Package ‘mosaicData’

November 5, 2023

Type Package
Title Project MOSAIC Data Sets
Version 0.20.4
Depends R (>= 4.1.0)
Suggests lattice, mosaic, reshape2, ggplot2, dplyr, tidyr, ggformula
Author Randall Pruim <rpruim@calvin.edu>, Daniel Kaplan
                   <kaplan@macalester.edu>, Nicholas Horton <nhorton@amherst.edu>
Maintainer Randall Pruim <rpruim@calvin.edu>
Description Data sets from Project MOSAIC (<http://www.mosaic-web.org>) used
to teach mathematics, statistics, computation and modeling. Funded by the
NSF, Project MOSAIC is a community of educators working to tie together
aspects of quantitative work that students in science, technology,
engineering and mathematics will need in their professional lives, but
which are usually taught in isolation, if at all.
License GPL (>= 2)
LazyLoad yes
LazyData yes
RoxygenNote 7.2.3
Encoding UTF-8
URL https://github.com/ProjectMOSAIC/mosaicData
BugReports https://github.com/ProjectMOSAIC/mosaicData/issues
NeedsCompilation no
Repository CRAN
Date/Publication 2023-11-05 05:50:02 UTC

R topics documented:

  Alcohol ................................................................. 2
  Birthdays ............................................................ 3
Description

These data provide per capita alcohol consumption values for many countries in 2005 and 2008. There are also a few countries for which there are data in other years.

Usage

data(Alcohol)

Format

A data frame with 411 observations on the following variables.

country  country name
year      year
alcohol   estimated per capita alcohol consumption for adults (15+) in litres pure alcohol
Birthdays

Source

Gapminder (https://www.gapminder.org/data/)

Examples

data(Alcohol)
# There are only a few observations in years other than 2005 and 2008
subset(Alcohol, !year %in% c(2005,2008))

Birthdays


Description

A day by day record of the number of births in each US State.

Usage

data(Birthdays)

Format

A data frame with 374221 observations on the following variables.

state  state where child was born
year   year (1969-1988)
month  month (1-12)
day    day of month
date   date as a date object
wday   Day of week (ordered factor)
births number of births

See Also

Births, Births78, Births2015, BirthsSSA, BirthsCDC for data sets that are aggregated at the level of the entire country.

Examples

data(Birthdays)
if (require(mosaic)) {
  MI <- Birthdays |> filter(state == "MI")
  gf_point(births ~ date, Birthdays, data = MI)
  gf_line(births ~ date, Birthdays, data = MI, color = ~ wday)
  gf_line(births ~ date,
            data = Birthdays |> group_by(date) |> summarise(births = sum(births)))
}
Births

US Births

Description

Number of births in the United States. There are several data sets covering different date ranges and obtaining data from different sources.

Usage

   data(Births)
   data(Births78)
   data(Births2015)
   data(BirthsSSA)
   data(BirthsCDC)

Format

A data.frame with the following 8 variables.

   date  Date
   births Number of births on date (integer)
   wday  Day of week (ordered factor)
   year  Year (integer)
   month Month (integer)
   day_of_year  Day of year (integer)
   day_of_month Day of month (integer)
   day_of_week  Day of week (integer)

Details

There are some overlapping dates in the various data sets, but the number of births does not always agree due to the different sources of the data. See the examples.

Source


• Data source for BirthsSSA US Social Security Administration, as curated at https://github.com/fivethirtyeight/data/tree/master/births
Births

- Data source for BirthsCDC US Centers for Disease Control, as curated at https://github.com/fivethirtyeight/data/tree/master/births

See Also

Birthdays for a data set aggregated at the state level.

Examples

data(Births78)
data(Births2015)
data(Births)
data(BirthsSSA)
data(BirthsCDC)

# date ranges for the different data sets
lapply(
  list(Births = Births, Births78 = Births78, Biths2015 = Births2015, BirthsSSA = BirthsSSA, BirthsCDC = BirthsCDC),
  function(x) range(x$date))

range(Births78$date)
range(Births2015$date)
range(Births$date)
range(BirthsSSA$date)
range(BirthsCDC$date)

# Births and Births78 have slightly different numbers of births

if(require(ggplot2)) {
  ggplot(data = Births, aes(x = date, y = births, colour = ~ wday)) +
  stat_smooth(se = FALSE, alpha = 0.8, geom = "line")
  ggplot(data = Births, aes(x = day_of_year, y = births, colour = ~ wday)) +
  geom_point(size = 0.4, alpha = 0.5) +
  stat_smooth(se = FALSE, geom = "line", alpha = 0.6, size = 1.5)
  if (require(dplyr)) {
    ggplot(
      data = bind_cols(Births |> filter(year == 1978),
                       Births78 |> rename(births78 = births)),
      aes(x = births - births78)
    ) +
    geom_histogram(binwidth = 1)
  }
}

if(require(ggplot2)) {
  ggplot(data = Births, aes(x = date, y = births, colour = ~ wday)) +
  stat_smooth(se = FALSE, alpha = 0.8, geom = "line")
  ggplot(data = Births, aes(x = day_of_year, y = births, colour = ~ wday)) +
  geom_point(size = 0.4, alpha = 0.5) +
  stat_smooth(se = FALSE, geom = "line", alpha = 0.6, size = 1.5)
  if (require(dplyr)) {

Cards

**Standard Deck of Cards**

**Description**

A character vector with two or three character representations of each card in a standard 52-card deck.

**Usage**

```r
Cards
```

**Details**

The 2 of clubs is represented as "2C", while the 10 of diamonds is "10D".

**Examples**

```r
if (require(mosaic)) {
  deal(Cards, 13)  # bridge hand
  deal(Cards, 5)   # poker hand
  shuffle(Cards)   # shuffled deck
}
```
Description

Temperature of a mug of water as it cools

Usage

data(CoolingWater)

Format

A data frame with 222 observations of the following variables.

time  time in minutes

temp  temperature in Celsius

Details

The water was poured into a mug and a temperature probe inserted into the water with a few seconds of the pour.

Source

These data were collected Stan Wagon to help his mathematical modeling students explore Newton’s Law of Cooling and the ways that the law is really only an approximation. More about Stan: http://stanwagon.com.

Examples

data(CoolingWater)
if (require(ggformula)) {
  gf_point(temp ~ time, data = CoolingWater, alpha = 0.5)
}
Countries

Description

A data frame containing country names as used by Gapminder and the maps package to facilitate conversation between the two.

Usage

data(Countries)

Format

A data frame with 258 observations on the following variables.

- maptools region name http://mappinghacks.com/ data sets
- gapminder country name in Gapminder data sets
- maps region name in maps data sets

Details

The "countries" in the maps data include several other geographic regions (bodies of water, islands belonging to other countries, Hawaii, etc.) that are not countries. Furthermore, the maps countries do not include many of the countries that have been created since ca. 2000. The mapping is therefore many-to-many, and also includes some NAs when there is no appropriate mapping. Bodies of water in the maps data, for example, are not assigned a country in the Gapminder.

Examples

data(Countries)
subset(Countries, maps=="Yugoslavia")  # Where has Yugoslavia gone?
subset(Countries, is.na(gapminder))     # Things from maps with no Gapminder equivalent
subset(Countries, is.na(maps))          # Things from Gapminder with no maps equivalent

CPS85

Data from the 1985 Current Population Survey (CPS85)

Description

The Current Population Survey (CPS) is used to supplement census information between census years. These data consist of a random sample of persons from the CPS85, with information on wages and other characteristics of the workers, including sex, number of years of education, years of work experience, occupational status, region of residence and union membership.
Usage

data(CPS85)

Format

A data frame with 534 observations on the following variables.

wage  wage (US dollars per hour)
educ  number of years of education
race  a factor with levels NW (nonwhite) or W (white)
sex   a factor with levels F M
hispanic  a factor with levels Hisp NH
south  a factor with levels NS S
married a factor with levels Married Single
exper  number of years of work experience (inferred from age and educ)
union  a factor with levels Not Union
age   age in years
sector  a factor with levels clerical const manag manuf other prof sales service

Details

Data are from 1985. The data file is recoded from the original, which had entirely numerical codes.

Source

Data are from https://dasl.datadescription.com

References


Examples

data(CPS85)
### Dimes

**Description**

Weights of a sample of dimes.

**Usage**

```r
data(Dimes)
```

**Format**

A data frame with 30 observations on the following 2 variables.

- `mass`: mass of dime in grams
- `year`: year the dime was minted

**Details**

These data were collected on a sample taken from a large sack of dimes for the purpose of estimating the total number of dimes in the sack based on the weights of the individual dimes.

**Source**

Data were collected by Michael Stob.

---

### Galton

**Description**

In the 1880’s, Francis Galton was developing ways to quantify the heritability of traits. As part of this work, he collected data on the heights of adult children and their parents.

**Usage**

```r
data(Galton)
```
**Format**

A data frame with 898 observations on the following variables.

- **family**: a factor with levels for each family
- **father**: the father’s height (in inches)
- **mother**: the mother’s height (in inches)
- **sex**: the child’s sex: F or M
- **height**: the child’s height as an adult (in inches)
- **nkids**: the number of adult children in the family, or, at least, the number whose heights Galton recorded.

**Details**

Entries were deleted for those children whose heights were not recorded numerically by Galton, who sometimes used entries such as “tall”, “short”, “idiotic”, “deformed” and so on.

**Source**

The data were transcribed by J.A. Hanley who has published them at [http://www.medicine.mcgill.ca/epidemiology/hanley/galton/](http://www.medicine.mcgill.ca/epidemiology/hanley/galton/)

**References**


**Examples**

```r
data(Galton)
```

---

**Gestation**

*Data from the Child Health and Development Studies*

**Description**

Birth weight, date, and gestational period collected as part of the Child Health and Development Studies in 1961 and 1962. Information about the baby’s parents — age, education, height, weight, and whether the mother smoked is also recorded.

**Usage**

```r
data(Gestation)
```
Gestation

Format

A data frame with 1236 observations on the following variables.

- id: identification number
- plurality: all "single fetus" in this data set
- outcome: all "live birth" (survived at least 28 days) in this data set
- date: birth date where 1096=January 1, 1961
- gestation: length of gestation (in days)
- wt: birth weight (in ounces)
- parity: total number of previous pregnancies (including fetal deaths and still births)
- sex: "male"
- race: mother's race: "asian", "black", "mex", "mixed", or "white"
- age: mother's age in years at termination of pregnancy
- ed: mother's education
- ht: mother's height in inches to the last completed inch
- wt.1: mother's prepregnancy weight (in pounds)
- drace: father's race
- dage: father's age (in years)
- ded: father's education
- dht: father's height in inches to the last completed inch
- dwt: father's weight (in pounds)
- marital: marital status
- inc: family yearly income in $2500 increments
- smoke: does mother smoke? (never, smokes now, until current pregnancy, once did, not now)
- time: time since quitting smoking (never smoked, still smokes, during current preg, within 1 year, 1 to 2 years ago, 2 to 3 years ago, 3 to 4 years ago, 5 to 9 years ago, 10+ years ago, quit and don't know)
- number: number of cigarettes smoked per day for past and current smokers (never, 1-4, 5-9, 10-14, 15-19, 20-29, 30-39, 40-60, 60+, smoke but don't know)

Details

The data were presented by Nolan and Speed to address the question of whether there is a link between maternal smoking and the baby's health for male births.

Source

The book by Nolan and Speed describes the data in more detail and provides an Internet site for accessing them: https://www.stat.berkeley.edu/users/statlabs/

References

Examples

data(Gestation)

GoosePermits

Goose Permit Study

Description

237 hunters were each offered one of 11 cash amounts (bids) ranging from $1 to $200 in return for their goose permits. Hunters returned either their permit or the cash.

Usage

data(GoosePermits)

Format

A data.frame with 11 observations on the following 3 variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bid</td>
<td>amount offered for permit (US $) (numeric)</td>
</tr>
<tr>
<td>keep</td>
<td>number of hunters who kept the permit and returned the cash (numeric)</td>
</tr>
<tr>
<td>sell</td>
<td>number of hunters who kept the cash and returned the permit (numeric)</td>
</tr>
</tbody>
</table>

Source


Examples

data(GoosePermits)

```r
  goose.model <-
  glm( cbind(keep, sell) ~ log(bid), data = GoosePermits, family = binomial())
  if (require(ggformula)) {
    y.hat <- makeFun(goose.model)
    gf_point( (keep/(keep+sell)) ~ bid, data = GoosePermits, ylim = c(0,1.05)) |> 
    gf_fun(y.hat(b) ~ b, add = TRUE, color = "red", alpha = 0.5) 
  }
```
HeatX

Data from a heat exchanger laboratory

Description

These data were collected by engineering students at Calvin College. The apparatus consists of concentric pipes insulated from the environment so that as nearly as can be managed the only heat exchange is between the hot and cold water.

Usage

data(HeatX)

Format

A data frame with 6 observations on the following variables.

- **trial**: trial number
- **T.cold.in**: temperature (C) of the cold water as it enters the apparatus
- **T.cold.out**: temperature (C) of the cold water as it leaves the apparatus
- **m.cold**: flow rate (L/min) of the cold water
- **T.hot.in**: temperature (C) of the hot water as it enters the apparatus
- **T.hot.out**: temperature (C) of the hot water as it leaves the apparatus
- **m.hot**: flow rate (L/min) of the hot water

Examples

```r
# We can test for heat exchange with the environment by checking to see if the
# heat gained by the cold water matches the heat lost by the hot water.
C_p <- 4.182 / 60  # / 60 because measuring m in L/min
HeatX2 <-
dplyr::mutate(HeatX,
  Q.cold = m.cold * C_p * (T.cold.out - T.cold.in),
  Q.hot = m.hot * C_p * (T.hot.out - T.hot.in),
  Q.env = Q.cold + Q.hot
)
if (require(ggformula)) {
  gf_jitter("" ~ Q.env, data = HeatX2, alpha = 0.6, size = 4,
    width = 0, height = 0.1, seed = 123) |> 
  gf_labs(y = "")
}
if (require(mosaic)) {
  t.test(~Q.env, data = HeatX2)
}
```
Description

The HELP study was a clinical trial for adult inpatients recruited from a detoxification unit. Patients with no primary care physician were randomized to receive a multidisciplinary assessment and a brief motivational intervention or usual care, with the goal of linking them to primary medical care.

Usage

data(HELPfull)

Format

A data frame with 1472 observations on the following variables.

ID  Subject ID
A10  Marital Status (1=Married, 2=Remarried, 3=Widowed, 4= Separated, 5=Divorced, 6=Never Married
A11A  Do you currently have a living mother? (0=No, 1= Yes
A11B  Do you currently have a living father? (0=No, 1=Yes
A11C  Do you currently have siblings? (0=No, 1=Yes
A11D  Do you currently have a partner (0=No, 1=Yes)
A11E  Do you currently have children? (0=No, 1=Yes)
A12B_REC  Hollingshead category (recode) (0=Cat 1,2,3, 1=Cat 4,5,6, 2=Cat 7,8,9)
A12B  Hollingshead categories (1=Major profess, 2=Lesser professional, 3=Minor professional, 4=Clerical/sales, 5=Skilled manual, 6=Semi-skilled, 7=Unskilled, 8= Homemaker, 9=No occupation)
A13  Usual employment pattern in last 6 months (1=Full time, 2=Part time, 3=Student, 4=Unemployed, 5=Control envir)
A14A  Lived alone-last 6 months (0=No, 1=Yes)
A14B  Lived with a partner-last 6 months (0=No, 1=Yes
A14C  Lived with parent(s)-last 6 months (0=No, 1=Yes)
A14D  Lived with children-last 6 months (0=No, 1=Yes)
A14E  Lived with other family-last 6 months (0=No, 1=Yes
A14F  Lived with friend(s)-last 6 months (0=No, 1=Yes
A14G_T a factor with levels 1/2 WAY HOUSE 3/4 HOUSE ANCHOR INN ARMY ASSOCIATES BOARDERS BOYFRIENDS MOM CORRECTIONAL FACILIT CRACK HOUSE DEALER ENTRE FAMILIA FENWOOD GAVIN HSE GIRLFRIENDS DAUGHTE GIRLFRIENDS SON GIRLFRIENDS DAUGHTER GROUP HOME HALF-WAY HOUSE HALFWAY HOUSE HALFWAY HOUSES HALFWAY HSE HOLDING UNIT HOME BORDER HOMELESS HOMELESS SHELTER IN JAIL IN PROGRAMS INCARCERATED JAIL JAIL
HALFWAY HOUSE JAIL, SHELTER JAIL, STREET JAIL/PROGRAM JAIL/SHELTER JAILS/LANDLADY  
LANDLORD LODGING HOUSE, MERIDIAN HOUSE, NURSING HOME ON THE STREET PARTNERS MOTHER  
PARTNERS CHILD/CHILDREN, PRGRAMS PRISON PROGRAM, MTHP PROGRAM  
ROOMMATES PROGRAM, SOBER HOUSE PROGRAM-RESIDENTIAL PROGRAM/HALFWAY HOU PROGRAM/JAIL  
PROGRAM/SHELTERS PROGRAM/SHELTERS PROGRAMS SUBSTANCE PROGRAMS/SHELTER  
PROGRAMS/SHELTERS PROGRAMS/SHELTERS DE PROJECT SOAR RESIDENTIAL FACILITY RESIDENTIAL  
PROGRAM ROOMING HOUSE, ROOMING HOUSE (RELIG ROOMMATE, ROOMMATES AT TRANSIT  
RYAN HOUSE, SALVATION ARMY, SHELTER, HALFWAY HSE, SHELTER/HOTEL, SHELTER/PROGRAM  
SHELTERS, SHELTERS/HOSPITALS, SHELTERS/JAIL, SHELTERS/PROGRAMS, SHELTERS/STREETS  
SOFER HOUSE SOBER HOUSING, SOUTH BAY JAIL, STEPSOHN STREET, STREETS SUBSTANCE ABUSE  
TREATMENT PROGRAMS/VA, SHELTER  

A14G  Lived w/ other-last 6 months (0=No, 1=Yes)  
A15A  #nights in overnight shelter-last 6 months  
A15B  # nights on street-last 6 months  
A15C  # months in jail-last 6 months  
A16A  # months in overnight shelter-last 5 years  
A16B  # months on street-last 5 years  
A16C  # months in jail-last 5 years  
A17A  Received SSI – past 6 months (0=No, 1=Yes)  
A17B  Received SSDI – past 6 months (0=No, 1=Yes)  
A17C  Received AFDC – past 6 months (0=No, 1=Yes)  
A17D  Received EAEDC – past 6 months (0=No, 1=Yes)  
A17E  Received WIC – past 6 months (0=No, 1=Yes)  
A17F  Received unemployment benefits – past 6 months (0=No, 1=Yes)  
A17G  Received Workman’s Compensation – past 6 months (0=No, 1=Yes)  
A17H  Received Child Support – past 6 months (0=No, 1=Yes)  
A17I  Received other income – past 6 months (0=No, 1=Yes)  
A17I_T  a factor with levels  
DISABLED VETERAN, EBT (FOOD STAMPS), EMERGENCY FOOD STAMP, FOOD  
STAMP, FOOD STAMPS, FOOD STAMPS/VETERAN FOOD STAMPS/VETERANS INSURANCE SETTLEMENT  
PENSION, CHECK, SECTION 8, SERVICE, CONNECTED, DI, SOCIAL SECURITY, SSDI FOR SON  
SURVIVORS BENEFITS, TEMPORARY DISABILITY, VA BENEFITS, DISABILI, VA COMPENSATION  
VA, DISABILITY PENSIO, VETERAN, BENEFITS, VETERANS, SERVICES, VETERANS, AFFAIRS  
A18_REC1  Most money made in 1 year (recode) (0=$19,000 or less, 1=$20,000-$49,000, 2=$50,000  
more)  
A18_REC2  Most money made-continuous recode  
A18  Most money made in any 1 year-last 5 years (1=<5000, 2=5000-10000, 3=11000-19000,  
4=20000-29000, 5=30000-39000, 6=40000-49000, 7=50000+)  
A1  Gender (1=Male, 2=Female)  
A9  Years of education completed  
ABUSE2  Type of abuse (0=No abuse, 1=Physical only, 2=Sexual only, 3=Physical and sexual)  
ABUSE3  Type of abuse (0=No abuse, 1=Physical only, 2=Sexual +/- physical (0=No, 1=Yes
ABUSE  Abuse-physical or sexual (0=No abuse, 1=Family abuse, 2=Stranger only abuse)

AGE  Age in years

ALCOHOL  1st/2nd drug of choice=Alcohol (0=No, 1=Yes)

ALCQ_30  Total number drinks past 30 days

ALONE6M  Usually lived alone past 6 months (0=No, 1=Yes)

ALT_TRT  Alternative treatments (0=No, 1=Yes)

ANYSUBSTATUS  Used alcohol, heroin, or cocaine since leaving detox-6 months

ANY_INS  Did you have health insurance in past 6 months (0=No, 1=Yes)

ANY_UTIL  Any recent health utilization (0=No, 1=Yes)

ANY_VIS_CUMUL  Cumulative # visits to regular doctor’s office

ANY_VIS  # visits to regular doctor’s office–This time point

B10  Any physical/emotional problem interfere with social activities-last 4 weeks (1=All of the time, 2=Most of the time, 3=Some of the time, 4=A little of time, 5=None of the time)

B11A  I seem to get sick easier than other people (1=Definitely true, 2=Mostly true, 3=Don’t know, 4=Mostly false, 5=Definitely false)

B11B  I am as healthy as anybody I know (1=Definitely true, 2=Mostly true, 3=Don’t know, 4=Mostly false, 5=Definitely false)

B11C  I expect my health to get worse (1=Definitely true, 2=Mostly true, 3=Don’t know, 3=Mostly false, 5=Definitely false)

B11D  My health is excellent (1=Definitely true, 2=Mostly true, 3=Don’t know, 4=Mostly false, 5=Definitely false)

B1  In general, how is your health (1=Excellent, 2=Very Good, 3=Good, 4=Fair, 5=Poor)

B2  Compared to 1 year ago, how is your health now (1=Much better, 2=Somewhat better, 3=About the same, 4=Somewhat worse, 5=Much worse)

B3A  Does health limit you in vigorous activity (1=Limited a lot, 2=Limited a little, 3=Not limited)

B3B  Does your health limit you in moderate activity (1=Limited a lot, 2=Limited a little, 3=Not limited)

B3C  Does health limit you in lift/carry groceries (1=Limited a lot, 2=Limited a little, 3=Not limited)

B3D  Does health limit you in climb several stair flights (1=Limited a lot, 2=Limited a little, 3=Not limited)

B3E  Does health limit you in climb 1 stair flight (1=Limited a lot, 2=Limited a little, 3=Not limited)

B3F  Does health limit you in bend/kneel/stoop (1=Limited a lot, 2=Limited a little, 3=Not limited)

B3G  Does health limit you in walking >1 mile (1=Limited a lot, 2=Limited a little, 3=Not limited)

B3H  Does health limit you in walking several blocks (1=Limited a lot, 2=Limited a little, 3=Not limited)

B3I  Does health limit you in bathing/dressing self (1=Limited a lot, 2=Limited a little, 3=Not limited)

B4A  Cut down work/activity due to physical health-last 4 weeks (0=No, 1=Yes)

B4B  Accomplish less due to phys health-last 4 weeks (0=No, 1=Yes)
B4C  Lim wrk/act type due to phys health-last 4 weeks (0=No, 1=Yes)
B4D  Diff perf work due to phys health-last 4 weeks (0=No, 1=Yes)
B5A  Cut wrk/act time due to emot prbs-last 4 weeks (0=No, 1=Yes)
B5B  Accomplish ess due to emot prbs-last 4 weeks (0=No, 1=Yes)
B5C  <carefl w/wrk/act due to em prb-last 4 weeks (0=No, 1=Yes)
B6   Ext phys/em intf w/ norm soc act-last 4 weeks (1=Not al all, 2=Slightly, 3=Moderately, 4=Quite a bit, 5=Extremely)
B7   Amount of bodily pain – past 4 weeks (1=None, 2=Very mild, 3=Mild, 4=Moderate, 5=Severe, 6=Very severe)
B8   Amount of pain interfering with normal work-last 4 weeks (1=Not at all, 2=A little bit, 3=Moderately, 4=Quite a bit, 5=Extremely)
B9A  Did you feel full of pep – past 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
B9B  Have you been nervous – past 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
B9C  Felt nothing could cheer you-last 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
B9D  Have you felt calm/peaceful – past 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
B9E  Did you have a lot of energy – past 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
B9F  Did you feel downhearted – past 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
B9G  Did you feel worn out – past 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
B9H  Have you been a happy pers – past 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
B9I  Did you feel tired – past 4 weeks (1=All of the time, 2=Most of the time, 3=Good bit of the time, 4=Some of the time, 5=A little of time, 6=None of the time)
BIRTHPLC  Where born (recode) (0=USA, 1=Foreign)
BP   SF-36 pain index (0-100)
C1A  Tolf by MD had seix, epil, convuls (0=No, 1=Yes)
C1B  Told by MD had asthma, emphysema, chr lung dis (0=No, 1=Yes)
C1C  Told by MD had MI (0=No, 1=Yes)
C1D  Told by MD had CHF (0=No, 1=Yes)
C1E  Told by MD had other heart dis (req med) (0=No, 1=Yes)
C1F  Told by MD had HBP (0=No, 1=Yes)
C1G  Told by MD had chronic liver disease (0=No, 1=Yes)
C1H  Told by MD had kidney failure (0=No, 1=Yes)
C1I  Told by MD had chronic art, osteoarth (0=No, 1=Yes)
<table>
<thead>
<tr>
<th>Question</th>
<th>Response Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1J  Ever told by MD had peripheral neuropathy (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C1K  Ever told by MD had cancer (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C1L  Ever told by MD had diabetes (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C1M  Ever told by MD had stroke (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2A1 Have you ever had skin infections (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2A2 Have you had skin infections – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2B1 Have you ever had pneumonia (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2B2 Have you had pneumonia – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2C1 Have you ever had septic arthritis (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2C2 Have you had septic arthritis – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2D1 Have you ever had TB (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2D2 Have you had TB-last 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2E1 Have you ever had endocarditis (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2E2 Have you had endocarditis – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2F1 Have you ever had an ulcer (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2F2 Have you had an ulcer – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2G1 Have you ever had pancreatitis (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2G2 Have you had pancreatitis – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2H1 Ever had abdom pain req overnt hosp stay (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2H2 Abdom pain req ovrnt hosp stay-last 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2I1 Have you ever vomited blood (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2I2 Have you vomited blood – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2J1 Have you ever had hepatitis (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2J2 Have you had hepatitis – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2K1 Ever had blood clots in legs/lungs (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2K2 Blood clots in legs/lungs – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2L1 Have you ever had osteomyelitis (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2L2 Have you had osteomyelitis – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2M1 Chest pain using cocaine req ER/hosp (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2M2 Chest pain using coc req ER/hosp-last 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2N1 Have you ever had jaundice (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2N2 Have you had jaundice – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2O1 Lower back pain &gt; 3 months req med attn (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2O2 Lwr back pain &gt;3 months req med attention-last 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2P1 Ever had seizures or convulsions (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2P2 Had seizures or convulsions – past 6 months (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>C2Q1 Ever had drug/alcohol overdose requiring ER attention (0=No, 1=Yes)</td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Question</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>C2Q2</td>
<td>Drug/alcohol overdose req ER attn (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2R1</td>
<td>Have you ever had a gunshot wound (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2R2</td>
<td>Had a gunshot wound – past 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2S1</td>
<td>Have you ever had a stab wound (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2S2</td>
<td>Have you had a stab wound – past 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2T1</td>
<td>Ever had accident/falls req med attn (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2T2</td>
<td>Had accident/falls req med attn – past 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2U1</td>
<td>Ever had frac/disloc to bones/joints (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2U2</td>
<td>Fract/disloc to bones/joints – past 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2V1</td>
<td>Ever had injury from traffic accident (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2V2</td>
<td>Had injury from traffic accident – past 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2W1</td>
<td>Have you ever had a head injury (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C2W2</td>
<td>Have you had a head injury – past 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3A1</td>
<td>Have you ever had syphilis (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3A2</td>
<td># times had syphilis</td>
</tr>
<tr>
<td>C3A3</td>
<td>Have you had syphilis in last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3B1</td>
<td>Have you ever had gonorrhea (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3B2</td>
<td># times had gonorrhea</td>
</tr>
<tr>
<td>C3B3</td>
<td>Have you had gonorrhea in last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3C1</td>
<td>Have you ever had chlamydia (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3C2</td>
<td># of times had Chlamydia</td>
</tr>
<tr>
<td>C3C3</td>
<td>Have you had chlamydia in last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3E</td>
<td>Have you ever had genital herpes (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3F1</td>
<td>Have you ever had other STD’s (not HIV) (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3F2</td>
<td># of times had other STD’s (not HIV)</td>
</tr>
<tr>
<td>C3F_T</td>
<td>a factor with levels</td>
</tr>
<tr>
<td>C3G1</td>
<td>Have you ever been tested for HIV/AIDS (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3G2</td>
<td># times tested for HIV/AIDS</td>
</tr>
<tr>
<td>C3G3</td>
<td>Have you been tested for HIV/AIDS last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3G4</td>
<td>What was the result of last test (1=Positive, 2=Negative, 3=Refused, 4=Never got result, 5=Inconclusive</td>
</tr>
<tr>
<td>C3H1</td>
<td>Have you ever had PID (0=No, 1=Yes)</td>
</tr>
<tr>
<td>C3H2</td>
<td># of times had PID</td>
</tr>
</tbody>
</table>
C3H3 Have you had PID in last 6 months (0=No, 1=Yes)
C3I Have you ever had a Pap smear (0=No, 1=Yes)
C3J Have you had a Pap smear in last 3 years (0=No, 1=Yes)
C3K_M How many months pregnant
C3K Are you pregnant (0=No, 1=Yes)
CESD_CUT CES-D score > 21 y/n (0=No, 1=Yes)
C5_D CES-D score, measure of depressive symptoms, high scores are worse
CHR_6M Chronic medical conds/HIV – past 6m y/n (0=No, 1=Yes)
CHR_EVER Chronic medical conds/HIV-ever y/n (0=No, 1=Yes)
CHR_SUM Sum chronic medical conds/HIV ever
CNTRL InDUC-2L-Control score
COC_HER 1st/2nd drug of choice=cocaine or heroine (0=No, 1=Yes)
CUAD_C CUAD-Cocaine
CUAD_H CUAD-Heroin
CURPHYAB Current abuse-physical (0=No, 1=Yes)
CURPHYSEXAB Current abuse-physical or sexual (0=No abuse, 1=Physical only, 2=Sexual +/- physical)
CURSEXAB Current abuse-sexual (0=No, 1=Yes)
C_AU ASI-Composite score for alcohol use
C_DU ASI-Composite score for drug use
C_MS ASI-Composite medical status
D1 $ of times hospitalized for med probs
D2 Take prescription medication regularly for physical problem (0=No, 1=Yes)
D3_REC Any medical problems past 30d y/n (0=No, 1=Yes)
D3 # days had med probs-30 days bef detox
D4_REC Bothered by medical problems y/n (0=No, 1=Yes)
D4 How bother by med prob-30days bef detox (0=Not at all, 1=Slightly, 2=Moderately, 3=Considerably, 4=Extremely)
D5_REC Medical trmt is important y/n (0=No, 1=Yes)
D5 How import is trmt for these med probs (0=Not at all, 1=Slightly, 2=Moderately, 3=Considerably, 4=Extremely)
DAYSANYSUB time (days) from baseline to first alcohol, heroin, or cocaine since leaving detox-6m
DAYSDRINK Time (days) from baseline to first drink since leaving detox-6m
DAYSLINK Time (days) to linkage to primary care within 12 months (by administrative record)
DAYS_SINCE_BL # of days from baseline to current interview
DAYS_SINCE_PREV # of days from previous to current interview
DEAD a numeric vector
DEC_AM SOCRATES-Ambivalence-Decile
DEC_RE  SOCRATES-Recognition-Decile
DEC_TS  SOCRATES-Taking steps-Decile

DRINKSTATUS  Drank alcohol since leaving detox-6m

DRUGRISK  RAB-Drug risk total

E10A  have you been to med clinic-last 6 months (0=No, 1=Yes)
E10B1  Mental health treatment past 6m y/n (0=No, 1=Yes)
E10B1 # x visit ment health clin/prof-last 6 months
E10B2  Med clinic/private MD past 6m y/n (0=No, 1=Yes)
E10B2 # x visited med clinic/priv MD-last 6 months
E10C19  Visited private MD-last 6 months (0=No, 1=Yes)
E11A  Did you stay ovrnite/+ in hosp-last 6 months (0=No, 1=Yes)
E11B # times ovrnight/+ in hosp-last 6 months
E11C  Total # nights in hosp-last 6 months
E12A  Visited Hosp ER for med care – past 6 months (0=No, 1=Yes)
E12B # times visited hosp ER-last 6 months
E13  Tlt # visits to MDs-last 2 weeks bef detox
E14A  Recd trtmt from acupuncturist-last 6 months (0=No, 1=Yes)
E14B  Recd trtmt from chiropractor-last 6 months (0=No, 1=Yes)
E14C  Trtd by hol/herb/hom med prac-last 6 months (0=No, 1=Yes)
E14D  Recd trtmt from spirit healer-last 6 months (0=No, 1=Yes)
E14E  Have you had biofeedback-last 6 months (0=No, 1=Yes)
E14F  Have you underwent hypnosis-last 6 months (0=No, 1=Yes)
E14G  Received other treatment-last 6 months (0=No, 1=Yes)
E15A  Tried to get subst ab services-last 6 months (0=No, 1=Yes)
E15B  Always able to get subst ab servies (0=No, 1=Yes)
E15C10  My insurance didn’t cover services (0=No, 1=Yes)
E15C11  There were no beds available at the prog (0=No, 1=Yes)
E15C12  Other reason not get sub ab services (0=No, 1=Yes)
E15C1  I could not pay for services (0=No, 1=Yes)
E15C2  I did not know where to go for help (0=No, 1=Yes)
E15C3  Couldn’t get to services due to transp prob (0=No, 1=Yes)
E15C4  The offie/clinic hrs were inconvenient (0=No, 1=Yes)
E15C5  Didn’t speak/understd English well enough (0=No, 1=Yes)
E15C6  Afraid other might find out about prob (0=No, 1=Yes)
E15C7  My substance abuse interfered (0=No, 1=Yes)
E15C8  Didn’t have someone to watch my children (0=No, 1=Yes)
E15C9  I did not want to lose my job (0=No, 1=Yes)
E16A10 I do not want to lose my job (0=No, 1=Yes)
E16A11 My insurance doesn’t cover charges (0=No, 1=Yes)
E16A12 I do not feel I need a regular MD (0=No, 1=Yes)
E16A13 Other reasons don’t have regular MD (0=No, 1=Yes)
E16A14 I cannot pay for services (0=No, 1=Yes)
E16A15 I am not eligible for free care (0=No, 1=Yes)
E16A16 I do not know where to go (0=No, 1=Yes)
E16A17 Can’t get to services due to transp prob (0=No, 1=Yes)
E16A18 My insurance doesn’t cover charges (0=No, 1=Yes)
E16A19 I don’t have someone to watch my children (0=No, 1=Yes)
E16A20 Barrier to regular MD: dislike docs/system (0=No, 1=Yes)
E16A21 Barrier to regular MD: internal barriers (0=No, 1=Yes)
E16A22 Barrier to regular MD: red tape (0=No, 1=Yes)
E16A23 Barrier to regular MD: time restrictions (0=No, 1=Yes)
E16A24 I could not pay for services (0=No, 1=Yes)
E16A25 I did not know where to go for help (0=No, 1=Yes)
E16A26 I did not want to lose my job (0=No, 1=Yes)
E18C Couldn’t get to services due to transp prob (0=No, 1=Yes)
E18D The office/clinic hrs were inconvenient (0=No, 1=Yes)
E18E Afraid others might find out about prob (0=No, 1=Yes)
E18F My substance abuse interfered (0=No, 1=Yes)
E18G I did not want to lose my job (0=No, 1=Yes)
E18H My insurance didn’t cover services (0=No, 1=Yes)
E18I I do not need substance abuse services (0=No, 1=Yes)
E18J Other reason not get sub ab services (0=No, 1=Yes)
E2A Detox prog for alcohol or drug prob-last 6 months (0=No, 1=Yes)
E2B # times entered a detox prog-last 6 months
E2C # nights overnight in detox prg-last 6 months
E3A Holding unit for drug/alcohol prob-last 6 months (0=No, 1=Yes)
E3B # times in holding unity=last 6 months
E3C # total nights in holding unit-last 6 months
E4A In halfway hse/resid facil-last 6 months (0=No, 1=Yes)
E4B # times in hlfwy hse/res facil-last 6 months
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E4C</td>
<td>Ttl nites in hlfwy hse/res fac-last 6 months</td>
</tr>
<tr>
<td>E5A</td>
<td>In day trtmt prg for alcohol/drug-last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>E5B</td>
<td>Total # days in day trtmt prg-last 6 months</td>
</tr>
<tr>
<td>E6</td>
<td>In methadone maintenance prg-last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>E7A</td>
<td>Visit outpt prg subst ab couns-last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>E7B</td>
<td># visits outpt prg subst ab couns-last 6 months</td>
</tr>
<tr>
<td>E8A1</td>
<td>Saw MD/H care worker regarding alcohol/drugs-last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>E8A2</td>
<td>Saw Prst/Min/Rabbi re alcohol/drugs-last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>E8A3</td>
<td>Employ Asst Prg for alcohol/drug prb-last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>E8A4</td>
<td>Oth source cnsl for alcohol/drug prb-last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>E8A5</td>
<td>AA/NA/slfl-hlp for drug/alcohol/emot-last 6 months (0=No, 1=Yes)</td>
</tr>
<tr>
<td>E9B</td>
<td>How often attend AA/NA/slfl-hlp-last 6 months (1=Daily, 2=2-3 Times/week, 3=Weekly, 4=Evry 2 weeks, 5=Once/month)</td>
</tr>
</tbody>
</table>

**EPI_6M**

Episodic (C2A-C2O)-6m y/n (0=No, 1=Yes)

**EPI_6M**

Episodic (C2A-C2O,C2R-C2U, STD)-6m y/n (0=No, 1=Yes)

**EPI_SUM**

Sum episodic (C2A-C2O, C2R-C2U, STD)-6m

F1A  | Bothered by thngs not generally bothered by (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1B  | My appetite was poor (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1C  | Couldn’t shake blues evn w/fam+frnds hlp (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1D  | Felt I was just as good as other people (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1E  | Had trouble keeping mind on what doing (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1F  | I felt depressed (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1G  | I felt everything I did was an effort (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1H  | I felt hopeful about the future (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1I  | I thought my life had been a failure (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1J  | I felt fearful (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1K  | My sleep was restless (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1L  | I was happy (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
| F1M  | I talked less than usual (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time) |
F1N  I felt lonely (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time)
F10  People were unfriendly (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time)
F1P  I enjoyed life (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time)
F1Q  I had crying spells (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time)
F1R  I felt sad (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time)
F1S  I felt that people dislike me (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time)
F1T  I could not get going (0=Rarely/never, 1=Some of the time, 2=Occas/moderately, 3=Most of the time)

FAMABUSE  Family abuse-physical or sexual (0=No, 1=Yes)
FRML_SAT  Formal substance abuse treatment y/n (0=No, 1=Yes)
G1A_30  Diff contr viol beh-sig per last 30 days (0=No, 1=Yes)
G1A  Diff contr viol beh for sig time per evr (0=No, 1=Yes)
G1B_30  Had thoughts of suicide-last 30 days (0=No, 1=Yes)
G1B_REC  Suicidal thoughts past 30 days y/n (0=No, 1=Yes)
G1B  Ever had thoughts of suicide (0=No, 1=Yes)
G1C_30  Attempted suicide-last 30 days (0=No, 1=Yes)
G1C  Attempted suicide ever (0=No, 1=Yes)
G1D_30  Prescr med for psy/emot prob-last 30 days (0=No, 1=Yes)
G1D_REC  Prescribed psych meds past 30 days y/n (0=No, 1=Yes)
G1D  Prescr med for pst/emot prob ever (0=No, 1=Yes)
GH  SF-36 general health perceptions (0-100)
GOV_SUPP  Received government support past 6 m (0=No, 1=Yes)
GROUP  Randomization Group (0=Control, 1=Clinic)
H10_30  # days in last 30 bef detox used cannabis
H10_LT  # years regularly used cannabis
H10_PRB  Problem sub: marijuana, cannabis (0=No, 1=Yes)
H10_RT  Route of admin of cannabis (0=N/A. 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H11_30  # days in last 30 bef detox used halluc
H11_LT  # years regularly used hallucinogens
H11_PRB  Problem sub: hallucinogens (0=No, 1=Yes)
H11_RT  Route of admin of hallucinogens (0=N/A. 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H12_30  # days in last 30 bef detox used inhalant
H12_LT  # years regularly used inhalants
H12_PRB  Problem sub: inhalants (0=No, 1=Yes)
H12_RT  Route of admin of inhalants (0=N/A. 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H13_30  # days used >1 sub/day-last 30 bef detox
H13_LT  # years regularly used >1 subst/day
H13_RT  Route of admin of >1 subst/day (0=N/A. 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H14  According to interviewer, which substance is main problem (0=No problem, 1=Alcohol, 2=Alcohol to intox, 3=Heroin 4=Methadone, 5=Other opiate/analg, 6=Barbituates, 7=Sed/hyp/tranq, 8=Cocaine, 9=Amphetamines, 10=Marij/cannabis, 15=Alcohol and one or more drug, 16=More than one drug
H15A  # times had alcohol DTs
H15B  # times overdosed on drugs
H16A  $ spent on alcohol-last 30 days bef detox
H16B  $ spent on drugs-last 30 days bef detox
H17A  # days had alcohol prob-last 30 days bef det
H17B  # days had drug prob-last 30 days bef det
H18A  How troubled by alcohol probs-last 30 days (0=Not at all, 1=Slightly, 2=Moderately, 3=Considerably, 4=Extremely)
H18B  How troubled by drug probs-last 30 days (0=Not at all, 1=Slightly, 2=Moderately, 3=Considerably, 4=Extremely)
H19A  How important is treatment for alcohol problems now (0=Not at all, 1=Slightly, 2=Moderately, 3=Considerably, 4=Extremely)
H19B  How important is trtmnt for drug probs now (0=Not at all, 1=Slightly, 2=Moderately, 3=Considerably, 4=Extremely)
H1_30  # days in past 30 bef detox used alcohol
H1_LT  # years regularly used alcohol
H1_RT  Route of administration use alcohol (0=N/A. 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H2_30  #days in 3- bef detox use alcohol to intox
H2_LT  # years regularly used alcohol to intox
H2_PRB Problem sub: alcohol to intox (0=No, 1=Yes)
H2_RT  Route of admin use alcohol to intox (0=N/A. 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H3_30  # days in past 30 bef detox used heroin
H3_LT  # years regularly used heroin
H3_PRB Problem sub: heroin (0=No, 1=Yes)
H3_RT  Route of administration of heroin (0=N/A. 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H4_30  # days used methadone-last 30 bef detox
H4_LT  # years regularly used methadone
H4_PRB  Problem sub: methadone (0=No, 1=Yes)
H4_RT  Route of administration of methadone (0=N/A, 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H5_30  # days used opiates/analgesics last 30 before detox
H5_LT  # years regularly used other opiates/analgesics
H5_PRB Problem sub: other opiates/analgesics (0=No, 1=Yes)
H5_RT  Route of admin of other opiates/analgesics (0=N/A, 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H6_30  # days in past 30 before detox used barbiturates
H6_LT  # years regularly used barbiturates
H6_PRB Problem sub: barbiturates (0=No, 1=Yes)
H6_RT  Route of admin of barbiturates (0=N/A, 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H7_30  # days used sedative/hypnotic/tranquilizer last 30 before detox
H7_LT  # years regularly used sedative/hypnotic/tranquilizer
H7_PRB Problem sub: sedative/hypnotic/tranquilizer (0=No, 1=Yes)
H7_RT  Route of admin of sedative/hypnotic/tranquilizer (0=N/A, 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H8_30  # days in last 30 before detox used cocaine
H8_LT  # years regularly used cocaine
H8_PRB Problem sub: cocaine (0=No, 1=Yes)
H8_RT  Route of admin of cocaine (0=N/A, 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
H9_30  # days in last 30 before detox used amphetamines
H9_LT  # years regularly used amphetamines
H9_PRB Problem sub: amphetamines (0=No, 1=Yes)
H9_RT  Route of admin of amphetamines (0=N/A, 1=Oral, 2=Nasal, 3=Smoking, 4=Non-IV injection, 5=IV)
HOMELESS  Homeless-shelter/street past 6 m (0=No, 1=Yes)
HS_GRAD High school graduate (0=No, 1=Yes)
HT  Raw SF-36 health transition item
I1  Avg # drinks in last 30 days bef detox
I2  Most drank any 1 day in last 30 bef detox
I3  On days used heroin, avg # bags used
I4  Most bags heroin used any 1 day – 30 before det
I5  Avg $ amt of heroin used per day
I6A  On days used cocaine, avg # bags used
I6B  On days used cocaine, avg # rocks used
I7A  Most bags cocaine use any 1 day-30 bef det

HELPlfull
I7B Most rcks cocaine use any 1 day-30 bef det
I8 Avg $ amt of cocaine used per day
IMPUL2 Inventory of Drug Use Consequences InDUC-2L-Impulse control-Raw (w/o M23)
IMPUL Inventory of Drug Use Consequences InDUL-2L-Impulse control-Raw
INDTOT2 InDUC-2L-Total drlnC-Raw- w/o M23 and M48
INDTOT InDUC-2LTotal drlnC sore-Raw
INTER InDUC-2L-Interpersonal-Raw
INTRA InDUC-2L-Intrapersonal-Raw
INT_TIME1 # of months from baseline to current interview
INT_TIME2 # of months from previous to current interview
J10A Get physically sick when stop using heroin (0=No, 1=Yes)
J10B Ever use heroin to prevent getting sick (0=No, 1=Yes)
J1 Ever don’t stop using cocaine when should (0=No, 1=Yes)
J2 Ever tried to cut down on cocaine (0=No, 1=Yes)
J3 Does cocaine take up a lot of your time (0=No, 1=Yes)
J4 Need use > cocaine to get some feeling (0=No, 1=Yes)
J5A Get physically sick when stop using cocaine (0=No, 1=Yes)
J5B Ever use cocaine to prevent getting sick (0=No, 1=Yes)
J6 Ever don’t stop using heroin when should (0=No, 1=Yes)
J7 Ever tried to cut down on heroin (0=No, 1=Yes)
J8 Does heroin take up a lot of your time (0=No, 1=Yes)
J9 Need use > heroin to get some feeling (0=No, 1=Yes)
JAIL_5YR Any jail time past 5 years y/n (0=No, 1=Yes)
JAIL_MOS Total months in jail past 5 years
K1 Do you currently smoke cigarettes (1=Yes-every day, 2=Yes-some days, 3=No-former smoker, 4=No-never>100 cigarettes
K2 Avg # cigarettes smoked per day
K3 Considering quitting cigarettes within next 6 months (0=No, 1=Yes)
L10 Have had blkouts as result of drinkng (0=No, never, 1=Sometimes, 2=Often, 3=Alm evry time drink)
L11 Do you carry bottle or keep close by (0=No, 1=Some of the time, 2=Most of the time)
L12 After abstin end up drink heavily again (0=No, 1=Sometimes, 2=Almost evry time)
L13 Passed out due to drinking-last 12 months (0=No, 1=Once, 2=More than once)
L14 Had convuls following period of drinkng (0=No, 1=Once, 2=Several times)
L15 Do you drink throughout the day (0=No, 1=Yes)
L16 After drinking heavily was thinking unclear (0=No, 1=Yes, few hrs, 2=Yes,1-2 days, 3=Yes, many days)
L17 Did drinking felt heart beat rapidly (0=No, 1=Once, 2=Several times)
L18  Do you constantly think about drinking/alcohol (0=No, 1=Yes)
L19  D/t drinking heard things not there (0=No, 1=Once, 2=Several times)
L20  Had weird/fright sensations when drinking (0=No, 1=Once or twice, 2=Often)
L21  When drinking felt things rawl not there (0=No, 1=Once, 2=Several times)
L22  With respect to blackouts (0=Never had one, 1=Had for <1hr, 2=Had several hrs, 3=Had for day+)
L23  Ever tried to cut down on drinking & failed (0=No, 1=Once, 2=Several times)
L24  Do you gulp drinks (0=No, 1=Yes)
L25  After taking 1 or 2 drinks can you stop (0=No, 1=Yes)
L26  Often have hangovers Sun or Mon mornings (0=No, 1=Yes)
L27  Have you had the shakes when sobering (0=No, 1=Sometimes, 2=Alm evry time drink)
L28  Do you get physically sick as reslt of drinking (0=No, 1=Sometimes, 2=Alm evry time drink)
L29  have you had the DTs (0=No, 1=Once, 2=Several times)
L30  When drink do you stumble/stagger/weave (0=No, 1=Sometimes, 2=Often)
L31  D/t drinking felt overly hot/sweaty (0=No, 1=Once, 2=Several times)
L32  As result of drinking saw things not there (0=No, 1=Once, 2=Several times)
L33  Panic because fear not have drink if need it (0=No, 1=Yes)
L34  Using alcohol/1 drug caused > use othr drugs (0=No, 1=Yes)
L35  I have been sick/vomited aft alcohol/drug use (0=No, 1=Yes)
L36  I have been unhappy because of alcohol/drug use (0=No, 1=Yes)
L37  Lost weight/eaten poorly due to alcohol/drug use (0=No, 1=Yes)
L38  Fail to do what expected due to alcohol/drug use (0=No, 1=Yes)
L39  Using alcohol/drugs has helped me to relax (0=No, 1=Yes)
L40  Felt guilt/ashamed because of my alcohol drug use (0=No, 1=Yes)
L41  Said/done emarras thngs when on alcohol/drug (0=No, 1=Yes)
L42  Personality changed for worse on alcohol/drug (0=No, 1=Yes)
L43  Taken foolish risk when using alcohol/drugs (0=No, 1=Yes)
L44  Had hangover/felt bad aftr using alcohol/drugs (0=No, 1=Yes)
L45  Gotten into trouble because of alcohol/drug use (0=No, 1=Yes)
L46  Said cruel things while using alcohol/drugs (0=No, 1=Yes)
L47  Done impuls thngs regret due to alcohol/drug use (0=No, 1=Yes)
L48  Gotten in physical fights when use alcohol/drugs (0=No, 1=Yes)
L49  My physical health was harmed by alcohol/drug use (0=No, 1=Yes)
L50  Using alcohol/drug helped me have more + outlook (0=No, 1=Yes)
L51  I have had money probs because of my alcohol/drug use (0=No, 1=Yes)
M27 My love relat harmed due to my alcohol/drug use (0=No, 1=Yes)
M28 Smoked tobacco more when using alcohol/drugs (0=No, 1=Yes)
M29 My physical appearance harmed by alcohol/drug use (0=No, 1=Yes)
M2 Felt bad about self because of alcohol/drug use (0=No, 1=Yes)
M30 My family hurt because of my alcohol drug use (0=No, 1=Yes)
M31 Close relationsp damaged due to alcohol/drug use (0=No, 1=Yes)
M32 Spent time in jail because of my alcohol/drug use (0=No, 1=Yes)
M33 My sex life suffered due to my alcohol/drug use (0=No, 1=Yes)
M34 Lost interst in activity due to my alcohol/drug use (0=No, 1=Yes)
M35 Soc life> enjoyable when using alcohol/drug (0=No, 1=Yes)
M36 Spirit/moral life harmed by alcohol/drug use (0=No, 1=Yes)
M37 Not had kind life want due to alcohol/drug use (0=No, 1=Yes)
M38 My alcohol/drug use in way of personal growth (0=No, 1=Yes)
M39 My alcohol/drug use damaged soc life/reputat (0=No, 1=Yes)
M3 Missed days wrk/sch because of alcohol/drug use (0=No, 1=Yes)
M30 Spent/lost too much $ because alcohol/drug use (0=No, 1=Yes)
M41 Arrested for DUI of alcohol or oth drugs (0=No, 1=Yes)
M42 Arrested for offenses rel to alcohol/drug use (0=No, 1=Yes)
M43 Lost marriage/love relat due to alcohol/drug use (0=No, 1=Yes)
M44 Susp/fired/left job/sch due to alcohol/drug use (0=No, 1=Yes)
M45 I used drugs moderately w/o having prob (0=No, 1=Yes)
M46 I have lost a friend due to my alcohol/drug use (0=No, 1=Yes)
M47 Had an accident while using alcohol/drugs (0=No, 1=Yes)
M48 Physically hurt/injured/burned when using alcohol/drugs (0=No, 1=Yes)
M49 I injured someone while using alcohol/drugs (0=No, 1=Yes)
M50 Damaged things/prop when using alcohol/drugs (0=No, 1=Yes)
M5 I have enjoyed drinking/using drugs (0=No, 1=Yes)
M6 Qual of work suffered because of alcohol/drug use (0=No, 1=Yes)
M7 Parenting ability harmed by alcohol/drug use (0=No, 1=Yes)
M8 Trouble sleeping/nightmares aft alcohol/drugs (0=No, 1=Yes)
M9 Driven motor veh while undr inf alcohol/drugs (0=No, 1=Yes)
MAR_STAT Marital status (recode) (0=Married, 1=Not married)
MCS Standardized mental component scale-00
MD_LANG Lang prefer to speak to MD (recode) (0=English, 1=Other lang)
MH SF-36 mental health index (0-100)
MMSEC MMSEC
N1A My friends give me the moral support I need (0=No, 1=Yes)
N1B Most people closer to friends than I am (0=No, 1=Yes)
N1C My friends enjoy hearing what I think (0=No, 1=Yes)
N1D I rely on my friends for emot support (0=No, 1=Yes)
N1E Friend go to when down w/o feel funny later (0=No, 1=Yes)
N1F Frnds and I open re what thnk about things (0=No, 1=Yes)
N1G My friends sensitive to my pers needs (0=No, 1=Yes)
N1H My friends good at helping me solve probs (0=No, 1=Yes)
N1I have deep sharing relat w/ a # of frnds (0=No, 1=Yes)
N1J When confide in frnds makes me uncomfort (0=No, 1=Yes)
N1K My friends seek me out for companionship (0=No, 1=Yes)
N1L Not have as int relat w/frnds as others (0=No, 1=Yes)
N1M Recent good idea how to do somethng frm frnd (0=No, 1=Yes)
N1N I wish my friends were much different (0=No, 1=Yes)
N2A My family gives me the moral support I need (0=No, 1=Yes)
N2B Good ideas of how do/make thngs from fam (0=No, 1=Yes)
N2C Most peop closer to their fam than I am (0=No, 1=Yes)
N2D When confide make close fam membs uncomf (0=No, 1=Yes)
N2E My fam enjoys hearing about what I think (0=No, 1=Yes)
N2F Membs of my fam share many of my intrsts (0=No, 1=Yes)
N2G I rely on my fam for emot support (0=No, 1=Yes)
N2H Fam memb go to when down w/o feel funny (0=No, 1=Yes)
N2I Fam and I open about what thnk about thngs (0=No, 1=Yes)
N2J My fam is sensitive to my personal needs (0=No, 1=Yes)
N2K Fam memb good at helping me solve probs (0=No, 1=Yes)
N2L Have deep sharing relat w/# of fam membs (0=No, 1=Yes)
N2M Makes me uncomf to confide in fam membs (0=No, 1=Yes)
N2N I wish my family were much different (0=No, 1=Yes)
NUM_BARR # of perceived barriers to linkage
NUM_INTERVALS Number of 6-month intervals from previous to current interview
O1A # people spend tx w/who drink alcohol (1=None, 2= A few, 3=About half, 4= Most, 5=All)
O1B_REC Family/friends heavy drinkers y/n (0=No, 1=Yes)
O1B # people spend tx w/who are heavy drinkrs (1=None, 2= A few, 3=About half, 4= Most, 5=All)
O1C_REC Family/friends use drugs y/n (0=No, 1=Yes)
O1C # people spend tx w/who use drugs (1=None, 2= A few, 3=About half, 4= Most, 5=All)
O1D_REC Family/friends support abst. y/n (0=No, 1=Yes)
O1D # peop spend tx w/who supprt your abstin (1=None, 2= A few, 3=About half, 4= Most, 5=All)
02_REC Live-in partner drinks/drugs y/n (0=No, 1=Yes)
02 Does live-in part/spouse drink/use drugs (0=No, 1=Yes, 2=N/A)
P1A Physical abuse/assault by family members/person I know (0=No, 1=Yes, 7=Not sure)
P1B Age first physically assaulted by person I know
P1C Physically assaulted by person I know-last 6 months (0=No, 1=Yes)
P2A Physical abuse/assault by stranger (0=No, 1=Yes, 7=Not sure)
P2B Age first physically assaulted by stranger
P2C Physically assaulted by stranger-last 6 months (0=No, 1=Yes)
P3 Using drugs/alcohol when physically assaulted (1=Don’t know, 2=Never, 3=Some cases, 4=Most cases, 5=All cases, 9=Never assaulted)
P4 Person who physically assaulted you was using alcohol/drugs (1=Don’t know, 2=Never, 3=Some cases, 4=Most cases, 5=All cases, 9=Never assaulted)
P5A Sexual abuse/assault by family member/person you know (0=No, 1=Yes, 7=Not sure)
P5B Age first sexually assaulted by person you know
P5C Sexually assaulted by person you know-last 6 months (0=No, 1=Yes)
P6A Sexual abuse/assault by stranger (0=No, 1=Yes, 7=Not sure)
P6B Age first sexually assaulted by stranger
P6C Sexually assaulted by stranger-last 6 months (0=No, 1=Yes)
P7 Using drugs/alcohol when sexually assaulted (1=Don’t know, 2=Never, 3=Some cases, 4=Most cases, 5=All cases, 9=Never assaulted)
P8 Person who sexually assaulted you using alcohol/drugs (1=Don’t know, 2=Never, 3=Some cases, 4=Most cases, 5=All cases, 9=Never assaulted)
PCP_ID a numeric vector
PCS Standardized physical component scale-00
PC_REC7 Primary cared received: Linked & # visits (0=Not linked, 1=Linked, 1 visit, 2=Linked, 2 visits, 3=Linked, 3 visits, 4=Linked, 4 visits, 5=Linked, 5 visits, 6=Linked, 6+visits)
PC_REC Primary care received: Linked & # visits (0=Not linked, 1=Linked, 1 visit, 2=Linked, 2+ visits)
PF SF-36 physical functioning (0-100)
PHSXABUS Any abuse (0=No, 1=Yes)
PHYABUSE Physical abuse-stranger or family (0=No, 1=Yes)
PHYS2 InDUC-2L-Physical 9Raw (w/o M48)
PHYS InDUC-2L-Physical-Raw
POLYSUB Polysubstance abuser y/n (0=No, 1=Yes)
PREV_TIME Previous interview time
PRIMLANG First language (recode) (0=English, 1=Other lang)
PRIMSUB2 First drug of choice (no marijuana) (0=None, 1=Alcohol, 2=Cocaine, 3=Heroin, 4=Barbituates, 5=Benzos, 6=Marijuana, 7=Methadone, 8=Opiates)
PRIM_SUB First drug of choice (0=None, 1=Alcohol, 2=Cocaine, 3=Heroin, 4=Barbituates, 5=Benzo, 6=Marijuana, 7=Methadone, 8=Opiates)
PSS_FA Perceived social support-family
PSS_FR Perceived social support-friends
Q10 How would you describe yourself (0=Straight, 1=Gay/bisexual)
Q11 # men had sex w/in past 6 months (0=0 men, 1=1 man, 2=2-3 men, 3=4+ men
Q12 # women had sex w/in past 6 months (0=0 women, 1=1 woman, 2=2-3 women, 3=4+ women
Q13 # times had sex In past 6 months (0=Never, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q14 How often had sex to get drugs last 6 months (0=Never, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q15 How often given drugs to have sex last 6 months (0=Never, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q16 How often were you paid for sex last 6 months (0=Never, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q17 How often you pay pers for sex last 6 months (0=Never, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q18 How often use condoms during sex last 6 months (0=No sex/always, 1=Most of the time, 2=Some of the time, 3=None of the time)
Q19 Condoms are too much of a hassle to use (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)
Q1A Have you ever injected drugs (0=No, 1=Yes)
Q1B Have you injected drugs last 6 months (0=No, 1=Yes)
Q20 Safer sex is always your responsibility (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)
Q2 Have you shared needles/works last 6 months (0=No/Not shot up, 3=Yes)
Q3 # people shared needles w/past 6 months (0=No/Not shot up, 1=1 other person, 2=2-3 diff people, 3=4/+ diff people)
Q4 How often been to shoot gall/hse last 6 months (0=Never, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q5 How often been to crack house last 6 months (0=Never, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q6 How often shared rinse-water last 6 months (0=Nevr/Not shot up, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q7 How often shared a cooker last 6 months (0=Nevr/Not shot up, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q8 How often shared a cotton last 6 months (0=Nevr/Not shot up, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
Q9 How often use syringe to div drugs last 6 months (0=Nevr/Not shot up, 1=Few times or less, 2=Few times/month, 3=Once or more/week)
R1A I really want to change my alcohol/drug use (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)
Sometimes I wonder if I’m an alcohol/addict (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

Id I don’t change alcohol/drug probs will worsen (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I started making changes in alcohol/drug use (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

Was using too much but managed to change (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I wonder if my alcohol/drug use hurting othrs (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I started making changes in alcohol/drug use (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I am a prob drinker or have drug prob (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

Already doing thngs to change alcohol/drug use (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I have changed use-trying to not slip back (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I have a serious problem w/ alcohol/drugs (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I wonder if I’m in control of alcohol/drug use (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

My alcohol/drug use is causing a lot of harm (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

Actively cutting down/stopping alcohol/drug use (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

Want help to not go back to alcohol/drugs (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I know that I have an alcohol/drug problem (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I wonder if I use alcohol/drugs too much (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I am an alcoholic or drug addict (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

I am working hard to change alcohol/drug use (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

Some changes-want help from going back (1=Strongly disagree, 2=Disagree, 3= Agree, 4=Strongly agree)

RAB SCALE  RAB scale sore

RACE2  Race (recode) (1=White, 2=Minority)

RACE  Race (recode) (1=Afr Amer/Black, 2=White, 3=Hispanic, 4=Other)

RAWBP  Raw SF-36 pain index

RAWGH  Raw SF-36 general health perceptions
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAWMH</td>
<td>Raw SF-36 mental health index</td>
</tr>
<tr>
<td>RAWPF</td>
<td>Raw SF-36 physical functioning</td>
</tr>
<tr>
<td>RAWRE</td>
<td>Raw SF-36 role-emotional</td>
</tr>
<tr>
<td>RAWRP</td>
<td>Raw SF-36 role-physical</td>
</tr>
<tr>
<td>RAWSF</td>
<td>Raw SF-36 social functioning</td>
</tr>
<tr>
<td>RAWVT</td>
<td>Raw SF-36 vitality</td>
</tr>
<tr>
<td>RAW_ADS</td>
<td>ADS score</td>
</tr>
<tr>
<td>RAW_AM</td>
<td>SOCRATES-Ambivalence-Raw</td>
</tr>
<tr>
<td>RAW_RE</td>
<td>SOCRATES-Recognition-Raw</td>
</tr>
<tr>
<td>RAW_TS</td>
<td>SOCRATES-Taking steps-Raw</td>
</tr>
<tr>
<td>RCT_LINK</td>
<td>Did subject link to primary care (RCT)–This time point (0=No, 1=Yes)</td>
</tr>
<tr>
<td>REALM2</td>
<td>REALM score (dichotomous) (1=0-60, 2=61-66)</td>
</tr>
<tr>
<td>REALM3</td>
<td>REALM score (categorical) (1=0-44), 2=45-60), 3=61-66)</td>
</tr>
<tr>
<td>REG_MD</td>
<td>Did subject report having regular doctor–This time point (0=No, 1=Yes)</td>
</tr>
<tr>
<td>RE</td>
<td>SF-36 role-emotional (0-100)</td>
</tr>
<tr>
<td>RP</td>
<td>SF-36 role physical (0-100)</td>
</tr>
<tr>
<td>S1A</td>
<td>At interview pt obviously depressed/withdrawn (0=No, 1=Yes)</td>
</tr>
<tr>
<td>S1B</td>
<td>At interview pt obviously hostile (0=No, 1=Yes)</td>
</tr>
<tr>
<td>S1C</td>
<td>At interview patientt obviously anxious/nervous (0=No, 1=Yes)</td>
</tr>
<tr>
<td>S1D</td>
<td>Trouble with real tst/thght dis/par at interview (0=No, 1=Yes)</td>
</tr>
<tr>
<td>S1E</td>
<td>At interview pt trbl w/ compr/concen/rememb (0=No, 1=Yes)</td>
</tr>
<tr>
<td>S1F</td>
<td>At interview pt had suicidal thoughts (0=No, 1=Yes)</td>
</tr>
<tr>
<td>SATREAT</td>
<td>Any BSAS substance abuse this time point (0=No, 1=Yes)</td>
</tr>
<tr>
<td>SECD_SUB</td>
<td>Second drug of choice (0=None, 1=Alcohol, 3=Cocaine, 3=Heroine, 4=Barbituates,</td>
</tr>
<tr>
<td></td>
<td>5=Benzos, 6=Marijuana, 7=Methadone, 8=Opiates)</td>
</tr>
<tr>
<td>SER_INJ</td>
<td>Recent (6m) serious injury y/n (0=No, 1=Yes)</td>
</tr>
<tr>
<td>SEXABUSE</td>
<td>Sexual abuse-stranger or family (0=No, 1=Yes)</td>
</tr>
<tr>
<td>SEXRISK</td>
<td>RAB-Sex risk total</td>
</tr>
<tr>
<td>SF</td>
<td>SF-36 social functioning (0-100)</td>
</tr>
<tr>
<td>SMOKER</td>
<td>Current smoker (every/some days) y/n (0=No, 1=Yes)</td>
</tr>
<tr>
<td>SR</td>
<td>InDUC-2L-Social responsibility-Raw</td>
</tr>
<tr>
<td>STD_6M</td>
<td>Had an STD past 6m y/n (0=No, 1=Yes)</td>
</tr>
<tr>
<td>STD_EVER</td>
<td>Ever had an STD y/n (0=No, 1=Yes)</td>
</tr>
<tr>
<td>STRABUSE</td>
<td>Stranger abuse-physical or sexual (0=No, 1=Yes)</td>
</tr>
<tr>
<td>T1B</td>
<td># days in row continued to drink</td>
</tr>
<tr>
<td>T1C</td>
<td>Longest period abstain-last 6 months (alcohol)</td>
</tr>
</tbody>
</table>
T1 Have used alcohol since leaving River St. (0=No, 1=Yes)
T2 Have used heroin since leaving River St (0=No, 1=Yes)
T3 Have used cocaine since leaving River St (0=No, 1=Yes)
TIME Interview time point
TOTALRAB RAB-Total RAB sore
U10A # times been to regular MDs office-pst 6 months
U10B # times saw regular MD in office-pst 6 months
U10C # times saw oth prof in office-pst 6 months
U11 Rate convenience of MD office location (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
U12 Rate hours MD office open for medical appointments (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
U13 Usual wait for appointment when sick (unscheduled) (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
U14 Time wait for appointment to start at MD office (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
U15A Do you pay for any/all of MD visits (0=No, 1=Yes)
U15B How rate amt of $ you pay for MD visits (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
U16A Do you pay for any/all of prescript meds (0=No, 1=Yes)
U16B Rate amt $ pay for meds/prescript trtmnts (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
U17 Ever skip meds/trtmnts because too expensive (1=Yes, often, 2=Yes, occasionally, 3=No, never)
U18A Ability to reach MC office by phone (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
U18B Ability to speak to MD by phone if need (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
U19 How often see regular MD when have regular check-up (1=Always, 2=Almost always, 3=A lot of the time, 4=Some of the time, 5=Almost never, 6=Never)
U20 When sick + go to MD how often see regular MD (1=Always, 2=Almost always, 3=A lot of the time, 4=Some of the time, 5=Almost never, 6=Never)
U21A How thorough MD exam to check health prb (1=Very poor, 2=Poor, 3=Fair, 4=Good, 5=Very good, 6=Excellent)
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Response Options</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>U21B</td>
<td>How often question if MD diagnosis right (1=Always, 2=Almost always, 3=A lot of the time, 4=Some of the time, 5=Almost never, 6=Never)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U22A</td>
<td>Thoroughness of MD questions re symptoms (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U22B</td>
<td>Attn MD gives to what you have to say (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U22C</td>
<td>MD explanations of health problems/treatments need (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U22D</td>
<td>MD instructions re symptom report/further care (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U22E</td>
<td>MD advice in decisions about your care (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U22F</td>
<td>How often leave MD office with unanswd quests (1=Always, 2=Almost always, 3=A lot of the time, 4=Some of the time, 5=Almost never, 6=Never)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U24A</td>
<td>Amount of time your MD spends with you (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U24B</td>
<td>MDs patience w/ your questions/worries (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U24C</td>
<td>MDs friendliness and warmth toward you (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U24D</td>
<td>MDs caring and concern for you (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U24E</td>
<td>MDs respect for you (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25A</td>
<td>Reg MD ever talked to you about smoking (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25B</td>
<td>Reg MD ever talked to you about alcohol use (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25C</td>
<td>Reg MD ever talk to you about seat belt use (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25D</td>
<td>Reg MD ever talked to you about diet (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25E</td>
<td>Reg MD ever talked to you about exercise (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25F</td>
<td>Reg MD ever talked to you about stress (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25G</td>
<td>Reg MD ever talked to you about safe sex (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25H</td>
<td>Reg MD ever talked to you about drug use (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U25I</td>
<td>Reg MD ever talked to you about HIV testing (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U26A</td>
<td>Cut/quit smoking because of MDs advice (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U26B</td>
<td>Tried to drink less alcohol because of MD advice (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U26C</td>
<td>Wore my seat belt more because of MDs advice (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U26D</td>
<td>Changed diet because of MDs advice (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U26E</td>
<td>Done more exercise because MDs advice (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U26F</td>
<td>Relax/reduce stress because of MDs advice (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U26G</td>
<td>Practiced safer sex because of MDs advice (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U26H</td>
<td>Tried to cut down/quit drugs because MD advice (0=No, 1=Yes)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
U26I  Got HIV tested because of MDs advice (0=No, 1=Yes)"
U27A  I can tell my MD anything (1=Strongly agree, 2= Agree, 3= Not sure, 4=Disagree, 5=Strongly disagree)"
U27B  My MD pretends to know things if not sure (1=Strongly agree, 2= Agree, 3= Not sure, 4=Disagree, 5=Strongly disagree)"
U27C  I trust my MDs judgment re my med care (1=Strongly agree, 2= Agree, 3= Not sure, 4=Disagree, 5=Strongly disagree)"
U27D  My MD cares > about < costs than my health (1=Strongly agree, 2= Agree, 3= Not sure, 4=Disagree, 5=Strongly disagree)"
U27E  My MD always tell truth about my health (1=Strongly agree, 2= Agree, 3= Not sure, 4=Disagree, 5=Strongly disagree)"
U27F  My MD cares as much as I about my health (1=Strongly agree, 2= Agree, 3= Not sure, 4=Disagree, 5=Strongly disagree)"
U27G  My MD would try to hide a mistake in trmt (1=Strongly agree, 2= Agree, 3= Not sure, 4=Disagree, 5=Strongly disagree)"
U28  How much to you trust this MD (0=Not at all, 1=1, 2=2, 3=3, 4=4, 5=5, 6=6, 7=7, 8=8, 9=9, 10=Completely)"
U29A  MDs knowledge of your entire med history (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)"
U29B  MD knowledge of your response-home/work/sch (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)"
U29C  MD knowledge of what worries you most-health (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)"
U29D  MDs knowledge of you as a person (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5=Very good, 6= Excellent)"
U2A  I cannot pay for services (0=No, 1=Yes)
U2B  I am not eligible for free care (0=No, 1=Yes)
U2C  I do not know where to go (0=No, 1=Yes)
U2D  Can’t get services due to transport probs (0=No, 1=Yes)
U2E  Office/clinic hours are inconvenient (0=No, 1=Yes)
U2F  I do not speak/understand English well (0=No, 1=Yes)
U2G  Afraid others discover health prb I have (0=No, 1=Yes)
U2H  My substance abuse interferes (0=No, 1=Yes)
U2I  I do not have a babysitter (0=No, 1=Yes)
U2J  I do not want to lose my job (0=No, 1=Yes)
U2K  My insurance does not cover services (0=No, 1=Yes)
U2L  Medical care is not important to me (0=No, 1=Yes)
U2M  I do not have time (0=No, 1=Yes)
U2N  Med staff do not treat me with respect (0=No, 1=Yes)
U2O  I do not trust my doctors or nurses (0=No, 1=Yes)
HELPfull

U2P Often been unsatisfied w/my med care (0=No, 1=Yes)
U2Q Other reason hard to get regular med care (0=No, 1=Yes)
U2R a factor with levels 7A B C D E F G H I J K L M N O P Q
U30 MD would know what want done if unconscious (1=Strongly agree, 2=Agree, 3=Not sure, 4=Disagree, 5=Strongly disagree)
U31 Oth MDs/RNs who play role in your care (0=No, 1=Yes) * 
U32A Their knowledge of you as a person (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U32B The quality of care they provide (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U32C Coordination between them and your regular MD (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U32D_T N/A, only my regular MD does this
U32D Their explanation of your health prbs/trmt need (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U33 Amt regular MD knows about care from others (1=Knows everything, 2=Knows almost everything, 3=Knows some things, 4=Knows very little, 5=Knows nothing)
U34 Has MD ever recommended you see MD specialists (0=No, 1=Yes)
U35A How helpful MD in deciding on specialist (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U35B How helpful MD getting appointment with specialist (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U35C MDs involvement when you trtd by specialist (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U35D MDs communication w/your specialists/oth MDs (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U35E MD help in explain what specialists said (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U35F Quality of specialists MD sent you to (1=Very poor, 2= Poor, 3=Fair, 4=Good, 5= Very good, 6= Excellent)
U36 How many minutes to get to MDs office (1=<15, 2=16-30, 3=31-60, 4=More than 60)
U37 When sick+call how long take to see you (1=Same day, 2=Next day, 3=In 2-3 days, 4=In 4-5 days, 5=In >5 days)
U38 How many minutes late appointment usually begin (1= None, 2=<5 minutes, 3=6-10 minutes, 4=11-20 minutes, 5=21-30 minutes, 6=31-45 minutes, 7=>45 minutes)
U39 How satisfied are you w/your regular MD (1=Completely satisfied, 2=Very satisfied, 3=Somewhat satisfied, 4=Neither, 5=Somewhat dissatisfied, 6=Very dissatisfied, 7=Completely dissatisfied)
U3A Has MD evr talked to you about drug use (0=No, 1=Yes)
U3B Has MD evr talked to you about alcohol use (0=No, 1=Yes)
U4  Is there an MD you consider your regular MD (0=No, 1=Yes)
U5  Have you seen any MDs in last 6 months (0=No, 1=Yes)
U6A  Would you go to this MD if med prb not emergency (0=No, 1=Yes)
U6B  Think one of these could be your regular MD (0=No, 1=Yes)
U7A_T a factor with levels  ARTHRITIS DOCTOR CHIROPRACTOR COCAINE STUDY DETOX DOCTOR DO
EAR DOCTOR EAR SPECIALIST EAR, NOSE, & THROAT. EAR/NOSE/THROAT ENT FAMILY PHYSICIAN
GENERAL MEDICINE GENERAL PRACTICE GENERAL PRACTITIONER HEAD & NECK SPECIALIST HERBAL/HOMEOPATHIC/ACUPUNCTURE ID DOCTOR MAYBE GENERAL PRACTITIONER
MEDICAL STUDENT NEUROLOGIST NURSE NURSE PRACTITIONER NURSE PRACTITIONER ONCOLOGIST
PRENATAL PRIMARY CARE PRIMARY CARE DOCTOR PRIMARY CARE THERAPIST UROLOGIST WOMENS CLINIC BMC
U7A  What type of MD is your regular MD/this MD (1=OB/GYN, 2=Family medicine, 3=Pediatri-
cian, 4=Adolescent medicine, 5=Internal medicine, 6=AIDS doctor, 7=Asthma doctor, 8=Pul-
monary doctor, 9=Cardiologist, 10=Gastroen)
U8A  Only saw this person once (=Only saw once)
U8B  Saw this person for < 6 months (1 = <6 months)
U8C  Saw this person for 6 months - 1 year (2 = Between 6 months & 1 year)
U8D  Saw this person for 1-2 years (3 = 1-2 years)
U8E  Saw this person for 3-5 years (4 = 3-5 years)
U8F  Saw this person for more than 5 years (5 = >5 years)
UNEMPLOY Usually unemployed last 6 months (0=No, 1=Yes)
V1  Ever needed to drink much more to get effect (0=No, 1=Yes)
V2  Evr find alcohol had < effect than once did (0=No, 1=Yes)
VT  SF-36 vitality 0-100)
Z1  Breath Alcohol Concentration:1st test
Z2  Breath Alcohol Concentration:2nd test

Details

Eligible subjects were adults, who spoke Spanish or English, reported alcohol, heroin or cocaine
as their first or second drug of choice, resided in proximity to the primary care clinic to which
they would be referred or were homeless. Patients with established primary care relationships they
planned to continue, significant dementia, specific plans to leave the Boston area that would prevent
research participation, failure to provide contact information for tracking purposes, or pregnancy
were excluded.

Subjects were interviewed at baseline during their detoxification stay and follow-up interviews were
undertaken every 6 months for 2 years. A variety of continuous, count, discrete, and survival time
predictors and outcomes were collected at each of these five occasions.

This dataset is a superset of the HELPmiss and HELPrct datasets which include far fewer variables.
Full details of the survey instruments are available at the following link.

Source

https://nhorton.people.amherst.edu/help/
References

See Also
HELPrct, and HELPmiss.

Examples
dataHELPfull)

---

HELPmiss  Health Evaluation and Linkage to Primary Care

Description
The HELP study was a clinical trial for adult inpatients recruited from a detoxification unit. Patients with no primary care physician were randomized to receive a multidisciplinary assessment and a brief motivational intervention or usual care, with the goal of linking them to primary medical care.

Usage
dataHELPmiss)

Format
Data frame with 470 observations on the following variables.

- age  subject age at baseline (in years)
- anysub use of any substance post-detox: a factor with levels no yes
- cesd  Center for Epidemiologic Studies Depression measure of depressive symptoms at baseline (higher scores indicate more symptoms)
- d1  lifetime number of hospitalizations for medical problems (measured at baseline)
- daysanysub  time (in days) to first use of any substance post-detox
- dayslink  time (in days) to linkage to primary care
- drugrisk  Risk Assessment Battery drug risk scale at baseline
- e2b  number of times in past 6 months entered a detox program (measured at baseline)
- female  0 for male, 1 for female
- sex a factor with levels male female
- g1b  experienced serious thoughts of suicide in last 30 days (measured at baseline): a factor with levels no yes
eligible subjects were adults, who spoke spanish or english, reported alcohol, heroin or cocaine as their first or second drug of choice, resided in proximity to the primary care clinic to which they would be referred or were homeless. patients with established primary care relationships they planned to continue, significant dementia, specific plans to leave the boston area that would prevent research participation, failure to provide contact information for tracking purposes, or pregnancy were excluded.

subjects were interviewed at baseline during their detoxification stay and follow-up interviews were undertaken every 6 months for 2 years. a variety of continuous, count, discrete, and survival time predictors and outcomes were collected at each of these five occasions.

this dataset is a superset of the helpRct data with 17 subjects with partially observed data on some of the baseline variables. this is a subset of the helpfull data which includes 5 timepoints and many additional variables.
References


See Also

HELPrct, and HELPfull.

Examples

data(HELPmiss)

---

**Description**

The HELP study was a clinical trial for adult inpatients recruited from a detoxification unit. Patients with no primary care physician were randomized to receive a multidisciplinary assessment and a brief motivational intervention or usual care, with the goal of linking them to primary medical care.

**Usage**

data(HELPrct)

**Format**

Data frame with 453 observations on the following variables.

- **age**: subject age at baseline (in years)
- **anysub**: use of any substance post-detox: a factor with levels no yes
- **cesd**: Center for Epidemiologic Studies Depression measure at baseline (high scores indicate more depressive symptoms)
- **d1**: lifetime number of hospitalizations for medical problems (measured at baseline)
- **hospitalizations**: lifetime number of hospitalizations for medical problems (measured at baseline)
- **daysanysub**: time (in days) to first use of any substance post-detox
- **dayslink**: time (in days) to linkage to primary care
- **drugrisk**: Risk Assessment Battery drug risk scale at baseline
- **e2b**: number of times in past 6 months entered a detox program (measured at baseline)
- **female**: 0 for male, 1 for female
- **sex**: a factor with levels male female
g1b experienced serious thoughts of suicide in last 30 days (measured at baseline): a factor with levels no yes

homeless housing status: a factor with levels housed homeless

i1 average number of drinks (standard units) consumed per day, in the past 30 days (measured at baseline)

i2 maximum number of drinks (standard units) consumed per day, in the past 30 days (measured at baseline)

id subject identifier

indtot Inventory of Drug Use Consequences (InDUC) total score (measured at baseline)

linkstatus post-detox linkage to primary care (0 = no, 1 = yes)

link post-detox linkage to primary care: no yes

mcs SF-36 Mental Component Score (measured at baseline, lower scores indicate worse status)

pcs SF-36 Physical Component Score (measured at baseline, lower scores indicate worse status)

pss_fr perceived social support by friends (measured at baseline, higher scores indicate more support)

racegrp race/ethnicity: levels black hispanic other white

satreat any BSAS substance abuse treatment at baseline: no yes

sexrisk Risk Assessment Battery sex risk score (measured at baseline)

substance primary substance of abuse: alcohol cocaine heroin

treat randomized to HELP clinic: no yes

Details

Eligible subjects were adults, who spoke Spanish or English, reported alcohol, heroin or cocaine as their first or second drug of choice, resided in proximity to the primary care clinic to which they would be referred or were homeless. Patients with established primary care relationships they planned to continue, significant dementia, specific plans to leave the Boston area that would prevent research participation, failure to provide contact information for tracking purposes, or pregnancy were excluded.

Subjects were interviewed at baseline during their detoxification stay and follow-up interviews were undertaken every 6 months for 2 years. A variety of continuous, count, discrete, and survival time predictors and outcomes were collected at each of these five occasions.

This data set is a subset of the HELPmiss data set restricted to the 453 subjects who were fully observed on the age, cesd, d1, female, sex, g1b, homeless, i1, i2, indtot, mcs, pcs, pss_fr, racegrp, satreat, substance, treat, and sexrisk variables. (There is some missingness in the other variables.) HELPmiss contains 17 additional subjects with partially observed data on some of these baseline variables. This is also a subset of the HELPfull data which includes 5 timepoints and many additional variables.

Note

The code HELPPrct data set was originally named code HELP but has been renamed to avoid confusion with the code help function.
Source

https://nhorton.people.amherst.edu/help/

References


See Also

HELPmiss, and HELPfull.

Examples

```r
data(KidsFeet)
```

---

**KidsFeet**  
*Foot measurements in children*

**Description**

These data were collected by a statistician, Mary C. Meyer, in a fourth grade classroom in Ann Arbor, MI, in October 1997. They are a convenience sample — the kids who were in the fourth grade.

**Usage**

```r
data(KidsFeet)
```

**Format**

A data frame with 39 observations on the following variables.

- **name**: a factor with levels corresponding to the name of each child
- **birthmonth**: the month of birth
- **birthyear**: the year of birth
- **length**: length of longer foot (in cm)
- **width**: width of longer foot (in cm)
- **sex**: a factor with levels B G
- **biggerfoot**: a factor with levels L R
- **domhand**: a factor with levels L R
Details

Quoted from the source: "From a very young age, shoes for boys tend to be wider than shoes for girls. Is this because boys have wider feet, or because it is assumed that girls, even in elementary school, are willing to sacrifice comfort for fashion? To assess the former, a statistician measures kids’ feet."

References


Examples

data(KidsFeet)

---

Marriage  
Marriage records

Description

Marriage records from the Mobile County, Alabama, probate court.

Usage

data(Marriage)

Format

A data frame with 98 observations on the following variables.

bookpageID  a factor with levels for each book and page (unique identifier)
appdate  date on which the application was filed
ceremonydate  date of the ceremony
delay  number of days between the application and the ceremony
officialTitle  a factor with levels BISHOP CATHOLIC PRIEST CHIEF CLERK CIRCUIT JUDGE ELDER MARRIAGE OFFICIAL MINISTER PASTOR REVEREND
person  a factor with levels Bride Groom
dob  a factor with levels corresponding to the date of birth of the person
age  age of the person (in years)
race  a factor with levels American Indian Black Hispanic White
prevcount  the number of previous marriages of the person, as listed on the application
prevconc  the way the last marriage ended, as listed on the application
hs  the number of years of high school education, as listed on the application
college  the number of years College education, as listed on the application. Where no number was listed, this field was left blank, unless less than 12 years High School was reported, in which case it was entered as 0.

dayOfBirth  the day of birth, as a number from 1 to 365 counting from January 1

sign  the astrological sign, with levels Aquarius Aries Cancer Capricorn Gemini Leo Libra Pisces Sagittarius Scorpio Taurus Virgo

Details

The calculation of the astrological sign may not correctly sort people directly on the borders between signs. This variable is not part of the original record.

Source

The records were collected through http://www.mobilecounty.org/probatecourt/recordssearch.htm

Examples

data(Marriage)

<table>
<thead>
<tr>
<th>Mites</th>
<th>Mites and Wilt Disease</th>
</tr>
</thead>
</table>

Description

Data from an experiment to test whether exposure to mites protects against Wilt Disease in cotton plants.

Usage

data(Mites)

Format

A data frame with 47 observations on the following variables.

- treatment  a factor with levels mites and no mites
- outcome  a factor with levels wilt and no wilt

Details

Researchers suspected that attack of a plant by one organism induced resistance to subsequent attack by a different organism. Individually potted cotton plants were randomly allocated to two groups: infestation by spider mites or no infestation. After two weeks the mites were dutifully removed by a conscientious research assistant, and both groups were inoculated with Verticillium, a fungus that causes Wilt disease. More information can be found at https://www.causeweb.org/cause/webinar/activity/2010-01/.
Source


Examples

```r
data(Mites)
if (require(mosaic)) {
  tally(~ treatment + outcome, data=Mites)
  tally(~ outcome | treatment, format="percent", data=Mites)
}
```

---

**RailTrail**

*Volume of Users of a Rail Trail*

**Description**

The Pioneer Valley Planning Commission (PVPC) collected data north of Chestnut Street in Florence, MA for ninety days from April 5, 2005 to November 15, 2005. Data collectors set up a laser sensor, with breaks in the laser beam recording when a rail-trail user passed the data collection station.

**Usage**

```r
data(RailTrail)
```

**Format**

A data frame with 90 observations on the following variables.

- `hightemp` daily high temperature (in degrees Fahrenheit)
- `lowtemp` daily low temperature (in degrees Fahrenheit)
- `avgtemp` average of daily low and daily high temperature (in degrees Fahrenheit)
- `spring` indicator of whether the season was Spring
- `summer` indicator of whether the season was Summer
- `fall` indicator of whether the season was Fall
- `cloudcover` measure of cloud cover (in oktas)
- `precip` measure of precipitation (in inches)
- `volume` estimated number of trail users that day (number of breaks recorded)
- `weekday` logical indicator of whether the day was a non-holiday weekday
- `dayType` one of "weekday" or "weekend"
Details

There is a potential for error when two users trigger the infrared beam at exactly the same time since the counter would only logs one of the crossings. The collectors left the motion detector out during the winter, but because the counter drops data when the temperature falls below 14 degrees Fahrenheit, there is no data for the cold winter months.

Source

Pioneer Valley Planning Commission

References

http://www.fvgreenway.org/pdfs/Northampton-Bikepath-Volume-Counts%20_05_LTA.pdf

Examples

data(RailTrail)

---

Riders Volume of Users of a Massachusetts Rail Trail

Description

The Pioneer Valley Planning Commission (PVPC) collected data north of Chestnut Street in Florence, MA for ninety days from April 5, 2005 to November 15, 2005. Data collectors set up a laser sensor, with breaks in the laser beam recording when a rail-trail user passed the data collection station.

Usage

data(Riders)

Format

A data frame with 90 observations on the following 12 variables.

date date of data collection (POSIXct)
day a factor with levels Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday.
highT high temperature for the day (in degrees Fahrenheit)
lowT low temperature for the day (in degrees Fahrenheit)
hi shorter name for highT
lo shorter name for lowT
precip inches of precipitation
clouds measure of cloud cover (in oktas)
riders estimated number of trail crossings that day (number of breaks recorded)
cr shorter name for riders
weekday type of day: a factor with levels N (weekend or holiday) Y (non-holiday weekday)
wday shorter name for weekday

Details

There is a potential for error when two users trigger the infrared beam at exactly the same time since the counter would only log one of the crossings. The collectors left the motion detector out during the winter, but because the counter drops data when the temperature falls below 14 degrees Fahrenheit, there are no data for the coldest winter months.

Source


References


Examples

data(Riders)
str(Riders)

SaratogaHouses

Houses in Saratoga County (2006)

Description

Data on houses in Saratoga County, New York, USA in 2006

Usage

data(SaratogaHouses)

Format

A data frame with 1728 observations on the following 16 variables.

price price (US dollars)
lotSize size of lot (acres)
age age of house (years)
landValue value of land (US dollars)
livingArea  living are (square feet)
pctCollege  percent of neighborhood that graduated college
bedrooms  number of bedrooms
fireplaces  number of fireplaces
bathrooms  number of bathrooms (half bathrooms have no shower or tub)
rooms  number of rooms
heating  type of heating system
fuel  fuel used for heating
sewer  type of sewer system
waterfront  whether property includes waterfront
newConstruction  whether the property is a new construction
centralAir  whether the house has central air

Source

Data collected by Candice Corvetti and used in the "Stat 101" case study "How much is a Fireplace Worth". See also https://www.saratogacountyny.gov/departments/real-property-tax-service-agency/

Description

SAT data assembled for a statistics education journal article on the link between SAT scores and measures of educational expenditures

Usage

data(SAT)

Format

A data frame with 50 observations on the following variables.

state  a factor with names of each state
expend  expenditure per pupil in average daily attendance in public elementary and secondary schools, 1994-95 (in thousands of US dollars)
ratio  average pupil/teacher ratio in public elementary and secondary schools, Fall 1994
salary  estimated average annual salary of teachers in public elementary and secondary schools, 1994-95 (in thousands of US dollars)
frac  percentage of all eligible students taking the SAT, 1994-95
verbal  average verbal SAT score, 1994-95
math  average math SAT score, 1994-95
sat  average total SAT score, 1994-95
Source

http://www.amstat.org/publications/jse/secure/v7n2/datasets.guber.cfm

References


Examples

```r
data(SAT)
if (require(ggformula)) {
  gf_point(sat ~ expend, data = SAT, color = "blue", alpha = 0.5) |
  gf_lm()
  gf_text(sat ~ expend, data = SAT, label = ~ abbreviate(SAT$state, 3),
          inherit = FALSE)
}
```

SnowGR

*Snowfall data for Grand Rapids, MI*

Description

Official snowfall data by month and season for Grand Rapids, MI, going back to 1893.

Usage

```r
data(SnowGR)
```

Format

A data frame with 119 observations of the following variables.

- **SeasonStart**: Year in which season started (July is start of season)
- **SeasonEnd**: Year in which season ended (June is end of season)
- **Jul**: Inches of snow in July
- **Aug**: Inches of snow in August
- **Sep**: Inches of snow in September
- **Oct**: Inches of snow in October
- **Nov**: Inches of snow in November
- **Dec**: Inches of snow in December
- **Jan**: Inches of snow in January
- **Feb**: Inches of snow in February
- **Mar**: Inches of snow in March
- **Apr**: Inches of snow in April
SwimRecords

100 m Swimming World Records

Description

World records for men and women over time from 1905 through 2004.

Usage

data(SwimRecords)

Format

A data frame with 62 observations of the following variables.

- time: time (in seconds) of the world record
- year: Year in which the record was set
- sex: a factor with levels M and F

May  Inches of snow in May
Jun  Inches of snow in June
Total  Inches of snow for entire season (July-June)

Source

These data were compiled by Laura Kapitula from data available from NOAA. The original URL used (http://www.crh.noaa.gov/grr/climate/data/grr/snowfall/) is no longer in service.

Examples

data(SnowGR)
if (require(ggformula)) {
  df_stats(~ Total, data = SnowGR)
gf_histogram(~ Total, data = SnowGR)
gf_point(Total ~ SeasonStart, data = SnowGR) |>  
gf_smooth()
}

if (require(tidyverse) & require(dplyr)) {
  Snow2 <- SnowGR |>  
pivot_longer(Jul:Total, names_to = "month", values_to = "snowfall") |>  
filter(month != "Total") |>  
mutate(month = factor(month, levels = unique(month)))
gf_violin(snowfall ~ month, data = Snow2, scale = "width")
}
Examples

```r
data(SwimRecords)
if (require(ggformula)) {

gf_point(time ~ year, data = SwimRecords, color = ~ sex)
}
```

---

**TenMileRace**

---

**Cherry Blossom Race**

Description

The Cherry Blossom 10 Mile Run is a road race held in Washington, D.C. in April each year. (The name comes from the famous cherry trees that are in bloom in April in Washington.) The results of this race are published. This data frame contains the results from the 2005 race.

Usage

```r
data(TenMileRace)
```

Format

A data frame with 8636 observations on the following variables.

- **state**: State of residence of runner.
- **time**: Official time from starting gun to finish line.
- **net**: The recorded time (in seconds) from when the runner crossed the starting line to when the runner crossed the finish line. This is generally less than the official time because of the large number of runners in the race: it takes time to reach the starting line after the gun has gone off.
- **age**: Age of runner in years.
- **sex**: A factor with levels `F` `M`.

Examples

```r
data(TenMileRace)
if (require(ggformula)) {

gf_point(net ~ age | sex, data = TenMileRace, color = ~sex, alpha = 0.1) |> 
gf_density2d(color = "gray40")

lm(net ~ age + sex, data = TenMileRace)
}
```
Utilities

Utility bills

Description

Data from utility bills at a residence. Utilities2 is a similar data set with some additional variables.

Usage

data(Utilities)

Format

A data frame containing 117 observations for the following variables.

- month  month (coded as a number)
- day    day of month on which bill was calculated
- year   year of bill
- temp   average temperature (F) for billing period
- kwh    electricity usage (kwh)
- ccf    gas usage (ccf)
- thermsPerDay  a numeric vector
- billingDays  number of billing days in billing period
- totalbill  total bill (in dollars)
- gasbill   gas bill (in dollars)
- elecbill  electric bill (in dollars)
- notes    notes about the billing period

Source


See Also

Utilities2.

Examples

data(Utilities)
if (require(ggformula)) {
  gf_point(gasbill ~ temp, data = Utilities)
}
Utilities2

Utility bills

Description

Data from utility bills at a private residence. This is an augmented version of Utilities.

Usage

data(Utilities2)

Format

A data frame containing 117 observations for the following variables.

- `month`  month (coded as a number)
- `day`    day of month on which bill was calculated
- `year`   year of bill
- `temp`   average temperature (F) for billing period
- `kwh`    electricity usage (kwh)
- `ccf`    gas usage (ccf)
- `thermsPerDay` a numeric vector
- `billingDays` number of billing days in billing period
- `totalbill` total bill (in dollars)
- `gasbill` gas bill (in dollars)
- `elecbill` electric bill (in dollars)
- `notes` notes about the billing period
- `ccfpday` average gas usage per day (Utilities2 only)
- `kwhpday` average electric usage per day (Utilities2 only)
- `gasbillpday` gas bill divided by billing days (Utilities2 only)
- `elecbillpday` electric bill divided by billing days a numeric vector (Utilities2 only)
- `totalbillpday` total bill divided by billing days a numeric vector (Utilities2 only)
- `therms` `thermsPerDay` * `billingDays` (Utilities2 only)
- `monthsSinceY2K` months since 2000 (Utilities2 only)

Source


See Also

Utilities.
Examples

data(Utilities2)
if (require(ggformula)) {
  gf_point(gasbillpday ~ temp, data = Utilities2)
}

Description

2016-17 weather in several cities

Usage

data(Weather)

Format

A data frame with weather-related variables for several world cities.

city  City name.
date  Date.
year  Numeric year.
month Numeric month.
day  Numeric day.
high_temp, avg_temp, low_temp  High, average, and low temperature for the day in degrees F.
high_dewpt, avg_dewpt, low_dewpt  High, average, and low dew point for the day in degrees F.
high_humidity, avg_humidity, low_humidity  High, average, and low relative humidity.
high_hg, avg_hg, low_hg  High, average, and low sea level pressure in inches of mercury.
high_vis, avg_vis, low_vis  High, average, and low visibility for the day in miles.
high_wind, avg_wind, low_wind  High, average, and low wind speed for the day in mph.
precip  Precipitation for the day – a character value; T means "trace amount".
events  Character string naming weather events on the day (Rain, Fog, Snow, etc.)

Source

These data were downloaded from WeatherUnderground in January 2018.
Examples

```r
if (require(dplyr)) {
  Weather |>
  group_by(city, year) |>
  summarise(
    min_temp = min(low_temp),
    max_temp = max(high_temp)
  )
}

if (require(ggformula)) {
  Weather |>
  gf_linerange(low_temp + high_temp ~ date | city ~ ..,
    color = ~ (high_temp + low_temp) / 2, show.legend = FALSE) |>
  gf_refine(scale_color_gradientn(colors = rev(rainbow(5))))
}
```

Whickham

Data from the Whickham survey

Description

Data on age, smoking, and mortality from a one-in-six survey of the electoral roll in Whickham, a mixed urban and rural district near Newcastle upon Tyne, in the UK. The survey was conducted in 1972-1974 to study heart disease and thyroid disease. A follow-up on those in the survey was conducted twenty years later.

Usage

```r
data(Whickham)
```

Format

A data frame with 1314 observations on women for the following variables:

- **outcome**: survival status after 20 years: a factor with levels Alive Dead
- **smoker**: smoking status at baseline: a factor with levels No Yes
- **age**: age (in years) at the time of the first survey

Details

This dataset contains a subset of the survey sample: women who were classified as current smokers or as never having smoked. The data were synthesized from the summary description tables given in the Appleton et al al paper.
References


Examples

data(Whickham)
Index

* datasets

- Cards, 6
- CoolingWater, 7
- Countries, 8
- Galton, 10
- Gestation, 11
- HeatX, 14
- HELPfull, 15
- HELPmiss, 41
- HELPrct, 43
- KidsFeet, 45
- Marriage, 46
- Mites, 47
- SAT, 51
- SnowGR, 52
- SwimRecords, 53
- TenMileRace, 54
- Utilities, 55
- Utilities2, 56
- Weather, 57
- Whickham, 58

- Alcohol, 2
- Birthdays, 3
- Births, 4
- Births2015 (Births), 4
- Births78 (Births), 4
- BirthsCDC (Births), 4
- BirthsSSA (Births), 4

- card (Cards), 6
- Cards, 6
- CoolingWater, 7
- Countries, 8
- CPS85, 8

- Dimes, 10
- Galton, 10
- Gestation, 11

- GoosePermits, 13
- HeatX, 14
- HELPfull, 15, 43–45
- HELPmiss, 41, 44, 45
- HELPrct, 41, 43, 43
- KidsFeet, 45
- Marriage, 46
- Mites, 47
- RailTrail, 48
- Riders, 49
- SaratogaHouses, 50
- SAT, 51
- SnowGR, 52
- SwimRecords, 53
- TenMileRace, 54
- Utilities, 55, 56
- Utilities2, 55, 56
- Weather, 57
- Whickham, 58