

Cleaning HPWRA input data for HAPI Project

Kelsey Brock - Last Updated Feb 28, 2020

```
In [1]: # ignore warnings to compile the final notebook
import warnings
warnings.filterwarnings('ignore')

In [2]: import pandas as pd

Cleaning HPWRA Data

In [3]: # import csv file of HPWRA data with analysis friendly column headings
CompiledData = pd.read_csv("UltimateDataCompilation.csv", sep=';',
                            names = ["Filename",
                                     "1.01_domesticated", "1.02_naturalized_grown", "1.03_weedy_races",
                                     "2.01_clim_match", "2.02_climmatch_qual",
                                     "3.01_beyond_native", "3.02_disturbance_weed",
                                     "3.03_agri_forestry_weed", "3.04_enviro_weed",
                                     "3.05_congener", "4.01_spyiny", "4.02_allopathic",
                                     "4.03_parasitic", "4.04_unpalatable", "4.05_toxic", "4.06_alternate_host",
                                     "4.07_allergies", "4.08_fire_hazard", "4.09_shade_tolerant", "4.10_tolerates_soilcond",
                                     "4.11_climber", "4.12_forms_thickets", "5.01_aquatic", "5.02_grass",
                                     "5.03_nitrogen_fixer", "5.04_geophyte", "5.01_repro_failure", "5.02_viable_seed",
                                     "5.03_hybridizes", "5.04_selfcompatible", "5.05_special_pollinators", "5.06_vegetative_repr",
                                     "6.01",
                                     "6.02",
                                     "6.03",
                                     "6.04",
                                     "6.05",
                                     "6.06",
                                     "6.07_minimum_gen_time", "7.01_unintentional_dispersal",
                                     "7.02_intentional_dispersal",
                                     "7.03_contaminant_dispersal",
                                     "7.04_wind_dispersal", "7.05_water_dispersal", "7.06_bird_dispersal",
                                     "7.07_animal_dispersal", "7.08_survive_gut", "8.01_prolific_seeder",
                                     "8.02_propagule_bank", "8.03_herbicide_controlled",
                                     "8.04_tolerates_mutilation", "8.05_local_enemies", "manual_score"]])

In [4]: # visual check to make sure everything loaded okay
CompiledData.head(5)

Out[4]:
   Filename  1.01_domesticated  1.02_naturalized_grown  1.03_weedy_races  2.01_clim_match  2.02_climmatch_qual  2.03_broad_clim  2.04_similar_clim
0      File_name              1.01              1.02              1.03              2.01              2.02              2.03              2
1  Abelia_x_grandiflora.xls              y              n              n              1              2              y              y
2  Acacia_auroculiformis.xls              N              Y              N              2              2              y              y
3  Acacia_confusa.xls              N              Y              N              2              2              N              N
4  Acacia_crassecarpa.xls              n              y              n              2              2              n              n
5 rows x 51 columns

In [5]: # removing the old names in the first column
HPWRA1 = CompiledData.iloc[1:]

In [12]: HPWRA1.head(5)

Out[12]:
   Filename  1.01_domesticated  1.02_naturalized_grown  1.03_weedy_races  2.01_clim_match  2.02_climmatch_qual  2.03_broad_clim  2.04_similar_clim
1  Abelia_x_grandiflora.xls              Yes              n              n              1              2              y              y
2  Acacia_auroculiformis.xls              No              y              N              2              2              y              y
3  Acacia_confusa.xls              No              Y              N              2              2              N              N
4  Acacia_crassecarpa.xls              No              y              n              2              2              n              n
5  Acacia_tenuicarpa.xls              No              y              n              2              2              y              y
6  Acacia_trophilata.xls              No              y              n              1              1              n              n
7  Acacia_mangium.xls              No              NaN              NaN              2              2              n              n
8  Acacia_mearnsii.xls              No              y              n              1              2              y              y
9  Acacia_melanocorymbos.xls              No              y              n              1              2              y              y
10 Acacia_nitida.xls              No              y              NaN              2              2              n              n
11 Acacia_puramidensis.xls              No              y              n              2              2              NaN              NaN
12 Acacia_pycnantha.xls              NaN              NaN              NaN              1              2              n              n
13 Acalypha_godseffiana.xls              No              n              n              2              1              n              n
14 Acalypha_hispida.xls              No              y              n              2              2              n              n
15 Acalypha_weddensiana.xls              No              n              n              2              2              n              n
16 Acnella_grandiflora.xls              No              NaN              NaN              2              2              NaN              NaN
17 Aconitoghe_wrightii.xls              No              n              n              2              2              n              n
18 Adansonia_digitata.xls              No              n              n              2              2              n              n
19 Adenanthera_pavonina.xls              No              y              n              2              2              n              n
20 Adenium_obesum.xls              No              n              n              2              2              n              n
21 Aechmea_bianchelliana.xls              No              n              n              2              2              n              n
22 Aechmea_fasciculata.xls              No              n              n              2              2              n              n
23 Aeschynomene_americanica.xls              No              y              n              2              2              n              n
24 Aesculus_hippocastanum.xls              No              n              n              2              2              y              y
25 Agaveparitis_alfarandis.xls              No              NaN              NaN              1              1              y              y
25 rows x 51 columns

In [7]: # Let's return the # of rows
print("Number of Assessments in Data Set = " + str(len(HPWRA1)))
Number of Assessments in Data Set = 2068

In [8]: HPWRA1.shape
Out[8]: (2068, 51)

Manual Data Cleaning

the following changes were made to the UltimateDataCompilation.csv file in Microsoft Excel:
• Some of the missing scores and mislabeled data set were modified.
• re-entered Tabebuia bertonii.xls - was added one column to the left
• re-entered Brya ebenus.pdf - all missing data columns were skipped.
• re-entered Archontophoenix alexandriae.xls - contained 0s instead of y/n
• re-entered Eucalyptus grandis.xls - there was an error on the original file
```

```
In [27]: # need a dictionary
yesno_dict = {'no':'No', 'n':'No', ' ' n':'No', 'n ':'No', 'N':'No', 'N ':'No', 'N ':'No', 'yes':'Yes',
              'y':'Yes', 'y ':'Yes', 'Y':'Yes', 'Y ':'Yes', 'Y ':'Yes', 'Y ':'Yes',
              ' ':None, ' ':None, '0':'None', '0 ':'None', 'None':'None', 'None':'None'}

In [11]: # Global plant history Q's: let's change y/n questions to binary 0s and 1s
HPWRA1["1.01_domesticated"].replace(yesno_dict, inplace=True)

In [13]: HPWRA1["1.02_naturalized_grown"].replace(yesno_dict, inplace=True)
HPWRA1["1.03_weedy_races"].replace(yesno_dict, inplace=True)

In [14]: #check:
HPWRA1["1.01_domesticated"].unique()

Out[14]: array(['Yes', 'No', nan], dtype=object)

In [15]: HPWRA1["1.02_naturalized_grown"].unique()

Out[15]: array(['No', 'Yes', nan], dtype=object)

In [16]: HPWRA1["1.03_weedy_races"].unique()

Out[16]: array(['No', nan, 'Yes'], dtype=object)

In [17]: # we need dictionary for questions about climate suitability
lowmedhigh_dict = {"Low": "Low", "Intermediate": "Intermediate", "Int": "Intermediate", "High": "High",
                   "0": "Low", "1": "Intermediate", "2": "High"}

# Climate suitability Q's: let's change the answers to 0s, 1s and 2s.
HPWRA1["2.01_clim_match"].replace(lowmedhigh_dict, inplace=True)
HPWRA1["2.02_climmatch_qual"].replace(lowmedhigh_dict, inplace=True)

In [18]: #check: there should be only 0,1,2s and NaN left
HPWRA1["2.01_clim_match"].unique()

Out[18]: array(['Intermediate', 'High', 'Low', nan], dtype=object)

In [19]: HPWRA1["2.02_climmatch_qual"].unique()

Out[19]: array(['High', 'Intermediate', 'Low', nan], dtype=object)

In [28]: # More climate suitability Q's, but these ones are y/n: let's change y/n questions to 0s, 1s
HPWRA1["2.03_broad_clim"].replace(yesno_dict, inplace=True)
HPWRA1["2.04_similar_clim"].replace(yesno_dict, inplace=True)
HPWRA1["2.05_repeat_intro"].replace(yesno_dict, inplace=True)

In [21]: #check: make sure there's nothing but 0s,1s, and NaN left
HPWRA1["2.03_broad_clim"].unique()

Out[21]: array(['Yes', 'No', nan], dtype=object)

In [22]: HPWRA1["2.04_similar_clim"].unique()

Out[22]: array(['No', 'Yes', nan], dtype=object)

In [28]: HPWRA1["2.05_repeat_intro"].unique()

Out[28]: array(['Yes', 'No', nan, None], dtype=object)

In [24]: # Q's about how a plant has behaved elsewhere in the world: let's change y/n to 1/0
HPWRA1["3.01_beyond_native"].replace(yesno_dict, inplace=True)
HPWRA1["3.02_disturbance_weed"].replace(yesno_dict, inplace=True)
HPWRA1["3.03_agri_forestry_weed"].replace(yesno_dict, inplace=True)
HPWRA1["3.04_enviro_weed"].replace(yesno_dict, inplace=True)
HPWRA1["3.05_congener"].replace(yesno_dict, inplace=True)

In [25]: #check: make sure there's nothing but 0s,1s, and NaN left
HPWRA1["3.01_beyond_native"].unique()

Out[25]: array(['No', 'Yes', nan], dtype=object)

In [26]: HPWRA1["3.02_disturbance_weed"].unique()

Out[26]: array(['No', 'Yes', nan, None], dtype=object)

In [29]: HPWRA1["3.03_agri_forestry_weed"].unique()

Out[29]: array(['No', 'Yes', nan, None], dtype=object)

In [30]: HPWRA1["3.04_enviro_weed"].unique()

Out[30]: array(['No', 'Yes', nan, None], dtype=object)

In [31]: HPWRA1["3.05_congener"].unique()

Out[31]: array(['No', 'Yes', nan], dtype=object)

In [32]: # Q's about undesirable traits: let's convert y/n to binary 1/0s
HPWRA1["4.01_spyiny"].replace(yesno_dict, inplace=True)
HPWRA1["4.02_allopathic"].replace(yesno_dict, inplace=True)
HPWRA1["4.03_parasitic"].replace(yesno_dict, inplace=True)
HPWRA1["4.04_unpalatable"].replace(yesno_dict, inplace=True)
HPWRA1["4.05_toxic"].replace(yesno_dict, inplace=True)
HPWRA1["4.06_alternate_host"].replace(yesno_dict, inplace=True)
HPWRA1["4.07_allergies"].replace(yesno_dict, inplace=True)
HPWRA1["4.08_fire_hazard"].replace(yesno_dict, inplace=True)
HPWRA1["4.09_shade_tolerant"].replace(yesno_dict, inplace=True)
HPWRA1["4.10_tolerates_soilcond"].replace(yesno_dict, inplace=True)
HPWRA1["4.11_climber"].replace(yesno_dict, inplace=True)
HPWRA1["4.12_forms_thickets"].replace(yesno_dict, inplace=True)

In [33]: #check: make sure there's nothing but 0s,1s, and NaN left
HPWRA1["4.01_spyiny"].unique()

Out[33]: array(['No', 'Yes', nan], dtype=object)

In [34]: HPWRA1["4.02_allopathic"].unique()

Out[34]: array(['No', 'Yes', nan], dtype=object)

In [35]: HPWRA1["4.03_parasitic"].unique()

Out[35]: array(['No', 'Yes', nan], dtype=object)

In [36]: HPWRA1["4.04_unpalatable"].unique()

Out[36]: array(['Yes', nan, 'No', None], dtype=object)

In [37]: HPWRA1["4.05_toxic"].unique()

Out[37]: array(['No', nan, 'Yes'], dtype=object)

In [38]: HPWRA1["4.06_alternate_host"].unique()

Out[38]: array(['No', 'Yes', nan], dtype=object)

In [39]: HPWRA1["4.07_allergies"].unique()

Out[39]: array(['No', nan, 'Yes', None], dtype=object)

In [40]: HPWRA1["4.08_fire_hazard"].unique()

Out[40]: array(['No', nan, 'Yes', None], dtype=object)

In [41]: HPWRA1["4.09_shade_tolerant"].unique()

Out[41]: array(['Yes', nan, 'No', None], dtype=object)

In [42]: HPWRA1["4.10_tolerates_soilcond"].unique()

Out[42]: array(['Yes', 'No', nan], dtype=object)

In [43]: HPWRA1["4.11_climber"].unique()

Out[43]: array(['No', 'Yes', nan], dtype=object)

In [44]: HPWRA1["4.12_forms_thickets"].unique()

Out[44]: array(['No', nan, 'Yes'], dtype=object)

In [45]: # Q's about whether they'll alter habits: these are y/n and should be changed to 1/0
HPWRA1["5.01_aquatic"].replace(yesno_dict, inplace=True)
HPWRA1["5.02_grass"].replace(yesno_dict, inplace=True)
HPWRA1["5.03_nitrogen_fixer"].replace(yesno_dict, inplace=True)
HPWRA1["5.04_geophyte"].replace(yesno_dict, inplace=True)

In [46]: #check: make sure there's nothing but 0s,1s, and NaN left
HPWRA1["5.01_aquatic"].unique()

Out[46]: array(['No', nan, 'Yes'], dtype=object)

In [47]: HPWRA1["5.02_grass"].unique()

Out[47]: array(['No', 'Yes'], dtype=object)

In [48]: HPWRA1["5.03_nitrogen_fixer"].unique()

Out[48]: array(['No', 'Yes', nan], dtype=object)

In [49]: HPWRA1["5.04_geophyte"].unique()

Out[49]: array(['No', 'Yes', nan], dtype=object)

In [58]: # Q's about whether they'll establish in Hawaii: these are y/n and should be changed to 1/0
HPWRA1["6.01_repro_failure"].replace(yesno_dict, inplace=True)
HPWRA1["6.02_viable_seed"].replace(yesno_dict, inplace=True)
HPWRA1["6.03_hybridizes"].replace(yesno_dict, inplace=True)
HPWRA1["6.04_selfcompatible"].replace(yesno_dict, inplace=True)
HPWRA1["6.05_special_pollinators"].replace(yesno_dict, inplace=True)
HPWRA1["6.06_vegetative_repro"].replace(yesno_dict, inplace=True)

In [51]: #check: make sure there's nothing but 0s,1s, and NaN left
HPWRA1["6.01_repro_failure"].unique()

Out[51]: array([nan, 'No', 'Yes'], dtype=object)

In [52]: HPWRA1["6.02_viable_seed"].unique()

Out[52]: array(['No', 'Yes', nan], dtype=object)

In [53]: HPWRA1["6.03_hybridizes"].unique()

Out[53]: array([nan, 'Yes', 'No', None], dtype=object)

In [54]: HPWRA1["6.04_selfcompatible"].unique()

Out[54]: array(['No', nan, 'Yes'], dtype=object)

In [55]: HPWRA1["6.05_special_pollinators"].unique()

Out[55]: array([nan, 'No', 'Yes'], dtype=object)

In [56]: HPWRA1["6.06_vegetative_repro"].unique()

Out[56]: array(['No', 'Yes', nan], dtype=object)

Minimum generation time question

• This is a question about how long a species takes to meet maturity. I binned these values according to Gordon et al 2010 (Guidance for addressing the Australian Weed Risk Assessment Questions, Plant Protection Quarterly 25(2): 56-74 where 1 = 0-2 years; 2 = 2-4 years; 3 = 4-10 years; 4 = >10 years)

In [58]: # making the dictionary
lifespans_dict = {"0":"0-2 years", "1":"0-2 years", "1'":"0-2 years", "1' year":"0-2 years", "2-Jan":"0-2 years", "3or 2":"0-2 years", "1.5-2":"0-2 years", "1.5-2.5":"0-2 years", "3-1":"0-2 years", "2 or 3":"2-4 years", "2-4 years", "2 or 3 years":"2-4 years", "3-Feb":"2-4 years", "3-1":"2-4 years", "4-1":"2-4 years", "5-Mar":"2-4 years", "9","3-1":"2-4 years", "4-1":"4-10 years", "4-1":"4-10 years", ">4":"4-10 years", "4-1":"4-10 years", "4-1":"4-10 years", "5-1":"4-10 years", "5-1":"4-10 years", "6-1":"4-10 years", "7-1":"4-10 years", "7-1":"4-10 years", "8-1":"4-10 years", "9-1":"4-10 years", "10-1":">10 years", "15-1":">10 years", "19-1":">10 years", "20-1":">10 years", "30-1":">10 years", "n":None}

# binning the responses
HPWRA1["6.07_minimum_gen_time"].replace(lifespans_dict, inplace=True)

In [59]: #check: make sure there's nothing but 1,4 and NaN left
HPWRA1["6.07_minimum_gen_time"].unique()

Out[59]: array([nan, '2-4 years', '4-10 years', '0-2 years', '>10 years'], dtype=object)

In [60]: # Q's about how easily a plant is dispersed: these are y/n and should be changed to 1/0
HPWRA1["7.01_unintentional_dispersal"].replace(yesno_dict, inplace=True)
HPWRA1["7.02_intentional_dispersal"].replace(yesno_dict, inplace=True)
HPWRA1["7.03_contaminant_dispersal"].replace(yesno_dict, inplace=True)
HPWRA1["7.04_wind_dispersal"].replace(yesno_dict, inplace=True)
HPWRA1["7.05_water_dispersal"].replace(yesno_dict, inplace=True)
HPWRA1["7.06_bird_dispersal"].replace(yesno_dict, inplace=True)
HPWRA1["7.07_animal_dispersal"].replace(yesno_dict, inplace=True)
HPWRA1["7.08_survive_gut"].replace(yesno_dict, inplace=True)

In [61]: #check: make sure there's nothing but 0s,1s, and NaN left
HPWRA1["7.01_unintentional_dispersal"].unique()

Out[61]: array(['No', 'Yes', nan], dtype=object)

In [62]: HPWRA1["7.02_intentional_dispersal"].unique()

Out[62]: array(['Yes', 'No', nan], dtype=object)

In [63]: HPWRA1["7.03_contaminant_dispersal"].unique()

Out[63]: array(['No', nan, 'Yes'], dtype=object)

In [64]: HPWRA1["7.04_wind_dispersal"].unique()

Out[64]: array(['No', nan, 'Yes'], dtype=object)

In [65]: HPWRA1["7.05_water_dispersal"].unique()

Out[65]: array(['No', nan, 'Yes'], dtype=object)

In [66]: HPWRA1["7.06_bird_dispersal"].unique()

Out[66]: array(['No', 'Yes', nan], dtype=object)

In [67]: HPWRA1["7.07_animal_dispersal"].unique()

Out[67]: array(['No', 'Yes', nan], dtype=object)

In [68]: HPWRA1["7.08_survive_gut"].unique()

Out[68]: array(['No', nan, 'Yes', None], dtype=object)

In [69]: # Q's about how easily a plant is controlled: these are y/n and should be changed to 1/0
HPWRA1["8.01_prolific_seeder"].replace(yesno_dict, inplace=True)
HPWRA1["8.02_propagule_bank"].replace(yesno_dict, inplace=True)
HPWRA1["8.03_herbicide_controlled"].replace(yesno_dict, inplace=True)
HPWRA1["8.04_tolerates_mutilation"].replace(yesno_dict, inplace=True)
HPWRA1["8.05_local_enemies"].replace(yesno_dict, inplace=True)

In [78]: #check: make sure there's nothing but 0s,1s, and NaN left
HPWRA1["8.01_prolific_seeder"].unique()

Out[78]: array(['No', None, nan, 'Yes'], dtype=object)

In [71]: HPWRA1["8.02_propagule_bank"].unique()

Out[71]: array(['No', 'Yes', nan, None], dtype=object)

In [72]: HPWRA1["8.03_herbicide_controlled"].unique()

Out[72]: array([nan, 'Yes', 'No', None], dtype=object)

In [73]: HPWRA1["8.04_tolerates_mutilation"].unique()

Out[73]: array(['Yes', nan, 'No'], dtype=object)

In [74]: HPWRA1["8.05_local_enemies"].unique()

Out[74]: array([nan, 'No', 'Yes', None], dtype=object)

Adding date of assessment and final risk category

The date each assessment was conducted, and the categorical risk assessment was not present on each of the scraped pdfs, so we need to add this information (downloadable here https://sites.google.com/site/weedriskassessment/home). Thankfully, each spreadsheet contains a "Filename" column, so we can match the datasets by these values

In [75]: # read in the necessary fields from the summary csv
#fields = ["Genus","Species","Synonyms","Common name","WRA score","WRA rating","WRA designation","Date","File name"]
HPWRA_summary = pd.read_csv("All_HPWRA_Risk.csv", sep=';', n=, usecols=fields)

In [76]: # Check: make sure it loaded okay and that the newest assessed species are on there
HPWRA_summary.tail()

Out[76]:
   Family      Taxa      Genus      Species  Synonyms  Common name  WRA_score  WRA_rating  WRA_designation  Date      ...  7.05_water_dispersal
2036  Solanaceae  Lycium  Lycium  barbarum              Lycium barbarum  dog berry, hairyberry, Lycium vulgare  15.0      High Risk      H (HPWRA)  8/28/2019  Lycium_barb...
2037  Apocynaceae  Strophanthus  Strophanthus  amboensis  Strophanthus gieskei  elephant vine, wind-blown pisonope  1.0      Evaluate      Evaluate  9/4/2019  Strophanthus_ambon...
2038  Malvaceae  Abroma  augusta  Abroma  augusta  Abroma fastuosum, Antennaria arguta  devil's cotton  4.0      Evaluate      Evaluate  9/9/2019  Abroma_aug...
2039  Dicotyledonaceae  Dicksonia  quercifolia  Dicksonia  squarrosa  Trichomanes squarrosum  harsh tree fern, high tree fern  18.0      High Risk      H (HPWRA)  9/11/2019  Dicksonia_sq...
2040  Myrtaceae  Syzygium  polyanthum  Syzygium  polyanthum  Eugenia polyantha  bayleaf, bayleaf, Indonesian bayleaf  3.0      High Risk      H (HPWRA)  9/13/2019  Syzygium_polyan...

In [77]: #check: how many assessments have been completed?
print("Number of Species that Have been Assessed = " + str(len(HPWRA_summary)))
Number of Species that Have been Assessed = 2041

• The number of assessments scraped from pdf and xls does not match up
• This is because some species have more than one assessment if it has been updated in recent years
• Must be careful to ensure that the Assessment we use is the most recent

In [78]: var1 = len(CompiledData) - len(HPWRA_summary)
print("Number of Species that have more than 1 assessment = " + str(var1))
Number of Species that have more than 1 assessment = 28

In [79]: #merging the
HPWRAall = pd.merge(HPWRA_summary, HPWRA1, on="Filename", how="left")

In [80]: HPWRAall.shape
Out[80]: (2048, 62)

In [81]: HPWRAall.to_csv("HPWRAall.csv", encoding='utf-8-sig', index=False)

In [82]: HPWRAall.head(5)

Out[82]:
   Family      Taxa      Genus      Species  Synonyms  Common name  WRA_score  WRA_rating  WRA_designation  Date      ...  7.05_water_dispersal
0  Fabaceae  Acacia  auriculiformis  Acacia  auriculiformis  NaN  black wattle  13.0      High Risk      H (HPWRA)  10/7/2002  ...  NaN
1  Fabaceae  Acacia  confusa  Acacia  confusa  NaN  Formosan acacia  10.0      High Risk      H (Newsw)  10/7/2002  ...  No
2  Fabaceae  Acacia  melanoxylon  Acacia  melanoxylon  NaN  Australian blackwood  12.0      High Risk      H (HPWRA)  10/7/2002  ...  NaN
3  Euphorbiaceae  Acalypha  hispida  Acalypha  hispida  NaN  Australia park  2.0      Low Risk      L (HPWRA)  10/7/2002  ...  No
4  Euphorbiaceae  Acalypha  wilkesiana  Acalypha  wilkesiana  NaN  blackbox plant  2.0      Low Risk      L (HPWRA)  10/7/2002  ...  No
5 rows x 62 columns

In [83]: # how many assessments didn't get scraped and aren't included in our dataset?
missing_assessments = HPWRAall[HPWRAall["manual_score"] != null].count()

Out[83]: (4, 62)

• The following are species that have HPWRA that could not be included because the original data file is corrupted in some way:

In [84]: missing_assessments.head(5)

Out[84]:
   Family      Taxa      Genus      Species  Synonyms  Common name  WRA_score  WRA_rating  WRA_designation  Date      ...  7.05_water_dispersal
292  Anacardiaceae  Anethrum  hookeri  Anethrum  hookeri  Naif  birds nest arbutum  4.0      Low Risk      L (HPWRA)  9/8/2004  ...  ...
725  Zingiberaceae  Kaempferia  galanga  Kaempferia  galanga  Naif  galanga  3.0      Low Risk      L (HPWRA)  7/2/2009  ...  ...
1122  Fabaceae  Leucaena  KXZ  Leucaena  KXZ  Naif  KXZ  3.0      Evaluate      Evaluate  3/9/2012  ...  ...
1123  Fabaceae  Leucaena  Wundergraz  Leucaena  Wundergraz  Naif  Wundergraz  7.0      High Risk      H (HPWRA)  3/12/2012  ...  ...
4 rows x 62 columns
```