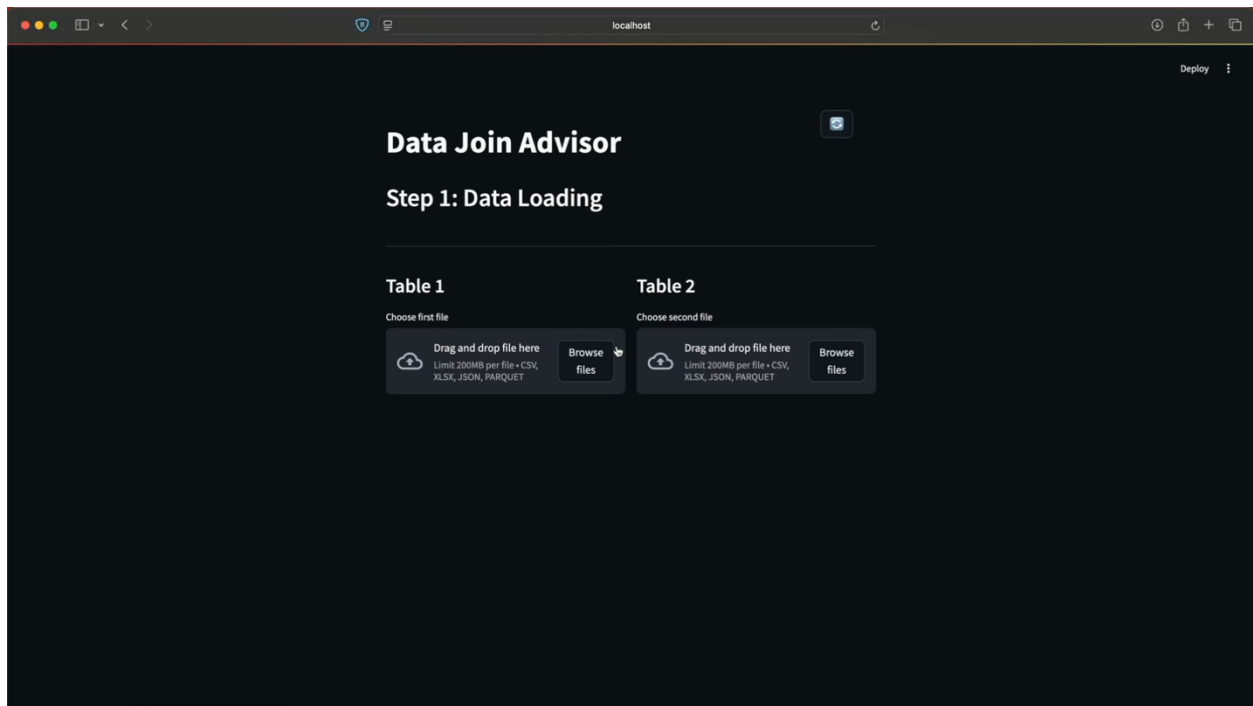


Join Agent Tutorial

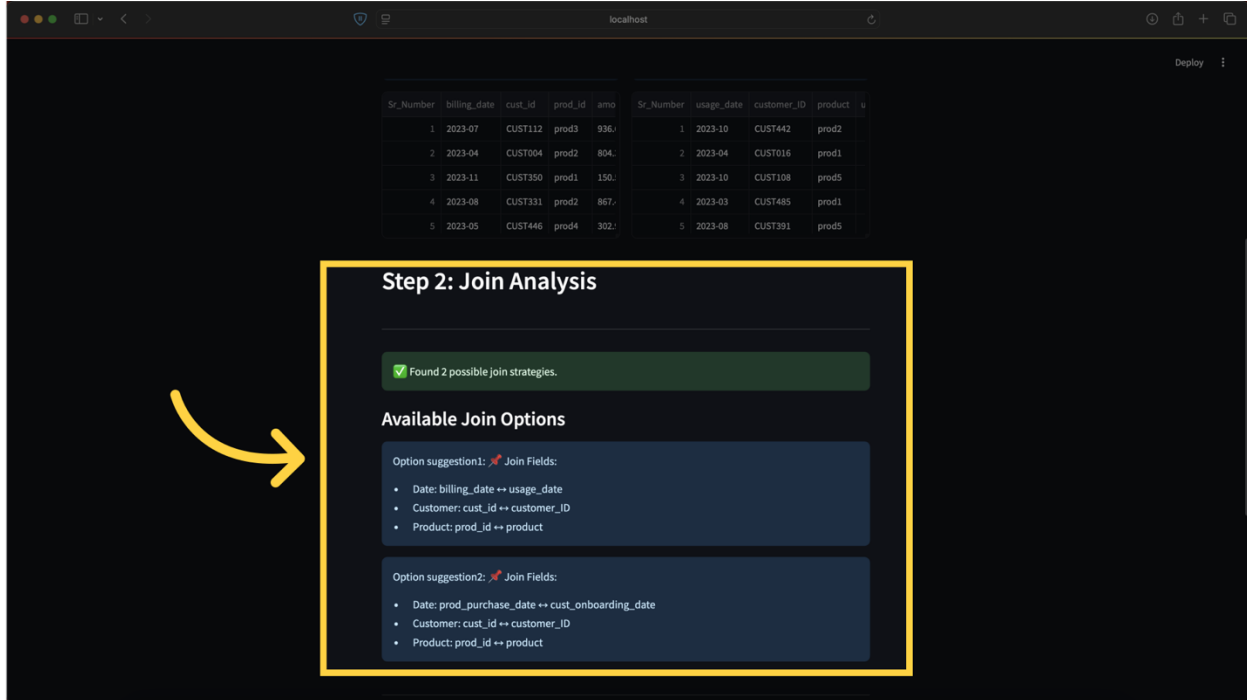
1. Introduction

The Join agent will take two datasets and merge them together in a no-code, end to end fashion, assisting you in your data unification efforts. Upload your two datasets to kick off the end-to-end process.



2. Join Analysis

As you are confirming the uploaded files, the Join agent generates strategies that can be used to merge your two files together. Keep these in mind as we proceed to the next step.



The screenshot shows a web application interface with a dark theme. At the top, there's a browser window header with 'localhost' and a 'Deploy' button. Below this, two data tables are displayed side-by-side. The left table has columns: Sr_Number, billing_date, cust_id, prod_id, amo. The right table has columns: Sr_Number, usage_date, customer_ID, product, u. A yellow arrow points from the left towards the 'Step 2: Join Analysis' panel, which is highlighted with a yellow border. This panel contains a green status bar indicating 'Found 2 possible join strategies.' and a section titled 'Available Join Options' with two suggestions, each listing common fields for joining.

Sr_Number	billing_date	cust_id	prod_id	amo
1	2023-07	CUST112	prod3	936
2	2023-04	CUST004	prod2	804
3	2023-11	CUST350	prod1	150
4	2023-08	CUST331	prod2	867
5	2023-05	CUST446	prod4	302

Sr_Number	usage_date	customer_ID	product	u
1	2023-10	CUST442	prod2	
2	2023-04	CUST016	prod1	
3	2023-10	CUST108	prod5	
4	2023-03	CUST485	prod1	
5	2023-08	CUST391	prod5	

Step 2: Join Analysis

✓ Found 2 possible join strategies.

Available Join Options

Option suggestion1: ✖ Join Fields:

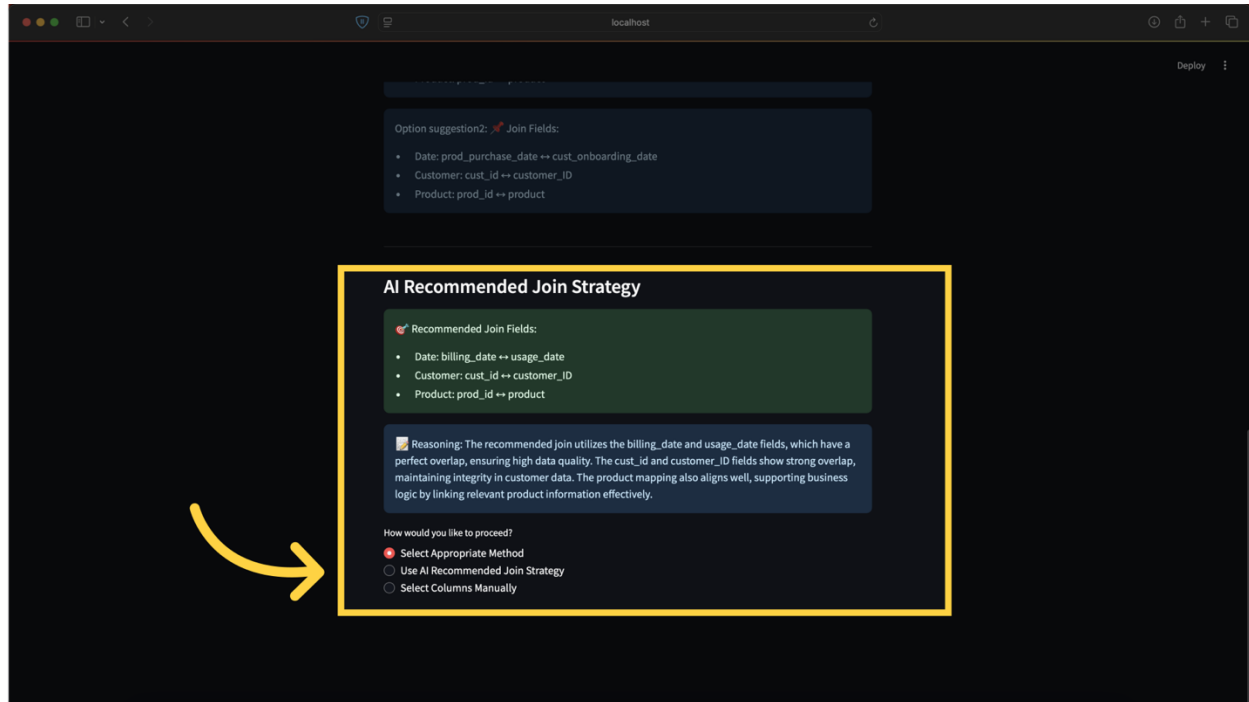
- Date: billing_date ↔ usage_date
- Customer: cust_id ↔ customer_ID
- Product: prod_id ↔ product

Option suggestion2: ✖ Join Fields:

- Date: prod_purchase_date ↔ cust_onboarding_date
- Customer: cust_id ↔ customer_ID
- Product: prod_id ↔ product

3. Choosing a Join Strategy

The Join agent will recommend a join strategy out of the list it generated and provide reasons for its selection. If you agree with the strategy and reasoning, you can click the respective button to use the generated strategy. Otherwise, we can click the button below to manually select the join fields.



4. Creating a Manual Join Strategy

If you wish to use a custom join strategy, input the respective fields in the dropdowns.

The screenshot shows a web application interface for creating a manual join strategy. The interface is dark-themed and includes a browser window with the address bar showing 'localhost'. The main content area displays a join strategy configuration for 'Product: prod_id ↔ product'. A reasoning box explains the recommended join strategy. Below this, a section titled 'How would you like to proceed?' offers two options: 'Select Appropriate Method' and 'Select Columns Manually'. The 'Select Columns Manually' option is selected, and a yellow arrow points to the 'Manual Column Mapping' section. This section contains two columns of dropdowns for mapping fields from Table 1 and Table 2. The 'Customer ID Column (Required)' and 'Date Column (Required)' fields are highlighted with a yellow box, and a warning message at the bottom states: 'Please select required mapping fields (Customer ID and Date)'.

Product: prod_id ↔ product

Reasoning: The recommended join utilizes the billing_date and usage_date fields, which have a perfect overlap, ensuring high data quality. The cust_id and customer_id fields show strong overlap, maintaining integrity in customer data. The product mapping also aligns well, supporting business logic by linking relevant product information effectively.

How would you like to proceed?

☐ Select Appropriate Method

☒ Select Columns Manually

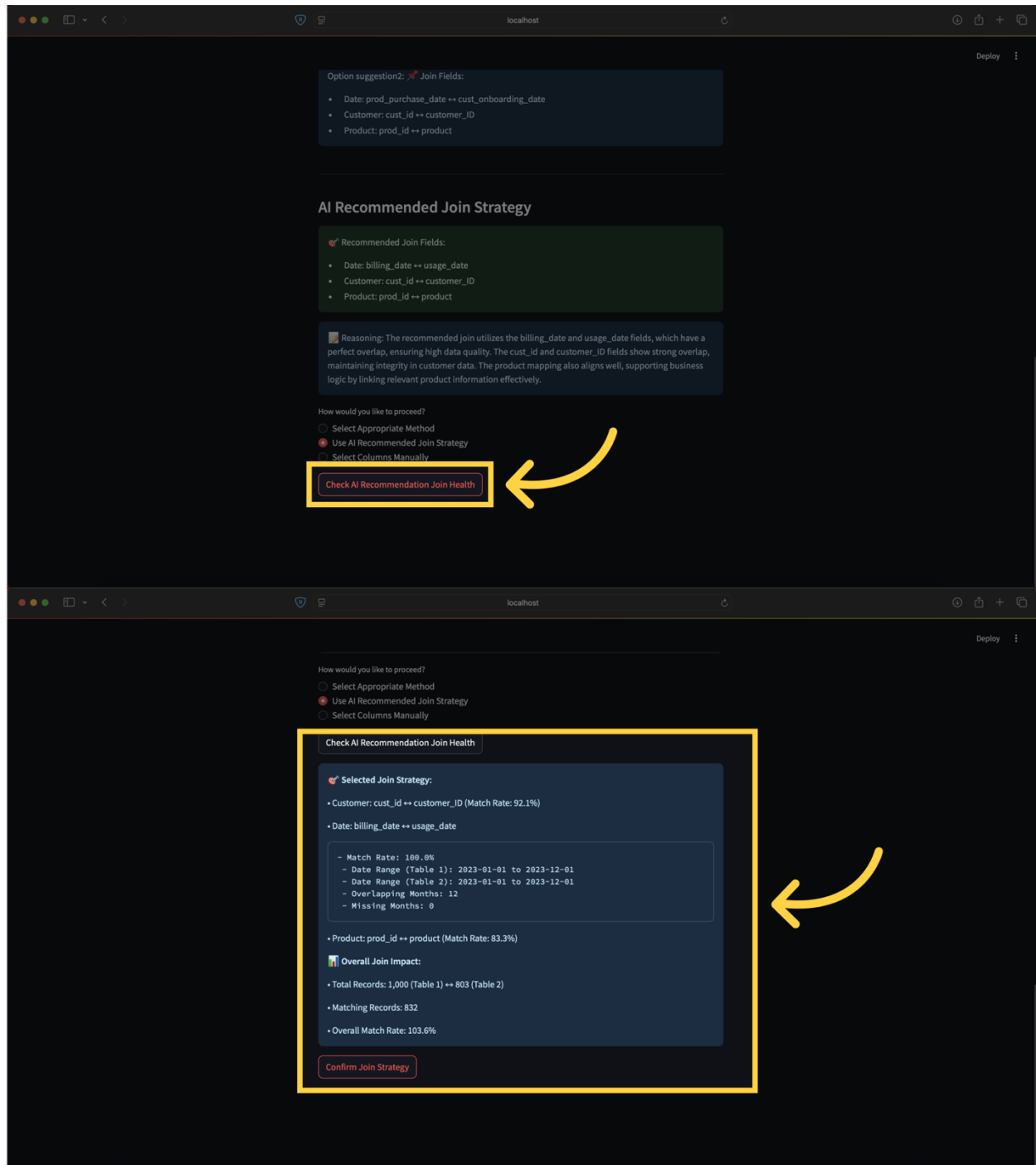
Manual Column Mapping

Table 1 Columns	Table 2 Columns
Customer ID Column (Required)	Customer ID Column (Required)
Date Column (Required)	Date Column (Required)
Product Column (Optional)	Product Column (Optional)

Please select required mapping fields (Customer ID and Date)

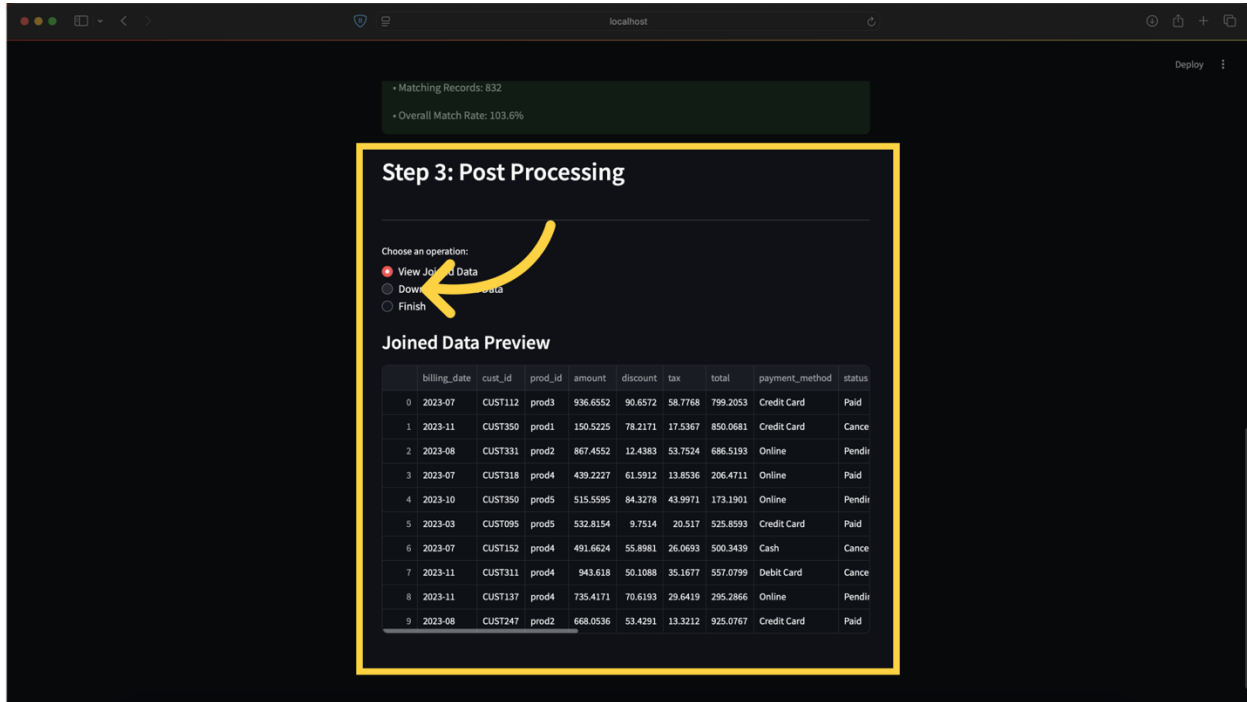
5. Join Strategy Confirmation and Join Health Checks

After confirming either the AI recommended or manual join strategy, the Join agent will provide you with join health statistics



6. Final

Upon confirmation, you will get a preview of your joined data and can download it.



The screenshot shows a web application interface with a dark theme. At the top, there's a status bar with 'localhost' and a 'Deploy' button. Below this, a green box displays 'Matching Records: 832' and 'Overall Match Rate: 103.6%'. The main section is titled 'Step 3: Post Processing'. Under the heading 'Choose an operation:', there are three radio buttons: 'View Joined Data' (selected), 'Download Data', and 'Finish'. A yellow arrow points to the 'View Joined Data' button. Below this, a 'Joined Data Preview' table is shown, displaying 10 rows of data with columns: billing_date, cust_id, prod_id, amount, discount, tax, total, payment_method, and status.

	billing_date	cust_id	prod_id	amount	discount	tax	total	payment_method	status
0	2023-07	CUST112	prod3	936.6552	90.6572	58.7768	799.2053	Credit Card	Paid
1	2023-11	CUST350	prod1	150.5225	78.2171	17.5367	850.0681	Credit Card	Cancel
2	2023-08	CUST331	prod2	867.4552	12.4383	53.7524	686.5193	Online	Pending
3	2023-07	CUST318	prod4	439.2227	61.5912	13.8536	206.4711	Online	Paid
4	2023-10	CUST350	prod5	515.5595	84.3278	43.9971	173.1901	Online	Pending
5	2023-03	CUST095	prod5	532.8154	9.7514	20.517	525.8593	Credit Card	Paid
6	2023-07	CUST152	prod4	491.6624	55.8981	26.0693	500.3439	Cash	Cancel
7	2023-11	CUST311	prod4	943.618	50.1088	35.1677	557.0799	Debit Card	Cancel
8	2023-11	CUST137	prod4	735.4171	70.6193	29.6419	295.2866	Online	Pending
9	2023-08	CUST247	prod2	668.0536	53.4291	13.3212	925.0767	Credit Card	Paid