

PI2: Interactive Visualization Interface Generation for SQL Analysis in Notebook

Introduction

Visualization is an integral part of the data analysis workflow, but specifying and tweaking visualizations in code is tedious.

Existing Notebooks:

- Do not support visualization interfaces (e.g. zoom/pan)
- Require manual effort for widget layout and creation
- Limited query parameter widget mapping interaction

PI2 Notebook Extension:

- Automatically generates complete interactive interfaces from SQL queries.

Notebook Extension

We implement a notebook extension that integrates PI2 with Jupyter Lab (Figure 1).

Extension Workflow:

1. Select queries for including in interface generation by checking their checkbox
2. Click “Generate Interfaces”
3. “Generated Interfaces” on the right will now display current interface
 - a. Version tabs along the top allow users to revisit previous interfaces
 - b. Input query logs retained for reproducibility

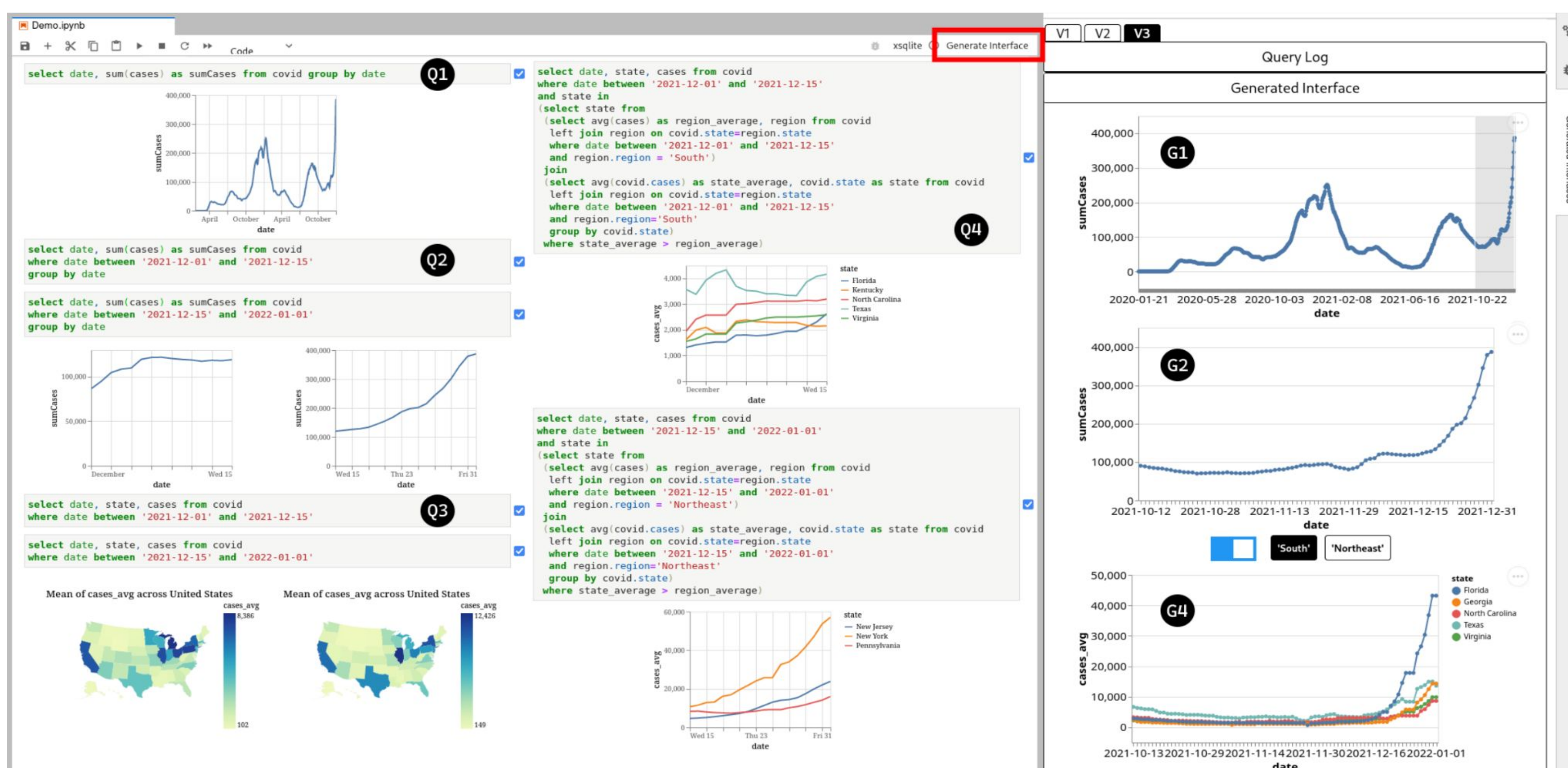


Figure 1. An example workflow for analysis utilizing PI2. SQL notebook analysis on left is used as input. Generated interface on right has multiple visualizations, linked visualization interactions, and query configuration widgets. The user can brush over the top graph to change the date range of the lower two graphs, and configure the query with the buttons and slider.

Interface Generation

PI2 transforms an input sequence of queries into an interactive interface by searching the space of possible interfaces (Figure 2):

1. Parse input queries into Diftrees
2. Map Diftrees to a candidate interface, accounting for user’s screen size
3. Score interface using a usability cost metric
4. Apply Diftree transformation rules to generate new candidate interfaces and repeat from step 2. Transformations preserve or increase the expressiveness of the Diftrees, allowing PI2 to generalize from the input queries.

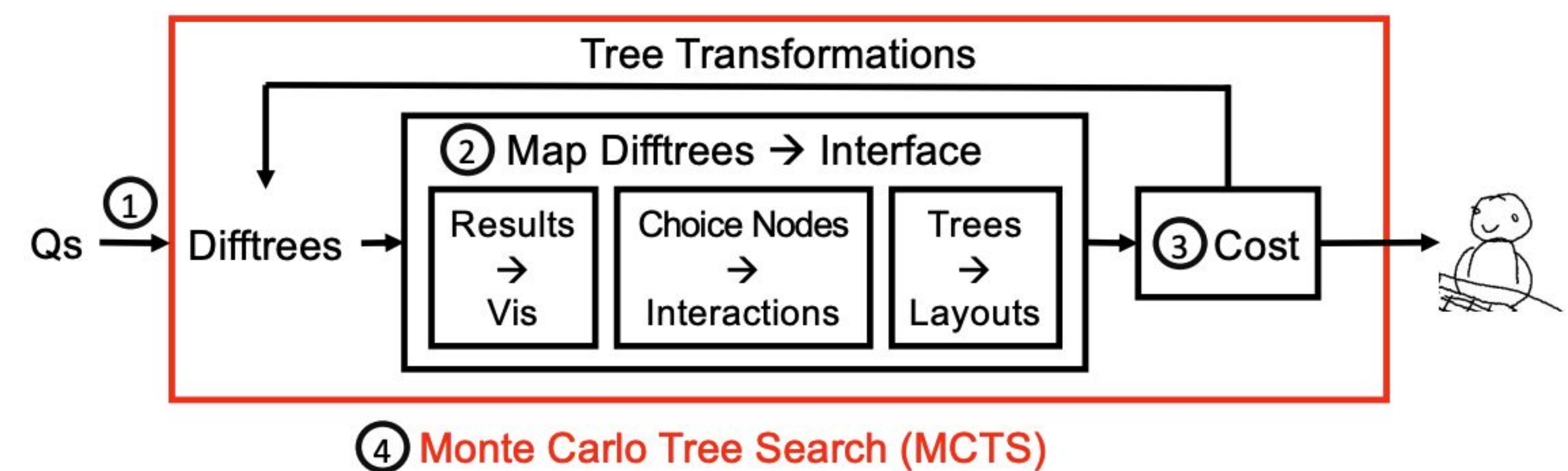


Figure 2. PI2 interface generation pipeline. The search space of possible Diftrees and interfaces is large, so we adopt Monte Carlo Tree Search, which balances exploration of new states with exploitation of known good states.

- Interface mapping uses the schema of the Diftree to choose appropriate visualizations and maps choice nodes to widgets and visualization interactions
- Interactions with widgets and visualizations change the values of choice nodes, transforming the underlying query.

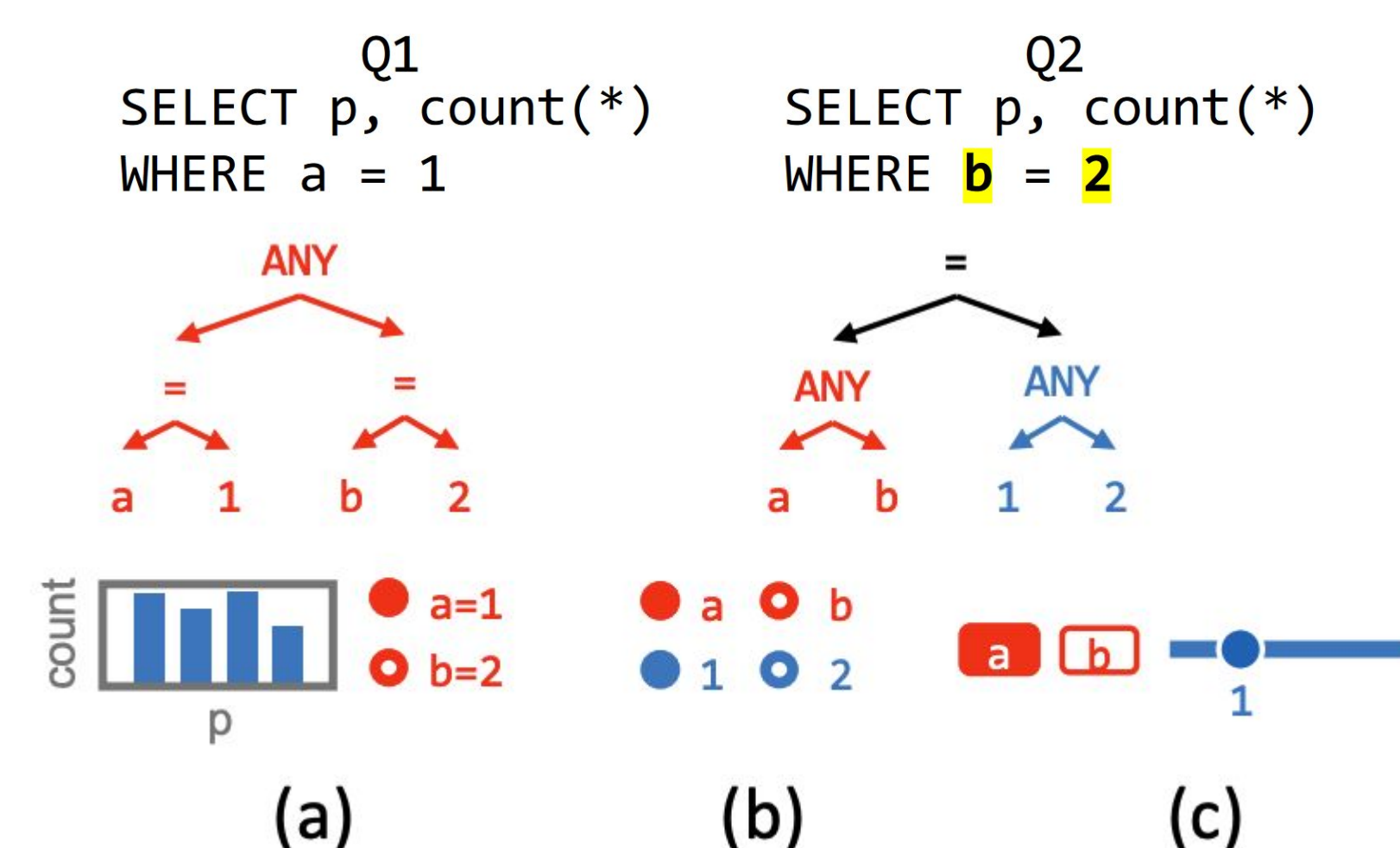


Figure 3. Example queries and their corresponding Diftrees and interfaces. ANY is an example of a choice node that can select any one of its children. Different Diftree structures give rise to different interface mappings.

References

- Yiru Chen and Eugene Wu. 2021. PI2: Generating Visual Analysis Interfaces From Queries. arXiv:2107.08203 [cs.DB]
- Krzysztof Gajos and Daniel S Weld. 2004. SUPPLE: automatically generating user interfaces. In Proceedings of the 9th international conference on Intelligent user interfaces. 93–100.