Q1)

Seizures → dependent variable (continuous)

age → independent variable (categorical)

Hours Slept → independent variable (continuous)

```
>> Researcher = [1;1;1;1;1;2;2;2;2;2;3;3;3;3;3;4;4;4;4;
   4]
>> Seizures = [5;16;4;2;9;96;52;39;22;33;1;2;5;4;2.6;8;4;
   2;6.8;8]
>> Age = ['0';'Y';'Y';'Y';'0';'Y';'0';'Y';'Y';'Y';'Y';'Y';'Y';'0';'0';'0';
   '0';'0';'0';'0']
>> HoursSlept = [5;8;9;1;0;15;15;13;19;10;12;1;5;6;8;3;4
   8;6;9]
```

A)

LINEAR

```
>> T = table(Seizures, Age, HoursSlept)
>> LM = fitlm(T, 'Seizures ~ Age + HoursSlept')
```

```
>> LM.ModelCriterion
J
AIC : 178.33
BIC : 181.32
```

```
Linear regression model:
    Seizures ~ 1 + Age + HoursSlept

Estimated Coefficients:
                   Estimate       SE        tStat      pValue
                   ________    ______    _______    _________

    (Intercept)    -7.1667      8.5125    -0.8419     0.41153
    Age_y           4.0179      9.118     0.44065    0.66502
    HoursSlept      2.7042     0.91681     2.9495    0.008968

Number of observations: 20, Error degrees of freedom: 17
Root Mean Squared Error: 19.5
R-squared: 0.383,  Adjusted R-Squared: 0.31
F-statistic vs. constant model: 5.27, p-value = 0.0165
```

→ ignore ; not important for our needs

there is a statistically significant difference btwn seizures and hours slept

there is NOT a statistically significant difference btwn Seizures and Age

MIXED EFFECTS

```
>> T2 = table(Seizures, Age, HoursSlept, Researcher)
>> LME = fitlme(T2, 'Seizures ~ Age + HoursSlept +
   (1|Researcher)')
```

```
Linear mixed-effects model fit by ML

Model information:
    Number of observations             20
    Fixed effects coefficients          3
    Random effects coefficients         4
    Covariance parameters               2

Formula:
    Seizures ~ 1 + Age + HoursSlept + (1 | Researcher)

Model fit statistics:
    AIC       BIC       LogLikelihood    Deviance
    181.75    186.73    -85.875          171.75

Fixed effects coefficients (95% CIs):
    Name                  Estimate    SE         tStat      DF    pValue     Lower       Upper
    {'(Intercept)'}       5.8389       10.533    0.55436    17    0.58656     -16.383    28.061
    {'Age_y'      }       1.3127       8.1808    0.16046    17    0.87441     -15.947    18.573
    {'HoursSlept' }       1.2197      0.94805     1.2865    17     0.2155    -0.78051    3.2199

Random effects covariance parameters (95% CIs):
Group: Researcher (4 Levels)
    Name1                 Name2                 Type          Estimate    Lower     Upper
    {'(Intercept)'}       {'(Intercept)'}       {'std'}       12.219      2.9285    50.98

Group: Error
    Name              Estimate    Lower     Upper
    {'Res Std'}       15.367      10.321    22.88
```

there is NOT a statistically significant difference btwn seizures and age

there is NOT a statistically significant difference btwn seizures and hours slept.

B)

1B) There is not a clear-cut conclusion to this question. Comparing AIC/BIC for the two models, the linear regression model is the better model as seen by a smaller AIC/BIC ratio, but the linear mixed-effects model contradicts the linear model in one critical aspect – the relationship between the number of seizures and hours slept. The linear model illustrates that there is a significant difference between number of seizures and hours slept, while the linear mixed effect model shows that there is not a significant difference between number of seizures and hours slept. Due to this, there is not a clear-cut answer to know if hours asleep can predict the number of seizures one will have. On the other hand, both models do agree that age cannot predict number of seizures. My conclusion would be different if I had used the linear regression model since I would not have known there could be some type of prediction between hours asleep and number of seizures (as shown by the linear mixed effects model). As a researcher, I would conclude that age cannot predict number of seizures and that there is a potential predicting ability between hours of sleep and number of seizures – more data and statistical analysis need to be applied to show a correlation.

Q2)

Summary of Article:

- CAT therapy → neurotic clients
  - 'uptake vs. credibility'
- worse than CBT
  - breaking boundaries btwn EBP and PBE
- CBT → very useful in convincing EBP and PBE across many psychological disorders
  - ↳ a family of approaches

A) - equivalence paradox (CAT excluded)

- **Methods (Problems)**
  - **no random assignment** → there should have been random assignment b/c this would rule out bias, BUT since clients were allocated due to the OPINION of the screening therapist, this puts bias since the screening therapist might prefer CBT over CAT or vice versa. There is now bias that can no longer be controlled.
  - **no standardized procedures** in the screening/allocation decisions → I believe they should have used a standardized screening test in determining what service the patient should receive b/c this would exclude bias and assure that all patients were fairly and accurately assessed.
  - full assessment of their mental health
  - using trained professionals for CBT and CAT service
  - self reports + therapist assessment section
  - using well-grounded tests (Brief Symptom Inventory + Beck Depression Inventory II + Inventory of Interpersonal Problems)

no critique ← ↳ no critique; this is what they should be doing

no critique → using these measures are good, but I wish they explained more + stated if they used one of a combination of them + which ones

B) Focus on **MANOVA** → used when there are multiple dependent variables or collections from the same individual

✓ Effectiveness of services

2B) I agree with their use of MANOVA because MANOVA is used when there are multiple dependent variable or collections from the same individual. They are measuring the outcomes of five measures for each service, in other words, five dependent variables for each independent variable (type of therapy). Therefore, I believe MANOVA was the correct statistical test to use since we are seeing how does the type of therapy and duration (short or long), which are both independent variables, affect their BDI-II, IIP-32, BSI-GSI, BSI-PSDI, and BST-PST scores. From the MANOVAs, they concluded the following:

a) There was a statistically significant improvement (alpha-value 0.05 and the p-value being below this) for CAT, CBT, and PCT from assessment to termination score
b) Short-Term CBT → greater rate of improvement than PCT or CAT clients
c) With medium therapies, there was a decrease in the CBT group compared to the CAT or PCT group

I do agree with their MANOVA conclusions to an extent. This is because I do not have the raw data, and therefore, cannot run a MANOVA test myself to determine the p-value, but I can say that they are using the proper jargon when the p-value is less than the alpha value. The part I do not agree with it is 'approaching statistical significance' because although the p-value is 0.085 and close to 0.05, this should not be a statement that is made. Other than that error, I think they did a great job explaining the conclusions across services, analyzing measures individually, and explaining between the groups – excellent job discussing all one can get out from the MANOVA!

Q3)

BREAKDOWN: STEVE → shy + withdrawn / meek + tidy

Male Farmers : Male Librarians → 20:1 → Order + structure

What is the probability of **shy given librarian**?

P(A): probability of being shy

P(B): probability of being a librarian

P(B|A): probability of being a librarian given being shy

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

P(A): 0.20 * from kidshealth.org

P(B): 0.05 * given

P(B|A): 0.01 * made up

$$P(A|B) = \frac{0.01 \times 0.20}{0.05} = 0.04 \rightarrow 4\%$$

What is the probability of **shy given farmer**?

P(A): probability of being shy

P(C): probability of being a farmer

P(C|A): probability of being a farmer given being shy

$$P(A|C) = \frac{P(C|A) \times P(A)}{P(C)}$$

P(A): 0.20 * from kidshealth.org

P(C): 0.95 * given

P(C|A): 0.99 * made up

$$P(A|C) = \frac{0.90 \times 0.20}{0.95} = 0.21 \rightarrow 21\%$$

3) Daniel Kahneman's book – 'Thinking Fast and Slow' – is a classic; however, not everything he says in the book is accurate nor reputable. In his example using Steve, Kahneman illustrates to the reader that we are driven by the subconscious mind, and often we fall into a trap of conjunction fallacy. Given the question, most people would think that Steve would be a librarian, but according to Kahneman, he is *more likely* to be a farmer since the ratio of the number of male farmers to male librarians is 20:1, going against people's pre-conceived stereotypes. However, I disagree with Kahneman. I believe that if people's stereotype of shyness is connected to a farmer over librarian, they would have picked farmer in the first place, regardless of the extra information Kahneman provides. Using Baye's Theorem, I oppose Kahneman's strategy, but not end result. The probability of being shy given that one is a librarian is 4%, and the probability of being shy given that one is a farmer is 21%. Due to these results, this characteristic of one being shy is more likely to be associated with a farmer than someone who is a librarian. Due to this, more people are likely to listen to their gut instinct and state that given this person's characteristics, Steve is farmer – without the extra information that Kahneman provides. By sticking to their heuristics, they do not need the 20:1 statistical evidence that Kahneman supplies to make the correct assertion.

Q4)

Florida (Hazard): $0.02 + t$

Tennessee (Hazard): $1.1t$

Formula For Survival $\longrightarrow$ $S(t) = e^{\left(-\int_0^t h(u)\,du\right)}$

For Florida:

$$S(t) = e^{-\int_0^t (0.02+u)\,du} = e^{-(0.50t^2 + 0.02t)}$$

$S(1.5) = 0.3151 =$ 31.51% of survivors past 1.5 yrs.

For Tennessee:

$$S(t) = e^{-\int_0^t 1.1u\,du} = e^{(-0.55t^2)}$$

$S(1.5) = 0.2901 \rightarrow$ 29.01% of survivors past 1.5 yrs.

Out of 100 ducks, 38 survived $\rightarrow$ 38% of surving past 1.5 yrs.

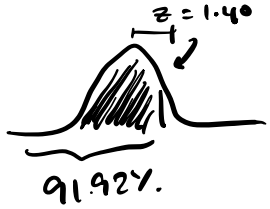
TN: 29.01%

FL: 31.51%

I'm 91.92% sure that the group of ducks were in Florida

Use Z-Test For Proportions: (one sided)

↳ how sure are you that they were in Florida?



measured survival rate: 38%

known survival rate: 31.51%

$z = 1.40 \rightarrow 0.9192 \rightarrow$ 91.92%

$$z = \frac{p - r}{\sqrt{r(1-r)/n}} = \frac{0.38 - 0.3151}{[0.3151(1-0.3151)/100]^{1/2}} = 1.40$$