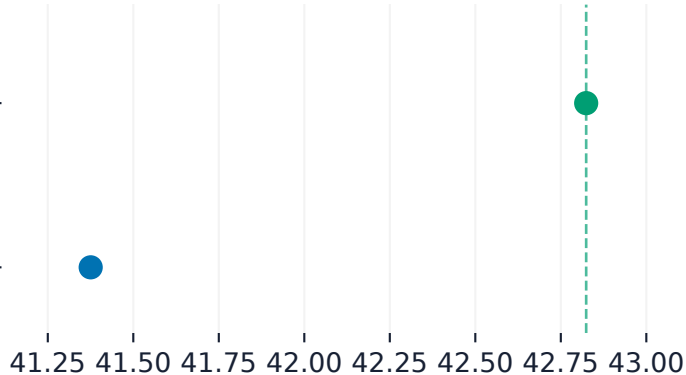


Task mean, %

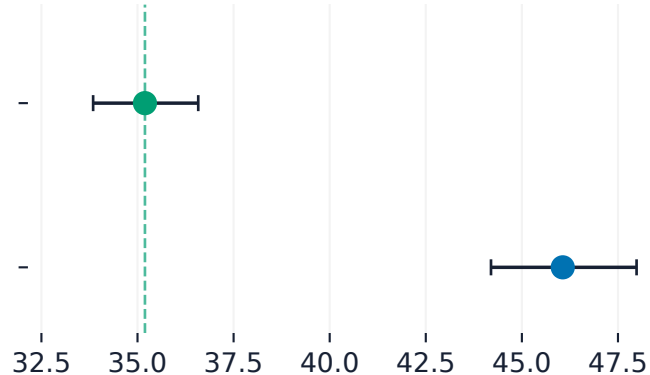
Conditional PPL

Q dense -

Zarya (AR + MDM) -



higher is better



lower is better · 95% CI