

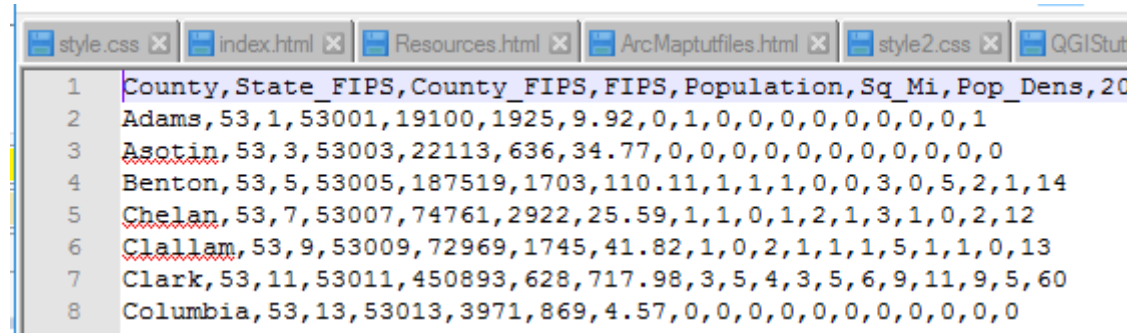
I have a .csv file, and need to assign a data type to each column. How do I create a .csvt file?

For this exercise, you will need an IDE

- Notepad++ comes pre-installed with most Windows Computers
- My personal favorite is [KomodoEdit](#)
- **DO NOT** use Notepad

What is a .csv file?

- CSV stands for “Comma Separated Value”
- Download the related spreadsheet VBAcc.csv, and open it in regular Notepad or Notepad++
- It kind of looks like the spreadsheet, but the values are separated by commas
- A program like Excel or QGIS understands how to arrange this data because it knows that each data field is so many commas and entries apart

A screenshot of a web browser window displaying a CSV file. The browser has several tabs open: style.css, index.html, Resources.html, ArcMaptutfiles.html, style2.css, and QGIS. The main content area shows a table with 8 rows of data. The first row is a header with 14 columns. The following rows contain data for various counties in Washington state, including Adams, Asotin, Benton, Chelan, Clallam, Clark, and Columbia. The data includes FIPS codes, population, square miles, and population density.

1	County	State	FIPS	County_FIPS	FIPS	Population	Sq_Mi	Pop_Dens	20	21	22	23	24	25
2	Adams	53	1	53001	19100	1925	9.92	0	1	0	0	0	0	0
3	Asotin	53	3	53003	22113	636	34.77	0	0	0	0	0	0	0
4	Benton	53	5	53005	187519	1703	110.11	1	1	1	0	0	3	0
5	Chelan	53	7	53007	74761	2922	25.59	1	1	0	1	2	1	3
6	Clallam	53	9	53009	72969	1745	41.82	1	0	2	1	1	1	5
7	Clark	53	11	53011	450893	628	717.98	3	5	4	3	5	6	9
8	Columbia	53	13	53013	3971	869	4.57	0	0	0	0	0	0	0

What is a .csvt file?

- A .csvt file helps QGIS understand which types of data are stored in a column
- It is an adjacent file
 - it is stored in the same folder with the same name as the related .csv file
 - DP03.csv and DP03.csvt live together

Fundamental Data Types: NOIR

- Nominal – names, in no particular order
 - Often categorical
- Ordinal
 - Naming indicates some sort of arrangement in a series or grouping
- Interval
 - Always numerical
 - Gaps between data points are interpretable, such as temperature
- Ratio
 - Continuous data
 - Gaps between data points are not important
 - Basic counts

Which data types does QGIS recognize?

- QGIS understands the following classes
 - Whole Number (**Integer** – good for counting people or number of objects)
 - Whole Number (64 bit integer – don't worry about this)
 - Decimal Number (**Real** – good for continuous data that doesn't have clean integers, such as surface area)
 - Text (**String**)
 - **Date**
 - **Time**
 - **Boolean** (TRUE or FALSE?)
- The **emboldened** text is what you will use to assign the data classification in your .csvt file

Open the VBAcc.csv file in Excel, or your preferred spreadsheet platform

- The spreadsheet has 18 columns, and we'll need to assign a data type to each one before importing it into QGIS
- County all text, so it will be assigned **String**
- The FIPS codes can be assigned an integer or a string, but since it is a categorical class of data, it is often classed as **String**
- Population represents people, so it will be an **Integer**
- Square Miles and Population Density both have decimals, so they will be **Real**
- The tabular data of events are whole – you can't have half a car accident – they are **Integer**

Open a new file in Notepad++ or KomodoEdit

- Save the file as “VBAcc.csvt”
- Now to assign categories...
- Note how there are no spaces between entries in the VBAcc.csv file
 - We will repeat this design with our .csvt file
- For county (column 1) type “String”
 - Must be capitalized at the front, must have quotes
- For State_FIPS, County_FIPS, and FIPS, type “String”
- Separate these first four entries with commas

It should look like this...

“String”,“String”,“String”,“String”,

Next, add the indicator for population, square miles, and population density

“String”, “String”, “String”, “String”, “Integer”, “Real”, “Real”,

Finally, the indicators for the tabular data from 2004-2013, and the total

- "String","String","String","String","Integer","Real","Real","Integer",
"Integer", "Integer", "Integer", "Integer", "Integer", "Integer",
"Integer", "Integer", "Integer", "Integer"

...and that's it. Make sure that you have the proper number of entries for each column, save, and then you can import your .csv file to QGIS