Hw_9

Pranita

2025-04-21

name: Pranita Chaudhury
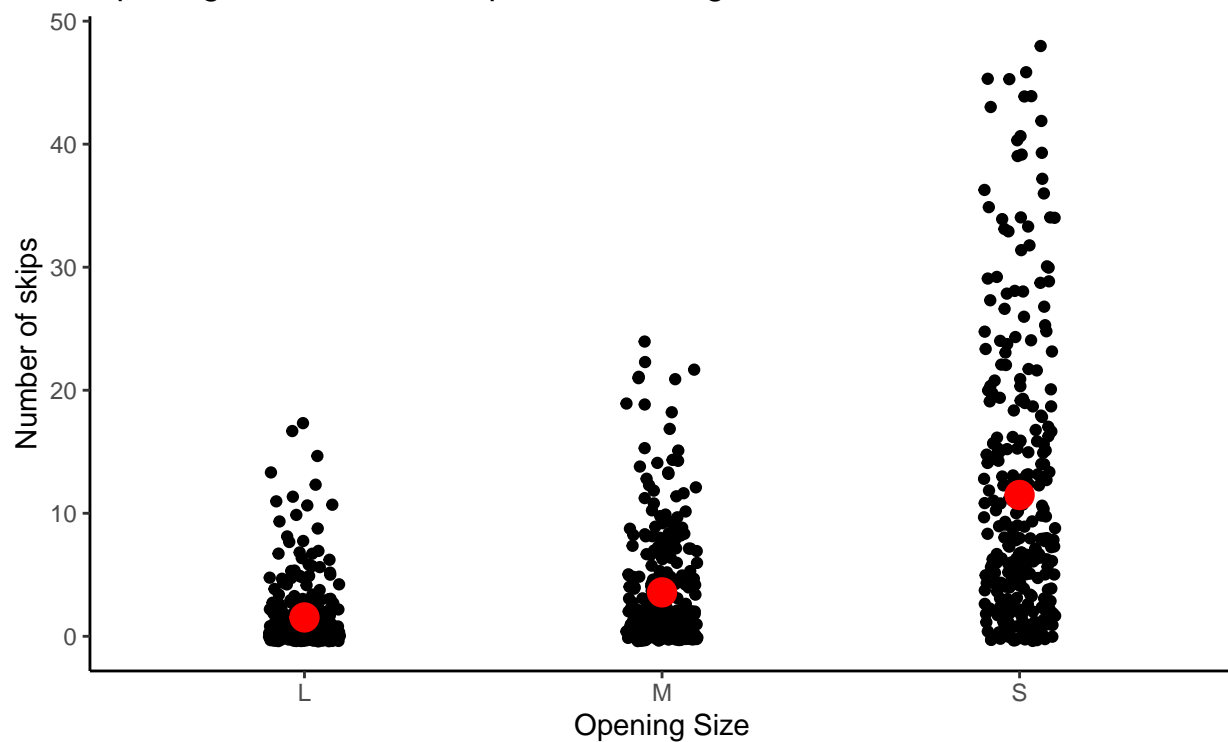
UT ID: pc28377

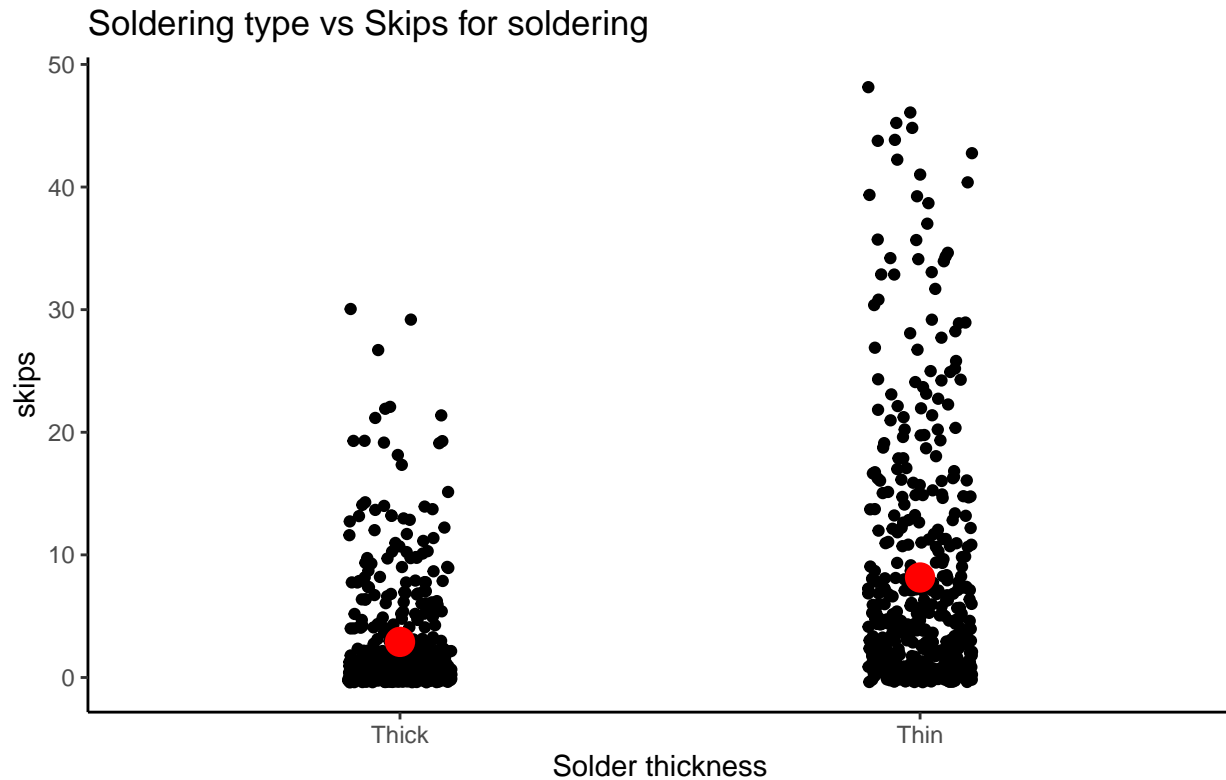**github: https://github.com/PranitaChau/Hw_9**

Problem 1

Part A

Opening vs Number of skips for soldering

Comparing small, medium, and large openings on solder gun.
The average skips for small openings are 11, for medium openings are 4 and for large openings are 2.

```
## L M  S
## 2 4 11
```

Soldering type vs Skips for soldering

Comparing thick and thin alloys used on solder guns.
It can be noted that the average skips for thick alloys are 3, while it is 8 for thin alloys.

```
## Thick  Thin
##     3     8
```

Part B

```
##         (Intercept)              OpeningM                OpeningS          SolderThin
##                   0                     2                       5                   2
## OpeningM:SolderThin OpeningS:SolderThin
##                  -1                  10

##        Intercept OpeningM OpeningS SolderThin OpeningM.SolderThin
## 2.5%   0.2160989 1.908409 4.242557   1.825323           -1.942035
## 97.5% 0.5356789 3.220736 6.242959   2.884647            0.273208
##        OpeningS.SolderThin
## 2.5%              7.44273
## 97.5%            11.37511
```

Regression model: Skips = 2(Opening_Medium) + 5(Opening_Small) + 2(SolderThin) - (Opening_Medium)(SolderThin) + 10 (Opening_Small)(SolderThin)

Part C

The baseline number of skips is 0, which represents boards with a large opening and thick solder. This intercept also refers to all interactions involving these baseline levels, since their coefficients are set to zero. The main effect for medium Openings is 2 skips, this is the effect of OpeningM in isolation. The main effect for small Openings is 5 skips, this is the effect of OpeningS in isolation. The main effect for thin solders is 2 skips, this is the effect of SolderThin in isolation. The interaction effect for OpeningM and SolderThin is -1 skips. In other words, circuits which had both medium openings and thin solders used average 1 less skips than

what you would expect from summing the individual "isolated" effects of the two variables. The interaction effect for OpeningS and SolderThin is 10 skips. In other words, circuits which had both small openings and thin solders used average 10 more skips than what you would expect from summing the individual "isolated" effects of the two variables.

Part D

If I had to recommend a combination of opening size and solder thickness to AT&T based on this analysis I would recommend they use a large opening and a thick solder since this is the most likely to result in 0 skips, which is the lowest number of skips possible.
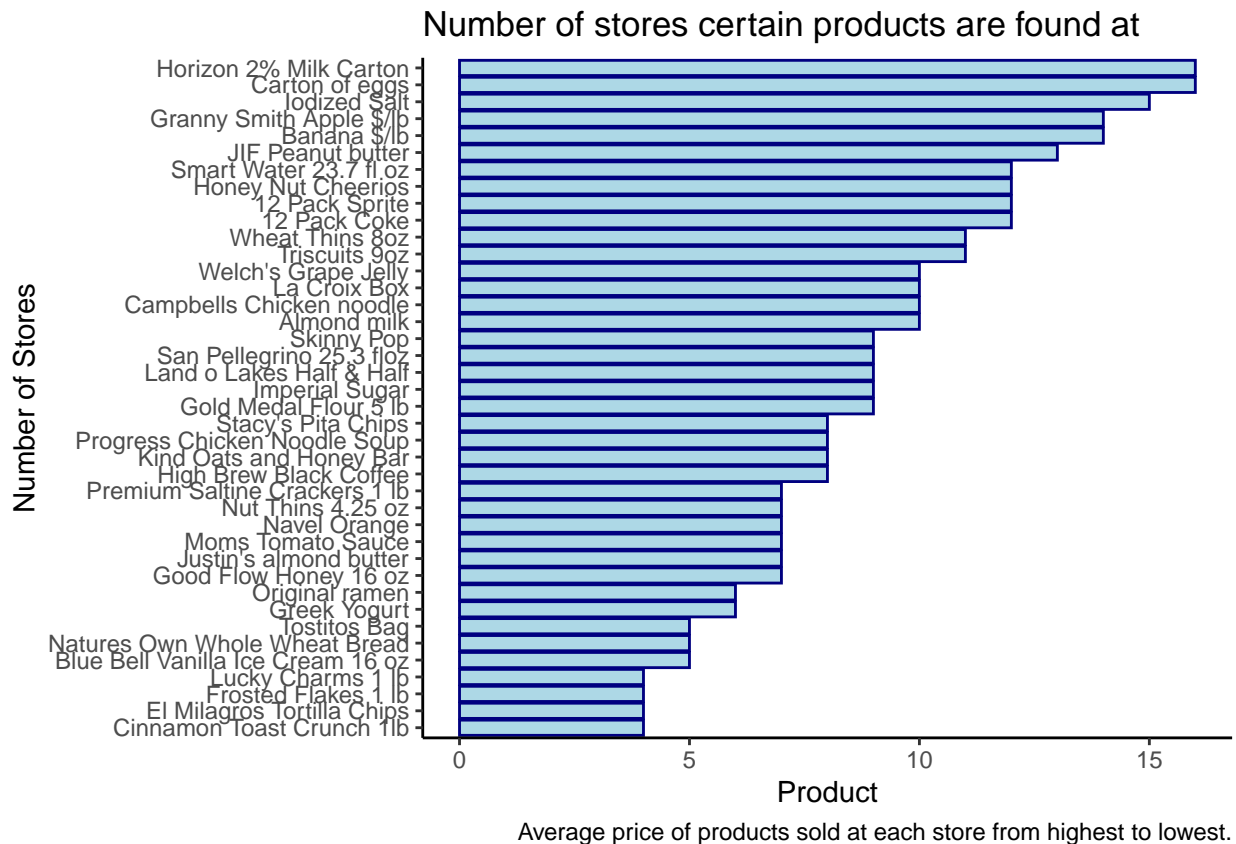
Problem 2

Part A
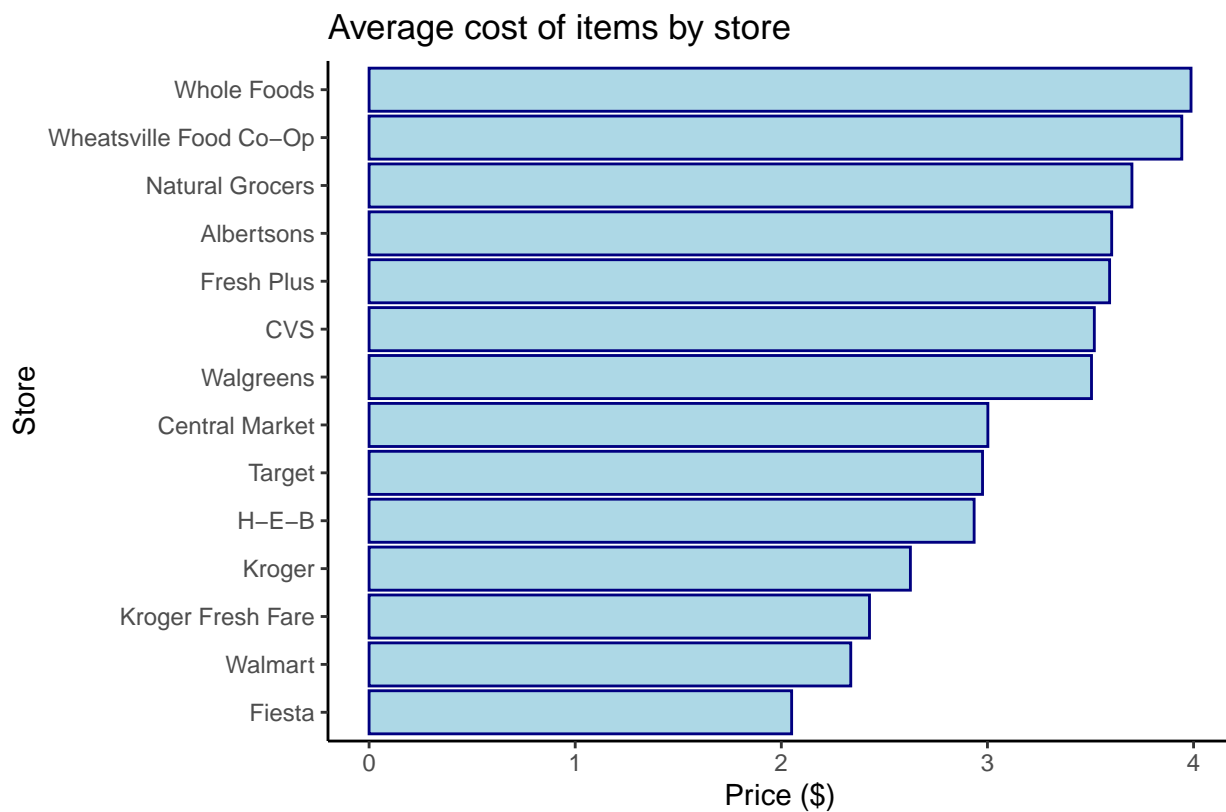
Average price of products sold at each store from highest to lowest.

We can see price differences of $0-$2 when comparing all products in the store, and the specific difference would depend no whch stores are being compared.

Part B

Number of stores certain products are found at

Average price of products sold at each store from highest to lowest.

Part C

```
## # A tibble: 6 x 7
##   term                  estimate std_error statistic p_value lower_ci upper_ci
##   <chr>                    <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
## 1 Product: Welch's Grape~  -2.94      0.33     -8.93       0    -3.58    -2.29
## 2 Product: Wheat Thins 8~  -2.27      0.32     -7.08       0     -2.9    -1.64
## 3 Type: Grocery            -0.66      0.13      -5.2       0    -0.92    -0.41
## 4 Type: High-end Grocery    -0.3      0.15     -1.97    0.05    -0.59        0
## 5 Type: Natural            -0.09      0.16     -0.58    0.56     -0.4     0.22
## 6 Type: Small Format       -0.44      0.15      -2.9       0    -0.75    -0.14
```

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

```
## # A tibble: 13 x 7
##   term                 estimate std_error statistic p_value lower_ci upper_ci
##   <chr>                   <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
## 1 Store: Walmart          -0.99      0.23     -4.25       0    -1.45    -0.53
## 2 Store: Kroger Fresh F~   -0.9      0.23     -3.87       0    -1.36    -0.44
## 3 Store: Fiesta           -0.7      0.27     -2.61    0.01    -1.23    -0.17
## 4 Store: Kroger           -0.7      0.23     -3.01       0    -1.16    -0.24
## 5 Store: H-E-B           -0.65      0.15     -4.25       0    -0.95    -0.35
## 6 Store: Central Market  -0.57      0.18     -3.24       0    -0.92    -0.23
```

```
##  7 Store: Target           -0.37      0.19    -1.97     0.05    -0.75     0
##  8 Store: Natural Grocers  -0.08      0.2     -0.41     0.68    -0.47     0.31
##  9 Store: Fresh Plus       -0.04      0.16    -0.22     0.82    -0.35     0.28
## 10 Store: CVS               0.19      0.18     1.06     0.29    -0.17     0.55
## 11 Store: Walgreens         0.22      0.18     1.19     0.23    -0.14     0.57
## 12 Store: Wheatsville Fo~   0.29      0.18     1.62     0.11    -0.06     0.64
## 13 Store: Whole Foods       0.36      0.18     2.06     0.04     0.02     0.71

## # A tibble: 13 x 7
##    term               estimate std_error statistic p_value lower_ci upper_ci
##    <chr>                 <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
##  1 Store: Whole Foods     0.36      0.18     2.06     0.04     0.02     0.71
##  2 Store: Wheatsville Fo~ 0.29      0.18     1.62     0.11    -0.06     0.64
##  3 Store: Walgreens       0.22      0.18     1.19     0.23    -0.14     0.57
##  4 Store: CVS             0.19      0.18     1.06     0.29    -0.17     0.55
##  5 Store: Fresh Plus     -0.04      0.16    -0.22     0.82    -0.35     0.28
##  6 Store: Natural Grocers -0.08      0.2     -0.41     0.68    -0.47     0.31
##  7 Store: Target         -0.37      0.19    -1.97     0.05    -0.75     0
##  8 Store: Central Market -0.57      0.18    -3.24     0       -0.92    -0.23
##  9 Store: H-E-B          -0.65      0.15    -4.25     0       -0.95    -0.35
## 10 Store: Fiesta         -0.7       0.27    -2.61     0.01    -1.23    -0.17
## 11 Store: Kroger         -0.7       0.23    -3.01     0       -1.16    -0.24
## 12 Store: Kroger Fresh F~ -0.9       0.23    -3.87     0       -1.36    -0.44
## 13 Store: Walmart        -0.99      0.23    -4.25     0       -1.45    -0.53
```

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

Part E

```
## # 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              5.11      0.22     23.0     0        4.67     5.55
##  2 Store: Albertsons      0.65      0.15      4.25    0        0.35     0.95
##  3 Store: Central Market  0.07      0.17      0.44    0.66    -0.25     0.4
##  4 Store: CVS             0.84      0.17      4.89    0        0.5      1.18
##  5 Store: Fiesta         -0.06      0.26     -0.22    0.83    -0.57     0.45
##  6 Store: Fresh Plus      0.61      0.15      4.08    0        0.32     0.9
##  7 Store: Kroger         -0.06      0.22     -0.26    0.8     -0.49     0.38
##  8 Store: Kroger Fresh F~ -0.26      0.22     -1.15    0.25    -0.69     0.18
##  9 Store: Natural Grocers 0.56      0.19      3.02    0        0.2      0.93
## 10 Store: Target          0.27      0.18      1.52    0.13    -0.08     0.63
## # i 43 more rows
```

Central Market on average charges $0.07 more for the same products as HEB. This difference is not a lot especially when compared to the other stores, so we can say that by relative difference Central Market is not price discriminating when looking at HEB prices.

Part F

```
## # A tibble: 1 x 7
##   term      estimate std_error statistic p_value lower_ci upper_ci
##   <chr>        <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
## 1 Income10K   -0.01      0.01      -1.4    0.16    -0.03     0.01
```

```
## # Standardization method: refit
##
## Parameter | Std. Coef. |        95% CI
## --------------------------------------
## Income10K |      -0.03 | [-0.07, 0.01]
```

Based on the sign of the Income10K variable, customers in poorer ZIP codes seem to pay more. I know thing because according to the model customers in poorer ZIP codes seem to pay 1 cent more for the same product on average for every $10,000 decrease in income. Since the confidence interval of -0.03 to 0.01 includes zero we cannot say this with certainty, but based on the sign of the variable alone we can say that as the income goes up the price goes down.

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.

Problem 3

Part A

This is true, as shown in the first graph (Fig A1) which compares the percent of minority residents in a ZIP code against the number of FAIR policies passed. This evidence shows that there is a positively related relationship with the number of FAIR policies and percentage of minority residents, so when the the minority percentage in a ZIP code goes up so does the number of FAIR policies passed. Additionally the simple linear regression performed below shows an r squared value of 0.5, so the data moderately fits the given equation indicating a positive moderately strong linear relationship.

Part B

This statement is false. The linear regression model and graph (Figure B1 and the model underneath) only show an interaction between the age of the house and the minority percentage, there is no association in any of the given evidence that points to an interaction with these variables and number of FAIR policies.

Part C

This statement is true. In the Figure C1 the relationship between minority percentage and the number of FAIR policies for ever 100 households is shown. From this we can gather than there is a positive liner relationship between the two, and since the red line (high fire rick) is sightly steeper we can rather that it on average has a more FAIR policies enacted as the minority percentage goes up. The minority:fire_riskLow represents the effect of both being in a minority ZIP code and having a low fire risk, and since the regression model does not show the variable for being in a minority ZIP code and having a high fire risk (minority:fire_riskHigh), we can assume it is the intercept. Since the minority:fire_riskLow has a negative estimate it is lower than the intercept (minority:fire_riskHigh) meaning that high fire risk ZIP codes have more FAIR policies enacted per 100 households.

Part D

This statement is false, adding income to the regression model changes the coefficients for the variables as well as the intercept, showing that income does have an effect on the association between FAIR policy uptake and the minority percentage.

Part E

This statement is true. After controlling for other variables minority stil has a positive sign so it is postively related, and a small p value of 0.006 showing that it is statistically significant. This shows that after controlling

for other variables in the dataset the number of FAIR policies still has a positve association with the minority percentage of a ZIP code.