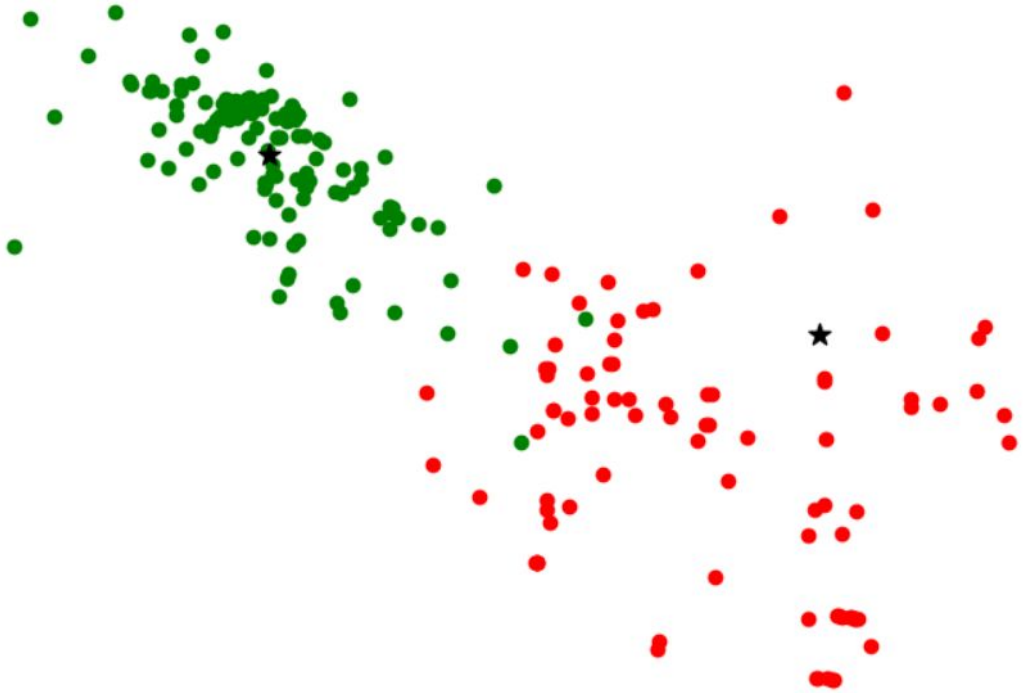
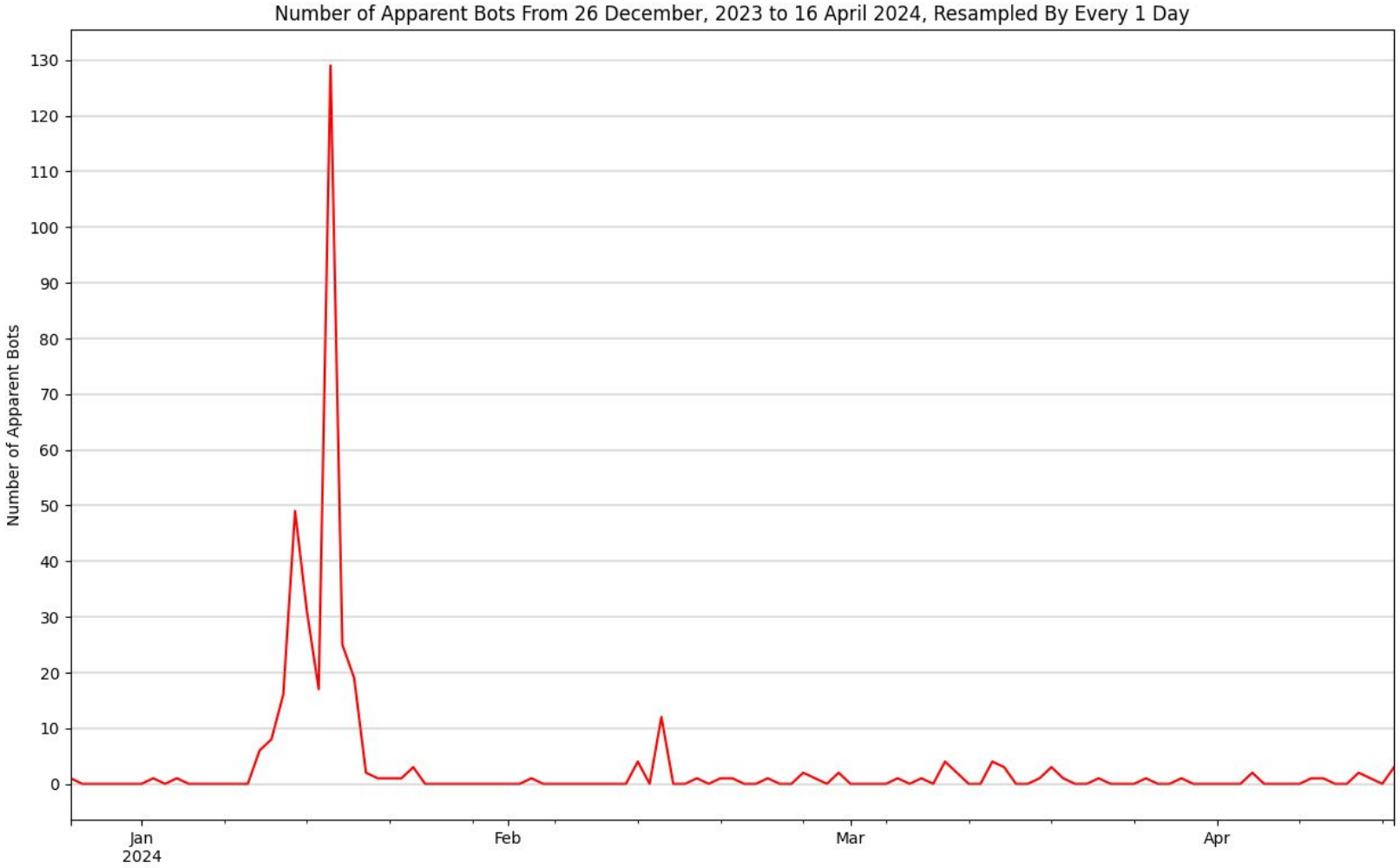
How Unsupervised Learning Can Help Besample Distinguish Bot Accounts From Non-Bot Accounts



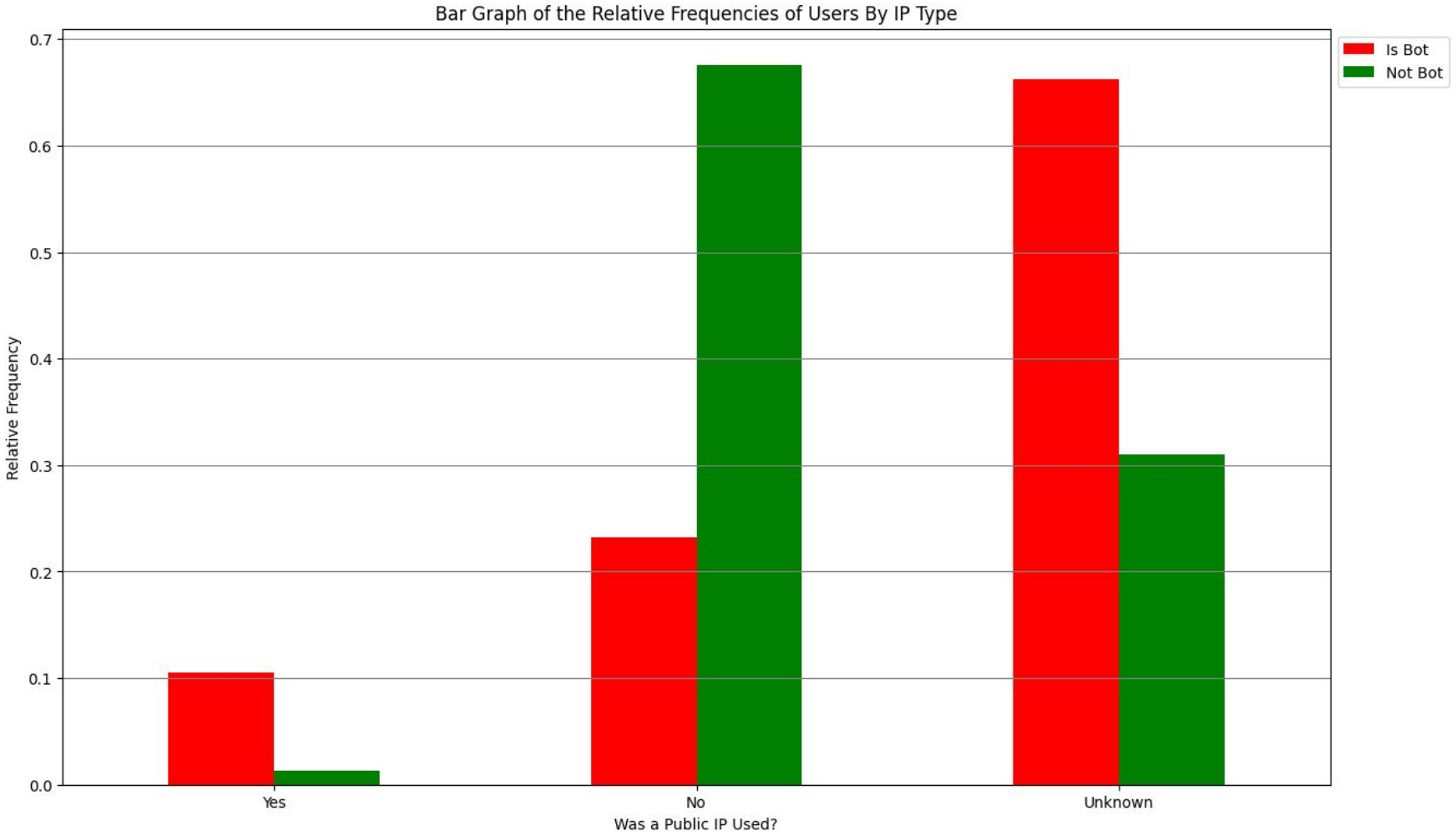
By: Christopher Steven Lewicki

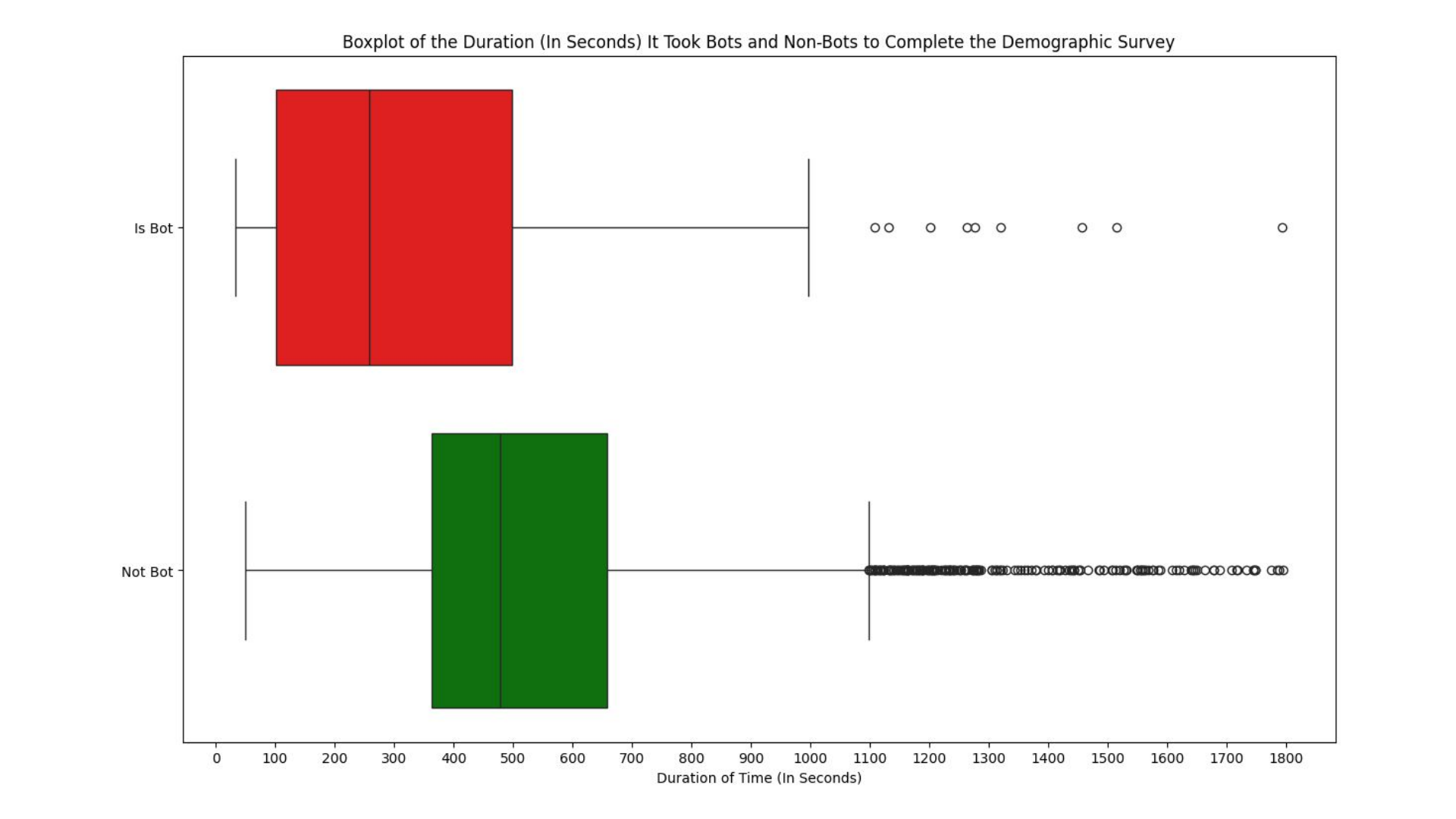

Number of Apparent Bots From 26 December, 2023 to 16 April 2024, Resampled By Every 1 Day
Number of Apparent Bots
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Jan
2024
Feb
Mar
Apr

My Guiding Questions Throughout the Project:

(1) In what ways do the bot users behave differently from the human users?

(2) Is it possible to distinguish between users who are likely to be bots from users who are likely not without using the **isBot** column?


Bar Graph of the Relative Frequencies of Users By IP Type
Is Bot
Not Bot
Relative Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Yes
No
Unknown
Was a Public IP Used?

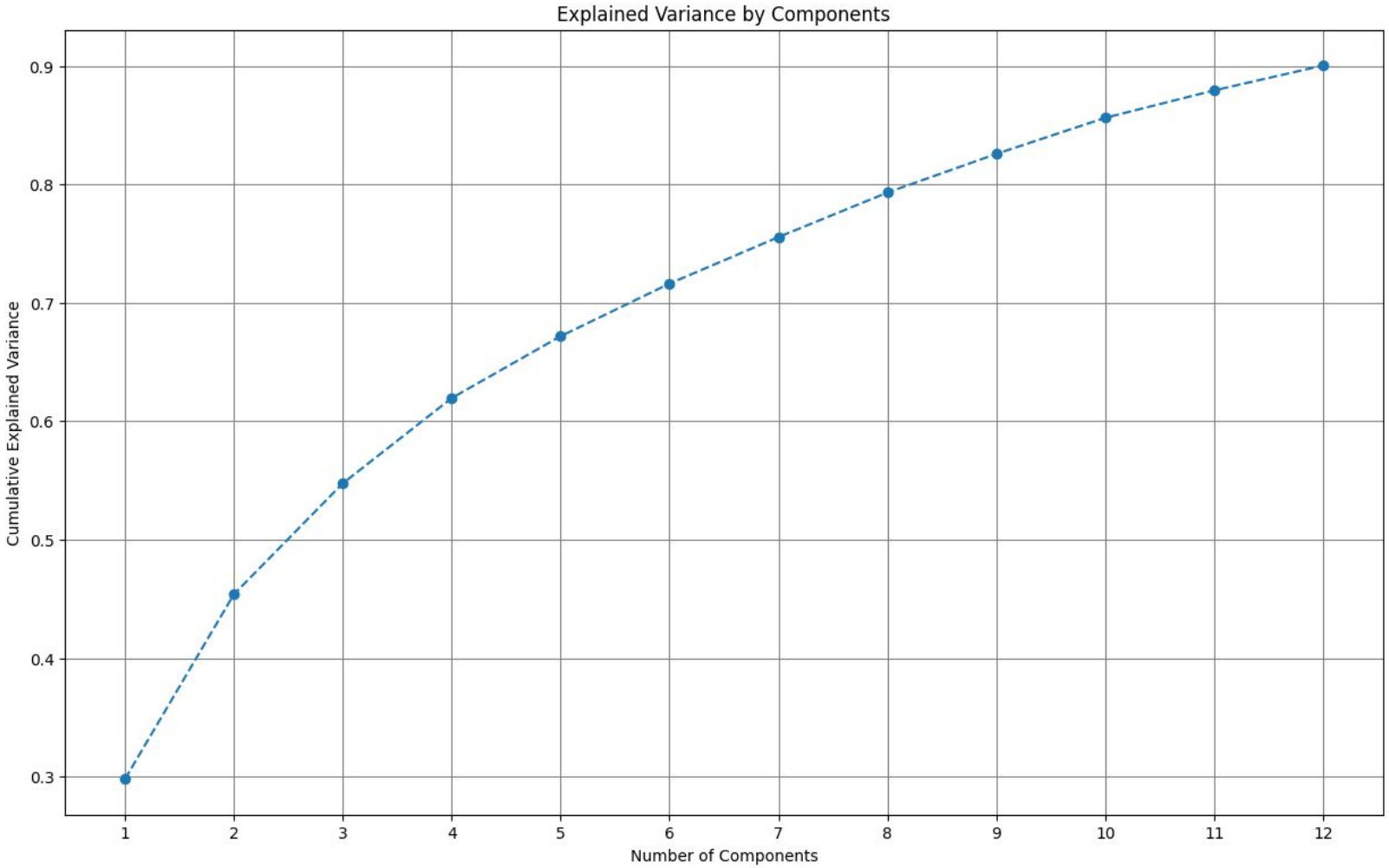

Boxplot of the Duration (In Seconds) It Took Bots and Non-Bots to Complete the Demographic Survey
Is Bot
Not Bot
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
Duration of Time (In Seconds)

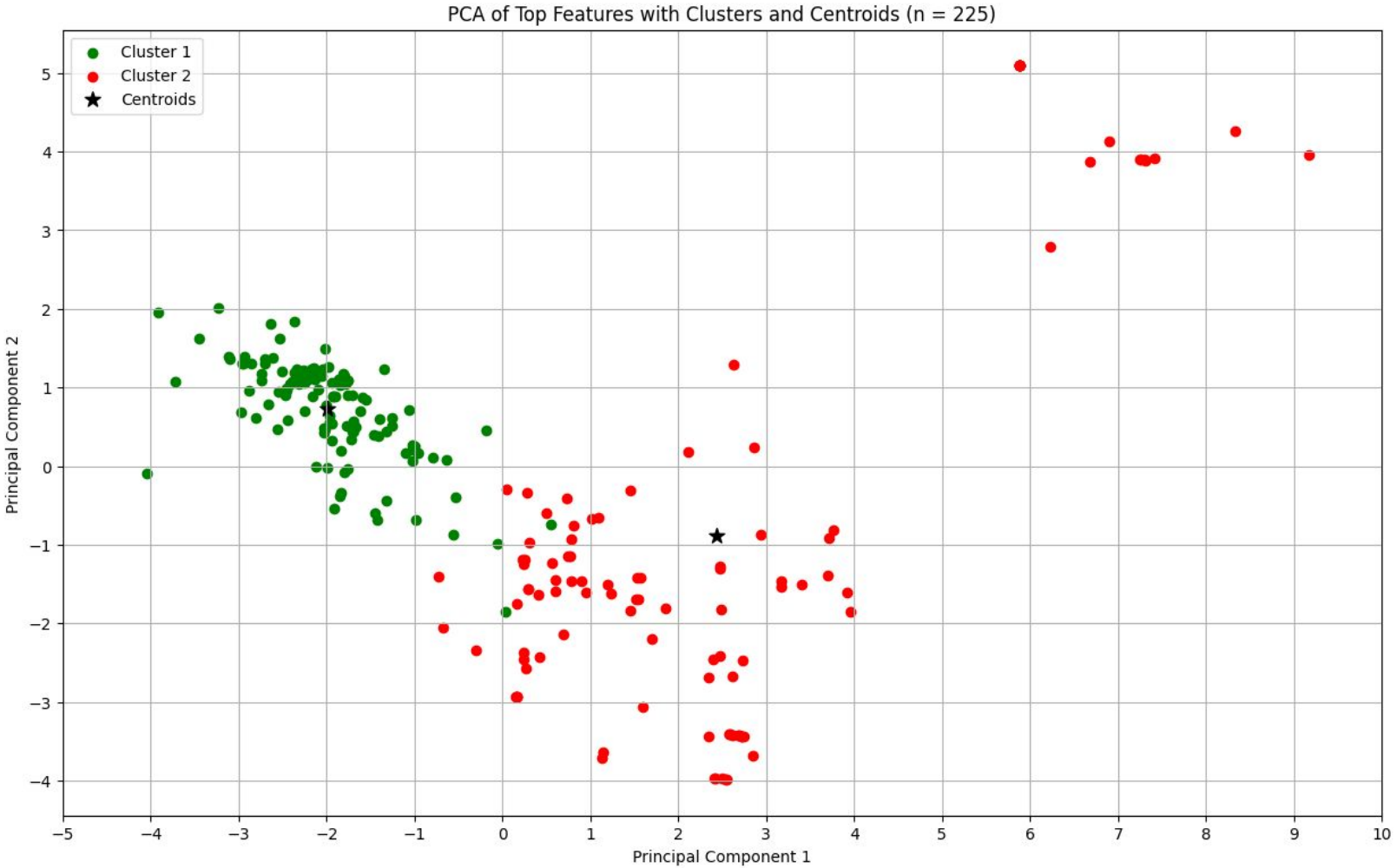
Features I Used That Were Already Give:

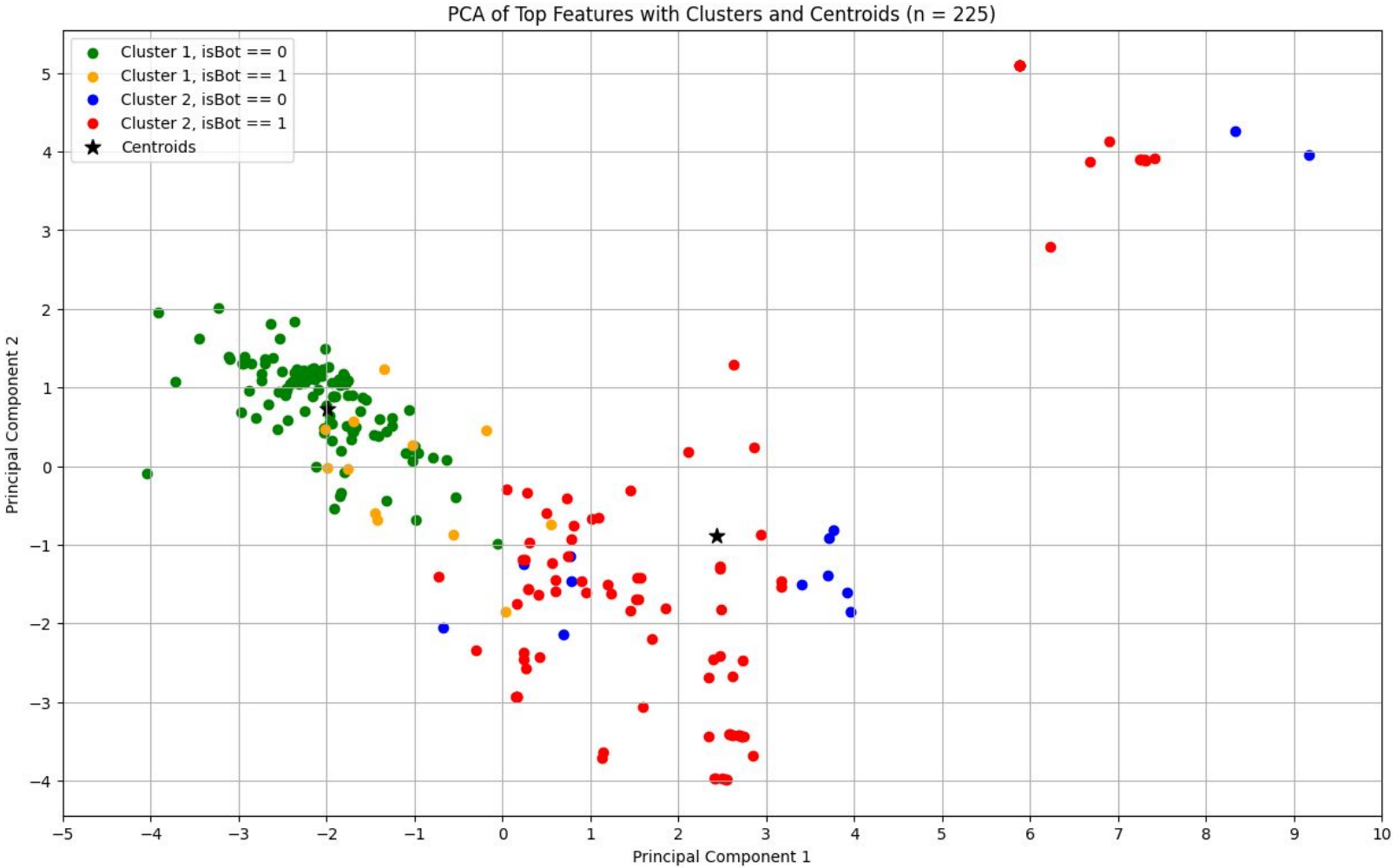
age_bracket	attentive	demo_duration
demo_speedrun	fraud_score	gender
lang	public_ip	status

Features I Used That Were Engineered:

age_contradictory	avg_recaptcha_score	balance_is_outlier
campaign_is_ad	cf_equals_ra	country_contradictory
from_ukraine	hours_after_cf	hours_after_ra
kids_contradictory	lc_is_qualtrics	top_fraud_score
top_ip	top_referral_ancestor	total_events
unusual_age		


Explained Variance by Components
Cumulative Explained Variance
Number of Components
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
2
3
4
5
6
7
8
9
10
11
12


PCA of Top Features with Clusters and Centroids (n = 225)
Cluster 1
Cluster 2
Centroids
Principal Component 2
Principal Component 1
5
4
3
2
1
0
−1
−2
−3
−4
−5
−4
−3
−2
−1
0
1
2
3
4
5
6
7
8
9
10


PCA of Top Features with Clusters and Centroids (n = 225)
Cluster 1, isBot == 0
Cluster 1, isBot == 1
Cluster 2, isBot == 0
Cluster 2, isBot == 1
Centroids
Principal Component 2
Principal Component 1
5
4
3
2
1
0
−1
−2
−3
−4
−5
−4
−3
−2
−1
0
1
2
3
4
5
6
7
8
9
10

likely_bot = 0

A user is likely not a bot if ALL of the following is true:

- avg_recaptcha_score >= 0.5
- demo_duration > 240
- demo_speedrun = 0
- fraud_score < 30
- top_ip = 0
- unusual_age = 0

likely_bot = 1

A user is likely a bot if ANY of the following is true:

- avg_recaptcha_score < 0.5
- demo_duration <= 240
- demo_speedrun = 1
- fraud_score >= 30
- top_ip = 1
- unusual_age = 1

Additionally, 6 referral ancestors had over 50% of the corresponding users flagged as a bot by **isBot**. All users connected to one of these 6 referral ancestors had their **likely_bot** value set equal to 1.

Lastly, I used the results of a coworker's text analysis to identify referral ancestors whose users had unusually similar written responses (and/or other suspicious characteristics), and set their **likely_bot** value to 1.

Value Counts

	isBot	likely_bot
1	370	818
0	4699	1291
–1	N/A	2960

INTERSECTION VALUES:

The number of users who were flagged as bots by both **likely_bot** and **isBot** is 240.

The number of users who were flagged as bots by **likely_bot** but not **isBot** is 578.

INTERSECTION VALUES (Continued)**:**

The number of users who were flagged as bots by **isBot** but not **likely_bot** is 0.

The number of users who were flagged as bots by **isBot** but **likely_bot** is uncertain is 130.

Regarding the 130 users who were flagged as bots by **isBot** but **likely_bot** is uncertain did any…

- Have contradictory answers?
- Have an unknown IP address?
- Have a repeated **came_from**, **referral_ancestor**, and/or **ip** value?
- Created their account between January 13th and January 19th, the week of the “army of bots”?

Value Counts (Updated)

	isBot	**likely_bot**
1	370	959
0	4699	1291
–1	N/A	2819

UPDATED INTERSECTION VALUES:

The number of users who were flagged as bots by both **likely_bot** and **isBot** is 364.

The number of users who were flagged as bots by **likely_bot** but not **isBot** is 595.

UPDATED INTERSECTION VALUES (Continued)**:**

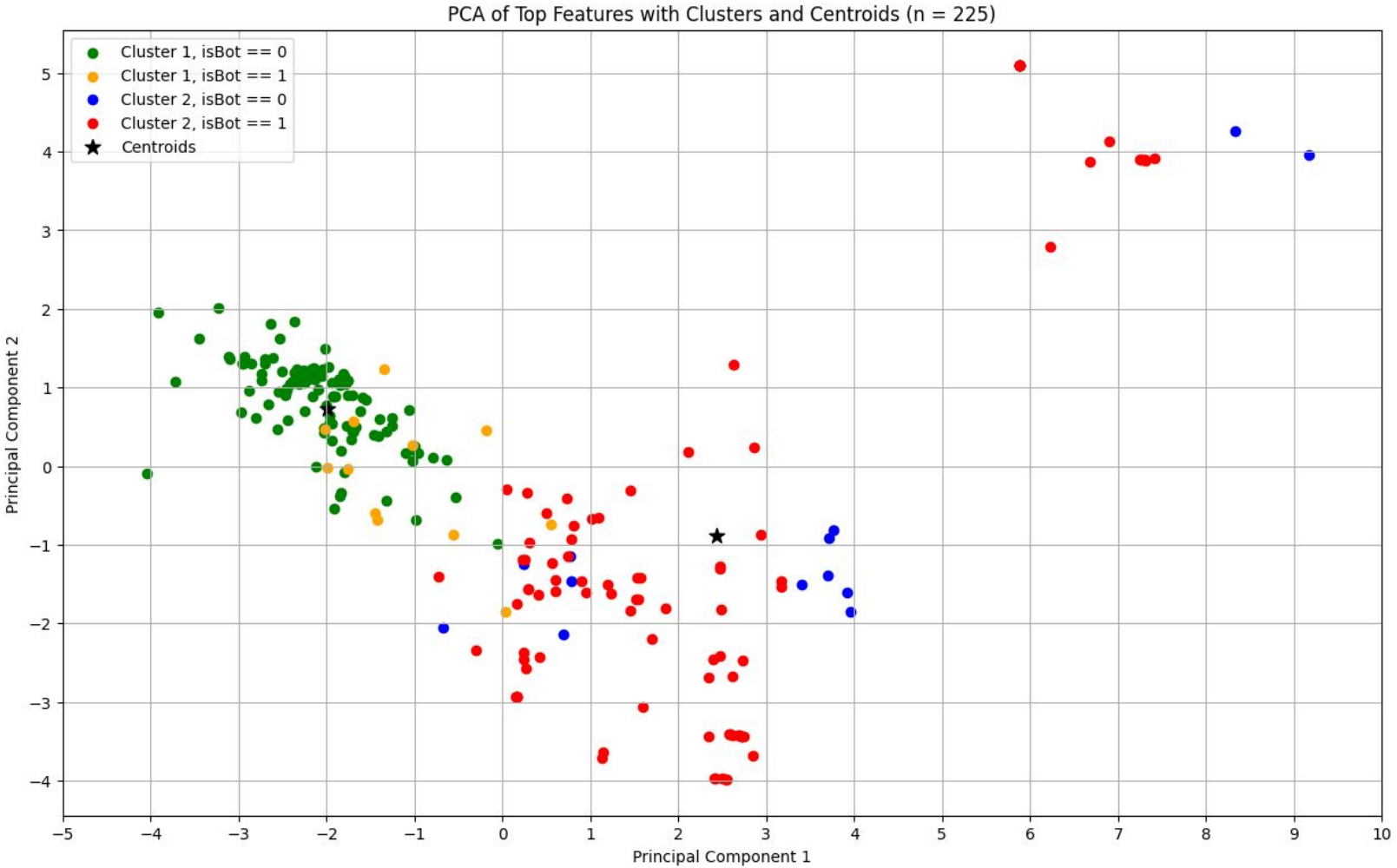
The number of users who were flagged as bots by **isBot** but not **likely_bot** is 0.

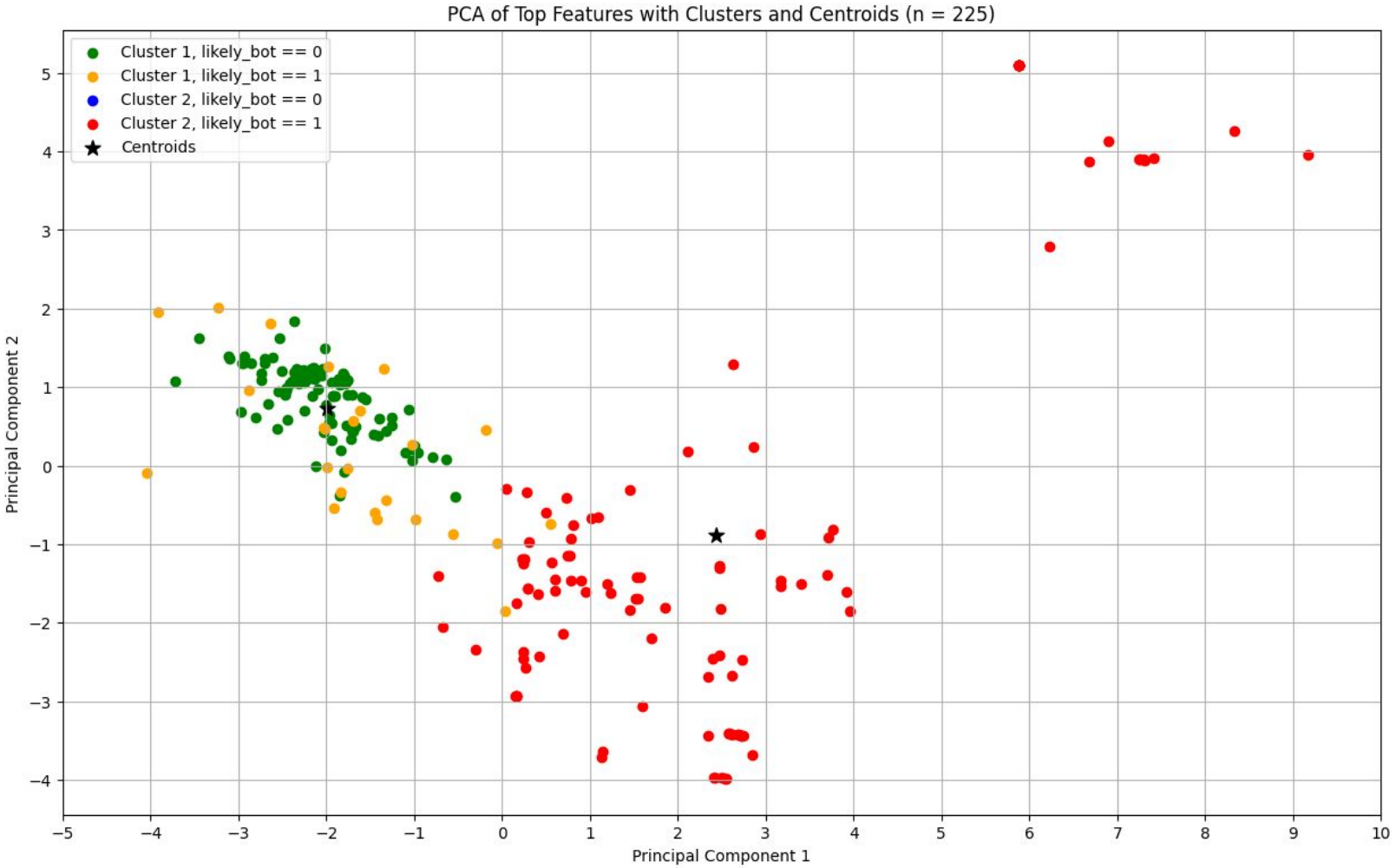
The number of users who were flagged as bots by **isBot** but **likely_bot** is uncertain is 6.

Random Sample (n = 225)**:**

- Outlier **demo_duration** values were dropped.
- Users where **likely_bot** = –1 were ignored.

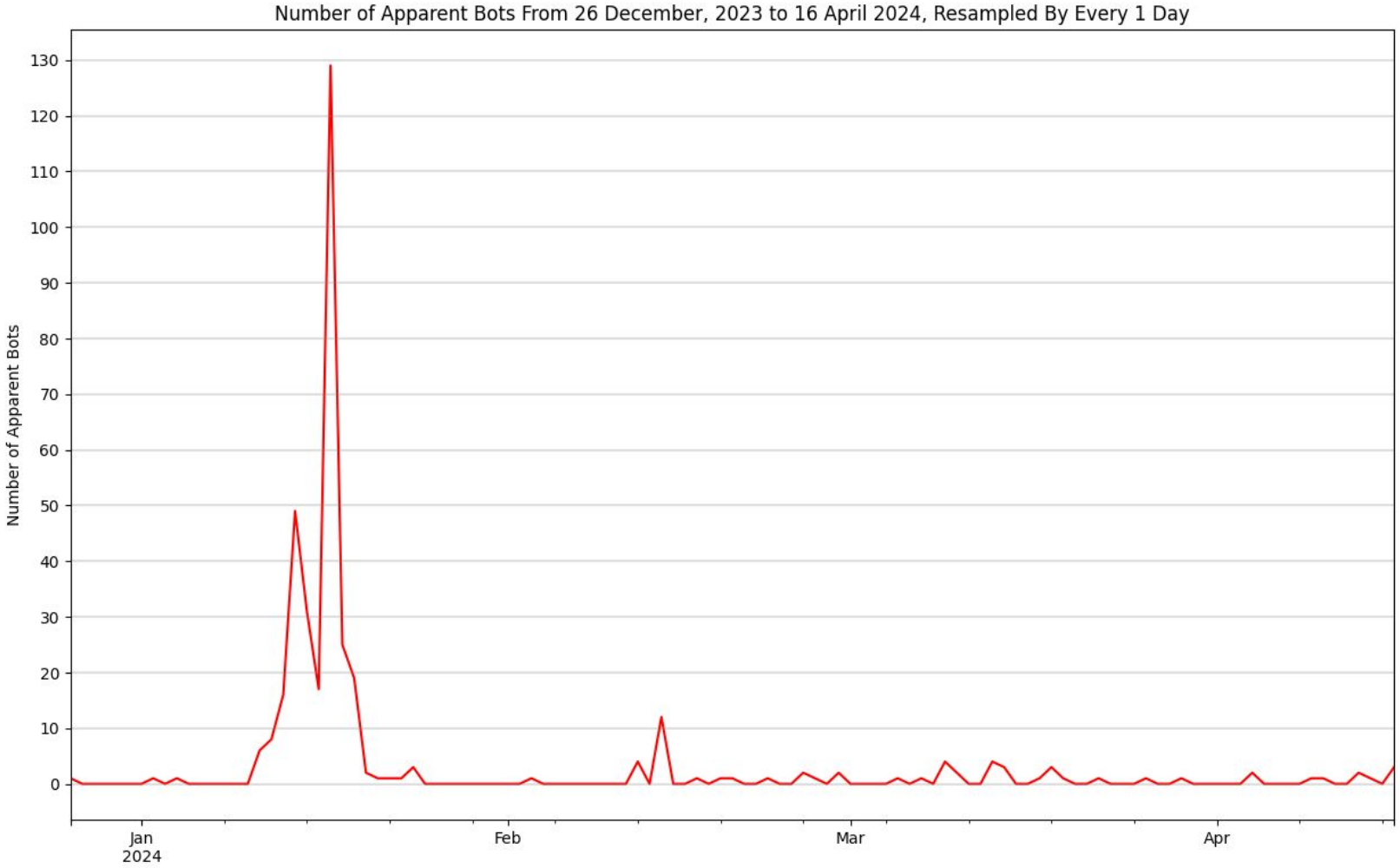
```
isBot
0    125
1    100
Name: count, dtype: int64

likely_bot
1    126
0     99
Name: count, dtype: int64
```

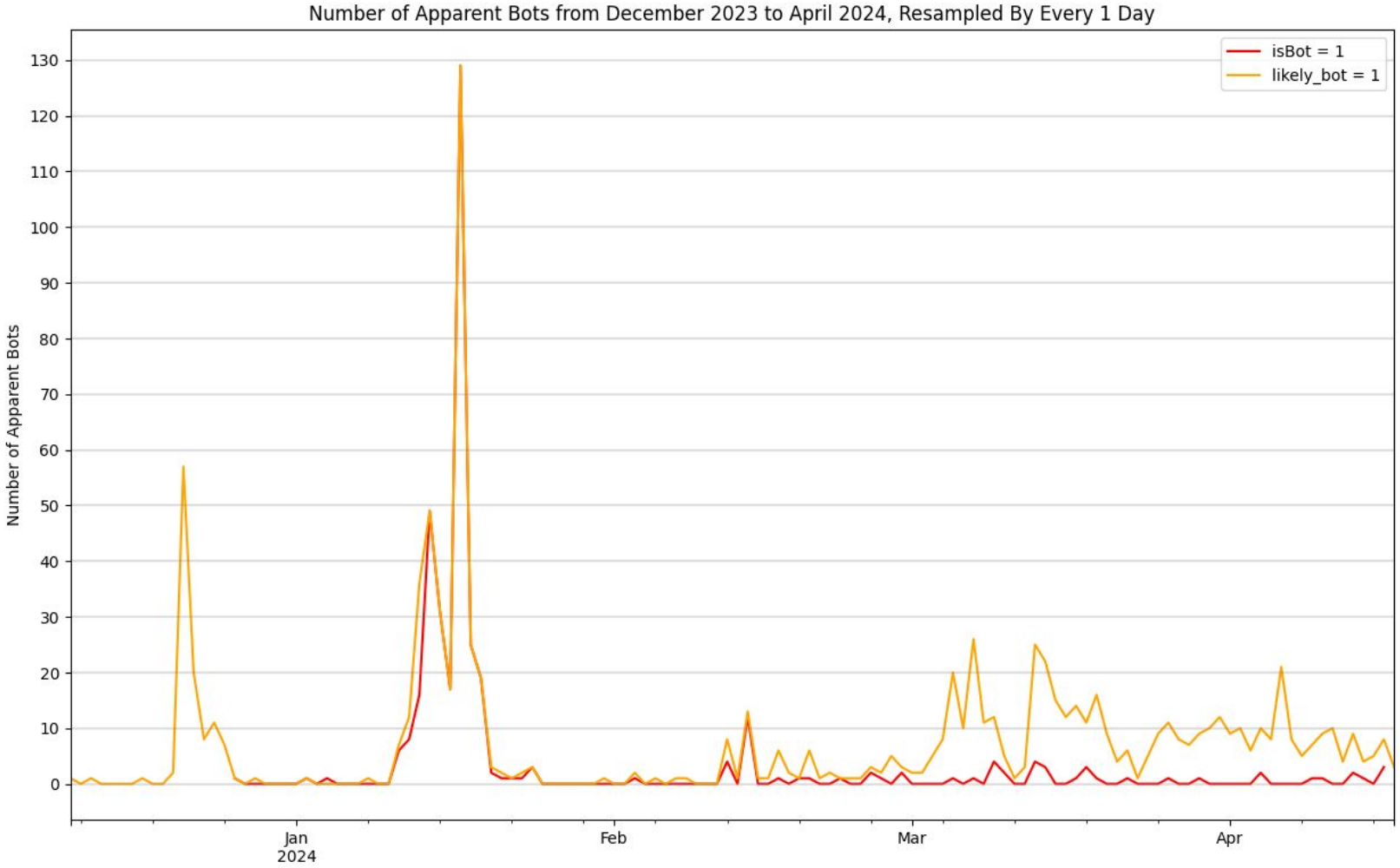

PCA of Top Features with Clusters and Centroids (n = 225)
Cluster 1, isBot == 0
Cluster 1, isBot == 1
Cluster 2, isBot == 0
Cluster 2, isBot == 1
Centroids
Principal Component 2
Principal Component 1
5
4
3
2
1
0
−1
−2
−3
−4
−5
−4
−3
−2
−1
0
1
2
3
4
5
6
7
8
9
10


PCA of Top Features with Clusters and Centroids (n = 225)
Cluster 1, likely_bot == 0
Cluster 1, likely_bot == 1
Cluster 2, likely_bot == 0
Cluster 2, likely_bot == 1
Centroids
Principal Component 2
Principal Component 1
5
4
3
2
1
0
−1
−2
−3
−4
−5
−4
−3
−2
−1
0
1
2
3
4
5
6
7
8
9
10

feature	pc1_abs
country_contradictory	0.275697
top_fraud_score	0.270417
attentive	0.268216
avg_recaptcha_score	0.265793
status	0.251329
unusual_age	0.249143
from_ukraine	0.247564
top_referral_ancestor	0.239463
age_contradictory	0.225589
cf_equals_ra	0.219797
age_bracket	0.219753
public_ip	0.219116
demo_duration	0.210450
lc_is_qualtrics	0.210145
balance_is_outlier	0.184234
campaign_is_ad	0.169557
total_events	0.160354
hours_after_ra	0.145243
kids_contradictory	0.131341
hours_after_cf	0.126660
gender	0.104938
top_ip	0.099401
fraud_score	0.088161
demo_speedrun	0.080193
lang	0.050665

	centroid_1	centroid_2
age_bracket	1.54032	0.59406
age_contradictory	0.15323	-0.77228
attentive	0.94355	0.33663
avg_recaptcha_score	0.88144	-0.45644
balance_is_outlier	0.52419	0.04950
campaign_is_ad	0.38710	-0.46535
cf_equals_ra	-0.93548	-0.37624
country_contradictory	0.02419	-0.97030
demo_duration	519.03226	272.90099
demo_speedrun	0.05645	0.19802
fraud_score	3.68548	-1.00000
from_ukraine	0.95161	0.57426
gender	0.16935	0.29703
hours_after_cf	1.15685	42.18564
hours_after_ra	1.23535	64.80033
kids_contradictory	-0.34677	-0.65347
lang	1.87903	2.00000
lc_is_qualtrics	0.57258	0.02970
public_ip	0.05645	-0.64356
status	3.00000	2.45545
top_fraud_score	-0.04032	-1.00000
top_ip	0.03226	0.11881
top_referral_ancestor	-0.92742	-0.01980
total_events	3.08065	0.14851
unusual_age	0.00000	-0.16832


Number of Apparent Bots From 26 December, 2023 to 16 April 2024, Resampled By Every 1 Day
Number of Apparent Bots
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Jan
2024
Feb
Mar
Apr


Number of Apparent Bots from December 2023 to April 2024, Resampled By Every 1 Day
isBot = 1
likely_bot = 1
Number of Apparent Bots
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Jan
2024
Feb
Mar
Apr