



Dating Apps Matching Algorithms



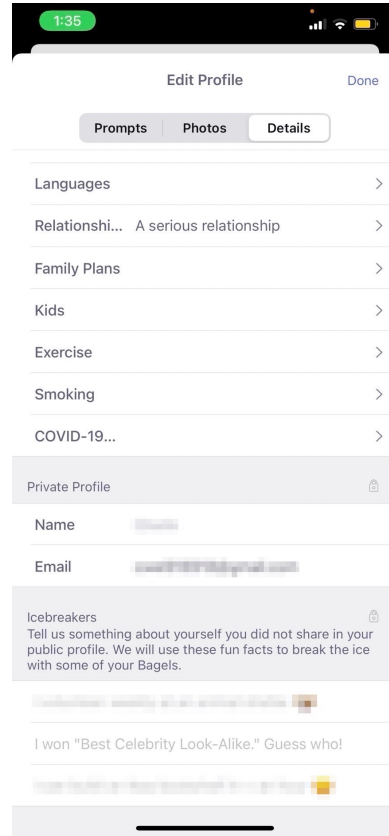
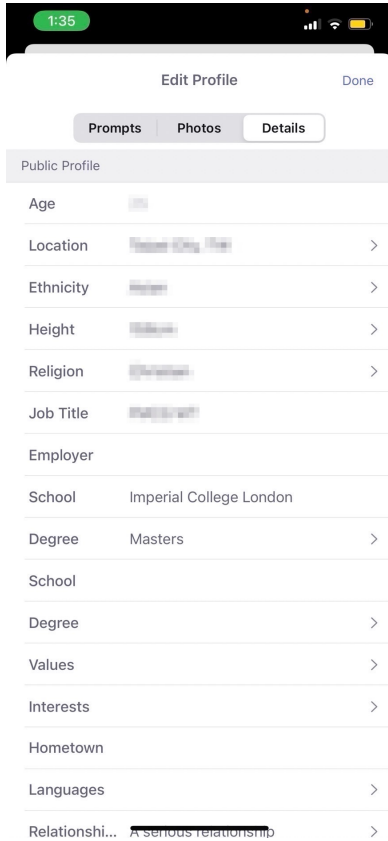


OkCupid & Hinge: Quick Facts

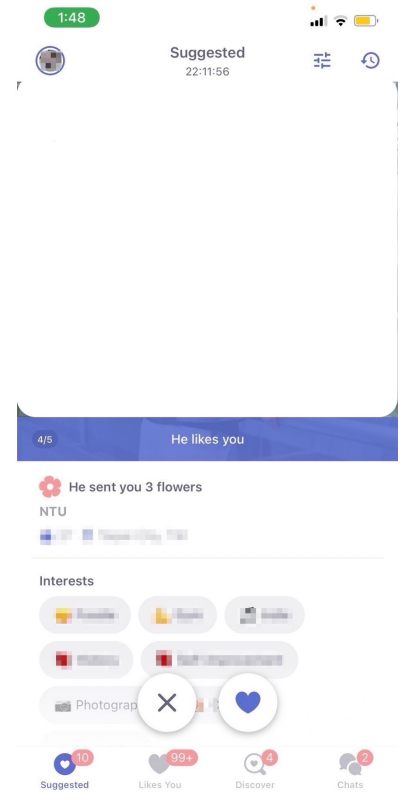
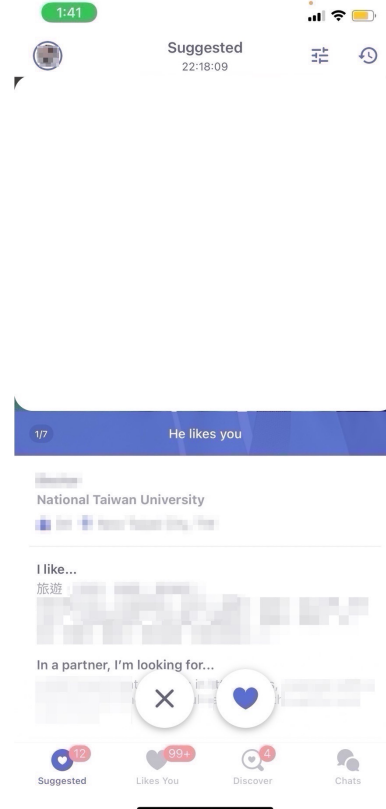
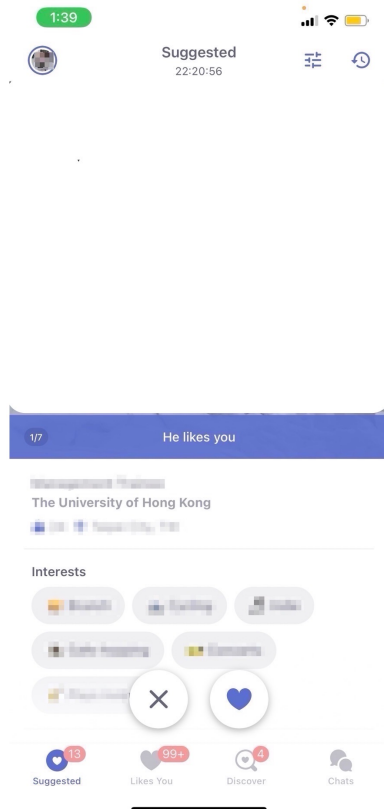
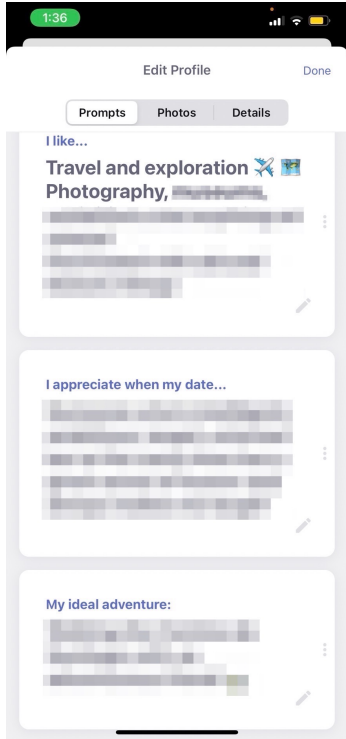
- ♥ U.S.-based online dating apps.
- ♥ Both owned by Match Group.
- ♥ OkCupid: founded in 2003 by a group of Harvard Math majors.
- ♥ Hinge: founded in 2012 by Justin McLeod.
- ♥ OkCupid: match percentage.
- ♥ Hinge: Gale-Shapley algorithm.



Coffee Meets Bagel



Is Algorithm the Culprit?





OkCupid: Match Questions



Question 1.

a) Should partners know one another's passwords?

- Yes
- No

b) How would you like your potential match to answer the question?

- Yes
- No

c) How important the question is to you?

- Irrelevant
- A little important
- Somewhat important
- Very important
- Mandatory

Question 2.

a) Would you be open to a long distance relationship?

- Yes, for the right person we can make it work.
- No, long distance is not my thing.

b) How would you like your potential match to answer the question?

- Yes, for the right person we can make it work.
- No, long distance is not my thing.

c) How important the question is to you?

- Irrelevant
- A little important
- Somewhat important
- Very important
- Mandatory

Question 3.

a) Do you enjoy planning a date?

- Yes, I am always the person with date ideas.
- No, I prefer to have my partner do the heavy lifting.

b) How would you like your potential match to answer the question?

- Yes, I am always the person with date ideas.
- No, I prefer to have my partner do the heavy lifting.

c) How important the question is to you?

- Irrelevant
- A little important
- Somewhat important
- Very important
- Mandatory

♥ Around 4,000 questions in the system.

♥ More questions answered => more accurate match result.

♥ n=3 questions in our case.

What Goes Into an OkCupid Match Question?

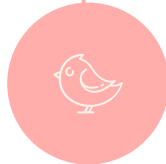
Your own answer



How you'd like your match to answer

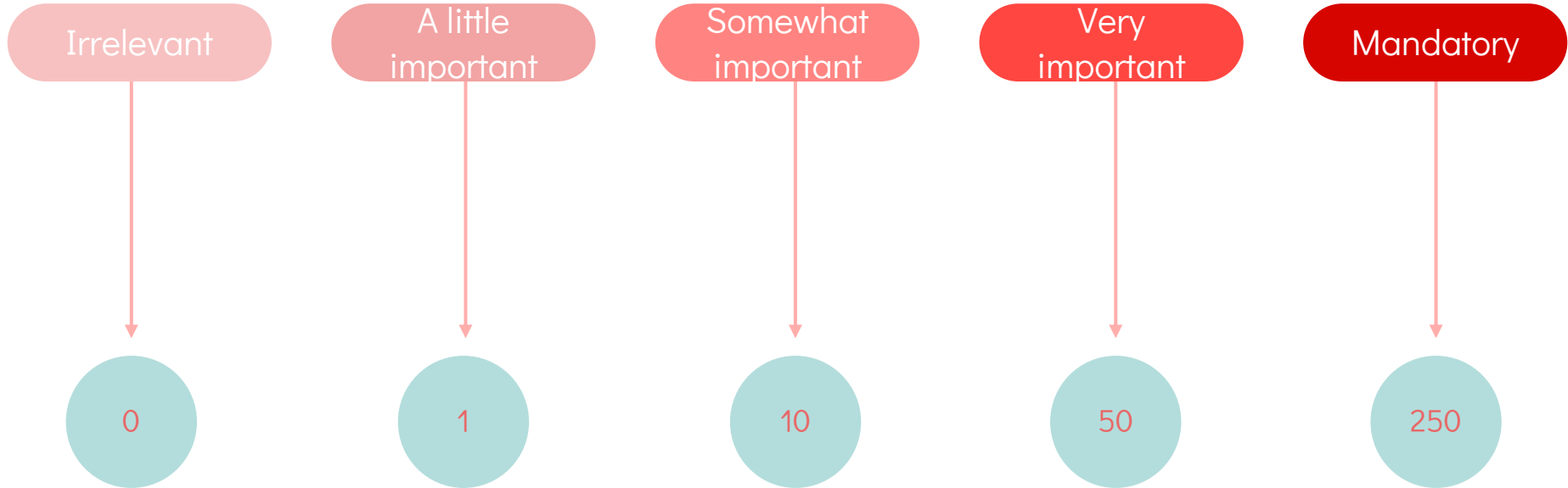


How important the question is to you





How Important the Question is To You?





Algorithm



For person A/B:

Add up importance weights (IW) for the n questions answered in common.

If A/B's response agrees with how A/B anticipates A/B to respond, add IW to the running count, otherwise, do nothing.

Divide the number obtained from step 2 by that obtained from step 1; the result represents A/B's percent satisfaction (PS).

Repeat step 1 through step 3 for another person.

Take the n^{th} root of the product of the two percent satisfactions obtained, i.e., find $\sqrt[n]{PS \text{ for } A \times PS \text{ for } B}$.

Adjust for the MoE and *Voilà!!*

Person A:

Question 1.

a) Should partners know one another's passwords?

- Yes
 No

b) How would you like your potential match to answer the question?

- Yes
 No

c) How important the question is to you?

- Irrelevant (0)
 A little important (1)
 Somewhat important (10)
 Very important (50)
 Mandatory (250)

Question 2.

a) Would you be open to a long distance relationship?

- Yes, for the right person we can make it work.
 No, long distance is not my thing.

b) How would you like your potential match to answer the question?

- Yes, for the right person we can make it work.
 No, long distance is not my thing.

c) How important the question is to you?

- Irrelevant (0)
 A little important (1)
 Somewhat important (10)
 Very important (50)
 Mandatory (250)

Question 3.

a) Do you enjoy planning a date?

- Yes, I am always the person with date ideas.
 No, I prefer to have my partner do the heavy lifting.

b) How would you like your potential match to answer the question?

- Yes, I am always the person with date ideas.
 No, I prefer to have my partner do the heavy lifting.

c) How important the question is to you?

- Irrelevant (0)
 A little important (1)
 Somewhat important (10)
 Very important (50)
 Mandatory (250)

Person B:

Question 1.

a) Should partners know one another's passwords?

- Yes
 No

b) How would you like your potential match to answer the question?

- Yes
 No

c) How important the question is to you?

- Irrelevant (0)
 A little important (1)
 Somewhat important (10)
 Very important (50)
 Mandatory (250)

Question 2.

a) Would you be open to a long distance relationship?

- Yes, for the right person we can make it work.
 No, long distance is not my thing.

b) How would you like your potential match to answer the question?

- Yes, for the right person we can make it work.
 No, long distance is not my thing.

c) How important the question is to you?

- Irrelevant (0)
 A little important (1)
 Somewhat important (10)
 Very important (50)
 Mandatory (250)

Question 3.

a) Do you enjoy planning a date?

- Yes, I am always the person with date ideas.
 No, I prefer to have my partner do the heavy lifting.

b) How would you like your potential match to answer the question?

- Yes, I am always the person with date ideas.
 No, I prefer to have my partner do the heavy lifting.

c) How important the question is to you?

- Irrelevant (0)
 A little important (1)
 Somewhat important (10)
 Very important (50)
 Mandatory (250)



Why Take the n^{th} Root?



n	Match Percentage
1	48%
2	69%
3	78%
4	83%
5	86%
6	88%
7	90%
8	91%
9	92%
10	93%
50	98.54%
100	99.27%
1000	99.93%
4000	99.98%



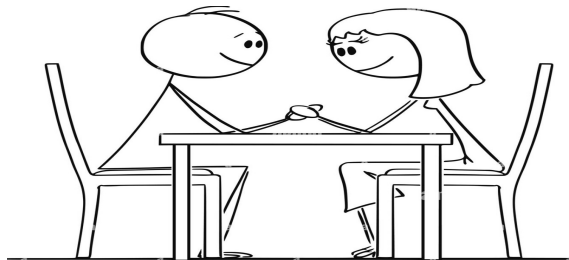
Encourages users to answer more questions!!



More questions answered in common
⇒ more confident
⇒ higher match percentage



Kinda Like Geometric Mean...



50% match

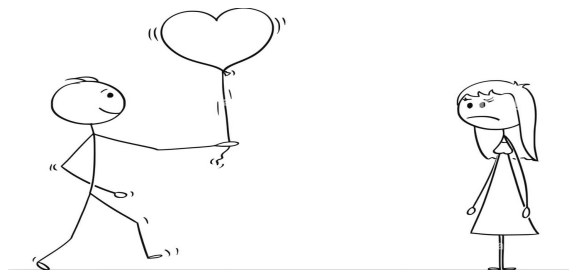
50% match

Arithmetic Mean

$$\frac{0.5 + 0.5}{2} = 0.5 = 50\%$$

Geometric Mean

$$\sqrt{0.5 \cdot 0.5} = 0.5 = 50\%$$



99% match

1% match

Arithmetic Mean

$$\frac{0.99 + 0.01}{2} = 0.5 = 50\%$$

Geometric Mean

$$\sqrt{0.99 \cdot 0.01} \cong 0.1 = 10\%$$



Margin of Error



n	Margin of Error	Highest attainable match percentage
1	100%	0%
2	50%	50%
3	33%	67%
4	25%	75%
5	20%	80%
10	10%	90%
50	2%	98%
100	1%	99%
1000	0.1%	99.9%
4000	0.025%	99.975%



Adjusted for the “margin of error.”



$\frac{1}{n}$, where $n := \#$ of questions answered in common.



Conservative approach: subtract off the MoE.



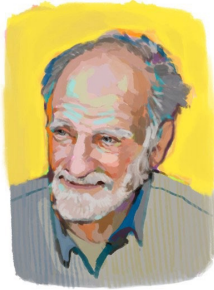
Get close to 100% by answering more questions.

Hinge: Gale-Shapley Algorithm

- ♥ Named after David Gale and Lloyd Shapley.
- ♥ College Admissions and the Stability of Marriage; 1962.
- ♥ Won Nobel in Economics.
- ♥ Hinge: “it’s all about pairing people who are likely to mutually like one another.”
- ♥ Profile information, prompt choices, likes/dislikes, “deal breaker” feature all help rank your preferences of potential matches.



David Gale
PROFESSOR, UC BERKELEY



Lloyd Shapley
PROFESSOR EMERITUS, UCLA



The Stable Marriage Problem



Def: A marriage is unstable if there are two women W_1 and W_2 who are married to men M_1 and M_2 , respectively, although W_2 prefers M_1 to M_2 and M_1 prefers W_2 to W_1 .



Gale-Shapley Algorithm

Preliminary: each man and woman rank the opposite sex based on his/her preferences.

Day 1: each woman proposes to her first choice of man. If a man receives more than one proposals, he rejects all but his first choice, and they get tentatively engaged.

Day 2: each rejected woman proposes to her second choice and tentative engagements are made as in the previous step.

Continue in this fashion until every man gets a proposal, or if all women are either tentatively engaged or rejected by all men.

For Never was a Story of More Algorithmic Complexity than This of Juliet and her Romeo...

Juliet	Rosaline	Romeo	Paris
Romeo	Romeo	Juliet	Juliet
Paris	Paris	Rosaline	Rosaline



It can be Shown



Thm. The maximum number of stages for pairing n couples is $n^2 - 2n + 2$, i.e., $O(n^2)$ time complexity.

Thm. The algorithm assigns stable marriages.

Thm. Those who make the proposals are assigned the best possible partners, whereas those who are proposed to are assigned the worst possible partners by the algorithm.

Much Ado About Nothing?



Is there a positive correlation between match percentage and dating success?



Does someone that fits your “ideal type” really makes a good partner?



Does match score really matter if most people on the platform are dating within their own race anyway?

Racial Bias?

2014

DateHookup "Let's Meet" Scores

		ASIAN women	BLACK women	LATINA women	WHITE women
men rating women	ASIAN men rating...	11%	-24%	9%	4%
	BLACK men rating...	7%	-9%	9%	-7%
	LATINO men rating...	12%	-27%	10%	6%
	WHITE men rating...	18%	-30%	6%	5%

		ASIAN men	BLACK men	LATINO men	WHITE men
women rating men	ASIAN women rating...	insuff.	-21%	9%	13%
	BLACK women rating...	5%	13%	-6%	-12%
	LATINA women rating...	-11%	-9%	14%	7%
	WHITE women rating...	-8%	-16%	4%	19%

Links



https://www.youtube.com/watch?v=m9PiPIRuy6E&ab_channel=TED-Ed



<https://blogs.ams.org/mathgradblog/2016/06/08/okcupid-math-online-dating/>



<https://medium.com/swlh/dating-data-an-overview-of-the-algorithm-afb9f0c08e2c>



[https://daily.jstor.org/dont-fall-in-love-okcupid/#:~:text=But%20perhaps%20I%20shouldn't,percent%2C%20seems%20to%20me\).](https://daily.jstor.org/dont-fall-in-love-okcupid/#:~:text=But%20perhaps%20I%20shouldn't,percent%2C%20seems%20to%20me).)



https://wiki.ubc.ca/Course:CPSC522/Analyzing_online_dating_trends_with_Weka



https://www.statista.com/topics/7443/online-dating/#topicHeader__wrapper



<https://www.gwern.net/docs/psychology/okcupid/raceandattraction20092014.html>



https://www.youtube.com/watch?v=LfTV6rIyhdo&ab_channel=Numberphile2