# Introduction to Mixture and Group-Based Trajectory Models:

## Dataset Used for Exercises

## Full resource and contact: https://www.ncrm.ac.uk/resources/online/all/?id=20826

The dataset provided with the exercises is called:

**NLS\_extract\_pb00\_08.dat**

This is an extract of data from the NLSY79 Child Survey, see: <https://www.nlsinfo.org/content/cohorts/nlsy79-children/intro-to-the-sample/nlsy79-childyoung-adult-sample-introduction>.

This survey included children of the NLSY79 study female respondents. These children were assessed and interviewed every two years from 1986 through 2014. The assessments measure cognitive ability, temperament, motor and social development, behaviour problems, and self-competence of the children as well as the quality of their home environment.

In the extract from this dataset, I focused on raw scores of problem behaviour collected every two years between 2000 and 2008, together with other participants’ characteristics and assessments.

The participants in the dataset are N = 3,076 who had provided *at least two* valid assessments of problem behaviour between 2000 and 2008.

The main variables of interest in this extract are the following:

|  |
| --- |
| cpubid\_xrnd: The child ID in the study. |
| behpr00: Raw behaviour problems index: total raw score (2000) |
| behpr02: Raw behaviour problems index: total raw score (2002) |
| behpr04: Raw behaviour problems index: total raw score (2004) |
| behpr06: Raw behaviour problems index: total raw score (2006) |
| behpr08: Raw behaviour problems index: total raw score (2008) |
| rbehpr00: Behaviour problems index 2000: Recoded into 8 ordered categorical scores |
| rbehpr02: Behaviour problems index 2002: Recoded into 8 ordered categorical scores |
| rbehpr04: Behaviour problems index 2004: Recoded into 8 ordered categorical scores |
| rbehpr06: Behaviour problems index 2006: Recoded into 8 ordered categorical scores |
| rbehpr08: Behaviour problems index 2008: Recoded into 8 ordered categorical scores |
| rbehpr00: Behaviour problems index 2000: Recoded into 8 ordered categorical scores |
| agech00: Age (in months) of child in 2000 |
| agech02: Age (in months) of child in 2002 |
| agech04: Age (in months) of child in 2004 |
| agech06: Age (in months) of child in 2006 |
| agech08: Age (in months) of child in 2008 |
| ageyr00: Age (in years) of child in 2000; Age is centred at 5 years |
| ageyr02: Age (in years) of child in 2002; Age is centred at 5 years |
| ageyr04: Age (in years) of child in 2004; Age is centred at 5 years |
| ageyr06: Age (in years) of child in 2006; Age is centred at 5 years |
| ageyr08: Age (in years) of child in 2008; Age is centred at 5 years |
| male: A dummy variable to indicate whether the child was reported of being male rather than other |
| bla\_his: A dummy variable to indicate whether the child ethnicity was indicated as being Black or Hispanic, rather than other. |
| mathz2010: Standardized Piat Math total score in 2010 |
| recogz2010: Standardized Piat Reading Recognition total score in 2010 |

The raw scores of variables **behpr00** to **behpr08** were recoded in variables **rbehpr00** to **rbehpr08** to obtain ordered categorical scores ranging from 0 to 7. The rules used for categorising were as follow:

|  |  |
| --- | --- |
| behpr\*\* Raw Score | rbehpr\*\* Categorical Score |
| 0 | 0 |
| 1 to 24 | 1 |
| 25 to 49 | 2 |
| 50 to 74 | 3 |
| 75 to 99 | 4 |
| 100 to 124 | 5 |
| 125 to 174 | 6 |
| 174 to 280 | 7 |