



Analyze Function

Similar Data Search

Similar Data Search

1. 데이터 검색 환경 세팅

The screenshot displays two side-by-side views of the Similar Data Search interface for the 'Berkely Deep Drive' dataset.

Left View (Display View):

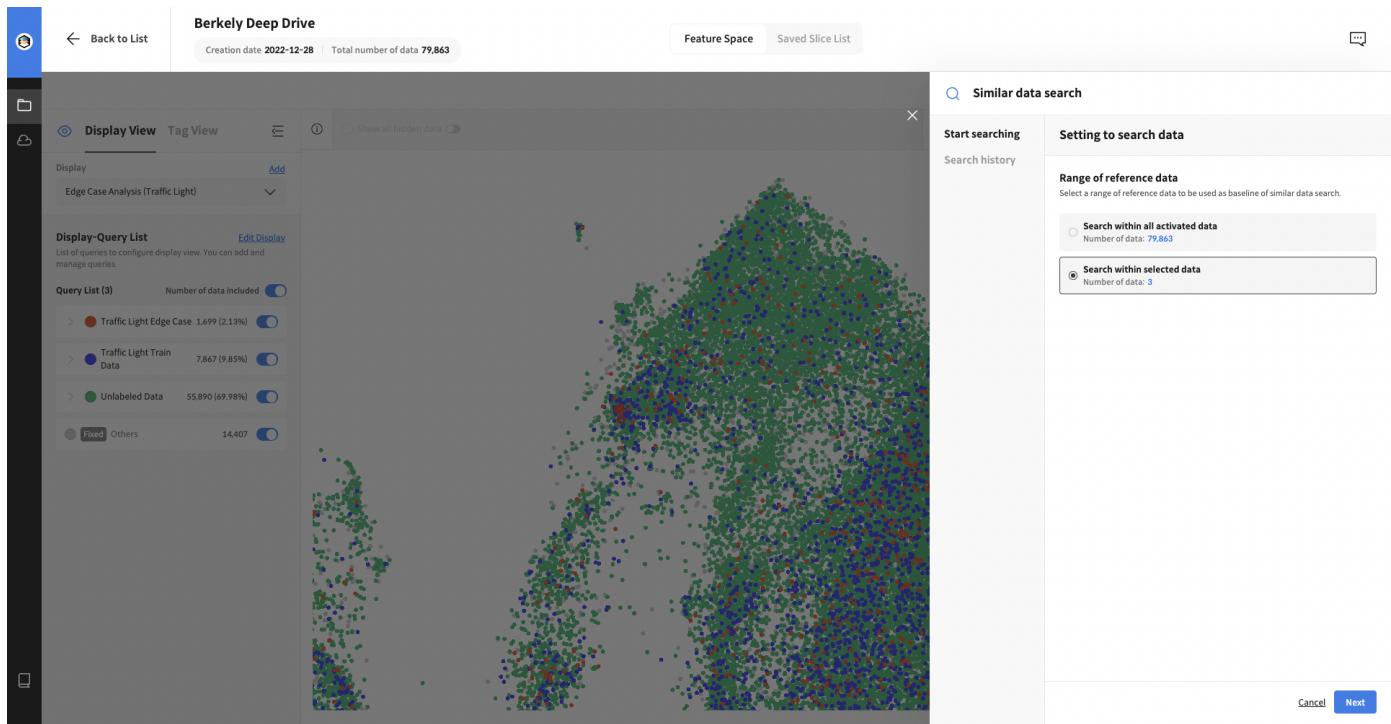
- Header:** Berkely Deep Drive, Creation date 2022-12-28, Total number of data 79,863.
- Toolbars:** Feature Space, Saved Slice List.
- Display Options:** Display View (selected), Tag View, Show all hidden data.
- Dataset Data List:** Shows 3 selected data items, with options to show hidden data or selected data only.
- Display-Query List:** Shows three categories: Traffic Light Edge Case (1,699, 2.13%), Traffic Light Train (7,867, 9.85%), and Unlabeled Data (55,890, 69.98%).
- Scatter Plot:** A large scatter plot showing data points, with a prominent cluster of pink points representing the selected data.

Right View (Tag View):

- Header:** Berkely Deep Drive, Creation date 2022-12-28, Total number of data 79,863.
- Toolbars:** Feature Space, Saved Slice List.
- Display Options:** Display View, Tag View (selected), Show all hidden data.
- Dataset Data List:** Shows 0 selected data items, with options to show hidden data or selected data only.
- Tag Options:** Shows one added tag: Case1 (3 data points).
- Scatter Plot:** A large scatter plot showing data points, with a few pink points representing the tagged data.

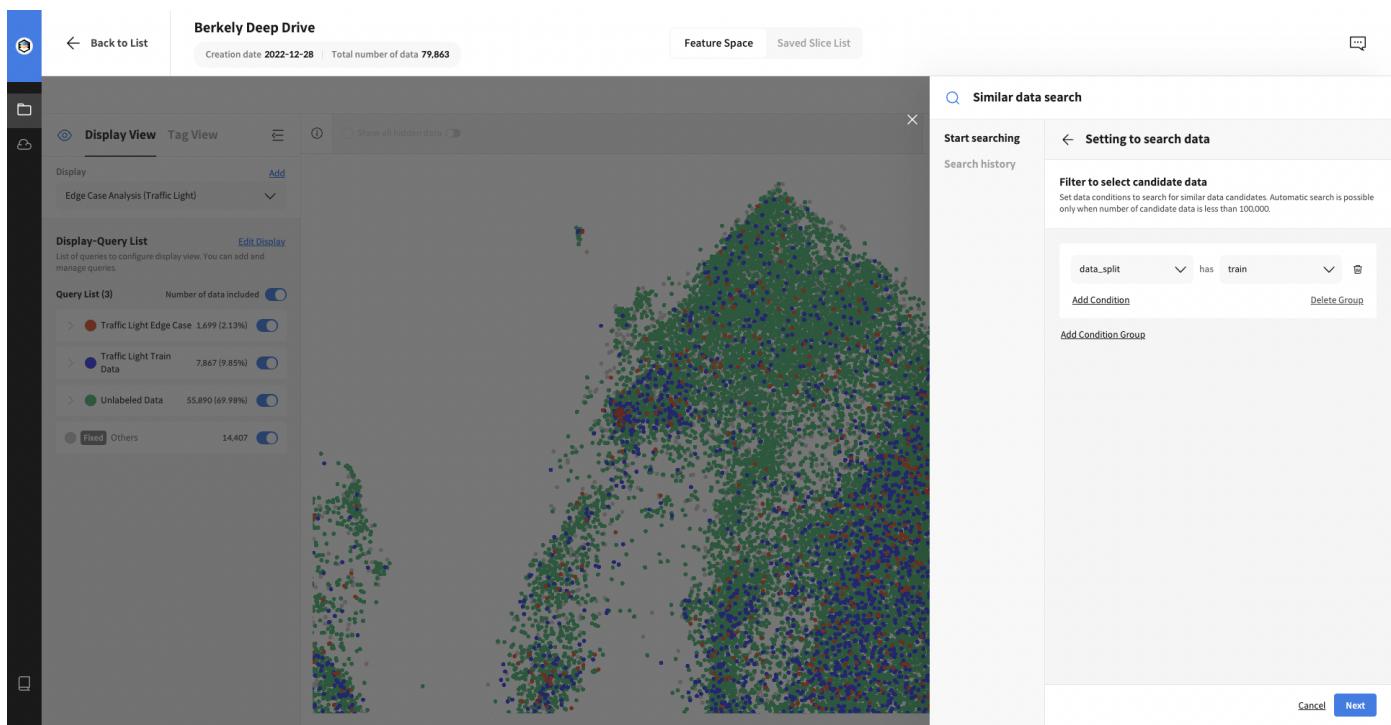
검색 대상이 되는 Reference Data(기준 데이터)를 선택하거나, Reference Data가 지정된 태그 뷰 화면에서

우측 상단 **`Similar data search`** 버튼을 클릭해 주세요.



하단 **Search within Selected data** 항목을 선택하면 지정한 데이터와 유사한 데이터를 검색합니다(기본).

`Search within all activated data` 항목을 선택하면 전체 데이터 특성이 종합적으로 반영된 데이터를 검색합니다.



Candidate data filter 를 설정해 검색 범위가 되는 데이터 군을 지정할 수 있습니다.

해당 작업을 건너 뛰면 자동으로 Reference Data 제외한 전체 데이터에서 유사 데이터를 검색합니다.

The screenshot shows the 'Similar data search' interface. On the left, there's a 'Tag' section with a 'Case1' entry. The main area features a large, grayscale heatmap visualization. On the right, there are two tabs: 'Start searching' and 'Setting to search data'. Under 'Setting to search data', it says 'Calculation is completed!' and provides statistics: Number of reference data (3), Number of candidate data (13,973), and Number of available candidate data (13,972). At the bottom right, there are 'Cancel' and 'Start similar data search' buttons.

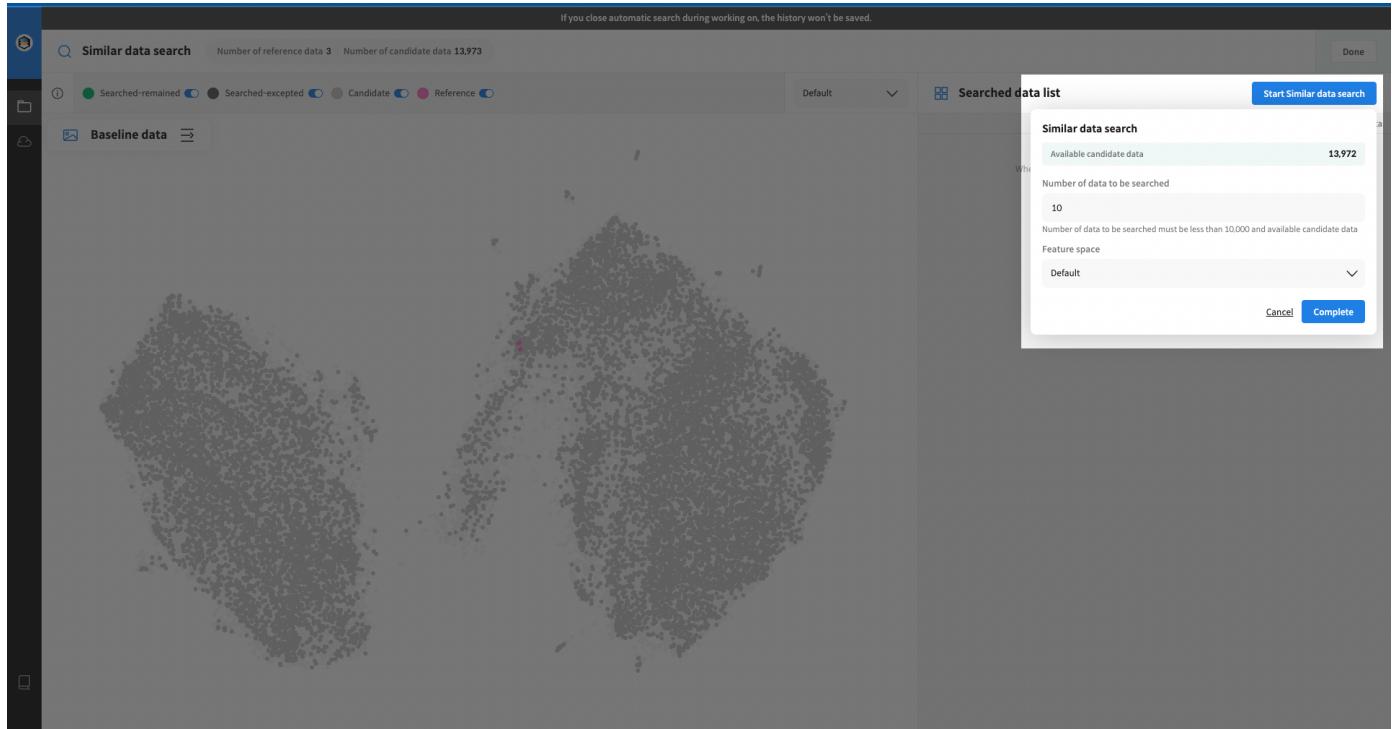
Reference Data와 Candidate Data 설정을 검증한 다음 **Start similar data search** 버튼을 누르면 검색 환경 세팅이 완료됩니다.

2. 유사 데이터 검색

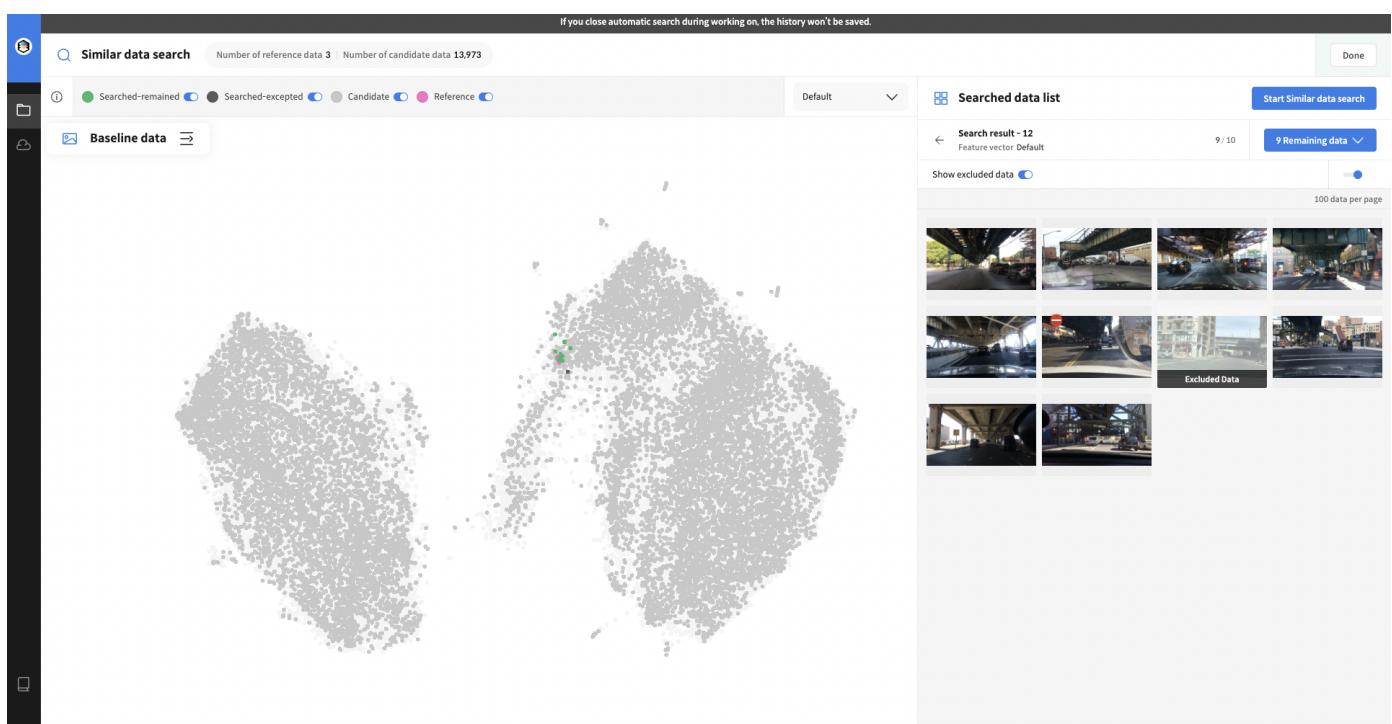
The screenshot shows the 'Similar data search' interface. On the left, there's a 'Tag' section with a 'Case1' entry. The main area features a large, grayscale heatmap visualization. On the right, the 'Search history' tab is selected, displaying a list of previous searches:

Reference	Candidate	Date	Action
Reference 3	13,973	2023-01-26 11:05:35	View search
Reference 1	29,393	2023-01-26 07:34:51	View search
Reference 1	79,863	2023-01-25 18:11:21	View search
Reference 1	55,890	2023-01-25 18:02:45	View search
Reference 1	13,973	2023-01-25 18:02:31	View search
Reference 1	79,863	2023-01-25 14:42:57	View search
Reference 1	55,890	2023-01-25 13:23:00	View search
Reference 1	55,890	2023-01-24 21:47:16	View search

Search History에서 검색 환경으로 진입해주세요.

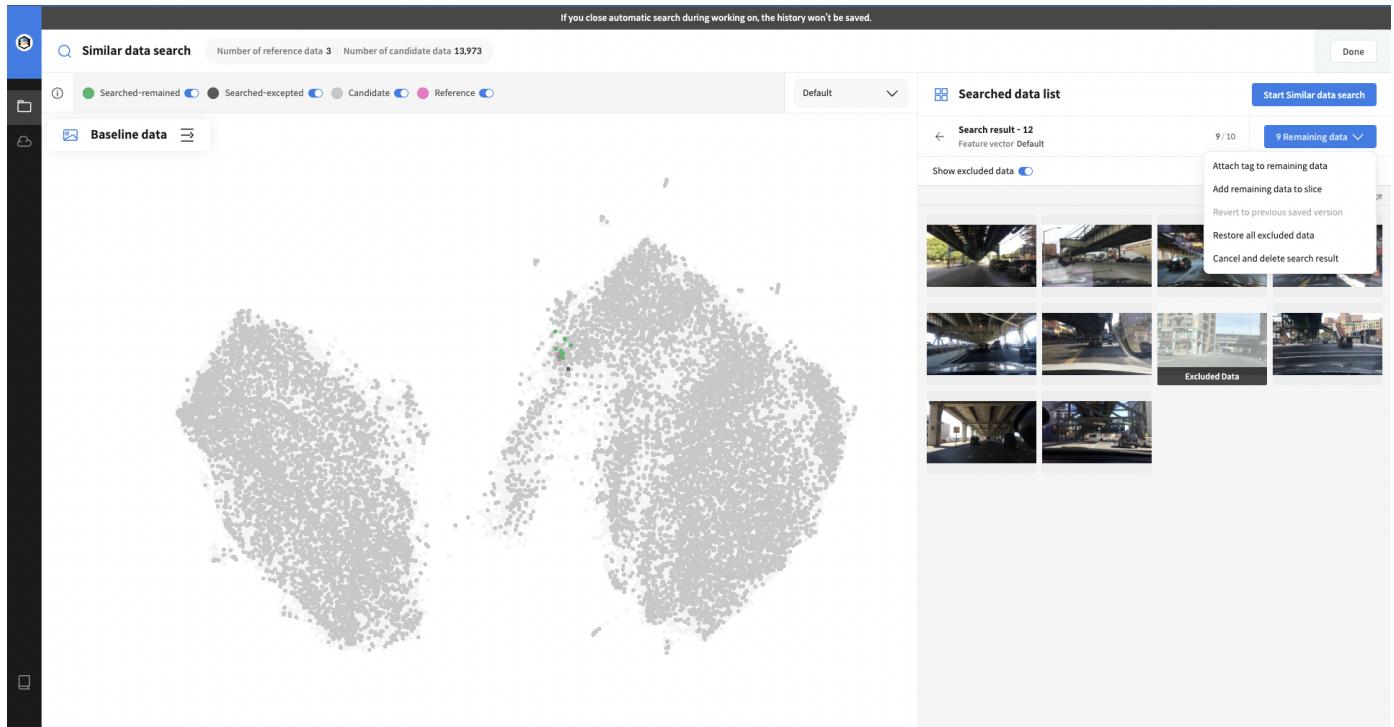


Start similar data search 버튼을 통해 원하는 검색 결과 수량과, 검색 기반이 되는 피쳐 스페이스를 선택해주세요.



검색이 완료되면 의도와 맞지 않았던 검색 결과를 제외하고 추가 검색을 진행 할 수 있습니다.

검색 결과 내에서 선택한 보정 데이터를 활용해 다음 검색에서 더 좋은 검색 결과를 제공합니다.



각 검색 결과에 데이터마다 Tag를 붙이거나, Slice에 추가할 수 있습니다.

 [Edit this page](#)