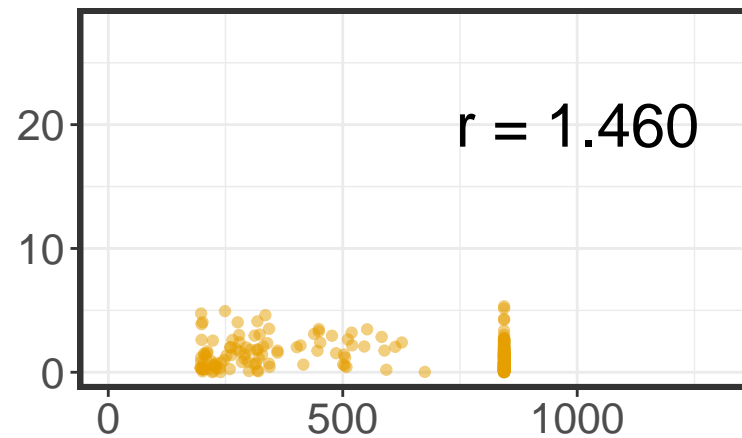
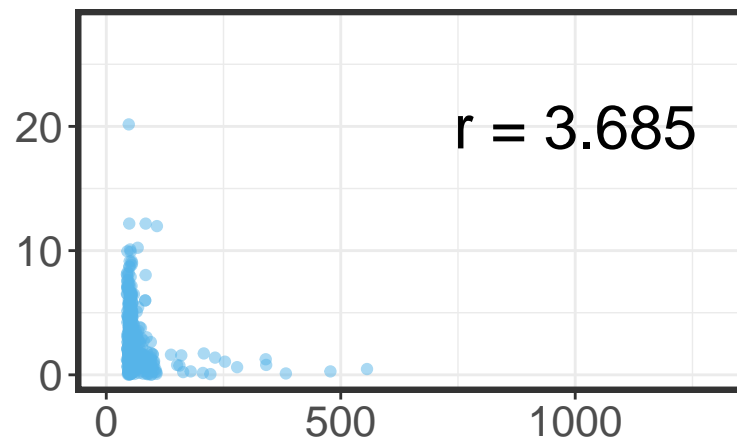


Score Recovery Error

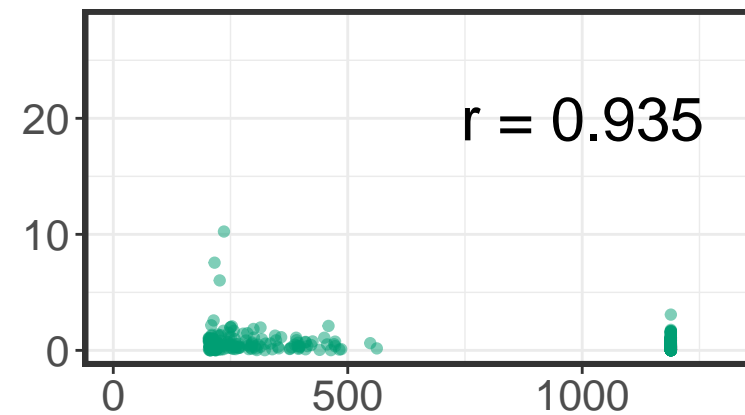
ARC



GSM8K

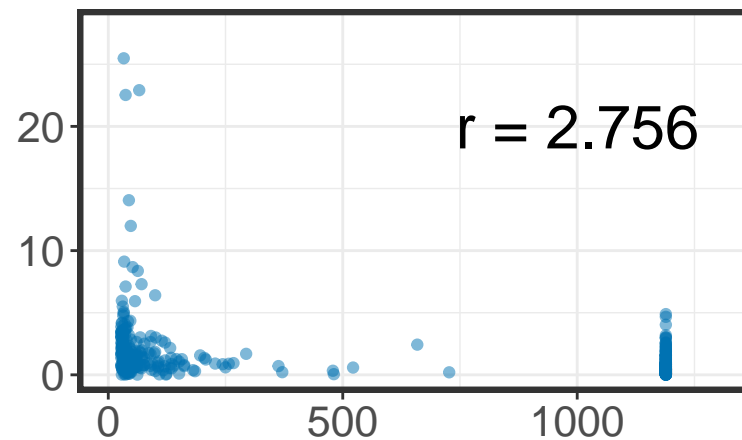


HellaSwag

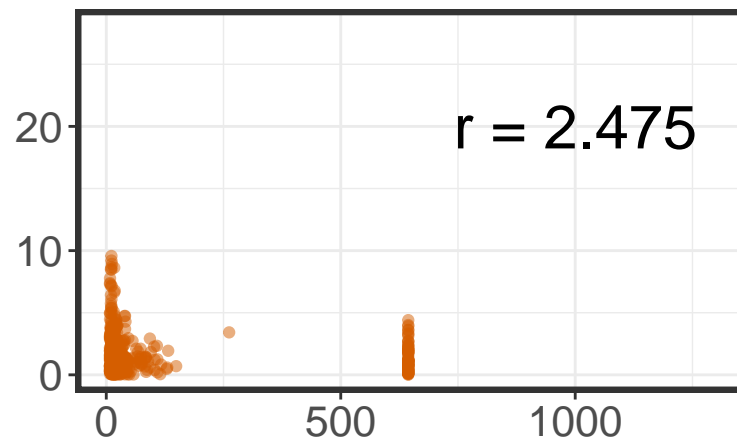


Score Recovery Error

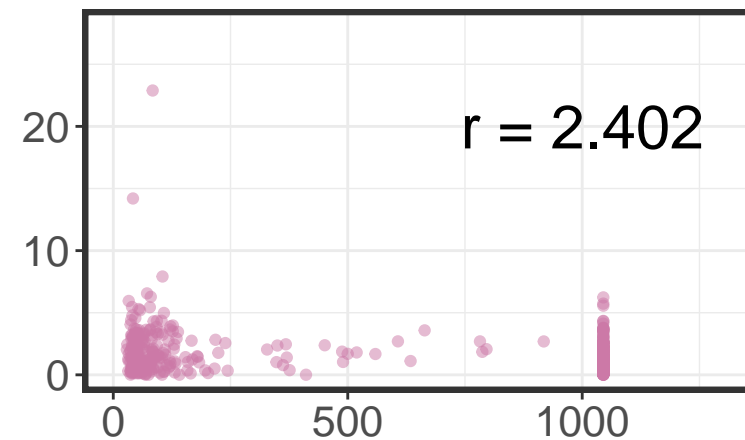
MMLU



TruthfulQA



Winogrande



Number of Items Administered

Number of Items Administered

Number of Items Administered