Americian Youth Smoking Rate

Yeyunfei Jiang (1004170354)

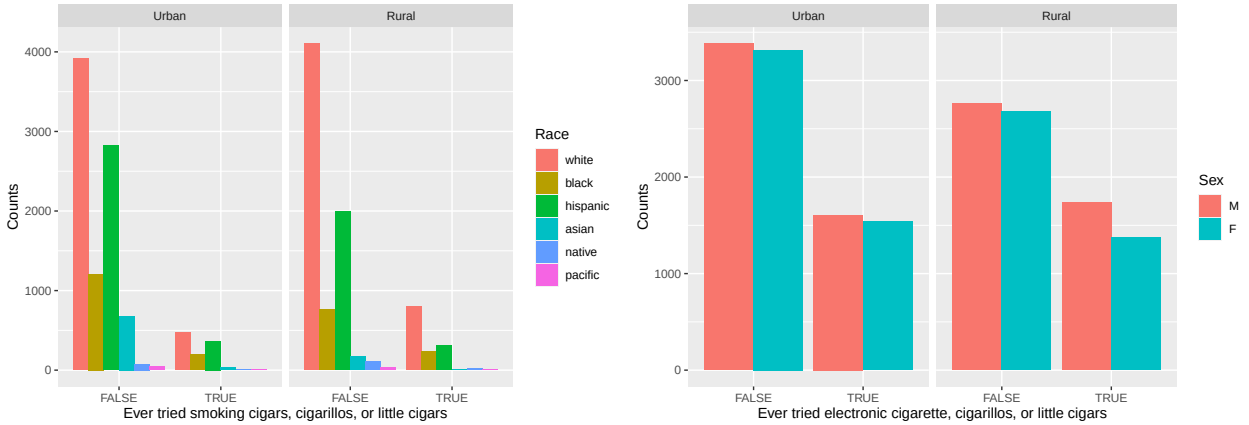
2020/10/2

Introduction:

Smoking causes cancer, heart disease, stroke, lung diseases and diabetes, which is a major health concern for youth.The research question of interest concerns the usage of cigar among white, black and hispanic Americians, and the likelihood of having used a electronic cigarettes for two individuals of the different sexes. Two hypothesis are made as follows:

1. Smoking of cigars, cigarillos or little cigars is no more common amongst Americans of European ancestry than for Hispanic-Americans and African-Americans, once one accounts for the fact that white Americans more likely to live in rural areas and cigar smoking is a rural phenomenon.
2. The likelihood of having used a electronic cigarettes on at least one occasion is the same for two individuals of the different sexes, provided their age, ethnicity, and other demographic characteristics are similar.

The dataset used in this study is from 2019 American National Youth Tobacco Survey, which contains over 18000 survey from Americian youth of different races.



Method:

The two models used for this research are both binomial generlized linear model with logit link.

$$log(\frac{\mu_i}{1-\mu_i}) = X_i\beta$$

Where μ_i is the probabilty of the ith student ever tried smoking cigars, cigarillos, or little cigars in first Model, and the probability of used an electronic cigarette in second model. The covariates in both model are gender(male, female), age(centered at 16), RuralUrban(Rural/Urban) and Race(white, asian, hispanic, native, black or pacific).

Results:

Table 4 shows the parameter estimate and corresponding p-value for the Cigar model. And table 5 shows the 95% Confidence interval for the parameters. Where baseline probability is a 16 years old white male live in urban areas ever tried smoking cigars.

Table 1: Summary of coefficient

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.562	0.046	-34.103	0.000
Raceblack	0.429	0.064	6.702	0.000
Racehispanic	-0.057	0.053	-1.066	0.286
Raceasian	-1.262	0.172	-7.355	0.000
Racenative	0.283	0.205	1.377	0.168
Racepacific	0.432	0.280	1.541	0.123
RuralUrbanRural	0.401	0.046	8.796	0.000
SexF	-0.381	0.046	-8.318	0.000
Agec	0.374	0.012	31.324	0.000

Table 2: Normal scale parameter estimate and 95% Confidence interval

	est	2.5 %	97.5 %
Racewhite	0.173	0.161	0.186
Raceblack	0.606	0.575	0.635
Racehispanic	0.486	0.460	0.512
Raceasian	0.221	0.166	0.281
Racenative	0.570	0.466	0.662
Racepacific	0.606	0.463	0.722
RuralUrbanRural	0.599	0.577	0.620
SexF	0.406	0.384	0.428
Agec	0.592	0.587	0.598

Table 6 shows the parameter estimate and corresponding p-value for the E cigarette model. And table 7 shows the 95% Confidence interval for the parameters. Where baseline probability is a 16 years old white male live in urban areas ever used electronic cigarettes.

Table 3: Summary of coefficient

	Estimate	Std. Error	z value	Pr(>|z|)
SexM	-0.139	0.034	-4.084	0.000
SexF	-0.199	0.035	-5.726	0.000
RuralUrbanRural	0.131	0.033	3.930	0.000
Raceblack	-0.515	0.054	-9.614	0.000
Racehispanic	-0.089	0.038	-2.369	0.018
Raceasian	-1.009	0.092	-11.023	0.000
Racenative	0.062	0.153	0.406	0.685
Racepacific	0.237	0.215	1.102	0.270
Agec	0.337	0.008	39.901	0.000

Table 4: Normal scale parameter estimate and 95% Confidence interval

	est	2.5 %	97.5 %
SexM	0.465	0.449	0.482
SexF	0.450	0.433	0.467
RuralUrbanRural	0.533	0.516	0.549
Raceblack	0.374	0.350	0.399
Racehispanic	0.478	0.459	0.496
Raceasian	0.267	0.233	0.303
Racenative	0.516	0.440	0.589
Racepacific	0.559	0.452	0.658
Agec	0.583	0.579	0.587

Discussion:

It is worth noting that, this dataset the model uses only contains youth that between 11 and 19 years old, which can be bias for a measure of all Americians. Which means it can only test the hypothesis in youth group.

As shown in Table 2 , a 95% confidence interval for a 16 years old white male live in urban areas ever tried cigars is (0.161, 0.186). Keep other variable constent, the CI for black and hispanic youth is (0.575, 0.635) and (0.46, 0.512) corresponding. Apparently, the probability of black and hispanic youth has a much higher probability of cigar usage compared to white youth. Hence. there is evidence to support the first hypothesis, namely, Smoking of cigars is no more common amongst Americans of Europeanancestry than for Hispanic-Americans and African-Americans.

As shown in Table 4, a 95% confidence interval for a 16 years old white male live in urban areas ever tried e-cigarettes is (0.449, 0.482). Keep other variable constent, the CI for female is (0.433, 0.467). There is maximal a 5 percentage difference in the probability between male and female when the age, ethnicity, and other demographic characteristics are similar, which can be considered as a small deference. Hence, there is evidence to support the second hypothesis, namely, the likelihood of having used a electronic cigarettes on at least one occasion is the same for two individuals of the different sexes, provided their age, ethnicity, and other demographic characteristics are similar.

Summary of Conclusion:

The study modelled the relation between each race, gender, age with the proportion of cigars and e-cigarettes usage among youth by using the 2019 American National Youth Tobacco Survey .It revealed that Smoking of cigars is no longer common in Americans of European ancestry than for Hispanic-Americans and African-Americans. It also revealed that given age, ethnicity, and other demographic characteristics similar, the probability of male and female you of tried e-cigarettes is very close to each other with a maximal percentage difference of 5.