Running Job Basics

Contents:

Configure Job

When you are ready to test your recipe against the entire dataset, click **Run** in the Transformer page. In the Run Job page, you specify the output formats and any compression to apply. Unless you are working with a large dataset, compression is unneeded for this basic walkthrough.

> **Tip:** Optionally, you can disable generating a visual profile of your results. While the visual profile is very useful for examining issues in your recipe and iterating, it is a resource-intensive process. If you are working with large datasets that do not require additional debugging, you can consider disabling the profiling of your results. For more information, see *Overview of Visual Profiling*.

> **Tip:** Depending on your product configuration, you may have multiple running environments available to you. In most cases, you should choose to use the default running environment, which is selected for you based on the size of the dataset.

For more information on the job execution options, see *Run Job Page*.

Run Job

To queue the specified job for execution, click **Run**.

The job is queued up for processing.

You can track progress in the Job Details page.

- If visual profiling was enabled for the job, click the Profile tab.
- When the job is completed, you can access results in the Output destinations tab.
- For more information, see *Job Details Page*.

Iterate

In the Profile tab of the Job Details page, you can review the effects of the transformation recipe across the entire dataset. Statistics and data histograms provide overall visibility into the quality of your transformation recipe.

All data

95% valid values   1% mismatching values   4% missing values

9 columns   38 rows   5 data types

Results profile by column

# FMID	ABC MarketName	Primary_Website_or_URL	ABC Address	ABC Location_Description	[ ] Seasons
Valid 38	Valid 38	Valid 22	Valid 38	Valid 38	Valid 38
Mismatched 0	Mismatched 0	Mismatched 4	Mismatched 0	Mismatched 0	Mismatched 0
Empty 0	Empty 0	Empty 12	Empty 0	Empty 0	Empty 0
	Top 20 values	Top 14 values	Top 20 values	Top 7 values	Top 20 values
	58th and Chester Farmer… 1	http://www.grownyc.org 6	507 Harrison Street,Kal… 1	Other 16	{"Date":"","Time":""} 115
	57th Street Greenmarket 1	http://www.foodtrustmar… 4	3rd St N (Hwy 11) besid… 1	Private business parki… 8	{"Date":"","Time":"Tu… 3
	52nd and Haverford Farm… 1	http://www.sandlercente… 1	3rd & Curry Street,Cars… 1	Local government build… 7	{"Date":"06/17/2013 t… 1
	3rd Street N (hwy 11) b… 1	http://www.peoplesfoodc… 1	3808 North Meridian Str… 1	Faith-based instituti… 3	{"Date":"06/06/2013 t… 1
	3rd & Curry St. Farmers… 1	http://www.pcfma.com/ma… 1	362 Plymouth Street,Abi… 1	Closed-off public stre… 2	{"Date":"06/06/2013 t… 1
	38th & Meridian Farmers… 1	http://www.iatp.org/min… 1	29th and Wharton Street… 1	On a farm from: a bar… 1	{"Date":"06/06/2013 t… 1
	33rd and Diamond Farmer… 1	http://www.foodtrustmar… 1	27 W Potomac Street,Bru… 1	Co-located with wholes… 1	{"Date":"05/11/2013 t… 1
	32nd Street/Waverly Far… 1	http://www.experimental… 1	2349 S Hwy 127,Russell … 1		{"Date":"05/11/2013 t… 1
	3 French Hens French Co… 1	http://www.carsonfarmer… 1	22nd and Tasker Streets… 1		{"Date":"05/04/2013 t… 1
	29th and Wharton Farmer… 1	http://www.abquptowngro… 1	201 Market Street,Virgi… 1		{"Date":"05/01/2013 t… 1
	25th Avenue Farmers' Ma… 1	http://www.abingtonsage… 1	1st Ave - E 92nd & 93 S… 1		{"Date":"04/20/2013 t… 1
	22nd and Tasker Farmers… 1	http://www.abingdonfarm… 1	194 W 25th Avenue,San M… 1		{"Date":"04/20/2013 t… 1
	2012 Wood County Farmer… 1	http://www.6701burnetro… 1	1622 6th St NE,Minneapo… 1		{"Date":"01/01/2013 t… 1
	17th Ave Market 1	http://www.3frenchhensm… 1	1400 U Street NW,Washin… 1		{"Date":"01/01/2013 t… 1
	175th Street Greenmarket 1		12th & Brandywine Stree… 1		

1.000M 002M 1.004M 1.006M .008M

Minimum	1,000,059
Lower quartile	1,001,683
Median	1,004,332
Upper quartile	1,005,310
Maximum	1,008,391

Figure: Visual Profile

See _Job Details Page_.

Use the links in the Job Details page to resume working on your dataset sample and the related recipe, generating jobs when you think you are done, until you have generated the appropriate dataset.