HUSO 2020 Oct. 18 - 22, 2020 Porto, Portugal

An Investigation of Twitter Users Who Disclosed Their Personal Profile Items in Their Tweets Honestly

Yasuhiko Watanabe, Hiromu Nishimura, Yuuya Chikuki, Kunihiro Nakajima, and Yoshihiro Okada

watanabe@rins.ryukoku.ac.jp

Ryukoku University

presenter information

- Doctor of Informatics
- Lecturer, Ryukoku University, Kyoto, Japan

Yasuhiko Watanabe

our research interest

- communications in SNS
- user behavior analysis
- trust and security in SNS

background

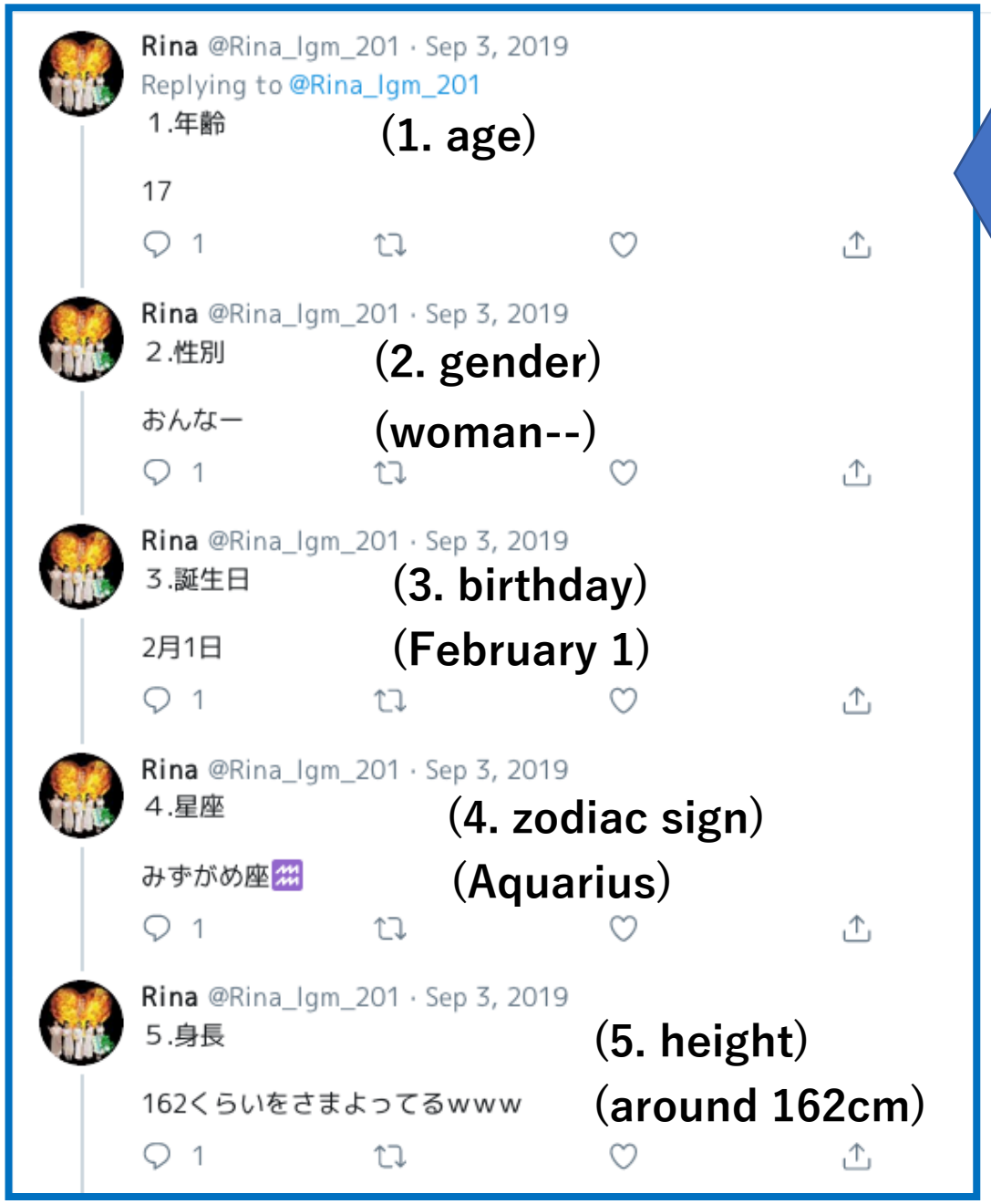
some SNS users willingly disclose their personal information

SNS users
- real name users
- non-real name users

Reliability of real name & non-real name users

Many of us think that
personal information disclosed by

- real name users → reliable
- non-real name users → unreliable



(1. age)

(2. gender)

(woman--)

(3. birthday)

(February 1)

(4. zodiac sign)

(Aquarius)

(5. height)

(around 162cm)

An example:

a SNS user, Rina, disclose her personal information

Many of us do not think
Rina disclosed
her personal info. honestly

[reason]
difflcult to check
whether she disclosed
her personal info. honestly

Our idea

check whether non-real name account users disclosed their personal information honestly

Our method

- collect & analyze statistically
  tweets where non-real name account users disclosed their personal information

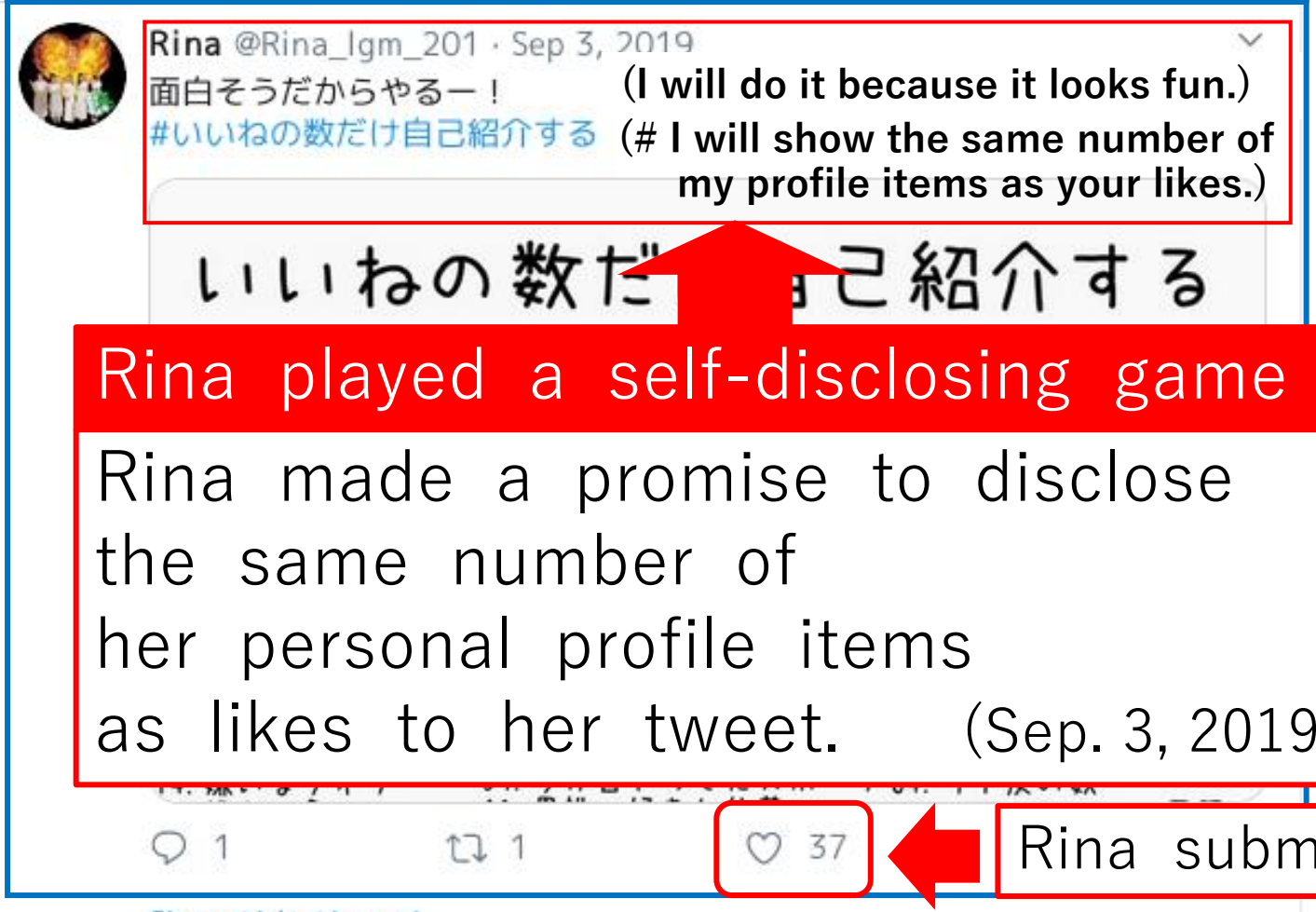
Investigation targets

tweets promising to disclose submitters' personal information

self-replies disclosing submitters' personal info.

Rina's self-disclosing game

tweets promising to disclose submitters' personal information



Rina played a self-disclosing game

Rina made a promise to disclose the same number of her personal profile items as likes to her tweet. (Sep. 3, 2019)

self-replies disclosing submitters' personal info.

Rina submitted 35 replies on Sep. 3-9, 2019

self-disclosing games

tweets promising to disclose submitters' personal information

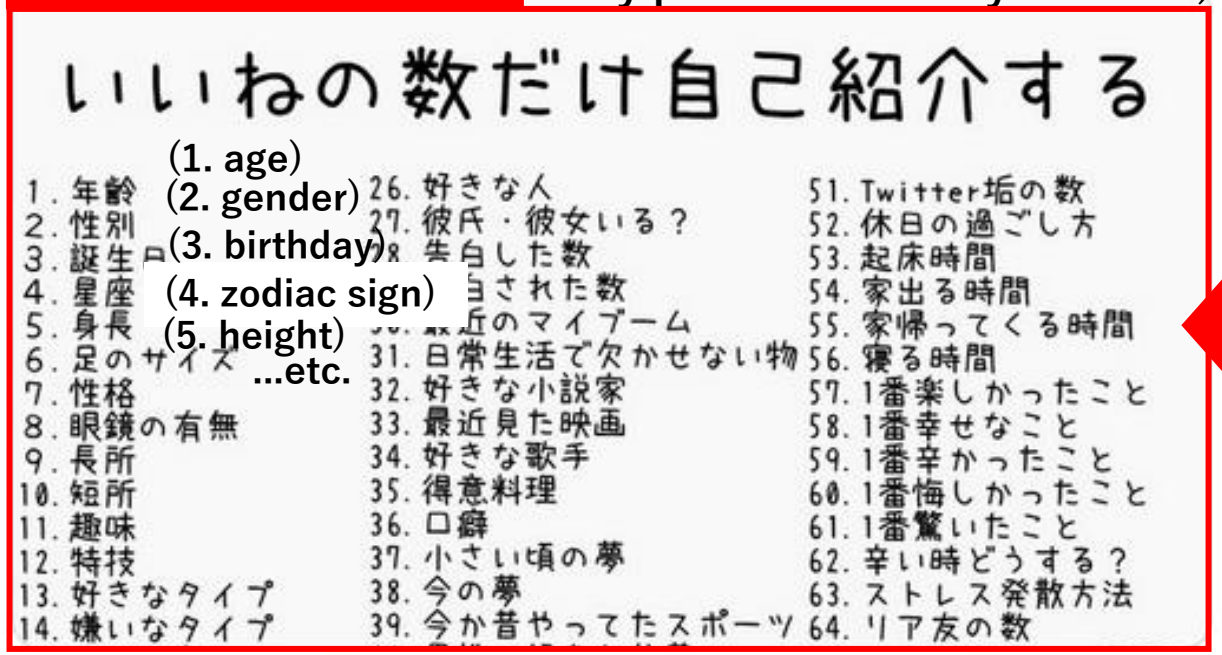
Rina @Rina_lgm_201 · Sep 3, 2019

面白そうだからやるー！

(I will do it because it looks fun.)

(# I will show the same number of my profile items as your likes.)

an attached image



1 1 37

Show this thread

self-replies disclosing submitters' personal info.

Rina @Rina_lgm_201 · Sep 3, 2019

2月1日 (February 1)

1

many users played a self-disclosing game

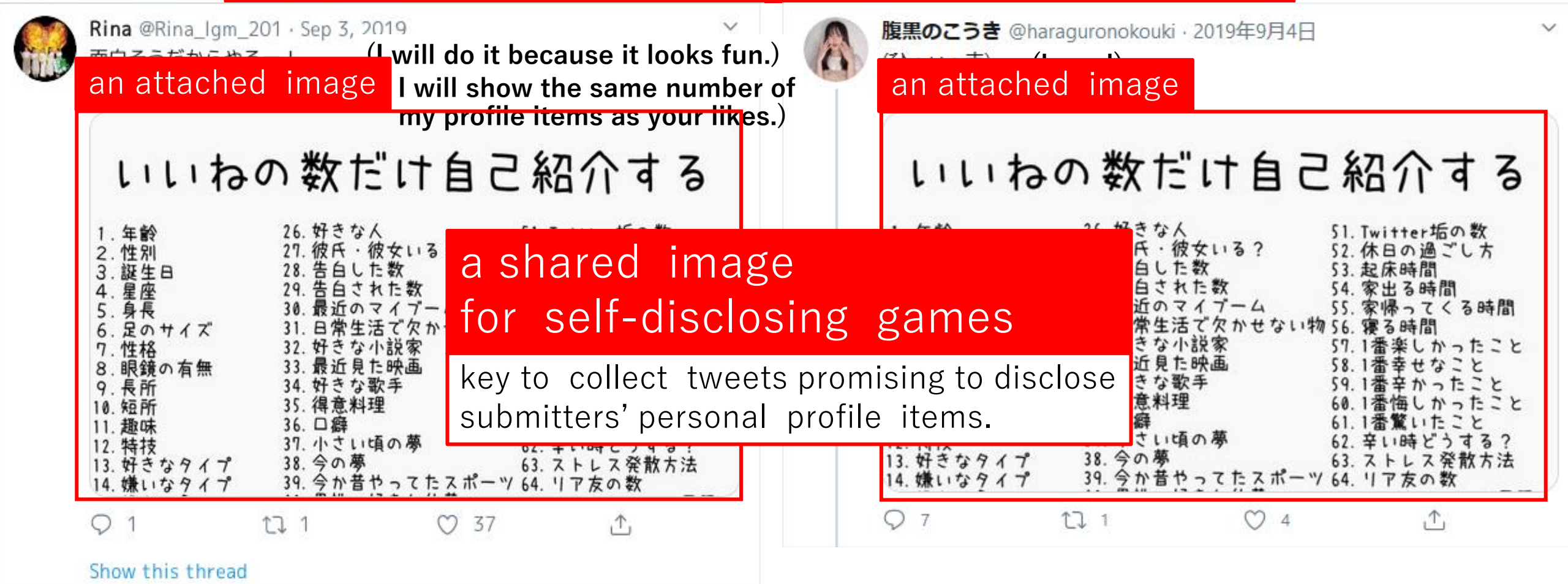
they made promises to disclose the same number of their personal profile items as likes to their tweets.

personal profile items list

they attached the same image to their tweets, a list of personal profile items that they promised to disclose.

Twigaten (https://twigaten.204504byse.info)

Twigaten helps us to collect tweets to which the same image is attached.



a shared image for self-disclosing games

key to collect tweets promising to disclose submitters' personal profile items.

Investigation data

565 Japanese tweets promising to disclose submitters' personal profile items

- collection date: Nov. 20, 2019.
- submission date: from Oct. 3, 2018 to Nov. 20, 2019.

personal profile items for investigation

- genders
- ages
- heights

submitters' genders

men	156
women	282
unclear	27
not disclosed	100

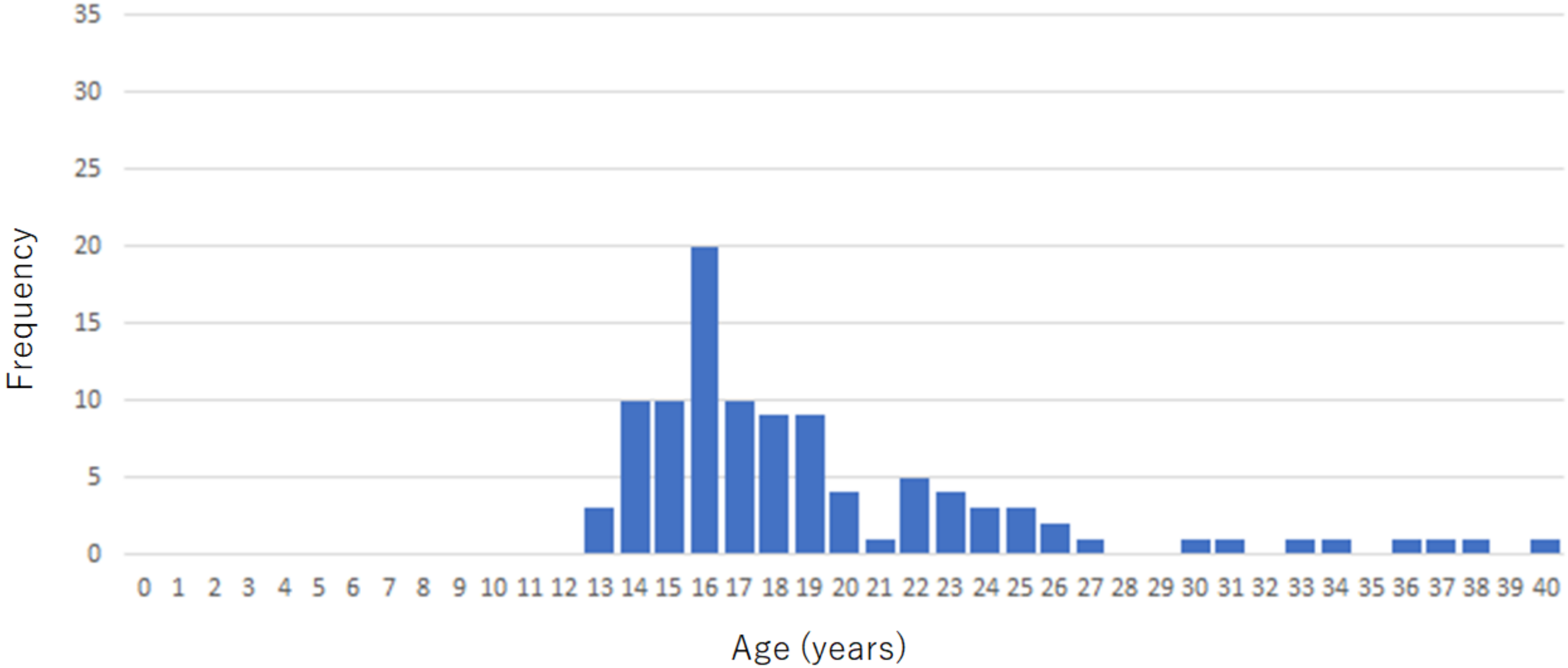
submitters’ ages

submitters’ ages	
clear	276
unclear	60
not disclosed	229

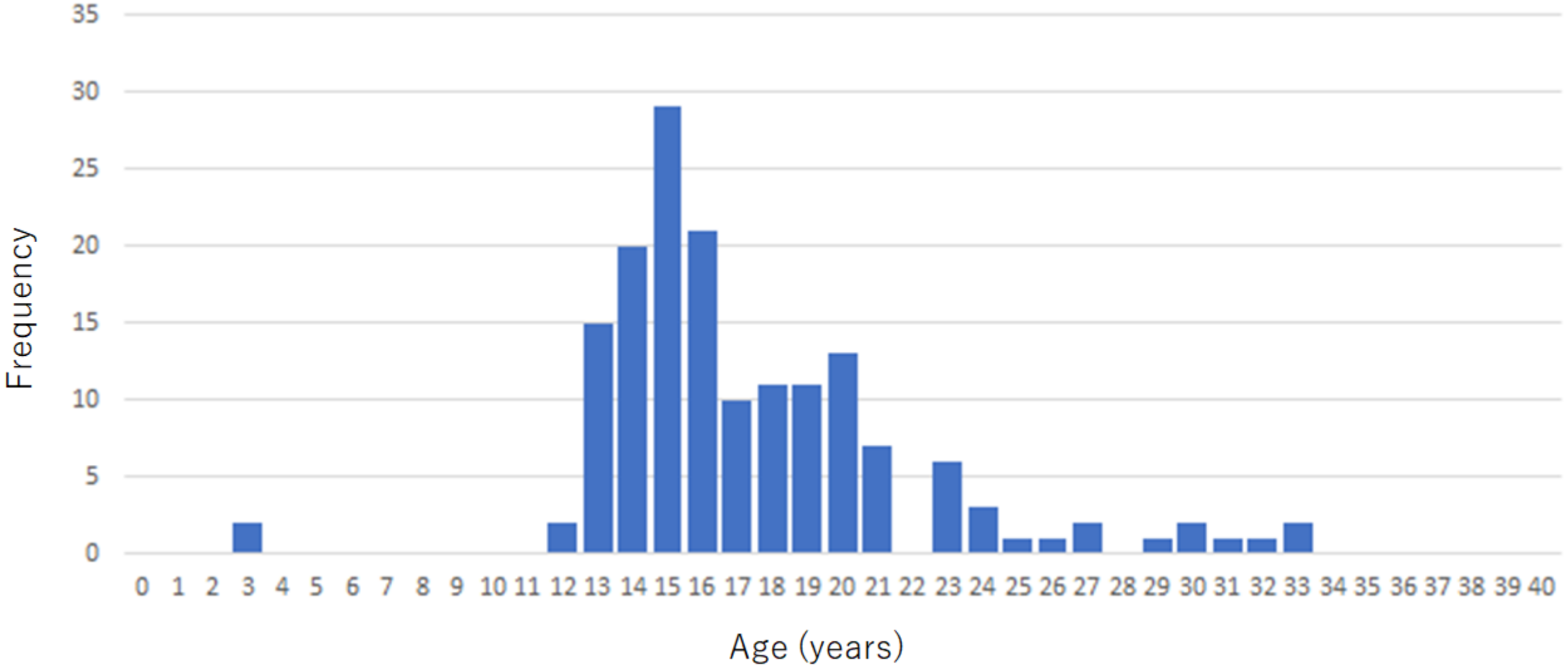
Among the 276 tweets, submitters’ genders

men	102
women	161
unclear/ not disclosed	13

The number of submitters who disclosed that they were men by age



The number of submitters who disclosed that they were women by age

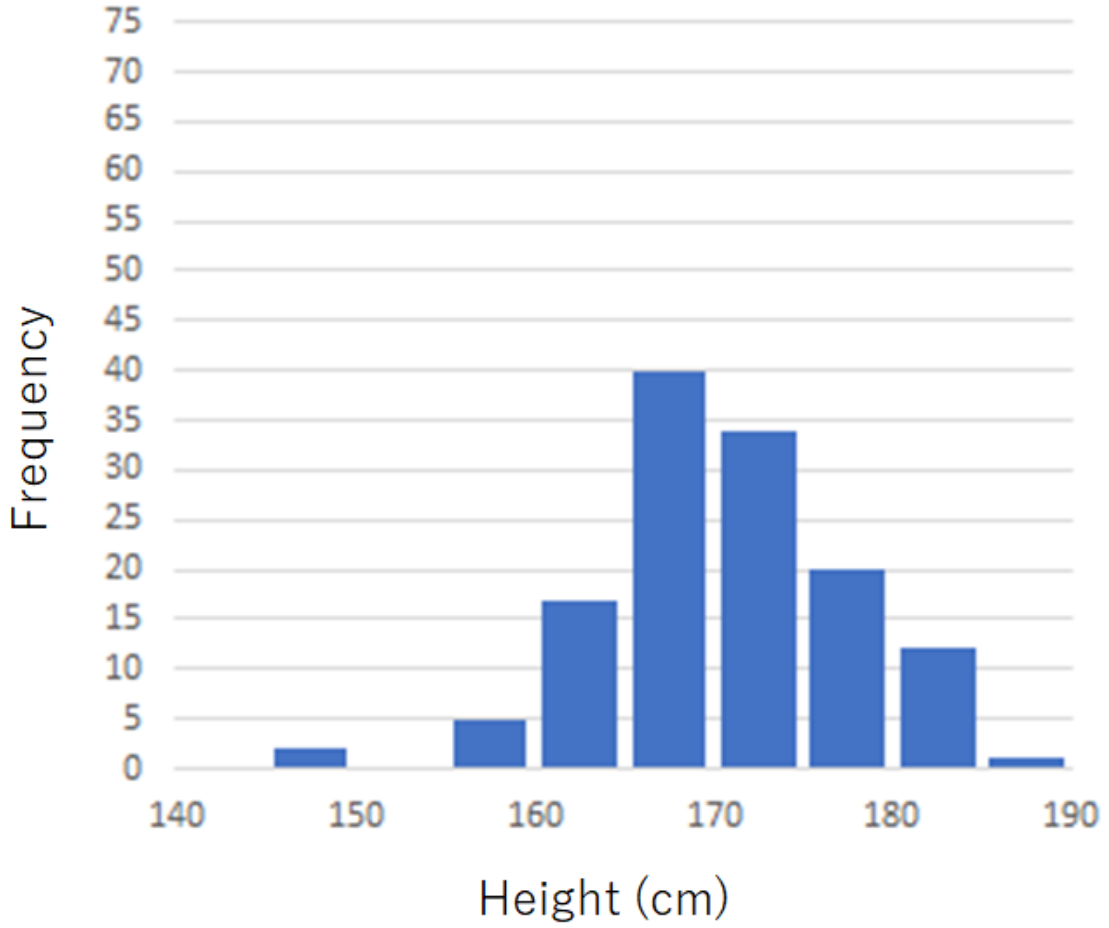


submitters’ heights

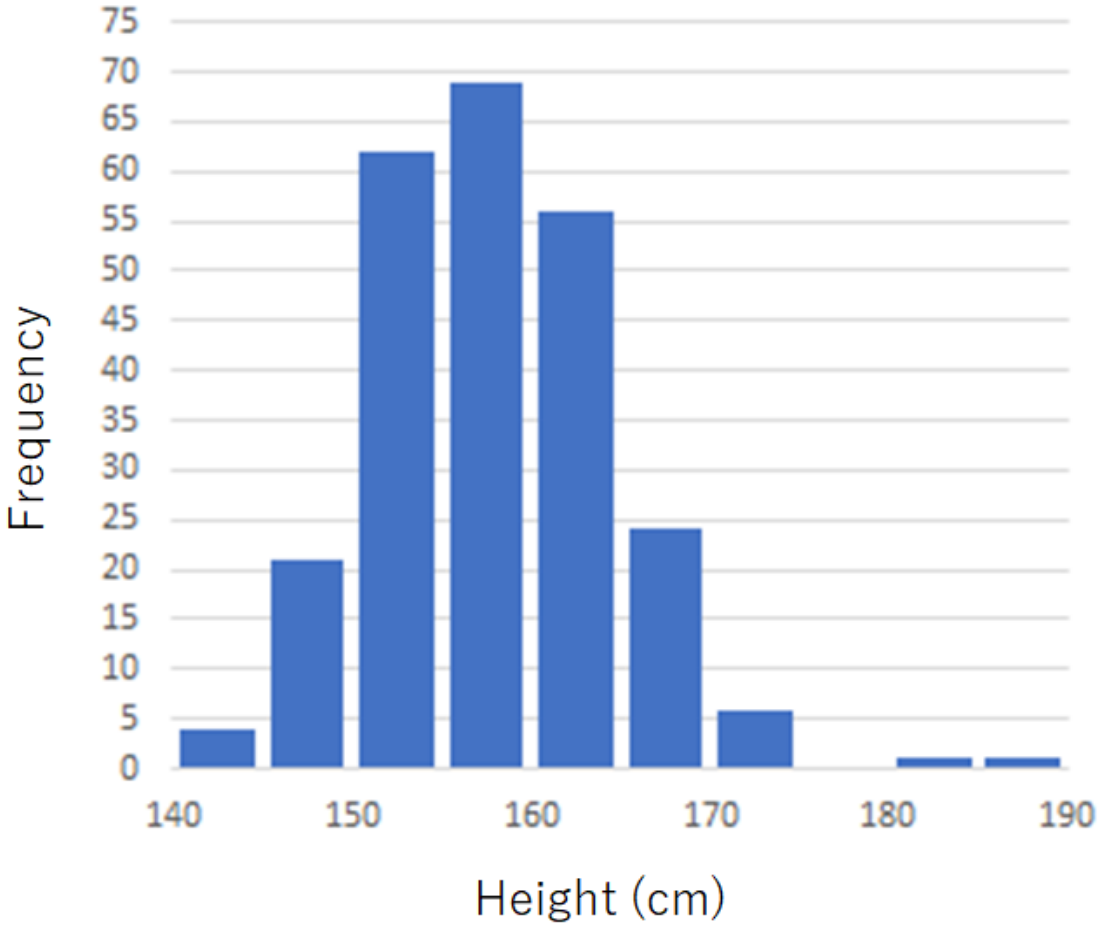
submitters’ heights	
clear	401
unclear	8
not disclosed	156

Among the 401 tweets, submitters’ genders

men	131
women	244
unclear/ not disclosed	26



(a) the histogram of submitters' heights (disclosed genders: men).



(b) the histogram of submitters' heights (disclosed genders: women).

problem

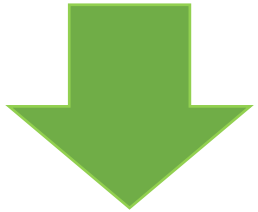
difficult to determine whether an individual submitter disclosed his/her personal information honestly

our idea

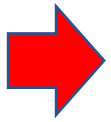
examine whether submitters disclosed their personal information honestly

our idea

our heights follow a normal (Gaussian) distribution



if most of submitters disclose
their ages, genders, and heights honestly,



their heights would follow a normal distribution

targets of statistical analysis

- 37 submitters who disclosed their genders (men), ages (15-17 years old), and heights clearly
- 60 submitters who disclosed their genders (women), ages (14-16 years old), and heights clearly.

men aged 15-17 and women aged 14-17
were the most popular segments in the submitters' ages.

statistical analysis methods

Shapiro-Wilk test

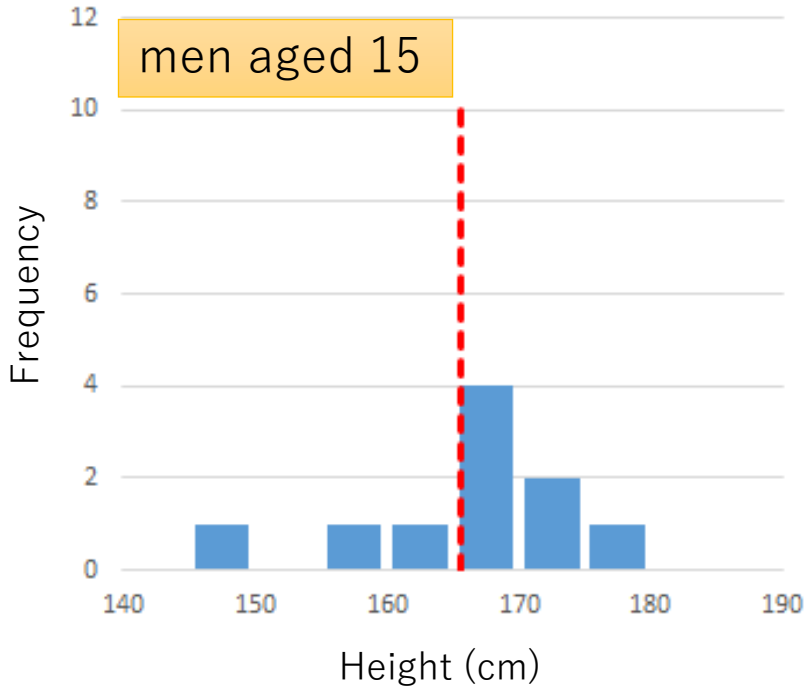
whether submitters' heights follow a normal distribution ?

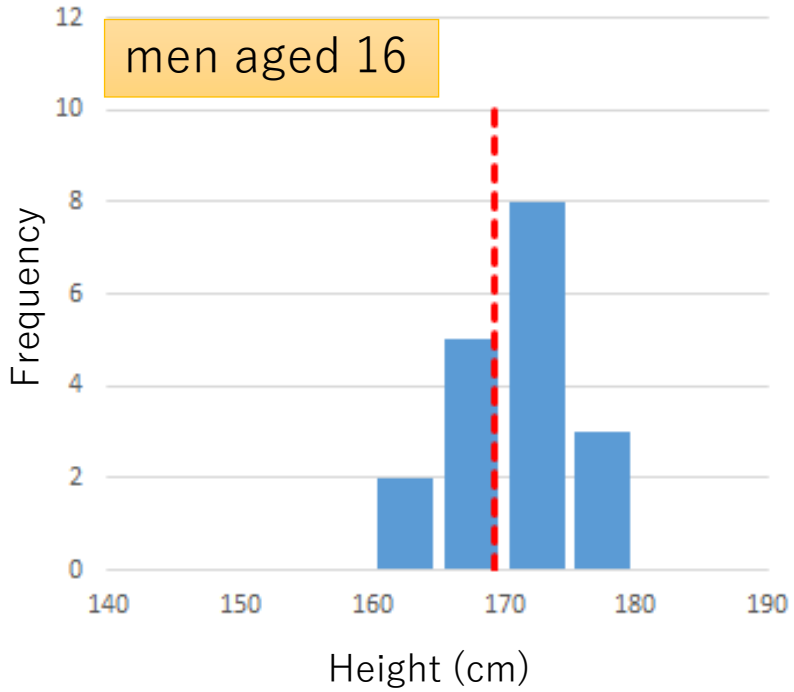
Welch's test

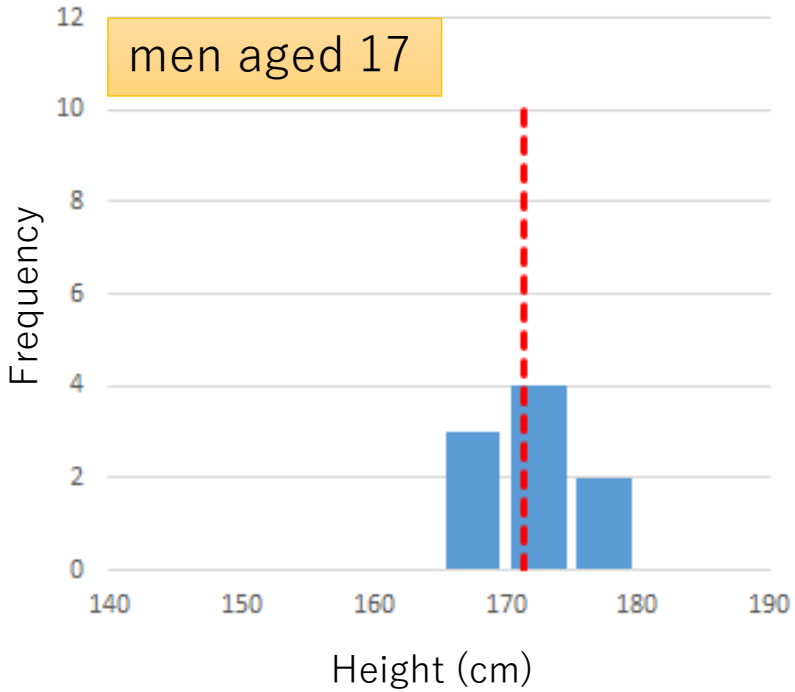
whether the average of submitters' heights
are equal to the national average height in Japan?

men aged 15-17

Shapiro-Wilk test: normal distribution ?







gender	age	sample size	average	SD	Shapiro-Wilk test W value	Shapiro-Wilk test p-value
men	15	10	165.5	7.58	0.885	0.147
men	16	18	169.2	4.08	0.929	0.190
men	17	9	171.3	3.84	0.977	0.946

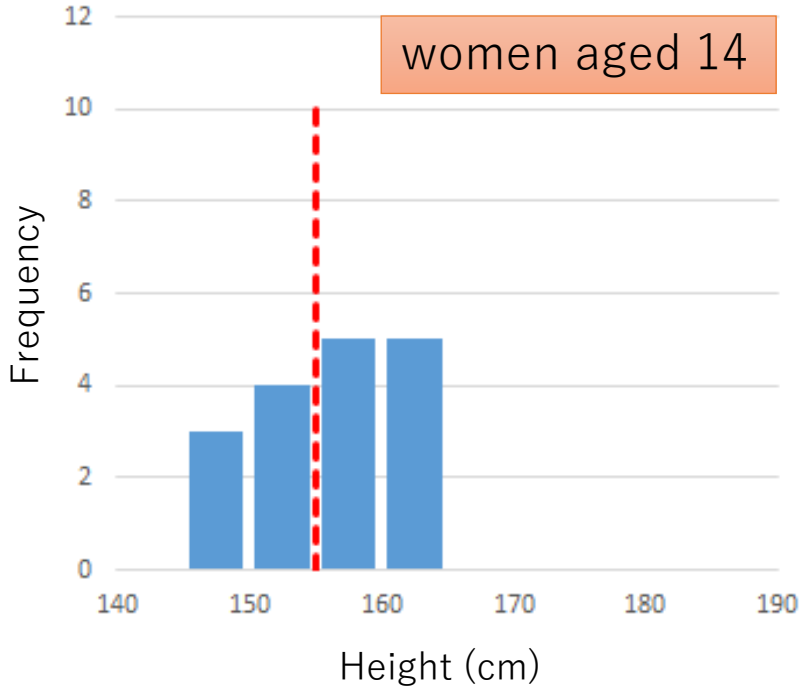
men aged 15-17

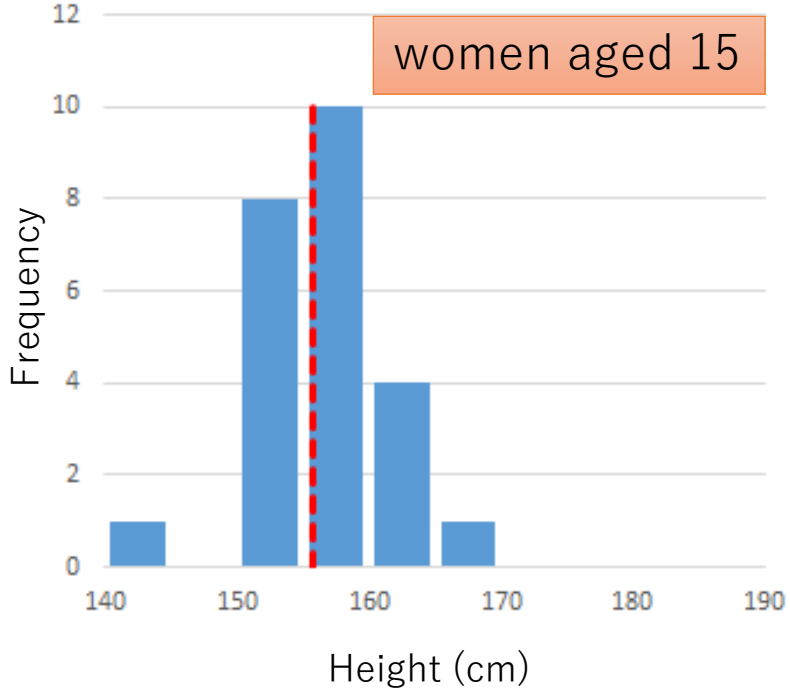
Yes!
normal distribution

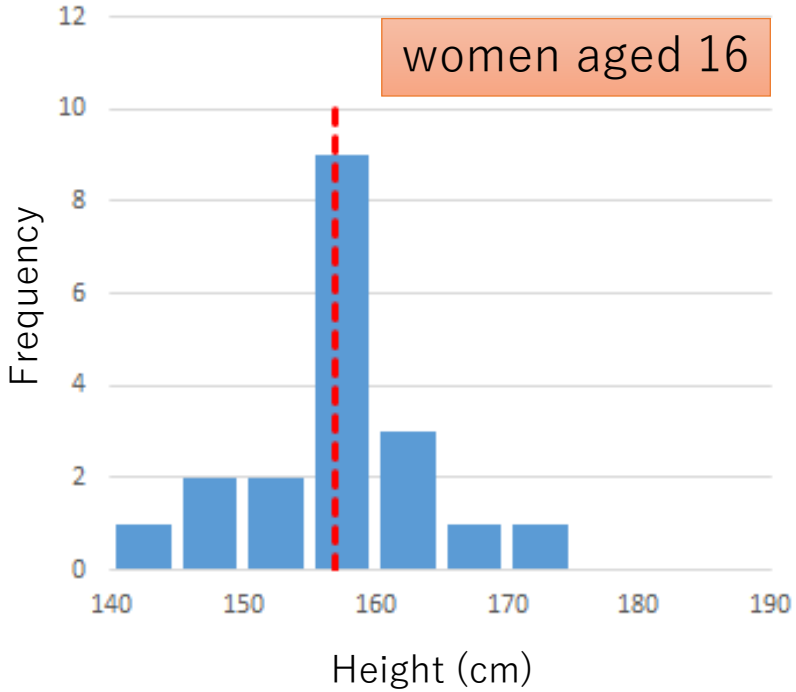
p-value > 0.05

women aged 14-16

Shapiro-Wilk test: normal distribution ?







gender	age	sample size	average	SD	Shapiro-Wilk test W value	Shapiro-Wilk test p-value
women	14	17	155.0	6.11	0.933	0.244
women	15	24	155.7	4.99	0.971	0.697
women	16	19	156.9	6.54	0.961	0.587

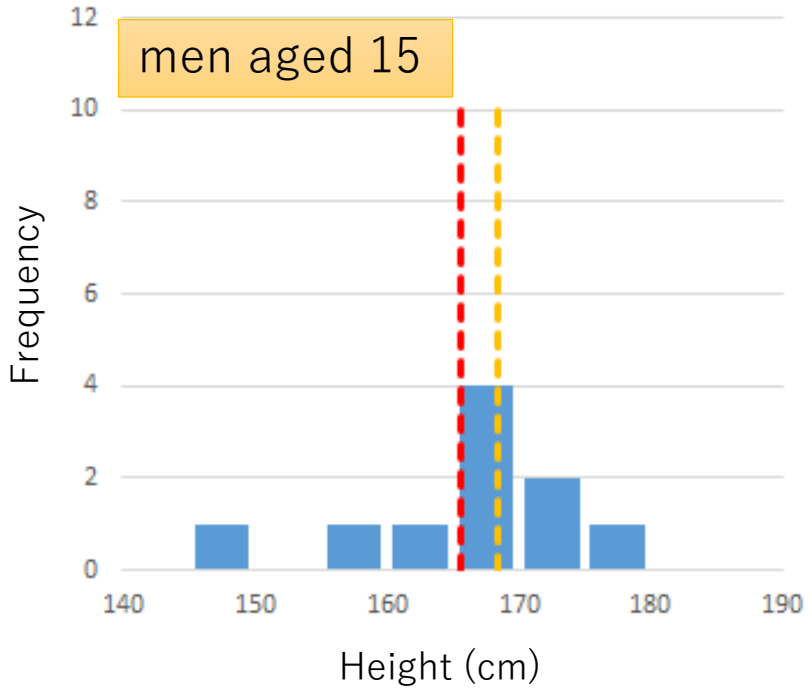
women aged 14-16

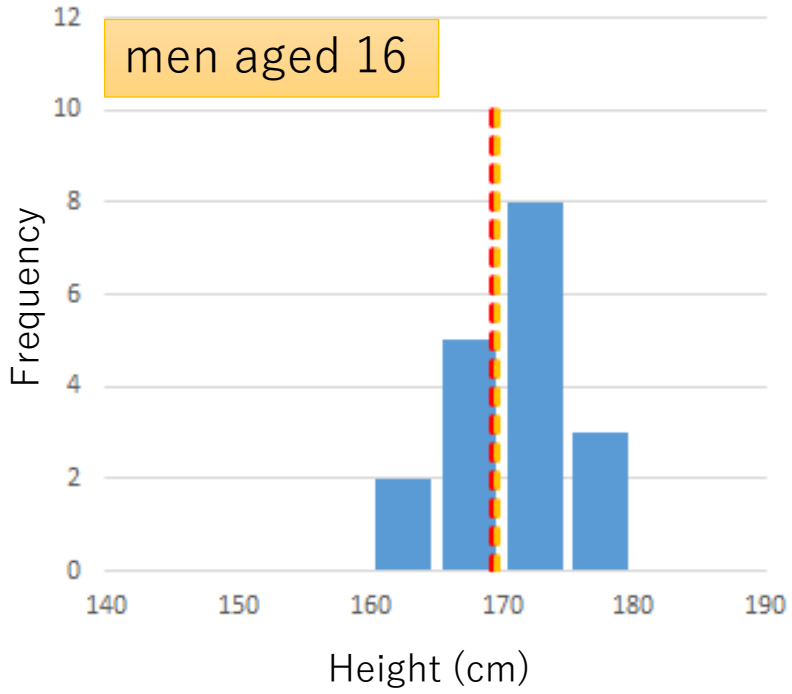
Yes!
normal distribution

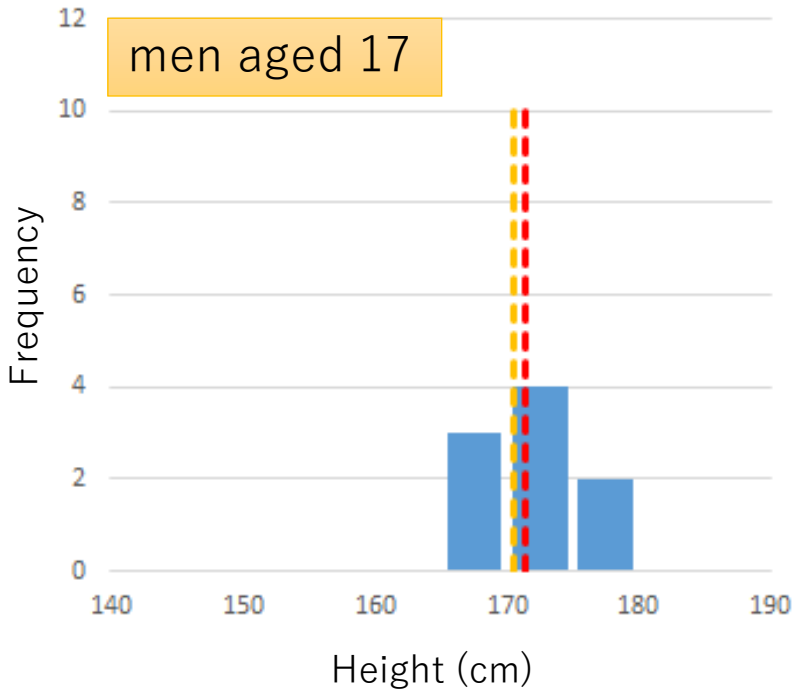
p-value > 0.05

men aged 15-17

Welch's test: equal to JPN national average ?







Welch's test

men aged 15-17

gender	age	sample size	average	sample size JPN	average JPN	Deg. of freedom	p-value
men	15	10	165.5	1411	168.37	9.07	0.262
men	16	18	169.2	1428	169.59	17.84	0.708
men	17	9	171.3	1427	170.46	8.23	0.518

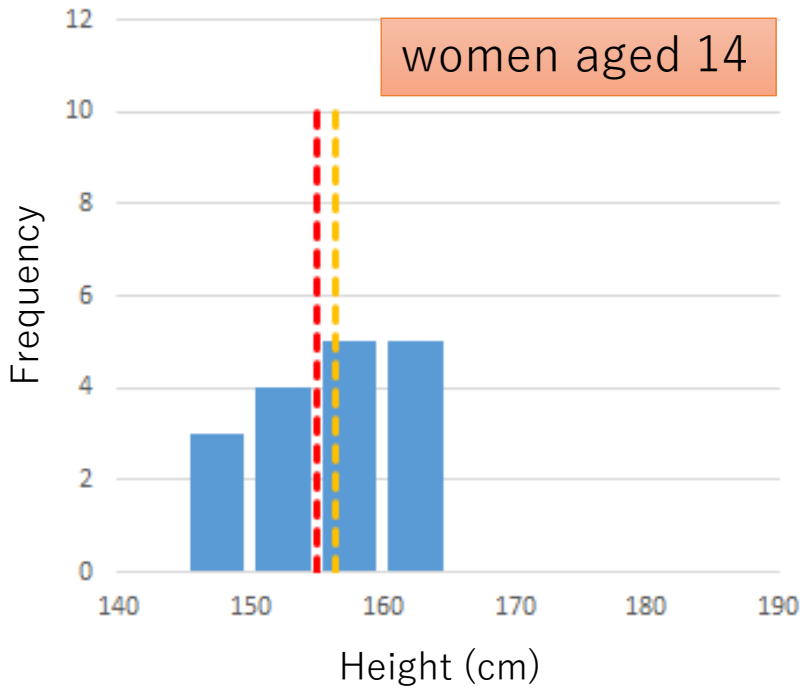
Yes!

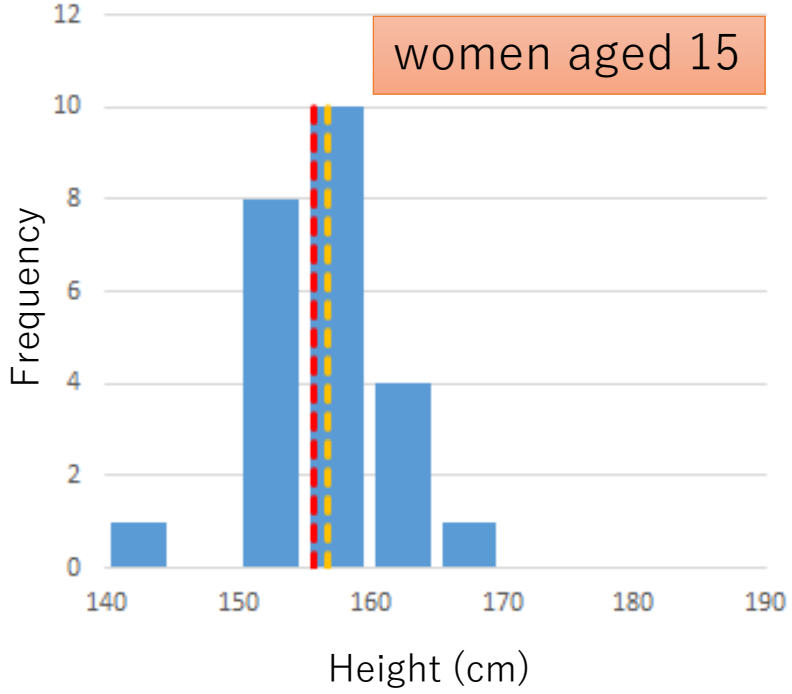
equal to JPN average

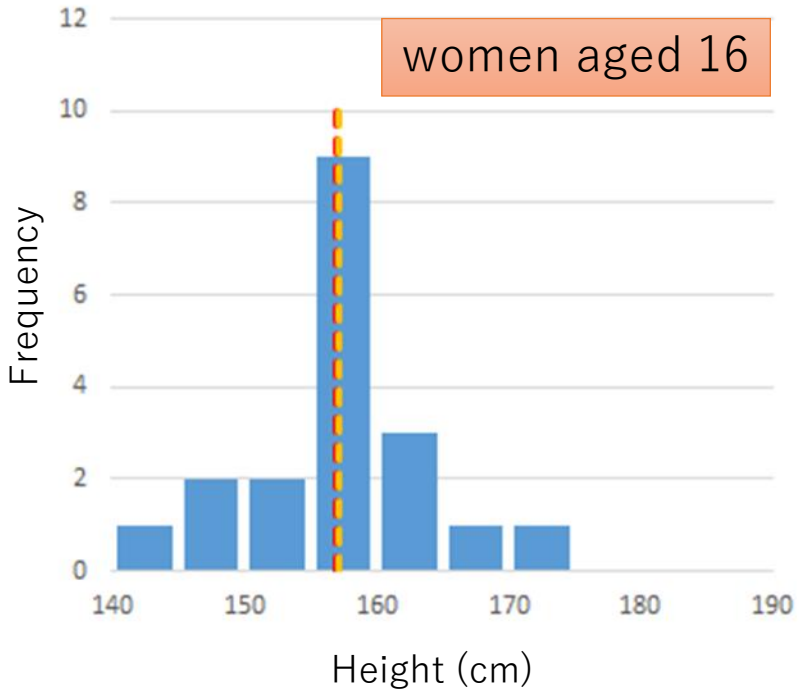
p-value > 0.05

women aged 14-16

Welch's test: equal to JPN national average ?







Welch's test

women aged 14-16

gender	age	sample size	average	sample size JPN	average JPN	Deg. of freedom	p-value
women	14	17	155.0	1386	156.36	16.29	0.374
women	15	24	155.7	1413	156.76	23.91	0.300
women	16	19	156.9	1419	157.16	18.30	0.860

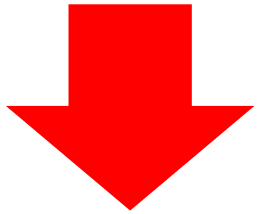
Yes!

equal to JPN average

p-value > 0.05

the analysis results

The results of the Shapiro-Wilk test and Welch's test rarely happened when many submitters disclosed their ages, genders, and heights dishonestly



most of the submitters disclosed
their ages, genders, and heights honestly

discussions

- Age, gender, and height were important personal information.
- It is likely that they disclosed not only their ages, genders, and heights but also other personal profile items honestly.
- ‘make a promise’ has an affect on submitters’ minds to disclose their personal information honestly ?

conclusion

- we investigated tweets disclosing
  submitters' personal profile items and
  analyzed submitters' ages, genders, and heights
  statistically.

- the results of our statistical analysis showed

  it is likely that most of the submitters disclosed
  their personal profile items honestly

future works

- statistical analysis on tweets
  in languages other than Japanese.

- whether submitters were concerned about
  their privacy and security risks caused by submitting
  tweets disclosing their personal profile items
  honestly.