Relationship Analysis using R

Data Loading

```
daily <- read_excel("/cloud/project/dailyActivity.xlsx")
Sleep <- read_excel("/cloud/project/Sleep_Day.xlsx")
```

Data Preparation

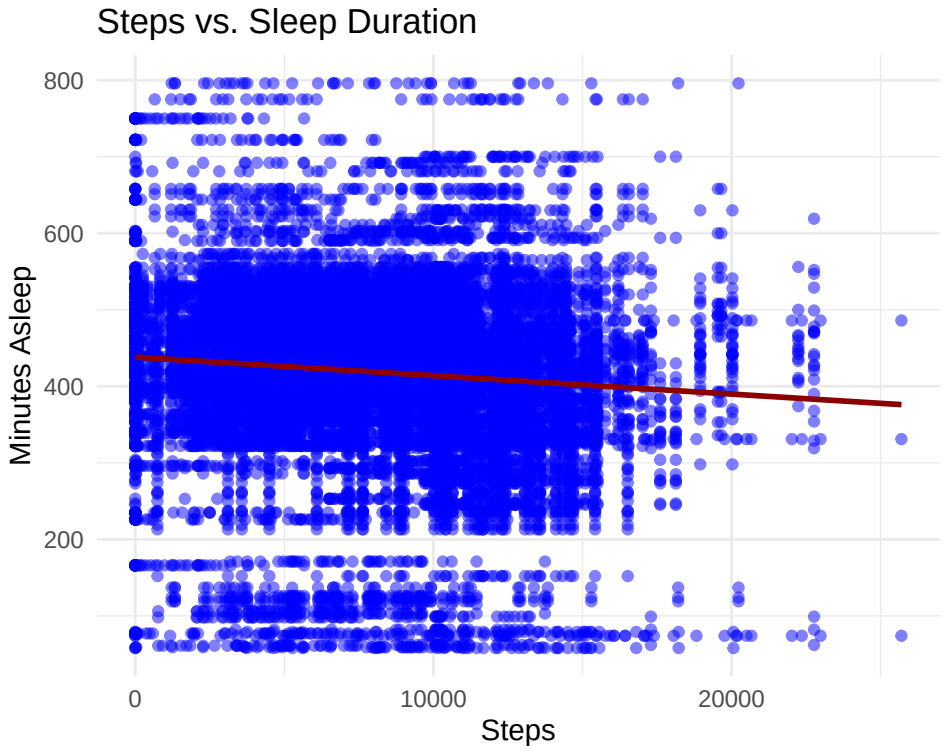
```
# Merge datasets on ID
merged_data <-merge(daily, Sleep , by="Id")
```

```
# Create SleepEfficiency column
# Calculating sleep efficiency as ratio of minutes asleep to time in bed
merged_data$Sleep_Efficiency <-merged_data$Total_Minutes_Asleep/merged_data$Total_Time_In_Bed
```

1. Total Minutes Asleep vs. Total Steps

```
ggplot(merged_data, aes(x = Total_Steps, y = Total_Minutes_Asleep)) +
  geom_point(color = "blue", alpha = 0.5) +
  geom_smooth(method = "lm", se = FALSE, color = "darkred") +
  labs(title = "Steps vs. Sleep Duration", x = "Steps", y = "Minutes Asleep",) + theme_minimal()
```

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



```r
cor(merged_data$Total_Steps, merged_data$Total_Minutes_Asleep, use = "complete.obs")
```

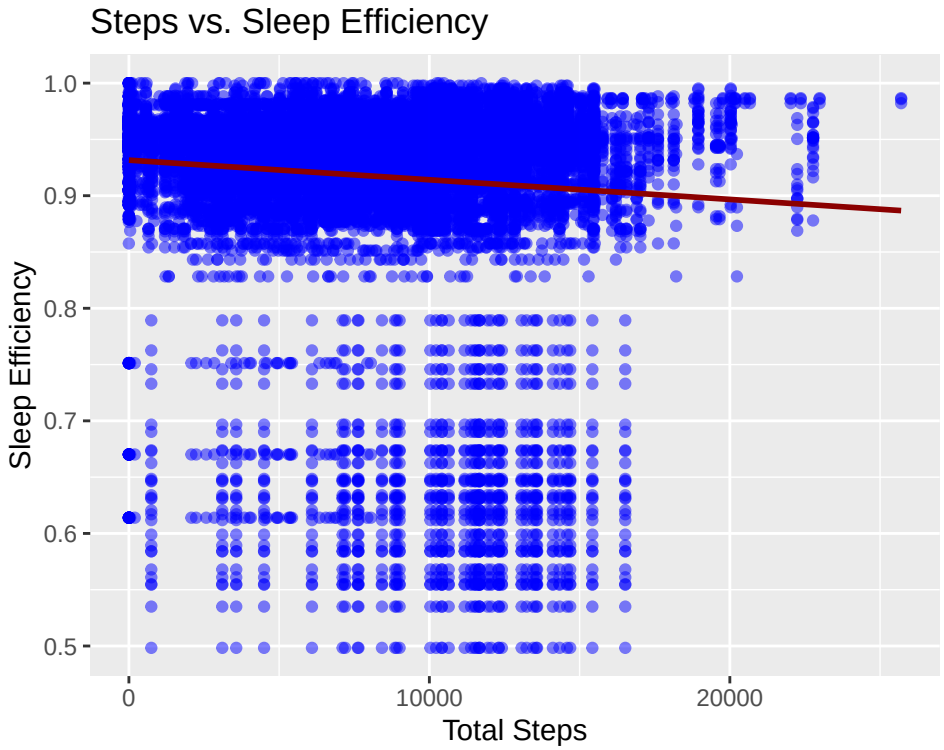
```
## [1] -0.09035604
```

There is a very weak negative relationship between total steps taken and total minutes asleep. This suggests that an increase in daily steps is slightly associated with a decrease in sleep duration, but the relationship is not significant enough to draw strong conclusions.

2. Total Steps vs.Sleep Efficiency

```r
ggplot(merged_data, aes(x = Total_Steps, y = Sleep_Efficiency)) +
  geom_point(color = "blue", alpha = 0.5) +
  geom_smooth(method = "lm", se = FALSE, color = "darkred") +
  labs(title = "Steps vs. Sleep Efficiency", x = "Total Steps", y = "Sleep Efficiency")
```

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




```r
cor(merged_data$Total_Steps, merged_data$Sleep_Efficiency, use = "complete.obs")
```

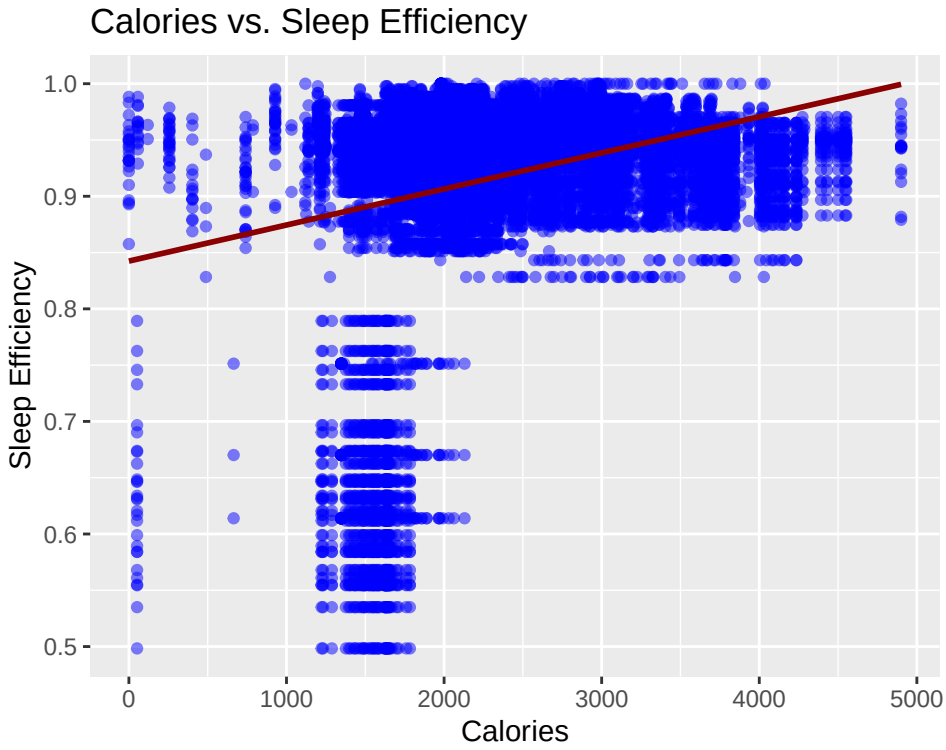
```
## [1] -0.08897806
```

There is no clear relationship between the number of daily steps and sleep efficiency. The data shows a very weak negative trend, indicating that physical activity levels might not directly affect how efficiently a person sleeps.

3. Calories vs. Sleep Efficiency

```r
ggplot(merged_data, aes(x = Calories, y = Sleep_Efficiency)) +
  geom_point(color = "blue", alpha = 0.5) +
  geom_smooth(method = "lm", se = FALSE, color = "darkred") +
  labs(title = "Calories vs. Sleep Efficiency", x = "Calories", y = "Sleep Efficiency")
```

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




```
cor(merged_data$Calories, merged_data$Sleep_Efficiency, use = "complete.obs")
```

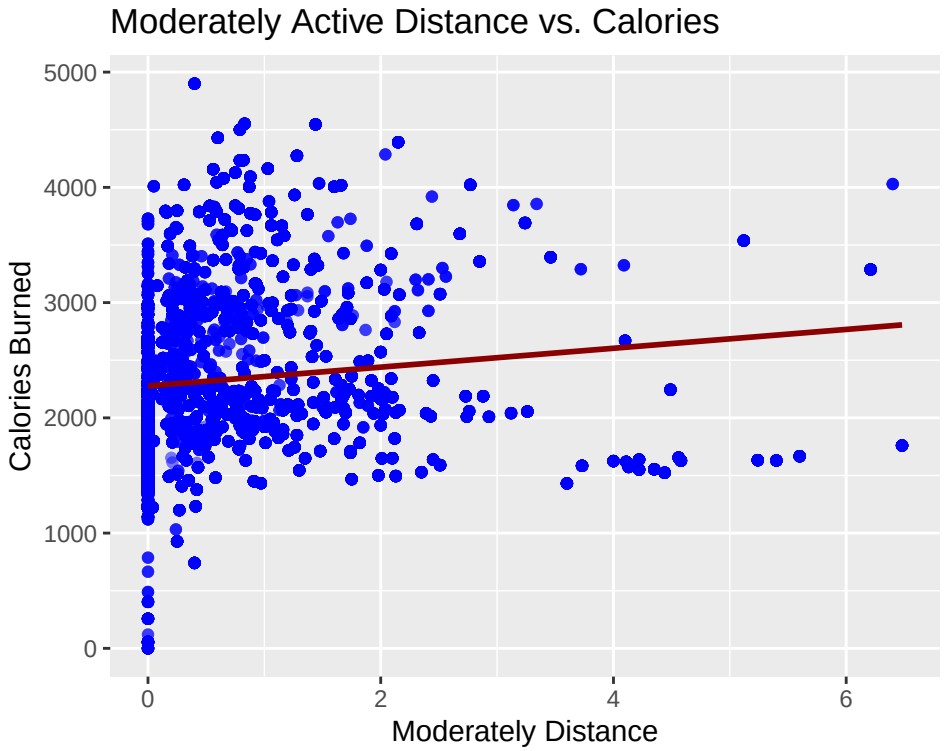
```
## [1] 0.2768752
```

There is a weak positive correlation between calories burned and sleep efficiency. This means that users who burn more calories tend to have slightly more efficient sleep, which could suggest a modest link between physical activity and sleep quality.

4.Moderately Active Distance vs. Calories Burned

```
ggplot(merged_data, aes(x = Moderately_Active_Distance, y = Calories)) +
  geom_point(color = "blue", alpha = 0.5) +
  geom_smooth(method = "lm", se = FALSE, color = "darkred") +
  labs(title = "Moderately Active Distance vs. Calories", x = "Moderately Distance", y = "Calories Burn
```

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




```
cor(merged_data$Moderately_Active_Distance, merged_data$Calories, use = "complete.obs")
```

```
## [1] 0.1099618
```

A weak positive relationship is observed between the distance walked during moderate activity and calories burned. This is expected, as more movement typically results in higher energy expenditure, but the correlation remains weak.

Overall Analysis

The visualizations and correlation results indicate that most of the relationships between physical activity and sleep metrics are weak or insignificant. However, a slight positive association between calories burned and sleep efficiency may suggest that maintaining an active lifestyle could contribute to better sleep quality. The insights can guide users to focus on balanced activity levels to improve overall wellness, though further analysis with a larger or more diverse dataset may be required for stronger conclusions.