

Data211ProjectFinal

Lykyteng Lysongtseng

2025-11-12

Daily Water Intake Project Final

Hypothesis 1. H_o : amount of daily water intake equal to 3000 ml H_A : amount of daily water intake is not equal to 3000 ml Hypothesis 2. μ_o : mean amount of water intake of workout day is equal to mean amount of water intake on non-workout day. μ_A : mean amount of water intake of workout day is not equal to mean amount of water intake on non-workout day.

```
library(readxl)
Daily_water_intake<-read_xlsx("C:/Users/lykyt/OneDrive/Daily water intake.xlsx")
```

```
mean(Daily_water_intake$water_intake)
```

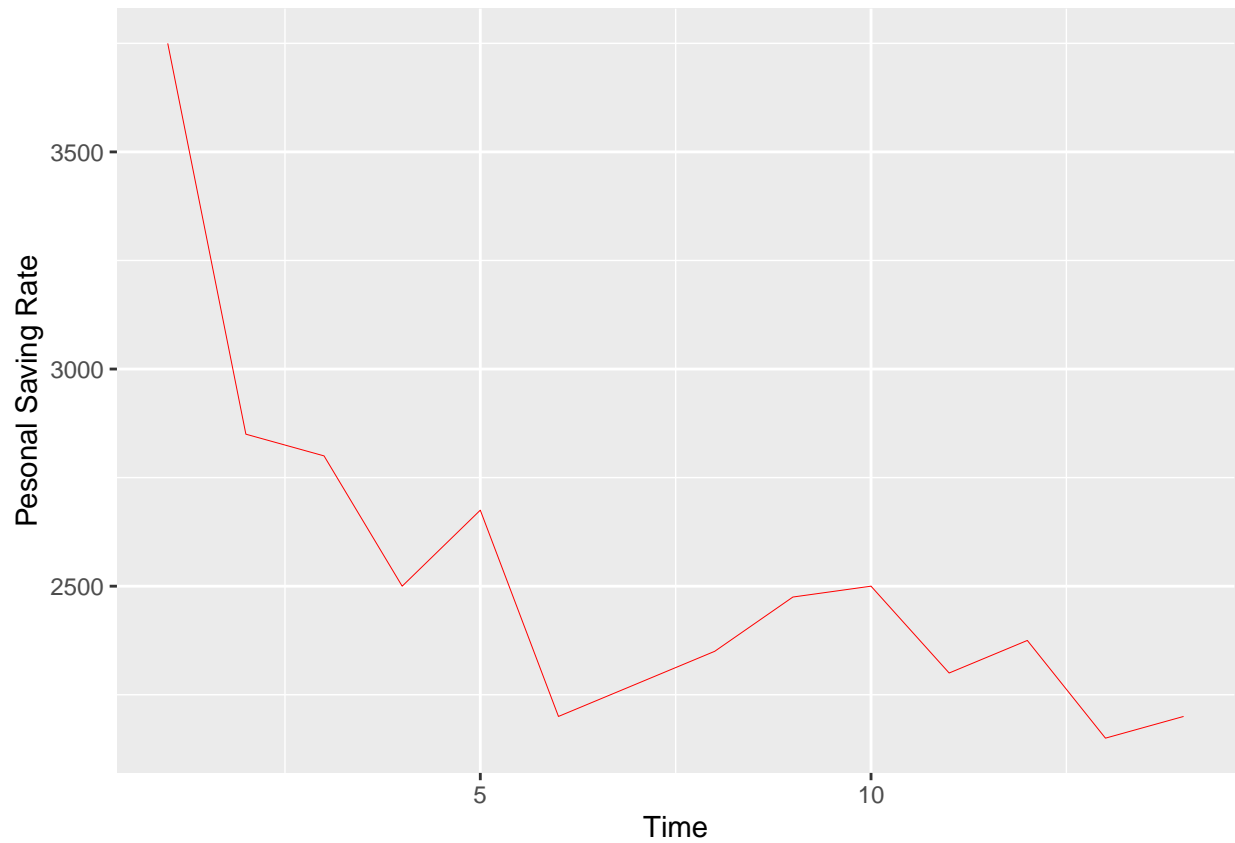
```
## [1] 2528.571
```

```
library(ggplot2)
```

```
ggplot(data = Daily_water_intake,aes(x=Day,y=water_intake))+
  geom_line(color="red", size=0.05, lineype="dotdash")+
  labs(x="Time",y="Pesonal Saving Rate")
```

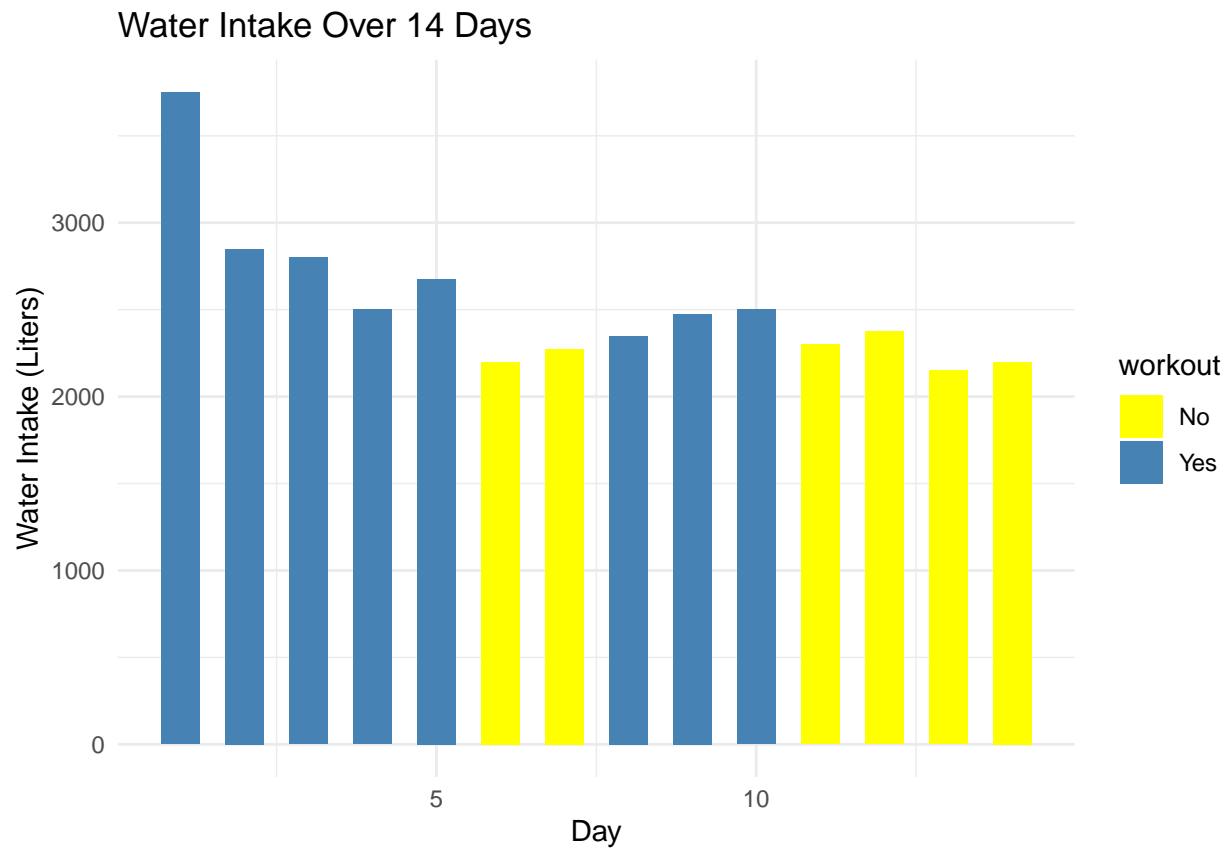
```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

```
## Warning in geom_line(color = "red", size = 0.05, lineype = "dotdash"): Ignoring
## unknown parameters: 'lineype'
```

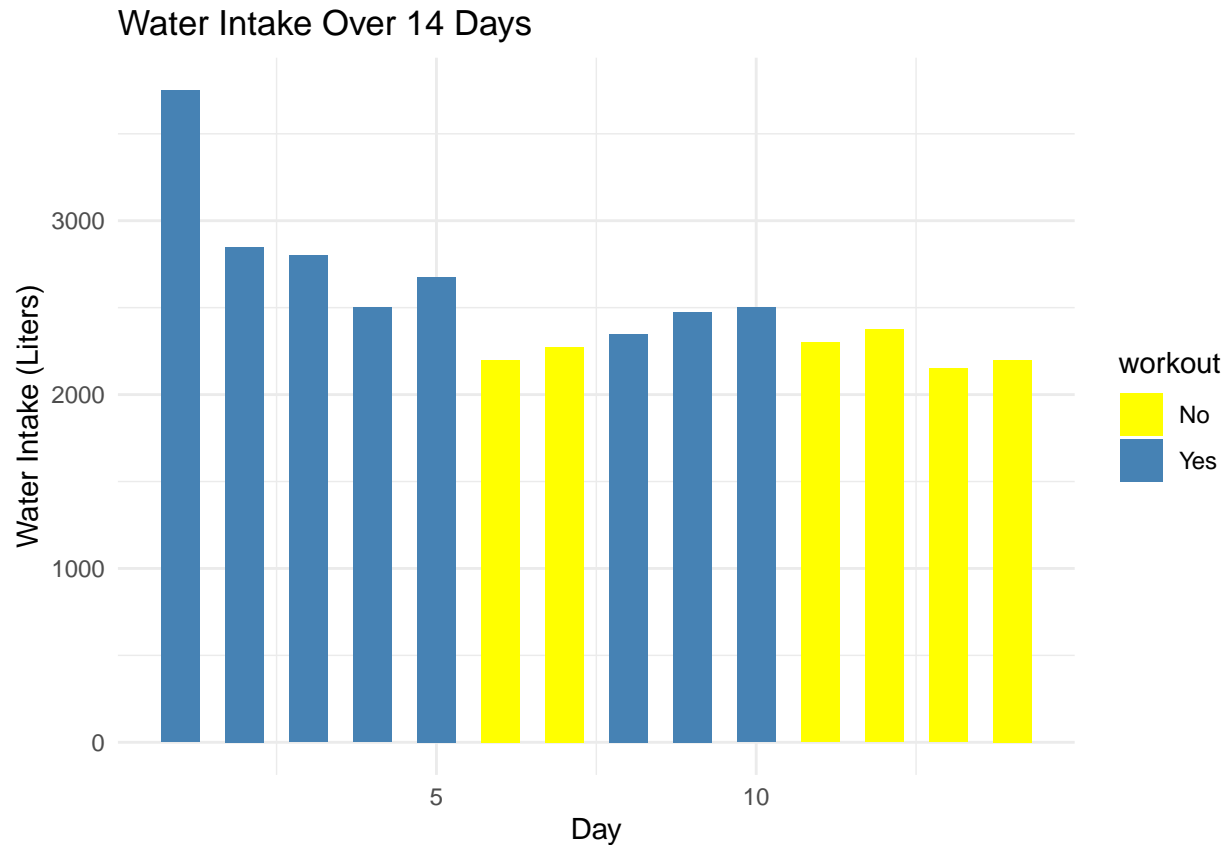


Histogram

```
ggplot(data = Daily_water_intake, aes(x = Day, y = water_intake, fill = workout)) +
  geom_col(width = 0.6) +
  labs(title = "Water Intake Over 14 Days",
        x = "Day",
        y = "Water Intake (Liters)") +
  scale_fill_manual(values = c("Yes" = "steelblue", "No" = "yellow")) +
  theme_minimal()
```



```
ggplot(data = Daily_water_intake, aes(x = Day, y = water_intake, fill = workout)) +  
  geom_col(width = 0.6) +  
  labs(title = "Water Intake Over 14 Days",  
        x = "Day",  
        y = "Water Intake (Liters)") +  
  scale_fill_manual(values = c("Yes" = "steelblue", "No" = "yellow")) +  
  theme_minimal()
```



Data Wrangling

```
workout_data <- subset(Daily_water_intake, workout == "Yes")
nonworkout_data <- subset(Daily_water_intake, workout == "No")
```

```
mean(workout_data$water_intake, na.rm = TRUE)
```

```
## [1] 2737.5
```

```
mean(nonworkout_data$water_intake, na.rm = TRUE)
```

```
## [1] 2250
```

```
library(dplyr)
```

```
##
```

```
## 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
```

```
Daily_water_intake %>%
  group_by(workout) %>%
  summarise(mean_intake = mean(water_intake, na.rm = TRUE))
```

```
## # A tibble: 2 x 2
##   workout mean_intake
##   <chr>      <dbl>
## 1 No        2250
## 2 Yes       2738.
```

```
table(Daily_water_intake$workout)
```

```
##
## No Yes
## 6 8
```

```
t.test(Daily_water_intake$water_intake, mu = 3, alternative = "two.sided")
```

```
##
## One Sample t-test
##
## data: Daily_water_intake$water_intake
## t = 22.836, df = 13, p-value = 7.083e-12
## alternative hypothesis: true mean is not equal to 3
## 95 percent confidence interval:
## 2289.646 2767.496
## sample estimates:
## mean of x
## 2528.571
```

$p - \text{value} \approx 0 < \alpha = 0.05$ So we reject H_0 . At 5% significance level we have sufficient evidence to conclude mean water intake is significantly lower than recommended 3000 mL.

```
Daily_water_intake$workout <- factor(Daily_water_intake$workout, levels = c("No", "Yes"))
t.test(water_intake ~ workout, data = Daily_water_intake, alternative = "two.sided")
```

```
##
## Welch Two Sample t-test
##
## data: water_intake by workout
## t = -3.0396, df = 7.6326, p-value = 0.01699
## alternative hypothesis: true difference in means between group No and group Yes is not equal to 0
## 95 percent confidence interval:
## -860.4695 -114.5305
## sample estimates:
## mean in group No mean in group Yes
## 2250.0 2737.5
```

$p\text{-value} \approx 0 < \alpha = 0.05$ So we reject H_0 . At 5% significance level we have sufficient evidence to conclude daily water intake is dependent on my workout days. The level of water intake on workout days are significantly different versus the water intake on non-workout days. I also found that the mean of water intake working out was closer to the daily water intake. More study would have to be done in terms of the length and types of workouts, height and weight, and other factors that could have contributed to water intake.