

Homework 9

Prady Kandi, EID: prk599

2025-04-21

Contents

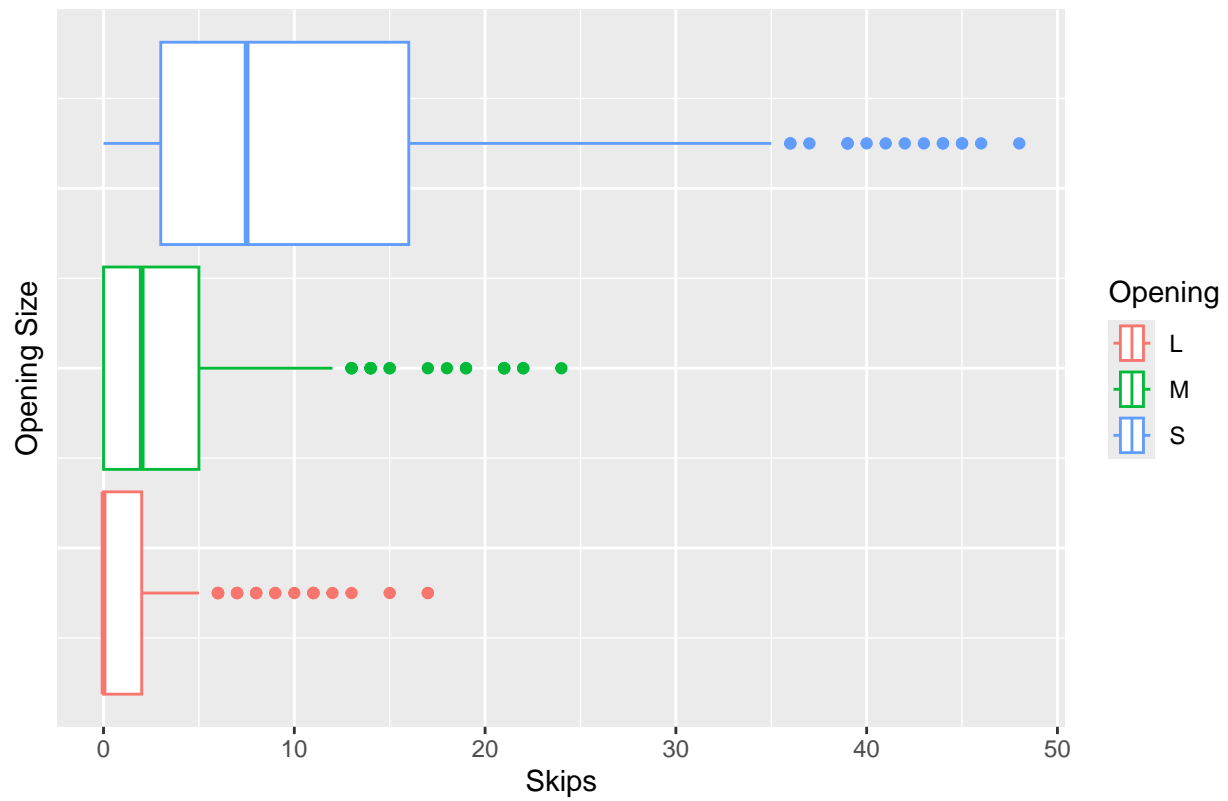
Problem 1	2
Part A	2
Part B	3
Part C	3
Part D	3
Problem 2	4
Part A	4
Part B	5
Part C	5
Part D	5
Part E	6
Part F	6
Problem 3	6

The link to the Github repo containing the R file can be found [here](#).

Problem 1

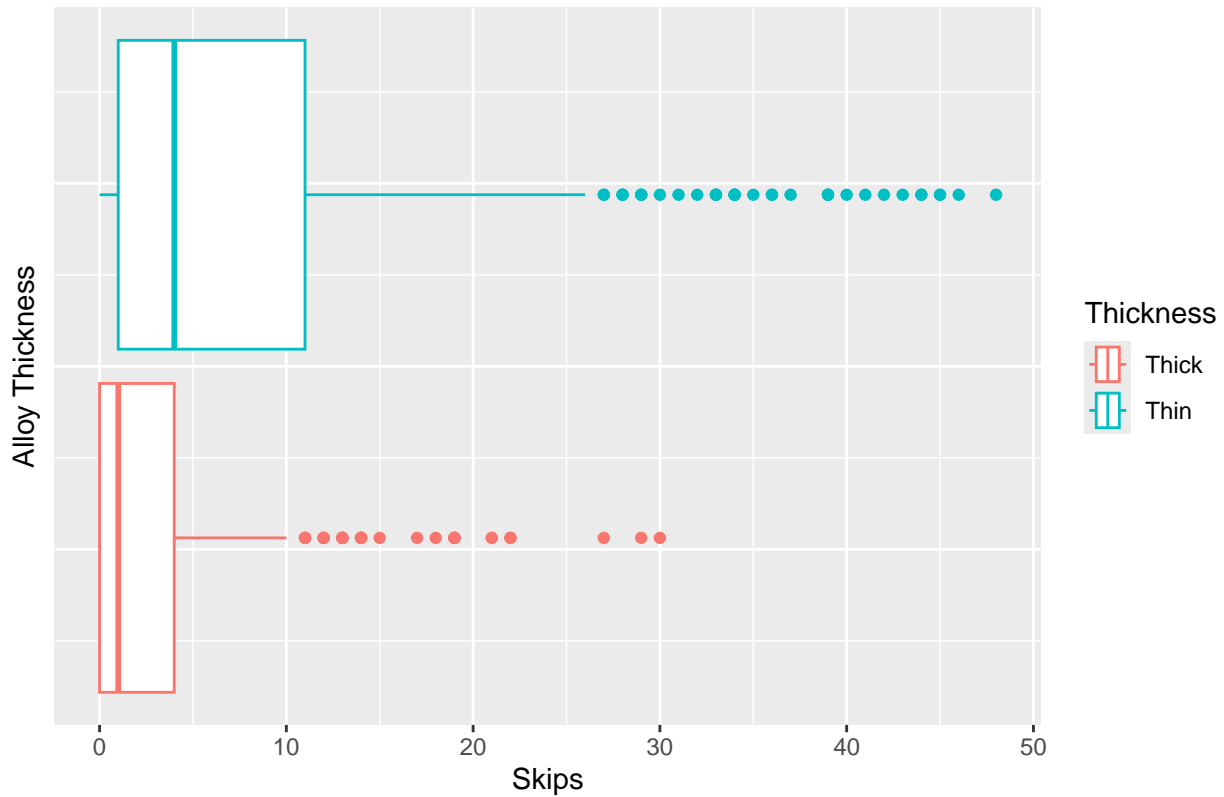
Part A

Distribution of Skips by Opening Size



Caption: Graph showing the distribution of skips by opening size

Distribution of Skips by Alloy Thickness



Caption: Graph showing the distribution of skips by soldering thickness

Part B

Term	Estimate	Lower_Bound	Upper_Bound
intercept	0.39	-0.63	1.41
Opening: M	2.41	0.96	3.85
Opening: S	5.13	3.68	6.57
Solder: Thin	2.28	0.84	3.72
Opening: M:SolderThin	-0.74	-2.78	1.30
Opening: S:SolderThin	9.65	7.61	11.70

Part C

The baseline level of skips, where the opening is large and the alloy is thick, is 0.39. If we keep the baseline of a thick alloy but use a medium opening, an additional 2.41 skips is expected, and 5.13 skips if a small opening is used. However, if a large opening is used with a thin alloy, an extra 2.28 skips can be expected to occur. With a medium opening and thin alloy, there will actually be an average of -0.74 less skips, but a small opening with a thin alloy yields the worst results, adding an average of 9.65 skips.

Part D

When setting a regression model with a large opening and thick solder as a baseline, every other combination resulted in more skips except the Medium opening with a thin alloy. Thus, if AT&T's goal is to cut down on skips, the best combination would be to use a medium opening size and a thin solder thickness.

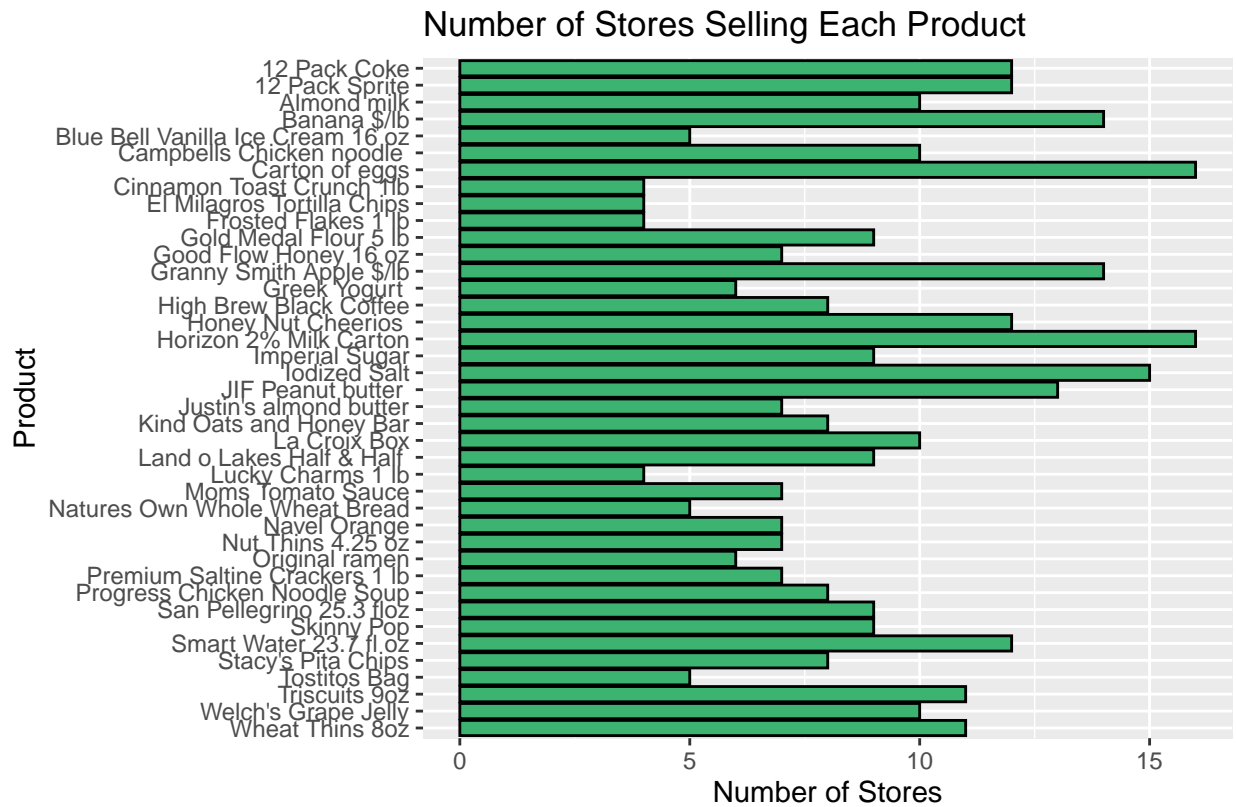
Problem 2

Part A



Graph displays the average price across all products at each store

Part B



This graph displays the availability of each product, i.e. how many stores have each product available

Part C

Store_Type	Estimate	Lower_Bound	Upper_Bound
Type: Convenience	0.66	0.41	0.92

Compared with ordinary grocery stores (like Albertsons, HEB, or Krogers), convenience stores charge somewhere between 41 and 92 cents more for the same product, with the mean being 66 cents.

Part D

Store	Estimate	Lower_Bound	Upper_Bound
Store: Whole Foods	0.36	0.02	0.71
Store: Wheatsville Food Co-Op	0.29	-0.06	0.64
Store: Walgreens	0.22	-0.14	0.57
Store: CVS	0.19	-0.17	0.55
Store: Fresh Plus	-0.04	-0.35	0.28
Store: Natural Grocers	-0.08	-0.47	0.31
Store: Target	-0.37	-0.75	0.00
Store: Central Market	-0.57	-0.92	-0.23
Store: H-E-B	-0.65	-0.95	-0.35
Store: Fiesta	-0.70	-1.23	-0.17
Store: Kroger	-0.70	-1.16	-0.24

Store	Estimate	Lower_Bound	Upper_Bound
Store: Kroger Fresh Fare	-0.90	-1.36	-0.44
Store: Walmart	-0.99	-1.45	-0.53

Observing the Estimate statistic, the stores with the 2 most expensive prices per product are Whole Foods and Wheatsville Food Co-Op, with a 36 and 29 cent increase above the baseline, respectively. Walmart and Kroger Fresh Fare have the lowest prices, with 99 and 90 cents below the baseline.

Part E

When examining the difference between Central Market's and HEB's prices, we want to account for difference in type of products being sold. After doing so with the previous model, the difference in true pricing between HEB and Central Market is 8 cents, with Central Market being 57 cents below baseline and HEB being 65 cents below. Among the 13 stores, the largest difference is 135 cents, so we can expect around a 10 cent difference between adjacent stores. Considering the difference of 7 cents between these two, we can say HEB and Central Market charge similar prices for the same product.

Part F

```
## Income10K
##      -0.014
```

The Income10K coefficient is negative, which indicates that poor people pay more for the same groceries than those who are richer. A negative coefficient shows that the regression line slopes down and to the right, which in the context of this graph is a higher income (x) leading to a lower price (y).

Parameter	Std_Coefficient	CI	CI_low	CI_high
Income10K	-0.0315971	0.95	-0.074066	0.0108718

A one-standard deviation increase in the income of a ZIP code seems to be associated with a .03 standard-deviation decrease in the price that consumers in that ZIP code expect to pay for the same product.

Problem 3

A. True. No other factors are present here, so looking at Figure A1 does in fact confirm that there is a correlation between minority percentage and FAIR policies. The 51.6% also demonstrates that the minority category makes up a majority of the variation of FAIR policies.

B. This statement is undecidable. From the previous statement, we know that minority and FAIR policies are positively correlated. In figure B1, we can also see a positive correlation between housing age and minority percentage. Unfortunately, there is not model or regression table that focuses on all three of these values at once, so it is impossible to determine if the two are synergistic, or one is simply an effect of the other. In order to confirm whether minority percentage and age have an interaction, a regression model contrasting these again price would be best.

C. False, the model_C table shows that both coefficients are near 0 in the regression model. They both have equal influence, that is, none. To further reinforce this, we can see the that the slope's of the red and blue line in Figure C1 are near identical.

D. This is false. When observing the differences between model_D1 and model_D2, the minority coefficient is almost completely gone in the second table, reduced to only 0.01. This indicates that it is no longer a substantial factor in the regression graph, however, in model_E, fire has a more substantial coefficient than

income, showing that income does not account for all of the difference. Even if minority disappears, income cannot accurately account for the difference.

E. True. Observing the coefficient in the “Multiple Predictors” table for minority, we can see it is still positive. The 95% interval also is above 0. This indicates that minority is still significant in predicting the FAIR policies, even after factoring everything else out.