The wedding formula :
un problème de probabilité résolu en détail avec des mathématiques du secondaire

Yvik Swan
Département de Mathématique
ULiège

Nous sommes en train de vivre une **révolution sociale** de grande envergure via l’automatisation et l’amélioration des compétences des intelligences artificielles.

De nombreux métiers sont **menacés d’obsolescence**.

Même les notions d’emploi, métier, chômage, richesse seront à repenser dans un avenir proche.

Parmi les compétences encore porteuses d’avenir, il y a consensus sur les **disciplines STEM :**

Science, **T**echnology, **E**ngineering, **M**athematics

Et, des 4, les **mathématiques** sont peut-être les plus cruciales.

best jobs 2016

All Images News Videos Maps More Settings Tools

About 321.000.000 results (0,62 seconds)

- Financial Analyst
- Financial Manager
- Actuary
- Mathematician
- Statistician

The 20 best jobs in business for 2016 - Business Insider
www.businessinsider.com/best-business-jobs-2016-1

About this result • Feedback

The Best Jobs In 2016 - Forbes
www.forbes.com/sites/karstenstrauss/2016/04/14/the-best-jobs-in-2016/
Apr 14, 2016 - It would seem that this year science and math are king on the **job** market. Positions that require skills in numbers and sciences are some of the ...

Best Jobs 2017

CAREERCAST

JOB SEEKERS: Sign Up Login

Search Jobs | Post Your Resume | Job Seeker Tools | Niche Job Networks | Jobs Rated | Career

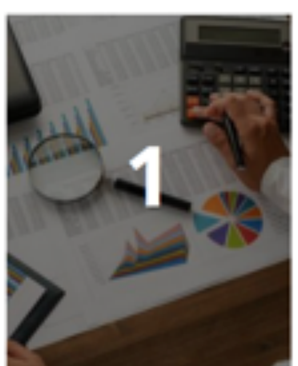
Find Jobs KEYWORDS: LOCATION:

in Share 208 Tweet Share

Jobs Rated Report 2017: Ranking 200 Jobs

By: CareerCast.com

Some of the careers ranking among the Top 10 best jobs of 2017 illustrate these changes. Careers in mathematics rank highly, with Statistician, Mathematician and Data Scientist all in the top 10 — but these jobs aren't just about numbers in the current landscape. A career in one of these mathematical fields can translate into work in healthcare, business, marketing, even entertainment.



Statistician

Overall Rating: 1/199 Median Salary: $80,110

Work Environment	Stress	Projected Growth
Very Good	Very Low	Very Good
4/199	39/199	3/199

View Raw Scores

Search Statistician Jobs

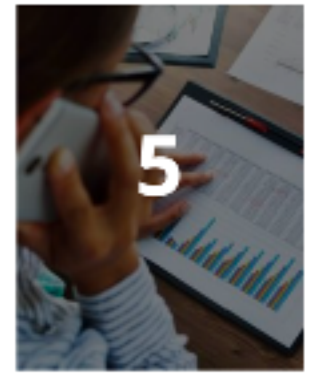
Operations Research Analyst

Overall Rating: 3/199 Median Salary: $79,200

Work Environment	Stress	Projected Growth
Very Good	Very Low	Very Good
35/199	10/199	6/199

View Raw Scores

Search Operations Research Analyst Jobs



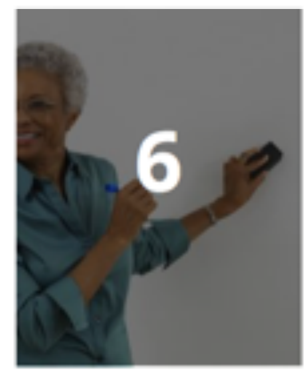
Data Scientist

Overall Rating: 5/199 Median Salary: $111,267

Work Environment	Stress	Projected Growth
Very Good	Very Low	Very Good
12/199	37/199	37/199

View Raw Scores

Search Data Scientist Jobs



University Professor

Overall Rating: 6/199 Median Salary: $72,416

Work Environment	Stress	Projected Growth
Very Good	Very Low	Very Good
1/199	6/199	39/199

View Raw Scores

Search University Professor Jobs

Mathematician

Overall Rating: 7/199 Median Salary: $111,298

Work Environment	Stress	Projected Growth
Good	Very Low	Very Good
66/199	28/199	18/199

View Raw Scores

Search Mathematician Jobs

Software Engineer

Overall Rating: 8/199 Median Salary: $100,690

Work Environment	Stress	Projected Growth
Good	Very Low	Very Good
53/199	24/199	32/199

View Raw Scores

Search Software Engineer Jobs

Dispose-t-on de mesures **objectives quantifiées** de l'importance **concrète** des mathématiques dans la société?

Oui.

France (2015)

Communiqué de presse

Quel est l'impact socio-économique des mathématiques en France ?

Le premier rapport consacré aux mathématiques en France confirme leur forte contribution à l'économie nationale et révèle/considère qu'elles joueront un rôle majeur pour relever les défis industriels et sociétaux de demain.

Paris, le 27 mai 2015 - Une étude afin de mesurer l'impact socio-économique des mathématiques en France a été commanditée par l'Agence pour les Mathématiques en Interaction avec l'Entreprise et la Société (AMIES), en partenariat avec la Fondation Sciences Mathématiques de Paris (FSMP) et la Fondation Mathématique Jacques Hadamard (FMJH), et en association avec les Labex de mathématiques.

Réalisée par le cabinet de conseil en stratégie CMI, cette étude réalisée pour la première fois en

http://www.agence-maths-entreprises.fr/a/sites/default/files/Eisem/20150527_Communiqu%C3%A9_de_presse_EISEM_v2.pdf

Conclusions

> D'un point de vue économique

- La valeur ajoutée apportée par les mathématiques en France représente 285 milliards d'euros, soit 15% de la valeur ajoutée française.
- Le nombre d'emplois impactés directement par les mathématiques en France s'élève à 2,4 millions, soit 9% du nombre total d'emplois en 2012, tous secteurs d'activités confondus.
- Les relations directes avec l'industrie, dans le cadre de collaborations contractuelles avec des laboratoires de mathématiques, restent encore peu significatives mais la grande transversalité des mathématiques permet de résoudre des problématiques industrielles de plus en plus complexes.
- Le rôle des mathématiques dans le développement industriel, notamment leur contribution aux technologies clés, est appelé à se renforcer.
- On estime que 30% des emplois mobilisant des mathématiques résultent de l'application directe des résultats de la recherche et des outils mathématiques.
- L'impact des mathématiques en France est très comparable à celui constaté au Royaume-Uni.

Hollande (2014)

Rapport Deloitte

Wiskunde nodig om arbeidsmarkt te redden

Wiskunde verdient jaarlijks 160 miljard voor Nederland BV

Het gaat goed met de Nederlandse wiskunde, maar er zijn ook urgente zorgen. Platform Wiskunde Nederland (PWN) heeft voor het ministerie van Economische zaken het rapport "Mathematical sciences and their value for Dutch economy" (Deloitte) laten maken. Dit rapport laat de bijdrage van wiskunde aan de economie zien en kijkt naar de toekomst. Wanneer we besluiten om niet verregaand te investeren in wiskunde kampt Nederland over dertig jaar met een overschot aan hoger en middelbaar opgeleiden die geen passend werk kunnen vinden.

https://www.platformwiskunde.nl/rapport-deloitte/

De tijd dat machines grote delen van ons dagelijkse leven gaan overnemen lijkt verre toekomst, maar is misschien wel dichterbij dan iedereen denkt. In de afgelopen twee decennia is technologie elk jaar verdubbeld in intelligentie en die trend zet zich de komende jaren voort. Dit zal ertoe leiden dat langzaam, maar gestaag, steeds meer werkgelegenheid wordt overgenomen door technologie.
Wat deze trend anders maakt dan de mechanisatie- en automatiseringsbewegingen die sinds de industriële revolutie onze economie typeerden, is dat niet de klassieke lage-lonen werkgelegenheid vervangen gaat worden, maar juist de werkgelegenheid in de middenklasse en bij hoger opgeleiden, de kurk waar de arbeidsmarkt op drijft.

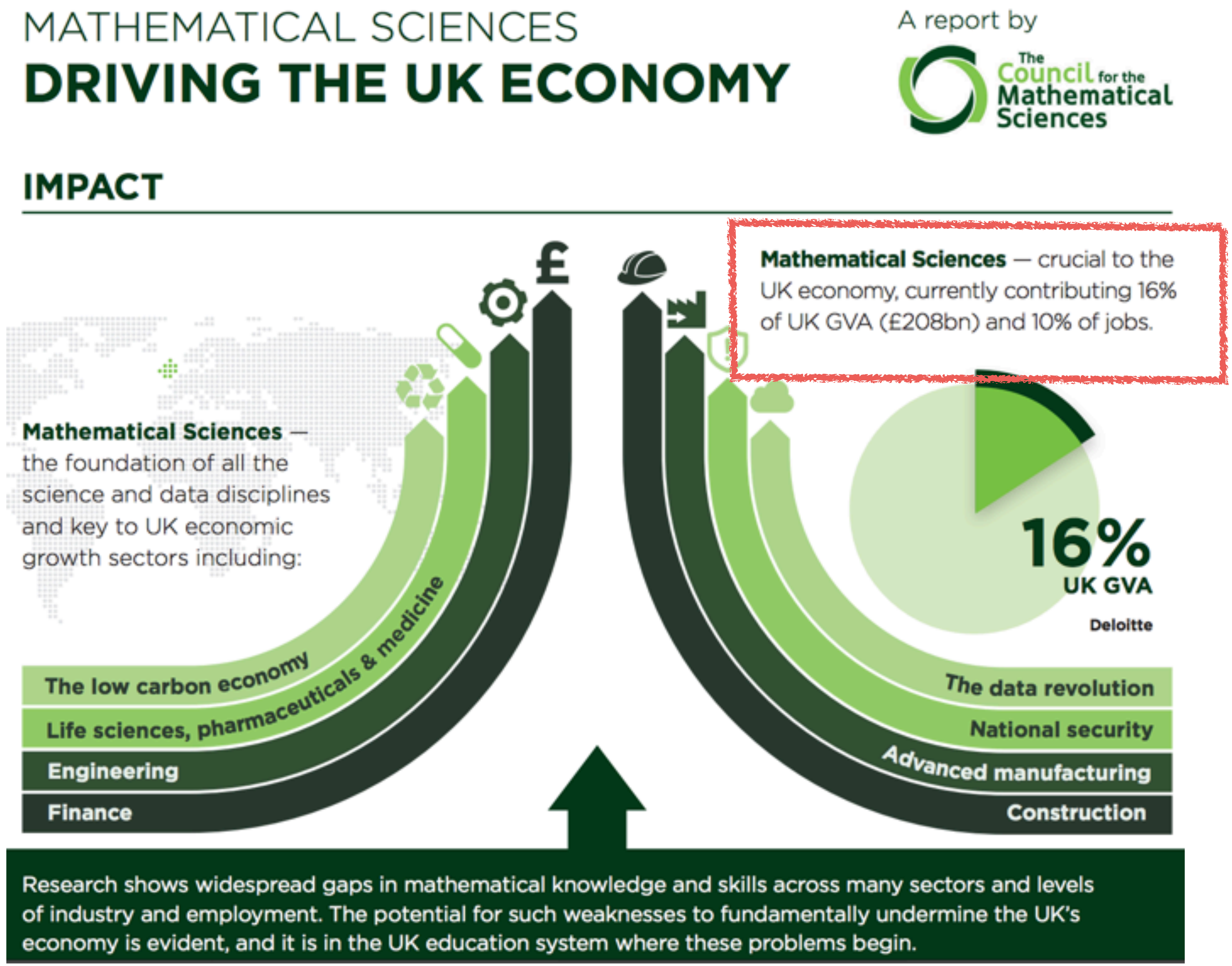
De oplossing voor de middenklasse: wiskunde. Een goede basis in wiskunde is een belangrijke pijler onder het succes van de Nederlandse economie.

Met de explosieve toename van data, onder andere door het internet, met steeds snellere computers en betere rekenmethodes, zal het belang van deze pijler in de komende jaren alleen maar toenemen. Het blijkt dan ook dat wiskundige kennis sterk samenhangt met de internationale concurrentiekracht van een land.

Nu gebruiken hoger opgeleide Nederlanders de tijd van 900.000 full time banen aan wiskunde in het dagelijkse werk. Zij variëren van wetenschappers die doorlopend met wiskunde bezig zijn, tot bankiers die wiskunde een deel van hun tijd gebruiken bijvoorbeeld om activa te waarderen en artsen die soms wiskunde nodig hebben om de resultaten van medische onderzoeken te kunnen interpreteren. Als een afgeleide van deze diensten en producten hebben nog eens 1,4 miljoen Nederlanders een baan.

Daarmee dragen gebruikers van wiskunde bij aan een kwart van de Nederlandse werkgelegenheid en, omdat zij relatief goed betaald worden, zorgen zij voor circa 30% van het bruto nationaal product.

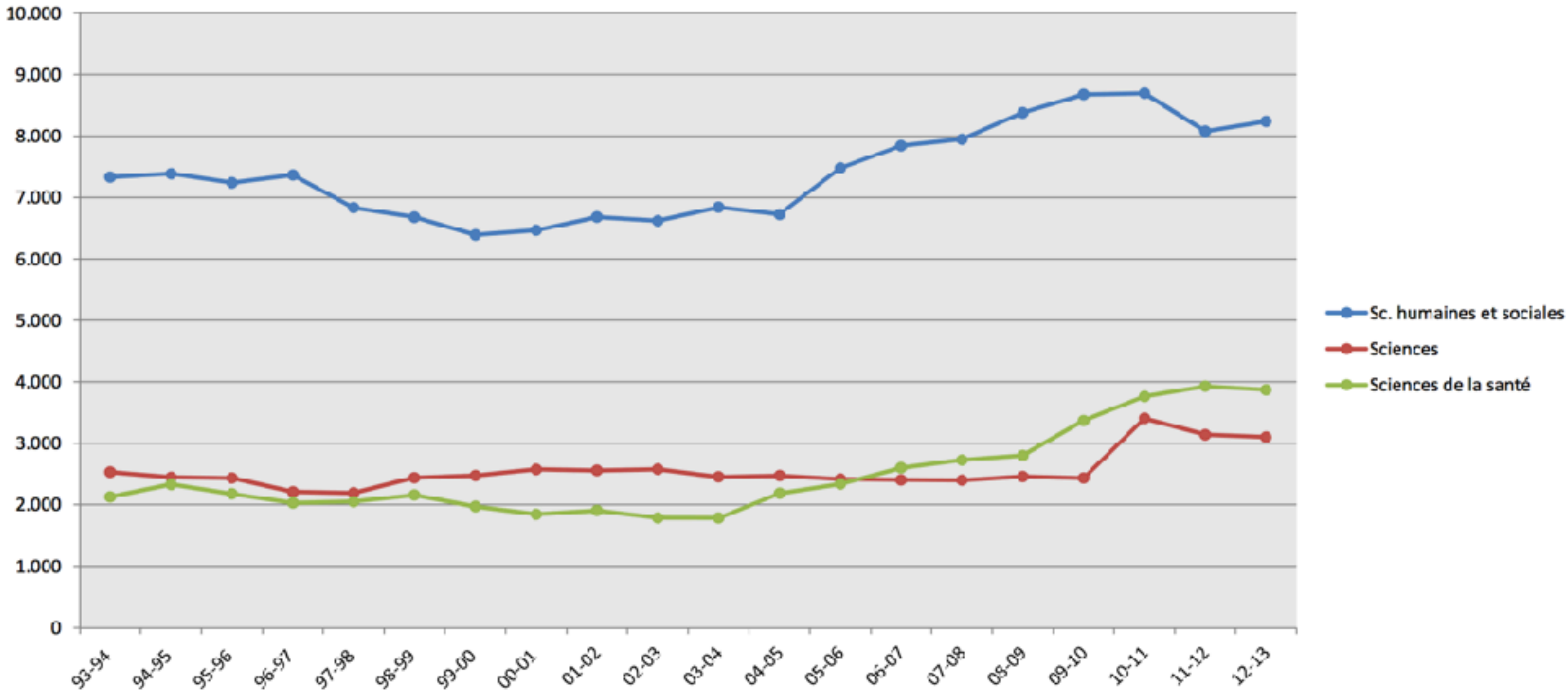
UK (2014)



http://www.cms.ac.uk/files/News/article_56a756668046f6.43082264.pdf

Pourtant le nombre d’étudiants en STEM n’augmente pas partout.

Evolution du nombre total d'étudiants de première génération universitaire par secteur d'études



http://www.cref.be/annuaires/apercu-historique

La Libre.be

Journal Alertez-nous Je me connecte

Pourquoi les sciences n'intéressent plus les étudiants ?

ABONNÉS **ISABELLE LEMAIRE** Publié le vendredi 13 mars 2015 à 07h13 - Mis à jour le vendredi 13 mars 2015 à 07h13

BELGIQUE Le refrain est connu : le désengouement pour les filières d'études scientifiques est une réalité depuis des années en Belgique francophone. Les chiffres d'inscriptions en première Bac pour l'année académique 2014-2015 le prouvent encore. A l'ULg, seuls 5,6 % des nouveaux inscrits ont choisi une des branches de la faculté des sciences (chimie, biologie, physique, mathématiques, géographie et géologie) et 5,4 % les filières de la faculté des sciences appliquées (ingénieur et informatique). A l'ULB (chiffres provisoires), le nombre d'inscriptions dans les filières des sciences et techniques est plus élevé. Elles recouvrent un peu plus de 20 % du total et sont en légère hausse.

A l'UCL, on ne se bouscule pas sur les bancs de la faculté des sciences : 52 étudiants sont inscrits en biologie, 40 en physique, 37 en chimie et 17 en mathématiques. Par contre, l'Ecole polytechnique de Louvain a enregistré une augmentation de 10 % du nombre d'inscriptions à l'examen d'entrée en ingénieur civil en 2014. Du jamais vu en 42 années d'existence.

Pour remédier à ce désintérêt il faut bien entendu qu'une prise de conscience du monde politique mène à des réformes bien dirigées et cohérentes.

En Hollande, la ***Platform Wiskunde*** créée en 2010 sert d'intermédiaire en toutes matières liées à la mathématique.

Le nombre d'élèves dans les filières mathématiques a explosé en Hollande ces quelques dernières années.

Il y a une discussion en Flandre sur la création d'un outil similaire.

Nous pouvons également agir à notre niveau : à l’ULg nous proposons

- Math à modeler
- Math en Jeans
- Un professeur de l'ULg dans vos classes
- …

Même si ces initiatives sont certainement utiles, elles ne sont pas suffisantes.

En particulier elles préservent les maths dans leur “tour d’ivoire” et ne renseignent pas vraiment sur les métiers.

Ces métiers sont difficiles à expliquer car **abstraits** : quelle est exactement l'occupation d'un actuaire, d'un analyste, d'un quant, d'un data scientist, d'un bio-statisticien, … ?

It was business as usual.

Les mathématiques utilisées par ces métiers sont souvent **trop avancées** pour qu'on puisse les utiliser pour justifier les formules que doivent apprendre les élèves.

Et donc les mathématiques deviennent, chez beaucoup, obscures et sèches :

les stats se résument à des calculs de moyenne,
les probas à des lancers de dé.

Dans cet exposé je vais tenter d'apporter une petite pierre aux fondations de votre cathédrale, en décrivant un problème

ludique

facile à expliquer mais difficile à résoudre

avec des applications pratiques évidentes

dont la résolution nécessite d'utiliser de nombreuses notions de math accessibles aux étudiants du secondaire :

Loi normale, Dénombrements, Loi des probabilités totales, Intégrales, Optimisation, Logarithme, Approximation et limite, …

Quel problème?

Celui de trouver l'amour

Version 1 : statistique des relations humaines

Où nous allons démontrer de façon irréfutable
qu'il est presque impossible de trouver l'amour.

Why I Will Never Have A Girlfriend

Tristan Miller
German Research Center for Artificial Intelligence*
Erwin-Schrödinger-Straße 57
67663 Kaiserslautern, Germany
tristan.miller@dfki.de

20 December 1999

Abstract

Informal empirical and anecdotal evidence from the (male) scientific community has long pointed to the difficulty in securing decent, long-term female companionship. To date, however, no one has published a rigorous study of the matter. In this essay, the author investigates himself as a case study and presents a proof, using simple statistical calculus, of why it is impossible to find a girlfriend.

Why don't I have a girlfriend?

Not the author, though. I, for one, refuse to spend my life brooding over my lack of luck with women. While I'll be the first to admit that my chances of ever entering into a meaningful relationship with someone special are practically non-existent, I staunchly refuse to admit that it has anything to do with some inherent problem with *me*. Instead, I am convinced that the situation can be readily explained in purely scientific terms, using nothing more than demographics and some elementary statistical calculus.

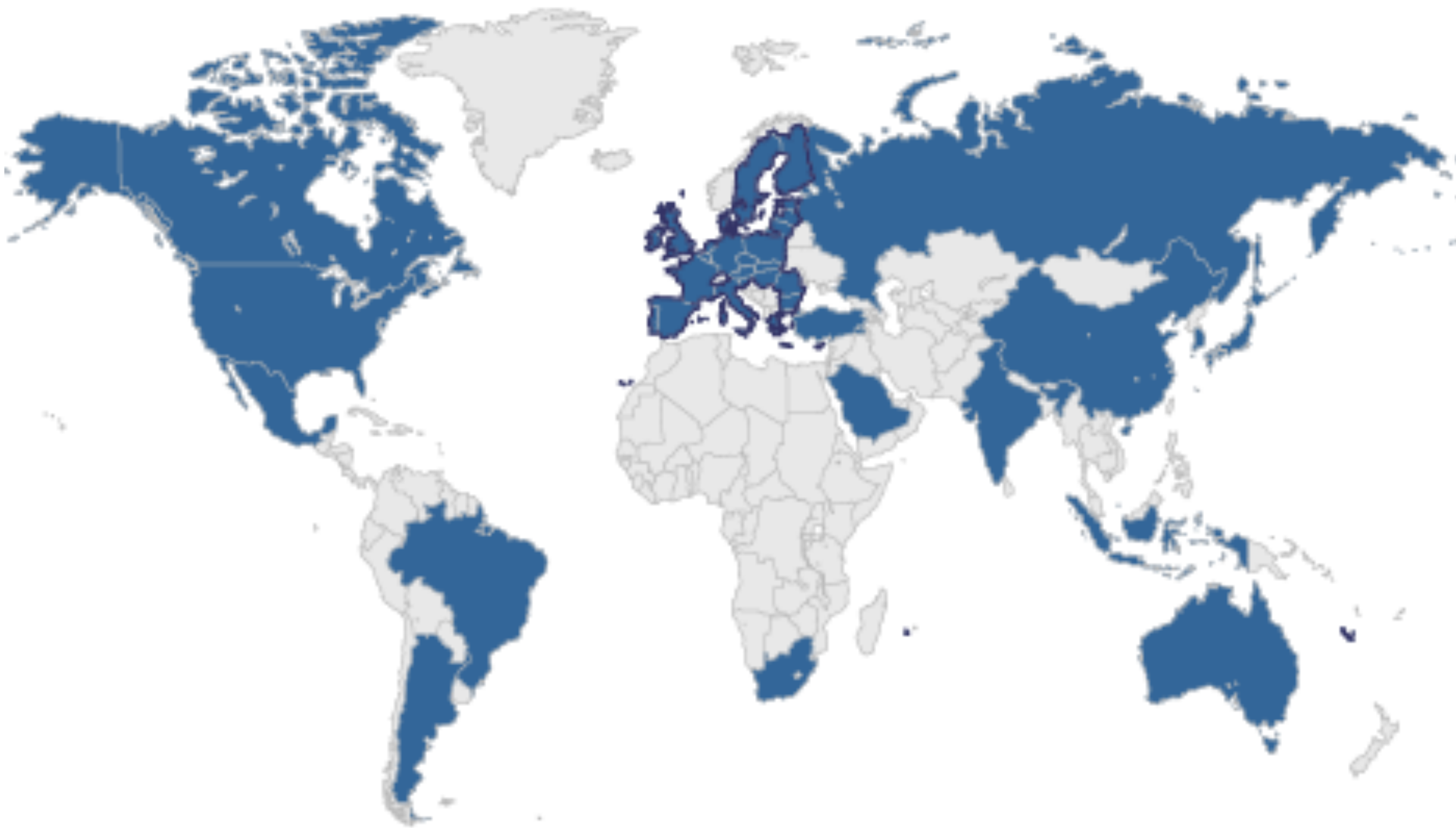
Lest anyone suspect that my standards for women are too high, let me allay those fears

Début mars 2017,
Wikipédia m’informe que nous étions plus de
7 488 669 636
êtres humains sur cette terre

Discrimination par le genre.

En se plaçant dans la peau d’une fille de 18 ans,
ceci laisse à peu près
3 477 830 000
prétendants (hommes) à disposition.

Discrimination géographique.



Si on restreint aux pays du G20 il en reste 2 327 570 000.

Discrimination par l’âge.

Il faut éliminer les trop jeunes et les trop vieux.

Cela laisse 1 598 040 000 mâles entre 15 et 64 ans.

Si de plus on suppose toutes les tranches d’âge équitablement représentées et qu’on se restreint aux 18-25 ans, ceci nous laisse avec

$$1\ 598\ 040\ 000\left(\frac{8}{50}\right)=255\ 686\ 400$$

prétendants “admissibles”.

Discrimination au facies.

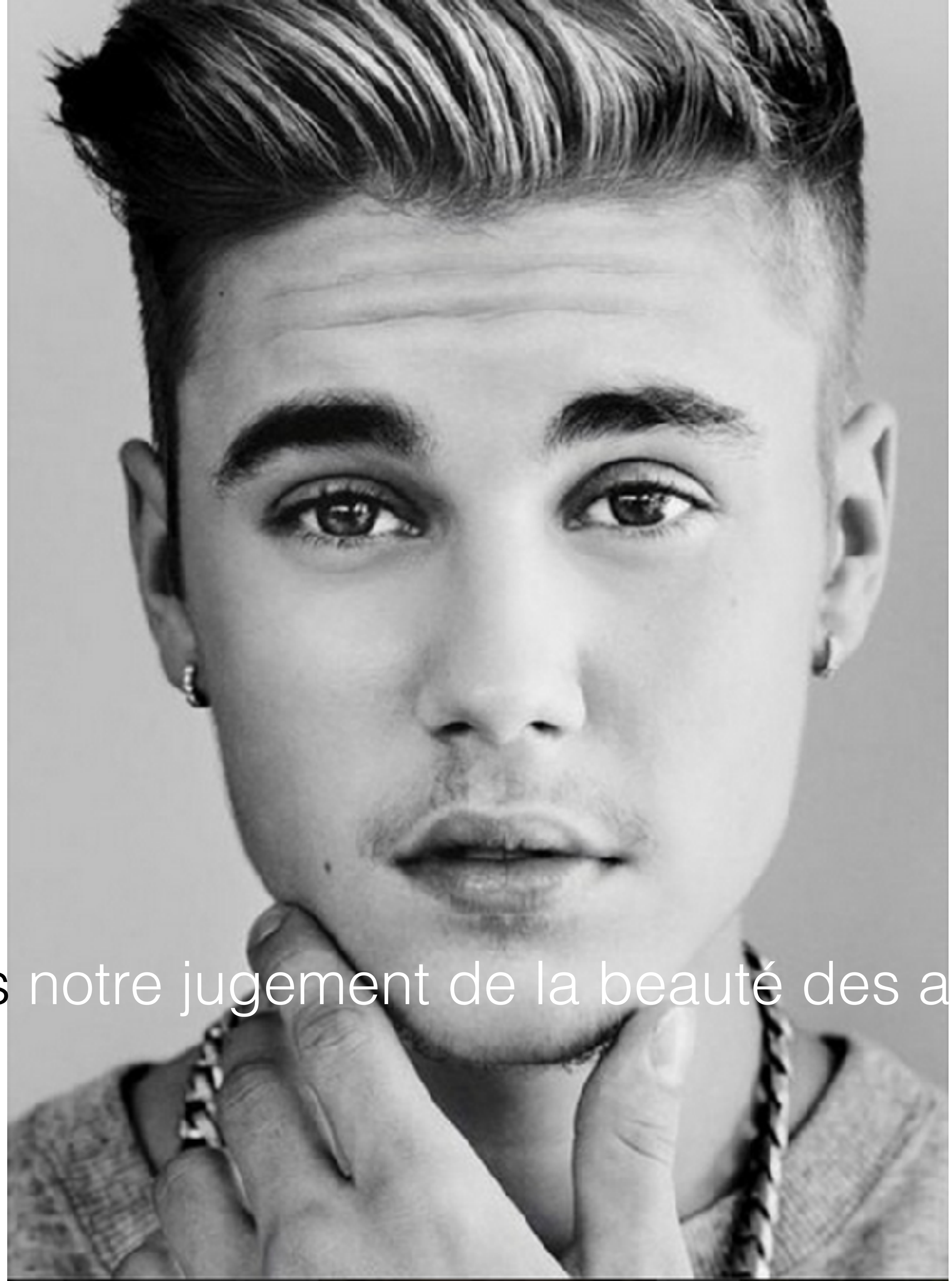
Il faut également que l’élu soit plaisant.

La beauté, comme tout le reste, comporte différents éléments d’ordre purement mathématique

Le plus évident est un ingrédient d’ordre géométrique.


Dans la nature

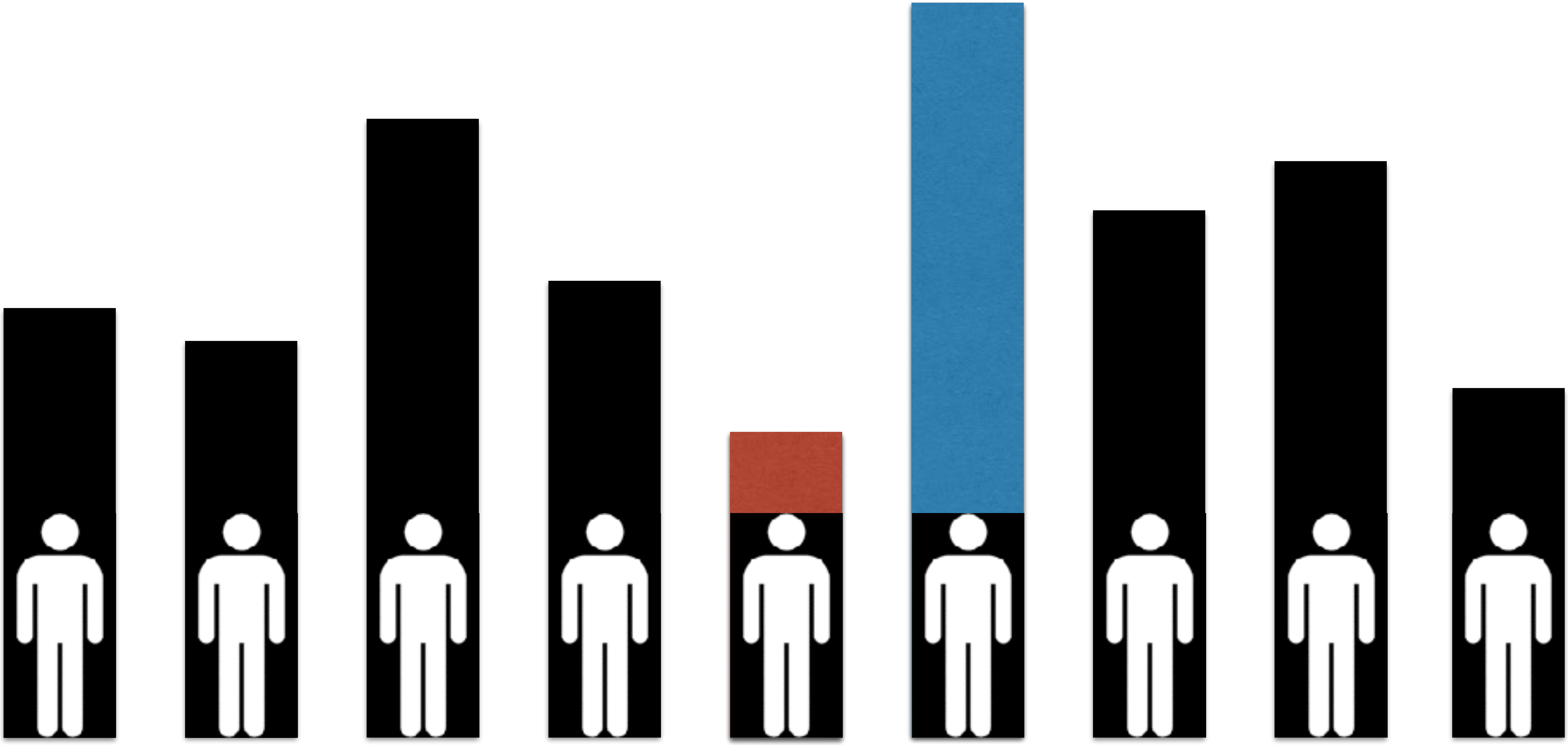

Dans l’art



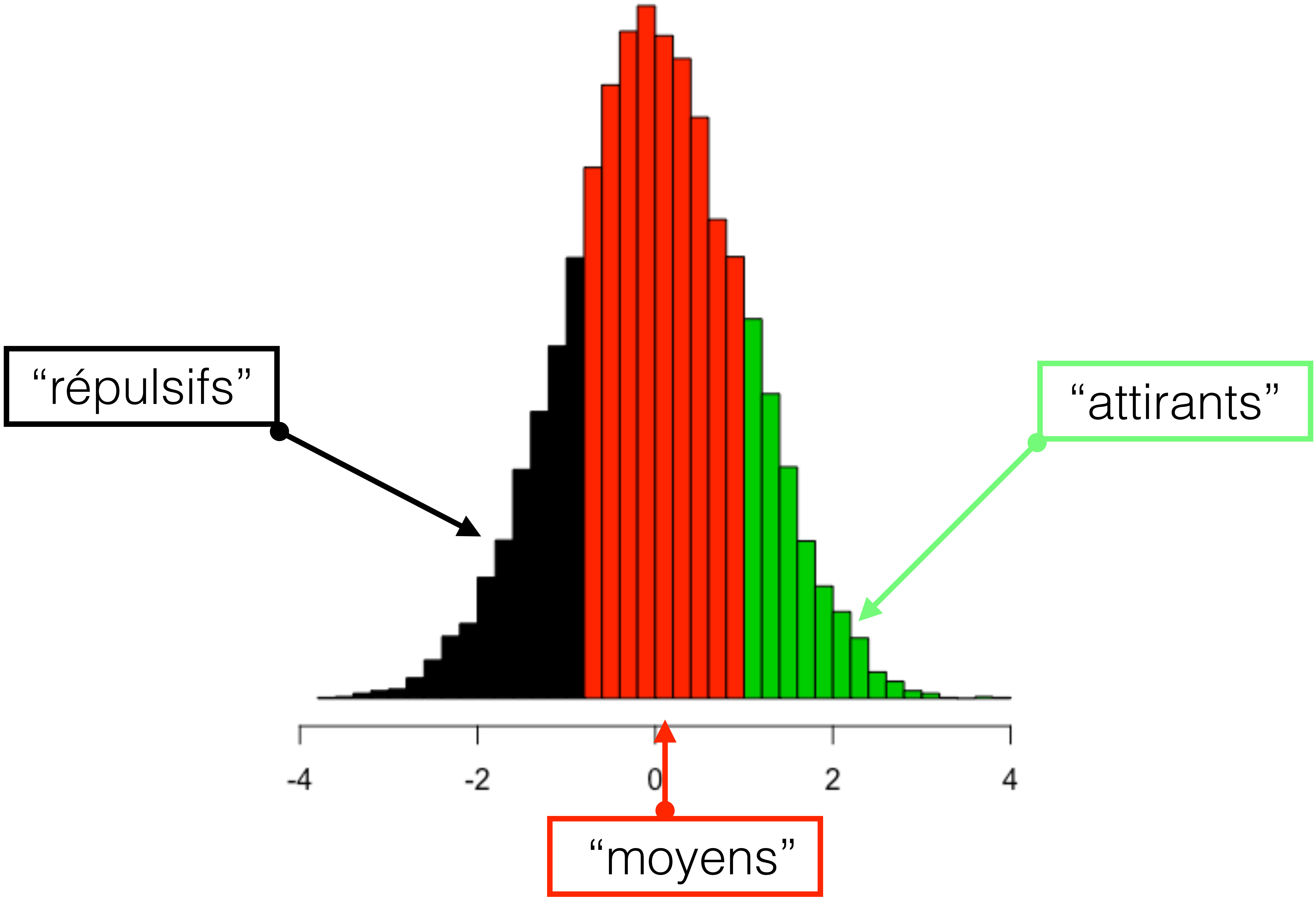
Dans notre jugement de la beauté des autres


Mais la beauté peut aussi être subjective

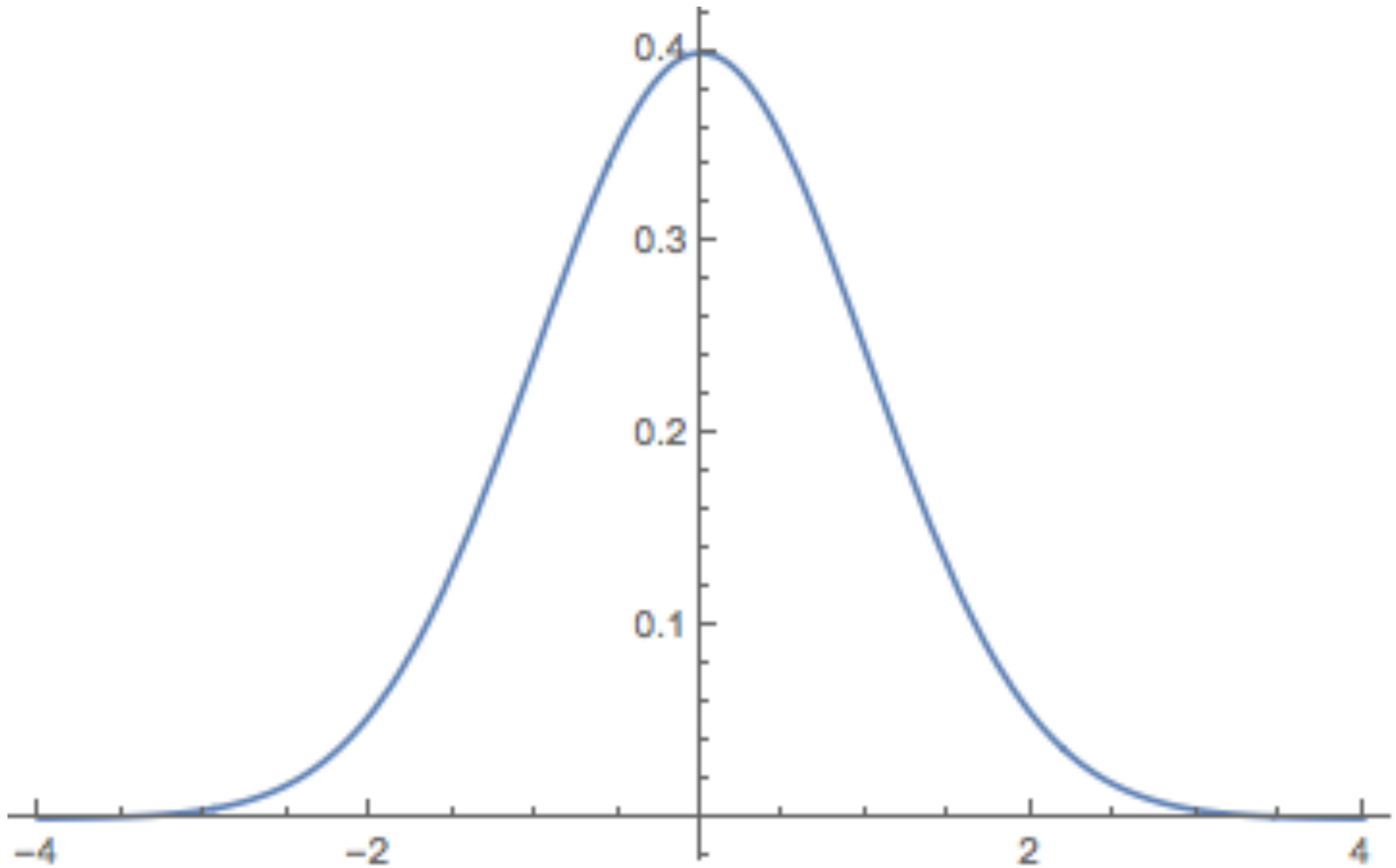
Il est raisonnable de supposer que chacun est en mesure de “classer” les prétendants.



On construit l'histogramme des candidats

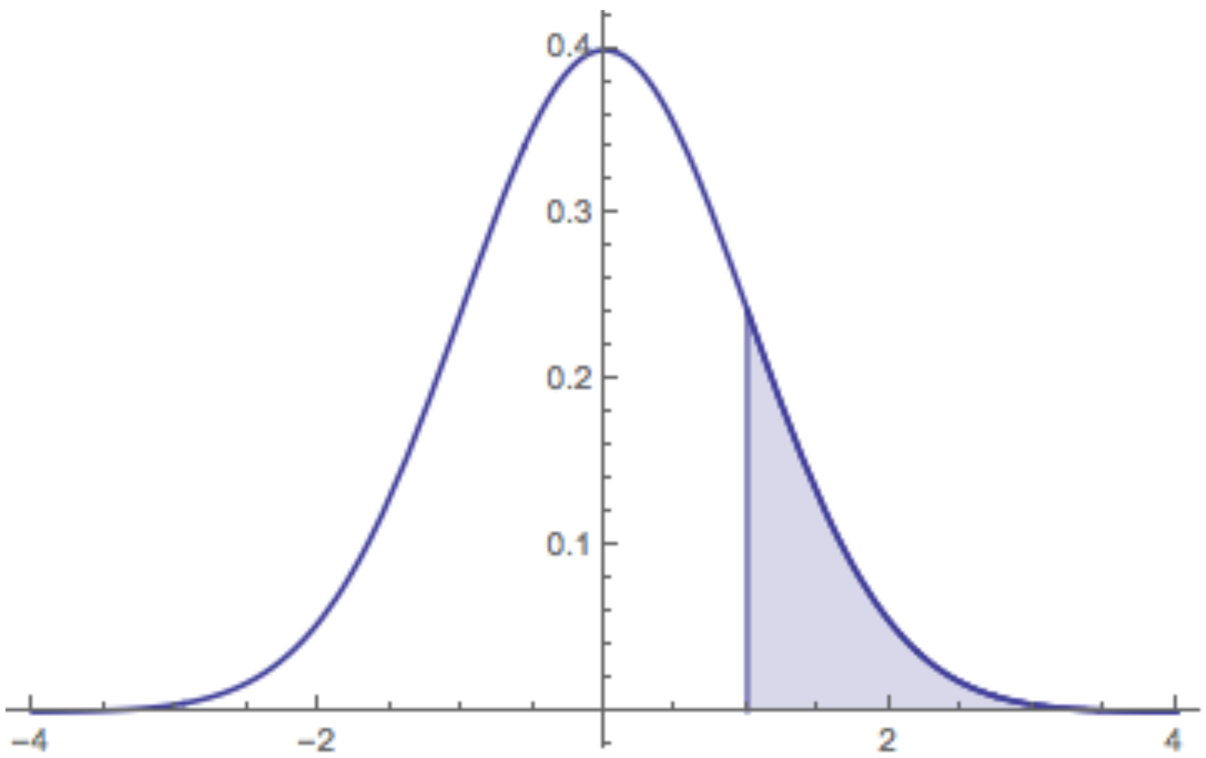


On modélise ça de façon à peu près réaliste via ce qu'on appelle une "distribution normale"



$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

La proportion d’acceptables est



$$\int_{1}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \approx 0.158$$

soit 16% du total.

Il nous reste donc 40 398 451 candidats.

Discrimination à la conversation.

Il faut également que le candidat soit intéressant.

En appliquant le même critère
on conclut qu’il y a également 16% qui nous convient.

Il nous reste 6 382 955 candidats

Discrimination à la moralité.

Excluons également,
pour la bonne moralité de notre expérience,
les personnes déjà impliquées dans une relation sérieuse.

Nous ne conservons que 64% de la population restante
(source "Gallup (US)" 19-28 ans).

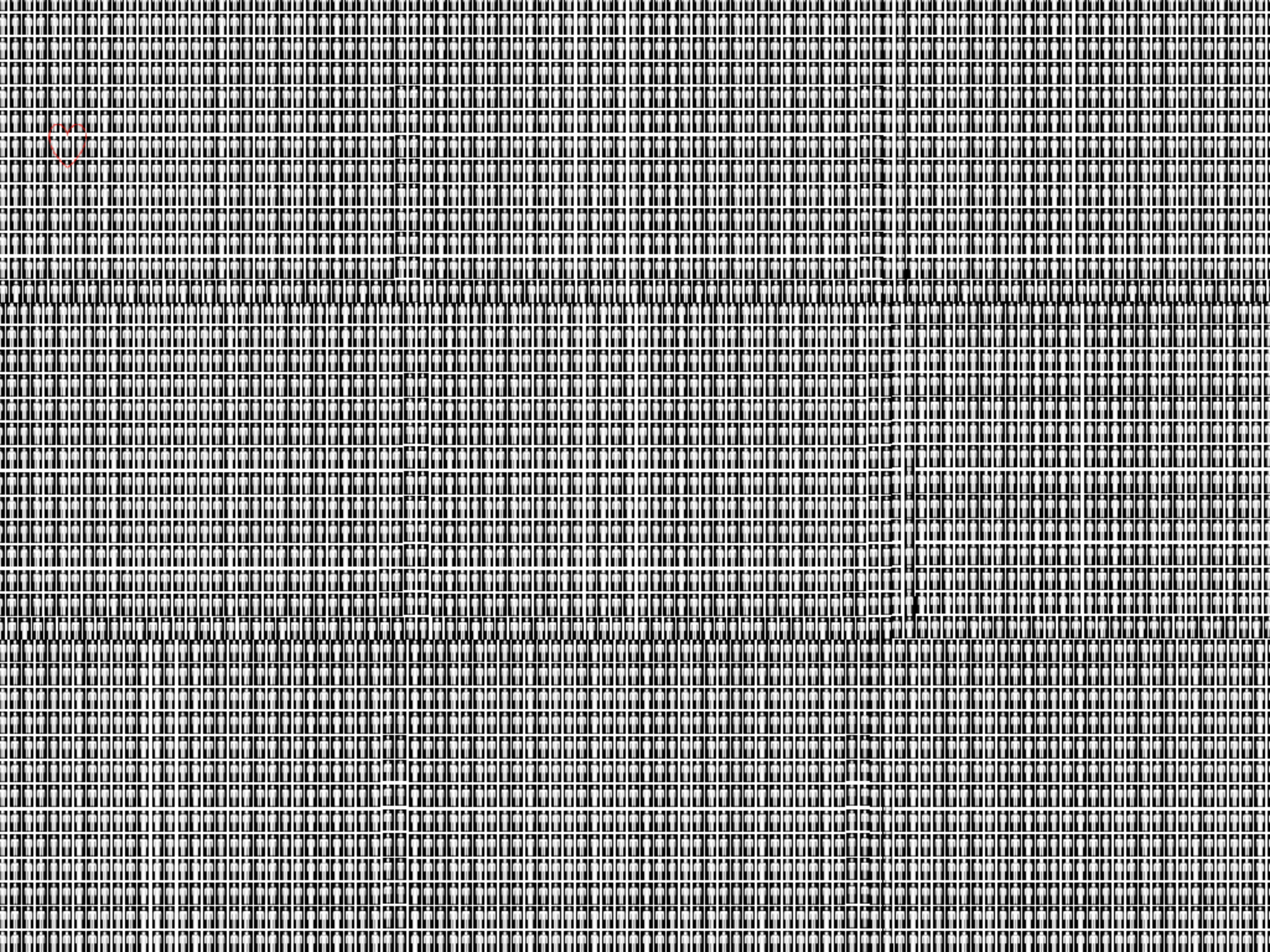
Il reste 4 085 091 candidats

Discrimination à la modernité.

Etant donné nos moeurs progressifs, il y a une exigence de réciprocité des sentiments.

En appliquant toujours le modèle normal,
on ne conserve que 2,5% des candidats

soit 101 980,2 personnes.



On a 101 980,2 personnes.

En imaginant les interviewer l’un après l’autre,

même en ne consacrant que 20 minutes par personne,

il faudrait 34 000 heures!

Donc, en y consacrant 8 heures par jour 5 jours sur 7,

cela prendra un peu plus de 16 années de tous les essayer!!

Conclusion : il est impossible de trouver l’amour.

Version 2 : magie de David Copperfield

(et de Michel Rigo)

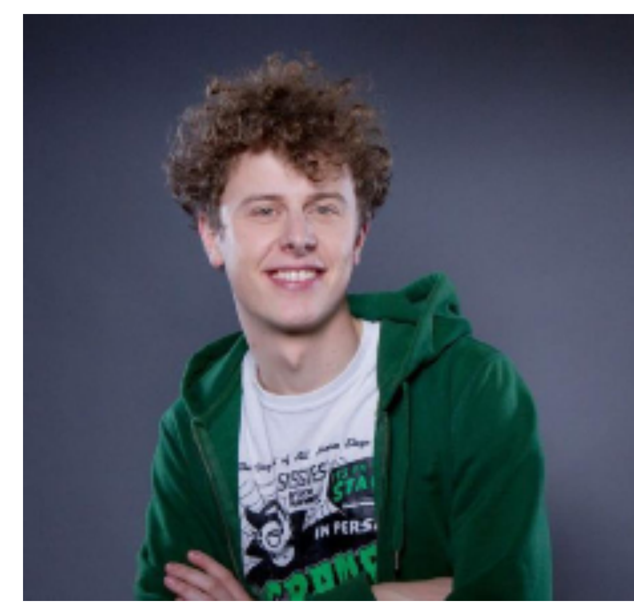
Où nous allons voir comment le destin
peut parfois intercéder pour mener vers l'élu.

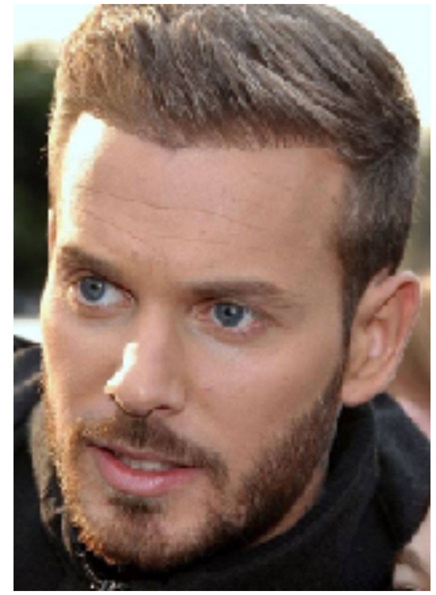
Voici 8 personnalités préférées des “jeunes”:



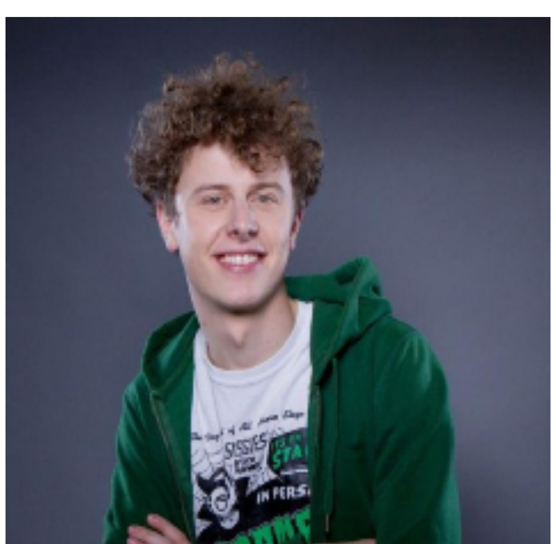
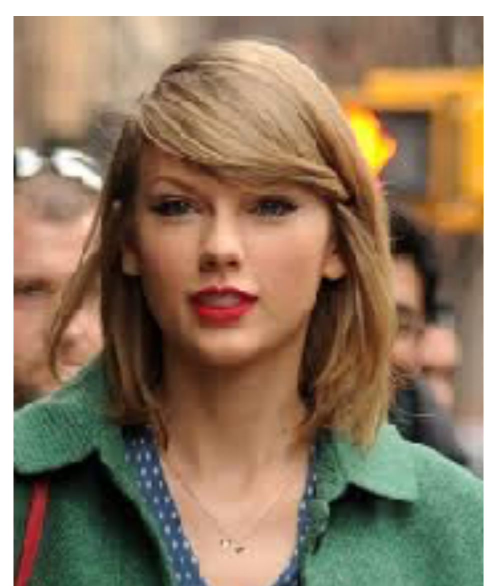
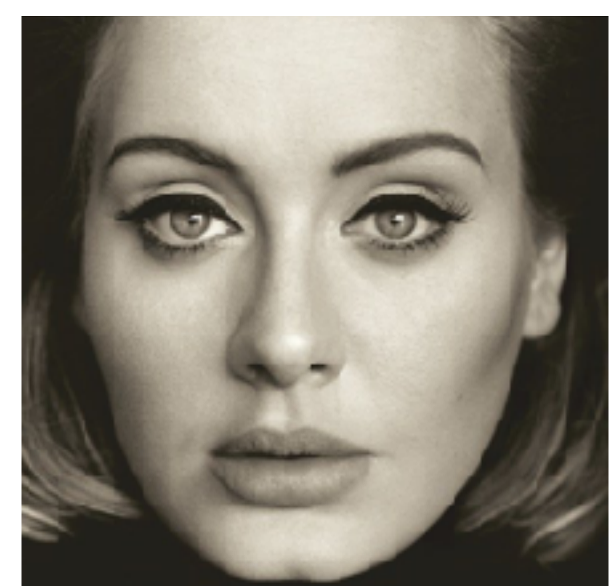























Love is blind

Version 3 : problème de secrétaire

Où nous allons voir qu'il est possible d'optimiser.

Imaginons que le destin n’ait pas porté son attention sur vous :
il vous faut donc choisir.

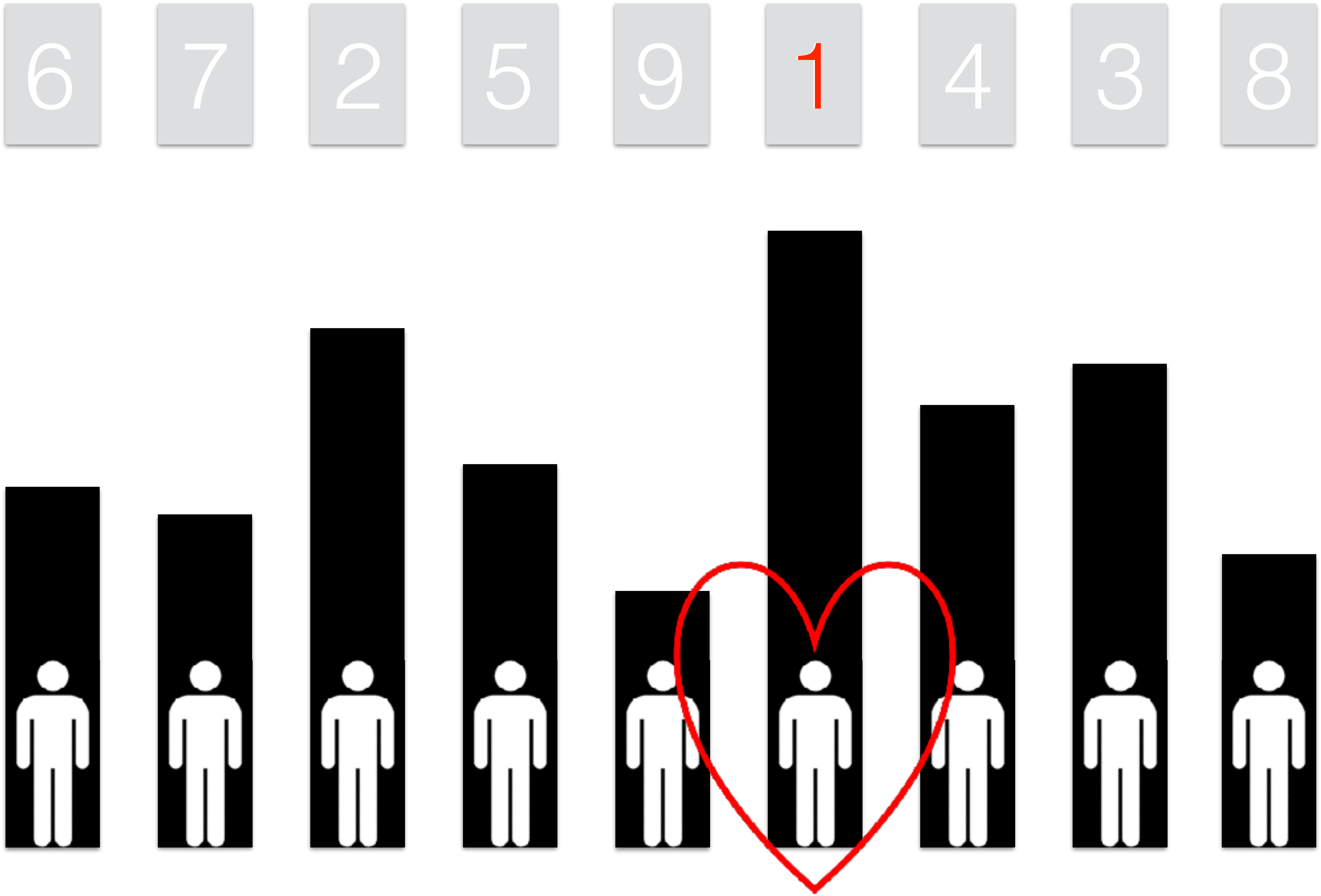
Vous allez “essayer” les candidat(e)s
l’un(e) après l’autre et n’en choisir qu’un.

On va imposer les règles suivantes :

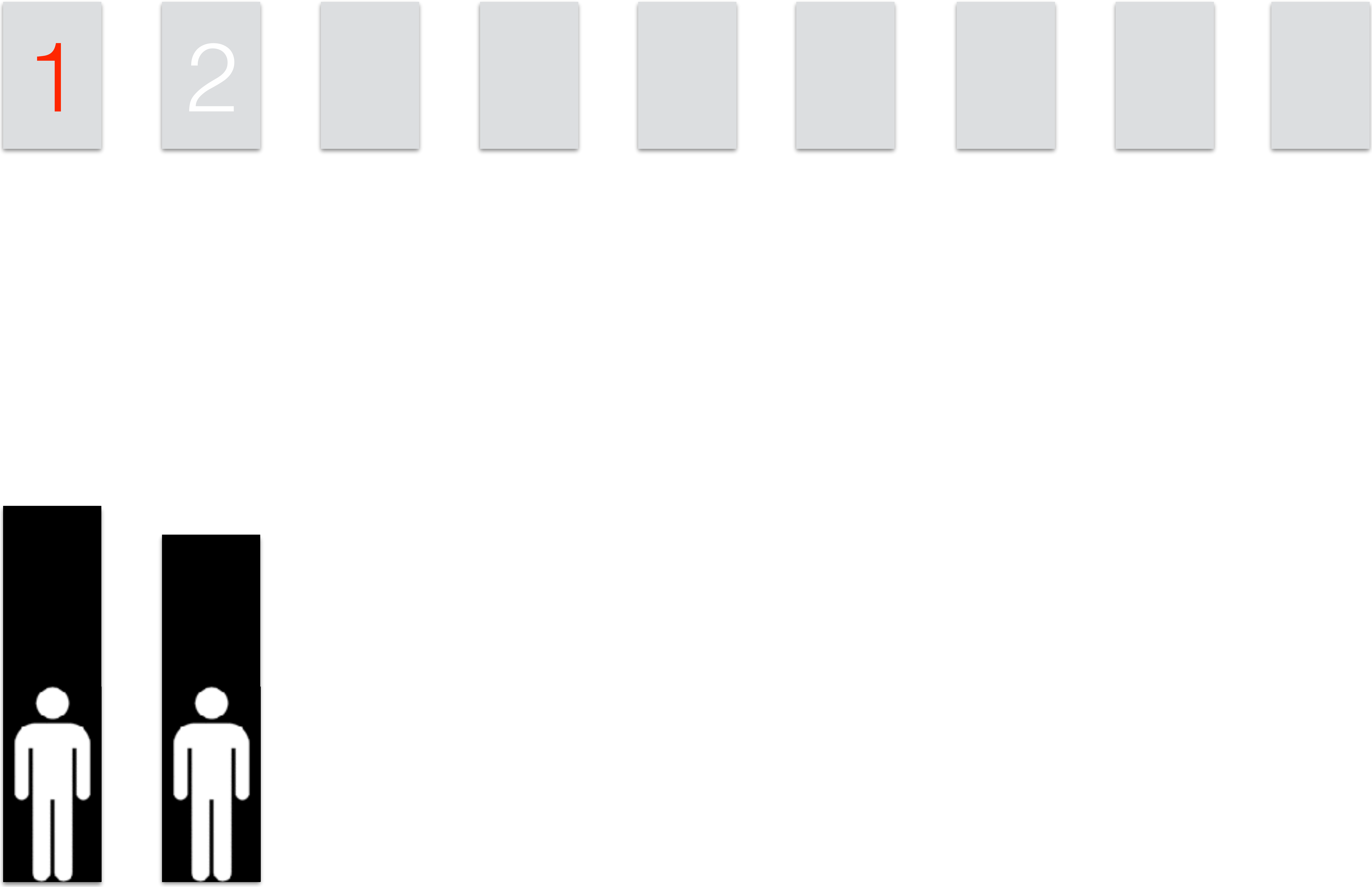
- Vous êtes en mesure de classer les candidat(e)s.
- Pas de rappel possible.
- Pas de changement d'avis non plus.

Question :

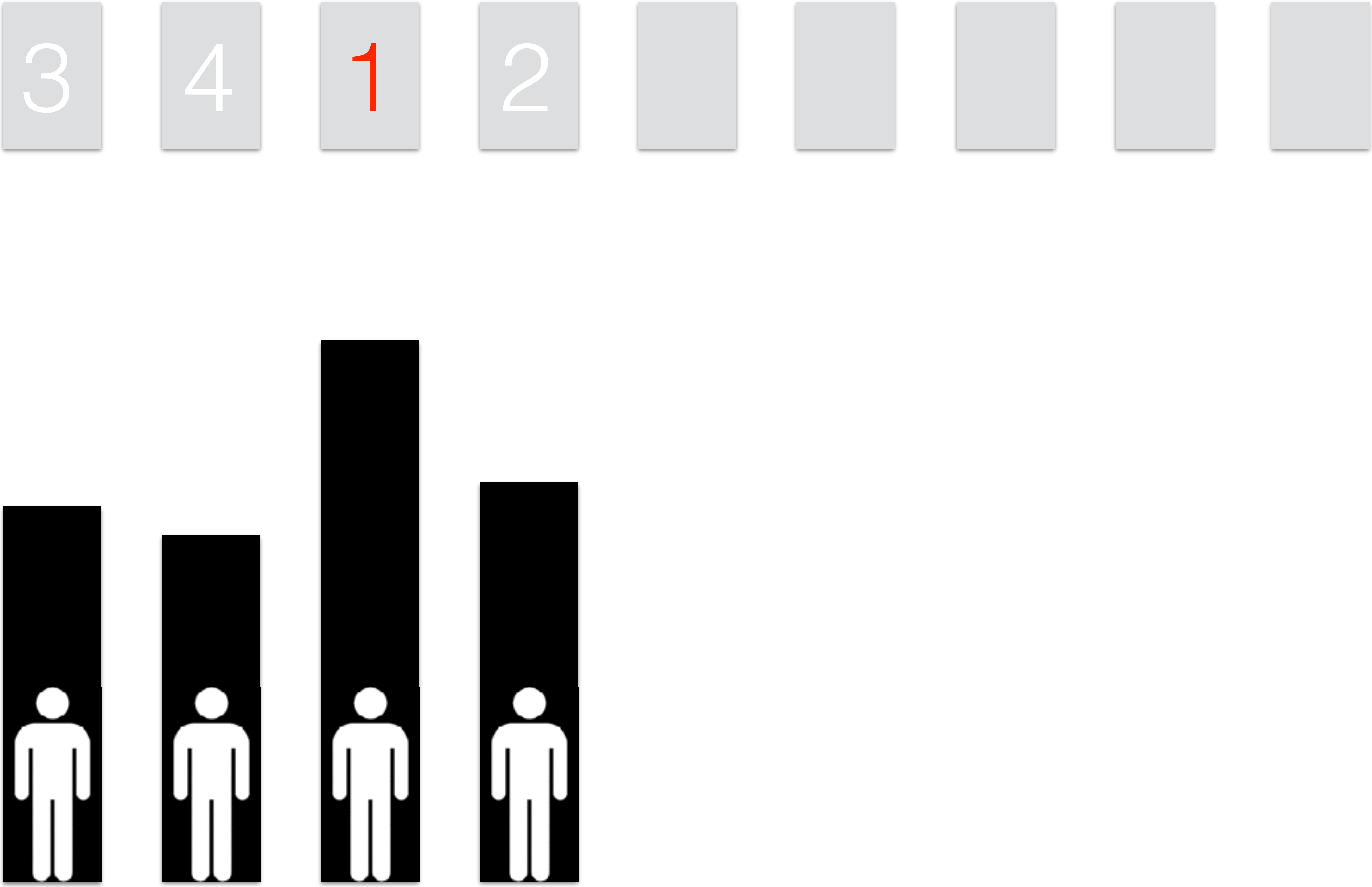
Si l'objectif est d'obtenir le/la meilleur(e) candidat(e), est-il possible de travailler de façon “intelligente”?

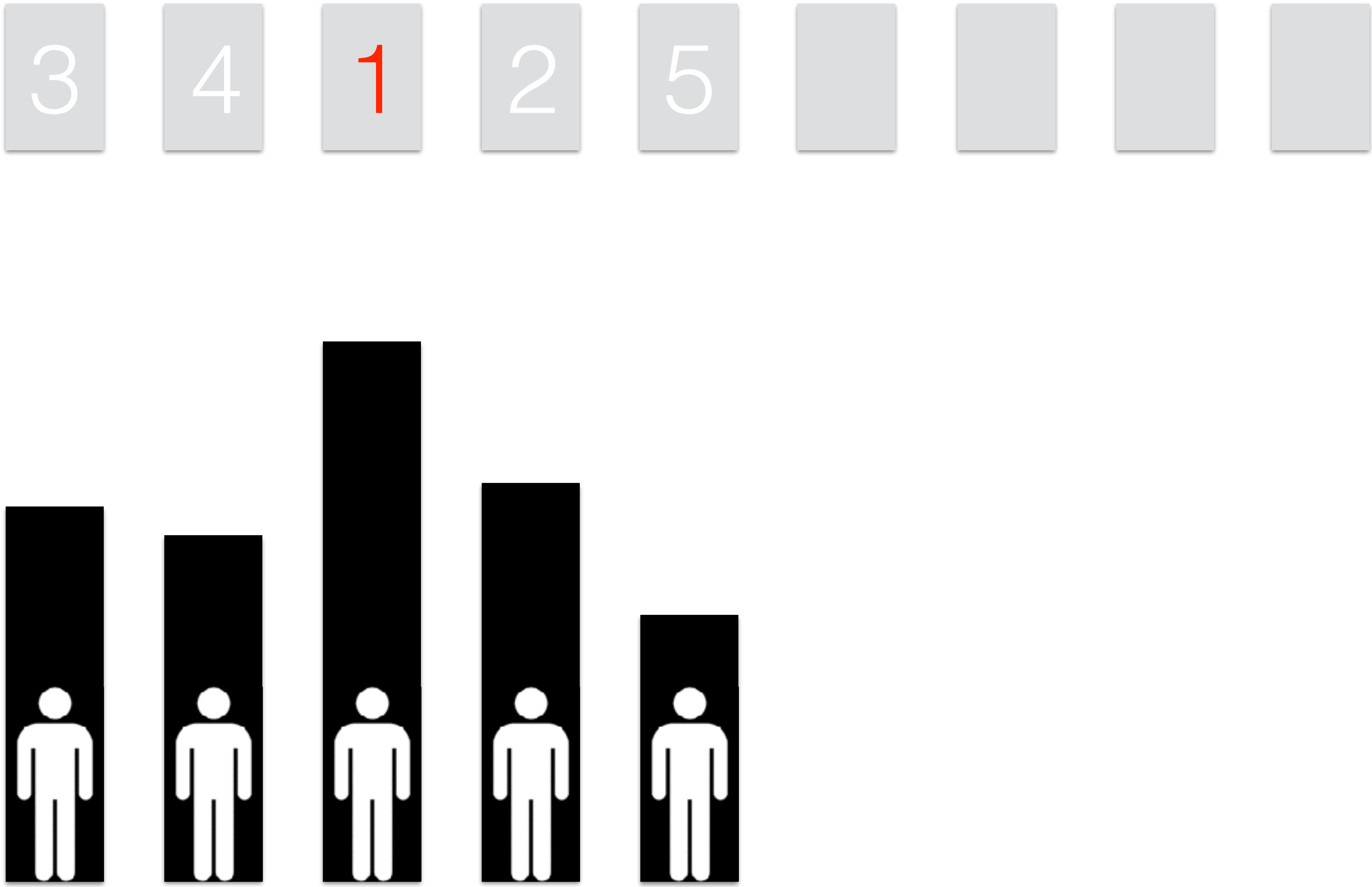

6
7
2
5
9
1
4
3
8

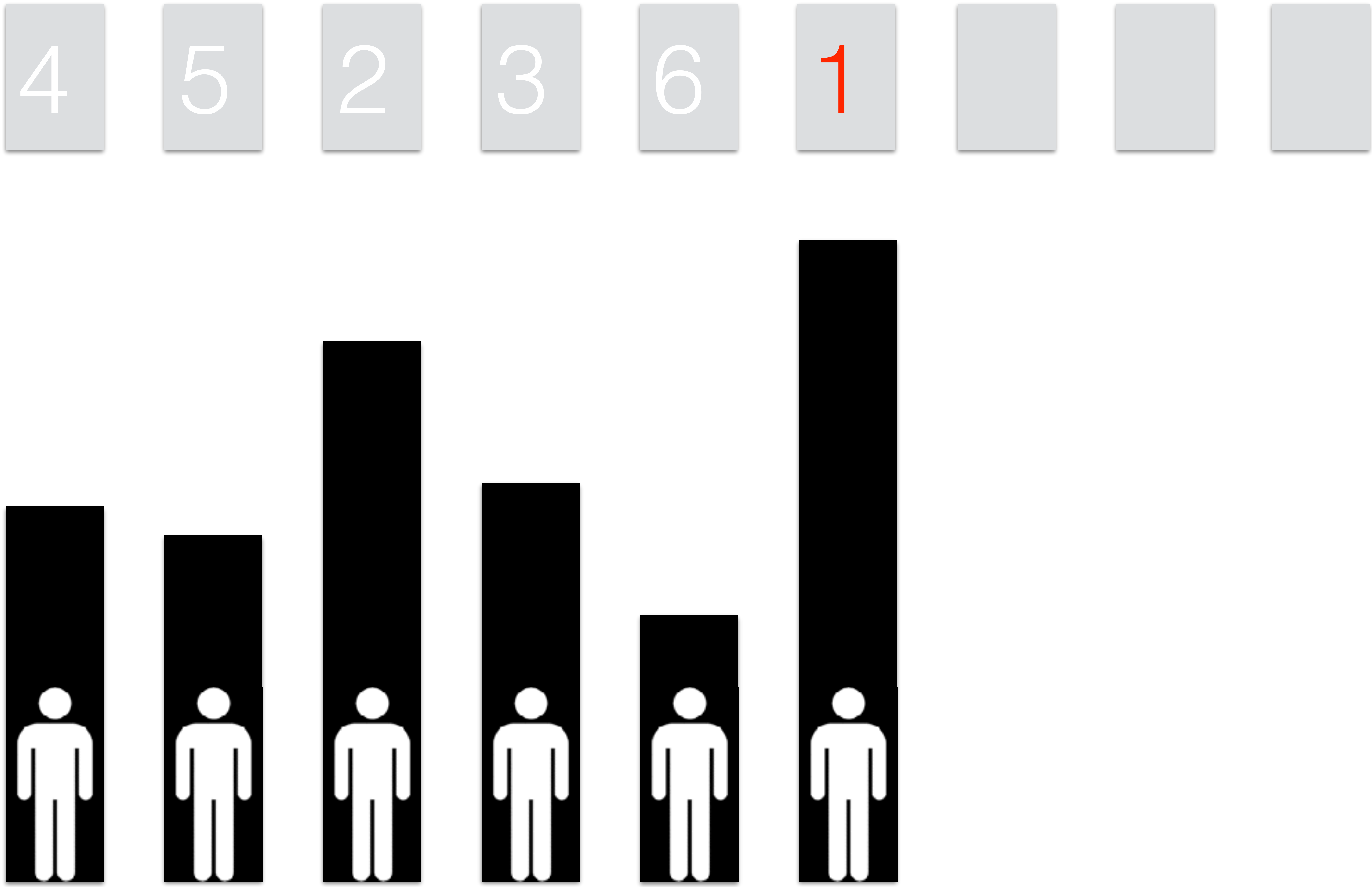

1

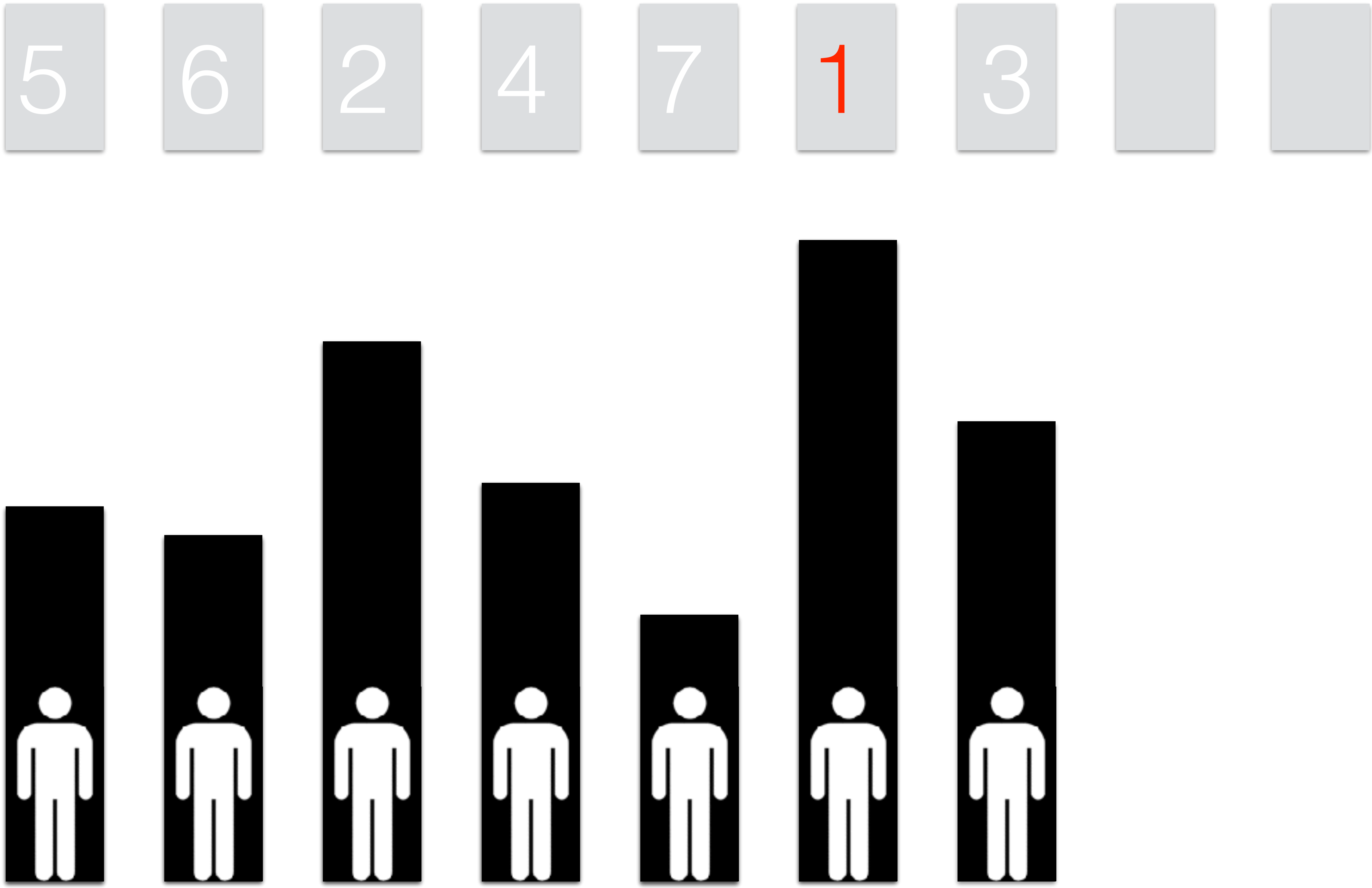

1
2

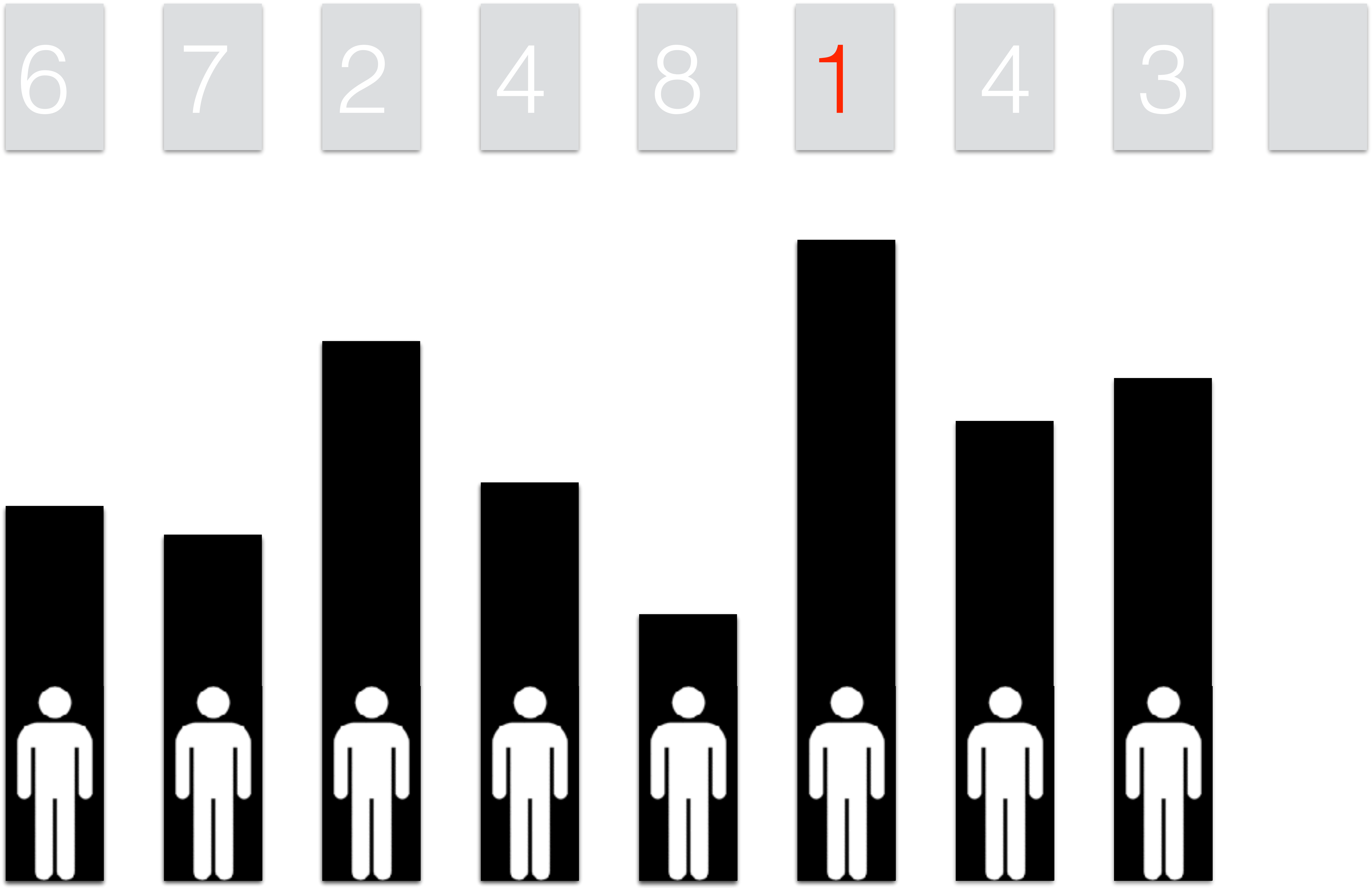

2
3
1

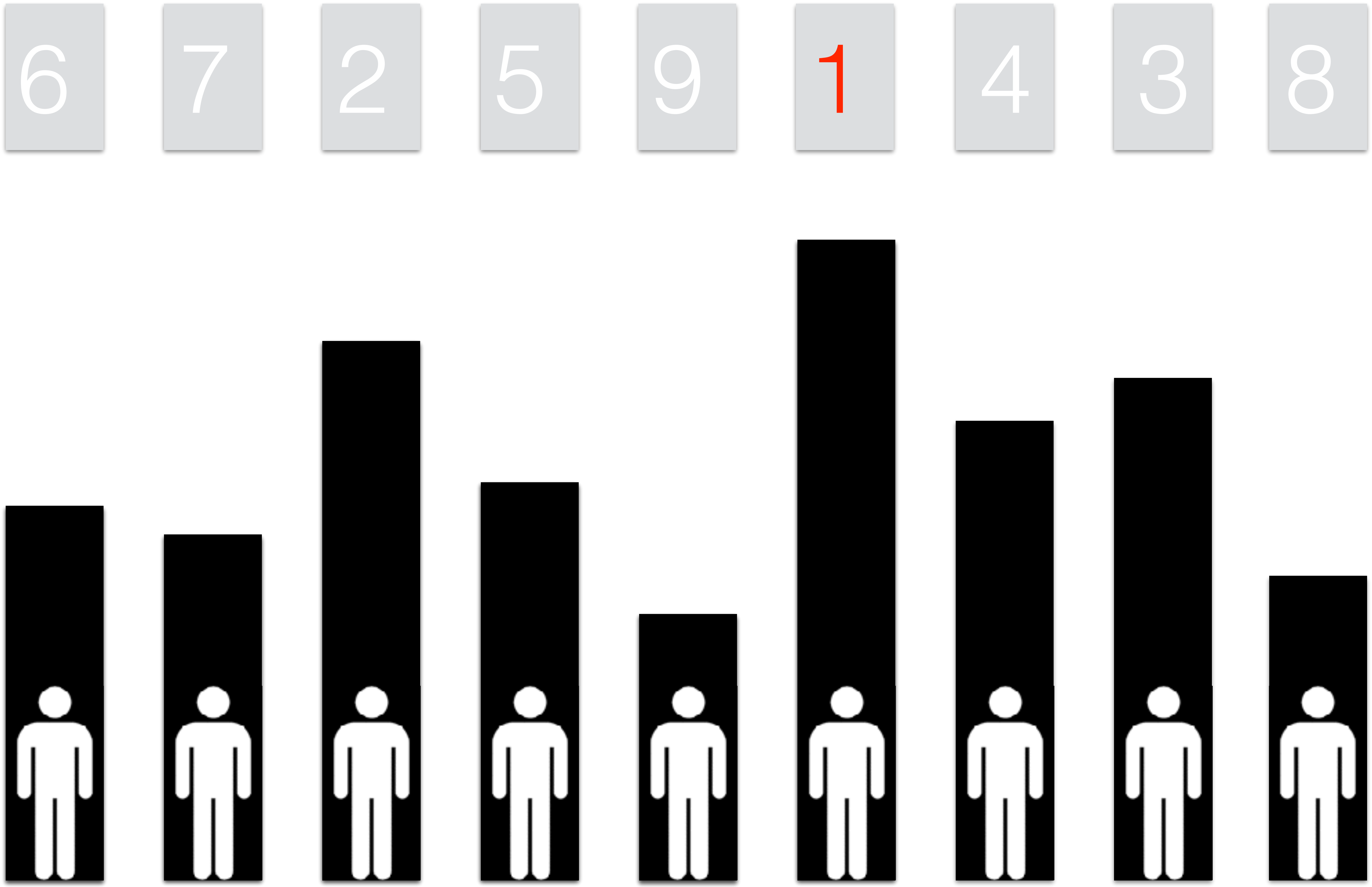

3
4
1
2


3
4
1
2
5


4
5
2
3
6
1


5
6
2
4
7
1
3


6
7
2
4
8
1
4
3


6
7
2
5
9
1
4
3
8

Question :

Existe-t-il une stratégie qui permette d’optimiser la sélection?

Stratégie 1 : prendre le premier venu

On décide arbitrairement de choisir le premier venu.
Pourrait-ce être une bonne stratégie?

Pour le savoir il faut calculer la probabilité de gagner…

Répétons l'expérience 10 fois,
en retenant à chaque fois le premier

7	1	5	8	6	4	9	2	3
8	3	1	4	2	9	7	6	5
1	5	9	2	7	3	6	8	4
2	3	9	7	6	8	5	4	1
5	2	3	9	7	8	4	6	1
1	3	8	9	6	5	7	4	2
3	4	5	2	9	6	8	7	1
9	1	6	3	5	7	2	4	8
8	6	7	1	4	9	5	2	3
7	9	2	5	1	4	6	3	8

Répétons l'expérience 10 fois,
en retenant à chaque fois le premier

7	1	5	8	6	4	9	2	3
8	3	1	4	2	9	7	6	5
1	5	9	2	7	3	6	8	4
2	3	9	7	6	8	5	4	1
5	2	3	9	7	8	4	6	1
1	3	8	9	6	5	7	4	2
3	4	5	2	9	6	8	7	1
9	1	6	3	5	7	2	4	8
8	6	7	1	4	9	5	2	3
7	9	2	5	1	4	6	3	8

20%?

En fait, si on a 9 candidats, la probabilité que le 1er soit le meilleur est

$$\frac{1}{9} = 0,111111\ldots \approx 11\%$$

Il n'y a pas non plus de différence entre choisir le premier, le deuxième ou même le neuvième.

Avec la stratégie 1 on a donc

$$P = \frac{1}{9} \approx 11\%$$

Si on a 100 candidats, la probabilité de gagner est

$$P = \frac{1}{100} = 1\%$$

Si on a 1000 candidats, la probabilité de gagner est

$$P = \frac{1}{1000} = 0{,}1\%$$

Nous, on a 101 980,2 candidats, et donc une probabilité de

$$P = \frac{1}{101\,980{,}2} \approx 0{,}00001 = 0{,}01\%$$

c’est-à-dire à peu près aucune chance…

Toutes ces stratégies triviales vont mener à une probabilité quasi nulle de trouver le meilleur.

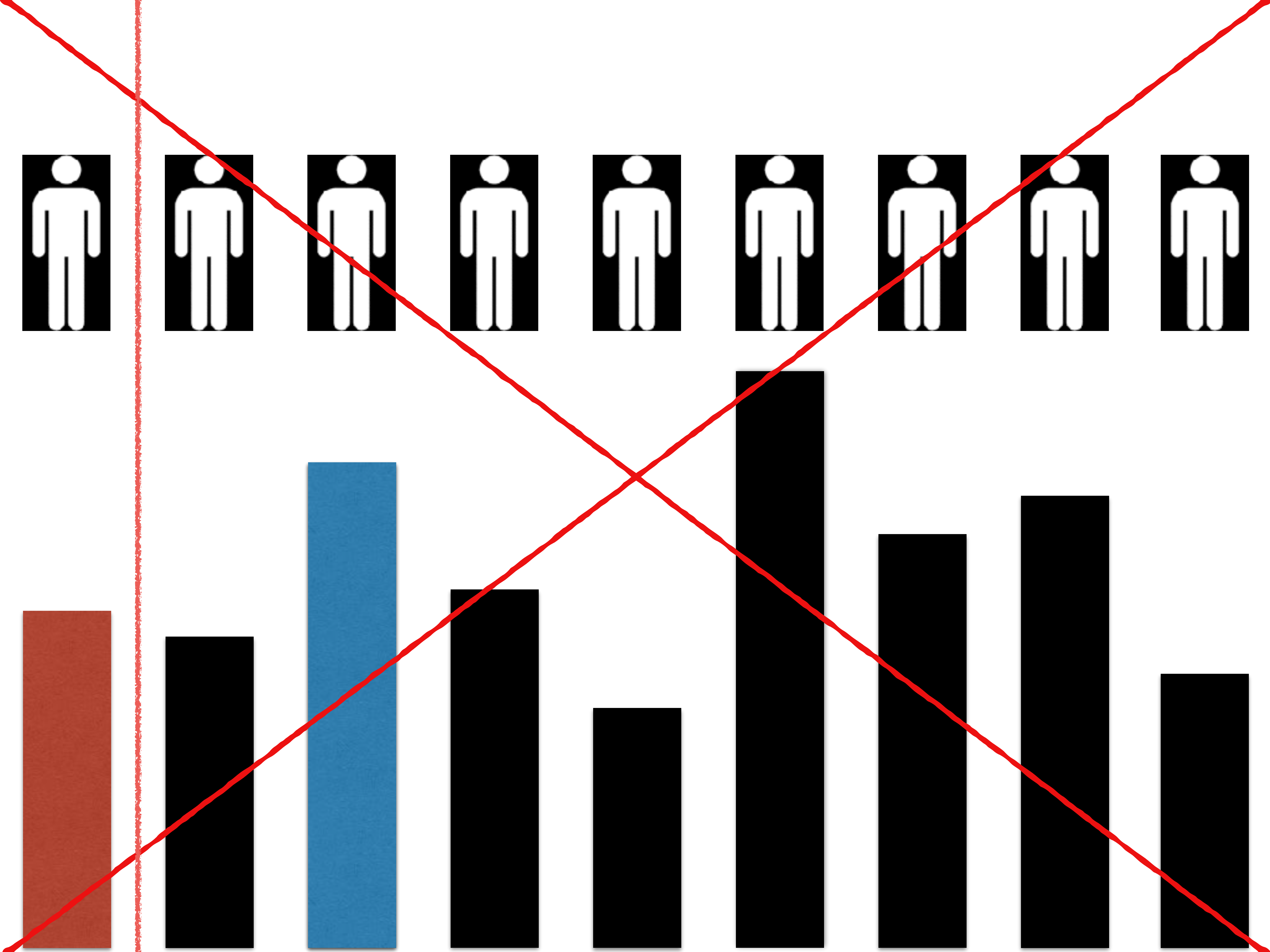
"There are no simple answers to complex problems."
(V.M. Manfredi)

Stratégie 2 : commencer par explorer le marché.

On observe d’abord, en rejetant un certain nombre de candidats puis on prend le meilleur qui suit.

Quelle est la probabilité de gagner ?

Commençons avec le cas simple : on jette le premier.



Quelle est la probabilité de gagner avec cette stratégie?

Répétons l’expérience 10 fois,
en appliquant chaque fois cette nouvelle stratégie

7	1	5	8	6	4	9	2	3
8	3	1	4	2	9	7	6	5
1	5	9	2	7	3	6	8	4
2	3	9	7	6	8	5	4	1
5	2	3	9	7	8	4	6	1
1	3	8	9	6	5	7	4	2
3	4	5	2	9	6	8	7	1
9	1	6	3	5	7	2	4	8
8	6	7	1	4	9	5	2	3
7	9	2	5	1	4	6	3	8

Répétons l'expérience 10 fois,
en appliquant chaque fois cette nouvelle stratégie

7	1	5	8	6	4	9	2	3
8	3	1	4	2	9	7	6	5
1	5	9	2	7	3	6	8	4
2	3	9	7	6	8	5	4	1
5	2	3	9	7	8	4	6	1
1	3	8	9	6	5	7	4	2
3	4	5	2	9	6	8	7	1
9	1	6	3	5	7	2	4	8
8	6	7	1	4	9	5	2	3
7	9	2	5	1	4	6	3	8

Proba = 30%?

La probabilité de gagner estimée semble proche de 30%.

Serait-ce imaginable qu'avec ce simple changement de stratégie on triple nos chances de gagner?

Il nous faut calculer la probabilité de gagner.

Nous allons décomposer le calcul
en fonction de la position du meilleur candidat.

Si le meilleur est en première position

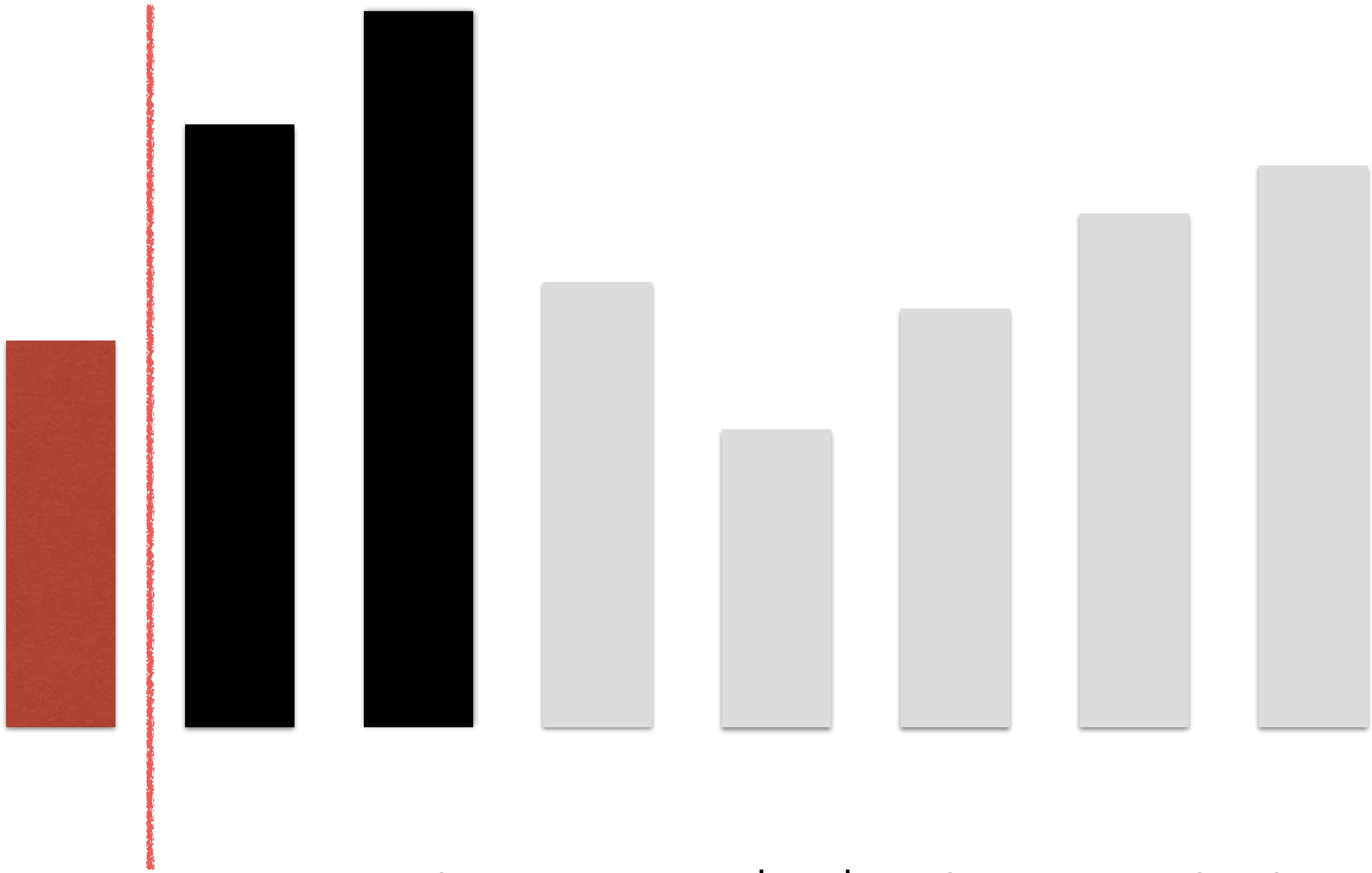
on n’a aucune chance de gagner!

Si le meilleur est en deuxième position

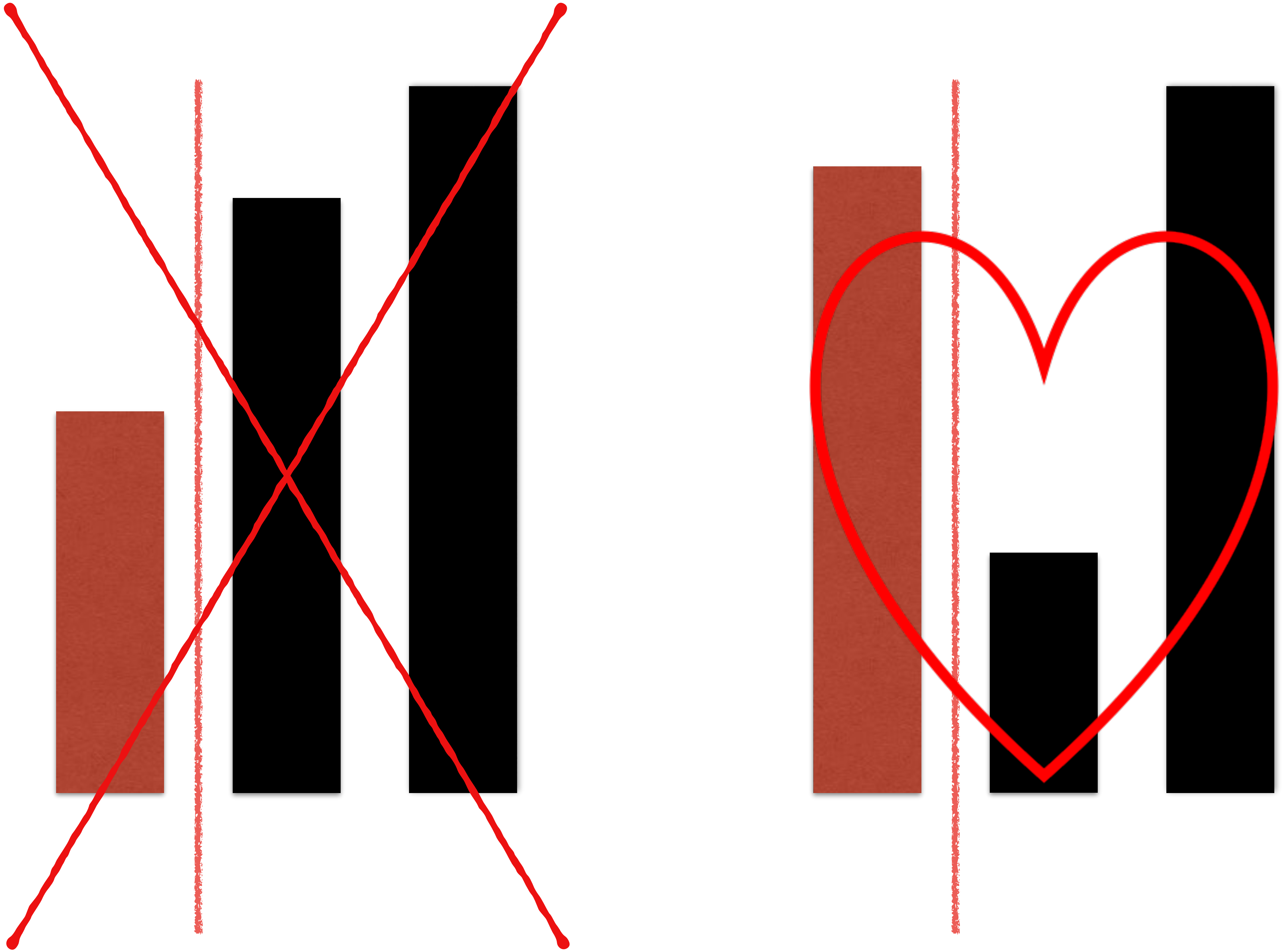
on est certain de gagner!

Si le meilleur est en troisième position

Si le meilleur est en troisième position

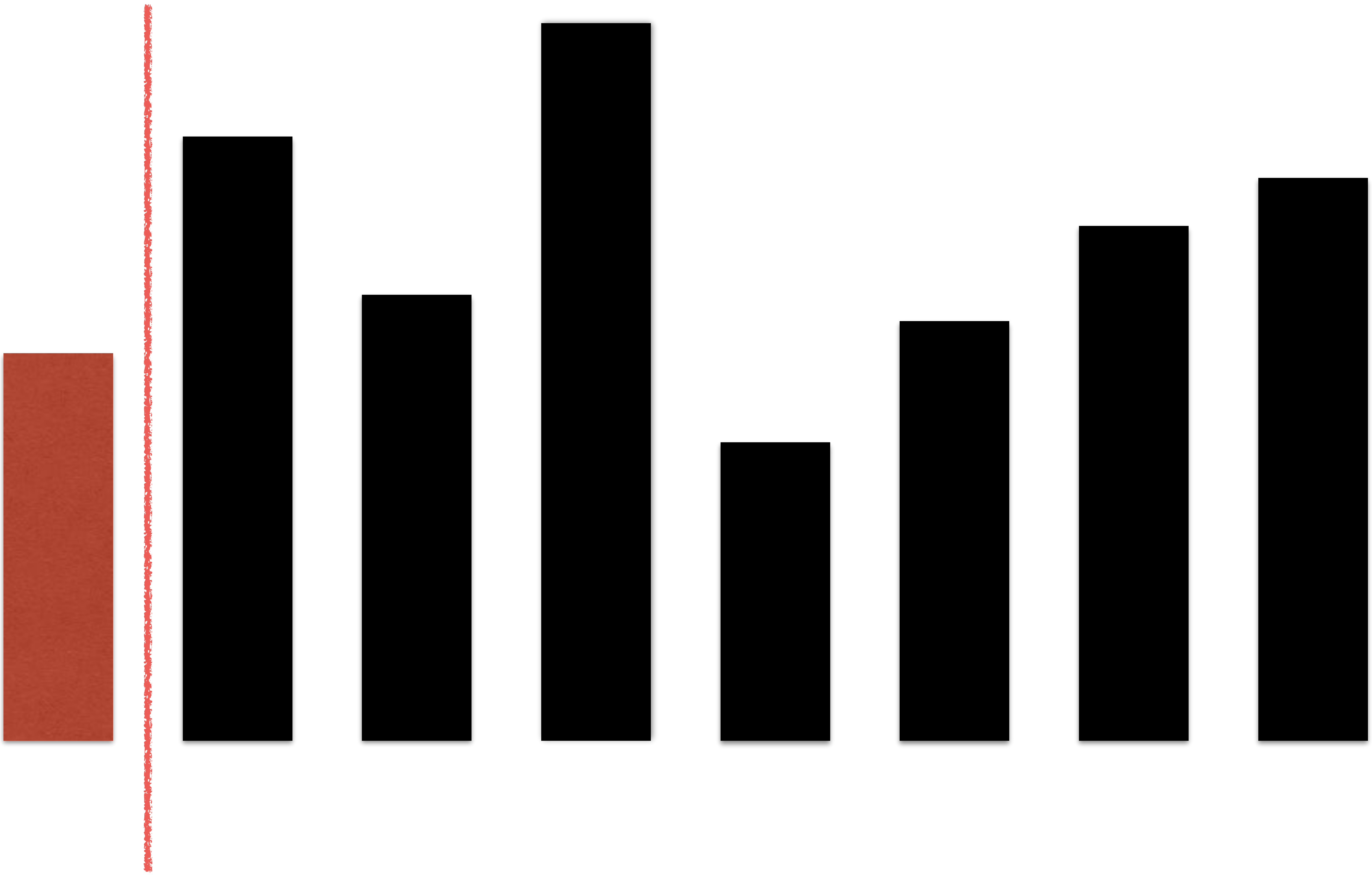


tous ceux qui suivent ne comptent pas…

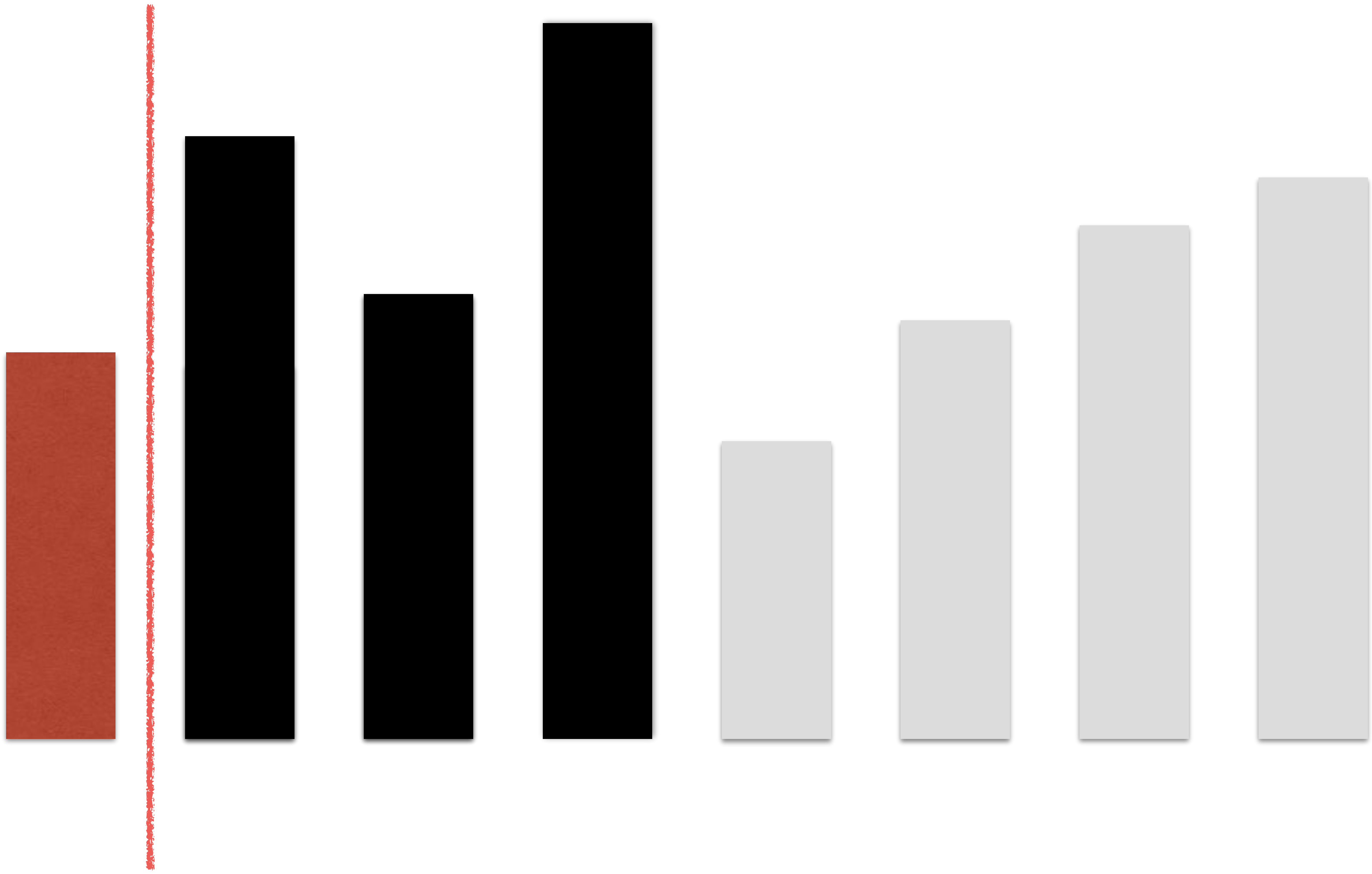


on gagne avec une chance sur deux !

Si le meilleur est en quatrième position



Si le meilleur est en quatrième position



on gagne avec une chance sur trois !

Si le meilleur est en cinquième position

on gagne avec une chance sur quatre !

Si le meilleur est en sixième position

on gagne avec une chance sur cinq !

etc.

Position du Meilleur	**Chance de gagner**
1	0
2	1/1
3	1/2
4	1/3
5	1/4
6	1/5
7	1/6
8	1/7
9	1/8

$$P_9 = \frac{1}{9}\times 0 + \frac{1}{9}\times 1 + \frac{1}{9}\times\frac{1}{2} + \ldots + \frac{1}{9}\times\frac{1}{8}$$

$$= \frac{1}{9}\left(1 + \frac{1}{2} + \ldots + \frac{1}{8}\right)$$

$$\approx 0.301$$

Aussi, si on a 9 candidats, la stratégie 2 est bien meilleure que la stratégie 1!!

Donc $$P_9 = \frac{1}{9}\times 0 + \frac{1}{9}\times 1 + \frac{1}{9}\times\frac{1}{2} + ... + \frac{1}{9}\times\frac{1}{8}$$

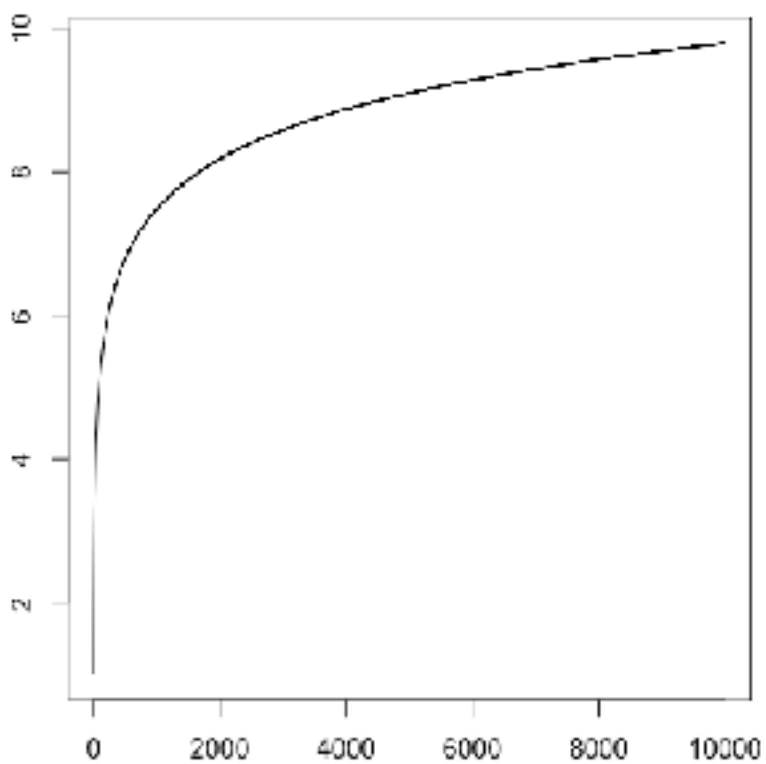
Et si on a plus de candidats?

Par le même type d'argument on calcule :

$$P_n = \frac{1}{n}\times 0 + \frac{1}{n}\times 1 + \frac{1}{n}\times\frac{1}{2} + \frac{1}{n}\times\frac{1}{3} + ... + \frac{1}{n}\times\frac{1}{n-1}$$
$$= \frac{1}{n}\left(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n-1}\right)$$

Ce n'est pas très parlant…
à quoi ressemble cette fonction?

La fonction $1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n-1}$ ressemble à $\ln(n)$



n	$1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n-1}$
10	2,928968
100	5,187378
1000	7,485471
10000	9,787606
100000	12,09015

Conclusion : $P_n = \frac{1}{n}\left(1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n-1}\right) \approx \frac{\ln(n)}{n}$

On calcule

n	100	1000	10000	1000000	1000000000
P_n	0,051	0,007	0.0009	0.0001	0.00001

Conclusion :

Cette stratégie est encore toute pourrie
et on ne va pas trouver l'amour avec ça non plus!

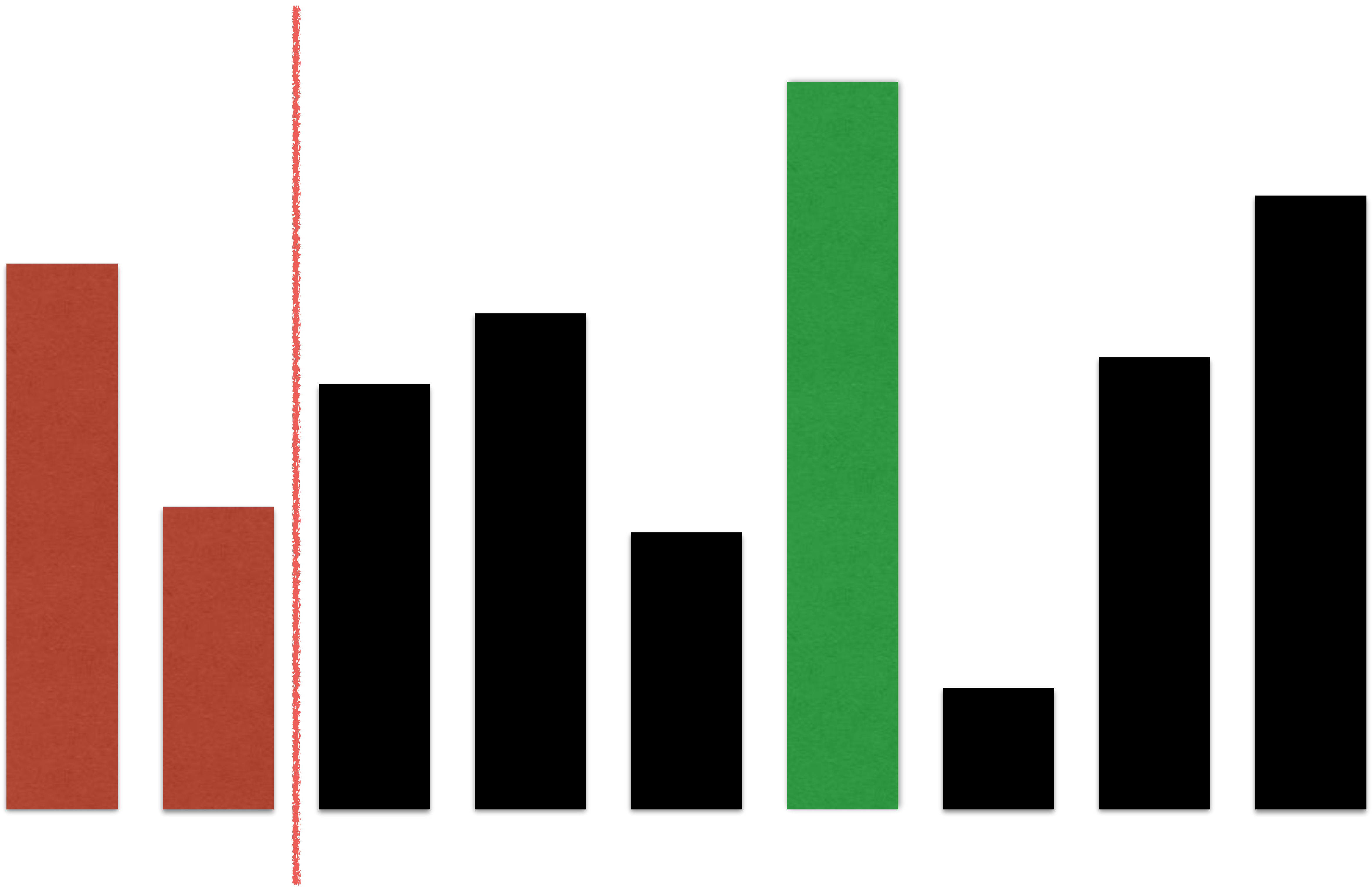
La stratégie précédente était bonne
pour les “petites” valeurs de n,

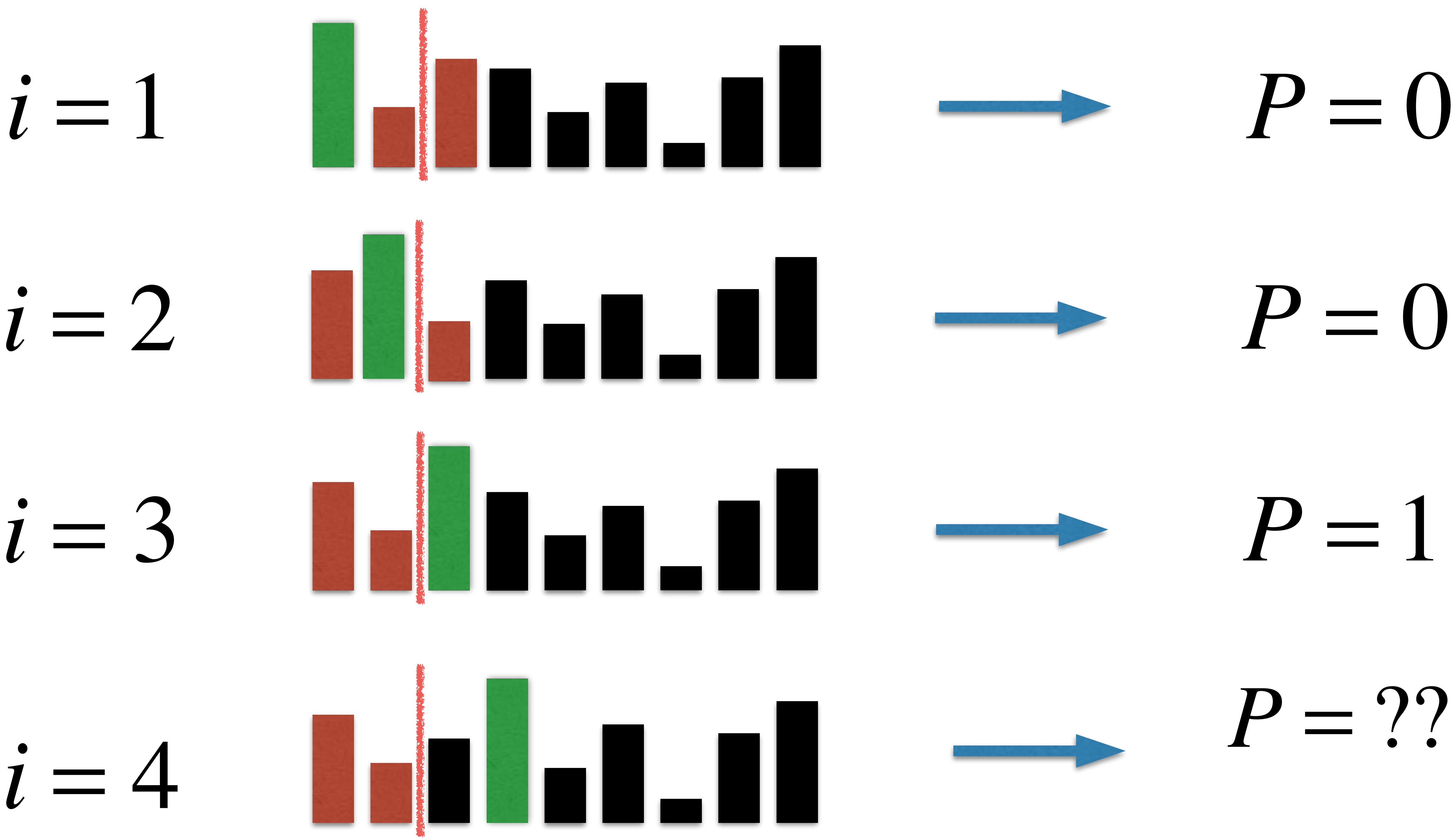
mais toute pourrie pour les “grandes” valeurs de n.

On devrait en fait **rejeter une proportion** du total.

Stratégie : rejeter les r premiers avec r = x * n
et prendre le meilleur qui vient après.

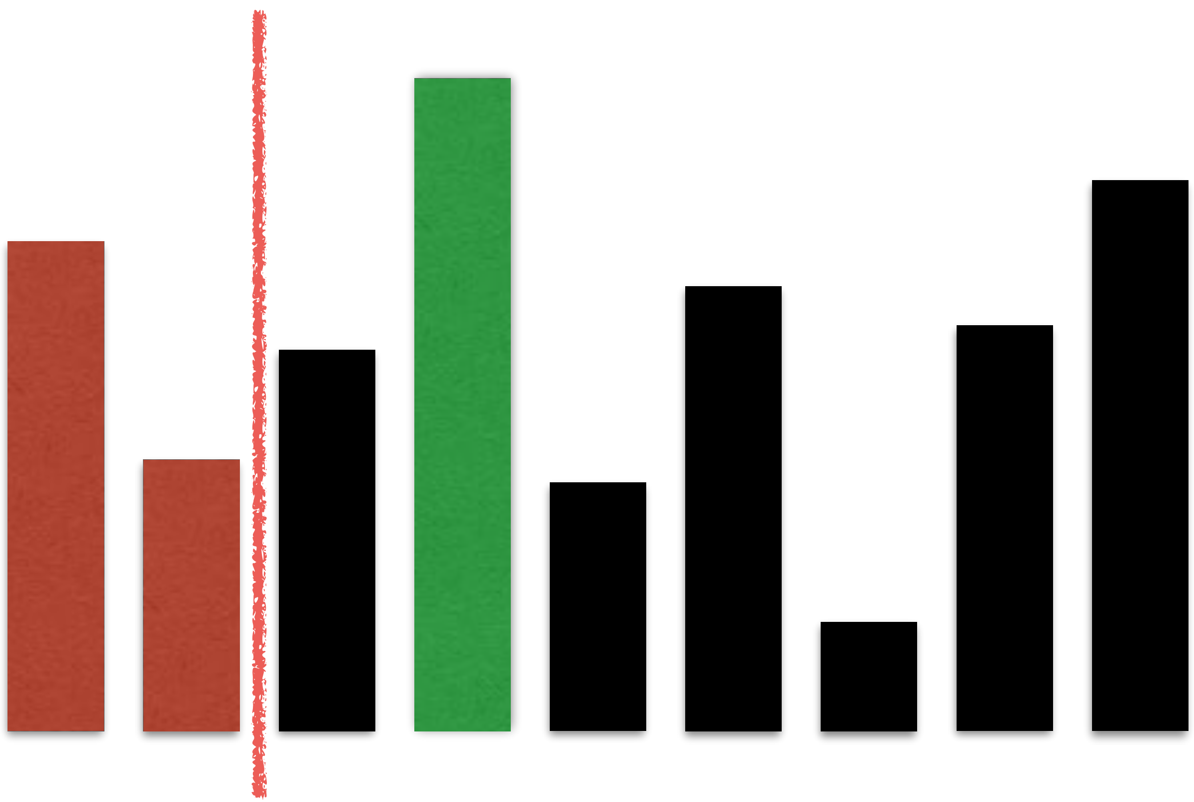
r = 2



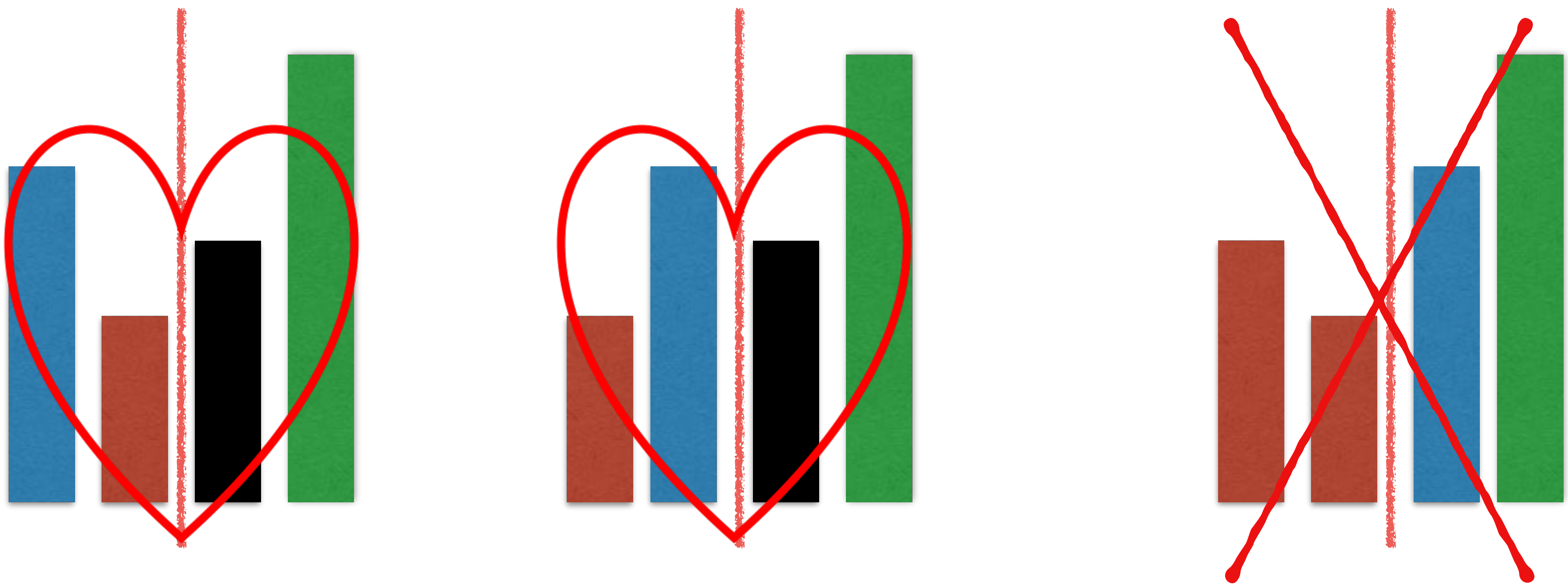


$i = 1$
$P = 0$
$i = 2$
$P = 0$
$i = 3$
$P = 1$
$i = 4$
$P = ??$

r=2, n=9, i=4

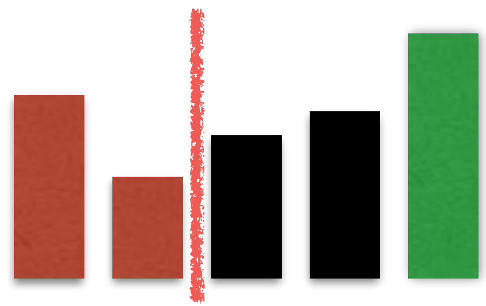


r=2, n=9, i=4



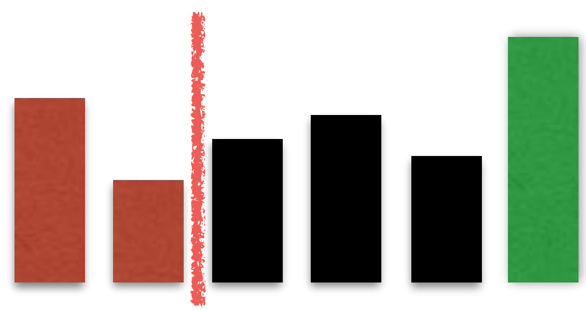
$$i = 4 \quad \longrightarrow \quad P = {}^{2}\!/_{3}$$

$i = 5$



$\longrightarrow$ $P = {}^{2}\!/_{4}$

$i = 6$



$\longrightarrow$ $P = {}^{2}\!/_{5}$

Etc.

r=2, n=9

Position du Meilleur	Chance de gagner
1	0
2	0
3	1
4	2/3
5	2/4
6	2/5
7	2/6
8	2/7
9	2/8

$$P_9(2) = \frac{2}{9}\left(\frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{8}\right)$$

r=3, n=9

Position du Meilleur	Chance de gagner
1	0
2	0
3	0
4	1
5	3/4
6	3/5
7	3/6
8	3/7
9	3/8

$$P_9(3) = \frac{3}{9}\left(\frac{1}{3} + \ldots + \frac{1}{8}\right)$$

De façon générale, la probabilité de gagner est

$$P_n(r) = \sum_{i=1}^{n} P(\text{gagner} \mid \text{meilleur} = \text{ième})P(\text{meilleur} = \text{ième})$$

Si i= 1, …, r :

$$P(\text{gagner} \mid \text{meilleur} = \text{ième}) = 0$$

Si i= r+1, …, n :

$$P(\text{gagner} \mid \text{meilleur} = \text{ième}) = \frac{r \times (i-2)!}{(i-1)!} = \frac{r}{i-1}$$

Conclusion :

$$P_n(r) = \sum_{i=r+1}^{n} \frac{r}{i-1}\frac{1}{n} = \frac{r}{n}\sum_{i=r}^{n-1}\frac{1}{i}$$

Théorème (Lindley, 1961)
Pour chaque n il existe un indice optimal r pour lequel cette probabilité est la meilleure possible.

Reste à calculer le “r optimal” ainsi que d’estimer nos chances de trouver l’amour avec ça.

Pour ça on va utiliser l’approximation

$$\sum_{i=r}^{n-1} \frac{1}{i} \approx \int_{i=r}^{n} \frac{1}{t} dt = \ln(n) - \ln(r)$$

qui est valide “lorsque n est grand”.

Avec cette approximation on a

$$P_n(r) = \frac{r}{n}\sum_{i=r}^{n-1}\frac{1}{i} \approx \frac{r}{n}(\ln(n) - \ln(r))$$

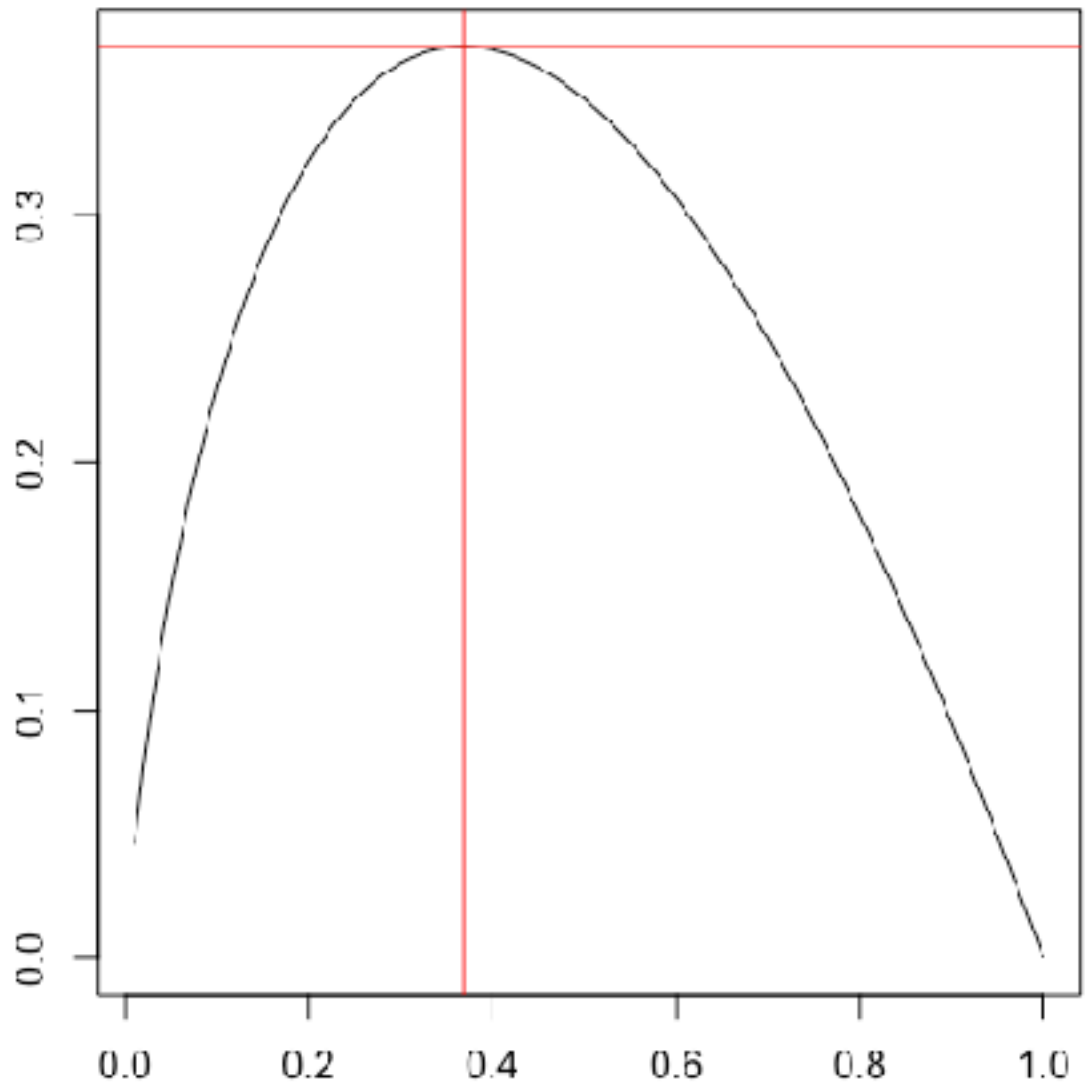
ou encore, si $r = n \times x$

$$\frac{r}{n}(\ln(n) - \ln(r)) = x(\ln(n) - \ln(n) - \ln(x))$$

donc

$$P_n(n \times x) \approx -x\ln(x)$$

La probabilité de gain pour n grand ressemble donc à :



$$f(x) = -x\ln(x)$$

On optimise :

$f'(x_0) = -\ln(x_0) - 1 = 0$ uniquement pour $x_0 = 1/e$

CONCLUSION

La stratégie optimale est :

- on rejette les 1/e premiers, ie. 36% des premiers
- on prend le premier acceptable qui suit

et avec ça nos chances de gagner sont

$$-1/e * \ln(1/e) = 1/e = 36\%$$

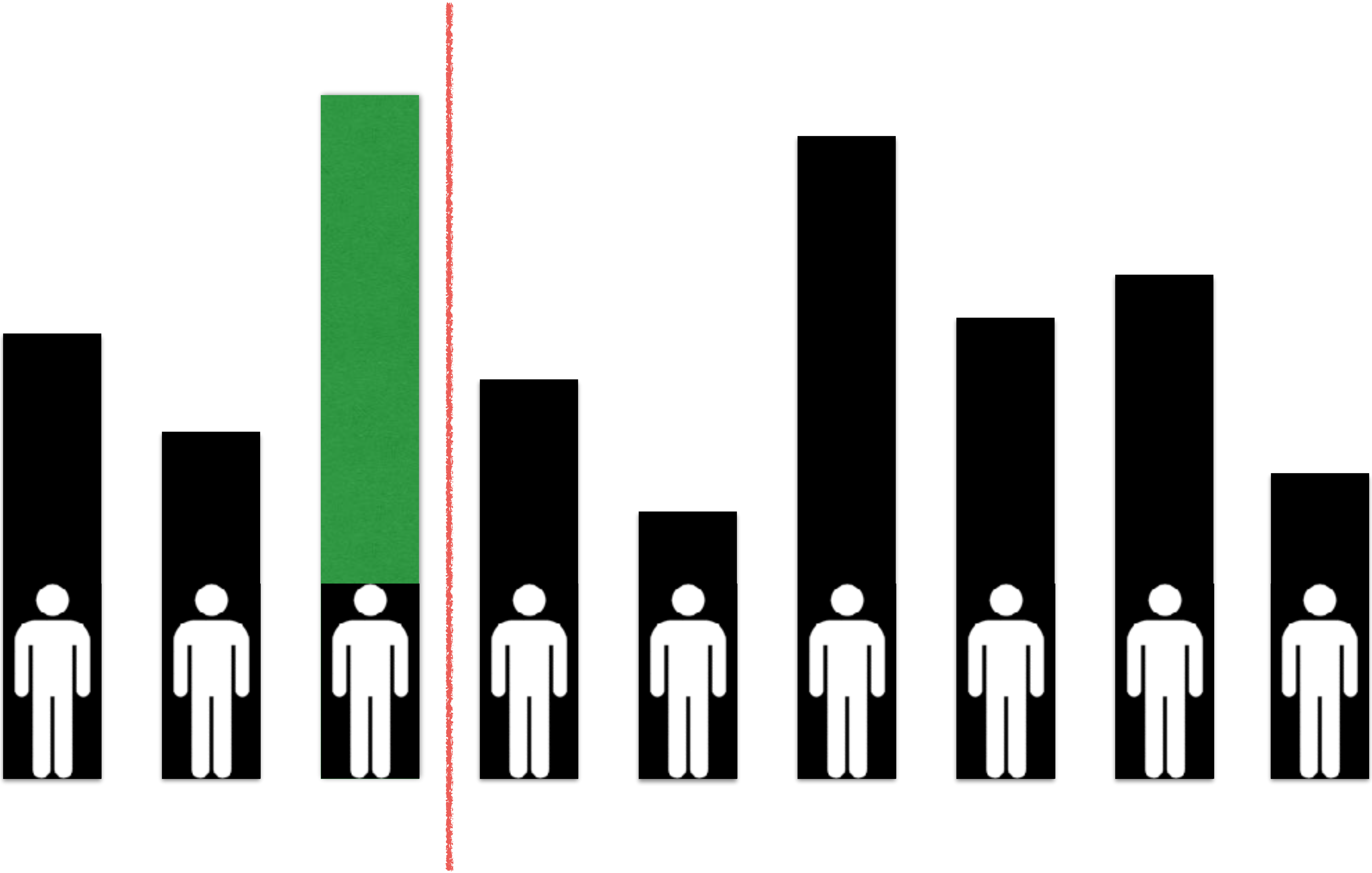
Pour 100000 on est passé de 0,00001 à 0,36 !!

Mission accomplie :

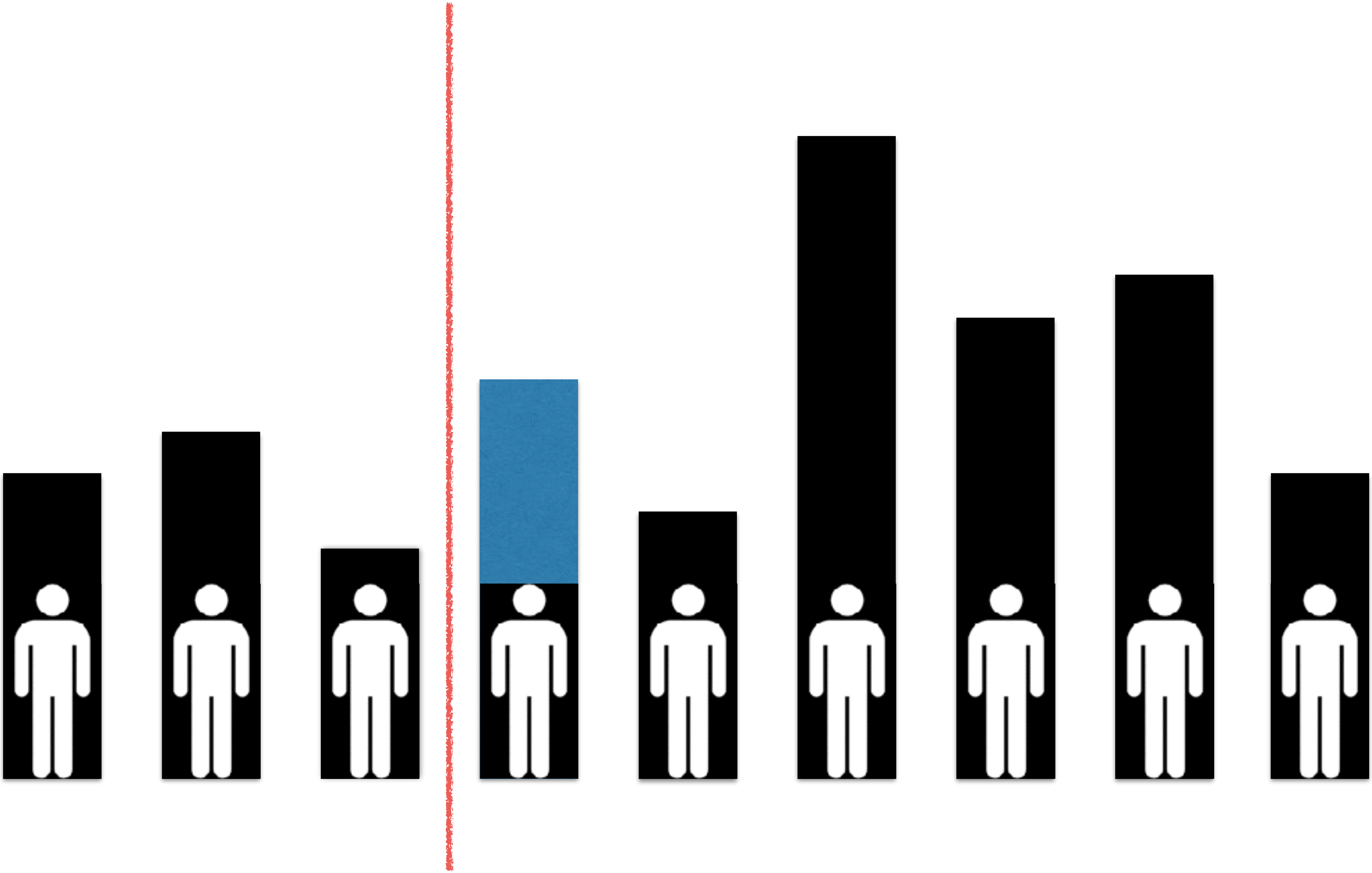
on a la stratégie optimale pour trouver l’amour.

Disclaimer :
je n’ai pas dit que vous alliez le trouver!

Il y a de grandes chances de terminer seul…



Il y a de grandes chances de terminer déçu.



La première apparition “sur papier” de ce problème date des années 1960 (dans Scientific American).

Et la première solution publiée date de 1961.

Le problème porte une collection de noms plus ou moins équivalents:

- Problème de secrétaire
- Problème de la dot
- Jeu de “Googol”
- …

Statistical Science
1989, Vol. 4, No. 3, 282–296

Who Solved the Secretary Problem?

Thomas S. Ferguson

Abstract. In Martin Gardner's *Mathematical Games* column in the February 1960 issue of *Scientific American*, there appeared a simple problem that has come to be known today as the Secretary Problem, or the Marriage Problem. It has since been taken up and developed by many eminent probabilists and statisticians and has been extended and generalized in many different directions so that now one can say that it constitutes a "field" within mathematics-probability-optimization. The object of this article is partly historical (to give a fresh view of the origins of the problem, touching upon Cayley and Kepler), partly review of the field (listing the subfields of recent interest), partly serious (to answer the question posed in the title), and partly entertainment. The contents of this paper were first given as the Allen T. Craig lecture at the University of Iowa, 1988.

Key words and phrases: Secretary problem, marriage problem, search problem, relative ranks, stopping times, minimax rules.

1. INTRODUCTION

In the late 1950's and early 1960's there appeared a simple, partly recreational, problem known as the secretary problem, or the marriage problem, or the dowry problem, that made its way around the mathematical community. The problem has a certain appeal. think you will find the journey interesting, and the conclusion surprising.

2. STATEMENT OF THE PROBLEM

The reader's first reaction to the title might well be to ask, "Which secretary problem?". After all, as I

Ce type de problème est riche en belles mathématiques

et aussi en applications pratiques :

- achat de maison,
- allocation de resources,
- traitements médicaux,
- choix de parking,
- finance,
- …

De nombreuses variantes sont possibles :

- “avec” ou “sans” information
- nombre d’offres connu ou inconnu
- instants d’arrivée fixes ou aléatoires
- assouplissement des critères de sélection
- problème du parking
- …

c’est le sujet de la seconde partie de l’exposé