exercise 1

2024-01-15

```r
# Load the necessary library for data manipulation
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
# Read the CSV file
df <- read.csv("performance_data.csv")
```

```r
#Performance change calculate
df$worker1_diff <- c(NA, diff(df$worker1))
df$worker2_diff <- c(NA, diff(df$worker2))
df$worker3_diff <- c(NA, diff(df$worker3))
```

```r
# Step 1: Filter rows where 'w2_intervention' is 'A'
filtered_df_w2 <- df[df$w2_intervention == 'A', ]

# Step 2: Calculate the count of rows where 'worker2' difference is greater than 0
count_greater_than_0 <- sum(filtered_df_w2$worker2_diff > 0, na.rm = TRUE)

# Step 3: Divide the count from step 2 by the total number of rows where 'w2_intervention' is 'A'
total_count_A <- nrow(filtered_df_w2)
percentage_w2 <- count_greater_than_0 / total_count_A
```

```r
#Worker 3
filtered_df_w3 <- df[df$w3_intervention == 'B', ]

count_greater_than_0 <- sum(filtered_df_w3$worker2_diff > 0, na.rm = TRUE)

total_count_A <- nrow(filtered_df_w3)
percentage_w3 <- count_greater_than_0 / total_count_A
```

```r
filtered_df_w2 <- df[df$w2_intervention == 'A', ]

count_greater_than_0 <- sum(filtered_df_w2$worker1_diff > 0, na.rm = TRUE)

total_count_A <- nrow(filtered_df_w2)
percentage_w1_A_days <- count_greater_than_0 / total_count_A
```

```r
filtered_df_w3 <- df[df$w3_intervention == 'B', ]

count_greater_than_0 <- sum(filtered_df_w3$worker1_diff > 0, na.rm = TRUE)

total_count_A <- nrow(filtered_df_w3)
percentage_w1_B_days <- count_greater_than_0 / total_count_A
```

```r
cat("Number of days performance increased when intervention applied to worker 2 %:", as.character(percer
```

```
## Number of days performance increased when intervention applied to worker 2 %: 0.923076923076923
```

```r
cat("Number of days performance increased when intervention applied to worker 3 %:", as.character(percer
```

```
## Number of days performance increased when intervention applied to worker 3 %: 0.461538461538462
```

```r
cat("Number of days performance increased when intervention A applied to worker 1 %:", as.character(per
```

```
## Number of days performance increased when intervention A applied to worker 1 %: 0.538461538461538
```

```r
cat("Number of days performance increased when intervention B applied to worker 1 %:", as.character(per
```

```
## Number of days performance increased when intervention B applied to worker 1 %: 0.615384615384615
```