

Determining the encoding of a CSV file

The Problem

When uploading or importing a CSV file PANDA raises an error related to the encoding of the file.

Describe your dataset

Your file is currently being uploaded to PANDA. While you wait please take a moment to provide some information about it, such as when and how it was acquired.

100%

This CSV file contains characters that are not utf-8 encoded. You need to input the correct encoding in order to import data from this file. Try again

Encodings are a complex subject, which we won't try to explain here. (Joel Spolsky has written an [excellent primer](#) if you want the technical details.) The important thing to know is that PANDA can not infer from a CSV file what encoding it is in. If the file contains characters which are not supported by the default encoding (known as `utf-8`) this can cause PANDA to generate errors.

The default is right in the large majority of cases, but when it isn't right figuring out the correct encoding can be very tricky.

The Solution

Always ask your source what the encoding of the file is. Input the correct encoding after you select the CSV file to upload.

Select your file

Start by selecting a **CSV (.csv)** or **Excel (.xls or .xlsx)** file from your computer:

activities.csv

Advanced CSV options

Character encoding 

utf-8

If you have no way of finding out the correct encoding of the file, then try the following encodings, in this order:

- utf-8
- iso-8859-1 (also known as latin-1)
(This is the encoding of all census data and much other data produced by government entities.)
- utf-16

If none of these work the likelihood you are going to determine the encoding without additional information from the source is very low. In theory you may be able to guess the encoding based on the language of the author, however this not a recommended practice.

PANDA Docs

Tutorial

Uploading a dataset

Searching for data

Searching for datasets

Subscribing to search results

Problems + Solutions

Importing fixed-width files

Importing Access files

Importing DBF files

Importing unusual CSV file formats

Storing related files

Searching for names

Searching for phone numbers

Determining the encoding of a CSV file

Handling unsupported Excel features

Dealing with two-digit year columns

Seeing PANDA data on a map

Sorting PANDA data

More questions?

Join the PANDA User Group »

Administrator?

Read the admin docs »