[Excel] – The original data format is changed

Last updated by | Jackie Huang | Jan 4, 2022 at 12:24 AM PST
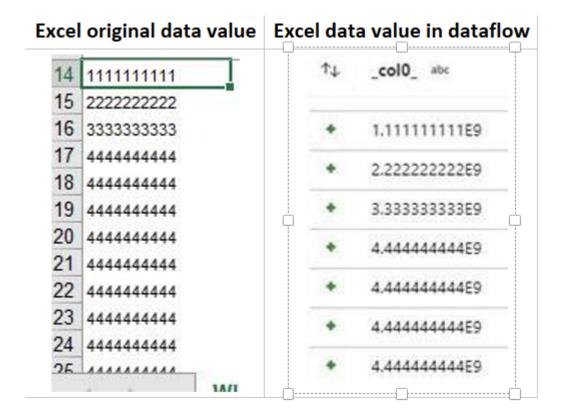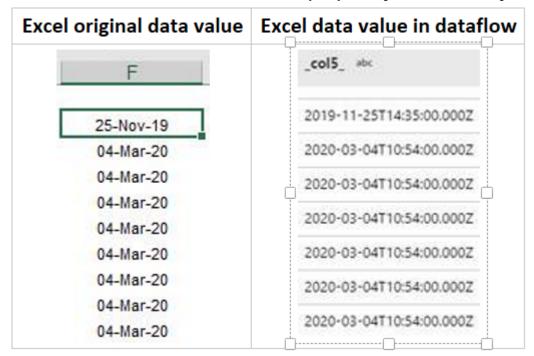

Contents

Issue

When customer uses Excel as source in dataflow (preview data, debug/trigger run...), he may find following problems:

1. The numeric value is parsed to scientific notation and the original value is lost, for example:

2. The timestamp value is parsed to another format, for example:

Excel original data value	Excel data value in dataflow
F	_col5_ abc
25-Nov-19	2019-11-25T14:35:00.000Z
04-Mar-20	2020-03-04T10:54:00.000Z
04-Mar-20	2020-03-04T10:54:00.000Z
04-Mar-20	2020-03-04T10:54:00.000Z
04-Mar-20	2020-03-04T10:54:00.000Z
04-Mar-20	2020-03-04T10:54:00.000Z
04-Mar-20	2020-03-04T10:54:00.000Z
04-Mar-20	
04-Mar-20	

Root Cause

For ADF excel, it does not have physical schema like csv, so we always set its schema as string data type on dataset level. When we start to load and read customer's excel data and convert each cell value into string, we use highest accuracy to avoid data lose.
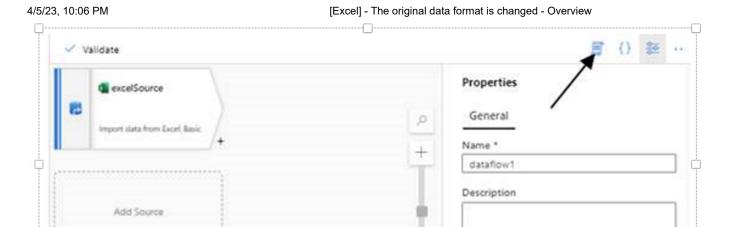
- For numeric cell value, its type is always double, so the value will be expressed scientific notation (cell.
- For timestamp cell value, we always use highest accuracy format (yyyy-MM-dd'T'HH:mm:ss.SSS'Z') to get its st

Resolution

If customer want to get the original value without format change, he can use following ways to achieve it:
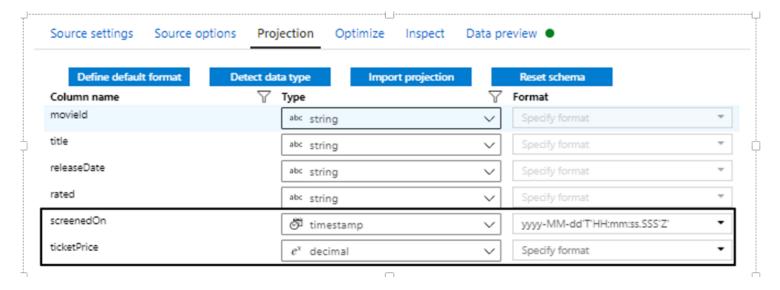
Solution-1: Login ADF portal and switch to your data factory and edit the script by changing related column to your expected data type (Note: it should be spark type in dataflow).

Data flow name   dataflow1

```
1   source(output(
2       movieId as string,
3       title as string,
4       releaseDate as string,
5       rated as string,
6       screenedOn as timestamp 'yyyy-MM-dd\'T\'HH:mm:ss.SSS\'Z\'',
7       ticketPrice as decimal(20,0)
8   ),
9   allowSchemaDrift: true,
10  validateSchema: false) ~> excelSource
```

Solution-2: Login ADF portal and switch to your data factory and change the source column type under "Projection" (Note: related UX support is ongoing now)

Additional Information:

- Icm Reference: N/A
- Author: Zhangyi Yu
- Reviewer: Zhangyi Yu; Shawn Xiao
- Keywords:

How good have you found this content?