Practicum 1 Analysis

Peter Claussen

Ben Derenge

Stephanie Liebl

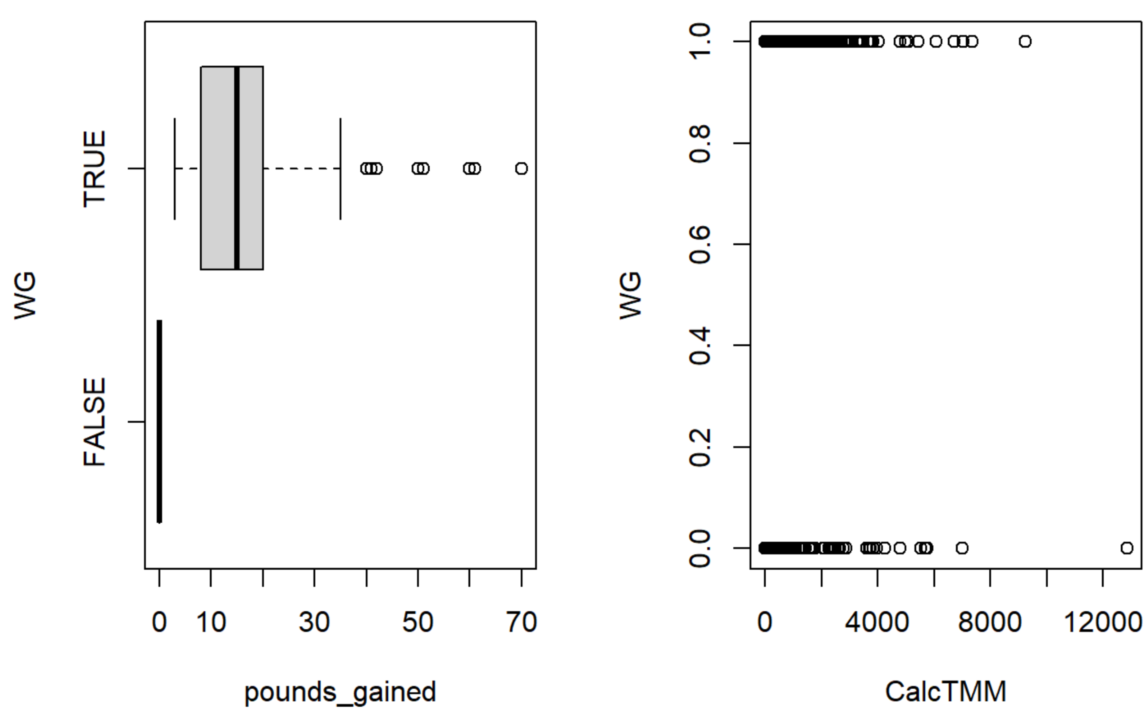
10/5/2021

Data Processing

Our initial step in processing the 'Practicum 1 Data' data set was to remove any rows where the `Snumber` column was "NA". We identified the `Snumber` column as a unique identifier, and thus were only concerned with nonempty rows. Initially, the data set had 392 rows. After removing removing the 'NA' identifiers, we were left with 352 rows.

Next, we verified that the column `Total_Met_Min` had correct computations. We created a separate column `CalcTMM` to check against the values in `Total_Met_min` . In our observations of the `shift` variable, we discovered that there existed an 'other' category in addition to the missing values category. We decided to refrain from combining these categories and created a 'missing' category for values of `shift` that were blank.

We will consider two subsets for analysis. First, we create a data table that has appropriate values for `weightgain` . This will be the larger of the two data sets. Our original data set contained a binary 'yes/no' column titled `weightgain` . This column was missing data for four rows, and thus, we subset the data to obtain a data table with values for `weightgain` in all rows. This data table left us with 348 rows. It should be noted that the original data set contains a numeric `pounds_gained` column as well. This column had many missing values, but we were able to impute a '0' if we knew that row had a 'No' for `weightgain` . Lastly, we observed a single missing value for the column `CalcTMM` , and so we once more subset the data to obtain a table with values of `CalcTMM` for all rows.



Analysis of Binary Response (WG)

(SA1) Does *total metabolic minutes* have an effect on *weight gain*?

The client provided us two specific aims to address in our analysis. We will begin with the first specific aim, which we will refer to as SA1. SA1 says, “Does total metabolic minutes have an effect on weight gain?”. To address this question, we began by creating a simple linear regression model of our binary weight gain versus calculated total metabolic minutes. The summary output of this model suggests that `CalcTMM` has very little effect on weight gain, as the p-value for `CalcTMM` is 0.21, quite large.

```
SA1.model1 <- glm(WG ~ CalcTMM, data=gained.dat,family = binomial)
tab_model(SA1.model1)
```

	WG		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	2.43	1.81 – 3.29	**<0.001**
CalcTMM	1.00	1.00 – 1.00	0.210
Observations	347		
R^2 Tjur	0.005		

(SA2) Does *shift* have an effect on *weight gain*?

The second specific aim given to us by the client says, “Does shift have an effect on weight gain?”. To address this question, we began by creating a simple linear regression model of our binary weight gain versus shift. From the summary output of this model, we see that no value of shift seems to have much significance in the model. The p-value for each of these values is greater than 0.1, suggesting each has little effect, if any, on weight gain.

```
SA2.model1 <- glm(WG ~ shift, data=gained.dat,family = binomial)
tab_model(SA2.model1)
```

	WG		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	1.73	0.84 – 3.75	0.149
shift [8am]	1.29	0.54 – 2.97	0.553
shift [9am]	0.96	0.38 – 2.40	0.939
shift [10am]	1.79	0.66 – 4.84	0.250
shift [11am]	0.98	0.37 – 2.56	0.962
shift [12pm]	1.04	0.28 – 4.12	0.951
shift [1pm]	0.96	0.20 – 5.44	0.965
shift [2pm]	2.32	0.58 – 11.81	0.262
shift [other]	2.17	0.60 – 9.10	0.253
Observations	347		
R^2 Tjur	0.015		

Model 2 Interactions

After observing the models representing weight gain versus total met minutes and shift, respectively, we sought to examine whether `shift` * `CalcTMM` had an effect on weight gain. Using a generalized linear model, we found that the interaction between shift and calculated total metabolic minutes did not have an effect on weight gain.

```
SA12.model2 <- glm(WG ~ shift*CalcTMM, data=gained.dat,family = binomial)
tab_model(SA12.model2)
```

	WG		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	2.89	0.96 – 10.01	0.071
shift [8am]	1.09	0.28 – 3.73	0.893
shift [9am]	0.51	0.13 – 1.85	0.318
shift [10am]	2.00	0.40 – 10.27	0.395
shift [11am]	0.54	0.12 – 2.13	0.387
shift [12pm]	0.80	0.11 – 6.24	0.821
shift [1pm]	1.08	0.08 – 28.50	0.956
shift [2pm]	1.28	0.15 – 13.07	0.825
shift [other]	0.37	0.04 – 2.94	0.347
CalcTMM	1.00	1.00 – 1.00	0.230
shift [8am] * CalcTMM	1.00	1.00 – 1.00	0.754
shift [9am] * CalcTMM	1.00	1.00 – 1.00	0.178
shift [10am] * CalcTMM	1.00	1.00 – 1.00	0.881
shift [11am] * CalcTMM	1.00	1.00 – 1.00	0.269
shift [12pm] * CalcTMM	1.00	1.00 – 1.00	0.767
shift [1pm] * CalcTMM	1.00	1.00 – 1.00	0.857
shift [2pm] * CalcTMM	1.00	1.00 – 1.00	0.421
shift [other] * CalcTMM	1.00	1.00 – 1.00	0.118
Observations	347		
R^2 Tjur	0.049		

Model 3 SA1 and 2 plus anthropometric variables

The original data set included anthropometric variables such as `gender` , `Age` , `height` , and `BMI` . We wanted to inspect whether such variables effect weight gain, and so we created a generalized linear model including them. We began by including just variables for gender and age, and then height. The p-values for these variables in both models are large, suggesting little effect on weight gain.

```
subset3a.dat <- gained.dat[which(complete.cases(gained.dat[, c("gender", "Age", "shift", "CalcTMM")])),]
subset3b.dat <- gained.dat[which(complete.cases(gained.dat[, c("gender", "Age", "height","shift", "CalcTMM"
)])),]

SA12.model3a <- glm(WG ~ gender + Age + shift + CalcTMM, data=subset3a.dat,family = binomial)
SA12.model3b <- glm(WG ~ gender + Age + height +shift + CalcTMM, data=subset3b.dat,family = binomial)

tab_model(SA12.model3a, SA12.model3b)
```

		WG			**WG**	
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	1.11	0.03 – 37.32	0.951	0.81	0.00 – 651.97	0.950
gender [Female]	3.75	0.14 – 104.44	0.370	3.88	0.14 – 109.26	0.364
gender [Male]	2.24	0.08 – 63.28	0.589	2.20	0.08 – 61.64	0.597
Age	0.99	0.97 – 1.02	0.628	1.00	0.97 – 1.02	0.848
shift [8am]	0.86	0.32 – 2.16	0.761	0.90	0.33 – 2.26	0.822
shift [9am]	0.68	0.24 – 1.87	0.467	0.71	0.24 – 1.95	0.509
shift [10am]	1.19	0.39 – 3.51	0.749	1.27	0.41 – 3.80	0.672
shift [11am]	0.82	0.27 – 2.42	0.728	0.84	0.27 – 2.46	0.747
shift [12pm]	0.66	0.16 – 2.81	0.565	0.66	0.16 – 2.82	0.568
shift [1pm]	0.64	0.12 – 3.78	0.599	0.65	0.12 – 3.84	0.610
shift [2pm]	3.81	0.54 – 77.44	0.245	3.84	0.55 – 78.35	0.243
shift [other]	2.04	0.46 – 11.59	0.372	1.93	0.43 – 11.04	0.416
CalcTMM	1.00	1.00 – 1.00	0.142	1.00	1.00 – 1.00	0.162
height				1.00	0.92 – 1.09	0.953
Observations	319			313		
R^2 Tjur	0.041			0.041		

Model 4 Partition CalcTMM into components

In the background information given to us by our client, it stated that the `Total_Met_min` column was calculated by a combination of `Vig.ex.Time` , `Mod.ex.Time` , and `Walk.ex.Time` . As noted above, total met minutes as a whole does not have a large effect on weight gain, we wanted to determine whether its individual components do.

```
subset4a.dat <- gained.dat[which(complete.cases(gained.dat[, c("gender", "Age", "shift", "Vig.ex.Time", "Mo
d.ex.time", "Walk.ex.Time")])),]
subset4b.dat <- gained.dat[which(complete.cases(gained.dat[, c("gender", "Age", "height", "shift", "Vig.ex.
Time", "Mod.ex.time", "Walk.ex.Time")])),]

SA12.model4a <- glm(WG ~ gender + Age + shift + Vig.ex.Time  + Mod.ex.time + Walk.ex.Time, data=subset4a.da
t, family = binomial)
SA12.model4b <- glm(WG ~ gender + Age + height + shift + Vig.ex.Time  + Mod.ex.time + Walk.ex.Time, data=su
bset4b.dat, family = binomial)

tab_model(SA12.model4a, SA12.model4b)
```

		WG			**WG**	
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	1.33	0.04 – 47.37	0.865	0.95	0.00 – 811.19	0.987
gender [Female]	3.11	0.11 – 91.66	0.456	3.26	0.11 – 98.21	0.443
gender [Male]	1.87	0.06 – 55.68	0.683	1.85	0.06 – 55.00	0.687
Age	0.99	0.97 – 1.02	0.609	1.00	0.97 – 1.02	0.836
shift [8am]	0.88	0.33 – 2.23	0.800	0.92	0.34 – 2.33	0.867
shift [9am]	0.68	0.23 – 1.87	0.462	0.71	0.24 – 1.97	0.515
shift [10am]	1.23	0.40 – 3.63	0.712	1.31	0.42 – 3.94	0.636
shift [11am]	0.85	0.28 – 2.52	0.776	0.87	0.28 – 2.57	0.796
shift [12pm]	0.66	0.16 – 2.81	0.567	0.66	0.16 – 2.81	0.566
shift [1pm]	0.64	0.12 – 3.80	0.602	0.65	0.12 – 3.86	0.615
shift [2pm]	3.82	0.53 – 78.50	0.248	3.77	0.52 – 77.63	0.254
shift [other]	2.07	0.46 – 11.78	0.365	1.95	0.43 – 11.22	0.408
Vig.ex.Time	1.00	1.00 – 1.00	0.241	1.00	1.00 – 1.00	0.279
Mod.ex.time	1.00	1.00 – 1.00	0.956	1.00	1.00 – 1.00	0.939
Walk.ex.Time	1.00	1.00 – 1.00	0.664	1.00	1.00 – 1.00	0.559
height				1.00	0.92 – 1.09	0.953
Observations	319			313		
R^2 Tjur	0.041			0.042		

Model 5 - Model 4 plus BMI and initial body weight

Next, we will observe if BMI effects weight gain or if the components of BMI effect weight gain. We will do this by adding to Model 4. Using the columns for body weight, height, and pounds gained, we were able to obtain `initial_BMI` and `initial_bweight` columns.

```
gained.dat['initial_bweight'] <- gained.dat$bweight - gained.dat$pounds_gained
gained.dat['initial_BMI'] <- (gained.dat$initial_bweight / (gained.dat$height)^2)*703
```

In the following models, we used the anthropometric variables, the individual total met minute components in addition to the initial BMI and initial body weight variables, respectively. We know that BMI is calculated using height and weight, so we avoided using all three (`initial_BMI` , `height` , `initial_bweight`) in a model to avoid confounding. We see that in the first model, `initial_BMI` has a small p-value and in the second model, `initial_bweight` has a small p-value. This suggests that the variables effect weight gain in their respective model.

```
subset5a.dat <- gained.dat[which(complete.cases(gained.dat[, c("gender", "Age", "shift", "Vig.ex.Time", "Mo
d.ex.time", "Walk.ex.Time", "initial_BMI")])),]
subset5b.dat <- gained.dat[which(complete.cases(gained.dat[, c("gender", "Age", "height", "shift", "Vig.ex.
Time", "Mod.ex.time", "Walk.ex.Time", "initial_bweight")])),]

SA12.model5a <- glm(WG ~ gender + Age + shift + Vig.ex.Time  + Mod.ex.time + Walk.ex.Time + initial_BMI, da
ta=subset5a.dat, family = binomial)
SA12.model5b <- glm(WG ~ gender + Age + height + shift + Vig.ex.Time  + Mod.ex.time + Walk.ex.Time +initial
_bweight, data=subset5b.dat, family = binomial)

tab_model(SA12.model5a, SA12.model5b)
```

		WG			**WG**	
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	1.45	0.03 – 65.22	0.835	0.07	0.00 – 141.47	0.501
gender [Female]	4.98	0.16 – 151.80	0.299	5.19	0.16 – 163.64	0.295
gender [Male]	3.08	0.10 – 95.62	0.471	3.10	0.10 – 96.24	0.468
Age	1.01	0.98 – 1.04	0.424	1.01	0.98 – 1.04	0.444
shift [8am]	0.76	0.25 – 2.15	0.613	0.76	0.25 – 2.15	0.612
shift [9am]	1.22	0.34 – 4.23	0.755	1.21	0.34 – 4.19	0.763
shift [10am]	1.72	0.48 – 6.14	0.396	1.70	0.47 – 6.07	0.408
shift [11am]	0.76	0.22 – 2.54	0.653	0.76	0.22 – 2.55	0.656
shift [12pm]	1.42	0.28 – 8.36	0.675	1.37	0.27 – 8.01	0.709
shift [1pm]	1.07	0.15 – 9.75	0.946	1.06	0.15 – 9.53	0.957
shift [2pm]	8245909.55	0.00 – NA	0.987	8548819.34	0.00 – NA	0.987
shift [other]	2.02	0.41 – 12.37	0.407	2.02	0.41 – 12.42	0.407
Vig.ex.Time	1.00	1.00 – 1.00	0.734	1.00	1.00 – 1.00	0.757
Mod.ex.time	1.00	1.00 – 1.00	0.666	1.00	1.00 – 1.00	0.649
Walk.ex.Time	1.00	1.00 – 1.00	0.594	1.00	1.00 – 1.00	0.605
initial_BMI	0.95	0.90 – 1.00	0.055			
height				1.04	0.95 – 1.16	0.401
initial_bweight				0.99	0.98 – 1.00	0.078
Observations	242			242		
R^2 Tjur	0.073			0.071		

Lastly, we utilized the function ‘stepAIC’ to find the simplest model. We called this function twice–once with the model consisting of the anthropometric variables, the total met minutes components, shift, and initial body weight, and once with the model consisting of the same variables but instead of initial weight, we have initial BMI. The stepAIC function suggest that the simplest model that includes initial BMI uses variables `gender` and `initial_BMI` as predictors. The simplest model that includes the initial body weight uses just `initial_bweight` as a predictor.

```
tab_model(best.modela, best.modelb)
```

	WG			WG		
Predictors	*Odds Ratios*	*CI*	*p*	*Odds Ratios*	*CI*	*p*
(Intercept)	2.42	0.08 – 71.46	0.565	6.52	2.15 – 20.51	**0.001**
gender [Female]	2.94	0.11 – 75.51	0.450			
gender [Male]	1.61	0.06 – 41.68	0.741			
initial_BMI	0.96	0.92 – 1.01	0.132			
initial_bweight				0.99	0.99 – 1.00	**0.047**
Observations	242			242		
R^2 Tjur	0.032			0.016		

```
AIC(best.modela,best.modelb)
```

```
##             df      AIC
## best.modela  4 301.6251
## best.modelb  2 301.2257
```

Thus, we cannot state, to statistical significance, that either specific aim 1 or aim 2 are true. We can, however, suggest that initial BMI and gender are better predictors of weight gain than either Total MET Minutes or shift. See the summary of the recommended model below.

We find the best, simplest model for `weightgained` includes only the single predictor, `initial_bodyweight` .

Model 6

Lastly, we wanted to observe the binomial models created from `initial_BMI` x `shift` x `CalcTMM` as well as `initial_bweight` x `shift` x `CalcTMM` . We then ran the stepAIC() function through both models, and obtained a lower AIC value for the model using `initial_BMI` x `shift` x `CalcTMM` . This suggests this model is simpler than the other. In this model, the following predictors had p-values <0.1: shift10am:CalcTMM, initial_BMI:shift10am:CalcTMM. This suggests that, at significance level 0.1, these products are significant in predicting whether an employee has gained weight.

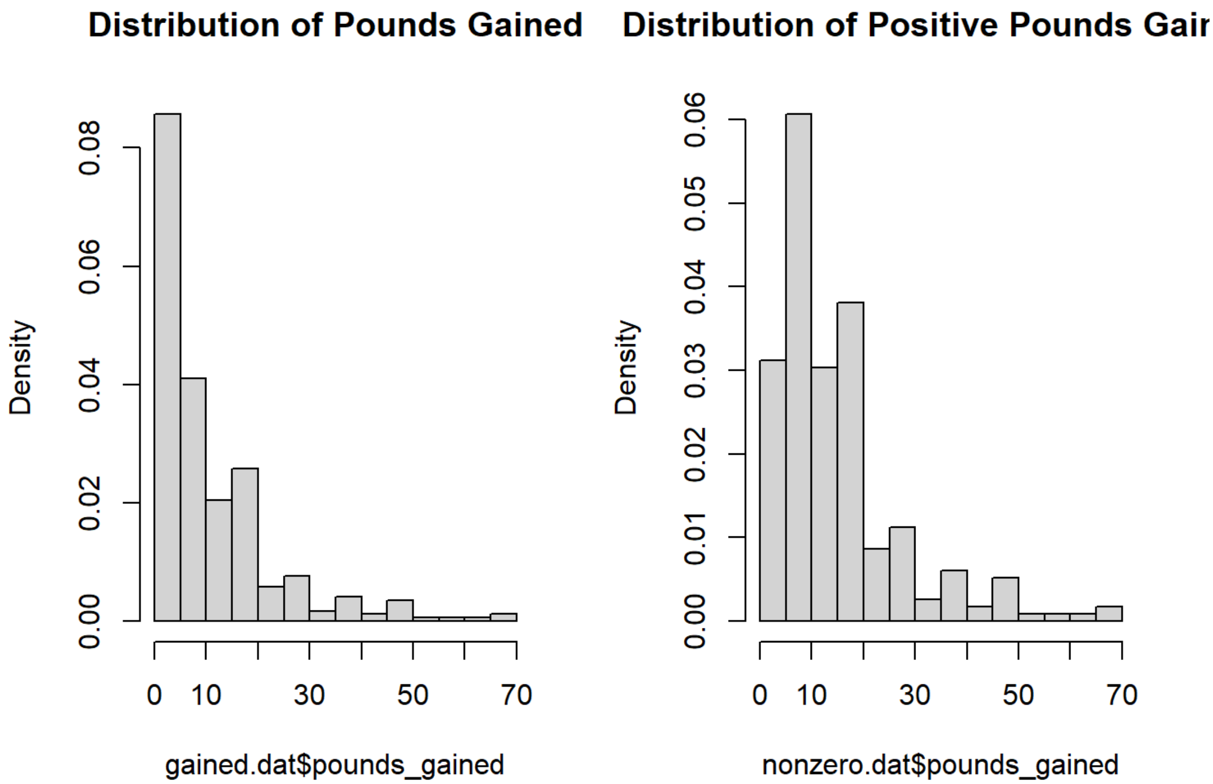
```
## Start:  AIC=325.08
## WG ~ initial_BMI * shift * CalcTMM
##
##                             Df Deviance    AIC
## <none>                           253.08 325.08
## - initial_BMI:shift:CalcTMM  8   270.92 326.92
```

```
##
## Call:  glm(formula = WG ~ initial_BMI * shift * CalcTMM, family = binomial,
##     data = subset6a.dat)
##
## Coefficients:
##                  (Intercept)                     initial_BMI
##                   -2.628e-01                       2.560e-02
##                     shift8am                        shift9am
##                    2.106e+00                       1.581e+00
##                    shift10am                       shift11am
##                    3.699e+00                       2.334e+00
##                    shift12pm                        shift1pm
##                   -3.378e+01                      -1.231e+02
##                     shift2pm                      shiftother
##                    9.971e+00                      -9.214e-01
##                      CalcTMM            initial_BMI:shift8am
##                    7.085e-03                      -6.524e-02
##         initial_BMI:shift9am           initial_BMI:shift10am
##                   -4.026e-02                      -4.925e-02
##        initial_BMI:shift11am           initial_BMI:shift12pm
##                   -1.263e-01                       1.555e+00
##         initial_BMI:shift1pm            initial_BMI:shift2pm
##                    5.311e+00                      -4.352e-01
##       initial_BMI:shiftother             initial_BMI:CalcTMM
##                    5.928e-03                      -2.447e-04
##             shift8am:CalcTMM                shift9am:CalcTMM
##                   -7.153e-03                      -5.916e-03
##            shift10am:CalcTMM               shift11am:CalcTMM
##                   -1.260e-02                      -6.443e-04
##            shift12pm:CalcTMM                shift1pm:CalcTMM
##                    3.956e-02                       6.347e-02
##             shift2pm:CalcTMM              shiftother:CalcTMM
##                   -1.103e-02                       3.146e-04
##   initial_BMI:shift8am:CalcTMM    initial_BMI:shift9am:CalcTMM
##                    2.398e-04                       1.948e-04
##  initial_BMI:shift10am:CalcTMM   initial_BMI:shift11am:CalcTMM
##                    4.277e-04                       5.416e-05
##  initial_BMI:shift12pm:CalcTMM    initial_BMI:shift1pm:CalcTMM
##                   -1.679e-03                      -2.472e-03
##   initial_BMI:shift2pm:CalcTMM  initial_BMI:shiftother:CalcTMM
##                    4.828e-04                       6.881e-05
##
## Degrees of Freedom: 248 Total (i.e. Null);  213 Residual
## Null Deviance:       311.1
## Residual Deviance: 253.1     AIC: 325.1
```

Analysis of Continuous Response (pounds gained)

(SA1) Does *total metabolic minutes* have an effect on *weight gain*?

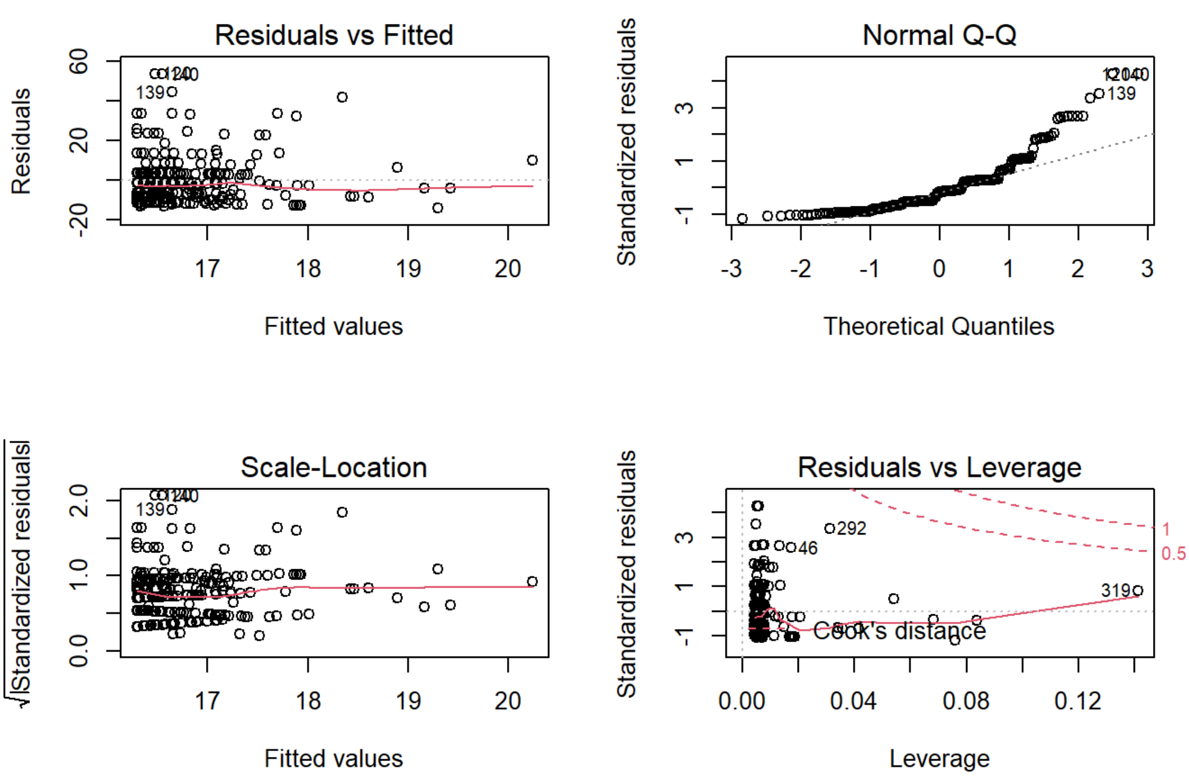
We have information about net pounds gained. We assume that when `weightgained` is false, we can substitute a value of 0 for `pounds_gained`. This allows us to analyze a full data set; otherwise, we limit our observations. It is worth noting that we may be oversimplifying cases were `pounds_gained` may be negative, thus creating a censored or zero-inflated data set. Therefore, we consider all possible data, and the subset where pounds gained is non-zero.



Pounds gained is highly skewed, even when zero observations are excluded.

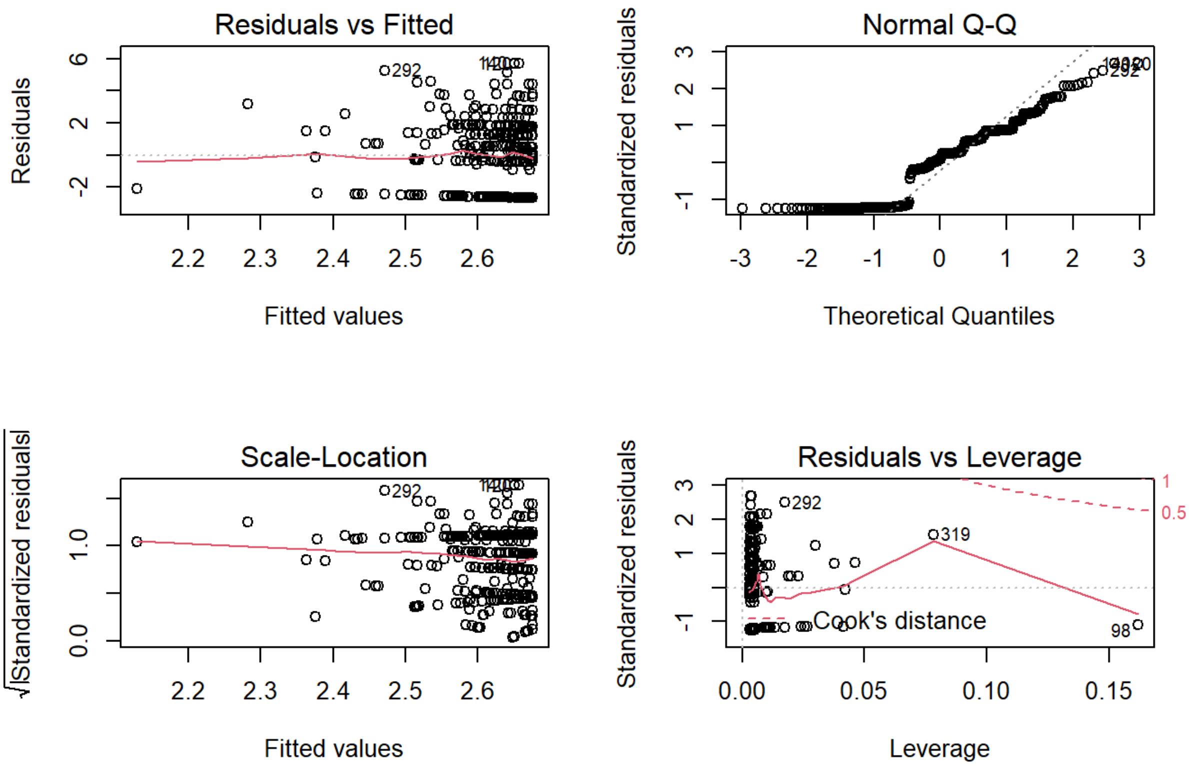
Linear model, gaussian errors

A linear model, with the assumption of normally distributed errors, may not be appropriate. We include a linear model here for reference only; this is not the recommended analysis.



Square-root transform

The distribution of weight gain is highly left-skewed. We may correct this by applying a square-root transformation. We apply this to both the full data (with extra 0s) and the data limited to nonzero weight gain.The square-root transformation does appear to improve upon the skewness of the data, although the residuals are not clearly normally distributed.

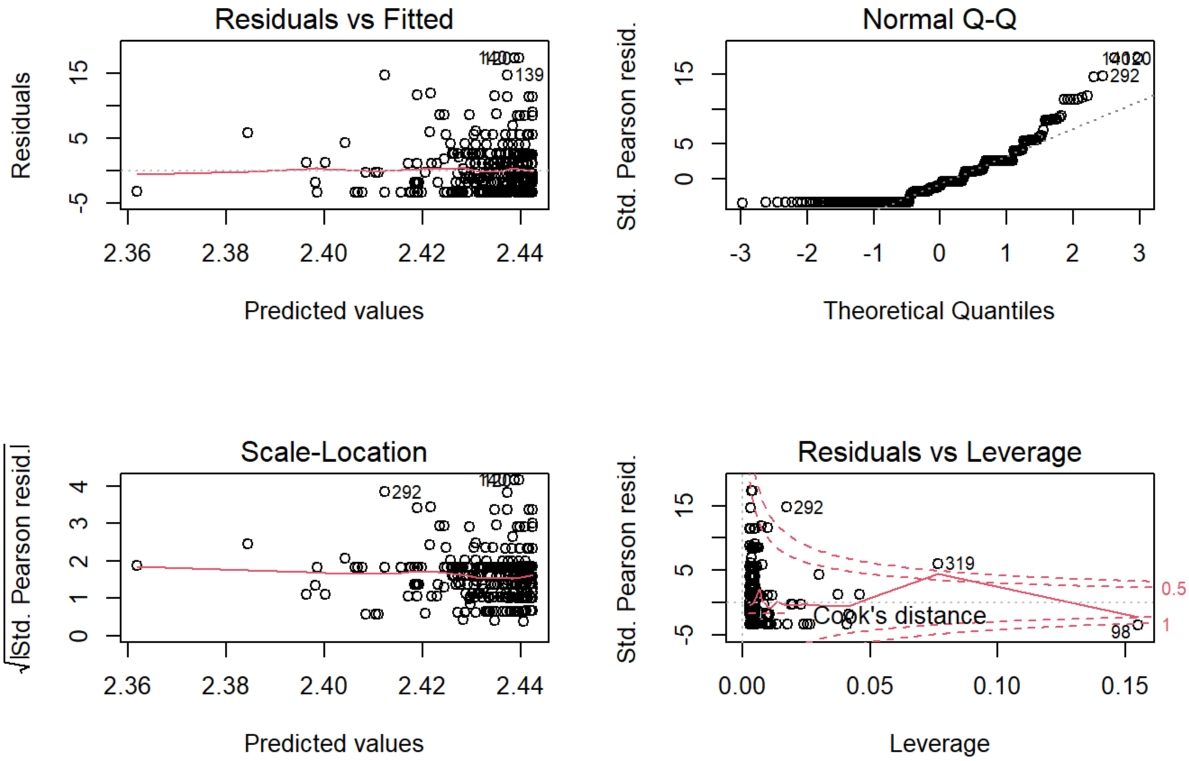


Log transformation

The log transformation is also commonly used to correct skewed data. The log-transform, however, is not defined for 0 values. We do not recommend analyzing these data using a log transformation.

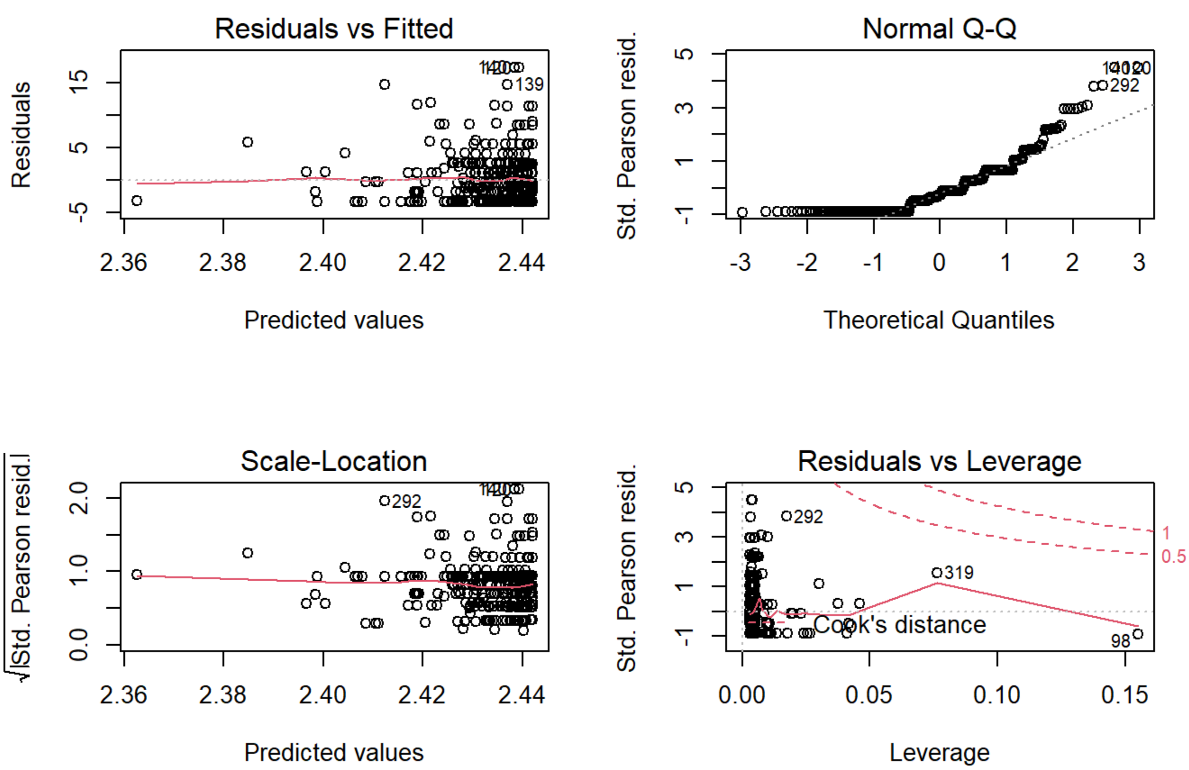
Poisson regression

The square root transformation is commonly applied to count data. This suggests a possible Poisson model for pounds gained. Poisson regression requires integer values. We'll round pounds gained for this. This poisson model provides a similar improvement on residual errors as did the square root transform. Thus, a Poisson model may be recommended for these data.



Quasi-poisson

We can, alternatively, fit a quasi-poisson family. This does not require integer values. This also provides a dispersion parameter that may help account for excess 0s.



Gamma

The shape of the distribution of weight gain suggests a gamma distribution. However, the gamma distribution is not defined for 0 values. We do not recommend this gamma model.

Zero-inflated Poisson

Of the statistical models considered to this point, the Poisson distribution family seems most suitable for these data. We now consider a zero-inflated Poisson (ZIP) analysis.

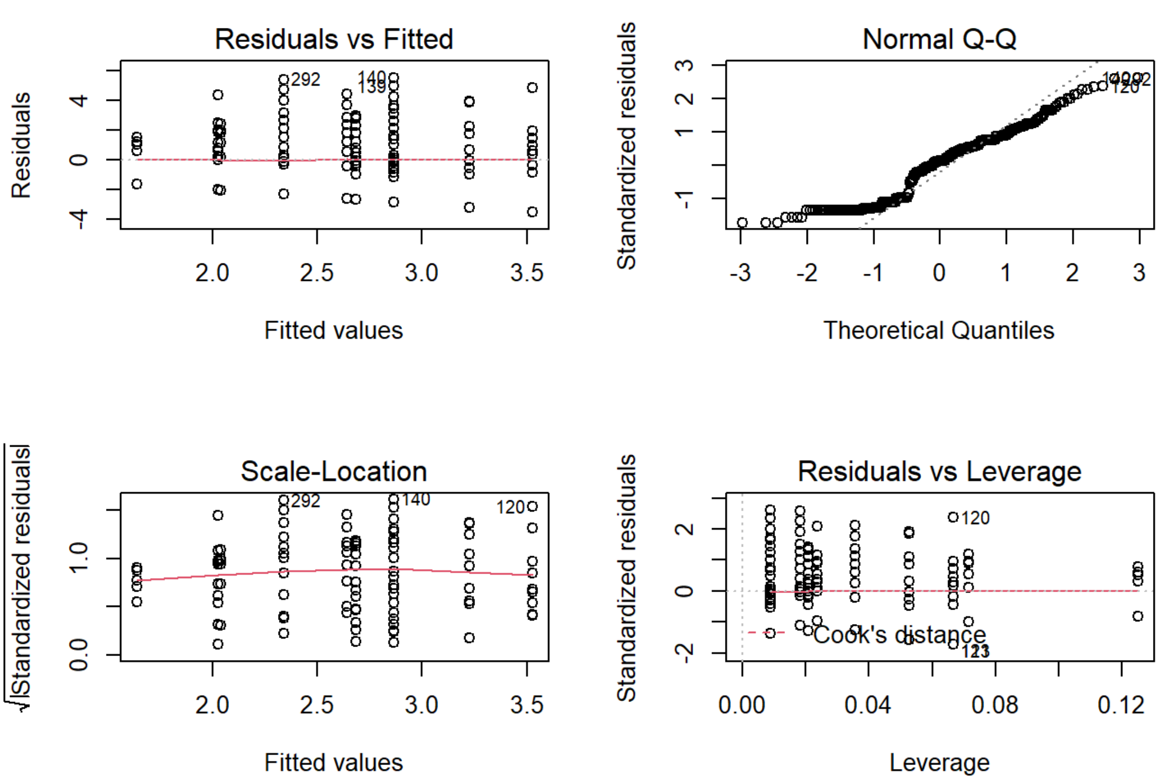
Briefly, ZIP defines a conditional probability model. The first stage is modeled as binomial - weight gain is either false (0 pounds gained) or true (a non-zero pounds gained value), with a defined probability. Then, conditional on weight gain being true, the remaining values are fit to a Poisson distribution. This is computed in R using the `pscl` library:

		LBS	
Predictors	*Incidence Rate Ratios*	*CI*	*p*
Count Model			
(Intercept)	16.32	NaN – NaN	NaN
CalcTMM	1.00	NaN – NaN	NaN
Zero-Inflated Model			
(Intercept)	0.43	0.30 – 0.61	**<0.001**
CalcTMM	1.00	1.00 – 1.00	0.519
Observations	341		
R^2 / R^2 adjusted	0.004 / -0.002		

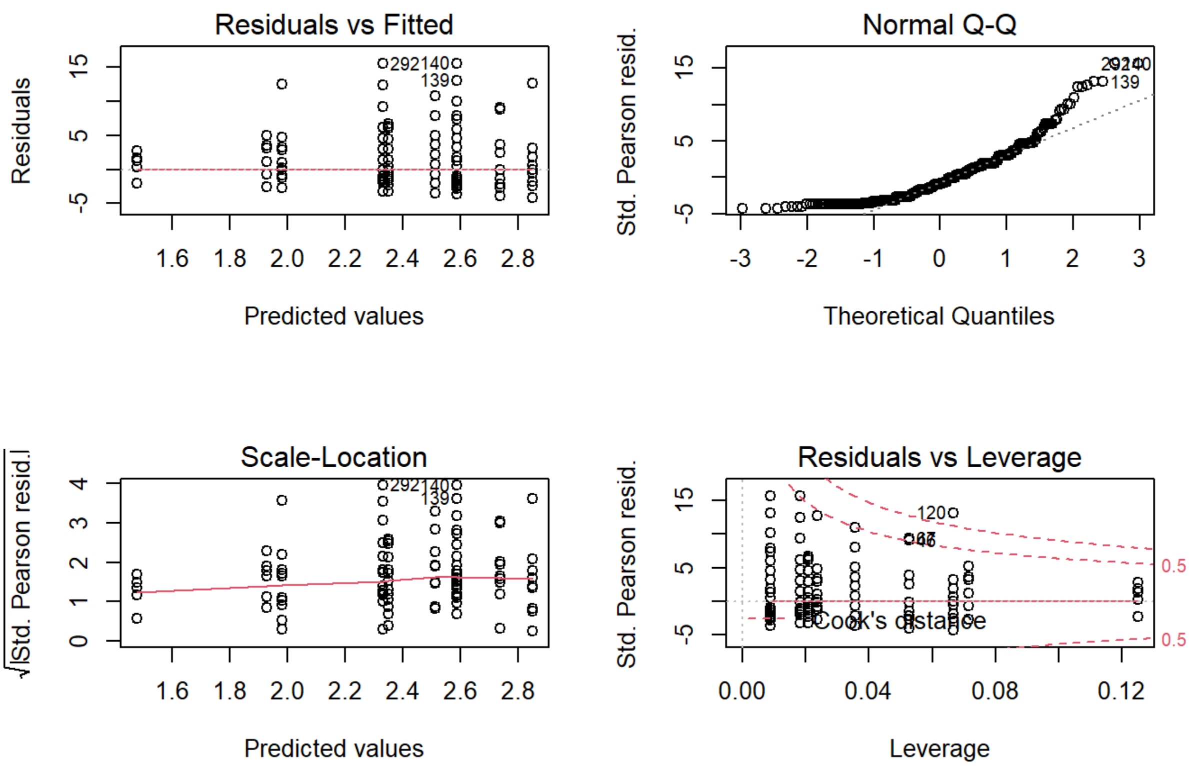
(SA2) Does *shift* have an effect on *weight gain*?

We repeat the analysis of different statistical distributions from above, using *shift* as a predictor variable.

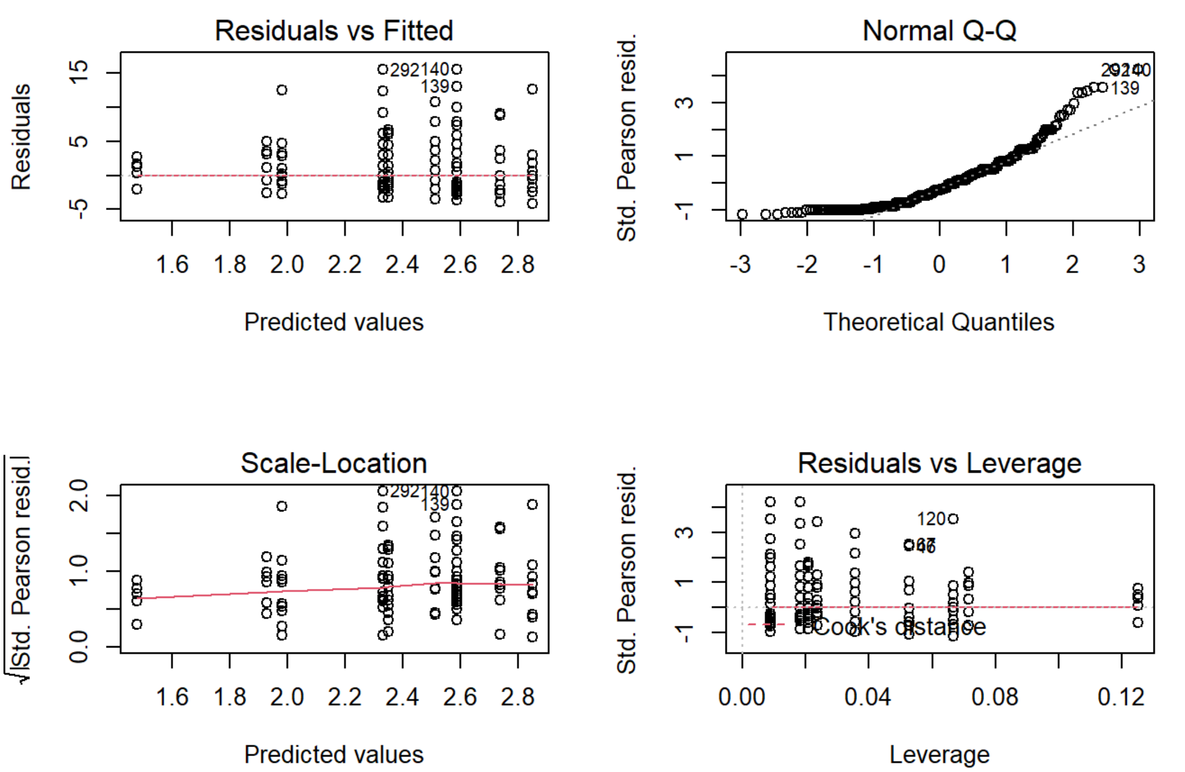
Square Root



Poisson



Quasi Poisson



Zero Inflated

	LBS		
Predictors	*Incidence Rate Ratios*	*CI*	*p*
Count Model			
(Intercept)	20.29	18.26 – 22.55	**<0.001**
shift [8am]	0.95	0.85 – 1.07	0.419
shift [9am]	0.82	0.72 – 0.94	**0.003**
shift [10am]	0.69	0.60 – 0.79	**<0.001**
shift [11am]	0.58	0.49 – 0.67	**<0.001**
shift [12pm]	0.53	0.42 – 0.66	**<0.001**
shift [1pm]	0.34	0.24 – 0.49	**<0.001**
shift [2pm]	1.06	0.91 – 1.25	0.454
shift [other]	0.96	0.82 – 1.12	0.631
Zero-Inflated Model			
(Intercept)	0.65	0.30 – 1.38	0.261
shift [8am]	0.70	0.30 – 1.66	0.419
shift [9am]	0.95	0.38 – 2.43	0.922
shift [10am]	0.52	0.19 – 1.40	0.194
shift [11am]	0.95	0.36 – 2.54	0.920
shift [12pm]	0.86	0.23 – 3.25	0.822
shift [1pm]	0.93	0.18 – 4.69	0.927
shift [2pm]	0.39	0.09 – 1.69	0.206
shift [other]	0.41	0.11 – 1.57	0.194
Observations	341		
R^2 / R^2 adjusted	0.795 / 0.789		

Best Model (Initial Bodyweight)

We now consider the different statistical families in the analysis of initial body weight and pounds gained, in the context of a zero-inflated poisson model.

```
SA12.best.zero <- zeroinfl(LBS ~ initial_bweight, data = gained.dat)
tab_model(SA12.best.poisson, SA12.best.quasi, SA12.best.zero)
```

	LBS			pounds_gained			LBS		
Predictors	*Incidence Rate Ratios*	*CI*	*p*	*Incidence Rate Ratios*	*CI*	*p*	*Incidence Rate Ratios*	*CI*	*p*
(Intercept)	10.15	8.77 – 11.74	**<0.001**	10.14	5.69 – 18.09	**<0.001**			

initial_bweight	1.00	1.00 – 1.00	0.059	1.00	1.00 – 1.00	0.633			
(Intercept)							10.63	9.20 – 12.29	**<0.001**
initial_bweight							1.00	1.00 – 1.00	**<0.001**
Zero-Inflated Model									
(Intercept)							0.17	0.06 – 0.50	**0.001**
initial_bweight							1.01	1.00 – 1.01	0.053
Observations	264			264			264		
R^2 Nagelkerke	0.013			0.013			0.056 / 0.049		

Conclusions and Recommendations

- The specific aims stated for this project are partially supportED with these data. Specifically, we find no significant effect of calculated *Total MET-Minutes* (`CalcTMM`) or *shift* on `weightgain` as a binomial response, using logistic regression models.
- However, we do find a statistical significant effect of `shift` on *pounds gained*. This result is most strongly suggested using a zero-inflated poisson model to account for the individuals reporting 0 pounds gained, and weakly supported using a quasi-poisson model to account for the excess 0 values when no weight gain is reported.
- The logistic regression model suggest that *initial body weight* or *initial BMI* are possible predictors. This was identified from a step-wise model selection algorithm implemented using the `stepAIC` function is R. We considered other combinations of variables, but the best, simplest model included only BMI or initial body weight.
- When interactions among `shift` , `CalcTMM` and `initial_BMI` are included in the model, there is a slight (but not significant at $p<0.05$) suggestion that weight gain may differ among shifts. Thus, initial body weight or initial BMI may be a confounding factor that influences the two variables identified in the specific aims. We note, however, that there were ~100 observations that did not have initial body weight or initial BMI, so this may warrant further investigation, and greater care should be taken when collecting data.