Meeting Notes 11/12

Chansoo Song

11/12/2018

```
# Read Data
df = read.csv('../data/master.csv')

# Data Exported from QGIS (schools merged with zones)
################################
qgis_exports = list.files('../data/qgis_exports/')
qgis_list = list()
qgis_list = lapply(paste('../data/qgis_exports/',qgis_exports,sep='/'),
                   read.csv, stringsAsFactors = F)

# Then combine list of data frames into single dataframe
qgis_df = rbind.fill(qgis_list)
cols_to_keep = c('DBN','esid_no','Year')

# Subset Dataframe
qgis_df = qgis_df[qgis_df$math==1,cols_to_keep]

# Filter / Subset Dataset
################################
master = df %>%
  filter(math == 1,
         Grade == 4) %>%
  mutate(year_sch = paste(DBN,Year,sep='_'))
```

```
## Warning: package 'bindrcpp' was built under R version 3.4.4
```

```
master = master %>%
  left_join(qgis_df, by=c('DBN','Year'))
```

```
## Warning: Column `DBN` joining factor and character vector, coercing into
## character vector
```

```
# Standardize Scores for Math / All Grades
mean_scores_year = tapply(master$Mean.Scale.Score,master$Year,mean)
sd_year = tapply(master$Mean.Scale.Score,master$Year,sd)
years = names(mean_scores_year)

for(i in 1:length(years)){
  master$Mean.Scale.Score[master$Year==years[i]] =
    (master$Mean.Scale.Score[master$Year==years[i]] - mean_scores_year[i]) / sd_year[i]
}
```

```
################################
# Model
################################

# Charter Count
temp = master %>%
  group_by(GEOGRAPHICAL_DISTRICT_CODE, Year) %>%
```

```
  dplyr::summarize(charter_count = sum(charter))

master2 = master %>%
  filter(charter == 0) %>%
  left_join(temp, by=c('Year','GEOGRAPHICAL_DISTRICT_CODE'))

pooled.1 = lm(Mean.Scale.Score ~ charter_count, data = master2)

mm.mod1 = lmer(Mean.Scale.Score ~ (1 | GEOGRAPHICAL_DISTRICT_CODE),
               data = master2)
```

```
## Warning: 'rBind' is deprecated.
##  Since R version 3.2.0, base's rbind() should work fine with S4 objects
```

```
mm.mod2 = lmer(Mean.Scale.Score ~ (1 | DBN),
               data = master2)
mm.mod3 = lmer(Mean.Scale.Score ~ (1 | esid_no),
               data = master2)
mm.mod4 = lmer(Mean.Scale.Score ~ (1 | GEOGRAPHICAL_DISTRICT_CODE/DBN),
               data = master2)
mm.mod5 = lmer(Mean.Scale.Score ~ charter_count + (1 | GEOGRAPHICAL_DISTRICT_CODE),
               data = master2)
mm.mod6 = lmer(Mean.Scale.Score ~ charter_count + (1 | DBN),
               data = master2)
mm.mod7 = lmer(Mean.Scale.Score ~ charter_count + (1 | esid_no),
               data = master2)
mm.mod8 = lmer(Mean.Scale.Score ~ charter_count + (1 | GEOGRAPHICAL_DISTRICT_CODE/esid_no/DBN),
               data = master2)
mm.mod9 = lmer(Mean.Scale.Score ~ charter_count + Poverty + Disabled + Ell + Asian +
                 Black + Hispanic + (1 | GEOGRAPHICAL_DISTRICT_CODE/esid_no/DBN),
               data = master2)
mm.mod10 = lmer(Mean.Scale.Score ~ charter_count + Poverty + Disabled + Ell + Asian + Black +
                  Hispanic + (1 | GEOGRAPHICAL_DISTRICT_CODE) + (charter_count | esid_no/DBN),
                data = master2)

summary(pooled.1)
```

```
##
## Call:
## lm(formula = Mean.Scale.Score ~ charter_count, data = master2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6361 -0.7195 -0.0722  0.6467  4.3535
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.322252   0.021210   15.19   <2e-16 ***
## charter_count -0.133544   0.005466  -24.43   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.962 on 4392 degrees of freedom
## Multiple R-squared:  0.1197, Adjusted R-squared:  0.1195
```

```
## F-statistic:   597 on 1 and 4392 DF,  p-value: < 2.2e-16
```

```
summary(mm.mod1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Mean.Scale.Score ~ (1 | GEOGRAPHICAL_DISTRICT_CODE)
##    Data: master2
##
## REML criterion at convergence: 10465.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4092 -0.6426 -0.0362  0.6190  4.6131
##
## Random effects:
##  Groups                      Name        Variance Std.Dev.
##  GEOGRAPHICAL_DISTRICT_CODE (Intercept) 0.4560   0.6752
##  Residual                               0.6128   0.7828
## Number of obs: 4394, groups:  GEOGRAPHICAL_DISTRICT_CODE, 32
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  -0.1335     0.1200  -1.113
```

```
summary(mm.mod2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Mean.Scale.Score ~ (1 | DBN)
##    Data: master2
##
## REML criterion at convergence: 6431.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7519 -0.5581  0.0146  0.5760  5.5621
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  DBN      (Intercept) 0.9252   0.9619
##  Residual             0.1333   0.3651
## Number of obs: 4394, groups:  DBN, 762
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) -0.04689    0.03531  -1.328
```

```
summary(mm.mod3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Mean.Scale.Score ~ (1 | esid_no)
##    Data: master2
##
## REML criterion at convergence: 7000.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -4.5085 -0.4822  0.0083  0.4617  6.3726
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  esid_no  (Intercept) 0.8477   0.9207
##  Residual             0.2314   0.4810
## Number of obs: 3694, groups:  esid_no, 638
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) -0.02181    0.03745  -0.582
```

```
summary(mm.mod4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Mean.Scale.Score ~ (1 | GEOGRAPHICAL_DISTRICT_CODE/DBN)
##    Data: master2
##
## REML criterion at convergence: 6086.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7551 -0.5655  0.0198  0.5680  5.6426
##
## Random effects:
##  Groups                          Name        Variance Std.Dev.
##  DBN:GEOGRAPHICAL_DISTRICT_CODE (Intercept) 0.5089   0.7133
##  GEOGRAPHICAL_DISTRICT_CODE     (Intercept) 0.4259   0.6526
##  Residual                                   0.1334   0.3652
## Number of obs: 4394, groups:
## DBN:GEOGRAPHICAL_DISTRICT_CODE, 762; GEOGRAPHICAL_DISTRICT_CODE, 32
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  -0.1236     0.1186  -1.042
```

```
summary(mm.mod5)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Mean.Scale.Score ~ charter_count + (1 | GEOGRAPHICAL_DISTRICT_CODE)
##    Data: master2
##
## REML criterion at convergence: 10471.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4241 -0.6403 -0.0347  0.6194  4.6041
##
## Random effects:
##  Groups                      Name        Variance Std.Dev.
##  GEOGRAPHICAL_DISTRICT_CODE (Intercept) 0.4327   0.6578
##  Residual                               0.6130   0.7829
## Number of obs: 4394, groups:  GEOGRAPHICAL_DISTRICT_CODE, 32
##
```

```
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)   -0.09224    0.12276  -0.751
## charter_count -0.01283    0.01164  -1.103
##
## Correlation of Fixed Effects:
##             (Intr)
## charter_cnt -0.304
```

```
summary(mm.mod6)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Mean.Scale.Score ~ charter_count + (1 | DBN)
##    Data: master2
##
## REML criterion at convergence: 6413.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7841 -0.5554  0.0189  0.5725  5.7886
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  DBN      (Intercept) 0.8788   0.9374
##  Residual             0.1337   0.3657
## Number of obs: 4394, groups:  DBN, 762
##
## Fixed effects:
##                Estimate Std. Error t value
## (Intercept)    0.032202   0.037508   0.859
## charter_count -0.027882   0.005234  -5.327
##
## Correlation of Fixed Effects:
##             (Intr)
## charter_cnt -0.396
```

```
summary(mm.mod7)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Mean.Scale.Score ~ charter_count + (1 | esid_no)
##    Data: master2
##
## REML criterion at convergence: 6967.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5494 -0.4740  0.0130  0.4554  6.3190
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  esid_no  (Intercept) 0.7618   0.8728
##  Residual             0.2332   0.4829
## Number of obs: 3694, groups:  esid_no, 638
##
## Fixed effects:
```

```
##                Estimate Std. Error t value
## (Intercept)    0.121334   0.041645   2.914
## charter_count -0.051158   0.007709  -6.636
##
## Correlation of Fixed Effects:
##             (Intr)
## charter_cnt -0.518
```

```
summary(mm.mod8)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Mean.Scale.Score ~ charter_count + (1 | GEOGRAPHICAL_DISTRICT_CODE/esid_no/DBN)
##    Data: master2
##
## REML criterion at convergence: 5291.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9138 -0.5370  0.0206  0.5531  5.3120
##
## Random effects:
##  Groups                                   Name        Variance Std.Dev.
##  DBN:(esid_no:GEOGRAPHICAL_DISTRICT_CODE) (Intercept) 0.43073  0.6563
##  esid_no:GEOGRAPHICAL_DISTRICT_CODE       (Intercept) 0.09446  0.3073
##  GEOGRAPHICAL_DISTRICT_CODE               (Intercept) 0.39570  0.6290
##  Residual                                             0.12754  0.3571
## Number of obs: 3694, groups:
## DBN:(esid_no:GEOGRAPHICAL_DISTRICT_CODE), 776; esid_no:GEOGRAPHICAL_DISTRICT_CODE, 640; GEOGRAPHICAL
##
## Fixed effects:
##                Estimate Std. Error t value
## (Intercept)   -0.077412   0.117723  -0.658
## charter_count -0.014393   0.006656  -2.163
##
## Correlation of Fixed Effects:
##             (Intr)
## charter_cnt -0.191
```

```
summary(mm.mod9)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Mean.Scale.Score ~ charter_count + Poverty + Disabled + Ell +
##     Asian + Black + Hispanic + (1 | GEOGRAPHICAL_DISTRICT_CODE/esid_no/DBN)
##    Data: master2
##
## REML criterion at convergence: 4978.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8736 -0.5403  0.0185  0.5624  5.3150
##
## Random effects:
##  Groups                                   Name        Variance Std.Dev.
##  DBN:(esid_no:GEOGRAPHICAL_DISTRICT_CODE) (Intercept) 0.2975   0.5455
```

```
##  esid_no:GEOGRAPHICAL_DISTRICT_CODE       (Intercept) 0.1100   0.3316
##  GEOGRAPHICAL_DISTRICT_CODE               (Intercept) 0.1948   0.4414
##  Residual                                             0.1289   0.3590
## Number of obs: 3521, groups:
## DBN:(esid_no:GEOGRAPHICAL_DISTRICT_CODE), 751; esid_no:GEOGRAPHICAL_DISTRICT_CODE, 618; GEOGRAPHICAL
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)    0.2372767  0.1016807   2.334
## charter_count -0.0199828  0.0068211  -2.930
## Poverty       -0.0001388  0.0002029  -0.684
## Disabled      -0.0025801  0.0004948  -5.214
## Ell           -0.0036046  0.0003468 -10.395
## Asian          0.0028580  0.0002308  12.384
## Black         -0.0007874  0.0002376  -3.314
## Hispanic       0.0009863  0.0002538   3.886
##
## Correlation of Fixed Effects:
##             (Intr) chrtr_ Povrty Disbld Ell    Asian  Black
## charter_cnt -0.281
## Poverty     -0.121  0.006
## Disabled    -0.160 -0.102 -0.062
## Ell          0.040 -0.043 -0.091  0.066
## Asian       -0.065  0.058 -0.450 -0.148 -0.423
## Black       -0.108  0.145 -0.602 -0.253  0.063  0.401
## Hispanic    -0.055  0.094 -0.611 -0.357 -0.372  0.487  0.453
```

```
summary(mm.mod10)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Mean.Scale.Score ~ charter_count + Poverty + Disabled + Ell +
##     Asian + Black + Hispanic + (1 | GEOGRAPHICAL_DISTRICT_CODE) +
##     (charter_count | esid_no/DBN)
##    Data: master2
##
## REML criterion at convergence: 4951.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8259 -0.5342  0.0176  0.5559  4.1839
##
## Random effects:
##  Groups                     Name          Variance Std.Dev. Corr
##  DBN:esid_no                (Intercept)   0.292752 0.54107
##                             charter_count 0.001368 0.03699  -0.08
##  esid_no                    (Intercept)   0.052478 0.22908
##                             charter_count 0.002341 0.04838  0.17
##  GEOGRAPHICAL_DISTRICT_CODE (Intercept)   0.192604 0.43887
##  Residual                                 0.125836 0.35473
## Number of obs: 3521, groups:
## DBN:esid_no, 751; esid_no, 616; GEOGRAPHICAL_DISTRICT_CODE, 32
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)    2.632e-01  1.006e-01   2.615
```

```
## charter_count -2.536e-02  8.159e-03  -3.109
## Poverty       -7.551e-05  1.987e-04  -0.380
## Disabled      -2.448e-03  4.883e-04  -5.013
## Ell           -3.524e-03  3.384e-04 -10.414
## Asian          2.737e-03  2.241e-04  12.215
## Black         -9.269e-04  2.332e-04  -3.974
## Hispanic       8.596e-04  2.499e-04   3.440
##
## Correlation of Fixed Effects:
##             (Intr) chrtr_ Povrty Disbld Ell    Asian  Black
## charter_cnt -0.268
## Poverty     -0.118  0.000
## Disabled    -0.161 -0.090 -0.067
## Ell          0.043 -0.046 -0.083  0.072
## Asian       -0.062  0.057 -0.457 -0.151 -0.438
## Black       -0.115  0.137 -0.594 -0.251  0.054  0.406
## Hispanic    -0.059  0.090 -0.612 -0.353 -0.380  0.499  0.449
```