Homework 4

Aditya Prasad - arp4759

2024-04-25

Github: https://github.com/aprasad-1/Homework10.git

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

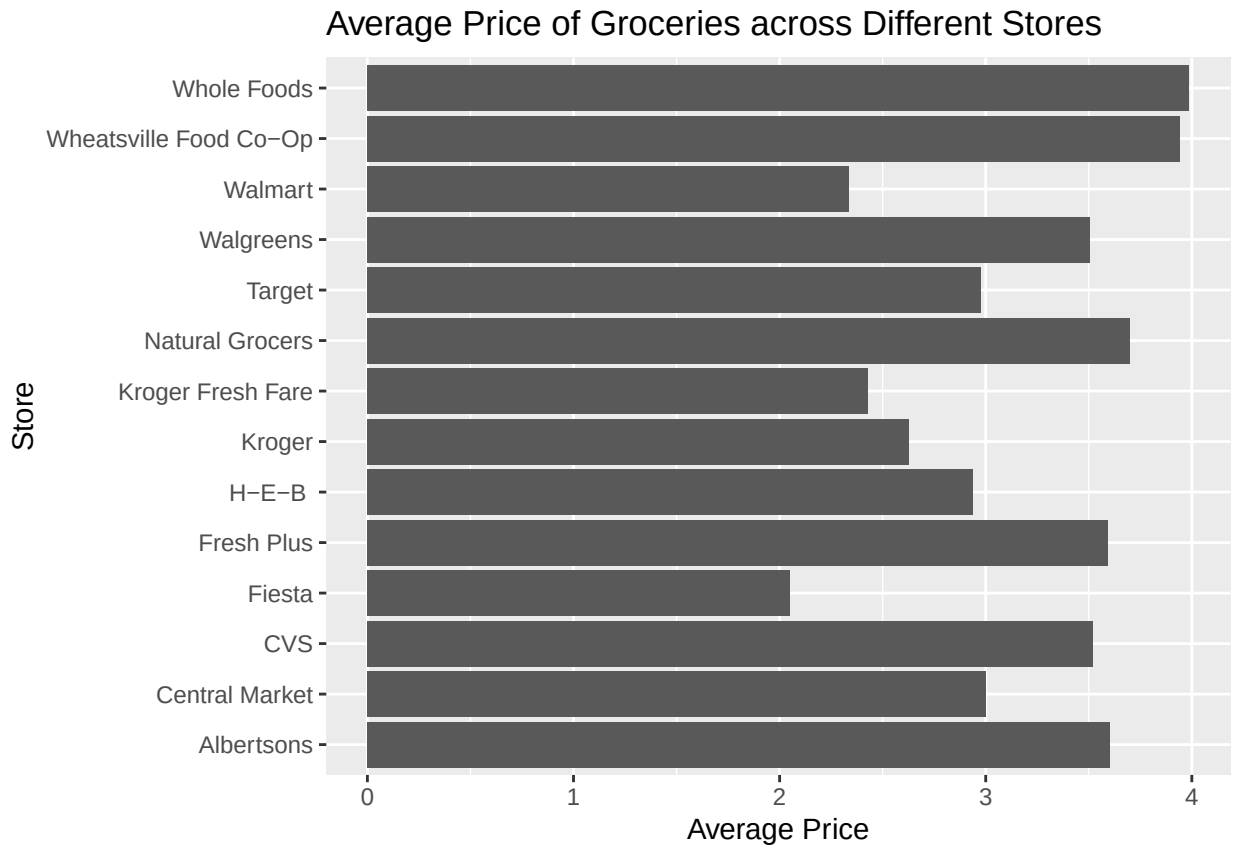
```
## # A tibble: 5 x 7
##   term      estimate std_error statistic p_value lower_ci upper_ci
##   <chr>        <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
## 1 intercept   -0.125     0.633    -0.197   0.845   -1.41     1.16
## 2 minority     0.008     0.003     2.92    0.006    0.003    0.014
## 3 fire         0.022     0.009     2.48    0.018    0.004    0.04
## 4 age          0.006     0.004     1.50    0.143   -0.002    0.013
## 5 income      -0.016     0.038    -0.418   0.679   -0.094    0.062
```

Question: Is there an association between the number of FAIR policies and the racial/ethic composition of a ZIP code adjusting for the fire, age, andincome variables?

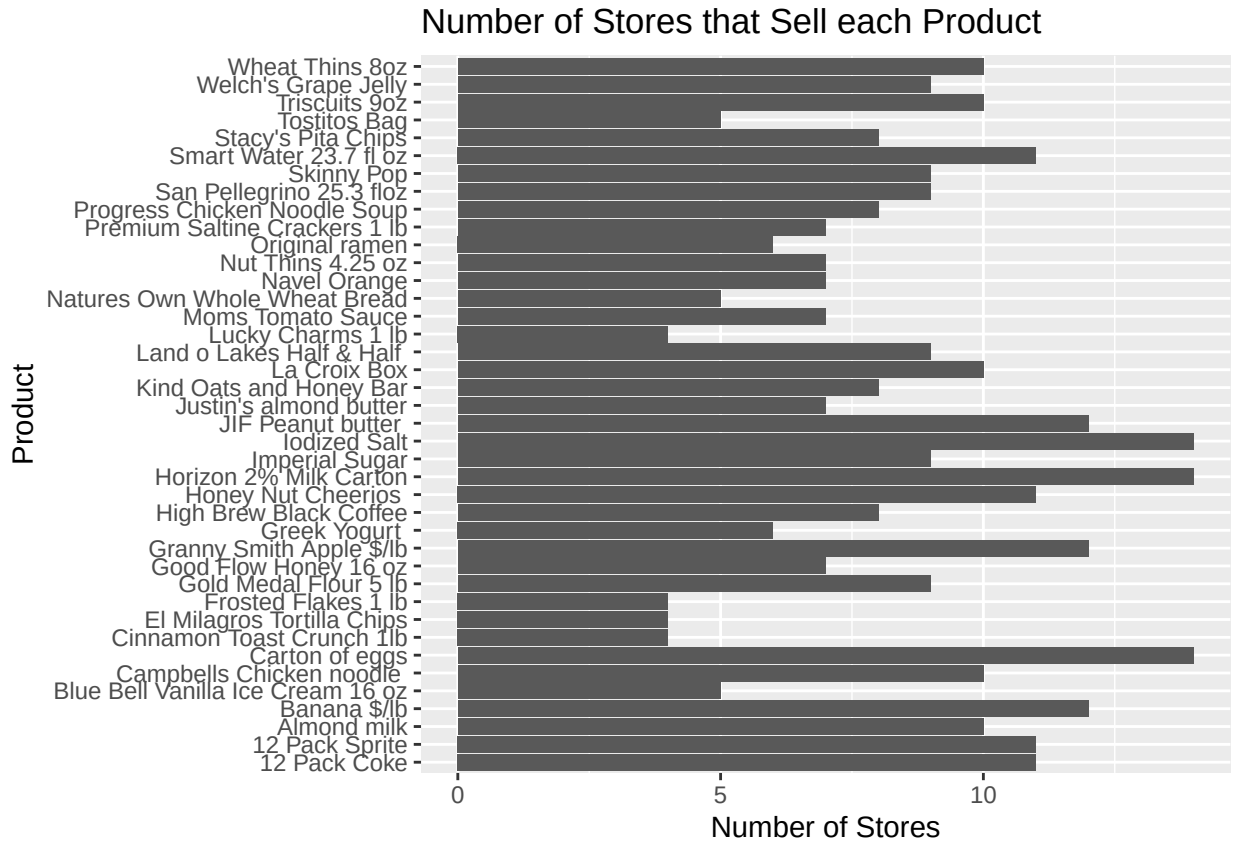
Approach: In order to answer this question, I used the linear model function to buid a regression model and then found the coefficients of the model and the confidence intervals for it as well.

Results: The estimated coefficient between the number of FAIR policies and minority groups is 0.08. The p-value is 0.006 and the 95% confidence interval is (0.003,0.014)

Conculsion: We are 95% confident that and additional 1% of minorities increases new fair programs by 0.08 per 100 housing units in that zip code since 0.008 is in our confidence interval. Additionally, we can conclude that no partial relationship between new policies and minorities is highly unlikely due to the p-value being 0.006, making it statistically significant.



The bar graph shows the average price of groceries across different grocery stores.



This bar graph shows the number of stores that sell each product.

```
## # A tibble: 44 x 7
##    term                  estimate std_error statistic p_value lower_ci upper_ci
##    <chr>                    <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
##  1 "intercept"               5.92     0.243    24.4     0        5.45     6.40
##  2 "Product: 12 Pack Spr~   -0.018    0.313    -0.059   0.953   -0.634    0.597
##  3 "Product: Almond milk"   -2.20     0.329    -6.70    0       -2.85    -1.56
##  4 "Product: Banana $/lb"   -4.88     0.303   -16.1     0       -5.48    -4.29
##  5 "Product: Blue Bell V~   -2.95     0.409    -7.20    0       -3.75    -2.14
##  6 "Product: Campbells C~   -3.46     0.329   -10.5     0       -4.11    -2.82
##  7 "Product: Carton of e~   -3.00     0.293   -10.2     0       -3.58    -2.42
##  8 "Product: Cinnamon To~   -0.997    0.444    -2.25    0.025   -1.87    -0.124
##  9 "Product: El Milagros~   -2.04     0.444    -4.59    0       -2.91    -1.16
## 10 "Product: Frosted Fla~   -1.25     0.444    -2.82    0.005   -2.12    -0.379
## # i 34 more rows
```

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.

```
## # A tibble: 53 x 7
##    term                  estimate std_error statistic p_value lower_ci upper_ci
##    <chr>                    <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
##  1 "intercept"               5.76     0.234    24.6     0        5.30     6.22
##  2 "Product: 12 Pack Spr~   -0.018    0.287    -0.064   0.949   -0.583    0.546
##  3 "Product: Almond milk"   -2.29     0.303    -7.56    0       -2.89    -1.70
```

```
##  4 "Product: Banana $/lb"   -4.86      0.278   -17.5     0       -5.40    -4.31
##  5 "Product: Blue Bell V~   -3.08      0.377    -8.16    0       -3.82    -2.34
##  6 "Product: Campbells C~   -3.55      0.303   -11.7     0       -4.15    -2.96
##  7 "Product: Carton of e~   -2.98      0.269   -11.1     0       -3.51    -2.45
##  8 "Product: Cinnamon To~   -1.17      0.409    -2.87    0.004   -1.98    -0.368
##  9 "Product: El Milagros~   -1.91      0.412    -4.64    0       -2.72    -1.1
## 10 "Product: Frosted Fla~   -1.43      0.409    -3.49    0.001   -2.23    -0.623
## # i 43 more rows
```

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

The estimated coefficent of Central market is -0.573 and the estimated coefficient of H-E-B is -0.646. The confindence interval for Central Market is (-0.922, -0.225) and the confidence interval for H-E-B is (-0.945, -0.347), which are very similar. The difference between the coefficents and confidence intervals is not as significant compared to other grocery stores. This shows the reason Central Market charges more is because it sells different products.

```
## # A tibble: 41 x 7
##    term                    estimate std_error statistic p_value lower_ci upper_ci
##    <chr>                      <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
##  1 "intercept"                5.62      0.249    22.6     0        5.13     6.11
##  2 "Product: 12 Pack Spr~    -0.018     0.328    -0.056   0.955   -0.664    0.627
##  3 "Product: Almond milk"    -2.11      0.345    -6.13    0       -2.79    -1.44
##  4 "Product: Banana $/lb"    -4.91      0.316   -15.5     0       -5.53    -4.29
##  5 "Product: Blue Bell V~    -2.91      0.429    -6.78    0       -3.75    -2.06
##  6 "Product: Campbells C~    -3.37      0.345    -9.78    0       -4.05    -2.70
##  7 "Product: Carton of e~    -2.97      0.307    -9.68    0       -3.58    -2.37
##  8 "Product: Cinnamon To~    -1.20      0.465    -2.57    0.011   -2.11    -0.281
##  9 "Product: El Milagros~    -2         0.464    -4.31    0       -2.91    -1.09
## 10 "Product: Frosted Fla~    -1.45      0.465    -3.12    0.002   -2.36    -0.536
## # i 31 more rows
```

```
## # Standardization method: refit
##
## Parameter                                | Std. Coef. |         95% CI
## ------------------------------------------------------------------------
## (Intercept)                              |       1.08 | [ 0.86,  1.31]
## Product12 Pack Sprite                    |  -9.03e-03 | [-0.33,  0.31]
## ProductAlmond milk                       |      -1.04 | [-1.37, -0.71]
## ProductBanana $/lb                       |      -2.42 | [-2.72, -2.11]
## ProductBlue Bell Vanilla Ice Cream 16 oz |      -1.43 | [-1.85, -1.02]
## ProductCampbells Chicken noodle          |      -1.66 | [-1.99, -1.33]
## ProductCarton of eggs                    |      -1.46 | [-1.76, -1.17]
## ProductCinnamon Toast Crunch 1lb         |      -0.59 | [-1.04, -0.14]
## ProductEl Milagros Tortilla Chips        |      -0.98 | [-1.43, -0.53]
## ProductFrosted Flakes 1 lb               |      -0.71 | [-1.16, -0.26]
## ProductGold Medal Flour 5 lb             |      -1.03 | [-1.38, -0.69]
## ProductGood Flow Honey 16 oz             |       0.52 | [ 0.15,  0.89]
## ProductGranny Smith Apple $/lb           |      -1.85 | [-2.15, -1.54]
## ProductGreek Yogurt                      |      -1.93 | [-2.32, -1.54]
## ProductHigh Brew Black Coffee            |      -1.39 | [-1.75, -1.03]
## ProductHoney Nut Cheerios                |      -0.83 | [-1.15, -0.52]
```

```
## ProductHorizon 2% Milk Carton           |      -0.53 | [-0.83, -0.23]
## ProductImperial Sugar                   |      -1.19 | [-1.53, -0.85]
## ProductIodized Salt                     |      -1.89 | [-2.19, -1.59]
## ProductJIF Peanut butter                |      -1.35 | [-1.67, -1.04]
## ProductJustin's almond butter           |       3.38 | [ 3.01,  3.75]
## ProductKind Oats and Honey Bar          |      -0.83 | [-1.19, -0.47]
## ProductLa Croix Box                     |      -0.48 | [-0.82, -0.15]
## ProductLand o Lakes Half & Half         |      -1.56 | [-1.91, -1.22]
## ProductLucky Charms 1 lb                |      -0.84 | [-1.29, -0.39]
## ProductMoms Tomato Sauce                |       0.74 | [ 0.37,  1.11]
## ProductNatures Own Whole Wheat Bread    |      -1.22 | [-1.63, -0.80]
## ProductNavel Orange                     |      -1.92 | [-2.29, -1.55]
## ProductNut Thins 4.25 oz                |      -1.19 | [-1.56, -0.82]
## ProductOriginal ramen                   |      -2.45 | [-2.84, -2.06]
## ProductPremium Saltine Crackers 1 lb    |      -1.11 | [-1.48, -0.74]
## ProductProgress Chicken Noodle Soup     |      -1.49 | [-1.85, -1.13]
## ProductSan Pellegrino 25.3 floz         |      -1.74 | [-2.08, -1.39]
## ProductSkinny Pop                       |      -0.96 | [-1.31, -0.62]
## ProductSmart Water 23.7 fl oz           |      -1.84 | [-2.16, -1.52]
## ProductStacy's Pita Chips               |      -0.83 | [-1.18, -0.47]
## ProductTostitos Bag                     |      -0.81 | [-1.22, -0.39]
## ProductTriscuits 9oz                    |      -1.13 | [-1.45, -0.80]
## ProductWelch's Grape Jelly              |      -1.48 | [-1.82, -1.15]
## ProductWheat Thins 8oz                  |      -1.13 | [-1.46, -0.81]
## Income10K                               |      -0.03 | [-0.07,  0.01]
```

Because the coeffient of the Income10k is -0.014, poorer people pay more for the same product since the coefficient is negative. This is because as income increases each 1k dollars, the price of the product decreases by 0.014 dollars. A one-standard deviation increase in the income of a ZIP code seems to be associated with a -0.03 standard-deviation change in the price that consumers in that ZIP code expect to pay for the same product.