



#TOT554518



ANALYZING LGBTIQ+ SUPPORT THROUGH SOCIAL MEDIA DATA

Matthew Kusen

OUTLINE

1
Background

2 Research
Methodology
& Data

3 Data
Analysis

Conclusion
5 &
Going
Forward

4 Key Findings &
Recommendations

BACKGROUND

- **LGBTIQ+ Stigma, Data and Discrimination**
- **Types of Digital Trace data**
- **Demographic Research using Social Media**
- **Sample Size using social media data is not representative**



TERMINOLOGY

- **LGBTIQ+:** an acronym to represent Lesbian, Gay, Bisexual, Transgender, Intersex, and Queer Identity and Experience.

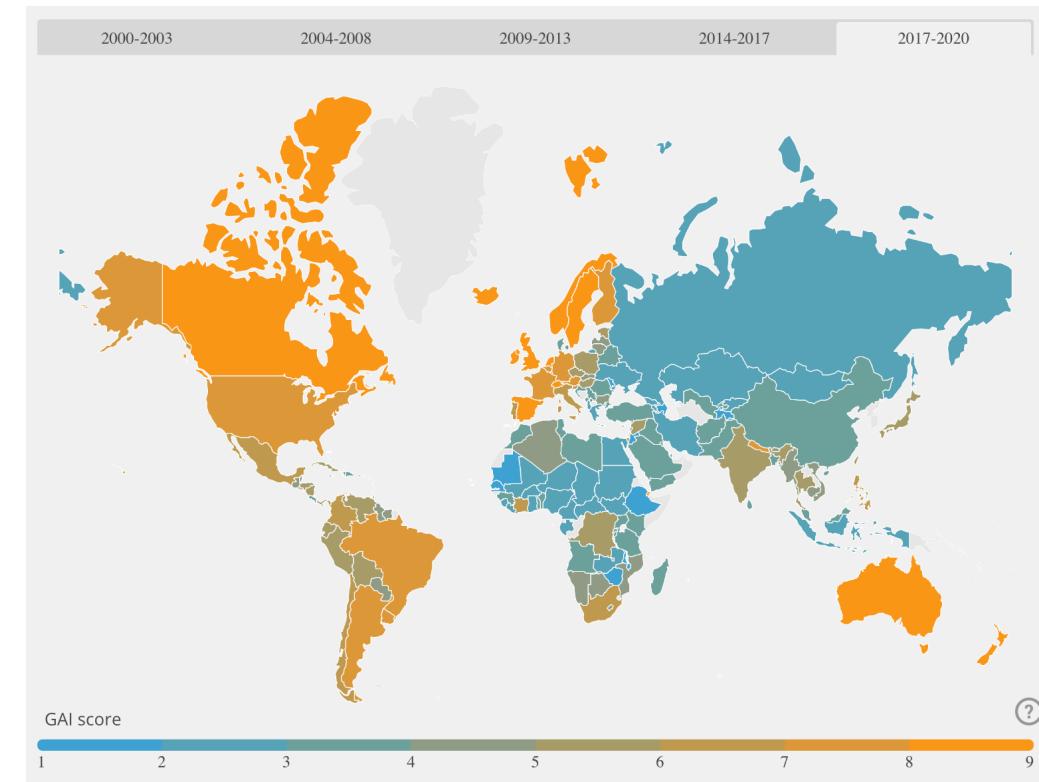


BACKGROUND

LGBTIQ+ Stigma, Data and Discrimination

GLOBALLY

- Social attitudes can heavily influence whether LGBT people are accepted or rejected by employers, family members, teachers, clergy, and society as a whole.
- The Global Acceptance Index (GAI) uses data from more than 2,750 surveys conducted over 30 years to measure changing levels of acceptance toward LGBT people and issues in 174 countries. The GAI provides each nation with a score of 0 to 10 to indicate their level of acceptance



<https://williamsinstitute.law.ucla.edu/projects/gai/>

BACKGROUND

LGBTIQ+ Stigma, Data and Discrimination

Some 68 countries still criminalize homosexuality, most of them majority-Muslim nations in the Middle East, Southeast Asia, and Africa.

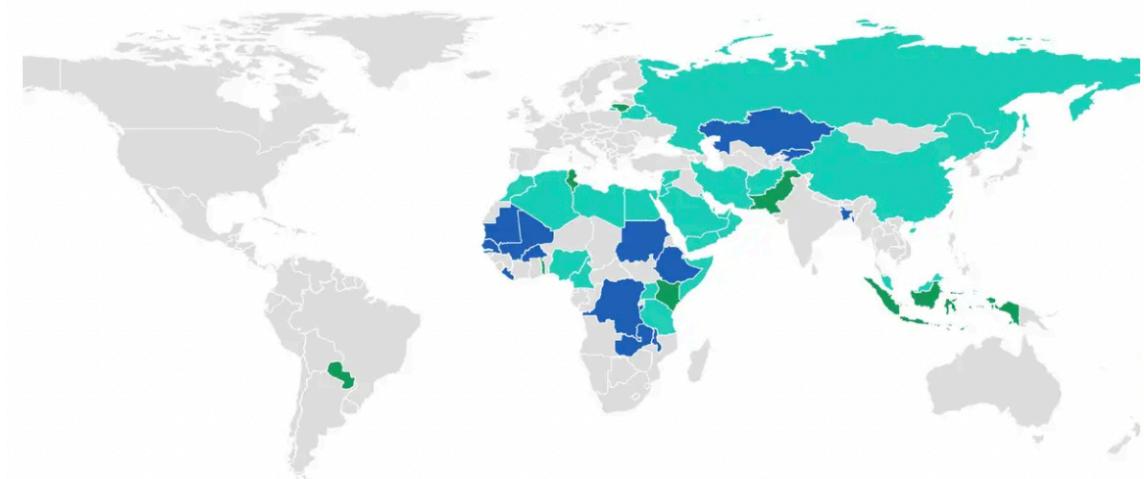


BUSINESS INSIDER

Even where homosexuality is legal, there are laws in place that make living openly difficult.

Countries with discriminatory laws based on sexual orientation

■ Promotion/morality laws ■ Barriers to LGBTQ organizations ■ Both



Source: ILGA

BUSINESS INSIDER

Shayanne Gal/Business Insider

<https://ilga.org/maps-sexual-orientation-laws>

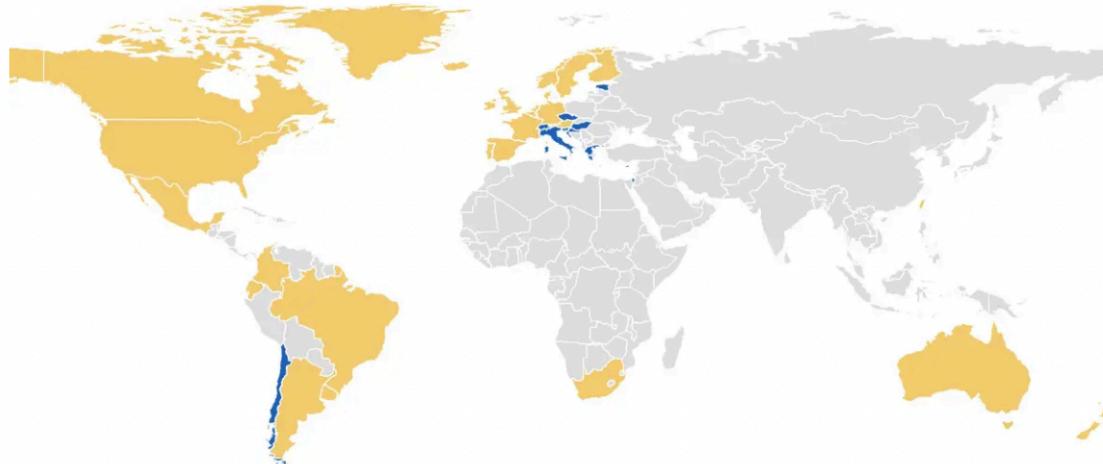
BACKGROUND

LGBTIQ+ Stigma, Data and Discrimination

Only 28 countries have legalized same-sex marriage.

Countries that recognize same-sex marriage

Marriage legal Civil Union only



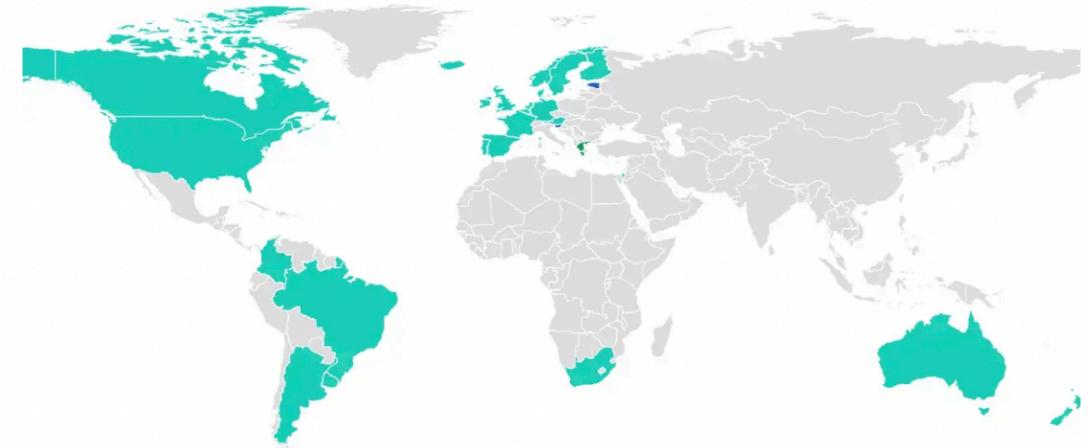
Sources: ILGA; Fortune

Shayanne Gal/Business Insider

Few countries outside of Europe and the Americas allow same-sex couples to adopt children.

Countries that allow same-sex couple adoption

Joint-adoption only Second parent adoption only Both



Source: ILGA

Shayanne Gal/Business Insider

BUSINESS INSIDER

BUSINESS INSIDER

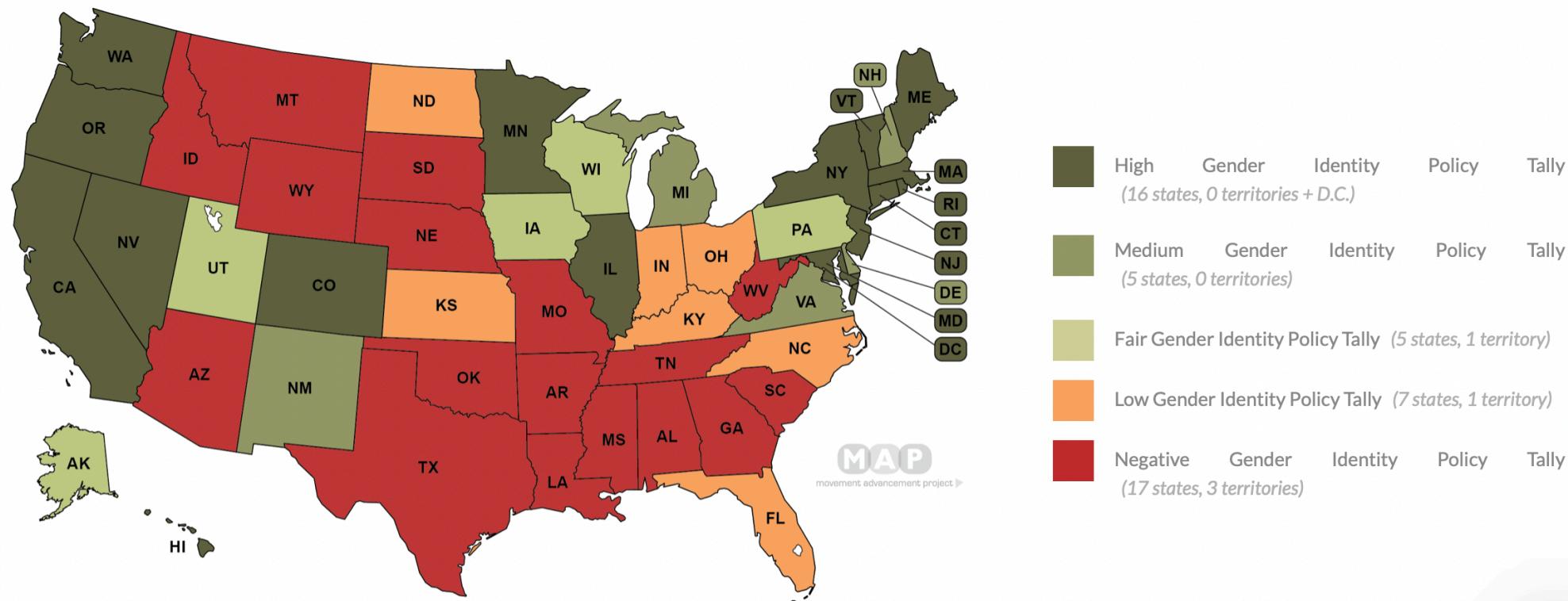
LGBTIQ+ ESTIMATES U.S.A.



- *Not included in census and most demographic surveys*
- *UCLA Williams Institute provides estimates of LGBTIQ population in the U.S. based on Gallup Survey.*
- *Key estimates (US Population):*
 - **11.3 million LGBT (4.5% of population)**
 - **1.39 million transgender people in the U.S. (2020 data)**
 - **1.2 million LGBTIQ+ people in the U.S. identify as nonbinary (11% of LGBTIQ+)**
 - *June 2021*

BACKGROUND

LGBTIQ+ Stigma, Data and Discrimination

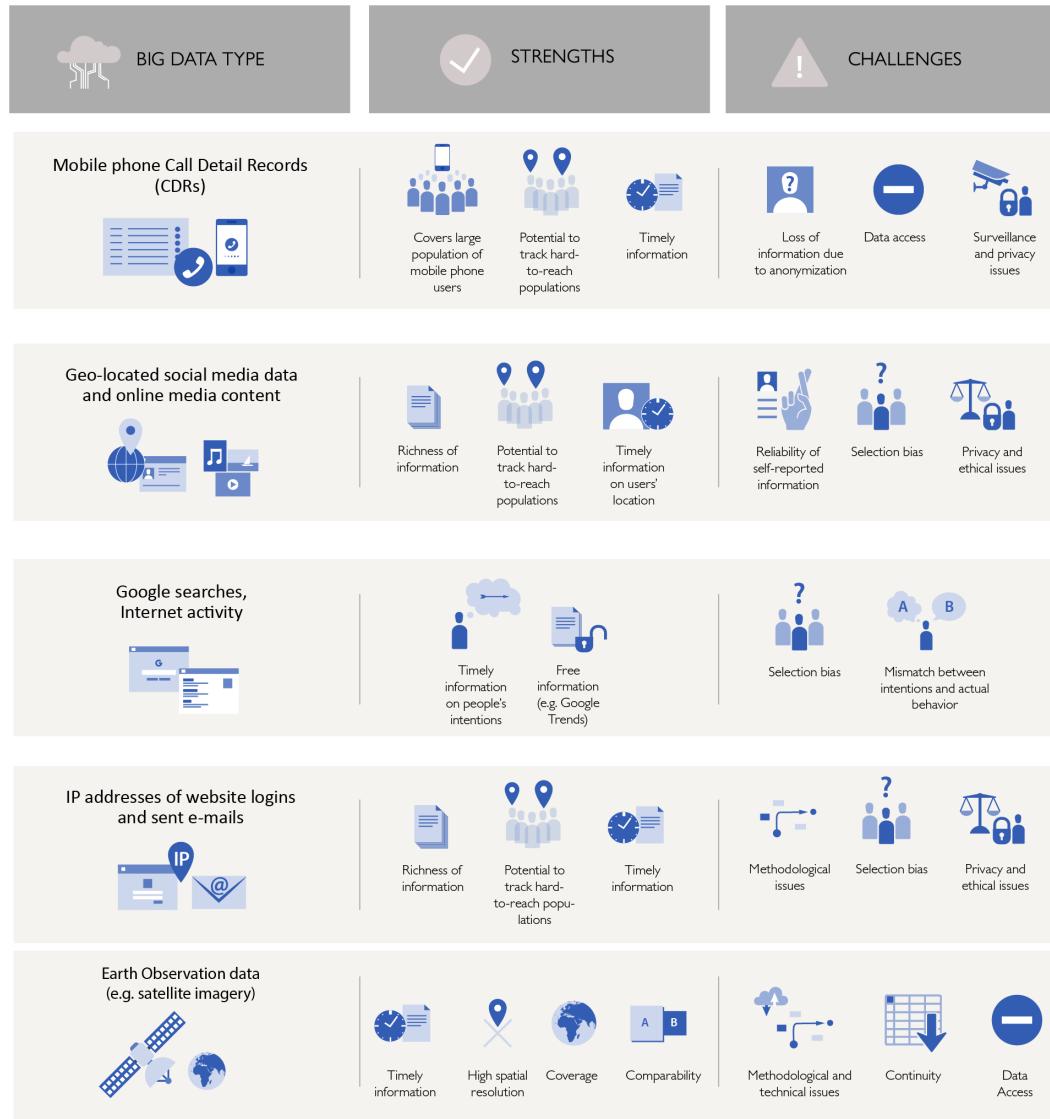


Movement Advancement Project. "Equality Maps: Snapshot: LGBTQ Equality By State." <https://www.lgbtmap.org/equality-maps/>. Accessed 09/19/2021.

TYPES OF DIGITAL TRACE DATA

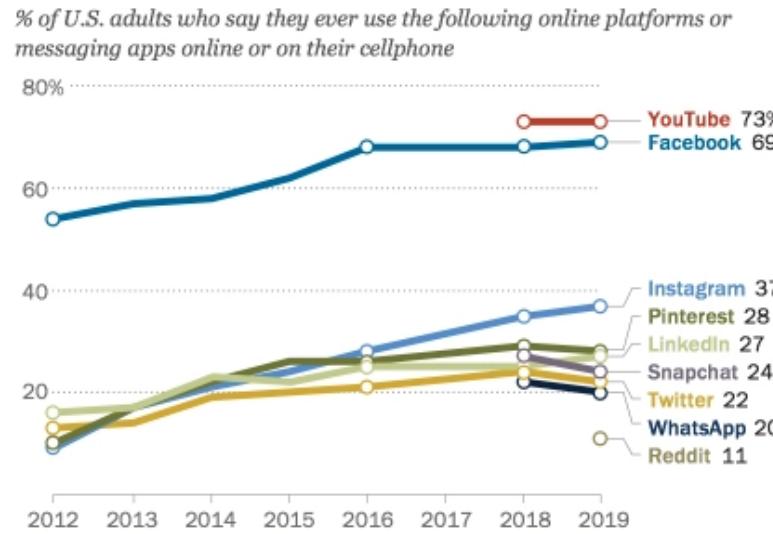
The potential of big data and innovative data sources

Filling the gaps of traditional data sources on migration (censuses, surveys and administrative sources) and contributing to improved understanding of various aspects of migration. A number of studies have shown such potential – but there are significant challenges too.



SAMPLE IS NOT REPRESENTATIVE

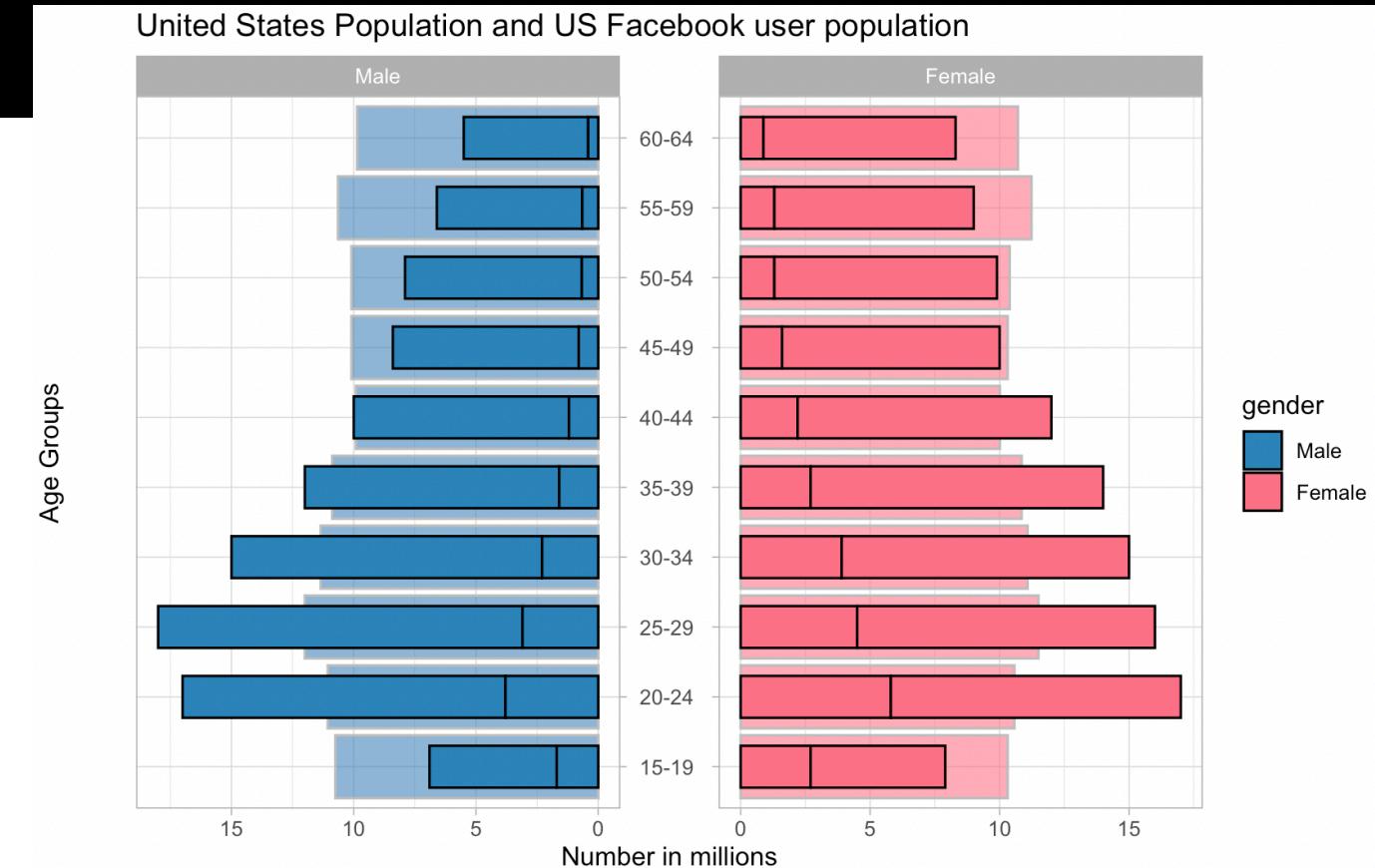
Facebook, YouTube continue to be the most widely used online platforms among U.S. adults



Note: Pre-2018 telephone poll data is not available for YouTube, Snapchat and WhatsApp. Comparable trend data is not available for Reddit.

Source: Survey conducted Jan. 8-Feb. 7, 2019.

PEW RESEARCH CENTER



Source: Facebook.com user data and US Census Population Estimates

Dark shading/smaller box: As of September 13, 2021

Light shading/larger frame: As of June 2020

RESEARCH

- To analyze and understand support for LGBTQ+ populations in the United States through Social Media Data.
- Real time data (available as of date of extraction) based on user digital trace data (geolocation).



RESEARCH

Question: Are women and younger age groups more likely to supportive and show support for LGBTIQ+ community on social media?

- **Hypothesis 1:** Women are more likely than men to be supportive of LGBTIQ+ community members
- **Hypothesis 2:** Younger people are more likely to be supportive of the LGBTIQ+ community on social media compared to older persons.



DATA: FACEBOOK & INSTAGRAM



Audience
Define who you want to see your ads. [Learn More](#)

Create New Audience [Use Saved Audience](#) ▾

Custom Audiences [Add a previously created Custom or Lookalike Audience](#)

[Exclude](#) | [Create New](#) ▾

Locations [Location:](#) United States

[Include](#) ▾ Search Locations [Browse](#)

Age 18 ▾ 64 ▾

Gender All Men Women

Languages Search Languages

Detailed Targeting [Include people who match](#) [Include people who don't match](#)

[Add demographics, interests or behaviors](#) | [Suggestions](#)

Interests

- Business and industry
- Entertainment
- Family and relationships
- Fitness and wellness
- Food and drink
- Hobbies and activities
- Shopping and fashion
- Sports and outdoors

Placements
Show your ads to the right people

Edit Audience

Audience Name LGBT support in USA [Create New](#)

Custom Audiences Search existing audiences

Exclude

Locations [People living in or recently in this location](#)

United States United States

[Include](#) ▾ Search Locations [Browse](#)

Age 18 ▾ 64 ▾

Potential Audience:
Potential Reach: 27,000,000 people [Learn More](#)

Audience Details:

- Location:
 - United States
- Age:
 - 18 - 64
- People Who Match:
 - Interests: LGBT culture, LGBT community or LGBT social movements

Data is provided as aggregate count data rounded up for privacy reasons

VARIABLES OF INTEREST



Outcome variable

Y = LGBTQ+ Support (binary categorical variable)

The behavior classification use the following three terms used:

"LGBT community", "LGBT culture", "LGBT social movements".

1. Coded as “Supportive” are Facebook & Instagram users who have demonstrated support towards the LGBTQ+ community through Facebook’s behavior.
2. Coded as “No support identified” are all users who have no record of behavior classification. May include people who are neutral as well as people who are anti-LGBTQ+ support.

Dependent variable

- **X₁ = Age groups (categorical variable) by Generational group**

- Age groups information is collected and categorized by 10 year age grouping to allow for comparison across groups.
 - **Measurement:** 15-64 (5-year age group categories)
 - Gen Z (15-24)
 - Millennials (25-34)
 - Gen X – Younger (35-44)
 - Gen X – Older (45-54)
 - Baby Boomers (55-64)

- **X₂ = Gender (binary categorical variable)**

- Gender is collected for male and female Facebook users. It is not possible to include 3rd gender options as Facebook does not allow for Facebook marketing towards non-binary or gender non-conforming individuals.
- **X₃** (Interaction of X₁ and X₂)

VARIABLES OF INTEREST



| Variable | Definition | Response Options | Type of Variable |
|--------------|---|---|--|
| LGBT support | Demonstrated support of LGBTIQ+ movements on Facebook and Instagram | 1) Supportive 2) No support identified | Binary / Categorical |
| Age Group | Age group based on Generation categories | 1) Gen Z (age 15-24) 2) <u>Millenials</u> (age 25-34) 3) Gen X - younger (age 35-44) 4) Gen X – older (age 45-54) 5) Baby Boomers (age 55-64) | Categorical <u>10 year age groups</u> |
| Gender | Gender | 1) Male 2) Female | Binary / Categorical |
| “Weights” | Estimated population size for each observation | Discrete value (>0) | |

DATA COLLECTION PROCESS



Create a Saved Audience

Audience Name

Name your audience

Custom Audiences

Search existing audiences

Create New ▾

Potential Audience:

Potential Reach: 15,000,000 people ⓘ

Audience Details:

- Location:
 - United States
- Age:
 - 15 - 19

Exclude

Locations

People living in or recently in this location

United States

United States

Include ▾ Search Locations Browse

Drop Pin

Add Locations in Bulk

Age

15 ▾ 19 ▾

<https://business.facebook.com/adsmanager/>

INPUT INTO DATABASE



Exploratory data analysis.R df Population Pyramid Facebook vs C... Filter

| | level | lgbt_support | gender | age_group | n_millions | POP | country | date_extracted | source |
|----|----------|--------------|--------|-----------|------------|---------|---------|----------------|----------|
| 1 | National | Yes | Male | 15-19 | 1.70 | 1700000 | USA | 2021-09-13 | facebook |
| 2 | National | Yes | Male | 20-24 | 3.80 | 3800000 | USA | 2021-09-13 | facebook |
| 3 | National | Yes | Male | 25-29 | 3.10 | 3100000 | USA | 2021-09-13 | facebook |
| 4 | National | Yes | Male | 30-34 | 2.30 | 2300000 | USA | 2021-09-13 | facebook |
| 5 | National | Yes | Male | 35-39 | 1.60 | 1600000 | USA | 2021-09-13 | facebook |
| 6 | National | Yes | Male | 40-44 | 1.20 | 1200000 | USA | 2021-09-13 | facebook |
| 7 | National | Yes | Male | 45-49 | 0.80 | 800000 | USA | 2021-09-13 | facebook |
| 8 | National | Yes | Male | 50-54 | 0.68 | 680000 | USA | 2021-09-13 | facebook |
| 9 | National | Yes | Male | 55-59 | 0.66 | 660000 | USA | 2021-09-13 | facebook |
| 10 | National | Yes | Male | 60-64 | 0.42 | 420000 | USA | 2021-09-13 | facebook |
| 11 | National | Yes | Female | 15-19 | 2.70 | 2700000 | USA | 2021-09-13 | facebook |
| 12 | National | Yes | Female | 20-24 | 5.80 | 5800000 | USA | 2021-09-13 | facebook |
| 13 | National | Yes | Female | 25-29 | 4.50 | 4500000 | USA | 2021-09-13 | facebook |
| 14 | National | Yes | Female | 30-34 | 3.90 | 3900000 | USA | 2021-09-13 | facebook |
| 15 | National | Yes | Female | 35-39 | 2.70 | 2700000 | USA | 2021-09-13 | facebook |
| 16 | National | Yes | Female | 40-44 | 2.20 | 2200000 | USA | 2021-09-13 | facebook |

Showing 1 to 16 of 60 entries, 9 total columns

DATA



Count = 226.4 million

| Data summary | |
|------------------------|-----|
| Name | df2 |
| Number of rows | 40 |
| Number of columns | 9 |
| <hr/> | |
| Column type frequency: | |
| character | 1 |
| Date | 1 |
| factor | 5 |
| numeric | 2 |

Variable type: Date

| skim_variable | n_missing | complete_rate | min | max | median | n_unique |
|----------------|-----------|---------------|------------|------------|------------|----------|
| date_extracted | 0 | 1 | 2021-09-13 | 2021-09-13 | 2021-09-13 | 1 |

Variable type: factor

| skim_variable | n_missing | complete_rate | ordered | n_unique | top_counts |
|---------------|-----------|---------------|---------|----------|--------------------------------|
| level | 0 | 1 | FALSE | 1 | Nat: 40, Pop: 0 |
| lgbt_support | 0 | 1 | FALSE | 2 | Yes: 20, No : 20 |
| gender | 0 | 1 | FALSE | 2 | Mal: 20, Fem: 20 |
| age_group | 0 | 1 | FALSE | 10 | 15-: 4, 20-: 4, 25-: 4, 30-: 4 |
| country | 0 | 1 | FALSE | 1 | USA: 40 |

Variable type: numeric

| skim_variable | n_missing | complete_rate | mean | sd | p0 | p25 | p50 | p75 | p100 | hist |
|---------------|-----------|---------------|------------|------------|-----------|------------|------------|------------|------------|------|
| n_millions | 0 | 1 | 5.66 | 4.19 | 0.42 | 1.67 | 5.14 | 8.65 | 14.9 | |
| POP | 0 | 1 | 5660000.00 | 4186209.35 | 420000.00 | 1675000.00 | 5140000.00 | 8650000.00 | 14900000.0 | |

DATA ANALYSIS



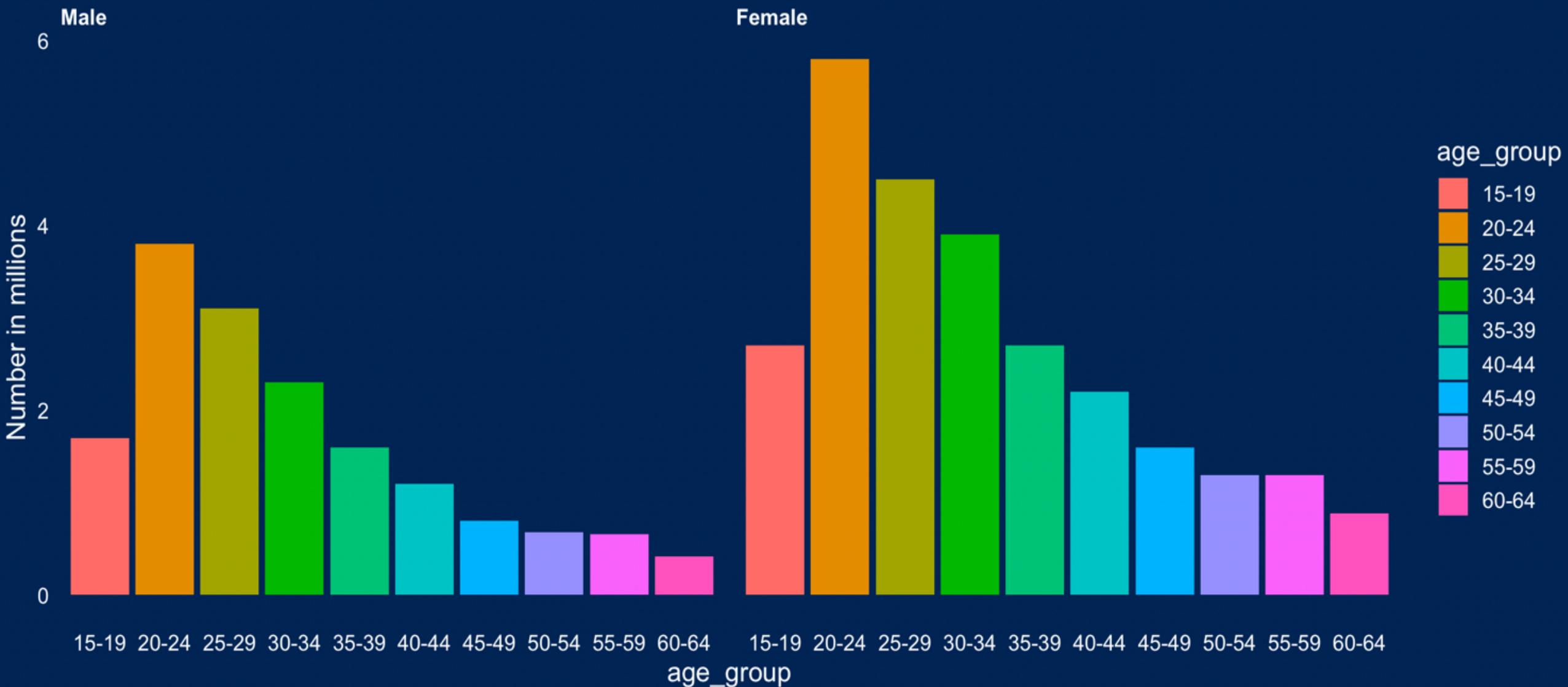
- Descriptive Statistics
- Descriptive univariate & bivariate data visualization and analysis

Multivariate statistical model

- Logistic Regression Analysis
 - Used to analyze binary outcome variable
 - Multivariate statistical analysis
 - Calculate odds ratio
 - Interaction terms



Demonstrated LGBTIQ+ Support by age and sex



source: Instagram and Facebook users in U.S.A.
September 19, 2021

DATA ANALYSIS

**Table 1: Descriptive Statistics of the LGBTIQ Supportive Statistics
(source: Facebook & Instagram)**

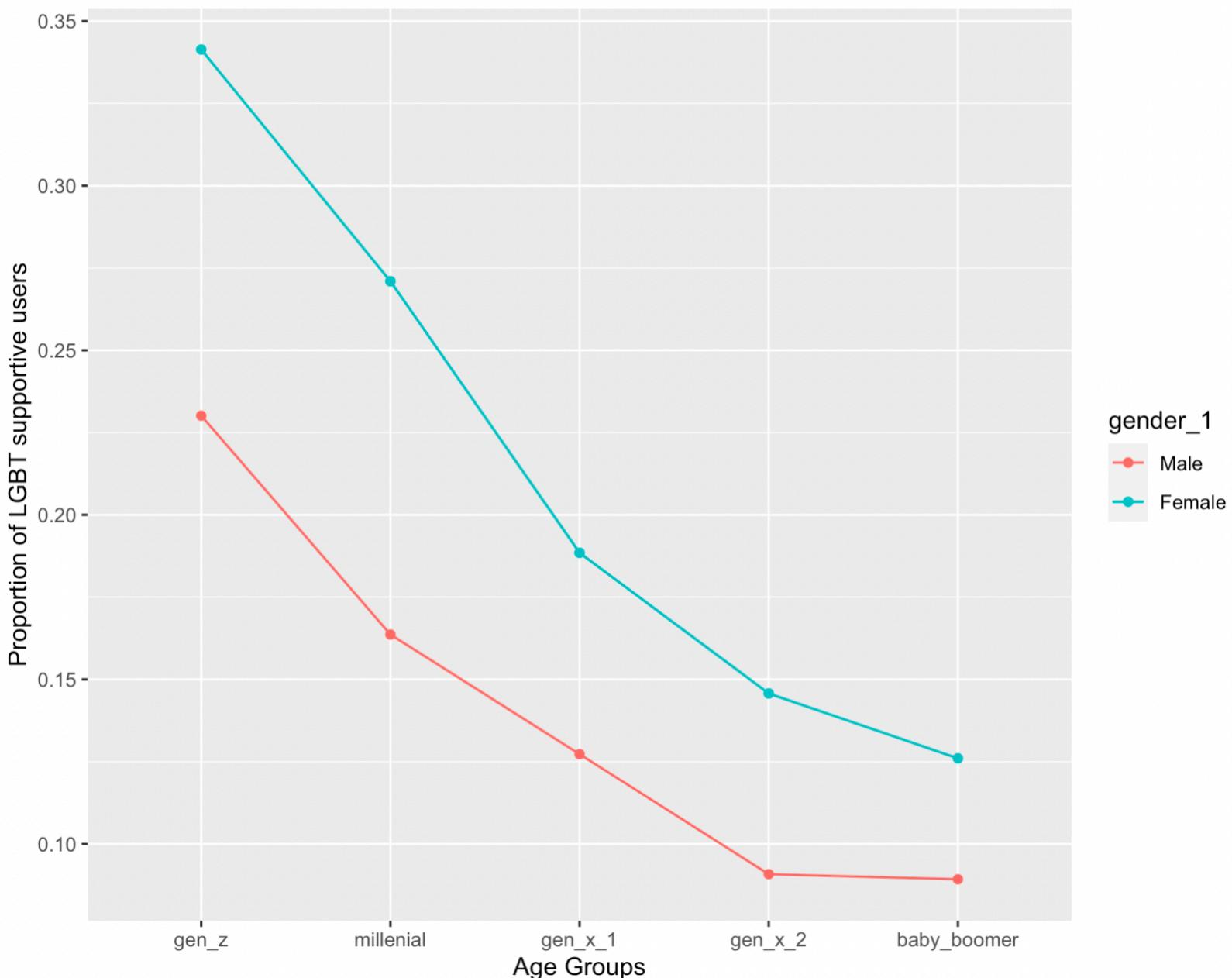
| Variable | Count (in millions) | Percentage (%) |
|---------------------------|---------------------|----------------|
| LGBT Support | | |
| - Supportive | 43.1 | 19.1 |
| - No Support Identified | 183 | 80.9 |
| Gender | | |
| - Men | 107 | 47.4 |
| - Women | 119 | 52.6 |
| Age | | |
| - 15-24 (Gen Z) | 48.8 | 21.5 |
| - 25-34 (Millennials) | 64 | 28.2 |
| - 35-44 (Gen X – group 1) | 48 | 21.2 |
| - 45-54 (Gen X – group 2) | 36.2 | 15.9 |
| - 55-64 (Baby Boomers) | 39.4 | 12.9 |
| TOTAL | 226.4 | |

Table 2: Bivariate Statistics of Facebook data collected comparing LGBT support (Y) with X1 (gender) and X2 (age).

| Variable | Count (in millions) | Count (in millions) | Chi Squared Test (in millions) |
|---------------------------|---------------------|------------------------|--------------------------------|
| LGBT Supportive | LGBT Supportive | No Evidence of Support | |
| Gender | | | 2 *** |
| - Men | 16.3 (15%) | 91 (85%) | |
| - Women | 26.9 (22.6%) | 92.2 (77.4%) | |
| Age | | | 5.82*** |
| - 15-24 (Gen Z) | 14 | 34.8 | |
| - 25-34 (Millennials) | 13.8 | 50.2 | |
| - 35-44 (Gen X – group 1) | 7.7 | 40.3 | |
| - 45-54 (Gen X – group 2) | 4.38 | 31.82 | |
| - 55-64 (Baby Boomers) | 3.26 | 26.14 | |
| TOTAL | 43.14 (19%) | 138.26 (81%) | |

*** p-value < 0.0000000000000022

Proportion of Facebook and Instagram Users demonstrating support for LGBTIQ+ community

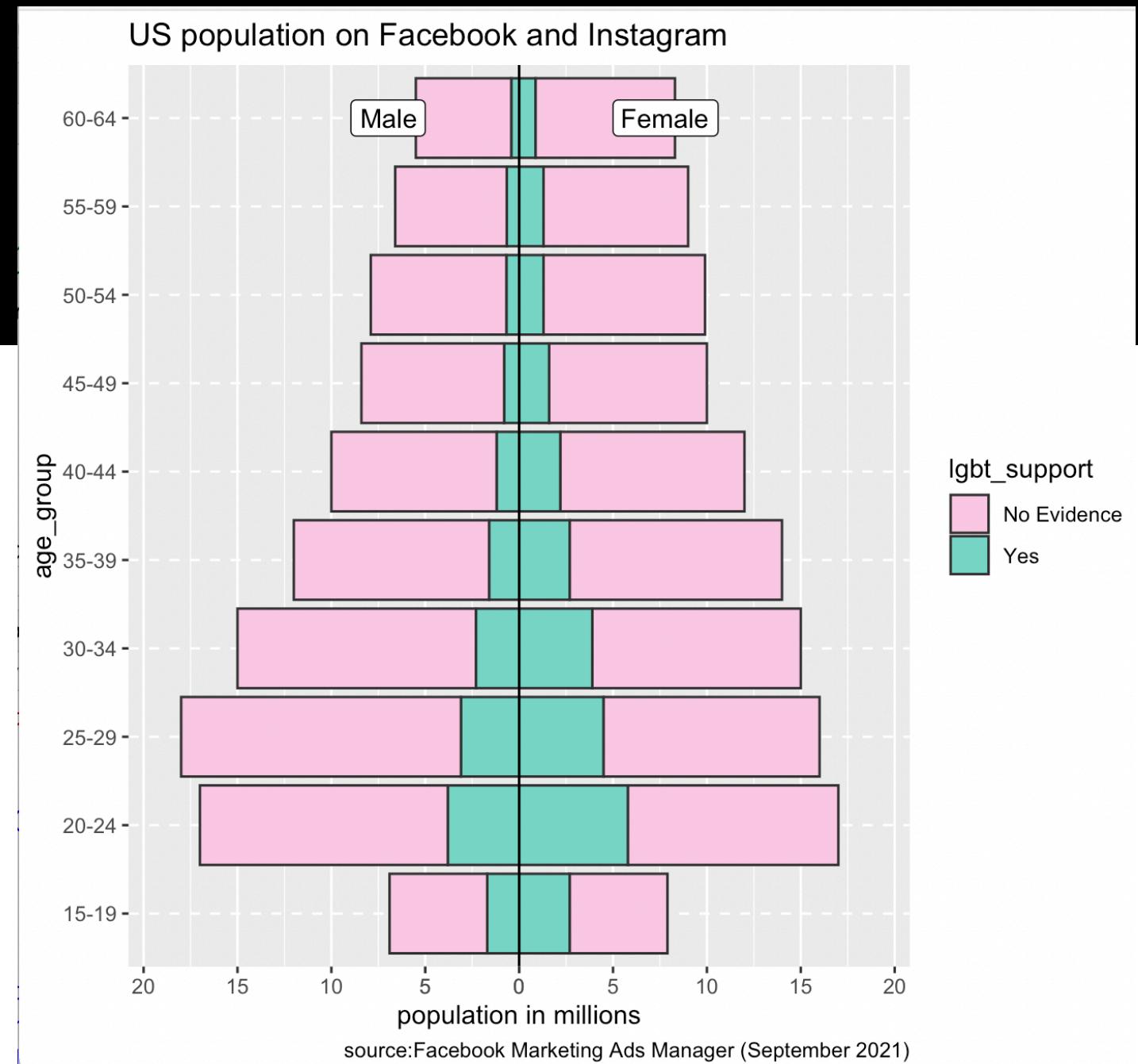


Source: Facebook marketing data - September 2021

POPULATION PYRAMID

Demonstrating LGBTIQ+ support on Facebook or Instagram.

- By gender & Age



MULTIVARIATE ANALYSIS

LOGISTIC REGRESSION



- Assumes binomial distribution
 - Random Component
 - Systematic Component
- Method of Maximum Likelihood Estimation
- Using logarithm form

$$\begin{cases} y_m \sim \text{Bin}(\mu, m) & (\text{random component}) \\ \log \frac{\mu}{1-\mu} = \beta_0 + \beta_1 x & (\text{systematic component}). \end{cases}$$

Systematic Component (includes interaction terms)*

$$\begin{aligned} \text{Log(odds)} = & \beta_0 + \beta_1 x_1 + \beta_2 x_{2a} + \beta_3 x_{2b} + \beta_4 x_{2c} + \beta_4 x_{2e} + \beta_6 x_{2f} + \beta_7 x_{2g} + \beta_8 x_{2h} + \beta_9 x_{2i} + \beta_{10} x_3 + \\ & \beta_{11} x_{3a} + \beta_{12} x_{3b} + \beta_{12} x_{3c} + \beta_{14} x_{3e} + \beta_{15} x_{3f} + \beta_{16} x_{3g} + \beta_{17} x_{3h} + \beta_{18} x_{3i} \end{aligned}$$

*X2 and X3 include categorical variable
with 9 levels against



REGRESSION RESULTS



Table 4: Logistic Regression Results

| Dependent Variable: LGBT Support | <i>Log of Odds</i> LGBT Support | <i>Odds Ratio</i> LGBT Support |
|---|---|-----------------------------------|
| Gender - Female | 0.550*** (0.001) | 1.734*** (0.001) |
| age_group - millenial | -0.424*** (0.001) | 0.655*** (0.001) |
| age_group - gen_x_young | -0.718*** (0.001) | 0.488*** (0.001) |
| age_group - gen_x_older | -1.096*** (0.001) | 0.334*** (0.001) |
| age_group - baby_boomer | -1.115*** (0.001) | 0.328*** (0.001) |
| Gender – Female : age_group - millenial | 0.091*** (0.001) | 1.096*** (0.001) |
| Gender – Female : age_group - gen_x_young | -0.085*** (0.001) | 0.918*** (0.001) |
| Gender – Female : age_group – Gen X - Older | -0.015*** (0.001) | 0.985*** (0.001) |
| Gender – Female : age_group – Baby Boomer | -0.164*** (0.001) | 0.848*** (0.001) |
| Constant | -1.208*** (0.0005) | 0.299*** (0.0005) |
| Observations | 10 | |
| Log Likelihood | -83.563 | |
| Akaike Inf. Crit. | 187.126 | |
| <i>Note:</i> | *p<0.1; **p<0.05; ***p<0.0000000000000002 | |

KEY FINDINGS



- The odds of a female being supportive of LGBTIQ+ on IG and FB is **1.7 times higher** than their male counterparts.
- All age groups are less likely to be supportive of LGBTIQ+ on IG and FB compared with youngest cohort, Gen Z.
- Gen Z (ages 15-24) males have a higher odds compared with Females from Gen X and Baby Boomer generations to support for LGBTIQ+ on social media.
Baby boomer females were .848 times the odds of a Gen Z to show support for LGBTIQ+.
- All of the results are significant at the $p < 0.0000000000000002$ level

IMPLICATIONS / RECOMMENDATIONS



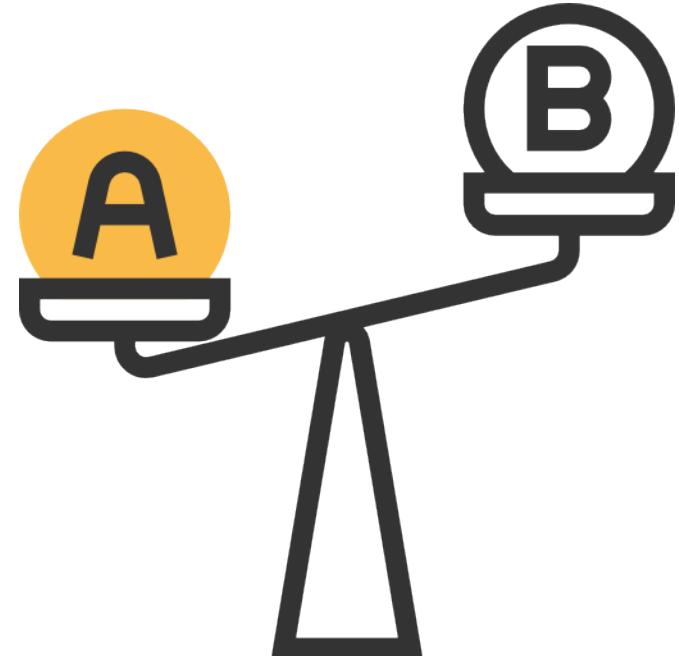
- 1. More is needed to be done to increase LGBTIQ+ support in the United States and increase number of “allies” who show support.**
- 2. More behavior change communication (BCC) strategies required to increase LGBTIQ+ support.**
 - Suggest to conduct sub-national studies.
 - To better understand where the lowest levels of support are and more investment may be required on community activism, LGBTIQ+ programs, and BCC strategies.

COMPARISON TO OTHER STUDIES



1. **Herek 2002: Gender Gaps in Public Opinion about Lesbians and Gay Men**
 - Found that males were less likely to be supportive than females towards gay and lesbian communities (n=1,335)

2. **Kaufman 2021: Attitudes toward LGBT marriage and legal protections (2015 data)**
 - Attitudes are generally supportive of LGBT rights though a substantial minority (about 40%) of American adults favor the rights of small businesses to discriminate against LGBT individuals.

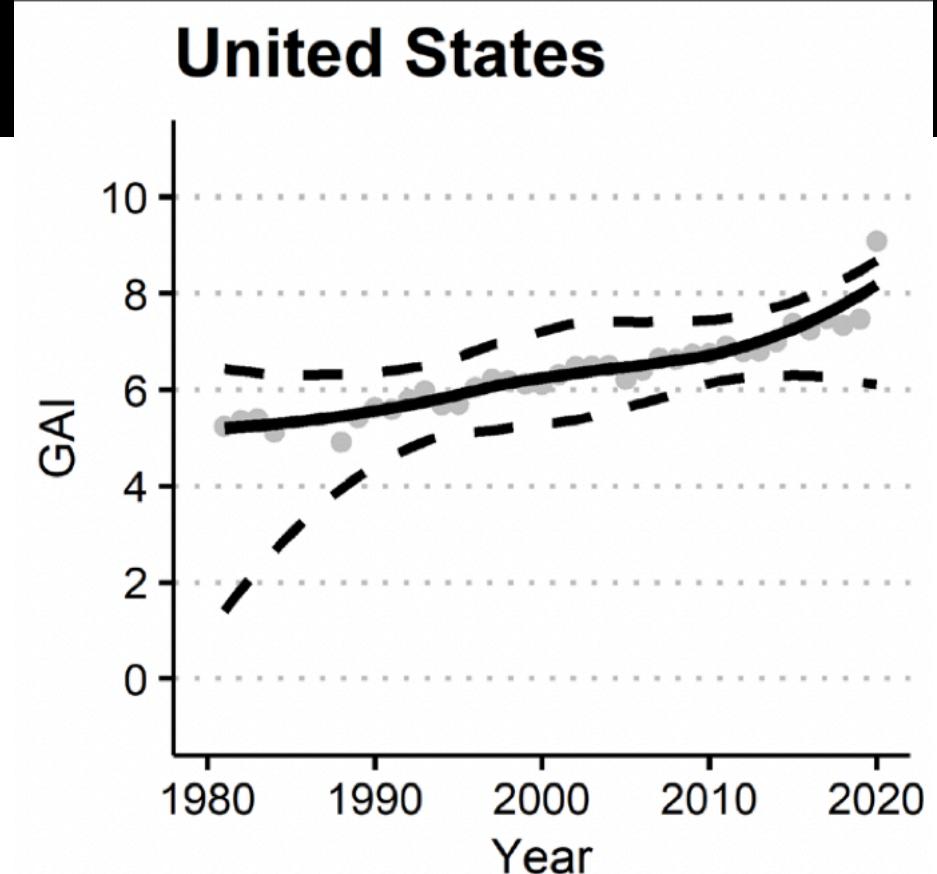


COMPARISON TO OTHER STUDIES

LGBTIQ

Global Acceptance Index (1980-2020)

- “Acceptance is the extent to which LGBTI people are seen in ways that are positive and inclusive, both with respect to an individual’s opinions about LGBTI people and with regard to an individual’s position on LGBTI policies”.



Social Acceptance of LGBTI People in 175 Countries and Locations, 1981 to 2020

Williams Institute - <https://williamsinstitute.law.ucla.edu/projects/gai/>

LIMITATIONS



1. **Data collection:** If using Advanced Programming Interface (API) to extract data, it is complicated and requires programming knowledge.
2. **Data availability:** Facebook may change what data is available free of charge from the Facebook Marketing APIs.
For example, in the past research has looked at sexuality disclosure datapoint through “interested in” on Facebook, which is no longer available. (Gilroy 2021)
3. **Variable Selection:** Not extremely flexible in terms of what demographics and behaviors are available to extract as this is decided and determined by Facebook and their Algorithms for the purposes of Marketing and advertisements
4. **Data Analysis:** Social media research is not representative of the population (Alexander et al., 2019) and it generally tends to be younger than the overall population

STRENGTHS

1. Ability to quickly collect data from Facebook and Instagram, in real-time, and at scale (sample size = 226,000,000 persons)
2. Research methodology also allows for agility in terms of ability for more complex data collection and statistical analyses such as a longitudinal and time series analysis.
3. Promising opportunities in the field of social science / demography and is currently under utilized within the field.



GOING FORWARD

1. Increase multisectoral partnerships to be established and maintained between Research Institutes, Governments and Private Sector companies, and in particular large social media and technology companies such as Facebook, Tik Tok, Google and others. (*this could increase data availability & variables available*)
2. Suggested future study: Collect data in multiple contexts and locations to understand comparative differences among internet users in different countries.
 - 3 billion Facebook users worldwide.



CONCLUSION

- While public opinion is changing and some countries especially in the developed world have instituted progressive LGBTQ+ support policies and laws, much more is still needed to be done.
- This research shows us that while we see increasing levels of support of LGBTQ+ persons, demonstrated support through social and online spaces is still relatively small at less than 20% of Americans.



thank
you

