AboriginalPeoplesSurvey

Timothy Leung

2023-01-11

Contents

1 Abstract

To gain a cursory understanding of the current landscape of healthcare accessibility for Indigenous folks in Canada, I did an environmental scan with Statistics Canada's open-source data. With my exploratory analysis, I intend to lay a foundation for future supervised classification studies about factors that supersede the populations' inaccessibility of equitable healthcare. Yet, limited by the format of the questionnaires, the analysis can be no more than a preliminary one. The correlates of the variable of interest- "needed (healthcare) but not received" (expressed need; Statistics Canada, 2014, p.111; 2020, p.765), appeared sparsed and semantically trivial. Indeed, a multiple-factor analysis (MFA) indicated that the expressed need and its correlates failed to drive the dominant patterns of the datasets. I infer that the vast amount of labels for non-responsiveness in the surveys yielded auxiliary levels in the measurements, resulting in near-meaningless observations. Therefore, I suggest we rework the data collection.

2 Introduction

Statistics Canada conducts the Indigenous Peoples Surveys- formerly named the "Aboriginal Peoples Survey" (APS; Statistics Canada, 2022b, para.3) until the most recent iteration, by surveying First Nations people living off reserve, Métis and Inuit living in Canada. Six cycles have been performed in 1991, 2001, 2006, 2012, 2017, and 2022, with the latest to be concluded by March 2023 (Statistics Canada, 2009a, b, 2011, 2014, 2020, 2022a). The national surveys composite of topics and questions changed over time according to the researchers' focus. However, only the cycles since 2012 specifically studied the lack of accessibility to healthcare. Thus, as our primary interest is the expressed need for healthcare for Indigenous folks, only the 2012 and 2017 surveys are relevant.

3 Method

The 2017 APS is 20849 by 355, and the 2012 APS is 24803 by 326. All items of the APSs were nominal. Especially given the presence of nonresponsive levels, I cannot derive meaning from the orders of any item's labels. For that reason, I cannot utilize commonly-used methods in analyzing relations.

3.1 Data screening

Across the 2012 and 2017 APSs, there are multiple occurrences of item- and unit-nonresponses. The design of the measurement allowed for participants- units, to give any of the following nonresponses for a survey question-item: "Don't know," "Refusal," and "Not stated" for the 2012 APS denoted as (7, 8, 9) or (97, 98, 99) and "Valid skip," "Don't know," "Refusal," and "Not stated" for the 2017 APS marked as (6, 7, 8, 9) or (96, 97, 98, 99) (Statistics Canada, 2014, pp.33-968; 2020, pp.9-227). As demonstrated in Figure 1, only 30% of the 2012 and 2017 APS items had no nonresponse. Particularly in the 2017 APS, a majority of the items had at least 43% being unusable, while a majority of the participants did not give meaningful answers to at least 43% of the questions. The prominence of the item- and unit-nonresponses make imputation impossible.

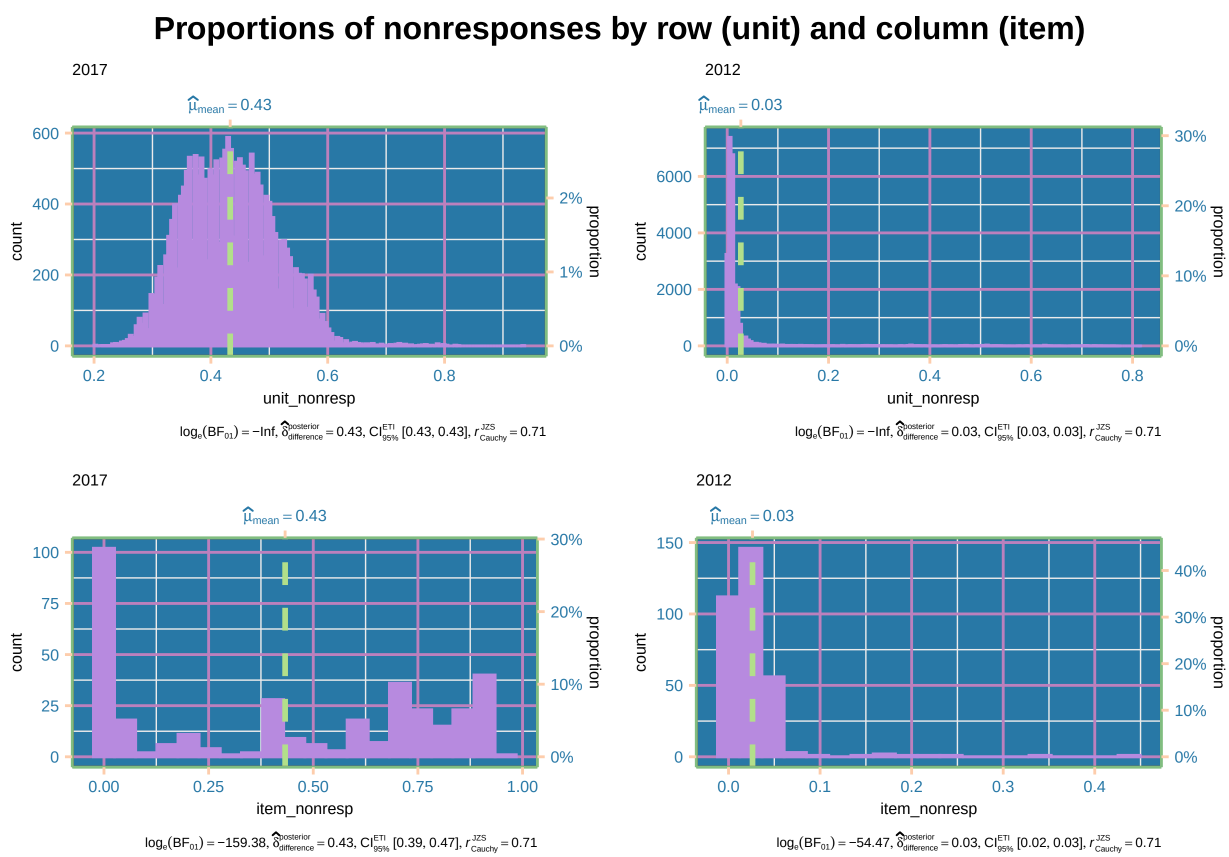

Figure 1: nonresponses

3.2 Procedure

3.2.1 Cramer's V with bias-correction

I will use Cramer's V in place of correlation as all variables in all data sets are nominal.

$$\begin{cases} \hat{\phi}^2 = \sum_{i\in[1,r]} \sum_{j\in[1,c]} \frac{1}{p_{i+}\cdot p_{+j}} \cdot \left(p_{ij} - p_{i+}\cdot p_{+j}\right)^2, \\ \tilde{r} = r - \frac{(r-1)^2}{n-1}, \\ \tilde{c} = c - \frac{(c-1)^2}{n-1}, \\ \tilde{\phi}^2_+ = \max\left(0, \hat{\phi}^2 - \frac{(r-1)\cdot(c-1)}{n-1}\right), \\ \tilde{V} = \left(\frac{\tilde{\phi}^2_+}{\min(\tilde{r},\tilde{c})-1}\right)^{\frac{1}{2}} \end{cases} \qquad \text{(Bergsma, 2013, pp.2,3).}$$

As a rule of thumb, as long as $\tilde{V} \geq 0.4$, we can consider the association between the two respective nominal variables to be at least moderate (Lee, 2016, p.560).

Here, I will use the statistics as a proxy of feature importance related to our variable of interest, expressed need, so that I can perform factor analyses with a focus on only the variables that reasonably correlate to it.

3.2.2 Multiple-factor analysis (MFA)

Here, I aim to reduce the dimensionality of the datasets and depict factors contributing to the dominant patterns. If some of the correlates of the expressed need are among the elements, it may be justifiable to use them for further classification analysis. Limited by the fact that all variables are nominal, I cannot use Principal Component Analysis (PCA) for the variables' lack of variance structures (Finnstats, 2022). Instead, I will apply Multiple-factor analysis (MFA), in which I gather the variables by groups, compute some variation of PCA, give weights to each PCA according to their inertia and augment them, analyzing the global structures (Abdi & Valentin, 2007). In the context of the APSs, given that Statistics Canada already grouped the items by topics, I naturally followed their format.

Guiding the decision about the number of dimensions to attend to, there is no method better than an arbitrary yet educated guess, namely the Elbow method. Ideally, we accept the minimal number of dimensions that cumulatively explain most of the variance and mark a sharp change in the decrease of variance per increased dimension. The cumulative percentage of explained variance should be at least 50% to be satisfactory (Hair et al., 2009, p.108).

Table 1: display of the top 25 stat of 2017(left) and 2012(right)

	Cramers V		Cramers V
GH2_30	1.0000	GH2_06	1.0000
GH2_35E	0.7070	GH2_07NA	0.7068
GH2_35NA	0.7070	GH2_07PR	0.7068
GH2_35PR	0.7070	GH2_07E	0.7068
GH2_35OT	0.7070	GH2_07OT	0.7068
GH2_40A	0.7070	GH2_08A	0.7059
GH2_40C	0.7070	GH2_08C	0.7059
GH2_40D	0.7070	GH2_08D	0.7059
GH2_40F	0.7070	GH2_08F	0.7059
GH2_40OT	0.7070	GH2_08OT	0.7059
DVISN_DF	0.6178	DCONSULT	0.6592
GH2_20P	0.6085	GH2_03A	0.6129
GH2_15C	0.6080	GH2_03C	0.6124
GH2_15B	0.6046	GH2_03B	0.6112
DUNK_DF	0.5902	GH2_03D	0.6012
DFLXB_DF	0.5821	GH2_04G	0.5861
GH2_15A	0.5814	GH2_01	0.5644
DHRNG_DF	0.5811	CC2_01	0.5493
DMOB_DF	0.5810	GH1_01	0.5491
DDEXT_DF	0.5785	INJ_01	0.5480
DVISN_FL	0.5670	CC2_02	0.5466
DMOB_FL	0.5665	CC2_07	0.5459
GH2_05	0.5627	CC2_04	0.5452
DFLXB_FL	0.5546	CC2_03A	0.5427
DUNK_FL	0.5527	CC2_03B	0.5419

4 Analysis

4.1 Cramer's V

The responses to the expressed need item- “GH2-30” in 2017 and “GH2_06” in 2012 strongly relate to only the respondents’ disability, health issues, recent or chronic, and interactions with the healthcare system. Perhaps, somewhat indicative is that those who expressed healthcare inaccessibility also seemed to suggest that the healthcare system had too high of a cost and too long of wait times. The variables on cost were “GH2_35E” in 2017 and “GH2_07E” in 2012, and those on wait time were “GH2_35NA” in 2017 (see Table 1).I would argue that none of the variables related to the expressed need is semantically insightful. People who personally experienced the flaws in the healthcare system in their environment and had pressing demands for medical support would naturally align with the advocacy for better accessibility.

4.2 MFA

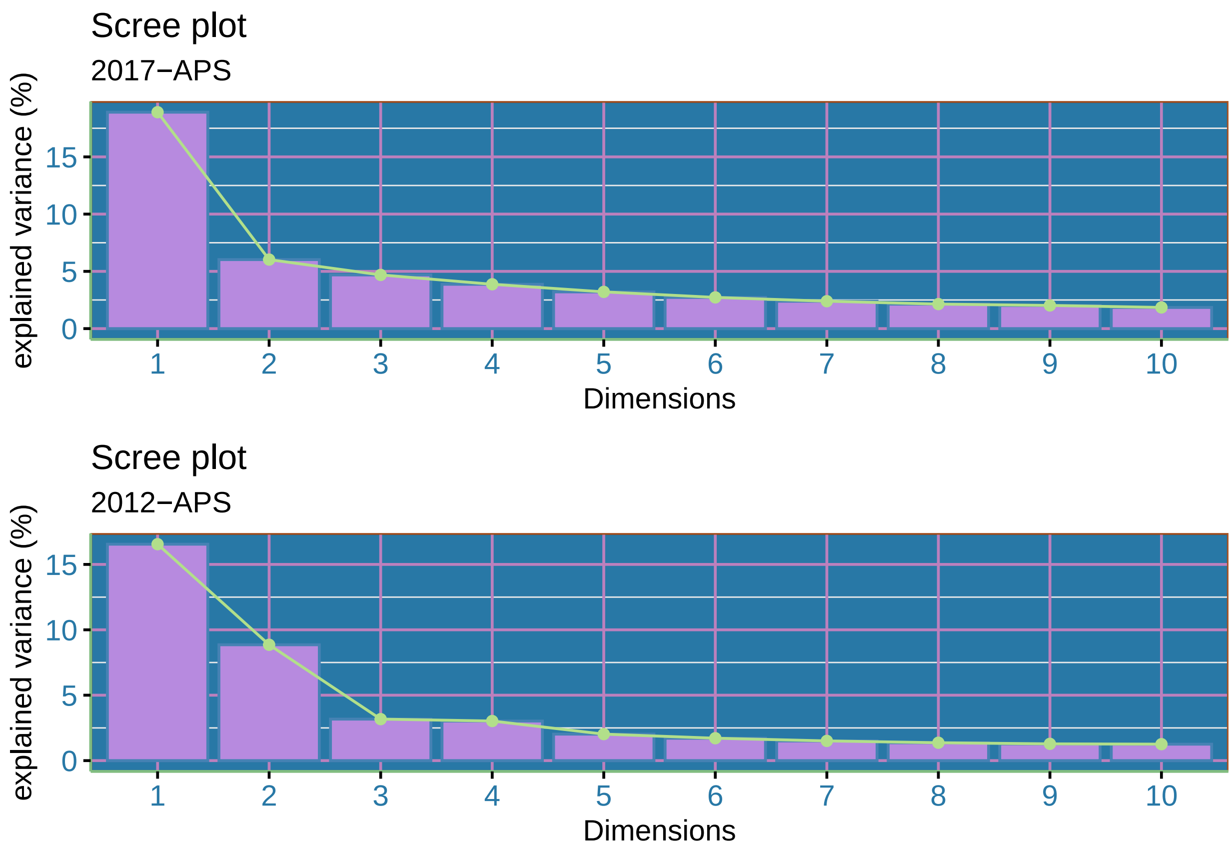

Figure 2: elbow method

According to the elbow method (see Figure 2), the rate of decrease in the percentage of explained variance per dimension increased changed the most drastically at (Dim.3, Dim.4) for the year (2017, 2012) respectively, so it would stand to reason that we should consider the same number of dimensions in the following Multiple Factor Analysis.

Readers should note that despite having chosen a relatively cumbersome number of dimensions, neither of the MFAs covered a good portion of the variances (see Table 2).

Upon dissecting the MFA findings, I can draw some themes by group, but I am uncomfortable assigning them individual variables or relating any of them to the expressed need variable. In the 2017 APS (see Figure 3), the top five groups contributing to the first dimension were all disclosure of disabilities, namely “DUNK,” “DMOB,” “DDEXT,” “DFLXB,” and “DVISN.” These are Disability indicators, including unknown conditions, mobility, dexterity, flexibility and optical issues (Statistics Canada, 2020, p.6). The top five groups contributing to the next dimension mainly were on education attainment, namely “PSS,” “CPSA,” “EHS,” “SMK,” and “EDNT.” Those are “[having] obtained education above high school,” “currently attending post-secondary [education],” “[having] completed [a] high school diploma,” smoking habits and having expressed desire of but not having obtained an education level. The third dimension’s top five contributors returned to the theme of disclosed disabilities, such as “DFLXB,” “DMOB,” “DPAIN,” “DDEXT,” and “DCOGN,” most of which were elaborated above except for “DCOGN-” cognitive disabilities.

In the 2012 APS (see Figure 6), the first theme appeared to be a mixture of social capital and health indicators: “CC2,” “SMK,” “GH1,” “GH2,” and “CS-” chronic health conditions, smoking habits, general self-perceived health and accessibility of healthcare and support network. The second theme of the 2012 APS remained somewhat scattered. They could be about participants’ employment and connection with their heritage: “DEMPSTAT,” “LMAM,” “LM,” ”DTRACTYR,” and “DTRACTDO-” employment status, conditions, mobility, and habits and interest in engaging with traditional ceremonies. The third theme appeared clearly as the social determinants of health, namely “FS,”

Table 2: 2017(top); 2012(bottom)-elbow

	eigenvalue	variance.percent	cumulative.variance.percent
Dim.1	13.172540	18.901426	18.90143
Dim.2	4.200629	6.027530	24.92896
Dim.3	3.266873	4.687673	29.61663
Dim.4	2.698035	3.871441	33.48807
Dim.5	2.235497	3.207740	36.69581
Dim.6	1.898916	2.724776	39.42059
Dim.7	1.667491	2.392701	41.81329
Dim.8	1.491914	2.140764	43.95405
Dim.9	1.409274	2.022184	45.97624
Dim.10	1.286747	1.846367	47.82260
Dim.11	1.174313	1.685035	49.50764
	eigenvalue	variance.percent	cumulative.variance.percent
Dim.1	20.057112	16.5461235	16.54612
Dim.2	10.735838	8.8565344	25.40266
Dim.3	3.844689	3.1716783	28.57434
Dim.4	3.668969	3.0267179	31.60105
Dim.5	2.453540	2.0240488	33.62510
Dim.6	2.074799	1.7116061	35.33671
Dim.7	1.824396	1.5050364	36.84175
Dim.8	1.655468	1.3656788	38.20742
Dim.9	1.552167	1.2804610	39.48789
Dim.10	1.523197	1.2565617	40.74445
Dim.11	1.476636	1.2181519	41.96260
Dim.12	1.384902	1.1424751	43.10507
Dim.13	1.272820	1.0500132	44.15509
Dim.14	1.262822	1.0417656	45.19685
Dim.15	1.174968	0.9692901	46.16614
Dim.16	1.155609	0.9533199	47.11946
Dim.17	1.126359	0.9291906	48.04865
Dim.18	1.099157	0.9067497	48.95540
Dim.19	1.078575	0.8897709	49.84517

“CS,” “LMAM,” “HOU,” and “MH-” food security, social support network, employment status, dwelling condition and mental health. The fourth and last dimension we chose to examine did not appear informative. All groups are duplicates from the previous except for “DMJH” and “DFTPT-” whether or not a respondent held multiple jobs and worked full-time.

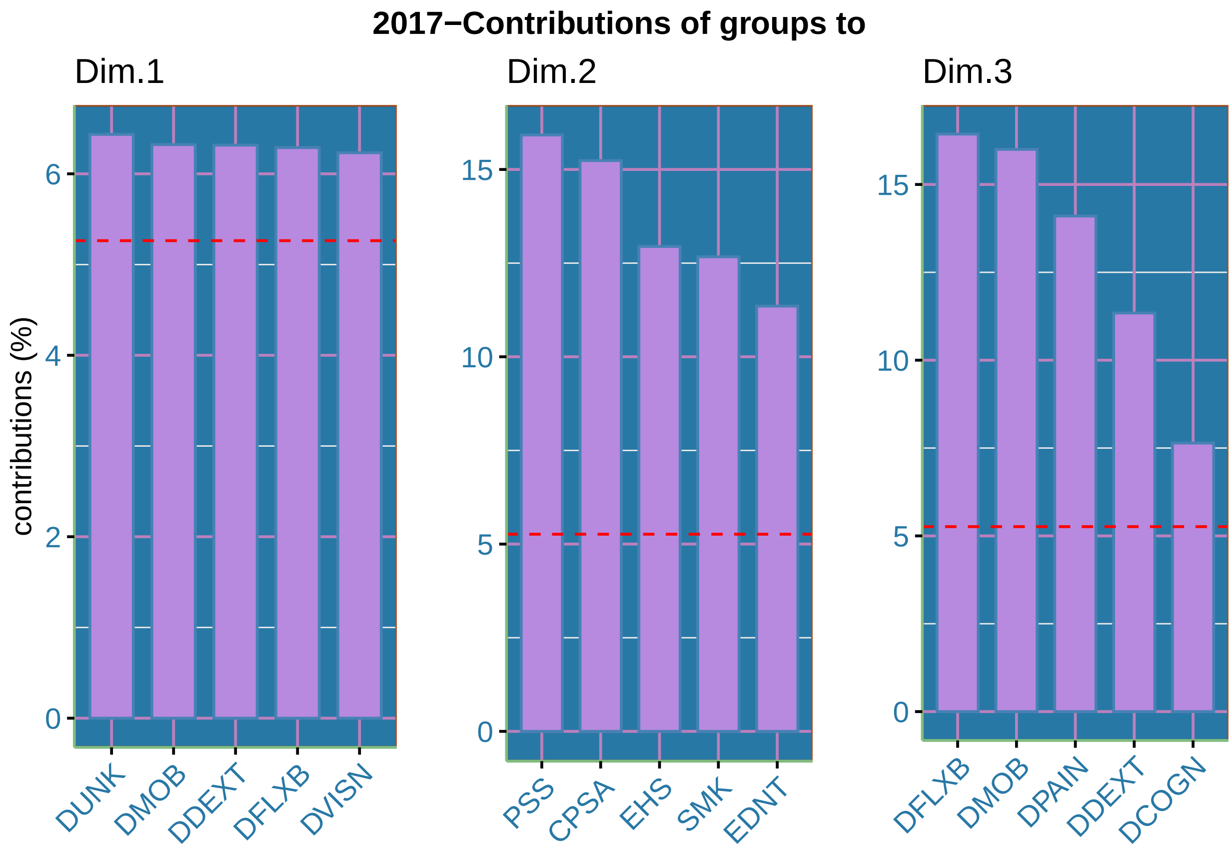

Figure 3: 2017-MFA-group

The patterns arising from individual (qualitative) variables were downright absurd. Most of the variables linked to the 2012 and 2017 APSs’ dimensions of interest were on the levels of nonresponse (8= ”Refusal” and 9= ”Not stated”; see Figures 4 and 7). Although we could make a case that in the 2017 APSs’ Dim2 vs. Dim3 biplot-Figure 5, the expressed need’s meaningful levels- “Yes” and “No” significantly contributed to the variability of the third dimension, by which I established to be the disclosed disabilities theme, I caution readers that the explained variance of Dim3 was extremely low (4.7%). We could also notice that Dim3 and Dim4 of 2012 APS had a high positive correlation due to the near-45-degree slope (see Figure 8), but for the same reason, plus the fact that the variability linked to the “Refusal” nonresponse was the most dominant, the observation might not be sound. Otherwise, the biplots from the 2012 and 2017 APSs indicated that the examined dimensions captured mostly only the variabilities of the expressed need’s nonresponse levels.

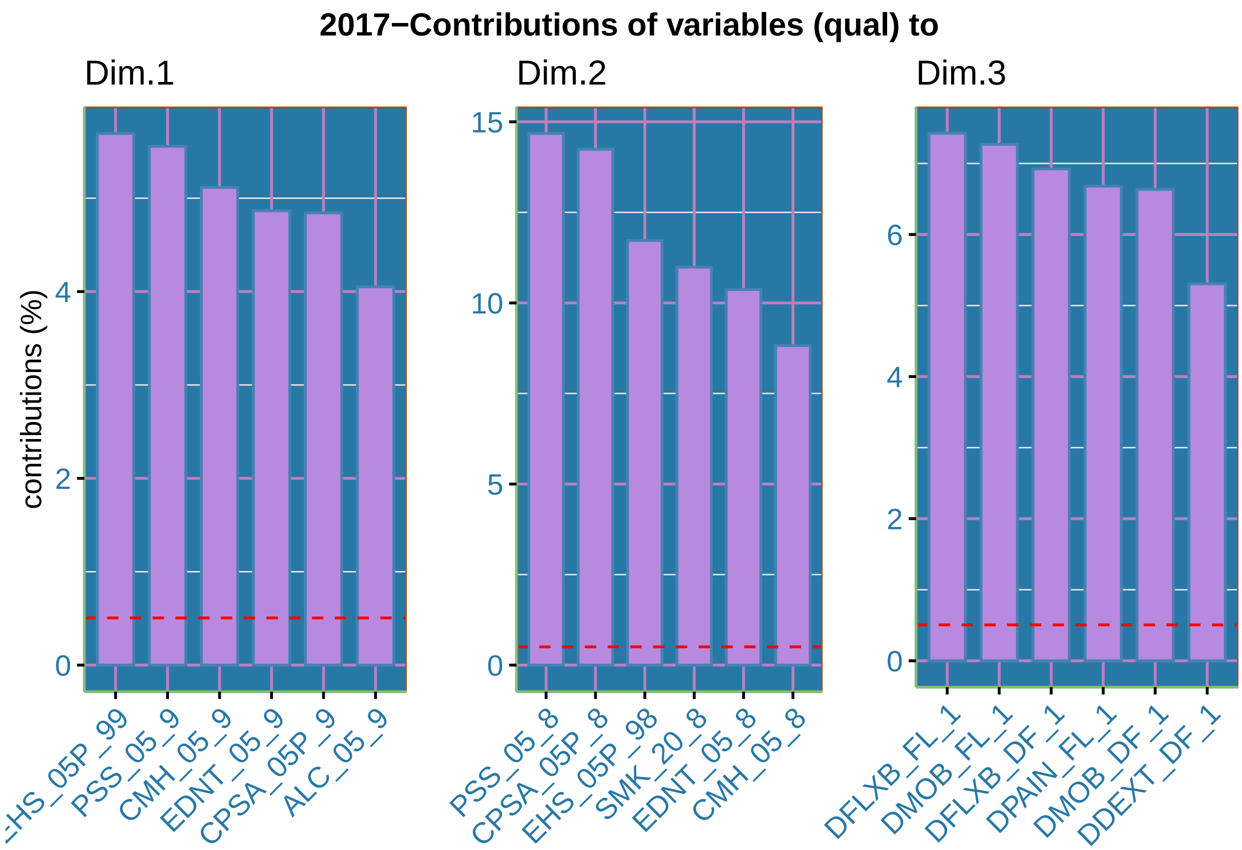

Figure 4: 2017-MFA-ind vari

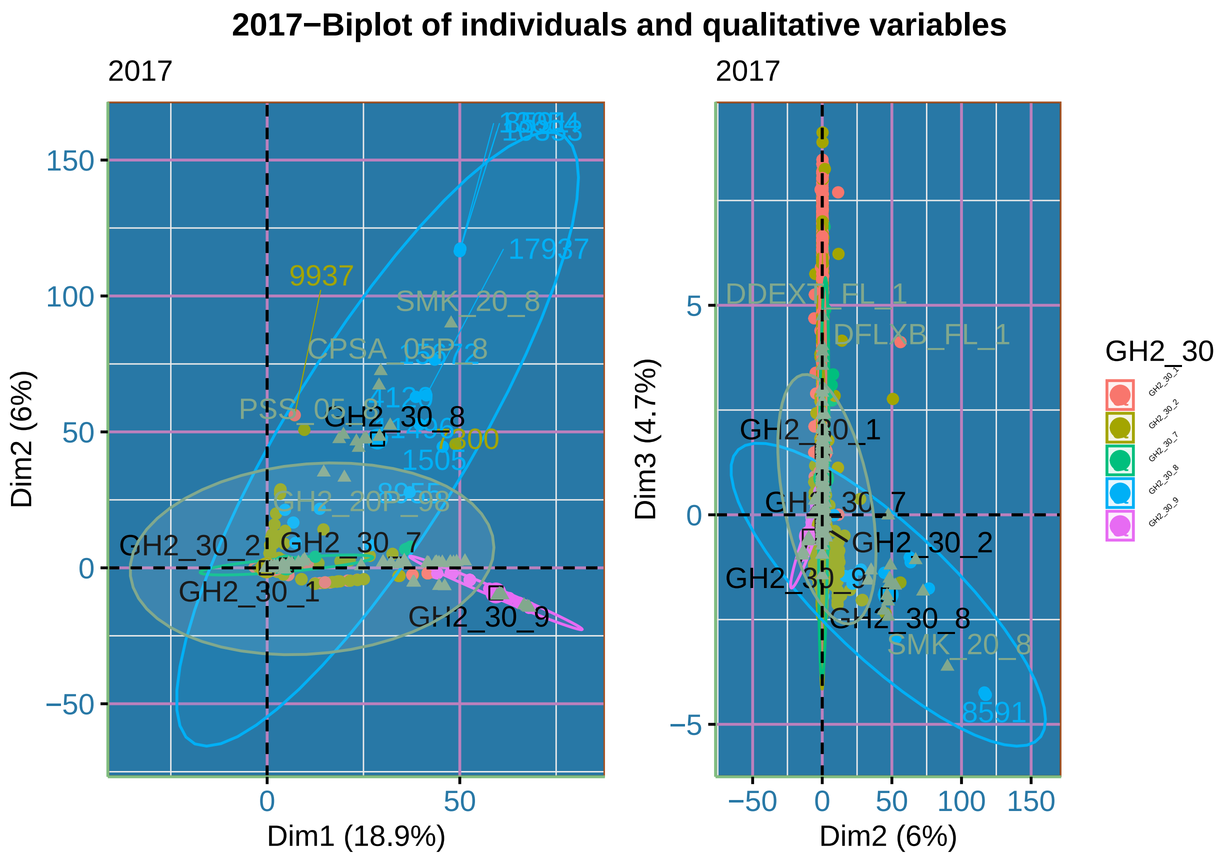

Figure 5: 2017-MFA-expressed need on dim

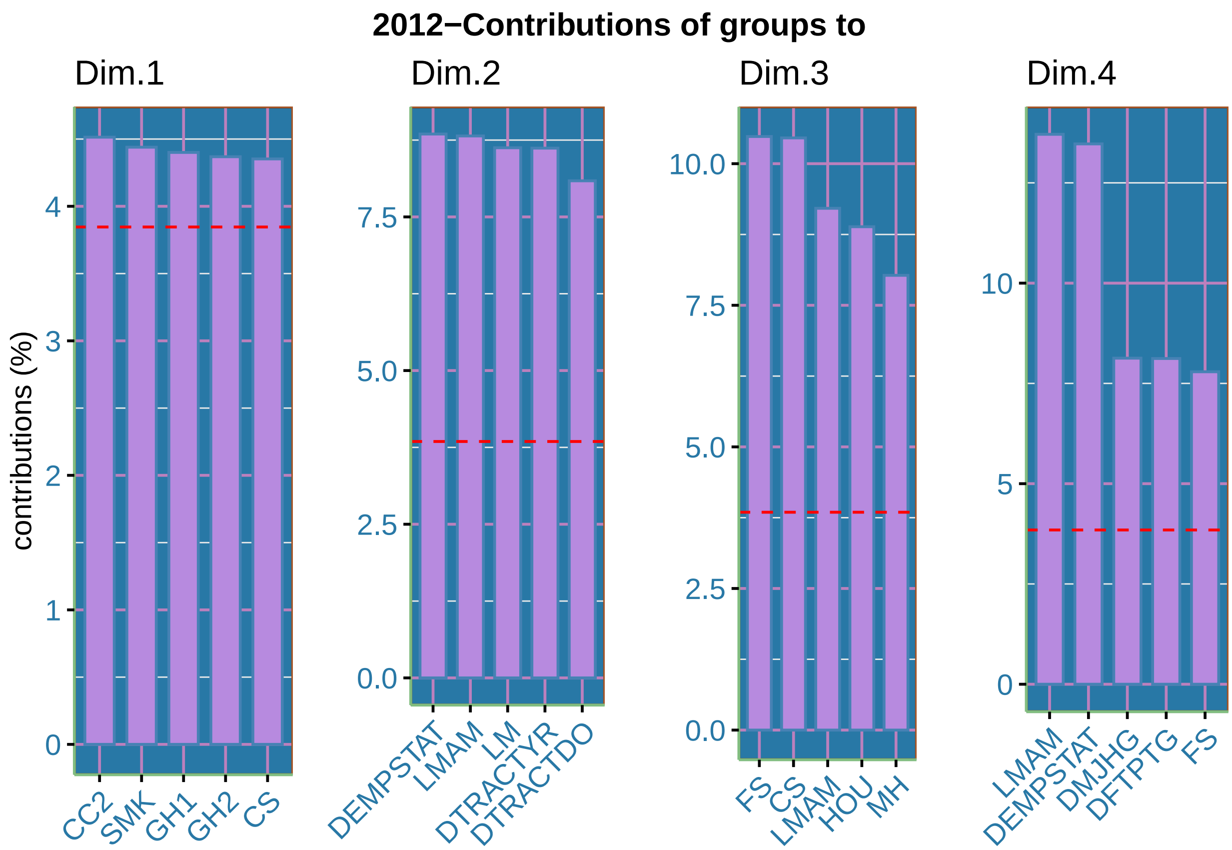

Figure 6: 2012-MFA-group

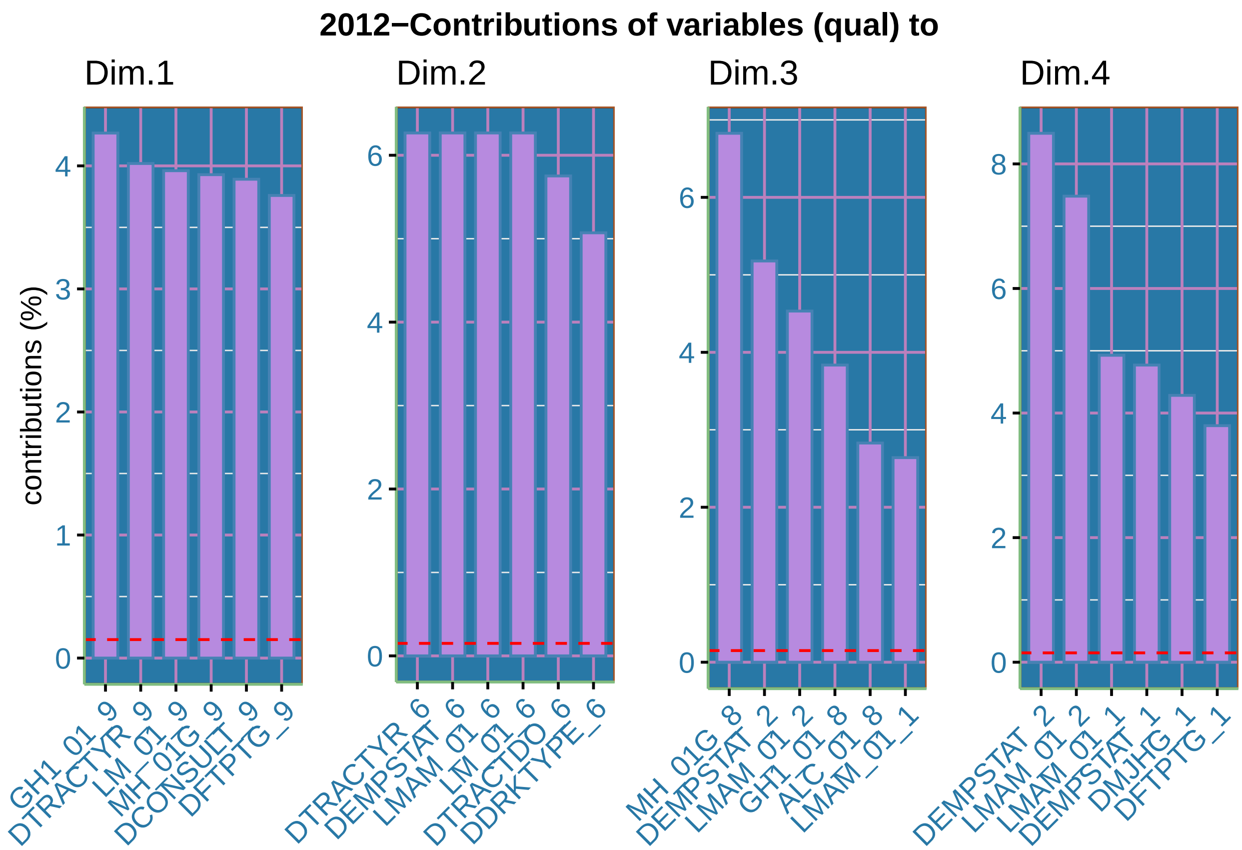

Figure 7: 2012-MFA-ind vari

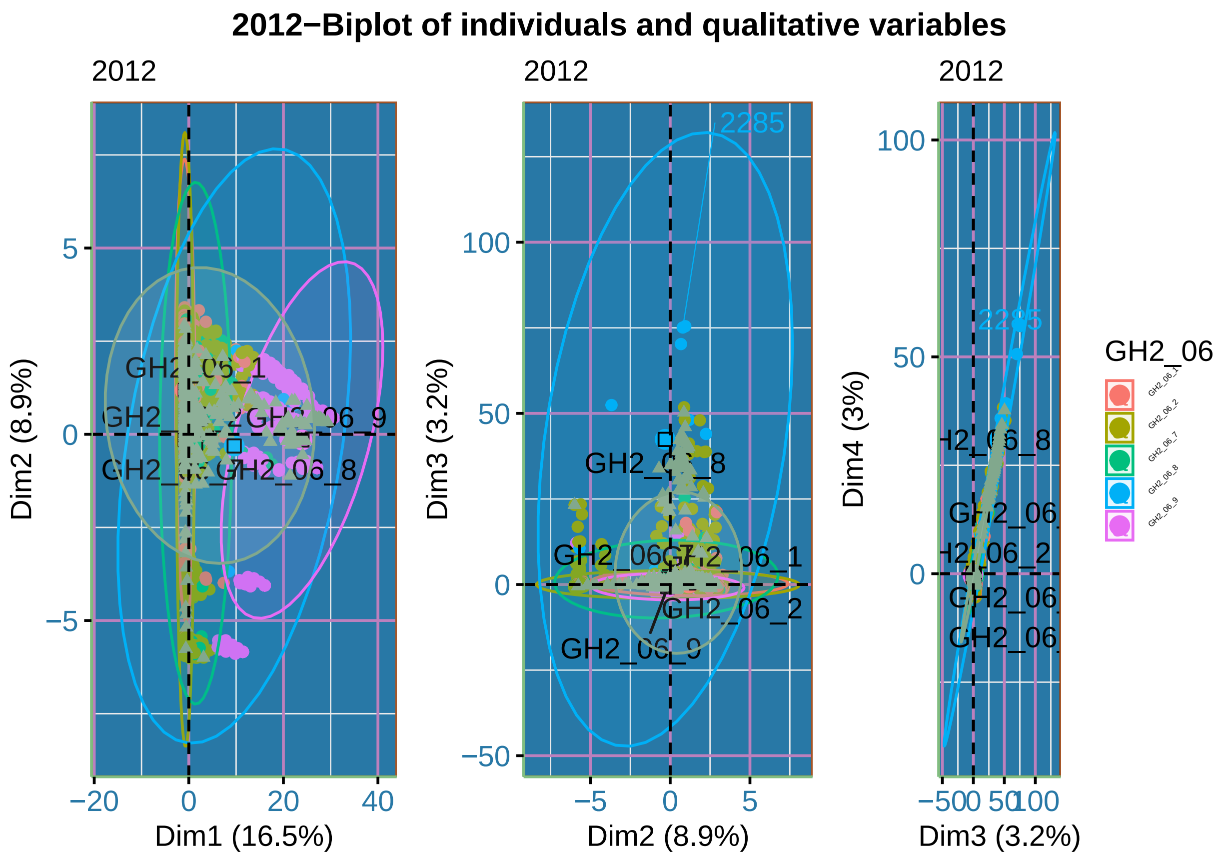

Figure 8: 2012-MFA-expressed need on dim

5 Conclusion

In summary, I caution against drawing hasty conclusions about the Indigenous populations' expressed needs for more accessible healthcare out of the data at hand. We might find evidence that those who displayed the expressed needs might also have experienced flaws in the healthcare system serving them, namely being too costly or tardy. We might see hints of correlations among themes, namely social determinants of health and employment conditions (Dim3 vs Dim4 in 2012 APS) or explanatory relations between expressed needs for healthcare accessibility and the topic of disclosed disabilities (GH2_30 to Dim3 in 2017 APS). However, solely based on the 2012 and 2017 APS datasets, we must examine such findings incredulously as the supporting facts that empower the observations are non-existent.

The datasets might have correctly piqued our interest in the questions about who may be allies of improving Indigenous healthcare accessibility, but they certainly have not answered how the Indigenous populations want the changes to bring about or by whom. In a more focused study, we shall hone in on folks who have already expressed the need and tune into their perspectives from their priorities.

I remark that we genuinely cannot improve the data quality from the given datasets. It is not uncommon for data analysts to arbitrarily collapse some levels across all items in a survey after the data collection is finalized. Even with justifications, it is a controversial method in the statistics field because the action is an offence to methodological conventions, for it undermines the interval-scaled properties of latent variables regardless of the observed variables' scales and makes the items' distributional properties unreliable (Grimbeek et al., 2005, p.3), not to mention that it bounds to lose information from the responses (DiStefano et al., 2021, p.11). Proponents of the method believe it facilitates model-based analysis (DiStefano et al., 2021, pp.2,11) and reduces redundancy (Van Dusen & Nissen, 2022, p.2), improving the analysis' intelligibility (Grimbeek et al., 2005, p.2). In particular, Van Dusen and Nissen (2022) justified their choice of survey-level combination using point-biserial correlation coefficients under the assumption that their Likert-scaled data had underlying continuous distributions and the condition that all items had the same number of levels. In our situation, neither our data was continuous- it was nominal, nor did the questions all share the same number of points- they ranged from 2 to about 20,000. In the Appendix, I will demonstrate how we would have justified the collapse of survey levels if the situation allowed.

References

Abdi, H., & Valentin, D. (2007). Multiple factor analysis (MFA). *Encyclopedia of measurement and statistics*, 657-663.

Attali, Y., & Fraenkel, T. (2000). The point-biserial as a discrimination index for distractors in multiple-choice items: Deficiencies in usage and an alternative. *Journal of Educational Measurement, 37*(1), 77-86.

Bergsma, W. (2013). A bias-correction for Cramér's V and Tschuprow's T. *Journal of the Korean Statistical Society, 42*(3), 323-328.

DiStefano, C., Shi, D., & Morgan, G. B. (2021). Collapsing categories is often more advantageous than modeling sparse data: investigations in the CFA framework. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(2), 237-249.

Finnstats. (2022, November 20). *PCA for Categorical Variables in R. R-bloggers*. https://www.r-bloggers.com/2022/11/pca-for-categorical-variables-in-r/

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Allyn & Bacon.

Grimbeek, P., Bryer, F., Beamish, W., & D'Netto, M. (2005, January). Use of data collapsing strategies to identify latent variables in CHP questionnaire data. In *Proceedings of 3rd International Conference on Cognition, Language, and Special Education. Brisbane: Griffith Research Online* (pp. 125-39).

Hair, J., Anderson, R., & Babin, B. (2009). *Multivariate Data Analysis*. Prentice Hall.

Lee, D. K. (2016). Alternatives to P value: confidence interval and effect size. *Korean journal of anesthesiology, 69*(6), 555-562.

Pagès, J. (2014). *Multiple factor analysis by example using R*. CRC Press.

Statistics Canada. (2009a). *Aboriginal Peoples Survey, 1991*. Abacus Data Network. https://abacus.library.ubc.ca/dataset.xhtml?persistentId=hdl:11272.1/AB2/J8PQNR

Statistics Canada. (2009b). *Aboriginal Peoples Survey, 2001*. Abacus Data Network. https://abacus.library.ubc.ca/dataset.xhtml?persistentId=hdl:11272.1/AB2/OAZOFP

Statistics Canada. (2011). *Aboriginal Peoples Survey, 2006 [2009]*. Abacus Data Network. https://abacus.library.ubc.ca/dataset.xhtml?persistentId=hdl:11272.1/AB2/1LW8BC

Statistics Canada. (2014). *Aboriginal Peoples Survey, 2012*. Abacus Data Network. https://abacus.library.ubc.ca/dataset.xhtml?persistentId=hdl:11272.1/AB2/0T6DYF

Statistics Canada. (2020). *Aboriginal Peoples Survey, 2017*. Abacus Data Network. https://abacus.library.ubc.ca/dataset.xhtml?persistentId=hdl:11272.1/AB2/U0UXQJ

Statistics Canada. (2022a, November 25). *Indigenous Peoples Survey (IPS)*. https://www.statcan.gc.ca/en/survey/household/3250

Statistics Canada. (2022b, December 14). *Surveys about Indigenous Peoples*. https://www.sac-isc.gc.ca/eng/1321384019753/1605196725394

Van Dusen, B., & Nissen, J. M. (2020, January). Criteria for collapsing rating scale responses: A case study of the CLASS. *In 2019 Physics Education Research Conference Proceedings.*

6 Appendix

Novel Justification for rating-scale combination

Van Dusen and Nissen (2020) explained that we could collapse a set of survey levels if their distributions do not differ significantly from one to another. We back such a judgement by reading the means or density curves of the levels' point-biserial correlation. Instead of examing the levels' validity, a point-biserial correlation is usually calculated in item analyses (or item discriminations) in examining the items' validity- precisely, it is to correlate respondents' performance on an item (right or wrong) with their performance on the criterion (how many answers does one score as correct). The formula is:

$$r_{pb}({}_{[resp]}x,{}_{[level]}\phi) = \frac{\overline{\#{-}item_y{=}key_\phi\ per\ resp_x} - \overline{\#{-}item_y{\neq}key_\phi\ per\ resp_x}}{S(\#{-}item_y{=}key_\phi\ per\ resp_x)}$$

$$\cdot \sqrt{\frac{1}{\{\#{-}resp\}} \cdot \{\# - resp \mid item_y = key_\phi\} \cdot \{\# - resp \mid item_y \neq key_\phi\}}$$

(Attali & Fraenkel, 2000, p.77), where
$r_{pb}(x,\phi) \cdot \sqrt{\frac{\{\#-resp\}-2}{1-r_{pb}^2(x,\phi)}} \sim \mathsf{T}(\{\# - \text{resp}\} - 2)$ under $H_0 \colon r_{pb} = 0$ (Glass & Hopkins, 1996, pp.134, 364). The current, novel method re-appropriates the said correlation for the levels' distribution, thus expanding the construct's substance.

More fundamentally, the current method alters the structure of the point-biserial correlation's construct. Only Van Dusen's and Nissen's (2020) article explicitly applied the R package, "itemanalysis" for such a task. To investigate its methodology, I have summarized their R code as the following formulae:
$\underline{Let}$ there be X $\#-$respondents, Y $\#-$items, and Φ $\#-$possible options.
with ${}_{[data\ matrix]}D_{X\times Y}$, ${}_{[key\ vector]}\vec{K}_{1\times Y}$, ${}_{[options=possible\ levels]}\vec{\Phi}_{1\times\Phi}$,

$$\begin{cases} \vec{S}_{0,(1\times X)}(\phi,y) = \vec{t}_{y,\forall x\in X}(\vec{D}_y = \phi) & \ldots Perf_{\text{"}item\text{"}=option} = ID(resp \mid item_y\ match\ option_\phi)(\phi,y), \\ \mathbf{M}_{X\times Y} = \{\vec{t}_{x,(1\times Y)}(\vec{D}_y = K_y)\}_{\forall x\in X}, & \\ \tilde{S}_{t,(1\times X)}(y) = \frac{1}{Y-1}\cdot\sum_{y'\in Y}\mathbf{M}_{y'\setminus y} & \ldots Perf_{Criterion} = ID(\sum_x \# - items_y\ match\ key_y)(y), \\ \vec{S}_{t,(1\times X)} = \frac{1}{Y}\cdot\sum_{y'\in Y}\mathbf{M}_{y'}. & \end{cases}$$

$$\forall(\phi,y)\in\Phi\times Y,\ Cor_{item\ analysis}(\phi,y) = \begin{cases} cor\left(S_0(\phi,y),\tilde{S}_t(y)\right) & if\ correction, \\ cor\left(S_0(\phi,y),S_t\right) & o.w. \end{cases}$$

Obviously, $\forall c\in\mathbf{R},\ cor(A,c\cdot B) = \frac{Cov(A,c\cdot B)}{\sqrt{V(A)\cdot V(c\cdot B)}} = \frac{c\cdot Cov(A,B)}{\sqrt{c^2\cdot V(A)\cdot V(B)}} = sgn(c)\cdot Cor(A,B)$,

so $\forall(\phi,y)\in\Phi\times Y,\ Cor_{item\ analysis}(\phi,y) = \begin{cases} cor\left(S_0(\phi,y),\sum_{y'\in Y}\mathbf{M}_{y'\setminus y}\right) & if\ correction, \\ cor\left(S_0(\phi,y),\sum_{y'\in Y}\mathbf{M}_{y'}\right) & o.w. \end{cases}$

Despite having the method laid out, there misses one crucial piece of information: How do we derive a set of "correct" answer keys in a survey? Van Dusen and Nissen (2020) did not clarify what "key" in the Item Analysis with which we should compare each item's responses, so I came up with two methods to extract the "overall score" (Van Dusen & Nissen, 2020, p.4): (1) by assuming that the item-wise responses are continuous, take the means of responses for the columns and transform the "means" back into the format of the possible levels; (2) by treating the responses as nominal, only count the item-wise median responses as the key.

Take our situation for example- the distributions of the levels 1 to 5 are almost identical to one another in 2012 APS and 2017 APS (see Figures 9 and 10). Additionally, the levels' point-biserial correlation means are (-0.063, -0.048, 0.106, -0.016, -0.012, -0.18) in 2017 and (-0.063, -0.048, 0.106, -0.016, -0.012, -0.18) in 2012, justifying combining level 1 with 2 and level 4 with 5. For readers' references, the said correlation method assuming the responses as nominal produced means more or less the same: They are (-0.086, -0.067, 0.139, -0.013, -0.01, -0.172) in 2017 and (-0.086, -0.067, 0.139, -0.013, -0.01, -0.172) in 2012.

Here, Glass' and Hopkins' (1996) statistics would be $\forall(\phi,y)\in\Phi\times Y,\ r_{pb}(\phi,y)\cdot\sqrt{\frac{Y-2}{1-r_{pb}^2(\phi,y)}} \sim \mathsf{T}(Y-2)$. Markedly, the statistics are no longer independent because, unlike responses per item, the items per respondent came from the same entity. Instead of taking averages for the survey items and comparing those statistics to the means of T-distributed random variables (i.e., $\overline{r_{pb}(x,\phi)}_x \sim_{iid} \frac{1}{X}\cdot\sum_x \mathsf{T}(X-2)$), it is only reasonable to compare the distributions of each survey level to the upper and lower critical values of $\mathsf{T}(Y-2)$. In our example, only the point-biserial correlations of levels 6 and 3 are significant at the confidence level, $\alpha = .05$ (see Figures 11 and 12).

Table 3: (”average”)first 6 items’ point-biserial correlations of 2017(top) and 2012(bottom)

	1	2	3	4	5	6
AGE_YRSG	-0.072	-0.083	-0.069	-0.001	0.042	0.140
DSIZEHH	-0.005	0.069	-0.006	-0.004	-0.063	-0.029
LMAM_02	0.025	-0.135	0.126	-0.021	-0.017	-0.042
LMA2_04	-0.213	-0.011	0.136	-0.046	-0.012	-0.055
DJOBTGR	-0.038	-0.014	-0.003	0.121	-0.089	0.014
DNOCSKIL	0.029	0.059	0.052	0.003	-0.102	-0.037

	1	2	3	4	5	6
DSIZHHGG	0.085	0.161	0.018	-0.024	-0.147	-0.125
MOB_02A	0.125	0.218	-0.297	-0.015	-0.004	-0.033
MOB_02B	0.049	0.276	-0.284	-0.014	-0.003	-0.032
MOB_02C	0.214	0.154	-0.291	-0.014	-0.003	-0.032
MOB_02D	0.045	0.271	-0.286	-0.014	-0.003	-0.032
MOB_02HS	0.042	0.275	-0.285	-0.014	-0.003	-0.032

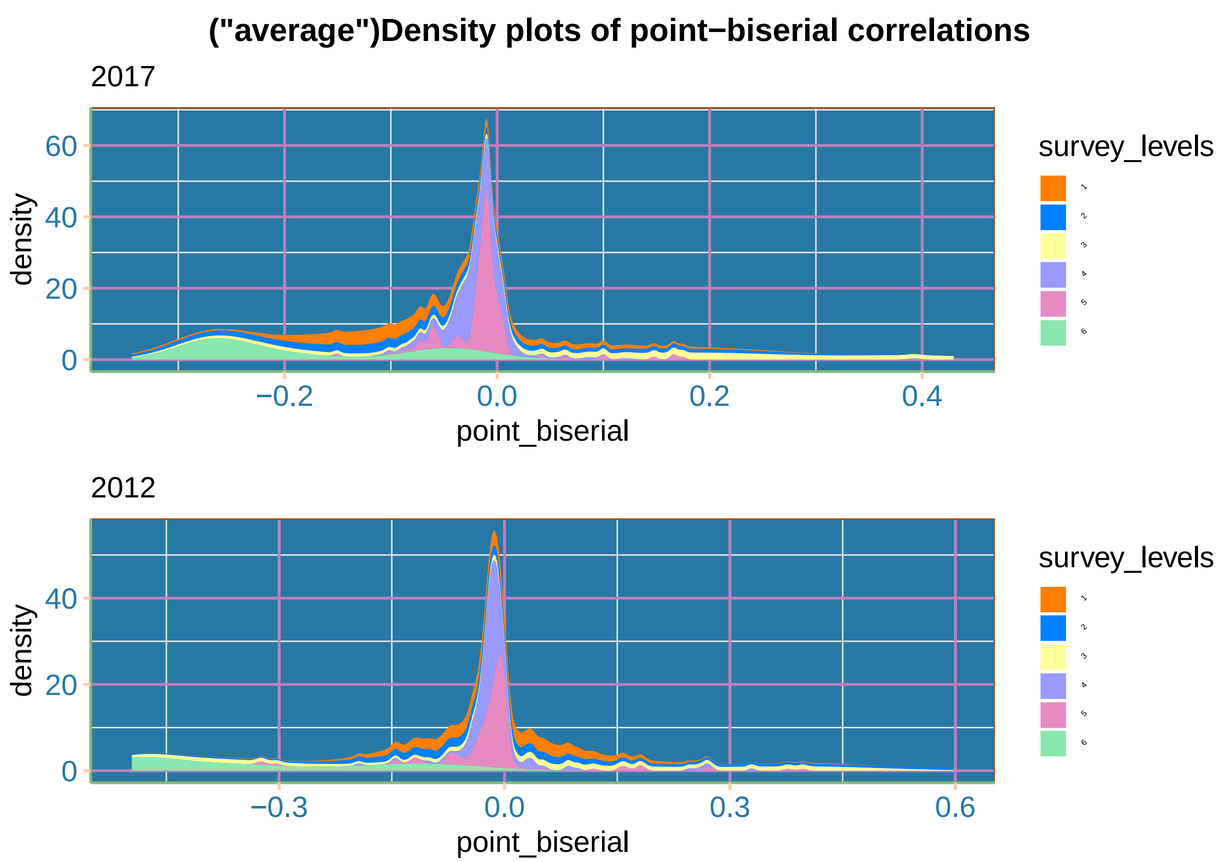

Figure 9: Demonstration-pb correlation for rating-scale combination

Table 4: (”median”)first 6 items’ point-biserial correlations of 2017(top) and 2012(bottom)

	1	2	3	4	5	6
AGE_YRSG	0.003	-0.005	-0.074	-0.055	0.003	0.094
DSIZEHH	-0.015	0.043	-0.002	0.003	-0.035	-0.024
LMAM_02	0.018	-0.093	0.087	-0.018	-0.019	-0.039
LMA2_04	-0.200	0.021	0.097	-0.039	-0.011	-0.051
DJOBTGR	0.006	0.013	0.015	0.082	-0.106	0.032
DNOCSKIL	-0.044	0.042	0.094	0.081	-0.117	-0.032

	1	2	3	4	5	6
DSIZHHGG	0.106	0.187	0.011	-0.026	-0.171	-0.134
MOB_02A	0.128	0.216	-0.298	-0.012	-0.004	-0.028
MOB_02B	0.051	0.276	-0.285	-0.011	-0.003	-0.028
MOB_02C	0.204	0.162	-0.292	-0.012	-0.004	-0.028
MOB_02D	0.031	0.278	-0.288	-0.012	-0.003	-0.028
MOB_02HS	0.047	0.274	-0.286	-0.011	-0.003	-0.028

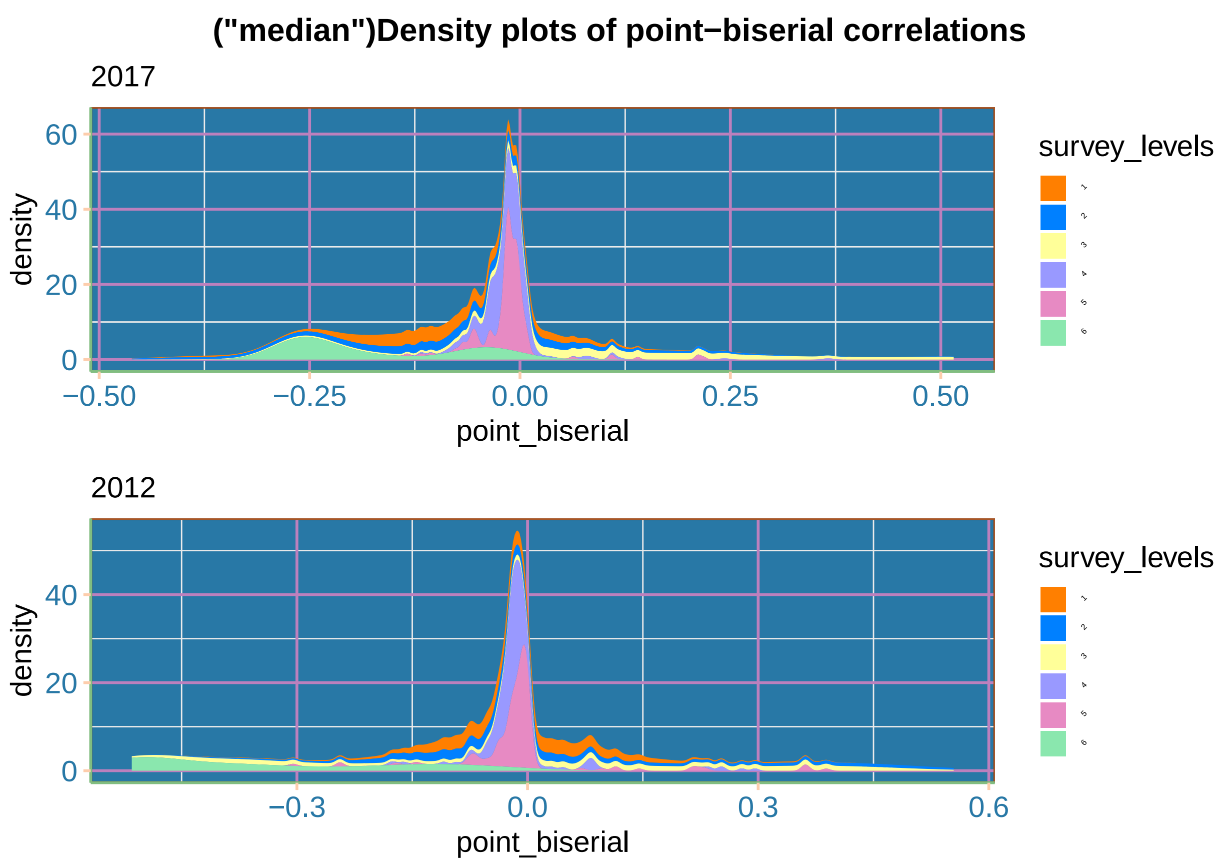

Figure 10: Demonstration-pb correlation for rating-scale combination

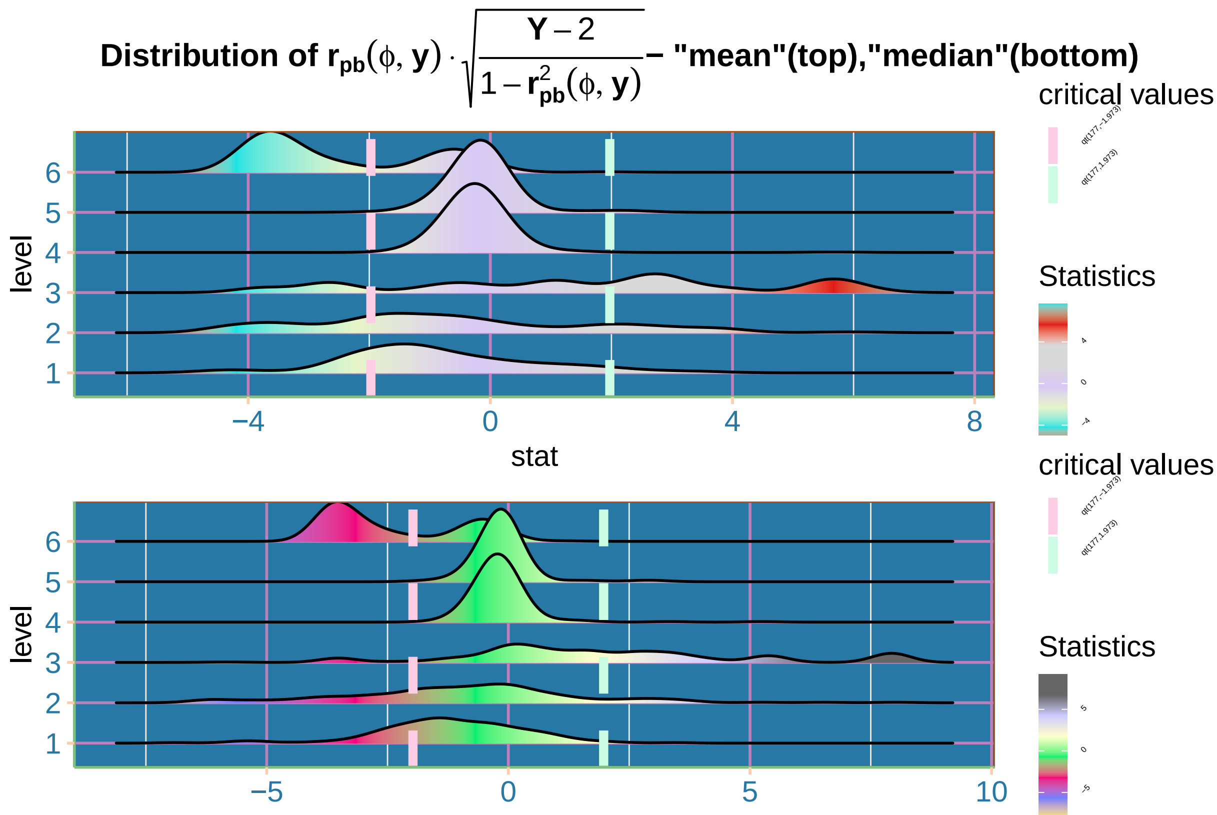

Figure 11: 2017-Stat(r_pb) distribution

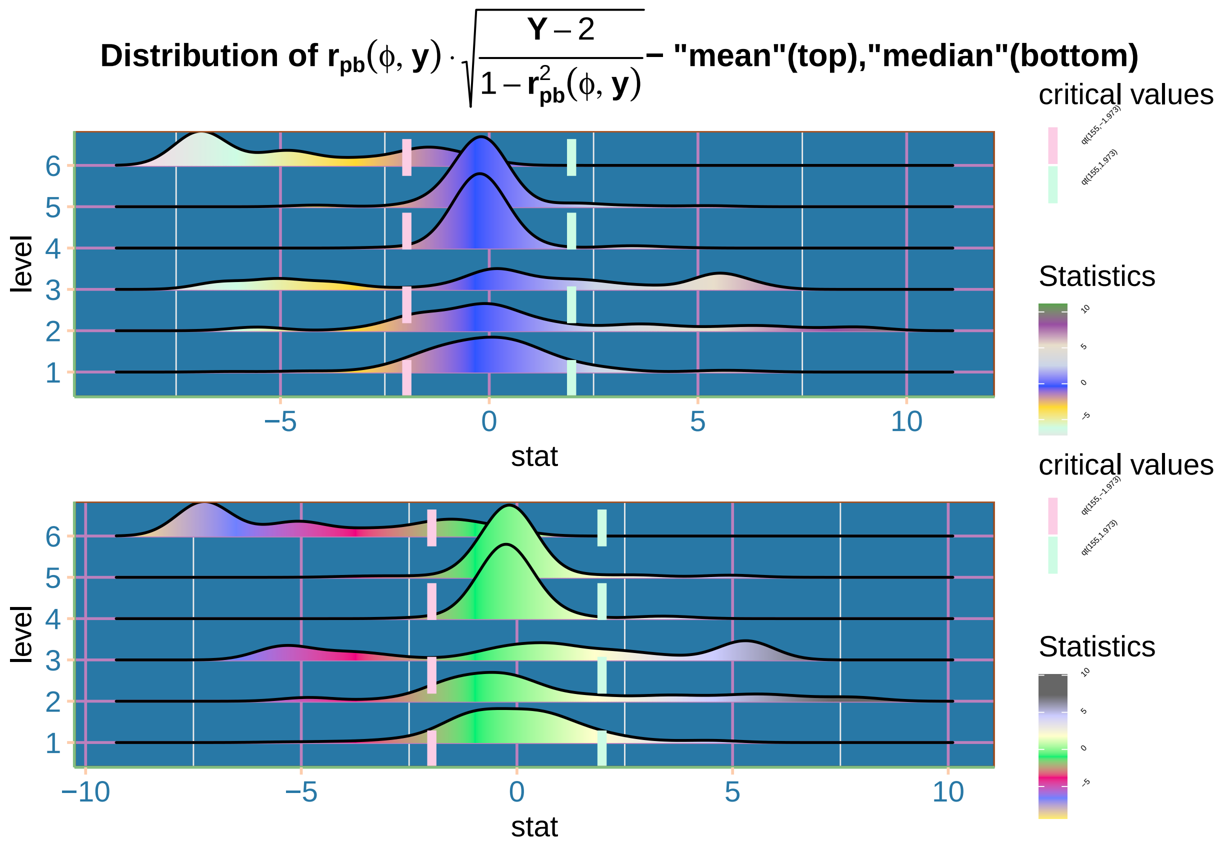

Figure 12: 2012-Stat(r_pb) distribution