
Sprint Review

Data File Formats & Conversion for N00bs



devleague

Sprint Objective

Introduce marketers to key file formats, show how they are applicable in the field, and demonstrate basic file conversions.

What will we cover today?

- 1.) Examine three file formats: CSV, XML, JSON
- 2.) Determine why each file type is important.
- 3.) How Python is used to convert files.
- 4.) Research scripts to convert CSV files to XML, and CSV files to JSON.

What are key distinguishing features of the different file formats?

Hierarchies

Size

**Human
Readability**

File hierarchy comparison

CSV

Cersei, Arya, Tyrion

JSON

```
{"name":"Cersei","age":"36"},  
{"name":"Arya","age":"14"},  
{"name":"Tyrion","age":"34"}
```

XML

```
<person>  
    <name> Cersei</name>  
    <age> 36 </age>  
</person>  
<person>  
    <name> Arya</name>  
    <age> 14 </age>  
</person>  
<person>  
    <name> Tyrion</name>  
    <age> 34 </age>  
</person>
```

File size comparison



creator	credit	date	description
Robert Can-Gilbridge	Granted by the World Progress Administration	1915	One of a pair of marble at the
Sam Brown	City and County of Nevada Purchase	1915	Stone figure of Jesus Christ
Robert Can-Gilbridge	Granted by the World Progress Administration	1915	One of a pair of marble at the
Charles Dickey		1915	Three sets of double stone and
Kate Kelly	Gift of the Amelia Sartori Memorial Commission	1915	A cast bronze commemorative
Raymond Louis Blomgren	Granted by the World Progress Administration	1915	The stone has relief carvings and
Charles May Fraser		1915	A relief depicting various acts
Charles Watson		1915	A stylized sculpture of a group
Norman Roth		1915	A roughly carved "C" form which represents "territory," representing
Raymond Louis Blomgren	Granted by the World Progress Administration	1915	One of a pair of four-sided marble tablets of a historical monument with a
Shirley		1915	A relief cast bronze bust of Jesus Christ (Jesus, Mary, City and County of Nevada)
Michael Woldrich	City and County of Nevada Purchase	1915	Stone sculpture of Indiana Brinkman of a Hawaiian woman carrying a
Jack Brinkman	City and County of Nevada Purchase	1915	A 1/2 size stone sculpture of a seated male wearing the newspaper with three men and
Ernest Schick	City and County of Nevada Purchase	1917	The sculpture represents one of the passengers of the New York's Columbus
	Gift of Nevada Memorial Park	1915	Stone base-relief plaque of a woman carrying a girl accompanied by a
	Gift from the City of Kailashang, China	1915	A pair of stone Chinese guardian lion figures. Located at Dr. Sun Yat-sen Memorial
Paul Henrichs	City and County of Nevada Purchase	1915	Life size stainless steel figure of a sea turtle (Turtle). Located at Dr. Sun Yat-sen
St. Stefan	City and County of Nevada Purchase	1914	Three bronze medallions each representing an aspect of Chinese philosophy: rock
Dennis Naguchi	City and County of Nevada Purchase	1917	An abstract nonrepresentational steel sculpture composed of three sculptural an
Jack Brinkman	Gift of Jack Brinkman	1915	Three concrete figures depicted on iron tubes, one male and one female, in
	Gift from the City of Nagasaki, Japan	1915	A bronze bell shaped with a stone supported by four steel posts with a stone
George W. Williams	Gift of William Williams of America	1915	Stone replica of the Liberty Bell. Located in front of the Western Museum Building
Karl Rupp	Gift of the Nevada State Board	1915	The Northern Nevada Nevada Memorial is a Japanese sculpture depicting the
Carl Rupp	Gift of the Japanese Society of Hawaii with the Delta Japanese Community Group	1915	Life size standing figure of Dr. Jose Protacio Rizal, national hero of the Philippine
Masakazu Naguchi	Gift of the City of Hiroshima, Japan	1914	Replica of the Bell of Hiroshima, Japan's Peace Memorial Bell. Located on College
Edward Shewell	Gift of the Downtown Improvement Association	1917	Five life-size male concrete statues depicting Hawaiian mythology. Like their
	Gift of the Governor of Hokkaido, Japan	1915	Stone triangular statue in the Japanese Garden. Located at the Foster
after Governor Ono	Gift of the Governor of Kanagawa Prefecture	1915	Three stone figures depicted on iron tubes, one male and one female, in
Bob New		1915	Three figures surrounding a granite pedestal copied from an original dated

Public_Art.csv

93 kb



Public_Art.json

114 kb



Public_Art.xml

163kb

When would you use these types of files?

- CSV: Transferring large amounts of data.
- JSON: Web applications.
- XML: Document markups.

How does this apply to marketers?

- CSV: Pulling large data sets out of CRMs
- JSON: Extracting data from social media platforms
- XML: RSS feeds for Google alerts about the business

File Conversion Process

- 1.) Determine which file types are being converted.
- 2.) Research Python scripts online.
- 3.) Build scripts in Atom.
- 4.) Execute scripts in command line/terminal.

Python & Command Line Basics

- 1.) The command line is used in tandem with a text-editor like Atom.
- 2.) Scripts can be built in Atom, and executed in the command line.

CSV to XML Script

```
Public_Art.json  movie_test_file.json  Public-Art-CSV-to-JSON-converting-scri...
8  # 40,001,5.0
9  # 9,ghi,6.3
10 # 76,def,99
11
12 import csv
13
14 csvFile = 'Public_Art.csv'
15 xmlFile = 'Public_Art.xml'
16
17 csvData = csv.reader(open(csvFile))
18 xmlData = open(xmlFile, 'w')
19 xmlData.write('<?xml version="1.0"?>' + "\n")
20 # there must be only one top-level tag
21 xmlData.write('<csv_data>' + "\n")
22
23 rowNum = 0
24 for row in csvData:
25     if rowNum == 0:
26         tags = row
27         # replace spaces w/ underscores in tag names
28         for i in range(len(tags)):
29             tags[i] = tags[i].replace(' ', '_')
30     else:
31         xmlData.write('<row>' + "\n")
32         for i in range(len(tags)):
33             xmlData.write('    ' + '<' + tags[i] + '>' \
34                 + row[i] + '</' + tags[i] + '>' + "\n")
35         xmlData.write('</row>' + "\n")
36
37     rowNum += 1
38
39 xmlData.write('</csv_data>' + "\n")
40 xmlData.close()
```

CSV to XML Script, Translated

- 1.) Read existing CSV file.
- 2.) Open a new writeable XML file.
- 3.) Translate CSV row headers into tags.
- 4.) Add brackets around the tags for each column of information to format as XML.

CSV to XML Output

```
<?xml version="1.0"?>
<csv_data>
  <row>
    <creator>Robert Lee Eskridge</creator>
    <credit>Funded by the Works Progress Administration</credit>
    <date>1935</date>
    <description>One of a pair of murals at the Lester McCoy Pavilion at Ala Mona Regional Park. A Works Progress Administration mural depicting the Makahiki Festival (a festival of dance and fertility) being presented with ho'okupu (tribute). In the distance kahuna (priests) guard the kapa (banners).
    <locationname>Lester McCoy Pavilion</locationname>
    <imagefile>http://hiculturearts.pastperfect-online.com/34250images/001/19300101.JPG</imagefile>
    <objectid>1930.01.01</objectid>
    <discipline>Mural</discipline>
    <title>The Makahiki Festival - The Makai Mural</title>
    <access>Limited</access>
    <latitude>21.290824</latitude>
    <longitude>-157.85131</longitude>
    <location>(21.290824, -157.85131)</location>
    <thumb></thumb>
  </row>
```

CSV to JSON Script

```
1  import pandas as pd
2  df = pd.read_csv('Public_Art.csv')
3  # any operations on dataframe df
4  df.to_json('Public_Art.json')
5  print ("Conversion complete!")
6
```

CSV to JSON Script, Translated

- 1.) Import the pandas package.
- 2.) Read the CSV file.
- 3.) Convert to JSON.
- 4.) Print a “function complete!” line in the terminal once this script has been executed.

CSV to JSON Output

Bonus: Use the Atom-Beautify package to “prettify” JSON outputs!

```
1  {  
2    "creator": {  
3      "0": "Robert Lee Eskridge",  
4      "1": "Sean Browne",  
5      "2": "Robert Lee Eskridge",  
6      "3": "Charles Dickey",  
7      "4": "Kate Kelly",  
8      "5": "Marguerite Louis Blasingame",  
9      "6": "Juliette May Fraser",  
10     "7": "Charles Watson",  
11     "8": "Yoshinari Kochi",  
12     "9": "Marguerite Louis Blasingame",  
13     "10": "Giordano",  
14     "11": "Michael Weidenbach",  
15     "12": "Jodi Endicott",  
16     "13": "Ernest DeCoito",  
17     "14": null,  
18     "15": null,
```

Fin