



## MIDUS 3 Project 2:

### National Study of *Daily Experiences*

#### README

---

Address correspondence to David M. Almeida, Department of Human Development and Family Studies, The Pennsylvania State University, 403 Biobehavioral Health Building, University Park, Pennsylvania, 16802. Email: [dalmeida@psu.edu](mailto:dalmeida@psu.edu). This project was supported by grants from the National Institute on Aging (P01-AG020166 and R01-AG019239) to conduct a longitudinal follow-up of the MIDUS (Midlife in the U.S.) investigation. The original study was supported by the John D. and Catherine T. MacArthur Foundation Research Network on Successful Midlife Development.

**README file for**  
**MIDUS 3 Project 2**  
**National Study of Daily Experiences**

**NOTE: Please read through this document carefully  
prior to using the data and documentation.**

The purpose of this memo is to provide basic information about the MIDUS 3 Project 2 (also known as National Study of Daily Experiences or NSDE) data and documentation files.

**1. What Data File Is Available?**

The M3P2 (NSDE 3) dataset: *M3\_P2\_DailyDiary\_N1236\_20220323.sav*

(For details about MIDUS 3 Project 2, see item 3 below.)

**2. What is the Structure of the MIDUS 3, Project 2 Dataset?**

The dataset is a “stacked” or “person-day” dataset (see below) comprised of 8 days of diary data resulting in 618 variables from 1,236 participants. The dataset contains cases from the Core National and Milwaukee samples. The variable SAMPLMAJ provides more detail on Core subsamples.

**The Person-Day Dataset**

The NSDE daily data file is structured as a “person-day” dataset such that each row of data corresponds to a single day for a given individual. Data regarding days are “nested” within individuals over time. Micro-level information (data collected on a given day) comes from a macro-unit, in this case a single individual. In contrast to typical multivariate datasets, where each row of values corresponds to one single individual and that individual’s values for some set of variables, each row of a person-day dataset corresponds to an individual’s values for some set of variables on that day. Figures 1 and 2 depict the traditional multivariate and person-day datasets, respectively. Both represent data from 5 participants, actually, the same 5 participants (note the *same* IDs). The figures below depict data from two components of a study.

First, Figure 1 shows data about participants' age, gender, and depression scores (CESD, The Center for Epidemiologic Studies Depression Scale) taken at the baseline assessment (i.e., MIDUS 1 Project 1). These data can be seen as characteristics that vary across individuals in the study (traditional individual differences research).

*Figure 1.* Traditional multivariate dataset.

| ID  | Age | Gender | CESD |
|-----|-----|--------|------|
| 101 | 45  | M      | 4    |
| 102 | 86  | F      | 7    |
| 103 | 37  | F      | 11   |
| 104 | 72  | M      | 8    |
| 105 | 66  | M      | 18   |

Figure 2 depicts data collected from the same participants assessed on three consecutive days. Participants were asked whether they experienced a stressor on that day (variable "Any Stress"; Yes = 1, No = 2), as well as their negative affect (Neg. Affect) over the past 24 hours (sum of NA items). The day of assessment is indicated by the variable "DAY".

*Figure 2.* Person-day dataset (3 days of assessment).

| ID  | Day | Any Stress | Neg. Affect |
|-----|-----|------------|-------------|
| 101 | 1   | 1          | 7           |
| 101 | 2   | 2          | 5           |
| 101 | 3   | 1          | 7           |
| 102 | 1   | 1          | 11          |
| 102 | 2   | 1          | 15          |
| 102 | 3   | 2          | 10          |
| 103 | 1   | 1          | 9           |
| 103 | 2   | 1          | 9           |
| 103 | 3   | 1          | 8           |
| 104 | 1   | 2          | 5           |
| 104 | 2   | 2          | 6           |
| 104 | 3   | 1          | 9           |
| 105 | 1   | 2          | 5           |
| 105 | 2   | 2          | 5           |
| 105 | 3   | 2          | 6           |

Figure 2 shows participant 101 experienced a stressful event on the first and third days that they were assessed, whereas participant 105 did not report experiencing a stressor on any of the days. Figure 2 also shows that across all 5 participants, their negative affect varies from one day to the next. Figure 2 displays quite clearly the concept of the "nested" data structure. Notice that there are multiple observations for each individual, and these observations are organized by ID **and** DAY of assessment.

## Linking Traditional Multivariate and Person-Day Datasets

It may be of interest to link the data from the baseline assessment to the daily assessments to answer certain research questions (e.g., Are age and depressive symptomatology related to the likelihood of experiencing a stressor on any given day?). To answer these questions, we would need to merge the two datasets into one. Let us assume that the data from the traditional multivariate dataset (baseline variables, including age, gender, and CESD scores) are in a dataset called “baseline”, and the data from the daily component of the study are contained in a dataset called “daily”. SAS and SPSS codes to merge these two datasets into one are shown below:

### SAS Code

**Data** all; \*Create/name a dataset;  
**Merge** baseline daily; \*Tells SAS that the new dataset 'all' is going to be a combination of the 'baseline' and 'daily' datasets;  
**By** ID; \*Tells SAS that 'ID' is the relevant variable by which the data should be linked;  
**Run;**

### SPSS Code

```
MATCH FILES /FILE=*  
  /TABLE='C:\NSDE\aggr2.sav'  
  /BY ID  
EXECUTE.
```

When these commands are executed, the resulting dataset should look similar to Figure 3.

*Figure 3. Combined multivariate (baseline) and person-day (daily) datasets.*

| ID  | Day | Any Stress | Neg. Affect | Age | Gender | CESD |
|-----|-----|------------|-------------|-----|--------|------|
| 101 | 1   | 1          | 7           | 45  | M      | 4    |
| 101 | 2   | 2          | 5           | 45  | M      | 4    |
| 101 | 3   | 1          | 7           | 45  | M      | 4    |
| 102 | 1   | 1          | 11          | 86  | F      | 7    |
| 102 | 2   | 1          | 15          | 86  | F      | 7    |
| 102 | 3   | 2          | 10          | 86  | F      | 7    |
| 103 | 1   | 1          | 9           | 37  | F      | 11   |
| 103 | 2   | 1          | 9           | 37  | F      | 11   |
| 103 | 3   | 1          | 8           | 37  | F      | 11   |
| 104 | 1   | 2          | 5           | 72  | M      | 8    |
| 104 | 2   | 2          | 6           | 72  | M      | 8    |
| 104 | 3   | 1          | 9           | 72  | M      | 8    |
| 105 | 1   | 2          | 5           | 66  | M      | 18   |
| 105 | 2   | 2          | 5           | 66  | M      | 18   |
| 105 | 3   | 2          | 6           | 66  | M      | 18   |

Figure 3 shows that age, gender, and CESD variables from the “baseline” dataset have been appended to the “daily” dataset. Furthermore, notice that each participant’s age, gender, and CESD score have been appended to EACH row. Because age, gender, and CESD only were assessed once, each variable has a single value for each participant, and that value is included at each observation (“DAY”) for each participant (“ID”). Although age, gender, and CESD only were assessed once, the fact that they appear at each observation after merging the datasets is normal. Because these measures were not time-varying (i.e., collected more than once), in the context of this example, we would not expect them to have differing values across DAYS for any given participant. If a value did vary from day to day (say CESD for participant 103 was 11, 14, 11, for days 1 to 3, respectively), something went wrong in the data merging process.

### **3. What additional files are available?**

- *M3\_P2\_DailyDiaryInstrument\_20220310.docx*. Describes the Daily Diary interview.
- *M3\_P2\_StudyDescription\_20220310.docx*. Provides description of the study, sample, background and procedures.
- *M3\_P2\_Saliva\_20220310.docx*. Describes the procedures used to collect and assay saliva samples, and the variables produced.
- *M3\_P2 Scales\_20220310.docx*. Documents construction of scale scores.

### **4. ID Systems**

A 5-digit respondent identification variable (M2ID) has been created for cases in the MIDUS Core sample. This ID system is implemented to help maintain confidentiality of respondents and can be used throughout to merge and link the publicly available MIDUS Core sample datasets.