# **Metrics**

# Andriy Koval Tuesday, June 24, 2014

#### **Contents**

Netrics: labeling factors and exploring scales													
Data preliminaries		1											
Labeling Factor Levels		2											
Time metrics : Age, Period, Cohort		4											

## Metrics: labeling factors and exploring scales

Report explains how the response categories from NLSY97 questionnaire are labeled and demonstrates application of labeled factors in data operations and graphing.

#### **Data preliminaries**

```
KEY!BDATE, RS BIRTHDATE MONTH/
                        HOW OFTEN PR ATTEND CHURCH IN LAST
                       WHAT IS PRS CURRENT RELIGIOUS PREFER
                                  WHAT RELIGION WAS PR RAIS
                            RS AGE IN MONTHS AS OF INTERVIEW
                                         RS AGE AT INTERVIEW
        # DAYS PER WEEK TYPICALLY FAMILY DOES SOMETHING REL
          HOW OFTEN R ATTENDED WORSHIP SERVICE IN PAST 12 M
                        R DOES NOT NEED RELIGION FOR GOOD V
                             GOD NOTHING TO DO HAPPENS
R BELIEVES RELIGIOUS TEACHINGS ARE TO BE OBEYED EXACTLY AS WI
                                    R PRAYS MORE THAN ONCE
                                  R ASKS GOD HELP MAKE DEC
                        WHAT IS R'S CURRENT RELIGIOUS PREFER
                            R A BORN-AGAIN EVANGELICAL CHRIS
                        IMPORTANCE OF RELIGIOUS FAITH IN DAIL
                HOW OFTEN R FELT CALM AND PEACEFUL IN PAST
                     HOW OFTEN R FELT DOWN OR BLUE IN PAST
               HOW OFTEN R HAS BEEN A HAPPY PERSON IN PAST
```

PUBID, YOUTH CASE IDENTIFICATION

HOW OFTEN R DEPRESSED IN LAST

HOW OFTEN R HAS BEEN A **NERVOUS** PERSON IN PAST HOW MANY HOURS PER WEEK DOES R WATCH **TELE** HOW MANY HOURS PER WEEK DOES R USE A **COM** CURRENTLY HAVE ACCESS TO INTE

KEY!RACE\_ETHNICITY, COMBINED RACE AND ETHN
KEY!BDATE. RS BIRTHDATE MONTH/

KEY!SEX, RS GE

Figure 3.2 Databox slice of variables sel

Initial point of departure - the databox of the selected variables, described in the Methods chapter.

This databox corresponds to the dataset dsL produced by Derive\_dsL\_from\_Extract report, given in the Appendix.

dsL<-readRDS("./Data/Derived/dsL.rds")</pre>

_						Tim	e Inva	riant	1	Ti	me V	ariar	ıt																
sample	id	sex	race	bmonth	byear	attendPR	relprefPR	relraisedPR	year	agemon	ageyear	famrel	attend	values	todo	obeyed	pray	decisions	relpref	bornagain	faith	calm	blue	happy	depressed	nervous	tv	computer	internet
sample	IU	Sex	4	9	1981	attenurk			1997	-										_					NA NA		NA		
1	1	2				/	21	21	_	190	15	NA.	NA.	NA.	NA	NA	NA	NA.	NA	NA	NA	NA	NA	NA		NA.			NA
1	1	2	4	9	1981	7	21	21	1998	206	17	NA	NA.	NA.	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		NA	NA
1	1	2	4	9		7	21	21	1999	219	18	NA	NA.	NA.	NA.	NA.	NA	NA	NA	NA	NA	NA	NA.	NA	NA	NA		NA	NA
1	1	2	4	9	1981	7	21	21	2000	231	19	NA.	1	NA.	NA.	NA.	NA	NA	NA	NA	NA	3	3	3	3	3	NA	NA	NA
1	1	2	4	9	1981	7	21	21	2001	243	20	NA	6	NA.	NA.	NA.	NA	NA	NA	NA	NA	NA.	NA.	NA	NA	NA	NA	NA	NA
1	1	2	4	9	1981	7	21	21	2002	256	21	NA	2	1	1	1	0	1	NA	NA	NA	4	2	3	2	1	2	5	NA
1	1	2	4	9	1981	7	21	21	2003	266	22	NA	1	NA.	NA.	NA.	NA	NA	NA	NA	NA	NA.	NA	NA	NA	NA	NA	NA	1
1	1	2	4	9	1981	7	21	21	2004	279	23	NA	1	NA	NA.	NA	NA	NA	NA	NA	NA	4	1	4	1	1	NA	NA	0
1	1	2	4	9	1981	7	21	21	2005	290	24	NA	1	0	1	0	0	1	21	NA	NA	NA.	NA.	NA	NA	NA	NA	NA	1
1	1	2	4	9	1981	7	21	21	2006	302	25	NA	1	NA	NA.	NA	NA	NA.	NA	NA	NA	4	1	4	1	1	NA	NA.	1
1	1	2	4	9	1981	7	21	21	2007	313	26	NA	1	NA.	NA.	NA.	NA	NA.	NA	NA	NA	NA.	NA	NA	NA	NA.	2	6	1
1	1	2	4	9		7	21	21	2888	325	27	NA	1	0	1	0	8	1	21	NA	3	3	3	3	3	3		NA	1
1	1	2	4	9		7	21	21	2009	337	28	NA.	1	NA.	NA	NA	NA	NA.	NA	NA	NA	NA	NA	NA	NA	NA.	2		1
1		2	4	9		7	21	21	2010	350	29	NA.		NA NA	NA.	NA.	NA	NA.	NA	NA.	NA.	2	2	3	3	3	1		1
		-							_					0	4							-							
1	1	2	4	9	1981	7	21	21	2011	360	29	NA	1	-	1	0	0	1	21	NA	1	NA	NA	NA	NA	NA	NA		1
1	2	1	2	7	1982	NA.	NA	NA	1997	178	14	3	NA.	NA.	NA.	NA	NA	NA	NA	NA	NA	NA.	NA	NA	NA	NA		NA	NA
1	2	1	2	7	1982	NA.	NA	NA	1998	196	16	1	NA.	NA.	NA.	NA.	NA	NA	NA	NA	NA.	NA.	NA.	NA	NA	NA	NA	NA	NA

Figure 3.3 Generic dataset used in the current study, view for one respondent

### **Labeling Factor Levels**

Review of the item reference cards shows that initially, all items were recorded on some discrete scale, either counting occasions or assigning an intiger to a category of response. However, data were saved as numerical values or intigers

```
ds<- dsL[,1:(ncol(dsL)/2)]# selects the first half of variables
str(ds)</pre>
```

```
'data.frame':
               134745 obs. of 30 variables:
$ sample
                    1 1 1 1 1 1 1 1 1 1 ...
$ id
                    1 1 1 1 1 1 1 1 1 1 ...
$ sex
                    2 2 2 2 2 2 2 2 2 2 . . .
             : int
$ race
                   4 4 4 4 4 4 4 4 4 ...
$ bmonth
             : int
                    9 9 9 9 9 9 9 9 9 ...
$ byear
             : int
                    7 7 7 7 7 7 7 7 7 7 7 ...
$ attendPR
             : int
$ relprefPR
                    21 21 21 21 21 21 21 21 21 21 ...
            : int
$ relraisedPR: int
                    21 21 21 21 21 21 21 21 21 21 ...
$ year
                    1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 ...
$ agemon
                   190 206 219 231 243 256 266 279 290 302 ...
                   15 17 18 19 20 21 22 23 24 25 ...
$ ageyear
             : num
$ famrel
                    NA NA NA NA NA NA NA NA NA ...
             : num
$ attend
                   NA NA NA 1 6 2 1 1 1 1 ...
             : num
                   NA NA NA NA NA 1 NA NA O NA ...
$ values
             : num
$ todo
             : num NA NA NA NA NA 1 NA NA 1 NA ...
             : num NA NA NA NA NA 1 NA NA O NA ...
$ obeyed
$ pray
             : num NA NA NA NA NA O NA NA O NA ...
$ decisions
                   NA NA NA NA NA 1 NA NA 1 NA ...
$ relpref
             : num NA NA NA NA NA NA NA NA 21 NA ...
$ bornagain
             : num
                    NA NA NA NA NA NA NA NA NA ...
$ faith
                   NA NA NA NA NA NA NA NA NA ...
             : num
$ calm
             : num
                    NA NA NA 3 NA 4 NA 4 NA 4 \dots
$ blue
                   NA NA NA 3 NA 2 NA 1 NA 1 ...
             : num
$ happy
                   NA NA NA 3 NA 3 NA 4 NA 4 ...
             : num
                   NA NA NA 3 NA 2 NA 1 NA 1 ...
$ depressed
             : num
                   NA NA NA 3 NA 1 NA 1 NA 1 ...
$ nervous
             : num
$ tv
                    NA NA NA NA NA 2 NA NA NA NA ...
             : num
$ computer
             : num
                   NA NA NA NA NA S NA NA NA NA ...
                   NA NA NA NA NA 1 O 1 1 ...
$ internet
             : num
```

LabelingFactorLevels.R sourced at the end of Derive\_dsL\_from\_Extract matches numeric values with response labels from the questionnaire and adds to dsL copies of the variables, saved as labeled factors. For estimations routines such as Ime4 or

graphing functions such as ggplot, the data type (string, numeric, factor) is a meaningful input, so a quick access to both formats frequently proves useful. It is convenient to think that **dsL** contains only

```
ncol(dsL)/2
```

[1] 30

variables, but each of them has a double, a labeled factor.

#### str(dsL)

```
'data.frame':
               134745 obs. of 60 variables:
$ sample
              : int 1 1 1 1 1 1 1 1 1 1 ...
$ id
              : int
                    1 1 1 1 1 1 1 1 1 1 ...
$ sex
              : int
                     2 2 2 2 2 2 2 2 2 2 ...
$ race
              : int
                     4 4 4 4 4 4 4 4 4 ...
$ bmonth
              : int 999999999...
                    $ byear
              : int
$ attendPR
              : int
                     7 7 7 7 7 7 7 7 7 7 ...
                     21 21 21 21 21 21 21 21 21 21 ...
$ relprefPR
              : int
$ relraisedPR : int 21 21 21 21 21 21 21 21 21 21 ...
$ year
              : int 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 ...
                    190 206 219 231 243 256 266 279 290 302 ...
$ agemon
              : num
              : num 15 17 18 19 20 21 22 23 24 25 ...
$ ageyear
$ famrel
              : num NA NA NA NA NA NA NA NA NA ...
$ attend
              : num
                     NA NA NA 1 6 2 1 1 1 1 ...
$ values
              : num
                     NA NA NA NA 1 NA NA 0 NA ...
                     NA NA NA NA NA 1 NA NA 1 NA ...
$ todo
              : num
$ obeyed
              : num
                     NA NA NA NA 1 NA NA 0 NA ...
                     NA NA NA NA O NA NA O NA ...
$ pray
              : num
$ decisions
                     NA NA NA NA 1 NA NA 1 NA ...
              : num
$ relpref
                     NA NA NA NA NA NA NA 21 NA ...
              : num
$ bornagain
                     NA NA NA NA NA NA NA NA NA ...
              : num
$ faith
                     NA NA NA NA NA NA NA NA NA ...
              : num
$ calm
              : num
                     NA NA NA 3 NA 4 NA 4 NA 4 ...
$ blue
              : num NA NA NA 3 NA 2 NA 1 NA 1 ...
$ happy
              : num NA NA NA 3 NA 3 NA 4 NA 4 ...
              : num NA NA NA 3 NA 2 NA 1 NA 1 ...
$ depressed
              : num NA NA NA 3 NA 1 NA 1 NA 1 ...
$ nervous
$ tv
              : num
                     NA NA NA NA NA 2 NA NA NA NA ...
$ computer
                     NA NA NA NA NA 5 NA NA NA NA ...
              : num
              : num NA NA NA NA NA 1 0 1 1 ...
  internet
$ sampleF
              : Ord.factor w/ 2 levels "Cross-Sectional" < ..: 1 1 1 1 1 1 1 1 1 1 1 ...
              : Factor w/ 8983 levels "1", "2", "3", "4", ...: 1 1 1 1 1 1 1 1 1 1 1 ...
$ idF
              : Ord.factor w/ 3 levels "Male"<"Female"<...: 2 2 2 2 2 2 2 2 2 2 ...
$ sexF
$ raceF
              : Ord.factor w/ 4 levels "Black"<"Hispanic"<...: 4 4 4 4 4 4 4 4 4 4 ...
              : Ord.factor w/ 12 levels "Jan"<"Feb"<"Mar"<..: 9 9 9 9 9 9 9 9 9 ...
$ bmonthF
$ byearF
              : Factor w/ 5 levels "1980", "1981", ...: 2 2 2 2 2 2 2 2 2 2 ...
              : Ord.factor w/ 8 levels "Never"<"Once or Twice"<...: 7 7 7 7 7 7 7 7 7 7 7 7 ...
$ attendPRF
$ relprefPRF : Ord.factor w/ 33 levels "Catholic"<"Baptist"<..: 21 21 21 21 21 21 21 21 21 21 21 ...
$ relraisedPRF: Ord.factor w/ 33 levels "Catholic"<"Baptist"<...: 21 21 21 21 21 21 21 21 21 21 21 ...
              : Factor w/ 15 levels "1997", "1998", ...: 1 2 3 4 5 6 7 8 9 10 ...
$ yearF
            : Factor w/ 244 levels "146","147","148",...: 45 61 74 86 98 111 121 134 145 157 ...
$ agemonF
```

```
: Factor w/ 21 levels "12", "13", "14", ...: 4 6 7 8 9 10 11 12 13 14 ...
$ famrelF
           : Factor w/ 8 levels "O","1","2","3",...: NA ...
           : Ord.factor w/ 8 levels "Never"<"Once or Twice"<..: NA NA NA 1 6 2 1 1 1 1 ...
$ attendF
$ valuesF
          : Ord.factor w/ 2 levels "FALSE/less Religious" < .. : NA NA NA NA NA 2 NA NA 1 NA ...
$ todoF
          : Ord.factor w/ 2 levels "FALSE/less Religious" < ...: NA NA NA NA NA 2 NA NA 2 NA ...
$ obeyedF
          : Ord.factor w/ 2 levels "FALSE/less Religious" < ... NA NA NA NA NA 2 NA NA 1 NA ...
$ pravF
         : Ord.factor w/ 2 levels "FALSE/less Religious" <...: NA NA NA NA NA 1 NA NA 1 NA ...
$ decisionsF : Ord.factor w/ 2 levels "FALSE/less Religious" < ...: NA NA NA NA NA NA 2 NA 2 NA ...
: Ord.factor w/ 4 levels "All of the time"<...: NA \dots
$ calmF
           : Ord.factor w/ 4 levels "All of the time"<...: NA \dots
$ blueF
           : Ord.factor w/ 4 levels "All of the time"<... NA ...
$ happyF
: Ord.factor w/ 4 levels "All of the time"<...: NA ...
$ nervousF
$ tvF
            : Ord.factor w/ 6 levels "less than 2"<...: NA NA NA NA NA 2 NA NA NA NA NA ...
          : Ord.factor w/ 6 levels "None" < "less than 1" < ... NA NA NA NA NA 5 NA NA NA NA NA ...
$ computerF
           : Ord.factor w/ 2 levels "No"<"Yes": NA NA NA NA NA NA 2 1 2 2 ...
$ internetF
```

This give a certain flexibity to assemble needed dataset quickly and have access to factor labels. One can alternate between the raw metric and labeled factor by adding "F" suffix to the end of the variable name:

```
selectCols<-c("year","id","byear","attend","attendF") # select the columns with these names
ds<-dsL[,selectCols] # select all rows for the columns listed selectCols
print(ds[ds$id==1,]) # print all available data for respondent with ID 1</pre>
```

	year	id	byear	attend	$\mathtt{attendF}$
1	1997	1	1981	NA	<na></na>
2	1998	1	1981	NA	<na></na>
3	1999	1	1981	NA	<na></na>
4	2000	1	1981	1	Never
5	2001	1	1981	6	About once/week
6	2002	1	1981	2	Once or Twice
7	2003	1	1981	1	Never
8	2004	1	1981	1	Never
9	2005	1	1981	1	Never
10	2006	1	1981	1	Never
11	2007	1	1981	1	Never
12	2008	1	1981	1	Never
13	2009	1	1981	1	Never
14	2010	1	1981	1	Never
15	2011	1	1981	1	Never

Having quick access to factor labels will be especially useful during graph production. For the grammer rules of operations with relevant data see Data Manipulation Guide.

### Time metrics: Age, Period, Cohort

NLSY97 sample includes individuals from five cohorts, born between 1980 and 1984. The following graphics shows how birth cohort, age of respondents, and round of observation are related in NSLY97.

Wide	Age	in yea	rs																
age	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Born in 1980					1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
1981				1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	
1982			1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011		
1983	1	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011			
1984	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011				
															Wave				
Wide wave	Waves					2002	2002	2004	2005	2006	2007	2000	2000	2010	2011				
Born in 1980	17	18	1999	2000	21	22	23	24	25	26	27	28	29	30	31	1			
1981	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30				
1982	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29				
1983	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28				
1984	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27				
															Age				
	₋ong vave		rn in							ı	ong	ј Во	rn in						
	יעמער	100	00 10	101 1	002	1002	100/	1		_				201 1	1022	1022	100	1	
		13.				1983	1984				age	2 19	80 19	981 1	1982	1983			
Wave:	199	97 1	7	16	15	14	13		A	ge yea	age	2 19 3		981 1	1982	1983	199	7	
		97 1	.7						A		age ars 1	19 3 4			1982 1997		199 199	7	
	199 199	97 1 98 1 99 1	.7 .8 .9	16 17	15 16	14 15	13 14		A		age ars 1	19 3 4 5	80 19			1997	199 199 199	7 8 9	
	199 199 199	97 1 98 1 99 1	.7 .8 .9	16 17 18	15 16 17	14 15 16	13 14 15		A		age ars 1 1	19 3 4 5 6	80 19	.997	1997	1997 1998	199 199 199 200	7 8 9	
	199 199 199 200	97 1 98 1 99 1 00 2	.7 .8 .9 .9	16 17 18 19	15 16 17 18	14 15 16 17	13 14 15 16		A		age ers 1 1 1 1 1 1	19.3 3 4 5 6 6 7 19.8 19.8	80 19 1 1 1 1 1 1 1 1 1 1 1 1 1 9 1 1 1 1	.997 .998 .999	1997 1998 1999 2000	1997 1998 1999 2000 2001	199 199 199 200 200 200	7 8 9 0 1 2	
	199 199 199 200 200 200 200	97 1 98 1 99 1 00 2 01 2 02 2 03 2	.7 .8 .9 .9 .0 .11	16	15 16 17 18 19 20 21	14 15 16 17 18 19 20	13 14 15 16 17 18 19		A		age ars 1 1 1 1 1 1	19/3 3 4 4 5 5 6 6 7 19 8 19 9 19	1997 1998 1999 2	997 998 999	1997 1998 1999 2000 2001	1997 1998 1999 2000 2001 2002	199 199 199 200 200 200 200	7 8 9 0 1 2 3	
	199 199 200 200 200 200 200	97 1 98 1 99 1 00 2 01 2 02 2 03 2 04 2	8	16	15 16 17 18 19 20 21 22	14 15 16 17 18 19 20 21	13 14 15 16 17 18 19 20		A		age ars 1 1 1 1 1 1 2	19/3 4 5 6 7 19/8 8 19/9 19/0 20/0	1997 1 998 1 999 2	.997 .998 .999 .000	1997 1998 1999 2000 2001 2002	1997 1998 1999 2000 2001 2002 2003	199 199 199 200 200 200 200 200	7 8 9 0 1 1 2 3 4	
	199 199 200 200 200 200 200 200 200	97 1 98 1 99 1 00 2 01 2 02 2 03 2 04 2	.7 8 .9 .9 .0 .11 .22 .23 .24	16 17 18 19 20 21 22 23 24	15 16 17 18 19 20 21 22 23	14 15 16 17 18 19 20 21 22	13 14 15 16 17 18 19 20 21		A		age ars 1 1 1 1 1 2 2	199 3 4 5 6 7 199 8 199 199 190 200 1 200	1997 1 1998 1 1999 2 1999 2 1900 2	997 998 999 000 001	1997 1998 1999 2000 2001 2002 2003	1997 1998 1999 2000 2001 2002 2003 2004	199 199 199 200 200 200 200 200 200	7 8 9 0 1 1 2 3 4	
	199 199 200 200 200 200 200 200 200	97 1 98 1 99 1 00 2 01 2 02 2 03 2 04 2 06 2	.7 .8 .9 .0 .0 .11 .22 .3 .44 .25	16 17 18 19 20 21 22 23 24 25	15 16 17 18 19 20 21 22 23 24	14 15 16 17 18 19 20 21 22 23	13 14 15 16 17 18 19 20 21 22		A		age 1 1 1 1 1 1 2 2 2 2	199 3 4 5 5 6 7 199 8 199 190 200 11 200 20 20 20 20 20 20 20 20 20 20 20 20	1997 1998 1999 2000 2001 2002 2	997 998 999 000 001 002	1997 1998 1999 2000 2001 2002 2003 2004	1997 1998 1999 2000 2001 2002 2003 2004 2005	199 199 200 200 200 200 200 200 200 200	7 8 9 0 1 1 2 3 4 5 6	
	199 199 200 200 200 200 200 200 200 200 200	97 1 98 1 99 1 90 2 01 2 02 2 04 2 05 2 06 2	.7	16	15 16 17 18 19 20 21 22 23 24 25	14 15 16 17 18 19 20 21 22 23 24	13 14 15 16 17 18 19 20 21 22 23		A		age 1 1 1 1 1 1 1 2 2 2 2 2	19 3 4 5 6 7 19 8 19 9 19 0 20 1 20 2 20 3 3 20	997 1 998 1 999 2 900 2 900 2 900 2	.997 .998 .999 .000 .001 .002 .003	1997 1998 1999 2000 2001 2002 2003 2004 2005	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006	199 199 200 200 200 200 200 200 200 200 200 2	7 8 9 0 11 2 3 4 4 5 6	
	199 199 200 200 200 200 200 200 200 200 200 2	97 1 98 1 99 1 00 2 01 2 02 2 03 2 04 2 05 2 06 2	.7	16	15 16 17 18 19 20 21 22 23 24 25 26	14 15 16 17 18 19 20 21 22 23 24 25	13 14 15 16 17 18 19 20 21 22 23 24		A		age 1 1 1 1 1 1 1 2 2 2 2 2 2 2	199 3 4 5 6 7 199 8 199 0 200 1 200 2 200 3 200 4 200	1997 1 1998 1 1999 2 1999 2 1990 2 1900 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	997 998 999 000 001 002 003 004	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007	199 199 200 200 200 200 200 200 200 200 200 2	7 8 9 0 1 1 2 3 4 4 5 6	
	199 199 200 200 200 200 200 200 200 200 200 2	257 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	.7	16	15 16 17 18 19 20 21 22 23 24 25	14 15 16 17 18 19 20 21 22 23 24	13 14 15 16 17 18 19 20 21 22 23		A		age 1 1 1 1 1 1 1 2 2 2 2 2 2 2	19. 19. 19. 19. 19. 19. 19. 19. 19. 19.	19997 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	997 998 999 000 001 002 003 004 005	1997 1998 1999 2000 2001 2002 2003 2004 2005	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006	199 199 200 200 200 200 200 200 200 200 200 2	7 8 9 0 1 1 2 3 3 4 4 5 6 6 7 8 9	
	199 199 200 200 200 200 200 200 200 200 200 2	257 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	77 88 8 99 900 121 122 123 133 144 1455 166 1677 177 178 188 199 1900 1900 1900 1900 1900 1900	16	15 16 17 18 19 20 21 22 23 24 25 26 27	14 15 16 17 18 19 20 21 22 23 24 25 26	13 14 15 16 17 18 19 20 21 22 23 24 25		A		age ars 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2	19 19 19 19 19 19 19 19 19 19 19 19 19 1	19997 1998 1999 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	9997 9998 9999 0000 0001 0002 0003 0004 0005 0006 0007	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008	1999 1999 2000 2000 2000 2000 2000 2000	7 8 9 0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9	'ave
	199 199 200 200 200 200 200 200 200 200 200 2	257 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	77 88 8 99 900 121 122 123 133 144 1455 166 1677 177 178 188 199 1900 1900 1900 1900 1900 1900	16	15 16 17 18 19 20 21 22 23 24 25 26 27 28	14 15 16 17 18 19 20 21 22 23 24 25 26 27	133 144 155 166 177 188 199 200 211 222 233 244 255 26				age	19 19 19 19 19 19 19 19 19 19 19 19 19 1	1999 1999 2 2000 2	9997 9998 9999 0000 0001 0002 0003 0004 0005 0006 0007	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009	1999 1999 2000 2000 2000 2000 2000 2000	7 8 9 0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9	'ave
	199 199 200 200 200 200 200 200 200 200 200 2	257 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	77 88 8 99 900 121 122 123 133 144 1455 166 1677 177 178 188 199 1900 1900 1900 1900 1900 1900	16	15 16 17 18 19 20 21 22 23 24 25 26 27 28	14 15 16 17 18 19 20 21 22 23 24 25 26 27	133 144 155 166 177 188 199 200 211 222 233 244 255 26				age ars 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2	19 19 19 19 19 19 19 19 19 19 19 19 19 1	19997 1 19997 1 19999 2 19999 2 19990 2 19900	997 998 999 000 001 002 003 004 005 006 007 008 009	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010	1999 1999 2000 2000 2000 2000 2000 2000	7 8 9 0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9	'ave
	199 199 200 200 200 200 200 200 200 200 200 2	257 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	77 88 8 99 900 121 122 123 133 144 1455 166 1677 177 178 188 199 1900 1900 1900 1900 1900 1900	16	15 16 17 18 19 20 21 22 23 24 25 26 27 28	14 15 16 17 18 19 20 21 22 23 24 25 26 27	133 144 155 166 177 188 199 200 211 222 233 244 255 26				age ars 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2	19 19 19 19 19 19 19 19 19 19 19 19 19 1	19997 1 19997 1 19999 2 19999 2 19990 2 19900	9997 9998 9999 0000 0001 0002 0003 0004 0005 0006 0007 0008	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010	1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010	1999 1999 2000 2000 2000 2000 2000 2000	7 8 9 0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9	'ave

Figure 3.1 Overview of the age-period-cohort structure of NLSY97

NSLY97 contains static and dynamic indicators of age. Variables byear and bmonth were recorded once in 1997 (static) and contain respondents' birth year and birth month respectively. Two age variables were recorded continuously at each interview (dynamic): age at the time of the interview in months (agemon) and in years (ageyear).

```
ds<- dsL %>% dplyr::filter(id==25, year %in% c(1997:2011)) %>% dplyr::select(id,bmonthF,byear,year, age
mutate (age = (year-byear+1), ageD = agemon/12)
print(ds)
```

```
id bmonthF byear year agemon ageyear age ageD
   25
               1983 1997
                                          15 13.92
1
          Mar
                             167
                                      13
2
   25
          Mar
               1983 1998
                             188
                                      15
                                          16 15.67
3
  25
          Mar
               1983 1999
                             201
                                      16
                                          17 16.75
4
  25
          Mar
               1983 2000
                             214
                                      17
                                          18 17.83
  25
               1983 2001
                             226
                                      18 19 18.83
5
          Mar
6
   25
               1983 2002
                             236
                                      19
                                          20 19.67
          Mar
7
  25
          Mar
               1983 2003
                             254
                                      21
                                          21 21.17
                                          22 21.75
8
  25
          Mar
              1983 2004
                             261
                                      21
9
   25
               1983 2005
                                      22
                                          23 22.67
          Mar
                             272
10 25
          Mar
              1983 2006
                             284
                                      23 24 23.67
               1983 2007
                                      24 25 24.58
11 25
          Mar
                             295
12 25
          Mar
               1983 2008
                             307
                                      25 26 25.58
13 25
          Mar
               1983 2009
                             319
                                      26
                                         27 26.58
14 25
          Mar
              1983 2010
                             332
                                      27
                                          28 27.67
15 25
          Mar
               1983 2011
                             342
                                      28
                                          29 28.50
```

The variable ageyear records the full number of years a respondent reached at the time of the interview. Due to difficulties of administering the survey, time intervals between the waves could differ. For example, for one person id = 25 the age was recorded as 21 years for both 2003 and 2004 (see ageyear). However, when you examine age in months (agemon) you can see this rounding issue disappears, once a more precise scale is used. To avoid this potentially confusing peculiarity, age in years will be calculated as age = year  $\hat{a} \in \text{"byear} + 1$  or as (ageD = agemon/12). The suffix D in ageD refers to the fact that it was calculated from a dynamic age indicator.