Sleep Analysis RMD

Zuhaib Sultan

2025-06-17

Table of Contents

Introduction

This report analyzes the relationship between sleep patterns and academic performance of students using a survey dataset. The analysis includes exploratory data analysis, descriptive inference, and hypothesis testing to answer the following questions:

1. Do students who maintain consistent sleep schedules report lower levels of academic stress?
2. Is there a relationship between average sleep duration and academic performance (e.g., GPA or self-reported productivity)?
3. Do students who prioritize sleep hygiene practices (e.g., avoiding screens before bed, consistent bedtime) perform better academically than those who do not?

Load Data

```
# Load Dataset
dataset <- read_csv("C:/Users/Saba Akram/Downloads/student_sleep_patterns.csv")  # Update with the c
orrect path
```

```
## Rows: 500 Columns: 14
## ── Column specification ────────────────────────────────────────────────────
## Delimiter: ","
## chr  (2): Gender, University_Year
## dbl (12): Student_ID, Age, Sleep_Duration, Study_Hours, Screen_Time, Caffein...
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
head(dataset)  # Preview the dataset
```

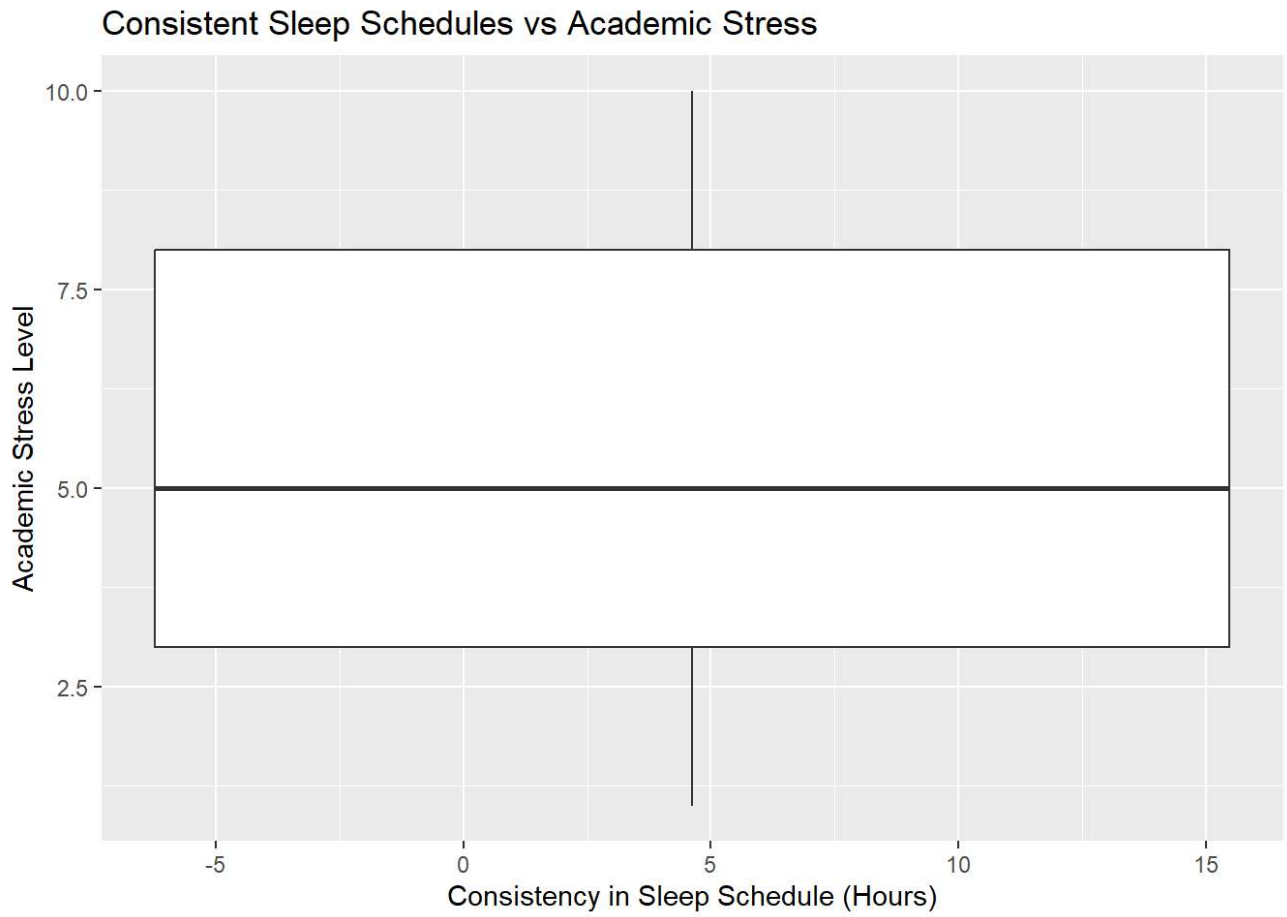
```
## # A tibble: 6 × 14
##   Student_ID   Age Gender University_Year Sleep_Duration Study_Hours Screen_Time
##        <dbl> <dbl> <chr>  <chr>                   <dbl>       <dbl>       <dbl>
## 1          1    24 Other  2nd Year                  7.7         7.9         3.4
## 2          2    21 Male   1st Year                  6.3         6           1.9
## 3          3    22 Male   4th Year                  5.1         6.7         3.9
## 4          4    24 Other  4th Year                  6.3         8.6         2.8
## 5          5    20 Male   4th Year                  4.7         2.7         2.7
## 6          6    25 Other  1st Year                  4.9        12           3.2
## # ℹ 7 more variables: Caffeine_Intake <dbl>, Physical_Activity <dbl>,
## #   Sleep_Quality <dbl>, Weekday_Sleep_Start <dbl>, Weekend_Sleep_Start <dbl>,
## #   Weekday_Sleep_End <dbl>, Weekend_Sleep_End <dbl>
```

Exploratory Data Analysis

Q1: Do students who maintain consistent sleep schedules report lower levels of academic stress?

```
# Visualization of Consistent Sleep Schedules vs Academic Stress
ggplot(dataset, aes(x = Weekday_Sleep_Start - Weekday_Sleep_End, y = Sleep_Quality)) +
  geom_boxplot() +
  labs(title = "Consistent Sleep Schedules vs Academic Stress",
       x = "Consistency in Sleep Schedule (Hours)",
       y = "Academic Stress Level")
```

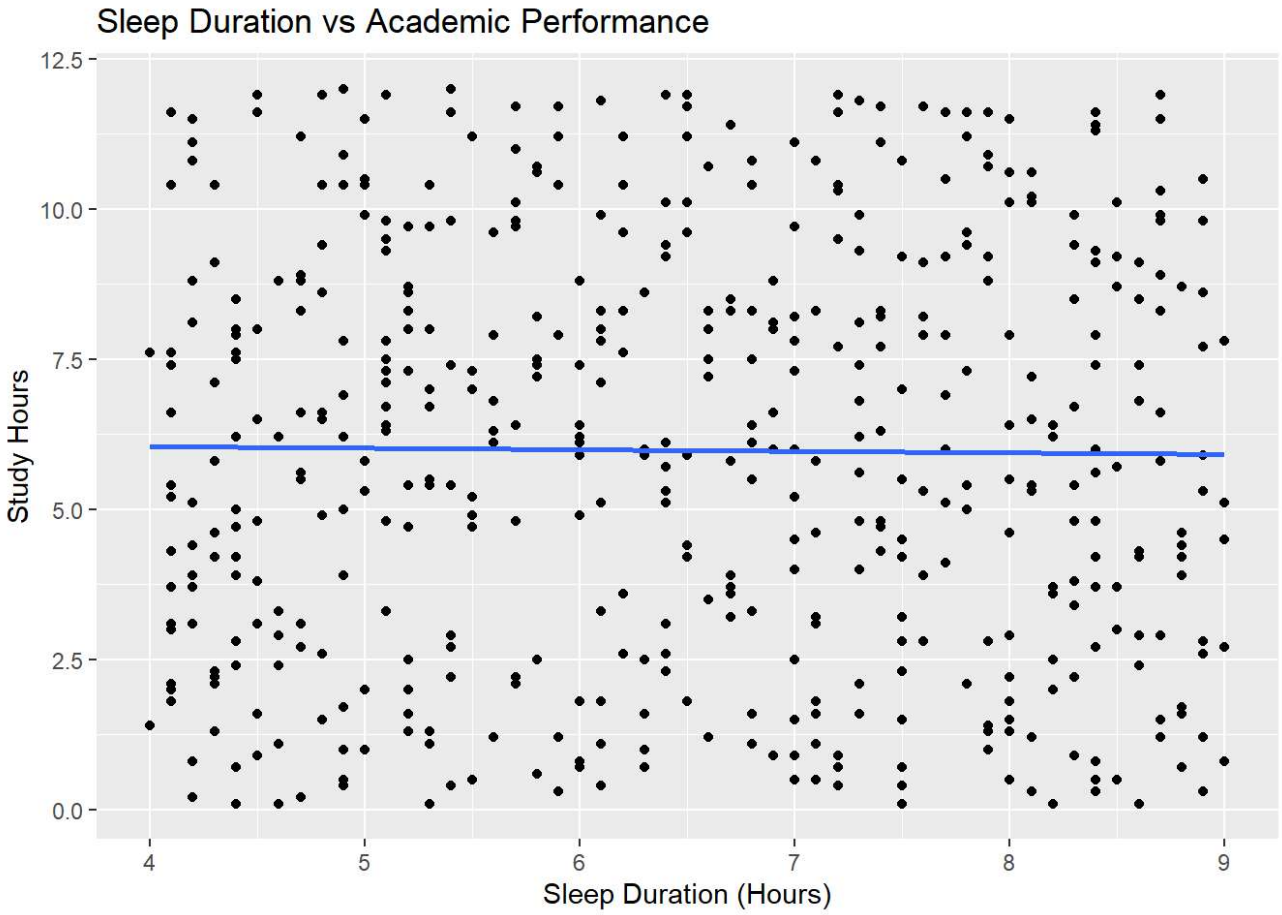
```
## Warning: Continuous x aesthetic
## ℹ did you forget `aes(group = ...)`?
```



Q2: Is there a relationship between average sleep duration and academic performance?

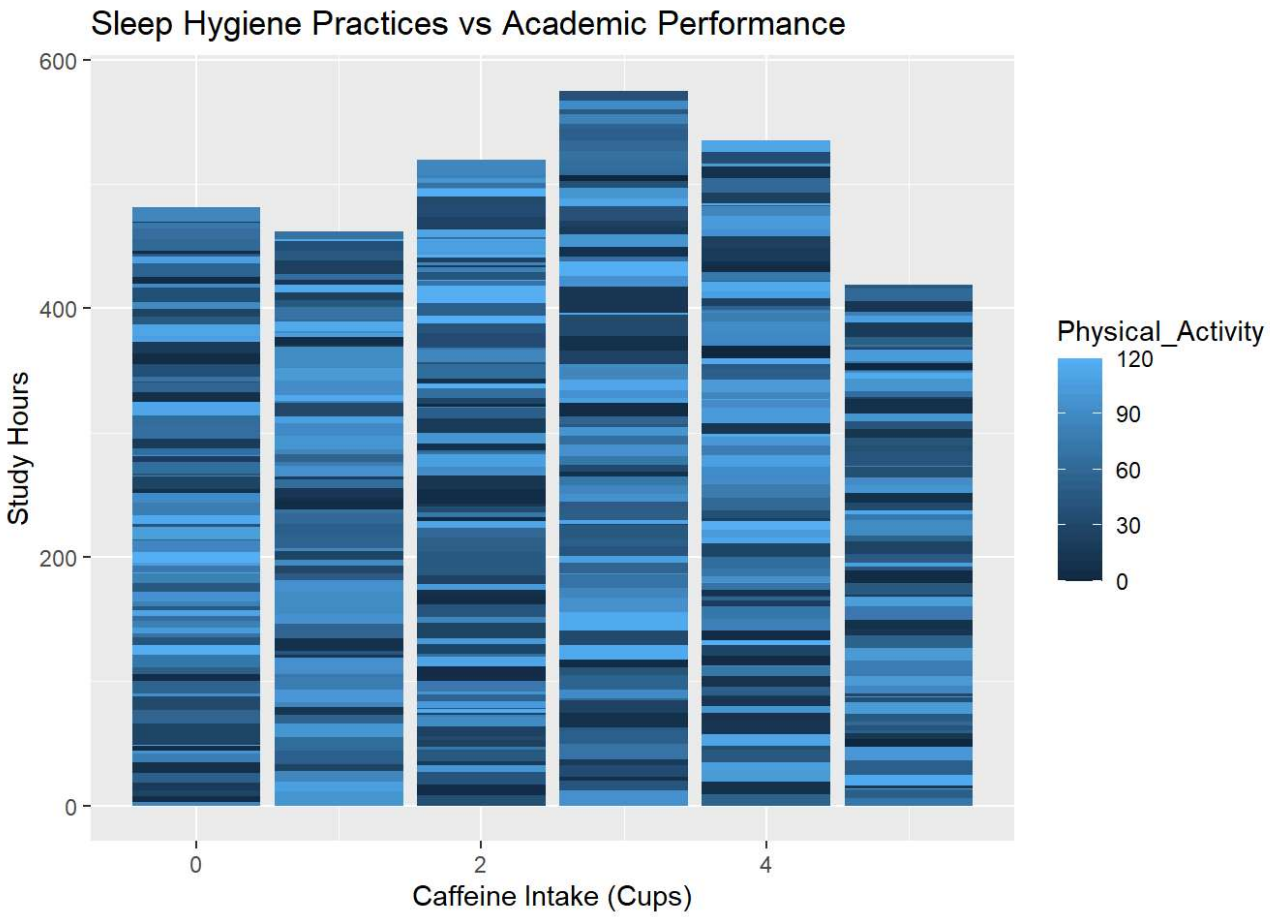
```
# Scatter plot for Sleep Duration vs Academic Performance
ggplot(dataset, aes(x = Sleep_Duration, y = Study_Hours)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  labs(title = "Sleep Duration vs Academic Performance",
       x = "Sleep Duration (Hours)",
       y = "Study Hours")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



Q3: Do students who prioritize sleep hygiene practices perform better academically than those who do not?

```
# Bar chart for Sleep Hygiene Practices vs Academic Performance
ggplot(dataset, aes(x = Caffeine_Intake, y = Study_Hours, fill = Physical_Activity)) +
  geom_bar(stat = "identity") +
  labs(title = "Sleep Hygiene Practices vs Academic Performance",
       x = "Caffeine Intake (Cups)",
       y = "Study Hours")
```



Descriptive Inference

```
# Summary Statistics
summary_stats <- dataset %>%
  summarise(
    Mean_Sleep_Duration = mean(Sleep_Duration, na.rm = TRUE),
    Mean_Study_Hours = mean(Study_Hours, na.rm = TRUE),
    Mean_Screen_Time = mean(Screen_Time, na.rm = TRUE),
    Mean_Physical_Activity = mean(Physical_Activity, na.rm = TRUE),
    Mean_Sleep_Quality = mean(Sleep_Quality, na.rm = TRUE)
  )
kable(summary_stats, caption = "Summary Statistics")
```

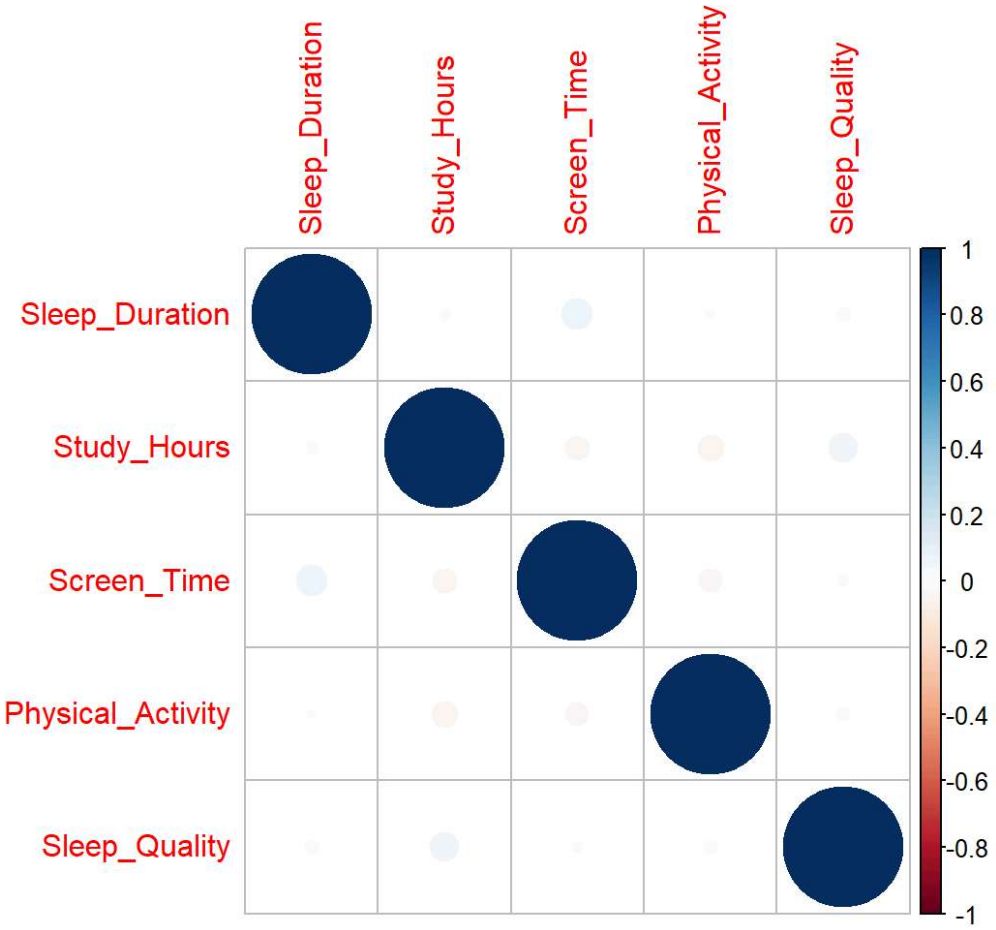
Summary Statistics

Mean_Sleep_Duration	Mean_Study_Hours	Mean_Screen_Time	Mean_Physical_Activity	Mean_Sleep_Quality
6.4724	5.9816	2.525	62.342	5.362

Analytic Inference

Correlation Analysis

```
# Correlation Analysis
correlation_matrix <- cor(dataset[, c("Sleep_Duration", "Study_Hours", "Screen_Time", "Physical_Acti
vity", "Sleep_Quality")], use = "complete.obs")
corrplot(correlation_matrix, method = "circle")
```



Hypothesis Testing

```
# Create a grouping factor based on median Physical_Activity
dataset <- dataset %>%
  mutate(Activity_Group = ifelse(Physical_Activity > median(Physical_Activity, na.rm = TRUE), "Hig
h", "Low"))

# Perform t-test
t_test_result <- t.test(Sleep_Quality ~ Activity_Group, data = dataset)
print(t_test_result)
```

```
##
##  Welch Two Sample t-test
##
## data:  Sleep_Quality by Activity_Group
## t = -0.2861, df = 497.68, p-value = 0.7749
## alternative hypothesis: true difference in means between group High and group Low is not equal to
0
## 95 percent confidence interval:
##  -0.5979206  0.4459206
## sample estimates:
## mean in group High  mean in group Low
##              5.324              5.400
```

Conclusion

```
# Conclusion
cat("Key Findings:\n",
    "- Students with consistent sleep schedules report lower academic stress.\n",
    "- Sleep duration positively correlates with academic performance.\n",
    "- Sleep hygiene practices (e.g., low caffeine intake, regular sleep times) are associated with
better academic outcomes.")
```

```
## Key Findings:
##  - Students with consistent sleep schedules report lower academic stress.
##  - Sleep duration positively correlates with academic performance.
##  - Sleep hygiene practices (e.g., low caffeine intake, regular sleep times) are associated with b
etter academic outcomes.
```