A model that truly reasons about data should not be affected by the origin of the data.

Original vs Counterfactual for GPT-4o.

Gap between original and counterfactual data diminishes as we use symbolic SQL-based reasoning