



**EdSurvey: Analysis of the Early Childhood Longitudinal Study,
Kindergarten Class of 2010-11**

Module Objectives

- Reintroduce EdSurvey to users
- Review Introduction to Using EdSurvey R for Analyzing Large-Scale Data Sets module before reviewing this module
- Review the EdSurvey: Analyzing Large-Scale Assessment Data – Using NAEP and TIMSS Data as Examples before reviewing this module
- Introduce users to using EdSurvey to analyze longitudinal data, with the ECLS-K:2011 data file

Install and Load EdSurvey

- Learn more by visiting the [EdSurvey Homepage](#)
- Use R Software with EdSurvey
- Install EdSurvey

```
install.packages("EdSurvey")
```
- Load EdSurvey

```
library(EdSurvey)
```

Data Processing

Data can be downloaded via an internet connection:

Step 1: Download ECLS-K:2011 data.

```
downloadECLS_K(years = c(2011), root = "C:/EdSurveyData")
```

Step 2: Unzip *ChildK5p.zip* file in the new “~/ECLSK/2011” folder directory as specified by the root argument.

Step 3: Read in ECLS-K:2011 data.

```
eclsk11 <- readECLS_K2011(path =  
"C:/EdSurveyData/ECLS_K/2011")
```

Note: reading data may take several hours to run the first time; subsequent calls to `readECLS_K2011` are stored on the user's drive for easy access and near instant retrieval.

Data Manipulation

- Data retrieval (for data manipulation)


```
df <- getData(data = eclsk11, varnames = c("x_chsex_r",
"x12sesl", "p5sumsch", "p5nhrprm", "w5cf5pf_50"),
addAttributes = TRUE, omittedLevels = FALSE)
```
- Data manipulation, Subset rows


```
df_subset <- subset(eclsk11, x_chsex_r %in% "2: FEMALE" &
(x_raceth_r %in% 1 | x_raceth_r %in% 2))
```

Data Manipulation, Recording

- Recode the levels of a categorical variable


```
## original values
table(df$p5nhrprm, useNA = "ifany")
##
##      2      3      4      5      6      7 <NA>
##  55    72   126    43    87    62 17729
df$p5nhrprm <- ifelse(df$p5sumsch == "2: NO", 0, df$p5nhrprm)
|
## recoded values
table(df$p5nhrprm, useNA = "ifany")
##
##      0      2      3      4      5      6      7 <NA>
## 3913   55    72   126    43    87    62 13816
```
- [Using EdSurvey to Analyze ECLS-K:2011 Data](#)

Analysis Function – Summary Table

```
es1 <- edsurveyTable(formula = ~ x_chsex_r + p9curmar, data =
eclsk11, weightVar = "w9c29p_9t90", varMethod = "jackknife")
```

```
es1$data
```

##	x_chsex_r	p9curmar	N	WTD_N	PCT	SE(PCT)
## 1	1: MALE	1: MARRIED (1)	2938	1367616.83	67.608642	1.1756039
## 2	1: MALE	2: SEPARATED (2)	151	86412.02	4.271810	0.3944507
## 3	1: MALE	3: DIVORCED OR WIDOWED (3, 4)	442	250607.34	12.388866	0.7625198
## 4	1: MALE	4: NEVER MARRIED (5)	425	273190.02	13.505250	0.9075561
## 5	1: MALE 5: CIVIL UNION/DOMESTIC PARTNERSHIP (6)		81	45017.01	2.225432	0.3368478
## 6	2: FEMALE	1: MARRIED (1)	2870	1319848.64	69.131210	1.0257652
## 7	2: FEMALE	2: SEPARATED (2)	143	80672.81	4.225491	0.4357400
## 8	2: FEMALE	3: DIVORCED OR WIDOWED (3, 4)	428	224738.15	11.771365	0.6138104
## 9	2: FEMALE	4: NEVER MARRIED (5)	385	237346.10	12.431746	0.7270084
## 10	2: FEMALE 5: CIVIL UNION/DOMESTIC PARTNERSHIP (6)		82	46587.90	2.440187	0.2406026

Analysis Function – Correlation

```
cor_pearson <- cor.sdf(x = "x9mscalk5", y = "x9povty_i", data
= eclsk11, method = "Pearson", weightVar = "w9c29p_9a0")
```

Correlation value:

```
cor_pearson$correlation
```

```
## [1] 0.369858
```

Correlation order:

```
cor_pearson$correlation
```

```
## [1] 0.369858
```

Analysis Function – Linear Regression

```
lm1 <- lm.sdf(formula = x9mscalk5 ~ x12sesl + x chsex r,  
              data = eclsk11, weightVar = "w9c29p_9t90")  
summary(lm1)  
##  
## Formula: x9mscalk5 ~ x12sesl + x chsex r  
##  
## Weight variable: 'w9c29p_9t90'  
## Variance method: jackknife  
## JK replicates: 80  
## full data n: 18174  
## n used: 7877  
##  
## Coefficients:  
##  
##           coef          se          t      dof Pr(>|t|)  
## (Intercept)  121.69721    0.36292 335.3305 47.291 < 2.2e-16 ***  
## x12sesl      9.14174      0.28348  32.2488 52.998 < 2.2e-16 ***  
## x_chsex_r2: FEMALE -2.17136    0.41993  -5.1707 38.322 7.619e-06 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Multiple R-squared: 0.1711
```

Module Summary

- Introduced users to using EdSurvey to analyze longitudinal data, with the ECLS-K:2011 data file
- Supplemented first two modules in this series: Introduction to EdSurvey and EdSurvey: Analyzing Large-Scale Assessment Data – Using NAEP and TIMSS Data as Examples

Module Resources

- [Using EdSurvey R for Analyzing Large-Scale Data Sets Module](#)
- [Analyzing Large-Scale Assessment Data – Using NAEP and TIMSS Data as Examples](#)
- [Install and Load EdSurvey](#)
- [EdSurvey Homepage](#)
- [Using EdSurvey to Analyze to ECLS-K:2011 Data](#)