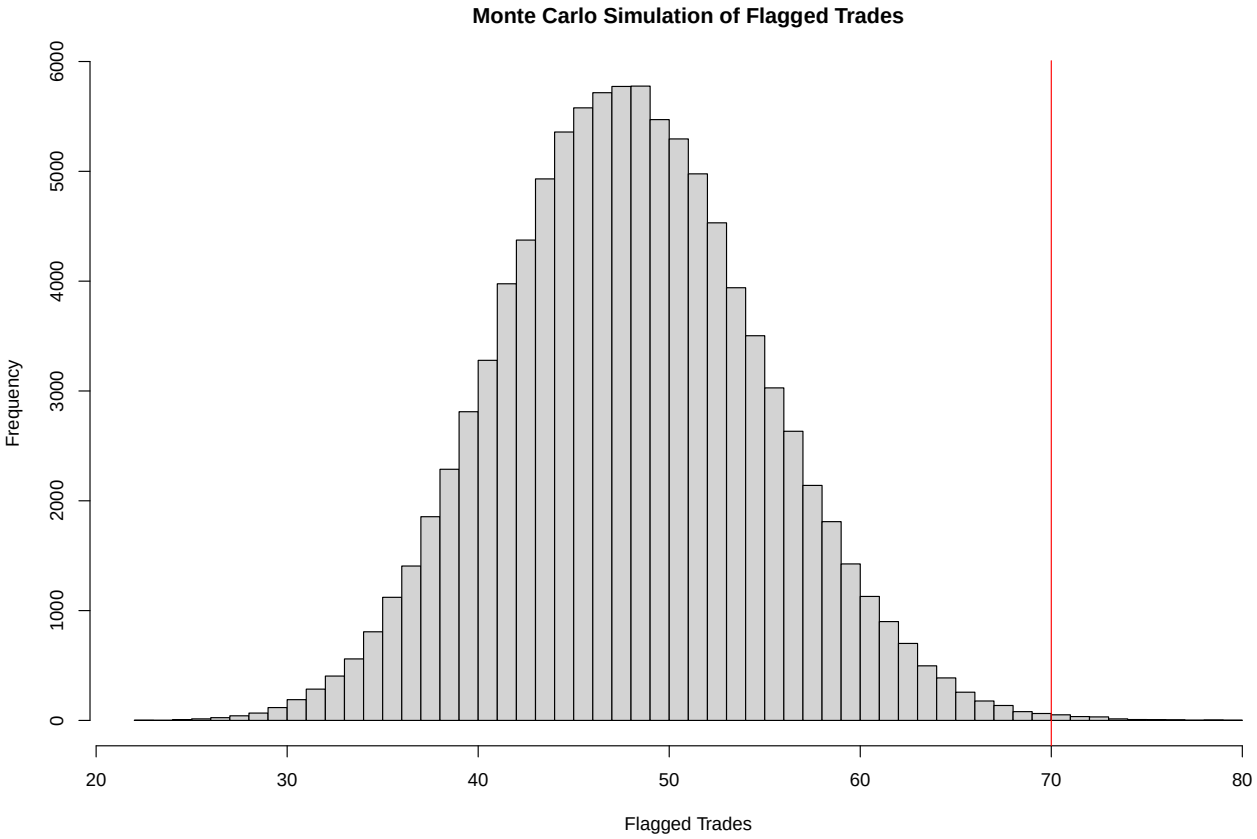
SDS315 HW 3

Samuel Li

2/12

https://github.com/SamuelLi1225/SDS_HW4

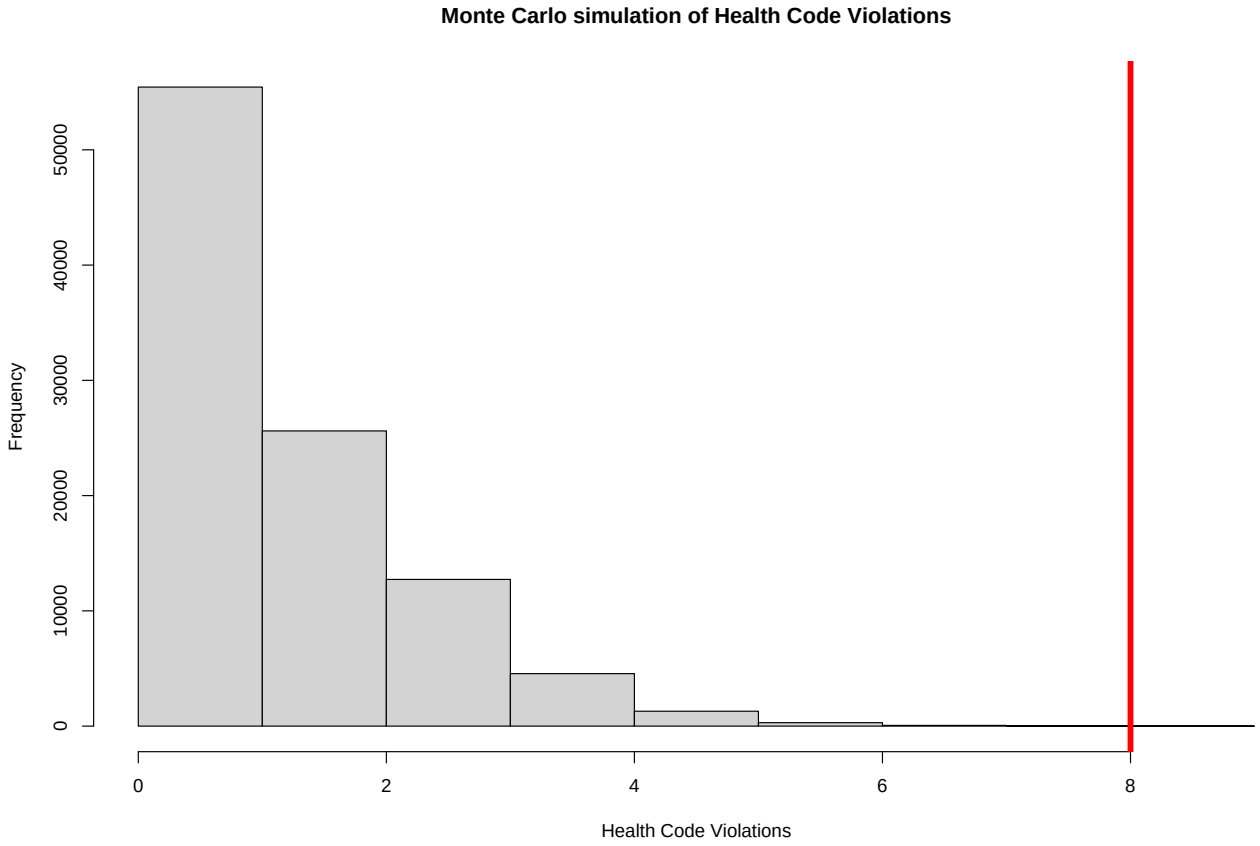
#Problem 1



1. Null Hypothesis: The 70 flagged trades out of 2021 from the Iron Bank employees were due to random variation. 2. Test Statistic: The number of flagged trades in the sample of 2021 trades. 3: Plot shown 4: P-value is 0.00181 5. The P-value is 0.00187 which is less than 0.05, so we reject the null hyptothesis. There was likely a suspicious trade.

#Problem 2

```
## [1] 0.00011
```



1. Null Hypothesis: The 8 health code violations out of the 50 health inspections of Gourment Bites were due to random variation. 2. Test Statistic: The number of health code violtions in the sample of the 50 health inspections 3: Plot shown 4: P-value is 0.00015 5. The P-value is 0.00015 which is less than 0.05, so we reject the null hyptothesis. There was likely health code violations.

#Problem 3

```
## [1] 72 60 48 36 24
```

```
##
##  Chi-squared test for given probabilities
##
## data:  observed_counts
## X-squared = 12.426, df = 4, p-value = 0.01445
```

H0/Null Hypothesis: The distribution of jurors picked by the judge is unbiased. Test Statistic: 12.426

P-Value: 0.01445, so we reject the null hypothesis. It is likely that the judge had bias when choosing the jurors

This suggests possible existence of systematic bias in jury selection. Other explanations could be that certain races were underrepresented because of a lack of registered people or too many exemptions.

We can investigate further by gathering more data from other judges, so that we can sample across a bigger sample size in order for a more accurate answer.

#Problem 4

```
## Mean: 26.95035
```

```
## Standard Deviation: 15.7733
```

```
## Min: 4.689436
```

```
## Max: 199.2418
```

```
##    Sentence_ID Chi_Squared  P_Value
## 1            1   22.930848 0.600573
## 2            2   13.051050 0.810893
## 3            3   46.285861 0.110130
## 4            4   23.546278 0.585433
## 5            5   23.676149 0.582221
## 6            6   96.452677 0.000005
## 7            7   28.271419 0.466626
## 8            8    9.635023 0.863846
## 9            9   44.928631 0.127187
## 10          10   49.960559 0.072309
```

```
## The sentence likely generated by the LLM is 6
```