



# DATATHON

Kick off:

Sat. Jan. 22nd, 12pm - 2pm EST

Presentations:

Sat. Jan. 29th, 12pm - 3pm EST

virtual

in support of




# #datathon Agenda

1. Overview of Data for Good
2. What to expect at a Datathon
3. Business problem
4. Overview of the data
5. Analysis Tips & Considerations

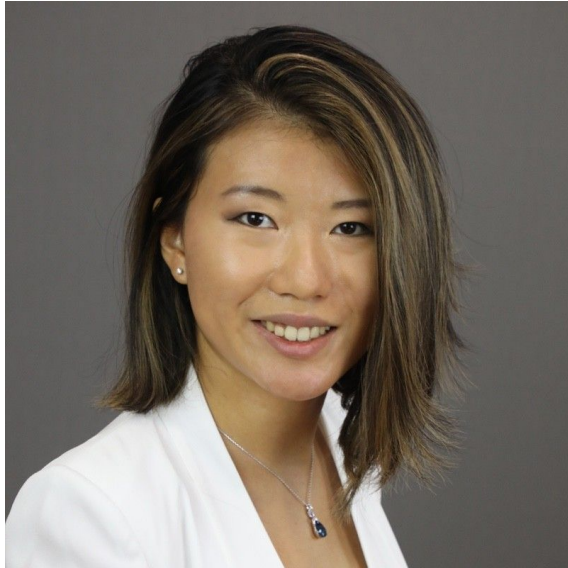


# What is dataforgood?



**We are a collective of passionate socially minded people who are empowering social change makers to be better by allowing their data to speak to them.**

# Data Ambassadors



**Helena Yu**



**Brendon Paul**



# Statistics Without Borders



**Emily Harville**  
Associate professor at  
Tulane University



**Cindy Tong**  
Director at Johnson &  
Johnson



**Rebecca Horne**  
PhD Candidate in  
Social Psychology



**Savita Ramaprasad**  
PhD in Econometrics

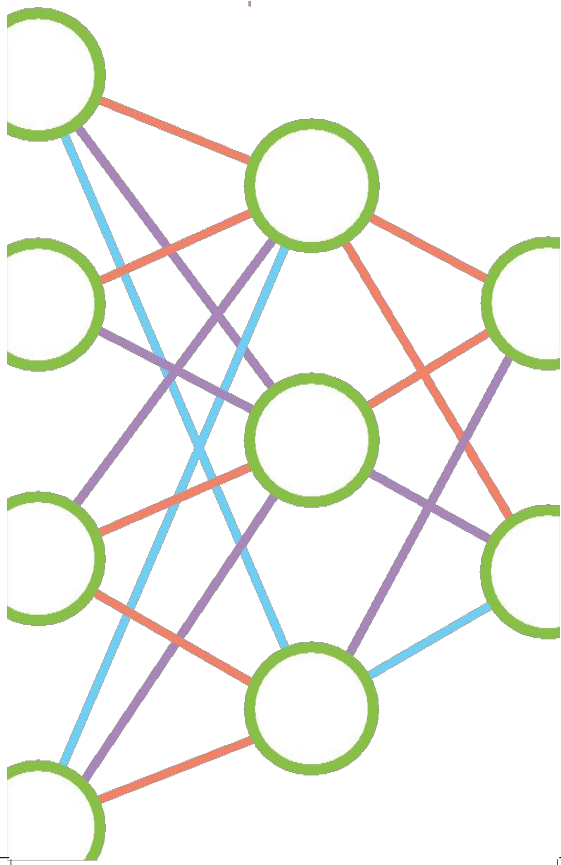
# Don't forget to follow us!



@dataforgoodwr



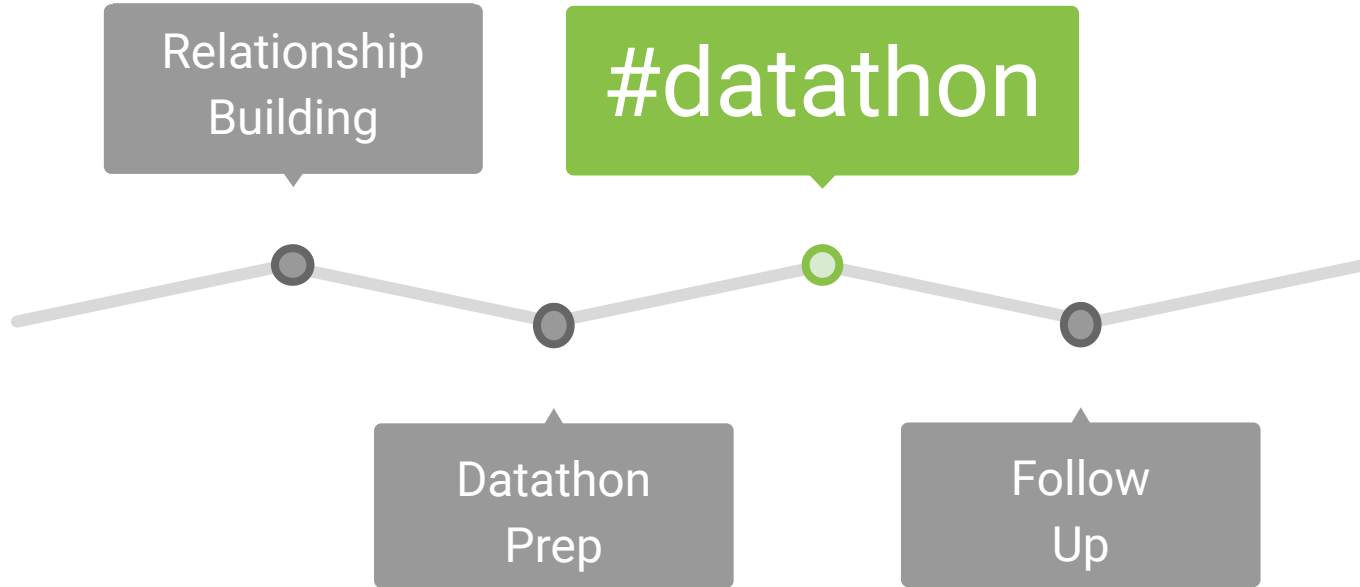
[www.linkedin.com/company/data-for-good-waterloo-region/](http://www.linkedin.com/company/data-for-good-waterloo-region/)



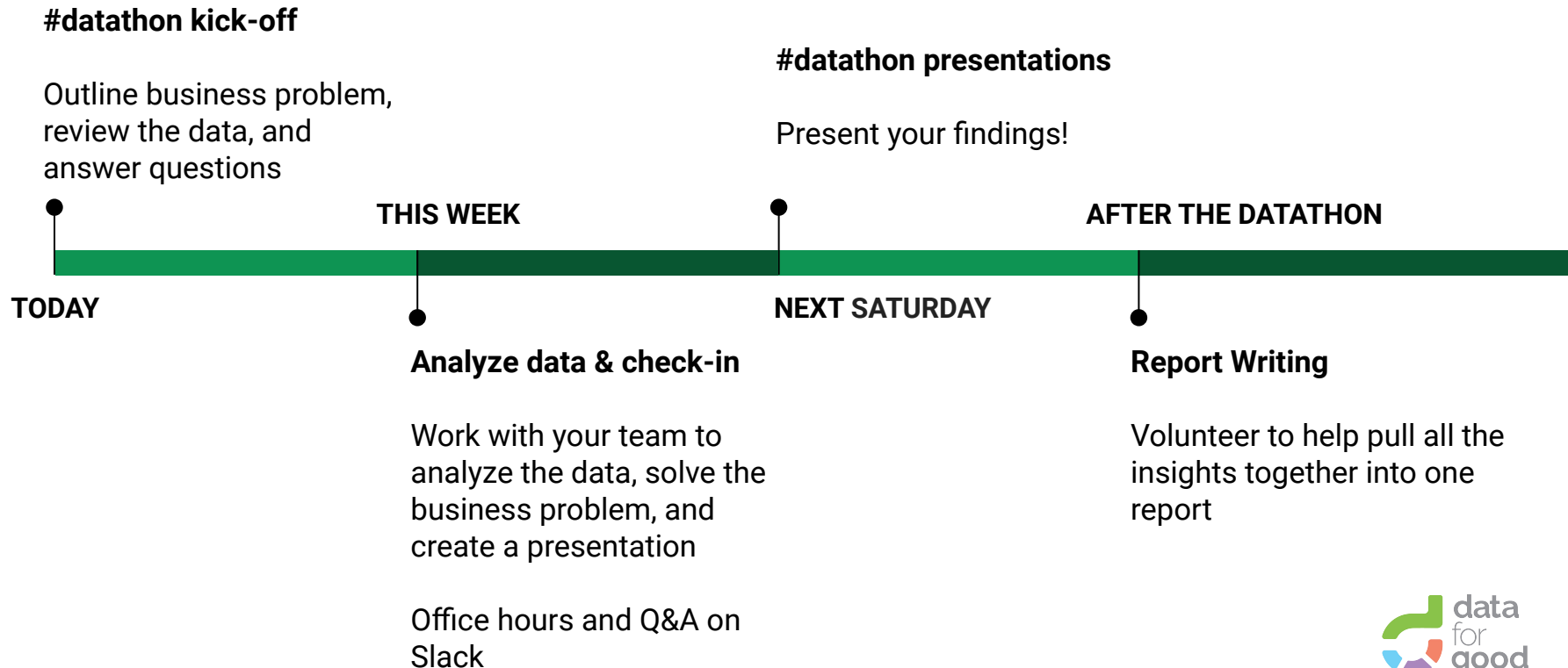
# Introduction to #datathon



# #datathon context



# #datathon timeline



# #datathon during the week

- We have set-up a slack channel for the event here, feel free to post questions throughout the week:  
<https://app.slack.com/client/TCSECUNQo/Co2TDBF98BZ>
- If you are working in a team, it is up to you & your team members to set up a method of communication (slack, email, google meet, etc.)
- Office hours are on google meet this week:
  - <https://meet.google.com/cta-tniw-ffg?hs=122&authuser=1>
  - Monday Jan 24 12pm-1pm EST (Melissa, Emily) & 8pm-9pm EST (Jeffrey) EST
  - Tuesday Jan 25 12pm-1pm EST (Sasha) & 6pm-7pm EST (Brendon, Cindy) EST
  - Wednesday Jan 26 8-9pm EST (Nadia)

# #datathon presentations

- **Upload your presentation** by **1130am EST on Saturday January 29th** to this google drive folder:  
<https://drive.google.com/drive/folders/1XGBuO17otgUeW6Vq8t72u7p2e4VINtou?usp=sharing>
  - Please submit details of methodology, sample, code, data visualization files as supplementary materials for review
- Keep your presentation business (Michelle) focused, and avoid technical details. Include assumptions & methodology as appendix content.
- Time limit will be maximum 5 minutes
- A sample presentation template can be found [here](#).
- We will organize the presentation order and let you know next Saturday



# Business Context



## Our Mission

Our mission is to mobilise and assemble the most impactful actions to defend women's human rights globally.





## Global Unity Principles

The Women's March Global Unity Principles provide a fleshed out framework for our vision. They are the guiding principles for the international movement and are the foundations from which we operate.

The Women's March on Washington convened a broad and diverse group of leaders to produce an intersectional platform known as the Unity Principles. Women's March Global Community organisers worked to ensure that the Unity Principles functioned across nations and borders. Representing a new understanding of the connected nature of our struggles and a vision of our collective liberation, the Unity Principles continue to be a framework for our movement.



## Global Unity Principles

- End Violence
- Reproductive Rights
- LGBTI+ Rights
- Workers' Rights
- Civil Rights
- Disability Rights
- Immigrant Rights
- Indigenous Rights
- Climate Justice

*for full details:* <https://womensmarchglobal.org/about/>



#### **THE GLOBAL NETWORK**

A safe and open environment for a global community of local activists to unite, coordinate, and mobilise.



#### **EDUCATION, TRAINING, + SUPPORT**

Free accessible resources for anyone, anywhere, to learn about the issues and discover ways to take action in their local communities.



#### **GRASSROOTS CAMPAIGNS + COALITIONS**

Elevate, amplify, and facilitate grassroots-led campaigns and build cross-sector coalitions to defend women's human rights.



WOMEN'S  
MARCH GLOBAL

WE  
COUNT.

# What is the Global Count

Every single woman counts, and should have the opportunity to unite and be counted at this turning point in history.

The '**Global Count**', a project of **Women's March Global** and a major ongoing campaign from the leading women's rights organisations that is polling women and gender-diverse people worldwide on the cultural, social, political, and economic barriers to their advancement.

The campaign aims to redefine how women's opinions, needs, and perspectives are taken into consideration by governments, institutions and international organisations. At the core of this campaign is the coming together and collaboration of like-minded movements and networks to centre the voices of women in the Global Reset and ensure that this data is extensive and representative enough to inform future efforts to progress gender equality.

**In 2021**, Women's March Global partnered with nearly a hundred movements and organisations to democratise the project of global gender equality in an unprecedented collective action: **Global Count**.

**Global Count** is a project that maps critical issues facing women and gender diverse people in order to redefine how women's opinions, needs, and perspectives inform policy agendas, and guide decision-making by governments, institutions, and international organisations.

**Global Count asked:**

1. What issues are most critical for women and gender-diverse people?
2. Which barriers to progress on these issues do they experience?
3. Which organisations are helpful in realising progress on these issues and barriers?
4. What does progress for women's human rights look like for you in 10 years?



# Peer Reviewers:

Adrianne Bradford, University of Maryland

Alena Sharday Maze, University of Maryland

Alexandra Pittman, Impact Mapper

Kimberlé Crenshaw, African American Policy Forum (AAPF)

Sarah Dickins, Plan International

# STEERING COMMITTEE



FACEBOOK



# STRATEGIC PARTNERS



# SUPPORTING PARTNERS



**Global Count is an opportunity to ensure women and gender-diverse people themselves define what is needed on policy and funding decisions that impact them, using gendered data.**

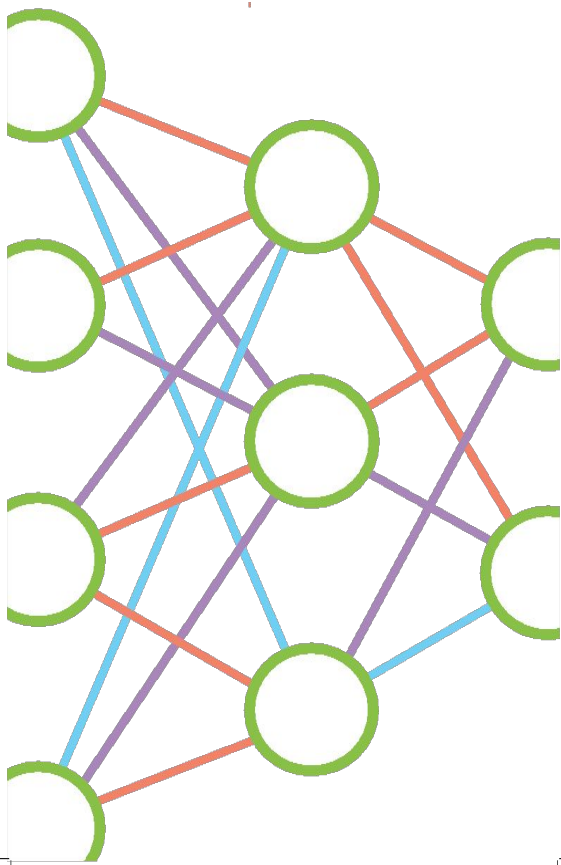
# Survey Description

- Survey was promoted using Facebook organic and advertising
  - Targeted to women and gender diverse people, by country, by language
- Survey captured using Momentiv Enterprise (SurveyMonkey) - online only
- Three survey versions:
  - Long
  - Medium (Mid)
  - Short
- Acquisition Period for dataset: Jan 21 - June 1, 2021
- Survey offered in 17 languages.
- Anonymous: No personally identifiable information was captured
- Survey captured using Momentiv Enterprise (SurveyMonkey) - online only
- ~32K survey responses were collected. ~8K were removed because they were blank.



# Survey Questions

- Demographics: Age, Gender Identity, Ethnic Identity, Country, City/Town
- Group: Personal
  - Which three issues are the most critical for you as a person?
  - For each issue area - which sector is the greatest barrier to progress in advocating or campaigning for issue XXXXX
  - What are up to three locally based organisations that you know of that are organising around these critical issues?
- Group: Regional (Long form only)
  - Which three issues most impact the city/town in which you live?
  - For this issue area - which sector is the greatest barrier to progress in advocating or campaigning for issue XXXXX
  - What are up to three locally based organisations that you know of that are organising around these critical issues?
- What does progress for women's human rights look like for you in 10 years?



# Business Problem

# Business Problem

## Descriptive Analysis for Critical Issues - General

1. *What are the most critical issues that survey respondents face*
  - 1a) globally
  - 1b) within each identity category (e.g. by country, gender identity, racial identity and by age)
  - 1c) within multiple identity categories (e.g. black transgender, young adult gender non-conforming/non-binary, older adult biracial/mixed race women, Asian transgender)

Global Count is also interested in understanding insights like: “For participants that reported Violence, Harassment, and Abuse as the most critical issue they face, what was the second most critical issue they faced?”

# Business Problem

## Descriptive Analysis for Critical Issues - Sub-Groups

2. *In which (sub)groups of survey respondents were certain issues reported as most critical?*

For example:

- In which *age/gender identity/racial identity/country group* was Sexual, Reproductive and Parental Health reported as the most critical issue?
- Which *racial identity + gender identity group combination* reported Sexual, Reproductive, and Parental Health as the most critical issue they were facing?
- Which (if any) racial identity group reported that Racial Equity was a critical issue that their local community faced, but not a critical issue that they personally faced?

# Business Problem

## Descriptive Analysis for Critical Issues - Combinations of Issues

*3. How often are certain combinations of critical issues reported by survey respondents?*

For example:

- What % of respondents reported both i) Racial Equity; and ii) Workers' Rights as 2 of the top 3 critical issues that they face?
- What % of respondents reported (i) Racial Equity, (ii) Economic and Workers' Rights, and (iii) Ending Violence, Harassment, and Abuse as the top 3 critical issues that they face?
- What % of respondents selected both "Ending Violence, Harassment, and Abuse" and "Reproductive Freedom and Bodily Autonomy"?

# Business Problem

## Descriptive Analysis for Barriers to Progress

*4. What were the biggest barriers to progress in advocating for specific issues, based on the survey responses?*

For example:

- What percentage of respondents reported political/cultural/social/education/technological barriers to ending Violence, Harassment, and Abuse?
- What was the primary barrier to advocating for Reproductive Freedom and Bodily Autonomy?
- What primary barriers do transgender men identify in advocating for LGBTQI+ Rights?
- What primary barriers do Hispanic or Latinx women identify in advocating for Racial Equality?



# Business Problem

## Text Analysis of Open-Ended Response Data on Women's Human Rights Progress

*5a. What were the most frequent themes that emerged from survey responses to what progress for women's human rights looks like in 10 years?*

*For example:*

- What can open-text answers tell us about FGM, child marriage, digital or online issues, sex trafficking or sexual violence?*
- What can open-text answers tell us about the role of the private sector to address job loss and livelihood burden on women around the world?*

*5b. How do these themes differ by segments of the survey population (age, gender, etc.)?*

# Business Problem

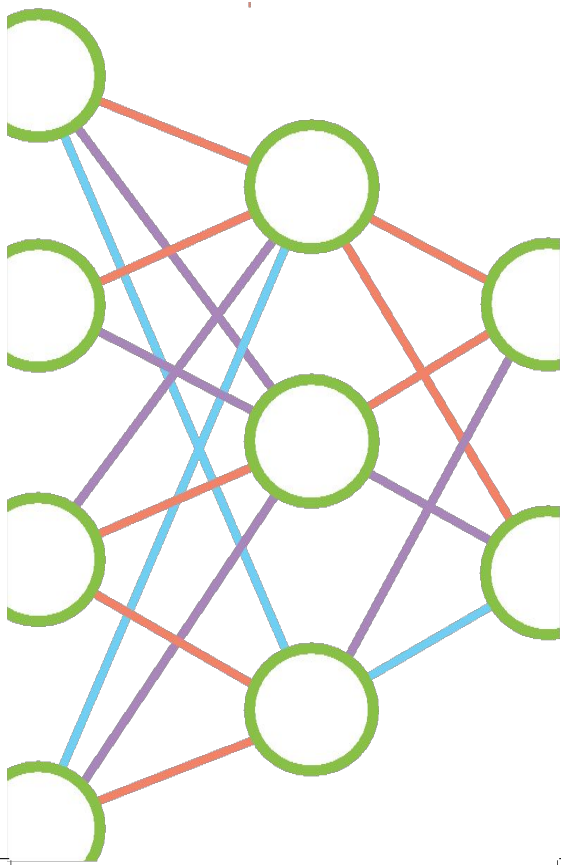
## Inferential Analysis

*Global Count is interested in being able to draw inferential conclusions based on the survey responses, with the understanding that the existing dataset is not representative of the target population.*

For example:

- What demographics make a survey respondent more likely to believe that Civil Rights and Freedoms is an important issue?
- Which critical issues are most likely to appear together in respondents' answers?

We encourage you to explore the analysis that you believe is feasible within your skill set and capabilities - such as weighting the survey responses - in order to approach these questions.



# Dataset Overview

# Data & Limitations

- Operations:
  - PROMOTION: The primary promotion method was Facebook Advertising
  - TARGETING: The FB Ad campaigns were targeted at women and gender-diverse individuals in specific countries.
  - LANGUAGE: The survey and ads were available in 17 languages.
  - ACQUISITION: The survey was conducted entirely online using the Momentiv (SurveyMonkey) platform.
  - The data was collected online using Survey Monkey between Jan 18ths and June 1st, 2021
- Limitations and Distortions
  - Respondents required active Facebook accounts in order to be targeted for advertising or to see the organic campaign. ~8K responses were removed because they were blank.
  - Respondents required computer literacy to view and respond to the campaign plus, participate in the survey
  - Sample collected is non-representative of women and gender diverse people globally
- Survey Mechanics
  - Three surveys: Long, Medium, and Short. All data is merged into a single dataset
    - Long: All questions were asked in a multi-page format
    - Medium: The “Issues affecting your region” question was omitted
    - Short: A brief, two-page experiment to increase completion rate.

# Additional data sets

## Facebook Survey on Gender Equality at Home, 2021

- 96,000 responses from FB users in 200 countries; covers gender norms, unpaid care etc
- publicly available dataset has country-gender level statistics for men, women

## Can it be linked to the Global Count survey?

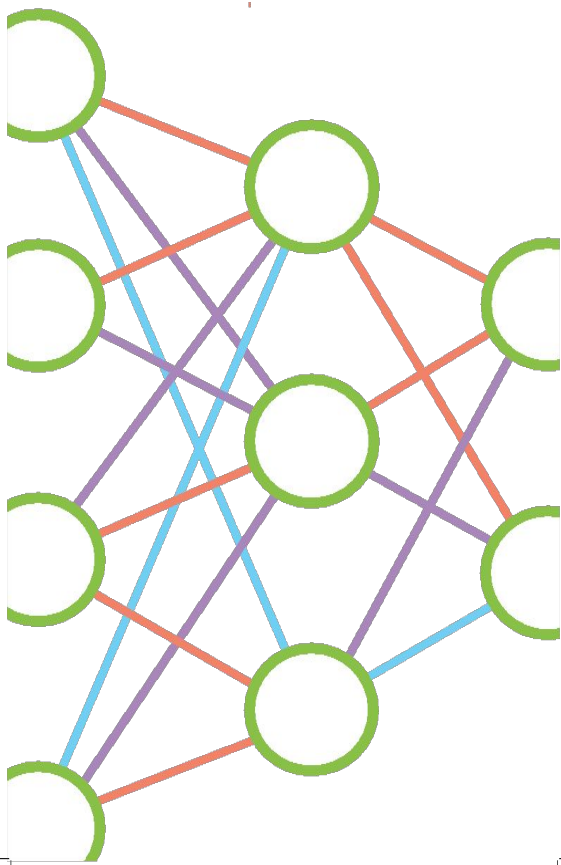
- Both surveys - online, conducted by FB; receive comparable num of responses
- Global Count's responses 32,000, from 173 countries (mostly women, gender diverse)
- can make the assumption that they both represent the same population for women
- from both datasets retain rows corresponding to women, aggregate Global Count on country, join on country

## Facebook Population Survey

- can be used to get population estimates of women, women of reproductive age etc for each city/country

## White Ribbon Alliance

- The question asked: what is your one request for quality reproductive and maternal healthcare services?
- Data is at individual level with info on country, age and their response.
- Not appropriate to link this survey to Global Count survey: only women, in 9 countries, offline component
- top requests for a country from WRA, and organizations working in the Sexual, Reproductive and Parental Health from the Global Count



# Analysis Tips & Considerations

# Analysis considerations & tips

## Understand the data

- What data are available? Are there missing data or duplicates? Predictors of incomplete surveys (e.g., certain subgroup more likely to not answer certain questions?)
- Determine size of entire sample and subgroups (age, identity, region, country)
- Summarize demographics of respondents; compare sample to other data sources to determine representativeness (overall and subpopulations)

# Analysis considerations & tips

## Statistical Analysis

- First, descriptively (cross-tabulate and/or calculate summary statistics) evaluate data before running any statistical tests or models. E.g., descriptively summarize top issues by group before calculating p-values comparing groups.
- The nature of analysis is exploratory rather than confirmatory, therefore, suggest focusing on descriptive analyses rather than comparative statistical tests (e.g., p-values)
- If statistical tests are performed, then determine whether the sample is sufficiently representative for the groups being compared, and whether small cells are created ( $<5$ ). Generally, conclusions drawn from small cells should be considered with caution.
- Tabulate across all groups of interest, but pull out responses to individual questions based on those tables (i.e., one question is “Which age group was most preoccupied with sexual & reproductive health/ climate justice?”. Create a table of top issues by age group across all issues and age groups, then interpret the responses based on that table.)
- Some research questions may benefit from adjustment for age, region, and so forth – again be careful of sample sizes here



# Analysis considerations & tips

## Data Visualization

- Keep your visualizations simple, only keep visual elements that are important in helping the viewer comprehend the information
- Label clearly: give title to the graph, have legible labels, use legends wisely, label axes.
- Use colors effectively: use the same colors for the same kind of data, limit the use of different colors
- Avoid deceiving with your visualizations: include complete data, same scales when showing plots side-by-side

Questions?