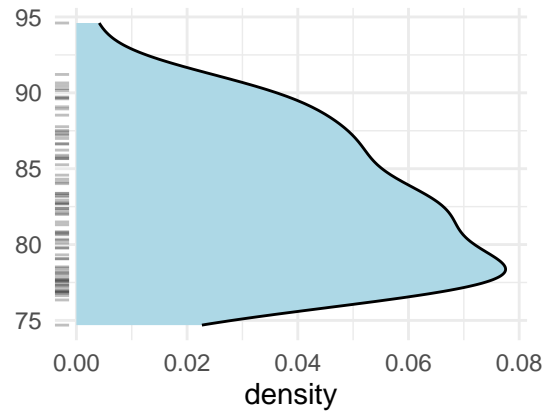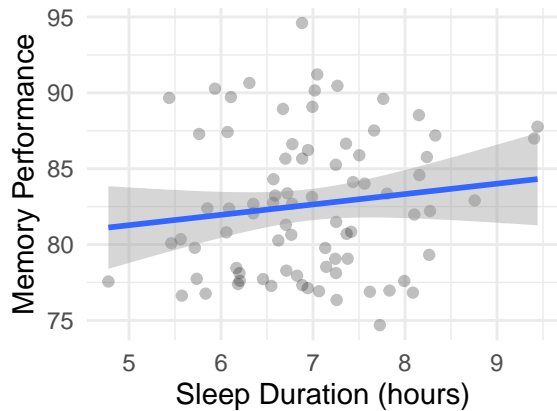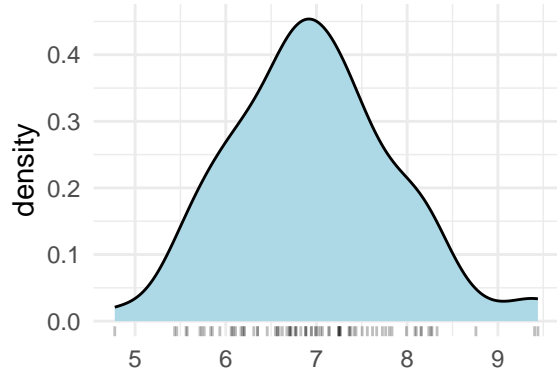# Homework 4

Author: Matthew J. Crossley

Last update: 23 May, 2024

Q1

Consider an investigation into the relationship between the duration of sleep and performance on a memory task. Below is a plot of the data collected, along with the predictions from the best fitting linear model.

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



```
##
## Call:
## lm(formula = memory_performance ~ sleep_duration, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.4508 -4.0698 -0.5393  3.2803 12.0390
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)      77.8549      4.0697    19.13    <2e-16 ***
## sleep_duration    0.6841      0.5800     1.18     0.242
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.637 on 78 degrees of freedom
## Multiple R-squared:  0.01752,    Adjusted R-squared:  0.004929
## F-statistic: 1.391 on 1 and 78 DF,  p-value: 0.2418
```

- Please write the equation of the best fitting linear model.

- Please list all random variables in the best fitting linear model and state how they are distributed.

- Please draw a pointrange plot illustrating the best fittings β coefficients with error bars showing SEM.

- What is Pearson's correlation coefficient between sleep duration and memory performance?

- Please write a few sentences reporting the results of this analysis.

Q2

Consider a study in which researchers collected data on participants' age and education level and examined how these variables predict performance on a memory task. Below is a plot of the data along with the predictions from the best-fitting multiple regression model.

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```
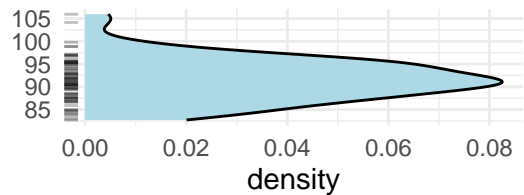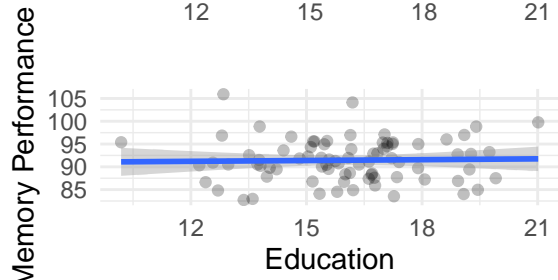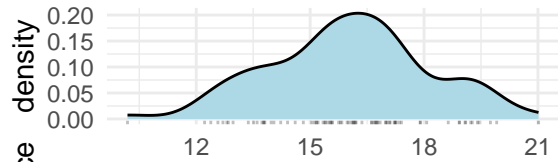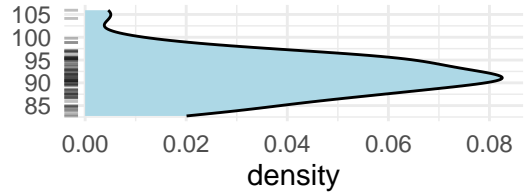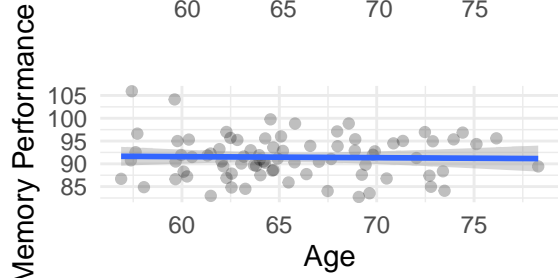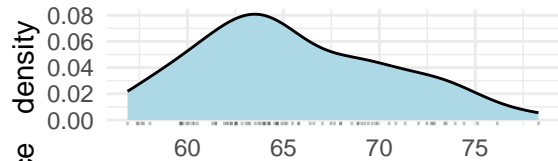
```
## 
## Call:
## lm(formula = memory_performance ~ age + education, data = d)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -8.4433 -3.4024 -0.1586  3.3605 14.4783 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 92.10810    7.64862  12.042   <2e-16 ***
## age         -0.02659    0.10573  -0.251    0.802    
## education    0.06764    0.25206   0.268    0.789    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 4.677 on 77 degrees of freedom
## Multiple R-squared:  0.001583,   Adjusted R-squared:  -0.02435 
## F-statistic: 0.06102 on 2 and 77 DF,  p-value: 0.9408
```

- Please write the equation of the best fitting linear model.

- Please list all random variables in the best fitting linear model and state how they are distributed.

- Please draw a pointrange plot illustrating the best fittings β coefficients with error bars showing SEM.

- Please draw a boxplot illustrating the residuals of the best fitting linear model. Please set the whiskers to extend to the min and max values. What random variable in your linear model does this plot tell you about?

- What is Pearson's correlation coefficient between (1) age and memory performance, and (2) education level and memory performance?

- What is the partial correlation between age and memory performance when controlling for education level?

- Please write a few sentences reporting the results of this analysis for an academic journal.

- Please write R code (using your pen or pencil) that generates the regression analysis reported above.