Regex Testing to Clean Data

Use Brackets to do Regex. VSCode confused.
https://regex101.com for testing.

Example URL:
"<img src=\"https://maps.psiee.psu.edu/PSU_ARB/px/rs/Abelia x grandiflora_2 (479).jpg\">"

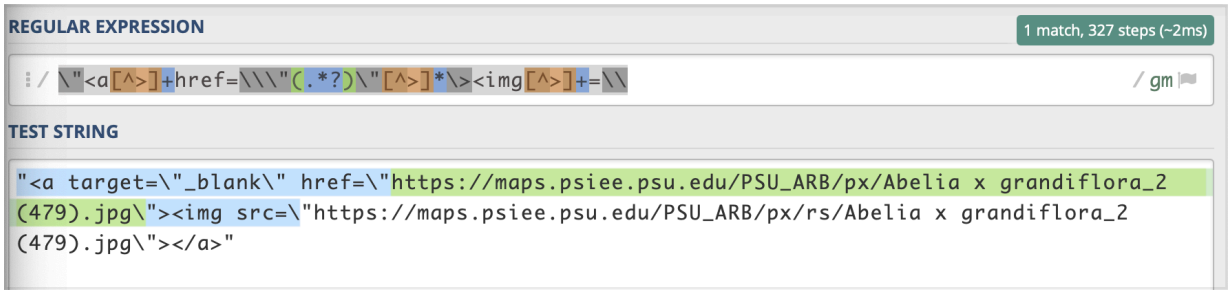
Removes up to the first https:// link.
\"<a[^>]+href=\\

Almost selects just the URL
\"(.*?)\"[^>]*

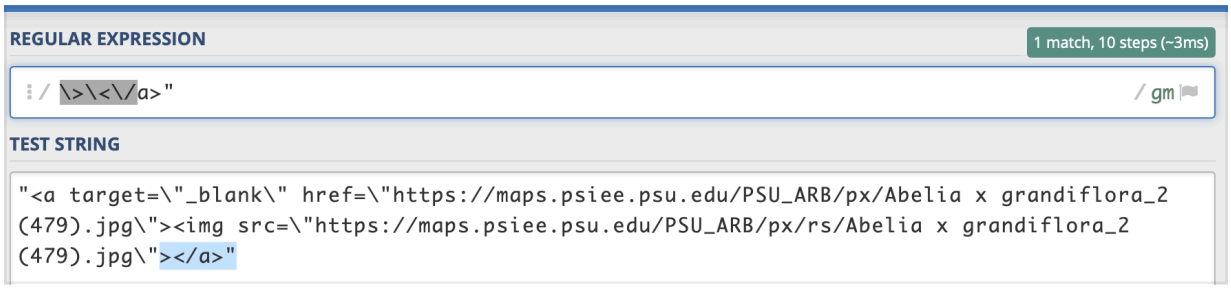
Selects the img src:
<img[^>]+=\\

Strategy: Remove the “a target” portion up to img src

Step 1:
Selects up to the first quote on the img src link:
\"<a[^>]+href=\\\"(.*?)\"[^>]*\><img[^>]+=\\



Step 2:
Remove the end <\a> tag and slash after jpg. Leaves the quote mark on the end.
\\"\>\<\/a>



Step3:
Find all spaces in the url strings.
(\ *?)

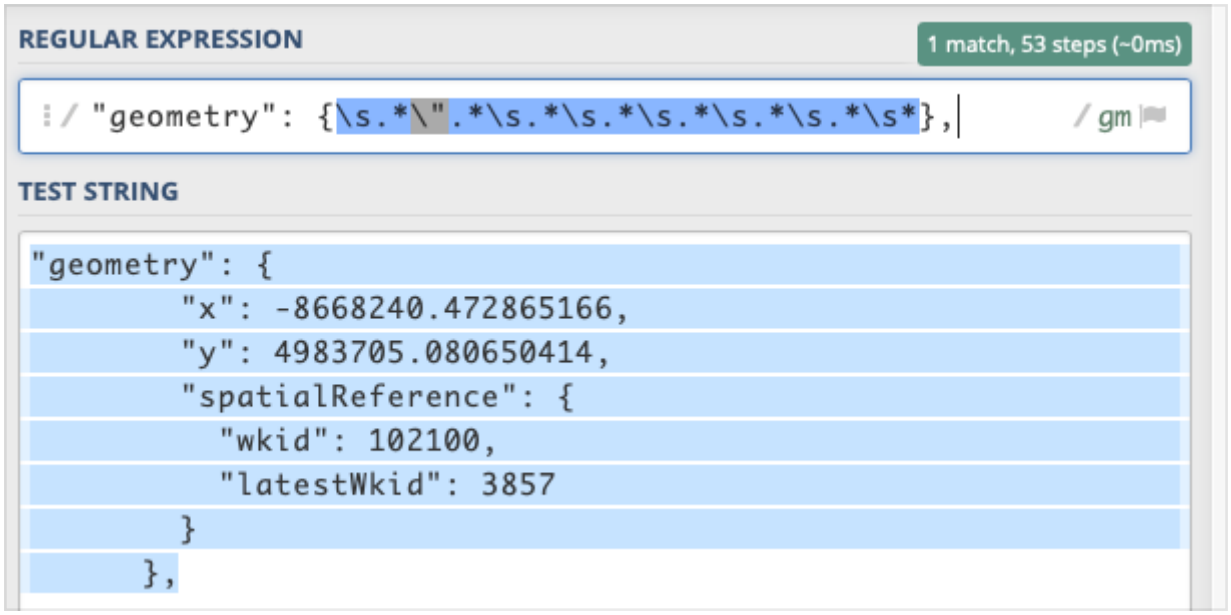
Then add “%20” in each of the spaces of the URL to make it a valid URL.

Replace with space

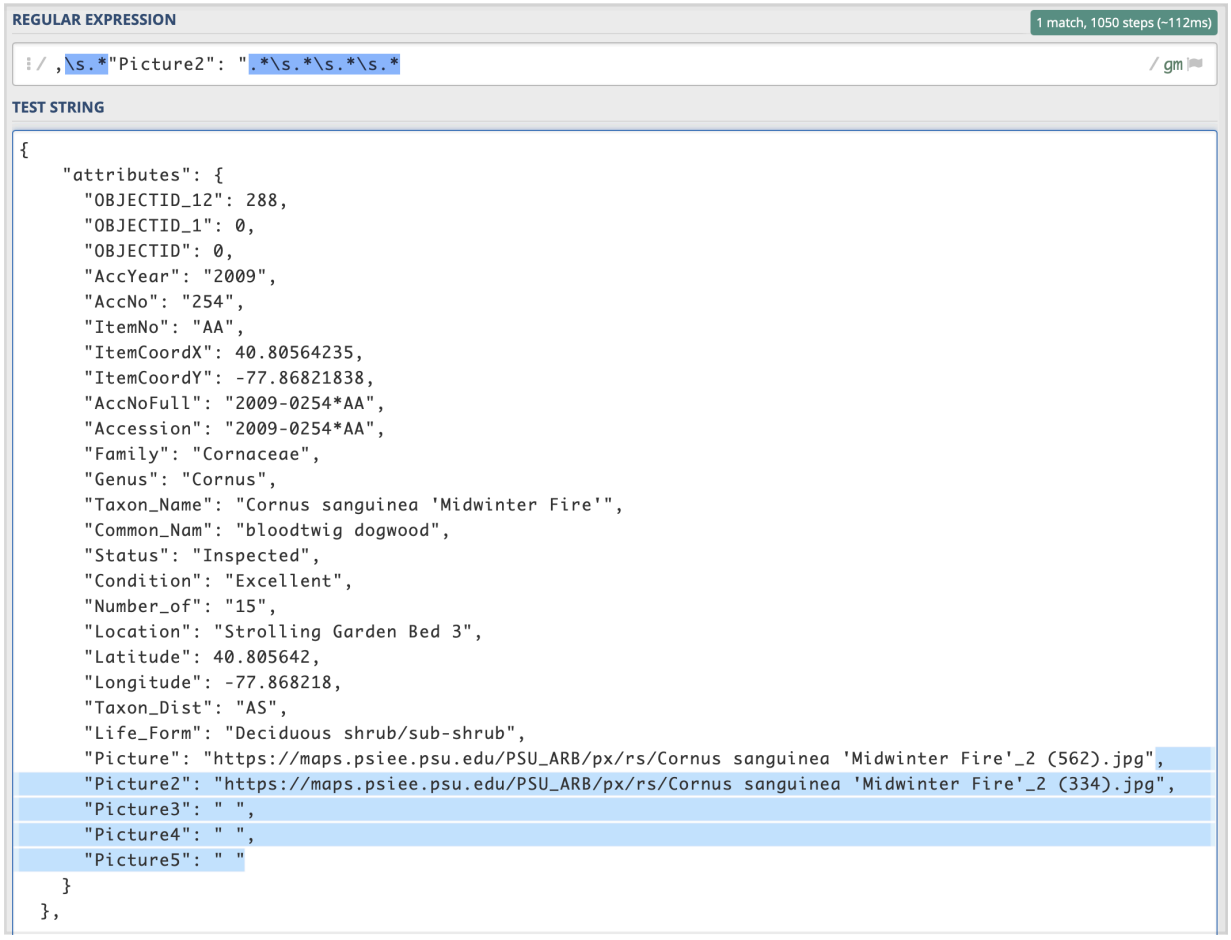
(*? \ *?)

Captures all dots:
(\.)

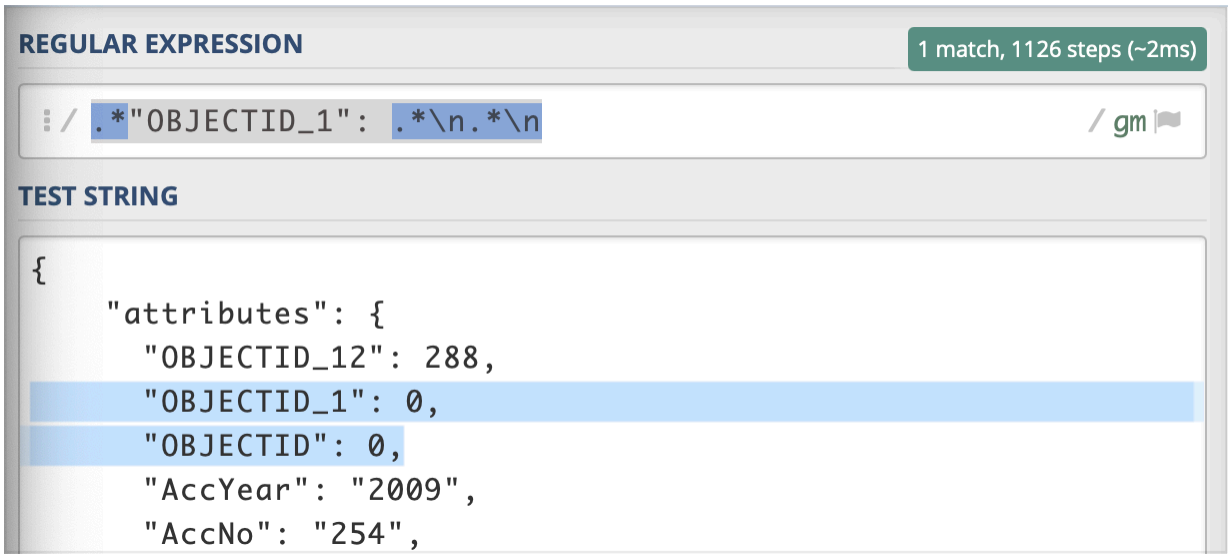
Working on clearing geometry in each
"geometry": {\s.*\".*\s.*\s.*\s.*\s.*\s*},



Removing Picture2-5, not needed, causes errors.
,\s.*"Picture2": ".*\n.*\n.*\n.*
Removes comma, JSON remains valid.



Found that OBJECTID and OBJECTID_1 have duplicate values for plants. Like 0 has 1057 matches. Not good for identifiable. Using OBJECTID_12 instead. Not removing.
NOTE: **Just ignoring for now.**
.*"OBJECTID_1": .*\n.*\n



Removing “Latitude, Longitude”, preferring to use ItemCoordX and Y due to accuracy of Coordinates. No need for duplicate values.

.*"Latitude": .*\n.*\n

REGULAR EXPRESSION

1 match, 1106 steps (~2ms)

/ .*"Latitude": .*\n.*\n / gm

TEST STRING

```
{
    "attributes": {
      "OBJECTID_12": 288,
      "OBJECTID_1": 0,
      "OBJECTID": 0,
      "AccYear": "2009",
      "AccNo": "254",
      "ItemNo": "AA",
      "ItemCoordX": 40.80564235,
      "ItemCoordY": -77.86821838,
      "AccNoFull": "2009-0254*AA",
      "Accession": "2009-0254*AA",
      "Family": "Cornaceae",
      "Genus": "Cornus",
      "Taxon_Name": "Cornus sanguinea 'Midwinter Fire'",
      "Common_Nam": "bloodtwig dogwood",
      "Status": "Inspected",
      "Condition": "Excellent",
      "Number_of": "15",
      "Location": "Strolling Garden Bed 3",
      "Latitude": 40.805642,
      "Longitude": -77.868218,
```