Healthcare Worker Visit Analysis

```
options(warn = -1)
suppressPackageStartupMessages(library(dplyr))
suppressPackageStartupMessages(library(ggplot2))
```

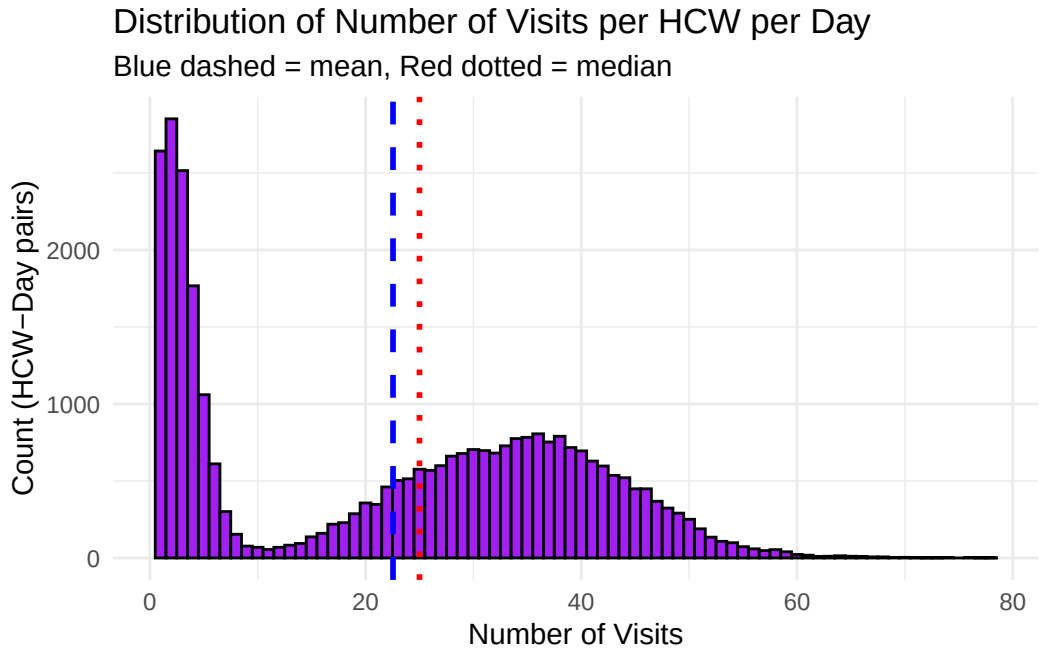
Healthcare Worker Visit Analysis

This document provides analysis of the visit rates and behaviors of health care worker agents (HCWs) in the simulation. All results and visualizations related to HCW visits will be presented here.

```
library(dplyr)

df1 <- read.table("visit_data.txt", header = TRUE, sep = ",", stringsAsFactors = FALSE)
df2 <- df1[df1$patientLocation == 'Ward', ]


df2$visitDay <- floor(df2$visitTime)

# Remove first 90 visitTime values
df2_filtered <- df2 %>% arrange(visitTime) %>% slice(-(1:90))


# Histogram: for each hcwId, each day, distribution of number of visits
visits_per_day <- df2_filtered %>% group_by(hcwId, visitDay) %>% summarise(n_visits = n())
```

```
`summarise()` has grouped output by 'hcwId'. You can override using the
`.groups` argument.
```

```
library(ggplot2)
ggplot(visits_per_day, aes(x = n_visits)) +
    geom_histogram(binwidth = 1, fill = "purple", color = "black") +
    geom_vline(aes(xintercept = mean(visits_per_day$n_visits)), color = "blue", linetype = "d

    geom_vline(aes(xintercept = median(visits_per_day$n_visits)), color = "red", linetype = "

    labs(title = "Distribution of Number of Visits per HCW per Day",
             x = "Number of Visits",
             y = "Count (HCW-Day pairs)",
             subtitle = "Blue dashed = mean, Red dotted = median") +
    theme_minimal()
```




```
# Text output for min, max, IQR
min_visits <- min(visits_per_day$n_visits)
max_visits <- max(visits_per_day$n_visits)
iqr_visits <- IQR(visits_per_day$n_visits)
cat("Min visits per HCW per day:", min_visits, "\n")
```

```
Min visits per HCW per day: 1
```

```r
cat("Max visits per HCW per day:", max_visits, "\n")
```

```
Max visits per HCW per day: 78
```

```r
cat("Interquartile range (IQR):", iqr_visits, "\n")
```

```
Interquartile range (IQR): 33
```

```r
nvisits <- nrow(df2)

# Filter for nurse visits
df_nurse_visits <- df2[df2$hcwType == 'NURSE', ]
df_nurses <- distinct(df_nurse_visits, hcwType, hcwId)
nurse_count <- nrow(df_nurses)
# Filter for other HCW types
df_doctor_visits <- df2[df2$hcwType == 'DOCTOR', ]
df_doctors <- distinct(df_doctor_visits, hcwType, hcwId)
doctor_count <- nrow(df_doctors)
df_pt_visits     <- df2[df2$hcwType == 'PT', ]
df_pts <- distinct(df_pt_visits, hcwType, hcwId)
pt_count <- nrow(df_pts)

df_ot_visits     <- df2[df2$hcwType == 'OT', ]
df_ots <- distinct(df_ot_visits, hcwType, hcwId)
ot_count <- nrow(df_ots)

df_rt_visits     <- df2[df2$hcwType == 'RT', ]
df_rts <- distinct(df_rt_visits, hcwType, hcwId)
rt_count <- nrow(df_rts)
```

Total patient visits by hcw type

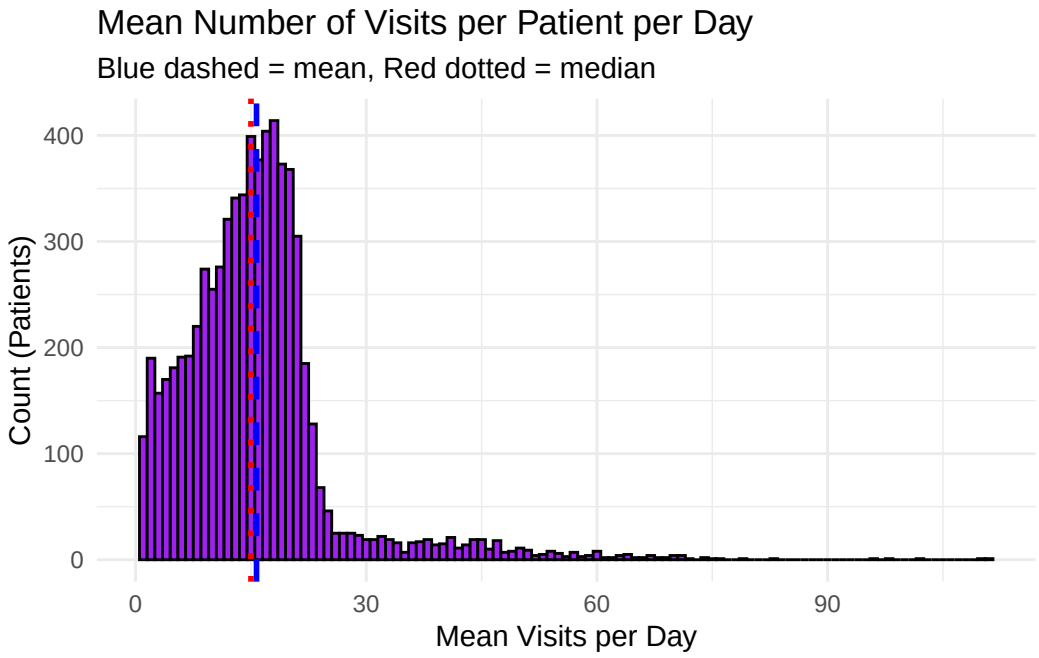
HCW Type	Total visits (365d)	mean/day
NURSE (40)	445962	30.5453425
DOCTOR (26)	244240	25.7365648
OT (12)	15174	3.4643836
PT (12)	11569	2.6413242
RT (12)	7384	1.6858447

Histogram: Mean Number of Visits per Patient per Day

```
# Calculate mean number of visits per patientId per day
visits_per_patient_day <- df2_filtered %>% group_by(patientId, visitDay) %>% summarise(n_visi
```

```
`summarise()` has grouped output by 'patientId'. You can override using the
`.groups` argument.
```

```
mean_visits <- visits_per_patient_day %>% group_by(patientId) %>% summarise(mean_visits = mea
ggplot(mean_visits, aes(x = mean_visits)) +
    geom_histogram(binwidth = 1, fill = "purple", color = "black") +
    geom_vline(xintercept = mean(mean_visits$mean_visits), color = "blue", linetype = "dashed
    geom_vline(xintercept = median(mean_visits$mean_visits), color = "red", linetype = "dotte
  labs(title = "Mean Number of Visits per Patient per Day",
      x = "Mean Visits per Day",
      y = "Count (Patients)",
      subtitle = "Blue dashed = mean, Red dotted = median") +
  theme_minimal()
```




```
min_mean_visits <- min(mean_visits$mean_visits)
max_mean_visits <- max(mean_visits$mean_visits)
```

```r
mean_mean_visits <- mean(mean_visits$mean_visits)
median_mean_visits <- median(mean_visits$mean_visits)
cat("Min  visits per patient per day:", min_mean_visits, "\n")
```

```
Min  visits per patient per day: 1
```

```r
cat("Max  visits per patient per day:", max_mean_visits, "\n")
```

```
Max  visits per patient per day: 111
```

```r
cat("Mean visits per patient per day:", mean_mean_visits, "\n")
```

```
Mean visits per patient per day: 15.71807
```

```r
cat("Median visits per patient per day:", median_mean_visits, "\n")
```

```
Median visits per patient per day: 15
```

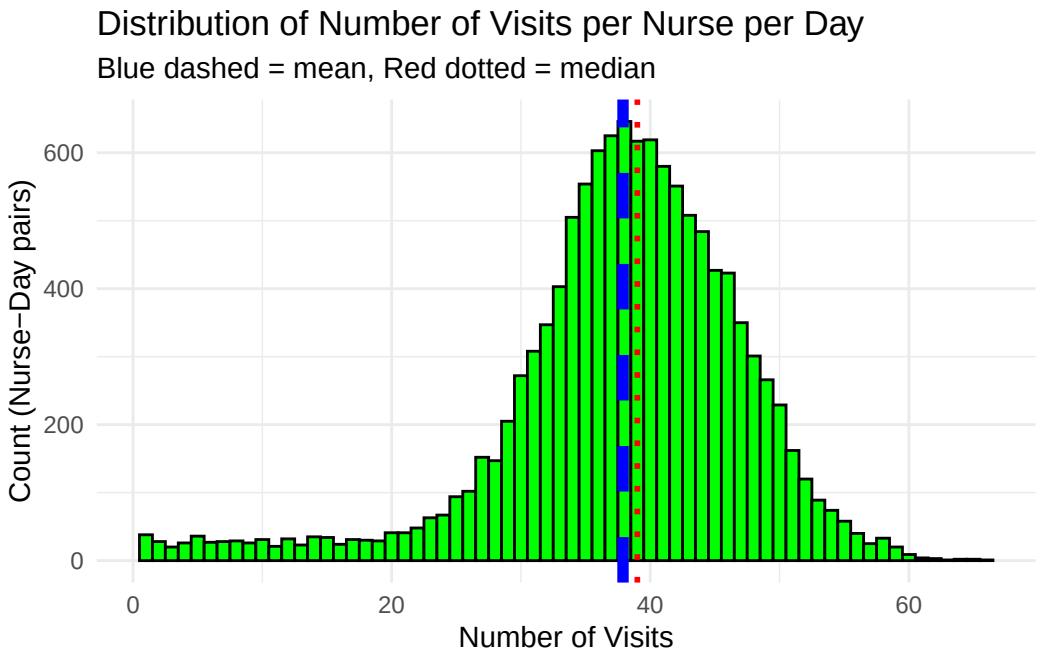
Visits per HCW per Day by Type

Distribution of Number of Visits per Nurse per Day

```r
visits_per_nurse_day <- df_nurse_visits %>% group_by(hcwId, visitDay) %>% summarise(n_visits
```

```
`summarise()` has grouped output by 'hcwId'. You can override using the
`.groups` argument.
```

```r
ggplot(visits_per_nurse_day, aes(x = n_visits)) +
  geom_histogram(binwidth = 1, fill = "green", color = "black") +
  geom_vline(aes(xintercept = mean(visits_per_nurse_day$n_visits)), color = "blue", linetype
  geom_vline(aes(xintercept = median(visits_per_nurse_day$n_visits)), color = "red", linetyp
  labs(title = "Distribution of Number of Visits per Nurse per Day",
       x = "Number of Visits",
       y = "Count (Nurse-Day pairs)",
       subtitle = "Blue dashed = mean, Red dotted = median") +
  theme_minimal()
```




```
min_nurse_visits <- min(visits_per_nurse_day$n_visits)
max_nurse_visits <- max(visits_per_nurse_day$n_visits)
mean_nurse_visits <- mean(visits_per_nurse_day$n_visits)
median_nurse_visits <- median(visits_per_nurse_day$n_visits)
```

Min visits per nurse per day: 1
Max visits per nurse per day: 66
Mean visits per nurse per day: 37.8929391
Median visits per nurse per day: 39

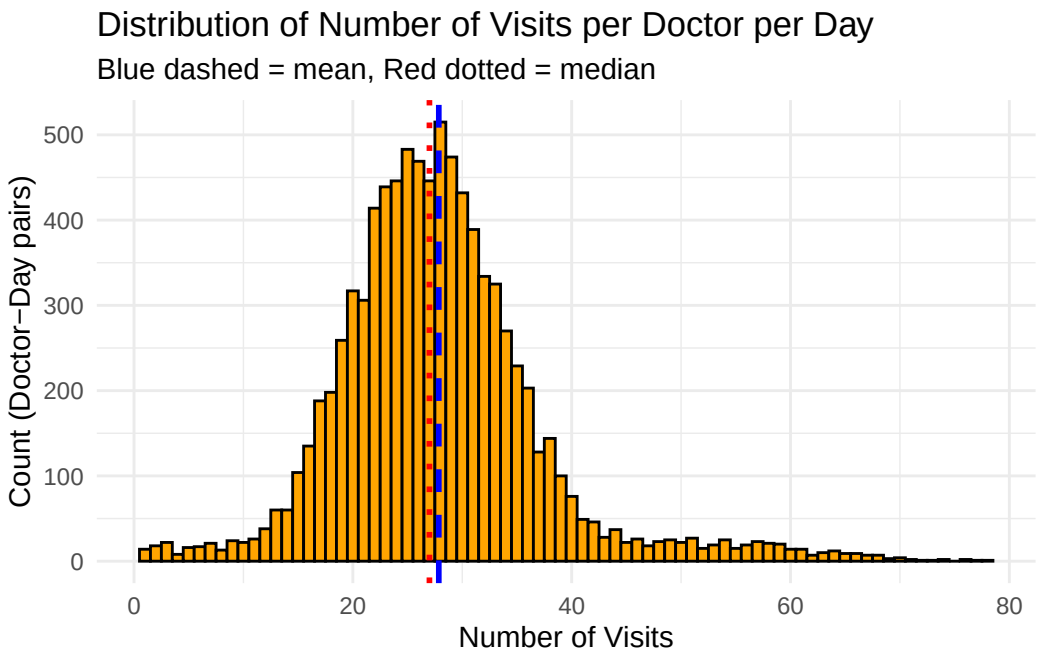
Distribution of Number of Visits per Doctor per Day

```
visits_per_doctor_day <- df_doctor_visits %>% group_by(hcwId, visitDay) %>% summarise(n_visi
```

```
`summarise()` has grouped output by 'hcwId'. You can override using the
`.groups` argument.
```

```
ggplot(visits_per_doctor_day, aes(x = n_visits)) +
    geom_histogram(binwidth = 1, fill = "orange", color = "black") +
    geom_vline(aes(xintercept = mean(visits_per_doctor_day$n_visits)), color = "blue", linety
```

```
    geom_vline(aes(xintercept = median(visits_per_doctor_day$n_visits)), color = "red", linet
    labs(title = "Distribution of Number of Visits per Doctor per Day",
             x = "Number of Visits",
             y = "Count (Doctor-Day pairs)",
             subtitle = "Blue dashed = mean, Red dotted = median") +
    theme_minimal()
```




```
min_doctor_visits <- min(visits_per_doctor_day$n_visits)
max_doctor_visits <- max(visits_per_doctor_day$n_visits)
mean_doctor_visits <- mean(visits_per_doctor_day$n_visits)
median_doctor_visits <- median(visits_per_doctor_day$n_visits)
```

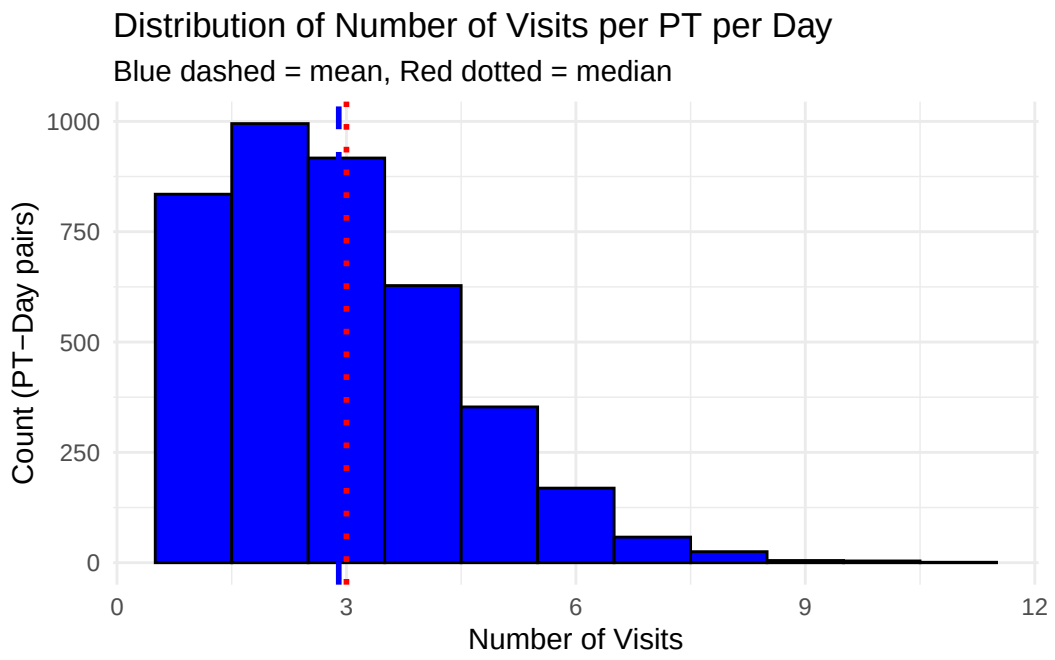
Min visits per doctor per day: 1
Max visits per doctor per day: 78
Mean visits per doctor per day: 27.8558394
Median visits per doctor per day: 27

Distribution of Number of Visits per PT per Day

```
visits_per_pt_day <- df_pt_visits %>% group_by(hcwId, visitDay) %>% summarise(n_visits = n()
```

```
`summarise()` has grouped output by 'hcwId'. You can override using the
`.groups` argument.
```

```
ggplot(visits_per_pt_day, aes(x = n_visits)) +
    geom_histogram(binwidth = 1, fill = "blue", color = "black") +
    geom_vline(aes(xintercept = mean(visits_per_pt_day$n_visits)), color = "blue", linetype =
    geom_vline(aes(xintercept = median(visits_per_pt_day$n_visits)), color = "red", linetype
    labs(title = "Distribution of Number of Visits per PT per Day",
             x = "Number of Visits",
             y = "Count (PT-Day pairs)",
             subtitle = "Blue dashed = mean, Red dotted = median") +
    theme_minimal()
```




```
min_pt_visits <- min(visits_per_pt_day$n_visits)
max_pt_visits <- max(visits_per_pt_day$n_visits)
mean_pt_visits <- mean(visits_per_pt_day$n_visits)
median_pt_visits <- median(visits_per_pt_day$n_visits)
```

Min visits per PT per day: 1
Max visits per PT per day: 11
Mean visits per PT per day: 2.8994987
Median visits per PT per day: 3

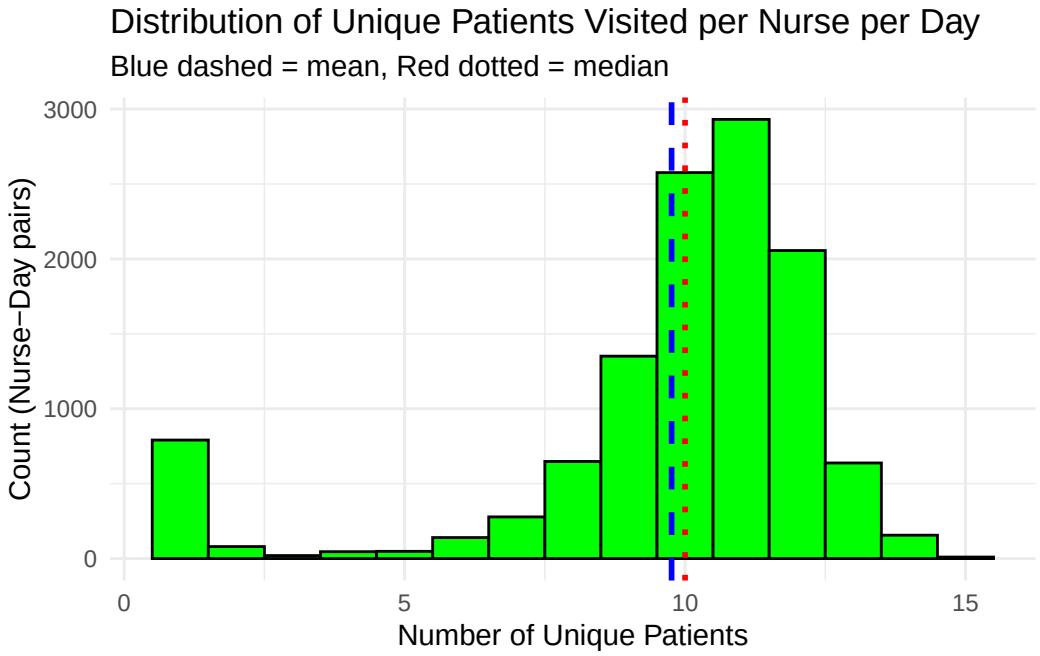
Distribution of Number of Unique Patients Visited per Day by HCW Type

Nurses

```
unique_patients_per_nurse_day <- df_nurse_visits %>% group_by(hcwId, visitDay) %>% summarise
```

```
`summarise()` has grouped output by 'hcwId'. You can override using the
`.groups` argument.
```

```
ggplot(unique_patients_per_nurse_day, aes(x = n_unique_patients)) +
    geom_histogram(binwidth = 1, fill = "green", color = "black") +
    geom_vline(aes(xintercept = mean(n_unique_patients)), color = "blue", linetype = "dashed"
    geom_vline(aes(xintercept = median(n_unique_patients)), color = "red", linetype = "dotte
    labs(title = "Distribution of Unique Patients Visited per Nurse per Day",
             x = "Number of Unique Patients",
             y = "Count (Nurse-Day pairs)",
             subtitle = "Blue dashed = mean, Red dotted = median") +
    theme_minimal()
```



```
min_nurse_patients <- min(unique_patients_per_nurse_day$n_unique_patients)
max_nurse_patients <- max(unique_patients_per_nurse_day$n_unique_patients)
mean_nurse_patients <- mean(unique_patients_per_nurse_day$n_unique_patients)
median_nurse_patients <- median(unique_patients_per_nurse_day$n_unique_patients)
```

Min unique patients per Nurse per day: 1
Max unique patients per Nurse per day: 15
Mean unique patients per Nurse per day: 9.7599626
Median unique patients per Nurse per day: 10

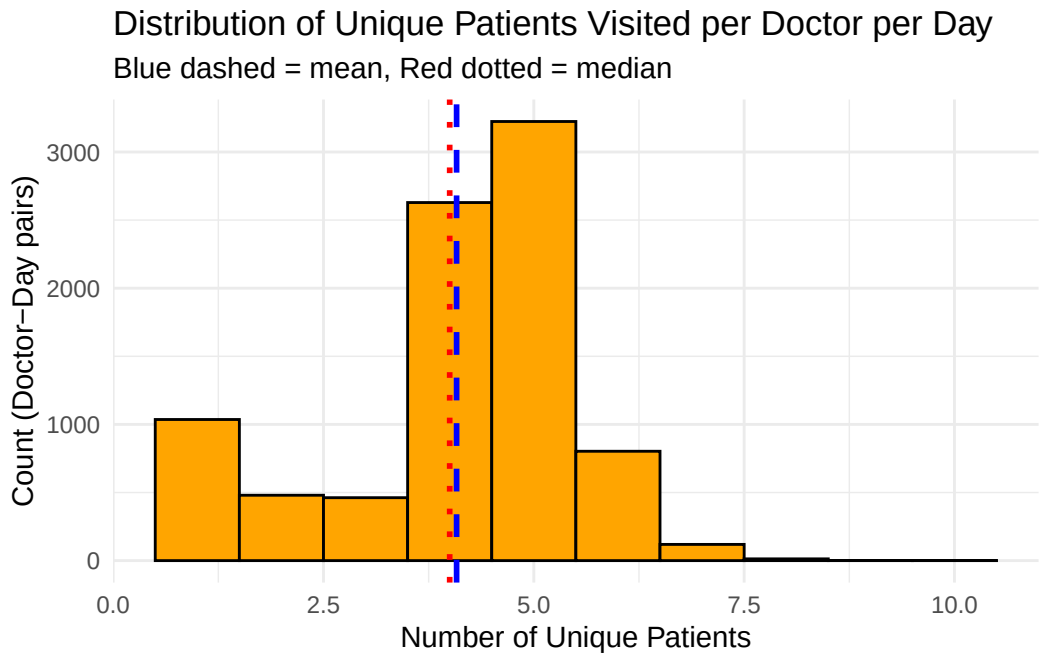
Unique Patients visited per day

Doctors

```
unique_patients_per_doctor_day <- df_doctor_visits %>% group_by(hcwId, visitDay) %>% summaris
```

```
`summarise()` has grouped output by 'hcwId'. You can override using the
`.groups` argument.
```

```
ggplot(unique_patients_per_doctor_day, aes(x = n_unique_patients)) +
    geom_histogram(binwidth = 1, fill = "orange", color = "black") +
    geom_vline(aes(xintercept = mean(n_unique_patients)), color = "blue", linetype = "dashed"
    geom_vline(aes(xintercept = median(n_unique_patients)), color = "red", linetype = "dotte
    labs(title = "Distribution of Unique Patients Visited per Doctor per Day",
             x = "Number of Unique Patients",
             y = "Count (Doctor-Day pairs)",
             subtitle = "Blue dashed = mean, Red dotted = median") +
    theme_minimal()
```




```
min_doctor_patients <- min(unique_patients_per_doctor_day$n_unique_patients)
max_doctor_patients <- max(unique_patients_per_doctor_day$n_unique_patients)
mean_doctor_patients <- mean(unique_patients_per_doctor_day$n_unique_patients)
median_doctor_patients <- median(unique_patients_per_doctor_day$n_unique_patients)
```

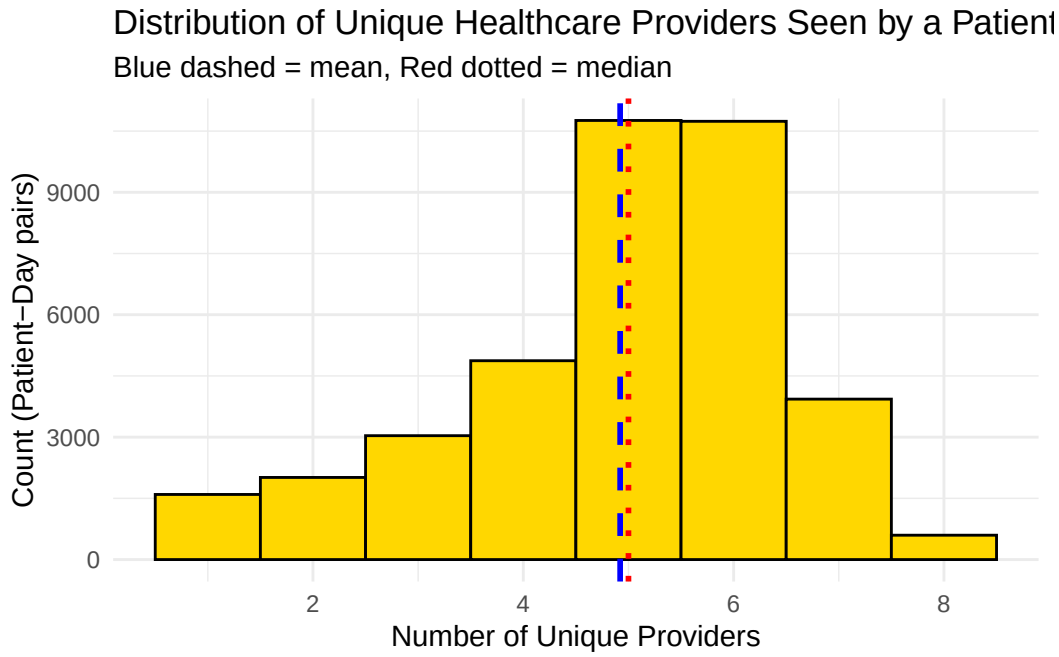
Min unique patients per Doctor per day: 1
Max unique patients per Doctor per day: 10
Mean unique patients per Doctor per day: 4.0821168
Median unique patients per Doctor per day: 4

Distribution of Number of Unique Healthcare Providers Seen by a Patient per Day

```
unique_providers_per_patient_day <- df2_filtered %>% group_by(patientId, visitDay) %>% summa
```

```
`summarise()` has grouped output by 'patientId'. You can override using the
`.groups` argument.
```

```r
ggplot(unique_providers_per_patient_day, aes(x = n_unique_providers)) +
    geom_histogram(binwidth = 1, fill = "gold", color = "black") +
    geom_vline(aes(xintercept = mean(n_unique_providers)), color = "blue", linetype = "dashed
    geom_vline(aes(xintercept = median(n_unique_providers)), color = "red", linetype = "dotte
    labs(title = "Distribution of Unique Healthcare Providers Seen by a Patient per Day",
             x = "Number of Unique Providers",
             y = "Count (Patient-Day pairs)",
             subtitle = "Blue dashed = mean, Red dotted = median") +
    theme_minimal()
```




```r
min_patient_providers <- min(unique_providers_per_patient_day$n_unique_providers)
max_patient_providers <- max(unique_providers_per_patient_day$n_unique_providers)
mean_patient_providers <- mean(unique_providers_per_patient_day$n_unique_providers)
median_patient_providers <- median(unique_providers_per_patient_day$n_unique_providers)
```

Min unique providers per patient per day: 1
Max unique providers per patient per day: 8
Mean unique providers per patient per day: 4.9208213
Median unique providers per patient per day: 5