HW1

Sebastian Kirkpatrick

2024-09-05

Question 1

a)

```
set.seed(20240823)
x <- rnorm(1000,0,1)

x_filtered <- x[x >= 1 | x <= 0]
length(x_filtered)
```

```
## [1] 666
```

```
mean(x_filtered)
```

```
## [1] -0.2779535
```

666 values remained with a mean of -0.278

b)

```
n_mean <- function(n) {
  x <- rnorm(n,0,1)
  x_filtered <- x[x >= 1 | x <= 0]
  mean <- mean(x_filtered)
  return(mean)
}

n_mean(100)
```

```
## [1] -0.004793571
```

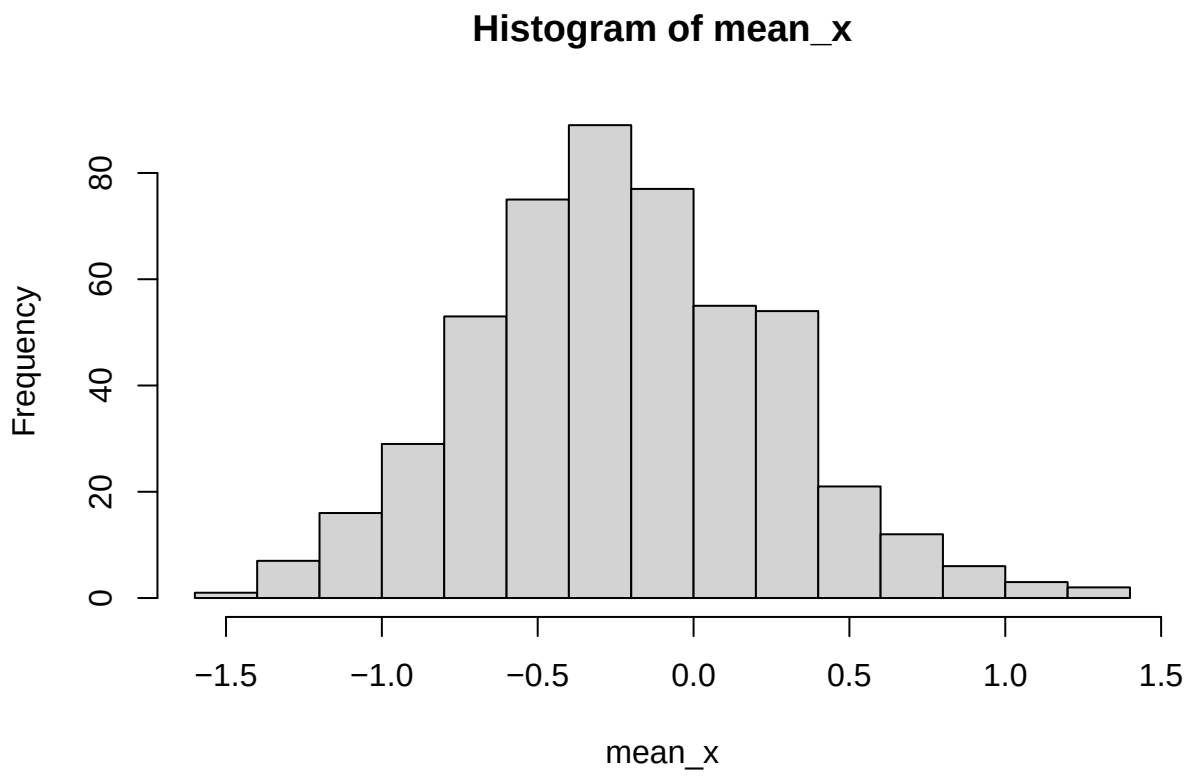
Mean value of -0.005

c)

```
n <- 500
mean_x <- rep(0, n)

for (i in 1:n) {
  mean_x[i] <- n_mean(10)
}
```

```r
hist(mean_x)
```




```r
max(mean_x)
```

```
## [1] 1.312787
```

```r
min(mean_x)
```

```
## [1] -1.454899
```

Maximum of 1.313, minimum of -1.455

Question 2

a)

```r
names(taylor_all_songs)
```

```
##  [1] "album_name"          "ep"                  "album_release"
##  [4] "track_number"        "track_name"          "artist"
##  [7] "featuring"           "bonus_track"         "promotional_release"
## [10] "single_release"      "track_release"       "danceability"
```

```
## [13] "energy"              "key"                 "loudness"
## [16] "mode"                "speechiness"         "acousticness"
## [19] "instrumentalness"    "liveness"            "valence"
## [22] "tempo"               "time_signature"      "duration_ms"
## [25] "explicit"            "key_name"            "mode_name"
## [28] "key_mode"            "lyrics"
```

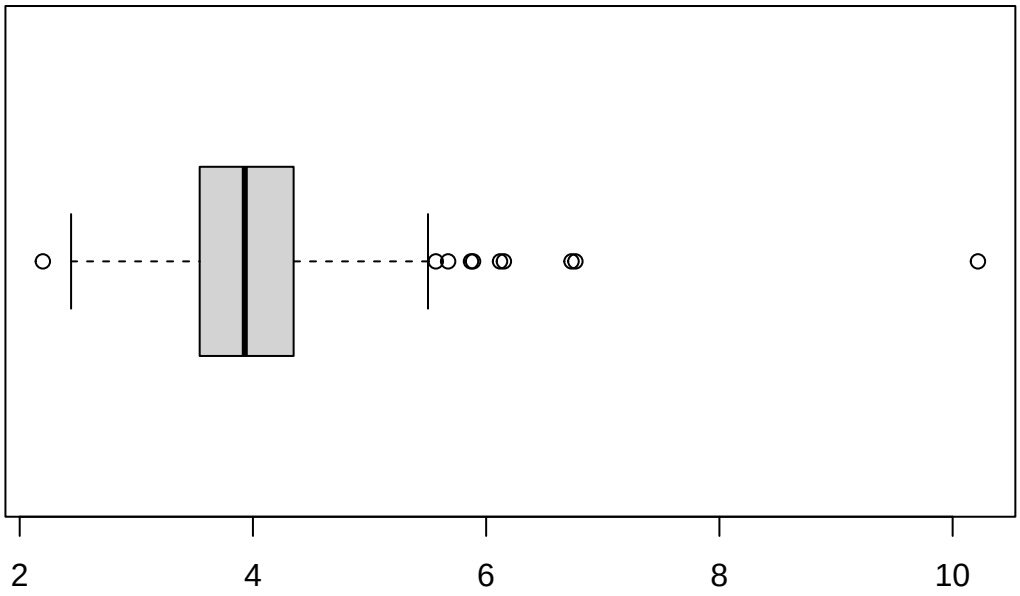
```
nrow(taylor_all_songs)
```

```
## [1] 356
```

29 variables and 356 observations, will not be naming them for you

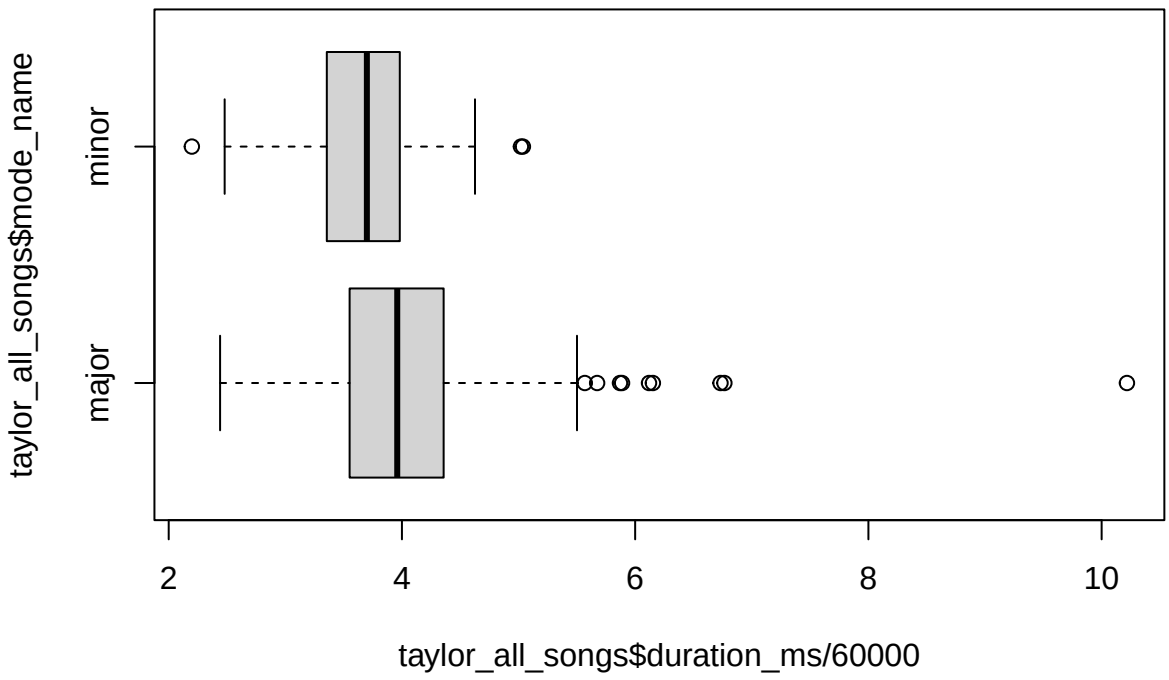
b)

```
boxplot(taylor_all_songs$duration_ms / 60000, horizontal = T)
```



c)

```
boxplot(taylor_all_songs$duration_ms / 60000 ~ taylor_all_songs$mode_name, horizontal = T)
```



d)

```
taylor_all_songs |>
  select(album_name, duration_ms) |>
  na.omit() |>
  group_by(album_name) |>
  summarize(total_mins = sum(duration_ms, na.rm=T) / 60000) |>
  arrange(-total_mins)
```

```
## # A tibble: 16 x 2
##    album_name                          total_mins
##    <chr>                                    <dbl>
##  1 Red (Taylor's Version)                   131.
##  2 THE TORTURED POETS DEPARTMENT            123.
##  3 Fearless (Taylor's Version)              107.
##  4 Speak Now (Taylor's Version)             105.
##  5 Red                                       90.2
##  6 Midnights                                 85.2
##  7 1989 (Taylor's Version)                   81.3
##  8 Speak Now                                 79.5
##  9 Fearless                                  79.4
## 10 evermore                                  69.1
## 11 folklore                                  67.1
## 12 Lover                                     61.9
## 13 1989                                      61.2
```

```
## 14 reputation                                55.8
## 15 Taylor Swift                              53.5
## 16 The Taylor Swift Holiday Collection       19.1
```

Red (Taylor's Version) is the longest with 130.7 minutes of music

Question 3

```
n <- 100000
results <- rep(0, n)

for (i in 1:n) {

initial_roll <- sum(sample(1:6,2, replace = TRUE))

if (initial_roll == 7 | initial_roll == 11) {
  result <- 1
} else {
  repeat {
    roll <- sum(sample(1:6, 2, replace = TRUE))
    if (roll == initial_roll) {
      result <- 1
      break
    } else if (roll == 7 | roll == 11) {
      result <- 0
      break
    }
  }
}

results[i] <- result
}

mean(results)
```

```
## [1] 0.46688
```

About a 46.7% chance to win this weird version of craps

Question 4

```
n <- 100000
roll_num <- rep(0, n)

for (i in 1:n) {
  rolls <- 0
  roll <- 0
  while (roll != 6) {
    roll <- sample(1:6, 1)
```

```
    rolls <- rolls + 1
  }
  roll_num[i] <- rolls
}

mean(roll_num)
```

```
## [1] 6.0367
```

On average, it takes 6 rolls to roll a six (oh my goodness this is groundbreaking!!!!!!!!)