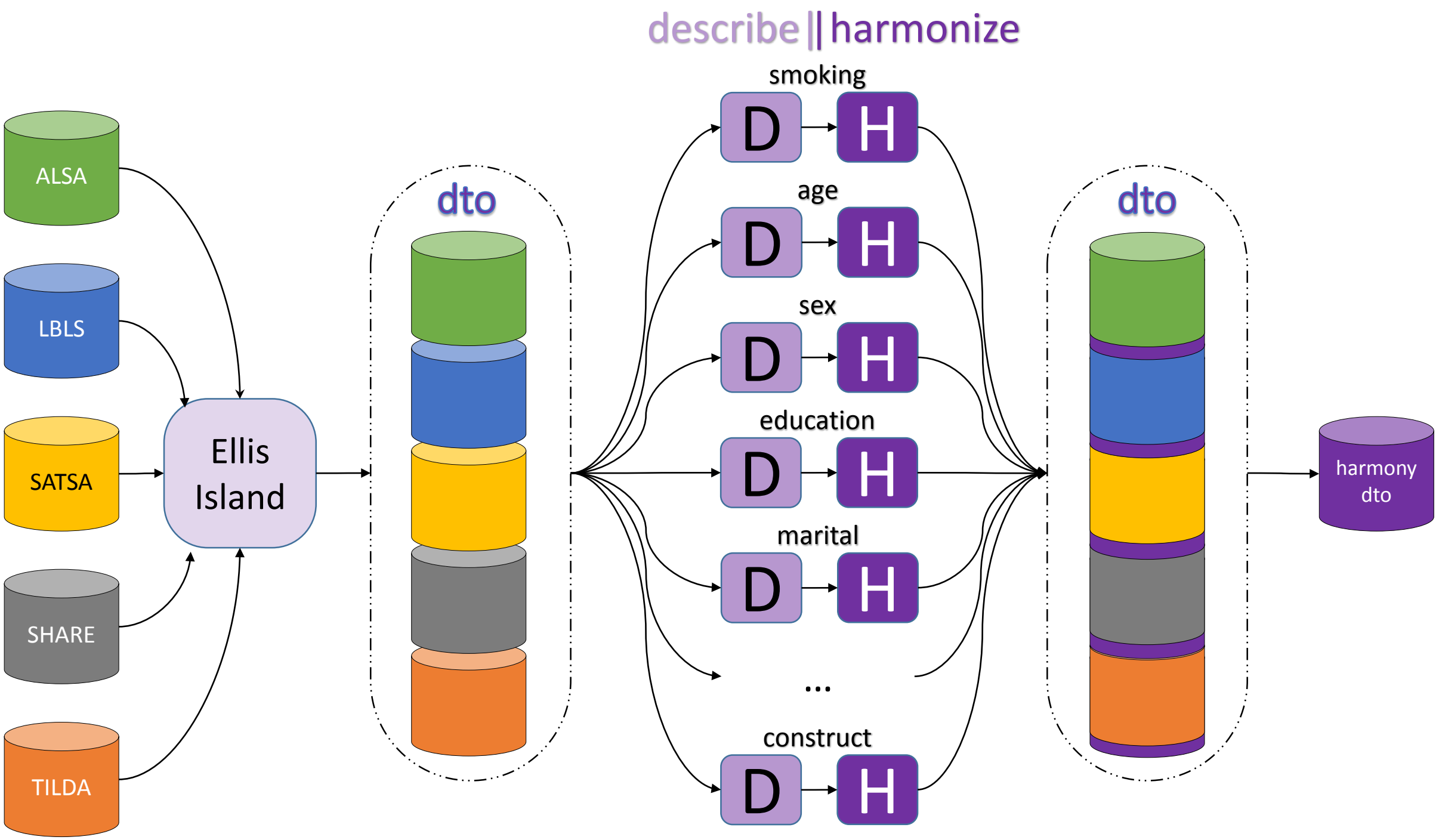
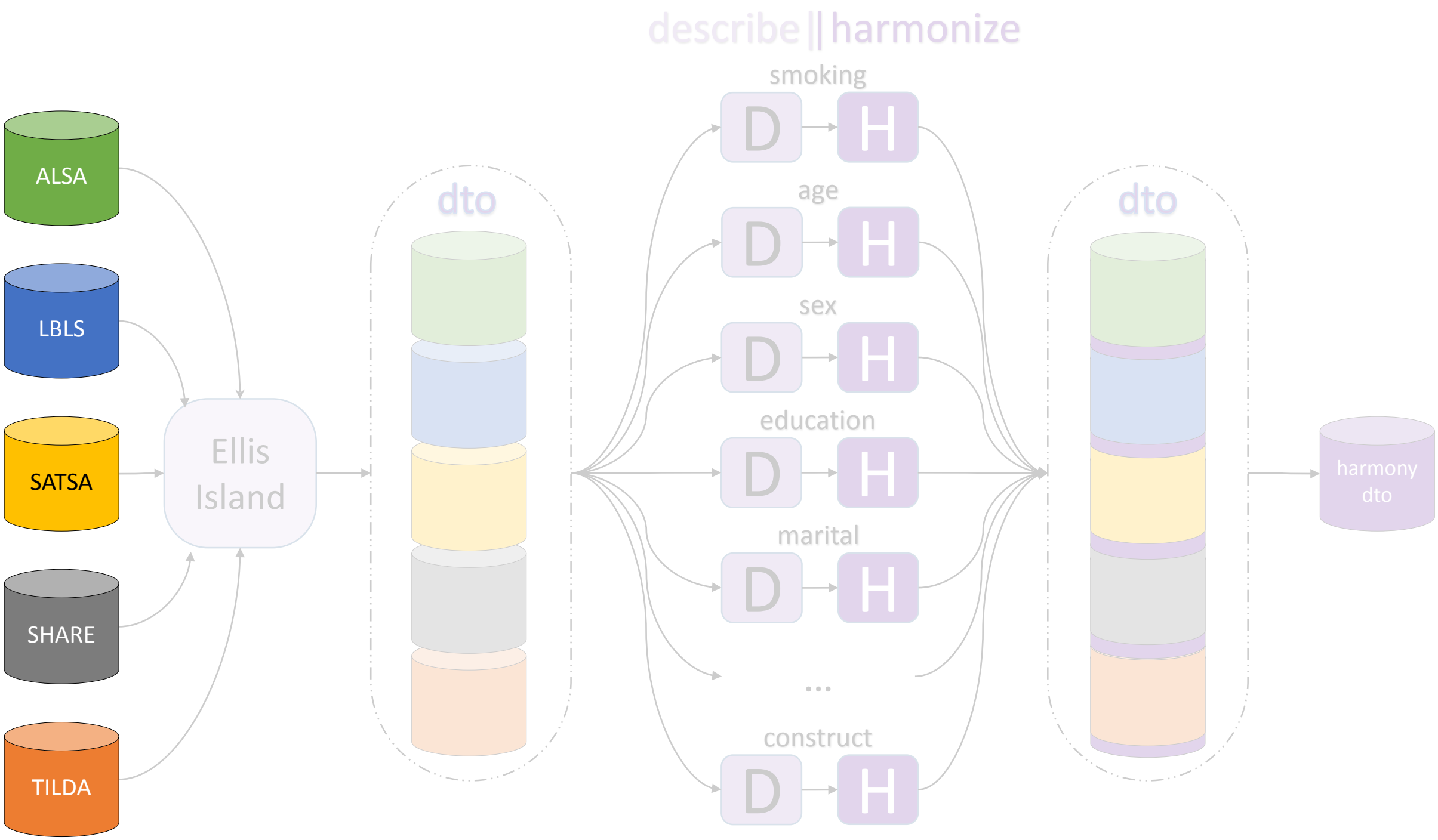
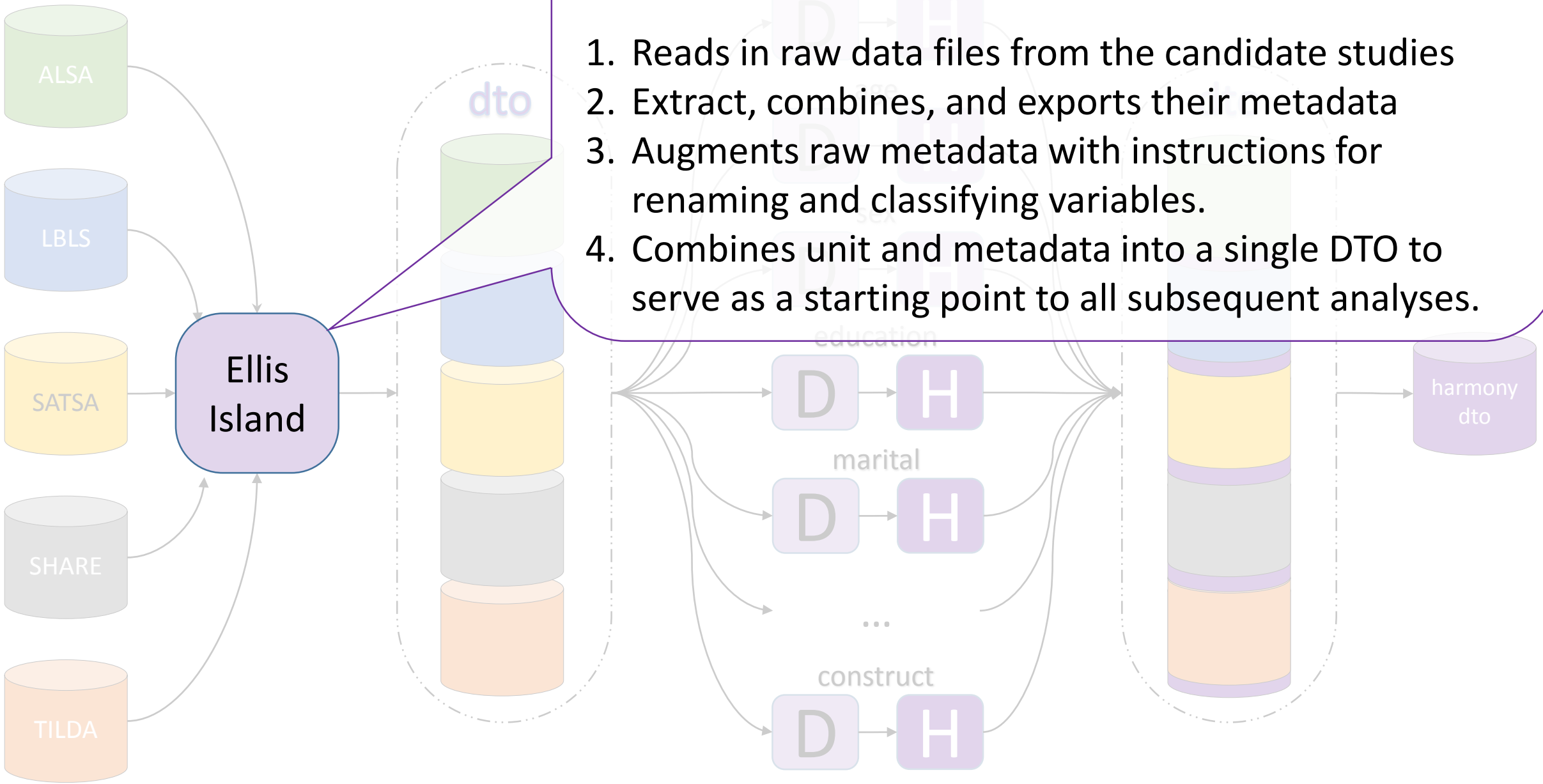
5 voices, 10 stories:

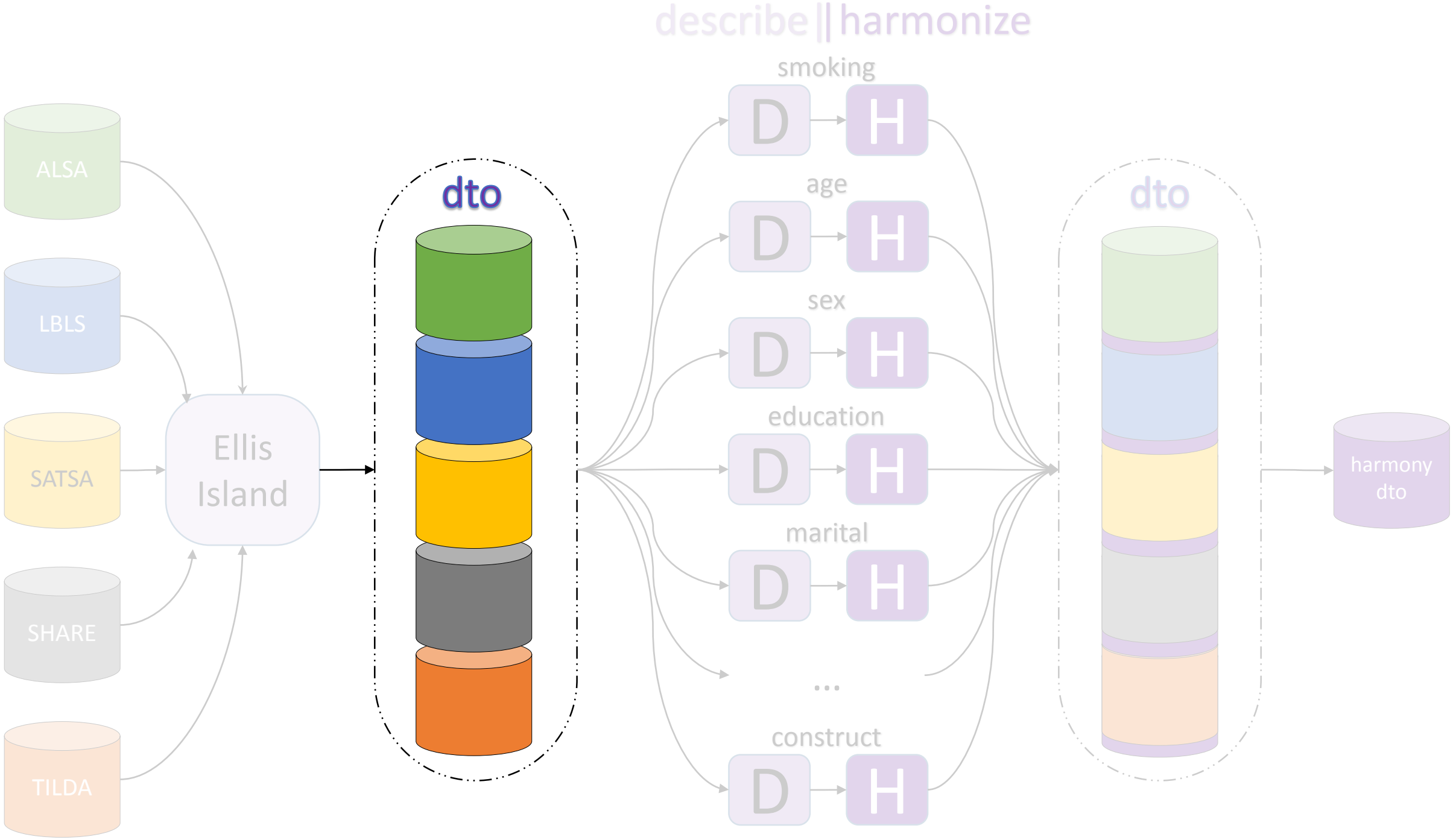
Groningen Harmonization Exercise

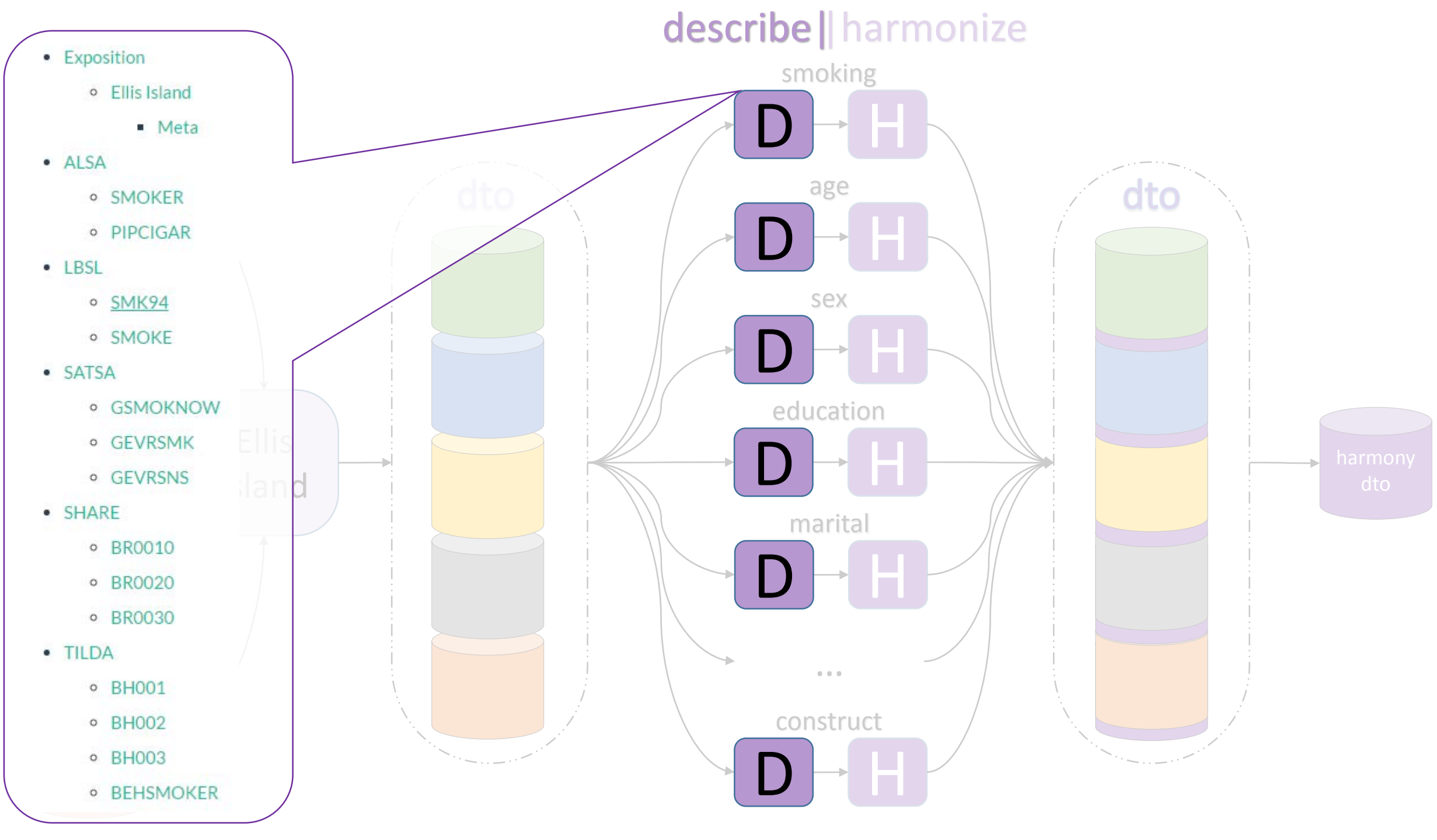
IALSA
Andrey Koval
2016-04-21
Groningen, NL

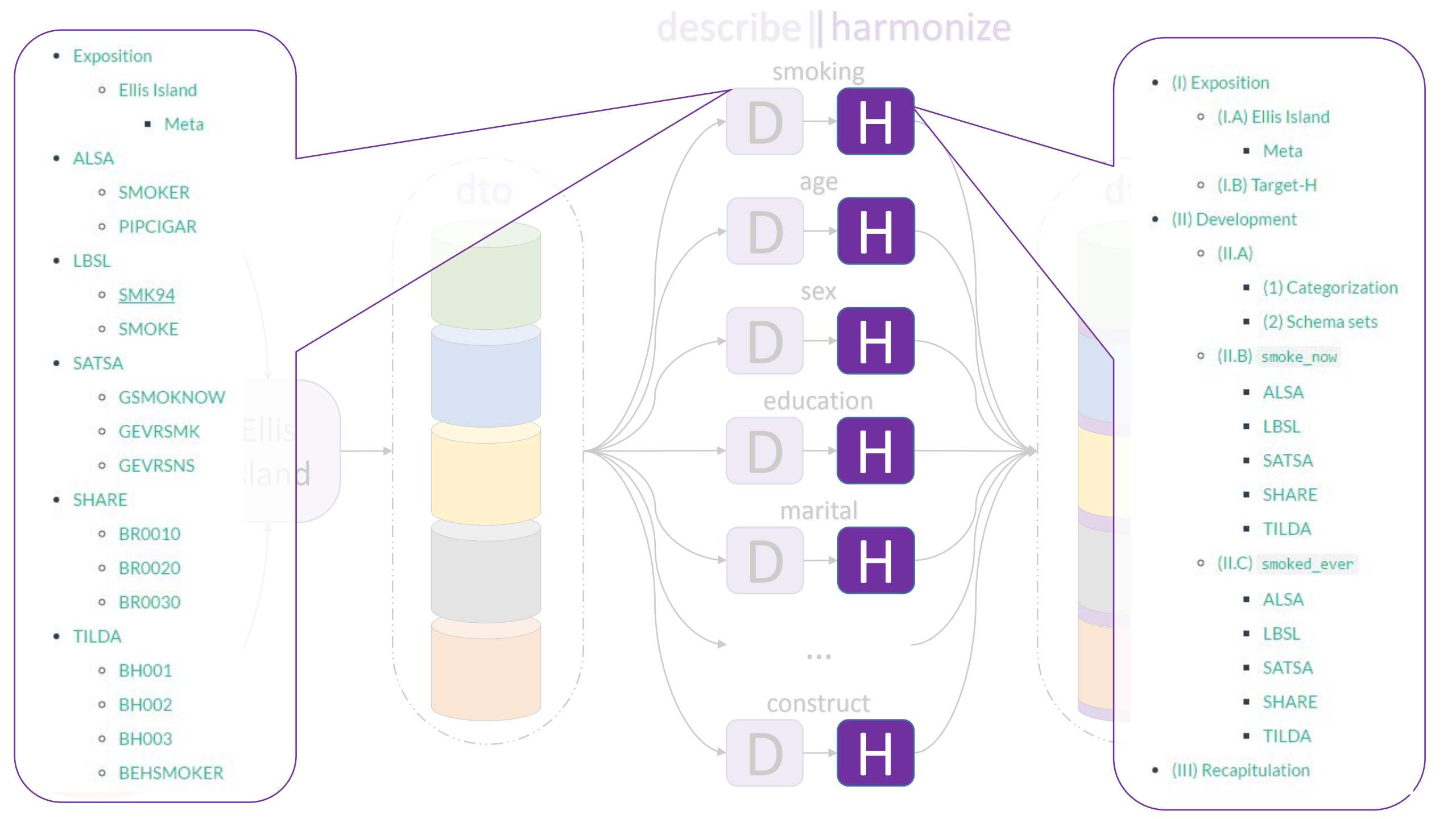

describe || harmonize
ALSA
LBLS
SATSA
SHARE
TILDA
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
dto
harmony dto


describe || harmonize
ALSA
LBLS
SATSA
SHARE
TILDA
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
dto
harmony dto


./manipulation/0-ellis-island.R
1. Reads in raw data files from the candidate studies
2. Extract, combines, and exports their metadata
3. Augments raw metadata with instructions for renaming and classifying variables.
4. Combines unit and metadata into a single DTO to serve as a starting point to all subsequent analyses.
ALSA
LBLS
SATSA
SHARE
TILDA
Ellis Island
dto
education
marital
construct
D
H
harmony dto

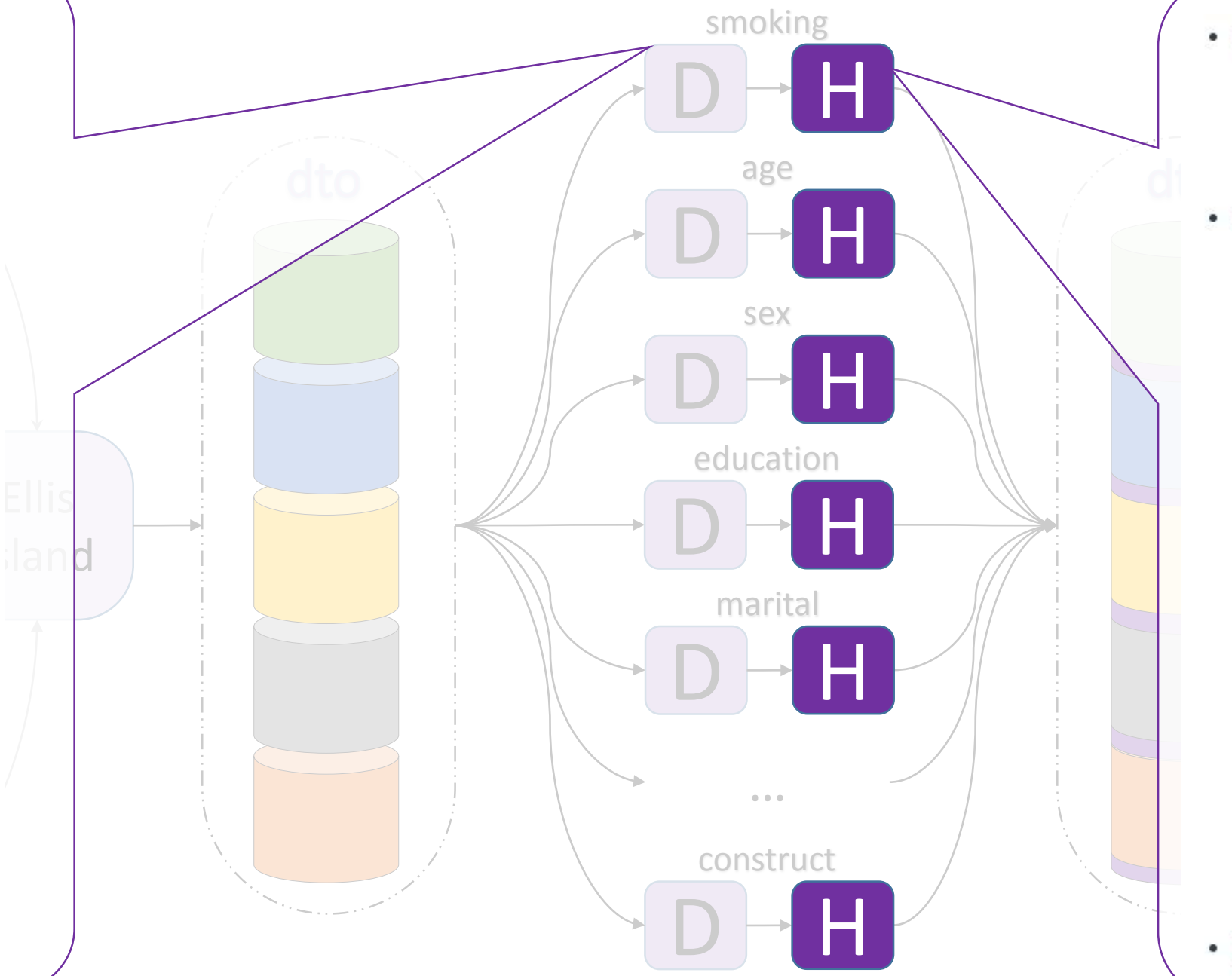

describe || harmonize
ALSA
LBLS
SATSA
SHARE
TILDA
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
dto
harmony dto


describe || harmonize
Exposition
Ellis Island
Meta
ALSA
SMOKER
PIPCIGAR
LBSL
SMK94
SMOKE
SATSA
GSMOKNOW
GEVRSMK
GEVRSNS
SHARE
BR0010
BR0020
BR0030
TILDA
BH001
BH002
BH003
BEHSMOKER
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
dto
harmony dto


describe || harmonize
Exposition
Ellis Island
Meta
ALSA
SMOKER
PIPCIGAR
LBSL
SMK94
SMOKE
SATSA
GSMOKNOW
GEVRSMK
GEVRSNS
SHARE
BR0010
BR0020
BR0030
TILDA
BH001
BH002
BH003
BEHSMOKER
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
(I) Exposition
(I.A) Ellis Island
Meta
(I.B) Target-H
(II) Development
(II.A)
(1) Categorization
(2) Schema sets
(II.B) smoke_now
ALSA
LBSL
SATSA
SHARE
TILDA
(II.C) smoked_ever
ALSA
LBSL
SATSA
SHARE
TILDA
(III) Recapitulation

prelude-describe || harmonize-sonata

- Exposition
  - Ellis Island
    - Meta
- ALSA
  - SMOKER
  - PIPCIGAR
- LBSL
  - SMK94
  - SMOKE
- SATSA
  - GSMOKNOW
  - GEVRSMK
  - GEVRSNS
- SHARE
  - BR0010
  - BR0020
  - BR0030
- TILDA
  - BH001
  - BH002
  - BH003
  - BEHSMOKER



- (I) Exposition
  - (I.A) Ellis Island
    - Meta
  - (I.B) Target-H
- (II) Development
  - (II.A)
    - (1) Categorization
    - (2) Schema sets
  - (II.B) `smoke_now`
    - ALSA
    - LBSL
    - SATSA
    - SHARE
    - TILDA
  - (II.C) `smoked_ever`
    - ALSA
    - LBSL
    - SATSA
    - SHARE
    - TILDA
- (III) Recapitulation

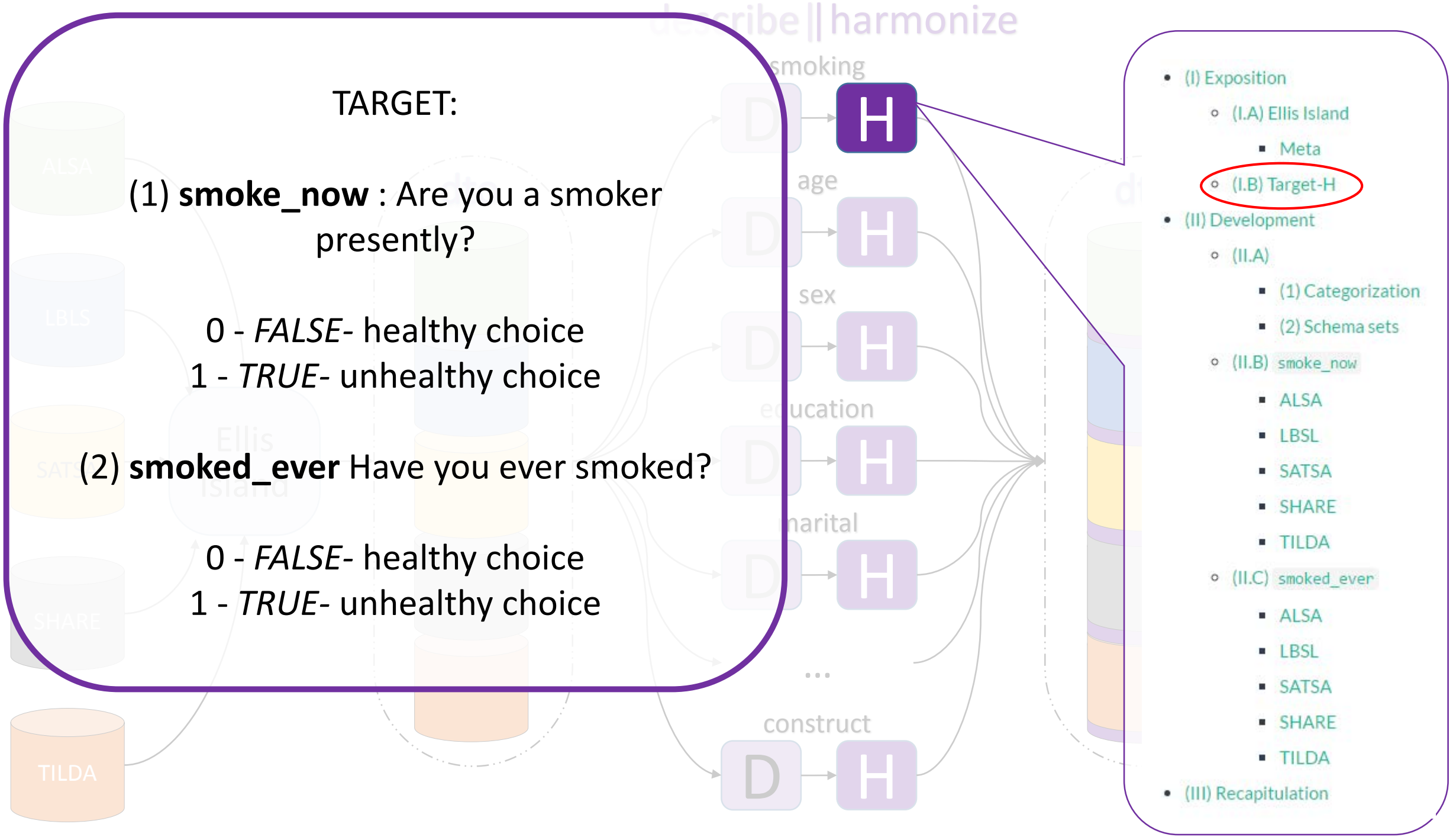
TARGET:

(1) **smoke_now** : Are you a smoker presently?

0 - *FALSE*- healthy choice
1 - *TRUE*- unhealthy choice

(2) **smoked_ever** Have you ever smoked?

0 - *FALSE*- healthy choice
1 - *TRUE*- unhealthy choice

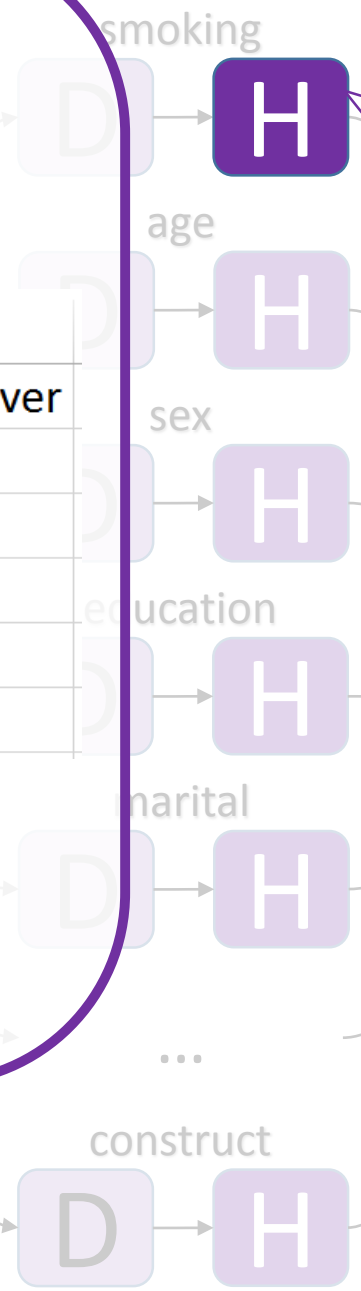


describe || harmonize

ALSA

	A	B	C	D	E	F
1		SMOKER	PIPCIGAR	count	smoke_now	smoked_ever
2		Yes	Yes	7	TRUE	TRUE
3		Yes	No	169	TRUE	TRUE
4		No	Yes	41	TRUE	TRUE
5		No	NA	1851	FALSE	FALSE
6		NA	NA	19	NA	NA

./data/meta/h-rules



- (I) Exposition
  - (I.A) Ellis Island
    - Meta
  - (I.B) Target-H
- (II) Development
  - (II.A)
    - (1) Categorization
    - (2) Schema sets
  - (II.B) `smoke_now`
    - ALSA
    - LBSL
    - SATSA
    - SHARE
    - TILDA
  - (II.C) `smoked_ever`
    - ALSA
    - LBSL
    - SATSA
    - SHARE
    - TILDA
- (III) Recapitulation

ialsa-2016-groningen

Maelstrom Harmonization Workshop. Assessing the impact of different harmonization procedures on the analysis results from several real datasets.

Welcome - letter of invitation - list of participants - summary of the agenda - directions and locations

Documentation : - Maelstrom Obiba Wiki - source data objects - resources and references - reproducibility instructions

Focus

> Habits, prevalence of, and factors associated with smoking among older adults.

Overview

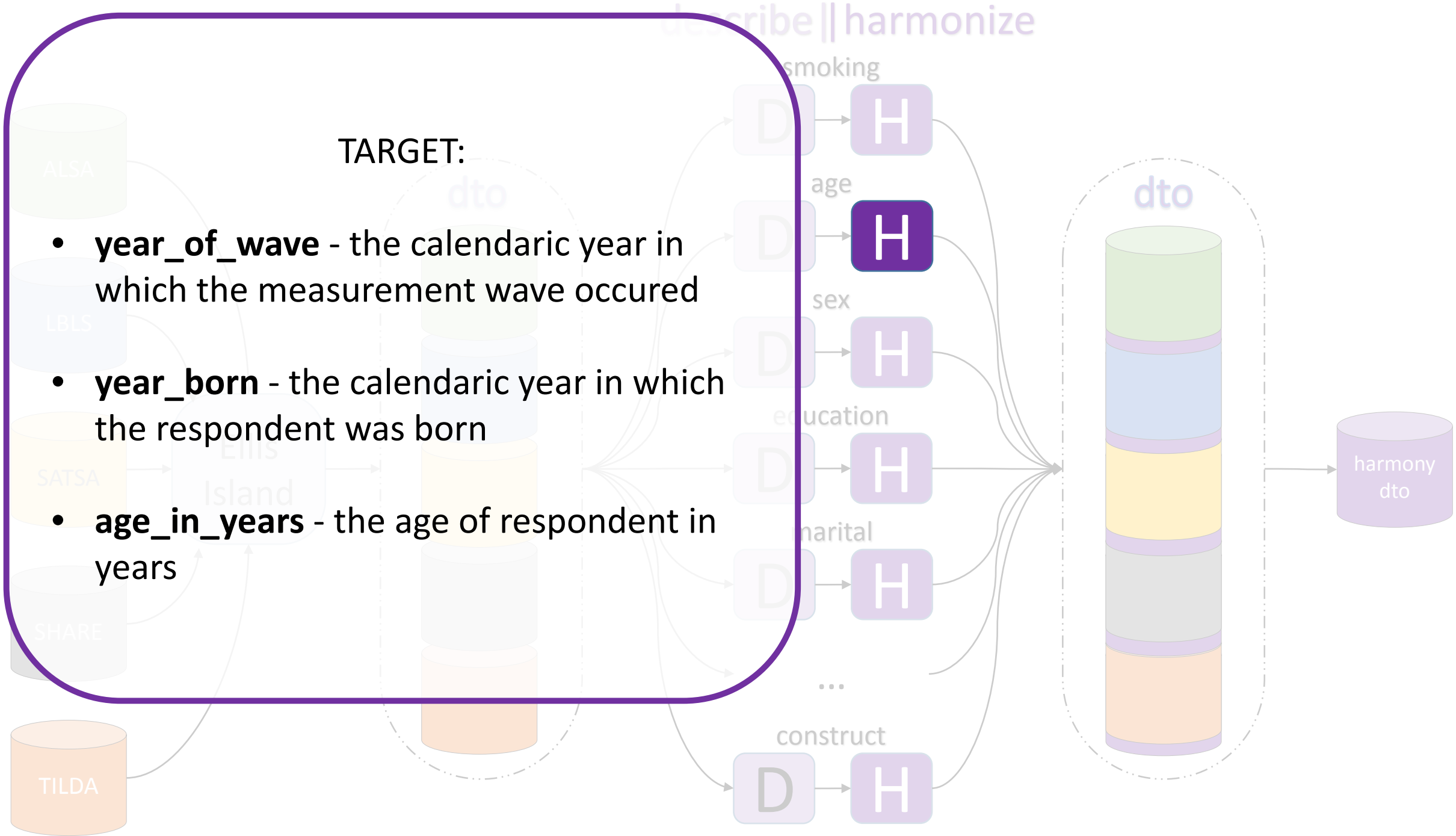
- Full list of variables and their labels used in the exercise, as extracted from the raw source files. This file was augmented with additional meta-data.
- Data provisioning documentation : the Ellis Island report.
- Harmonization rules : mapping pattern of responses to values of the harmonized variables.

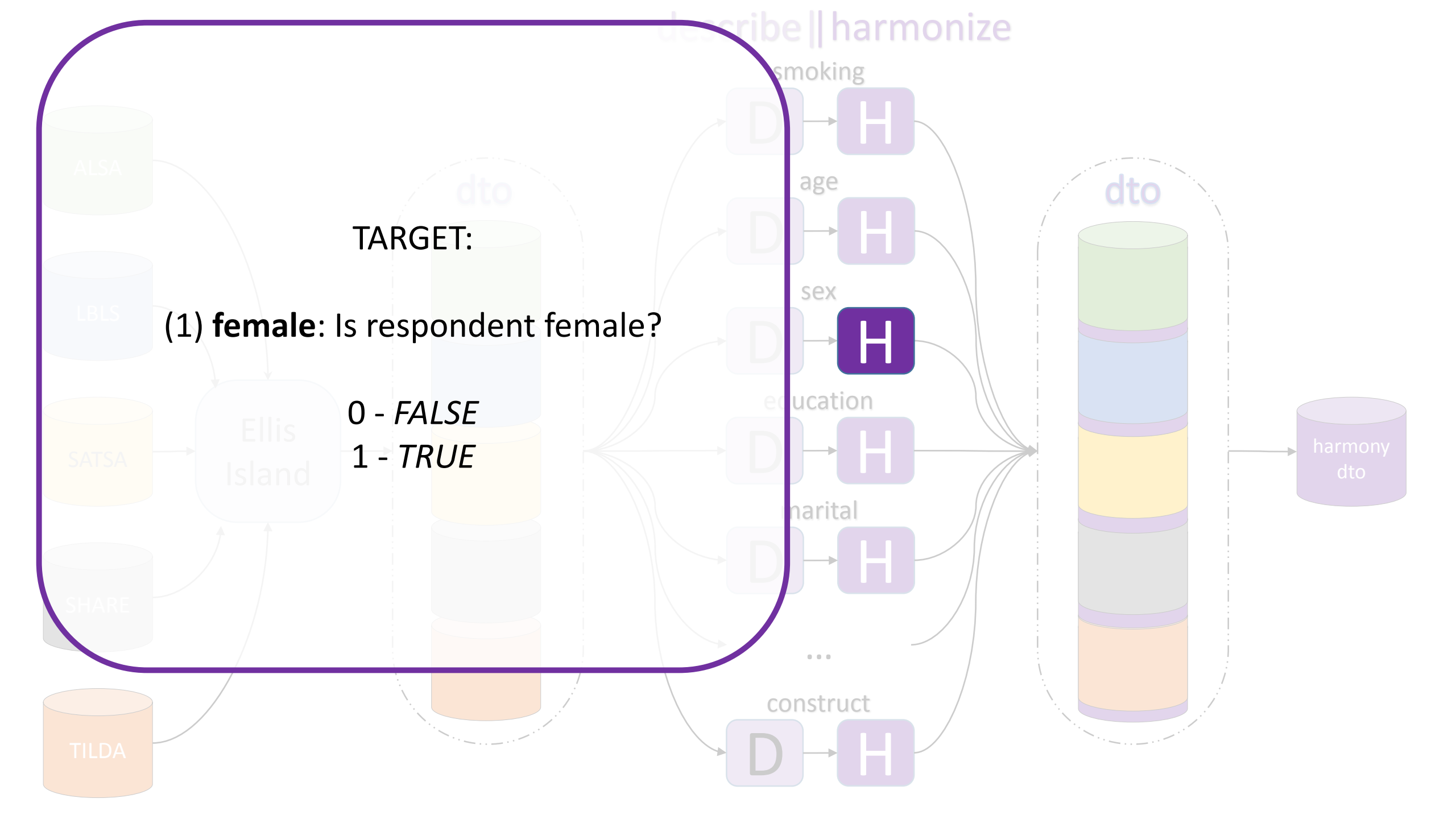
Measures and Harmonization

- Current smoking status (describe, harmonize)
- age (describe, harmonize)
- sex (describe, harmonize)
- marital status (describe, harmonize)
- educational level (describe, harmonize)
- working status (describe, harmonize)
- alcohol consumption (describe, harmonize)
- level of physical activity (describe, harmonize)
- perceived health (describe, harmonize)
- physique (describe, harmonize)

TARGET:

- **year_of_wave** - the calendaric year in which the measurement wave occured
- **year_born** - the calendaric year in which the respondent was born
- **age_in_years** - the age of respondent in years





TARGET:

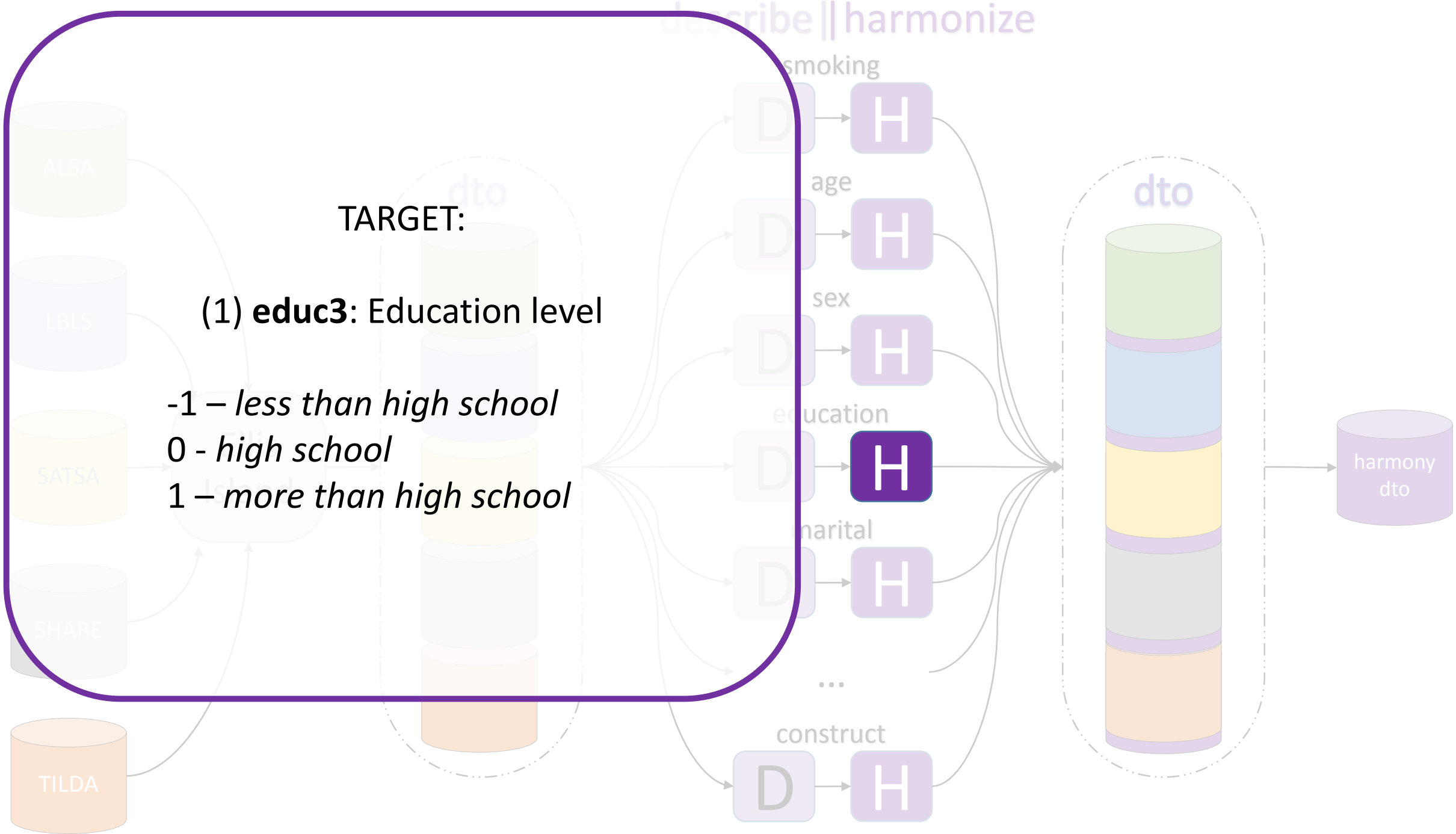
(1) **female**: Is respondent female?

0 - *FALSE*
1 - *TRUE*

TARGET:

(1) **educ3**: Education level

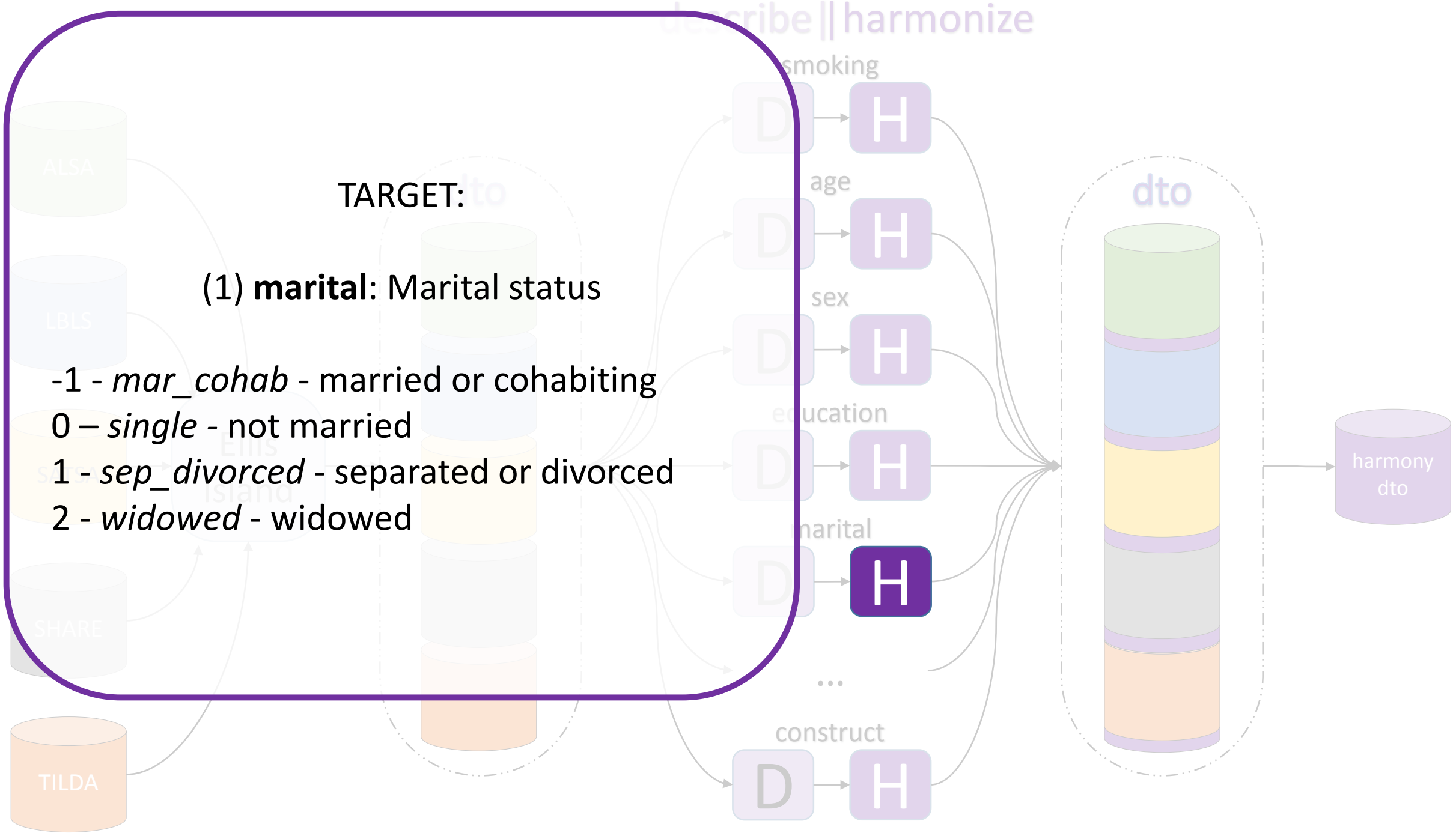
-1 – *less than high school*
0 - *high school*
1 – *more than high school*

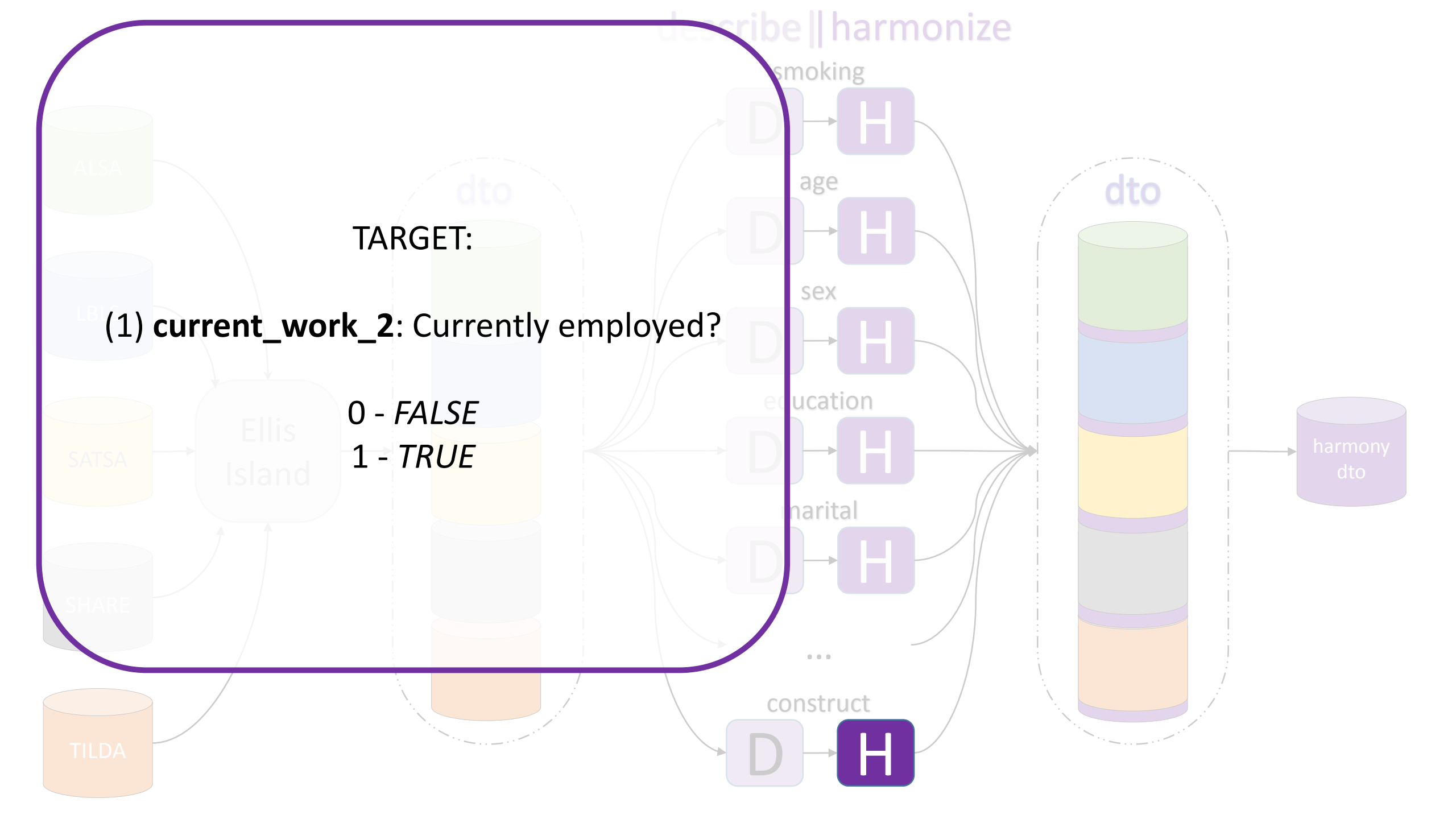


TARGET:

(1) **marital**: Marital status

-1 - *mar_cohab* - married or cohabiting
0 – *single* - not married
1 - *sep_divorced* - separated or divorced
2 - *widowed* - widowed

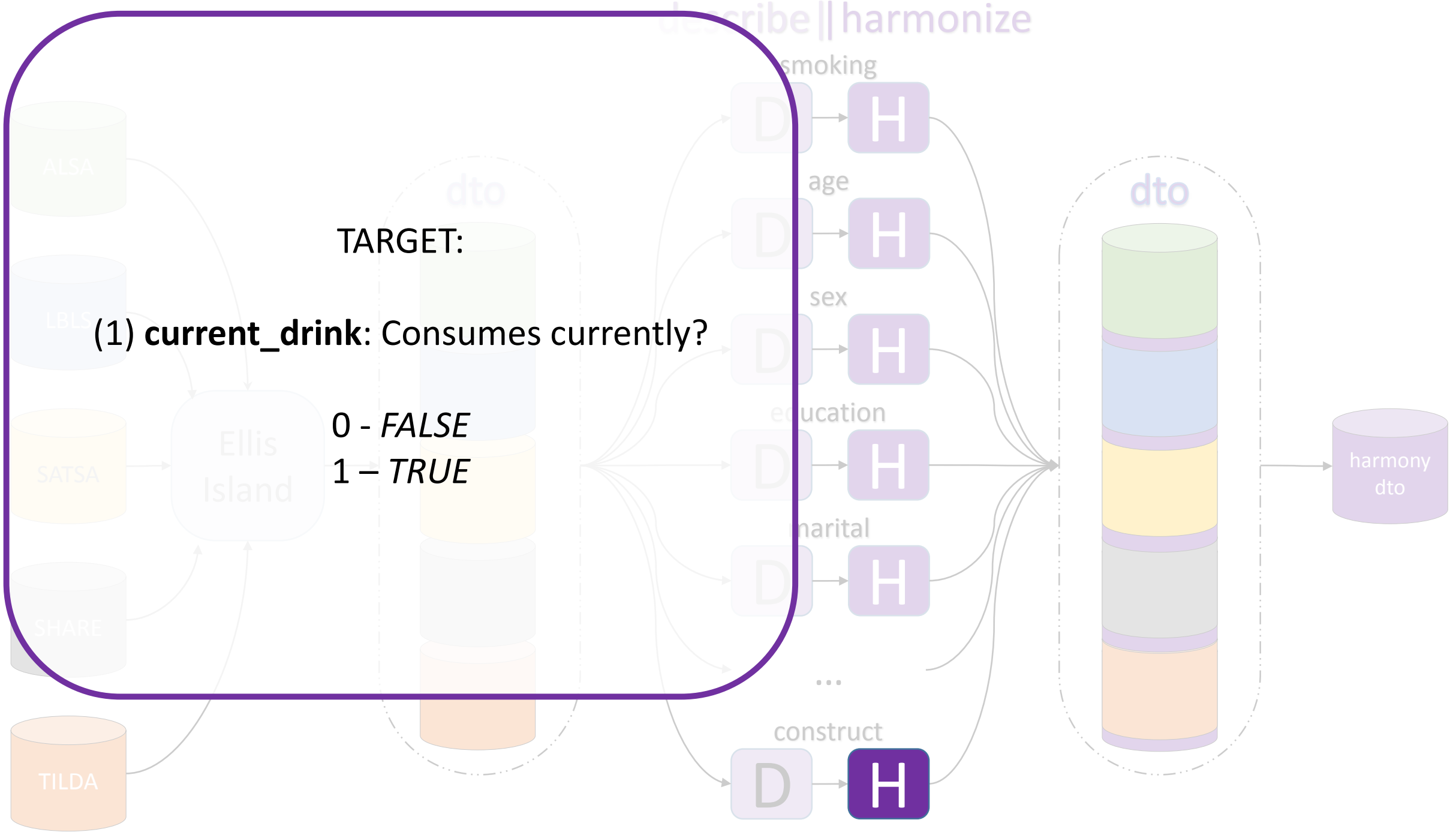




TARGET:

(1) **current_work_2**: Currently employed?

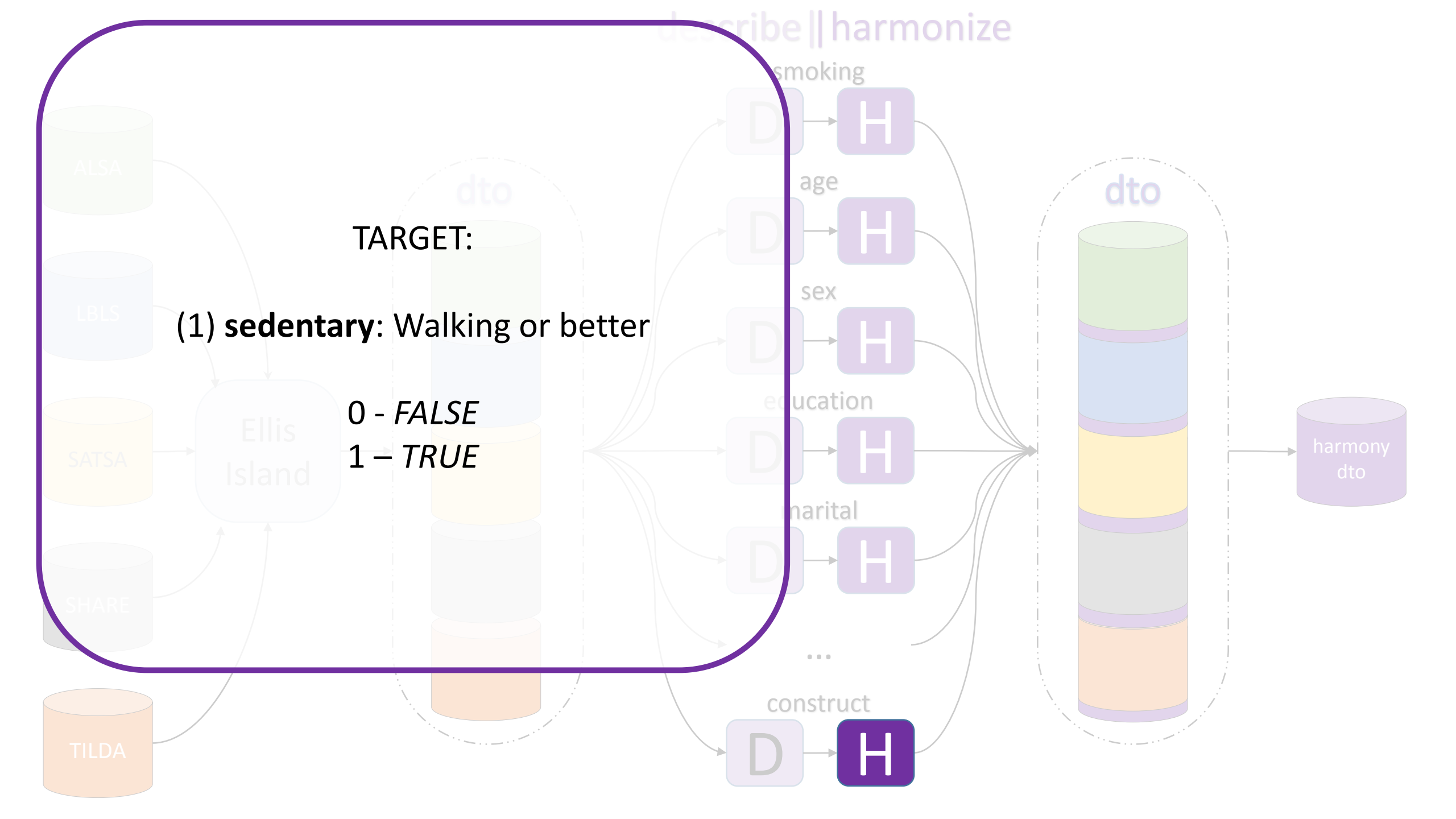
0 - *FALSE*
1 - *TRUE*



TARGET:

(1) **current_drink**: Consumes currently?

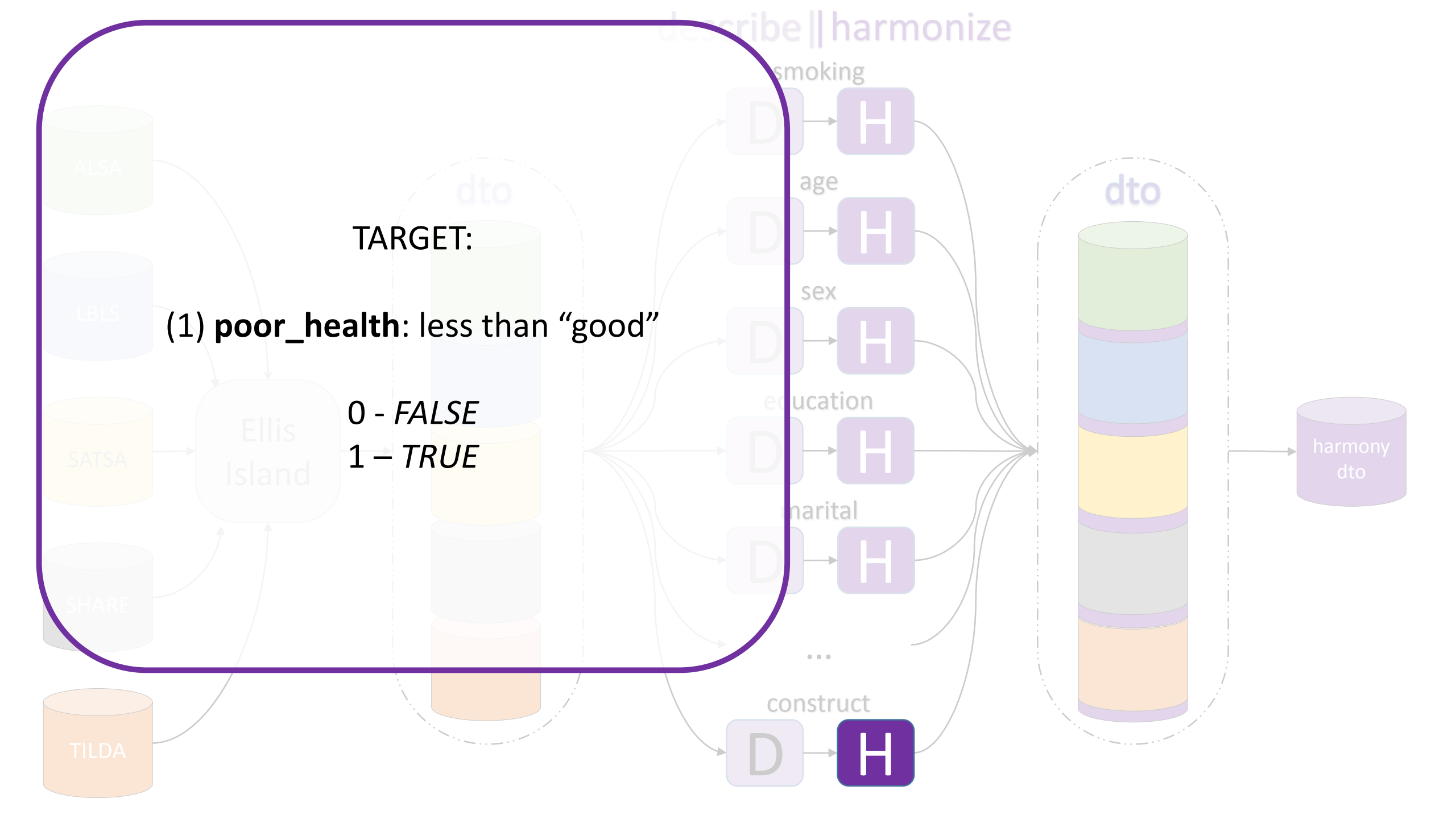
0 - *FALSE*
1 – *TRUE*



TARGET:

(1) **sedentary**: Walking or better

0 - *FALSE*
1 – *TRUE*



TARGET:

(1) **poor_health**: less than “good”

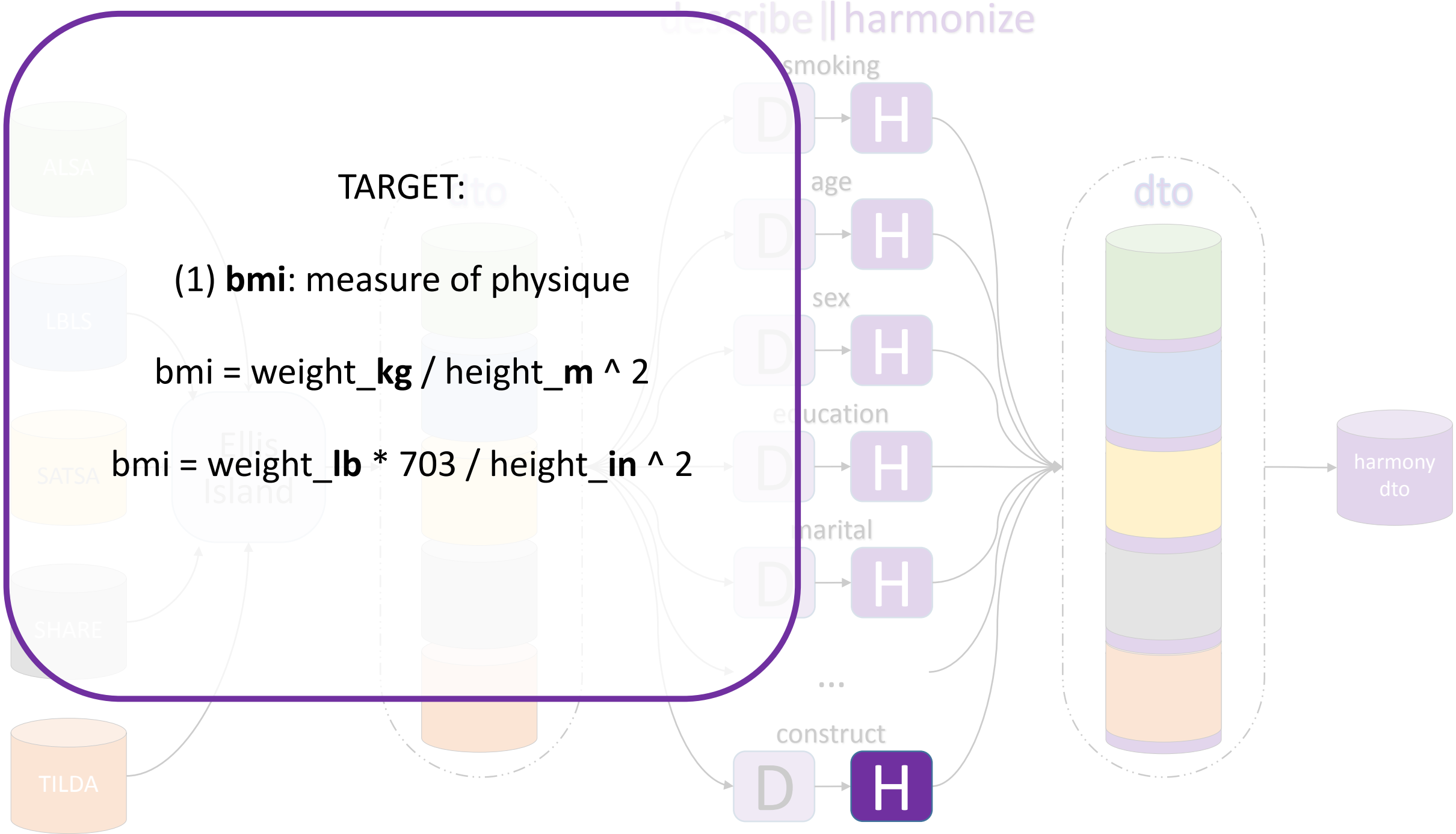
0 - *FALSE*
1 – *TRUE*

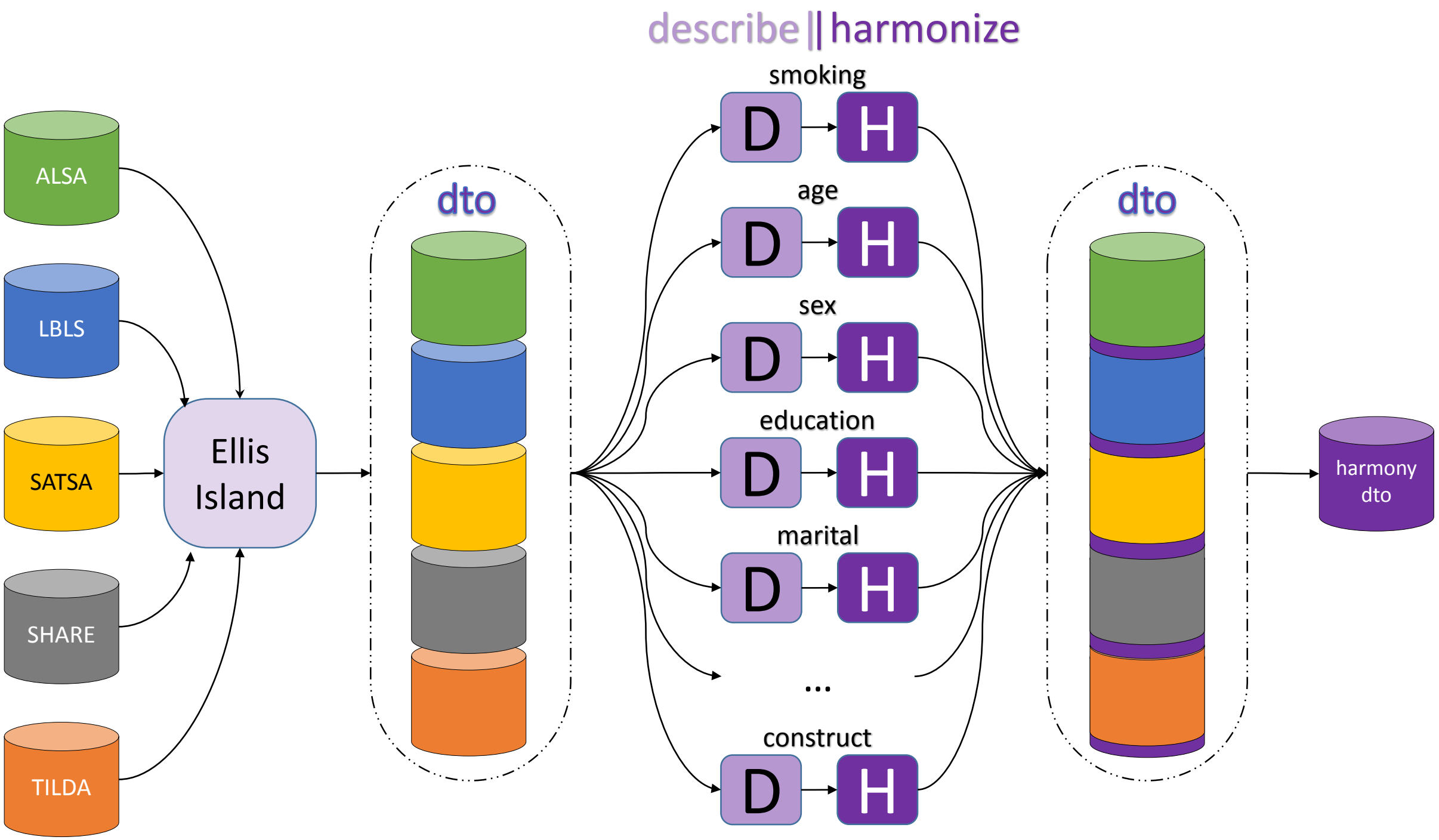
TARGET:

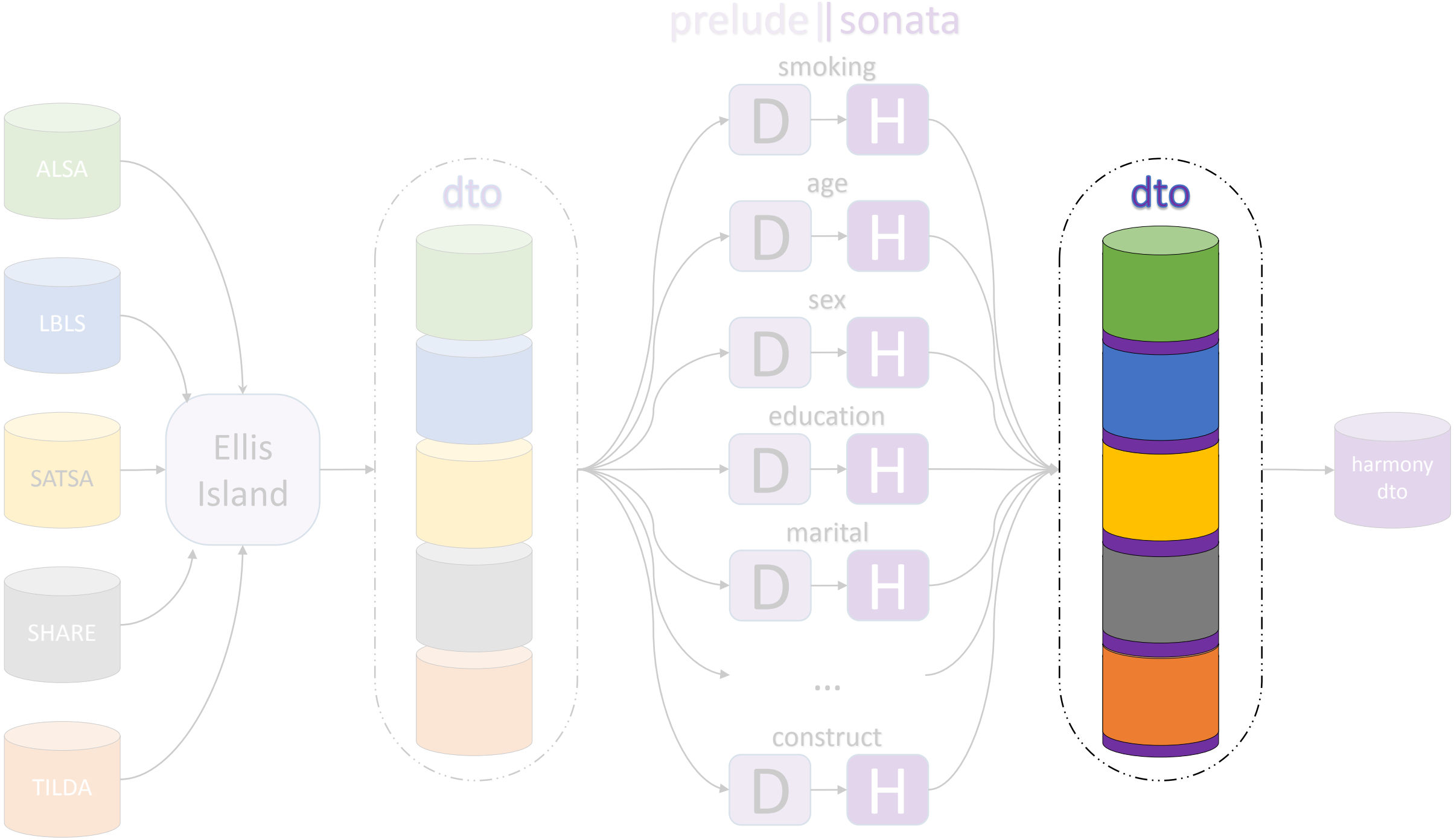
(1) **bmi**: measure of physique

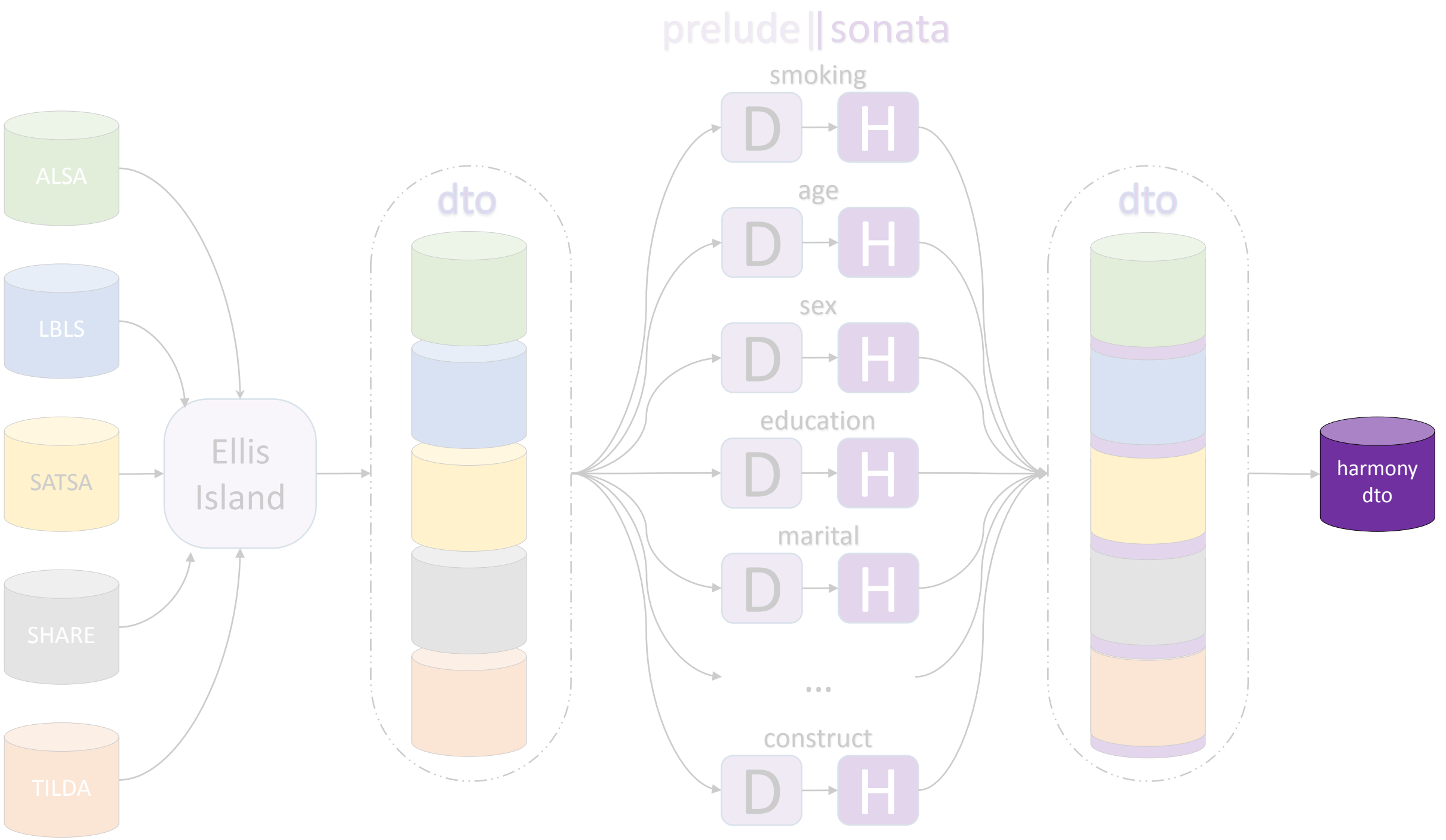
bmi = weight_**kg** / height_**m** ^ 2

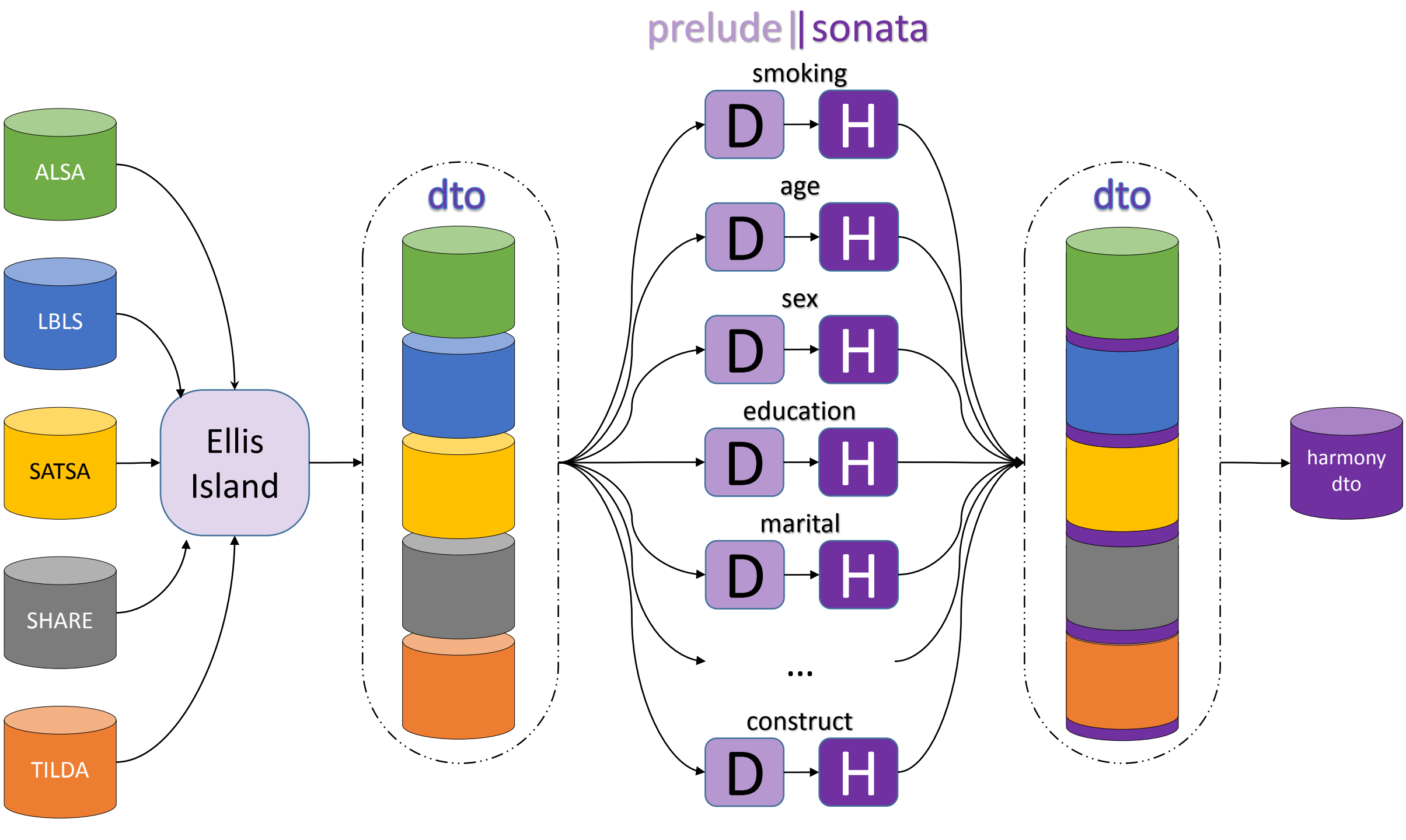
bmi = weight_**lb** * 703 / height_**in** ^ 2

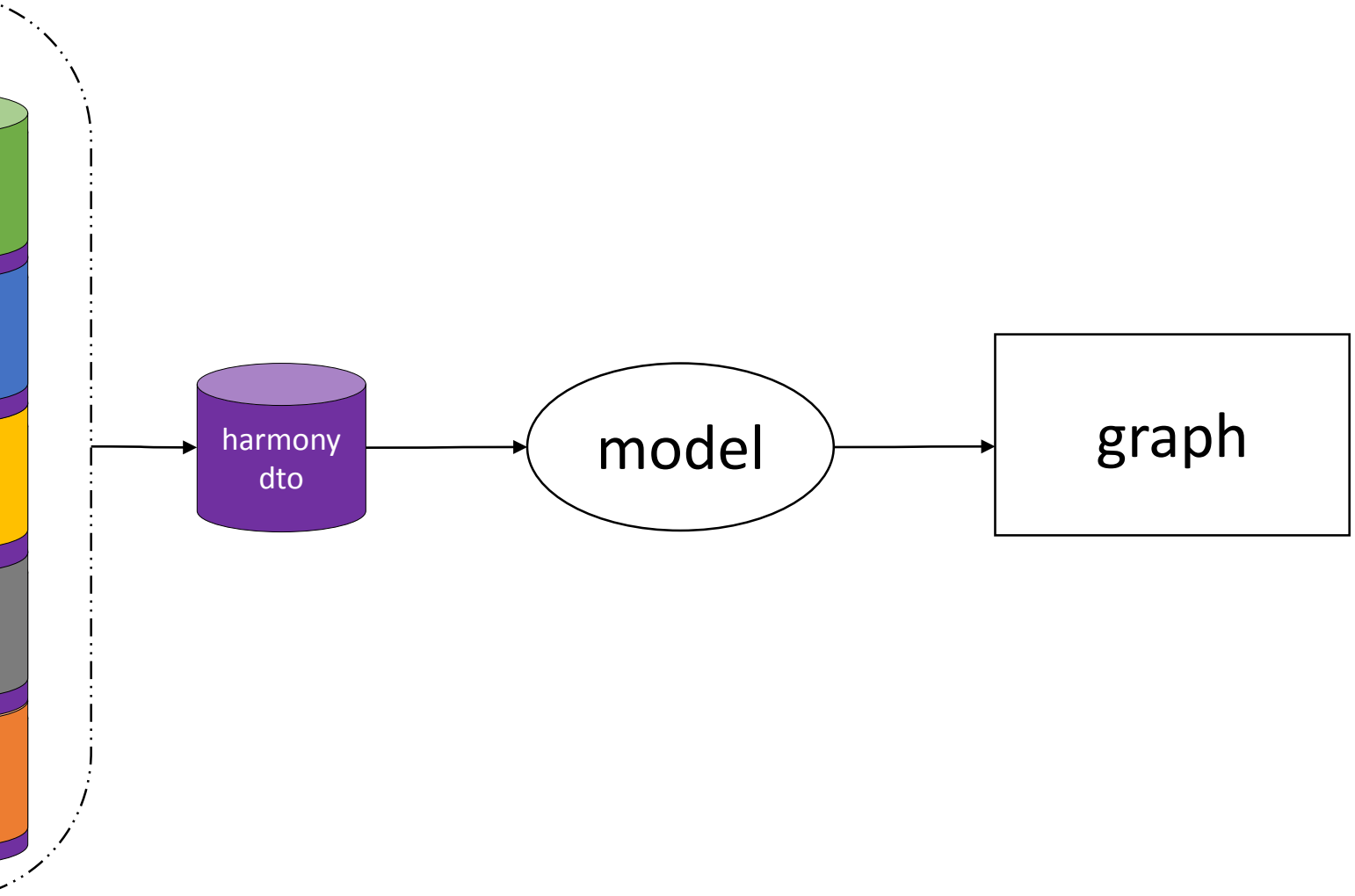


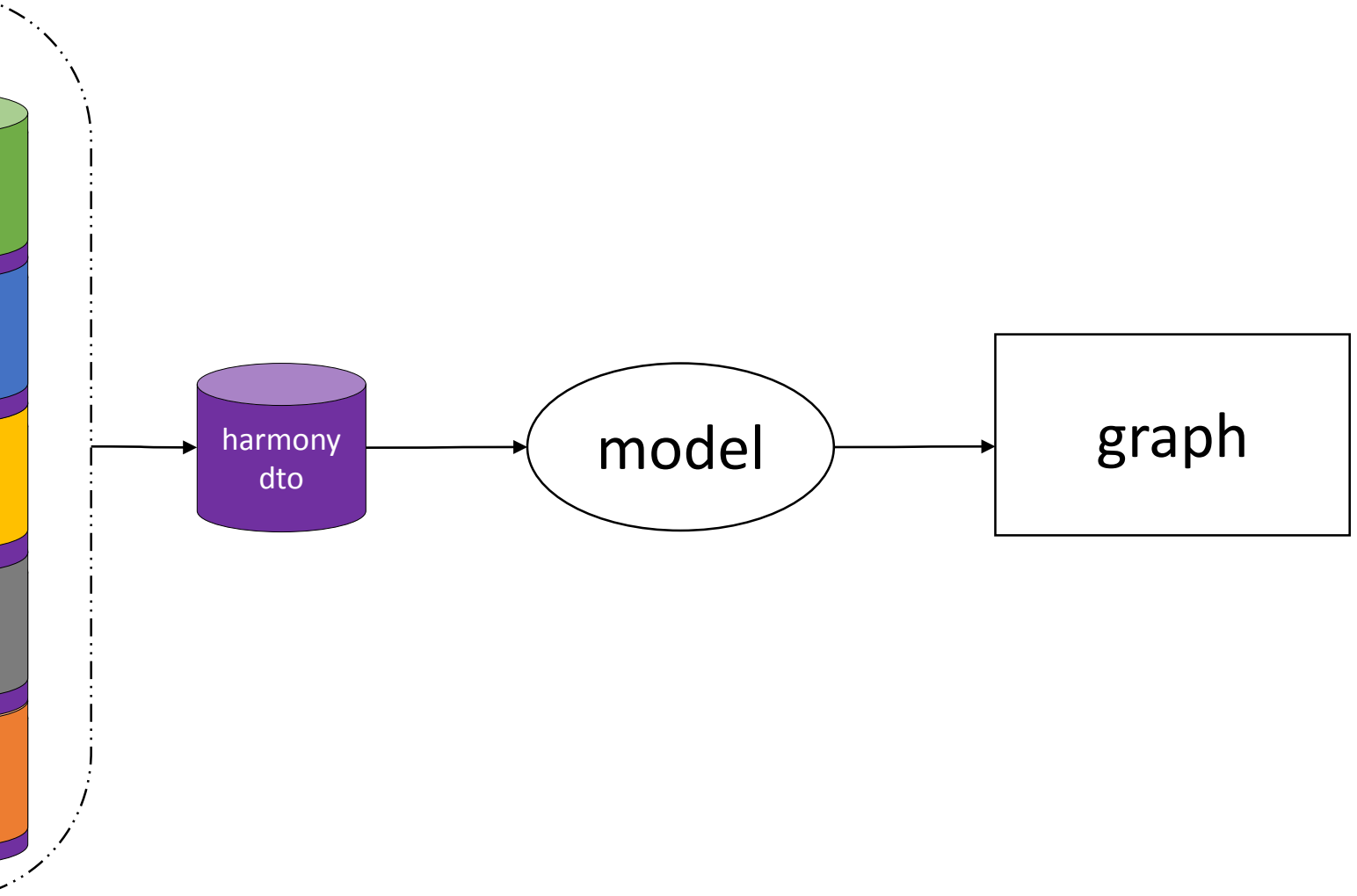

describe || harmonize
ALSA
LBLS
SATSA
SHARE
TILDA
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
dto
harmony dto


prelude||sonata
ALSA
LBLS
SATSA
SHARE
TILDA
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
dto
harmony dto


prelude||sonata
ALSA
LBLS
SATSA
SHARE
TILDA
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
dto
harmony dto

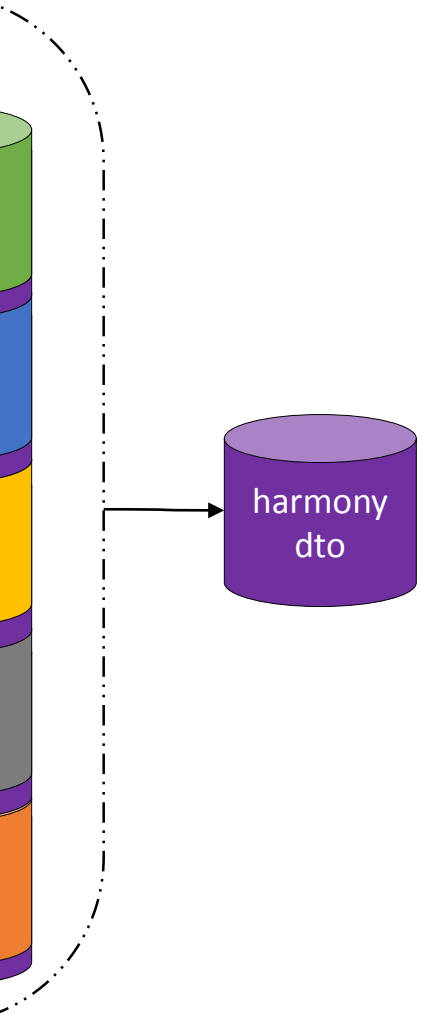

prelude || sonata
ALSA
LBLS
SATSA
SHARE
TILDA
Ellis Island
dto
smoking
age
sex
education
marital
...
construct
D
H
dto
harmony dto


harmony
dto
model
graph

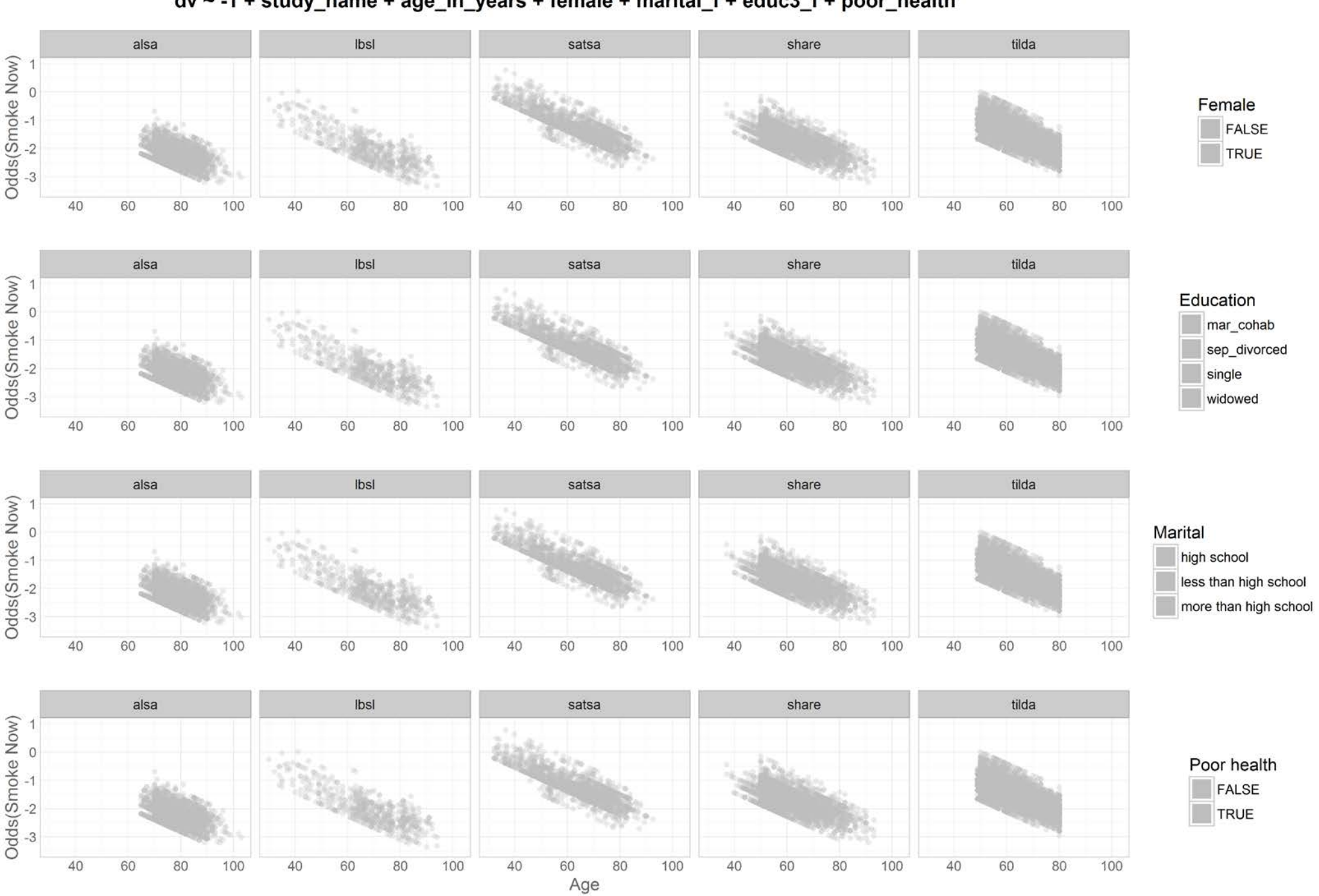


smoke_now

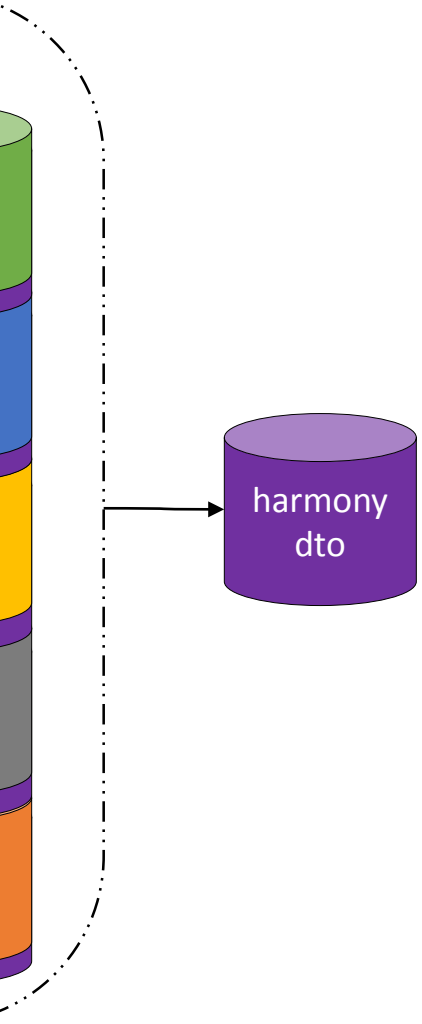
Odds | Prob



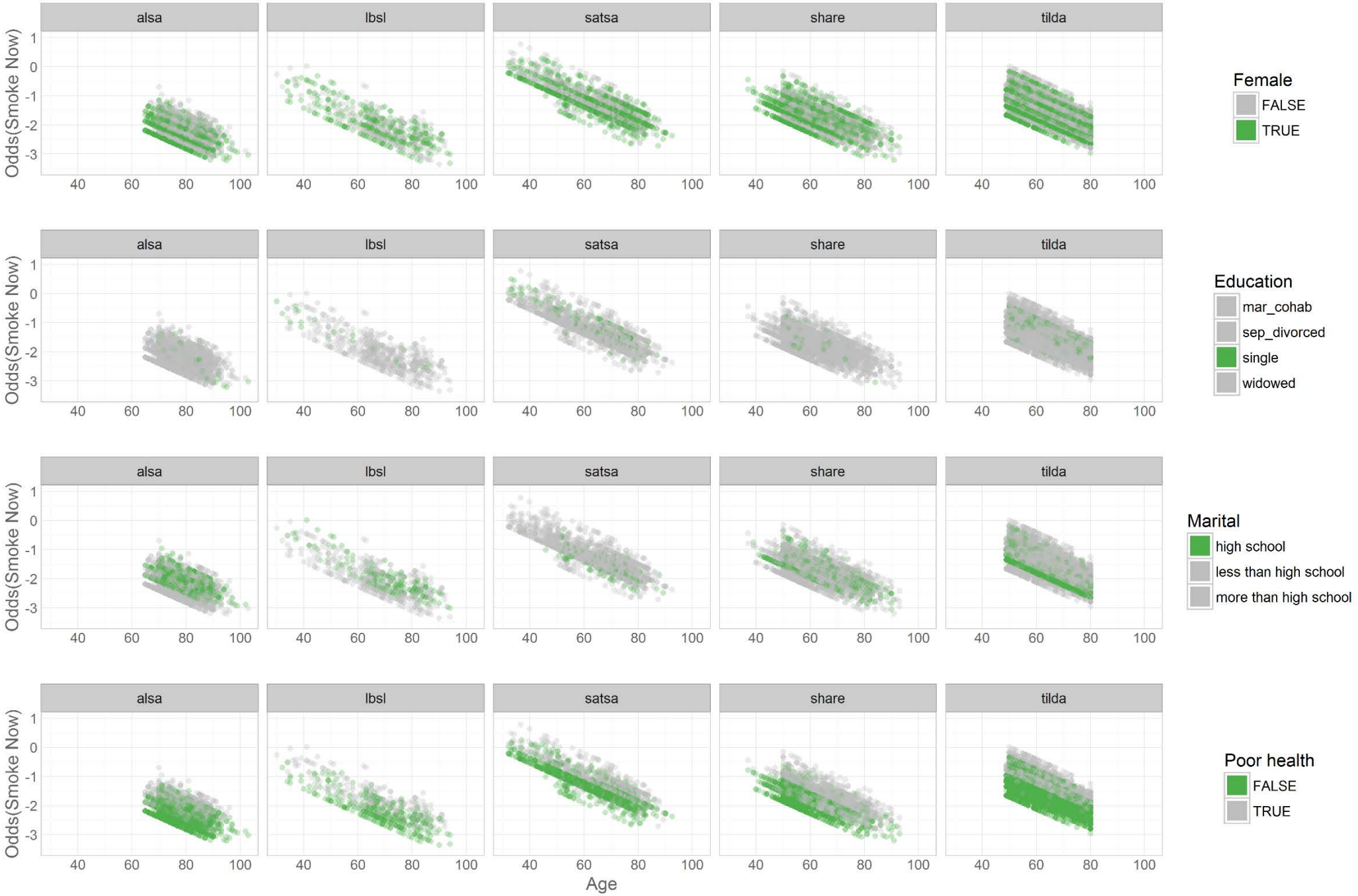
harmony dto



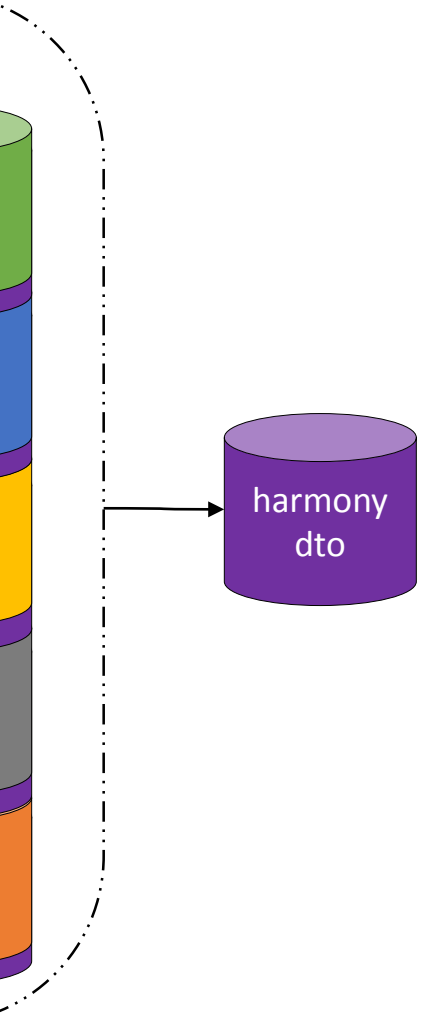
Odds | Prob



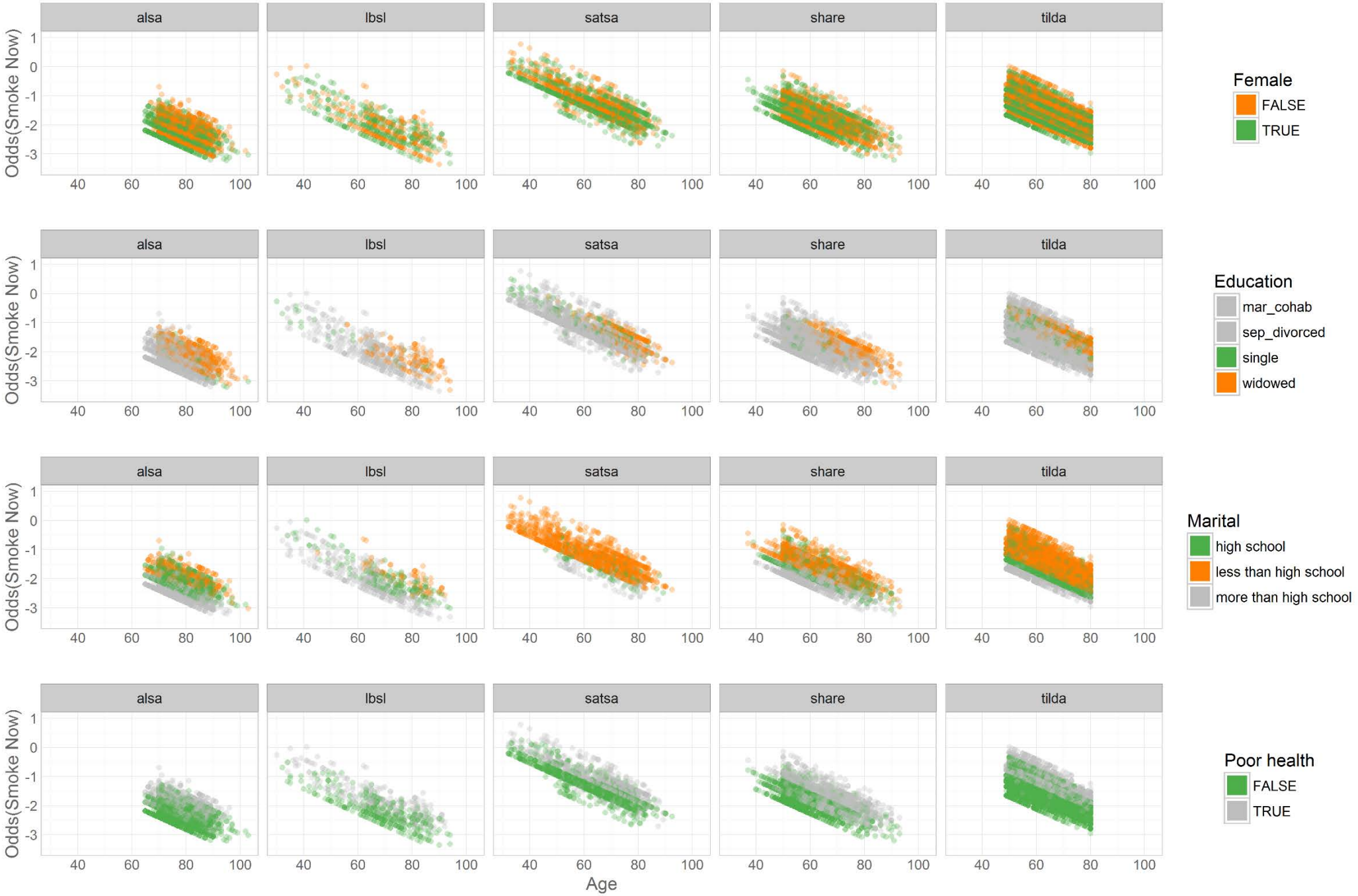
harmony dto



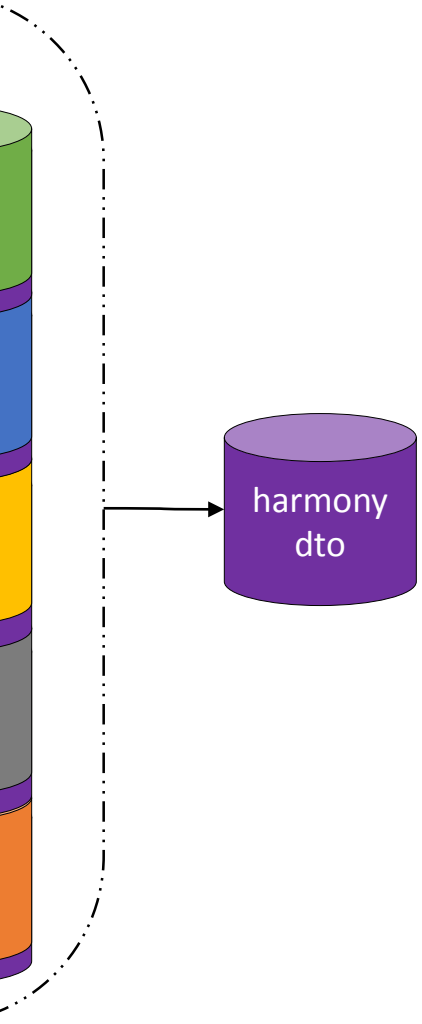
Odds | Prob



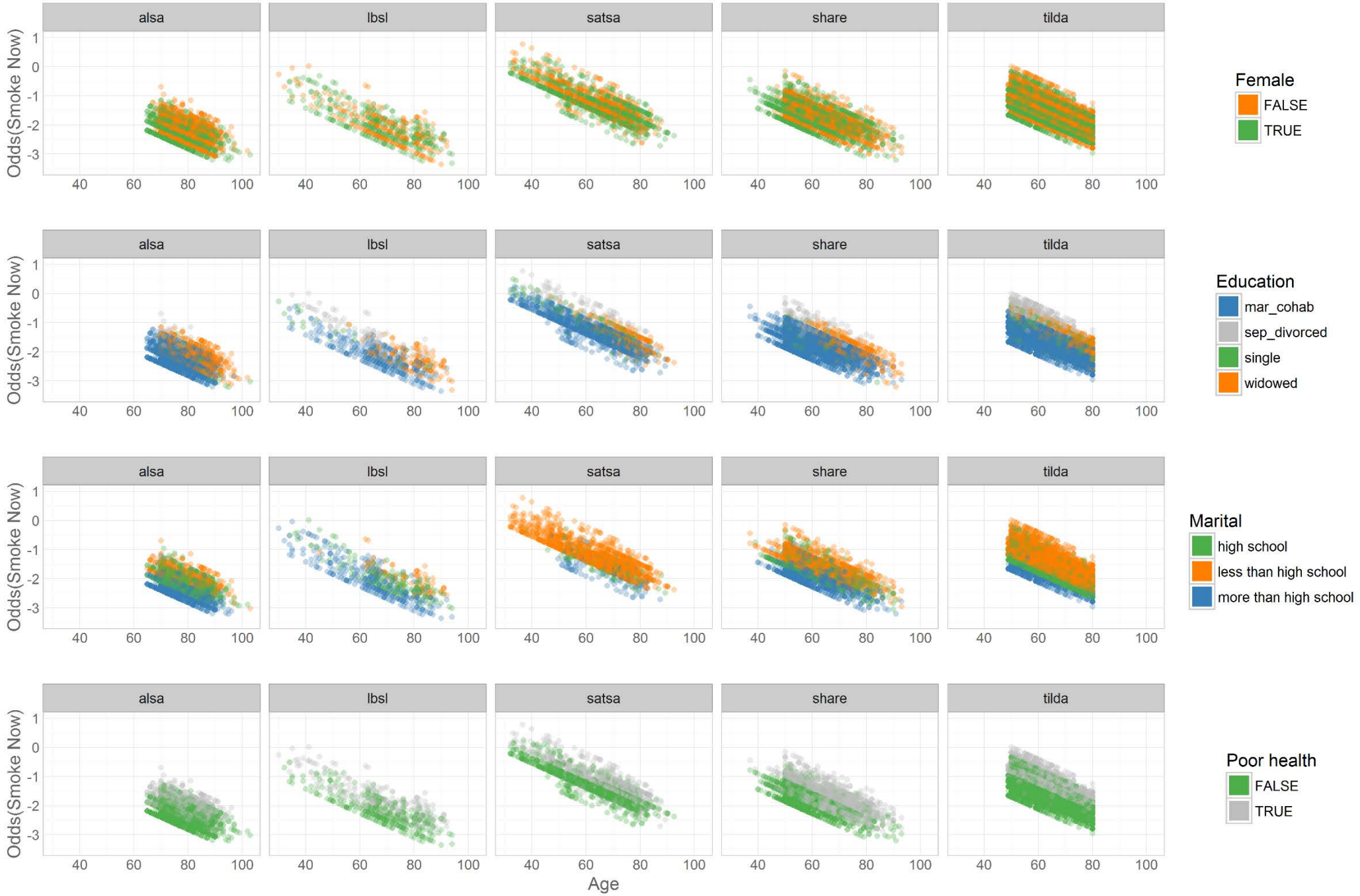
harmony dto



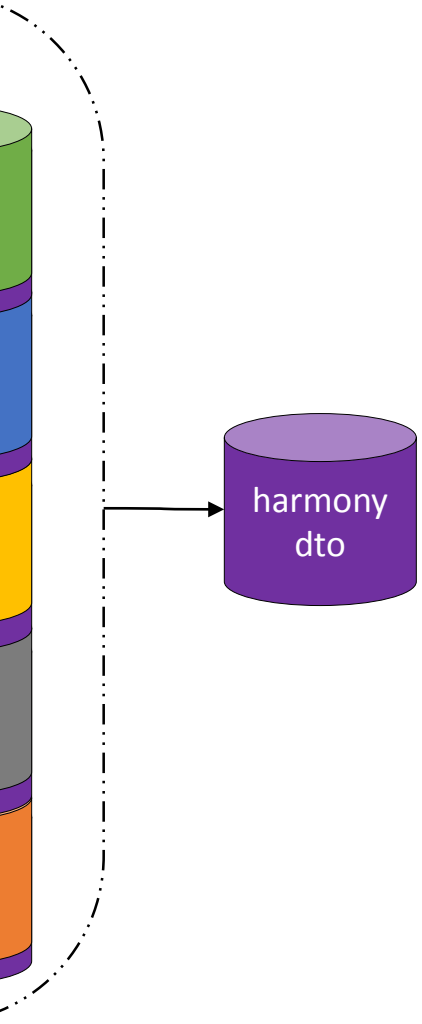
Odds | Prob



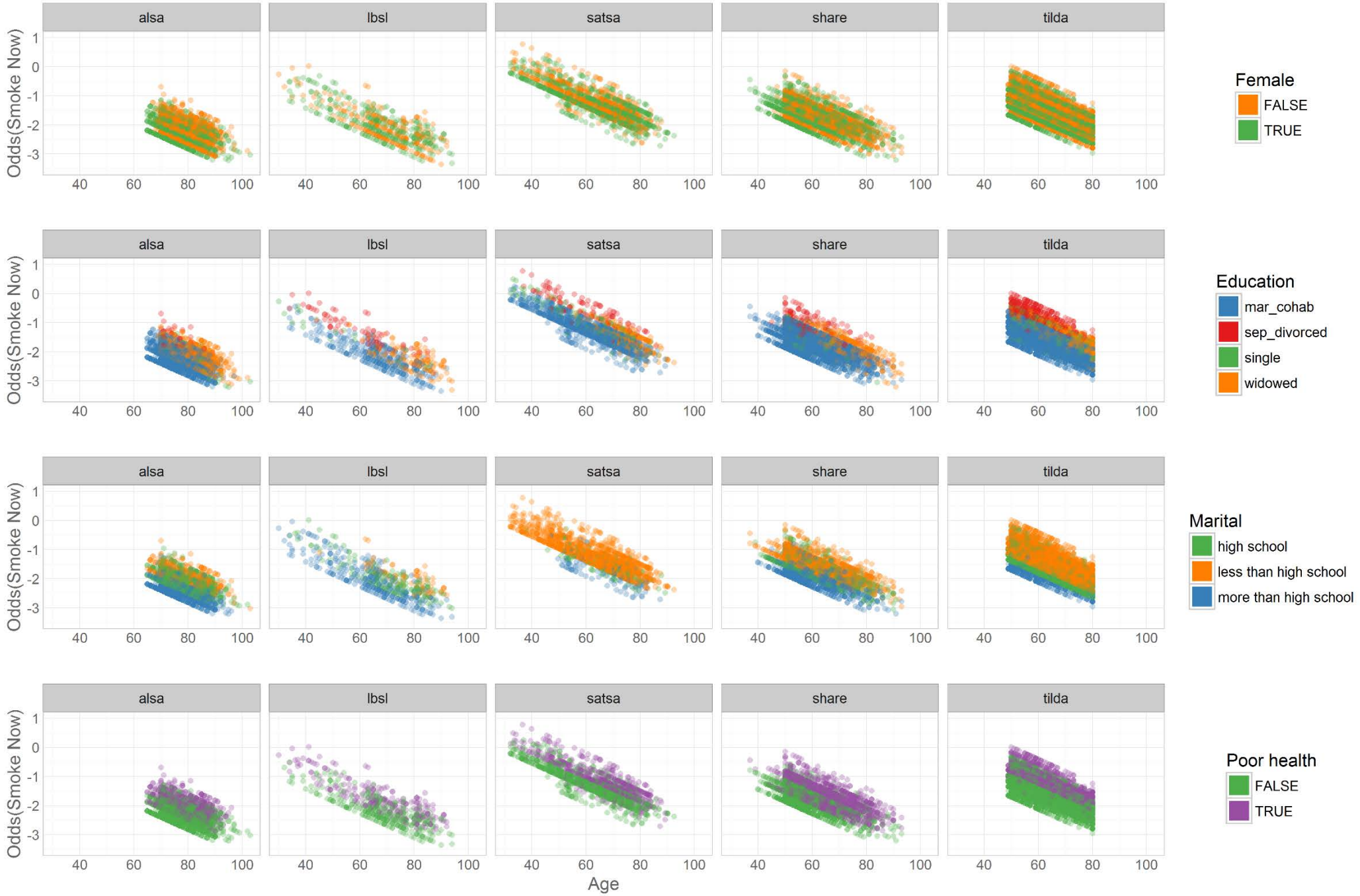
harmony dto



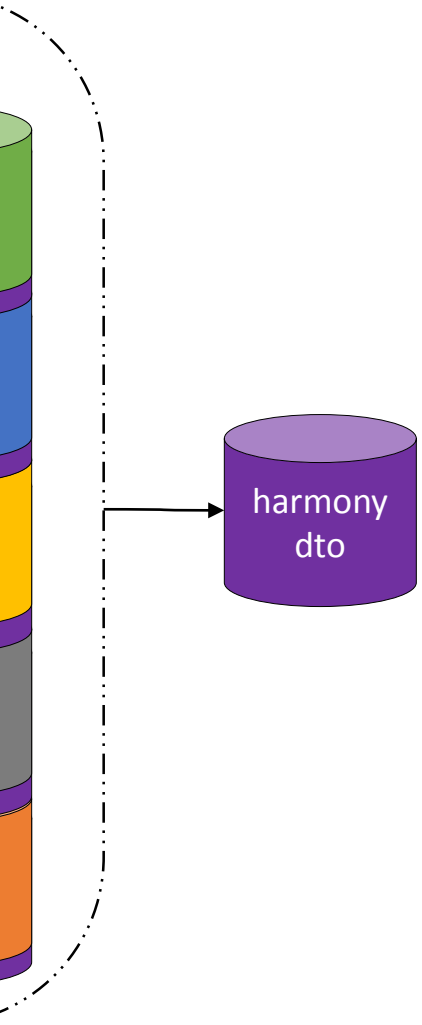
Odds | Prob



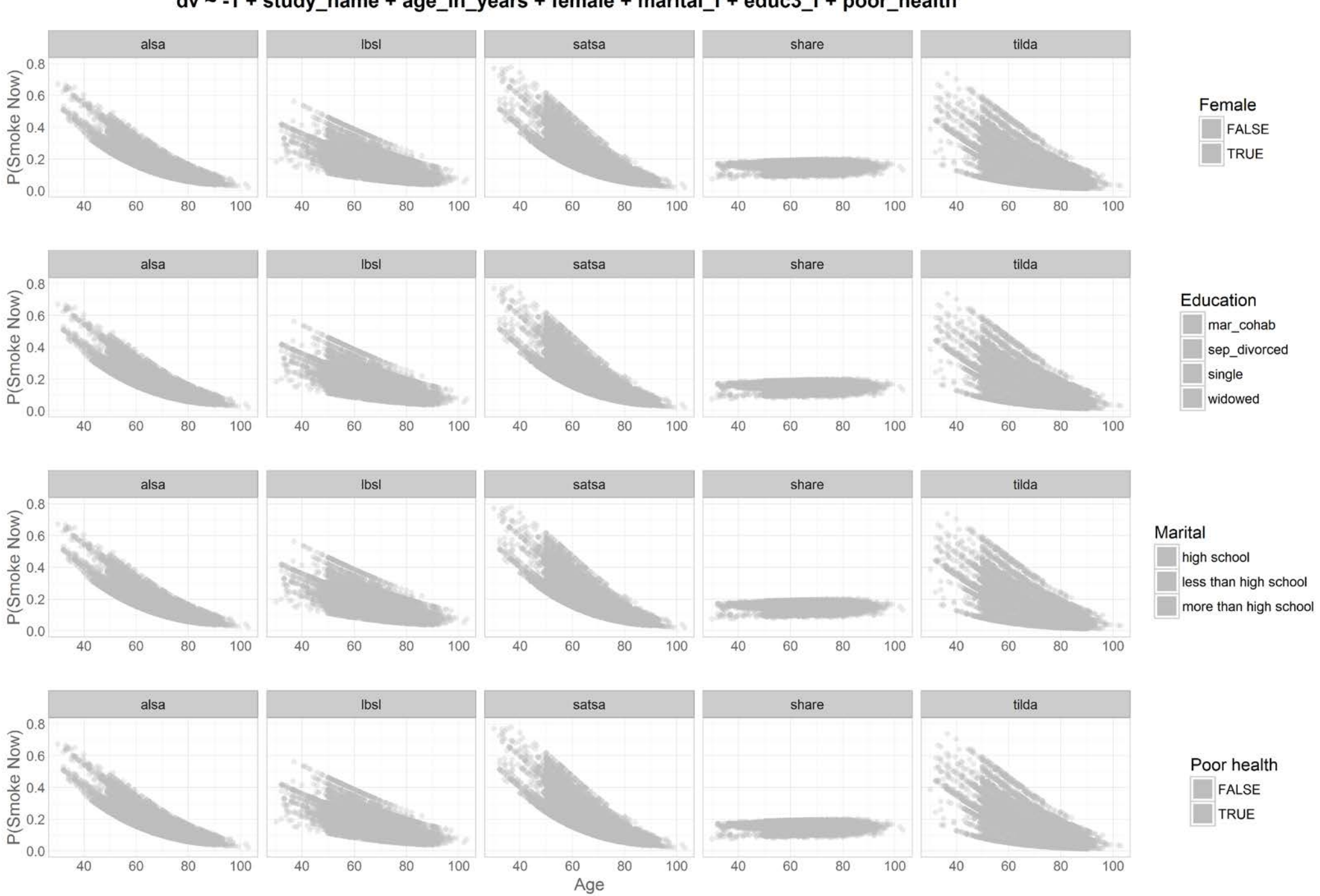
harmony dto



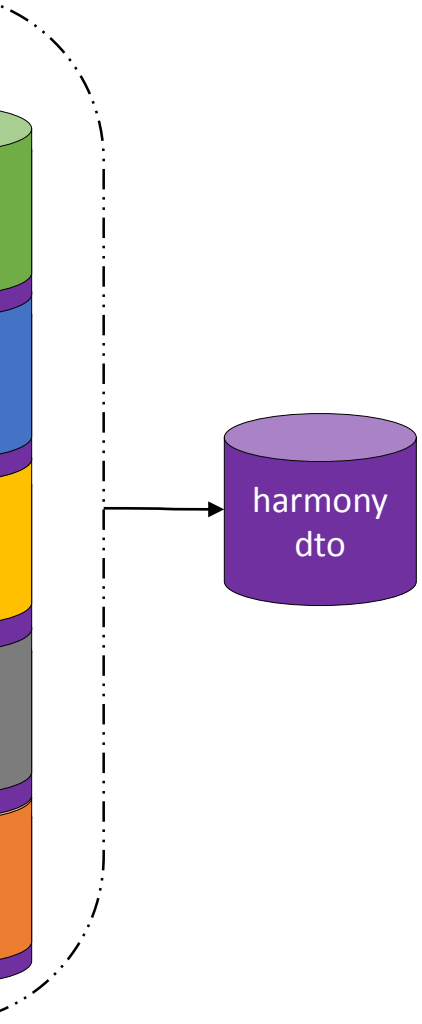
Odds | Prob



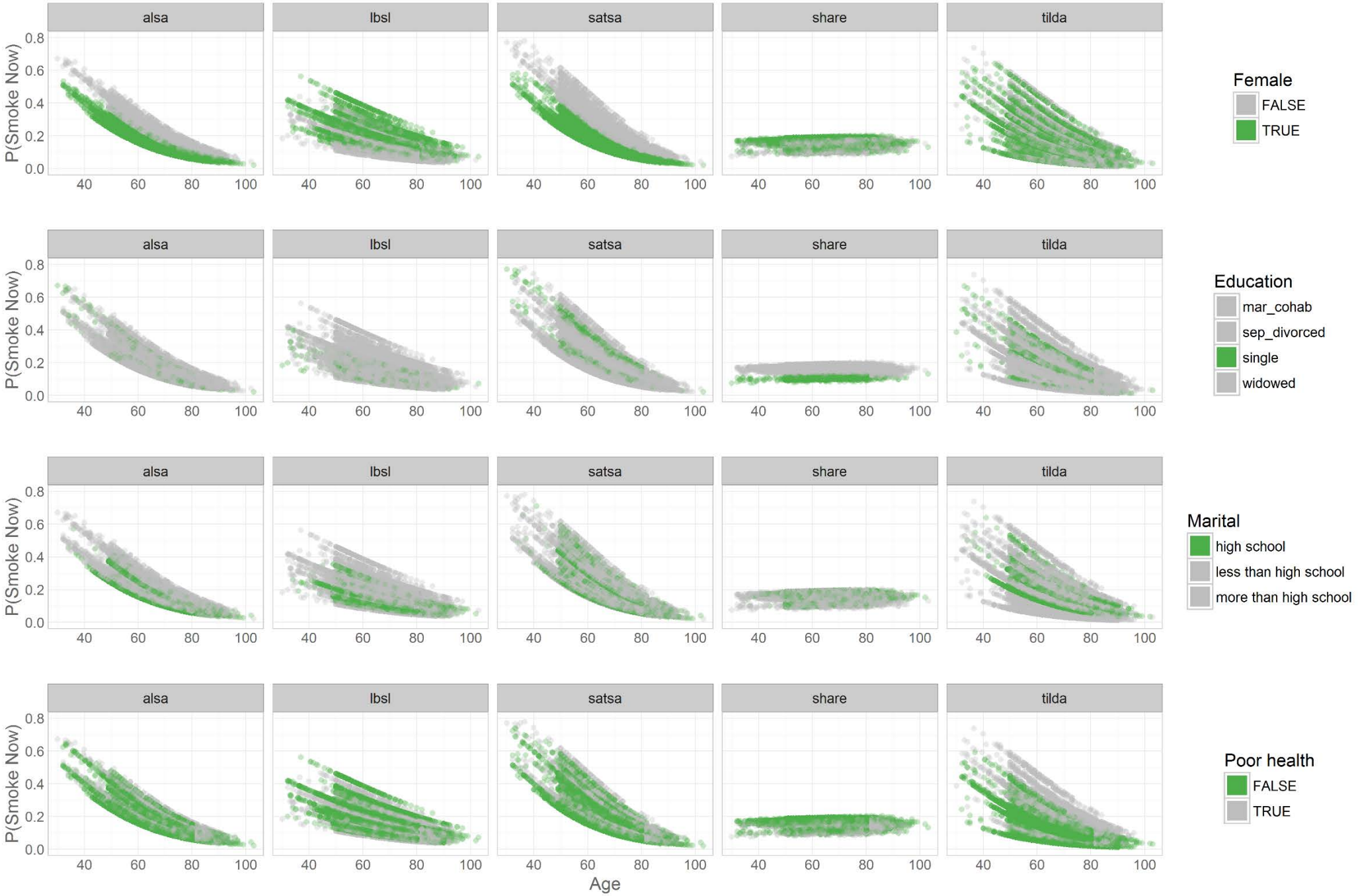
harmony dto



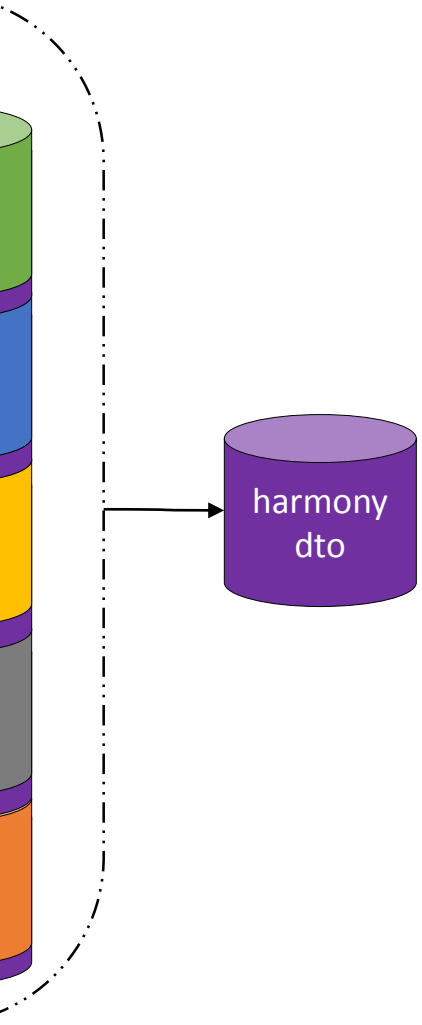
Odds | **Prob**



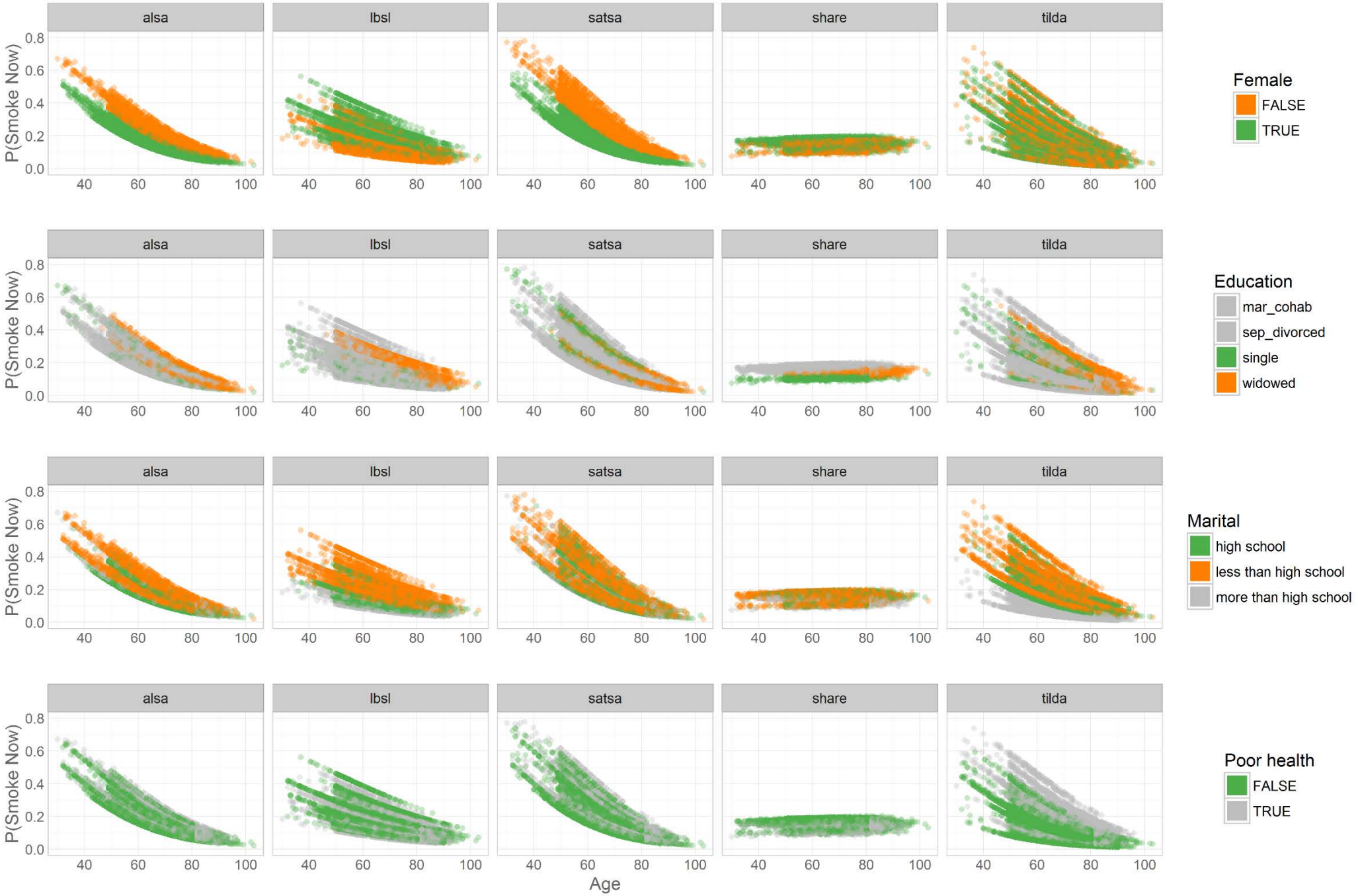
harmony dto



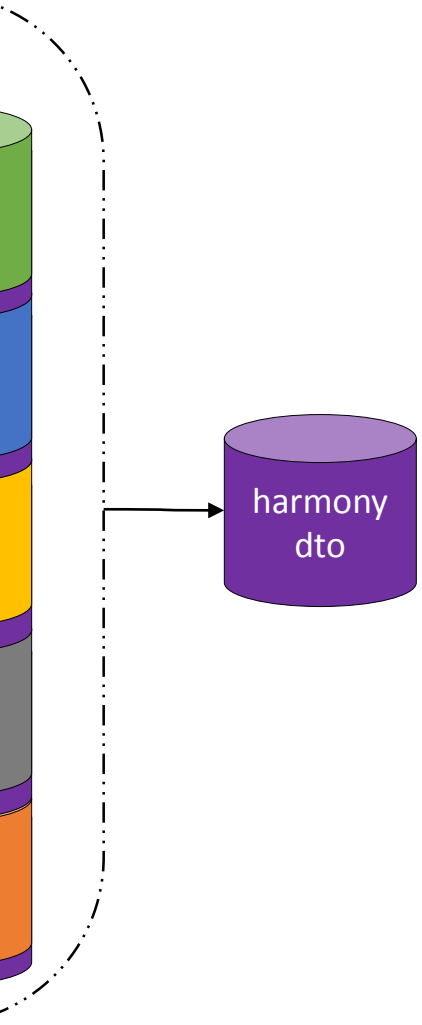
Odds | Prob



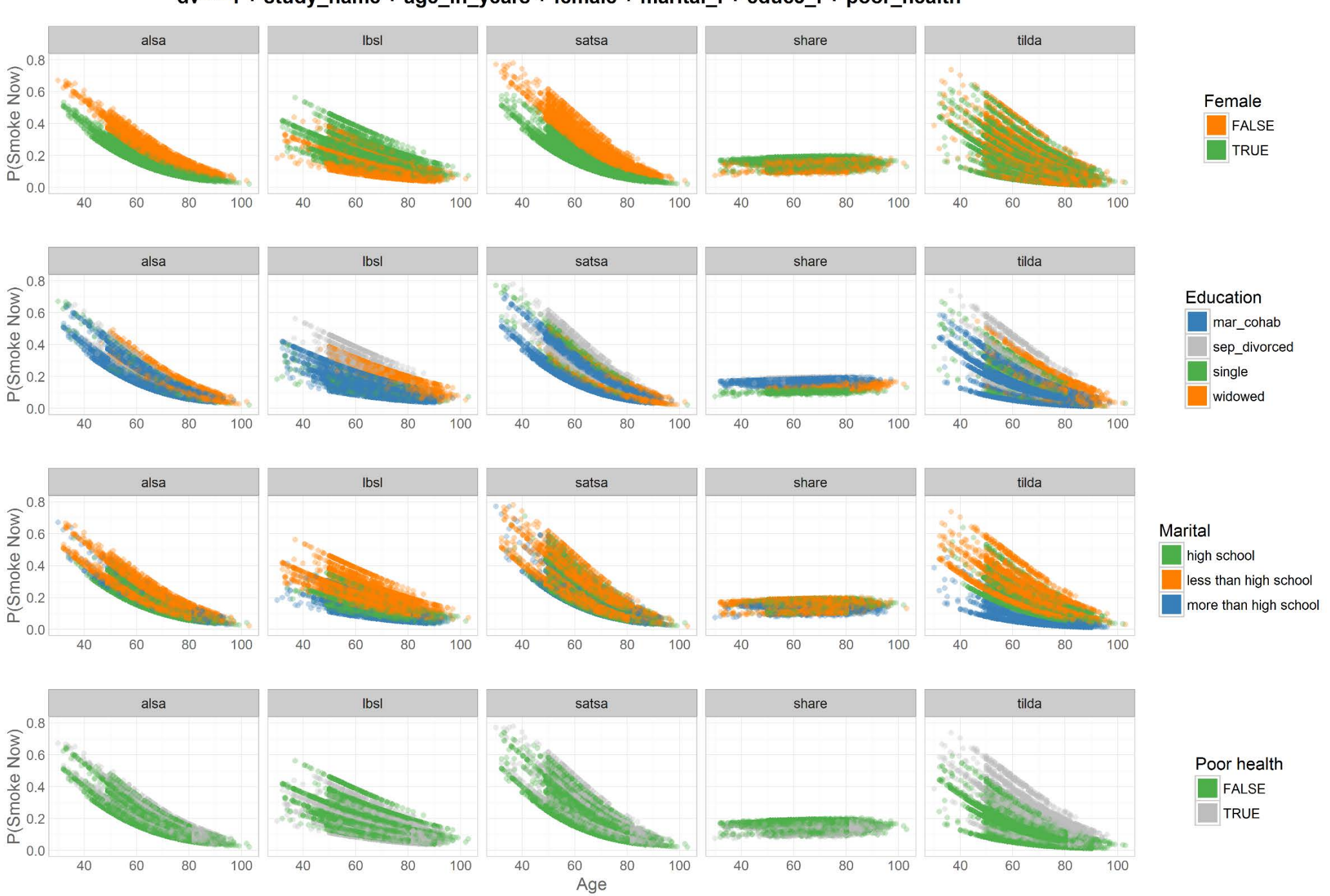
harmony dto



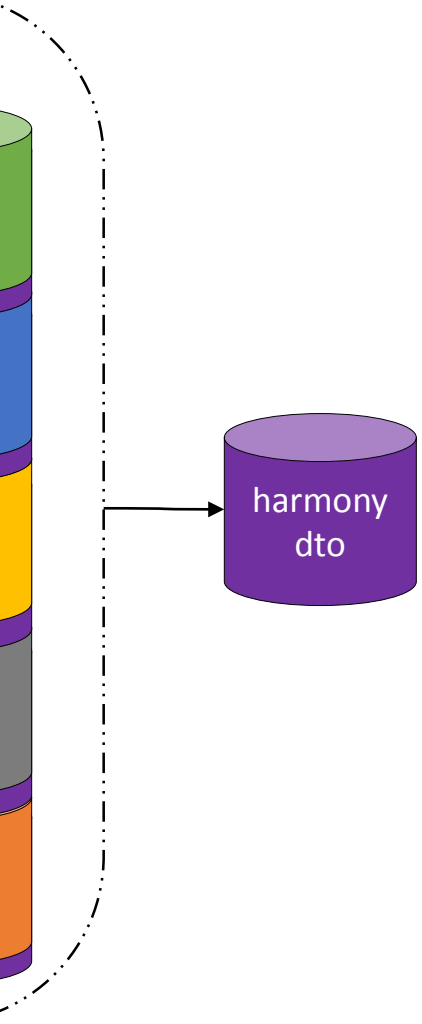
Odds | **Prob**



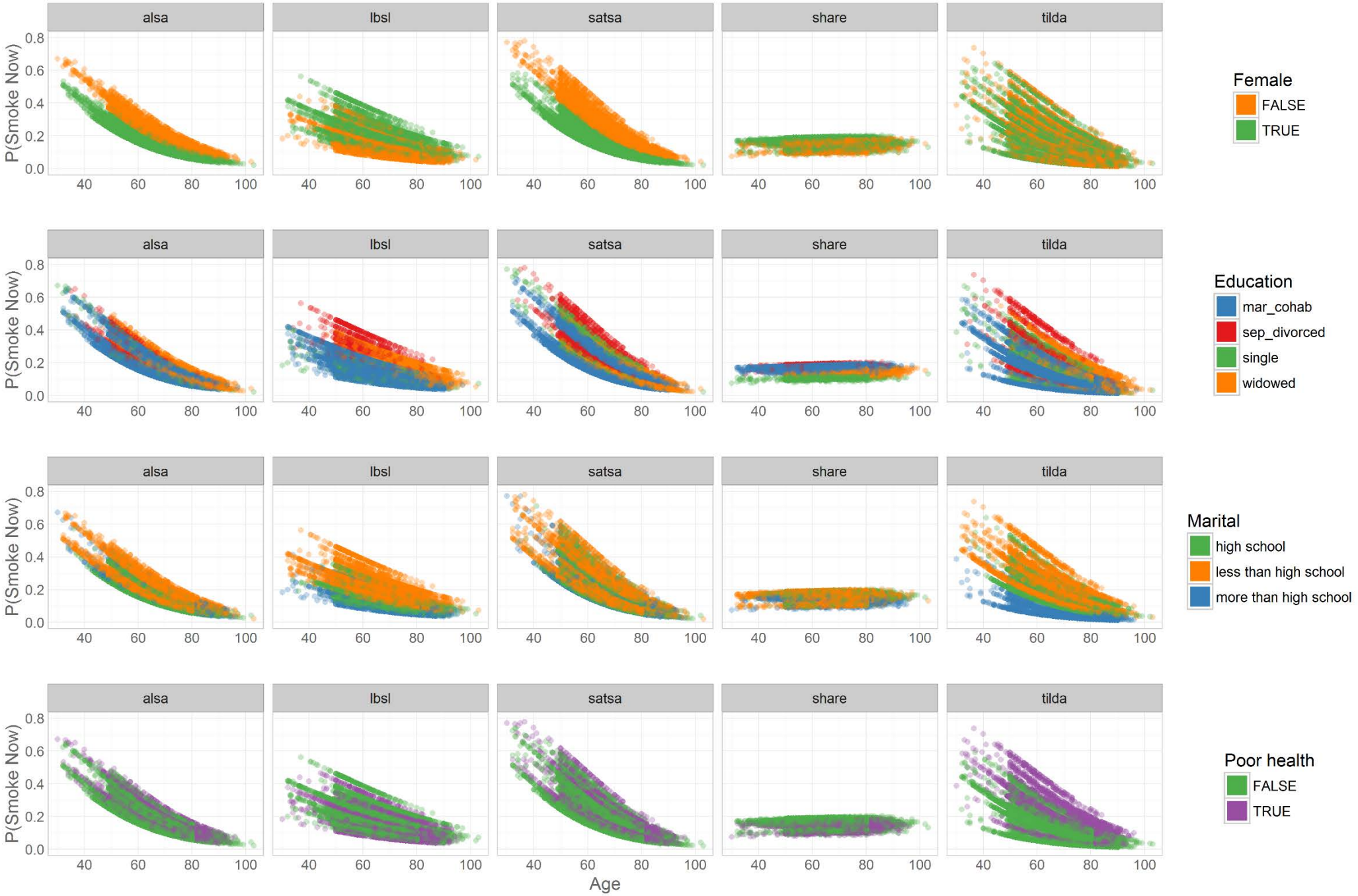
harmony
dto


dv ~ -1 + study_name + age_in_years + female + marital_f + educ3_f + poor_health
alsa
lbsl
satsa
share
tilda
P(Smoke Now)
0.0
0.2
0.4
0.6
0.8
40
60
80
100
Age
Female
FALSE
TRUE
Education
mar_cohab
sep_divorced
single
widowed
Marital
high school
less than high school
more than high school
Poor health
FALSE
TRUE

Odds | **Prob**

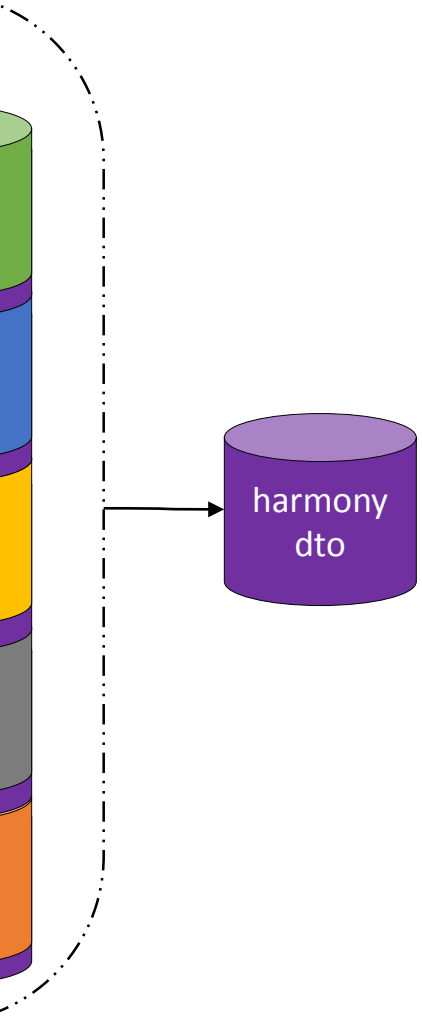


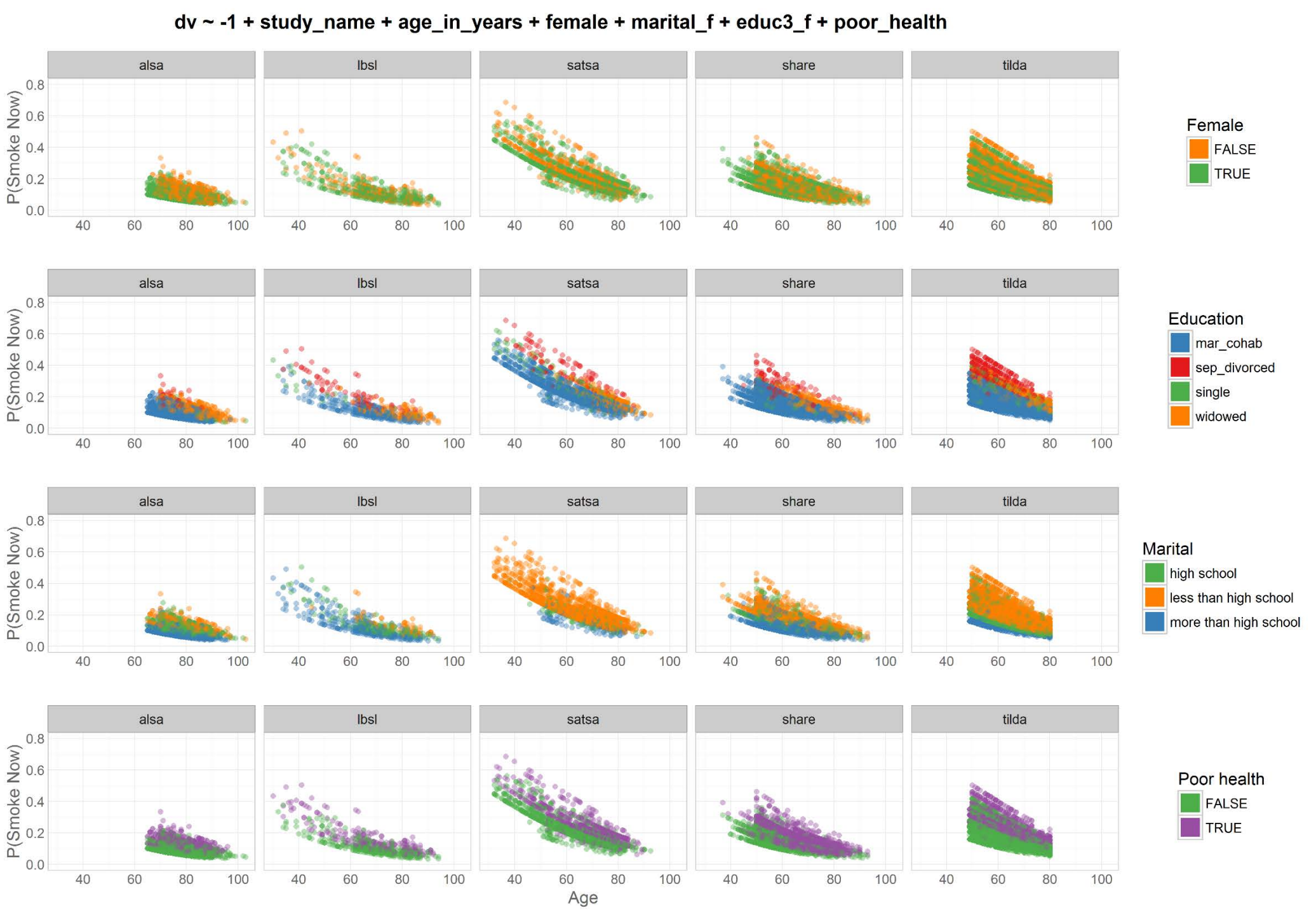
harmony dto



Odds | Prob

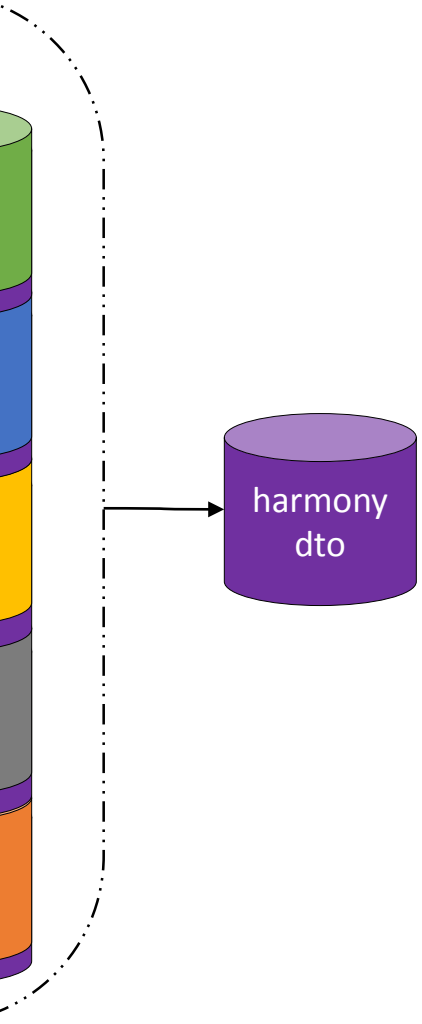
Study = factor



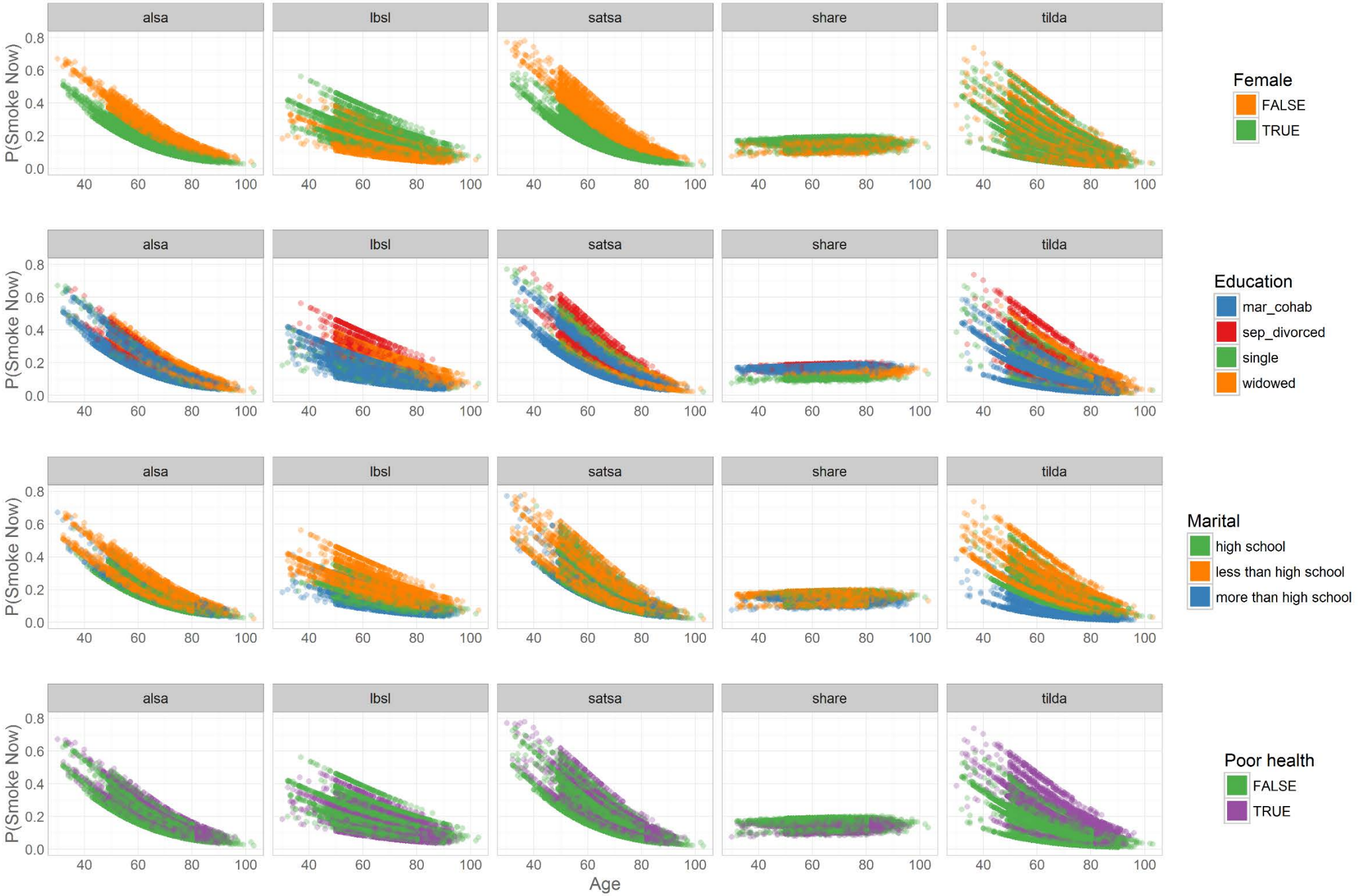


Odds | Prob

Study = cluster

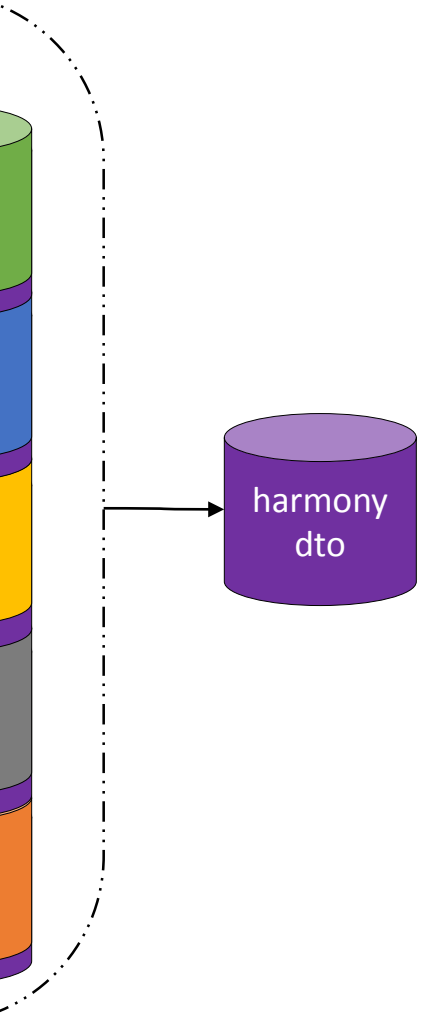


harmony dto



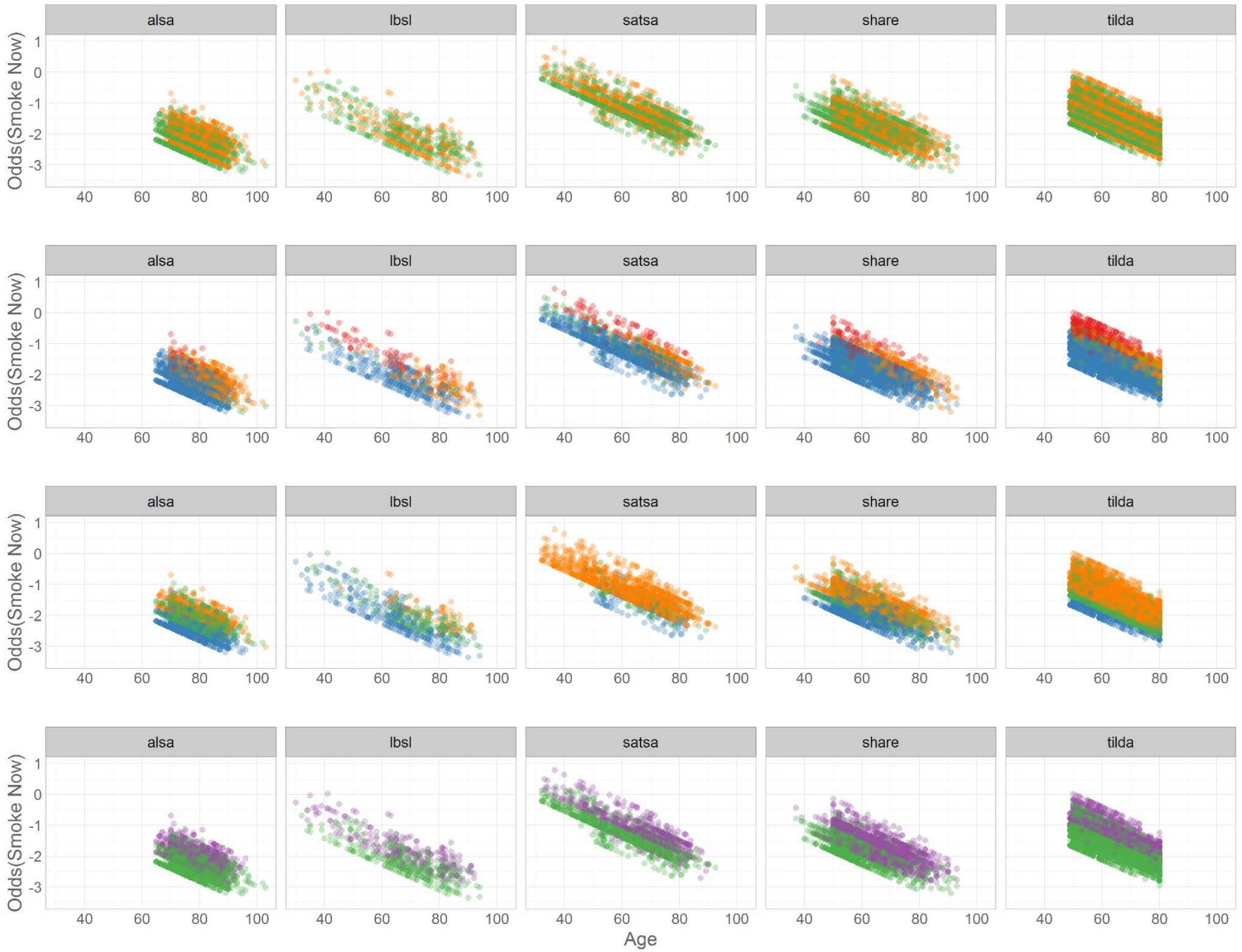
Odds | Prob

Study = factor



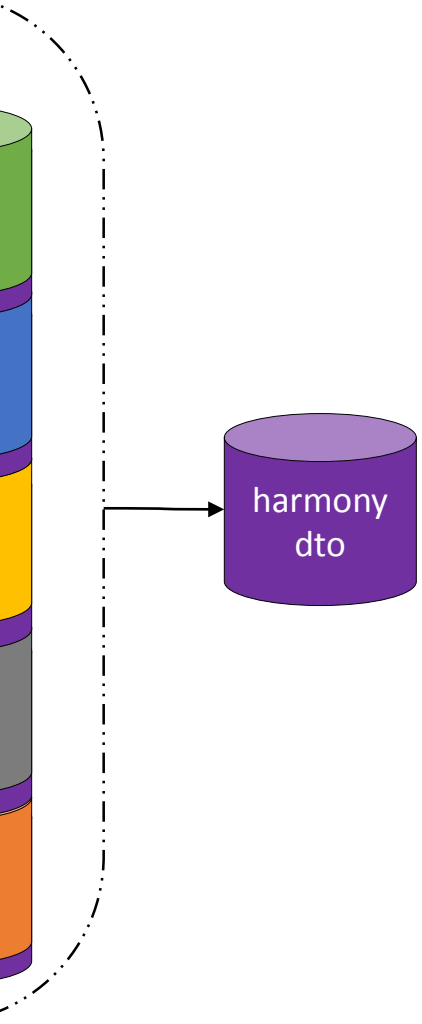
harmony dto

dv ~ -1 + study_name + age_in_years + female + marital_f + educ3_f + poor_health



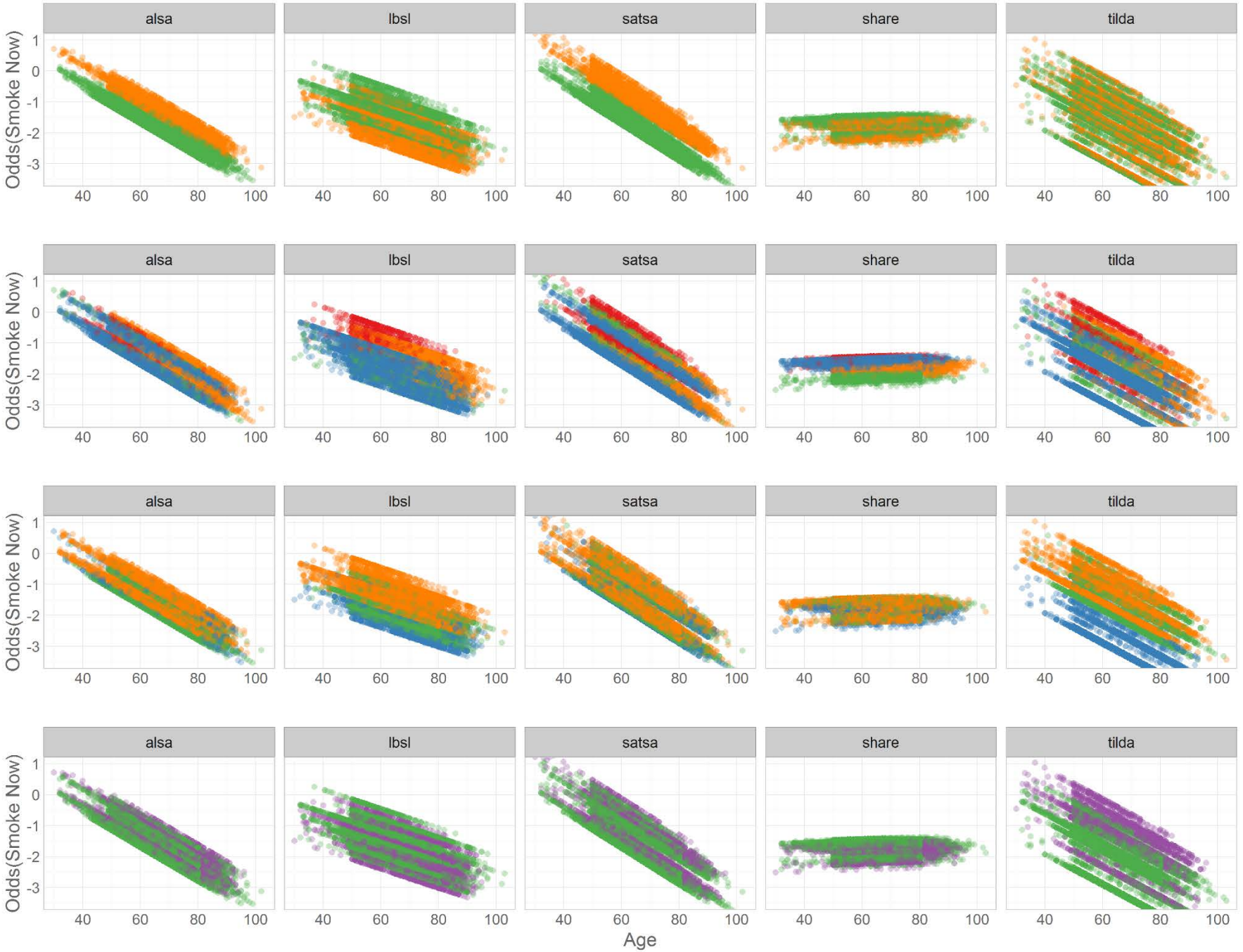
Odds | Prob

Study = cluster



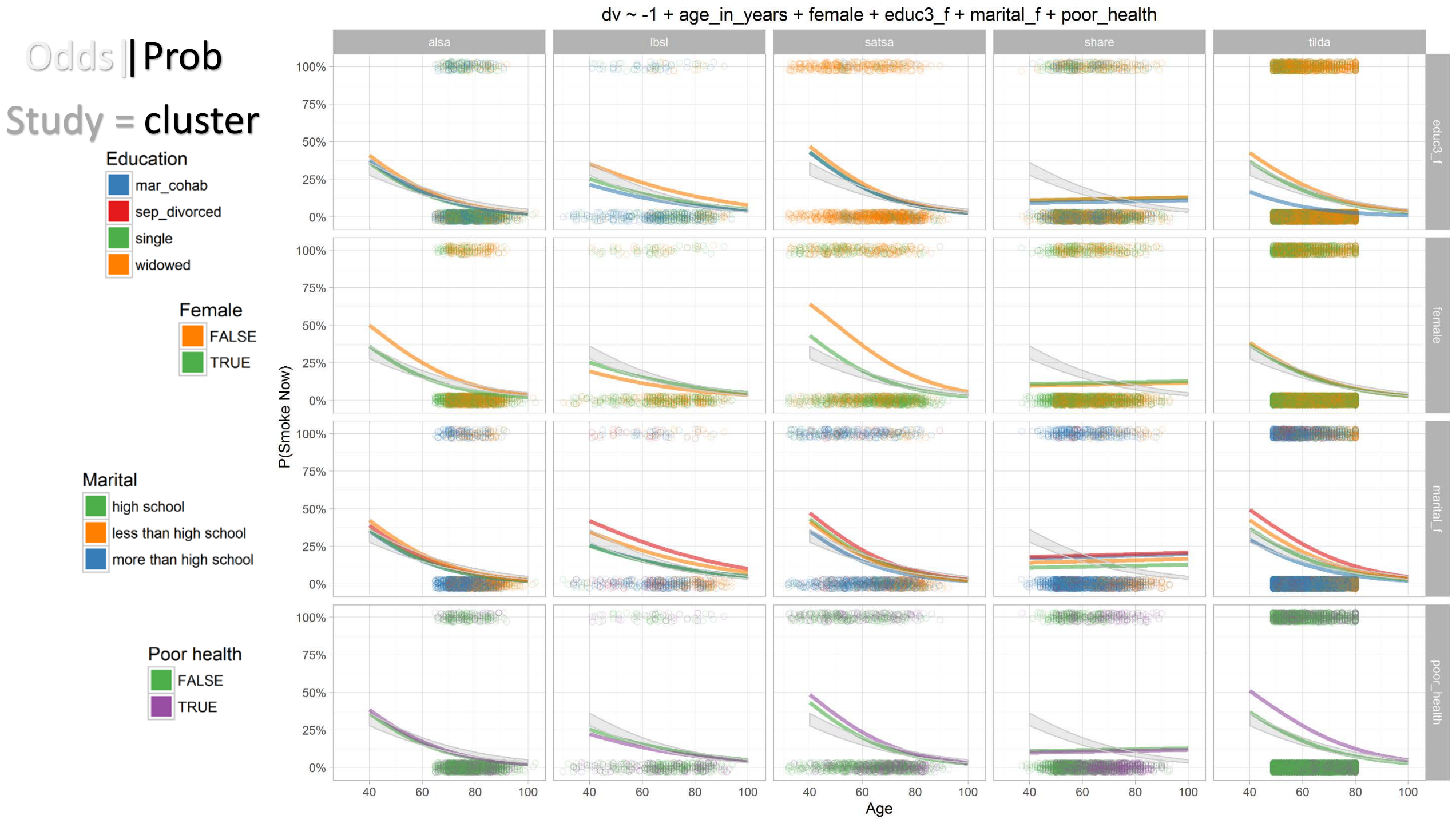
harmony dto

dv ~ -1 + study_name + age_in_years + female + marital_f + educ3_f + poor_health



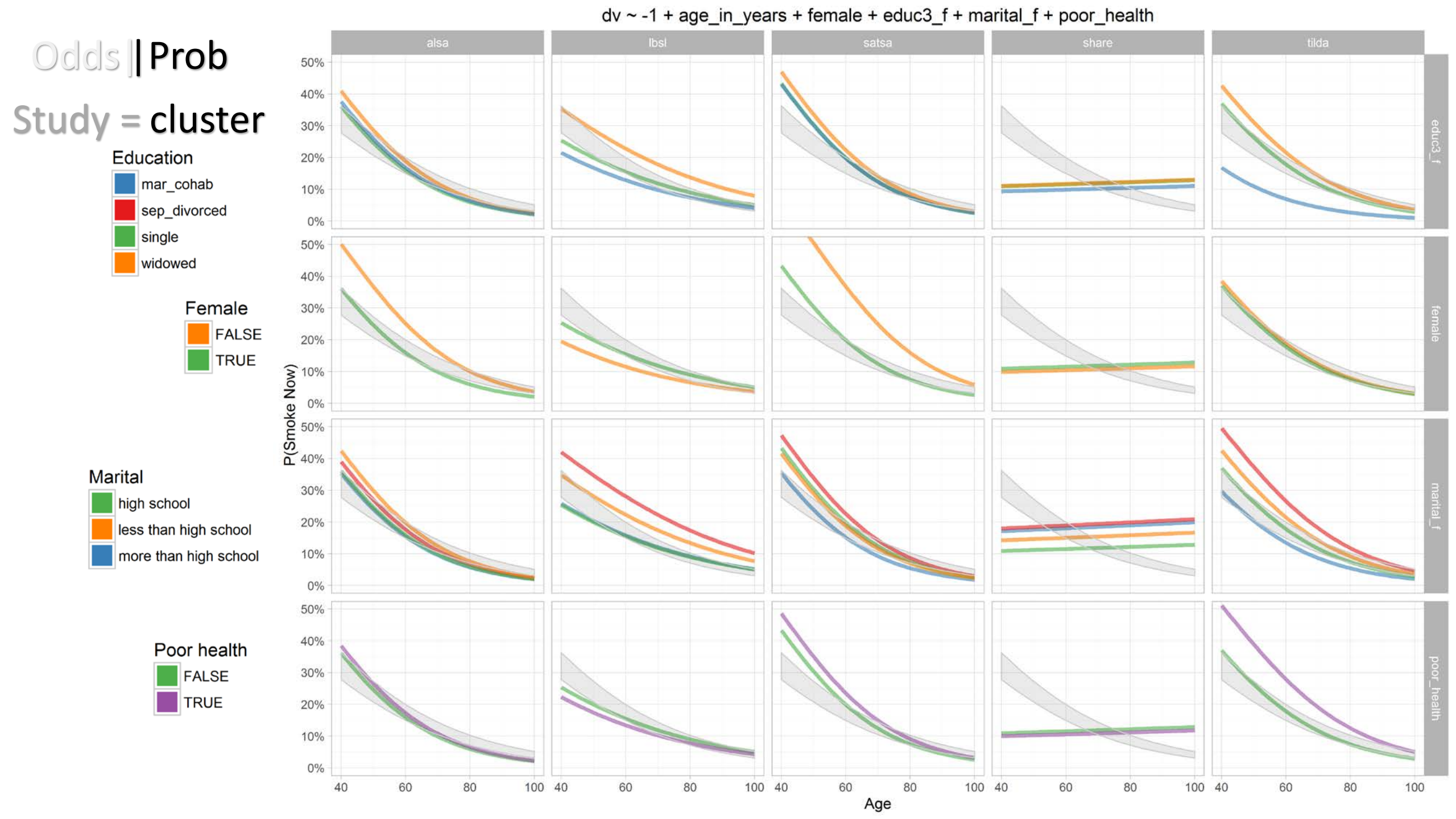
Odds | Prob

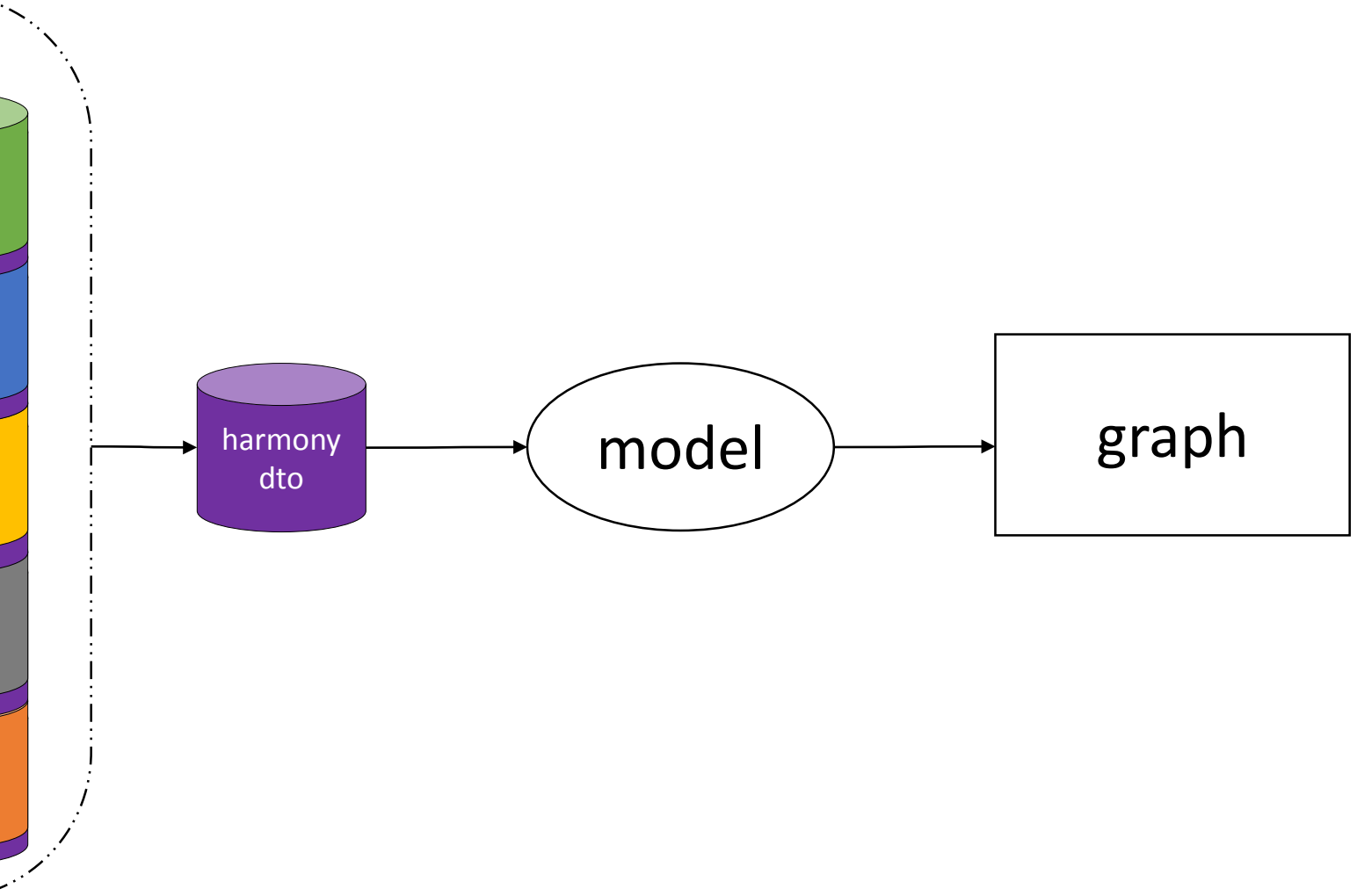
Study = cluster



Odds | Prob

Study = cluster

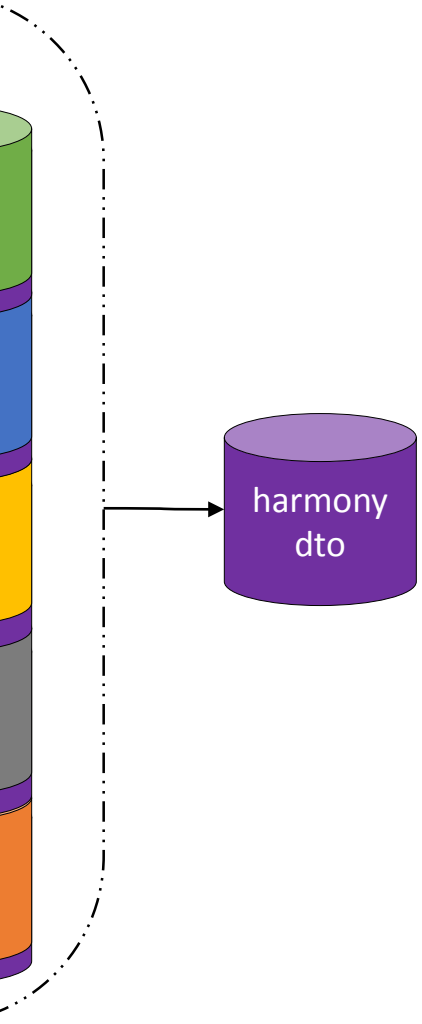


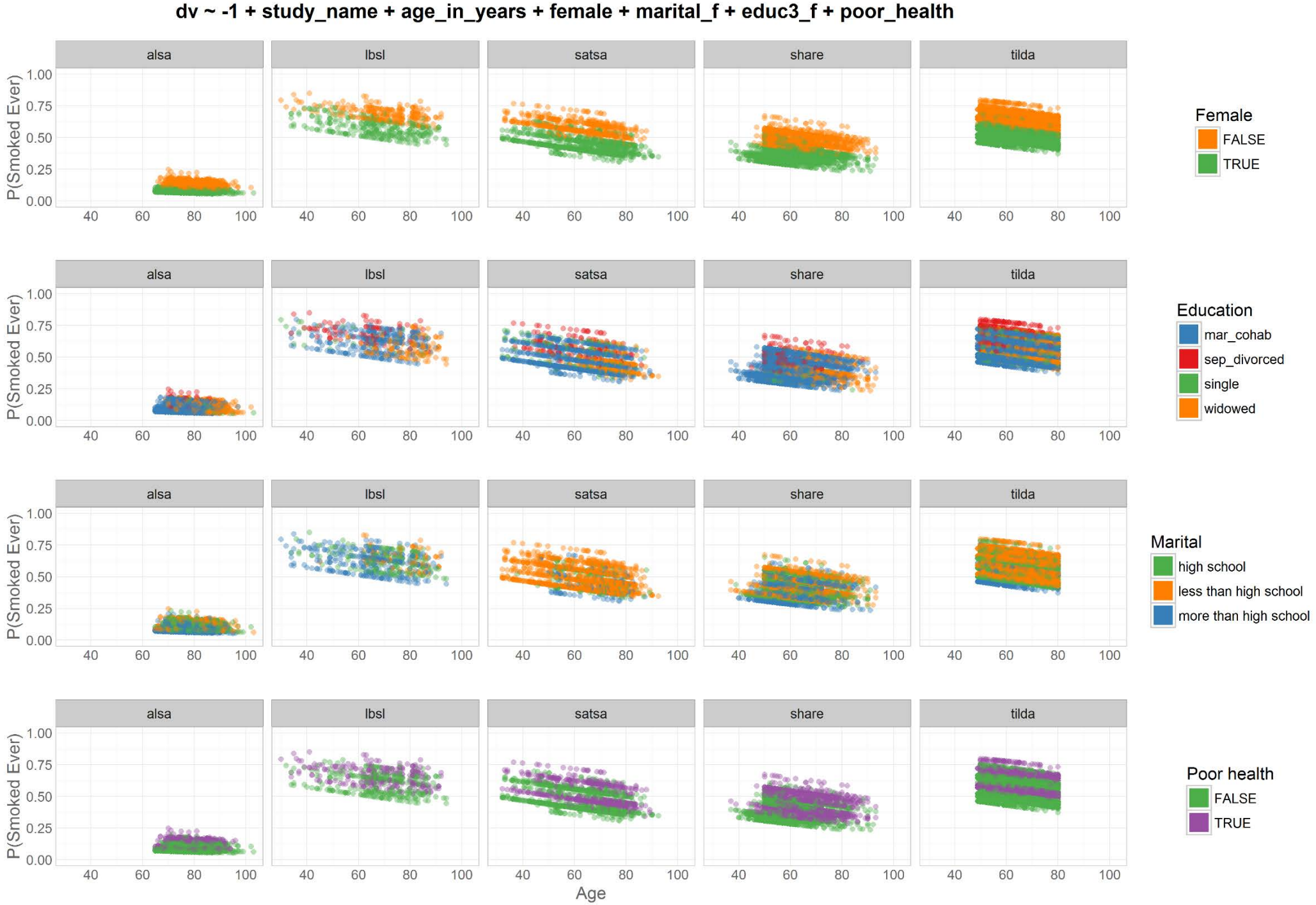


smoked_ever

Odds | Prob

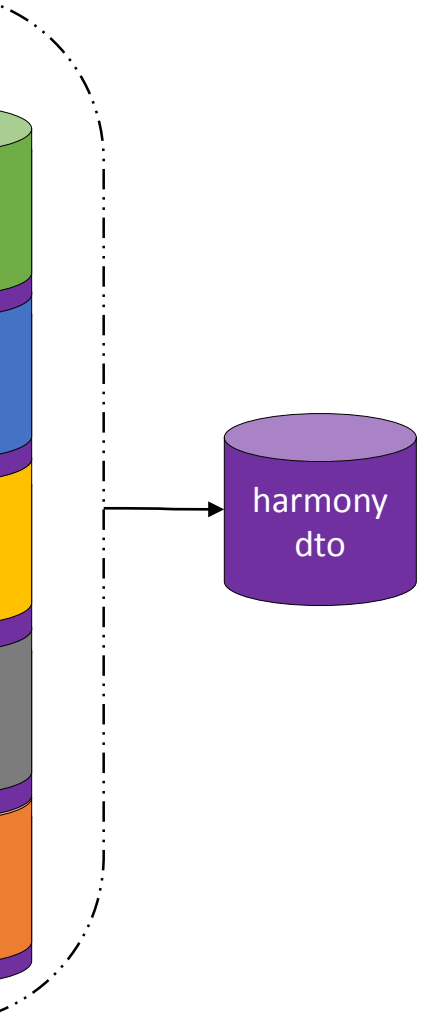
Study = factor



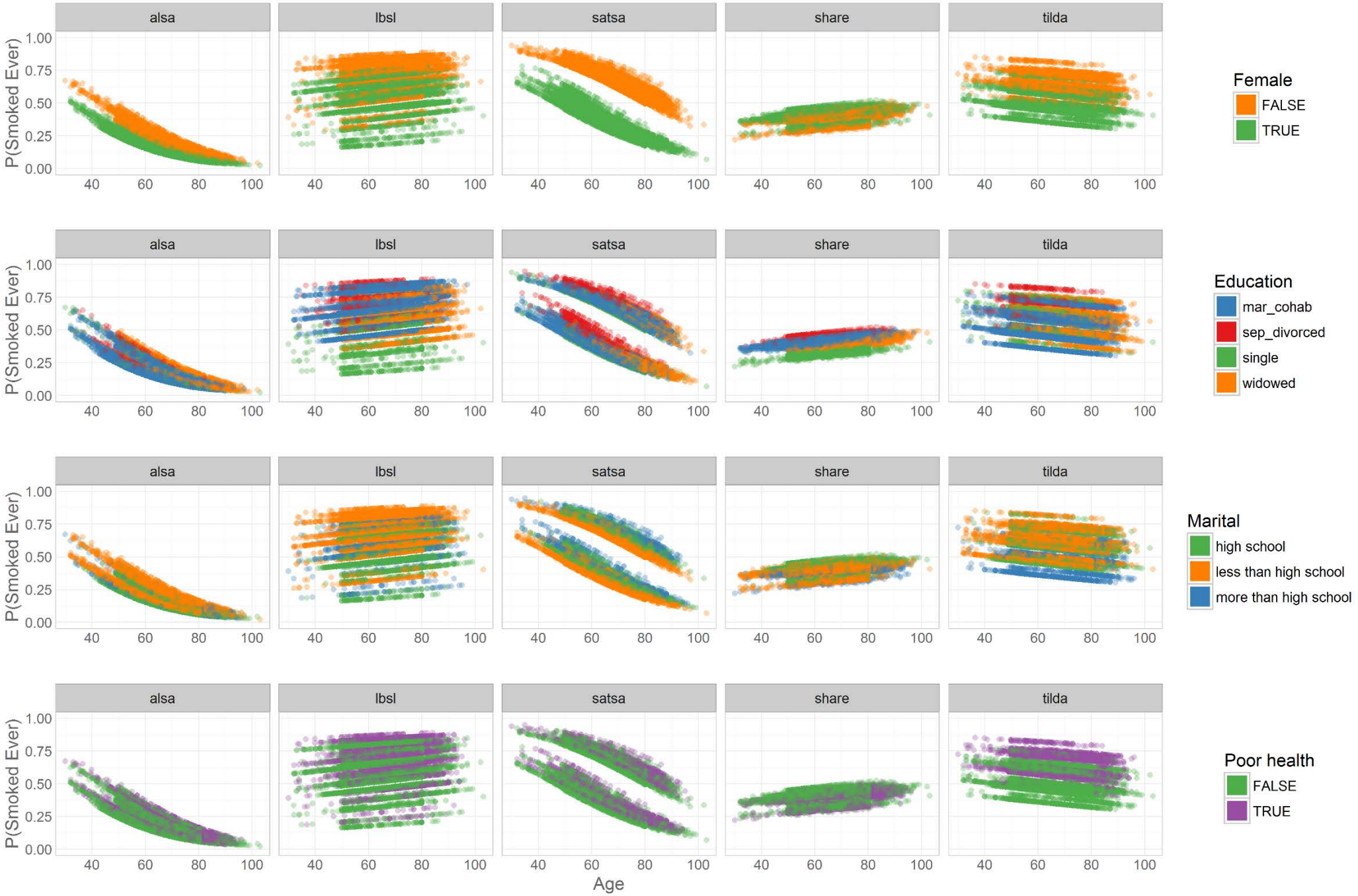


Odds | Prob

Study = cluster

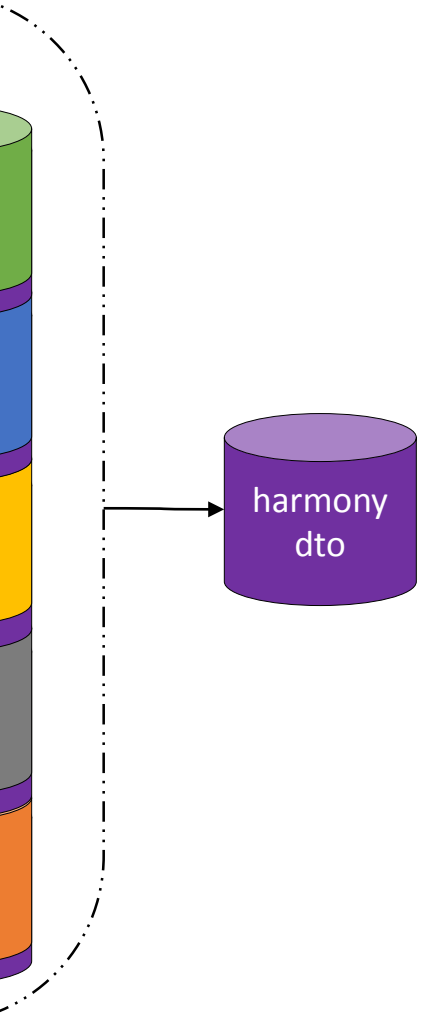


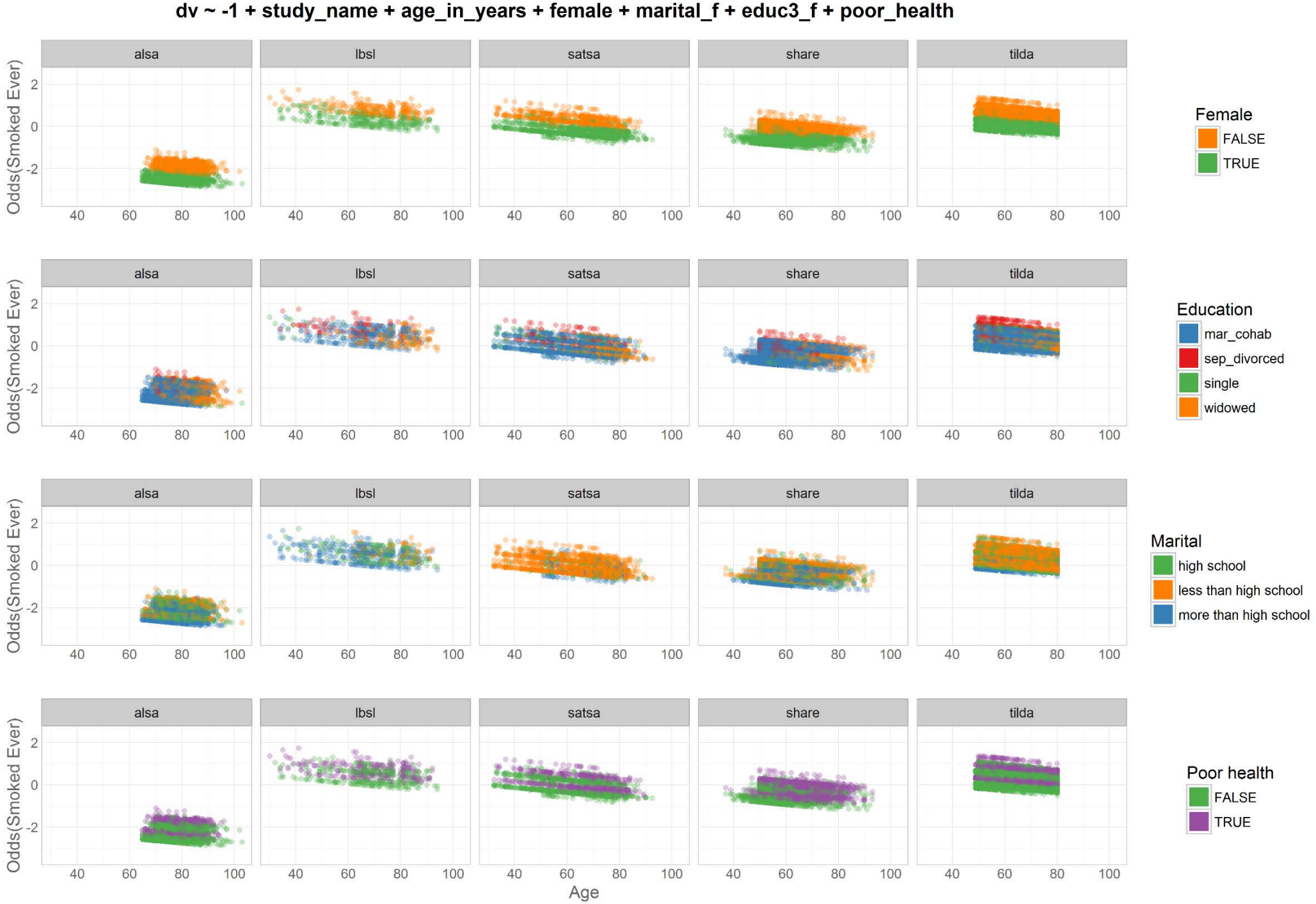
harmony dto



Odds | Prob

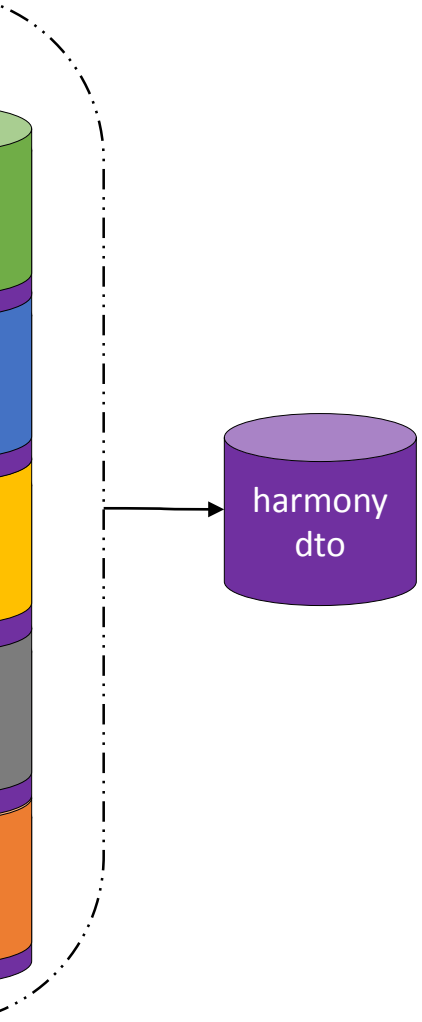
Study = factor

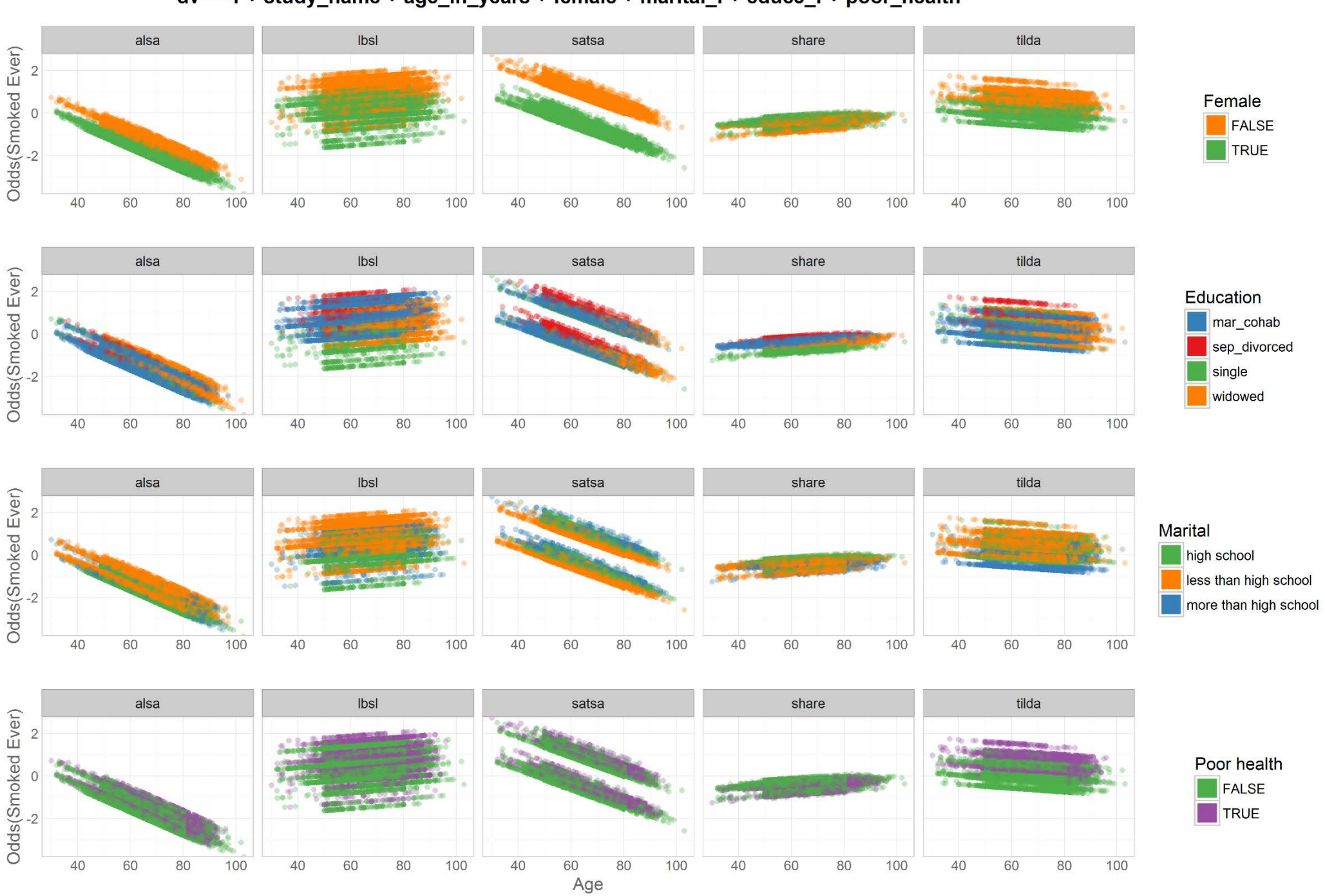




Odds | Prob

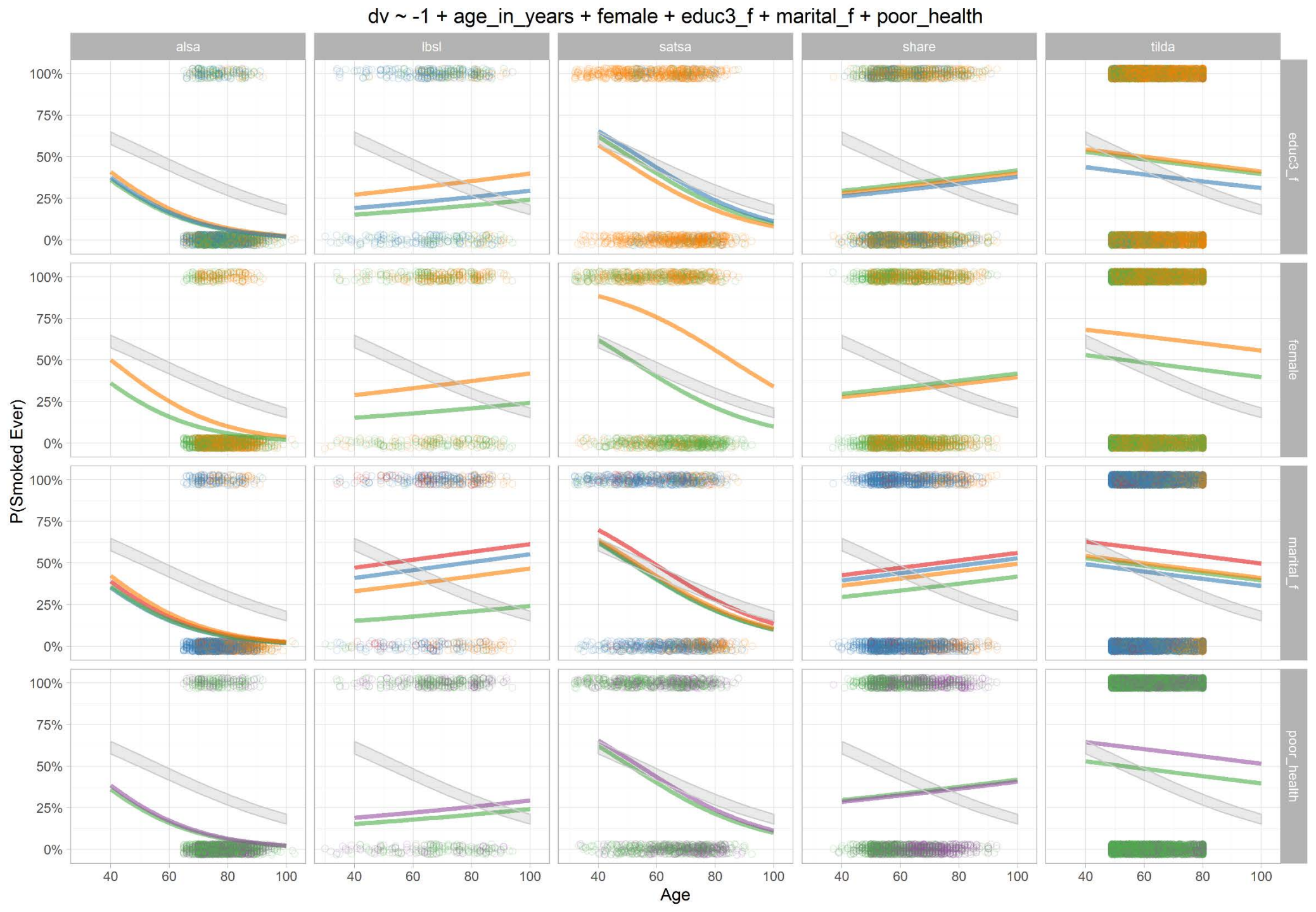
Study = cluster





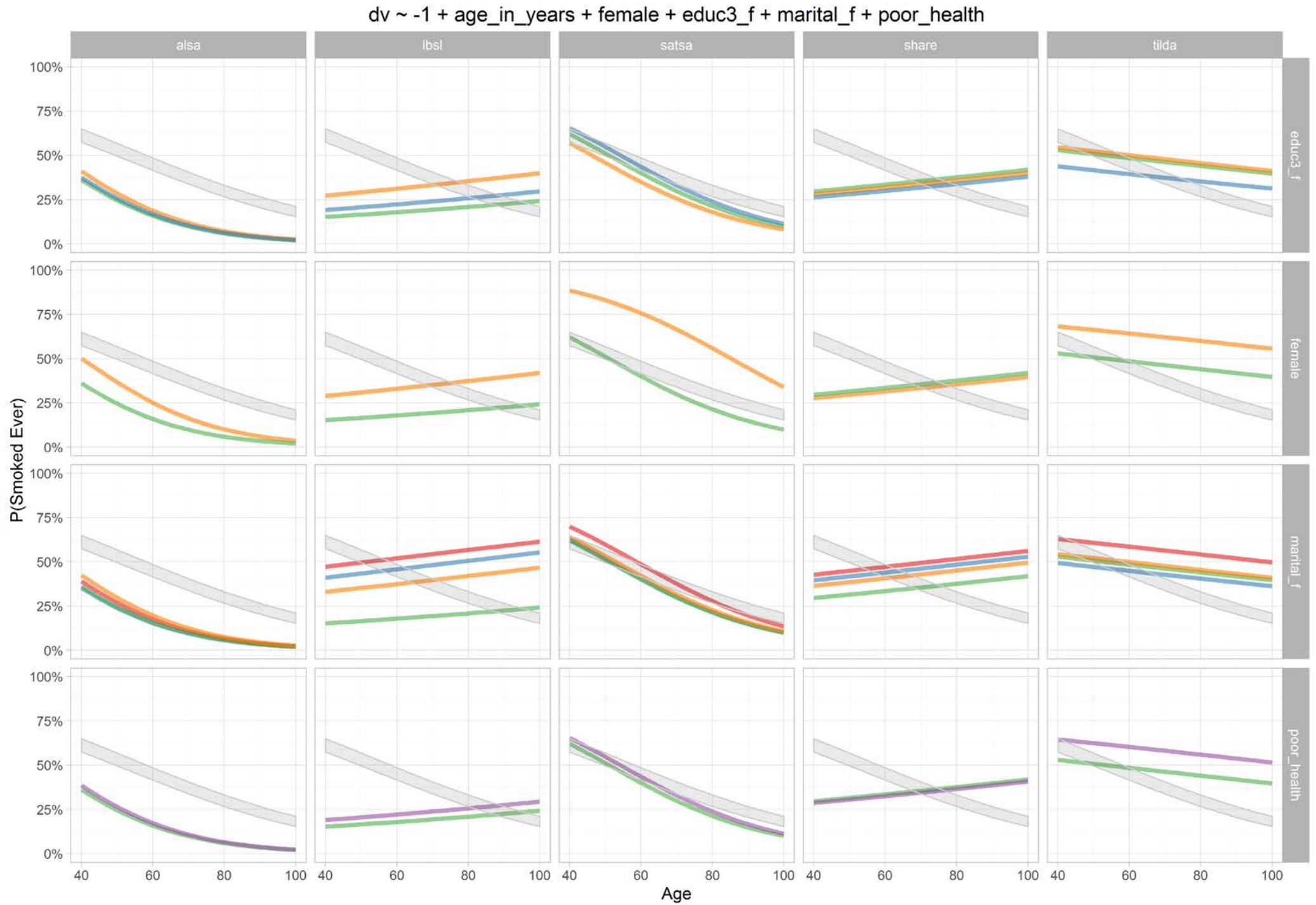
Odds | Prob

Study = cluster



Odds | Prob

Study = cluster



5 voices, 10 stories:

Groningen Harmonization Exercise

IALSA team

- Andrey Koval, University of Victoria
- Andrea Piccinin, University of Victoria
- Scott Hofer, University of Victoria
- Graciela Muniz, University College London

Special thanks

- William Beasley, University of Oklahoma