Intentional Walks

Ethan Kruger

```
after_walk <- read_excel("/Users/owner/Desktop/Intentional Walks/ibb2024(2).xlsx")
no_walk <- read_excel("/Users/owner/Desktop/Intentional Walks/non-ibb2024(2).xlsx")
```

Data Processing

```
after_walk = after_walk %>%
  mutate(total_bases = H + H2 + (2 * H3) + (3 * HR), OBP = (H + BB + HBP)/(AB + BB + HBP + SF)) %>%
  mutate(ID = row_number())

no_walk = no_walk %>%
  mutate(total_bases = H + H2 + (2 * H3) + (3 * HR), OBP = (H + BB + HBP)/(AB + BB + HBP + SF)) %>%
  mutate(ID = row_number())
```

Slugging Percentage Testing

```
slug_after_walk <- after_walk %>%
  filter(AB != 0)

slug_no_walk <- no_walk %>%
  filter(AB != 0)

print(paste("Slugging percentage for at-bats following an intentional walk:", format(round(mean(slug_af
```

```
## [1] "Slugging percentage for at-bats following an intentional walk: 0.414"
```

```
print(paste("Slugging percentage for all other at-bats:", format(round(mean(slug_no_walk$total_bases), 3
```

```
## [1] "Slugging percentage for all other at-bats: 0.400"
```

```
slug_t_test_result <- t.test(slug_after_walk$total_bases, slug_no_walk$total_bases, alternative = "grea
slug_t_test_result
```

```
##
##  Welch Two Sample t-test
##
## data:  slug_after_walk$total_bases and slug_no_walk$total_bases
## t = 0.35592, df = 432.25, p-value = 0.361
```

```
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.05207353         Inf
## sample estimates:
## mean of x mean of y
## 0.4139535 0.3996137
```

OBP testing

```
OBP_after_walk <- after_walk %>%
  filter(!is.nan(OBP))

OBP_no_walk <- no_walk %>%
  filter(!is.nan(OBP))

print(paste("On-base percentage for at-bats following an intentional walk:", format(round(mean(OBP_afte
```

```
## [1] "On-base percentage for at-bats following an intentional walk: 0.345"
```

```
print(paste("On-base percentage for all other at-bats:", format(round(mean(OBP_no_walk$OBP), 3), nsmall
```

```
## [1] "On-base percentage for all other at-bats: 0.314"
```

```
OBP_t_test_result <- t.test(OBP_after_walk$OBP, OBP_no_walk$OBP, alternative = "greater")
OBP_t_test_result
```

```
##
##  Welch Two Sample t-test
##
## data:  OBP_after_walk$OBP and OBP_no_walk$OBP
## t = 1.4246, df = 507.51, p-value = 0.07744
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.004733975          Inf
## sample estimates:
## mean of x mean of y
## 0.3445545 0.3143457
```

OPS Testing

```
calculate_group_metrics <- function(data) {
  # Total metrics
  total_H <- sum(data$H, na.rm = TRUE)
  total_BB <- sum(data$BB, na.rm = TRUE)
  total_HBP <- sum(data$HBP, na.rm = TRUE)
  total_AB <- sum(data$AB, na.rm = TRUE)
  total_SF <- sum(data$SF, na.rm = TRUE)
  total_H2 <- sum(data$H2, na.rm = TRUE)
```

```r
  total_H3 <- sum(data$H3, na.rm = TRUE)
  total_HR <- sum(data$HR, na.rm = TRUE)

  # Total bases calculation
  total_bases <- (total_H - total_H2 - total_H3 - total_HR) + (2 * total_H2) + (3 * total_H3) + (4 * to

  # Calculate OBP
  group_OBP <- (total_H + total_BB + total_HBP) / (total_AB + total_BB + total_HBP + total_SF)

  # Calculate SLG
  group_SLG <- total_bases / total_AB

  # Calculate group-level OPS
  group_OPS <- group_OBP + group_SLG

  return(list(OBP = group_OBP, SLG = group_SLG, OPS = group_OPS))
}

# Apply the function to both groups
group1_metrics <- calculate_group_metrics(after_walk)
group2_metrics <- calculate_group_metrics(no_walk)

print(paste("OPS for at-bats following an intentional walk:", format(round(mean(group1_metrics$OPS), 3)
```

```
## [1] "OPS for at-bats following an intentional walk: 0.759"
```

```r
print(paste("OPS for all other at-bats:", format(round(mean(group2_metrics$OPS), 3), nsmall = 3)))
```

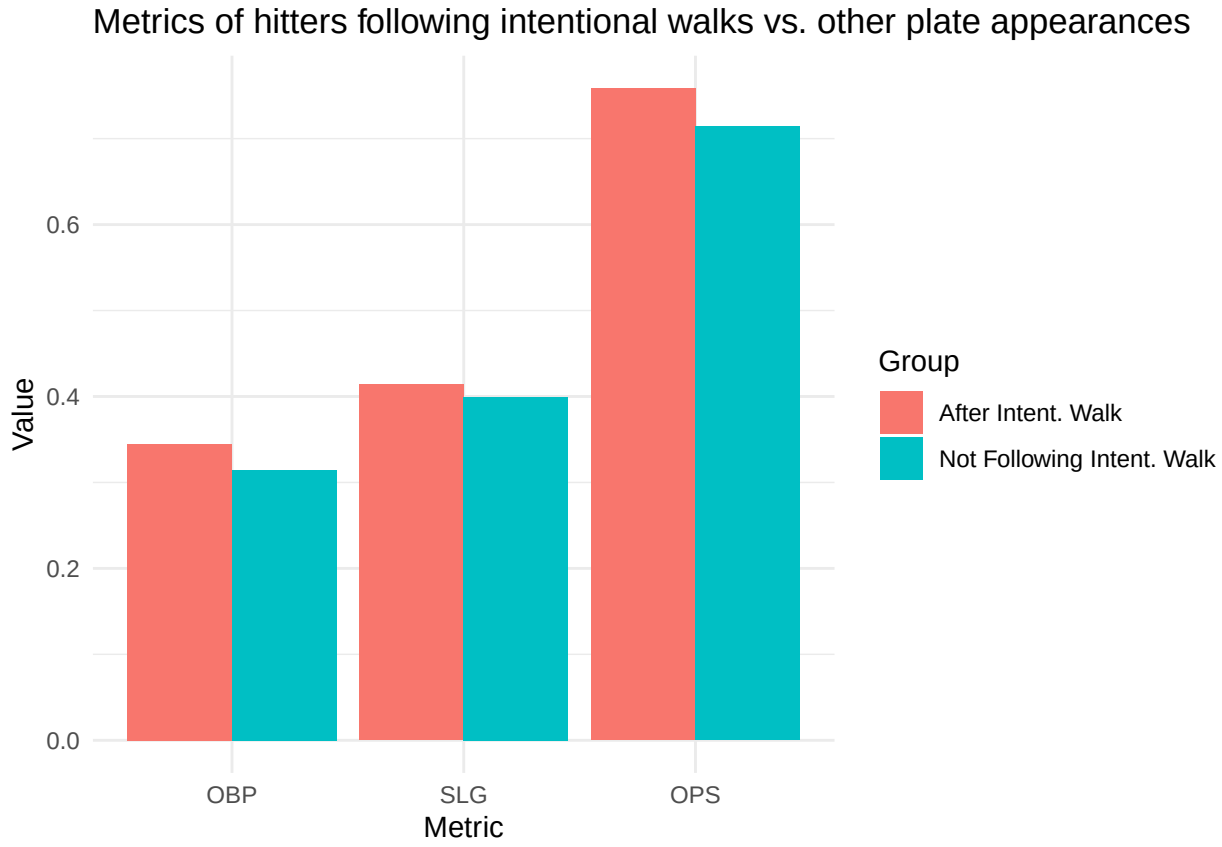
```
## [1] "OPS for all other at-bats: 0.714"
```

```r
library(ggplot2)

# Combine group metrics into a data frame for visualization
group_metrics <- data.frame(
  Metric = rep(c("OBP", "SLG", "OPS"), 2),
  Value = c(group1_metrics$OBP, group1_metrics$SLG, group1_metrics$OPS,
            group2_metrics$OBP, group2_metrics$SLG, group2_metrics$OPS),
  Group = rep(c("After Intent. Walk", "Not Following Intent. Walk"), each = 3)
)

group_metrics$Metric <- factor(group_metrics$Metric, levels = c("OBP", "SLG", "OPS"))

# Create a bar plot
ggplot(group_metrics, aes(x = Metric, y = Value, fill = Group)) +
  geom_bar(stat = "identity", position = position_dodge()) +
  labs(
    title = "Metrics of hitters following intentional walks vs. other plate appearances",
    x = "Metric",
    y = "Value"
  ) +
  theme_minimal()
```



Runs Created

```
RC_after_walk <- after_walk %>%
  mutate(RC = (total_bases * (H + BB))/ (AB + BB)) %>%
  na.omit()

RC_no_walk <- no_walk %>%
  mutate(RC = (total_bases * (H + BB))/ (AB + BB)) %>%
  na.omit()

print(paste("Runs Created for at-bats following an intentional walk:", format(round(mean(RC_after_walk$
```

```
## [1] "Runs Created for at-bats following an intentional walk: 0.378"
```

```
print(paste("Runs Created for all other at-bats:", format(round(mean(RC_no_walk$RC), 3), nsmall = 3)))
```

```
## [1] "Runs Created for all other at-bats: 0.365"
```

```
RC_t_test_result <- t.test(RC_after_walk$RC, RC_no_walk$RC, alternative = "greater")
RC_t_test_result
```

```
##
```

```
##  Welch Two Sample t-test
##
## data:  RC_after_walk$RC and RC_no_walk$RC
## t = 0.343, df = 473.56, p-value = 0.3659
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.0485125        Inf
## sample estimates:
## mean of x mean of y
## 0.3779193 0.3651692
```

```r
# Create a bar plot
RC_dataframe <- data.frame(
  Metric = c("After Intentional Walk", "Other At-bats"),
  Value = c(mean(RC_after_walk$RC), mean(RC_no_walk$RC)))

ggplot(RC_dataframe, aes(x = Metric, y = Value, fill = Metric)) +
  geom_bar(stat = "identity", position = position_dodge()) +
  labs(
    title = "Metrics of hitters following intentional walks vs. other plate appearances",
    x = "Metric",
    y = "Value"
  ) +
  scale_y_continuous(limits = c(0, max(RC_dataframe$Value) + 0.1)) +
  scale_fill_hue() +
  theme_minimal()
```

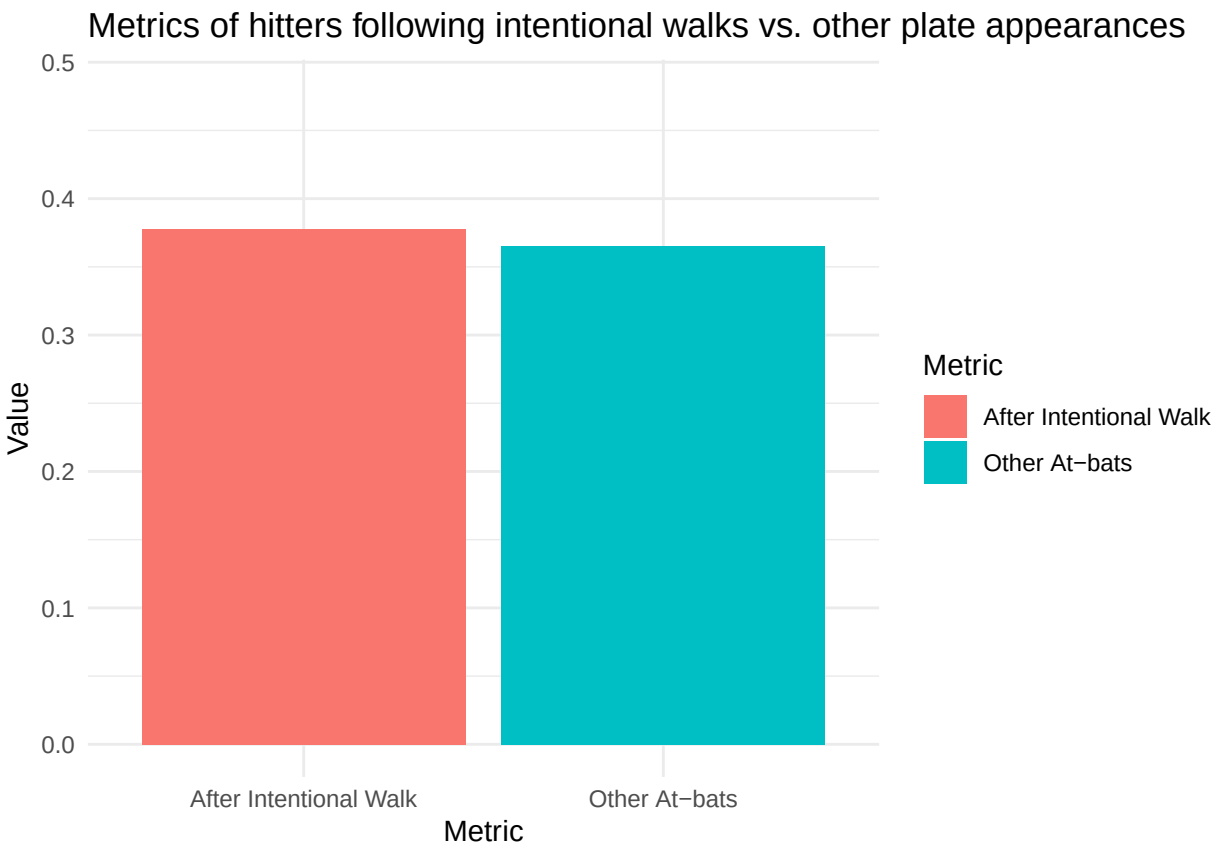


Plate Appearances per Home Run

```
HRPA_after_walk <- 1/mean(after_walk$HR)
HRPA_no_walk <- 1/mean(no_walk$HR)

print(paste("At-bats per home run following an intentional walk:", format(round(1/mean(after_walk$HR), 1
```

```
## [1] "At-bats per home run following an intentional walk: 42.25"
```

```
print(paste("At-bats per home run for all other at-bats:", format(round(1/mean(no_walk$HR), 2), nsmall =
```

```
## [1] "At-bats per home run for all other at-bats: 33.83"
```

```
HR_t_test_result <- t.test(after_walk$HR, no_walk$HR)
HR_t_test_result
```

```
##
##  Welch Two Sample t-test
##
## data:  after_walk$HR and no_walk$HR
## t = -0.86912, df = 510.59, p-value = 0.3852
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.019193382  0.007419988
## sample estimates:
##  mean of x  mean of y
## 0.02366864 0.02955534
```

Plate Appearances per Strikeout

```
print(paste("At-bats per strikeout following an intentional walk:", format(round(1/mean(after_walk$K), 3
```

```
## [1] "At-bats per strikeout following an intentional walk: 4.922"
```

```
print(paste("At-bats per strikeout for all other at-bats:", format(round(1/mean(no_walk$K), 3), nsmall =
```

```
## [1] "At-bats per strikeout for all other at-bats: 4.420"
```

```
K_t_test_result <- t.test(after_walk$K, no_walk$K, alternative = "less")
K_t_test_result
```

```
##
##  Welch Two Sample t-test
##
## data:  after_walk$K and no_walk$K
## t = -1.2897, df = 510, p-value = 0.09888
## alternative hypothesis: true difference in means is less than 0
```

```
## 95 percent confidence interval:
##         -Inf 0.006419325
## sample estimates:
## mean of x mean of y
## 0.2031558 0.2262688
```