PSC 103B - Lab 1 Assignment
Answer Key

Data

For this homework assignment, we will be using a subset of data from the Nerdy Personality Attributes Scale (NPAS). This is a real scale, and if you're interested, you can take the quiz yourself: https://openpsychometrics.org/tests/NPAS/

The original dataset had over 25,000 responses from all over the world. In order to make things a bit simpler, I chose 200 responses randomly from each continent (in order to prevent the responses from being too skewed towards one continent, such as the Americas). There are a bunch of variables in this dataset, but the ones we'll be interested in for this assignment are:

- `nerdy_scale`: Participant's average "nerdiness" score from the NPAS (1-5 scale)
- `TIPI1`: How extraverted or enthusiastic the participant rates themselves (1-7 scale)
- `TIPI3`: How dependable or self-disciplined the participant rates themselves (1-7 scale)
- `nerdy_selfreport`: Do you see yourself as someone who is nerdy? (1-7 scale)

Question 1

Please read the NPAS dataset (NPAS.csv) into R using read.csv(). Create a dataset that only contains the variables listed above (recall: how do you subset columns from a dataframe?) (.5 pt). Show the first 6 rows (.5 pt) and last 6 rows (.5 pt) of the new dataset. Show your code.

```
npas <- read.csv("data/NPAS.csv")
npas_subset = npas[ , c("nerdy_scale", "TIPI1", "TIPI3", "nerdy_selfreport")]

head(npas_subset, n=6)
```

```
  nerdy_scale TIPI1 TIPI3 nerdy_selfreport
1    3.307692     4     1                5
2    4.038462     3     4                6
3    3.653846     6     3                3
4    2.423077     2     2                4
5    3.576923     4     6                4
6    4.500000     2     3                7
```

```
tail(npas_subset, n=6)
```

```
     nerdy_scale TIPI1 TIPI3 nerdy_selfreport
995     3.692308     2     5                5
996     3.961538     1     5                6
997     4.000000     3     7                5
998     4.038462     5     7                7
999     3.384615     2     5                4
1000    4.076923     3     3                6
```

Question 2

Show the correlations among all the variables in the new dataset (1pt). Show your code.

```
round(cor(npas_subset, use = "complete.obs"), 3)
```

```
                 nerdy_scale  TIPI1   TIPI3 nerdy_selfreport
nerdy_scale            1.000 -0.262   0.020            0.532
TIPI1                 -0.262  1.000   0.069           -0.172
TIPI3                  0.020  0.069   1.000           -0.039
nerdy_selfreport       0.532 -0.172  -0.039            1.000
```

```
# Note: the round() function is not necessary, but I used it to make my output prettier
```

Question 3

Test whether the correlation between the participants' nerdiness scores on the NPAS (nerdy_scale) their level of extraversion is significant (1 pt). Report this correlation using APA format (.5 pt). Show your code.

```
cor.test(npas_subset$nerdy_scale, npas_subset$TIPI1,
         method = "pearson", alternative = "two.sided")
```

```
	Pearson's product-moment correlation

data:  npas_subset$nerdy_scale and npas_subset$TIPI1
t = -8.5754, df = 998, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3187849 -0.2032772
sample estimates:
      cor
-0.261969
```

We examined the relation between extraversion and participants' nerdiness score on the NPAS in a sample of 1000 participants. The Pearson's product-moment correlation between extraversion and nerdiness is negative, statistically significant, and small, $r = -0.26$, $t(998) = -8.58$, $p < .001$.
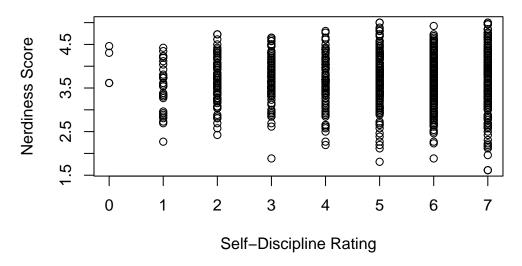
Question 4

Create a scatterplot to reveal the relationship between participants' level of self-discipline and their nerdiness score on the NPAS (make sure to display their level of self-discipline on the x- axis and nerdiness score on the y-axis, and include appropriate axis labels and a title) (1 pt). Use a few words to describe the general trend revealed from the plot (1 pt). Show your code.

Note: This plot might look a little weird (e.g., a bunch of straight lines), and that is because the level of self-discipline is an ordinal variable, so it cannot take on all possible values between 1 and 7, only the integers.

```
plot(npas_subset$TIPI3, npas_subset$nerdy_scale, xlab = "Self-Discipline Rating",
     ylab = "Nerdiness Score",
     main = "Relation Between Self-Discipline \nand Nerdiness")
```

3

Relation Between Self–Discipline
and Nerdiness

The plot suggests no correlation exists between nerdiness scores and self-discipline ratings, regardless of their level.

Question 5

Build up a linear model to test if level of self-discipline can significantly predict the nerdiness score (1 pt). Show your code and the summary of the model.

```
nerdy_model <- lm(nerdy_scale ~ TIPI3, data = npas_subset)
summary(nerdy_model)
```

```
Call:
lm(formula = nerdy_scale ~ TIPI3, data = npas_subset)

Residuals:
     Min       1Q   Median       3Q      Max
-2.09896 -0.34761  0.03014  0.40275  1.29937

Coefficients:
           Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 3.666318    0.055883  65.607    <2e-16 ***
TIPI3        0.006862    0.010742   0.639     0.523
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5744 on 998 degrees of freedom
Multiple R-squared:  0.0004087, Adjusted R-squared:  -0.0005929
F-statistic: 0.4081 on 1 and 998 DF,  p-value: 0.5231
```

Question 6

Interpret the intercept (.5 pt) and slope (.5 pt) of your model.

The intercept is 3.67, which means that when a person obtains a score of 0 on the self- discipline rating, their expected nerdiness score is 3.67.

The slope is 0.007, which means that for every 1-point increase on the rating of self-discipline, a participant's nerdiness score is expected to increase by 0.007.

Question 7

Report and interpret the (unadjusted) R^2 value of your model (1 pt).

The R^2 value is 0.0004, which means that approximately 0.04% of the variation in nerdiness scores is explained by participants' self-discipline scores.

Question 8

Report the results of this linear regression in APA format (.5 pt).

We ran a simple linear regression to determine whether self-discipline rating predicts nerdiness score on the NPAS. Self-discipline ratings did not significantly predict nerdiness scores ($b = 0.007$, 95% CI [-0.01, 0.03], $t(998) = 0.64$, $p = .523$). Self-discipline explains a statistically not significant and very weak proportion (0.04%) of variance in nerdiness scores.

Question 9

Assuming a person gives themselves a self-discipline rating of 4, what is their expected nerdiness score on the NPAS based on the linear model from Q5? (.5 pt).

$$\widehat{Nerdiness}_i = b_0 + b_1 \times SelfDiscipline_i$$
$$\widehat{Nerdiness}_i = 3.67 + 0.007 \times 4$$
$$\widehat{Nerdiness}_i = 3.70$$

Their expected nerdiness score would be 3.70