

Variation in Sleep Behavior across Mammal Species

Introduction:

I applied my knowledge on regression analysis to help understand how variation in sleep behavior across mammal species can be attributed to several species characteristics.

“All known mammal species spend at least part of each day sleeping, some species sleeping more than others. Sleep must serve some biological function, but why do sleep requirements vary so much from species to species? One approach to gathering information that might help understand this question is to study the dependence of hours of sleep on species characteristics”.(Cook and Weisberg, 1999)

About The Data Set:

The data set includes data on 51 different species from at least 13 different orders of mammals. The data set is adapted from the data set described in Allison and Cicchetti (1976) and includes the variables described in below. Each case in the data set represents a species of mammal; values of the variables for each mammal are average of characteristic values for the species.

Species: species of mammal

TS: Total sleep hrs/day

BodyWt: Body weight in kilograms

BrainWt: Brain weight in grams

Life: Maximum life span in years

GP: Gestation time in days

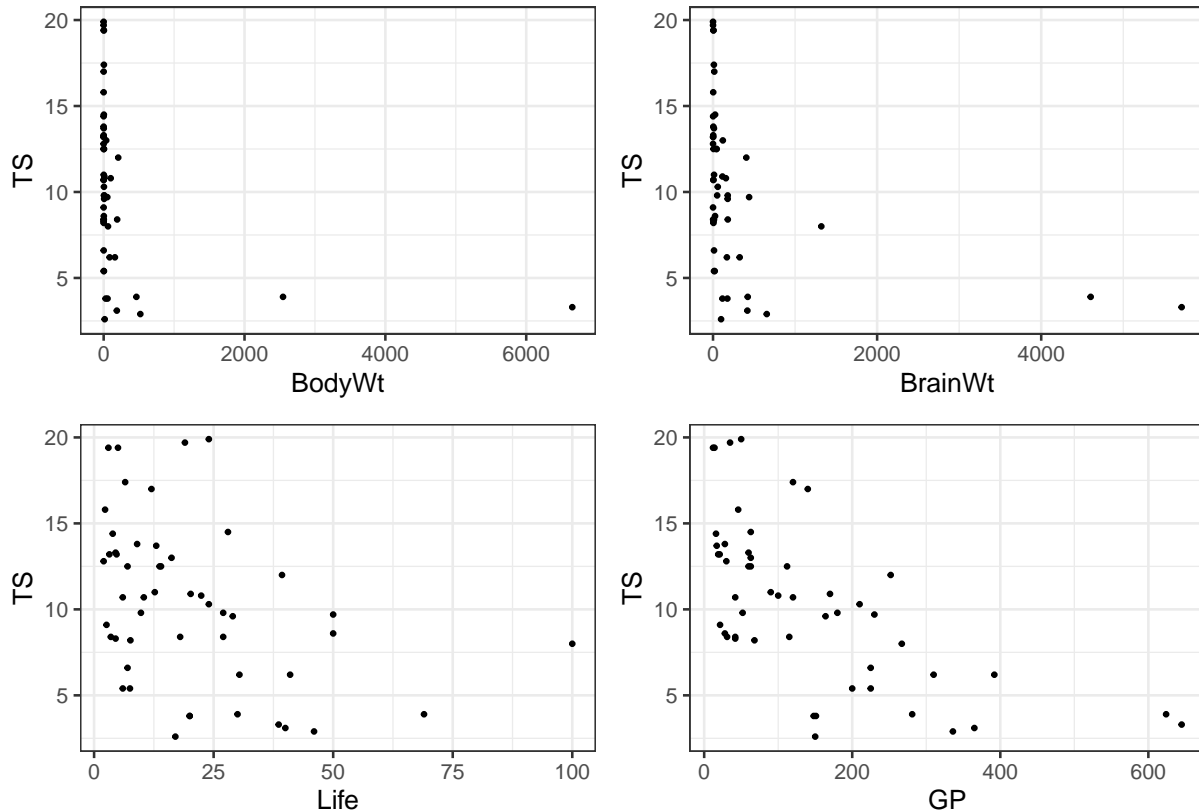
D: Danger index, D1 = relatively low danger from other animals, ..., D5 = highest level of danger from other animals

Part I:

I begin by doing an explanatory analysis of the data. I am interested in seeing how the response variables TS depends on the predictor variables.

```
# read in the dataset
sleepdata <- read.csv("~/Downloads/sleepdata.csv")
Orig_1 = sleepdata %>% ggplot(aes(x = BodyWt, y = TS)) + geom_point(size = 0.5) + theme_bw(10)
Orig_2 = sleepdata %>% ggplot(aes(x = BrainWt, y = TS)) + geom_point(size = 0.5) + theme_bw(10)
Orig_3 = sleepdata %>% ggplot(aes(x = Life, y = TS)) + geom_point(size = 0.5) + theme_bw(10)
Orig_4 = sleepdata %>% ggplot(aes(x = GP, y = TS)) + geom_point(size = 0.5) + theme_bw(10)

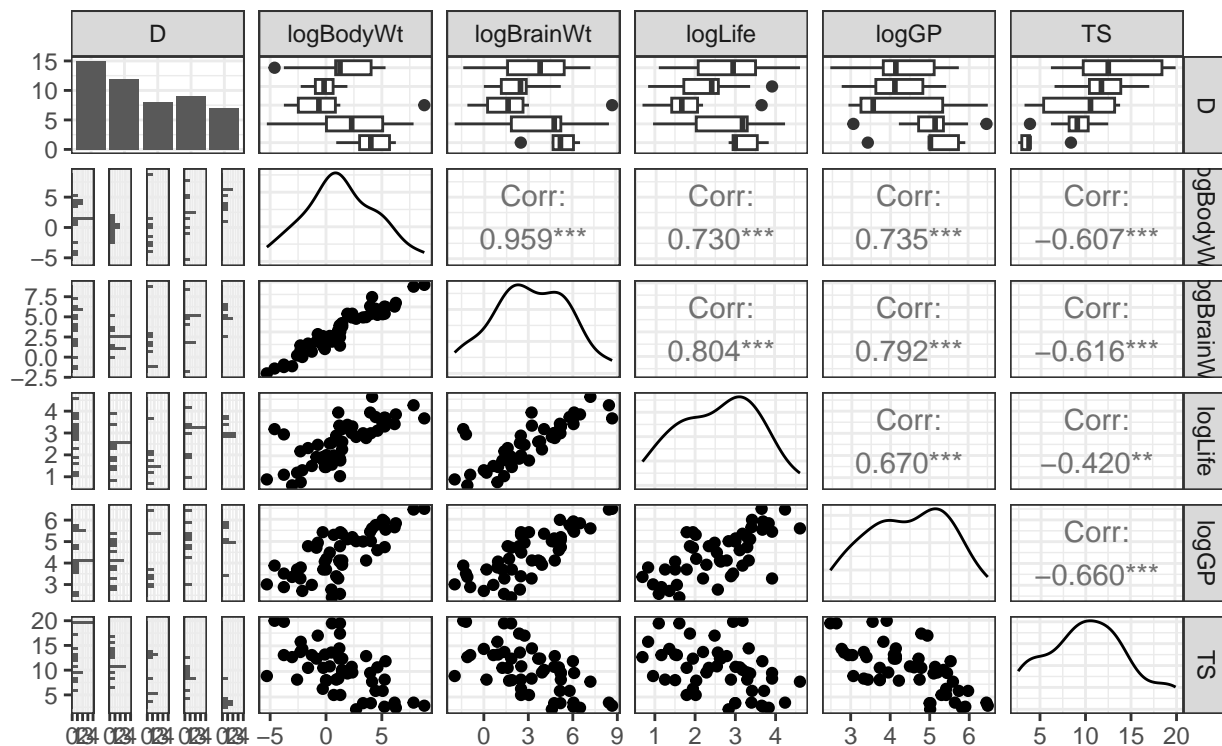
Orig_1 + Orig_2 + Orig_3 + Orig_4
```



Take a look at the plots of the predictor variables BodyWT, BrainWT, Life, and GP. It can be determined that the four predictor variables BodyWT, BrainWT, Life, GP must be transformed into log variables in order to make the plots for these more linear. As we can see with the original graph above the plots for these 4 predictor variables are nonlinear. By looking at the range of the variables it is determined that the predictor variables (BodyWT, BrainWT, Life, and GP) all range over more than one ranges of magnitude. For the response ,TS, variable the range is within one order of magnitude. The log rule clearly states that if the values of a variable range over more than one order of magnitude and the variables is strictly positive then we can replace the variables by its log. This is why we can replace the BodyWT, BrainWT, Life, and GP variables. However, since the response variable TS does not range over one order of magnitude it can not be change to a log variable.

```
sleepNew = sleepdata %>% mutate(logBodyWt = log(BodyWt), logBrainWt = log(BrainWt), logLife = log(Life))
sleepNew = sleepNew %>% dplyr::select(-BodyWt, -BrainWt, -Life, -GP, -species)
```

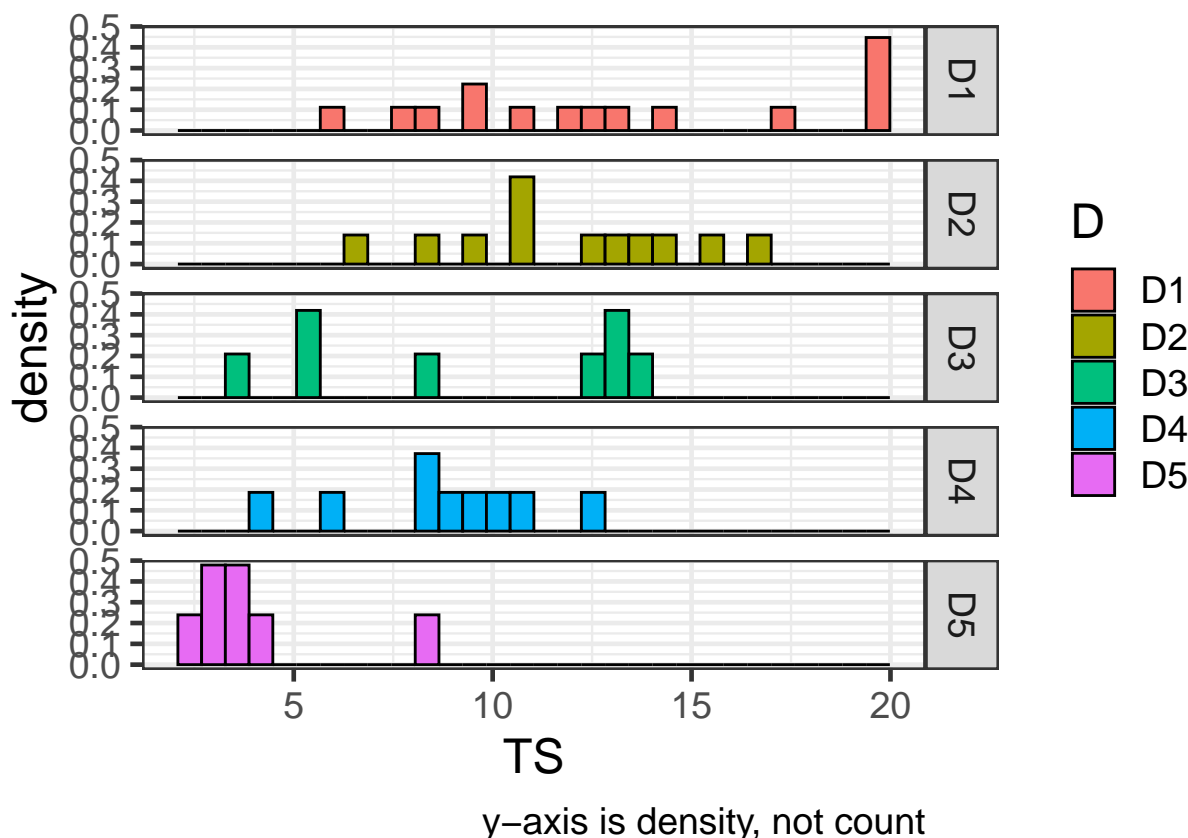
```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



```
sleepdata %>% ggplot(aes(x = TS)) +
  geom_histogram( aes(y = ..density.., fill = D), color="black") +
  facet_grid(D~.) + xlab("TS") + scale_fill_hue("D") +
  theme_bw(18) + labs(caption = "y-axis is density, not count")
```

```
## Warning: The dot-dot notation ('..density..') was deprecated in ggplot2 3.4.0.
## i Please use 'after_stat(density)' instead.
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



I began by transforming the predictor variables `BodyWt`, `BrainWt`, `Life`, and `GP` into log predictor variables. By doing this the relationship between the predictor variables and the response variable `TS` goes from a nonlinear relationship to ones that look to be more linear. .

`D` is a factor categorical variable, it can be ordered `D1`, `D2`, `D3`, `D4`, `D5`. If we break down the response variable (`TS`) for the sleep data grouped by the `D` factor variable. We will see as shown by the bar graph above that there is an average total sleep for animals that have a Danger index of 1, 2, 3, 4, and 5. Through the graph we can see that the average total sleep for animals decreases as the Danger index increases from 1 to 5. This means that animals with a higher danger index sleep on average less. So animals with `D5` danger index on average sleep less than animals with a `D4`. Animals with a `D4` danger index on average sleep less than animals with `D3`. Animals with `D3` danger index on average sleep less than animals with `D2`. Animals with the `D1` danger index on average sleep more than the other danger indexes.

Using the scatter plot matrix we are able to see the relationship between all the variables. There is a negatively associated relationships between the predictor variables ($\log(\text{BodyWt})$, $\log(\text{BrainWt})$, $\log(\text{Life})$, $\log(\text{GP})$, and `D`) and the response variable `TS`. This essentially means as body weight, brain weight, maximum life span, and gestation time increases the total sleep time for the animal decreases. There is a positive correlation between the predictor variables against each another. Essentially this says that as Body weight increases so does brain weight, maximum life span, and gestation time. And as Brain Weight increases so does Body weight, maximum life span, and gestation time. And so on for the other variables. So the predictor variables themselves are positively related to one another. As one predictor variable increases so do the others.

Part II:

I am interested in understanding the adjusted relationships. I consider the full model that uses main effects for $\log(\text{BodyWt})$, $\log(\text{BrainWt})$, $\log(\text{Life})$, $\log(\text{GP})$ and the categorical variable `D` to predict `TS`.

The statistical model is: $TS \sim \log(\text{BodyWT}) + \log(\text{BrainWT}) + \log(\text{Life}) + \log(\text{GP}) + D$

We will use “Wilkinson- Rogers” Notation to refer to the models with main effects for D (Danger index). $\beta_0 + \beta_1 U_{D2} + \beta_2 U_{D3} + \beta_3 U_{D4} + \beta_4 U_{D5}$

$\hat{\beta}_1 U_{D2} + \hat{\beta}_2 U_{D3} + \hat{\beta}_3 U_{D4} + \hat{\beta}_4 U_{D5}$ are all the dummy variable created for the D categorical variable.

so the main effects statistics model for $\log(\text{BodyWt})$, $\log(\text{BrainWt})$, $\log(\text{Life})$, $\log(\text{GP})$ and the categorical variable D to predict TS is: $\hat{E}(TS | GP, \text{BrainWT}, \text{BodyWT}, \text{Life}, D) = \hat{\beta}_0 + \hat{\beta}_1 U_{D2} + \hat{\beta}_2 U_{D3} + \hat{\beta}_3 U_{D4} + \hat{\beta}_4 U_{D5} + \hat{\beta}_5 \log(\text{GP}) + \hat{\beta}_6 \log(\text{BrainWT}) + \hat{\beta}_7 \log(\text{BodyWT}) + \hat{\beta}_8 \log(\text{Life})$

The summary table provided below shows us all the estimations of the regression coefficients and standard errors, it also gives the test statistic and the p value.

```
sleepMod = lm(TS ~ D + log(GP) + log(BrainWt)+ log(BodyWt) + log(Life), data = sleepdata)
summary(sleepMod)
```

```
##
## Call:
## lm(formula = TS ~ D + log(GP) + log(BrainWt) + log(BodyWt) +
##     log(Life), data = sleepdata)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.9584 -1.3524  0.2988  1.5238  6.5396
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   20.5578     2.5166   8.169 3.25e-10 ***
## DD2           -1.9756     1.1214  -1.762  0.08538 .
## DD3           -4.3761     1.3252  -3.302  0.00196 **
## DD4           -3.1764     1.2126  -2.620  0.01220 *
## DD5           -7.1567     1.3220  -5.413 2.76e-06 ***
## log(GP)       -1.4749     0.6467  -2.281  0.02769 *
## log(BrainWt)  -0.7988     0.7116  -1.123  0.26801
## log(BodyWt)    0.1012     0.4728   0.214  0.83155
## log(Life)      0.5756     0.7533   0.764  0.44903
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.737 on 42 degrees of freedom
## Multiple R-squared:  0.7127, Adjusted R-squared:  0.658
## F-statistic: 13.03 on 8 and 42 DF,  p-value: 3.259e-09
```

I want to assess whether the categorical predictor D is useful in this full model. Therefore I perform a hypothesis test.

The hypothesis test we are doing will remove the categorical variable D from the model to see if it is useful. The hypothesis being tested is:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

H_1 : At least one of the $\beta_1, \beta_2, \beta_3, \beta_4$ is not equal to 0.

We set $\beta_1, \beta_2, \beta_3, \beta_4$ all equal to 0 because we are looking at a reduced model with all effective D removed from the model. It wouldn't make sense to only set one dummy variable (for the categorical variable) to 0

because if we only set one equal to 0 that is just removing one of the levels of D from the model not the full effect of D. This is why my hypothesis is set the way it is above.

To do the hypothesis test we will calculate the F- statistic and the P-value of the F- statistic. Formula of F- statistic is: $\frac{(RSS_{NH} - RSS_{AH}) / (df_{NH} - df_{AH})}{(RSS_{AH} / df_{AH})} \geq 0$

The F- test statistic is always greater than or equal to 0. There should be a moderately large value of the F-statistic in order to indicate that there is a difference between the null and alternative hypothesis models.

RSS_{NH} : is the residual sum of squares under the null hypothesis (model without D as a predictor).

RSS_{AH} : is the residual sum of squares under the alternative hypothesis (model including D as a predictor).

df_{NH} : n- p

df_{AH} : n - (p + q) = n - p - q

$df_{NH} - df_{AH}$: The difference in degrees of freedom (also known as the numerator degrees of freedom) is just the number of additional regressor's, q, that are in the alternative model relative to the null hypothesis model.

If the null hypothesis is false we expect the $RSS_{AH} \ll RSS_{NH}$, Residual Sum of Squares under the Alternative Hypothesis will be much less than the Residual Sum of Squares under the Null Hypothesis. When this happens a large F value is seen, this is evidence against the null hypothesis. Next the P-value under the null should be calculated. The P- value is the probability of observing a test statistic larger than what is observed. We use the P- Value we calculate to determine if null hypothesis can be rejected or not.

Under the null model:

```
null.model = lm(TS ~ log(GP) + log(BrainWt) + log(BodyWt) +
  log(Life), data = sleepdata)

RSS.null = sum(resid(null.model)^2)
RSS.null
```

```
## [1] 562.526
```

```
df.null = null.model$df.residual
df.null
```

```
## [1] 46
```

Under the alternative model:

```
alt.model = lm(TS ~ D * log(GP) + log(BrainWt) + log(BodyWt) +
  log(Life), data = sleepdata)

RSS.alt = sum(resid(alt.model)^2)
RSS.alt
```

```
## [1] 275.8331
```

```
df.alt = alt.model$df.residual
df.alt
```

```
## [1] 38
```

F statistic:

```
Fstat = ((RSS.null-RSS.alt) / (df.null-df.alt)) / (RSS.alt/df.alt)
Fstat
```

```
## [1] 4.937011
```

P- Value:

```
pf(Fstat, df1 = df.null - df.alt, df2 = df.alt, lower.tail = F)
```

```
## [1] 0.000310935
```

In conclusion our null hypothesis was $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ and we tested that against our alternative hypothesis, H_1 : At least one of the $\beta_1, \beta_2, \beta_3, \beta_4$ is not equal to 0. We fit the two models and got our F-statistic = 4.937011. The corresponding P- value under the null hypothesis was, P- value = 0.000310935. So, we conclude that the data is not consistent with the null hypothesis and the categorical predictor D is useful.

3. Leaving D in the model, describe in sentence/paragraph form the relationship between the predictors $\log(\text{BodyWt})$ and $\log(\text{BrainWt})$ and the response TS in the context of the MLR model with all predictors. As part of your written answer, be sure to address the following items/questions:

- Use the fitted model to provide confidence intervals for the individual coefficients for $\log(\text{BodyWt})$ and $\log(\text{BrainWt})$ and explain what they mean in terms of the context of the problem.
- Contrast your interpretation of the relationships between $\log(\text{BodyWt})$ and TS (and also between $\log(\text{BrainWt})$ and TS) under the multiple linear regression model with your description of the pairwise relationships from Part I. Describe how the interpretations differ.

We are going to look at the $\log(\text{BodyWt})$ and fix the values of all the other predictor variables. In other words $\hat{E}(TS \mid \text{BodyWt} + 1, \text{BrainWt}^*, \text{GP}^*, \text{Life}^*, D^*) - \hat{E}(TS \mid \text{BodyWt}, \text{BrainWt}^*, \text{GP}^*, \text{Life}^*, D^*)$ (* just means we are fixing the value of these variables at something). When we simplify this expression down we will see that it simply equals $\hat{\beta}_1$. So, for 2 animals with the same number of estimated GP, BrainWt, Life, and D an increase of 1 kg in $\log(\text{BodyWt})$ is associated with $\hat{\beta}_1$ 0.1012 additional hours of total sleep per day.

For 2 animals with the same number of estimated GP, BodyWt, Life, and D an increase in 1 gram of BrainWt is associated with -0.7988 less hours of total sleep per day.

The 95% confidence interval for $\log(\text{BrainWt})$ is (-2.2347959, 0.6372426). Essentially this means that for the animals, a 1 gram difference in brain weight is associated with a difference in average total sleep of between -2.2347959 and 0.6372426 hours (95% level of confidence). The 95% confidence interval for $\log(\text{BodyWt})$ is (-0.8528798 1.0552550). Essential this means that for the animals, a 1 kilogram difference in body weight is associated with a difference in average total sleep between -0.8528798 and 1.0552550 hours (95% level of confidence).

In the pairwise model from part 1 when looking at the graph I assumed that there was a negative linear relationship between BodyWt and TS as well as BrainWt and TS. Basically saying that as the brain weight increases total sleep decrease and as body weight increases total sleep decreases. Taking a look at our linear model we see that we come up with a different interpretation. In fact based on the linear model we see with an increase in body weight there is an increase in total sleep per day. However, with brain weight we do see that an increase in brain weight results in a decrease in total sleep per day. Which is what we saw with our pairwise model.

95% Confidence intervals of individual coefficients:

```
confint(sleepMod, level = 0.95)
```

```
##              2.5 %      97.5 %
## (Intercept) 15.4791231 25.6364043
## DD2         -4.2386904  0.2874049
## DD3         -7.0505039 -1.7017911
## DD4         -5.6234979 -0.7293568
## DD5         -9.8246765 -4.4887901
## log(GP)      -2.7799142 -0.1699266
## log(BrainWt) -2.2347959  0.6372426
## log(BodyWt)  -0.8528798  1.0552550
## log(Life)    -0.9445323  2.0958053
```

Part III:

In this step I am looking for the best model to implement. Considering main effects only for `log(BodyWt)`, `log(BrainWt)`, `log(Life)`, `log(GP)` and the categorical variable `D`, use the `stepAIC` function in the R package `MASS` to implement forward selection and backward elimination to find a subset of the predictors that provides a good balance between model fit and model complexity.

The idea behind forward selection is that we start with an empty model (null model: has no predictors only the intercept) and then incrementally make it into a bigger model. You only incrementally increase as long as making the model bigger does not make it a worse model. In backwards elimination the idea is the same, however, you are now instead starting with a full model and removing predictors one at a time until making the model any smaller makes the model worse instead of better.

Forward Selection and Backward Elimination:

```
null = lm(TS ~ 1, data = sleepNew)
full = lm(TS ~ ., data = sleepNew)
n = dim(sleepNew)[1]

stepAIC(null, scope = list(upper = full), direction = "forward", k = 2)
```

```
## Start:  AIC=158.39
## TS ~ 1
##
##              Df Sum of Sq    RSS    AIC
## + logGP       1    476.84   618.03 131.23
## + logBrainWt   1    416.10   678.77 136.01
## + logBodyWt    1    402.97   691.90 136.99
## + D            4    474.16   620.71 137.45
## + logLife      1    193.38   901.49 150.48
## <none>                1094.87 158.40
##
## Step:  AIC=131.23
## TS ~ logGP
##
##              Df Sum of Sq    RSS    AIC
## + D            4    268.744 349.28 110.13
## + logBodyWt    1     35.151 582.88 130.24
## + logBrainWt   1     25.721 592.30 131.06
```



```

## <none>                618.03 131.23
## + logLife      1      0.941 617.09 133.15
##
## Step:  AIC=110.13
## TS ~ logGP + D
##
##           Df Sum of Sq    RSS    AIC
## + logBrainWt  1    30.3864 318.90 107.48
## + logBodyWt   1    24.8139 324.47 108.37
## <none>                349.28 110.13
## + logLife      1     2.0139 347.27 111.83
##
## Step:  AIC=107.48
## TS ~ logGP + D + logBrainWt
##
##           Df Sum of Sq    RSS    AIC
## <none>                318.90 107.48
## + logLife      1     4.0315 314.86 108.84
## + logBodyWt    1     0.0015 318.89 109.48

##
## Call:
## lm(formula = TS ~ logGP + D + logBrainWt, data = sleepNew)
##
## Coefficients:
## (Intercept)      logGP          DD2          DD3          DD4          DD5
##      21.2179      -1.4219      -2.1651      -4.7433      -3.2938      -7.1699
##  logBrainWt
##      -0.5319

```

Backwards Elimination:

```
stepAIC(full, direction = "backward", k = 2)
```

```

## Start:  AIC=110.78
## TS ~ D + logBodyWt + logBrainWt + logLife + logGP
##
##           Df Sum of Sq    RSS    AIC
## - logBodyWt  1     0.343 314.86 108.84
## - logLife    1     4.373 318.89 109.48
## - logBrainWt 1     9.436 323.96 110.29
## <none>                314.52 110.78
## - logGP      1    38.958 353.48 114.74
## - D          4   248.005 562.53 132.43
##
## Step:  AIC=108.84
## TS ~ D + logBrainWt + logLife + logGP
##
##           Df Sum of Sq    RSS    AIC
## - logLife    1     4.032 318.90 107.48
## <none>                314.86 108.84
## - logBrainWt 1    32.404 347.27 111.83
## - logGP      1    40.503 355.37 113.01

```

```
## - D          4    251.808 566.67 130.81
##
## Step:  AIC=107.48
## TS ~ D + logBrainWt + logGP
##
##           Df Sum of Sq    RSS    AIC
## <none>                318.90 107.48
## - logBrainWt  1      30.386 349.28 110.13
## - logGP       1      37.615 356.51 111.17
## - D           4     273.409 592.30 131.06

##
## Call:
## lm(formula = TS ~ D + logBrainWt + logGP, data = sleepNew)
##
## Coefficients:
## (Intercept)          DD2          DD3          DD4          DD5    logBrainWt
##      21.2179      -2.1651      -4.7433      -3.2938      -7.1699      -0.5319
##      logGP
##      -1.4219
```

I chose to do the AIC method for this problem. I also took a look at the BIC method, however, the results for both AIC and BIC are the same. I did both a backward and Forward AIC. If you look at the results above you can see the best models resulted from the backward elimination and forward selection. For Backwards elimination the model is $TS \sim D + \log BrainWt + \log GP$. For forward elimination the best model given to us was $TS \sim \log GP + D + \log BrainWt$. For both the backwards and forwards we see that the same predictor variables were used in the best model, so, the models are the same. Therefore, the model that provides a good balance between model fit and model complexity is the model that contains the 3 variables logGP, D, and logBrainWt. The mean function we get is $\hat{E}(TS \mid GP, BrainWT, D) = \hat{\beta}_0 + \hat{\beta}_1 U_{D2} + \hat{\beta}_2 U_{D3} + \hat{\beta}_3 U_{D4} + \hat{\beta}_4 U_{D5} + \hat{\beta}_5 \log(GP) + \hat{\beta}_6 \log(BrainWT)$.

Part IV:

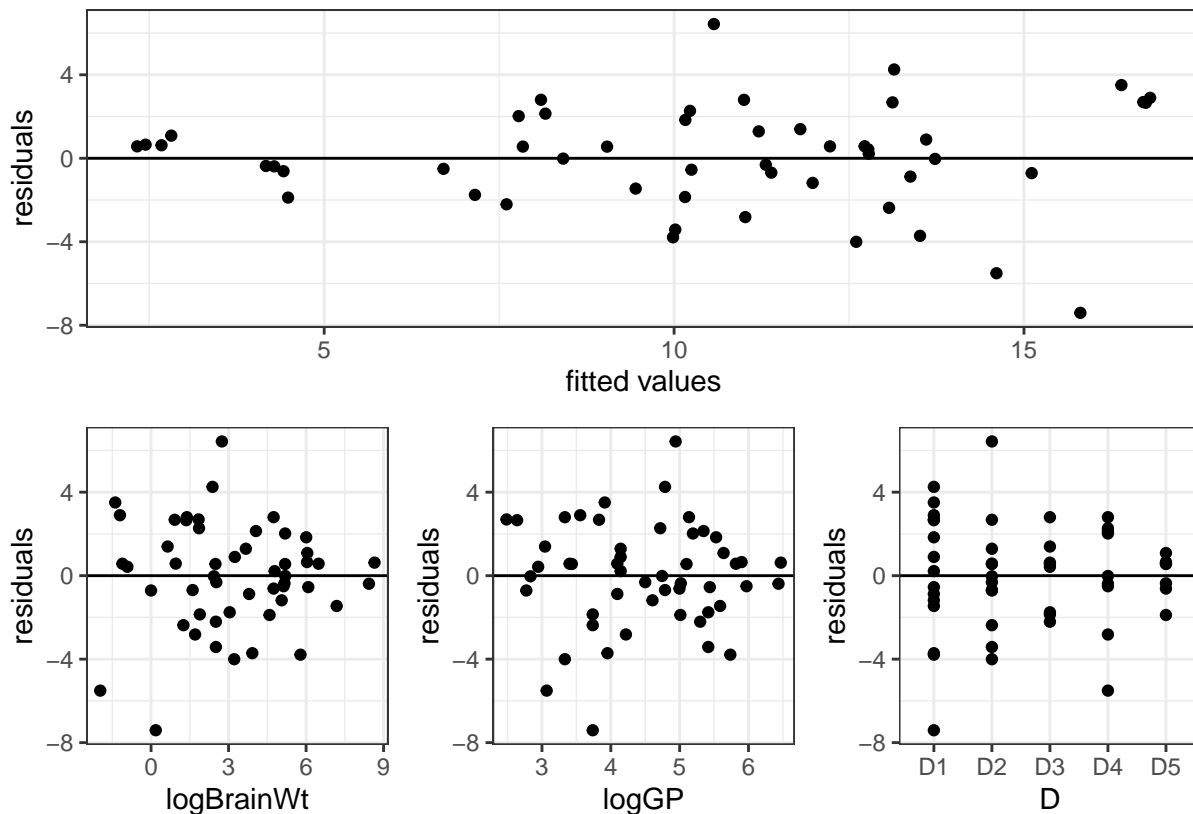
Consider the the model that uses main effects for $\log(BrainWt)$, $\log(GP)$ and the categorical variable D to predict sleep time. I will provide the appropriate plots of residuals to check the fit of the model. Use the plots to check whether the model assumptions seem reasonable, or whether I diagnose any lack of fit.

```
sleep.lm = lm(TS ~ D + logGP + logBrainWt, data = sleepNew)
summary(sleep.lm)
```

```
##
## Call:
## lm(formula = TS ~ D + logGP + logBrainWt, data = sleepNew)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.4064 -1.3164  0.2171  1.6172  6.4313
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   21.2179     2.2017   9.637 2.06e-12 ***
## DD2           -2.1651     1.0608  -2.041 0.047277 *
##
```

```
## DD3          -4.7433      1.2132   -3.910 0.000316 ***
## DD4          -3.2938      1.1832   -2.784 0.007889 **
## DD5          -7.1699      1.2766   -5.616 1.23e-06 ***
## logGP        -1.4219      0.6241   -2.278 0.027628 *
## logBrainWt   -0.5319      0.2597   -2.048 0.046599 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.692 on 44 degrees of freedom
## Multiple R-squared:  0.7087, Adjusted R-squared:  0.669
## F-statistic: 17.84 on 6 and 44 DF,  p-value: 2.353e-10
```

```
base = augment(sleep.lm) %>% ggplot(aes(y = .resid)) + geom_point() + theme_bw() + ylab("residuals") +
(base + aes(x = .fitted)) + xlab("fitted values") / ((base + aes(x = logBrainWt)) + (base + aes(x = logGP
```



I graphed the plots of the residuals \hat{e}_i versus the fitted values \hat{y}_i and the predictors x_{ij} . We are looking for plots of the residuals that have no discerning trend and there is no apparent change of variance. These are known as null plots. A trend or pattern in the residual plots indicate a lack of linearity. Looking at the residual plot for logBrainWt and logGP there doesn't seem to be any obvious trend remaining in the data. The residuals seem to be pretty equally spread around 0. Also when we look at the residual plots of logBrainWt and logGP the spread of the residuals is roughly equal so we can say that the constant variance assumption is met for these plots. These residuals are reasonable looking in terms of detecting evidence of any trend we may have missed.

For the residual plot of the D predictor variable we have points in a vertical line for each of the levels. The reason that this plot looks the way it does is because we are plotting the residual variables against a

categorical predictor variable. The points on the plot, however, are still pretty equally spread around 0. However, the spread of the residuals does not seem to be roughly equal at each level of D. There looks to be less residuals at the D5 level as compared to D1. So, this means that the constant variance assumption may not be met. If the constant variance assumption is not met that means the residuals for the D plot suffer from heteroscedasticity. This essentially means that there may be some trend for the D residuals we may have missed. If we look at the residual plot of fitted values we see the heteroscedasticity a lot more obviously. The residuals are less concentrated at the lower fitted values and more concentrated at the higher fitted values. So, though the values are centered around 0 we see evidence that the constant variance assumption was not met. There are ways to fix the violation of the constant variance, most commonly you would transform the response variable.

Part V:

Summary of the model.

In the model $TS \sim \log(GP) + \log(BrainWt) + D$ the estimated regression coefficients are $\hat{\beta}_0 = 21.2179$, $\hat{\beta}_1 = -2.1651$, $\hat{\beta}_2 = -4.7433$, $\hat{\beta}_3 = -3.2938$, $\hat{\beta}_4 = -7.1699$, $\hat{\beta}_5 = -1.4219$, $\hat{\beta}_6 = -0.5319$. Essentially this means for a given animal each additional g of BrainWt is associated with total sleep per day that is 41% lower. I got the 41% by doing $(e^{-0.5319} - 1) * 100$. This also means that for a given animal each additional day of gestation time is associated with total sleep per day that is $(e^{-1.4219} - 1) * 100 = 75\%$ lower.

Effects plot for BrainWt:

```
sleepMod = lm(TS ~ D + log(GP) + log(BrainWt), data = sleepdata)

GPFix = 320

# Grid of points for GP over the observed range
BrainWtgrid = seq(from = 0.005, to = 6654, length.out = 100)
# Fitted values when GP = GPFix and D = 1
yhatD1 = predict(sleepMod, newdata = data.frame(D= "D1",
                                                GP = GPFix,
                                                BrainWt = BrainWtgrid))

# Fitted values when GP = GPFix and D = 2
yhatD2 = predict(sleepMod, newdata = data.frame(D= "D2",
                                                GP = GPFix,
                                                BrainWt = BrainWtgrid))

# Fitted values when GP = GPFix and D = 3
yhatD3 = predict(sleepMod, newdata = data.frame(D= "D3",
                                                GP = GPFix,
                                                BrainWt = BrainWtgrid))

# Fitted values when GP = GPFix and D = 4
yhatD4 = predict(sleepMod, newdata = data.frame(D= "D4",
                                                GP = GPFix,
                                                BrainWt = BrainWtgrid))

# Fitted values when GP = GPFix and D = 5
yhatD5 = predict(sleepMod, newdata = data.frame(D= "D5",
                                                GP = GPFix,
                                                BrainWt = BrainWtgrid))

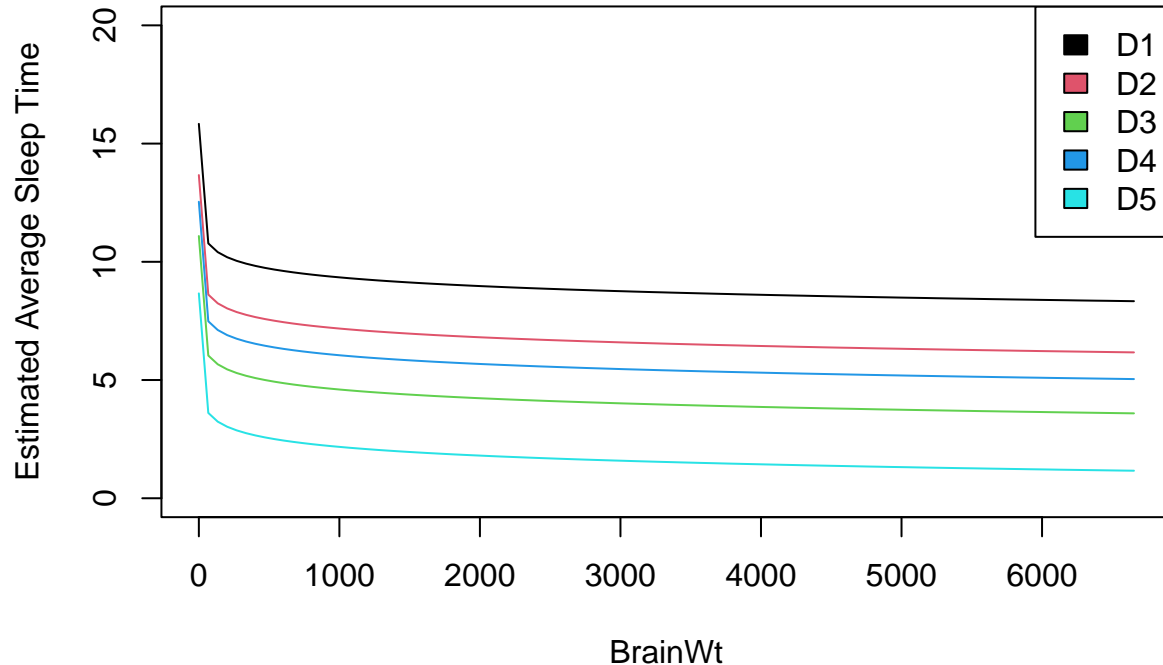
plot(BrainWtgrid, yhatD1, type = "l", xlab = "BrainWt", ylab = "Estimated Average Sleep Time",
     ylim = c(0,20))
lines(BrainWtgrid, yhatD2, col = 2)
lines(BrainWtgrid, yhatD3, col = 3)
```

```

lines(BrainWtgrid, yhatD4, col = 4)
lines(BrainWtgrid, yhatD5, col = 5)

legend(x = "topright", legend = c("D1", "D2", "D3", "D4", "D5"), fill = c(1, 2, 3, 4, 5))

```



The Effects plot for the BrainWt predictor variable has 5 curves each of which corresponds to the different fixed D level. The plot shows that as the predictor variable Brain weight increases there is a decrease in the response, estimated average sleep time. The decrease depends on the level of D. All of the curves are decreasing at the same rate but where the curve starts differs depending on the level of D. If we increase the value of the fixed GP the starting point of all the curves reduce, however, they still continue to decrease at the same rate. If we decrease the value of the fixed GP the starting points of all the curves increase, they still continue to decrease at the same rate. So, we can say that as Brain Weight increases at a fixed value of gestation period the curve at each level of D decreases. The starting point for each curve of different D values depends on what value we fix the gestation period at. The higher the gestation period the lower the start value of the curves.

Effects Plot for GP:

```

BrainWtFix = 317.15

# Grid of points for GP over the observed range
GPgrid = seq(from = 12, to = 645, length.out = 100)
# Fitted values when BrainWt = BrainWtFix and D = 1
yhatD1 = predict(sleepMod, newdata = data.frame(D= "D1",
                                                  BrainWt = BrainWtFix,
                                                  GP = GPgrid))

```

```

# Fitted values when BrainWt = BrainWtFix and D = 2
yhatD2 = predict(sleepMod, newdata = data.frame(D= "D2",
                                                BrainWt = BrainWtFix,
                                                GP = GPgrid))

# Fitted values when BrainWt = BrainWtFix and D = 3
yhatD3 = predict(sleepMod, newdata = data.frame(D= "D3",
                                                BrainWt = BrainWtFix,
                                                GP = GPgrid))

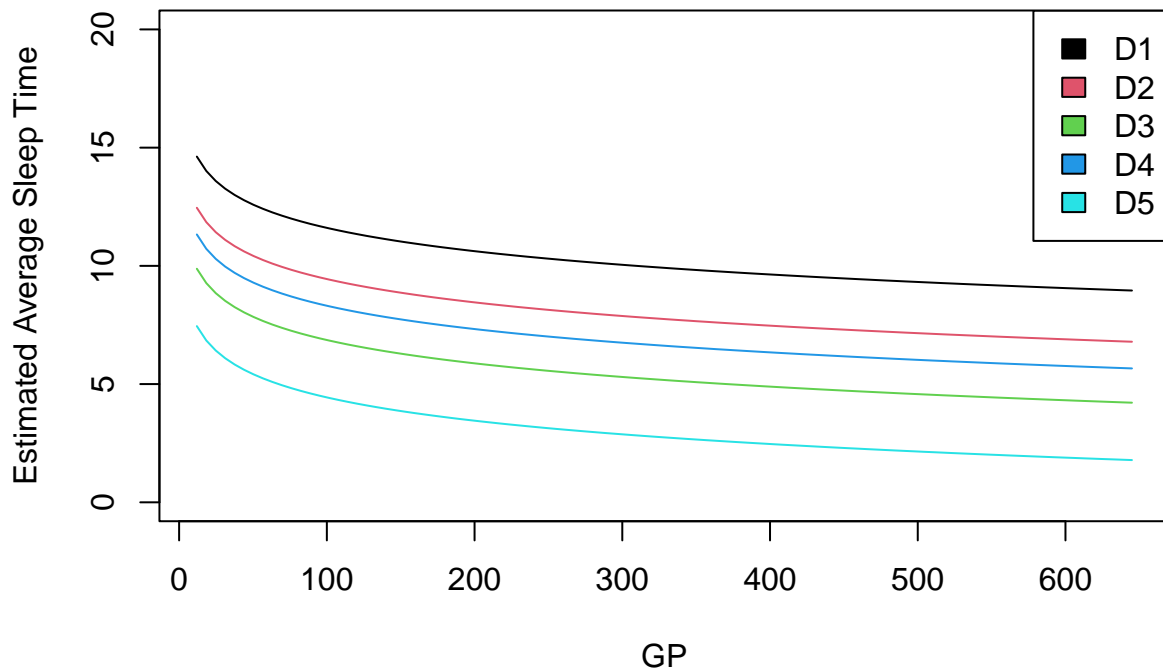
# Fitted values when BrainWt = BrainWtFix and D = 4
yhatD4 = predict(sleepMod, newdata = data.frame(D= "D4",
                                                BrainWt = BrainWtFix,
                                                GP = GPgrid))

# Fitted values when BrainWt = BrainWtFix and D = 5
yhatD5 = predict(sleepMod, newdata = data.frame(D= "D5",
                                                BrainWt = BrainWtFix,
                                                GP = GPgrid))

plot(GPgrid, yhatD1, type = "l", xlab = "GP", ylab = "Estimated Average Sleep Time",
     ylim = c(0,20))
lines(GPgrid, yhatD2, col = 2)
lines(GPgrid, yhatD3, col = 3)
lines(GPgrid, yhatD4, col = 4)
lines(GPgrid, yhatD5, col = 5)

legend(x = "topright", legend = c("D1", "D2", "D3", "D4", "D5"), fill = c(1, 2, 3, 4, 5))

```



The Effects plot for the GP predictor variable has 5 curves each of which corresponds to the different fixed

D level. The plot shows that as GP increases there is a decrease in the response estimated average sleep time. The decrease depends on the level D. All of the curves are decreasing at the same rate but where each curve starts differs depending on the level of D. If we increase the value of the fixed BrainWt the starting point of all the curves reduces, however, they still continue to decrease at the same rate. If we decrease the value of the fixed BrainWt the starting points of all the curves increase, they still continue to decrease at the same rate. So, we can say that as gestation period increases at a fixed value of brain weight the curve at each level of D decreases. The starting point for each curve of different D levels depends on what value we fix the brain weight at. The higher the brain weight the lower the start value of the curves.

References

- Allison, T. and Cicchetti, D. (1976). Sleep in mammals: Ecological and constitutional correlates. *Science*, 194, 732–734.
- Cook, R. D. and Weisberg, S. (1999). *Applied regression including computing and graphics*. John Wiley & Sons.