To evaluate the impact of the tobacco policy implemented in 2016 in all counties in California, the researchers utilized multilevel linear models, including both fixed and random effects. Fixed effects are the independent variables that are intentionally manipulated, while random effects account for random variation.

The models included a range of factors such as socio-economic factors (e.g., poverty, ethnicity, and education backgrounds) and policy factors (e.g., smoke-free policies, city licensing, pharmacy bans, emerging licenses, retail restrictions, flavored products, and fees). When including categorical variables as fixed effects in an lmer model, it is necessary to specify a reference level. The parameters estimated by the model are relative to that reference level. By default, the first level of a factor is used as the reference level, but the relevel() function can be used to specify a different reference level.

To identify the most significant factors, the researchers used p-values, estimated coefficients, and confidence intervals at a 5% confidence level. Factors with p-values less than 0.05, 0.01, and 0.001 were considered statistically significant and denoted as "\*", "\*\*", and "\*\*\*", respectively. The coefficients associated with these fixed effects represent the effect of each predictor variable on the outcome variable, controlling for the other predictor variables in the model.

The inclusion of key socio-economic factors provided valuable insights into the groups most affected by the policy and highlighted the underlying factors contributing to tobacco use. Further statistical approaches were adopted to analyze the impact of the tobacco policy in California. Firstly, a binary variable was used to represent the years before and after the treatment year. Secondly, interaction terms were considered to test the combined effect and significance of multiple factors, such as year, city license, pharmacy ban, and their interactions with socio-economic factors like poverty, ethnicity, and education background. Thirdly, to make the coefficients of estimates more interpretable, the response variable RetailerCount\_density was multiplied by 1000 in the linear mixed effects model.

The random effect in the lmer model was used to model the variation in the outcome variable that is not explained by the fixed effects but rather by the variation between different cities. The term (1 | City) specifies a random intercept for the variable City in the linear mixed effects model. Each city is assumed to have a different baseline level of RetailerCount\_density that is drawn from a normal distribution with mean zero and a certain variance. The random intercept allows for the possibility that the outcome variable may vary across different cities, even after accounting for the fixed effects included in the model.

The random effect captures the unobserved variation in the outcome variable that is specific to each city, which is not accounted for by the fixed effects. The random effect is used to model the correlation between observations within the same city and to estimate the variance of the random intercept, which represents the variation in the outcome variable between different cities.

By utilizing these statistical approaches, the researchers were able to obtain a more nuanced understanding of the impact of the tobacco policy in California. The inclusion of binary variables and interaction terms allowed the researchers to assess the impact of the policy over time and across different groups, providing valuable insights into the effectiveness of the tobacco policies in California. Additionally, the use of a random effect in the lmer model allowed for the modeling of the variation in the outcome variable between different cities that is not explained by the fixed effects.

Overall, the study highlights the importance of using a range of statistical approaches, including multilevel linear models, p-values, estimated coefficients, and confidence intervals, binary variables, interaction terms, and random effects, to effectively analyze complex data and obtain a comprehensive understanding of policy outcomes. The inclusion of key socio-economic factors in the analysis provided a comprehensive understanding of the impact of tobacco policies on tobacco use in California. By identifying the most significant factors contributing to tobacco use, policymakers can develop more effective policies to reduce tobacco use and its associated health risks.