INTERNATIONAL
Population Data Linkage
NETWORK

November 2020

USING REPRODUCIBLE DATA VISUALIZATIONS TO AUGMENT DECISION-MAKING DURING SUPPRESSION OF SMALL COUNTS

Andriy Koval
Assistant Professor
UNIVERSITY OF CENTRAL FLORIDA

Kate Smolina
Director
British Columbia BSERVATORY Population & Public Health
UBC

Anthony Leamon
Regional Epidemiologist
vancouver island health authority

IPDLN2020 is proudly sponsored by

GOLD SPONSORS

Australian Government Australian Institute of Health and Welfare
PHRN Population Health Research Network
SA·NT DataLink

SILVER SPONSORS

IC/ES Data Discovery Better Health
UNIVERSITY OF CALGARY O'Brien Institute for Public Health

BRONZE SPONSORS

University of South Australia
THE UNIVERSITY OF WESTERN AUSTRALIA

2020-11-10

CIHR IRSC

Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

Health System Impact Fellowship: Key Objectives

1

Support Impact-Oriented Career Paths

- Elevate the career readiness and expand the career options for doctoral fellows in HSPR through experiential learning and othe program enhancements (cohort + enriched competency training

2

Expand and Enrich the Traditional Training Environment

- Engage health system and related organizations in preparing a cadre of promising PhD graduates for impactful careers.

3

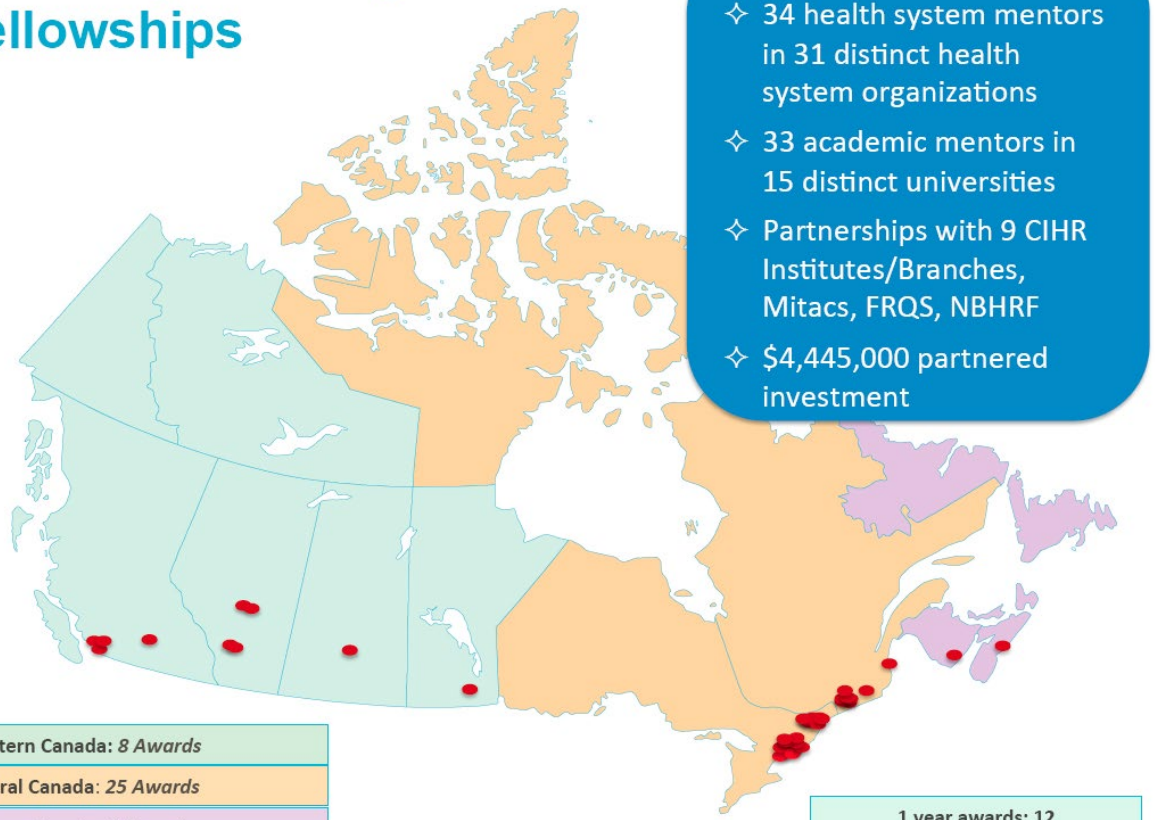
Increase Organizations' Awareness of the Value of PhD-trained Individuals

- Provide health system and related organizations with direct opportunities to realize and harness the benefits that PhD-trained individuals can bring to such organizations.

Health System Impact Fellowships

35 HSI Fellows:

- 34 health system mentors in 31 distinct health system organizations
- 33 academic mentors in 15 distinct universities
- Partnerships with 9 CIHR Institutes/Branches, Mitacs, FRQS, NBHRF
- $4,445,000 partnered investment



Western Canada: *8 Awards*
Central Canada: *25 Awards*
Eastern Canada: *2 Awards*

1 year awards: 12
2 year awards: 23

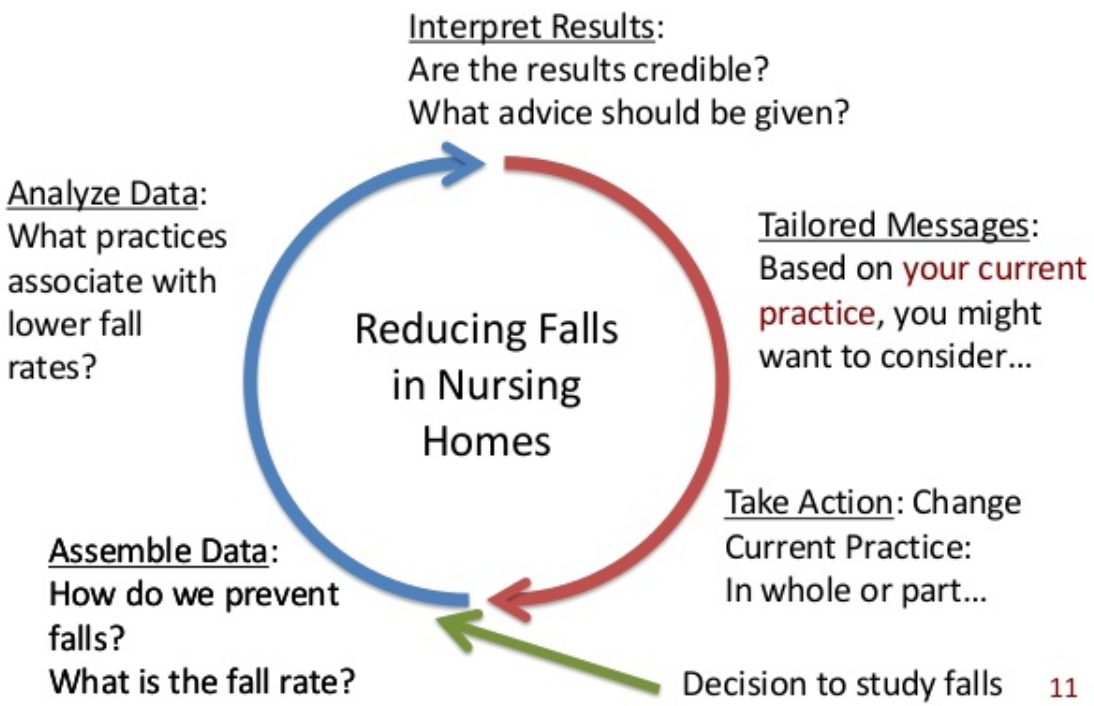
Dr. Meghan McMahon
IHSPR Associate Director

Dr. Robyn Tamblyn
IHSPR Scientific Director

CIHR IRSC

Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

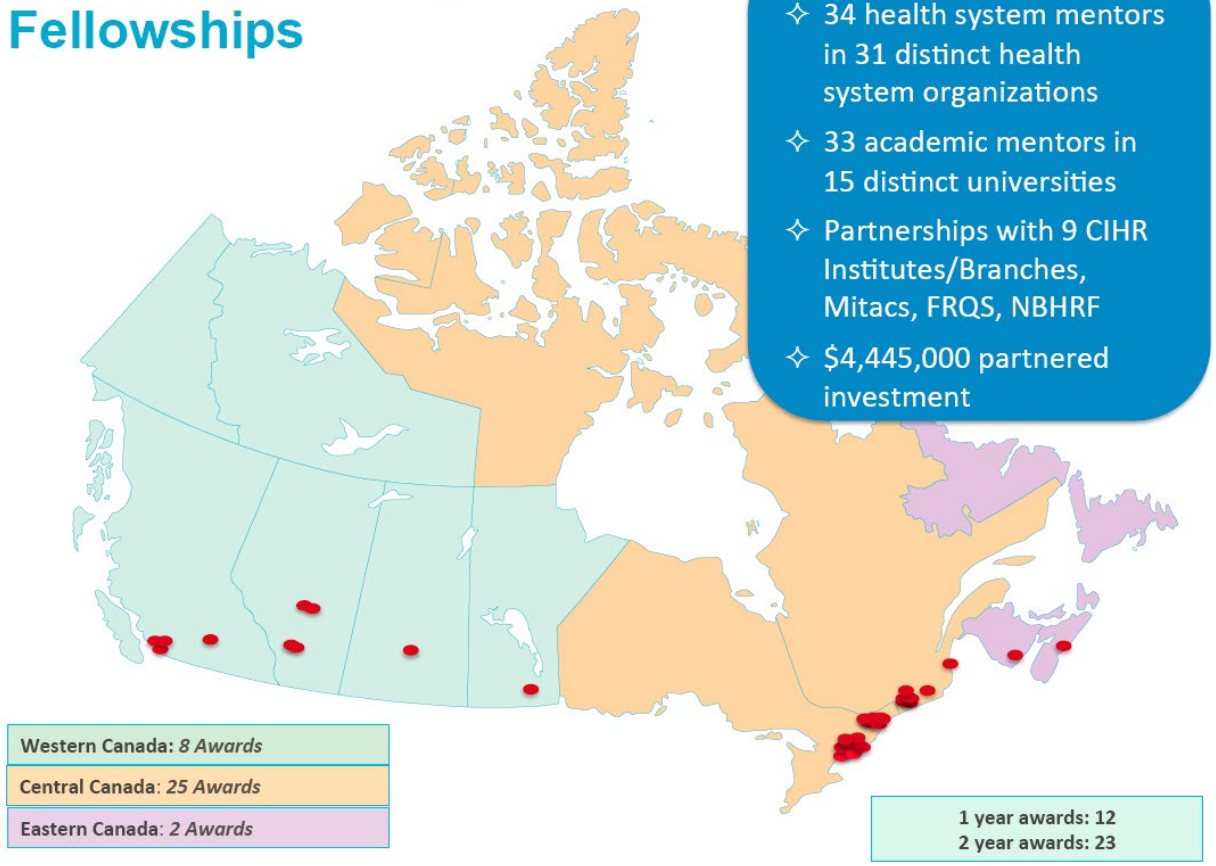
Example of A Virtuous Learning Cycle



Health System Impact Fellowships

35 HSI Fellows:

- 34 health system mentors in 31 distinct health system organizations
- 33 academic mentors in 15 distinct universities
- Partnerships with 9 CIHR Institutes/Branches, Mitacs, FRQS, NBHRF
- $4,445,000 partnered investment



The Learning Health System

Health systems--at any level of scale--become learning systems when they can, continuously and routinely, study and improve themselves

Source: Friedman et al., 2015

British Columbia
BSERVATORY
Population & Public Health

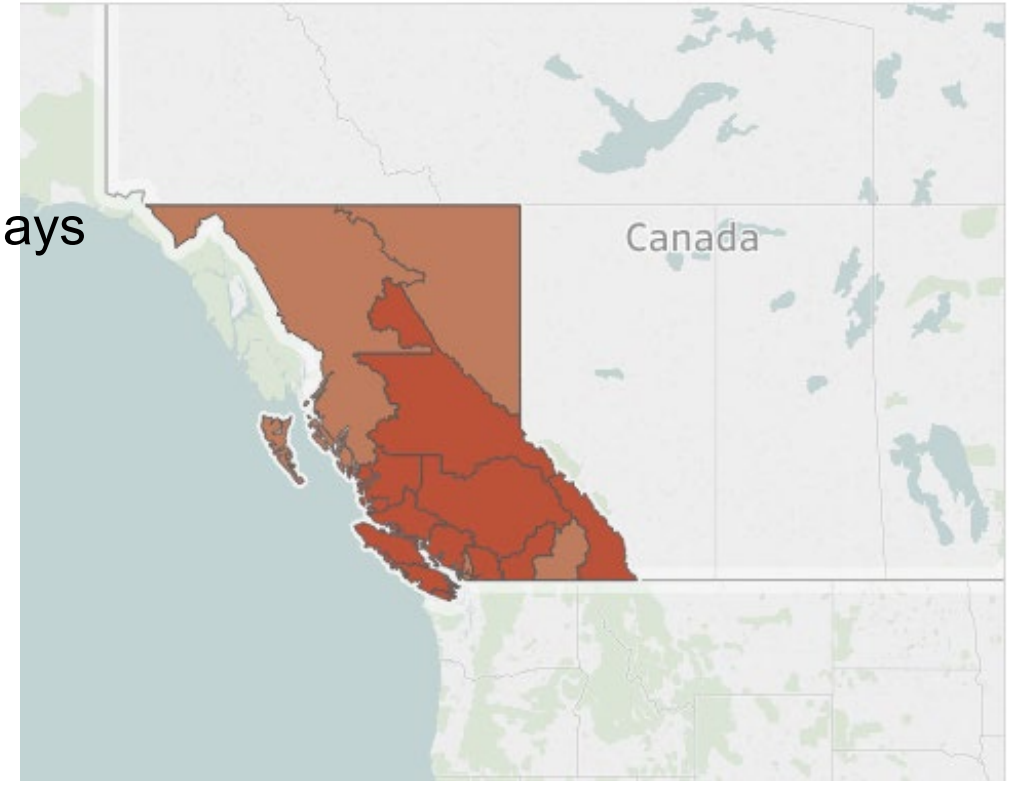
THE UNIVERSITY OF BRITISH COLUMBIA

How can I contribute to Learning Health Systems?

A. Popularize reproducible workflows

B. Bridge data sources and information displays

C. Build learning communities



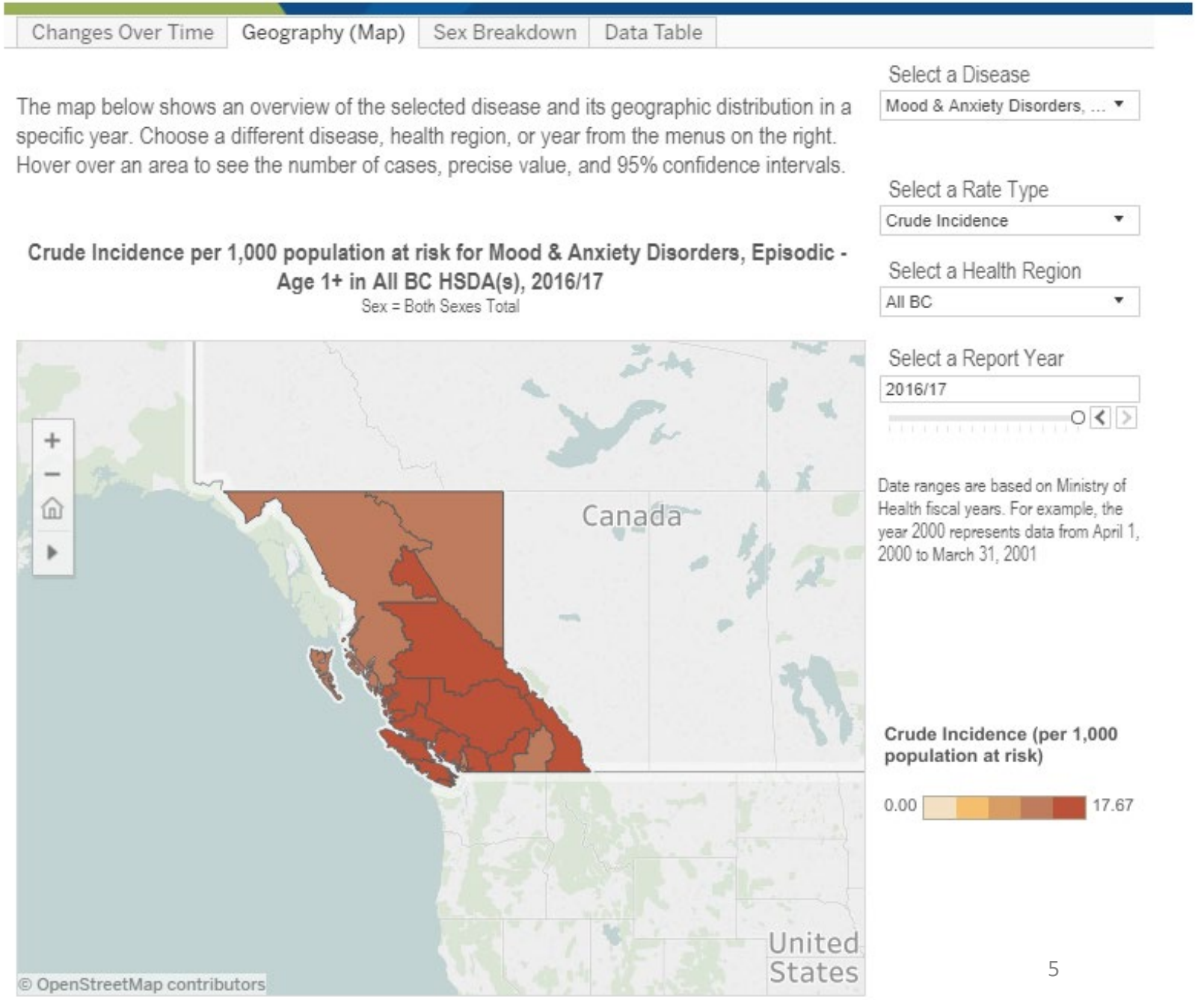
British Columbia BSERVATORY Population & Public Health

THE UNIVERSITY OF BRITISH COLUMBIA

http://www.bccdc.ca/health-professionals/data-reports/chronic-disease-dashboard

Hierarchy of administrative units in **BC**: Health Service Delivery Areas (**HSDA**) nested within Health Authorities (**HA**).

label_prov	label_ha	label_hsda
BC	Interior	East Kootenay
BC	Interior	Kootenay Boundary
BC	Interior	Okanagan
BC	Interior	Thompson Cariboo Shuswap
BC	Fraser	Fraser East
BC	Fraser	Fraser North
BC	Fraser	Fraser South
BC	Vancouver Coastal	Richmond
BC	Vancouver Coastal	Vancouver
BC	Vancouver Coastal	North Shore/Coast Garibaldi
BC	Vancouver Island	South Vancouver Island
BC	Vancouver Island	Central Vancouver Island
BC	Vancouver Island	North Vancouver Island
BC	Northern	Northwest
BC	Northern	Northern Interior
BC	Northern	Northeast



British Columbia BSERVATORY Population & Public Health

THE UNIVERSITY OF BRITISH COLUMBIA

Ministry Of Health

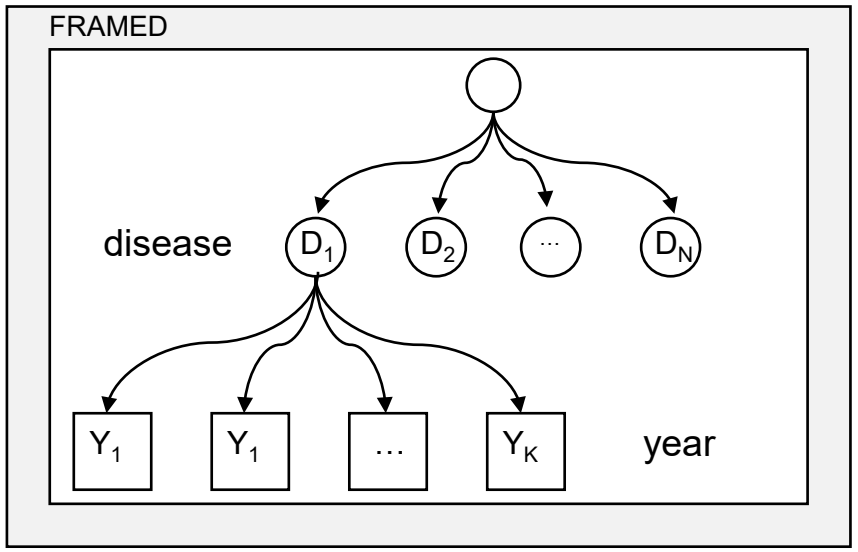
disease	year	sex	incase	region_id	region_label	desc_label
Flower Deafness	1995	F	104	0	BC	BC
Flower Deafness	1995	F	25	1	HA	Interior
Flower Deafness	1995	F	15	2	HA	Fraser
Flower Deafness	1995	F	20	3	HA	Vancouver Coastal
Flower Deafness	1995	F	19	4	HA	Vancouver Island
Flower Deafness	1995	F	25	5	HA	Northern
Flower Deafness	1995	F	9	11	HSDA	East Kootenay
Flower Deafness	1995	F	7	12	HSDA	Kootenay Boundary
Flower Deafness	1995	F	4	13	HSDA	Okanagan
Flower Deafness	1995	F	5	14	HSDA	Thompson Cariboo Shuswap
Flower Deafness	1995	F	1	21	HSDA	Fraser East
Flower Deafness	1995	F	8	22	HSDA	Fraser North
Flower Deafness	1995	F	4	23	HSDA	Fraser South
Flower Deafness	1995	F	5	31	HSDA	Richmond
Flower Deafness	1995	F	7	32	HSDA	Vancouver
Flower Deafness	1995	F	8	33	HSDA	North Shore/Coast Garibaldi
Flower Deafness	1995	F	8	41	HSDA	South Vancouver Island
Flower Deafness	1995	F	5	42	HSDA	Central Vancouver Island
Flower Deafness	1995	F	6	43	HSDA	North Vancouver Island
Flower Deafness	1995	F	7	51	HSDA	Northwest
Flower Deafness	1995	F	9	52	HSDA	Northern Interior
Flower Deafness	1995	F	9	53	HSDA	Northeast
Flower Deafness	1995	M	97	0	BC	BC
Flower Deafness	1995	M	22	1	HA	Interior
Flower Deafness	1995	M	42	2	HA	Fraser
Flower Deafness	1995	M	21	3	HA	Vancouver Coastal
Flower Deafness	1995	M	19	4	HA	Vancouver Island
Flower Deafness	1995	M	20	5	HA	Northern
Flower Deafness	1995	M	7	11	HSDA	East Kootenay
Flower Deafness	1995	M	7	12	HSDA	Kootenay Boundary

Must answer the questions:

1. *Is the cell value smaller than a threshold?*
2. *Can re-calculate from gender triplet?*
3. *Can re-calculate from higher-order total?*

Challenges with Manual method

1. Arduous
2. Time consuming
3. Prone to human error



FLOWER DEAFNESS - 1995

HA	HSDA	HSDA			HA			PROV		
Interior	East Kootenay	9	7	16	25	22	47	102	97	199
Interior	Kootenay Boundary	7	7	14	25	22	47	102	97	199
Interior	Okanagan	4	5	9	25	22	47	102	97	199
Interior	Thompson Cariboo Shuswap	5	3	8	25	22	47	102	97	199
Fraser	Fraser East	1	3	4	13	15	28	102	97	199
Fraser	Fraser North	8	4	12	13	15	28	102	97	199
Fraser	Fraser South	4	8	12	13	15	28	102	97	199
Vancouver Coastal	Richmond	5	3	8	20	21	41	102	97	199
Vancouver Coastal	Vancouver	7	9	16	20	21	41	102	97	199
Vancouver Coastal	North Shore/Coast Garibaldi	8	9	17	20	21	41	102	97	199
Vancouver Island	South Vancouver Island	8	7	15	19	19	38	102	97	199
Vancouver Island	Central Vancouver Island	5	5	10	19	19	38	102	97	199
Vancouver Island	North Vancouver Island	6	7	13	19	19	38	102	97	199
Northern	Northwest	7	8	15	25	20	45	102	97	199
Northern	Northern Interior	9	9	18	25	20	45	102	97	199
Northern	Northeast	9	3	12	25	20	45	102	97	199
		F	M	T	F	M	T	F	M	T

Approach in a nutshell

1. Split into frames
2. Apply redaction logic
3. Print a **graph** of each frame

Requirements for Automation

1. Reproducible
2. Verifiable
3. Extendable
4. Approachable

Reproducible scripts and vignettes available from:

https://github.com/IHACRU/suppress-for-release

LOGICAL TESTS

1. *Is the cell value smaller than a threshold?*
2. *Can re-calculate from gender triplet?*
3. *Can re-calculate from higher-order total?*

Preserved by targeting

FLOWER DEAFNESS - 1995

HA	HSDA	HSDA			HA			PROV		
		F	M	T	F	M	T	F	M	T
Interior	East Kootenay	9	7	16	25	22	47	102	97	199
Interior	Kootenay Boundary	7	7	14	25	22	47	102	97	199
Interior	Okanagan	4	5	9	25	22	47	102	97	199
Interior	Thompson Cariboo Shuswap	5	3	8	25	22	47	102	97	199
Fraser	Fraser East	1	3	4	13	15	28	102	97	199
Fraser	Fraser North	8	4	12	13	15	28	102	97	199
Fraser	Fraser South	4	8	12	13	15	28	102	97	199
Vancouver Coastal	Richmond	5	3	8	20	21	41	102	97	199
Vancouver Coastal	Vancouver	7	9	16	20	21	41	102	97	199
Vancouver Coastal	North Shore/Coast Garibaldi	8	9	17	20	21	41	102	97	199
Vancouver Island	South Vancouver Island	8	7	15	19	19	38	102	97	199
Vancouver Island	Central Vancouver Island	5	5	10	19	19	38	102	97	199
Vancouver Island	North Vancouver Island	6	7	13	19	19	38	102	97	199
Northern	Northwest	7	8	15	25	20	45	102	97	199
Northern	Northern Interior	9	9	18	25	20	45	102	97	199
Northern	Northeast	9	3	12	25	20	45	102	97	199

Reproducible scripts and vignettes available from:

https://github.com/IHACRU/suppress-for-release

LOGICAL TESTS

1. *Is the cell value smaller than a threshold?*
2. *Can re-calculate from gender triplet?*
3. *Can re-calculate from higher-order total?*

Preserved by targeting

FLOWER DEAFNESS - 1995

HA	HSDA	HSDA			HA			PROV		
Interior	East Kootenay	9	7	16	25	22	47	102	97	199
Interior	Kootenay Boundary	7	7	14	25	22	47	102	97	199
Interior	Okanagan	4	5	9	25	22	47	102	97	199
Interior	Thompson Cariboo Shuswap	5	3	8	25	22	47	102	97	199
Fraser	Fraser East	1	3	4	13	15	28	102	97	199
Fraser	Fraser North	8	4	12	13	15	28	102	97	199
Fraser	Fraser South	4	8	12	13	15	28	102	97	199
Vancouver Coastal	Richmond	5	3	8	20	21	41	102	97	199
Vancouver Coastal	Vancouver	7	9	16	20	21	41	102	97	199
Vancouver Coastal	North Shore/Coast Garibaldi	8	9	17	20	21	41	102	97	199
Vancouver Island	South Vancouver Island	8	7	15	19	19	38	102	97	199
Vancouver Island	Central Vancouver Island	5	5	10	19	19	38	102	97	199
Vancouver Island	North Vancouver Island	6	7	13	19	19	38	102	97	199
Northern	Northwest	7	8	15	25	20	45	102	97	199
Northern	Northern Interior	9	9	18	25	20	45	102	97	199
Northern	Northeast	9	3	12	25	20	45	102	97	199
		F	M	T	F	M	T	F	M	T

Reproducible scripts and vignettes available from:

https://github.com/IHACRU/suppress-for-release

LOGICAL TESTS

1. *Is the cell value smaller than a threshold?*
2. *Can re-calculate from gender triplet?*
3. *Can re-calculate from higher-order total?*

Preserved by targeting

FLOWER DEAFNESS - 1995 - draconian

HA	HSDA	HSDA			HA			PROV		
		F	M	T	F	M	T	F	M	T
Interior	East Kootenay	9	7	16	25	22	47	102	97	199
Interior	Kootenay Boundary	7	7	14	25	22	47	102	97	199
Interior	Okanagan	4	5	9	25	22	47	102	97	199
Interior	Thompson Cariboo Shuswap	5	3	8	25	22	47	102	97	199
Fraser	Fraser East	1	3	4	13	15	28	102	97	199
Fraser	Fraser North	8	4	12	13	15	28	102	97	199
Fraser	Fraser South	4	8	12	13	15	28	102	97	199
Vancouver Coastal	Richmond	5	3	8	20	21	41	102	97	199
Vancouver Coastal	Vancouver	7	9	16	20	21	41	102	97	199
Vancouver Coastal	North Shore/Coast Garibaldi	8	9	17	20	21	41	102	97	199
Vancouver Island	South Vancouver Island	8	7	15	19	19	38	102	97	199
Vancouver Island	Central Vancouver Island	5	5	10	19	19	38	102	97	199
Vancouver Island	North Vancouver Island	6	7	13	19	19	38	102	97	199
Northern	Northwest	7	8	15	25	20	45	102	97	199
Northern	Northern Interior	9	9	18	25	20	45	102	97	199
Northern	Northeast	9	3	12	25	20	45	102	97	199

Reproducible scripts and vignettes available from:

https://github.com/IHACRU/suppress-for-release

LOGICAL TESTS

1. *Is the cell value smaller than a threshold?*
2. *Can re-calculate from gender triplet?*
3. *Can re-calculate from higher-order total?*

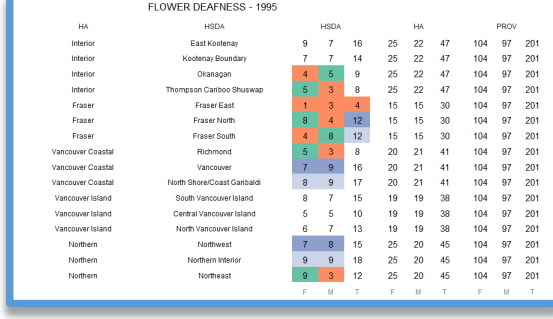
Preserved by targeting

FLOWER DEAFNESS - 1995 - targeted

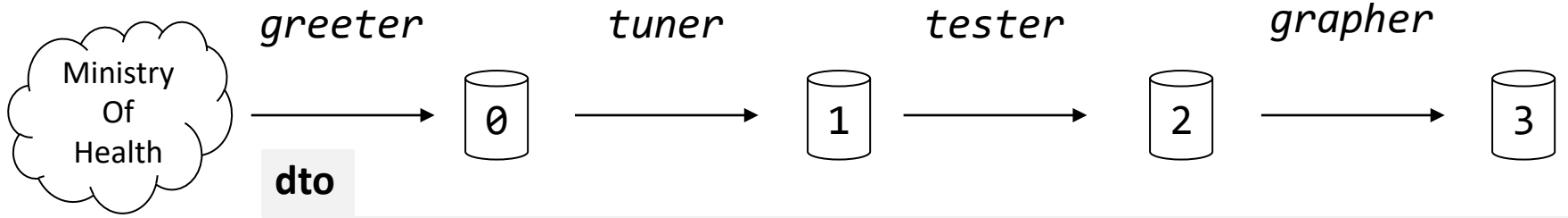
HA	HSDA	HSDA			HA			PROV		
Interior	East Kootenay	9	7	16	25	22	47	102	97	199
Interior	Kootenay Boundary	7	7	14	25	22	47	102	97	199
Interior	Okanagan	4	5	9	25	22	47	102	97	199
Interior	Thompson Cariboo Shuswap	5	3	8	25	22	47	102	97	199
Fraser	Fraser East	1	3	4	13	15	28	102	97	199
Fraser	Fraser North	8	4	12	13	15	28	102	97	199
Fraser	Fraser South	4	8	12	13	15	28	102	97	199
Vancouver Coastal	Richmond	5	3	8	20	21	41	102	97	199
Vancouver Coastal	Vancouver	7	9	16	20	21	41	102	97	199
Vancouver Coastal	North Shore/Coast Garibaldi	8	9	17	20	21	41	102	97	199
Vancouver Island	South Vancouver Island	8	7	15	19	19	38	102	97	199
Vancouver Island	Central Vancouver Island	5	5	10	19	19	38	102	97	199
Vancouver Island	North Vancouver Island	6	7	13	19	19	38	102	97	199
Northern	Northwest	7	8	15	25	20	45	102	97	199
Northern	Northern Interior	9	9	18	25	20	45	102	97	199
Northern	Northeast	9	3	12	25	20	45	102	97	199
		F	M	T	F	M	T	F	M	T

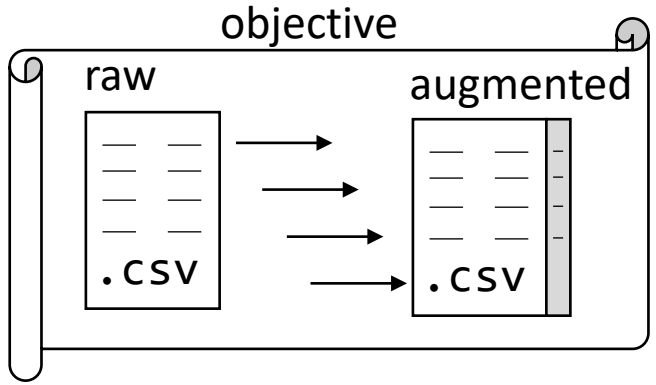
Reproducible scripts and vignettes available from:

https://github.com/IHACRU/suppress-for-release



Workflow for `./suppress-for-release/manipulation/`





data **t**ransfer **o**bject

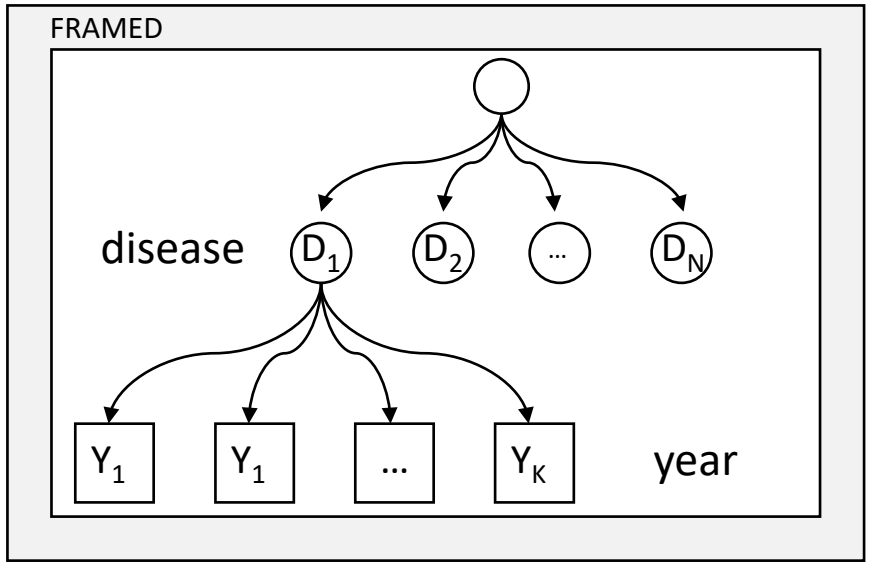
- ○ list object
- ▪ data frame
- ❖ graph

dto

PRODUCED COMPONETNS

- ▪ raw [L]
- ▪ meta [L]
- ▪ target [W]
- ○ FRAMED
  - ▪ raw [L]
- ○ FRAMED
  - ▪ cleaned [L]
  - ▪ tuned [W]
- ○ FRAMED
  - ▪ test 1 [W]
  - ▪ test 2 [W]
  - ▪ test 3 [W]
- ○ FRAMED
  - ▪ redacted [L]
  - ❖ plotted [W]
- ▪ augmented [L]

Workflow maps help you remember key concepts



[Long]

disease	year	sex	incase	region_id	region_label	desc_label
Flower Deafness	1995	F	104	0	BC	BC
Flower Deafness	1995	F	25	1	HA	Interior
Flower Deafness	1995	F	15	2	HA	Fraser
Flower Deafness	1995	F	20	3	HA	Vancouver Coastal
Flower Deafness	1995	F	19	4	HA	Vancouver Island
Flower Deafness	1995	F	25	5	HA	Northern
Flower Deafness	1995	F	9	11	HSDA	East Kootenay
Flower Deafness	1995	F	7	12	HSDA	Kootenay Boundary
Flower Deafness	1995	F	4	13	HSDA	Okanagan
Flower Deafness	1995	F	5	14	HSDA	Thompson Cariboo Shuswap
Flower Deafness	1995	F	1	21	HSDA	Fraser East
Flower Deafness	1995	F	8	22	HSDA	Fraser North
Flower Deafness	1995	F	4	23	HSDA	Fraser South
Flower Deafness	1995	F	5	31	HSDA	Richmond
Flower Deafness	1995	F	7	32	HSDA	Vancouver
Flower Deafness	1995	F	8	33	HSDA	North Shore/Coast Garibaldi
Flower Deafness	1995	F	8	41	HSDA	South Vancouver Island
Flower Deafness	1995	F	5	42	HSDA	Central Vancouver Island
Flower Deafness	1995	F	6	43	HSDA	North Vancouver Island
Flower Deafness	1995	F	7	51	HSDA	Northwest
Flower Deafness	1995	F	9	52	HSDA	Northern Interior
Flower Deafness	1995	F	9	53	HSDA	Northeast
Flower Deafness	1995	M	97	0	BC	BC
Flower Deafness	1995	M	22	1	HA	Interior
Flower Deafness	1995	M	42	2	HA	Fraser
Flower Deafness	1995	M	21	3	HA	Vancouver Coastal
Flower Deafness	1995	M	19	4	HA	Vancouver Island
Flower Deafness	1995	M	20	5	HA	Northern
Flower Deafness	1995	M	7	11	HSDA	East Kootenay
Flower Deafness	1995	M	7	12	HSDA	Kootenay Boundary

[Wide]

disease	year	label_prov	label_ha	label_hsda	HSDA_F	HSDA_M	HSDA_T	HA_F	HA_M	HA_T	BC_F	BC_M	BC_T
Flower Deafness	1995	BC	Interior	East Kootenay	9	7	16	25	22	47	104	97	201
Flower Deafness	1995	BC	Interior	Kootenay Boundary	7	7	14	25	22	47	104	97	201
Flower Deafness	1995	BC	Interior	Okanagan	4	5	9	25	22	47	104	97	201
Flower Deafness	1995	BC	Interior	Thompson Cariboo Shuswap	5	3	8	25	22	47	104	97	201
Flower Deafness	1995	BC	Fraser	Fraser East	1	3	4	15	15	30	104	97	201
Flower Deafness	1995	BC	Fraser	Fraser North	8	4	12	15	15	30	104	97	201
Flower Deafness	1995	BC	Fraser	Fraser South	4	8	12	15	15	30	104	97	201
Flower Deafness	1995	BC	Vancouver Coastal	Richmond	5	3	8	20	21	41	104	97	201
Flower Deafness	1995	BC	Vancouver Coastal	Vancouver	7	9	16	20	21	41	104	97	201
Flower Deafness	1995	BC	Vancouver Coastal	North Shore/Coast Garibaldi	8	9	17	20	21	41	104	97	201
Flower Deafness	1995	BC	Vancouver Island	South Vancouver Island	8	7	15	19	19	38	104	97	201
Flower Deafness	1995	BC	Vancouver Island	Central Vancouver Island	5	5	10	19	19	38	104	97	201
Flower Deafness	1995	BC	Vancouver Island	North Vancouver Island	6	7	13	19	19	38	104	97	201
Flower Deafness	1995	BC	Northern	Northwest	7	8	15	25	20	45	104	97	201
Flower Deafness	1995	BC	Northern	Northern Interior	9	9	18	25	20	45	104	97	201
Flower Deafness	1995	BC	Northern	Northeast	9	3	12	25	20	45	104	97	201

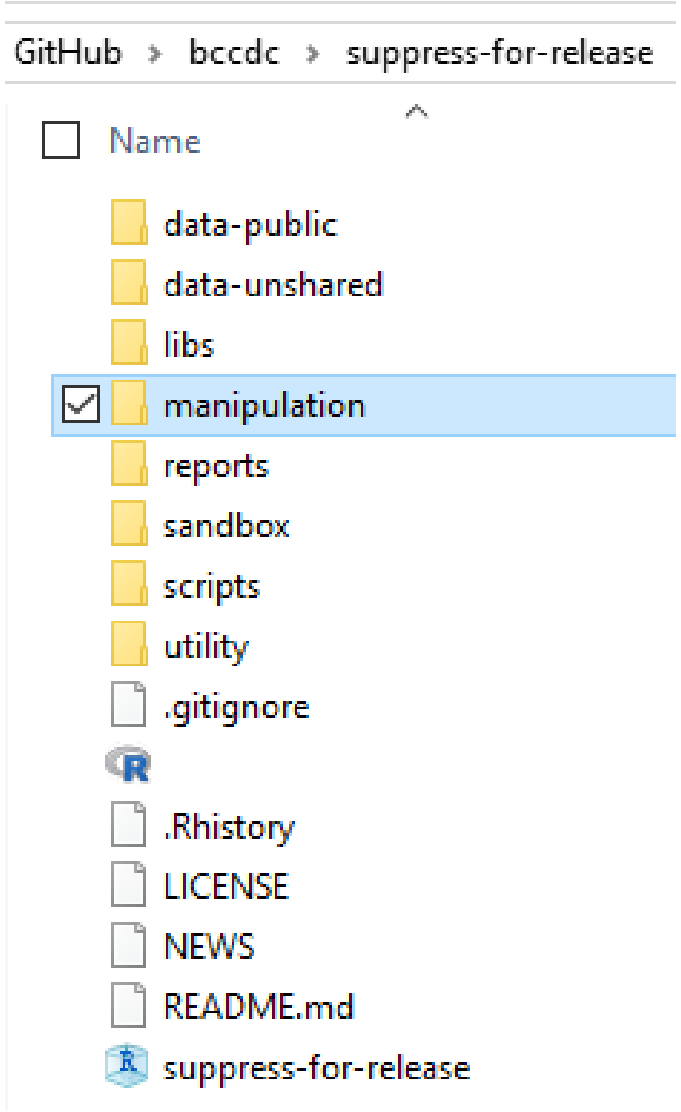
LOGICAL TESTS

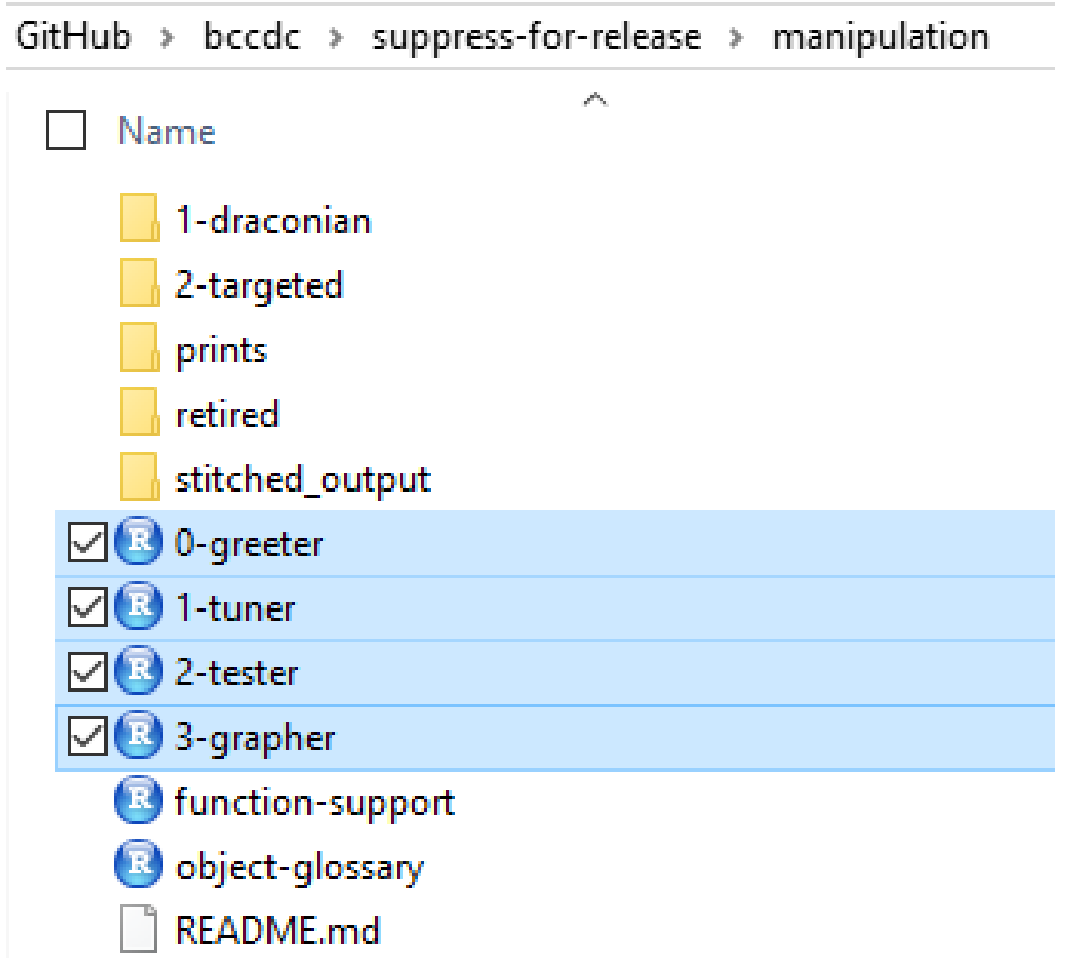
1. *Is the cell value smaller than a threshold?*
2. *Can re-calculate from gender triplet?*
3. *Can re-calculate from higher-order total?*

Preserved by targeting

Reproducible scripts and vignettes available from:

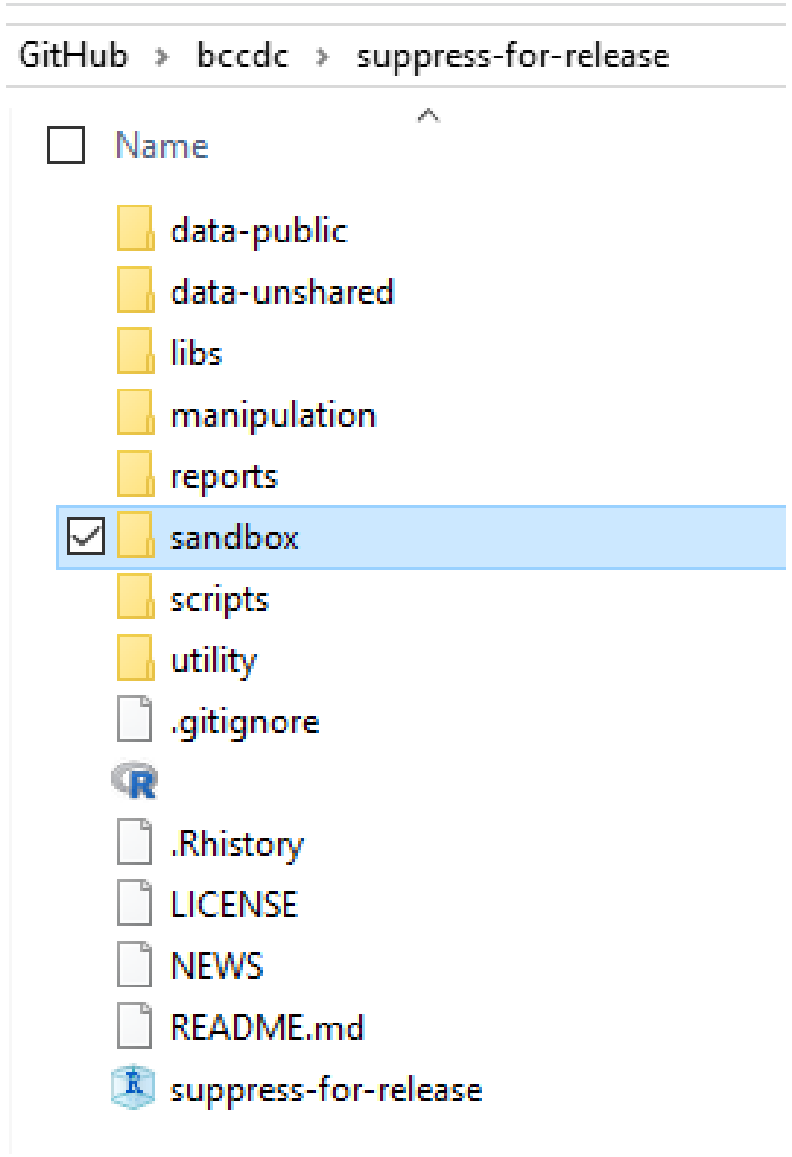
https://github.com/IHACRU/suppress-for-release

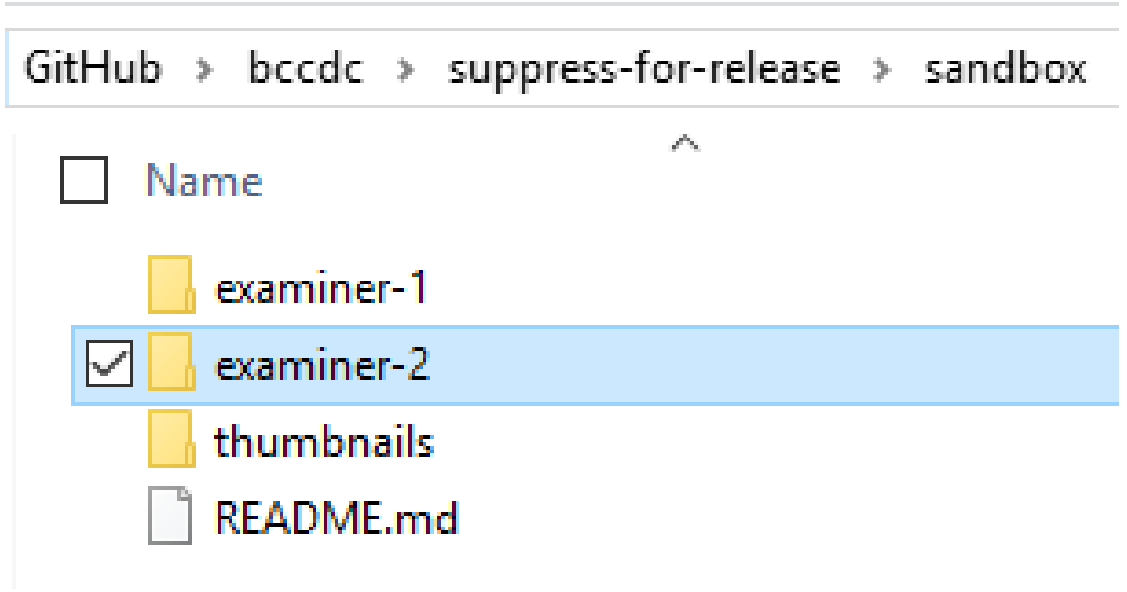


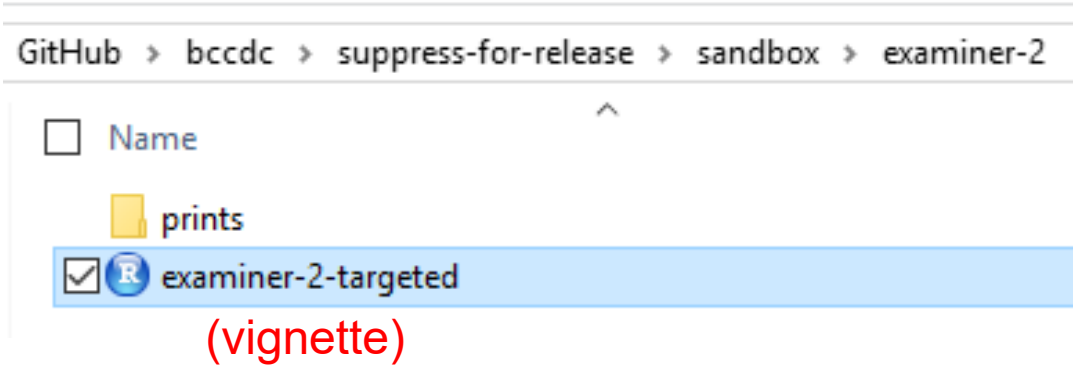


Reproducible scripts and vignettes available from:

https://github.com/IHACRU/suppress-for-release

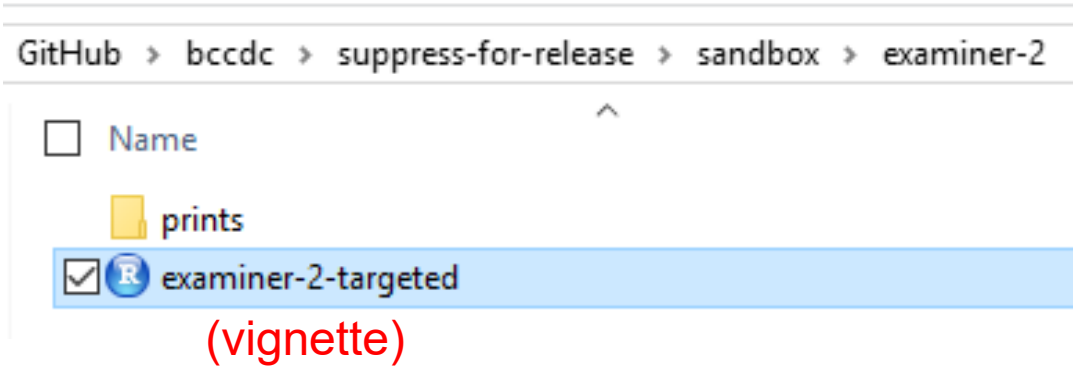




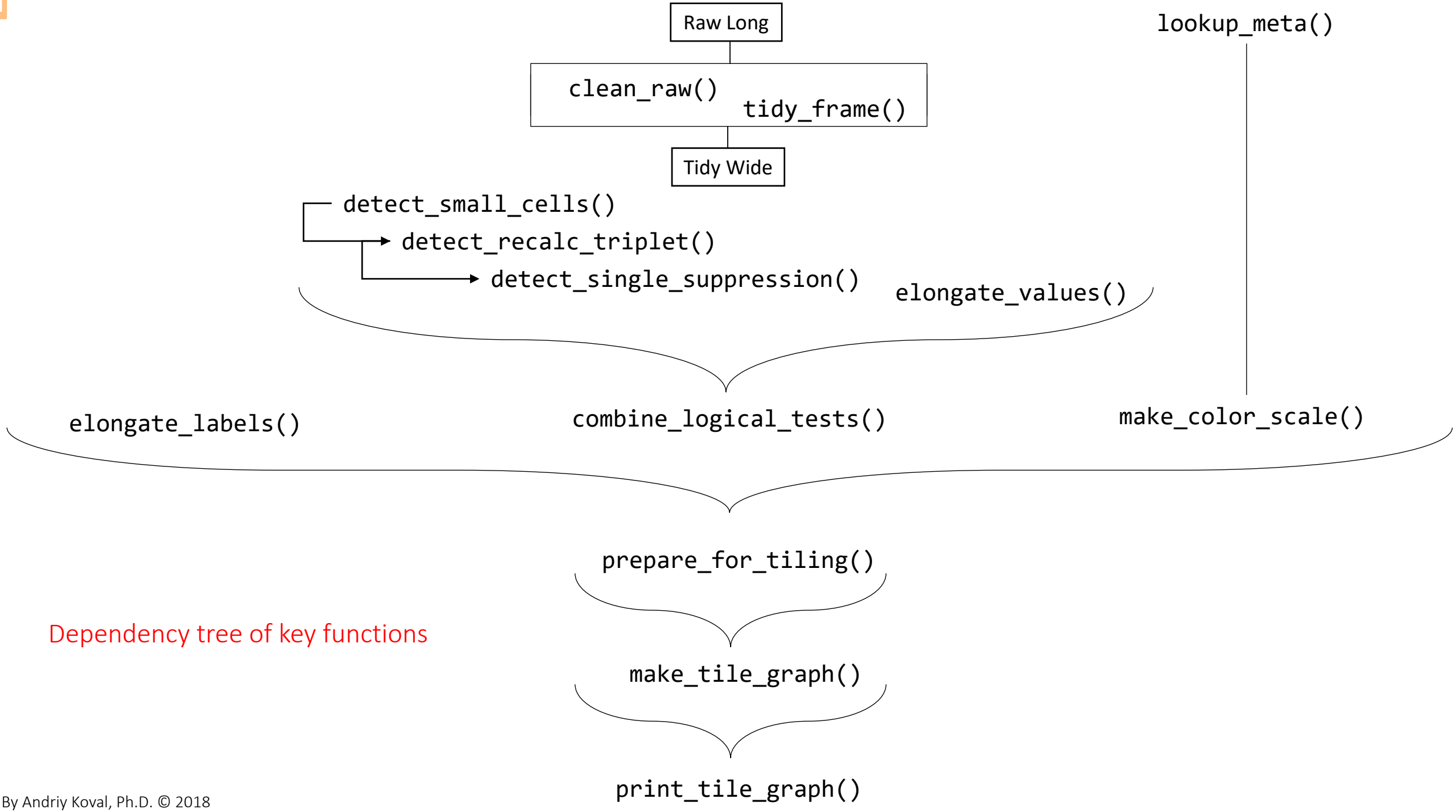


(vignette)

```r
# TEST 1: What cells are `too small` ( < 5)
# Censor 1: What cells should be suppressed as "too small"?
d1_small_cell <- df %>% detect_small_cell()
# creates a replica of the data, with count values are replaced by TRUE/FALSE according to test

# TEST 2: What cells can help calculate the suppressed cells from the same triple?
# because we need to remove them, otherwise they make recalculation possible.
# Censor 2: What triples should be suppressed? (eg. F-M-T)
# reverse calculate from:
d2_recalc_from_triplet <- df %>% detect_recalc_triplet()

# TEST 3: Is this is the only triplet that is being suppressed in a higher order block?
# because if yes, recalculation is possible
# Censor 3: What cells should be suppressed as those that could be calculated from higher order count?
d3_single_suppression <- df %>% detect_single_suppression()

# ---- service-functions ------------------------

# function to elongate the VALUE (count) in the smallest decision frame
d_long_values <-  df %>% elongate_values()

# function to elongate the LABEL (name) in the smallest decision frame
d_long_labels <- df %>% elongate_labels(c("label_prov", "label_ha","label_hsda"))

# create color scale to highlight suppression decisions
d_colors <- bc_health_map %>% make_color_scale()

# apply sequential logical tests to suppress desired cells
d_combined_tests <- df %>% combine_logical_tests()

# ---- graphing-functions ------------------------

# prepare the context for suppression = smallest decision frame
# create a list object containing required data in required shape to generate graphs
l <- df %>% prepare_for_tiling(bc_health_map)

# generate a graph of a single logical test
df %>% make_tile_graph(bc_health_map)

# it is very useful to segregate how
# (1) a plot is assembled with graphing script from how
# (2) a plot is committed to a hard digital form (PNG, JPG, PDF)
# can help us avoid going insane from trying to make it look right/useful on paper/screen
# there are many decision about the appearance of the plot that needs to be scripted
df %>% print_tile_graph(bc_health_map, path_folder = "./sandbox/examiner-2/prints/", size = 3)

# so far, df referred to a single Data Frame = a context for a single suppression decision
# we can use a wrapper function to loop through  a large number of frames
```



(vignette)



Dependency tree of key functions

British Columbia BSERVATORY Population & Public Health | UBC THE UNIVERSITY OF BRITISH COLUMBIA

Conclusions

- Use case for reproducible workflows
- Transparent and Extendable
- Workflow maps + Dependency trees
- Approachable = Learning community
- Open-source = democratic + global collaboration

https://github.com/IHACRU/suppress-for-release

INTERNATIONAL
Population Data Linkage
NETWORK
November 2020

USING REPRODUCIBLE DATA VISUALIZATIONS TO AUGMENT DECISION-MAKING DURING SUPPRESSION OF SMALL COUNTS

Andriy Koval
Assistant Professor
UNIVERSITY OF CENTRAL FLORIDA

Kate Smolina
Director
British Columbia BSERVATORY Population & Public Health
UBC

Anthony Leamon
Regional Epidemiologist
vancouver island health authority

Thank you!

andriy.koval@ucf.edu
http://andriy.rbind.io
@andkovpro