Real vs. Synthetic Data for column 'source'