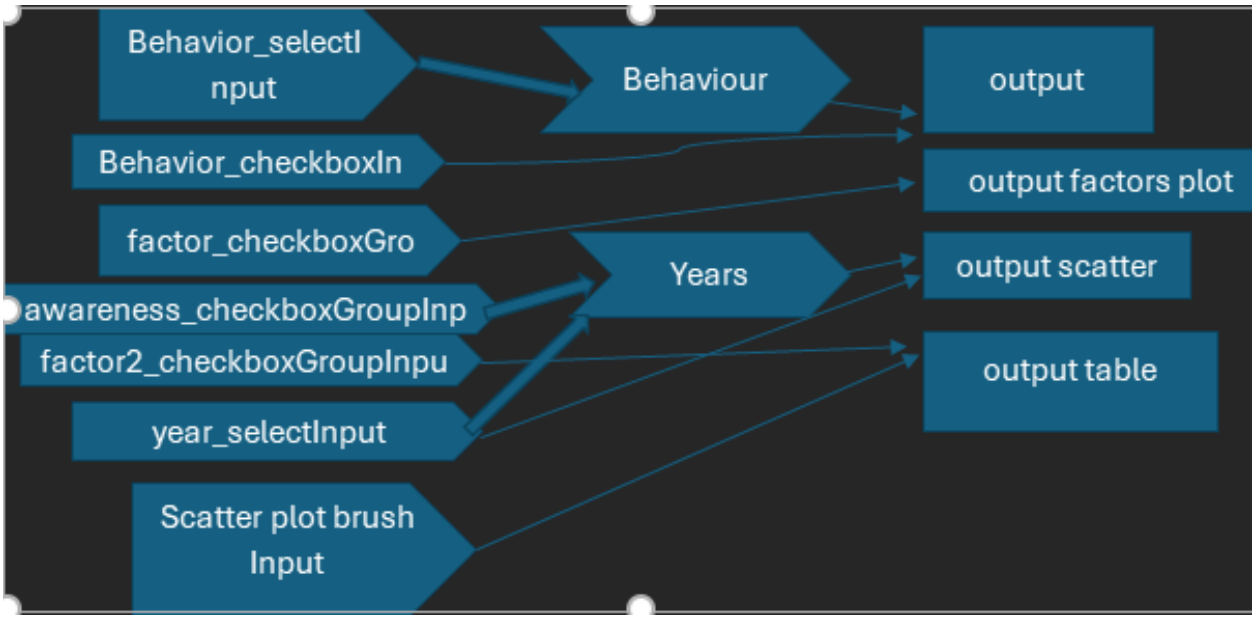

Figure 1: Reactive Graph

5 Future Enhancements

Moving forward, the analysis will be expanded to examine the specific impact of individual influencing factors on substance use rates. Additional features will be integrated into the Shiny application to provide deeper insights, aiming to create a robust tool for policy development and behavioral research.

6 Links to Code and Visualization

- **Code Repository:** Link to Code
- **Interactive Visualization:** Link to Shiny App