Retrieving Item Ordering with Binary Data

"Comparing Ways of Obtaining Candidates Orderings from Approval Ballots"

Théo Delemazure
Paris Dauphine University – PSL

Chris Dong
Technical University of Munich

Dominik Peters
Paris Dauphine University – PSL

Magdaléna Tydrichová
CentraleSupélec

Introduction: Input data

Among the following candidates, which do you approve?

	 Ann	 Bob	 Cora	 Dan
4 ×	✓	✓		✓
4 ×	✓	✓		
3 ×	✓			✓
1 ×	✓	✓	✓	
1 ×		✓	✓	

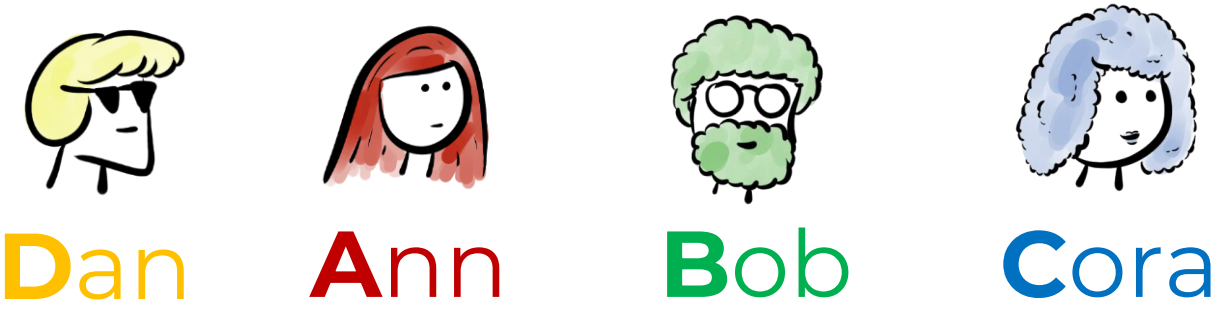
Introduction: Input data

Among the following candidates, which do you approve?

	 Dan	 Ann	 Bob	 Cora
4 x	✓	✓	✓	
4 x		✓	✓	
3 x	✓	✓		
1 x		✓	✓	✓
1 x			✓	✓

Introduction: Input data

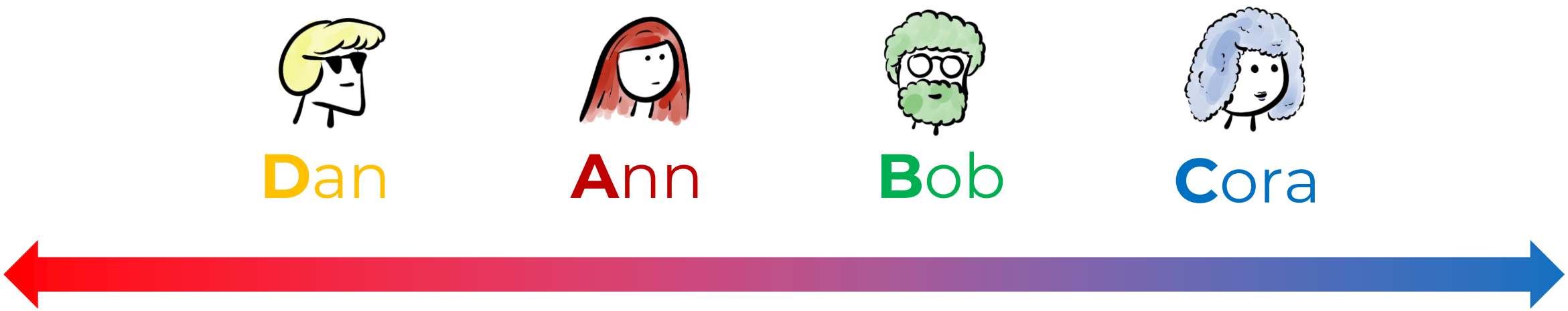
Among the following candidates, which do you approve?



	Dan	Ann	Bob	Cora
4 x	✓	✓	✓	
4 x		✓	✓	
3 x	✓	✓		
1 x		✓	✓	✓
1 x			✓	✓

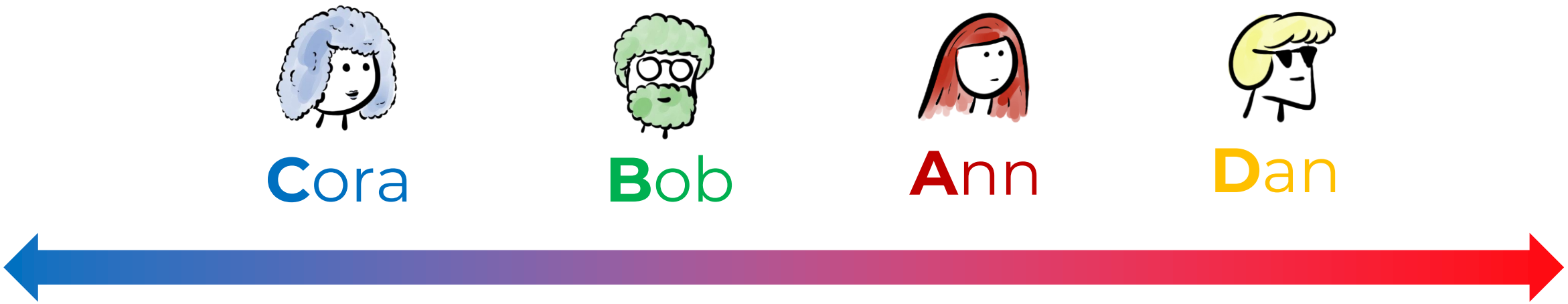
Introduction: Output

(one of) The correct axis is:



Introduction: Output

(one of) The correct axis is:



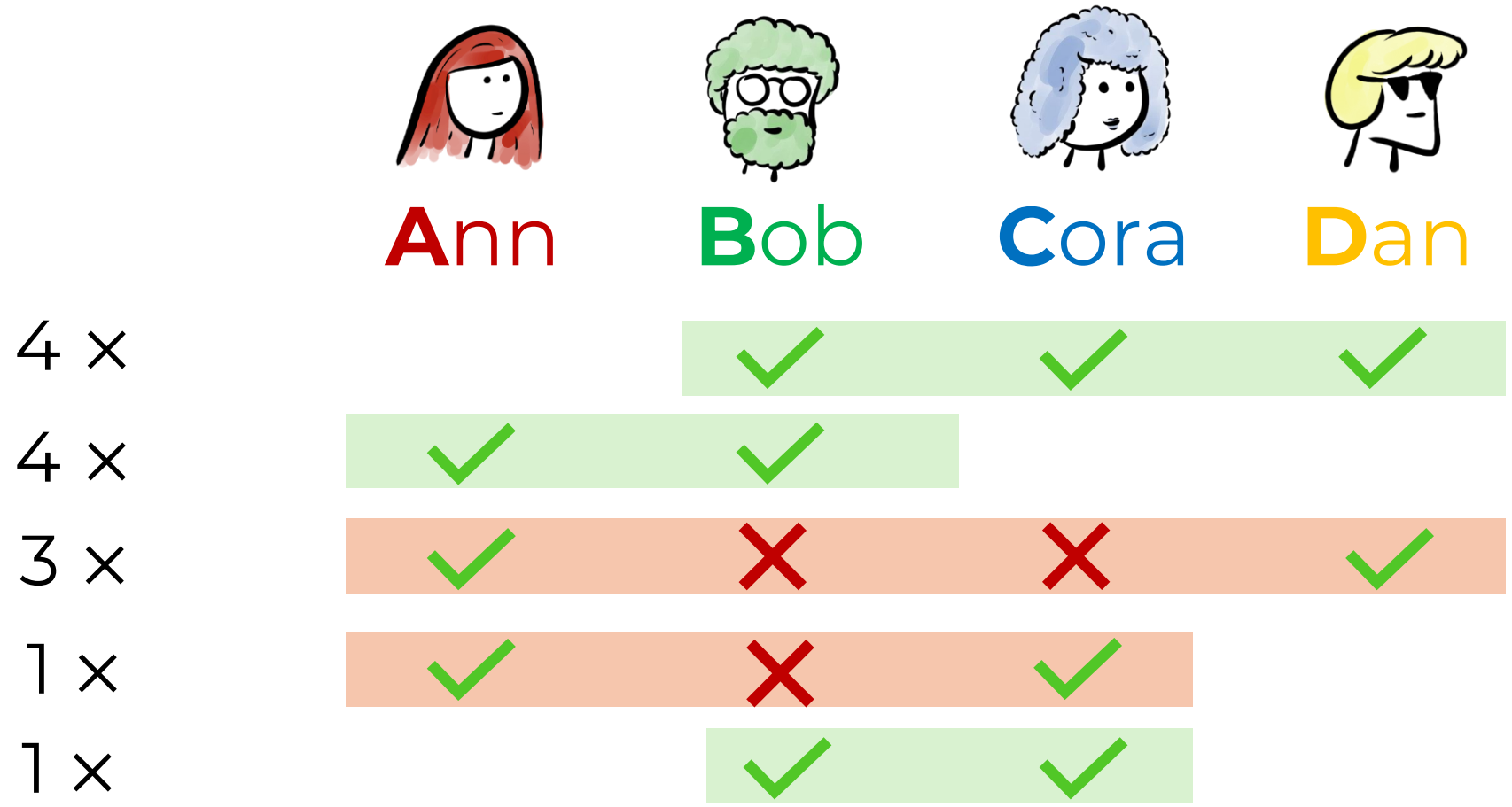
Introduction: The problem

Among the following candidates, which do you approve?

	 Ann	 Bob	 Cora	 Dan
4 x		✓	✓	✓
4 x	✓	✓		
3 x	✓			✓
1 x	✓		✓	
1 x		✓	✓	

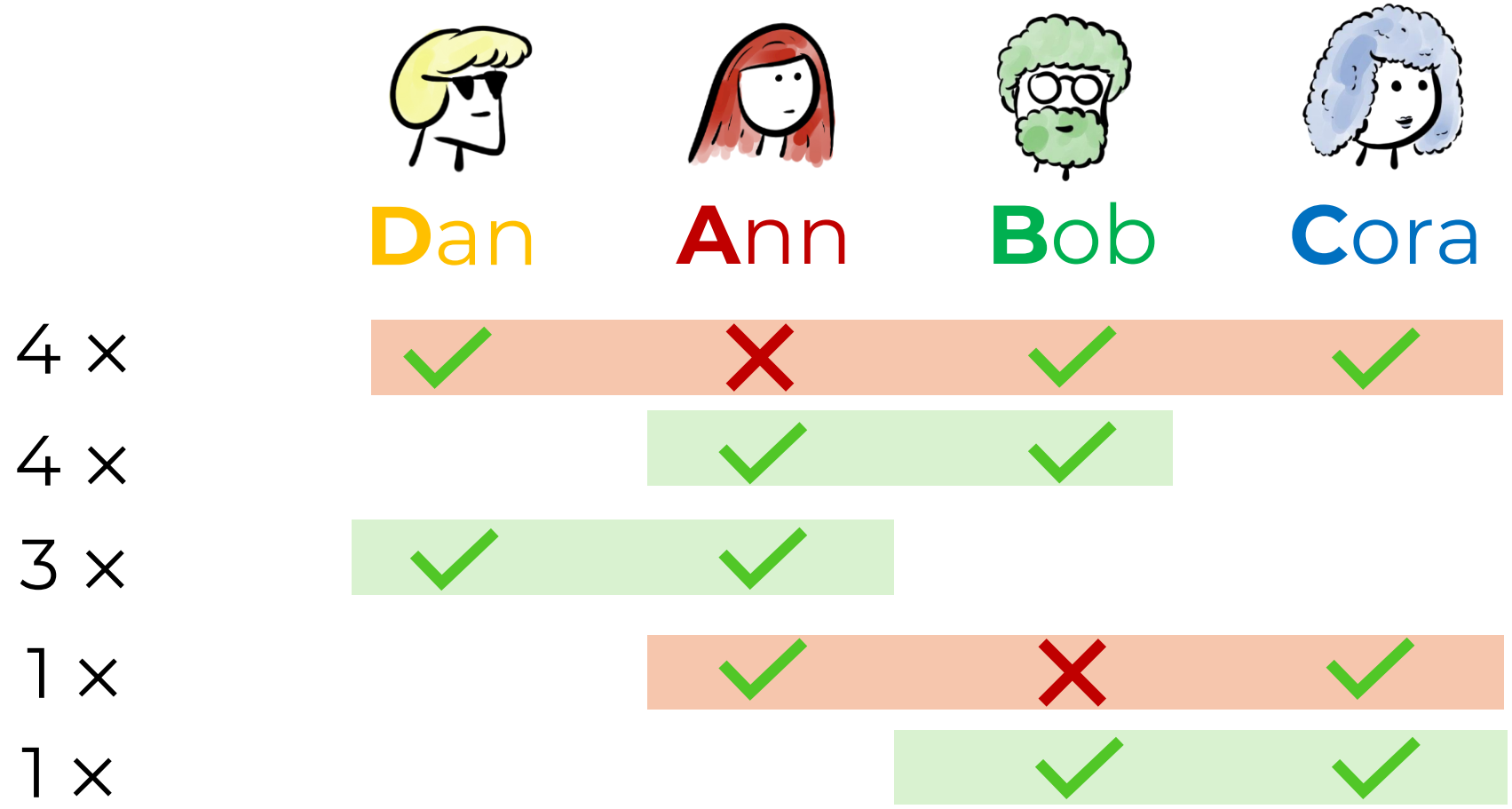
Introduction: The problem

Among the following candidates, which do you approve?



Introduction: The problem

Among the following candidates, which do you approve?



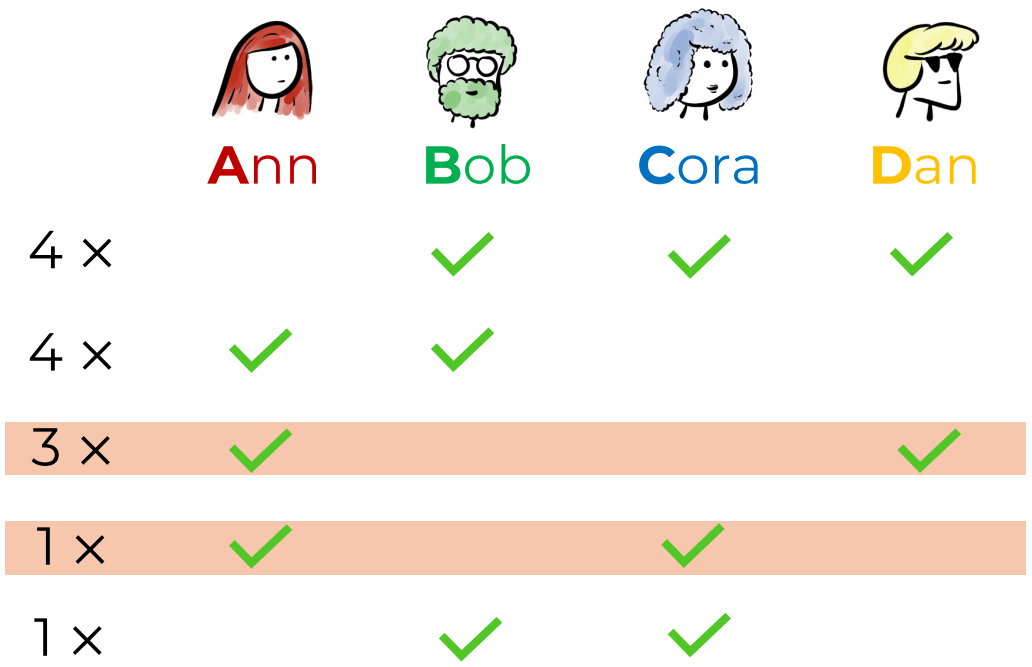
Question

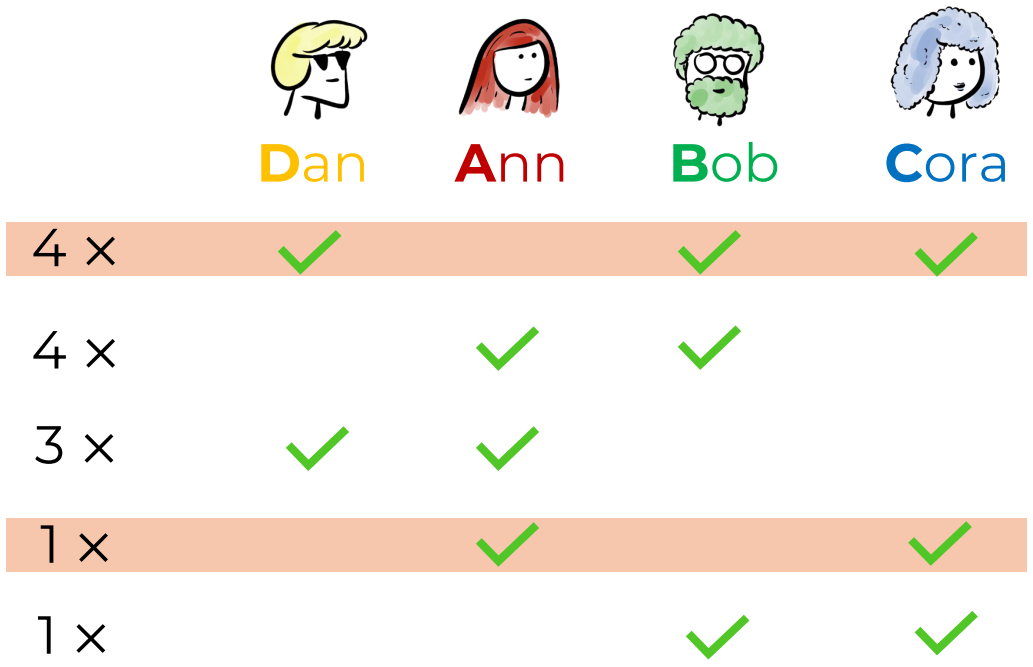
When there is *no perfect axis*, what methods can we use to find the near-axes?

Five methods

Solution 1 : Voter Deletion

The optimal axis is the one for which we need to delete the least number of voters to get a *linear* profile.





Five methods

Voter Deletion

Minimizes how many votes are not interval.

?

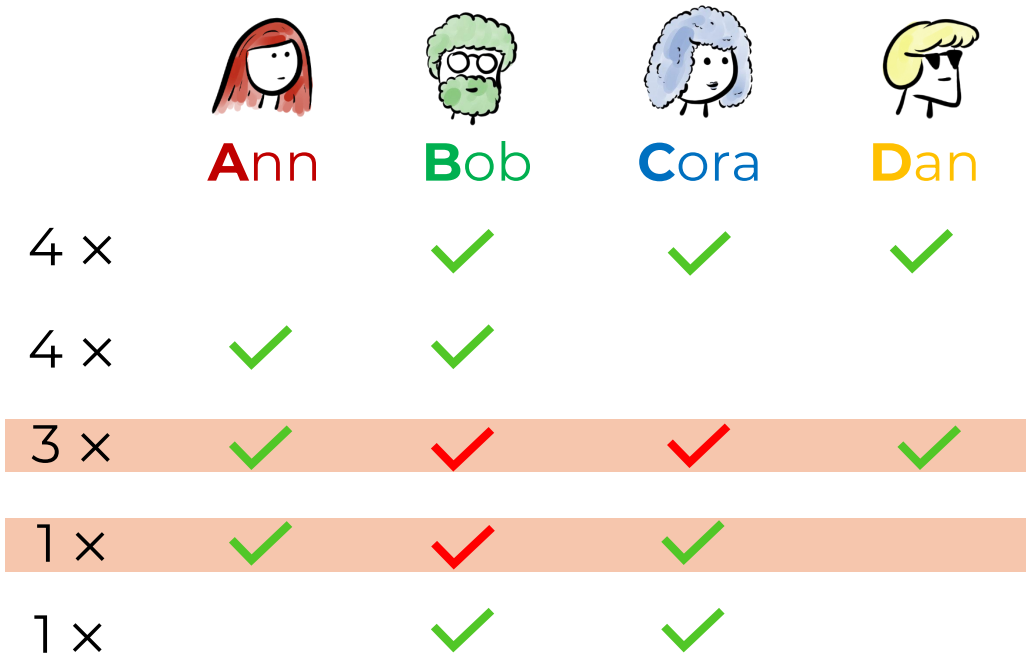
?

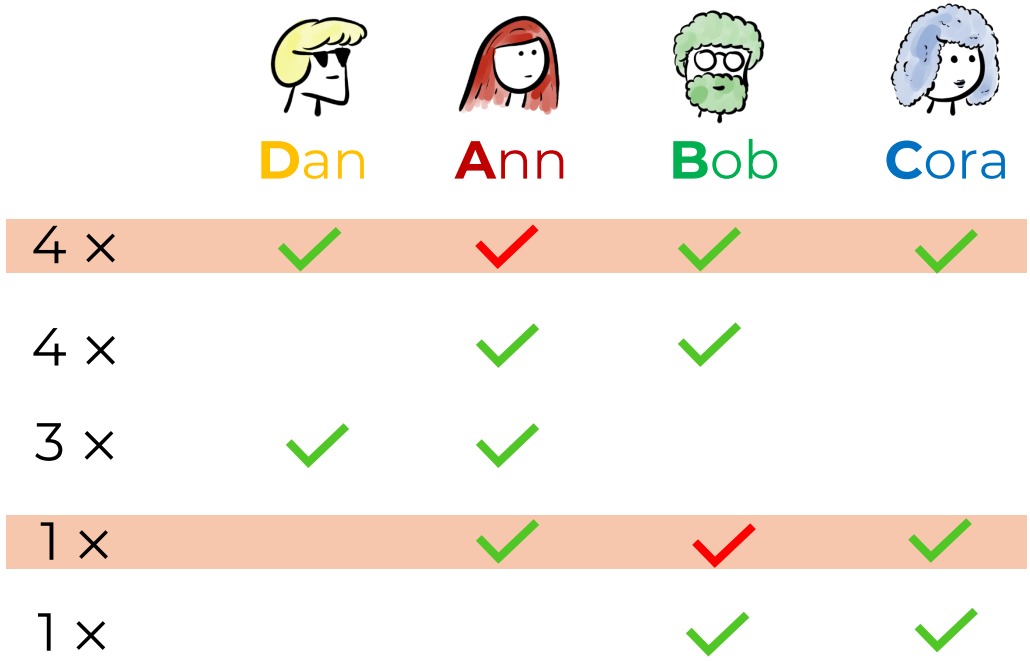
?

?

Solution 2 : Ballot Completion

The optimal axis is the one for which we need to add the least number of ✓ to get a *linear* profile.





Five methods

Voter Deletion
Minimizes how many votes are not interval.

Ballot Completion
Minimizes how many ✓ need to be added.

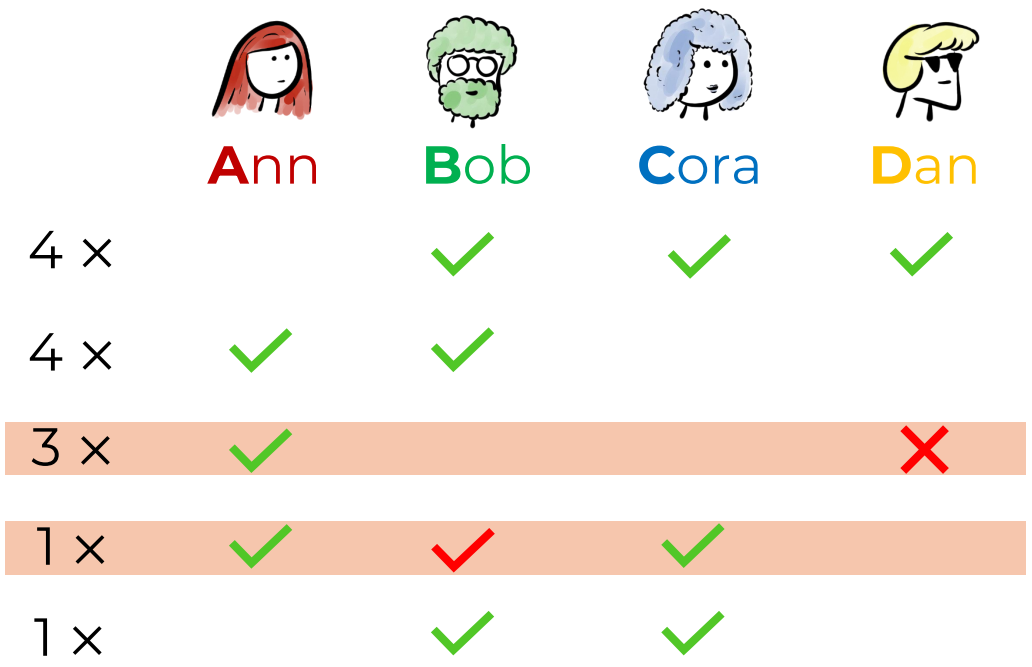
?

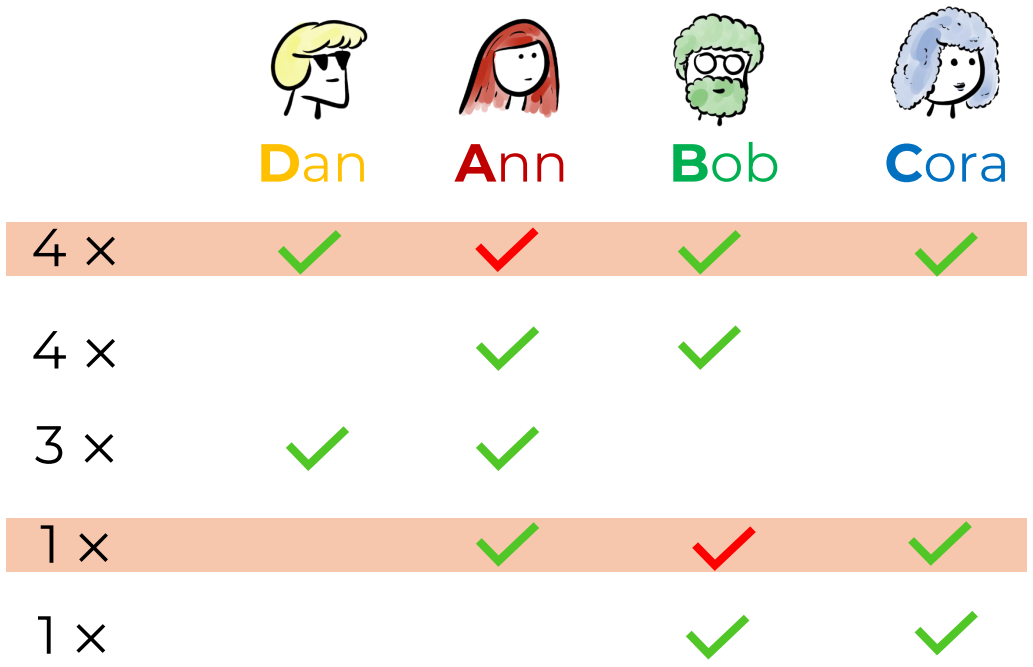
?

?

Solution 3 : Minimum Flips

The optimal axis is the one for which we need to **add/remove** the least number of ✓ / ✓ to get a *linear* profile.





Five methods

Voter Deletion

Minimizes how many votes are not interval.

Ballot Completion

Minimizes how many ✓ need to be added.

Minimum Flips

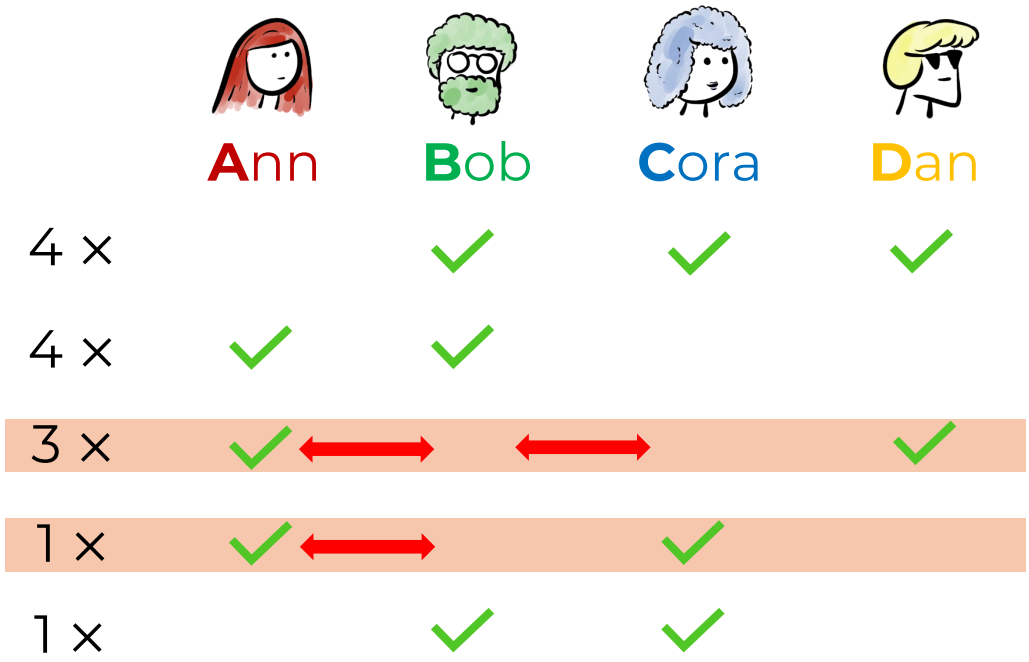
Minimizes how many ✓ need to be added/removed.

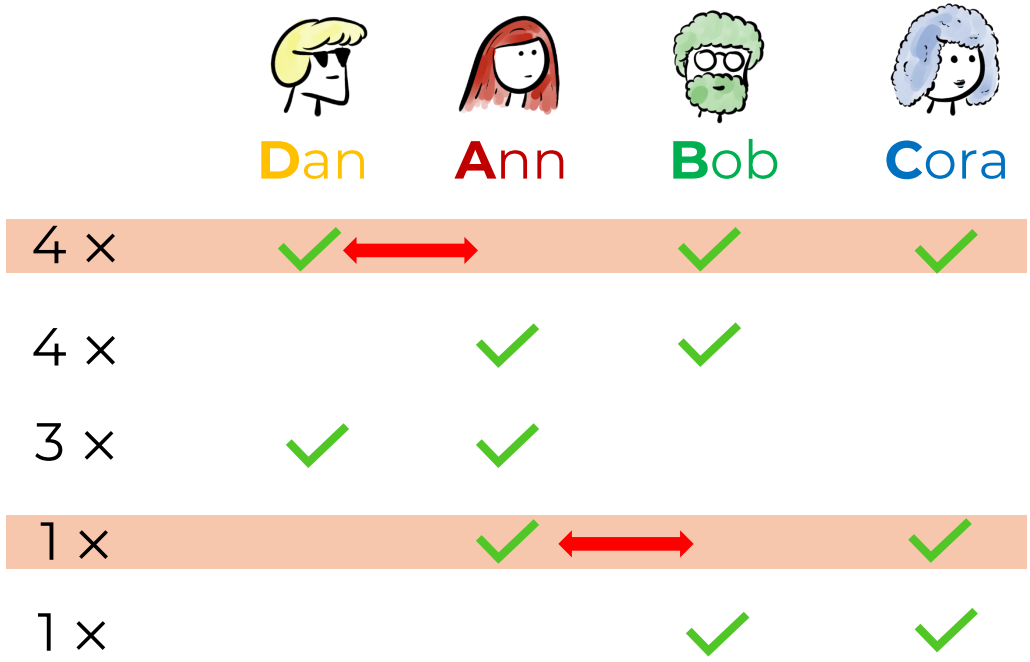
?

?

Solution 4 : Minimum Swaps

The optimal axis is the one for which we need to **swaps** the least number of candidates for each voter to get a *linear* profile.





Five methods

Voter Deletion

Minimizes how many votes are not interval.

Ballot Completion

Minimizes how many ✓ need to be added.

?

Minimum Flips

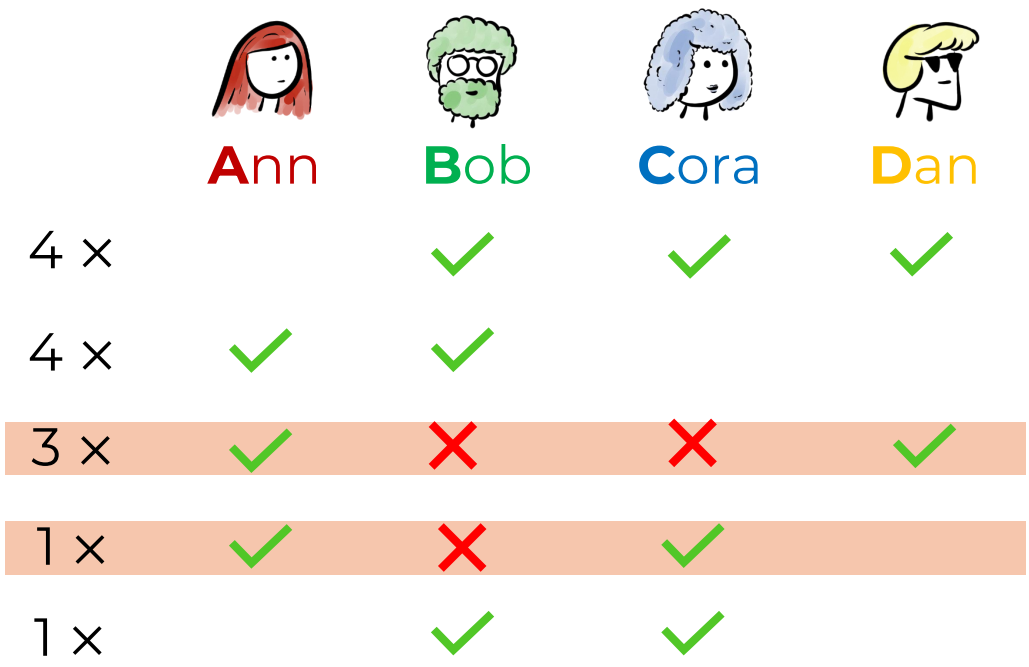
Minimizes how many ✓ need to be added/removed.

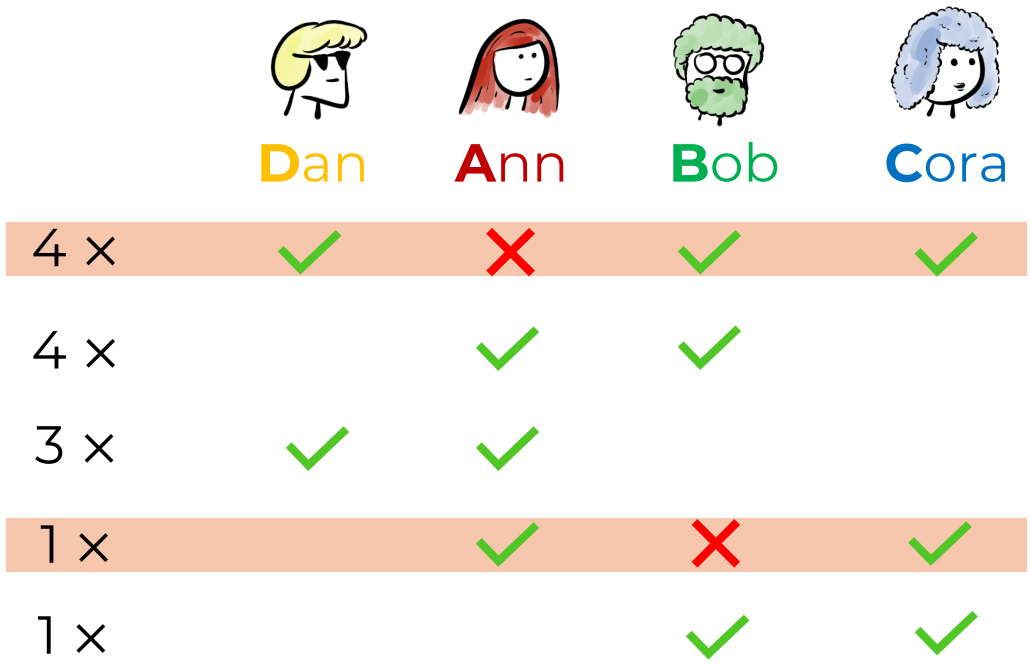
Minimum Swaps

Minimizes the number of swaps on the axis.

Solution 5 : Forbidden Triples

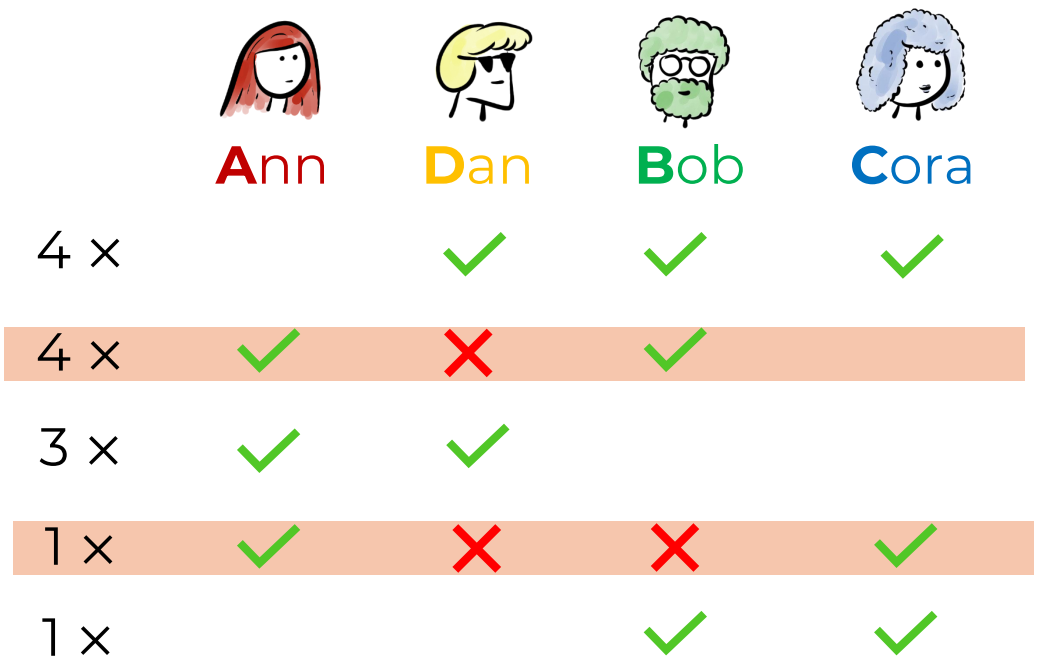
The optimal axis is the one which minimizes the number of triplets (✓, ✗, ✓).

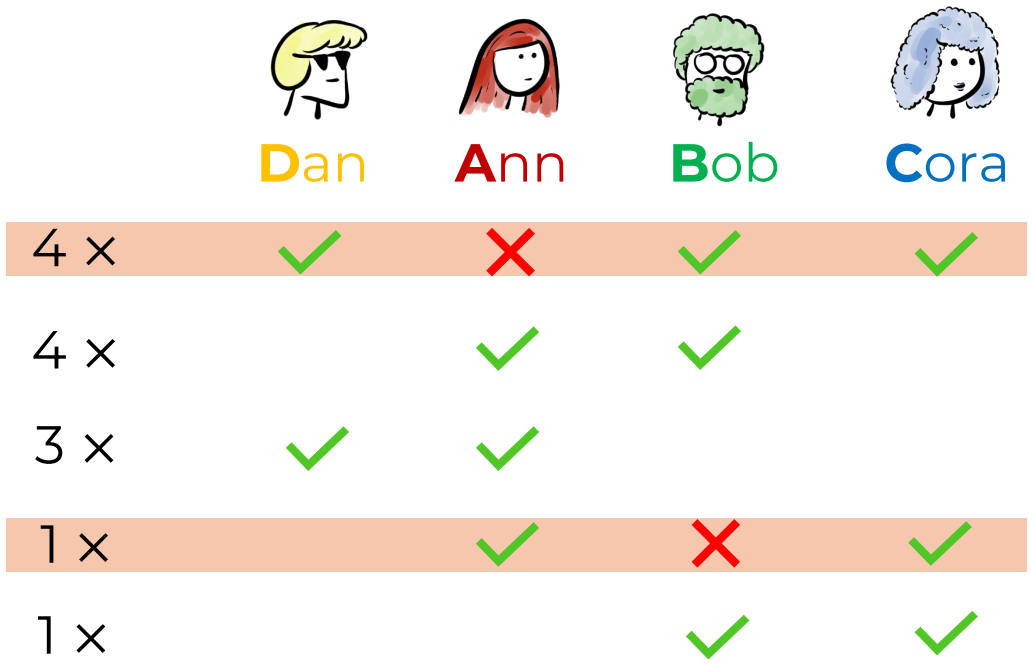




Solution 5 : Forbidden Triples

The optimal axis is the one which minimizes the number of triplets (✓, ✗, ✓).





Five methods

Voter Deletion

Minimizes how many votes are not interval.

Ballot Completion

Minimizes how many ✓ need to be added.

Forbidden Triples

Minimizes the number of triplets (✓, ✗, ✓).

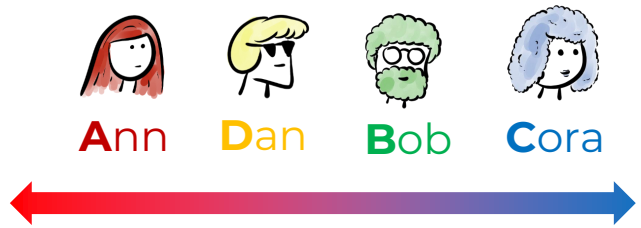
Minimum Flips

Minimizes how many ✓ need to be added/removed.

Minimum Swaps

Minimizes the number of swaps on the axis.

Five methods



Voter Deletion

Minimizes how many votes are not interval.

Ballot Completion

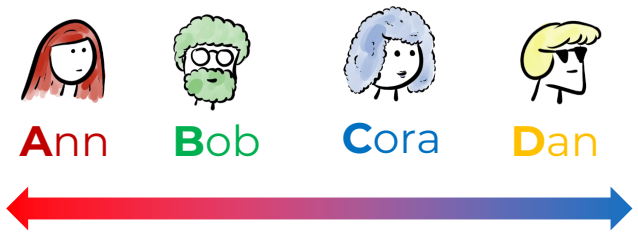
Minimizes how many ✓ need to be added.

Forbidden Triples

Minimizes the number of triplets (✓, ✗, ✓).

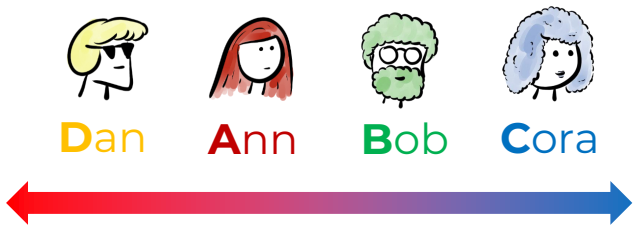
Minimum Flips

Minimizes how many ✓ need to be added/removed.



Minimum Swaps

Minimizes the number of swaps on the axis.



Question

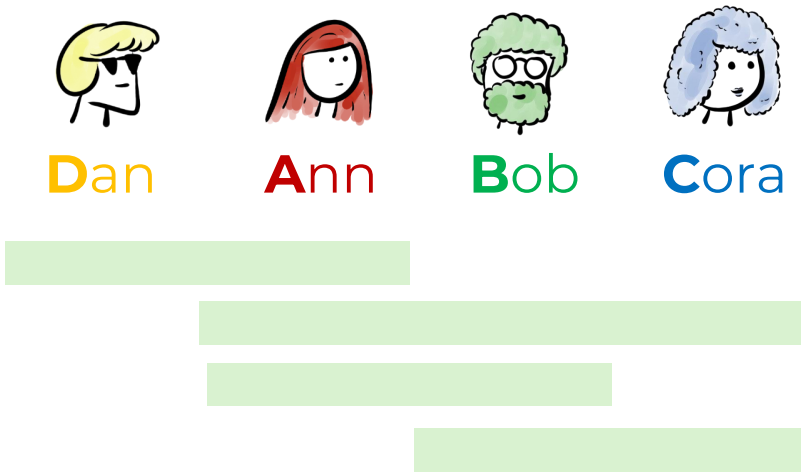
If different methods return different axes, which method(s) should we use?

Introducing...

The axiomatic method

Axioms: basic example

*If there exist a perfect axis for the profile**, this axis should be chosen.***



Axioms: clones

	 Dan	 Ann	 Bob	 Bobby	 Cora
4 x	✓		✓	✓	✓
4 x		✓	✓	✓	
3 x	✓	✓			
1 x		✓			✓
1 x			✓	✓	✓

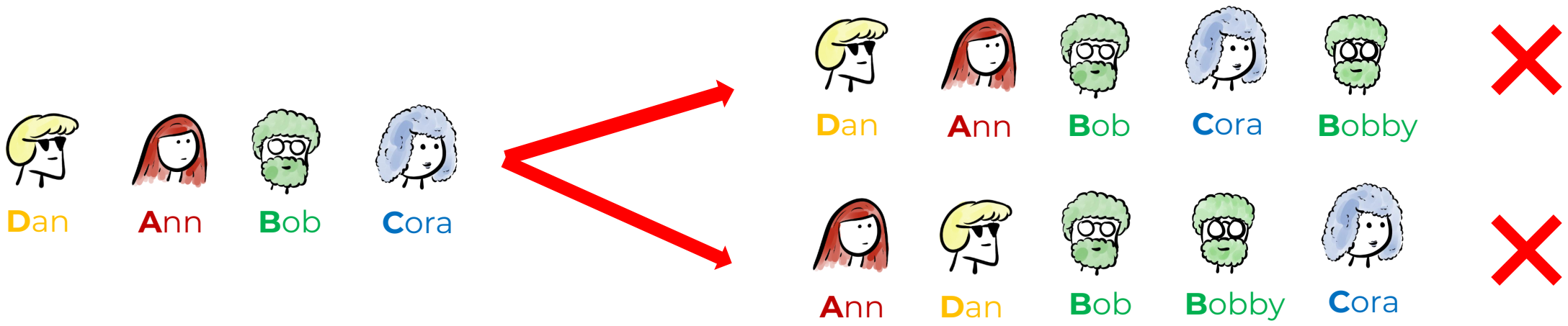
Axioms: clones

Clone-proximity

Clones should be next to each other on the axis.

Resistance to cloning

Adding a clone should not change the order of the other candidates on the axis.



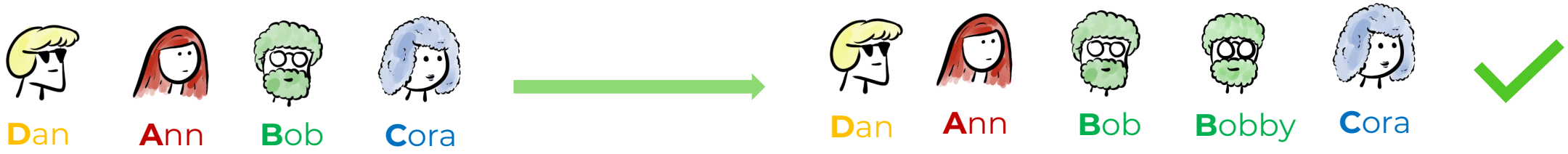
Axioms: clones

Clone-proximity

Clones should be next to each other on the axis.

Resistance to cloning

Adding a clone should not change the order of the other candidates on the axis.



Axioms: clones

Clone-proximity

Clones should be next to each other on the axis.

Resistance to cloning

Adding a clone should not change the order of the other candidates on the axis.

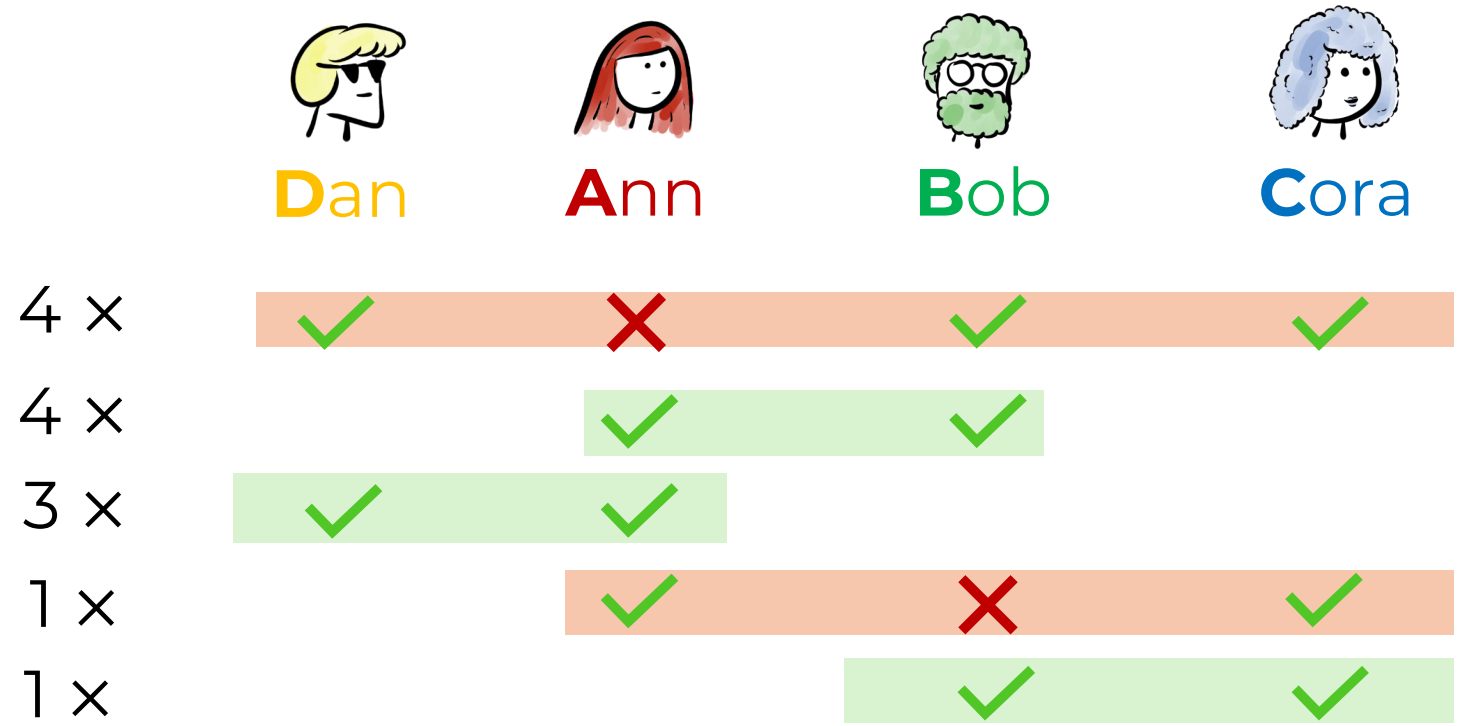
Impossibility Result

No rule can always satisfy both axioms.

Axioms

Ballot Monotonicity

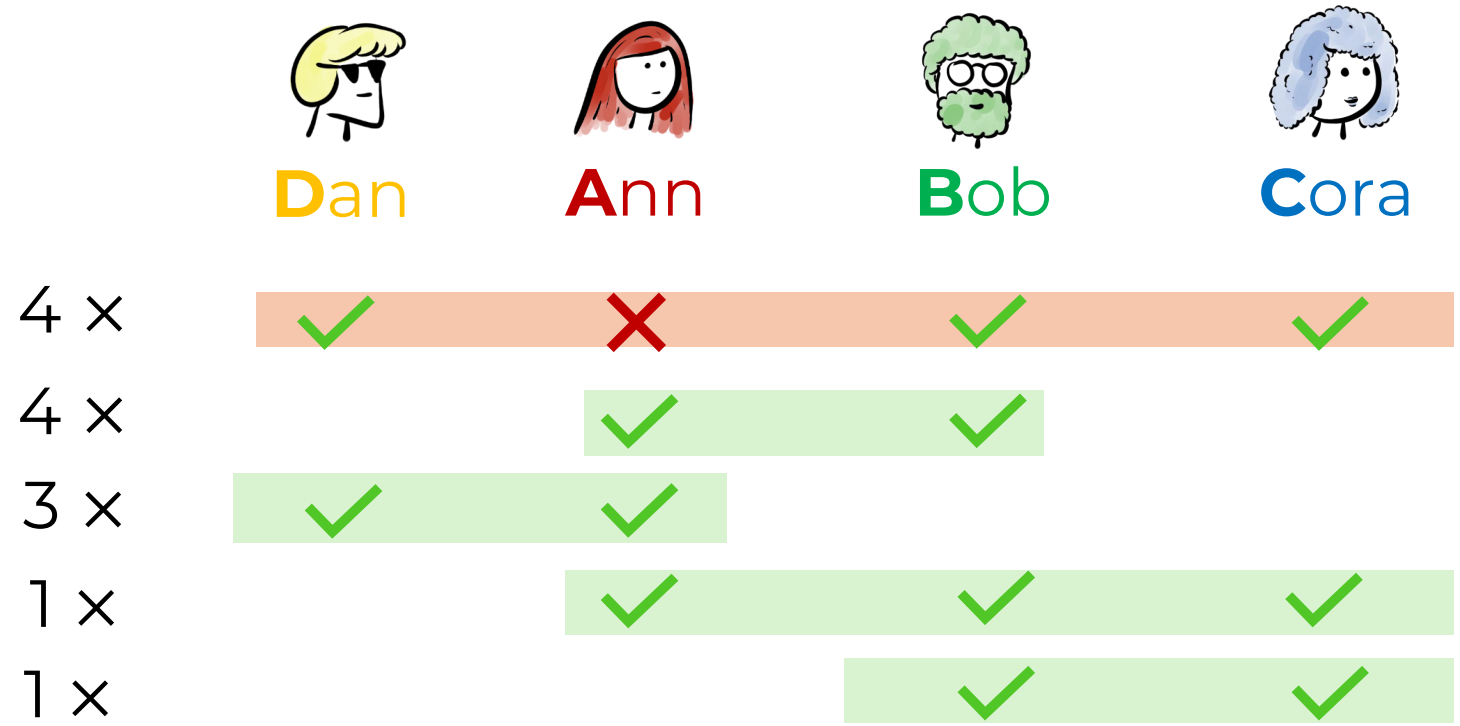
If we add approvals to the ballot of a voter to turn it into an **interval** of the optimal axis, this axis is still selected.



Axioms

Ballot Monotonicity

If we add approvals to the ballot of a voter to turn it into an **interval** of the optimal axis, this axis is still selected.



Axioms

Ballot Monotonicity

If we add approvals to the ballot of a voter to turn it into an **interval** of the optimal axis, this axis is still selected.

Characterization Result

Voter Deletion **is the only rule** to satisfy Ballot Monotonicity and Resistance to cloning.

Axioms satisfied

Voter Deletion

Minimizes how many votes are not interval.

- ✓ Ballot monotonicity
- ✓ Resistance to cloning

Ballot Completion

Minimizes how many ✓ need to be added.

- ✓ Ballot monotonicity
- ✓ Clearance

Forbidden Triples

Minimizes the number of triplets (✓, ✗, ✓).

- ✓ Clearance
- ✓ Veto-centrism
- ✓ Clone proximity

Minimum Flips

Minimizes how many ✓ need to be added/removed.

- ✗ Nothing

Minimum Swaps

Minimizes the number of swaps on the axis.

- ✓ Clearance
- ✓ Veto-centrism

Introducing...

Experiments

Experiments

Synthetic Data

We used different probabilistic models to generate data.

Real Data

We gather and collect approval data from various sources.

Experiments: France

Source of the data

"Voter Autrement : The online experiment"

Experiments: France

Source of the data

"Voter Autrement : The online experiment"

Baseline axes

Poll institutes

LE SECOND CHOIX DES ELECTEURS HÉSITANTS

Question : « Si finalement vous ne deviez pas voter pour [Nom du candidat], quel serait votre choix au premier tour ? »
*(Base : Inscrits, certains d'aller voter qui ont exprimé une intention de vote mais **qui ne sont pas certains de leur choix**)*

CANDIDAT CHOISI EN 1er ▼ / SECOND CHOIX →	Philippe Poutou	Nathalie Arthaud	Jean-Luc Mélenchon	Fabien Roussel	Yannick Jadot	Anne Hidalgo	Emmanuel Macron	Valérie Pécresse	Jean Lassalle	Nicolas Dupont-Aignan	Marine Le Pen	Eric Zemmour	Non exprimé	TOTAL
Jean-Luc Mélenchon	4	1		9	23	8	20	2	5	3	12	3	10	100
Fabien Roussel	4	1	35		15	4	16	2	2	3	3	4	11	100
Yannick Jadot	3	1	35	2		9	33	6	2	2	1	1	5	100
Emmanuel Macron	1		8	6	11	7		29	3	3	16	3	14	100

Experiments: France

Source of the data

"Voter Autrement : The online experiment"

Baseline axes

Poll institutes

Results

Experiments: US

Source of the data

Opinions of the justices of the Supreme court of the USA.

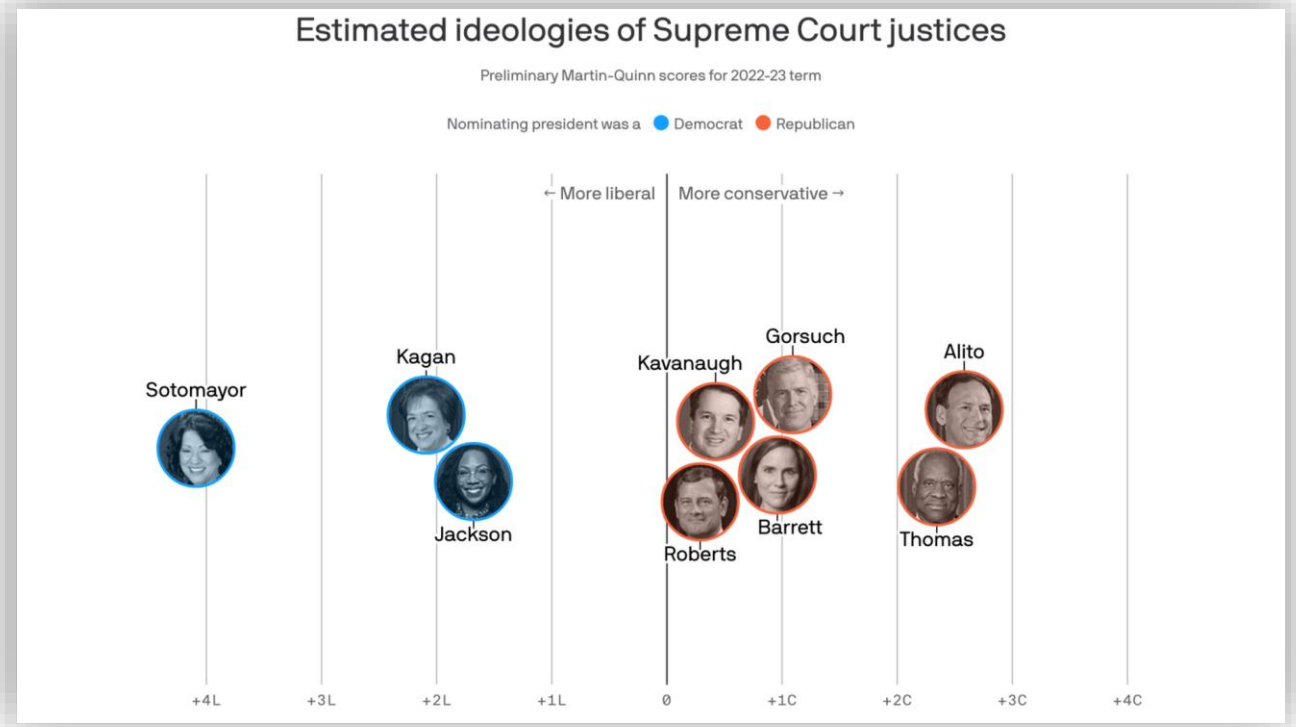
Experiments: US

Source of the data

Opinions of the justices of the Supreme court of the USA.

Baseline axes

Using the Martin-Quinn method



Experiments: US

Source of the data

Opinions of the justices of the Supreme court of the USA.

Baseline axes

Using the Martin-Quinn method

Results

Forbidden Triple on the 2021 term

	Voter Deletion	*Min. Flips*	*Ballot Completion*	*Mini. Swaps*	*Forbidden Triples*
Correct median (65 years)	**54%**	**58%**	**57%**	**65%**	**66%**

Experiments: TierMaker

Source of the data

Tierlist maker website

Experiments: TierMaker

Source of the data

Tierlist maker website

Baseline axes

I’m just having fun here

Experiments: TierMaker

Source of the data

Tierlist maker website

Baseline axes

I'm just having fun here

Results (Ballot Completion)

Solo < II < I < Rogue One < III < V < IV < VI < VIII < VII < IX

Experiments: TierMaker

Source of the data

Tierlist maker website

Baseline axes

I'm just having fun here

Results (Ballot Completion)

Solo < II < I < Rogue One < III < V < IV < VI < VIII < VII < IX

Mar < Apr < May < Aug < Jun < Jul < Dec < Oct < Nov < Sep < Jan < Feb

Experiments: TierMaker

Source of the data

Tierlist maker website

Baseline axes

I'm just having fun here

Results (Ballot Completion)

Solo < II < I < Rogue One < III < V < IV < VI < VIII < VII < IX

Mar < Apr < May < Aug < Jun < Jul < Dec < Oct < Nov < Sep < Jan < Feb

5 < 10 < 1 < 2 < 4 < 3 < 7 < 9 < 6 < 8

Experiments: TierMaker

Source of the data

Tierlist maker website

Baseline axes

I'm just having fun here

Results (Ballot Completion)

Solo < II < I < Rogue One < III < V < IV < VI < VIII < VII < IX

Mar < Apr < May < Aug < Jun < Jul < Dec < Oct < Nov < Sep < Jan < Feb

5 < 10 < 1 < 2 < 4 < 3 < 7 < 9 < 6 < 8

Experiments: TierMaker

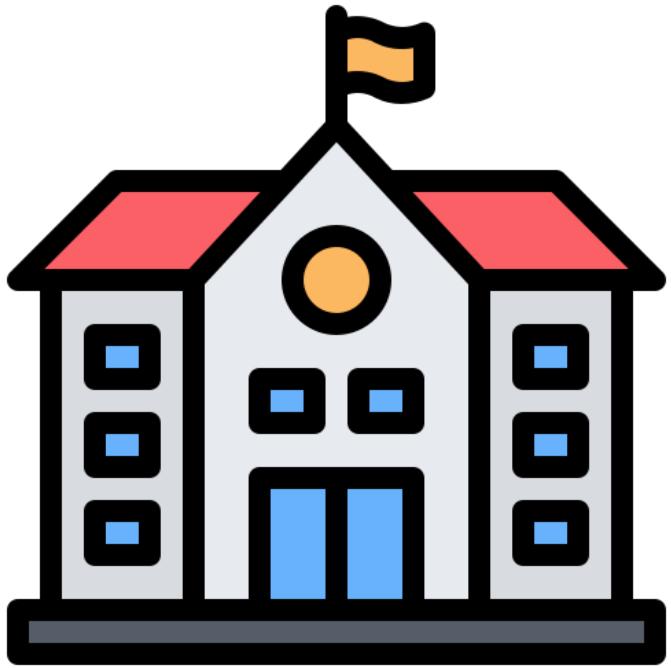
Source of the data

Tierlist maker website

Baseline axes

I'm just having fun here

Results (Ballot Completion)



Geometry
Chemistry
Physics
Math
Biology
Technology

Music
Art
Physical Education
History
Social Studies
Foreign Language
Literature

Question

When do we need to reconstruct such axes?

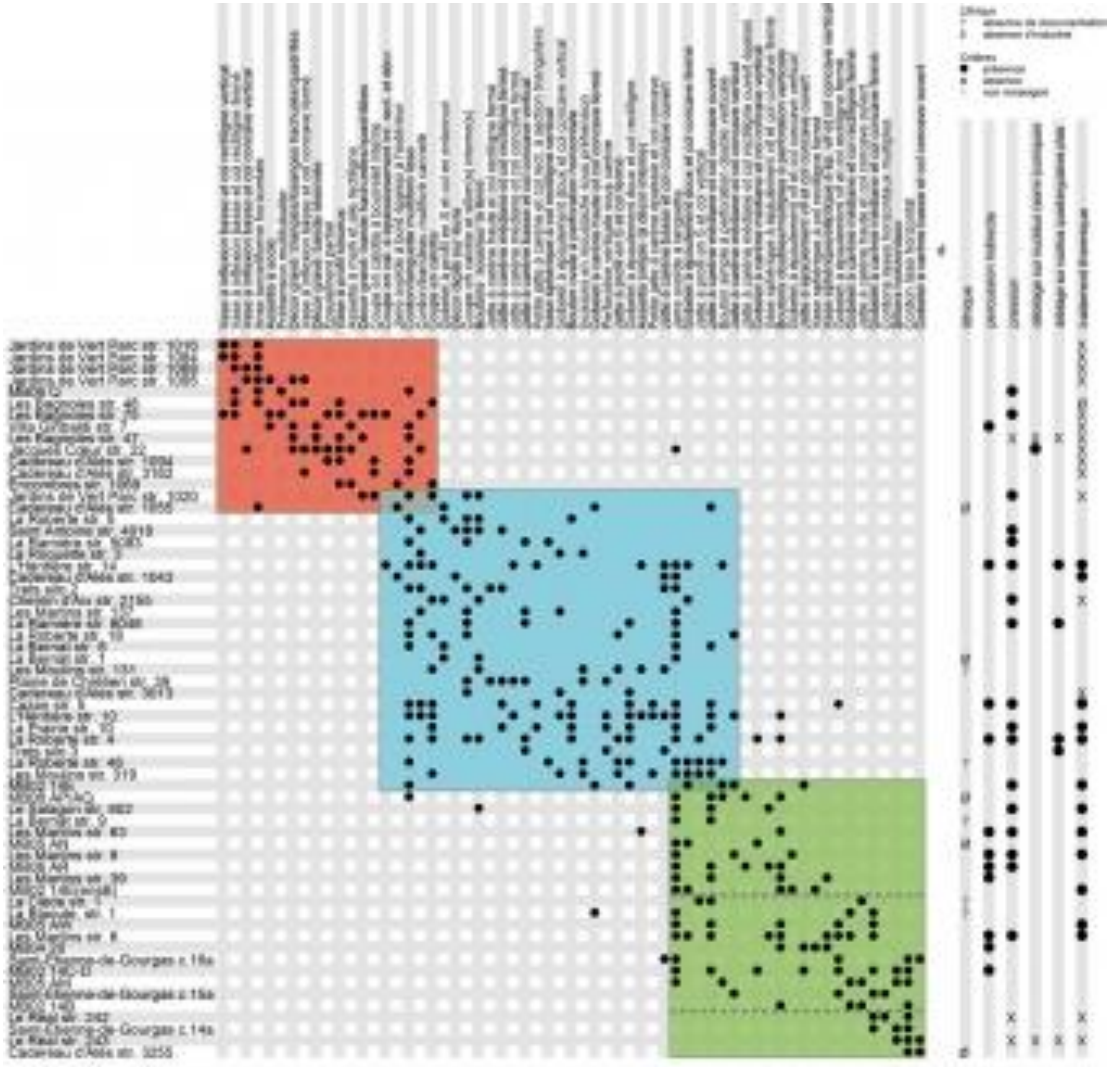
Motivation: Political science

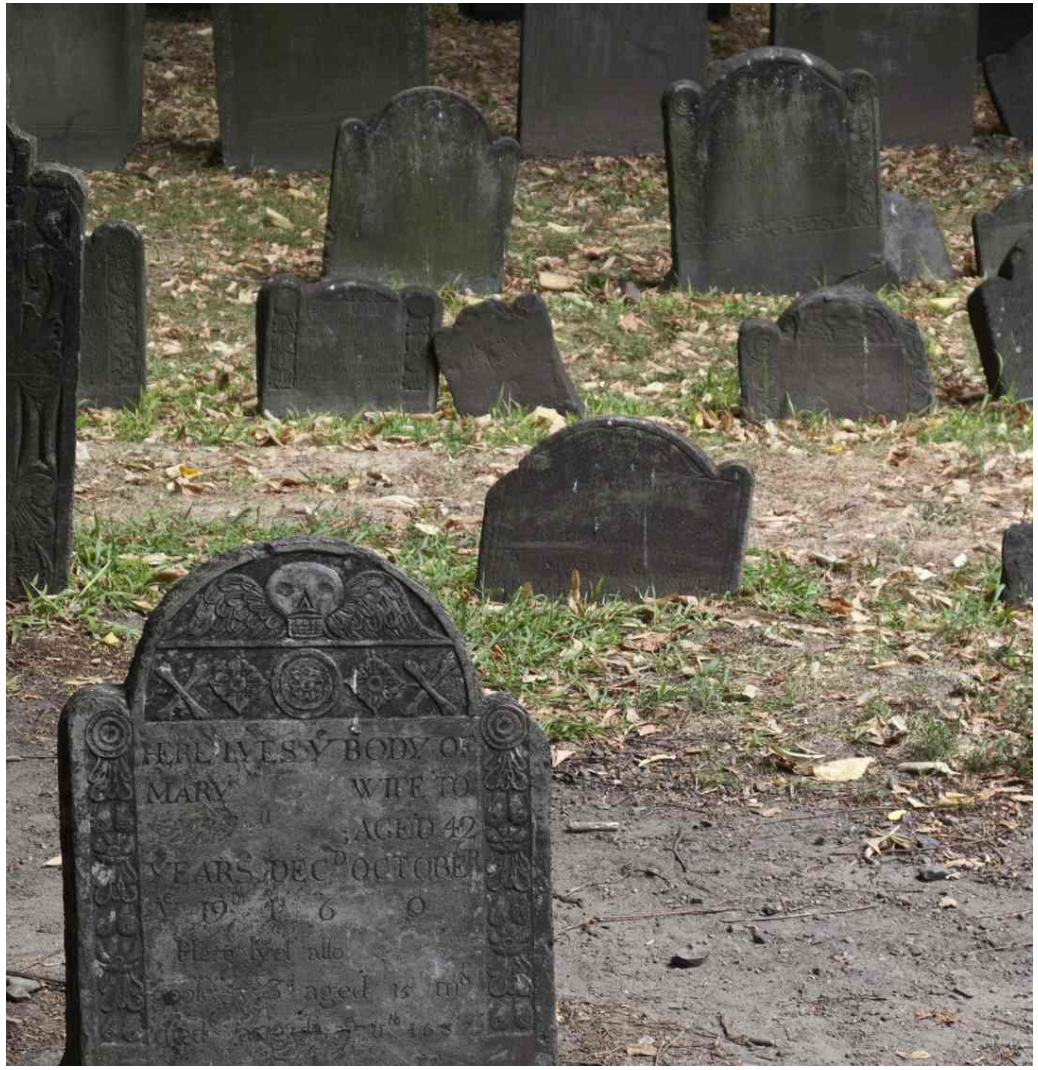
Ce que le vote par approbation révèle des préférences des électeurs français

Isabelle Lebon, Antoinette Baujard, Frédéric Gavrel, Herrade Igersheim, Jean-François Laslier

DANS REVUE ÉCONOMIQUE **2017/6 (Vol. 68)**, PAGES 1063 À 1076

Motivation: Seriation (archeology)





Motivation: Scheduling

	Lundi	Mardi	Mercredi	Jeudi
8h30-10h	*Arrivée*	*Petit-déjeuner*		
10h-12h30		*Session méthodologie*	*Session thématique 3 :* **Complexité** Camille Richer Henrik Agbaryan Sofia Vazquez Alferez *Chair :* Pierre Cazals	*Session thématique 5 :* **Modéliser l'humain** Louise Dupuis Sarra Tajouri *Chair :* Felipe Garrido
12h30-14h	*Déjeuner*			
14h-16h30	*Session thématique 1 :* **Adversarité** Roxane Cohen Lucas Gnecco Heredia Ahmad Qadeib Alban *Chair :* Matthieu Hervouin	*Session thématique 2* : **Reconstruire l'information** Houria Braikia Nicolas Fayard *Chair :* Tahar Allouche	*Session thématique 4* : **Applications réelles** Thibault de Surrel Lola Martin Charles Nourry *Chair :* Céline Béji	*Départ*
16h30-17h	Introduction Installation	Discussions/rump session	Discussions/rump session	
17h-19h30				
19h30-21h	*Dîner*			

Conclusion

We want to reconstruct: an ordering of items.
Example: political axis, chronological axis, optimal spatial ordering.

What we know: binary information on these items.
Example: preferences of voters, items particularities.

How to do it: we study 5 deterministic methods.
Specifically: Voter Deletion, Minimum Flips, Ballot Completion, Minimum Swaps, Forbidden Triples.

How to compare them: The axiomatic method and experiments.
Axiomatic results: Axioms satisfactions, Impossibility theorems, characterizations.
Experiments: On synthetic (with various models) and real (from various sources) data.

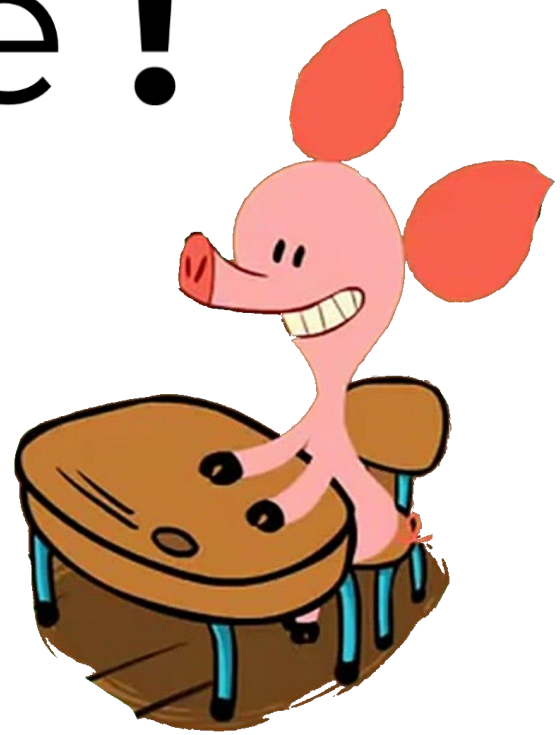
More questions

What if... the binary data contains *uncertainty*?

What if... we want a *more expressive structure* than a linear ordering (circular, two dimensional, with positions...)?

What if... we want to be able to *compute the result fast*, and with a lot of data?

Question time!



Now giving the floor to the **great**, the **beautiful**, the **magnificent**...

NICOLAS FAYARD !