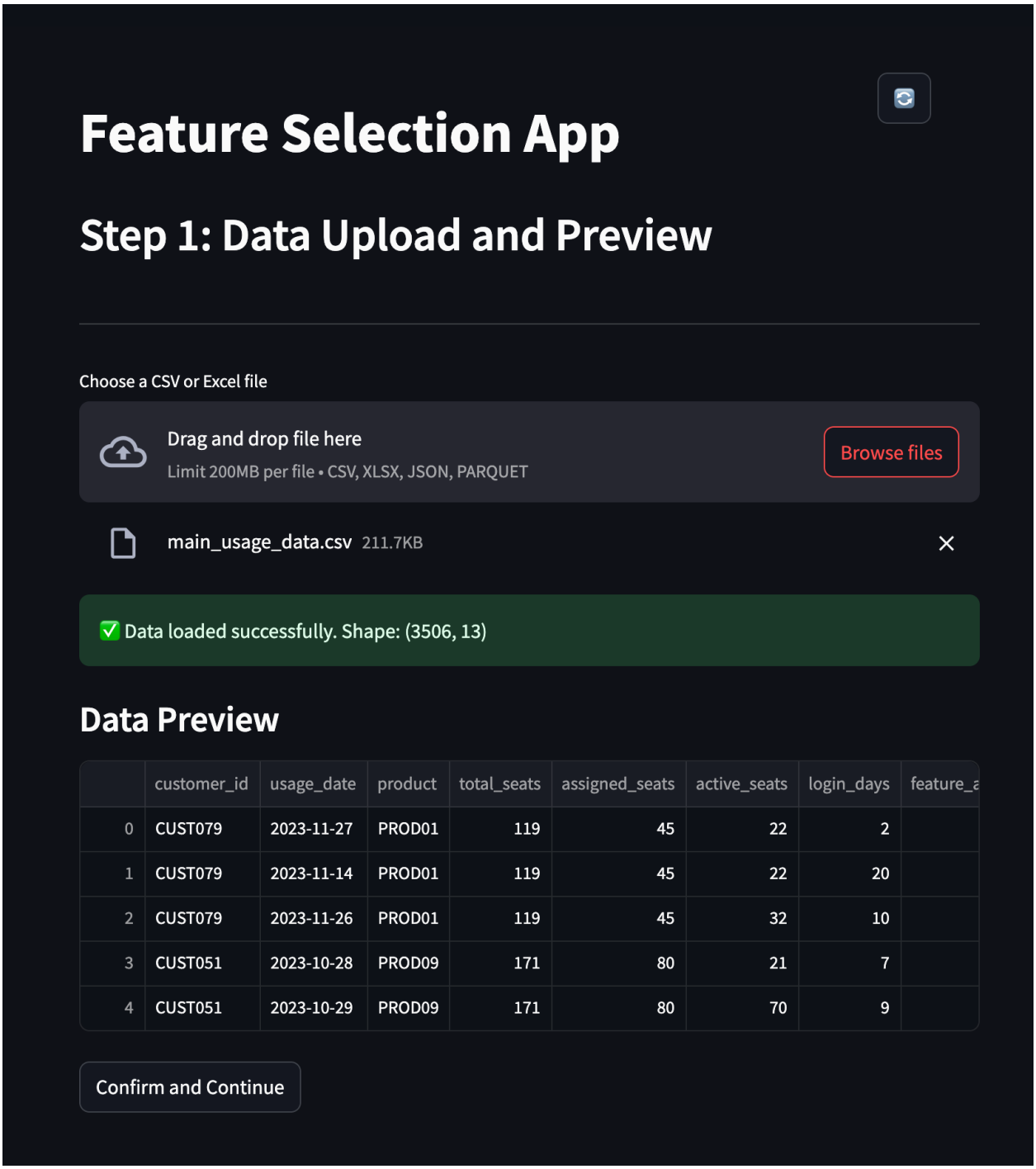
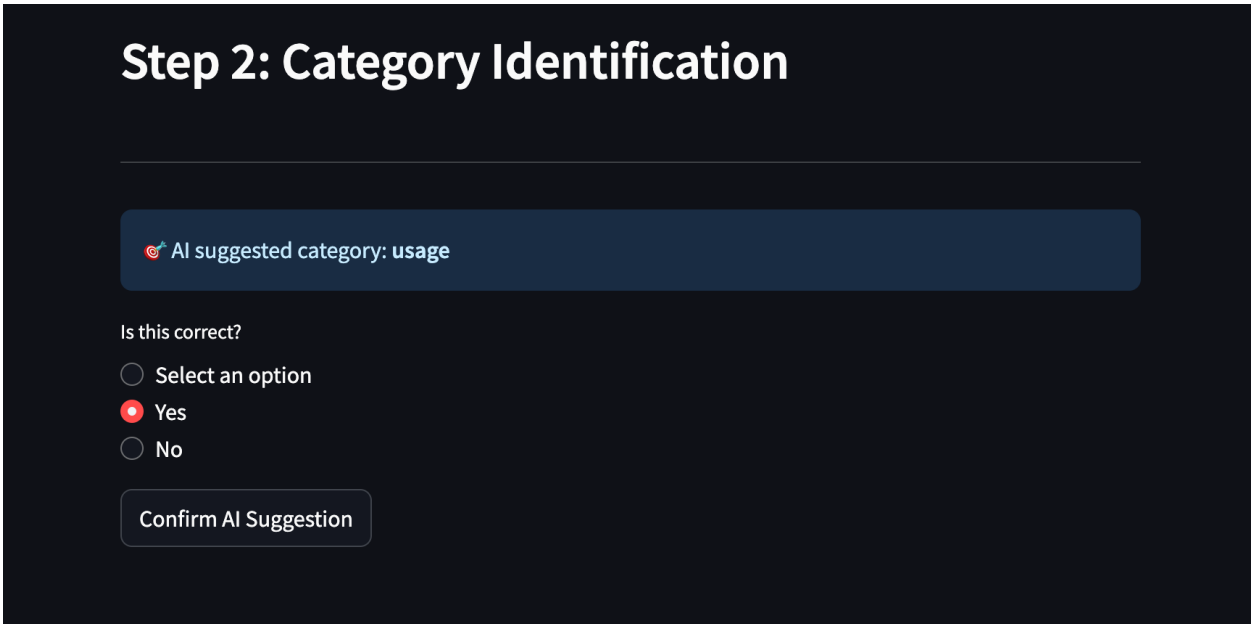
Feature Agent Tutorial

1. Today we will walk you through StepFunction's Feature Agent. This agent will create tuned inputs for AI models from your data. Start by uploading the data you wish to transform.



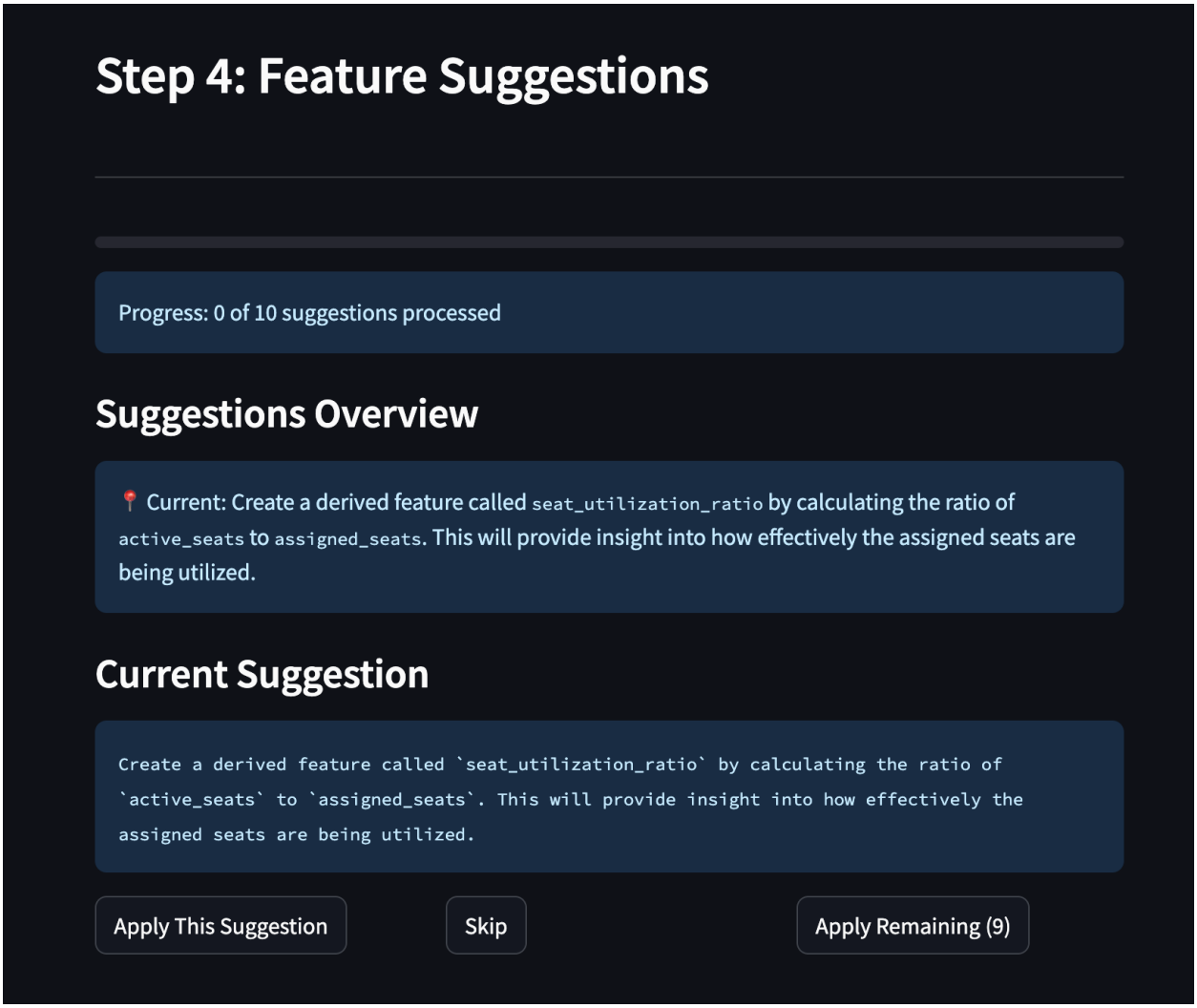
	customer_id	usage_date	product	total_seats	assigned_seats	active_seats	login_days	feature_a
0	CUST079	2023-11-27	PROD01	119	45	22	2	
1	CUST079	2023-11-14	PROD01	119	45	22	20	
2	CUST079	2023-11-26	PROD01	119	45	32	10	
3	CUST051	2023-10-28	PROD09	171	80	21	7	
4	CUST051	2023-10-29	PROD09	171	80	70	9	

2. Next, the Feature Agent will attempt to identify a category to describe your data. The reason it does this is because it will apply category-specific transformations to your data. If your category differs from the one the Agent identified, you can manually select the category of data you have uploaded.

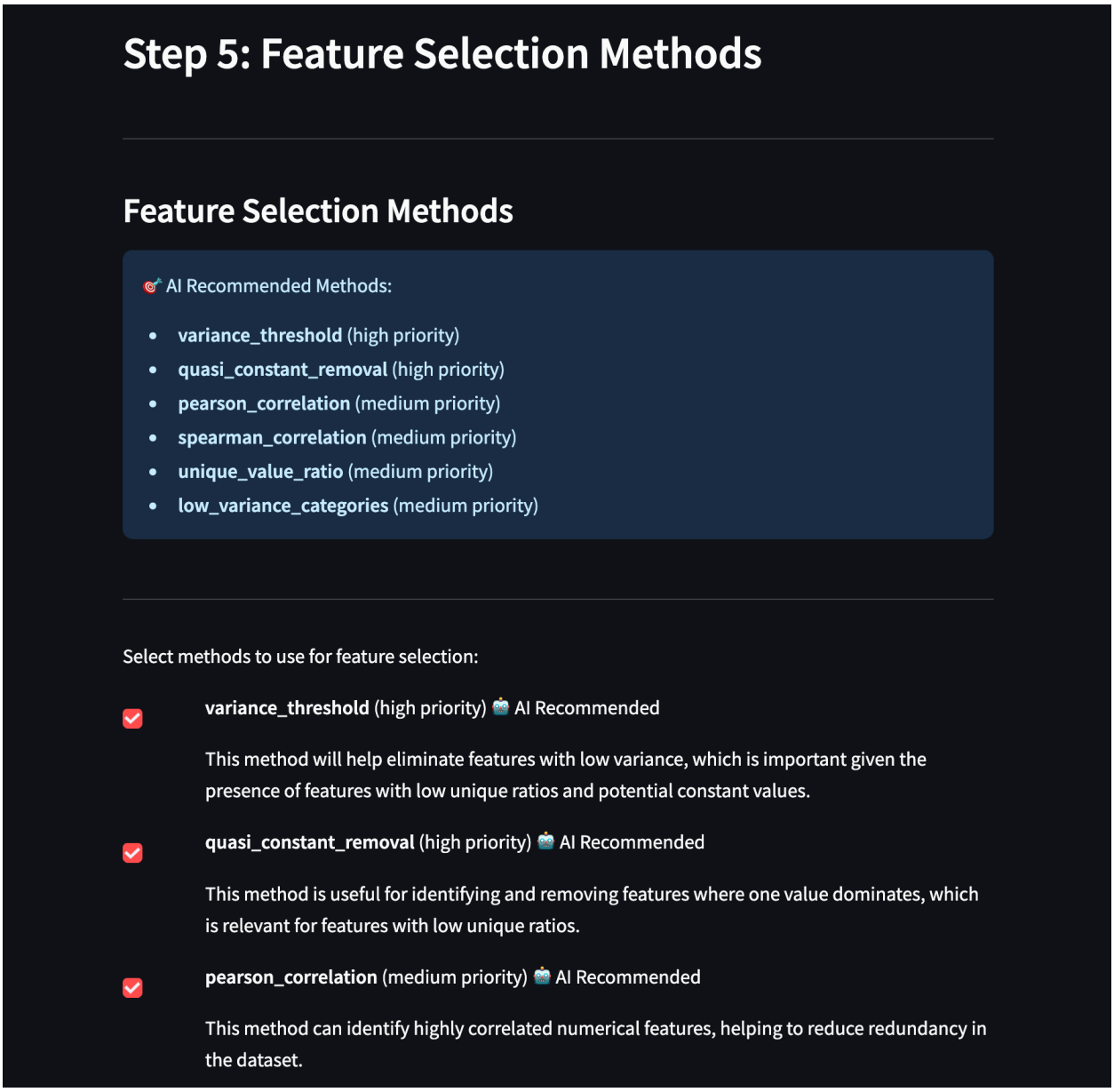


3. As a continuation of the prior step, the Feature Agent will suggest some mapping fields to help enforce which columns will and will not be used for transformations in the next step. If you wish to modify them, click the manual selection button and map them via a dropdown.

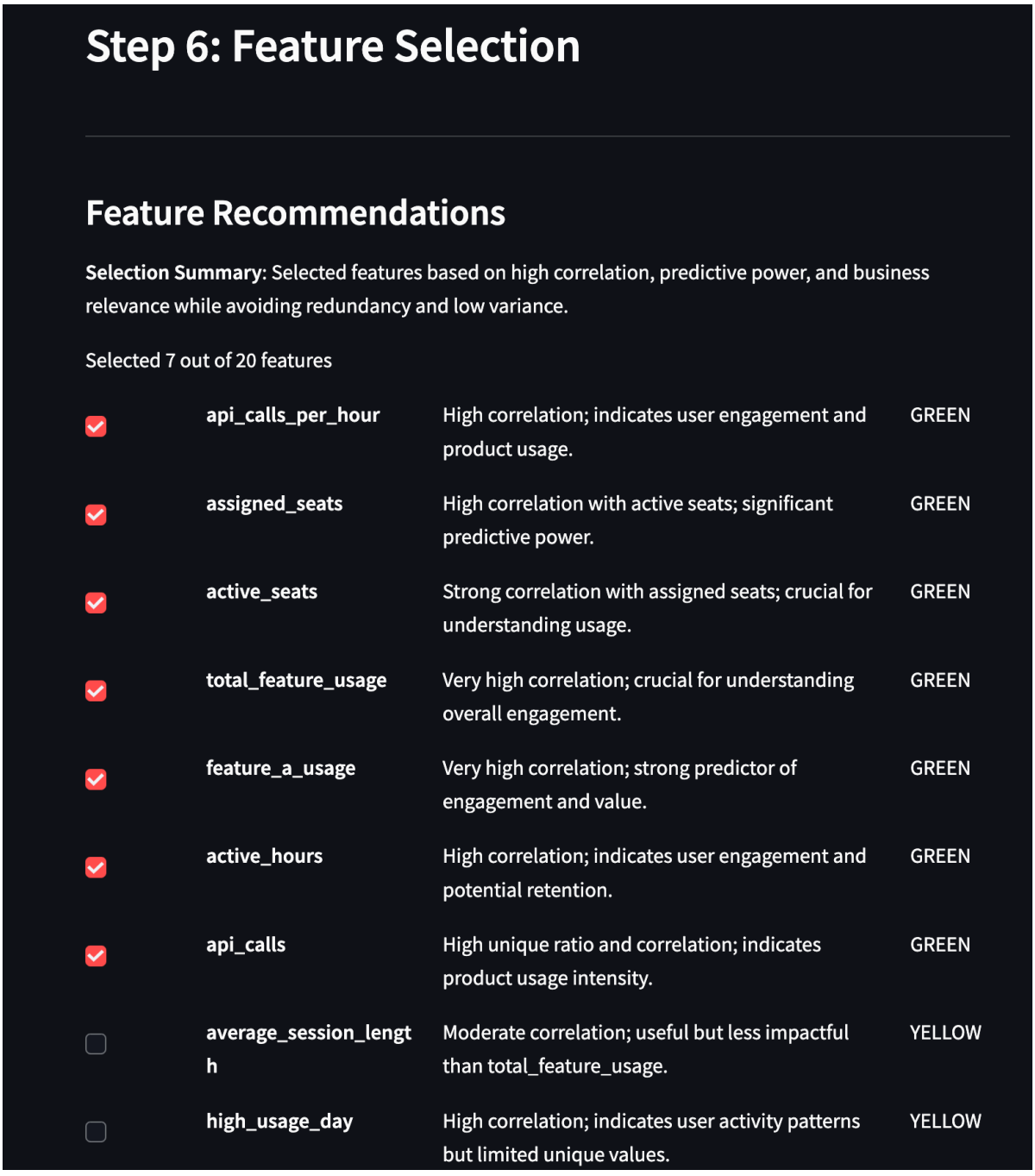
4. After categorization and field mapping, we will move onto the individual feature suggestions. The Agent will provide you with different transformations it will apply to as a final prep before feature creation. You can review each suggestion and its description one-by-one, or if you don't want to apply a given suggestion, you can always skip it. Finally, you can apply all suggestions without reviewing via the bottom right button.



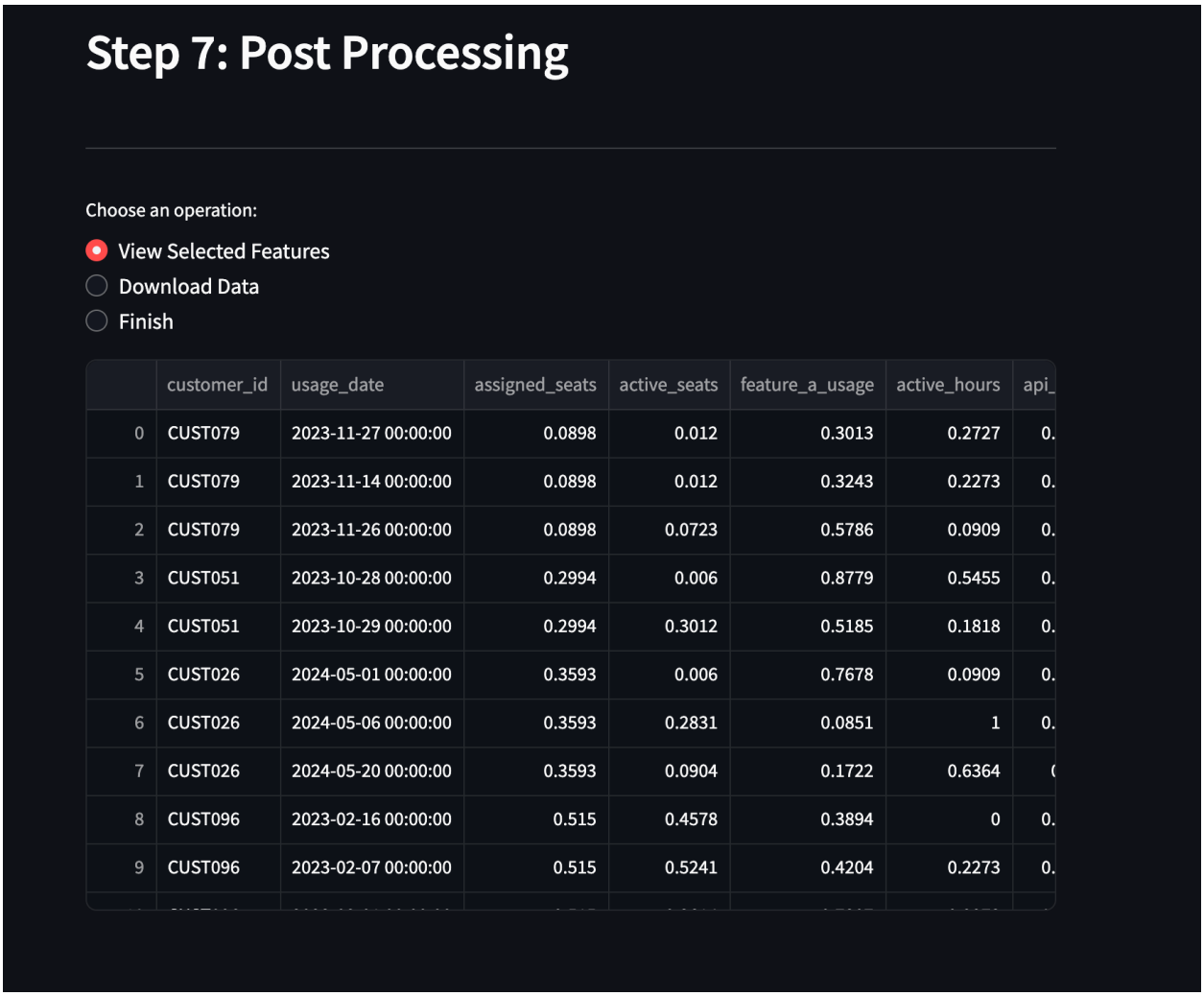
5. After creating features, the Agent will apply statistical transformations on them to help normalize the created features and remove the features that are not as representative of the data. You will be presented with multiple suggested statistical methods to follow. If you want to overrule them, you can manually select and deselect them via the checkboxes.



6. The Agent will now present you with the features it created, and it will recommend which ones you should choose. The features are organized by name, what they represent, and their viability for use in AI models. By default, the Agent will recommend features with the Green designation (high benefit for models). Features with the Yellow designation indicate that they may have some viability while the Red designation indicates that the feature is not recommended for use. You can always modify the list based on your preferences via the checkbexes.



7. After selecting the features you want to add, the Agent will then apply them onto the data. Download it, and your data will now be ready for use in AI models.



	customer_id	usage_date	assigned_seats	active_seats	feature_a_usage	active_hours	api_
0	CUST079	2023-11-27 00:00:00	0.0898	0.012	0.3013	0.2727	0.
1	CUST079	2023-11-14 00:00:00	0.0898	0.012	0.3243	0.2273	0.
2	CUST079	2023-11-26 00:00:00	0.0898	0.0723	0.5786	0.0909	0.
3	CUST051	2023-10-28 00:00:00	0.2994	0.006	0.8779	0.5455	0.
4	CUST051	2023-10-29 00:00:00	0.2994	0.3012	0.5185	0.1818	0.
5	CUST026	2024-05-01 00:00:00	0.3593	0.006	0.7678	0.0909	0.
6	CUST026	2024-05-06 00:00:00	0.3593	0.2831	0.0851	1	0.
7	CUST026	2024-05-20 00:00:00	0.3593	0.0904	0.1722	0.6364	(
8	CUST096	2023-02-16 00:00:00	0.515	0.4578	0.3894	0	0.
9	CUST096	2023-02-07 00:00:00	0.515	0.5241	0.4204	0.2273	0.