10 TRAITS OF SUCCESSFUL DATERS

How to make someone like you in 4 minutes?

SPEED DATING

Hello

Goodbye

♡ **INTRODUCTION**

- Over 500 participants

- 21 sessions

- pre-, peri-, and post-event data

- Target variable: success of a participant in a given round

- Success: getting "likes" from over 50% of the opposite sex

01

**DATA CLEANING
&
FEATURE ENGINEERING**

02

EDA

03

**BASELINING
&
FEATURE SELECTION**

04

MODELING & RESULTS

01

DATA CLEANING
&
FEATURE ENGINEERING

SELECTED COLUMNS

iid   id   wave

age   race

from   gender

goal   date

imprace

imprelig

go_out

career_c

sports   tvsports   tv

exercise   art   movies

concerts   dining   hiking

museums   clubbing   yoga

gaming   reading   music

theater   shopping

attr3_1   sinc3_1

fun3_1   intel3_1

amb3_1

match_es

exphappy

WAVE 4, 18 MEN & 18 WOMEN

gender	from	age	dec_o	race	imprace	imprelig	...	shopping	yoga	exphappy	attr3_1	sinc3_1	fun3_1	intel3_1	amb3_1	match_es	popularity
1	Bangladesh	28.0	3	4.0	6.0	8.0	...	4.0	5.0	5.0	6.0	9.0	9.0	9.0	7.0	4.0	no
1	Long Island	26.0	2	2.0	7.0	5.0	...	3.0	3.0	6.0	7.0	9.0	8.0	10.0	7.0	2.0	no
1	Argentina	26.0	0	3.0	3.0	6.0	...	7.0	3.0	7.0	7.0	9.0	9.0	9.0	9.0	4.0	no
1	Brooklyn, NY	29.0	10	2.0	4.0	1.0	...	3.0	5.0	5.0	7.0	5.0	7.0	6.0	6.0	2.0	popular
1	Cincinnati, Ohio	39.0	2	2.0	1.0	1.0	...	2.0	7.0	5.0	8.0	8.0	8.0	8.0	8.0	5.0	no
1	New York	28.0	17	2.0	1.0	1.0	...	8.0	5.0	8.0	10.0	10.0	10.0	10.0	8.0	10.0	popular
1	Israel	30.0	5	6.0	8.0	9.0	...	6.0	2.0	5.0	7.0	8.0	8.0	9.0	6.0	4.0	no
1	India	22.0	0	4.0	3.0	2.0	...	8.0	4.0	7.0	8.0	8.0	8.0	9.0	9.0	5.0	no
1	Canada	32.0	2	2.0	2.0	6.0	...	5.0	2.0	6.0	8.0	5.0	7.0	10.0	9.0	8.0	no
1	Chicago	26.0	5	2.0	8.0	9.0	...	6.0	3.0	6.0	8.0	8.0	8.0	9.0	9.0	3.0	no
1	Colorado	27.0	9	2.0	6.0	8.0	...	3.0	7.0	8.0	9.0	8.0	7.0	9.0	9.0	3.0	popular
1	Texas	25.0	10	1.0	1.0	1.0	...	5.0	5.0	5.0	8.0	7.0	9.0	8.0	6.0	2.0	popular
1	Texas	22.0	5	3.0	1.0	3.0	...	5.0	3.0	5.0	4.0	7.0	3.0	9.0	2.0	4.0	no
1	Ann Arbor, MI	21.0	12	2.0	1.0	1.0	...	1.0	10.0	6.0	7.0	9.0	9.0	9.0	6.0	2.0	popular
1	Spain	24.0	10	2.0	1.0	1.0	...	2.0	2.0	5.0	4.0	8.0	3.0	7.0	2.0	3.0	popular
1	Hong Kong	25.0	8	4.0	1.0	1.0	...	5.0	3.0	5.0	6.0	9.0	8.0	9.0	8.0	2.0	no
1	Ann Arbor	28.0	5	2.0	4.0	6.0	...	3.0	6.0	6.0	7.0	9.0	8.0	8.0	7.0	3.0	no

02

EXPLORATORY DATA ANALYSIS

POPULARITY DISTRIBUTION

Number of participants who are classified as 'popular' VS those who are not.

	Popular	Less Popular
60		
40		
20		
0		
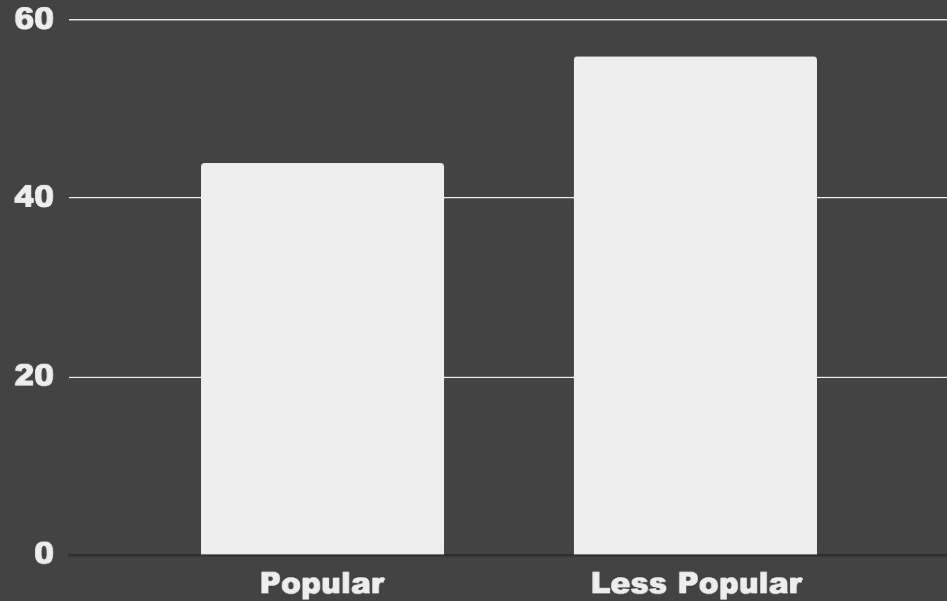
GENDER DISTRIBUTION

Among those who are 'popular', the male-to-female ratio is roughly 40:60
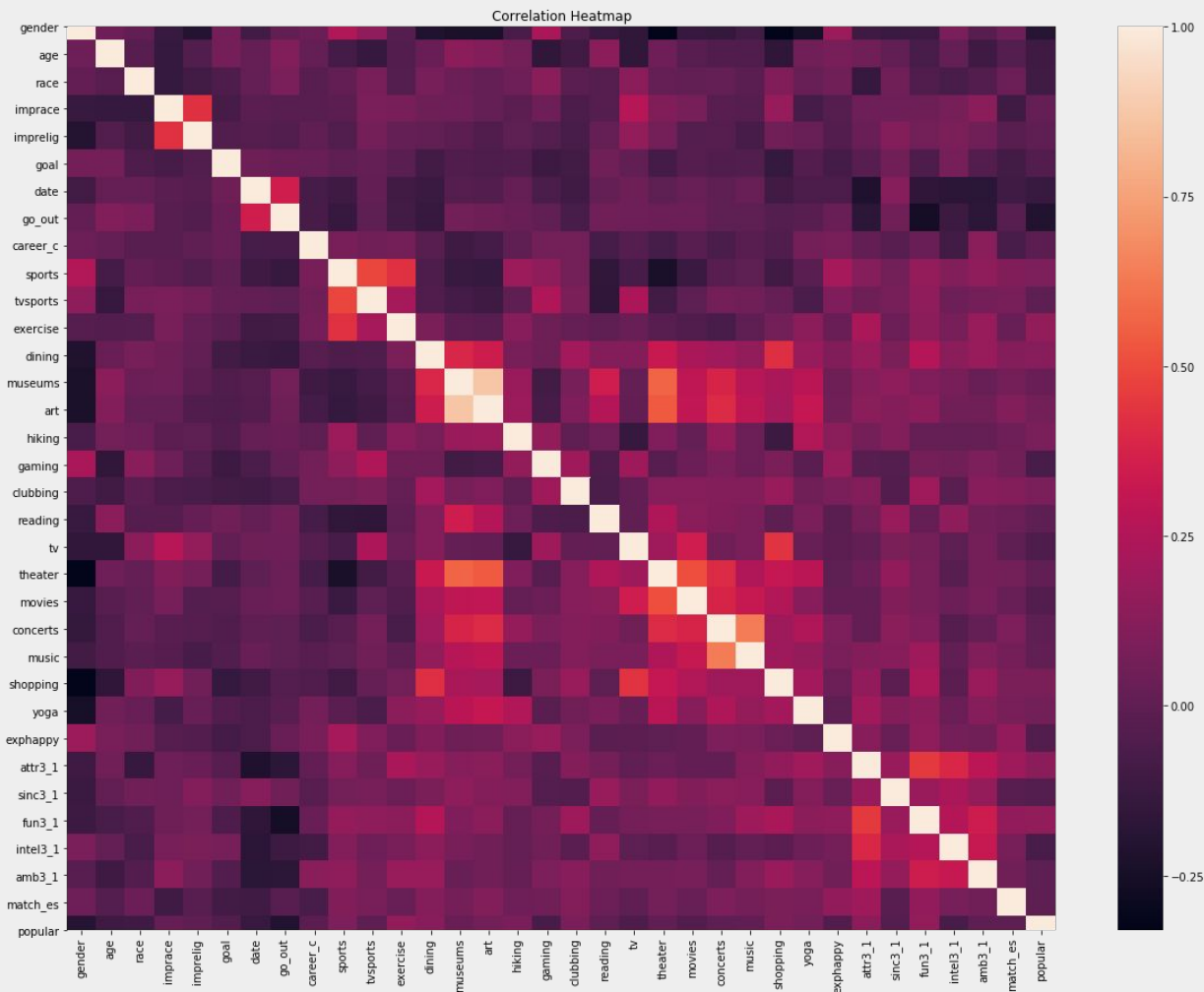
HEATMAP

A few moderately high correlations.

Dropped 'concerts', 'museums', and 'music'.

03

BASELINING & FEATURE SELECTION

COMPARISON

	LOGISTIC REGRESSION	RANDOM FOREST	XGBOOST
Validation Accuracy	0.6727	0.6090	0.5818
F1 score	0.5610	0.5274	0.5106

MOST IMPORTANT FEATURES

age

gender

goal

What do you want to get out of the speed dating experience

match_es

How many matches do you estimate you will get

date

How often do you go on dates

exphappy

how happy do you expect to be with the people you meet during the event

reading   **theater**   **art**

Rating out of 10 on those interests

attr3_1

How do you rate your own attractiveness

04

MODELING & RESULTS

MODEL SELECTED: LOGISTIC REGRESSION

70%
Accuracy

0.6292
F score - weighted average of recall and precision

NEXT STEPS

CHANGE THRESHOLD

Define success at
different rates

ONLINE DATING DATA

For higher relevance
in modern dating

OTHER MODELS

Explore other models
and multi-class
classification

But remember, at the end of the day, popularity has nothing to do with finding true love!

— ME

THANKS!

Presented by Linan Chen

Do you have any questions?

linanconviechen@gmail.com