



FAIRNESS IN COVID-19 PANDEMIC: SEXUAL, RACIAL, AND SOCIO-ECONOMICAL

Team_Baker: Luca Wang, Tracy Zhang, Charlotte Ma, Kerr Tan



PART 1: INTRODUCTION

*Please delete this section before downloading.

DATASET INTRODUCTION

Well-Being and Basic Needs Survey (WBNS)

- 2022-10-19 release date
- Nationally representative
- Internet-based
- non-elderly
- cross-sectional
- topics: health insurance, housing, food security, employment, family income etc

ICPSR | INTER-UNIVERSITY CONSORTIUM FOR
POLITICAL AND SOCIAL RESEARCH

ICPSR 38388

**Well-Being and Basic Needs
Survey, United States, 2020**

Stephen Zuckerman
Urban Institute

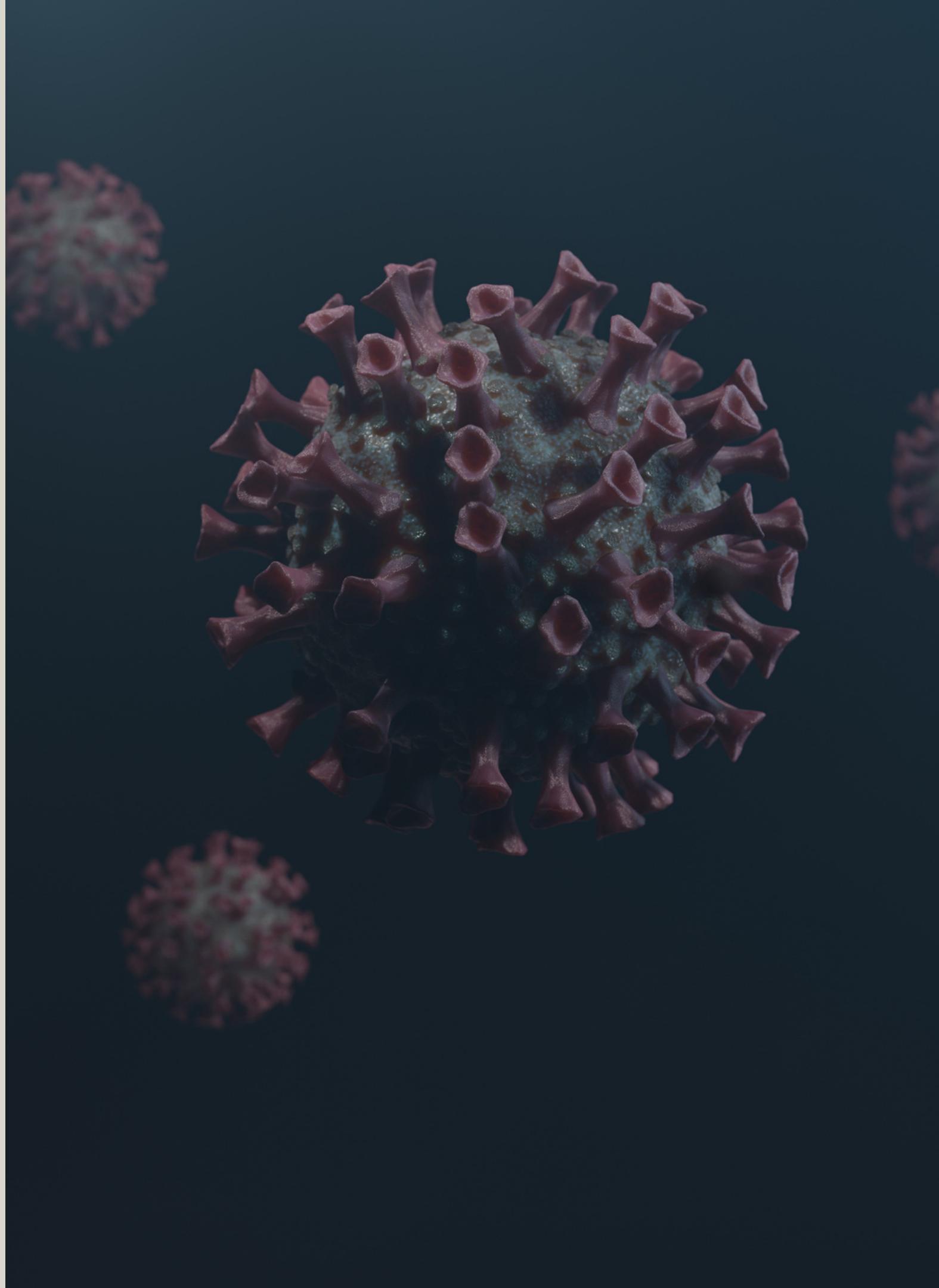
ICPSR Codebook for Public-Use Data

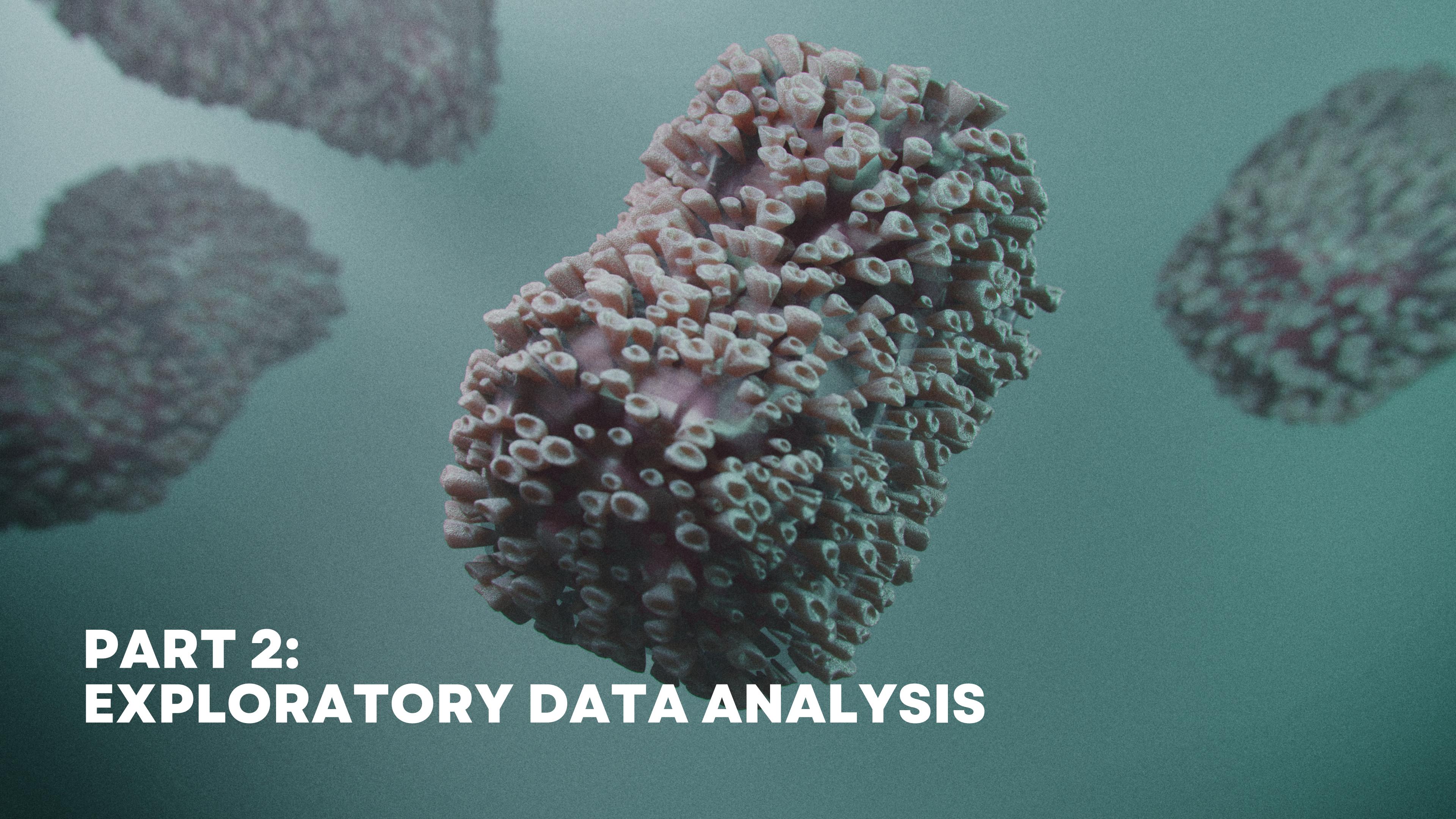
VARIABLES

- ID: Random Unique Identifier
- VARUNIT: Cluster grouping for variance estimation
- Questionnaire Answer: Answers to questions in survey
 - Yes, No, Refused to answer and "."
- Properties: Age, Gender, Income etc

CATEGORIES

- Administrative variables
- Survey questions
- Demographic variables
- Family income
- Insurance Status
- Marital hardship
- School enrollment and child care





PART 2: EXPLORATORY DATA ANALYSIS

BASIC INFORMATION

- RangeIndex: 7737 entries, 0 to 7736
- Columns: 392 entries, ID to ANY_CHCR_NOW_DK
- dtypes: float64(1), int64(188), object(203)

MISSING VALUE

	ID	VARUNIT	WEIGHT	Q1	Q2	Q3	Q5	Q5A	Q96B	Q97	...	ANY_SCH_OTH	ANY_CHCR_PRE_Y	ANY_CHCR_PRE_NO	ANY_CHCR_PRE_DK	ANY
0	17567	20	3.1469	3	2	1	0			4	...		0			
1	13051	16	1.4563	1		2	0	1		1	...		0			
2	14472	4	0.7231	3	2	1	0			4	...					
3	4115	9	0.9195	1		0	3				...					
4	2976	28	0.3872	1		1	6			2	...		0			

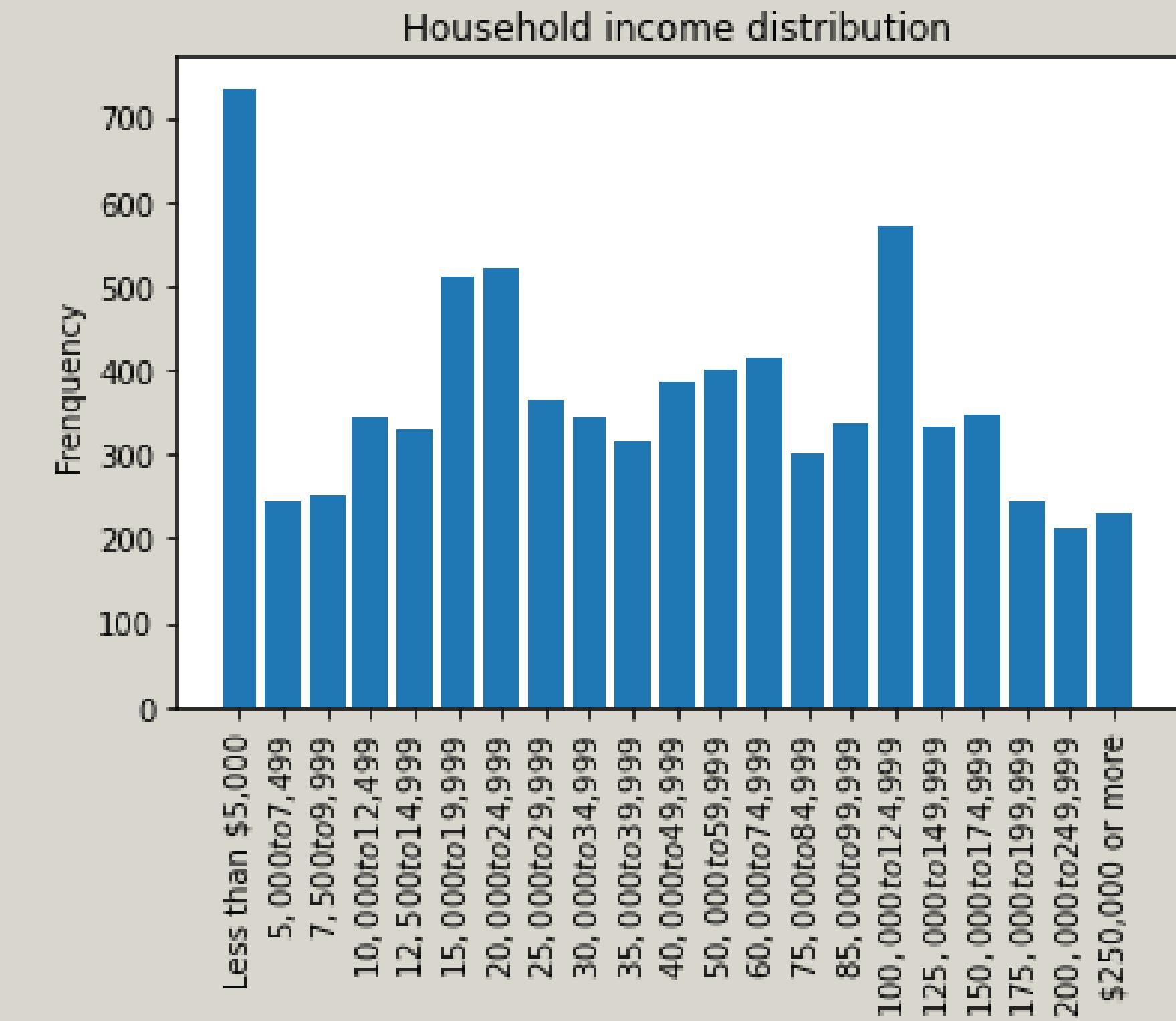
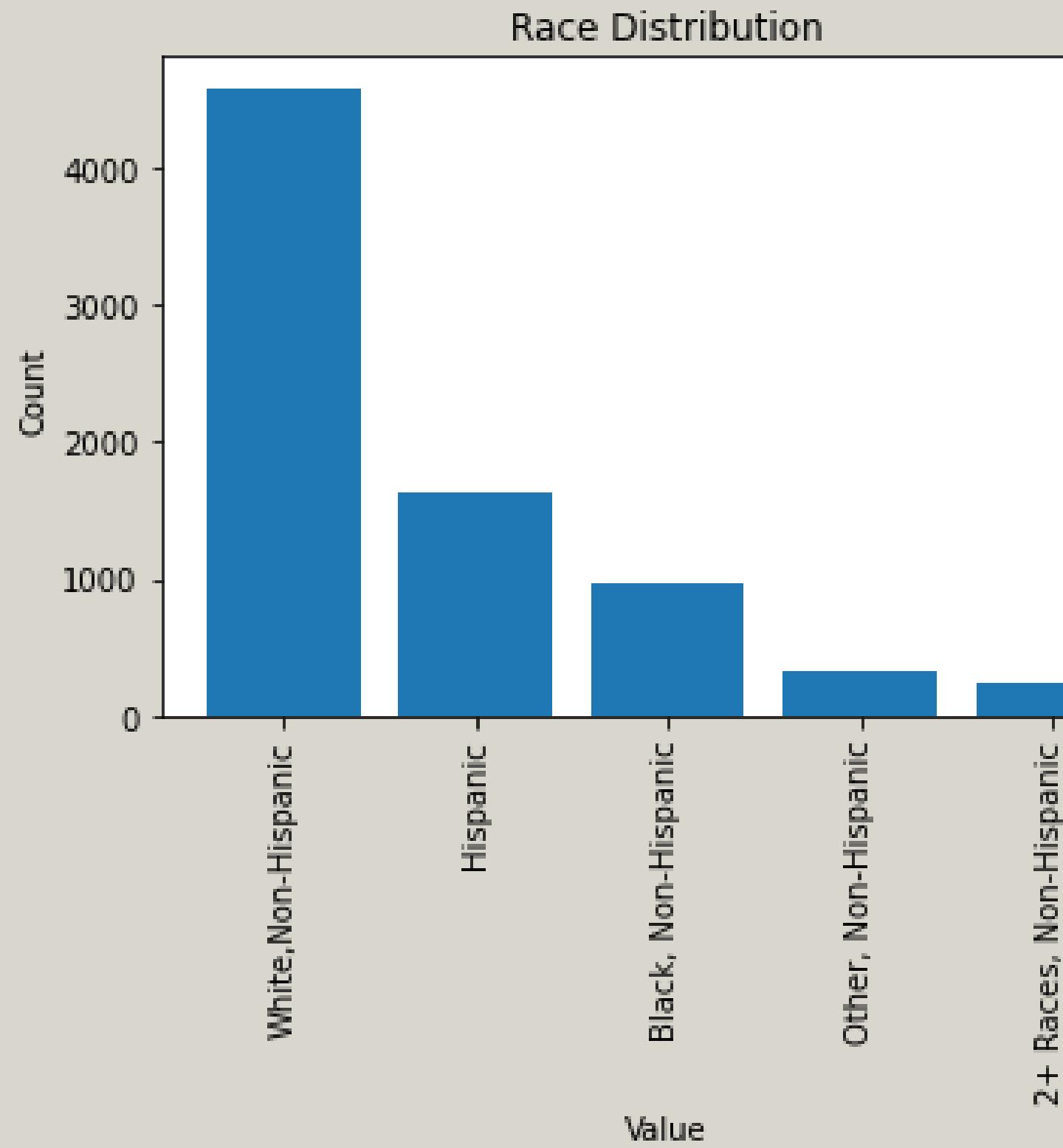
- special notation: "", "-1"
- measure:
 - changed "" to '-2' representing 'no-answer'
 - kept "-1" notation representing "refused"

DATA TRANSFORMATION

```
# convert data type to int (except for the first 3 cols)
for column in df.columns[3:]:
    df[column] = pd.to_numeric(df[column], errors='coerce').astype('Int64')
```

- most column variables are "object"
- measure: change to "Int64"

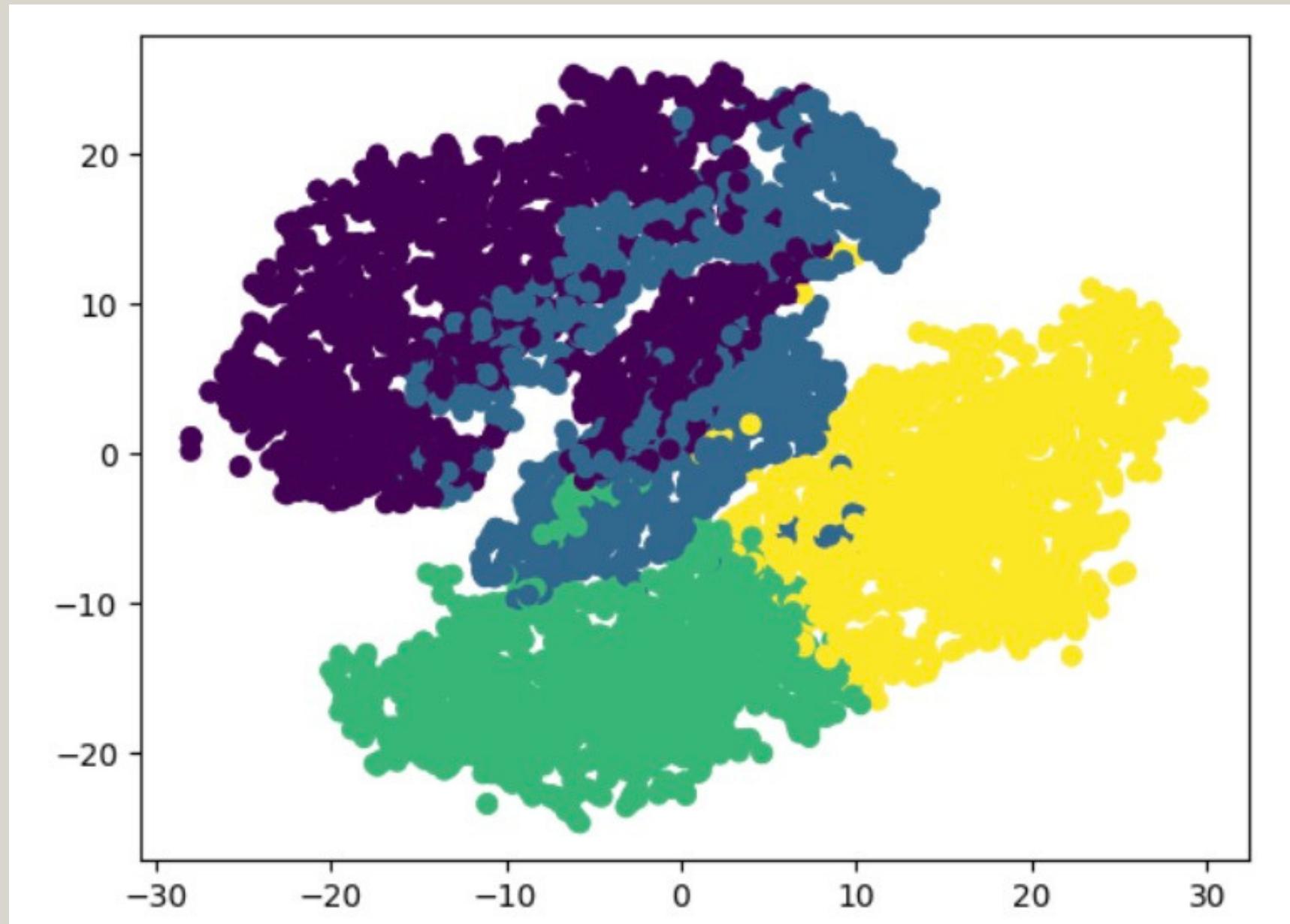
DEMOGRAPHIC DISTRIBUTIONS



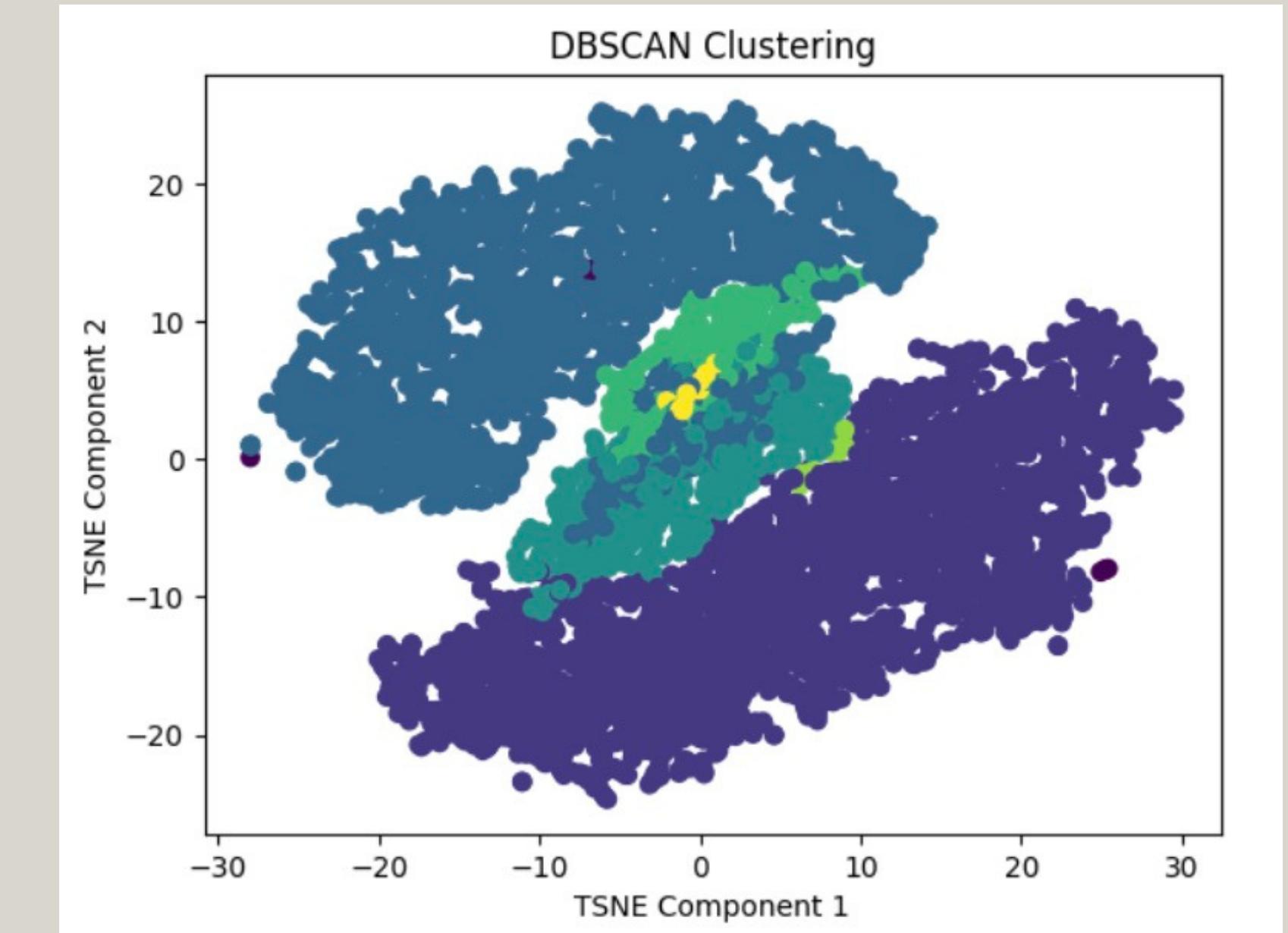
A photograph of two petri dishes containing bacterial cultures on a red agar medium. The dish on the left shows a dense, irregular growth of white bacterial colonies. The dish on the right shows a more organized, linear growth pattern of colonies, possibly indicating a different type of bacteria or a different growth condition. A person's hand is visible at the bottom, holding an open book, which is partially visible on the right side of the frame.

PART 3: IMPLEMENTATION

CLUSTERING (Previous Attempt) on t-SNE (2D Result)

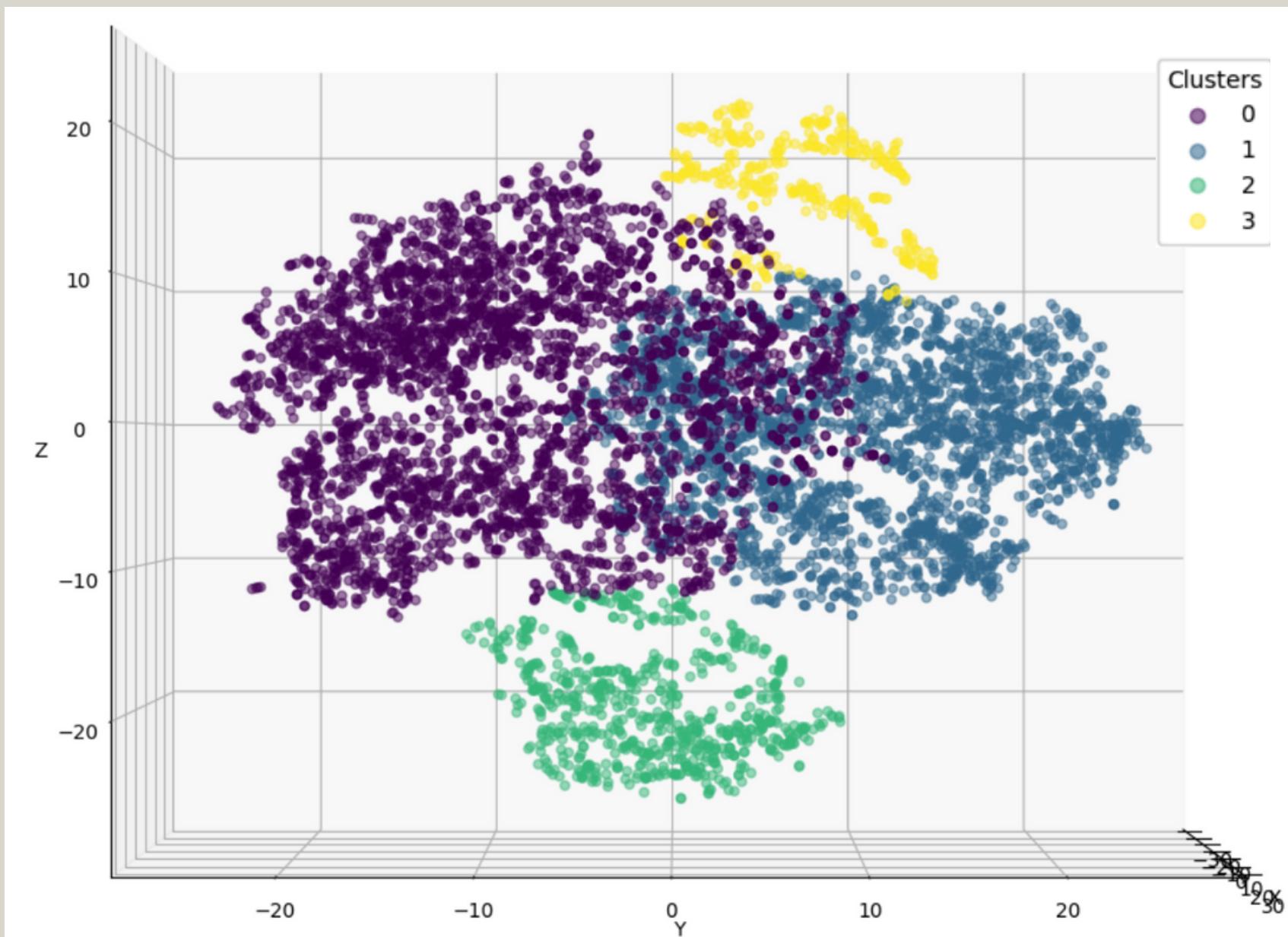
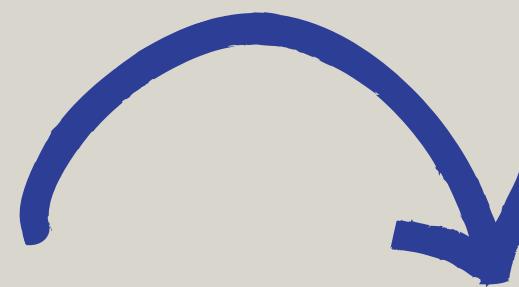


K MEANS

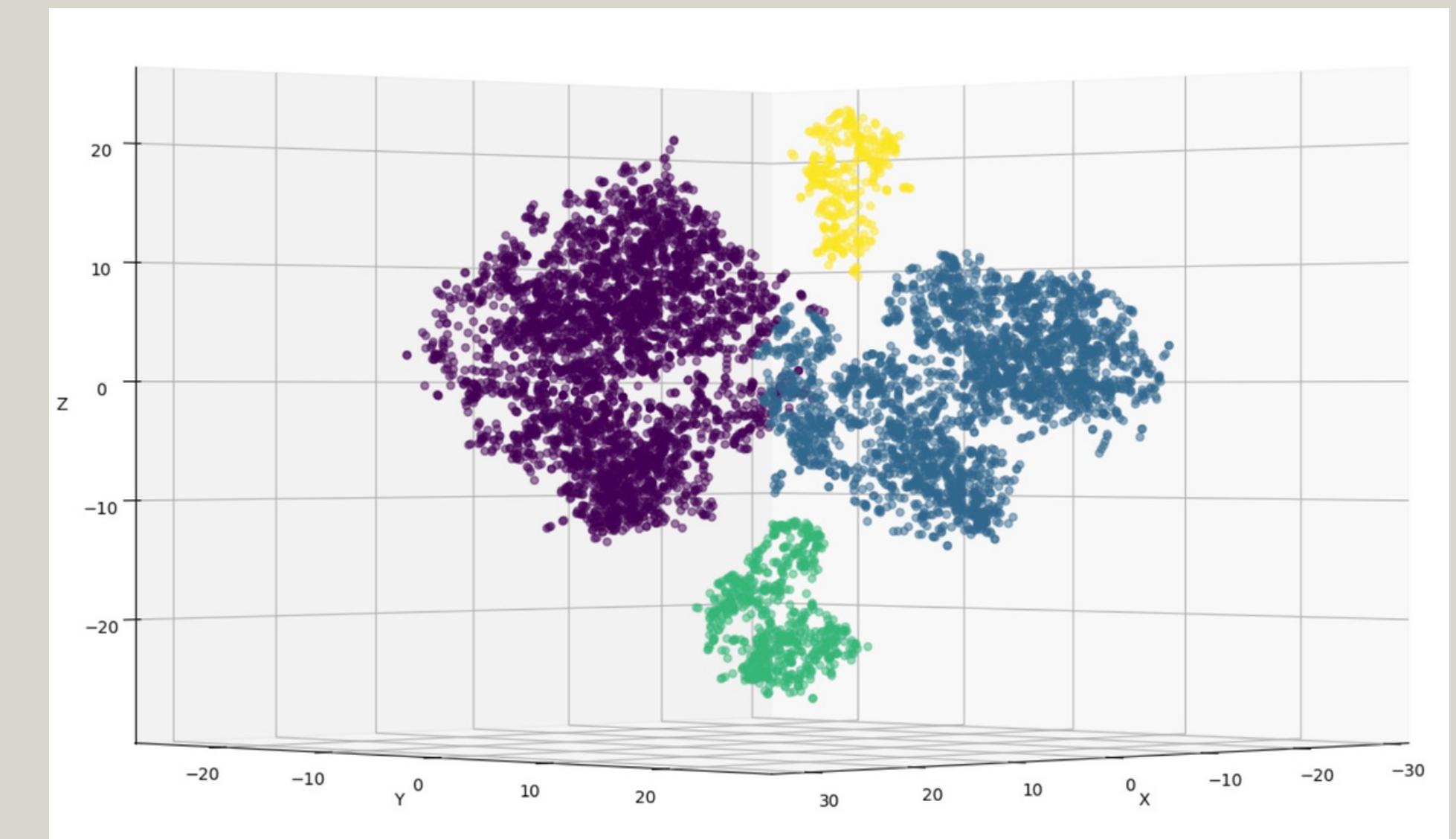


DBSCAN 2D

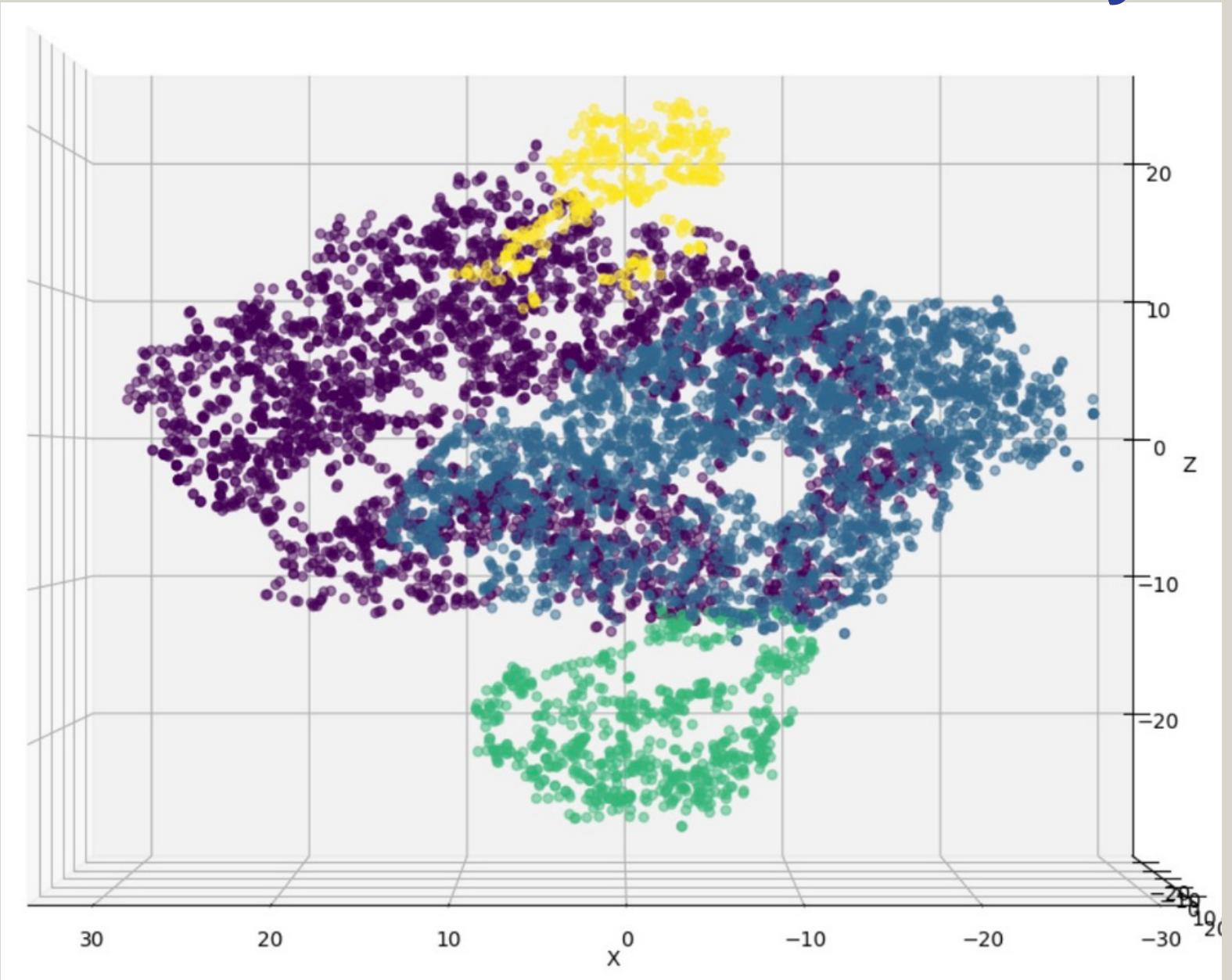
CLUSTERING 3D



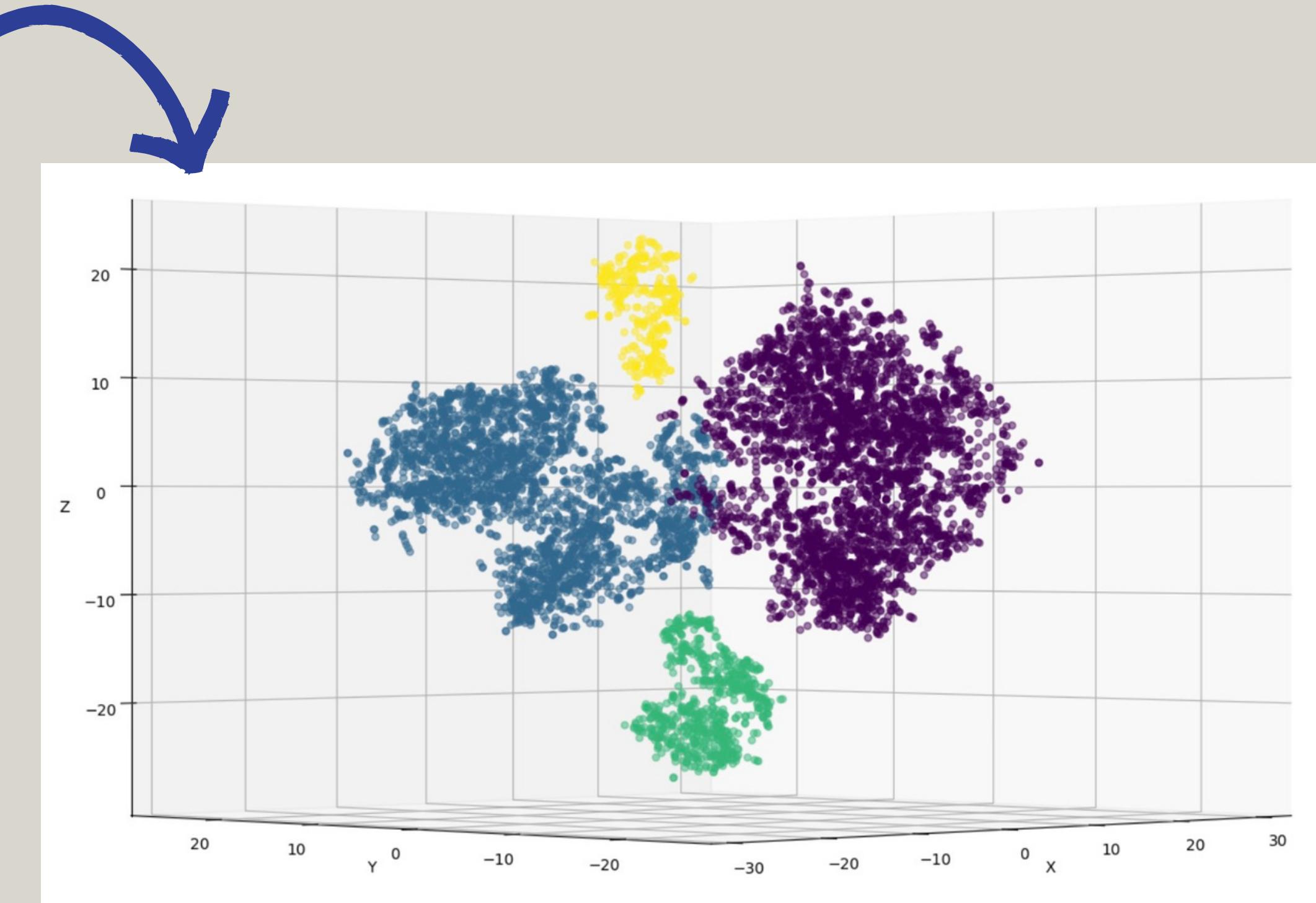
DBSCAN 3D DATA - ZY AXIS



DBSCAN 3D DATA - ZYX AXIS



DBSCAN 3D DATA - XZ AXIS



DBSCAN 3D DATA - ZYX AXIS

PART 4: OUTCOMES & DISCUSSION



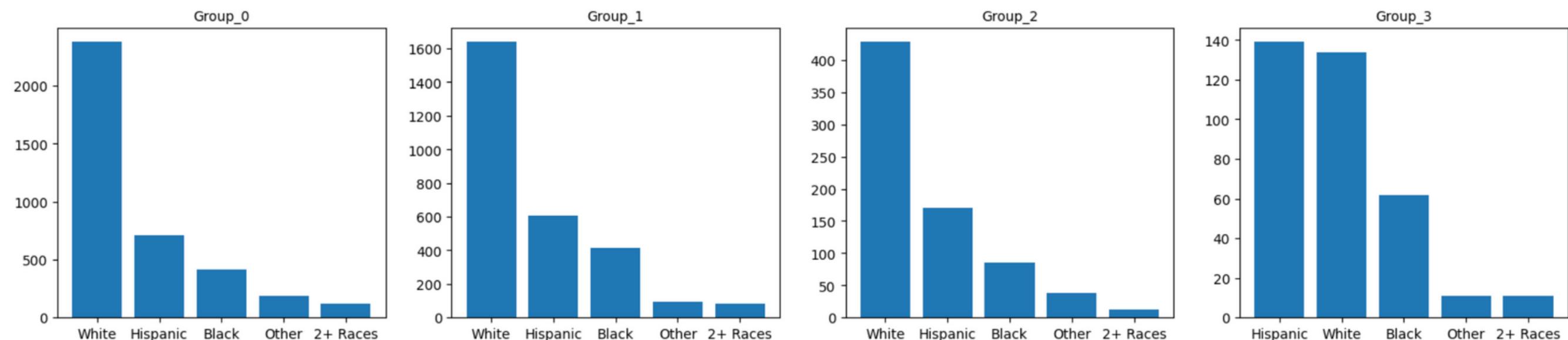
CLUSTERING OUTCOMES

	Q29	Q34	Q35	Q35C	Q35B	Q21	Q70_5	Q102	Q102C_3	Q102C_9	Q12_1	Q12_2	Q37A	Q37B	PPETHM	PPGENDER	PPINCIMP	label	size	
Group_0	3	2	2	-2	-2	3	2	1	2	3	2	2	2	-2	-2	1	1	16	0	3815
Group_1	2	2	2	-2	-2	3	2	1	2	3	2	2	2	-2	1	2	1	1	1	2831
Group_2	3	2	2	2	-2	3	2	2	2	3	2	2	2	-2	-2	1	1	11	2	734
Group_3	3	2	2	-2	-2	3	2	2	2	3	2	2	1	2	4	2	1	3	357	

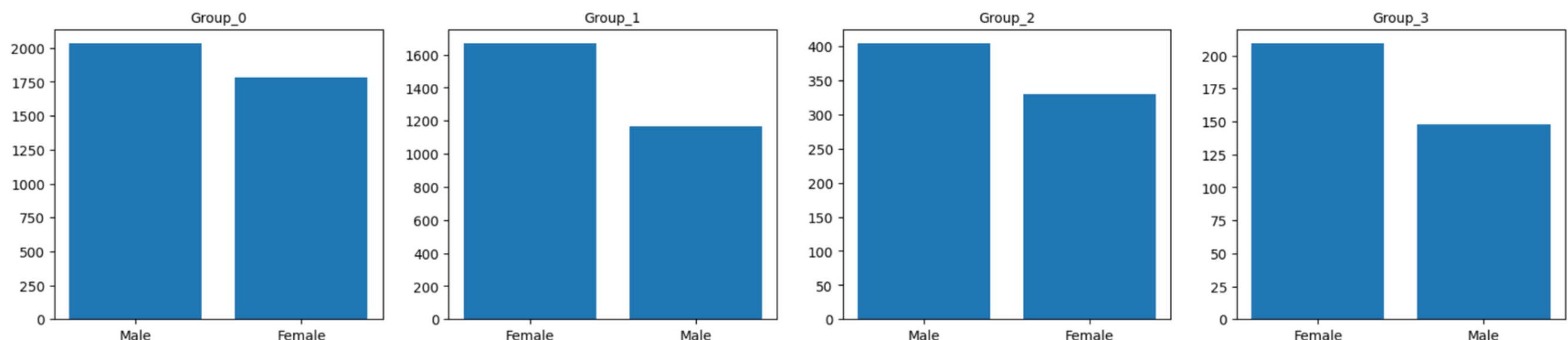
4 clusters with the most common answers for each question

NEXT>>> deeper analysis of each question

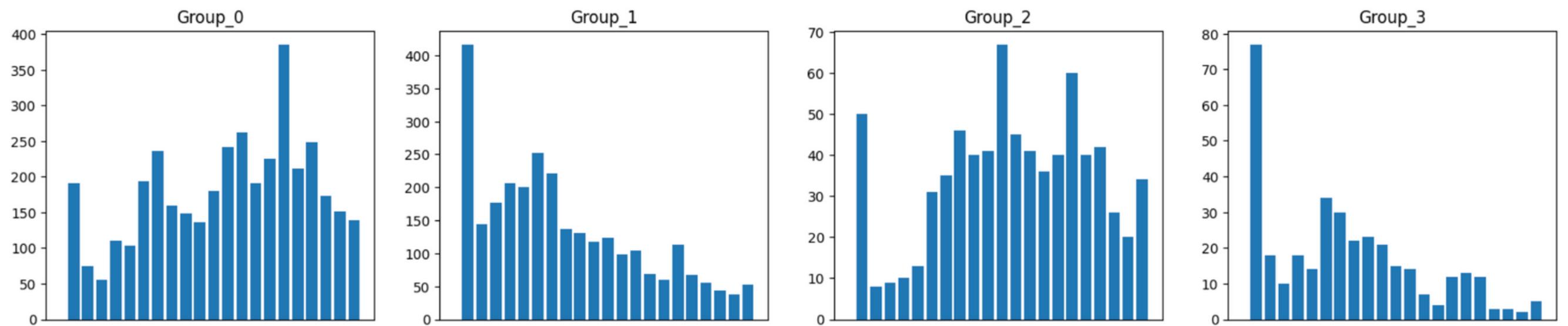
Distribution of Race



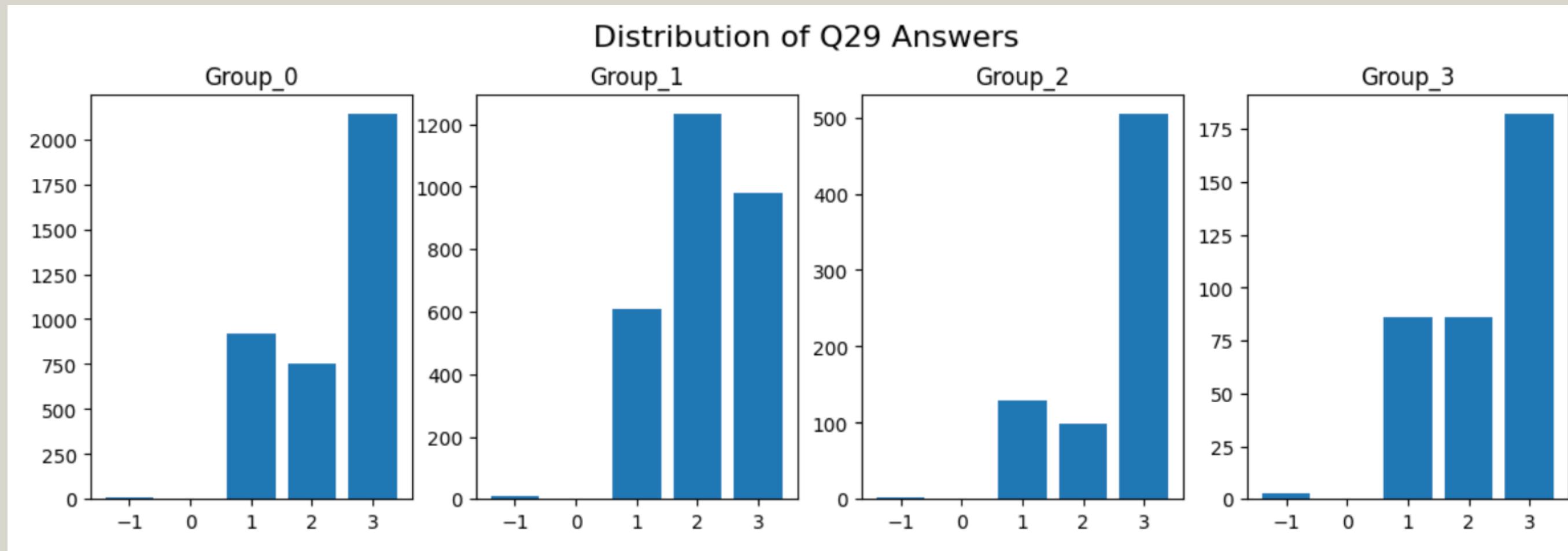
Distribution of Gender



Distribution of Income



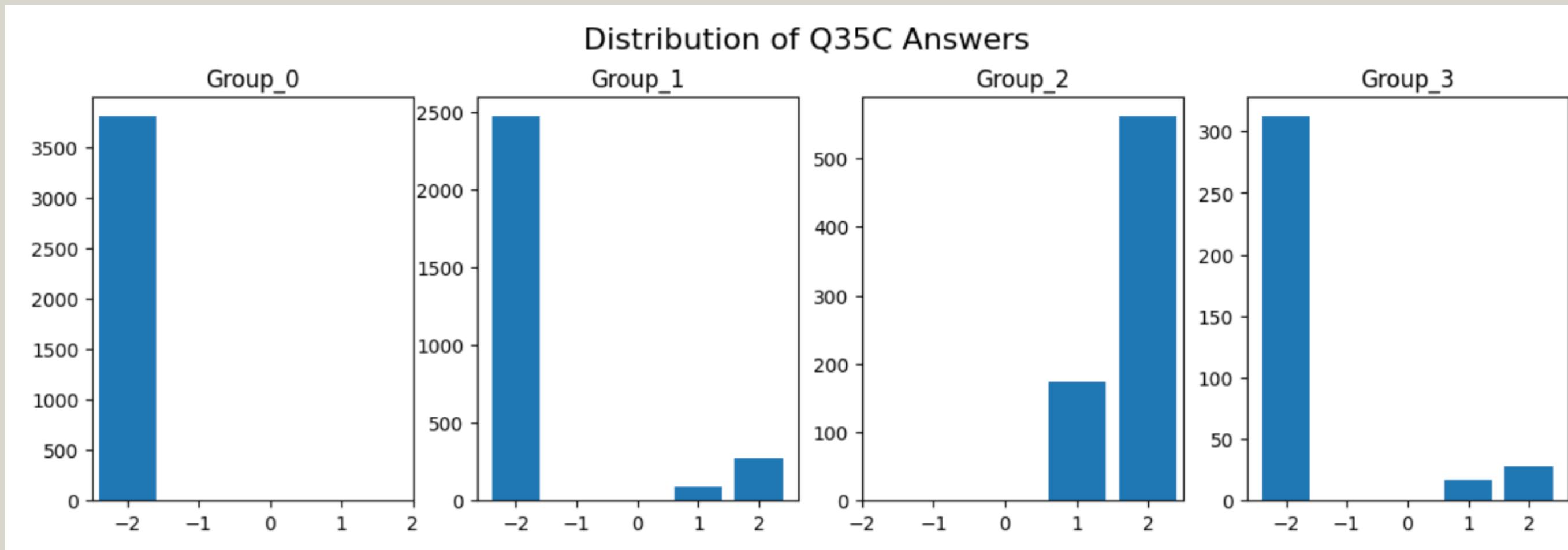
SPECIFIC QUESTIONS



Q29: Do you currently have a health condition that has lasted for a year or more or is expected to last for a year or more?

1: Yes, one condition; 2: Yes, more than 1 condition; 3: No -1: Refused

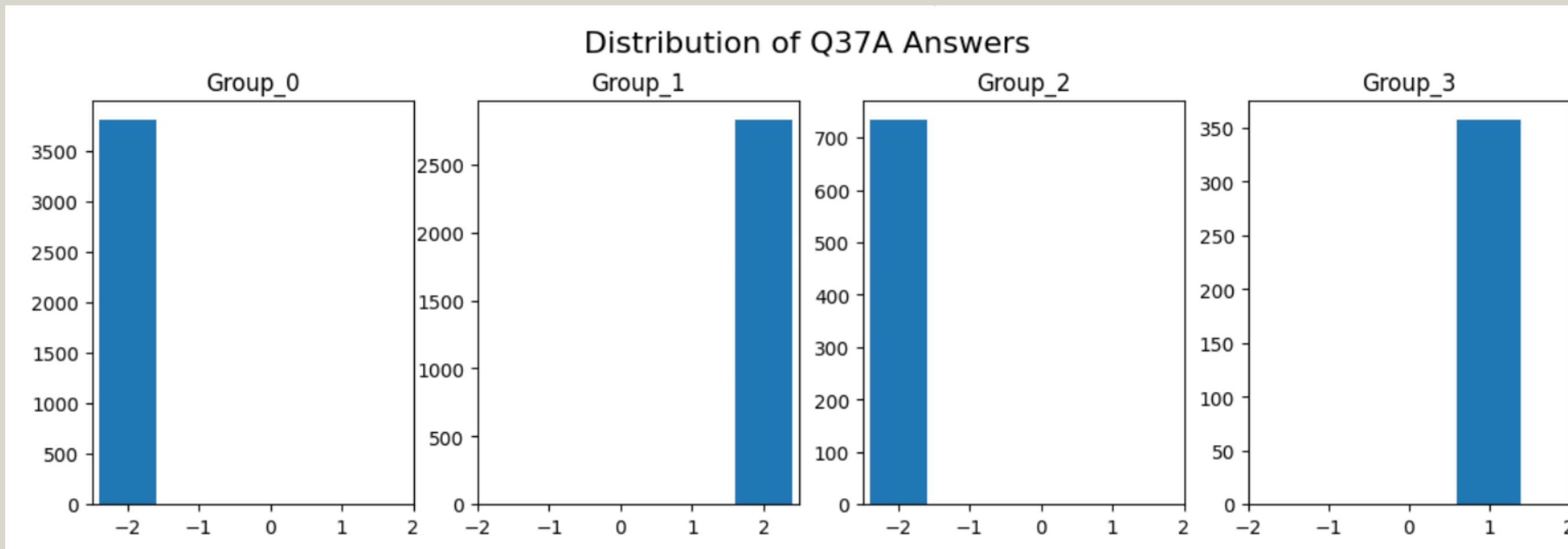
Only Group 1, which is mostly consisted of low-income, female, and White American, has 2 as the most answer to this question.



Q35C: Since the coronavirus outbreak began, have any of the children under 6 who are living with you missed a well-child check-up, that is, a general check-up when they were not sick or injured?

-1: Yes; -2: not answered; 2: No

Only Group 2, which is mostly consisted of male, above-average-income, and white, replied 2 the most



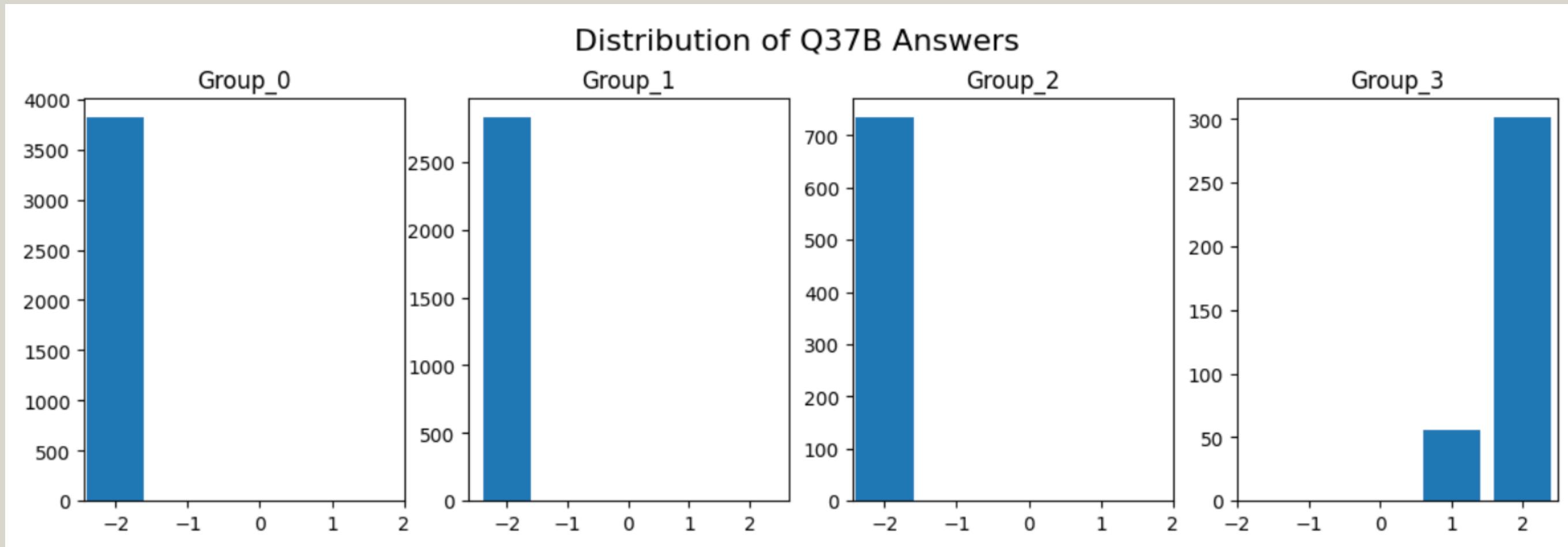
Q37A: Are you on temporary layoff from a job?

-2: not answered; -1: Refused; 1: Yes, 2: No

Group 0 and 2, which are mostly consisted of higher than average income male, mostly did not answer this question.

Group 1, which is mostly consisted of low income female and white, mostly said no.

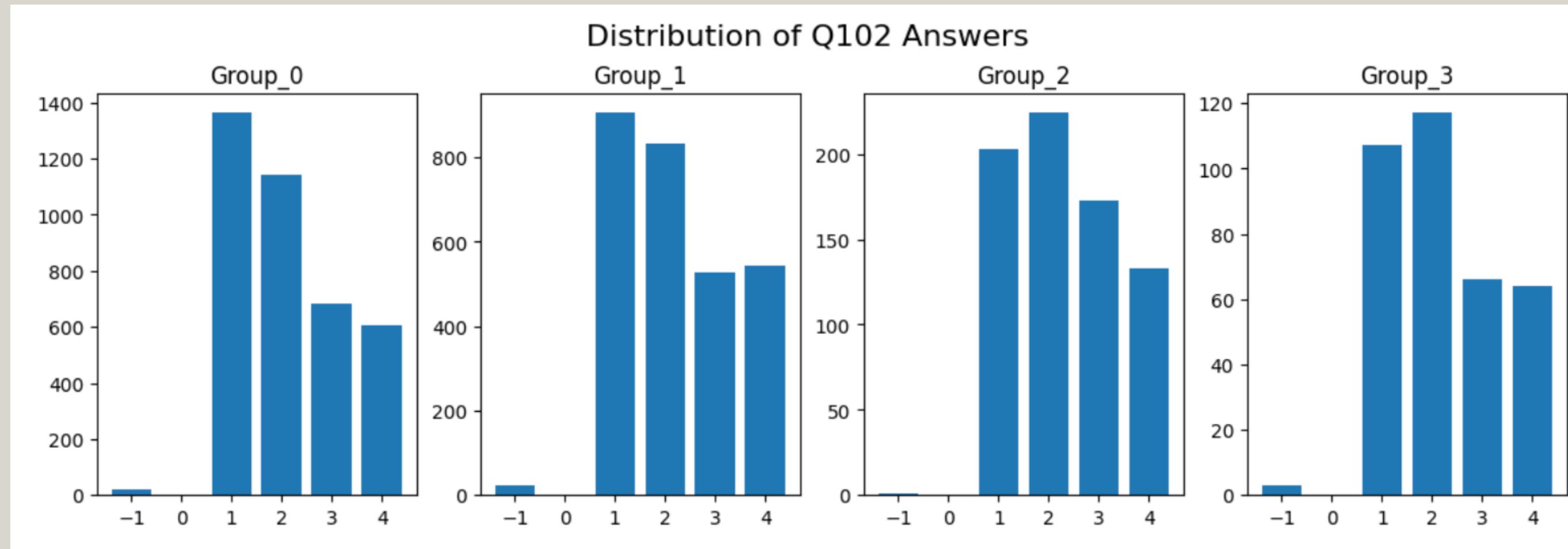
Group 3, which is also mostly consisted of low income female, but most are hispanic, but having a smaller size than Group 1, mostly said Yes.



Q37B: Has your employer either given you a date to return to work or indicated you will be recalled to work within the next 6 months?

-2: not answered; 2: No; 1:Yes; -1: Refused;

Group 0, 1 and 2 mostly did not answer to this question, which are all White Americans. However, Group 3, which is mostly consisted of low income female and hispanic, said No.

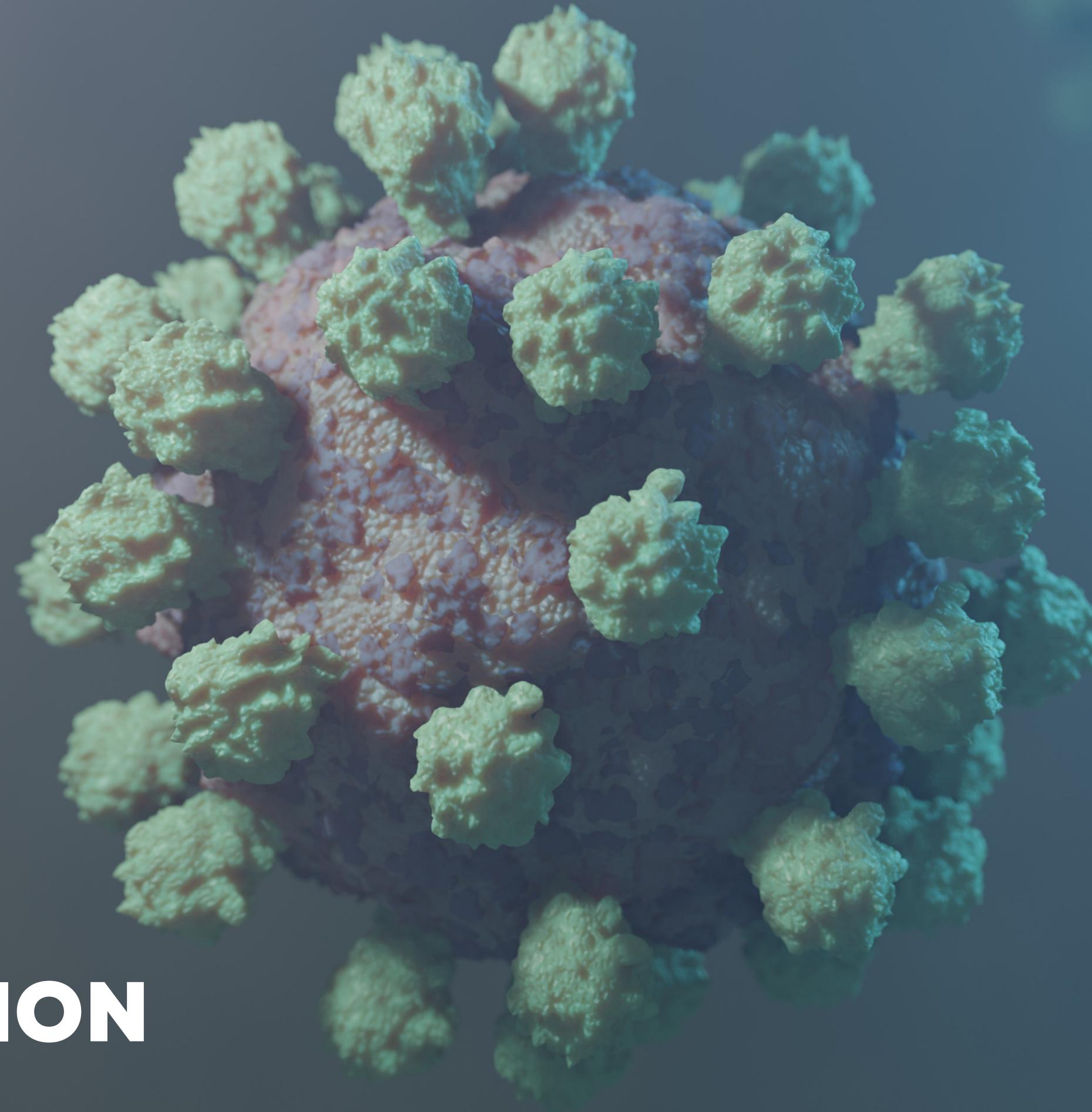


Q102: If a vaccine that protected you from the coronavirus was available for free to everyone who wanted it, how likely would you be to get it?

1: definitely get it; 2: probably get it; 3: Would probably NOT get the vaccine; 4: Would definitely NOT get the vaccine; -1: Refused

Group 0 and 1 mostly said 1; Group 2 and 3 mostly said 2, where group 0 and 1 are mostly white American, and group 2 and 3 have more hispanic

PART 5: CONCLUSION





REFERENCE

ZUCKERMAN, S. (2022, OCTOBER 19). WELL-BEING AND BASIC NEEDS SURVEY, UNITED STATES, 2020. WELL-BEING AND BASIC NEEDS SURVEY, UNITED STATES, 2020. RETRIEVED APRIL 30, 2023, FROM [HTTPS://WWW.ICPSR.UMICH.EDU/WEB/ICPSR/STUDIES/38388/DATADOCUMENTATION#](https://www.icpsr.umich.edu/web/ICPSR/studies/38388/datadocumentation#)