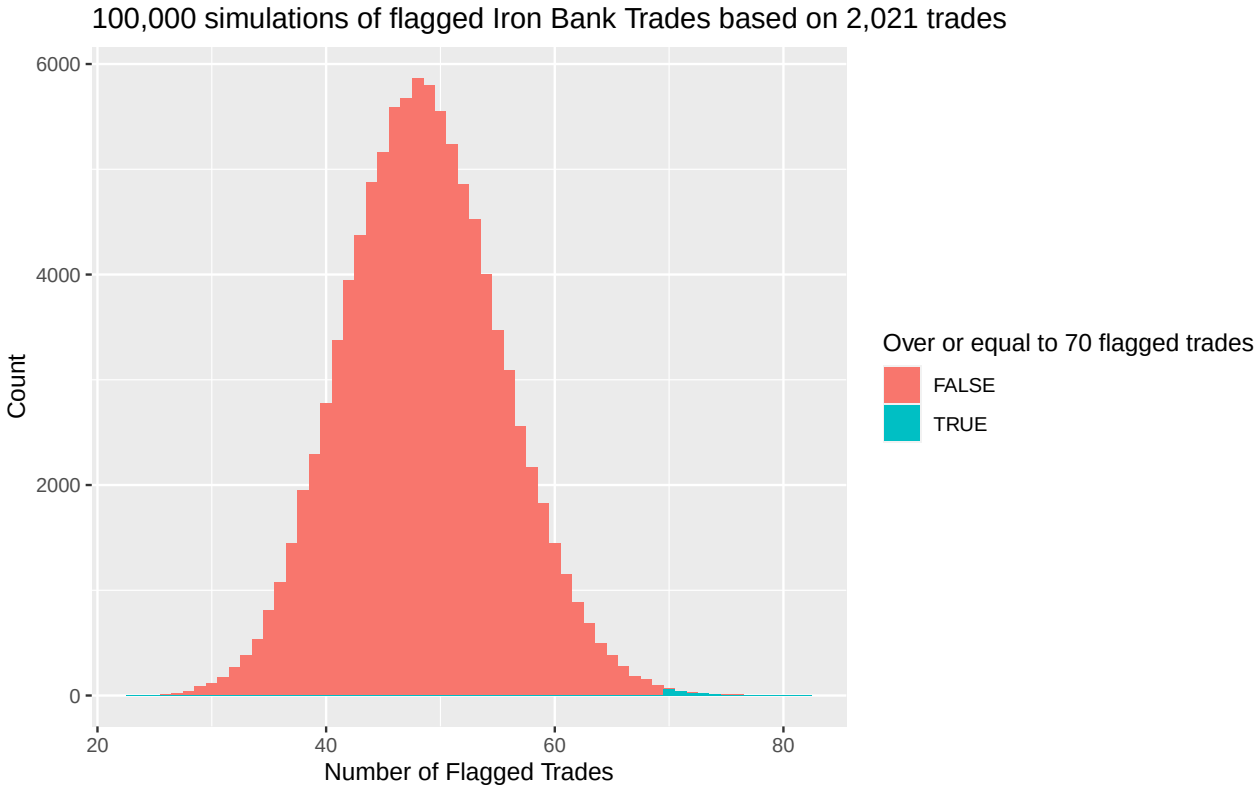
SDS HW 5

Gianluca Bollo (gb25625) - https://github.com/gianlucabollo/HW5-SDS315

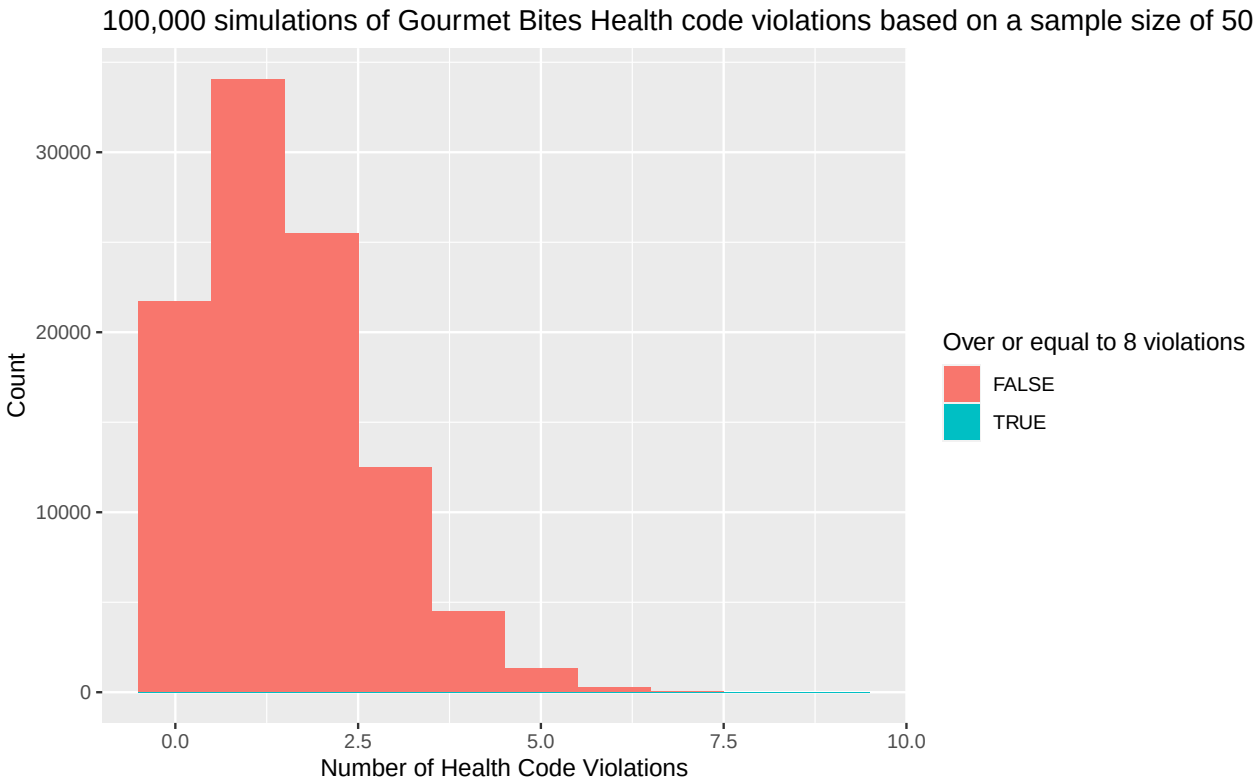
2/29/2024

Iron Bank Suspicious Trading Activity Investigation



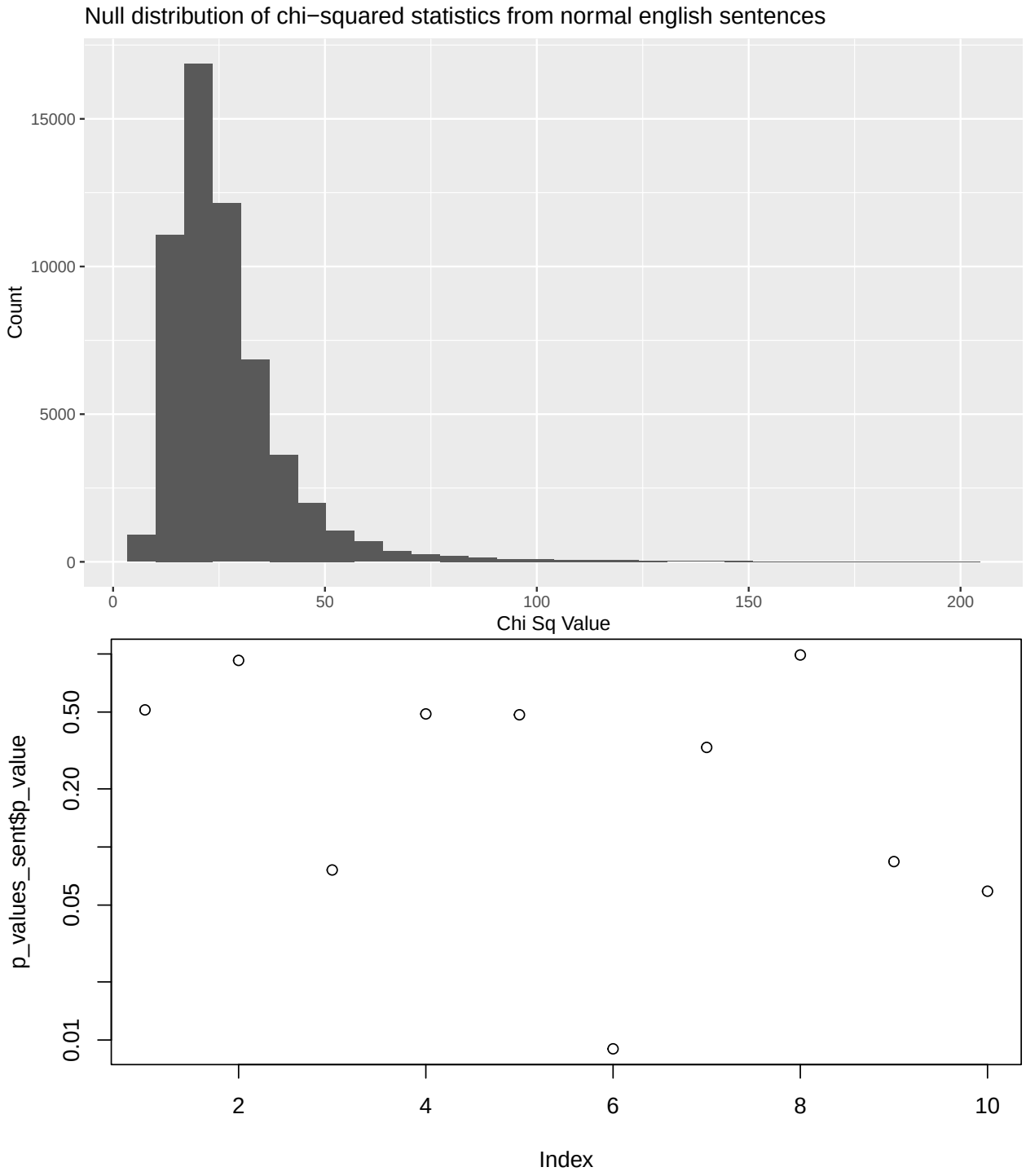
- The null hypothesis for this test is that the cluster of flagged Iron Bank trades was truly random and the data can be explained by the baseline probability of any legal trade being flagged (2.4%).
- The test statistic I used is the number of flagged Iron Bank Trades out of the total 2,021, which is 70. This test statistic is used to measure evidence in the data against the null hypothesis.
- The p-value of this test statistic is 0.00185.
- Due to the p-value of 0.00185 being very small and "statistically significant", the null hypothesis that the 70 flagged trades is consistent with the variability of the baseline trade-flag probability does not seem that very plausible, and there is likely something in need of explanation.

Gourmet Bites Health Code Vilation Investigation



- The null hypothesis for this test is that the cluster of Gourmet Bite locations with violations can be explained by random issues, consistent with the Health Department's baseline probability of any given restaurant inspection resulting in a health code violation (3.0%).
- The test statistic I used is the number of Gourmet Bites health code violations out of the total 50 locations, which is 8. This test statistic is used to measure evidence in the data against the null hypothesis.
- The p-value of this test statistic is 1.5×10^{-4}.
- Due to the p-value of 1.5×10^{-4} being very small and "statistically significant", the null hypothesis that the 8 health code violations is consistent with the Health Departments 3% baseline probability of a violation does not seem plausible. Therefore, the Health Department should decide to take action and further pursue this issue.

LLM Watermark Identification




```
## # A tibble: 10 x 1
##    p_value
##      <dbl>
##  1   0.513
##  2   0.926
##  3   0.076
##  4   0.489
```

```
##  5   0.484
##  6   0.009
##  7   0.328
##  8   0.988
##  9   0.084
## 10   0.059
```

The watermarked sentence made by a LLM is "Feeling vexed after an arduous and zany day at work, she hoped for a peaceful and quiet evening at home, cozying up after a quick dinner with some TV, or maybe a book on her upcoming visit to Auckland." I know this because when using the null distribution and the sentences' chi squared statistic as a test statistic, its p-value is 0.009, which is relatively smaller than the p-values of the other sentences. Is also satisfies being below the 0.05 threshold, therefore deeming it "significant". This can be visibly seen above in the scatter plot, with a logged y-axis.