

Review

- Image Processing with ... uh ... Processing
 - Image Segmentation (Thresholding)
 - Image Enhancement (Histogram Equalization)
 - Feature Extraction (Dilate/Erode)
 - Spatial Filtering
 - Sharpen, Edge Detection and Gaussian Blur Kernels
 - Processing Functions
 - tint(), blend(), filter()
 - Applications
- Strings
 - Objects that hold an array of chars
 - Making Strings
 - ASCII Encoding
 - String methods and functions
 - Comparing Strings
 - Building Strings
 - Splitting Strings into an array
 - Joining the elements of a String array into a String

WHO Tuberculosis Data

The screenshot shows a Mozilla Firefox browser window with the address bar displaying <http://www.who.int/tb/country/data/download/en/index.html>. The page features the WHO logo and a navigation menu with the following items: Home, Health topics, Data and statistics, Media centre, Publications, Countries, Programmes and projects (highlighted), and About WHO. A search bar is located below the navigation menu.

Tuberculosis (TB)

- Tuberculosis
- TB Topics index
- Stop TB Strategy
- DOTS expansion
- TB diagnostics and laboratories
- TB/HIV
- MDR/XDR-TB
- Health systems
- Public-Private Mix
- Affected people
- TB research
- TB data**
- TB publications
- About us

Download data as CSV files

Data provided by countries to WHO and estimates of TB burden generated by WHO for *Global Tuberculosis Control 2010* are available for download as [comma-separated value \(CSV\)](#) files. CSV files can be imported into many spreadsheet and database programs, and some spreadsheets such as Excel can open CSV files directly.

The first row in each CSV file contains variable names; find the definition of each variable in the data dictionary.

Definition of variables: [» Download the data dictionary](#)

WHO TB burden estimates: [» Download WHO estimates \[760 KB\]](#)

This includes WHO-generated estimates of TB mortality, prevalence, incidence (including incidence of HIV+TB) and case detection rate.

[E-mail](#)

<http://www.who.int/tb/country/data/download/en/index.html>

Microsoft Excel

Home Insert Page Layout Formulas Data Review View Developer

Normal Page Layout Page Break Preview Custom Views Full Screen

Workbook Views

Ruler Formula Bar Gridlines Headings Message Bar

Show/Hide

Zoom 100% Zoom to Selection

Zoom

New Window Arrange All Freeze Panes Hide Split View Side by Side Synchronous Scrolling Reset Window Position

Window

Save Workspace Switch Windows

Macros

A1 country

Book1:2

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	country	iso2	iso3	iso_numeric	g_whoregion	year	e_pop_num	e_prev_100k	e_prev_100k_lo	e_prev_100k_hi	e_prev_num	e_prev_num_lo	e_prev_num_hi	e_mort_exe	e_mort_exc_tbf
2	Afghanistan	AF	AFG	4	EMR	1990	12580412	452	196	754	57000	25000	95000	66	
3	Afghanistan	AF	AFG	4	EMR	1991	13427960	452	196	754	61000	26000	100000	66	
4	Afghanistan	AF	AFG	4	EMR	1992	14572340	452	196	754	66000	28000	110000	66	
5	Afghanistan	AF	AFG	4	EMR	1993	15861049	452	196	754	72000	31000	120000	66	
6	Afghanistan	AF	AFG	4	EMR	1994	17081664	452	196	754	77000	33000	130000	66	
7	Afghanistan	AF	AFG	4	EMR	1995	18083748	452	196	754	82000	35000	140000	66	
8	Afghanistan	AF	AFG	4	EMR	1996	18807500	452	196	754	85000	37000	140000	66	
9	Afghanistan	AF	AFG	4	EMR	1997	19303252	452	196	754	87000	38000	150000	66	
10	Afghanistan	AF	AFG	4	EMR	1998	19665668	452	196	754	89000	38000	150000	66	
11	Afghanistan	AF	AFG	4	EMR	1999	20041026	443	196	734	89000	39000	150000	63	

Book1:1

	A	B	C	D	E	F	G	H
1	variable_name	dataset	code_list	definition				
2	country	Country identification		Country or territory name				
3	iso_numeric	Country identification		ISO numeric country/territory code				
4	iso2	Country identification		ISO 2-character country/territory code				
5	iso3	Country identification		ISO 3-character country/territory code				
6	c_cdr	Estimates		Case detection rate (all forms), percent				
7	c_cdr_hi	Estimates		Case detection rate (all forms), percent, high bound				
8	c_cdr_lo	Estimates		Case detection rate (all forms), percent, low bound				
9	e_inc_100k	Estimates		Estimated incidence (all forms) per 100 000 population				
10	e_inc_100k_hi	Estimates		Estimated incidence (all forms) per 100 000 population, high bound				
11	e_inc_100k_lo	Estimates		Estimated incidence (all forms) per 100 000 population, low bound				

C:\BMC\CS110 Fall 2011\Examples\19\ParseFile1\data\reduced.csv - Notepad++

File Edit Search View Encoding Language Settings Macro Run TextFX Plugins Window ?

reduced.csv

```
1 country,year,e_pop_num,e_inc_100k
2 Afghanistan,1990,12580412,189
3 Albania,1990,3289483,24
4 Algeria,1990,25282516,38
5 American Samoa,1990,47108,21
6 Andorra,1990,52778,54
7 Angola,1990,10661460,205
8 Anguilla,1990,8402,24
9 Antigua and Barbuda,1990,61922,2
10 Argentina,1990,32497508,60
11 Armenia,1990,3544695,33
12 Australia,1990,17091250,7.4
13 Austria,1990,7670513,25
14 Azerbaijan,1990,7211692,110
15 Bahamas,1990,255603,22
16 Bahrain,1990,492952,40
17 Bangladesh,1990,115632152,225
18 Barbados,1990,259668,2.4
19 Belarus,1990,10259700,80
20 Belgium,1990,9933055,20
21 Belize,1990,189822,40
22 Benin,1990,4795088,77
23 Bermuda,1990,59796,0
24 Bhutan,1990,548774,308
25 Bolivia (Plurinational State of),1990,6670556,251
26 Bosnia and Herzegovina,1990,4308199,94
27 Botswana,1990,1351599,307
28 Brazil,1990,149570480,84
```

length : 122858 lines : 4215 Ln : 1 Col : 1 Sel : 0 Dos\Windows ANSI as UTF-8 INS

Loading Data From a File

- `loadStrings()` function
 - Reads all data and returns an array of Strings
 - Each String in the array is a separate line from the file

```
// ParseFile1

String[] data;
int count = 0;

void setup() {
  // Load data from a file as array of strings
  data = loadStrings("reduced.csv");
}

void draw() {
  // Continue printing data until run out
  if (count >= data.length) return;
  println(data[count]);
  count++;
}
```

Split a String based on a single or multiple separator chars

```
String s1 = "12, 34, 56";  
String[] as;
```

```
void setup() {  
    as = split(s1, ",");  
    //as = trim(as);  
    println( as );  
}
```

```
[0] "12"  
[1] " 34"  
[2] " 56"
```

```
String s1 = "Data: 12, 34, 56";  
String[] as;
```

```
void setup() {  
    as = splitTokens(s1, ":",");  
    //as = trim(as);  
    println( as );  
}
```

```
[0] "Data"  
[1] " 12"  
[2] " 34"  
[3] " 56"
```

```

// ParseFile2

String[] data;
Item[] items;
int count = 0;

void setup() {
  // Load data as array of strings
  data = loadStrings("reduced.csv");

  // Build object array
  items = new Item[data.length];
  for (int i=0; i<data.length; i++) {
    items[i] = new Item(data[i]);
  }
}

void draw() {
  // Continue printing data until run out
  if (count >= items.length) return;
  items[count].pr();
  count++;
}

class Item {
  String country; // Country name
  int year; // Year
  int pop; // Population
  int inc; // Incidences of TB
             // per 100,000

  Item(String line) {
    String[] data = split(line, ",");
    country = data[0];
    year = int(data[1]);
    pop = int(data[2]);
    inc = int(data[3]);
  }

  void pr() {
    String msg = "In " + year + ", " + country;
    msg += " had population " + pop;
    msg += " and TB incidences per 100k of " + inc;
    println(msg);
  }
}

```

File Selection

- `selectInput()` function
 - Displays a file chooser allowing user to select a file
 - Returns a complete path to selected file as a String

```
void draw() { }
```

```
void mousePressed()  
{  
  String filepath = selectInput("Please select a data file");  
  println(filepath);  
}
```

Also see `selectOutput()`

A PERIODIC TABLE OF VISUALIZATION METHODS

<p>Data Visualization Visual representations of quantitative data in schematic form (either with or without axes)</p>								<p>Strategy Visualization The systematic use of complementary visual representations in the analysis, development, formulation, communication</p>								<p>G graphic facilitation</p>																			
<p>C continuum</p>		<p>Information Visualization The use of interactive visual representations of data to amplify cognition. This means that the data is transformed into an image, it is mapped to screen space. The image can be changed by users as they proceed working with it</p>								<p>CLUSTERING</p>								<p>G graphic facilitation</p>																	
<p>Tb table</p>		<p>Ga cartesian coordinates</p>		<p>Concept Visualization Methods to elaborate (mostly) qualitative concepts, ideas, plans, and analyses.</p>								<p>Tm temple</p>		<p>St story template</p>		<p>Tr tree</p>		<p>Ct cartoon</p>																	
<p>Pi pie chart</p>		<p>L line chart</p>		<p>Roles</p>								<p>Cs concept skeleton</p>		<p>Br bridge</p>		<p>Fu funnel</p>		<p>Ri rich picture</p>																	
<p>B bar chart</p>		<p>Ac area chart</p>		<p>R radar chart</p>		<p>Pa parallel coordinates</p>		<p>Hy hyperbolic tree</p>		<p>Cy cycle diagram</p>		<p>T timeline</p>		<p>V venn diagram</p>		<p>Pm perspectives diagram</p>		<p>D dilemma diagram</p>		<p>Pr parameter ruler</p>		<p>Kn knowledge map</p>													
<p>Hi histogram</p>		<p>Sc scatterplot</p>		<p>Sa sankey diagram</p>		<p>In information lense</p>		<p>E entity relationship diagram</p>		<p>Pt petri net</p>		<p>Fl flow chart</p>		<p>Cl clustering</p>		<p>Lc layer chart</p>		<p>Fy minto pyramid technique</p>		<p>Ce cause-effect chains</p>		<p>Tl toulmin map</p>		<p>Dt decision tree</p>		<p>Cp cpm critical path method</p>		<p>Cf concept fan</p>		<p>Co concept map</p>		<p>Ic iceberg</p>		<p>Lm learning map</p>	
<p>Tk tukey box plot</p>		<p>Sp spectrogram</p>		<p>Da data map</p>		<p>Tp treemap</p>		<p>Cn cone tree</p>		<p>Sy system dyn./ simulation</p>		<p>Df data flow diagram</p>		<p>Se semantic network</p>		<p>So soft system modeling</p>		<p>Sn synergy map</p>		<p>Fo force field diagram</p>		<p>Ib ibis argumentation map</p>		<p>Pr process event chains</p>		<p>Pe pert chart</p>		<p>Ev evocative knowledge map</p>		<p>V vee diagram</p>		<p>Hh heaven 'n' bell chart</p>		<p>I infomural</p>	

Cy Process Visualization

Note: Depending on your location and connection speed it can take some time to load a pop-up picture.

version 1.5

© Ralph Lengler & Martin J. Eppler, www.visual-literacy.org

Jer Thorp. Artist/Educator - NYU

225 "random" numbers chosen and tweeted by 225 people

19	42	42	87	81	99	33	98	61	47	24	66
69	23	67	67	57	71	5	79	57	46	93	54
43	32	18	42	77	37	37	6	93	55	55	77
15	88	42	55	77	42	93	3	17	26	64	65
23	21	9	7	23	17	14	42	45	27	97	83
89	4	4	26	6	39	97	72	35	6	66	
19	2	72	81	37	47	66	17	12	52	74	
54	61	43	19	57	17	77	47	26	72	64	
69	99	64	88	67	1	36	2	60	27	73	
4	43	97	67	42	37	27	1	75	15	17	
13	59	32	78	40	15	64	77	11	1	17	
37	13	7	26	57	25	12	69	8	84	23	
66	42	14	33	17	97	25	57	1	81	97	
8	18	78	12	95	37	84	86	41	56	73	
78	60	21	39	28	17	83	69	12	74	37	
67	19	19	88	96	69	29	74	53	33	72	
32	81	72	72	73	39	52	97	77	77	41	
76	17	69	83	67	64	25	35	42	4	76	
13	36	2	37	52	47	43	25	66	7	6	
87	94	16	28	20	79	23	21	55	66	87	

Data Sources

<http://www.data.gov/>

<http://archive.ics.uci.edu/ml/>

<http://opendata.socrata.com/>

<http://www.reddit.com/r/datasets>

Data Source Lists

<http://www.quora.com/Data/Where-can-I-get-large-datasets-open-to-the-public>

[http://www.readwriteweb.com/archives/where to find open data on the.php](http://www.readwriteweb.com/archives/where-to-find-open-data-on-the.php)

Ideas for Visualizations

[http://www.visual-literacy.org/periodic table/periodic table.html](http://www.visual-literacy.org/periodic_table/periodic_table.html)

Examples

- BoxOffice
- BoxOffice2
- BoxOffice3
- PieSlice
- PieChart
- PieChart2
- PieChart3
- USMap
- USMap2
- RandomTweets1
- RandomTweets2
- BarTweets1
- BarTweets2
- BarTweets3
- GridTweets1
- GridTweets2

GapMinder

<http://www.gapminder.org>

<http://www.gapminder.org/videos/hans-rosling-on-cnn-us-in-a-converging-world/>

Hans Rosling

Karolinska Institutet

Stockholm, Sweden