

Using Spreadsheets to Manage and Analyze Data

Data & Donuts Series, Fondren Library
November 10, 2017
Monica Rivero | mpr1@rice.edu

- Best Practices for Managing your Data
- Tips for Using Excel Productively
- Importing Data
- Data Lists
- Advance Functions
- Data Analysis (Pivot Tables, Conditional Formatting)

Best practices for managing your data

Data entry guidelines:

- Place variable names in first row
- Use unique identifier per entity
- Format columns to match variable type (date, numeric, text, etc.)
- Enter string variables in a consistent manner (case, leading zeros, etc.)
- Do not leave any blank rows in the data range

Use in conjunction with a "Data Dictionary"

Keep track of changes made to your worksheet

Backup data periodically or before any major change.

Use descriptive filenames. Examples:

- WeatherData_2017-11-10.xlsx
- Census_data_v01.xlsx,
Census_data_v02.xlsx ,etc.

Sources:

Elliott, A C. (2006). Preparing data for analysis using Microsoft Excel. *Journal of investigative medicine*, 54(06), 334-341.

Hook, Les A., et al. 2010. Best Practices for Preparing Environmental Data Sets to Share and Archive. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. doi:10.3334/ORNLDAAC/BestPractices-2010 (<http://daac.ornl.gov/PI/BestPractices-2010.pdf>)

Exporting Data

Reasons to export data

Compatibility and interoperability

Import data to other tool or system

Publishing for reuse and supporting documentation

Supplemental datasets for Thesis and Dissertations

Method

Get rid of formulas – Use Copy > Paste special > Values only

Save each sheet as a separate file

Save as CSV format and Encode UTF-8 character set

Unicode.txt - Notepad

File Edit Format View Help

St??? ?p????st??, i? d?e???? p??t?p? Unicode st??e?e? st?? ??d???p???s? ???? i?? s?st?μ?i?? ??
 af?? p?? ???s?μ?p????ta? st?? p?a?i??, ?ste ?a ???e? d??at? ? ap????e?s? -st? μ??μ? e??? ?
 p????st?- ??ap?? ?e?μ???? ???? i?? ???ss?? s?μpe???αβα?μ???? ?a? s?μβ???? ep?st?μ??, ?p??
 μa??μat???, i?s??? ?i?

? ?a????s? i?? Unicode e??a? ??a f???d??? s??d?? af?? s??pe?e? ?a a?i??atst?se? ???e? i?? ?
 p????se? ??d???p???se?? s?????? ?a?a?i????, ?? ?p??e? ?????? pe?????sμ??? p?? i?? ?a?st????
 p??B??μat???? ?a ???s? se p????ss??? ?p????st??? s?st?μata.

?a?? ta te???? p??B??μata p?? ?????? pa???s?aste? i? Unicode ??e? ?a??e???e? ?? i? p?? p???e?
 s????? ?a?a?i???? ?a? ?? ? p??t?μ?te?? ??d??p???s? se p????ss??? ???sμ??? ???? p???sfata
 p??t?pa ?p?? i? XML, ?a??? ?a? ???sμ??? s?st?μat?? ?p?? ?e????????? s?st?μata, ????? ?????e?
 se? i? Unicode ??a ?a a?apa??st??? es?te???? ?e?μe??.

	A	B	C	D
		parent_sku	post_title	
		male Adaptor	Comp Female Adaptor	
		CFA	Comp Female Adaptor 16mm x	1
		CFA	Comp Female Adaptor 20mm x	1
		CFA	Comp Female Adaptor 20mm x	1
		CFA	Comp Female Adaptor 25mm x	1
		CFA	Comp Female Adaptor 25mm x	1
		CFA	Comp Female Adaptor 25mm x	1
		CFA	Comp Female Adaptor 32mm x	1
		CFA	Comp Female Adaptor 32mm x	1
		CFA	Comp Female Adaptor 40mm x	1

readme.txt x

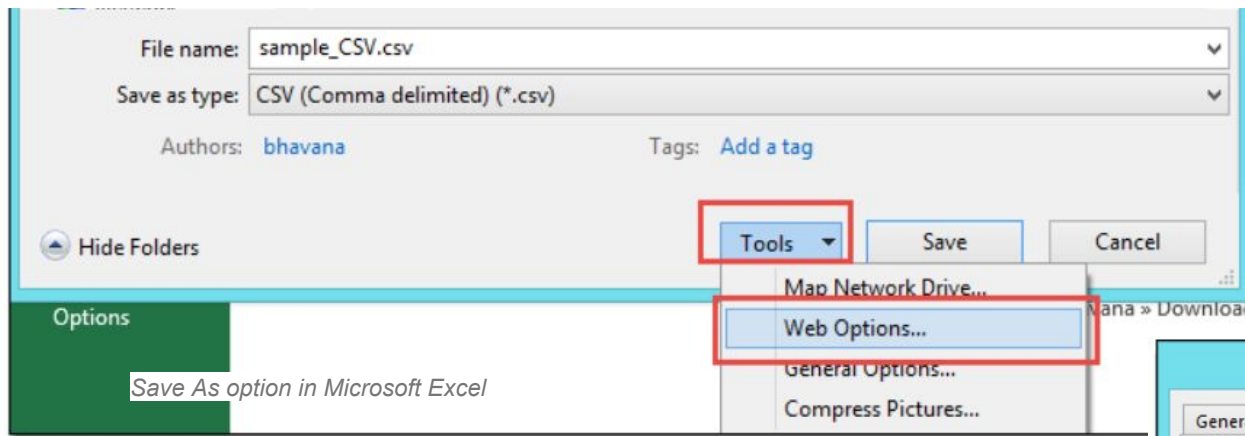
File was loaded in the wrong encoding: 'UTF-8' [Reload in another encoding](#)

?????? ?????????????? ?????????????? ?????????????? ?????????? Java
 ??????????? Java ???? ?????????????????, ???????????, ??????????? ? ?????????? ?????????????? ???????????

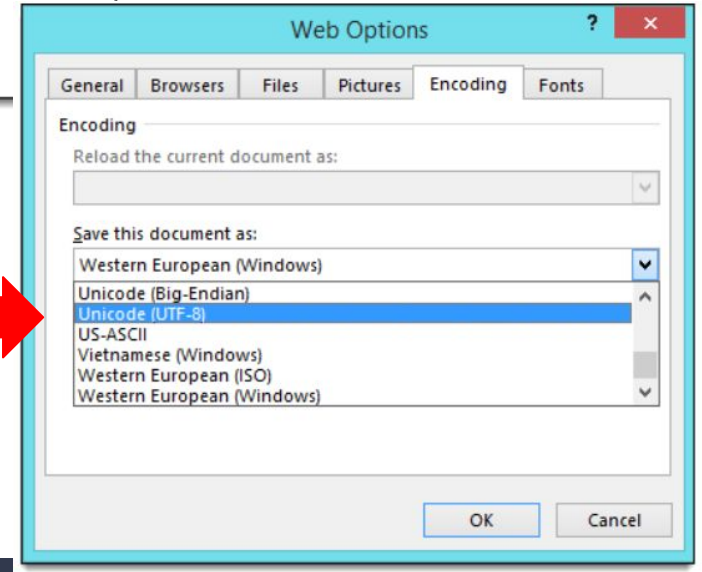
?????????? ?????????????? ?????????????? ?? ?????? ??????????? ? ??? ?????? ?????????????? ?? ?????? ?????
 ?????????? ??????????, ?????????????? ? ???-?????????? ? ?????????? ??????? ? ???-??????????

????????????? ?????????????? ?? ?????????? ?????????? ??? ?????????? ? ???????????, ???????????, ??????????, ??????????
 ?????????????? ?????????????? ??? ?????? ? ??????????????????? ?????? Java ??? ?????????? ???????????????????????
 ??????????? ?????????????????????????? ? ?????????????? ?????????????? ??? ??????????? ???????????, ??????????? ???

Examples of unencoded spreadsheets



Save As option in Microsoft Excel



Source: How to save CSV file as UTF-8 in Microsoft Excel?

<https://www.xadapter.com/how-to-save-csv-excel-file-as-utf-8-encoded/#save-csv-microsoft-excel>



Library / LibGuides / Data Resources / Sample Datasets

Data Resources: Sample Datasets

Links to a wide variety of data from local, state, national and international sources.

Introduction & Guide Scope

Specialty Topics

Comprehensive Tools

Census

Business

Sample Datasets

Datasets for Teaching and Learning

The following resources can be useful starting points for locating relevant datasets.

- [Tableau Sample Data Sets](#) - A changing sample of datasets for use in teaching and learning.
- [Awesome Public Datasets](#) - This curated list of datasets is arranged by discipline; the majority of the datasets are free.
- [Data.gov](#) - The home of the U.S. Government's open data.
- [FedStats](#) - This site provides access to the full range of official statistical information produced by the U.S. Government without having to know in advance which Federal agency produces which particular statistic.
- [NOAA Storm Events Database](#) - The Storm Events Database contains records documenting significant weather phenomena and meteorological events.
- [CORGIS Datasets Project](#) - Real-world datasets for subjects such as politics, education, literature, and construction.

http://libguides.rice.edu/data_resources/samples

Next will go through hands on examples using real data found from this list. See workshop handout for more details.

Attribution: List of data resources taken in part from Datasets for Teaching and Learning.

<https://www.lib.ncsu.edu/teaching-and-learning-datasets>