Homework 10

Isaiah Benny EID: ieb357

2024-04-25

Github

Problem 1

Question

The question that we are trying to answer is whether insurance companies in Chicago in the 1970s were redlining.

Approach

To answer this question, I will create a linear regression model to determine if the amount of minorities in a ZIP code is associated with the number of FAIR policies, adjusting for rates of fire, the proportion of old houses, and median income.

Results

term	estimate	std_error	statistic	p_value	lower_ci	upper_ci
intercept	-0.125	0.633	-0.197	0.845	-1.410	1.160
minority	0.008	0.003	2.918	0.006	0.003	0.014
fire	0.022	0.009	2.475	0.018	0.004	0.040
age	0.006	0.004	1.497	0.143	-0.002	0.013
income	-0.016	0.038	-0.418	0.679	-0.094	0.062

The table above displays the estimates of the linear model along with the confidence intervals for each coefficient.

Conclusion

We can conclude that an increase in the percentage of minorities in a ZIP code during this time likely has a positive effect on the number of FAIR plan policies and renewals when keeping fire rates, age, and median income constant. This is because the 95% confidence interval for the minority term contains only positive values. Since the individual effect of an increase in the percentage of minorities is an increase in FAIR

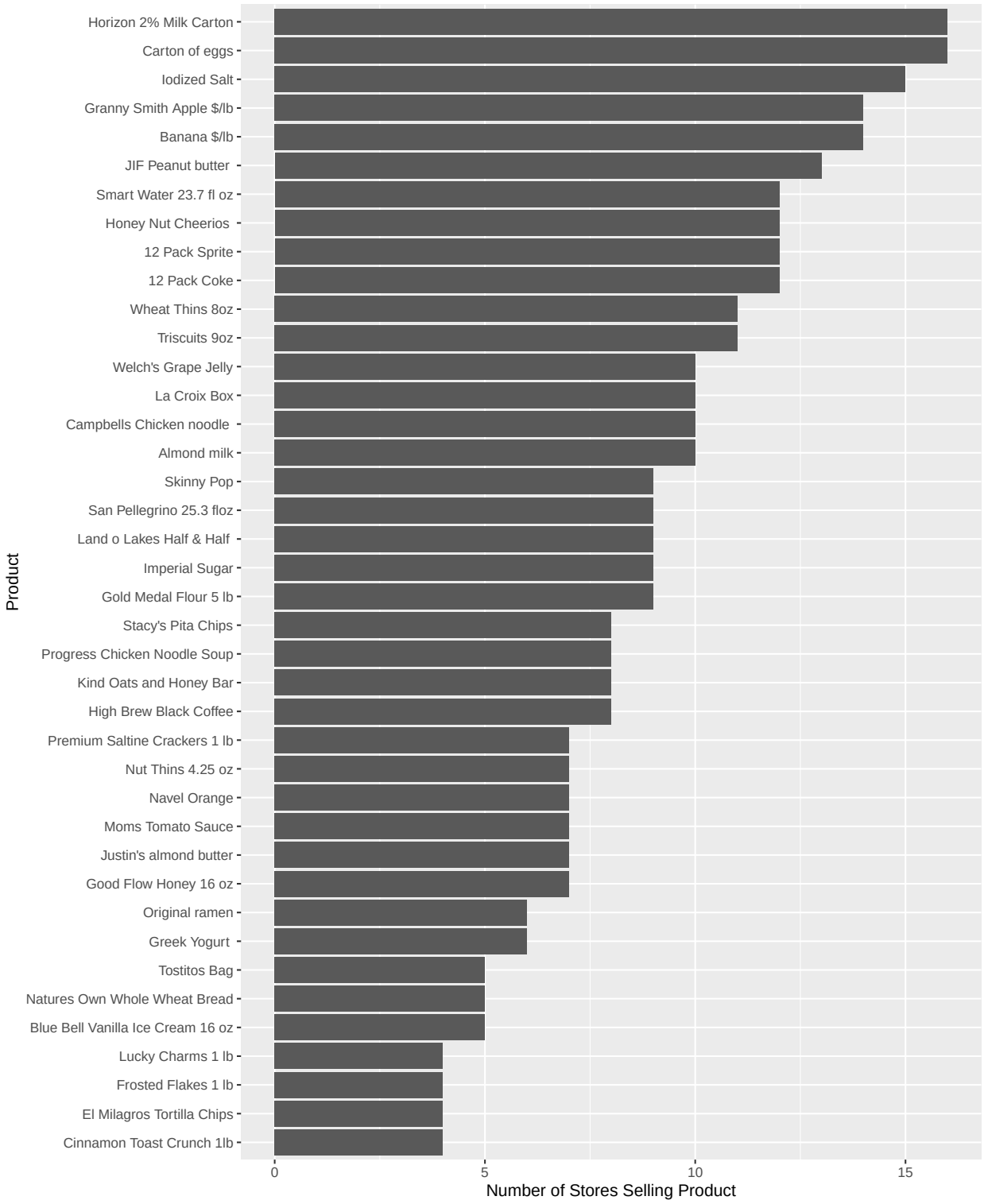
policies, this provides some evidence that private insurance policies during this time had denied service on the basis of race (redlining), because we have assumed that more FAIR policies implies a lack of access to the private insurance market. However, this evidence is not definitive, as there could be other impactful confounders we have not adjusted for.

Problem 2

Part A

This plot shows the average price of the products sold for each store in the dataset.

Part B



This plot shows each product in the dataset and the number of stores that sell that product.

Part C

Compared with ordinary grocery stores (like Albertsons, HEB, or Krogers), convenience stores charge somewhere between 0.41 and 0.92 dollars more for the same product.

Part D

The two stores that seem to charge the lowest prices when comparing the same product are Kroger Fresh Fare and Walmart. The two stores that appear to charge the highest prices when comparing the same product are Whole Foods and Wheatsville Food Co-Op.

Part E

term	estimate	std_error	statistic	p_value	lower_ci	upper_ci
Store: Central Market	-0.573	0.177	-3.240	0.001	-0.922	-0.225
Store: CVS	0.193	0.183	1.056	0.292	-0.167	0.553
Store: Fiesta	-0.703	0.269	-2.609	0.010	-1.234	-0.173
Store: Fresh Plus	-0.036	0.162	-0.223	0.823	-0.355	0.283
Store: H-E-B	-0.646	0.152	-4.249	0.000	-0.945	-0.347
Store: Kroger	-0.703	0.233	-3.014	0.003	-1.162	-0.244
Store: Kroger Fresh Fare	-0.902	0.233	-3.865	0.000	-1.361	-0.443
Store: Natural Grocers	-0.081	0.198	-0.411	0.681	-0.470	0.308
Store: Target	-0.373	0.190	-1.966	0.050	-0.747	0.000
Store: Walgreens	0.215	0.181	1.192	0.234	-0.140	0.571
Store: Walmart	-0.993	0.233	-4.254	0.000	-1.452	-0.533
Store: Wheatsville Food Co-Op	0.290	0.179	1.624	0.105	-0.061	0.642
Store: Whole Foods	0.364	0.177	2.062	0.040	0.017	0.712

The table above shows the coefficients for the store variables in the linear model from part D. Using this table, is appears as if Central Market charges a similar amount to HEB for the same product. This is because the difference between the estimates of these variables is 0.073, which is relatively small when comparing the differences between Central Market and HEB with other stores in the model.

Part F

The Income10k coefficient has a coefficient of -0.014, which is negative. This means that if we hold the product constant, as income increases, price decreases. Thus, consumers in poorer ZIP codes seem to pay more for the same product on average compared to consumers in richer ZIP codes.

A one-standard deviation increase in the income of a ZIP code seems to be associated with a -0.032 standard-deviation change in the price that consumers in that ZIP code expect to pay for the same product.