

How to Import .CSV UTF8 into Excel

Excel can open a .CSV file, but it defaults to incorrect parameters. MySQL databases are in character set UTF-8. Sometimes data can have line endings embedded in text fields. This will show how to import a .CSV file with these settings.

Step by Step Instructions

1. Start with an empty Excel sheet.
2. Data tab > Get and Transform Data > From Text / CSV > choose the CSV file > Import > Set File Origin to 65001: Unicode (UTF-8) > Transform Data

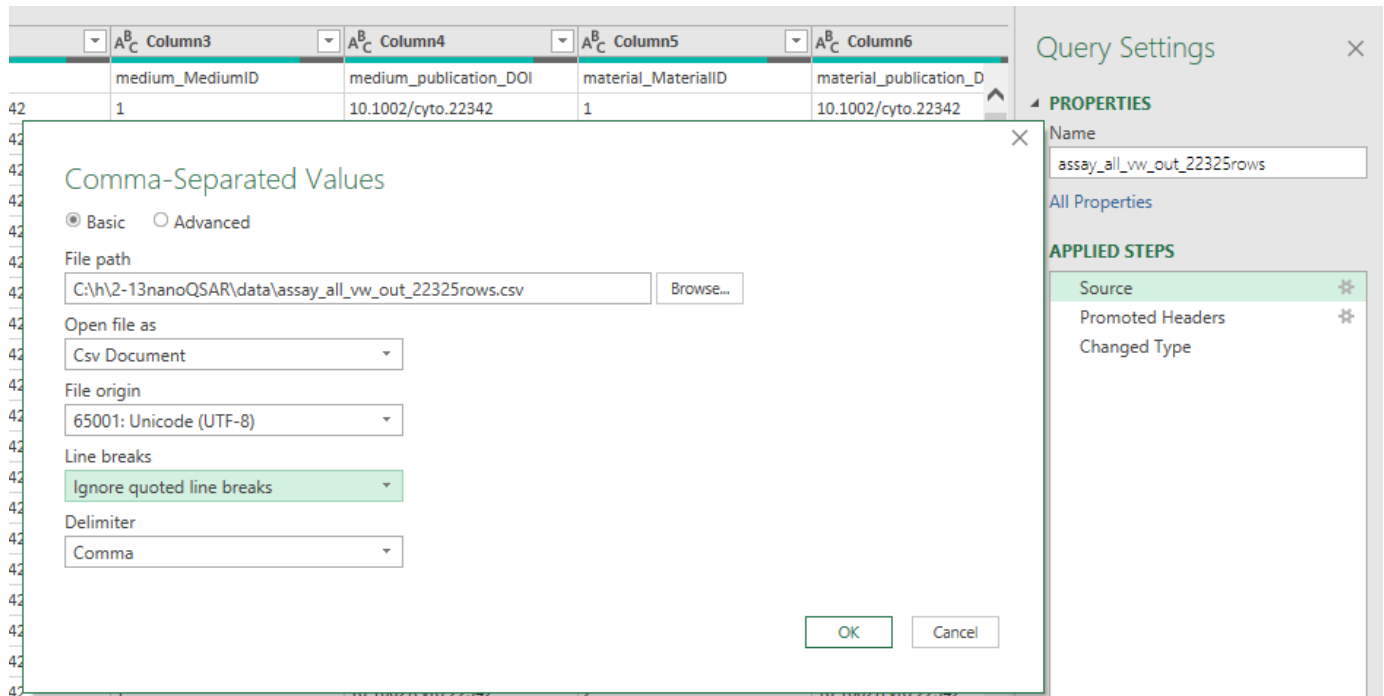
assay_all_vw_out_22325rows.csv

File Origin: 65001: Unicode (UTF-8) | Delimiter: Comma | Data Type Detection: Based on first 200 rows

assayID	publication_DOI	medium_MediumID	medium_publication_DOI	material_MaterialID	material_publication_DOI
1	10.1002/cyto.22342	1	10.1002/cyto.22342	1	10.1002/cyto.22342
2	10.1002/cyto.22342	1	10.1002/cyto.22342	1	10.1002/cyto.22342
3	10.1002/cyto.22342	1	10.1002/cyto.22342	1	10.1002/cyto.22342
4	10.1002/cyto.22342	1	10.1002/cyto.22342	1	10.1002/cyto.22342
5	10.1002/cyto.22342	1	10.1002/cyto.22342	2	10.1002/cyto.22342
6	10.1002/cyto.22342	1	10.1002/cyto.22342	2	10.1002/cyto.22342
7	10.1002/cyto.22342	1	10.1002/cyto.22342	2	10.1002/cyto.22342
8	10.1002/cyto.22342	1	10.1002/cyto.22342	2	10.1002/cyto.22342
9	10.1002/cyto.22342	1	10.1002/cyto.22342	3	10.1002/cyto.22342
10	10.1002/cyto.22342	1	10.1002/cyto.22342	3	10.1002/cyto.22342
11	10.1002/cyto.22342	1	10.1002/cyto.22342	3	10.1002/cyto.22342
12	10.1002/cyto.22342	1	10.1002/cyto.22342	3	10.1002/cyto.22342
13	10.1002/cyto.22342	1	10.1002/cyto.22342	1	10.1002/cyto.22342
14	10.1002/cyto.22342	1	10.1002/cyto.22342	1	10.1002/cyto.22342
15	10.1002/cyto.22342	1	10.1002/cyto.22342	1	10.1002/cyto.22342

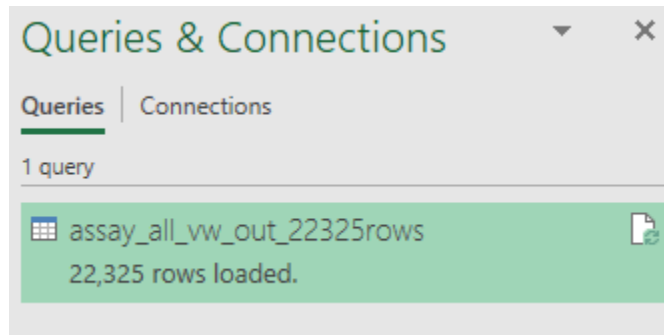
Load | Transform Data | Cancel

3. Applied Steps frame of the Transform Data window. > Source > Gear icon > Open file as CSV > file origin 65001: Unicode utf-8 > line breaks Ignore quoted line breaks > delimiter comma



4. Close and Load

5. Ensure there are no errors.



6. Save As a file with .XLSX extension.

Troubleshooting

- Greek characters are replaced with nonsense characters.
Choose the charset UTF8 which handles the Greek characters
- Import generates blank rows of numeric fields, but text field have portions of text data.
This is caused by embedded line returns in the text data. Transform the data with the Source Settings Ignore quoted line breaks. The Queries and Connections summary shows the error count.

AutoSave Off

Book1 -...

Helgen, Henry

FileHomeInsertPage LayoutFormulasDataReviewViewAdd-InsHelpPower QueryDesignQueries

Search

Table Name: assay_all_vw_out

Resize Table

Summarize with PivotTable

Remove Duplicates

Convert to Range

Insert Slicer

Export

Refresh

Table Style Options

Quick Styles

PropertiesToolsExternal Table DataTable Styles

A1

1

2

3

4

5

6

7

8

9

10

11

12

13

A

assayID

publication_DOI

1 10.1002/cyto.22342

2 10.1002/cyto.22342

3 10.1002/cyto.22342

4 10.1002/cyto.22342

5 10.1002/cyto.22342

6 10.1002/cyto.22342

7 10.1002/cyto.22342

8 10.1002/cyto.22342

9 10.1002/cyto.22342

10 10.1002/cyto.22342

11 10.1002/cyto.22342

12 10.1002/cyto.22342

Queries & Connections

QueriesConnections

1 query

assay_all_vw_out_22325rows

39,761 rows loaded. 17,436 errors.

100%