1. Think of 10 people you know and indicate how much you like them,
on a scale from 0 to 10 (0 means you do not like the person at all,
10 means you like the person extremely well)
This is your outcome variable Y
For example, Y=c(6, 7, 5, 9, 8, 8, 6, 9, 6, 10)
The covariate X indicates whether the person in question has a dog,
for example, X=c(0,0,0,1,1,0,1,0,0,0), with 0 for no dog and 1 for dog
YOU SHOULD USE YOUR OWN COLLECTED DATA, NOT MINE
Make sure there are at least two with a dog.
- Analyse the data with lm (Y ~ X).
- What is the proportion of variance explained by owning a dog. Underline.
- What is the t value of the slope of X? Underline.
- Also conduct a t test (two-tailed)
- What is the t value? Underline.
Show your code and output for the regression analysis and for the t test.

2. Predict GPA based on SATM and SATV using the College Success data set from JASP.
- What is the correlation between SATM and SATV? Underline the correlation.
- Write down the regression equation for the standardized variables (they are the standardized
regression coefficients).
For example (an arbitrary example), estimated sGPA = 0.4*sSATM + 0.1*sSATV (* denotes
multiplication, lower case s denotes standardized variable).
Underline the equation.

3. X = c (0, 4, 3, 5, 2, 4)
Center X, Standardize X, and report the mean and standard deviation of X, centered X (cX) and
standardized X (sX). Underline the three means and the three standard deviations and indicate
which means and standard deviations are equal.

4. Y = c (1, 0, 1, 2, 2, 3, 5, 4) pounds lossed
X1 = c (0, 0, 0, 0, 1, 1, 1, 1) X1=0 no diet, X1=1 diet
X2 = c (0, 0, 1, 1, 0, 0, 1, 1) X2=0 no increased PA, X2=1 increased PA
estimate the model with Y ~ X1*X2 (original X1 and X2)
also estimate the same model with centered X1 and X2
(i.e., with original Y and with centered X1 and X2)
- Make a table with two columns to report the two times four p values.
Underline when they are equal in both columns.
The first column is for the analysis with the data before centering.
The second column is for the analysis after the covariates are centered.

5. Y = 10 + 2X1 + 3X2 + 2X12   X12 = X1*X2   (* denotes multiplication)
Use the following combinations of values to determine Y:
X1=1   X2=2   Y=?

X1=1   X2=-2   Y=?
X1= -1   X2=2   Y=?
X1= -1 X2=-2   Y=?
Underline the four resulting values.
Explain how you can see the interaction in the four values.

6. What is the name of the experimental station in England where Ronald Fisher invented analysis of variance and experimental designs. Underline the name of the experimental station.

7. Y = c (4, 5, 6, 9, 8, 7)
X = c (0, 0, 0, 1, 1, 1)
conduct an ANOVA with Y as outcome variable and X as a factor with two levels
Show the output and underline F and p.

8. Is the following design orthogonal? Answer with yes or no and underline.

	B1	B2
A1	N=10	N=5
A2	N=10	N=15

Show the row-wise proportions (two times two). Underline the 2 x 2 proportions.
Show the column-wise proportions (two times two) and underline. Underline the 2 x 2 proportions.

9. Test the null hypothesis that the variance of X1 and X2 are equal
X1 = c (1, 2, 3, 4)
X2 = c (3, 6, 9)
What is the F value? Write down as $F (df1=?, df2=?) = ?$ and underline.
What is the p value? Underline.
Do you reject H0? yes of no. Underline.

10. A sample of 20 female participants in a study is subject to three conditions: A1, A2, A3, the levels of factor A. Another sample, of 30 male participants in the same study is subject to the same three conditions: A1, A2, A3, the same three levels of factor A. Because of the two samples, female and male participants, the design has a second factor, called B, with two levels: B1 (female participants), and B2 (male participants)
- Is this design a within-groups or a between-groups, or a combination of both?
Underline one of the following three: within, between, combination
- Is this an orthogonal design?
Underline yes or no.

11. Y=c(0, 1, 2, 1, 2, 3, 5, 6, 7)
A=c(1, 1, 1, 2, 2, 2, 3, 3, 3)
- Calculate the mean squares for A
- Calculate the mean squares of the residuals

Underline the two mean squares values and F.
Show your calculations.
Check with model=aov(Y~A) and show the output.

12. Y=c(3, 4, 6, 1, 2, 4, 2, 5, 8, 0, 2, 4, 0, 3, 3, 3, 6, 9)
A=c(1, 1, 1, 2, 2, 2, 3, 3, 3, 1, 1, 1, 2, 2, 2, 3, 3, 3)
B=c(1, 2, 3, 1, 2, 3, 1, 2, 3, 1, 2, 3, 1, 2, 3, 1, 2, 3)
What are the degrees of freedom in the denominator to calculate
- mean squares of A?
- mean squares of B?
- mean squares of the interaction A*B?
- mean squares of the residuals?
Underline your responses.

13. For the Singers data set, calculate:
- prior odds of the Gender + Pitch model, underline
- posterior odds of the Gender + Pitch model, underline
- the BF (BF subscript M) as a multiplication factor to obtain the posterior odds from the prior
odds of the Gender + Pitch model
Show your calculation of the BF (subscript M)

14. First, use a within-groups ANOVA to analyze the risk rating data without the data from
Person 5. In quiz 11, you were asked the same but without Persons 4 and 5. What are the three
MS values, for A, P, and residuals when just Person 5 is omitted? Show your calculations and
show and underline the three MS values,
all three on the same seperate line.
Second, use a between-groups ANOVA for the same data (again without Person 5), and show
that the SS residuals is the sum of two SS values from the within-groups ANOVA. Report the SS
residuals from the between-groups ANOVA and show it is the sum if two SS values from the
within-groups analysis.
Show the sum on a separate line and underline.

15. First omit the Q1 data from the risk rating data but keep the data from all five persons. Then
calculate for Person 1 what the result is of the decomposition in terms of latvar, Q2, and Q3 and
the two residuals (one for each of the two questions). Show and underline the five values (latvar,
Q2, Q3, two residuals) on the same separate line.

16. Here are data from three persons in two conditions
Y=c(2,2,3,3,4,5)   note that this line has changed on Friday at 5pm
A=c(1,1,1,2,2,2)
P=c(1,2,3,1,2,3)
Conduct a paired sample t test
Show and underline the t value on a separate line

17. Here are outcomes from four small independent groups
Y= c (1, 3, 3, 5, 5, 7, 7, 9)
A= c (1, 1, 2, 2, 3, 3, 4, 4)
Use dummy coding for regression (with lm) and report the F value
Use aov to check the F value
Show and underline the two F values on one separate line

17. The first question on quiz 4 was the following:
"In a sample of college students, the observed IQs are 117, 129, 120, 121
what is the estimated sd of the means (i.e., standard error of the mean)?"
Suppose the IQs had been 119, 131, 122, 123 instead.
- What are the two standard errors of the mean? Underline the two standard errors.
- Are they equal of different? Underline equal or different and explain why equal or different
referring to the standard deviations of the data.

18. In a study on the effect of a diet,
the weight losses in pounds for six adults are 3, 4, 2, 5, 1, 1 in one week time
H1: the diet makes you loose more than one pound in one week,
H0: the diet makes you loose one pound or less in one week, $\mu\_0=1$.
- Determine the p-value (one-tailed), and underline
(do not forget to use the mu value to test the null hypothesis)
You can either use the equation for t or the t.test ( ) function

19. This is too small a study, but suppose that for five women and five men of the same age and
with the same kind of job, the hourly wages are as follows:
women: 12, 13, 14, 14, 15;
men: 11, 12, 15, 16, 17.
Test the null hypothesis that the wages for men and women are equal.
- What are the t statistic and the p value? Underline your responses.

20. Y= c (70, 71, 72, 73, 75, 75, 76, 77, 77),
X= c (0, 0, 0, 0, 0, 1, 1, 1, 1).
Y: years of life,
men: X = 0, women: X = 1.
- What is the t statistic (underline) comparing women to men using the linear model?
The null hypothesis is that there is no difference.
Show your code and the full output.

21. Here are two pairs of samples. Are they paired samples?
- Two random samples of size 100 (200 students in total) from the same school. One sample of
female juniors and another of male juniors. Paired: yes or no?
- Two samples of size 100, with a randomly sampled male and female junior student from the
same school, which is each time one school from 100 different schools (200 students in total).
Paired: yes or no?

22. Here are hourly wages for married couples. The first of each pair is for the wife, the second is for the husband:
couple 1: 15, 17
couple 2: 14, 15
couple 3: 16, 16
couple 4: 14, 15
couple 5: 15, 14
couple 6: 16, 18
H1: the hourly wages of married women are lower than those of their husbands
H0: the hourly wages of married women are equal to or higher than those of their husbands
Use the t.test ( ) function with two sets of outcomes. What is the t value, what is the p value? Underline your responses and show your code.

23. For a similar data set of married couples:
women: 15, 15, 14, 16, 13
men: 15, 14, 16, 18, 17 (their corresponding husbands)
What is the standard error of the mean difference, and what are the t value and the p value, given that H1 is that women's wages are lower?

24. The average grade of a class of 16 students is 88/100, the sd is 4. What is the confidence interval of the mean of 88?

25. X = c (4, 6, 3, 5, 3)
Y = c (1, 0, 1, 3, 0)
Transform X and Y into z scores.
Calculate the mean of X and Y, the sd (s) of X and Y, and the correlation between X and Y. Do the same after the transformation.

Underline the two times five values.

26. Use the same data as for Q6a, the original ones (not the z scores). Predict Y based on X, show your code, and underline the regression coefficient (slope) of X

Q6c Determine the intercept and slope given
Y = c (15, 13, 14, 11, 10)
X = c (13, 12, 13, 10, 10)
for Y as dependent variable and X as independent variable, in two ways:
(1) using lm(Y~X), show your output and underline intercept and slope
(2) using equations 6.5 and 6.6. Show and underline your calculation

27. The z score of an X value is 3, the correlation between X and Y is 0.70.
What is the predicted z score of Y? Underline the predicted z score