

# Contributions

**Introduced TEMPTABQA-C**, a synthetic dataset for **precise and controlled evaluation** of temporal tabular reasoning across **diverse and complex scenarios**.

**Analyzed the limitations of direct prompting**, including reliance on **memorization**, sensitivity to **table size**, and difficulty with **multi-step and counterfactual reasoning**.

**Proposed a symbolic intermediate representation**, enabling LLMs to **generate and execute SQL queries** on structured schemas, enhancing **interpretability**, reducing **biases**, and improving **generalization**.

