HW_9

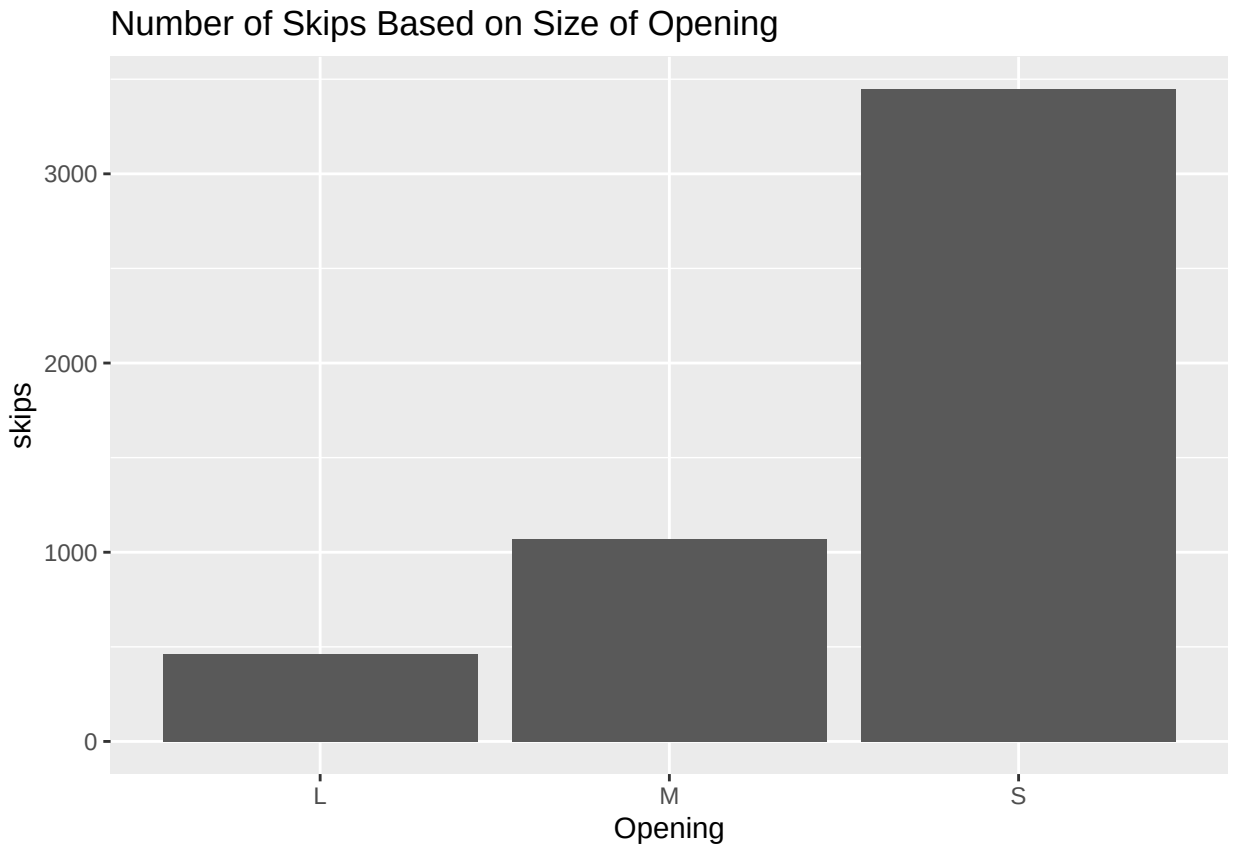
Maggie Ma

2025-04-19

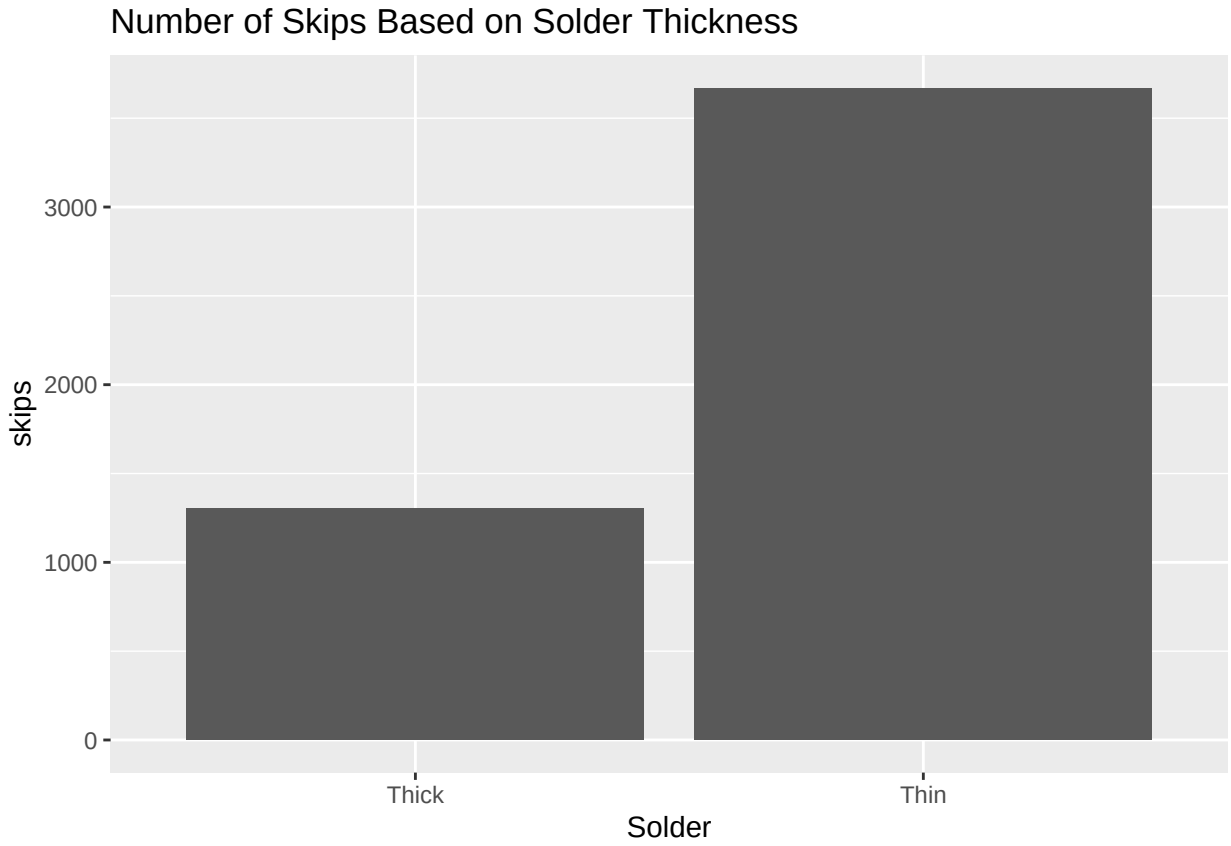
Maggie Ma, mm227339, Github Link: https://github.com/maggiema9966/Homework-9

Problem 1

Part A



This graph shows that as the opening size of the solder gun gets smaller, there is a higher rate of skipping while soldering. The number of skips for the medium opening is a little over twice the number of skips for the large opening, while the number of skips for the small opening is over 3 times the number of skips for the medium opening.



This graph shows that a thinner solder alloy has a higher rate of skipping while soldering. The number of skips for the thin solder alloy is over 3 times the number of skips for the thick solder alloy.

Part B

```
## # A tibble: 6 x 7
##   term                  estimate std_error statistic p_value lower_ci upper_ci
##   <chr>                    <dbl>     <dbl>     <dbl>   <dbl>    <dbl>    <dbl>
## 1 intercept                0.393     0.52      0.756   0.45    -0.628     1.42
## 2 Opening: M               2.41      0.736     3.27    0.001    0.962     3.85
## 3 Opening: S               5.13      0.736     6.97    0        3.68      6.57
## 4 Solder: Thin             2.28      0.736     3.10    0.002    0.836     3.72
## 5 Opening: M:SolderThin   -0.74      1.04     -0.711   0.477   -2.78      1.30
## 6 Opening: S:SolderThin    9.65      1.04      9.28    0        7.61     11.7
```

Part C

The baseline number of skips for thick solder and large opening is about 0.393, with a 95% confidence interval from -0.628 to 1.415. The main effect for the Opening variable when the Opening is M (medium) is about 2.407 with a 95% confidence interval from 0.962 to 3.851. This is the effect of when Opening is M in isolation. The main effect for the Opening variable when the Opening is S (small) is about 5.127 with a 95% confidence interval from 3.682 to 6.571. This is the effect of when Opening is S in isolation. The main effect for the Solder variable when the Solder is Thin is 2.280 with a 95% confidence interval from 0.836 to 3.724. This is the effect of when Solder is Thin in isolation. The interaction effect for when Opening is M and solder is thin is -0.740 skips with a 95% confidence interval from -2.782 to 1.302. In other words, soldering with a medium solder gun opening and a thin solder alloy yield average skipping numbers that are 0.740

skips less than what you would expect from summing the individual "isolated" effects of the two variables. The interaction effect for when Opening is S and solder is thin is 9.653 skips with a 95% confidence interval from 7.611 to 11.696. In other words, soldering with a small solder gun opening and a thin solder alloy yield average skipping numbers that are 0.740 skips less than what you would expect from summing the individual "isolated" effects of the two variables.

Part D

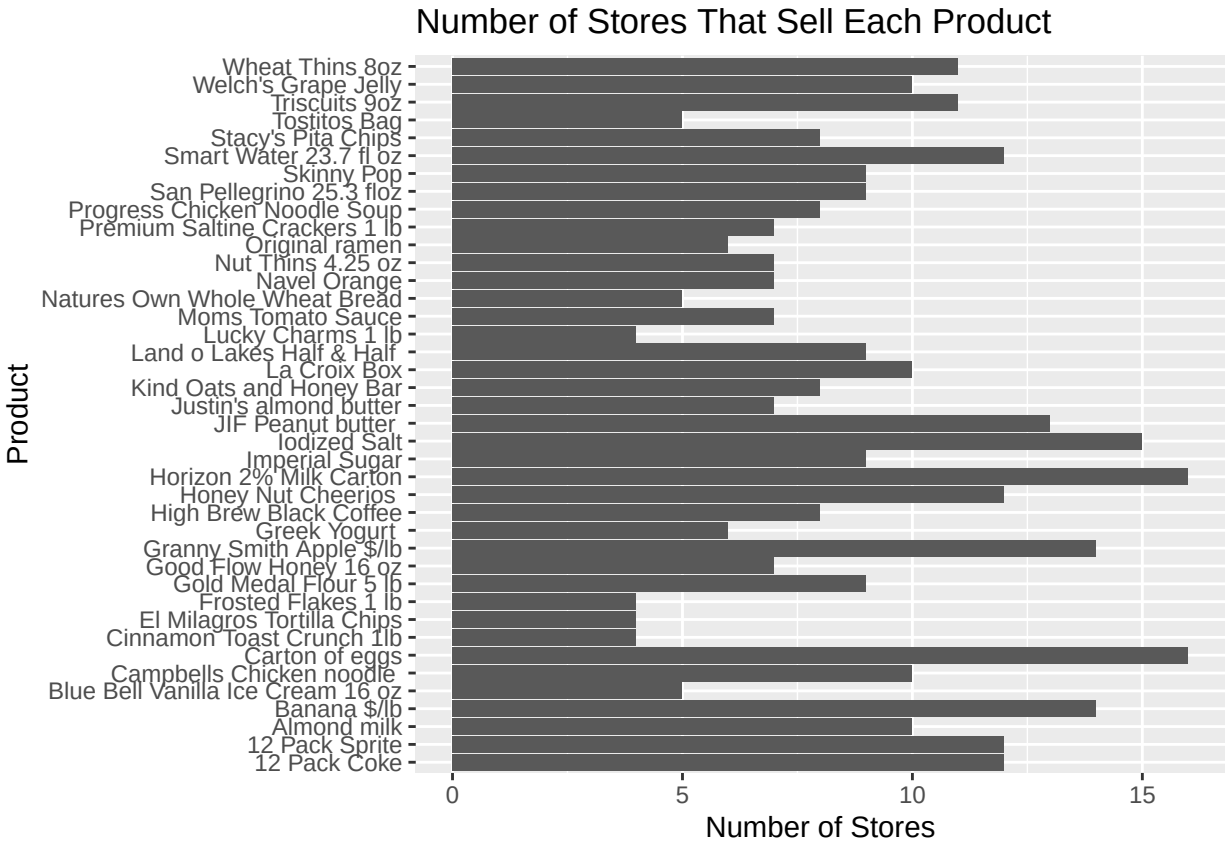
If I had to recommend a combination of Opening size and Solder thickness to AT&T based on this analysis, I would recommend a solder gun opening of size large and thick solder alloy. When the other effects (such as medium opening or thin solder) are added, they increase the number of skips, and even though the interaction between medium opening and thin solder would decrease the number of skips, the addition of the main effects of the medium opening and thin solder outweigh the small amount the interaction subtracts. Thus, the baseline of large opening and thick alloy is the best combination.

Problem 2

Part A

This graph shows the average price among all products at each store. A few notable stores that are on the more expensive side are Whole Foods and Wheatsville Food Co-op, while a few stores on the cheaper side are Fiesta and Walmart.

Part B



This graph shows the number of stores that sell each product. The most common products are milk and eggs with 16 stores, while the least common products are Lucky Charms, Frosted Flakes, tortilla chips, and Cinnamon Toast Crunch with 4 stores.

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 2 stores that seem to charge the lowest prices when comparing the same product are Kroger Fresh Fare and Walmart. The 2 stores that seem to charge the highest prices when comparing the same product are Wheatsville Food Co-op and Whole Foods.

Part E

I think Central Market charges a similar amount to HEB for the same product. Although it charges about 7 cents more (since HEB charges 0.64595932 less than Albertsons (our baseline) and Central Market charges 0.57338651 less than Albertsons), the entire range of possible charges is from -0.99254892 ($0.99 less than Albertsons) to 0.36415850 ($0.36 more than Albertsons). When we compare our values of -0.57338651 for Central Market and -0.64595932 for HEB, it is not as big of a difference compared to differences among other stores, as demonstrated by the range that it could differ.

Part F

Based on the negative sign of the Income10K coefficient, consumers in poorer ZIP codes seem to pay more for the same product, on average. The negative sign of the Income10k (-1.408973e-10) tells us that as income increases, people pay less for their groceries. Thus, if you have lower income (i.e. poorer ZIP codes), then you pay more for your groceries. 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.

Problem 3

A. ZIP codes with a higher percentage of minority residents tend to have more FAIR policies per 100 housing units.

TRUE. In figure A1, we can see that as the percentage of minority residents in a ZIP code increases, the number of FAIR policies per 100 housing units also increases. The simple linear regression also shows this, as it has an r^2 value of 0.5163908 (which shows a moderately strong relationship) and a positive slope of 0.014 (with a 95% confidence interval of 0.009 to 0.018, which doesn't include 0), indicating a positive trend between the percentage of minority residents in a ZIP code and the number of FAIR policies per 100 housing units.

B. The evidence suggests an interaction effect between minority percentage and the age of the housing stock in the way that these two variables are related to the number of FAIR policies in a ZIP code.

UNDECIDABLE. In the linear regression model model_b, there is no interaction effect shown as evidence, so we cannot say much about how the interaction effect between minority percentage and the age of the housing stock affects the number of FAIR policies in a ZIP code. If the linear regression model included the interaction between minority:age, then we could look at the confidence interval of the interaction variable to see whether there was a significant interaction effect.

C. The relationship between minority percentage and number of FAIR policies per 100 housing units is stronger in high-fire-risk ZIP codes than in low-fire-risk ZIP codes.

FALSE. The main effect of low fire risk has a 95% confidence interval of -0.995 to 0.109, which contains 0. The interaction between minority and low fire risk also has a 95% confidence interval that contains 0, from -0.012 to 0.01. Because these variables have confidence intervals that contain 0, it is possible that there is no effect from fire risk on the relationship between minority percentage and number of FAIR policies per 100 housing units.

D. Even without controlling for any other variables, income "explains away" all the association between minority percentage and FAIR policy uptake.

FALSE. A corrected statement would be that without controlling for any other variables, income "explains away" part of the association between minority percentage and FAIR policy uptake. Looking at the model_D2 linear regression model, although the income accounts for -0.074 for every dollar the income increases, the minority variable still explains 0.01 for every additional percentage of minority residents in a ZIP code, with a 95% confidence interval from 0.004 0.015 (which does not contain 0). Although it may not seem like a lot, the baseline is only 1.08, and when the minority percentages add up, it can still change the number of FAIR policies.

E. Minority percentage and number of FAIR policies are still associated at the ZIP code level, even after controlling for income, fire risk, and housing age.

TRUE. Looking at the model_E linear regression model, the minority variable still accounts for 0.008 for every additional percentage of minority residents in a ZIP code, with a confidence interval from 0.003 to 0.014 (which doesn't contain 0). Similarly to what was said above, when the minority percentages are high, it can still change the number of FAIR policies.