Homework 0

Brief Data Set Sumamry:

- Number of unique chicks: 50
- Total observations: 578
- Average number of observations per Chick: 11.56; 12 observations for those with complete data
- Number of Chicks with complete observation period: 45; number with last time before time = 21: 5
- Number of Chicks with skipped measurement: 0
- Number of observations for Chicks with incomplete data (maximum number of observation periods is 12):

Chick ID	N observed Periods	Last Observed Time
8	11	20
15	8	14
16	7	12
18	2	2
44	10	18

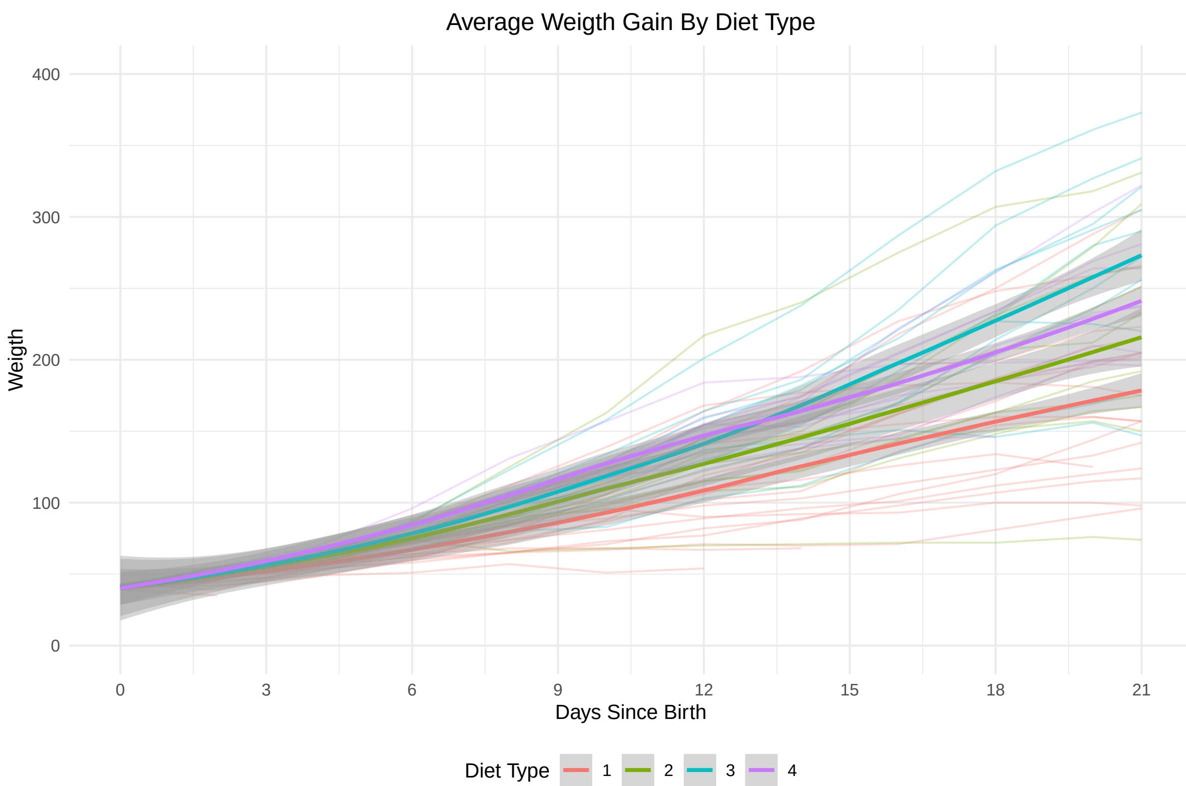

Figure 1: Average Growth Trend by Diet

Longitudinal Plot

Regression Analysis

Table 1: Linear Regression Model Estimates On Correlated Data

Model Term	Estimate	SE	T-value	P-value
Intercept	30.93	4.25	7.28	0.00
Diet 2 Factor	-2.30	7.27	-0.32	0.75
Diet 3 Factor	-12.68	7.27	-1.74	0.08
Diet 4 Factor	-0.14	7.29	-0.02	0.98
Time Slope	6.84	0.34	20.08	0.00
Diet 2 Factor	1.77	0.57	3.09	0.00
Diet 3 Factor	4.58	0.57	8.01	0.00
Diet 4 Factor	2.87	0.58	4.97	0.00

Takeaways

- Failure to account for proper correlation structures produces standard errors of regression lines that are way too small
  - Linear model assumptions state that the observations are independent, meaning that for chick with ID = 1 weight at time 15 is independent of weight weight at 7, which is incorrect
  - Correlation within the subject produces a strong linear trend within a subject, but the trend lines are different from subject to subject, which we are not able to account for using a linear regression model
- Perhaps the slope estimate itself is a valid results, just like in the case of highly correlated predictors. However, standard errors are definitely incorrect
- There is *some* effect of diet on weight change, with diet 3 producing the highest average weight gain, however, we are not able to make a proper statistical conclusion due to wrong standard error of model estimates