Sleep Analysis

Group 1: Cassidy Cubra, Jack Kenzakowski,
Finn Mokrzycki, Hilde Younce, Qais Youssef

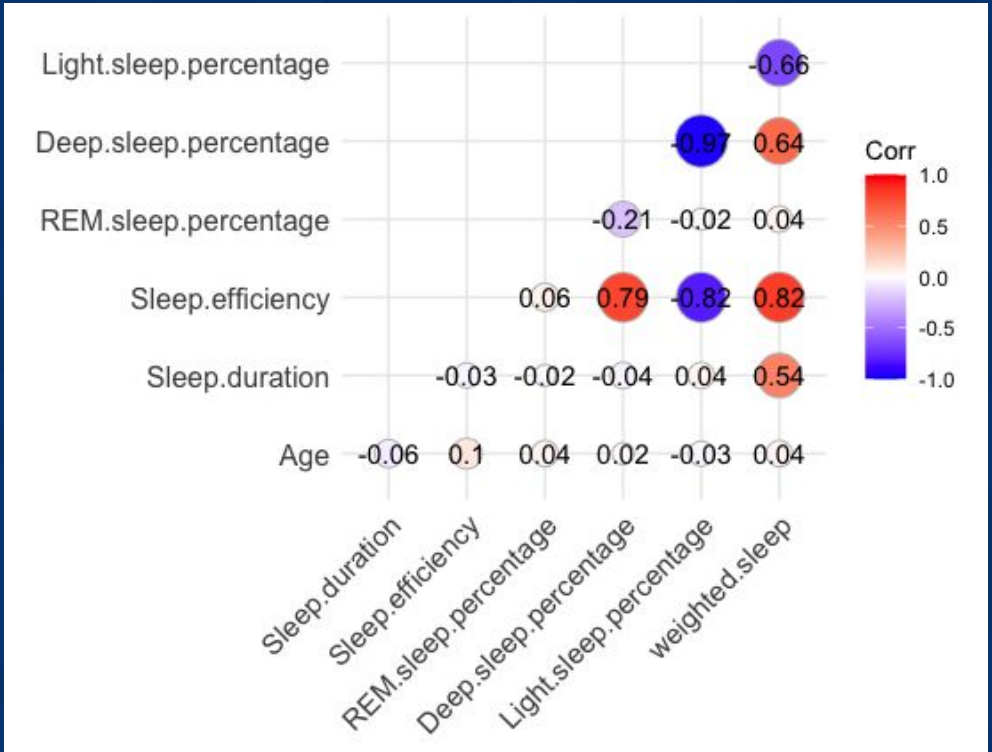
Data Summary

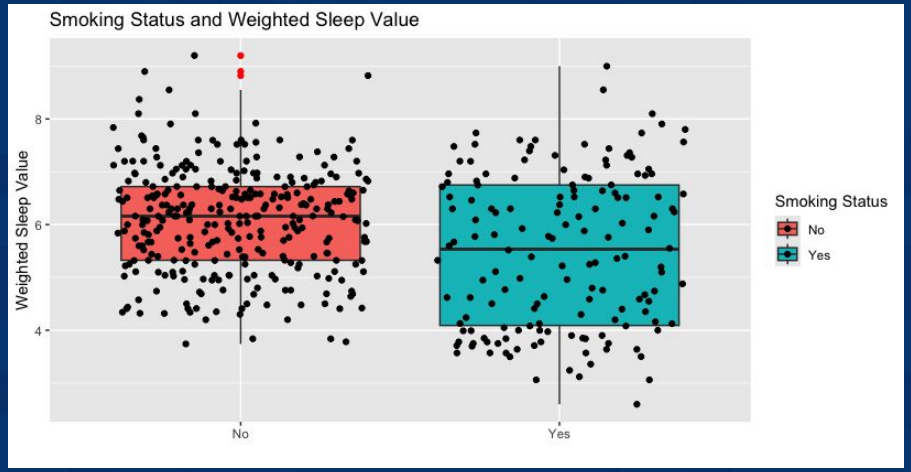
- Sleep Efficiency Dataset
- 452 observations
- Contains:
  - Age, Gender
  - Sleep Duration, Sleep efficiency, Weighted sleep
  - Caffeine consumption, Alcohol consumption, Smoking status, Exercise frequency

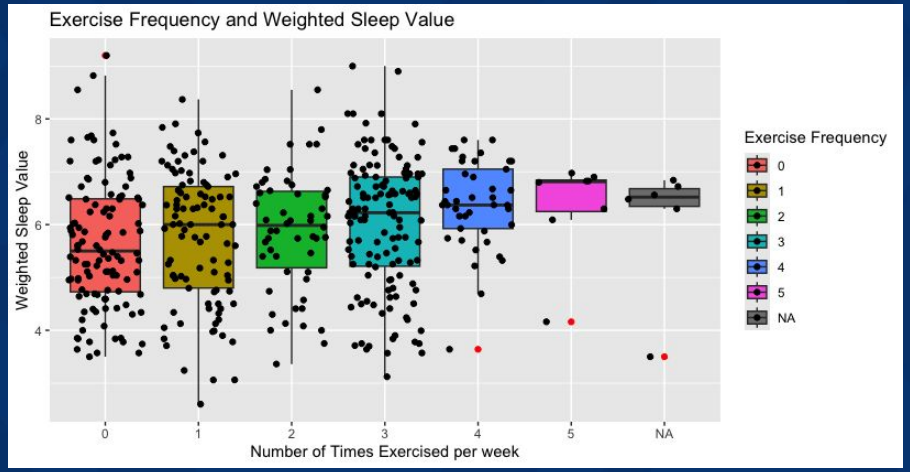
Research Questions

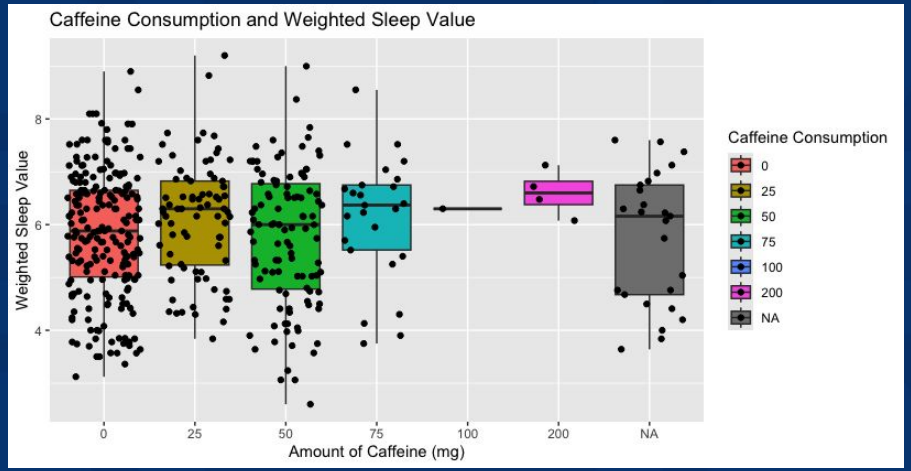
- How does consumption affect sleep duration and efficiency?
- Which behaviors have the largest impact on sleep?
- Can we predict how much sleep someone will get?
- Are college students getting enough sleep?

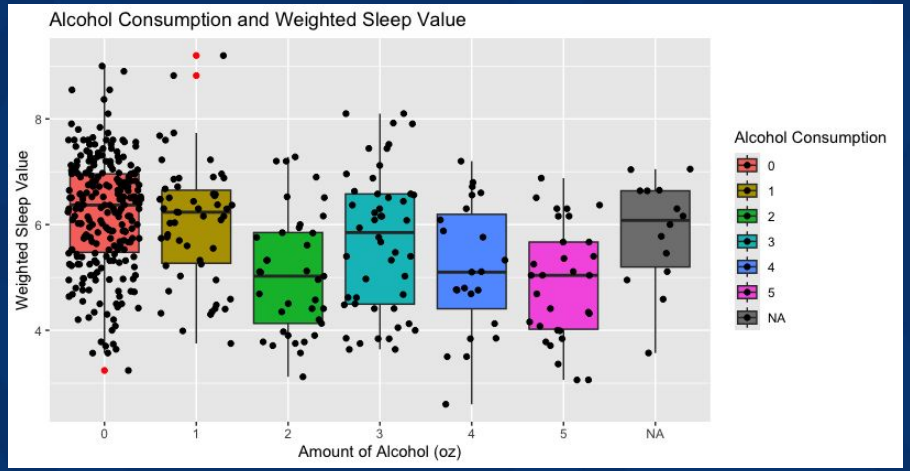
EDA & Visualizations




Smoking Status and Weighted Sleep Value
Weighted Sleep Value
8
6
4
No
Yes
Smoking Status
No
Yes


Exercise Frequency and Weighted Sleep Value
Weighted Sleep Value
8
6
4
0
1
2
3
4
5
NA
Number of Times Exercised per week
Exercise Frequency
0
1
2
3
4
5
NA


Caffeine Consumption and Weighted Sleep Value
Weighted Sleep Value
8
6
4
0
25
50
75
100
200
NA
Amount of Caffeine (mg)
Caffeine Consumption
0
25
50
75
100
200
NA


Alcohol Consumption and Weighted Sleep Value
Weighted Sleep Value
8
6
4
0
1
2
3
4
5
NA
Amount of Alcohol (oz)
Alcohol Consumption
0
1
2
3
4
5
NA

Multiple Linear Regression Model for Sleep Efficiency

```
## Call:
## lm(formula = weighted.sleep ~ Caffeine.consumption + Alcohol.consumption +
##     Smoking.status + Exercise.frequency, data = sleep)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4926 -0.7818 -0.0038  0.7429  3.3723
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)           6.0456482  0.1145083  52.797  < 2e-16 ***
## Caffeine.consumption  0.0007615  0.0018962   0.402    0.688
## Alcohol.consumption  -0.2370237  0.0341465  -6.941 1.57e-11 ***
## Smoking.statusYes    -0.5071377  0.1152799  -4.399 1.39e-05 ***
## Exercise.frequency    0.1560424  0.0377602   4.132 4.37e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

- Predicting weighted sleep using caffeine and alcohol consumption, smoking status, and exercise frequency
- Alcohol and smoking were the most significant predictors

Predictions

Case 1: Unhealthy 23 year old

```
caseA <- data.frame(Age=23, Caffeine.consumption=200, Alcohol.consumption=5,
                    Smoking.status="Yes", Exercise.frequency=0)

predict(lm.weightsleep, caseA, interval="prediction")
```

```
##        fit      lwr      upr
## 1 4.505685 2.221789 6.789581
```

Case 2: 50 year old health nut

```
caseB <- data.frame(Age=50, Caffeine.consumption=0, Alcohol.consumption=0,
                    Smoking.status="No", Exercise.frequency=5)

predict(lm.weightsleep, caseB, interval="prediction")
```

```
##       fit      lwr      upr
## 1 6.82586 4.649688 9.002032
```

Case 3: Average college student on a Friday night

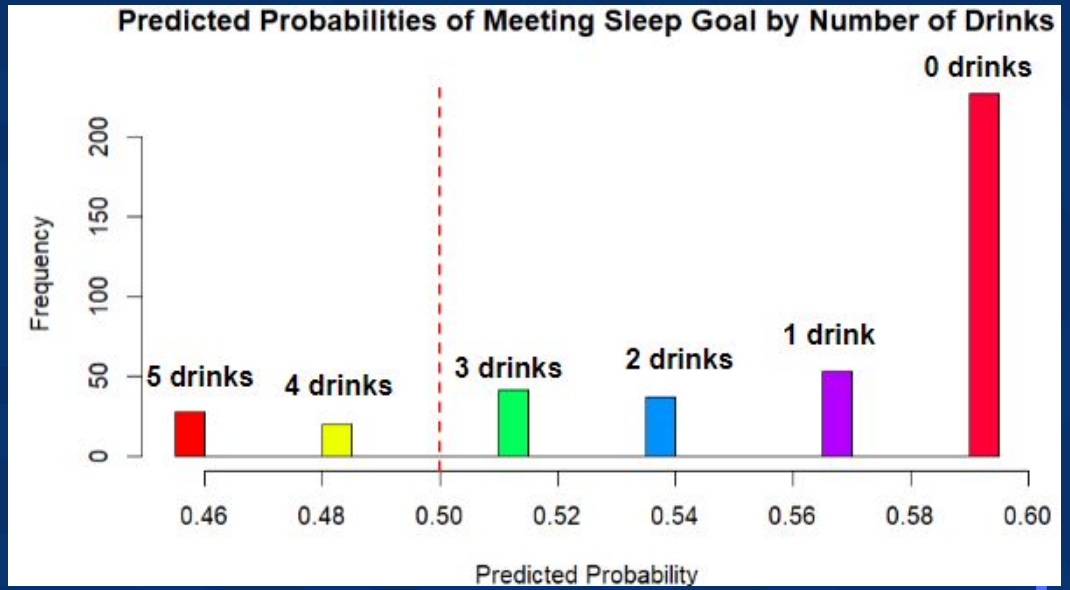
```
caseC <- data.frame(Age=22, Caffeine.consumption=100, Alcohol.consumption=3,
                    Smoking.status="No", Exercise.frequency=3)

predict(lm.weightsleep, caseC, interval="confidence")
```

```
##        fit      lwr      upr
## 1 5.878851 5.509233 6.248468
```

Logistic Regression Model

Did we meet our sleep goal of 7.5 hours?

- Used Number of Drinks to predict probability of meeting sleep goal



Questions?