**Voorstel Tijdschema:**

11 maart – Werkbare hypotheses af en datasets uitgekozen.

18 maart – eerste analyses geprobeerd

25 maart – gespeeld met palt veranderen

8 april – PPD gebruikt om data te genereren

15 april – smearing en PPD vergelijken

22 april – kijken naar pruning

29 april – goede pruning gevonden

6 mei – beste BART based BA GLMM met GLMM en multilevel BART vergelijken

13 mei - mooie figuren maken

20 mei – een hoop schrijven

27 mei – methods af

3 juni – results af

10 juni – discussion schrijven

17 juni – discussion af

24 juni – finishing touch

29 juni – eerste versie Thesis af

**Datasets**

**Yaohui Zhao;John Strauss;Gonghuan Yang; John Giles;Peifeng (Perry) Hu;  
Yisong Hu;Xiaoyan Lei;Albert Park;James P. Smith; Yafeng Wang (2013).  
China Health and Retirement Longitudinal Study, 2011-2012 National  
Baseline Users’ Guide, National School of Development, Peking  
University.**

We thank the China Center for Economic Research, the National School of  
Development of Peking University for providing the data.

(ECLS-K; National Center for Education Statistics, 2016)

**[CHAPTER 7-STREET SAFETY DATA](https://multilevel-analysis.sites.uu.nl/datasets/" \l "custom-collapse-0-6)**

A sample of 100 streets are selected, and on each street a random sample of 10 persons are asked how often they feel unsafe while walking that street. The question about feeling unsafe is asked using three answer categories: 1 = never, 2 = sometimes, 3 = often. Predictor variables are age and gender; street characteristics are an economic index (standardized Z-score) and a rating of the crowdedness of the street (7-point scale). File: Safety. Used in Chapter 7 on ordinal data.

<https://github.com/MultiLevelAnalysis/Datasets-third-edition-Multilevel-book/tree/master/chapter%207/Safety>

**[CHAPTER 7-EPILEPSY DATA](https://multilevel-analysis.sites.uu.nl/datasets/" \l "custom-collapse-0-7)**

The epilepsy data come from a study by Leppik et al. (1987). They have been analyzed by many authors, including Skrondal and Rabe-Hesketh (2004). The data come from a randomized controlled study on the effect of an anti-epileptic drug versus a placebo. It is a longitudinal design. For each patient the number of seizures was measured for a two-week baseline. Next, patients were randomized to the drug or the placebo condition. For four consecutive visits the clinic collected counts of epileptic seizures in the two weeks before the visit. The data set contains the following variables: count of seizures, treatment indicator, visit number, dummy for visit #4, log of age, log of baseline count. All predictors are grand mean centered. The data come from the GLLAMM homepage at: [www.gllamm.org/books](http://www.gllamm.org/books), used in Chapter 7 on count data.

<https://github.com/MultiLevelAnalysis/Datasets-third-edition-Multilevel-book/tree/master/chapter%207/Epilepsy>

**[CHAPTER 15-GALO DATA](https://multilevel-analysis.sites.uu.nl/datasets/" \l "custom-collapse-0-17)**

The GALO data in file *galo* are from an educational study by Schijf and Dronkers (1991). They are data from 1377 pupils within 58 schools. We have the following pupil-level variables: father’s occupational status, *focc*; father’s education, *feduc*; mother’s education, *meduc*; pupil sex, *sex*; the result of GALO school achievement test, *GALO*; and the teacher’s advice about secondary education, *advice*. At the school level we have only one variable: the school’s denomination, *denom*. Denomination is coded 1 = Protestant, 2 = nondenominational, 3 = Catholic (categories based on optimal scaling). The data file *galo* contains both complete and incomplete cases, and an indicator variable that specifies whether a specific case in the data file is complete or not.

<https://github.com/MultiLevelAnalysis/Datasets-third-edition-Multilevel-book/tree/master/chapter%2015>