Lab Repeated Measures 1 – Lab Key

1. Birthorder does not vary between clusters (all moms have births 1 to 5). Birthweights vary within clusters (moms) because the babies are different weights.
2. Lowess plot is about as straight as it could be. So treat as linear.
3. Table agrees well with plot. Progression in table with birthorder is similar to that from the lowess plot.
4. It appears statistically significant. But mother's age is increasing along with birth order. So could be confounding with age.
5. Now we can easily calculate the differences between the last and first birthweights and conduct a t-test.

   ```
   . generate bwdiff=bweight5-bweight1
   . ttest bwdiff=0
   ```

   How would you expect this to compare to the regression, given that we are ignoring the three intermediate births? How does it compare with regard to p-value for the test of birth order?

   Ans: We'd ordinarily expect that the t-test, which discards information, would be less precise and give a smaller t-statistic. In this case it doesn't (t=4.2 for t-test versus t=3.6 for regression). We are seeing the effect of ignoring the clustering, which the paired t-test takes advantage of. It gives a more precise analysis by utilizing within person differences.

6. Here is a regression of birth weight on birth order, taking account of the clustering on mom:

   ```
   mixed bweight birthord || momid:, reml
   ```

   How does the p-value compare to the t-test and the regression? Does it make sense?

   Ans: The statistic for testing the coefficient is larger (z=4.7 as compared to t=4.2 or 3.6 before) and hence has a smaller p-value. This would be expected because it accounts for the clustering (which improves precision) and uses all the data (which also improves precision). Note that the size of the coefficient is largely the same, it is the standard error that has decreased. `xtgee` analysis is virtually identical.

7. Georgia babies diagnostics:

   The quadratic terms for either birthorder or initial age are not statistically significant, so the linearity assumption is reasonable.

```
mixed bweight birthord initage c.birthord#c.birthord c.initage#c.initage || momid:, reml
------------------------------------------------------------------------------
             bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+--------------------------------------------------------------
           birthord |   110.9151   51.41255     2.16   0.031     10.14839    211.6819
            initage |   91.63336   71.97621     1.27   0.203    -49.43743    232.7041
                    |
c.birthord#c.birthord |  -10.71786   8.406851    -1.27   0.202    -27.19498    5.759269
```

```
                            |
  c.initage#c.initage |   -1.60959   1.771027    -0.91   0.363    -5.080739    1.861559
                            |
                _cons |   1825.055    710.924     2.57   0.010     431.6697    3218.441
------------------------------------------------------------------------------
```

8. The histogram and qnorm plot shows a distribution with a few clear outliers. So we should redo the analysis without them. If we repeat the analysis removing the outliers (Conservatively we could drop all of the ones with standardized residuals larger than 3 adding "if abs(resids)<3" to the command in Stata or filtering the dataset in R) the results are qualitatively similar.

   More specifically, rerunning the analysis without them changes the birth order coefficient from 47 to 40, a modest drop. It changes the initial age coefficient from 27 to 28, a minimal change. Qualitative conclusions are similar – both still exhibit highly statistically significant associations with birthweight.

9. How does this compare to the descriptive statistics?

   Ans: It gives exactly the same estimates and gives a formal statistical comparison of the differences. Not statistically significant (p=0.17).

10. Why did we need to include the interaction? How does this compare to the t-test?

    Ans: Since we are interested in the *difference in* the change over time between men and women that is the interaction of time and sex. The interaction effect is exactly the same as the difference in the change scores, with the same p-value.

11. How does this compare?

    Ans: Whether you analyze the 12 month value or the change score, the sex effect is the same (this is true in general). So in simple situations the more complicated longitudinal analysis does the sensible thing and is equivalent to an analysis of change scores.

    With multiple time points and/or missing data, however, the longitudinal analysis will be better.

    Adjusting for the baseline value is not a good idea when the baseline values are not equal or expected to be equal as is usually the case in an observational study. The analysis can no longer be interpreted as an analysis of the change in pain over time, which was the original question. An exception is a randomized trial where values should be the same at baseline. In such a case, the adjustment is innocuous to the overall estimates and may increase the power.