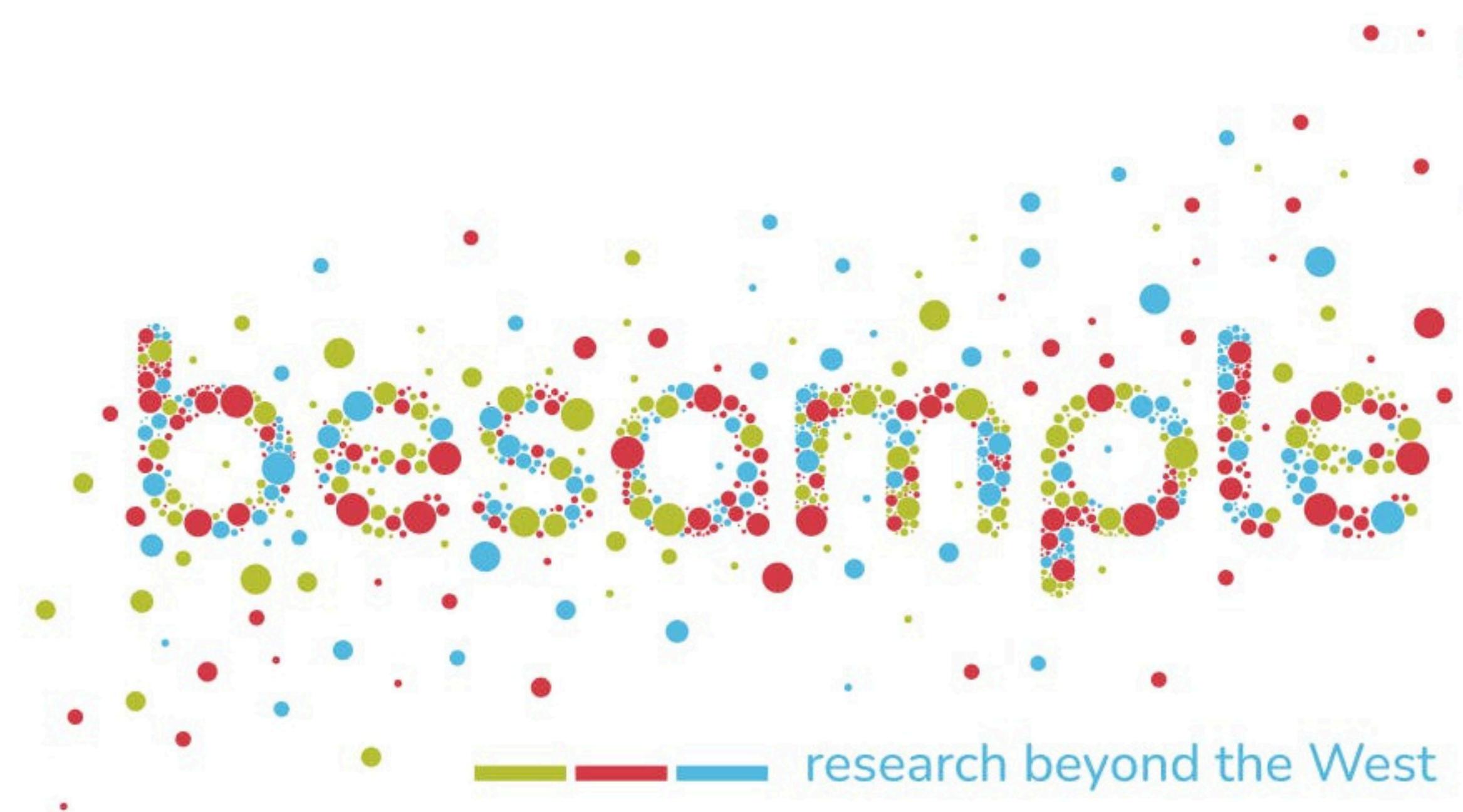


BESAMPLE

BOT DETECTION FOR ACCURATE SURVEY RESULTS

Zoey Espinoza
TripleTen Externship 2024



"BeSample is a new tool designed to empower human-focused scientists to study participants beyond Western, educated, industrialized, rich, democratic (WEIRD) populations. Ninety-six percent of research studies only use American or European participants, which does not accurately represent the diversity of humanity."

BeSample provides a fast, reliable, and affordable way for researchers to collect data directly from respondents in 42 countries across Africa, Asia, South America, and Eastern Europe."

PROJECT

Aims to improve the reliability of survey results across diverse global populations by addressing the issue of **automated bots, which compromises data integrity.**



**Detect and analyze data patterns
that signify potential bot activity or
suspicious behavior.**

METHODS

Datasets used:

- profile.csv
- users.csv
- users_event_log.csv

Surveys Examined:

- demography survey
- eyal survey



**Exploratory Data
Analysis and Data
Preprocessing**

**Correlation
Analysis**

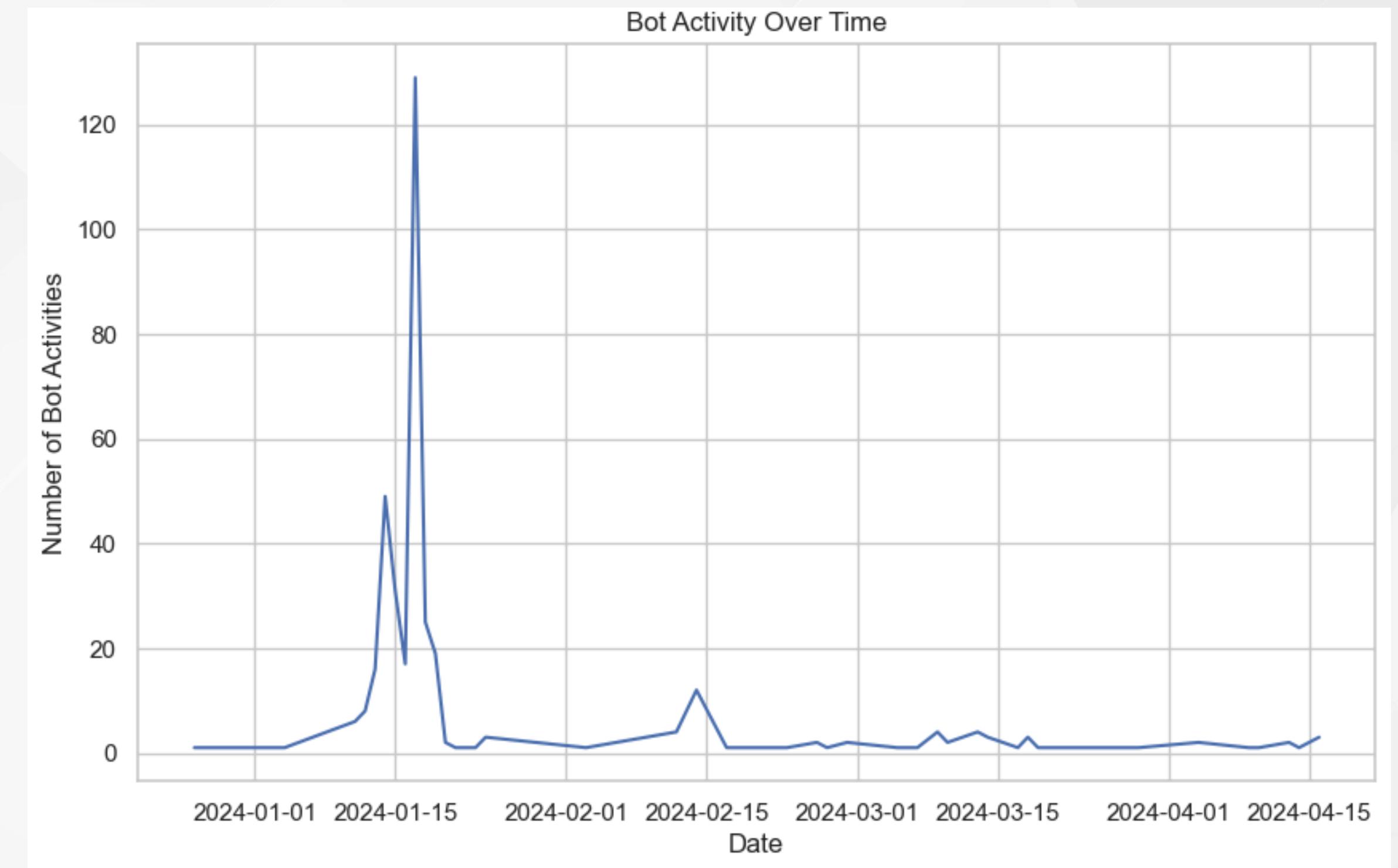
**Statistical Analysis
/ Hypothesis
Testing**

**Feature
Engineering and
Label Encoding**

Data Visualization

**Suggestions for
Future Proofing**

THE ATTACK BOT INFILTRATION

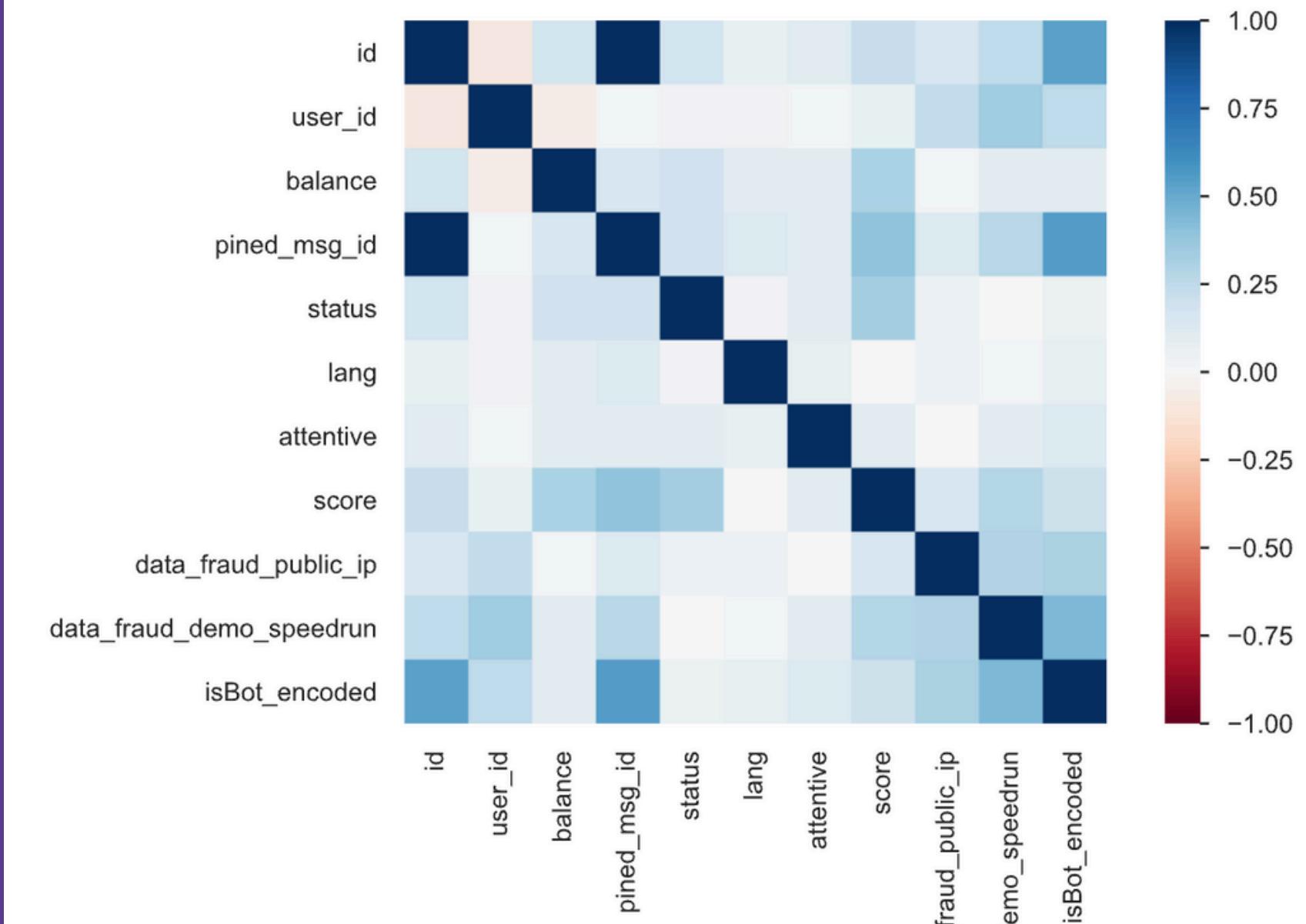


During the “Bot Attack”, between January 13th, 2024 – January 19th, 2024, **281 bots were labeled**. Then later another smaller attack **February 14th, 2024**, and some other activity resulting in a total of **370 labeled Bots**.

CORRELATION ANALYSIS

The "**isBot**" column is analyzed as the initial step in identifying potential bots among users to find patterns of automated activity.

data_fraud_public_ip and **data_fraud_demo_speedrun** have notable positive correlations with **isBot_encoded**, suggesting predictions for bot behavior.

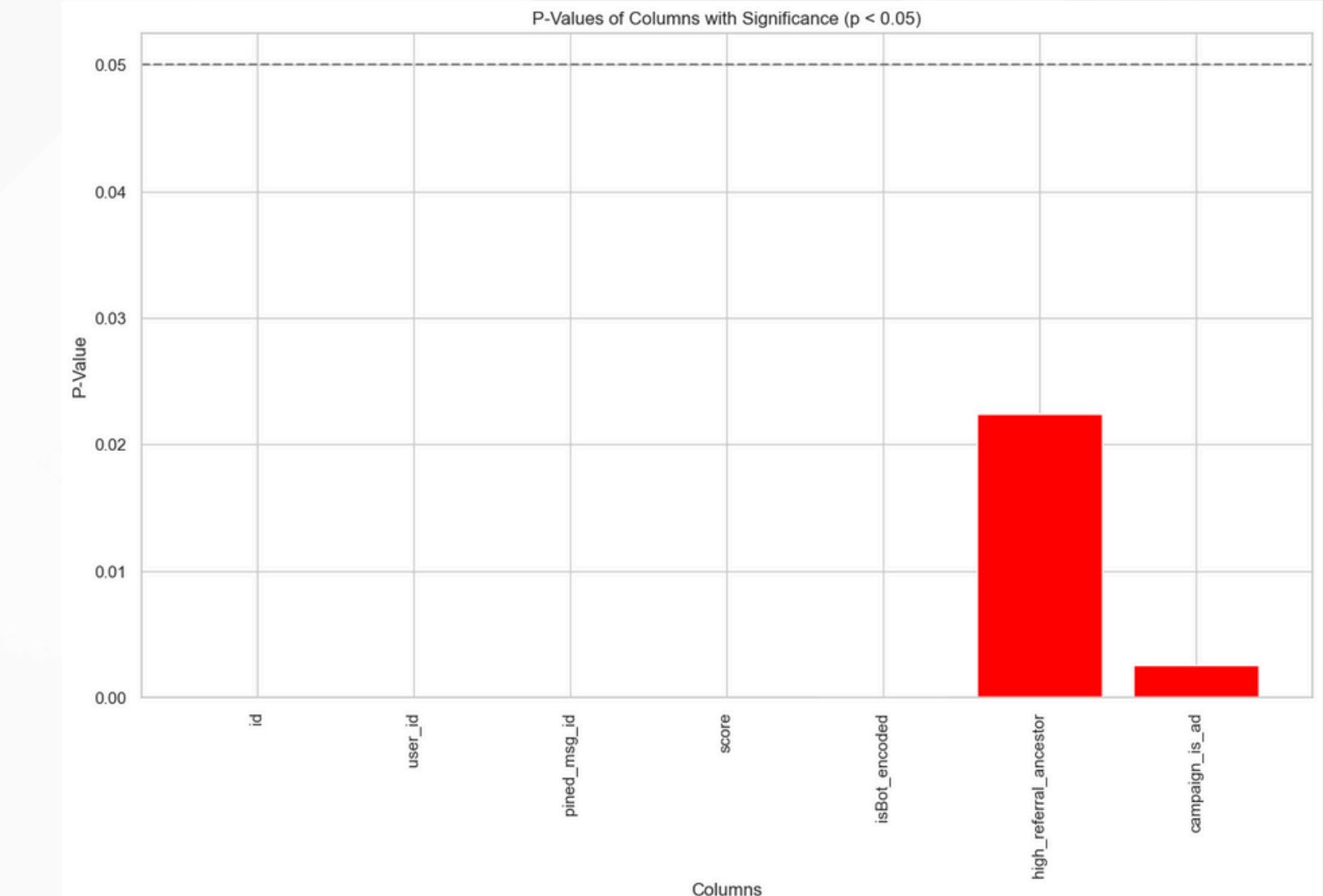


CHANGES AFTER THE BOT ATTACK?

Assuming data integrity pre-attack, we use **t-tests** to compare feature distributions **before and after 01-13-24**, revealing potential **anomalies**.

A p-value below 0.05 indicates significant change in the skewed data.

Primary influenced:
high_referral_ancestors
ad campaign referrals



IDENTIFYING UNUSUAL PATTERNS

BEHAVIORS THAT DEVIATE FROM THE NORM

Hard Facts

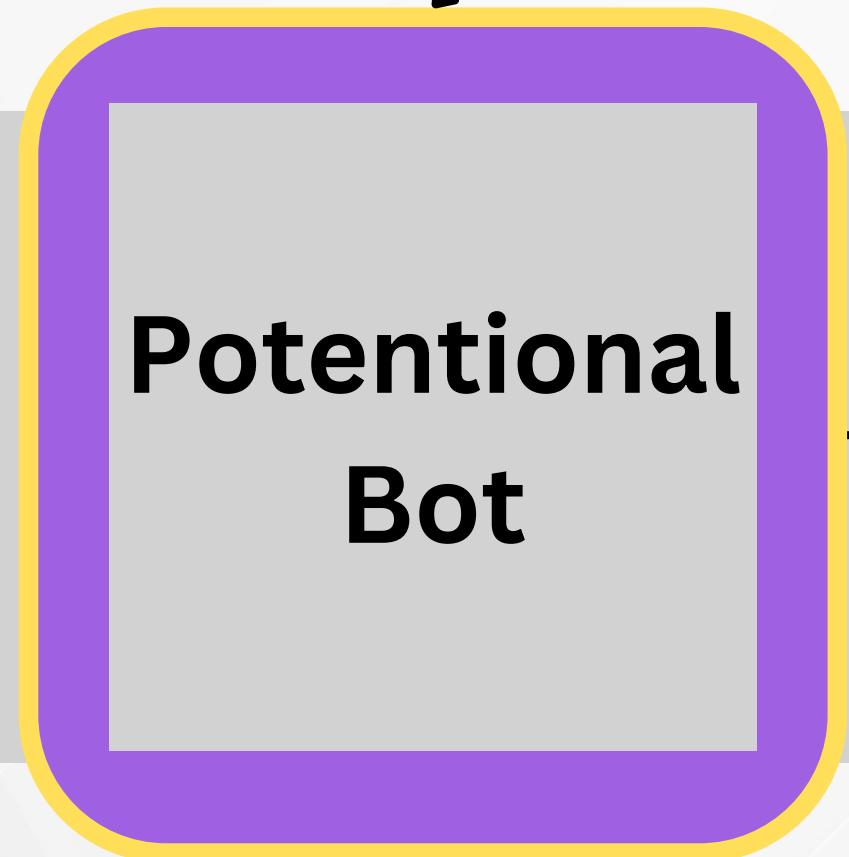
Evidence

Soft Facts

Ambiguity

Potential
Bot

Suspicious
User



HARD FACTS

Public IP Fraud: Flags fraudulent activity associated with public IP addresses.

Frequent IP: Identifies IP addresses used excessively, over the average of one IP per user, with 3 as a limit.

High Referral Ancestor: Indicates an abnormally high number of referrals from a single ancestor.

Fraud Score: Fraud score provided. Flagged above the upper bound IQR.

Speedrun Fraud: Flags users completing surveys too quickly, above upper IQR.

COLUMN	COUNT
public_ip	310
frequent_ip	260
high_referral_ancestor	237
fraud_score	39
speedrun	21

SOFT FACTS

- **Age:** Flags ages deemed too high (≤ 100) or too low (< 18).
- **Balance:** Identifies values above the upper bound of the Interquartile Range (IQR).
- **Duration:** Detects values below the lower bound of the IQR.
- **Recaptcha Score:** Flags values outside the IQR.

Survey question Inconsistencies:

- **Age:** Flags inconsistencies in reported age, age range, and birthdate, with a one-year leniency.
- **Country:** Detects variations in responses across 3 questions, comparing answers to those selecting "Ukraine".
- **Children:** Identifies inconsistencies between the number of children and those under 18.
- **Education:** Highlights disparities in responses to education-related questions.

COLUMN	COUNT
inconsistent_age	148
inconsistent_children	143
inconsistent_education	115
inconsistent_country	112
high_balance	68
high_recaptcha	21
age_flag	18
fast_duration	3

FEATURE ENGINEERING

```
columns_to_keep_from_users = [  
    'user_id', 'ip', 'status', 'created_at', 'attentive', 'balance',  
    'pined_msg_id', 'came_from', 'isBot',  
    'repeated_ip', 'high_referral_ancestor_encoded',  
    'data_fraud_referral_ancestor', 'data_fraud_public_ip', 'data_fraud_demo_speedrun']
```

```
columns_to_keep_from_profile = [  
    'user_id','country', 'fraud_score', 'contradictory_country_answers',  
    'demography_body_values.duration','eyal_answers_values.duration', 'education_mismatch', 'QID11_age',  
    'not_within_range_QID11', 'not_within_range_QID37', 'age_difference_exceeds_1', 'contradictory_children']
```

```
columns_to_keep_from_uel = ['user_id', 'event', 'recaptcha_score', 'proxy_used']
```

Condition: (actual labels are encoded with 1:True, 0:False)

```
condition = (  
    (df_profile['country'] == Ukraine) &  
    (df_profile['eyal_answers_labels.QID32'] == Ukraine) &  
    (df_profile['demography_body_labels.QID47'] == Yes, in Ukraine Now)  
)  
df_profile['contradictory_country_answers'] = ~condition
```

112

INCONSISTANT COUNTRY RESIDENCE

Demography Survey

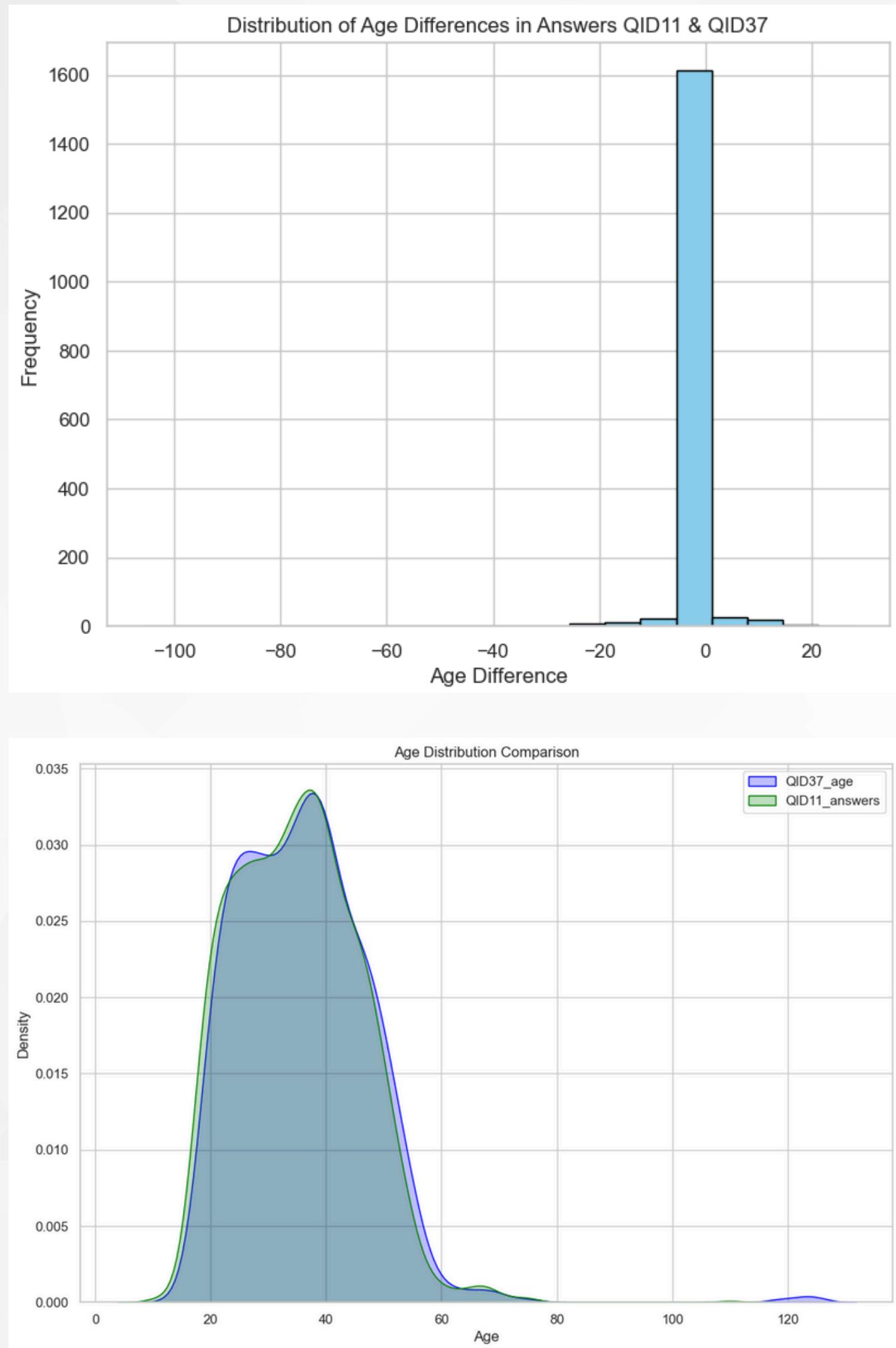
QID47 - "Are you in Ukraine now?"

QID15 - "What is your current country of residence?"

df_profile['country'] - dropdown selection

Eyal Survey

QID32 - "Which country do you currently reside in?"



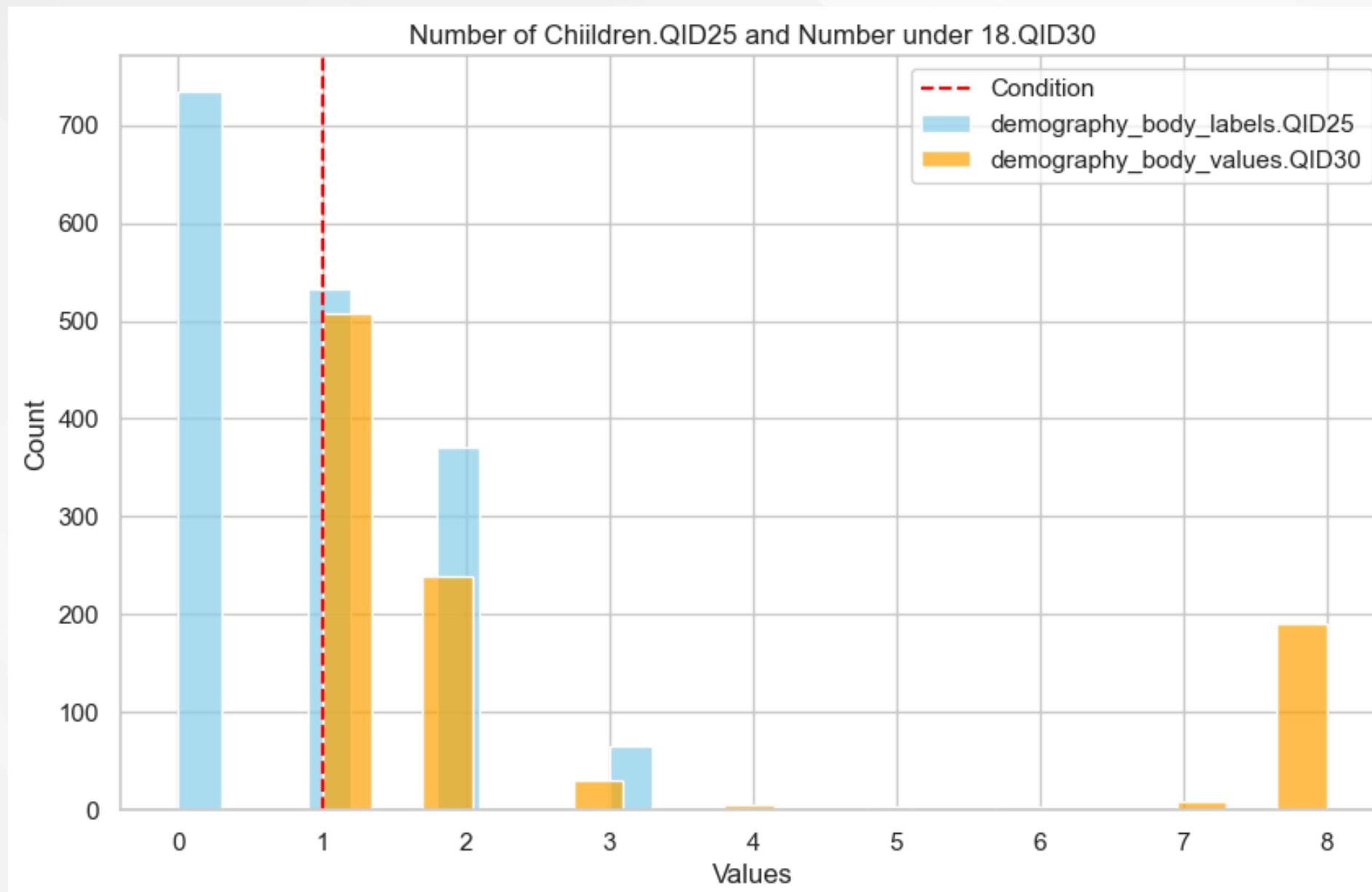
157

INCONSISTANT

AGE

Demography Survey
 QID14_Text - How old are you?
 Please enter your age in years.
 QID37 - When is your birthdate?
 QID46 - What is your age range?

Eyal Survey
 QID11 - "How old are you?"



143

INCONSISTANT **NUMBER OF CHILDREN**

Demography Survey
QID25 - How many children do you have?
QID30 - How many of them are under 18 years old?

Demography Survey

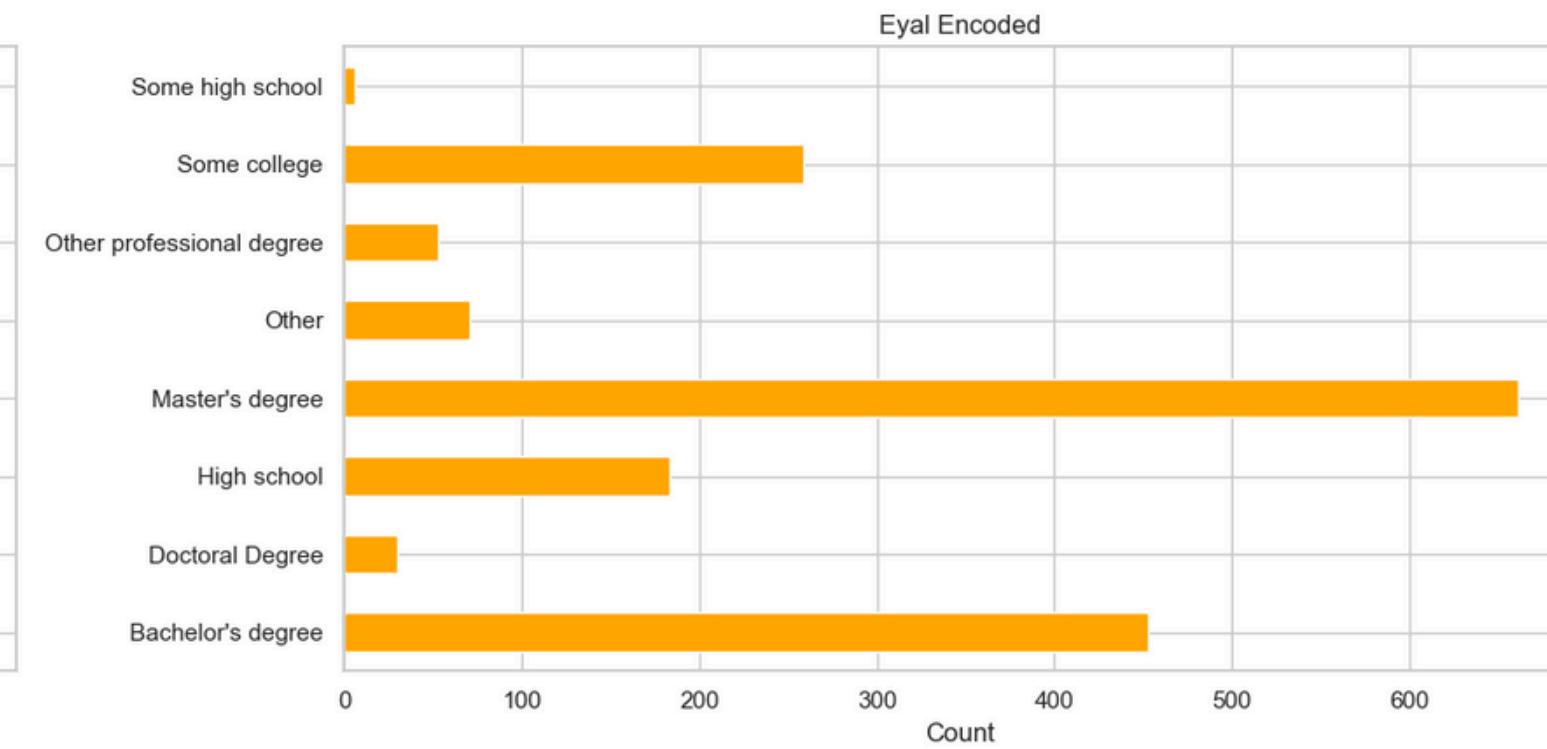
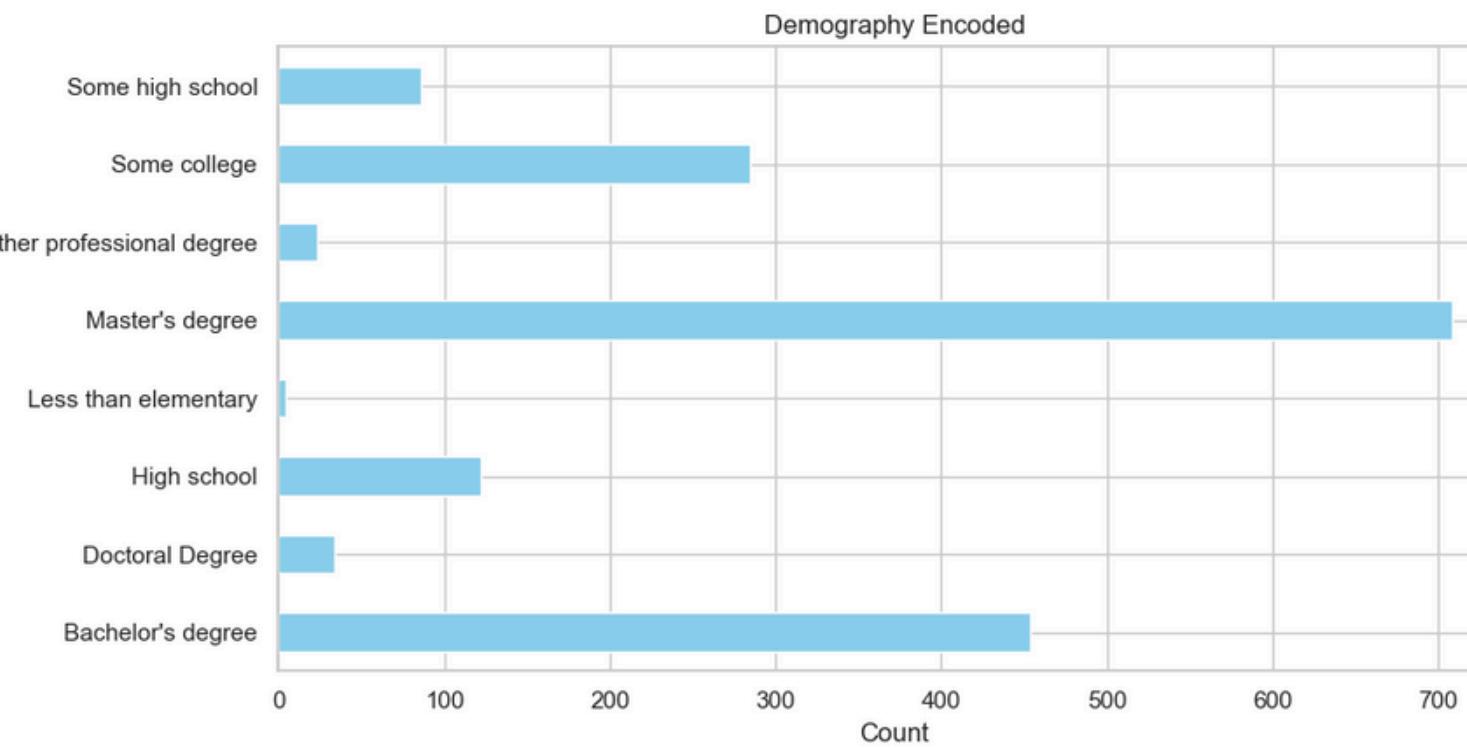
QID4 - What is the highest level of education you have received?

Eyal Survey

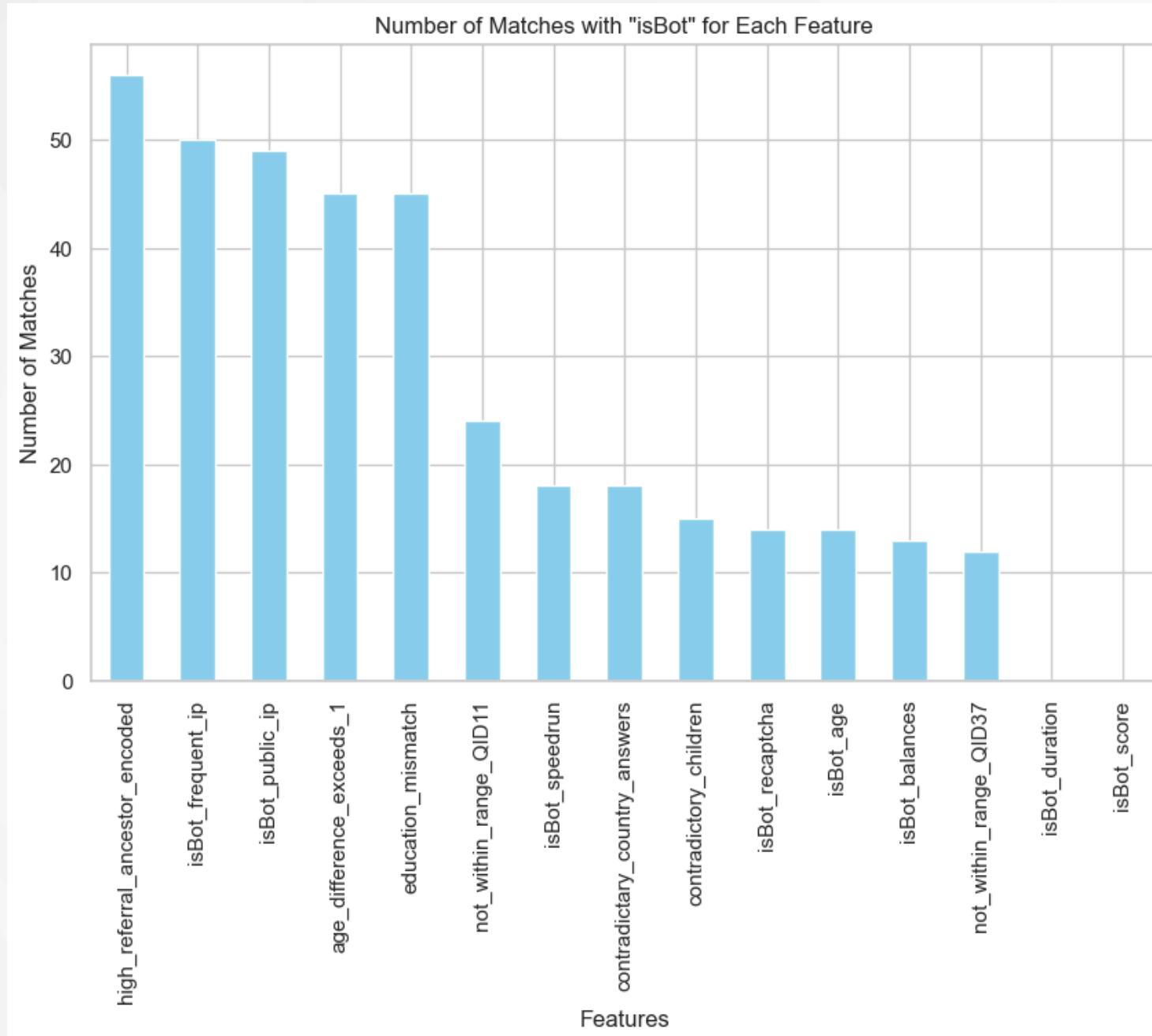
QID12 - What is the highest level of education you have received?

115

INCONSISTANT EDUCATION



CONDITIONS TO 'ISBOT' MATCHES



380
**SUSPICIOUS
USERS**
71
**USERS
ARE BOT
MATCHES**

UUID from top matches provided

SUGGESTIONS FOR FUTURE PROOFING

- Employ Two-Factor Authentication
- Automate unique survey links for participants to prevent bots from accessing multiple surveys with the same link.
- Include hidden items in a survey that will be seen by computers but not by respondents, so if completed it will flag as bot (honeypot).
- Apply rate limiting, IP address monitoring services, security (Cloudflare).
- Use third-party bot detection services or APIs.
- Utilize machine learning algorithms to continuously improve bot detection capabilities and stay ahead of threats.



Q&A

THANK YOU

Zoey Espinoza

Project link

linkedin.com/zoeypespinoza
github.com/zoeypespinoza