

Quantification and Prediction of Gender Identity

Zack Tyler, Joseph Paulson, and Harrison Winslow

Commitment: 10/10, 10/10, and 10/10

1. Description of the project

Gender identity tends to be quite complex and diverse, even among males and females. We will attempt to quantify and predict the gender identities of college students living in Utah using various traits that are not obviously associated with gender. Our goal is to achieve accurate prediction with as few traits as possible while also seeing if there are any interesting correlations. We will survey students at BYU and other universities, as well as non-students of college age, to get their traits and self-described gender identities.

2. What features the data set might include

The goal is for the questions to not be obvious associations (e.g. “Do you prefer dresses or pants?”). We will err on the side of too many features in the data collection so that we can pare them down during the tidying process. Most of our features will be categorical and one-hot encoded. Here are some examples of possible features (see the appendix for more):

GenderImportance	CollegeBelong	PassionIntensity	Neurodivergence	NumStickers
Very much	No	Obsessive	None	0-5

DragonInterest	Leadership	NumSiblings	BirthOrder	GenderIdentity
Low	Rarely	4	Youngest	Male

3. How and from where would the data set be gathered and labeled

The data set will be gathered from college-age adults (ages 18-25) from different campuses, jobs, and hangouts. Some of these will be friends, many will be random people as well. We will use an anonymous online survey application in order to collect the data so that the participant can feel safe to be honest. Each question will have an answer style relevant to the question. The data set will be labeled with the selected gender identity of each participant – Male, Female, or Non-Binary – which they will select themselves as one of the questions in the data set.

Appendix: Potential Survey Questions

#	Question	Potential Answers
1	How much do you relate to or find it easy to interact with those of your assigned gender at birth?	“Not at all” to “Very much”
2	How important is your gender to you?	“Not important” to “Very important”
3	If you are/were a student, what area is your major in?	Agriculture; Business; Education; Engineering; Humanities; Law; Life Sciences; Fine Arts and Communications; Nursing; Physical and Mathematical Sciences; Social Sciences; Other
4	If you are a member of The Church of Jesus Christ of Latter-day Saints, do you feel like you belong in it?	Yes; Somewhat; No; I’m not a member
5	Do you feel like you belong at your college?	Yes; Somewhat; No; I’m not a college student
6	Are you neurodivergent? (If multiple of the answers are true, select the one that most impacts your life.)	Yes, I have a developmental disorder (ASD, ADHD, etc.); Yes, I have a mental health disorder (anxiety, OCD, depression, bipolar, BPD, etc.); Yes, I have a different condition or disorder; I might be, but I’m not sure; No
7	How passionate are you about your interests or hobbies?	I’m not very engaged in any interests; I’m somewhat/occasionally engaged in my interests; I’m fairly/regularly engaged in my interests; I’m obsessive about my interests
8	How many stickers, pins, or keychains do you regularly display, such as on a laptop, water bottle, backpack, etc.?	0-5; 6-20; 21-50; more than 50
9	How interesting are dragons to you?	“Not interesting” to “Very interesting”
10	Do you like to be in charge? / Do you see yourself as a leader?	Never; Sometimes; Often; Always
11	How many close friends do you have?	0-1; 2-3; 4-6; 7+
12	Do you feel discriminated against based on your gender? / Do people treat you differently because of your gender?	Never; Sometimes; Often; Always
13	How many siblings do you have?	0, 1, 2, etc.
14	Where do you fall among your siblings?	Youngest; Middle; Oldest; Only child
15	How old are you?	17 or younger; 18-24; 25 or older
16	Do you lean more towards order or chaos?	Order; Chaos
17	The superior shape?	Square; Triangle; Circle; Star
18	(Optional) Do any of these identities apply to you?	Transgender Male; Transgender Female; Agender; Other (specify)