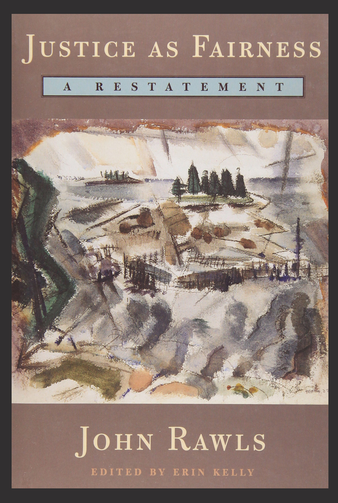
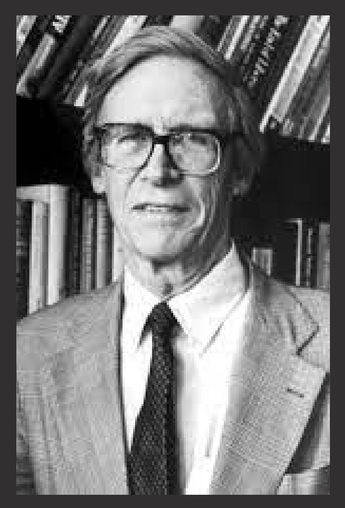
Can *Rawls' ideas of fairness* be embodied in *k-means clustering*?

It seems so!

> *... they [social and economic inequalities] are to be to the greatest benefit of the least-advantaged members of society.*
>
> – Rawls (Justice as Fairness)





John Rawls (1921–2002) was an influential 20th century moral and political philosopher in liberal tradition. He is frequently cited in courts of law and by politicians in the US and UK. His ideas of fairness are regarded as time-tested and a good mix of pragmatism and principledness.

1. *k*-means Clustering

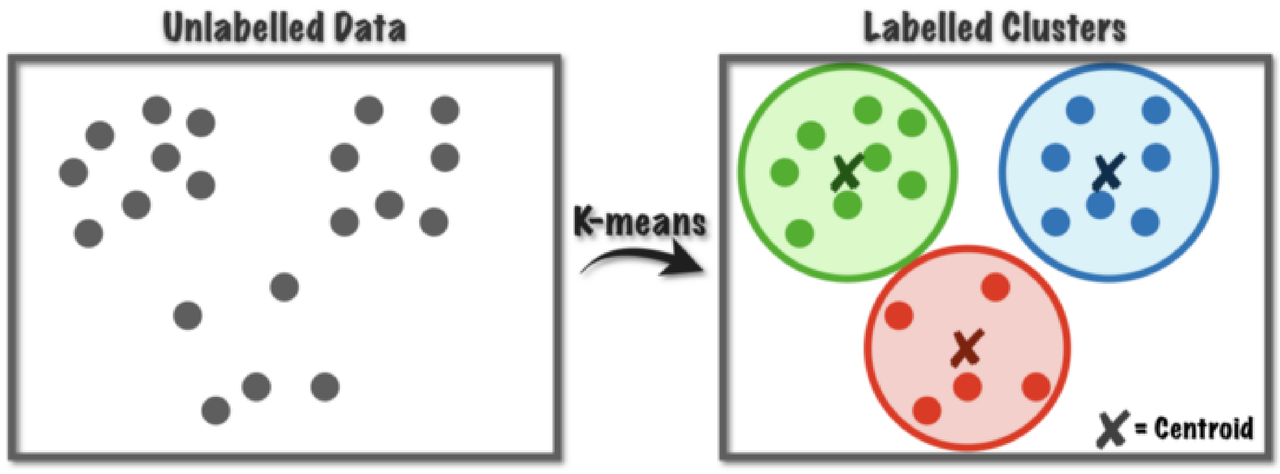


image taken from `shorturl.at/osyEW`

2. Fairness Notion

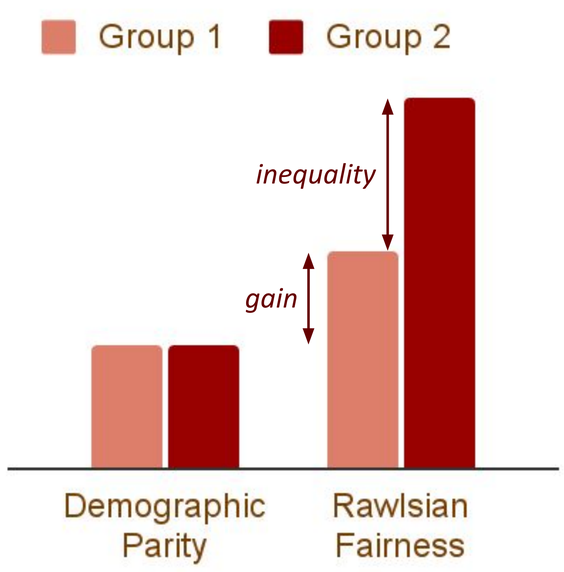
Do centroids represent groups 'fairly'?

(Popular notion) Demographic Parity

minimum inequality

Rawlsian Fairness

allow inequality, but represent least-advantaged group better



3. Research Questions

Is there a cluster assignment satisfying Rawlsian Fairness?

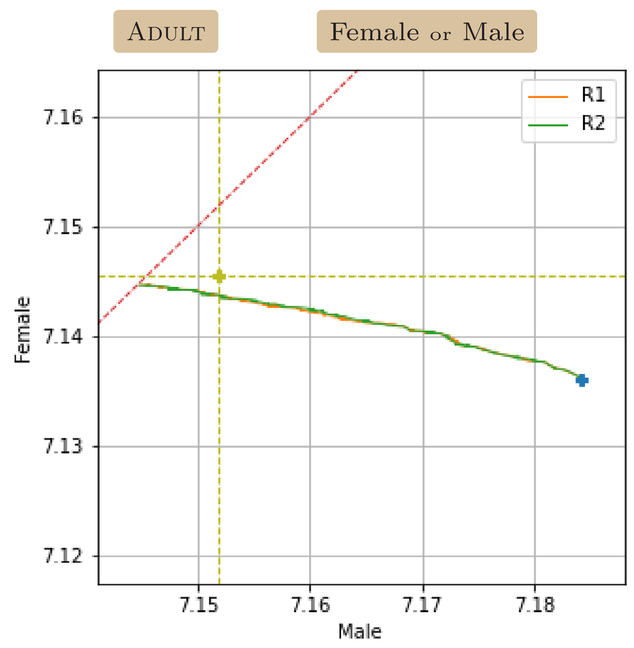
Yes, approximately

indicated by ✤ in the plot below

Can *k*-means clusters be perturbed to embody Rawlsian Fairness?

Yes, but can be improved

trajectories R1 and R2 moving from ✤ towards ✤ in the plot below



✤ *k*-means cluster assignment
✤ Rawlsian cluster assignment
--- Demographic Parity
— R1: reassign one data point at a time
— R2: reassign two data points at a time

Based on the paper '*Exploring Rawlsian Fairness for K-Means Clustering*' by **Stanley Simoes, Deepak P, Muiris MacCarthaigh** at ICDSE '21.

QUEEN'S UNIVERSITY BELFAST | CITI-GENS

Co-funded by the Horizon 2020 programme of the European Union

Stanley Simoes ✉ `ssimoes01@qub.ac.uk`