### Predictor

* **• Basic demographics (age, gender, education) demo.damon**
* **• Duration of sobriety prior to treatment demo.damon**
* **Specific SUDs (e.g., alcohol use disorder) SUDdiagnosis.damon**
* **Social Support (MSPSS) mspss.damon**
* Substance Use History,  **subuse.damon**
* **AA/NA Affiliation aana.damon**
* Stressful Life Experiences (LEC-5) **lec.damon**

**\***The structure of this dataframe is really bonkers. You probably want to use the variable “toyou\_total” which is a sum of event types that the patients endorsed as having happened to them. We sometimes also use “toyou\_wit\_total”, which is a similar sum score, but includes events that have happened to the participant AND events that the participant has witnessed.

* Childhood Experiences (ACE) **aces.damon**
* **Spiritual Experiences (Brief R-COPE) r\_cope.damon**

### Things to predict

* Quality of Life (WHOQOL-BREF): Evaluates general, physical, psychological health, social relationships, and environment.

**QOL.damon**

* **Commitment to Sobriety (CSS-5): Assesses dedication towards sobriety. change.damon**
* **Alcohol/Drug Craving (PACS):** Measures the frequency and intensity of cravings.

**craving.damon**

* • Treatment dropout

**demo.damon (dropout\_yn)**

* • Length of stay in treatment

**demo.damon (days\_in\_tx\_clean)**

* Impression of Change (PGIC)

**pgic.damon**

Things to be aware of:

* There are three data collection points: baseline (day 1), followup (day 30), discharge (last day of Tx). These are denoted by values of 0/1/2 in the “redcap\_event\_number” column. Some patients have more than three rows of data because they have returned to treatment multiple times, and thus have data from multiple baselines. I would suggest removing and rows that have values beginning with “readmission” in the column “redcap\_event\_name”. This will remove all readmission data.
* Some variables (like the dropout variable) are only populated on the first row for each participant, and may need to be filled downwards
* The key to merging the various datasets is merging by “record\_id” AND “redcap\_event\_name”. These columns should be included in every dataset (if not let me know!)