

# Exploring Discrimination Against LGB\* Individuals

Tiffany Feng  
Derek Wang  
Ryan Sheatz

\*: Those who identified as transgender were excluded from our analysis due to an insufficient number of observations. As such, our analysis only applies to lesbian, gay, and bisexual individuals.

# Research Question

How might one's gender, sexual orientation, ethnicity, whether one is openly LGBT on social networks, perception of societal acceptance, perception of how friendly a major organization is towards LGBT individuals, level of support of pro-LGBT actions and policies impact discrimination against LGB\* individuals?

Additionally\*\*, does the effect of sexual orientation on percent score of discrimination vary by ethnicity?

\*: Those who identified as transgender were excluded from our analysis due to an insufficient number of observations. As such, our analysis only applies to lesbian, gay, and bisexual individuals.

\*\* : For all other predictors held constant



An illustration featuring a young boy with dark hair, wearing a yellow t-shirt, orange shorts, and blue sneakers with white socks. He stands on a grey shadow on the left. A large yellow circle is in the center, containing the text 'Our Predictors and Response Variables' in bold dark blue. A dotted line starts from the top left, passes through the circle, and ends with a paper airplane on the right. There are also a small teal triangle, a grey cloud, and a yellow triangle in the background.

# **Our Predictors and Response Variables**

# Our Numerical Predictors

Numerical Predictors (0-100)	Definition	Detailed Explanation of Variable
LGBTsocacceptp	% score of how accept society is to LGBT individuals*	*: how much acceptance do gay/lesbian/bisexual men/bisexual women/transgender people have?
LGBTfriendlyp	% score on how friendly major organizations* are towards LGBT individuals.	*: democratic & republican party, the obama administration, the entertainment and professional sports industry, and the military)
LGBTpolicyp	% score on how important the support for LGBT policies* are	*: equal employment, AIDS/HIV research, civil unions, establishing organizations for LGBT youth)
LGBTsupportp	% score on actions that support the LGBT community	*: membership in LGBT community, buying products from companies that support LGBT, avoiding products from companies that do not support LGBT, LGBT rally attendance, donation status

- These numerical predictors are all survey scores.
- The dataset provided two versions of the survey score
  - raw score (so like a scale from 0-8).
  - same score converted to a percentage.
- Percentages are more intuitive
  - All out of 100
  - A high percentage score represents high support for the value being measured.

# Our Categorical Predictors

Categorical Predictors	Levels
gender	Male, Female (2)
sexualorientation	Gay, Lesbian, Bisexual, Transgender (4)
ethnicity	White, Black, Hispanic, Other (4)
revealLGBTsocialnetwork	Yes, No (2)

- *Gender*
  - men and women face different expectations.
- *sexualorientation*
  - each group's experiences are unique from one another.
- *ethnicity*
  - different ethnicities are treated differently in our society.
- *revealLGBTsocialnetwork*
  - reveals whether a person is “out” to society which may affect how they are treated

People's personal biases against these predictors are what cause them to discriminate.

# Our Response Variable

Outcome Variable	Type	Definition	Explanation of Variable
LGBTdiscriminationp	Numerical (0-100)	% score of being discriminated* against as an LGB individual	*: physical assault, slurs, poor service, unwelcome at place of worship, unfair unemployment, estrangement

- This variable is defined by actions that are meant to harm LGB\* individuals.

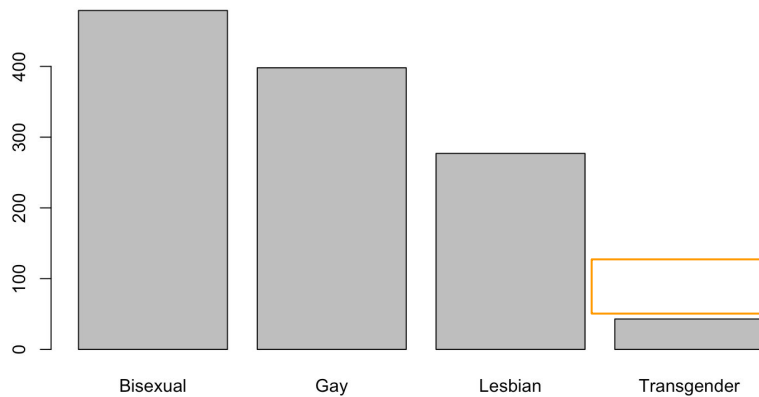
# Transformations

Variable	Type	Definition
sexualorientation	Categorical (4 levels, Gay/Lesbian/Bisexual/Transgender)	What is your sexual orientation?

Bisexual	Gay	Lesbian	Transgender
479	398	277	43

- There are not enough observations in the Transgender category
  - Greater variation in the data

Bar Chart of Sexual Orientation Before Removing Transgender Category



# About our dataset "lgbt.csv"



- Some variables (e.g. lgbt\_policyp) touch upon systemic factors that affect the LGBT community as a whole, while other variables (e.g. gender, ethnicity) affects how one person may treat another at the individual level.
- By investigating these factors, we can identify which variables influence discrimination against LGB\* individuals, and in doing so we can further investigate what can be done to reduce it.

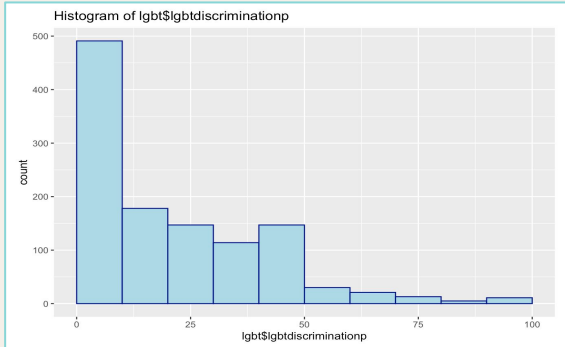


An illustration featuring a large yellow circle in the center with the text "Exploratory Data Analysis" in bold dark blue. To the left, a cartoon boy with dark hair, wearing a yellow shirt, orange shorts, and blue sneakers, stands with his arms crossed. A dotted line starts from the top left, curves around the yellow circle, and ends with a paper airplane on the right. Decorative elements include an orange shape in the top left, a teal triangle, a grey cloud, a yellow triangle, and a teal shape in the bottom right.

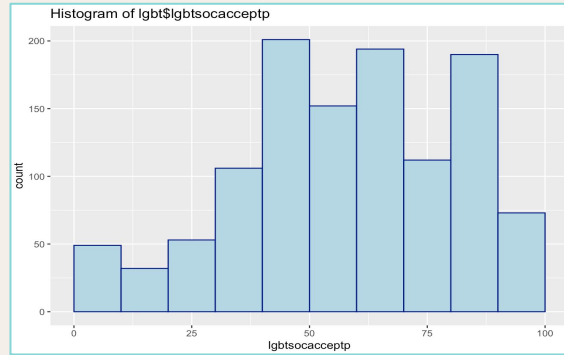
# **Exploratory Data Analysis**

# EDA: Histograms (Numerical)

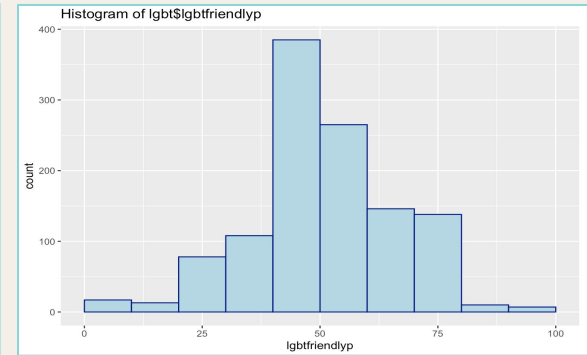
% score on the  
discrimination against LGBT



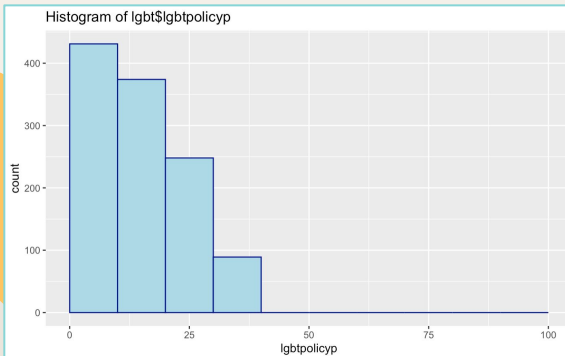
% score on the  
Social acceptance of LGBT



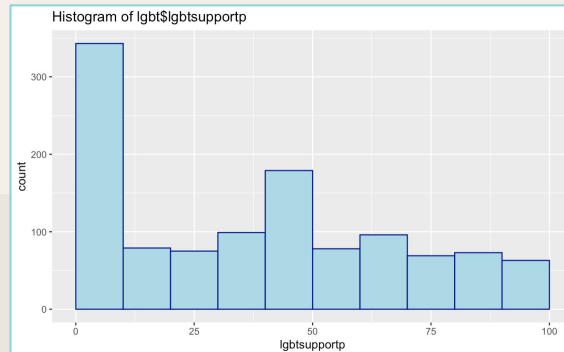
% score on the friendly  
treatment of LGBT



% score on the priority of support  
for policies related to LGBT



% score on supporting LGBT



- The histograms are non-problematic
  - no outstanding outliers or problems with the distribution of the variables.

# EDA: Frequency Tables (Categorical)

## gender

	Freq	%
female	640	53.56
male	555	46.44

## Sexual orientation

	Freq	%
bisexual	479	40.02
gay	398	33.25
lesbian	277	23.14
trans gender	43	3.59

## ethnicity

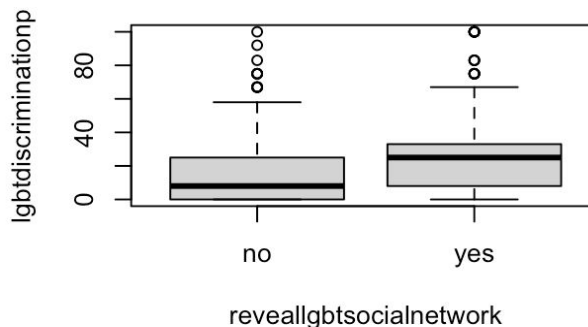
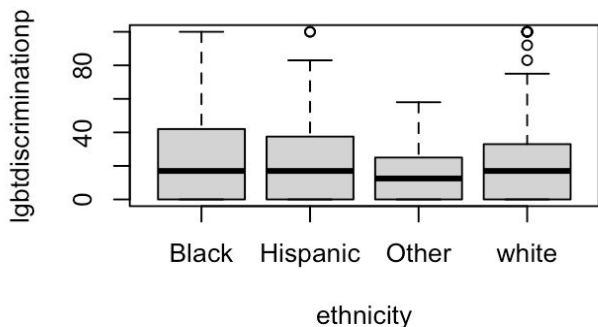
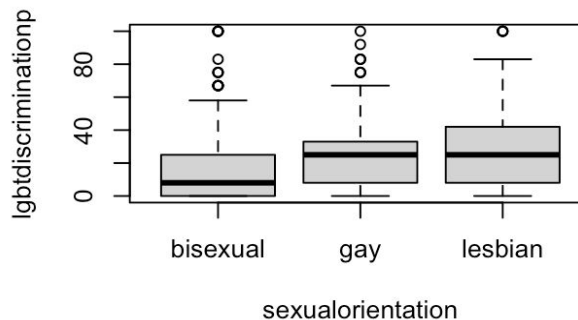
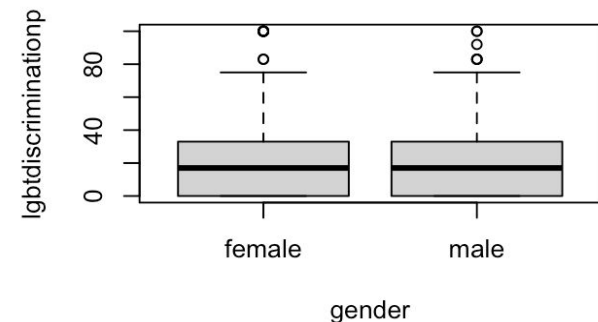
	Freq	%
Black	83	7.24
Hispanic	129	11.26
Other	33	2.88
white	901	78.62

## Have you ever revealed your sexual orientation on a social network?

	Freq	%
no	419	45.15
yes	509	54.85

- Due to the low number of transgender individuals, we changed these values to `NA`
- Decided not to transform or pool our ethnicity variable as we felt it was important to analyze the differences associated with each level.

# EDA: LGBTDiscriminationp (Outcome Variable) in Relation to Categorical Predictors



- Those who reported yes to *reveallgbtsocialnetwork* have a higher mean percent score of feeling discriminated against.
- Gay and lesbian individuals on average also seem to have a greater mean percent score in feeling discriminated against than bisexual individuals

# EDA: Contingency Tables

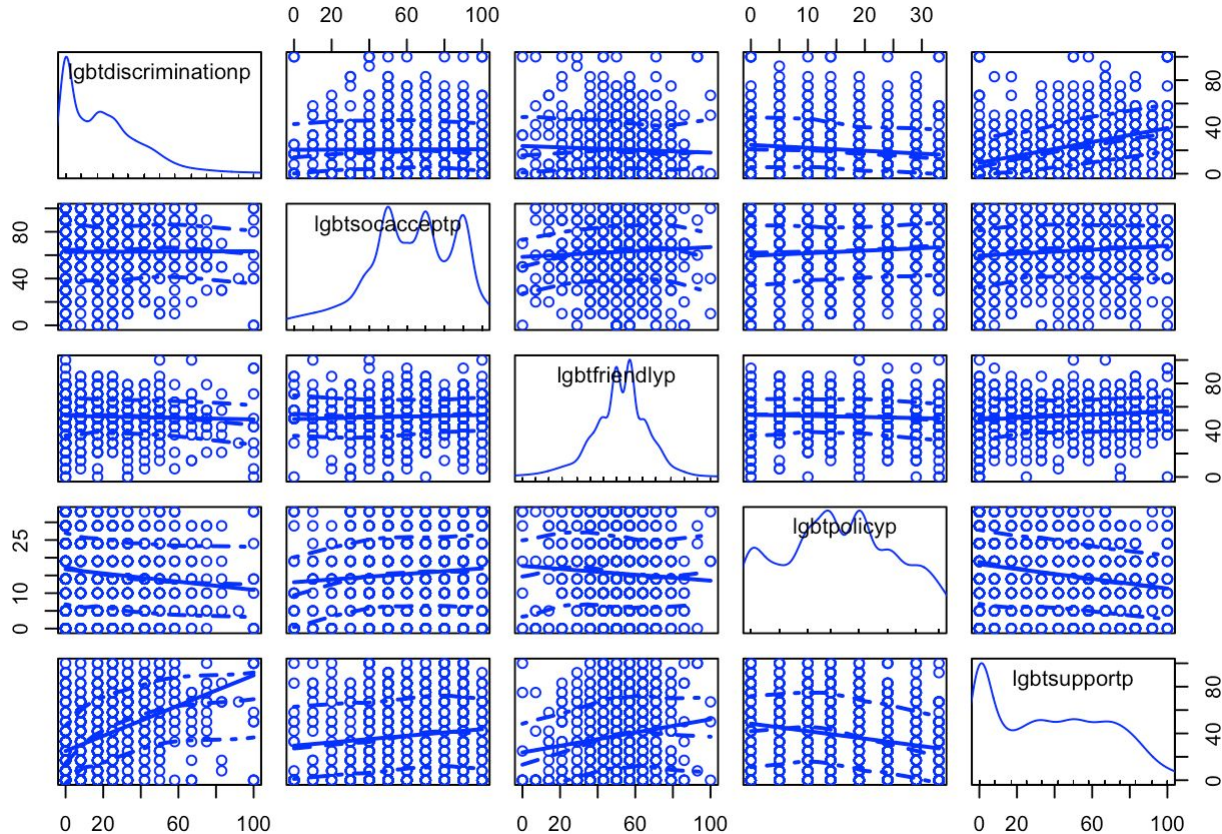
## Sexual orientation & Ethnicity

Data Frame: lgbt

	sexualorientation			
ethnicity	bisexual	gay	lesbian	Total
Black	39 ( 48.1%)	22 ( 27.2%)	20 ( 24.7%)	81 ( 100.0% )
Hispanic	41 ( 33.9%)	51 ( 42.1%)	29 ( 24.0%)	121 ( 100.0% )
Other	12 ( 36.4%)	16 ( 48.5%)	5 ( 15.2%)	33 ( 100.0% )
white	362 ( 41.6%)	295 ( 33.9%)	214 ( 24.6%)	871 ( 100.0% )
Total	454 ( 41.0%)	384 ( 34.7%)	268 ( 24.2%)	1106 ( 100.0% )

- Highlighted = sexual orientation with the largest proportion for each ethnicity.

# EDA: Scatterplot Matrix (Numerical)



- None of the scatterplots appear nonlinear, which is a good sign.

# EDA: Correlation Matrix (Numerical)

	Outcome variable				
	lgbtdiscriminationp	lgbtsocacceptp	lgbtfriendlyp	lgbtpolicyp	lgbtsupportp
lgbtdiscriminationp	1.0000	0.0052	-0.0421	-0.1208	0.4383
lgbtsocacceptp	0.0052	1.0000	0.0548	0.0925	0.1094
lgbtfriendlyp	-0.0421	0.0548	1.0000	-0.0639	0.1427
lgbtpolicyp	-0.1208	0.0925	-0.0639	1.0000	-0.2101
lgbtsupportp	0.4383	0.1094	0.1427	-0.2101	1.0000

Predictors

- We did not have any highly correlated variables, which is desirable.



The illustration features a young boy with dark hair, wearing a yellow t-shirt, orange shorts, and blue sneakers with white socks. He stands on a light gray ground with his arms crossed. To his right is a large, solid yellow circle. Inside this circle, the text "Our Final Model" is written in a bold, dark blue, sans-serif font. A dotted line starts from the top left, curves over the top of the yellow circle, and ends with a white paper airplane flying towards the right. The background is a light gray sky with a large orange shape in the top left, a small yellow triangle in the top right, a gray cloud, and a teal shape in the bottom right corner.

# **Our Final Model**



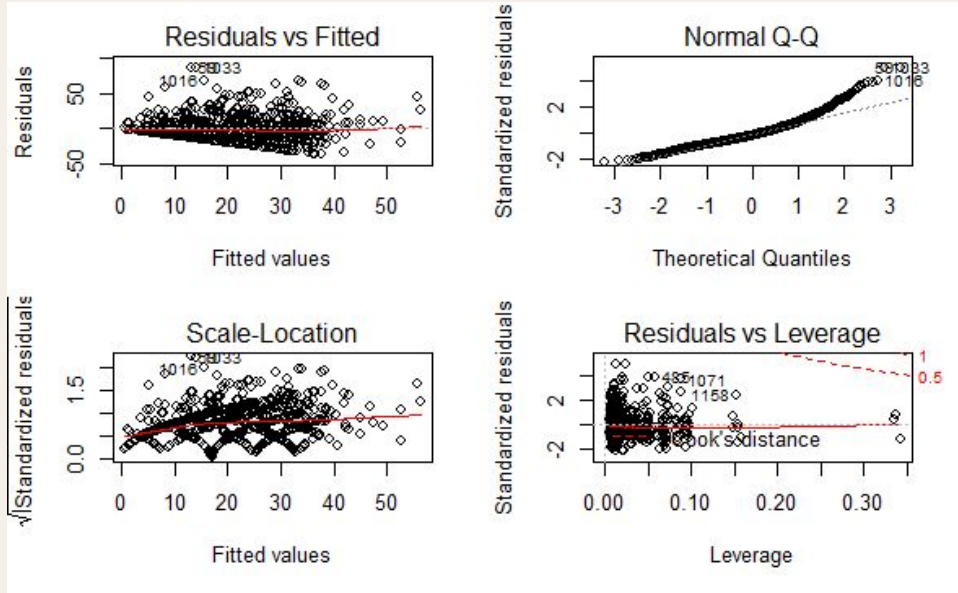
# Final Multiple Linear Regression Model

We chose a linear model because they are very easy to interpret.

Our model:

```
lgbt_model <- lm(lgbtdiscriminationp ~ gender + sexualorientation * ethnicity  
+ reveallgbtsocialnetwork #categorical predictors  
+ lgbsocacceptp + lgbtfriendlyp + lgbt Policyp  
+ lgbsupportp, data = lgbt) #numerical predictors
```

# Assumptions



- Linearity assumption satisfied
  - the residuals have no pattern.
- Normality assumption satisfied
  - most points are on the reference line.
- Homoscedasticity is satisfied
  - the variance is scattered
- Independence of residual errors assumption satisfied
  - assume that data was collected in a way that the observations are independent



# Multicollinearity

Variable	GVIF	Degrees of Freedom
gender	3.63	1
sexualorientation	272.55	2
ethnicity	19.51	3
reveallgbtsocialnetwork	1.17	1
lgbtsocacctp	1.05	1
lgbtfriendlyp	1.11	1
lgbtpolicyp	1.04	1
lgbtsupportp	1.29	1
sexualorientation:ethnicity	2091.84	6

- We do not need to worry about multicollinearity because most VIF values for predictors are close to 1 or below 5.
- Interaction in the model includes *sexualorientation* and *ethnicity* they appear to be very correlated in the VIF test, but we do not need to worry about multicollinearity.

# Interpretation of regression coefficients

- Our significant predictors explain a good amount of variation, especially in the humanities context.
- Our model R-squared value is 0.2645
  - About 26.45% of the variation in *lgbtdiscriminationp* can be explained by:
    - sexual orientation
    - whether or not sexual orientation has been revealed on a social network
    - actions that support the LGBT,
    - how friendly different organizations are towards LGBT
    - Interactions between sexual orientation and ethnicity

Variables Significant at a 0.05 Level	Regression Coefficient
sexualorientationgay	22.25
sexualorientationlesbian	16.27
reveallgbtsocialnetworkyes	4.68
lgbtfriendlyp	-0.13
lgbtsupportp	0.27
sexualorientationgay:ethnicityHispanic	-16.09
sexualorientationgay:ethnicityOther	-22.20
sexualorientationgay:ethnicitywhite	-18.40
sexualorientationlesbian:ethnicitywhite	-14.45

# Interpretation of regression coefficients

- The default baseline (intercept)
  - Black female bisexual individuals that are not openly “out” on social networks, when all the numerical predictors are at zero.
- Numerical Predictor Interpretation Example
  - For every 1 percentage point increase in how friendly organizations are towards LGBT, the the percentage score of discrimination against LGBT, on average, decreases by 0.13 points when holding all other factors constant.

Variables Significant at a 0.05 Level	Regression Coefficient
Categorical: sexualorientationgay	22.25
Categorical: sexualorientationlesbian	16.27
Categorical: reveallgbtsocialnetworkyes	4.68
Numerical: lgbtfriendlyp	-0.13
Numerical: lgbsupportp	0.27
sexualorientationgay:ethnicityHispanic	-16.09
sexualorientationgay:ethnicityOther	-22.20
sexualorientationgay:ethnicitywhite	-18.40
sexualorientationlesbian:ethnicitywhite	-14.45

# Interpretation of regression coefficients

- **Categorical Predictor Interpretation Example**
  - Compared to the baseline, we would expect gay individuals, on average, to score 22.25 percentage points higher in *lgbtdiscriminationp* when holding all other factors constant.
- **Interaction Term Interpretation Example**
  - Compared to the baseline, we would expect gay Hispanic individuals, on average, to score 16 percentage points lower in *lgbtdiscriminationp* when holding all other factors constant.

Variables Significant at a 0.05 Level	Regression Coefficient
Categorical: sexualorientationgay	22.25
Categorical: sexualorientationlesbian	16.27
Categorical: reveallgbtsocialnetworkyes	4.68
Numerical: lgbtfriendlyp	-0.13
Numerical: lgbsupportp	0.27
sexualorientationgay:ethnicityHispanic	-16.09
sexualorientationgay:ethnicityOther	-22.20
sexualorientationgay:ethnicitywhite	-18.40
sexualorientationlesbian:ethnicitywhite	-14.45



# **Full List of Coefficient Interpretations**

# Interpretation of numerical regression coefficients

- **lgtsocacceptp**
  - For every 1 percentage point increase in how much acceptance LGBT individuals have, the the percentage score of discrimination against LGBT, on average, decreases by 0.03 points when holding all other factors constant.
- **lgbtfriendlyp**
  - For every 1 percentage point increase in how friendly organizations are towards LGBT, the the percentage score of discrimination against LGBT, on average, decreases by 0.13 points when holding all other factors constant.
- **lgbtpolycyp**
  - For every 1 percentage point increase in support for LGBT policies, the the percentage score of discrimination against LGBT, on average, increases by 0.03 points when holding all other factors constant.
- **lgbtsupportp**
  - For every 1 percentage point increase in actions that support the LGBT communtiy, the the percentage score of discrimination against LGBT, on average, increases by 0.27 points when holding all other factors constant.



# Interpretation of categorical regression coefficients\*

- **gendermale**
  - Compared to the baseline, we would expect males, on average, to score 0.72 percentage points lower in *lgbtdiscriminationp* when holding all other factors constant.
- **sexualorientationgay**
  - Compared to the baseline, we would expect gay individuals, on average, to score 22.25 percentage points higher in *lgbtdiscriminationp* when holding all other factors constant.
- **sexualorientationlesbian**
  - Compared to the baseline, we would expect lesbian individuals, on average, to score 16.27 percentage points higher in *lgbtdiscriminationp* when holding all other factors constant.
- **ethnicityHispanic**
  - Compared to the baseline, we would expect Hispanic individuals, on average, to score 0.02 percentage points higher in *lgbtdiscriminationp* when holding all other factors constant.

\*: This slide and the next two slides were provided for grading purposes and was not presented in our video in the interest of time--as such the three interpretation examples in the previous slide are repeated here

# Interpretation of categorical regression coefficients

- **ethnicityOther**
  - Compared to the baseline, we would expect individuals that do not identify as White, Hispanic, or Black, to score, on average, 2.84 percentage points lower in *lgbtdiscriminationp* when holding all other factors constant.
- **ethnicitywhite**
  - Compared to the baseline, we would expect white individuals, on average, to score 2.22 percentage points lower in *lgbtdiscriminationp* when holding all other factors constant.
- **reveallgbtsocialnetworkyes**
  - Compared to the baseline, we would expect individuals who reveal their LGBT sexuality on social media, on average, to score 4.68 percentage points higher in *lgbtdiscriminationp* when holding all other factors constant.

# Interpretation of interaction terms

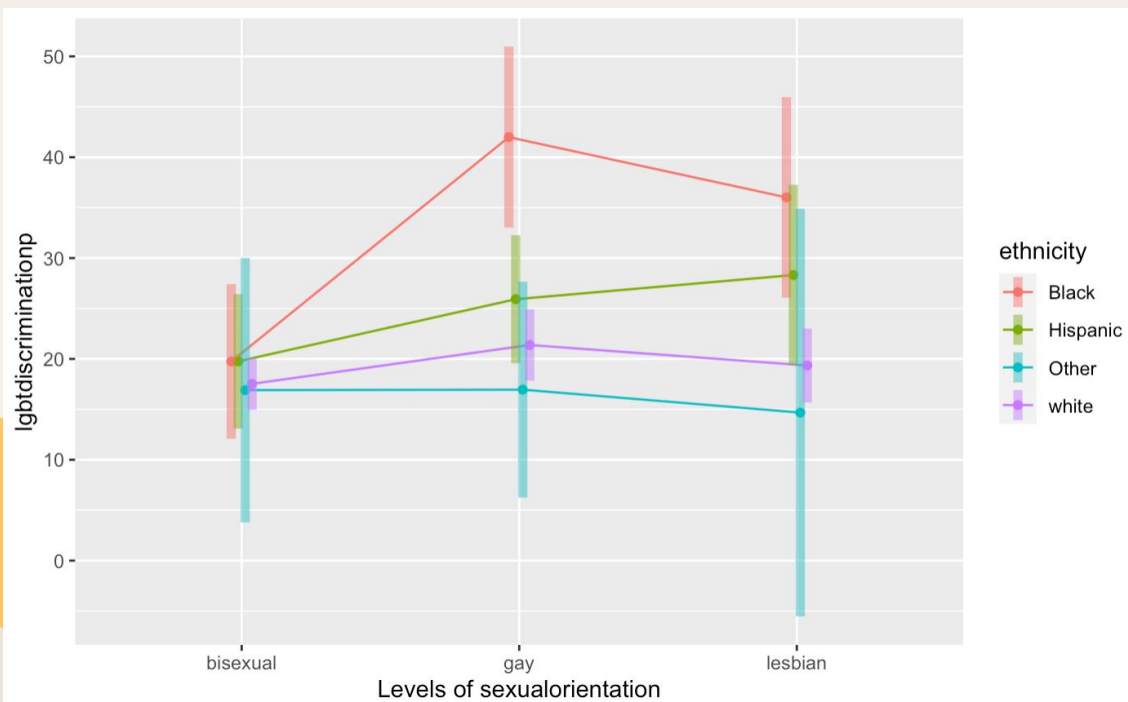
	Hispanic	Other
Gay	<ul style="list-style-type: none"><li>• <b>sexualorientationgay:ethnicityHispanic</b><ul style="list-style-type: none"><li>○ Compared to the baseline, we would expect gay Hispanic individuals, on average, to score 16 percentage points lower in <i>lgbtdiscriminationp</i> when holding all other factors constant.</li></ul></li></ul>	<ul style="list-style-type: none"><li>• <b>sexualorientationgay:ethnicityOther</b><ul style="list-style-type: none"><li>○ Compared to the baseline, we would expect gay individuals that do not identify as White, Hispanic, or Black, to score, on average, 22.2 percentage points lower in <i>lgbtdiscriminationp</i> when holding all other factors constant.</li></ul></li></ul>
Lesbian	<ul style="list-style-type: none"><li>• <b>sexualorientationlesbian:ethnicityHispanic</b><ul style="list-style-type: none"><li>○ Compared to the baseline, we would expect lesbian Hispanic individuals, on average, to score 7.72 percentage points lower in <i>lgbtdiscriminationp</i> when holding all other factors constant.</li></ul></li></ul>	<ul style="list-style-type: none"><li>• <b>sexualorientationlesbian:ethnicityOther</b><ul style="list-style-type: none"><li>○ Compared to the baseline, we would expect lesbian individuals that do not identify as White, Hispanic, or Black, to score, on average, 18.5 percentage points lower in <i>lgbtdiscriminationp</i> when holding all other factors constant.</li></ul></li></ul>

# Interpretation of interaction terms

	White
Gay	<ul style="list-style-type: none"><li>• <b>sexualorientationgay:ethnicityWhite</b><ul style="list-style-type: none"><li>○ Compared to the baseline, we would expect gay White individuals, on average, to score 18.4 percentage points lower in <i>lgbtdiscriminationp</i> when holding all other factors constant.</li></ul></li></ul>
Lesbian	<ul style="list-style-type: none"><li>• <b>sexualorientationlesbian:ethnicityWhite</b><ul style="list-style-type: none"><li>○ Compared to the baseline, we would expect lesbian White individuals, on average, to score 14.5 percentage points lower in <i>lgbtdiscriminationp</i> when holding all other factors constant.</li></ul></li></ul>

# Interaction Plots

Additionally, for all other predictors held constant, does the effect of sexual orientation on percent score of discrimination vary by ethnicity?



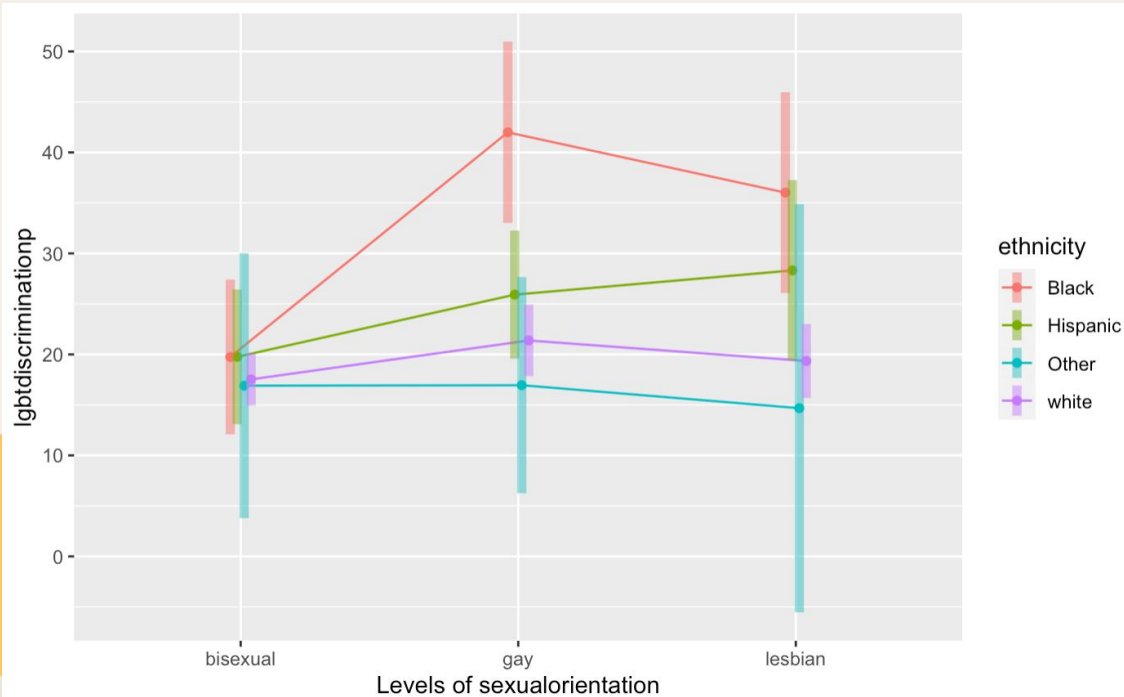
Interaction Effects Significant at a 0.05 Level	Regression Coefficient
sexualorientationgay:ethnicityHispanic	-16.09
sexualorientationgay:ethnicityOther	-22.20
sexualorientationgay:ethnicitywhite	-18.40
sexualorientationlesbian:ethnicitywhite	-14.45

The interaction effect was significant.

In this interaction plot, the lines are not parallel, which suggests that there is in fact an interaction effect between ethnicity and sexual orientation.

# Interaction Plots

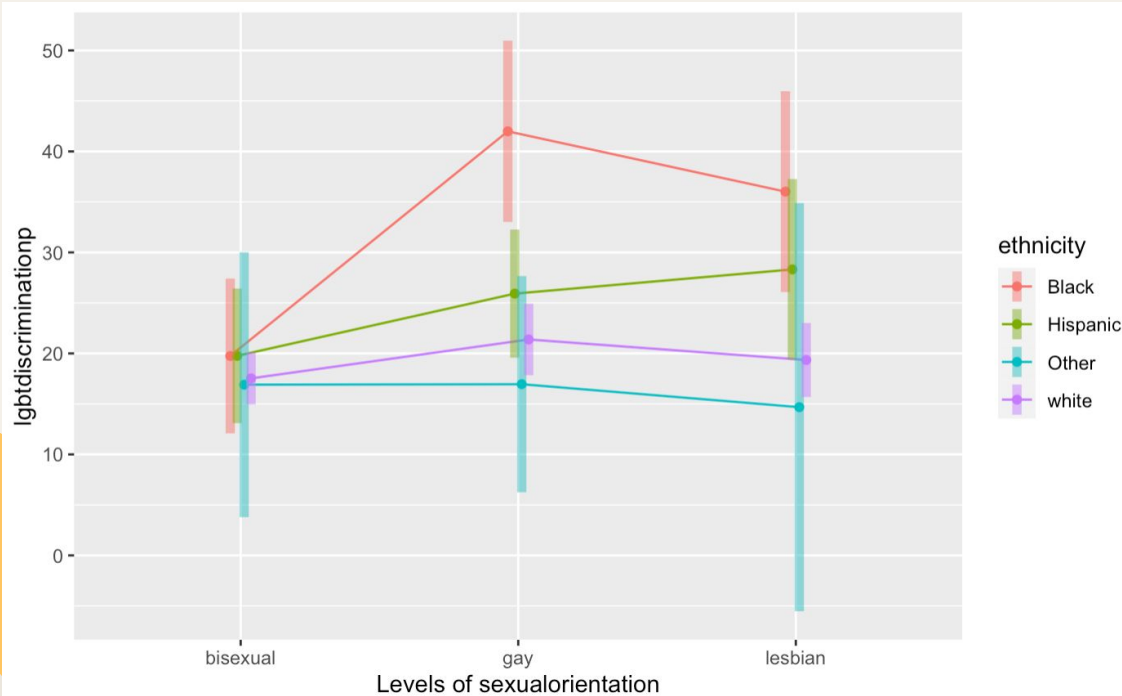
Additionally, for all other predictors held constant, does the effect of sexual orientation on percent score of discrimination vary by ethnicity?



- The effect of sexual orientation on percent score of discrimination varies for different ethnicities.
- The interaction effect is very significant for gay or lesbian individuals that are Black or Hispanic.

# Interaction Plots

Additionally, for all other predictors held constant, does the effect of sexual orientation on percent score of discrimination vary by ethnicity?



- Across all groups, gay Black individuals have the highest percentage score of discrimination.
- It is interesting to note that LGB Black and Hispanic individuals face more discrimination than people that identify as White.

# Interpretation of regression coefficients

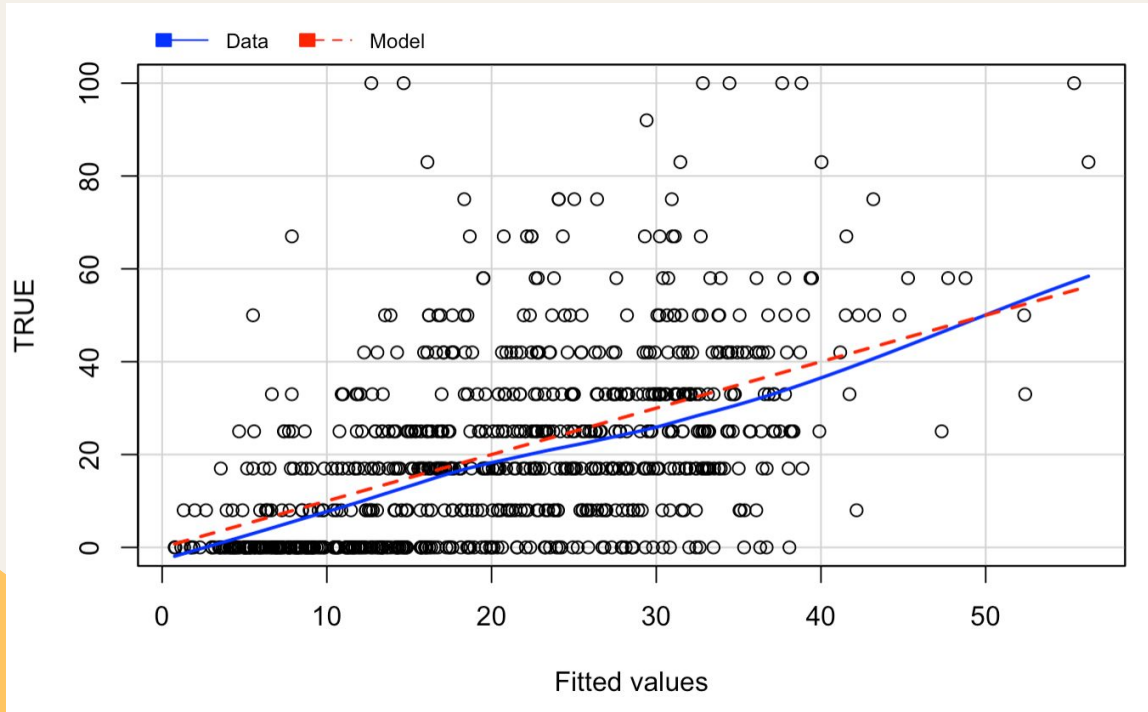
Characteristic	Beta	95% CI <sup>†</sup>	p-value
gender			
female	—	—	
male	-0.72	-5.5, 4.1	0.8
sexualorientation			
bisexual	—	—	
gay	22	10, 34	<0.001
lesbian	16	4.2, 28	0.009
ethnicity			
Black	—	—	
Hispanic	0.02	-9.8, 9.8	>0.9
Other	-2.8	-18, 12	0.7
white	-2.2	-9.9, 5.5	0.6

Characteristic	Beta	95% CI <sup>†</sup>	p-value
reveallgbtsocialnetwork			
no	—	—	
yes	4.7	2.0, 7.4	<0.001
lgbtsocacceptp	-0.03	-0.08, 0.03	0.3
lgbtfriendlyp	-0.13	-0.21, -0.04	0.003
lgbtpolicyp	0.03	-0.11, 0.16	0.7
lgbtsupportp	0.27	0.22, 0.32	<0.001
sexualorientation * ethnicity			
gay * Hispanic	-16	-30, -1.7	0.028
lesbian * Hispanic	-7.7	-24, 8.5	0.3
gay * Other	-22	-43, -1.9	0.032
lesbian * Other	-18	-45, 8.4	0.2
gay * white	-18	-30, -6.5	0.002
lesbian * white	-14	-27, -1.8	0.025

<sup>†</sup> CI = Confidence Interval



# Marginal Model Plot



- Our model is a good fit for the data



**Conclusion**

An illustration featuring a young boy with dark hair, wearing a yellow t-shirt, orange shorts, and blue sneakers with white socks. He stands on a grey shadow on the left side of the frame. In the center is a large yellow circle containing the word "Conclusion" in bold, dark blue text. A dotted line starts from the top left, curves over the circle, and ends with a white paper airplane flying towards the right. The background is light grey with various decorative elements: an orange shape in the top left, a grey cloud in the top right, a small yellow triangle in the top right, a small teal triangle to the left of the circle, and a teal shape in the bottom right corner.

# Conclusions

- We found that the amount of discrimination an LGB individual experiences is influenced by their sexual orientation, whether that person is out on social media, by actions that support the LGBT community and by how friendly major organizations are to LGBT individuals.
- Lesbian, gay, and bisexual individuals are met with implicit bias.
  - Ethnicity adds another layer of bias, as shown in the interaction plot
- According to Jerry Kang's 2016 speech on Race and Diversity, individuals that are perceived to be LGBT are vulnerable to implicit bias. Our model indicates that individuals that report 'Yes' on 'revealLGBTsocialnetwork' are expected, on average, to score higher on feeling discriminated against.

# Conclusions

- Efforts on campus towards being more inclusive and accepting of LGBT community
  - LGBTQ Resource Center
  - LGBTQ studies courses
  - LGBTQ Research Initiatives and Health Programs

**Thanks for listening!**

