Methodology and Technical Notes

May 28, 2020
Version 18 (Final)

Authors
Evan A. Krueger
Shahrzad Divsalar
Winston Luhur
Soon Kyu Choi
Ilan H. Meyer

Table of Contents

About the TransPop Study

The TransPop[1] study is the first national probability sample of transgender individuals in the United States (it also includes a comparative cisgender sample). A primary goal of this study was to provide researchers with a representative sample of transgender people in the United States. The study examines a variety of health-relevant domains including basic demographic characteristics, health outcomes and health behaviors, experiences with interpersonal and institutional discrimination, identity, and transition-related experiences.

In probability sampling, researchers target specific people at random (for example by sampling addresses or phone numbers). Each person in the target population has a known non-zero probability of being included in the sample. In comparison, in nonprobability sampling the researchers have no control on who is enrolled and who completes the study. Also, because typically a call goes out to people who meet a certain target characteristic (e.g., transgender respondents) and those people then may choose to participate, there may be bias related to who gets the call to participate and who actually participates. For example, if researchers advertise a study in a particular community outlet (e.g., Facebook, an LGBT community center), people who see this advertisement and respond to it may differ from people who do not participate in such community outlets or choose to not respond. In a probability sample, everyone is selected at random (e.g., if their address or phone number were selected). Of course, using either sampling approach, participation in the actual survey is voluntary, and some people may choose to not participate. But in probability sampling, researchers can correct for bias in responses by weighting the sample to be similar to the intended target population; in non-probability sampling researchers do not know who received the solicitation for the study therefore cannot readily correct for biases (Meyer & Wilson, 2009; Meyer, Marken, Russell, Frost, & Wilson, 2020).

Data Sources

Two data sources contribute to the TransPop study. The first source is the TransPop survey, which was administered to transgender people, including gender nonbinary people who identify as transgender (but not gender nonbinary individuals who did not identify as transgender). The second data source includes a national probability sample of cisgender respondents, collected for comparison to the TransPop sample.

---

[1]TransPop (U.S. Transgender Population Health Survey) is funded by a grant from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD grant R01HD090468). The TransPop investigators are: Ilan H. Meyer, Ph.D., (PI), Walter O. Bockting, Ph.D., Jody L. Herman, Ph.D., and Sari L. Reisner, Sc.D. (Co-Investigators, listed alphabetically).

We also combined the transgender and cisgender data sources to create one dataset that includes both samples. This dataset is described in more detail below.

Transgender Sample (N = 274)

The TransPop survey was collected in two periods, referred to as "TransPop1" (administered April 2016 – August 2016) and "TransPop2" (administered June 2017 – December 2018). Together, the two samples create the final TransPop sample.

Cisgender Sample (N = 1,162)

The cisgender companion survey was administered at two time points in February 19 – 23, 2018 and November 12 – December 10, 2018. (Because there are so many more cisgender individuals, the recruitment period for cisgender respondents was much shorter than it was for transgender respondents).

Transgender and Cisgender sample (N = 1,436)

The combined transgender and cisgender dataset includes all variables from the transgender and cisgender surveys. Variables that overlap are renamed to match the TransPop questionnaire. Variables that are only in the transgender survey, match the transgender questionnaire. Variables that are only in the cisgender survey match the cisgender questionnaire and have been relabeled with a prefix "cis".

This combined dataset was reweighted to allow researchers to make comparisons between transgender and cisgender respondents using data representative of each population. However, researchers should be aware that whole population estimates (estimates of a combined cisgender and transgender population) will be overwhelmed by the cisgender population, given that the transgender population is less than 5% of the total U.S. population. Reflecting this, the average sampling weight applied to cisgender respondents is 50 times larger than the average sampling weight applied to transgender respondents.

Three new variables were created for the combined transgender and cisgender dataset:

- **trans_cis**: indicates whether respondent identifies as transgender or cisgender
- **surveycompleted**: indicates which survey respondent completed
- **gender_identity**: indicates whether respondent identifies as cis man, cis woman, trans man (FTM), trans woman (MTF), or transgender gender nonbinary.

Please review the code books and crosswalk code book to learn more about these datasets.

Recruitment

Transgender Sample

TransPop participants were recruited and screened by Gallup, Inc., a survey research consulting company (http://www.gallup.com/), using two methodologies. The first was a probability sample of U.S. adults recruited using random digit dialing (RDD) to reach both cellphone and landline users. Following industry trends, Gallup later shifted to address-based sampling (ABS). ABS procedures included a mailed survey followed by a reminder mail. Respondents were provided, in addition to the mailed survey, a web link so they could self-administer the screener questionnaire on paper or online. Recruitment occurred during two periods: between April 2016 and August 2016 and between June 2017 and December 2018, with minimal changes in between the periods. Data from both recruitment periods were consolidated into one dataset.

Based on the recruitment period, Gallup asked two styles of screening questions for to identify gender minorities. After testing the questions after period 1, the questions were slightly revised in period 2 because that version was more efficient in reducing refusal and "Don't Know" answers from 7% in version 1 to 0.1% in version 2 (Meyer, Krueger, Marken, Reisner, Bockting, & Herman, *Comparing two versions of a 2-step assessment for identifying transgender respondents in a national sample of U.S. adults*, Paper presented at the 2017 meeting of the American Association of Public Opinion Research):

1. Period 1 (April 2016 – August 2016) respondents were first screened using the question of the Gallup U.S. Poll, "Do you, personally, identify as lesbian, gay, bisexual, or transgender?" Respondents who answered "yes," were then screened for eligibility. Transgender classification was based on a 2-step question, which, first, asked for respondent's sex assigned at birth ("What sex were you assigned at birth, on your original birth certificate?" With response options Female and Male) and then asked about the respondent's gender identity ("Which of the following terms best describes your current gender identity?" With response options, Woman, Man, Trans woman (Male-to-female), Trans man (Female-to-male), and Non-binary/Genderqueer). Responses from both of these questions were combined to classify people. Respondents were classified as transgender if their sex assigned at birth differed from gender identity and if they identified as transgender regardless of sex assigned at birth. That is, people who reported male sex at birth and identified their gender as "woman," or those who reported female sex at birth and identified their gender as "man" were classified as transgender. In addition, people who identified as "transgender" in the second-step question), regardless of sex at birth, were also classified as transgender.
2. Period 2 (June 2017 – December 2018) respondents were screened using the same scheme but using a revised 2-step questions. Respondents were asked a slightly different version of the question for their sex assigned at birth ("on your original birth certificate, was your sex assigned as female or male?") and gender identity ("do you currently describe yourself as a man, a woman, or transgender?"). Transgender respondents were further asked if they are "trans woman (male-to-

female),” “trans man (female-to-male),” or “non-binary/genderqueer.” Responses from the 2-step questions were combined to define the study population for period 2. Those who reported male sex assigned at birth and identified their gender as “woman,” those who reported female sex assigned at birth and identified their gender as “man”, and those who identified as “transgender” were defined as transgender.

Respondents were also screened for other eligibility requirements—adult over age 18, education above 6th grade, and conducted the interview in English (5% of Gallup respondents respond to the interview in Spanish but they were not interviewed for eligibility for TransPop).

Cisgender Sample

Cisgender participants were recruited and screened by Gallup, Inc., a survey research consulting company (http://www.gallup.com/), using two methodologies. The first was a probability sample of U.S. adults recruited using random digit dialing (RDD) to reach both cellphone and landline users. Following industry trends, Gallup later shifted to address-based sampling (ABS). ABS procedures included a mailed survey followed by a reminder mail. Respondents were provided, in addition to the mailed survey, a web link so they could self-administer the screener questionnaire online or on paper. Recruitment occurred in two periods: February 2018 using RDD and between November-December 2018 using ABS. Data from both recruitment periods were consolidated into one dataset. Because there are so many more cisgender individuals, the recruitment period for cisgender respondents was much shorter than it was for transgender respondents).

Respondents were screened for gender identity using the same methodology as the TransPop survey described above. Respondents who were not classified as transgender were invited to take to Cisgender survey. Respondents were also screened for other eligibility requirements—adult over age 18, education above 6th grade, and conducted the interview in English.

Procedure

Respondents who met the eligibility criteria were invited to participate in the TransPop and Cisgender studies. If they agreed, they were emailed or mailed a TransPop or Cisgender survey questionnaire to complete by self-administration. Participants responded to the TransPop or Cisgender survey by self-administering the study questionnaire either online via a link provided in an email or on paper via a mailed questionnaire returned in a pre-stamped preaddressed envelope. Along with their invitation to participate, potential respondents were sent a $25 gift certificate (an Amazon gift card with the emailed invitation and cash with the mailed questionnaire).

Participants read an information sheet (See Appendix 1) prior to beginning the survey and their consent was assumed if they completed the questions and submitted it to the

researchers. No signed consent forms were collected because it was determined that a signed consent form, it if were collected, would impose an unnecessary risk to the respondents' confidentiality. Identifying data were kept confidential at Gallup and separated from any data and were not made available to the investigators at any time.

The study protocol was reviewed and approved by the Gallup IRB, the UCLA IRB and the IRBs of collaborating institutions through reliance on UCLA IRB. Collaborating institutions have included Columbia University, the University of Texas at Austin, the University of California, Santa Cruz, the University of California, San Francisco, the University of Arizona, Surrey University, UK, and the University College London, UK.

Eligibility

Transgender Sample

Eligibility for the TransPop study was determined through 5 items, 3 screener items which already existed through Gallup, age, education requirement, and ability to conduct survey in English (Table 1). In addition, 2 screener items, transgender status and minimum completion of 6th grade, were added by the TransPop Study team (Table 2).

Table 1. Gallup items used to assess TransPop eligibility			
Measure	*Question Text*	*Response Options*	*TransPop Eligibility*
Age	*Please tell me your age*	*Open Ended*	*18+*
Education	*What is the highest level of school you have completed or the highest degree you have received?*	*Less than a high school diploma (Grades 1 through 11 or no schooling*	*If yes, then final education criteria assessed in subsequent education question, i.e., minimum 6th grade (below).*
		High school graduate (Grade 12 with diploma or GED certificate)	*Eligible*
		Technical, trade, vocational or business school or program after high school	*Eligible*

		Some college – college, university, or community college -- but no degree	*Eligible*
		Two year associate degree from a college, university, or community college	*Eligible*
		Four year bachelor's degree from a college or university (e.g., BS, BA, AB)	*Eligible*
		Some postgraduate or professional schooling after graduating college, but no postgraduate degree (e.g., some graduate school)	*Eligible*
		Postgraduate or professional degree, including master's, doctorate, medical, or law degree (e.g., MA, MS, PhD, MD, JD)	*Eligible*
Sexual orientation/ gender identity	*I have one final question we are asking only for statistical purposes. Do you, personally, identify as lesbian, gay, bisexual, or transgender?*	*Yes, do*	*Eligible. Only used as a screener in Period 1*
		No, do not	*Not eligible (Only used as a screener for Period 1)*
Survey	*Able to conduct survey in English*		*Eligible*
	Not able to conduct survey in English		*Not Eligible*

Table 2. Additional eligibility criteria added by TransPop Study Team

Measure	*Question Text*	*Response Options*	*TransPop Eligibility*
Education, 6th grade or higher	*What is the highest level of school you have completed? (Only asked of those selecting "Less than a high school diploma (Grades 1 through 11 or no schooling" on education)")*	*5th grade or lower*	*Not eligible*
		6th grade or higher	*Eligible*
Period 2 Transgender status*	*On your original birth certificate, was your sex assigned as female or male?*	*Female*	*Eligible if currently identify as "transgender" or assigned "female" at birth and currently identify as "man" or assigned "male" at birth and currently identify as "woman"*
		Male	
	Do you currently describe yourself as man, woman, or transgender?	*Man*	
		Woman	
		Transgender	

*see text for variation in Period 1 transgender eligibility questions

Cisgender Sample

Eligibility for the Cisgender study was determined through 5 items, 3 screener items which already existed through Gallup, age, education requirement, and ability to conduct survey in English (Table 3). In addition, 2 screener questions on gender identity (cisgender) status and minimum completion of 6th grade, were added by the TransPop Study team (Table 4).

Table 3. Gallup items used to assess Cisgender eligibility

Measure	*Question Text*	*Response Options*	*Cisgender Eligibility*
Age	*Please tell me your age*	*Open Ended*	*18+*

Education	What is the highest level of school you have completed or the highest degree you have received?	Less than a high school diploma (Grades 1 through 11 or no schooling	*If yes, then final education criteria assessed in subsequent education question, i.e., minimum 6th grade (below).*
		High school graduate (Grade 12 with diploma or GED certificate)	Eligible
		Technical, trade, vocational or business school or program after high school	Eligible
		Some college – college, university, or community college -- but no degree	Eligible
		Two year associate degree from a college, university, or community college	Eligible
		Four year bachelor's degree from a college or university (e.g., BS, BA, AB)	Eligible
		Some postgraduate or professional schooling after graduating college, but no postgraduate degree (e.g., some graduate school)	Eligible

11

Measure	Question Text	Response Options	Eligibility
		Postgraduate or professional degree, including master's, doctorate, medical, or law degree (e.g., MA, MS, PhD, MD, JD)	*Eligible*
Survey	*Able to conduct survey in English*		*Eligible*
	Not able to conduct survey in English		*Not Eligible*

Table 4. Additional eligibility criteria added by TransPop Study Team

Measure	Question Text	Response Options	Cisgender Eligibility
Education, 6th grade or higher	*What is the highest level of school you have completed? (Only asked of those selecting "Less than a high school diploma (Grades 1 through 11 or no schooling" on education)")*	*5th grade or lower*	*Not eligible*
		6th grade or higher	*Eligible*
Cisgender status	*On your original birth certificate, was your sex assigned as female or male?*	*Female*	*Eligible if currently identify as "female" or assigned "female" at birth, or currently identify as "man" or assigned "male" at birth.*
		Male	
	Do you currently describe yourself as man, woman, or transgender?	*Man*	
		Woman	
		Transgender	

Sample

Transgender Sample

Period 1

The Period 1 sample was recruited between April 2016 and August 2016. In total, 149,593 participants were screened by Gallup for inclusion in the study, of which 185

(0.12%) were classified as transgender and eligible to participate in the study. Of those eligible, 136 (73.5%) agreed to participate in the survey and of those, 56 (41.2%) completed the survey, for a total response rate of 30.3% (among those initially eligible).

Period 2

The Period 2 sample was recruited between June 2017 and December 2018. In total, 432,251 participants were screened by Gallup for inclusion in the study, of which 929 (0.21%) were classified as transgender and eligible to participate in the study. Of those eligible, 668 (71.9%) agreed to participate in the survey and of those, 270 (40.4%) completed the survey, for a total response rate of 28.7% (among those initially eligible).

The final TransPop sample size is 274.

For a more complete breakdown of the respondents recruited from the various Gallup polls, see Table 5.

Table 5. Recruitment from Gallup surveys (Period 2)						
	Gallup Politics & Economy Survey 2018	Gallup Sharecare Study 2018	Gallup Education Consumer Survey 2018	Gallup Sharecare Pilot 2017	Gallup Daily Poll 2017	Total
Total screened	75,749	41,735	128,806	14,504	171,457	432,251
Transgender	250	153	233	34	259	929
% Transgender	0.33%	0.37%	0.18%	0.23%	0.15%	0.21%
Agreed to participate	111	153	169	34	201	668

Originally 326 eligible respondents were included, 56 from period 1 and 270 from period 2. Of these 326 respondents, 52 (5 from period 1 and 47 from period 2) were dropped after determining these respondents misclassified as transgender. The decision to drop these respondents was determined on a case-by-case basis by comparing sex at birth with gender identity on both Gallup screener and the TransPop survey, additionally, by evaluating questions regarding transition (q245a-e and q246a-j), comfortability with the word transgender (q224), coming out milestone (q239), and name change (q242) questions.

TransPop Demographics

Table 6. TransPop Demographics (N = 274)	
	n (weighted %)
Age	
18–29	103 (48.9)

30-49	83 (33.0)
50-64	56 (13.0)
65 or older	29 (5.1)
Race	
Asian/Asian American	8 (4.5)
Black/African American	21 (9.5)
Hispanic/Latino/Spanish	8 (5.4)
Middle Eastern/North African	1 (0.6)
Native Hawaiian/Pacific Islander	4 (1.8)
White	187 (56.5)
American Indian/Alaskan Native	2 (0.3)
Multirace	42 (20.7)
Other	1 (0.8)
Gender Identity	
Man	41 (14.1)
Woman	54 (19.8)
Trans Man (FTM)	37 (16.8)
Trans Woman (MTF)	66 (18.0)
GNB	76 (31.3)
Assigned Sex at Birth	
Female	130 (54.3)
Male	144 (45.7)
Sexual Identity	
Straight/heterosexual	58 (17.6)
Lesbian	24 (8.4)
Gay	23 (8.5)
Bisexual	50 (18.9)
Queer	49 (18.1)
Same-gender loving	12 (4.0)
Other	21 (7.1)
Asexual spectrum	12 (5.4)
Pansexual	22 (12.0)
Education	
Less than high school diploma	6 (5.0)
High school degree or diploma	52 (39.1)
Technical/Vocational school	16 (3.4)
Some college	84 (27.7)
College graduate	60 (14.3)
Post graduate work or degree	52 (10.5)
Employment	
Work full-time for an employer	109 (36.0)
Work part-time for an employer	57 (22.5)
Self-employed	37 (14.7)
Unemployed but looking for work	24 (13.3)
Unemployed and have stopped looking for work	9 (3.7)
Not employed due to disability	32 (9.5)
Student	51 (19.0)

Retired	32 (4.9)
Homemaker or full-time parent	13 (7.1)
Other	9 (3.5)
Personal Income	
No income	15 (6.0)
$1 to $4,999	27 (15.9)
$5,000 to $9,999	35 (14.9)
$10,000 to $14,999	29 (7.6)
$15,000 to $19,999	31 (11.0)
$20,000 to $24,999	15 (5.8)
$25,000 to $29,999	13 (7.6)
$30,000 to $39,999	21 (7.0)
$40,000 to $49,999	20 (5.8)
$50,000 to $59,999	18 (5.7)
$60,000 to $74,999	10 (1.7)
$75,000 to $99,999	16 (2.8)
$100,000 to $149,999	14 (3.4)
$150,000 or more	10 (4.8)
Household Income	
No income	6 (1.9)
$1 to $4,999	14 (5.9)
$5,000 to $9,999	15 (6.2)
$10,000 to $14,999	23 (8.8)
$15,000 to $19,999	27 (9.3)
$20,000 to $24,999	8 (4.5)
$25,000 to $29,999	12 (5.7)
$30,000 to $39,999	26 (10.1)
$40,000 to $49,999	24 (7.9)
$50,000 to $59,999	30 (9.2)
$60,000 to $74,999	16 (7.3)
$75,000 to $99,999	24 (6.5)
$100,000 to $149,999	30 (10.2)
$150,000 or more	19 (6.6)
Poverty	
<100%	60 (25.7)
100-199%	46 (19.4)
200-299%	52 (20.1)
300%+	107 (34.8)
Urbanicity	
Non-urban	55 (16.4)
Urban	211 (83.6)
Census Region	
Northeast	54 (18.7)
Midwest	50 (19.9)
South	71 (30.0)
West	97 (31.4)

Cisgender Sample

The cisgender sample was recruited between February 19-23, 2018 and November 12-December 10, 2018. In total, 4,020 participants were screened by Gallup for inclusion in the study. In total, 87.6% (N = 3,523) were eligible to participate based on the final eligibility criteria. Of those eligible, 2,548 (72.3%) agreed to participate in the survey and of those, 1,184 (46.7%) completed the survey, for a total response rate of 33.6% (among those initially eligible). Of these, 22 respondents were dropped due to missing data leaving a final sample size of 1,162.

Cisgender Demographics

Table 7. Cisgender Demographics (N = 1,162)	
	n (weighted %)
Age	
18–29	90 (18.9)
30-49	252 (31.7)
50-64	352 (26.6)
65 or older	468 (22.8)
Race	
Asian/Asian American	21 (2.5)
Black/African American	58 (11.1)
Hispanic/Latino/Spanish	29 (4.9)
Middle Eastern/North African	4 (0.2)
Native Hawaiian/Pacific Islander	1 (0.0)
White	961 (72.3)
American Indian/Alaskan Native	5 (0.2)
Multirace	83 (8.8)
Gender Identity	
Woman	606 (52.0)
Man	556 (48.0)
Assigned Sex at Birth	
Female	606 (52.0)
Male	556 (48.0)
Sexual Identity	
Straight/heterosexual	1048 (90.1)
Lesbian	15 (1.2)
Gay	17 (1.4)
Bisexual	28 (4.3)
Queer	4 (0.5)
Same-gender loving	9 (1.0)
Other	3 (0.3)
Asexual spectrum	4 (0.6)
Pansexual	3 (0.6)
Education	
Less than high school diploma	26 (5.5)

High school degree or diploma	143 (26.4)
Technical/Vocational school	33 (2.2)
Some college	354 (29.4)
College graduate	301 (19.9)
Post graduate work or degree	305 (16.6)
Employment	
Work full-time for an employer	347 (35.0)
Work part-time for an employer	116 (13.0)
Self-employed in your own business, profession or trade, or operate a farm	131 (11.8)
Unemployed but looking for work	36 (5.2)
Unemployed and have stopped looking for work	11 (1.6)
Not employed due to disability	77 (6.8)
Student	30 (6.1)
Retired	446 (23.6)
Homemaker or full-time parent	55 (6.2)
Not listed above (please specify)	26 (1.9)
Personal Income	
No income	41 (6.1)
$1 to $4,999	40 (5.8)
$5,000 to $9,999	52 (6.7)
$10,000 to $14,999	80 (8.1)
$15,000 to $19,999	68 (6.5)
$20,000 to $24,999	65 (6.5)
$25,000 to $29,999	57 (3.6)
$30,000 to $39,999	95 (7.9)
$40,000 to $49,999	107 (8.6)
$50,000 to $59,999	89 (5.6)
$60,000 to $74,999	125 (9.9)
$75,000 to $99,999	130 (9.2)
$100,000 to $149,999	123 (8.1)
$150,000 or more	90 (7.2)
Household Income	
No income	14 (2.4)
$1 to $4,999	15 (2.2)
$5,000 to $9,999	33 (3.2)
$10,000 to $14,999	54 (5.3)
$15,000 to $19,999	47 (4.2)
$20,000 to $24,999	52 (5)
$25,000 to $29,999	42 (3.6)
$30,000 to $39,999	86 (10.4)
$40,000 to $49,999	82 (6.8)
$50,000 to $59,999	92 (5.5)
$60,000 to $74,999	124 (10.6)
$75,000 to $99,999	155 (11.8)
$100,000 to $149,999	191 (15.2)

$150,000 or more	175 (13.9)
Poverty	
<100%	111 (16.0)
100-199%	162 (15.8)
200-299%	167 (19.6)
300%+	645 (48.6)
Urbanicity	
Non-urban	267 (20.3)
Urban	895 (79.7)
Census Region	
Northeast	219 (17.7)
Midwest	292 (22.5)
South	378 (37.1)
West	273 (22.7)

Data Processing and Transformation

TransPop data set

The dataset consists of all questions asked on the TransPop1 and TransPop2 survey (variables labeled "q01," "q02," etc.). Variable numbers match the question numbers from the TransPop2 survey. For example, the question "Which of the following best describes your current sexual orientation?" was question 34 on the TransPop 1 survey, but it was question 32 on the TransPop2 survey. In the dataset, this question is labeled "q32."

By and large, the TransPop1 and TransPop2 surveys contained identical questions. In cases where the questions differed, however, the prefix "t1" was affixed to variables specific to the TransPop1 survey and the prefix "t2" was affixed to variables specific to the TransPop2 survey. In addition to the TransPop survey questions, key variables, such as demographics, from the Gallup survey are included in the final dataset (affixed with "g" prefix).

Cisgender data set

The dataset consists of all questions asked on the survey (variables labeled "q01," "q02," etc.). Variable numbers match the question numbers from the Cisgender survey. Key variables are also included in the dataset from the Gallup survey (affixed with "g" prefix).

New variable creation

Several variables were created using items from the TransPop and Cisgender surveys. The calculated variables are included in the final dataset. Each newly created variable is described below.

Survey complete

A variable (**surveycompleted)** was included that indicates whether TransPop respondents completed the first (TransPop1) or second (TransPop2) waves of data collection.

Sex assigned at birth and gender identity

TransPop:
Respondents' sex assigned at birth (variable: **sex**) and gender identity (variable: **gender**) were calculated based on their reported sex at birth (variable: t1q27, t2q28) and gender identity (variables: t1q28, t1q29, t2q29, t2q30). 9 respondents were missing one or both items on the survey. Due to variation in responses between screening and survey completion, the decision was made to replace *both* "sex" and "gender" with respondents' reported sex and gender at screening – on the Gallup survey – when either was missing. 18 additional respondents provided sex and gender responses at screening that classified them as transgender (FTM, MTF, or GNB), and therefore eligible to complete the TransPop survey. For example, on the survey, they provided responses that classified them as cisgender (e.g., reported female sex at birth + woman gender identity on the survey, whereas, on the screener, they reported female sex at birth + man gender identity). Since it is not possible to determine whether these respondents mistakenly identified their sex at birth or gender identity – either at screening or on the survey – or whether these changes reflect a shift in self-definition, these 18 respondents were reviewed and reclassified into one of the three eligible categories on a case-by-case basis, based upon their responses to several other questions throughout the survey. Both the calculated (**sex**, **gender**, **trans**) and "raw" variables used to create the calculated variables (t1q27, t1q28, t2q28, t2q29, t2q30) are included in the dataset.

Cisgender:
Respondents' sex assigned at birth (variable: **sex**) and gender identity (variable: **gender**) were calculated based on their reported sex at birth (variable: q28) and gender identity (variables: q29 and q30) on the survey.
20 respondents were missing sex at birth, and so the response provided at screening was assigned. 34 respondents were missing a gender identity, and so the response provided at screening was assigned. To maintain the "cleanest" comparison between transgender (TransPop) and cisgender respondents, the decision was made to drop 7 respondents from the Cisgender survey who did not consistently report sex and/or gender between screening and on the survey.

Age

TransPop:
Respondents were asked "in what year were you born?" (variable: q195), and a numeric age (variable: **age**) was calculated by subtracting birth year from the year in which the respondent completed the survey. 8 respondents were missing a value for age. Of them, 4 were missing a value for q195 and 4 had a value of "9999" for q195. For these 8 respondents, their age reported on the Gallup screener was assigned.

Cisgender:
Respondents were asked "in what year were you born?" (variable: q146), and a numeric age (variable: **age**) was calculated by subtracting birth year from the year in which the respondent completed the survey. 55 respondents were missing a value for age because they were missing a value for q146, and so their age reported on the Gallup screener was assigned. 1 additional respondent's calculated age was 17, but the age they reported in response to the screener question asked by Gallup was 18. Their calculated age was, then, recoded to 18.

Race

TransPop:
A 9-category race/ethnicity variable (**race**) was calculated based on responses from the TransPop survey (variables: q20_1 – q20_7). Respondents who selected more than one race/ethnicity from the 7 categories were categorized as "multiracial" (n=42). 3 respondents did not provide a race response on the survey, and so their race reported to Gallup at screening was assigned. Of these 3, 1 respondent was categorized into an "other" category, which corresponded to the "other" category in the Gallup screener race variable (**grace**).

Additionally, a race/ethnicity variable (**race_recode**), was created where respondents who indicated Latino/Hispanic were coded to Latino regardless of whether they had selected a race category in addition to Latino/Hispanic. The resulting categories are White, Black, Latino/Hispanic/Spanish origin, Asian, and Multirace, Native Hawaiian/Pacific Islander, and other. This variable was then recoded into a 5-category variable (**race_recode_cat5**) with fewer response options (White, Black, Latino/Hispanic/Spanish origin, Multirace, and other).

Cisgender:
A 9-category race/ethnicity variable (**race**) was calculated using responses from the TransPop survey (variables: q20_1 – q20_7). Respondents selecting more than one race/ethnicity were categorized as "multiracial" (n=84).
24 respondents did not provide a race response on the survey, and so their race reported to Gallup at screening was assigned. 3 remaining respondents did not provide a race – either on the Cisgender survey or at screening. Because race is considered a critical analytic variable, and it was used to impute other variables (described below), these three respondents were dropped from the dataset.
Additionally, a race/ethnicity variable (**race_recode**), was created where respondents who indicated Latino/Hispanic were coded to Latino regardless if they selected multiple categories. The resulting categories are White, Black, Latino, Asian, Multirace, Native Hawaiian/Pacific Islander, and other. Furthermore, a more restricted 5-category variable (**race_recode_cat5**) was derived from the variable **race_recode**.

Sexual identity

TransPop:

Two calculated sexual identity variables are included in the dataset. The first variable (**sexualid**) is equivalent to respondents' self-reported sexual identity on the survey (variable: q34). However, 59 respondents provided a write-in response (variable: q34verb). These 59 responses were placed into existing categories when possible (e.g., "Straight, with desire to be woman with a man" write-in response was placed into the "Heterosexual" identity category), and new categories were created for common write-in responses (e.g., pansexual). As such, the final variable (sexualid) contains more response categories than the original survey item (q34). When a respondent provided two identity labels in their write-in response, the first label was used for categorization purposes (e.g., "Bisexual and pansexual" was coded as "Bisexual"). However, asexual took precedence when a respondent wrote two identity labels, and one label was asexual. The remaining write-in responses that could not be easily categorized were kept in the "other" category. 3 respondents did not provide a sexual identity and were coded as missing. The resulting categorizations are shown in Table 9. A second calculated variable (**sexminid**) was also included, which specified whether respondents reported a heterosexual identity (n=58) or a sexual minority identity (n=213).

Table 8. Sexual identity write-in responses and resulting categorizations	
Resulting categorization (sexualid)	**Write-in response (q34_verb)**
Straight/heterosexual	Straight female, which has a strong relationship component as well as social acceptability. Female sexuality is not at all like male sexuality.
Straight/heterosexual	Straight, with desire to be woman with a man
Bisexual	Bisexual and Pansexual
Bisexual	Bisexual, Queer, and Same-gender Loving
Bisexual	I'm bisexual, but I want to invest my love in building community with women and women are who I seek as partners
Asexual spectrum	ASEXUAL
Asexual spectrum	ASEXUAL
Asexual spectrum	Asexual
Asexual spectrum	Asexual
Asexual spectrum	Asexual
Asexual spectrum	Asexual
Asexual spectrum	I'm asexual, but aesthetically attracted to other girls
Asexual spectrum	NON SEXUAL
Asexual spectrum	Pan-Aromantic Asexual*
Asexual spectrum	Queer Asexual*
Asexual spectrum	pansexual & asexual*

Pansexual	PANSEXUAL
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual
Pansexual	Pansexual/queer
Pansexual	Polysexual (like bi but includes 3rd genders)
Pansexual	pan-sexual
Pansexual	pansexual
Pansexual	pansexual
Pansexual	pansexual
Other	DUDE! PLEASE HELP!
Other	HUMAN
Other	I prefer men sexually but am celibate.
Other	LOVING WOMEN AS A MAN WOULD
Other	NONE OF THESE
Other	NOT SURE
Other	TRANSGENDER
Other	TRANSPECIES FELIDIAN
Other	Transgender
Other	trans woman that loves men
Other	transwoman married to transwoman
Other	x

*In order to capture all asexual individuals, we coded these individuals as *asexual* regardless of other labels.

Cisgender:
Two calculated sexual identity variables are included in the dataset. The first variable (**sexualid**) is equivalent to respondents' self-reported sexual identity on the survey

(variable: q32). However, 22 respondents who provided a write-in response were placed into existing categories when possible (e.g., "Straight, with desire to be woman with a man" write-in response was placed into the "Heterosexual" identity category), and new categories were created for common write-in responses (e.g., pansexual). As such, the final variable (sexualid) contains more response categories than the original survey item (q32). When a respondent provided two identity labels in their write-in response, the first label was used for categorization purposes (e.g., "Heterosexual biromantic" was coded as "Heterosexual"). The remaining write-in responses could not be easily categorized and were kept in the "other" category. 31 respondents did not provide a sexual identity and were coded as missing. The resulting categorizations are shown in Table 10. A second calculated variable (**sexminid**) was also included, which specified whether respondents reported a heterosexual identity (n=1,053) or a sexual minority identity (n=85).

Table 9. Sexual identity write-in responses and resulting categorizations	
Resulting categorization (sexualid)	**Write-in response (q32_verb)**
Straight/heterosexual	CELIBANT NO SEXUAL ACTIVITY
Straight/heterosexual	Heterosexual Biromantic
Straight/heterosexual	I have been married to the same male husband for over 50 years. I don't quite understand the term 'straight/heterosexual.
Straight/heterosexual	MARRIED TO WIFE 50 YEARS
Straight/heterosexual	MARRIED WOMAN
Straight/heterosexual	STRAIGHT
Straight/heterosexual	STRAIGHT
Straight/heterosexual	STRAIGHT
Straight/heterosexual	STRAIGHT
Straight/heterosexual	STRAIGHT LOVE MEN
Pansexual	Omni sexual
Pansexual	Pansexual
Pansexual	Pansexual
Asexual spectrum	Asexual
Asexual spectrum	Asexual
Asexual spectrum	Asexual
Asexual spectrum	asexual celibate
Other	ABSTINIST
Other	Bi-romantic, gray-a (ish)
Other	None of the above
Missing	WIDOW
Missing	NONE

Income

TransPop:
Two income variables: household income (variable: **hinc**) and personal income (variable: **pinc**) were included, which matched the variables included on the TransPop survey. The responses from TransPop 1 (t1q198 and t1q200, respectively) and TransPop 2 (t2q204 and t2q206, respectively) were collapsed. However, the response categories differed between the two surveys, and so the midpoints of the TransPop 1 categories were used when creating the final calculated variables (e.g., $720-5,999 was assigned as $2,640). 18 respondents did not indicate a household income on the survey. Of them, 10 provided a household income to Gallup, and so missing values were replaced accordingly. The remaining 8 missing responses were imputed using Predictive Mean Matching, described in detail in a later section. 11 respondents did not indicate a personal income on the survey. Because Gallup did not ask respondents to report a personal income, all 11 missing values were imputed using Predictive Mean Matching. Both un-imputed (hinc, pinc) and imputed (**hinc_i, pinc_i**) versions of the variables are included in the dataset.

Cisgender:
Two income variables: household income (variable: **hinc**) and personal income (variable: **pinc**) were included, constructed from q155 and q157, respectively.
95 respondents did not indicate a household income on the survey. Of them, 63 provided a response to Gallup, and so missing values were replaced accordingly. However, because Gallup used different age ranges for categorization, the midpoints of Gallup responses were used when assigning values to the final calculated variable (e.g., $720-5,999 was assigned as $2,640). The remaining 32 missing responses were imputed using Predictive Mean Matching. 105 respondents did not indicate a personal income on the survey. Because Gallup did not ask respondents report a personal income, all 105 missing values were imputed using Predictive Mean Matching. Both un-imputed (hinc, pinc) and imputed (**hinc_i, pinc_i**) versions of the variables are included in the dataset.

Poverty

TransPop:
Using weighted Census estimates for poverty thresholds in 2018 (US Census Bureau, 2018), respondents were categorized as either living in poverty (below 100% FPL) or not, based upon their reported household income (hinc), and the reported number of people living on that household income (q205) (constructed variable: **poverty**). Respondents were asked to indicated their household income within a range. To calculate poverty, we used the midpoint of that range to compare against poverty thresholds. Using the same approach, a second variable was created (**povertycat**), which placed respondents into 1 of 4 income categories: <100% FPL, 100-199% FPL, 200-299% FPL, ≥ 300% FPL. An imputed version of poverty (**poverty_i, povertycat_i**), using imputed household income (hinc_i) and reported number of people living on that household income (q205_i) variables is also in the dataset.

Cisgender:

Using weighted Census estimates for poverty thresholds in 2018 (US Census Bureau, 2018), respondents were categorized as either living in poverty (below 100% FPL) or not, based upon their reported household income (hinc_i), and the reported number of people living on that household income (q156) (constructed variable: **poverty**). Respondents were asked to indicated their household income within a range. To calculate poverty, we used the midpoint of that range to compare against poverty thresholds. Using the same approach, a second variable was created (**povertycat**), which placed respondents into 1 of 4 income categories: <100% FPL, 100-199% FPL, 200-299% FPL, ≥ 300% FPL. An imputed version of poverty (**poverty_i, povertycat_i**), using imputed household income (hinc_i) and reported number of people living on that household income (q156_i) variables is also in the dataset.

Education

TransPop & Cisgender:
Responses from the Gallup Daily Tracking Survey variable (**geducation**) were re-categorized into two additional variables with fewer response options: **geduc1** (high school or less, some college, college completed, more than college completed) and **geduc2** (high school or less, more than high school).

Geography

Urbanicity

TransPop:
Using respondents' zip codes, urbanicity scores were calculated using the USDA Rural-Urban Commuting Area coding system (USDA, 2013). RUCA scores are included in the dataset (variable: **gruca**). 2010 RUCA codes were used, and scores of 1-3 represent urban zip codes, while scores of greater than 3 represent non-urban zip codes (variable: **gurban**). 8 respondents' zip codes did not have a corresponding RUCA code or corresponding urbanicity score. These 8 values were imputed using Predictive Mean Matching. Both un-imputed (gruca, gurban) and imputed (**gruca_i, gurban_i**) versions of the variables are included in the dataset.

Cisgender:
Using respondents' zip codes, urbanicity scores were calculated using the USDA Rural-Urban Commuting Area coding system (USDA, 2013). RUCA scores are included in the dataset (variable: **gruca**). 2010 RUCA codes were used, and scores of 1-3 represent urban zip codes, while scores of greater than 3 represent non-urban zip codes (variable: **gurban**). 16 respondents' zip codes did not have a corresponding RUCA code or corresponding urbanicity score. These 16 values were imputed using Predictive Mean Matching. Both un-imputed (gruca, gurban) and imputed (**gruca_i, gurban_i**) versions of the variables are included in the dataset.

Census region and division

TransPop & Cisgender:
Using respondents' states of residence (**gzipstate**), respondents were assigned to their corresponding Census regions (**gcenreg**) and divisions (**gcendiv**) (US Census Bureau, 2015). There are 9 Census divisions:

1. New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont)
2. Middle Atlantic (New Jersey, New York, Pennsylvania)
3. East North Central (Indiana, Illinois, Michigan, Ohio, Wisconsin)
4. West North Central (Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota)
5. South Atlantic (Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia)
6. East South Central (Alabama, Kentucky, Mississippi, Tennessee)
7. West South Central (Arkansas, Louisiana, Oklahoma, Texas)
8. Mountain (Arizona, Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming)
9. Pacific (Alaska, California, Hawaii, Oregon, Washington)

There are 4 corresponding Census regions:

1. Northeast (New England, Middle Atlantic regions)
2. Midwest (East North Central, West North Central regions)
3. South (South Atlantic, East South Central, West South Central regions)
4. West (Mountain, Pacific regions)

Distance from an LGBT community health center

TransPop & Cisgender:
Distance from the respondents' residence to the nearest LGBT community health center (**gmilesaway**). This distance was calculated using geocoded health center data and respondents' zip codes (**gzipcode**), as described by Martos et al. (2017).

Scale creation

Several items from the TransPop study are part of validated scales, designed to measure constructs relevant to identity, stress, and health. Each of the scales within the TransPop survey have been calculated from individual variables, according to published instructions, detailed below. Two calculated variables are included in the dataset for each of the scales: an un-imputed version and an imputed version. The unimputed version has missing values for participants who were missing on one or more items that make up the scale. The imputed variable has no missing values. The steps taken to create each scale are described below.

Positive Health

Social Well-Being assessed one's "appraisal of one's circumstances and functioning in society," and serves as a measure of one's "social wellness" (Keyes, 1998). Keyes (1998) Social Well-Being scale consists of 15 items (TransPop: q04- q18, Cisgender: q04-q18; e.g. "I don't feel I belong to anything I'd call a community," "My community is a source of comfort," "I have something valuable to give to the world."), each rated on a 7-point Likert scale ranging from "strongly disagree" to "strongly agree." To create a scale variable, 8 of the 15 items (q04, q08, q11, q12, q14, q15, q16, q17) were reverse-coded then the scale was created as a mean score of each of the items within the scale. Lower values represent lower social well- being and higher values represent greater social well-being. Scale values range from 1 to 7.

There were two resulting variables: "**socialwb**" (calculated only from complete cases, in which no individual scale items were missing) and "**socialwb_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Satisfaction with Life (Satisfaction with Life Scale, SWLS) assessed respondents' global satisfaction with life "as a cognitive-judgmental process" (Diener et al., 1985). The scale consisted of 5 items (TransPop: q224-q228; Cisgender: q173-q177; e.g., "In most ways my life is close to ideal," "The conditions of my life are excellent," "I am satisfied with life."), each rated on a 7-point Likert scale ranging from "strongly disagree" to "strongly agree." The scale variable was created as a mean score of each of the items within the scale. Lower values represent less satisfaction with life and higher values represent greater satisfaction with life. Scale values range from 1 to 7.

There were two resulting variables: "**lifesat**" (calculated only from complete cases, in which no individual scale items were missing) and "**lifesat_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Identity

Multi-group Ethnic Identity was assessed using Phinney and Ong's (2007) revised Multi-group Ethnic Identity Measure (MEIM-R). MEIM-R assessed respondents' "investigation, learning, and commitment" to their race/ethnic identities (Phinney & Ong, 2007). The scale consisted of 6 items (TransPop: q22-q27, Cisgender: q22-q27; e.g., "I have spent time trying to find out more about my race/ethnic group, such as its history, traditions, and customs," and "I have a strong sense of belonging to my own race/ethnic group." Each item was rated on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree." The scale variable was created as a mean score of each of the items within the scale. Lower values represent less investigation, learning, and commitment to one's own race/ethnic identity, and higher values represent greater investigation, learning, and commitment. Scale values range from 1 to 5.

There were two resulting variables: "**meim**" (calculated only from complete cases, in which no individual scale items were missing) and "**meim_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Community connectedness (TransPop only), a 5-item scale described by Testa et al. (2015), assessed the desire for and strength of the transgender community affiliation among respondents. Scale items (q51-q55) included "I feel a part a community of people who share my gender identity," and "I feel connected to other people who share my gender identity" Responses were recorded on a 5-item scale ranging from "strongly disagree strongly" to "strongly agree." Of note, q54 and q55 were reverse-coded in order to bring it in line with other scale items. The scale variable was then created as a mean score of each of the items within the scale. Lower scores represented lower community connectedness, while higher scores represented greater community connectedness. Scale values range from 1 to 5.

There were two resulting variables: "**connectedness**" (calculated only from complete cases, in which no individual scale items were missing) and "**connectedness_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Non-affirmation of gender identity (TransPop only), a 6-item scale from Testa et al. (2015), assessed the degree to which respondents felt their gender identity was understood and accepted by others. Scale items (q45-q50) included "I have to repeatedly explain my gender identity to people or correct the pronouns people use," and "I have difficulty being perceived as my gender." Responses were recorded on a 5-point scale ranging from "strongly disagree" to "strongly agree." The scale was created as a mean score of each of the items within the scale. Lower values represent more affirmation, while higher values represent less affirmation (more non-affirmation). Scale values range from 1 to 5.

There were two resulting variables: "**nonaffirm**" (calculated only from complete cases, in which no individual scale items were missing) and "**nonaffirm_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Healthcare Access & Utilization

Healthcare Stereotype Threat (TransPop only), a 4-item scale modified from Abdou & Fingerhut's (2014) measure, assessed the degree to which respondents worried about being negatively judged by or confirming stereotypes about LGBT people with healthcare providers. Scale items (q77-q80) included "I worry about being negatively judged because of my sexual orientation or gender identity," and "I worry that evaluations of me may be negatively affected by my sexual orientation or gender identity." Responses were recorded on a 5-point scale ranging from "strongly disagree" to "strongly agree." The scale was created as a mean score of each of the items within the scale. Lower values represent less worry about being judged or confirming LGBT stereotypes, and higher values represent greater worry. Scale values range from 1 to 5.

There were two resulting variables: "**hcthreat**" (calculated only from complete cases, in which no individual scale items were missing) and "**hcthreat_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Health Outcomes

Mental Disability was assessed using the Kessler-6, a 6-item scale from the National Comorbidity Survey (Kessler et al., 2003). Scale items (TransPop: q100a-q100f, Cisgender: q61a-q61f) asked respondents how often, in the past 30 days, they had felt "nervous," "hopeless," "restless or fidgety," "so depressed that nothing could cheer you up," "that everything was an effort," and "worthless." Responses were recorded on a 5-point scale ranging from "all of the time" to "none of the time." All items were first reverse-coded so that "none of the time" had a value of 1 and "all of the time" had a value of 5. The scale was then created as the sum of all variables within the scale. Per scale creation instructions, respondents failing to answer any single item in the scale were recorded as "missing," on the resulting scale score. In addition, an imputed version of the scale was calculated in which missing individual scale items were imputed, and a final scale score was calculated for each respondent.

The resulting scales, named "**kessler6**" and "**kessler6_i**" had values ranging from 0 to 24.
There appear to be no clear standards for optimal K6 scoring. The unweighted scale has values in the range 0–24. The scoring rule used in most applications based on standard validation studies is to classify respondents with scores of 13–24 as having probable serious mental illness and those with scores of 0–12 as probably not having serious mental illness (Kessler et al., 2003). Furukawa and colleagues (Furukawa et al., 2003, 2008) have shown that this simple dichotomous scoring approach can be refined by using polychotomous rather than dichotomous scoring rules that collapse K6 scores into strata based on analysis of data in a clinical calibration study such that the observed prevalence of serious mental illness differs significantly across strata. For example, one such scoring rule might collapse K6 scores into strata with K6 score values of 0, 1–7, 8–12, 13–18, and 19–24, with respondents in each stratum assigned a predicted probability of serious mental illness based on the results of a clinical calibration study (Kessler et al., 2010).

Alcohol use was assessed using the Alcohol Use Disorder Identification Test (AUDIT-C), a 3-item scale designed to identify persons with hazardous drinking behavior, or who have active alcohol use disorders (Bush et al., 1998). The scale items (TransPop: q112-q114, Cisgender: q73-q75) and available responses were "how often do you have a drink containing alcohol?" (never [0 points], monthly or less [1 point], 2-4 times a month [2 points], 2-3 times a week [3 points], 4 or more times a week [4 points]), "how many standard drinks containing alcohol do you have on a typical day?" (none [0 points], 1 or 2 [0 points], 3 or 4 [1 point], 5 or 6 [2 points], 7 to 9 [3 points], 10 or more [4 points]), and "how often do you have six or more drinks on one occasion?" (never [0 points], less than monthly [1 point], monthly [2 points], weekly [3 points], daily or almost daily [4 points]). The scale was then created as the sum of all variables in the scale. Per scale creation

instructions, respondents failing to answer any single item in the scale were recorded as "missing," on the resulting scale score. In addition, an imputed version of the scale was calculated in which missing individual scale items were imputed, and a final scale score was calculated for each respondent.

The resulting scales, named "**auditc**," and "**auditc_i**" had values ranging from 0 to 12.

The recommending screening thresholds for the AUDIT-C questions to identify alcohol use disorders or risky drinking is 4 or more for men, and 3 or more for women (Frank et al., 2008). One study examined the AUDIT-C using the gender screening thresholds for the transgender population based on sex assigned at birth and then again based on binary gender identities, and did not find any difference in the results (Kidd et al., 2019). Thus, the paper applied the screening thresholds based on sex assigned at birth, as it is also unclear how to apply the screening thresholds to nonbinary gender identities.

Drug use was assessed using the Drug Use Disorders Identification Test (DUDIT), an 11-item scale designed to identify individuals with drug- related problems (Table 11; Berman et al., 2003). The scale was created as the sum of all variables (TransPop: q117-q127, Cisgender: q78-q88) in the scale. Per scale creation instructions, respondents failing to answer any single item in the scale were recorded as "missing," on the resulting scale score. In addition, an imputed version of the scale was calculated in which missing individual scale items were imputed, and a final scale score was calculated for each respondent.

The resulting scales, named "**dudit**," and "**dudit_i**" had values ranging from 0 to 44.

The suggested cut-off score for men with drug-related problems is a score of 6 or more, indicating probable drug-related problems, either substance abuse/harmful use or dependence. For women, the cut-off score is 2 or more. For both sexes, it is highly probable that a score of 25 or more indicates dependence on one or more drugs (Berman et al., 2003).

Table 10. Variables used for calculating DUDIT scale			
Variable	*Question Text*	*Response Options*	*Points*
q117/q78	*How often do you use drugs other than alcohol?*	*Never*	*0*
		Once a month or less often	*1*
		2-4 times a month	*2*
		2-3 times a week	*3*
		4 times a week or more often	*4*
q118/q79	*Do you use more than one type of drug on the same occasion?*	*Never*	*0*
		Once a month or less often	*1*
		2-4 times a month	*2*
		2-3 times a week	*3*
		4 times a week or more often	*4*
q119/q80	*How many times do you take drugs on a*	*0*	*0*

	typical day when you use drugs?	1-2	1
		3-4	2
		5-6	3
		7 or more	4
q120/q81	How often are you influenced heavily by drugs?	Never	0
		Once a month or less often	1
		2-4 times a month	2
		2-3 times a week	3
		4 times a week or more often	4
q121/q82	Over the past year, have you felt that your longing for drugs was so strong that you could not resist it?	Never	0
		Once a month or less often	1
		2-4 times a month	2
		2-3 times a week	3
		4 times a week or more often	4
q122/q83	Has it happened, over the past year, that you have not been able to stop taking drugs once you get started?	Never	0
		Once a month or less often	1
		2-4 times a month	2
		2-3 times a week	3
		4 times a week or more often	4
q123/q84	How often over the past year have you taken drugs and then neglected to do something you should have done?	Never	0
		Once a month or less often	1
		2-4 times a month	2
		2-3 times a week	3
		4 times a week or more often	4
q124/q85	How often over the past year have you needed to take a drug the morning after heavy drug use the day before?	Never	0
		Once a month or less often	1
		2-4 times a month	2
		2-3 times a week	3
		4 times a week or more often	4
q125/q86	How often over the past year have you had guilt feelings or a bad conscience because you used drugs?	Never	0
		Once a month or less often	1
		2-4 times a month	2
		2-3 times a week	3
		4 times a week or more often	4
q126/q87	Have you or anyone else been hurt (mentally or physically) because you used drugs?	No	0
		Yes, but not over the past year	2
		Yes, over the past year	4
q127/q88	Has a relative or a friend, a doctor or a nurse, or anyone else, been worried about your drug use or said to you that you should stop using drugs?	No	0
		Yes, but not over the past year	2
		Yes, over the past year	4

Stressors

Internalized transphobia (TransPop only) assessed the degree to which respondents accept stigma as a part of their own value systems (Testa et al., 2015). Scale items (q156-q161) included "I resent my transgender identity" and "I ask myself why I can't just be normal?" Responses were recorded on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree." The scale was created as a mean score of each of the items within the scale. Lower values represent less internalized transphobia and higher values represent greater internalized transphobia. Scale values range from 1 to 5.

There were two resulting variables: "**internalized**" (calculated only from complete cases, in which no individual scale items were missing) and "**internalized_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Everyday discrimination assessed chronic, relatively minor experiences of discrimination or unfair treatment (Williams et al., 1997). Scale items (TransPop: q171a-q171i, Cisgender: q121a-q121i) asked respondents who often the following things happened to them over the past year, including "you were treated with less courtesy than other people," "you were treated with less respect than other people," "and you were called names or insulted." Responses were recorded on a 4-point Likert scale ranging from "often" to "never." The scale was created as a mean score of each of the items within the scale. The resulting variable was reverse-coded so that lower values represent less everyday discrimination and higher values represent more everyday discrimination. Scale value range from 1 to 4.

There were two resulting variables: "**everyday**" (calculated only from complete cases, in which no individual scale items were missing) and "**everyday_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Chronic strains (Wheaton, 1999, abridged version). Scale items (TransPop: q173a-q173l, Cisgender: q123a-q123l) asked respondents to think about their lives currently, and to determine whether several statements "were not true", "somewhat true", or "very true". A "does not apply" response option was also provided. Questions included "you're trying to take on too many things at once," "your job often leaves you feeling both mentally and physically tired," "and you are alone too much." "Does not apply" was coded as 0. "Not true" was also coded as 0, "somewhat true" was coded as 1, and "very true" was coded as 2.

Childhood gender conformity (Zucker et al., 2006). Scale items (TransPop: q174-q177, Cisgender: q124-q127) included "as a child, my favorite toys and games were…," and "as a child, the characters on TV or in the movies that I imitated or admired were…" Response were recorded on a 5-point scale, with the wording of response options varying according to the question, but all ranged from "masculine" (e.g., 1= "always 'masculine,'" "always boys or men") to "feminine" (e.g., 5= "always 'feminine,'" "always girls or women"). "Neither" and "not applicable" responses were set as missing.

A preliminary score was assigned to each participant, and was calculated the mean score of all the values present within the scale for each individual. A final categorical score was then calculated for each participant, based on their sex at birth (male/female), using cutoff scores described in the table below. The resulting variable for the scale was named "**childgnc**." In addition, an imputed version of the scale was calculated in which missing individual scale items were imputed, and a final scale score was calculated for each respondent in the same manner (**childgnc_i**). The resulting categories of the scale are 1 "Top decile (most GNC)" 2 "Median-Top decile" 3 "< Median (least GNC)".

Table 11. Cutoff scores used to calculate childgnc		
	Lower cutoff	Upper cutoff
90th percentile or greater (most gender non-conforming in childhood)	**TransPop:** Assigned Female at Birth: 1.00 Assigned Male at Birth: 4.25 **Cisgender:** Assigned Female at Birth: 1.00 Assigned Male at Birth: 2.25	**TransPop:** Assigned Female at Birth: 1.25 Assigned Male at Birth: 5.00 **Cisgender:** Assigned Female at Birth: 3.00 Assigned Male at Birth: 5.00
Between 50th percentile and 90th percentile	**TransPop:** Assigned Female at Birth: 1.25 Assigned Male at Birth: 3.25 **Cisgender:** Assigned Female at Birth: 3.00 Assigned Male at Birth: 1.50	**TransPop:** Assigned Female at Birth: 2.25 Assigned Male at Birth: 4.25 **Cisgender:** Assigned Female at Birth: 3.75 Assigned Male at Birth: 2.25
Less than 50th percentile (least gender non-conforming in childhood)	**TransPop:** Assigned Female at Birth: 2.25 Assigned Male at Birth: 1.00 **Cisgender:** Assigned Female at Birth: 3.75 Assigned Male at Birth: 1.00	**TransPop:** Assigned Female at Birth: 5.00 Assigned Male at Birth: 3.25 **Cisgender:** Assigned Female at Birth: 5.00 Assigned Male at Birth: 1.50

Adverse childhood experiences (ACE) (CDC-BRFSS, 2010). Scale items (TransPop: q178-q188, Cisgender: q124-q127) asked respondents to "look back before you were 18 years of age," and included items such as "did you live with anyone who was depressed, mentally ill, or suicidal," and "how often did your parents or adults in your home ever slap, hit, kick, punch, or beat each other up?" Available response options ranged from dichotomous (yes/no) to 3-point Likert scales (never to more than once), depending on the question. "Don't know/not sure" and "refused" answer options were also available to respondents, where appropriate. To create a summary ACE score, all items were dichotomized (1= yes, event occurred at least once vs. 0=no, event never occurred) if not already dichotomized. Per published instruction (CDC, 2016), 8 subscores were created from the existing 11 items: presence of emotional abuse (**ace_emo:** q185), physical abuse (**ace_phy:** q184), sexual abuse (**ace_sex:** q186, q187, q188), household intimate partner violence (**ace_ipv:** q183), household substance use (**ace_sub:** q179, q180), household

33

mental illness (**ace_men:** q178), parental separation or divorce (**ace_sep:** q182), incarcerated household member (**ace_inc:** q181). A resulting final score was created as a sum score indicating the number of adverse childhood experiences respondents reported during childhood. Scale values for the resulting ACE measure (**ace**) range from 0 to 8. Respondents indicating "don't know" or "refused" on any single scale item were recorded as missing for that subscore(s), and the subsequent final score.

Missing individual scale items were also imputed using predictive mean matching, and individual subscores and the final scale score was calculated for each respondent (**ace, ace_emo_i, ace_phy_i, ace_sex_i, ace_ipv_i, ace_sub_i, ace_men_i, ace_sep_i, ace_inc_i**).

Gender identity non-disclosure (TransPop only), a 5-item scale from Testa et al. (2015), assessed the degree to which respondents avoid disclosing their gender identity to others. Scale items (q151-q155) included "I don't talk about certain experiences from my past or I change parts of what I will tell people," and "I modify my way of speaking." Responses were recorded on a 5-point scale ranging from "strongly disagree" to "strongly agree." The scale was created as a mean score of each of the items within the scale. Lower values represent less of an effort to avoid disclosure, while higher values represent a higher effort to avoid disclosure. Scale values range from 1 to 5.

There were two resulting variables: "**nondisclosure**" (calculated only from complete cases, in which no individual scale items were missing) and "**nondisclosure_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Negative expectations of the future (TransPop only), a 9-item scale from Testa et al. (2015), assessed the degree to which respondents expect they will not be understood or accepted because of their gender identity. Scale items (q191a-q191i) included "if I express my gender identity/history, others wouldn't accept me," and "if I express my gender identity/history, employers would not hire me." Responses were recorded on a 5-point scale ranging from "strongly disagree" to "strongly agree." The scale was created as a mean score of each of the items within the scale. Lower values represent expecting less rejection, while higher values represent expecting more rejection. Scale values range from 1 to 5.

There were two resulting variables: "**negexpfuture**" (calculated only from complete cases, in which no individual scale items were missing) and "**negexpfuture_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Social support

Social support was assessed utilizing the Multidimensional Scale of Perceived Social Support (Zimet et al., 1988). Scale items (TransPop: q192a-q192l, Cisgender: q141a-q141l) asked respondents to rate their levels of agreement with several items, including

"there is a special person who is around when I am in need," and "my family really tries to help me." Responses were recorded on a 7-point scale ranging from "very strongly disagree" to "very strongly agree." The scale was created as a mean score of each of the items within the scale. Lower values represent less perceived social support and higher values represent more perceived social support. Scale values range from 1 to 7.

There were two resulting variables: "**socsupport**" (calculated only from complete cases, in which no individual scale items were missing) and "**socsupport_i**" (missing individual scale items were imputed, and a final scale score was calculated for each respondent).

Additionally, 3 subscales were created, representing perceived social support from significant others (**socsupport_so**: q192a, q192b, q192e, q192j), family (**socsupport_fam**: q192c, q192d, q192h, q192k), and friends (**socsupport_fr**: q192f, q192g, q192i, q192l). Each subscale was similarly created as a mean score of each of the items within the subscale. Lower values represent less perceived social support and higher values represent more perceived social support. Subscale values range from 1 to 7.

Missing individual scale items were also imputed using predictive mean matching, and individual subscores were calculated for each respondent (**socsupport_so_i, socsupport_fam_i, socsupport_fr_i**).

Missing Data and Imputation

When possible, missing values on demographic characteristics were assigned from other known sources. (See Appendices 3 & 4 for missing values in the TransPop and Cisgender datasets, respectively).

TransPop

- Sex at birth and gender identity: 9 respondents were missing sex at birth and/or gender identity on the TransPop survey. Of them 5 were missing BOTH sex at birth and gender identity, 2 were missing sex at birth only, and 2 were missing gender identity only. For all 9 respondents, **both** sex at birth and gender identity were replaced with the values reported at screening, regardless of whether respondents were missing one or both values on the survey.
- 18 additional respondents provided responses to the sex at birth and gender identity items on the TransPop survey that would categorize them as cisgender (e.g., "male" sex at birth & "man" gender identity). Various items were reviewed throughout the survey for each of these 18 respondents, and missing values were imputed on a case-by-case basis.
- 4 respondents were missing a valid age on the TransPop survey (age). In each case, respondents were assigned the age reported to Gallup on the Gallup survey.
- 3 respondents were missing a race on the TransPop survey (race). Of them, 1 was assigned the race reported to Gallup at screening, and 2 were assigned the race reported to Gallup on the Gallup survey.

- 18 respondents were missing a household income. Of them, 10 could be assigned the household income reported to Gallup on the Gallup survey.

Cisgender

- 2 respondents were missing a valid age on the Cisgender survey (age). Both were assigned the age reported to Gallup on the Gallup survey.
- 30 respondents were missing a race on the Cisgender survey (race). Of them, 27 were assigned the race reported to Gallup at screening, and 3 were assigned the race reported to Gallup on the Gallup survey.
- 20 respondents were missing a sex at birth on the Cisgender survey (sex). All 20 were assigned the age reported to Gallup on the Gallup survey.
- 34 respondents were missing a current gender identity on the Generations survey (gender). All 20 were assigned the age reported to Gallup on the Gallup survey.
- 96 respondents were missing a household income. Of them, 63 could be assigned the household income reported to Gallup on the Gallup survey.

For the remaining missing values of demographic characteristic variables—household income, personal income, number of people in household, poverty, RUCA (rural-urban commuting area), and urbanicity—we did a single imputation by chained equations (fully conditional specification), using predictive mean matching (Little, 1988) to draw the imputed values. With predictive mean matching, regression is used to predict the missing value, and then a single value is randomly selected from the k observed values nearest to the predicted missing value from a donor pool of complete observations. We used donor pools of size k=5 according to Heitjan and Little (1991). When doing imputation by chained equations, each of the imputed variables serve as predictors in the imputation regression models for all other imputed variables. Additionally, age, race/ethnicity, and sex at birth, completed through other sources, were included in the imputation models to improve matching.

Predictive mean matching can be considered a more general form of hot-deck imputation, in which missing values are imputed by matching non-respondents to respondents only through categorical predictors. These matching-imputation methods are attractive because they recreate distributions well by using observed values for imputations and because they are somewhat more robust to misspecification of the imputation model (e.g. normality assumption violation) than parametric imputation methods (Morris et al., 2014). For each of the variables that were imputed, both the original/un-imputed and imputed versions are available in the dataset.

Defining missing values

True or "system missing" are denoted with a period "**.**" in the dataset. In addition, we denote the following extended missing values: **.a** denotes missing values arising from skip patterns in the questionnaire; **.b** denotes "don't know" or "cannot recall" responses; **.c** denotes "never" or "none"; **.d** denotes "does not apply"; **.e** denotes "refused".

Sample weighting information

In Stata, the sample weight can be applied to analyses using the "svy" command. Both the TransPop and Cisgender datasets have the same variable named **weight** as the weighting variable. To register the survey design of the data for analysis, use the following command: svyset _n [pweight=(weight)]

TransPop sample weighting

The Transpop 2 respondents were recruited from the following sources:

> Gallup Daily Politics and Economy survey (6/21/2017-12/10/2018) and Sharecare Wellbeing Index survey (6/21/2017-12/31/2017)

> The two tracks form the frame. All respondents are asked for permission to be re-contacted in the future. The starting weight is the daily national demographic weight.

> i. "RECONTACT_WT": The weighted profile of all respondents is compared to those who agree to be re-contacted (Frame vs Re-contact). The re-contact group is adjusted for non-response on age, education and race.

> ii. "OPT_IN_WT": Respondents who qualified for the Transpop survey are compared to those who opt-in to join the Transpop survey. The opt-in group is adjusted for non-response on gender, age, education, census region, race and ethnicity.

> iii. "Complete_WT": Respondents who completed and returned the Transpop survey are compared to the opt-in group. The complete group is adjusted for non-response on gender, age, education, census region, race and ethnicity.

A. Gallup-Sharecare Wellbeing Mail Survey (1/1/2018-12/1/2018) –
   a. Qualified respondents are asked for permission to be re-contacted for the Transpop survey. The starting weight is the monthly national demographic weight.
      i. "OPT_IN_WT": Respondents who qualified for the Transpop survey are compared to those who opt-in to join the Transpop survey. The opt-in group is adjusted for non-response on gender, age, education, census region, race and ethnicity.
      ii. "Complete_WT": Respondents who completed and returned the Transpop survey are compared to the opt-in group. The complete group is adjusted for non-response on age.

B. Strada Education Network Education Consumer Survey

a. The starting weight is 1 and no non-response adjustment is applied due to small number of completes.

C. Gallup-Sharecare Well-Being Index Pilot
    a. The starting weight is 1 and no non-response adjustment is applied due to small number of completes.

The four groups of completes are then combined together. The combined dataset is adjusted on gender, age, education, census region, race and ethnicity. Targets are based on the demographic profile of the respondents who qualified for the Transpop survey in the Gallup political and economic surveys and wellbeing surveys.

Gallup reweighted the total Transpop 1 and 2 studies using the steps and targets outlined in the Transpop 2 study description above.

Cisgender sample weighting

The Cisgender respondents were recruited from the 2018 Gallup Daily PE Track (2/12/2018-2/18/2018, 11/13/2018-12/2/2018). Respondents who qualified for the Cisgender survey formed the qualified group. Any adult (18 years of age or older) who was not identified as a member of the Transgender group and who also had at least 6th grade of school education was eligible (qualified). Qualified respondents who agreed to participate in the survey formed the opt-in group. The complete group consisted of respondents from the opt-in group who had completed the follow-up survey (the Cisgender survey).

Each respondent of the qualified group had an initial weight from the PE track which was weighted to represent the 18+ US population. The first non-response adjustment was to compare the qualified group to the opt-in group under the assumption that the opt-in group may be different in characteristics from the respondents who did not agree to join the Cisgender survey. Weighting factors were calculated for the opt-in group to represent the qualified group by accounting for members from the qualified group who refused to participate. Age and education were the two demographic characteristics that were adjusted in this step.

The next step was to account for the differences between those who had returned the completed surveys (the complete group) and those who had not (opt-in group). The non-response adjustment was based on gender, age, education, race, ethnicity and region. Non-response adjustment cells were created based on these variables and weighting factors were derived to account for non-response within each of those cells.

The final step was a post-stratification adjustment. Cisgender profile (or the target data for cisgender group) was created using the weighted estimates of the 2018 Gallup Daily Track. All qualified (cisgender) respondents of the 2018 daily track survey for the entire year were used to generate the target data for post-stratification weighting of the cisgender group. The complete group with the non-response adjusted weights (i.e., with

cumulative weights prior to the stage of post-stratification) as starting weights were then post stratified on gender, age education, race, ethnicity and region to create the final weight using the cisgender profile targets. Finally, minimal trimming of weights was carried out to eliminate some extreme weights.

Applying the sample weights

Sample weights are available for use with the data. When applied, results from analyses are generalizable to the U.S. population of transgender and cisgender adults aged 18 and above. The following instructions detail how to apply the final sample weight using Stata or SPSS.

STATA

In Stata, the sample weight can be applied to analyses using the "svy" command.

For example, if the procedure is a generalized linear model, use the same procedure as you would when analyzing non-complex survey data, but include the prefix **svy**: regress x y z.

To register the survey design of the data for analysis, use the following command:

> **svyset _n [pweight=(*weight variable name*)]**

> Copy the code above but replace "weight variable name" with the weight variable of the dataset you are using.

SPSS

In SPSS, first a Complex Samples Module is needed, this is not included with the base SPSS package. From there, create an "Analysis Plan" which contains survey design variable information and identifies the sampling weight. When running an analysis procedure, procedures that are found in the Complex Sample modules must be used and these are accessed through a link to the Analysis Plan that the analyst must create.

For example, if the procedure generalized linear model (GLM) is used to run a regression with data, with complex survey data, the analyst needs to create the Analysis Plan file and then use complex sample generalized linear model (CSGLM) to run the regression. CSGLM uses a syntax different from GLM.

The steps to analysis survey data in SPSS:
1. Complex Samples Module is needed
2. Create an Analysis Plan file using the code below:

> **CSPLAN ANALYSIS**

/PLAN FILE='*myplanfile***.csaplan'**
/PLANVARS ANALYSISWEIGHT=*weight variable name*
/DESIGN
/ESTIMATOR TYPE=WR.

Copy the code above, and replace'myplanfile.csaplan' with file name that makes sense for your project. Keep the .csaplan extension. Replace "weight variable name" with the weight variable in the dataset.

3. Use Complex Samples procedures to analyze the data

Point-and-click is another method to create the Analysis Plan file. To find step-by-step instructions on this approach, please see Appendix 6.

References

Abdou, C. M., & Fingerhut, A. W. (2014). Stereotype Threat Among Black and White Women in Health Care Settings. Cultural Diversity & Ethnic Minority Psychology, 20(3), 316–323. http://doi.org/10.1126/science.1249098.Sleep

Alcohol Use Disorder Identification Test (AUDIT-C). Retrieved from: http://www.integration.samhsa.gov/images/res/tool_auditc.pdf

Berman, A.H., Bergman, H., Palmstierna, T., & Schlyter, F. (2003). The Drug Use Disorders Identification Test (DUDIT) Manual. Retrieved from: http://www.paihdelinkki.fi/sites/default/files/duditmanual.pdf

Bostwick (2012) Assessing Bisexual Stigma and Mental Health Status: A Brief Report, Journal of Bisexuality, 12(2), 214-222.

Centers for Disease Control and Prevention (CDC). (2016, April 01). Violence Prevention. Retrieved from https://www.cdc.gov/violenceprevention/acestudy/ace_brfss.html

CDC-BRFSS (2010). Adverse Childhood Experiences (ACE) module. Retrieved from: http://www.acestudy.org/

Diener, E. ., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. Journal of Personality, 49(1), 71–75.

Frank, D., DeBenedetti, A.F., Volk, R.J., Williams, E.C., Kivlahan, D.R., & Bradley, K.A. (2008). Effectiveness of the AUDIT-C as a screening test for alcohol misuse in three race/ethnic groups. Journal of General Medicine, 23 (6), 781-787.

Frost, D. M., & Meyer, I. H. (2012). Measuring community connectedness among diverse sexual minority populations. Journal of Sex Research, 49(1), 36–49.

Furukawa T.A., Kawakami N., Saitoh M., Ono Y., Nakane Y., Nakamura Y., Tachimori H., Iwata N., Uda H., Nakane H., Watanabe M., Naganuma Y., Hata Y., Kobayashi M., Miyake Y., Takeshima T., Kikkawa T. (2008). The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res, 17, 152–158, DOI: 10.1002/mpr.257

Furukawa T.A., Kessler R.C., Slade T., Andrews G. (2003). The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. Psychol Med, 33, 357–362, DOI: 10.1017/S0033291702006700

Gallup Daily Methodology. (2015).

Herek (2008), Hate Crimes and Stigma-Related Experiences Among Sexual Minority Adults in the United States. Journal of Interpersonal Violence. Retrieved from: http://jiv.sagepub.com/content/early/2008/04/07/0886260508316477.full.pdf+html

Herek et al (2009), Internalized stigma among sexual minority adults: Insights from a social psychological perspective. Journal of Counseling Psychology, 56(1).

Heitjan, D.F, & Little, R.J.A. (1991). Multiple Imputation for the Fatal Accident Reporting System. Applied Statistics, 40(1), 13-29.

Kessler, R.C., Green, J. G., Gruber, M.J. et al. (2010). Screening for serious mental illness in the general population with the K6 screening scale: results from the WHO World Mental Health (WMH) survey initiative. Int. J. Methods Psychiatr. Res. 19(Supplement 1): 4–22, DOI: 10.1002/mpr.310

Kessler R.C., Barker P.R., Colpe L.J., Epstein J.F., Gfroerer J.C., Hiripi E., Howes M.J., Normand S.L., Manderscheid R.W., Walters E.E., Zaslavsky A.M. (2003). Screening for serious mental illness in the general population. Arch Gen Psychiatry, 60, 184 –189.

Keyes, C. L. (1998). Social Well-Being. Social Psychology Quarterly, 61(2), 121–140.

Kidd, J.D., Levin, F.R., Dolezal, C., Hughes, T., & Bockting, W.O. (2019). Understanding predictors of improvement in risky drinking in a U.S. multi-site, longitudinal cohort study of transgender individuals: Implications for culturally-tailored prevention and treatment efforts. Addictive Behaviors, 96.

Little, R.J.A. (1988). Missing-Data Adjustments in Large Surveys. Journal of Business & Economic Statistics, 6(3), 287-296.

Martos AJ, Wilson PA, Meyer IH (2017). Lesbian, gay, bisexual, and transgender (LGBT) health services in the United States: Origins, evolution, and contemporary landscape. PLoS ONE 12(7), e0180544. https://doi.org/10.1371/journal. pone.0180544

Mohr, J. J., & Kendra, M. S. (2011). Revision and extension of a multidimensional measure of sexual minority identity: the Lesbian, Gay, and Bisexual Identity Scale. Journal of Counseling Psychology, 58(2), 234–45. http://doi.org/10.1037/a0022858

Meyer, I. H., & Wilson, P. A. (2009). Sampling Lesbian, Gay, and Bisexual Populations. Journal of Counseling Psychology, 56(1), 23-31. doi:10.1037/a0014587

Meyer, I.H., Marken, S., Russell, S.T., Frost, D.M., Wilson, B.D.M. (2020). An innovative approach to the design of a national probability sample of sexual minority adults. *LGBT Health.*

Morris, T.P., White, I.R., & Royston, P. (2014). Tuning multiple imputation by predictive mean matching and local residual draws. BMC Medical Research Methodology, 14(75).

National Comorbidity Survey. Kessler 6 - Self Report Q1 (a)-(f). Retrieved from: http://www.integration.samhsa.gov/images/res/K6%20Questions.pdf

Phinney, J. S., & Ong, A. D. (2007). Conceptualization and measurement of ethnic identity: Current status and future directions. Journal of Counseling Psychology, 54(3), 271–281.

Testa, R.J., Habarth, J., Peta, J., Balsam, K., & Bockting, W. (2015). Development of the Gender Minority Stress and Resilience Measure. *Psychology of Sexual Orientation and Gender Diversity, 2*(1), 65.

US Census Bureau. (2015). 2010 Geographic Terms and Concepts - Census Divisions and Census Regions. Retrieved from https://www.census.gov/geo/reference/gtc/gtc_census_divreg.html

US Census Bureau. (2018). Poverty Thresholds. Retrieved from https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html

USDA. (2013). Rural-Urban Commuting Area Codes. Retrieved from https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx

Wheaton B. The nature of stressors. In: Horwitz AF, Scheid TL, editors. A handbook for the study of mental health: Social contexts, theories, and systems. Cambridge, UK: Cambridge University Press; 1999. pp. 176–197.

Williams, D.R., Yu, Y., Jackson, J.S. & Anderson, N.B (1997). Racial differences in
physical and mental health: Socioeconomic status, stress, and discrimination.
Journal of Health Psychology, 2(3).

Zimet, G.D., Dahlem, N.W., Zimet, S.G. & Farley, G.K. (1988). The Multidimensional
Scale of Perceived Social Support. Journal of Personality Assessment, 52, 30-41.
Retrieved from: http://www.yorku.ca/rokada/psyctest/socsupp.pdf

Zucker, K.J., Mitchell, J.N., Bradley, S.J., Tkachuk, J. Cantor, J.M. & Allin, S.M.(2006),
The Recalled Childhood Gender Identity/Gender Role Questionnaire:
Psychometric properties. Sex Roles, 54(7).

Appendices

Appendix 1: Information Sheet

TransPop Study

The TransPop survey is the first survey to examine the health and well-being of transgender individuals. The survey explores identity, stress, health outcomes, and health care among transgender individuals.

You were selected as a participant in this survey because you are 18 years or older and because you recently told Gallup you were willing to participate in this survey. Your participation in this research study is completely voluntary and you can skip any question you do not want to answer.

The information you provide will be kept confidential and will be kept separate from your identifying information including your name, email address, or home address. Information will only be reported in the aggregate.

As a token of our appreciation you will receive $25 for your participation in this survey.

If you have any questions about this research, may contact the study's primary investigator, Dr. Ilan Meyer at meyer@law.ucla.edu, call (310) 825-7270, or write to him at The Williams Institute UCLA School of Law, Box 951476, Los Angeles, CA 90095.

UCLA Office of the Human Research Protection Program (OHRPP):

If you have questions about your rights as a research subject, or you have concerns or suggestions and you want to talk to someone other than the researchers, you may contact the UCLA OHRPP by phone: (310) 206-2040; by email: participants@research.ucla.edu or by mail: Box 951406, Los Angeles, CA 90095-1406.

Cisgender Survey

Gallup is conducting an important study on behalf of the University of California at Los Angeles (UCLA) and we need your help.

You were selected as a participant in this survey because you are 18 years or older and because you recently told Gallup you were willing to participate in this survey. Your participation in this research study is completely voluntary and you can skip any question you do not want to answer. Because you were randomly selected to participate in our poll, your responses are important, and we cannot replace you with someone else. Your responses will represent thousands of people just like you who were no selected to participate.

The information you provide will be kept confidential and will be kept separate from your identifying information including your name, email address, or home address. Information will only be reported in the aggregate. As a token of our appreciation, we have enclosed $25.00.

If you have questions about this research, you may contact the study's primary investigator, Dr. Ilan Meyer by email: meyer@law.ucla.edu; by phone: (310)825-7270; or by mail: The Williams Institute UCLA School of Law, Box 951476, Los Angeles, CA 90095.

If you have any questions about your rights as a research subject, or you have concerns or suggestions and you want to talk to someone other than the researchers, you may contact the UCLA OHRPP by phone: (310)206-2040; by email: participants@research.ucla.edu or by mail: Box 951406, Los Angeles, CA 90095-1406.

If you need assistance, please contact Gallup Support at galluppoll@gallupmail.com or 1-888-297-8999.

Thank you in advance for helping Gallup achieve our mission of "Helping People be Heard."

Gallup, Inc.

Appendix 2: TransPop Imputed Scale reliability (Cronbach's α) by total sample and gender categorization

Scale	Total Sample	Gender Categorization		
		Trans Woman (Male-to-female)	Trans Man (Female-to-male)	Non-binary/Genderqueer
Social Well-Being	0.81	0.79	0.77	0.87
Satisfaction with Life	0.90	0.89	0.91	0.90
Multi-group Ethnic Identity	0.82	0.76	0.85	0.86
Community Connectedness	0.78	0.80	0.77	0.76
Non-affirmation of Gender Identity	0.93	0.92	0.95	0.88
Healthcare Stereotype Threat	0.90	0.92	0.87	0.87
Mental Disability	0.90	0.91	0.91	0.89
Alcohol Use	0.65	0.68	0.63	0.61
Drug Use	0.87	0.85	0.86	0.88
Internalized Transphobia	0.86	0.86	0.83	0.88
Everyday Discrimination	0.92	0.93	0.93	0.90
Childhood Gender Conformity	0.84	0.80	0.80	0.77
Adverse Childhood Experiences	0.75	0.76	0.74	0.75
Gender Identity Non-disclosure	0.75	0.77	0.72	0.77
Expectations of Rejection	0.91	0.92	0.92	0.89
Social Support	0.92	0.94	0.86	0.91

Appendix 3: Cisgender Survey Imputed Scale reliability (Cronbach's α) by total sample and sex at birth

Scale	Total Sample	Sex at Birth	
		Female	Male
Social Well-Being	0.79	0.78	0.81
Satisfaction with Life	0.90	0.90	0.91
Multi-group Ethnic Identity	0.85	0.85	0.85
Mental Disability	0.88	0.88	0.88
Alcohol Use	0.58	0.44	0.65
Drug Use	0.86	0.84	0.87
Everyday Discrimination	0.90	0.89	0.91
Childhood Gender Conformity	0.94	0.75	0.76
Adverse Childhood Experiences	0.76	0.79	0.73
Social Support	0.96	0.96	0.95

Appendix 4: Missing values for each variable in TransPop dataset

Variable	Missing	Total Sample	Percent Missing
studyid	0	274	0
weight	0	274	0
surveycomp~d	0	274	0
gresponden~e	0	274	0
q01	3	274	1.09
q02	16	274	5.84
q03	12	274	4.38
q04	1	274	0.36
q05	1	274	0.36
q06	0	274	0
q07	0	274	0
q08	0	274	0
q09	0	274	0
q11	1	274	0.36
q12	0	274	0
q13	0	274	0
q14	2	274	0.73
q15	2	274	0.73
q16	3	274	1.09
q17	2	274	0.73
q18	3	274	1.09
q19a	2	274	0.73
q19b	2	274	0.73
q19c	2	274	0.73
q19d	3	274	1.09
q20_1	261	274	95.26
q20_2	240	274	87.59
q20_3	249	274	90.88
q20_4	268	274	97.81
q20_5	264	274	96.35
q20_6	54	274	19.71
q20_7	249	274	90.88
q21_verb	97	274	0
q22	2	274	0.73
q23	2	274	0.73

Variable	Missing	Total Sample	Percent Missing
q24	4	274	1.46
q25	2	274	0.73
q26	3	274	1.09
q27	3	274	1.09
q31	0	274	0
q32	3	274	1.09
q33	2	274	0.73
q34	4	274	1.46
q34_verb	221	274	0
q35_1	131	274	47.81
q35_2	166	274	60.58
q35_3	233	274	85.04
q35_4	226	274	82.48
q35_5	200	274	72.99
q36a	14	274	5.11
q36b	13	274	4.74
q36c	38	274	13.87
q36d	35	274	12.77
q36e	41	274	14.96
q36f	44	274	16.06
q37	2	274	0.73
q38	125	274	45.62
q39	123	274	44.89
q40	123	274	44.89
q41	123	274	44.89
q42	2	274	0.73
q43	0	274	0
q44	3	274	1.09
q45	0	274	0
q46	1	274	0.36
q47	1	274	0.36
q48	0	274	0
q49	0	274	0
q50	1	274	0.36
q51	1	274	0.36
q52	1	274	0.36
q53	2	274	0.73

Variable	Missing	Total Sample	Percent Missing
q54	2	274	0.73
q55	1	274	0.36
q56	38	274	13.87
q57	39	274	14.23
q58	44	274	16.06
q59	0	274	0
q60	2	274	0.73
q61	2	274	0.73
q62_t_verb	260	274	0
q62a	148	274	54.01
q62b	149	274	54.38
q62c	153	274	55.84
q62d	151	274	55.11
q62e	227	274	82.85
q63_t_verb	253	274	0
q63a	138	274	50.36
q63b	138	274	50.36
q63c	140	274	51.09
q63d	138	274	50.36
q63e	140	274	51.09
q63f	139	274	50.73
q63g	140	274	51.09
q63h	140	274	51.09
q63i	139	274	50.73
q63j	213	274	77.74
q64	169	274	61.68
q65	3	274	1.09
q66	150	274	54.74
q67	143	274	52.19
q68	159	274	58.03
q69	2	274	0.73
q70	84	274	30.66
q73_1	261	274	95.26
q73_2	262	274	95.62
q74	9	274	3.28
q75	234	274	85.4
q76_1	251	274	91.61

Variable	Missing	Total Sample	Percent Missing
q76_2	271	274	98.91
q77	2	274	0.73
q78	3	274	1.09
q79	2	274	0.73
q80	2	274	0.73
q81_1	250	274	91.24
q81_2	181	274	66.06
q81_3	258	274	94.16
q81_4	225	274	82.12
q81_5	271	274	98.91
q81_6	250	274	91.24
q81_7	264	274	96.35
q81_8	232	274	84.67
q81_9	235	274	85.77
q81_10	267	274	97.45
q81_11	259	274	94.53
q81_12	273	274	99.64
q81_13	262	274	95.62
q81_t_verb	263	274	0
q82	53	274	19.34
q83_1	172	274	62.77
q83_2	159	274	58.03
q83_3	238	274	86.86
q83_4	255	274	93.07
q83_5	261	274	95.26
q83_t_verb	262	274	0
q84	94	274	34.31
q85	1	274	0.36
q86_verb	270	274	0
q88	5	274	1.82
q89	52	274	18.98
q90_1	209	274	76.28
q90_2	137	274	50
q90_3	113	274	41.24
q91	54	274	19.71
q92a	7	274	2.55
q92b	10	274	3.65

Variable	Missing	Total Sample	Percent Missing
q92c	11	274	4.01
q92d	7	274	2.55
q93	0	274	0
q94	89	274	32.48
q95	57	274	20.8
q96	90	274	32.85
q97_1	218	274	79.56
q97_2	209	274	76.28
q97_3	256	274	93.43
q97_4	269	274	98.18
q97_5	265	274	96.72
q97_6	268	274	97.81
q97_7	266	274	97.08
q97_8	227	274	82.85
q97_9	250	274	91.24
q97_10	261	274	95.26
q97_11	252	274	91.97
q97_12	241	274	87.96
q97_13	219	274	79.93
q97_14	262	274	95.62
q97_15	258	274	94.16
q97_16	234	274	85.4
q97_17	261	274	95.26
q97_18	261	274	95.26
q97_19	267	274	97.45
q97_20	269	274	98.18
q97_21	263	274	95.99
q97_22	254	274	92.7
q97_23	200	274	72.99
q98	0	274	0
q99	0	274	0
q10	0	274	0
q100a	1	274	0.36
q100b	1	274	0.36
q100c	1	274	0.36
q100d	1	274	0.36
q100e	1	274	0.36

Variable	Missing	Total Sample	Percent Missing
q100f	1	274	0.36
q101	1	274	0.36
q102	2	274	0.73
q103	2	274	0.73
q104	2	274	0.73
q105	4	274	1.46
q106	4	274	1.46
q107	12	274	4.38
q108	12	274	4.38
q109	12	274	4.38
q110	3	274	1.09
q111	4	274	1.46
q112	3	274	1.09
q113	2	274	0.73
q114	1	274	0.36
q115	1	274	0.36
q116	163	274	59.49
q117	1	274	0.36
q118	2	274	0.73
q119	3	274	1.09
q120	2	274	0.73
q121	1	274	0.36
q122	3	274	1.09
q123	3	274	1.09
q124	3	274	1.09
q125	5	274	1.82
q126	4	274	1.46
q127	4	274	1.46
q128	1	274	0.36
q129	243	274	88.69
q130	124	274	45.26
q131	125	274	45.62
q132	1	274	0.36
q133	233	274	85.04
q134	191	274	69.71
q135	191	274	69.71
q136	1	274	0.36

Variable	Missing	Total Sample	Percent Missing
q137	240	274	87.59
q138	141	274	51.46
q139	141	274	51.46
q140	1	274	0.36
q141	224	274	81.75
q142	238	274	86.86
q143	226	274	82.48
q144	227	274	82.85
q145	175	274	63.87
q146	2	274	0.73
q147	256	274	93.43
q148	160	274	58.39
q149	162	274	59.12
q150a	5	274	1.82
q150b	11	274	4.01
q150c	7	274	2.55
q150d	6	274	2.19
q150e	15	274	5.47
q150f	9	274	3.28
q151	5	274	1.82
q152	4	274	1.46
q153	5	274	1.82
q154	4	274	1.46
q155	5	274	1.82
q156	6	274	2.19
q157	6	274	2.19
q158	6	274	2.19
q159	7	274	2.55
q160	6	274	2.19
q161	9	274	3.28
q162a	3	274	1.09
q162b	4	274	1.46
q162c	3	274	1.09
q162d	3	274	1.09
q162e	5	274	1.82
q162f	3	274	1.09
q163_1	232	274	84.67

Variable	Missing	Total Sample	Percent Missing
q163_2	192	274	70.07
q163_3	189	274	68.98
q163_4	169	274	61.68
q163_5	244	274	89.05
q163_6	248	274	90.51
q163_7	196	274	71.53
q163_8	192	274	70.07
q163_9	260	274	94.89
q163_10	249	274	90.88
q164	2	274	0.73
q165	2	274	0.73
q166_1	254	274	92.7
q166_2	256	274	93.43
q166_3	238	274	86.86
q166_4	234	274	85.4
q166_5	264	274	96.35
q166_6	253	274	92.34
q166_7	248	274	90.51
q166_8	248	274	90.51
q166_9	271	274	98.91
q166_10	254	274	92.7
q167	3	274	1.09
q168_1	265	274	96.72
q168_2	266	274	97.08
q168_3	258	274	94.16
q168_4	263	274	95.99
q168_5	268	274	97.81
q168_6	250	274	91.24
q168_7	264	274	96.35
q168_8	267	274	97.45
q168_9	271	274	98.91
q168_10	266	274	97.08
q169a	5	274	1.82
q169b	2	274	0.73
q169c	5	274	1.82
q169d	3	274	1.09
q169e	3	274	1.09

Variable	Missing	Total Sample	Percent Missing
q169f	4	274	1.46
q169g	4	274	1.46
q169h	5	274	1.82
q169i	3	274	1.09
q169j	2	274	0.73
q169k	2	274	0.73
q170_1	231	274	84.31
q170_2	240	274	87.59
q170_3	203	274	74.09
q170_4	201	274	73.36
q170_5	257	274	93.8
q170_6	231	274	84.31
q170_7	230	274	83.94
q170_8	241	274	87.96
q170_9	264	274	96.35
q170_10	230	274	83.94
q171_10	264	274	96.35
q171a	2	274	0.73
q171b	3	274	1.09
q171c	3	274	1.09
q171d	3	274	1.09
q171e	4	274	1.46
q171f	2	274	0.73
q171g	4	274	1.46
q171h	5	274	1.82
q171i	2	274	0.73
q172_1	217	274	79.2
q172_2	207	274	75.55
q172_3	168	274	61.31
q172_4	166	274	60.58
q172_5	234	274	85.4
q172_6	234	274	85.4
q172_7	206	274	75.18
q172_8	204	274	74.45
q172_9	252	274	91.97
q172_10	234	274	85.4
q173a	4	274	1.46

Variable	Missing	Total Sample	Percent Missing
q173b	2	274	0.73
q173c	4	274	1.46
q173d	4	274	1.46
q173e	3	274	1.09
q173f	4	274	1.46
q173g	2	274	0.73
q173h	3	274	1.09
q173i	4	274	1.46
q173j	4	274	1.46
q173k	4	274	1.46
q173l	3	274	1.09
q174	2	274	0.73
q175	2	274	0.73
q176	5	274	1.82
q177	3	274	1.09
q178	4	274	1.46
q179	3	274	1.09
q180	4	274	1.46
q181	4	274	1.46
q182	2	274	0.73
q183	33	274	12.04
q184	28	274	10.22
q185	19	274	6.93
q186	28	274	10.22
q187	28	274	10.22
q188	26	274	9.49
q189	3	274	1.09
q190_1	231	274	84.31
q190_2	217	274	79.2
q190_3	240	274	87.59
q190_4	164	274	59.85
q190_5	249	274	90.88
q190_6	233	274	85.04
q190_7	210	274	76.64
q190_8	130	274	47.45
q190_9	260	274	94.89
q190_10	241	274	87.96

Variable	Missing	Total Sample	Percent Missing
q191a	5	274	1.82
q191b	4	274	1.46
q191c	5	274	1.82
q191d	4	274	1.46
q191e	5	274	1.82
q191f	5	274	1.82
q191g	6	274	2.19
q191h	6	274	2.19
q191i	5	274	1.82
q192a	4	274	1.46
q192b	4	274	1.46
q192c	3	274	1.09
q192d	5	274	1.82
q192e	4	274	1.46
q192f	3	274	1.09
q192g	4	274	1.46
q192h	4	274	1.46
q192i	3	274	1.09
q192j	5	274	1.82
q192k	4	274	1.46
q192l	3	274	1.09
q193_1	252	274	91.97
q193_2	233	274	85.04
q193_3	218	274	79.56
q193_4	256	274	93.43
q193_5	224	274	81.75
q193_6	255	274	93.07
q193_t_verb	265	274	0
q194	54	274	19.71
q195	8	274	2.92
q196	2	274	0.73
q197	2	274	0.73
q198	2	274	0.73
q199	3	274	1.09
q200	2	274	0.73
q201_1	254	274	92.7
q201_2	263	274	95.99

Variable	Missing	Total Sample	Percent Missing
q201_3	263	274	95.99
q201_4	238	274	86.86
q202_1	165	274	60.22
q202_2	217	274	79.2
q202_3	237	274	86.5
q202_4	250	274	91.24
q202_5	265	274	96.72
q202_6	242	274	88.32
q202_7	223	274	81.39
q202_8	242	274	88.32
q202_9	261	274	95.26
q202_10	0	274	0
q205	9	274	3.28
q205_i	0	274	0
q207	8	274	2.92
q208	2	274	0.73
q209_1	183	274	66.79
q209_2	142	274	51.82
q209_3	246	274	89.78
q209_4	218	274	79.56
q209_5	248	274	90.51
q209_6	262	274	95.62
q209_7	270	274	98.54
q209_8	271	274	98.91
q209_9	272	274	99.27
q209_10	262	274	95.62
q209_11	272	274	99.27
q209_12	265	274	96.72
q210	1	274	0.36
q211	2	274	0.73
q212	2	274	0.73
q213	1	274	0.36
q214	3	274	1.09
q215	244	274	89.05
q216	256	274	93.43
q217	256	274	93.43
q218	1	274	0.36

Variable	Missing	Total Sample	Percent Missing
q219_1	273	274	99.64
q219_2	263	274	95.99
q219_3	229	274	83.58
q219_4	261	274	95.26
q219_5	268	274	97.81
q220	2	274	0.73
q221_1	48	274	17.52
q221_2	100	274	36.5
q221_3	268	274	97.81
q222	1	274	0.36
q223	1	274	0.36
q224	1	274	0.36
q225	1	274	0.36
q226	2	274	0.73
q227	2	274	0.73
q228	2	274	0.73
t1q27	223	274	81.39
t1q28	274	274	100
t2q28	58	274	21.17
t2q29	60	274	21.9
t2q30	146	274	53.28
t2q71	101	274	36.86
t2q72	137	274	50
t2q86	128	274	46.72
t2q87	56	274	20.44
t1q198	225	274	82.12
t1q200	225	274	82.12
t2q203_1	93	274	33.94
t2q203_2	238	274	86.86
t2q203_3	272	274	99.27
t2q204	67	274	24.45
t2q206	60	274	21.9
gann_inc	75	274	27.37
gann_inc2	256	274	93.43
gd74	67	274	24.45
gd75	67	274	24.45
gd76	267	274	97.45

Variable	Missing	Total Sample	Percent Missing
geducation	4	274	1.46
gemploy~2010	146	274	53.28
gmsaname	0	274	0
grace	5	274	1.82
gzipcode	6	274	2.19
gzipstate	2	274	0.73
gd7	58	274	21.17
sex	0	274	0
gender	0	274	0
trans	0	274	0
gruca	8	274	2.92
gmilesaway	7	274	2.55
gcendiv	2	274	0.73
gcenreg	2	274	0.73
age	1	274	0.36
race	0	274	0
race_recode	0	274	0
race_recod~5	0	274	0
sexualid	3	274	1.09
sexminid	3	274	1.09
hinc	8	274	2.92
pinc	11	274	4.01
hinc_i	0	274	0
pinc_i	0	274	0
geduc1	4	274	1.46
geduc2	4	274	1.46
gruca_i	0	274	0
q191a_i	0	274	0
q191b_i	0	274	0
q191c_i	0	274	0
q191d_i	0	274	0
q191e_i	0	274	0
q191f_i	0	274	0
q191g_i	0	274	0
q191h_i	0	274	0
q191i_i	0	274	0
gurban	8	274	2.92

Variable	Missing	Total Sample	Percent Missing
gurban_i	0	274	0
poverty	15	274	5.47
poverty_i	9	274	3.28
povertycat	15	274	5.47
povertycat_i	9	274	3.28
socialwb	7	274	2.55
socialwb_i	0	274	0
meim	5	274	1.82
meim_i	0	274	0
kessler6	1	274	0.36
kessler6_i	0	274	0
auditc	4	274	1.46
auditc_i	0	274	0
dudit	8	274	2.92
dudit_i	0	274	0
everyday	10	274	3.65
everyday_i	0	274	0
childgnc	93	274	33.94
childgnc_i	0	274	0
socsupport	10	274	3.65
socsupport_i	0	274	0
socsupport~o	5	274	1.82
socsuppo~o_i	0	274	0
socsupport~m	7	274	2.55
socsuppo~m_i	0	274	0
socsupport~r	4	274	1.46
socsuppo~r_i	0	274	0
lifesat	4	274	1.46
lifesat_i	0	274	0
connectedn~s	3	274	1.09
connectedn~i	0	274	0
hcthreat	3	274	1.09
hcthreat_i	0	274	0
internalized	11	274	4.01
internaliz~i	0	274	0
nondisclos~e	7	274	2.55
nondisclos~i	0	274	0

Variable	Missing	Total Sample	Percent Missing
nonaffirm	3	274	1.09
nonaffirm_i	0	274	0
expectrej	10	274	3.65
expectrej_i	0	274	0
ace_emo	19	274	6.93
ace_phy	28	274	10.22
ace_sex	25	274	9.12
ace_ipv	33	274	12.04
ace_sub	5	274	1.82
ace_men	4	274	1.46
ace_sep	2	274	0.73
ace_inc	4	274	1.46
ace_emo_i	0	274	0
ace_phy_i	0	274	0
ace_sex_i	0	274	0
ace_ipv_i	0	274	0
ace_sub_i	0	274	0
ace_men_i	0	274	0
ace_sep_i	0	274	0
ace_inc_i	0	274	0
ace	72	274	26.28
ace_i	0	274	0

Appendix 5: Missing values for each variable in Cisgender dataset.

Variable	Missing	Total Sample	Percent Missing
studyid	0	1,162	0
weight	0	1,162	0
gmethod_type	0	1,162	0
gresponden~e	0	1,162	0
gcenreg	0	1,162	0
age	0	1,162	0
race	0	1,162	0
race_recode	0	1,162	0
race_recod~5	0	1,162	0
sexualid	31	1,162	2.67
sexminid	31	1,162	2.67
sex	0	1,162	0
gender	0	1,162	0
hinc	32	1,162	2.75
pinc	105	1,162	9.04
geduc1	0	1,162	0
geduc2	0	1,162	0
grace_i	0	1,162	0
hinc_i	0	1,162	0
pinc_i	0	1,162	0
gruca_i	0	1,162	0
gurban	16	1,162	1.38
gurban_i	0	1,162	0
poverty	94	1,162	8.09
poverty_i	77	1,162	6.63
povertycat	94	1,162	8.09
povertycat_i	77	1,162	6.63
gd25a	0	1,162	0
gd65	0	1,162	0
gd66	438	1,162	37.69
gd66a	1,015	1,162	87.35
gp20	0	1,162	0
gann_inc	0	1,162	0
gann_inc2	1,038	1,162	89.33

Variable	Missing	Total Sample	Percent Missing
gd1	0	1,162	0
gd15c	526	1,162	45.27
gd4_1	0	1,162	0
gd5	0	1,162	0
gd63	0	1,162	0
gd69	1,162	1,162	100
gd69_1	0	1,162	0
gd69_2	27	1,162	2.32
gd69_3	1,122	1,162	96.56
gd69_4	1,160	1,162	99.83
gd69_5	1,162	1,162	100
gd74	0	1,162	0
gd75	0	1,162	0
gd76	1,063	1,162	91.48
gd8b	588	1,162	50.6
gd9	0	1,162	0
geducation	0	1,162	0
gemploy~2010	1,137	1,162	97.85
gmsaname	0	1,162	0
gp1	0	1,162	0
gp2	741	1,162	63.77
grace	13	1,162	1.12
gsurvey	0	1,162	0
gwp10200	1,137	1,162	97.85
gwp10202	1,146	1,162	98.62
gwp10208	1,149	1,162	98.88
gwp10209	1,149	1,162	98.88
gwp10215	1,152	1,162	99.14
gwp10216	1,159	1,162	99.74
gwp119	588	1,162	50.6
gwp1223	0	1,162	0
gwp1225	1,150	1,162	98.97
ggzipcode	5	1,162	0.43
gzipstate	0	1,162	0
gruca	16	1,162	1.38
gmilesaway	8	1,162	0.69
gcendiv	0	1,162	0
q01	9	1,162	0.77

Variable	Missing	Total Sample	Percent Missing
q02	91	1,162	7.83
q03	73	1,162	6.28
q04	11	1,162	0.95
q05	10	1,162	0.86
q06	15	1,162	1.29
q07	16	1,162	1.38
q08	12	1,162	1.03
q09	13	1,162	1.12
q10	13	1,162	1.12
q11	12	1,162	1.03
q12	13	1,162	1.12
q13	9	1,162	0.77
q14	23	1,162	1.98
q15	31	1,162	2.67
q16	27	1,162	2.32
q17	29	1,162	2.5
q18	26	1,162	2.24
q19a	32	1,162	2.75
q19b	50	1,162	4.3
q19c	67	1,162	5.77
q19d	46	1,162	3.96
q20_1	1,127	1,162	96.99
q20_2	1,087	1,162	93.55
q20_3	1,099	1,162	94.58
q20_4	1,148	1,162	98.8
q20_5	1,155	1,162	99.4
q20_6	144	1,162	12.39
q20_7	1,111	1,162	95.61
q21	1,162	1,162	100
q21_verb	0	1,162	0
q22	30	1,162	2.58
q23	30	1,162	2.58
q24	32	1,162	2.75
q25	31	1,162	2.67
q26	27	1,162	2.32
q27	27	1,162	2.32
q28	20	1,162	1.72
q29	34	1,162	2.93

Variable	Missing	Total Sample	Percent Missing
q30	1,162	1,162	100
q31	38	1,162	3.27
q32	34	1,162	2.93
q32_verb	0	1,162	0
q33_1	694	1,162	59.72
q33_2	738	1,162	63.51
q33_3	1,157	1,162	99.57
q33_4	1,156	1,162	99.48
q33_5	908	1,162	78.14
q34a	72	1,162	6.2
q34b	77	1,162	6.63
q34c	103	1,162	8.86
q34d	106	1,162	9.12
q34e	110	1,162	9.47
q34f	114	1,162	9.81
q35	34	1,162	2.93
q36	361	1,162	31.07
q37	357	1,162	30.72
q38	355	1,162	30.55
q39	353	1,162	30.38
q40	24	1,162	2.07
q41	28	1,162	2.41
q42	31	1,162	2.67
q43	706	1,162	60.76
q44_1	691	1,162	59.47
q44_2	1,072	1,162	92.25
q45_1	1,115	1,162	95.96
q45_2	735	1,162	63.25
q45_3	985	1,162	84.77
q45_4	1,120	1,162	96.39
q45_5	1,156	1,162	99.48
q45_6	1,116	1,162	96.04
q45_7	1,048	1,162	90.19
q45_8	713	1,162	61.36
q45_9	1,081	1,162	93.03
q45_10	1,122	1,162	96.56
q45_11	1,105	1,162	95.09
q45_12	1,161	1,162	99.91

Variable	Missing	Total Sample	Percent Missing
q45_13	1,121	1,162	96.47
q45_verb	0	1,162	0
q46	31	1,162	2.67
q47_1	797	1,162	68.59
q47_2	312	1,162	26.85
q47_3	957	1,162	82.36
q47_4	1,065	1,162	91.65
q47_5	1,116	1,162	96.04
q48	127	1,162	10.93
q49	24	1,162	2.07
q50	238	1,162	20.48
q50_verb	0	1,162	0
q51	29	1,162	2.5
q52	26	1,162	2.24
q53a	113	1,162	9.72
q53b	166	1,162	14.29
q53c	172	1,162	14.8
q53d	75	1,162	6.45
q54	21	1,162	1.81
q55	31	1,162	2.67
q56	31	1,162	2.67
q57	43	1,162	3.7
q58_1	679	1,162	58.43
q58_2	774	1,162	66.61
q58_3	993	1,162	85.46
q58_4	1,138	1,162	97.93
q58_5	1,110	1,162	95.52
q58_6	1,130	1,162	97.25
q58_7	1,139	1,162	98.02
q58_8	1,046	1,162	90.02
q58_9	1,117	1,162	96.13
q58_10	1,015	1,162	87.35
q58_11	1,006	1,162	86.57
q58_12	1,010	1,162	86.92
q58_13	800	1,162	68.85
q58_14	1,123	1,162	96.64
q58_15	1,045	1,162	89.93
q58_16	985	1,162	84.77

Variable	Missing	Total Sample	Percent Missing
q58_17	1,143	1,162	98.36
q58_18	1,111	1,162	95.61
q58_19	1,143	1,162	98.36
q58_20	1,124	1,162	96.73
q58_21	1,159	1,162	99.74
q58_22	1,140	1,162	98.11
q58_23	949	1,162	81.67
q59	23	1,162	1.98
q60	22	1,162	1.89
q61a	28	1,162	2.41
q61b	30	1,162	2.58
q61c	30	1,162	2.58
q61d	27	1,162	2.32
q61e	27	1,162	2.32
q61f	27	1,162	2.32
q62	27	1,162	2.32
q63	30	1,162	2.58
q64	26	1,162	2.24
q65	26	1,162	2.24
q66	43	1,162	3.7
q67	60	1,162	5.16
q68	74	1,162	6.37
q69	72	1,162	6.2
q70	77	1,162	6.63
q71	43	1,162	3.7
q72	41	1,162	3.53
q73	35	1,162	3.01
q74	33	1,162	2.84
q75	33	1,162	2.84
q76	38	1,162	3.27
q77	656	1,162	56.45
q78	36	1,162	3.1
q79	39	1,162	3.36
q80	38	1,162	3.27
q81	38	1,162	3.27
q82	37	1,162	3.18
q83	43	1,162	3.7
q84	39	1,162	3.36

Variable	Missing	Total Sample	Percent Missing
q85	41	1,162	3.53
q86	41	1,162	3.53
q87	42	1,162	3.61
q88	43	1,162	3.7
q89	37	1,162	3.18
q90	1,003	1,162	86.32
q91	978	1,162	84.17
q92	977	1,162	84.08
q93	41	1,162	3.53
q94	1,084	1,162	93.29
q95	1,103	1,162	94.92
q96	1,104	1,162	95.01
q97	51	1,162	4.39
q98	1,032	1,162	88.81
q99	1,039	1,162	89.41
q100	1,041	1,162	89.59
q101	48	1,162	4.13
q102	1,149	1,162	98.88
q103	1,095	1,162	94.23
q104	1,149	1,162	98.88
q105	1,149	1,162	98.88
q106	1,084	1,162	93.29
q107	45	1,162	3.87
q108	1,117	1,162	96.13
q109	1,116	1,162	96.04
q110	1,117	1,162	96.13
q111a	93	1,162	8
q111b	92	1,162	7.92
q111c	102	1,162	8.78
q111d	94	1,162	8.09
q111e	116	1,162	9.98
q111f	121	1,162	10.41
q112a	41	1,162	3.53
q112b	36	1,162	3.1
q112c	44	1,162	3.79
q112d	42	1,162	3.61
q112e	39	1,162	3.36
q112f	42	1,162	3.61

Variable	Missing	Total Sample	Percent Missing
q113_1	1,006	1,162	86.57
q113_2	937	1,162	80.64
q113_3	1,162	1,162	100
q113_4	1,146	1,162	98.62
q113_5	1,081	1,162	93.03
q113_6	1,070	1,162	92.08
q113_7	1,131	1,162	97.33
q113_8	1,036	1,162	89.16
q113_9	1,137	1,162	97.85
q113_10	1,141	1,162	98.19
q114	40	1,162	3.44
q115	38	1,162	3.27
q116_1	1,061	1,162	91.31
q116_2	1,068	1,162	91.91
q116_3	1,161	1,162	99.91
q116_4	1,157	1,162	99.57
q116_5	1,122	1,162	96.56
q116_6	1,084	1,162	93.29
q116_7	1,153	1,162	99.23
q116_8	1,125	1,162	96.82
q116_9	1,142	1,162	98.28
q116_10	1,131	1,162	97.33
q117	39	1,162	3.36
q118_1	1,148	1,162	98.8
q118_2	1,154	1,162	99.31
q118_3	1,161	1,162	99.91
q118_4	1,161	1,162	99.91
q118_5	1,147	1,162	98.71
q118_6	1,122	1,162	96.56
q118_7	1,160	1,162	99.83
q118_8	1,162	1,162	100
q118_9	1,162	1,162	100
q118_10	1,161	1,162	99.91
q119a	38	1,162	3.27
q119b	44	1,162	3.79
q119c	43	1,162	3.7
q119d	44	1,162	3.79
q119e	49	1,162	4.22

Variable	Missing	Total Sample	Percent Missing
q119f	50	1,162	4.3
q119g	45	1,162	3.87
q119h	39	1,162	3.36
q119i	46	1,162	3.96
q119j	46	1,162	3.96
q119k	42	1,162	3.61
q120_1	1,070	1,162	92.08
q120_2	1,102	1,162	94.84
q120_3	1,162	1,162	100
q120_4	1,159	1,162	99.74
q120_5	1,128	1,162	97.07
q120_6	1,083	1,162	93.2
q120_7	1,158	1,162	99.66
q120_8	1,131	1,162	97.33
q120_9	1,148	1,162	98.8
q120_10	1,130	1,162	97.25
q121a	43	1,162	3.7
q121b	46	1,162	3.96
q121c	45	1,162	3.87
q121d	44	1,162	3.79
q121e	46	1,162	3.96
q121f	43	1,162	3.7
q121g	47	1,162	4.04
q121h	50	1,162	4.3
q121i	41	1,162	3.53
q122_1	937	1,162	80.64
q122_2	984	1,162	84.68
q122_3	1,162	1,162	100
q122_4	1,149	1,162	98.88
q122_5	1,053	1,162	90.62
q122_6	1,033	1,162	88.9
q122_7	1,150	1,162	98.97
q122_8	1,037	1,162	89.24
q122_9	1,123	1,162	96.64
q122_10	1,116	1,162	96.04
q123a	43	1,162	3.7
q123b	43	1,162	3.7
q123c	48	1,162	4.13

Variable	Missing	Total Sample	Percent Missing
q123d	45	1,162	3.87
q123e	50	1,162	4.3
q123f	47	1,162	4.04
q123g	44	1,162	3.79
q123h	47	1,162	4.04
q123i	46	1,162	3.96
q123j	45	1,162	3.87
q123k	47	1,162	4.04
q123l	44	1,162	3.79
q124	47	1,162	4.04
q125	52	1,162	4.48
q126	50	1,162	4.3
q127	51	1,162	4.39
q128	46	1,162	3.96
q129	48	1,162	4.13
q130	49	1,162	4.22
q131	50	1,162	4.3
q132	45	1,162	3.87
q133	130	1,162	11.19
q134	78	1,162	6.71
q135	147	1,162	12.65
q136	84	1,162	7.23
q137	83	1,162	7.14
q138	81	1,162	6.97
q139	45	1,162	3.87
q140_1	942	1,162	81.07
q140_2	1,048	1,162	90.19
q140_3	1,162	1,162	100
q140_4	1,134	1,162	97.59
q140_5	1,102	1,162	94.84
q140_6	1,069	1,162	92
q140_7	1,140	1,162	98.11
q140_8	816	1,162	70.22
q140_9	1,133	1,162	97.5
q140_10	1,139	1,162	98.02
q141a	50	1,162	4.3
q141b	50	1,162	4.3
q141c	48	1,162	4.13

Variable	Missing	Total Sample	Percent Missing
q141d	53	1,162	4.56
q141e	51	1,162	4.39
q141f	49	1,162	4.22
q141g	47	1,162	4.04
q141h	50	1,162	4.3
q141i	51	1,162	4.39
q141j	54	1,162	4.65
q141k	52	1,162	4.48
q141l	49	1,162	4.22
q142	48	1,162	4.13
q143_1	1,148	1,162	98.8
q143_2	1,091	1,162	93.89
q143_3	1,097	1,162	94.41
q143_4	898	1,162	77.28
q143_5	989	1,162	85.11
q143_6	1,154	1,162	99.31
q143_7	906	1,162	77.97
q144	76	1,162	6.54
q145	70	1,162	6.02
q146	55	1,162	4.73
q147	52	1,162	4.48
q47_verb	0	1,162	0
q148	50	1,162	4.3
q149	44	1,162	3.79
q150	49	1,162	4.22
q151	50	1,162	4.3
q152_1	949	1,162	81.67
q152_2	1,146	1,162	98.62
q152_3	1,019	1,162	87.69
q152_4	594	1,162	51.12
q153_1	815	1,162	70.14
q153_2	1,046	1,162	90.02
q153_3	1,031	1,162	88.73
q153_4	1,126	1,162	96.9
q153_5	1,151	1,162	99.05
q153_6	1,085	1,162	93.37
q153_7	1,132	1,162	97.42
q153_8	716	1,162	61.62

Variable	Missing	Total Sample	Percent Missing
q153_9	1,107	1,162	95.27
q153_10	1,136	1,162	97.76
q153_verb	0	1,162	0
q154_1	128	1,162	11.02
q154_2	1,089	1,162	93.72
q154_3	1,154	1,162	99.31
q155	95	1,162	8.18
q156	77	1,162	6.63
q156_i	0	1,162	0
q157	105	1,162	9.04
q158	92	1,162	7.92
q159	54	1,162	4.65
q160_1	345	1,162	29.69
q160_2	931	1,162	80.12
q160_3	1,109	1,162	95.44
q160_4	1,104	1,162	95.01
q160_5	1,139	1,162	98.02
q160_6	1,158	1,162	99.66
q160_7	1,160	1,162	99.83
q160_8	1,161	1,162	99.91
q160_9	1,156	1,162	99.48
q160_10	1,154	1,162	99.31
q160_11	1,162	1,162	100
q160_12	1,138	1,162	97.93
q161	48	1,162	4.13
q162	52	1,162	4.48
q163	52	1,162	4.48
q164	48	1,162	4.13
q165	56	1,162	4.82
q166	993	1,162	85.46
q167	54	1,162	4.65
q168_1	1,161	1,162	99.91
q168_2	1,162	1,162	100
q168_3	1,159	1,162	99.74
q168_4	1,162	1,162	100
q168_5	1,162	1,162	100
q169	59	1,162	5.08
q170_1	165	1,162	14.2

Variable	Missing	Total Sample	Percent Missing
q170_2	355	1,162	30.55
q170_3	1,068	1,162	91.91
q171	47	1,162	4.04
q172	48	1,162	4.13
q173	47	1,162	4.04
q174	49	1,162	4.22
q175	46	1,162	3.96
q176	46	1,162	3.96
q177	44	1,162	3.79
socialwb	87	1,162	7.49
socialwb_i	0	1,162	0
meim	45	1,162	3.87
meim_i	0	1,162	0
kessler6	47	1,162	4.04
kessler6_i	0	1,162	0
auditc	37	1,162	3.18
auditc_i	0	1,162	0
dudit	69	1,162	5.94
dudit_i	0	1,162	0
everyday	69	1,162	5.94
everyday_i	0	1,162	0
childgnc	410	1,162	35.28
childgnc_i	0	1,162	0
socsupport	87	1,162	7.49
socsupport_i	0	1,162	0
socsupport~o	65	1,162	5.59
socsuppo~o_i	0	1,162	0
socsupport~m	63	1,162	5.42
socsuppo~m_i	0	1,162	0
socsupport~r	59	1,162	5.08
socsuppo~r_i	0	1,162	0
ace_emo	147	1,162	12.65
ace_phy	78	1,162	6.71
ace_sex	85	1,162	7.31
ace_ipv	130	1,162	11.19
ace_sub	53	1,162	4.56
ace_men	46	1,162	3.96
ace_sep	45	1,162	3.87

Variable	Missing	Total Sample	Percent Missing
ace_inc	50	1,162	4.3
ace_emo_i	0	1,162	0
ace_phy_i	0	1,162	0
ace_sex_i	0	1,162	0
ace_ipv_i	0	1,162	0
ace_sub_i	0	1,162	0
ace_men_i	0	1,162	0
ace_sep_i	0	1,162	0
ace_inc_i	0	1,162	0
ace	266	1,162	22.89
ace_i	0	1,162	0
lifesat	57	1,162	4.91
lifesat_i	0	1,162	0

Appendix 6. Creating an analysis plan and regression analysis using point-and-click method in SPSS

1) Click on Analyze -> Complex Samples -> Prepare for Analysis
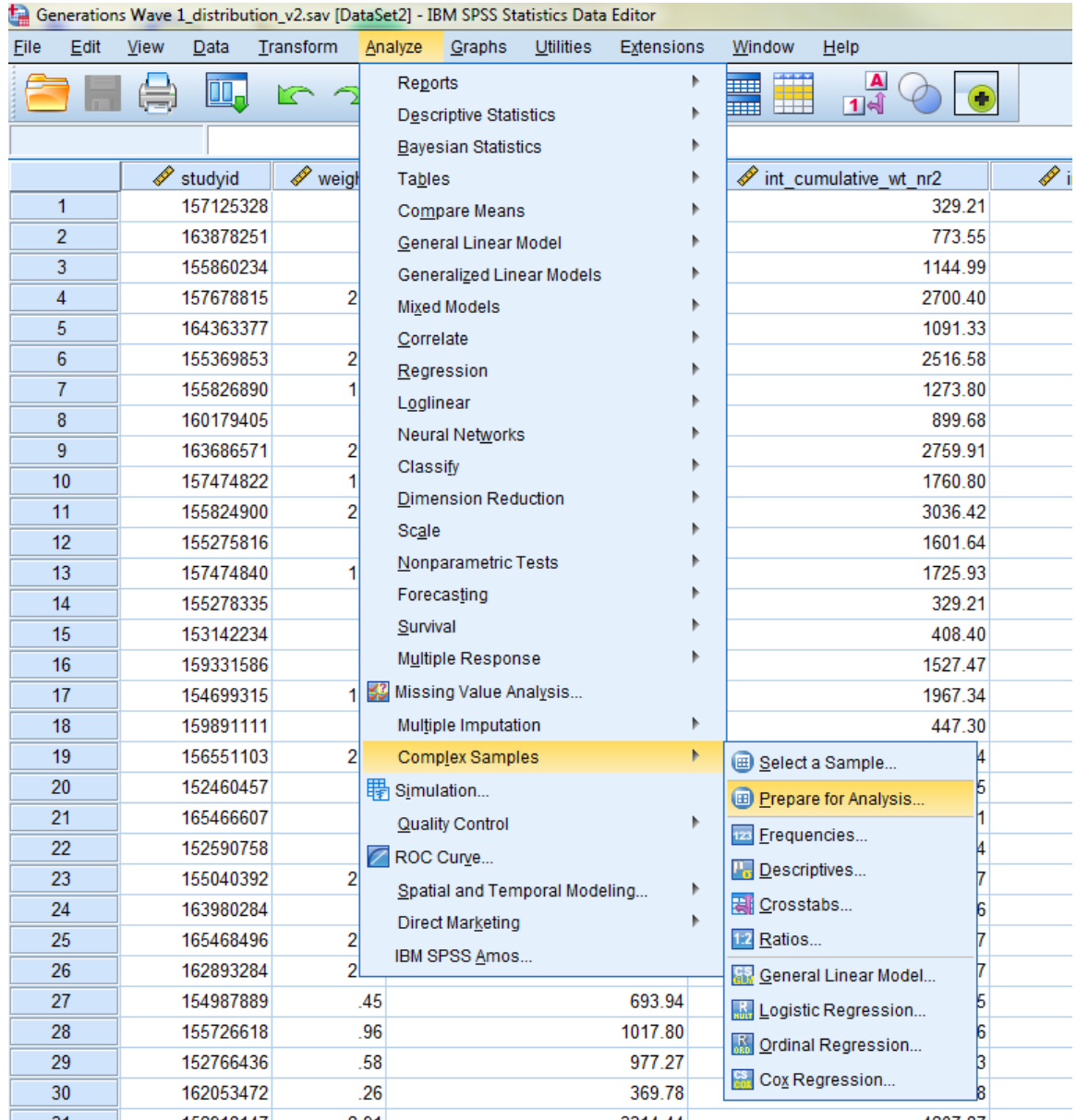
2) When the "Analysis Preparation Wizard" window appears, make sure that "Create a plan file" is selected and then click on "Next"

3) You will be creating a file (a .csaplan file) that contains the survey design information, which SPSS will use in its statistical modeling. For now, you just need to enter a name for this file. I named mine "Generations Wave 1 survey design". Click "Save" after entering a name.
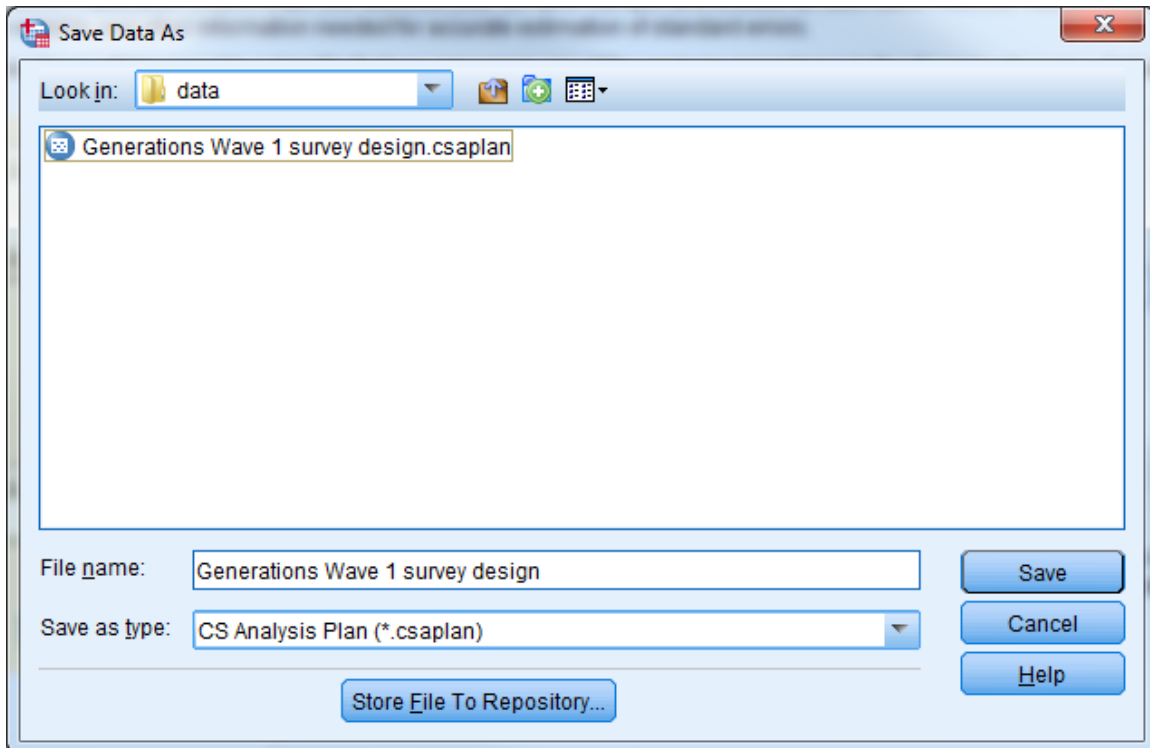
4) A window called "Analysis Preparation Wizard" should appear. Here you will specify the weighting variable, by moving the "weight" variable into "Sample Weight". Then you hit "Finish":
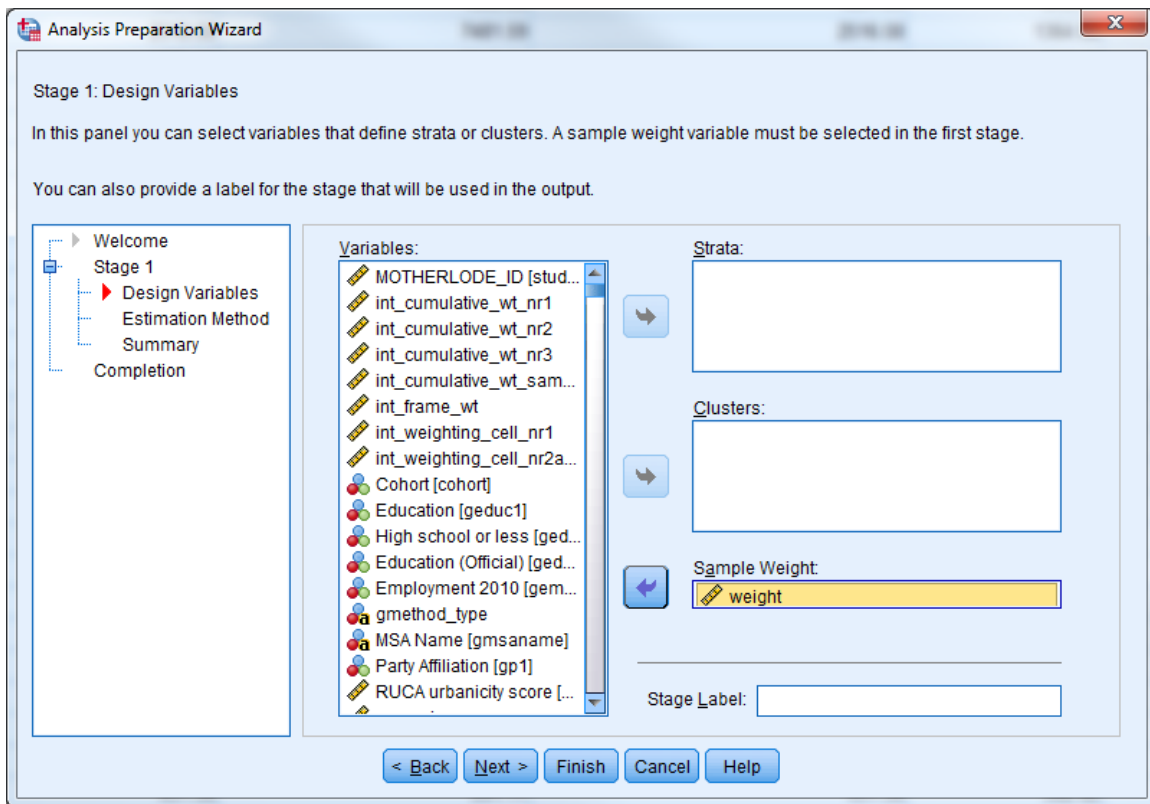

5) The survey design file has now been created. You can reuse it whenever you need to do a weighted analysis with this dataset.

6) Only the procedures listed under "Complex Design" can use this survey design file, so you are limited in the analyses available.

7) Let's do a linear regression. First select Complex Samples > General Linear Model. A new window will appear called "Complex Samples Plan for General Linear Model". In the "Plan" section of this window, you should see the survey design file that we just created. If you later reopen SPSS after shutting down, you'll need to Browse to find this file again (or just create it again, it's so easy). If the survey design file is specified, hit "Continue".
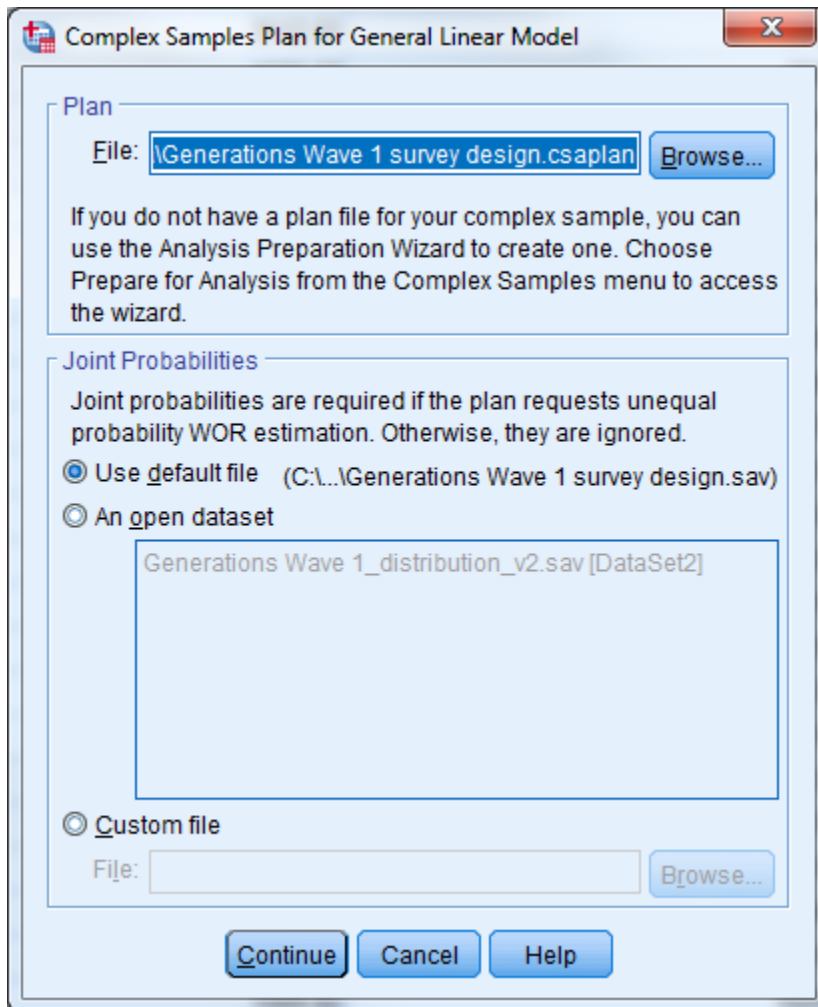
8) Enter regression variables as you normally would. Remember that "Factors" is for categorical predictors (last group will be the reference) and "Covariates" is for continuous predictors.   You can also specify a subpopulation filtering variable as well.   Also, click on "Statistics" and then select "Estimate", "Standard Error", "Confidence Interval", and "t-test", for standard regression output. Then click "Ok" when ready to run.
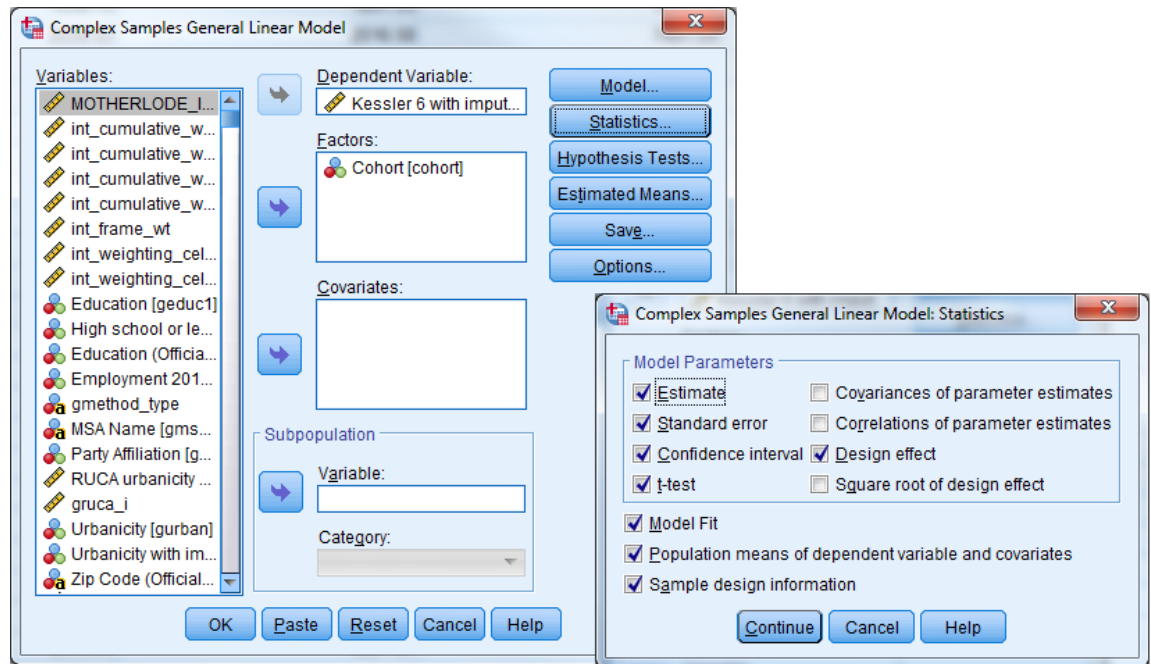
9) In the output, you should see new output like new tables called "Sample Design Information" and "Factor Information", with columns that use the word "Weighted". This will let you know that SPSS is using the survey design information in the statistical model.

Sample Design Information

		N
Unweighted Cases	Valid	1331
	Invalid	14
	Total	1345
Population Size		1326.795
Stage 1	Strata	1
	Units	1331
Sampling Design Degrees of Freedom		1330

Variable Information

	Mean

Dependent Variable	Kessler 6 with imputation	14.66

Factor Information

		Weighted Count	Weighted Percent
Cohort	Younger	805.159	60.7%
	Middle	281.642	21.2%
	Older	239.993	18.1%
Population Size		1326.795	100.0%

10) The regression parameter estimates table will look familiar, but may have an extra column called "Design Effect", which will be filled if the option is chosen.

Parameter Estimates[a]

Parameter	Estimate	Std. Error	95% Confidence Interval		Hypothesis Test			Design Effect
			Lower	Upper	t	df	Sig.	
(Intercept)	11.345	.255	10.845	11.845	44.496	1330.000	.000	.
[cohort=1]	4.726	.363	4.014	5.437	13.029	1330.000	.000	.
[cohort=2]	2.112	.471	1.188	3.037	4.482	1330.000	.000	.
[cohort=3]	.000[b]	.	.	.	.	.	.	.

a. Model: Kessler 6 with imputation = (Intercept) + cohort

b. Set to zero because this parameter is redundant.