# FSVM: Data Preparation
## Formatting and Taxonomic Grouping

Robert Ritson, WMI Research Associate

2021-07-30



## Formatting Raw Field Data and Rectifying Taxonomy

### Overview

Before we can start modeling species distributions from vegetation survey data with `fsvm`, we need to ensure that the data is properly formatted and that taxonomic classification is accurate. The package includes a variety of functions for accomplishing these tasks which will be illustrated in this document. As there are two major survey types used in the fine scale vegetation modeling data stream - line-point intercept and cover plot - each needs to be handled slightly differently, including interpolation of plots along/within transects. We also have to make sure that species are properly classified so our distribution models are accurate. While your data should be properly formatted prior to attempting to rectify taxonomy, we will illustrate the taxonomic tools first because of their importance to the modeling process. Next we will demonstrate geoprocessing helpers for LPI and COVER survey then end with `lpi2fsvm` which is a composite of all the necessary formatting and taxonomy functions for IDFG type Line-point intercept surveys.

### Taxonomy

Most field surveys use codes or symbols to represent species encountered in a plot or transect. We first need to translate these codes to proper scientific names. The most frequently used codes are those of the United States Department of Agriculture (https://plants.usda.gov/home), which can be interpreted from a look-up table using the `usda_resolve` function. Once we have scientific names for all observations, we need to rectify any synonymous names to ensure they are grouped together during modeling and that the most up-to-date accepted name is used. This is accomplished using tools from the R package `taxize` which includes an API to access the Integrated Taxonomic Information System (https://itis.gov) database maintained by the United States Geological Survey and the Smithsonian Institute. These are used by the function `rectify_taxa` to assign ITIS taxonomic serial numbers (TSNs) and modeling groups by genus (G1), species (G2), subspecies (G3), and variety (G4). Next, we can assign 'TaxonIDs' from IFWIS in order to associate our vegetation data with the IFWIS Taxonomic database using `ifwis_resolve`. Finally, we can classify forage species using `assign_forage` which identifies which species are considered food resources for mule deer, elk, moose, and/or sage-grouse.

Begin by installing the most up-to-date version of `fsvm` from GitLab and loading necessary packages.

```r
#Install latest version of `fsvm`
remotes::install_gitlab("idfg-r/fsvm_package", subdir = "pkg",auth_token = "oYfSyynwxTaobvGua9tF")
> Downloading GitLab repo idfg-r/fsvm_package@HEAD
> from URL https://gitlab.com/api/v4/projects/28272719/repository/archive.tar.gz?sha=HEAD
> utf8 (1.2.1 -> 1.2.2) [CRAN]
> curl (4.3.1 -> 4.3.2) [CRAN]
> Installing 2 packages: utf8, curl
> Installing packages into 'C:/Users/rritson/Documents/R/win-library/4.0'
> (as 'lib' is unspecified)
>
>    There is a binary version available but the source version is later:
>       binary source needs_compilation
> utf8  1.2.1  1.2.2              TRUE
>
>    Binaries will be installed
> package 'utf8' successfully unpacked and MD5 sums checked
> Warning: cannot remove prior installation of package 'utf8'
> Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
> \Users\rritson\Documents\R\win-library\4.0\00LOCK\utf8\libs\x64\utf8.dll to C:
> \Users\rritson\Documents\R\win-library\4.0\utf8\libs\x64\utf8.dll: Permission
> denied
> Warning: restored 'utf8'
> package 'curl' successfully unpacked and MD5 sums checked
> Warning: cannot remove prior installation of package 'curl'
> Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
> \Users\rritson\Documents\R\win-library\4.0\00LOCK\curl\libs\x64\curl.dll to C:
> \Users\rritson\Documents\R\win-library\4.0\curl\libs\x64\curl.dll: Permission
> denied
> Warning: restored 'curl'
>
> The downloaded binary packages are in
>    C:\Users\rritson\AppData\Local\Temp\RtmpaKNUsq\downloaded_packages
>          checking for file 'C:\Users\rritson\AppData\Local\Temp\RtmpaKNUsq\remotes18c415b01a8d\fsvm_pa
>       - preparing 'fsvm': (2s)
>    checking DESCRIPTION meta-information ...     checking DESCRIPTION meta-information ...   v  check
>       - excluding invalid files
>     Subdirectory 'R' contains invalid file names:
>      '.gitkeep'
>    Subdirectory 'man' contains invalid file names:
>      '.gitkeep'
>       - checking for LF line-endings in source and make files and shell scripts
>       - checking for empty or unneeded directories
>       - building 'fsvm_1.3.4.tar.gz'
>
>
> Installing package into 'C:/Users/rritson/Documents/R/win-library/4.0'
> (as 'lib' is unspecified)

#Load packages
lapply(c("fsvm","rgdal","sp","sf","ggplot2"),require,character.only=T)
> Loading required package: fsvm
> Loading required package: rgdal
> Loading required package: sp
```

```
> rgdal: version: 1.5-23, (SVN revision 1121)
> Geospatial Data Abstraction Library extensions to R successfully loaded
> Loaded GDAL runtime: GDAL 3.2.1, released 2020/12/29
> Path to GDAL shared files: C:/Users/rritson/Documents/R/win-library/4.0/rgdal/gdal
> GDAL binary built with GEOS: TRUE
> Loaded PROJ runtime: Rel. 7.2.1, January 1st, 2021, [PJ_VERSION: 721]
> Path to PROJ shared files: C:/Users/rritson/Documents/R/win-library/4.0/rgdal/proj
> PROJ CDN enabled: FALSE
> Linking to sp version:1.4-5
> To mute warnings of possible GDAL/OSR exportToProj4() degradation,
> use options("rgdal_show_exportToProj4_warnings"="none") before loading rgdal.
> Overwritten PROJ_LIB was C:/Users/rritson/Documents/R/win-library/4.0/rgdal/proj
> Loading required package: sf
> Warning: package 'sf' was built under R version 4.0.5
> Linking to GEOS 3.9.0, GDAL 3.2.1, PROJ 7.2.1
> Loading required package: ggplot2
> Warning: package 'ggplot2' was built under R version 4.0.5
> [[1]]
> [1] TRUE
>
> [[2]]
> [1] TRUE
>
> [[3]]
> [1] TRUE
>
> [[4]]
> [1] TRUE
>
> [[5]]
> [1] TRUE
```

### Resolving USDA PLANT Codes

usda_resolve can be used in two ways, translating scientific names into USDA PLANT Codes (resolve = "n2c") or the reverse (resolve = "c2n"). This function will typically be used to translate USDA PLANT Codes into scientific names as these are necessary for modeling. To illustrate this, we will first assign USDA codes to a subset of our field data using getSQLData.

```
## Load dummy data (first 50 records from SQL Database)
fsvm_dummy <- fsvm::getSQLData()[c(1:50),c(2:9)]

## Resolve Taxonomic Name to USDA Plant Code
fsvm_dummy_usda <- fsvm::usda_resolve(dat = fsvm_dummy,
                                      target = "SpeciesName",
                                      resolve = "n2c")

head(fsvm_dummy_usda)
>         TranKey        PlotKey            Source DataType SampleYear
> 1  CPNWH_872726   CPNWH_872726 CPNWH_ID_Herbarium Presence       1970
> 2  CPNWH_878330   CPNWH_878330 CPNWH_ID_Herbarium Presence       1970
> 3  CPNWH_925921   CPNWH_925921 CPNWH_ID_Herbarium Presence       1932
> 4 CPNWH_3266691 CPNWH_3266691 CPNWH_ID_Herbarium Presence       2019
> 5 CPNWH_3266690 CPNWH_3266690 CPNWH_ID_Herbarium Presence       2019
> 6 CPNWH_3266688 CPNWH_3266688 CPNWH_ID_Herbarium Presence       2019
```

```
>   PercentCover PlotArea Symbol
> 1            0        0  TRPE3
> 2            0        0 FRAFA2
> 3            0        0  AMALC
> 4            0        0 ARLA11
> 5            0        0   MYMU
> 6            0        0   OXCO
```

As "n2c" assigns USDA codes to the column 'Symbol', we specify that as our target column to translate back to scientific names.

```
## Resolve USDA Plant Codes to Scientific Name
fsvm_dummy_resolved <- fsvm::usda_resolve(dat = fsvm_dummy_usda, target = "Symbol",
                                          resolve = "c2n")

fsvm_dummy_resolved
>          TranKey         PlotKey              Source DataType SampleYear
> 1   CPNWH_872726   CPNWH_872726  CPNWH_ID_Herbarium Presence       1970
> 2   CPNWH_878330   CPNWH_878330  CPNWH_ID_Herbarium Presence       1970
> 3   CPNWH_925921   CPNWH_925921  CPNWH_ID_Herbarium Presence       1932
> 4  CPNWH_3266691 CPNWH_3266691  CPNWH_ID_Herbarium Presence       2019
> 5  CPNWH_3266690 CPNWH_3266690  CPNWH_ID_Herbarium Presence       2019
> 6  CPNWH_3266688 CPNWH_3266688  CPNWH_ID_Herbarium Presence       2019
> 7  CPNWH_3266689 CPNWH_3266689  CPNWH_ID_Herbarium Presence       2019
> 8  CPNWH_3266692 CPNWH_3266692  CPNWH_ID_Herbarium Presence       2019
> 9  CPNWH_3266693 CPNWH_3266693  CPNWH_ID_Herbarium Presence       2019
> 10 CPNWH_3266694 CPNWH_3266694  CPNWH_ID_Herbarium Presence       2019
> 11 CPNWH_3266695 CPNWH_3266695  CPNWH_ID_Herbarium Presence       2019
> 12 CPNWH_3266696 CPNWH_3266696  CPNWH_ID_Herbarium Presence       2019
> 13 CPNWH_3266699 CPNWH_3266699  CPNWH_ID_Herbarium Presence       2019
> 14 CPNWH_3266701 CPNWH_3266701  CPNWH_ID_Herbarium Presence       2019
> 15 CPNWH_3156331 CPNWH_3156331 CPNWH_EWU_Herbarium Presence       2017
> 16 CPNWH_3268235 CPNWH_3268235  CPNWH_ID_Herbarium Presence       2019
> 17 CPNWH_3268236 CPNWH_3268236  CPNWH_ID_Herbarium Presence       2019
> 18 CPNWH_3266700 CPNWH_3266700  CPNWH_ID_Herbarium Presence       2019
> 19   CPNWH_863840   CPNWH_863840  CPNWH_ID_Herbarium Presence     2003
> 20   CPNWH_983172   CPNWH_983172  CPNWH_ID_Herbarium Presence     2003
> 21   CPNWH_930682   CPNWH_930682  CPNWH_ID_Herbarium Presence     1969
> 22   CPNWH_177850   CPNWH_177850 CPNWH_WTU_Herbarium Presence     1975
> 23   CPNWH_876122   CPNWH_876122  CPNWH_ID_Herbarium Presence     1975
> 24   CPNWH_876141   CPNWH_876141  CPNWH_ID_Herbarium Presence     1975
> 25 CPNWH_2835498 CPNWH_2835498  CPNWH_ID_Herbarium Presence        0
> 26 CPNWH_2791501 CPNWH_2791501  CPNWH_ID_Herbarium Presence     1932
> 27   CPNWH_931479   CPNWH_931479  CPNWH_ID_Herbarium Presence     1938
> 28   CPNWH_996003   CPNWH_996003  CPNWH_ID_Herbarium Presence        0
> 29   CPNWH_926691   CPNWH_926691  CPNWH_ID_Herbarium Presence     1999
> 30   CPNWH_944766   CPNWH_944766  CPNWH_ID_Herbarium Presence     1984
> 31   CPNWH_979264   CPNWH_979264  CPNWH_ID_Herbarium Presence     1985
> 32 CPNWH_2821162 CPNWH_2821162  CPNWH_ID_Herbarium Presence     1912
> 33 CPNWH_3260092 CPNWH_3260092  CPNWH_ID_Herbarium Presence     2019
> 34 CPNWH_3260093 CPNWH_3260093  CPNWH_ID_Herbarium Presence     2019
> 35 CPNWH_3270703 CPNWH_3270703  CPNWH_ID_Herbarium Presence     2019
> 36 CPNWH_3429950 CPNWH_3429950 CPNWH_IDE_Herbarium Presence     1960
> 37 CPNWH_2790327 CPNWH_2790327  CPNWH_ID_Herbarium Presence     1960
> 38 CPNWH_3430676 CPNWH_3430676 CPNWH_IDE_Herbarium Presence     2001
```

```
> 39 CPNWH_2845930 CPNWH_2845930  CPNWH_ID_Herbarium Presence         1913
> 40 CPNWH_3245639 CPNWH_3245639 CPNWH_EWU_Herbarium Presence         1999
> 41 CPNWH_2816317 CPNWH_2816317  CPNWH_ID_Herbarium Presence         1912
> 42 CPNWH_2801952 CPNWH_2801952  CPNWH_ID_Herbarium Presence         1912
> 43 CPNWH_2816364 CPNWH_2816364  CPNWH_ID_Herbarium Presence         1912
> 44  CPNWH_870815  CPNWH_870815  CPNWH_ID_Herbarium Presence         1913
> 45 CPNWH_3432104 CPNWH_3432104 CPNWH_IDE_Herbarium Presence         1951
> 46 CPNWH_3435571 CPNWH_3435571 CPNWH_IDE_Herbarium Presence         1951
> 47  CPNWH_896823  CPNWH_896823  CPNWH_ID_Herbarium Presence         1913
> 48  CPNWH_928466  CPNWH_928466  CPNWH_ID_Herbarium Presence         1933
> 49  CPNWH_999039  CPNWH_999039  CPNWH_ID_Herbarium Presence         1914
> 50  CPNWH_868595  CPNWH_868595  CPNWH_ID_Herbarium Presence         1913
>    PercentCover PlotArea Symbol                        SpeciesName
> 1             0        0  TRPE3                 Trillium petiolatum
> 2             0        0 FRAFA2      Fritillaria affinis var. affinis
> 3             0        0  AMALC  Amelanchier alnifolia var. cusickii
> 4             0        0 ARLA11                Arceuthobium laricis
> 5             0        0   MYMU                    Mycelis muralis
> 6             0        0   OXCO                  Oxalis corniculata
> 7             0        0   <NA>                               <NA>
> 8             0        0  MEOF2                 Melissa officinalis
> 9             0        0  PRVUV      Prunella vulgaris ssp. vulgaris
> 10            0        0  CRCA3                   Crepis capillaris
> 11            0        0   RUUR                       Rubus ursinus
> 12            0        0  RUIDI           Rubus idaeus ssp. idaeus
> 13            0        0   <NA>                               <NA>
> 14            0        0  PHAU7                Phragmites australis
> 15            0        0  TRLA8                 Trifolium latifolium
> 16            0        0   SARI                   Sagittaria rigida
> 17            0        0  ZIPAP     Zizania palustris var. palustris
> 18            0        0  ARCA3           Arceuthobium campylopodum
> 19            0        0  AREL3               Arrhenatherum elatius
> 20            0        0 PEATA2 Penstemon attenuatus var. attenuatus
> 21            0        0 RORU82                    Rosa rubiginosa
> 22            0        0  ALCO2                   Allium columbianum
> 23            0        0  ALCO2                   Allium columbianum
> 24            0        0  ALCO2                   Allium columbianum
> 25            0        0 PAPA20                 Packera paupercula
> 26            0        0   CIUN                  Cirsium undulatum
> 27            0        0   RUPE                      Rubus pedatus
> 28            0        0  LICA2                  Ligusticum canbyi
> 29            0        0 CRSU16               Crataegus suksdorfii
> 30            0        0   ACPL                    Acer platanoides
> 31            0        0  MELO4                Mertensia longiflora
> 32            0        0 PACA15                       Packera cana
> 33            0        0  SCAN2                 Scleranthus annuus
> 34            0        0   <NA>                               <NA>
> 35            0        0  AMAR2             Ambrosia artemisiifolia
> 36            0        0   CHJU                  Chondrilla juncea
> 37            0        0   CHJU                  Chondrilla juncea
> 38            0        0  BEPE3                     Betula pendula
> 39            0        0  SYFOP Symphyotrichum foliaceum var. parryi
> 40            0        0   PHSP                     Phlox speciosa
```

```
> 41           0          0    GRNA                      Grindelia nana
> 42           0          0   COCA5                  Conyza canadensis
> 43           0          0    GRSQ                 Grindelia squarrosa
> 44           0          0    <NA>                               <NA>
> 45           0          0    EUES                   Euphorbia esula
> 46           0          0    EUES                   Euphorbia esula
> 47           0          0   BRAR5                    Bromus arvensis
> 48           0          0   POAR8                Potentilla argentea
> 49           0          0   PLMA4               Plectritis macrocera
> 50           0          0    POSE                        Poa secunda
```

**Rectifying Taxonomy**

Now that we presumably have Latin binomial names from our survey codes, we are ready to attempt `rectify_taxa` and assign our modeling groups. Especially with large data sets (>1000 unique species) this function can take a while to process. Due to occasional failures of the API to communicate with the ITIS online database, processing times can increase with each additional attempt required to query the database. Among the columns appended to the input data upon successful completion of this function are: "AcceptedName", "G1", "G2", "G3", "G4", "G1TSN", "G2TSN", "G3TSN", "G4TSN", "Taxa_Kingdom", "Taxa_Phylum", "Taxa_Class","Taxa_Order","Taxa_Family", and "Common_Name".

```r
#Load Dummy Data (first 50 records from SQL Database)
fsvm_dummy <- fsvm::getSQLData()[c(1:50),c(2:9)]

## Rectify taxonomic names
fsvm_dummy_rectified <- fsvm::rectify_taxa(dat = fsvm_dummy,
                                           scientific.name = "SpeciesName")
> [1] "Begin Rectifying Names..."
> [1] "Iteration 1 of 1"
> [1] "Assigning Accepted Name..."
> Loading required package: taxize
> [1] "Assigning Model Groups by Accepted Name..."
> [1] "Getting G1 TSN..."
> [1] "Getting G2 TSN..."
> [1] "Getting G3 TSN..."
> [1] "Getting G4 TSN..."
> [1] "Getting Hierarchy..."
> [1] "Joining Names and TSNs to Data..."
head(fsvm_dummy_rectified)
>        TranKey        PlotKey            Source DataType SampleYear
> 1  CPNWH_872726   CPNWH_872726 CPNWH_ID_Herbarium Presence       1970
> 2  CPNWH_878330   CPNWH_878330 CPNWH_ID_Herbarium Presence       1970
> 3  CPNWH_925921   CPNWH_925921 CPNWH_ID_Herbarium Presence       1932
> 4 CPNWH_3266691 CPNWH_3266691 CPNWH_ID_Herbarium Presence       2019
> 5 CPNWH_3266690 CPNWH_3266690 CPNWH_ID_Herbarium Presence       2019
> 6 CPNWH_3266688 CPNWH_3266688 CPNWH_ID_Herbarium Presence       2019
>   PercentCover PlotArea                       SpeciesName
> 1            0        0              Trillium petiolatum
> 2            0        0            Fritillaria lanceolata
> 3            0        0 Amelanchier alnifolia var. cusickii
> 4            0        0              Arceuthobium laricis
> 5            0        0                  Mycelis muralis
> 6            0        0                 Oxalis corniculata
>                  AcceptedName           G1                    G2    G3
```

```
> 1              Trillium petiolatum     Trillium       Trillium petiolatum <NA>
> 2 Fritillaria affinis var. affinis  Fritillaria        Fritillaria affinis <NA>
> 3             Amelanchier cusickii  Amelanchier        Amelanchier cusickii <NA>
> 4        Arceuthobium campylopodum Arceuthobium Arceuthobium campylopodum <NA>
> 5                 Mycelis muralis      Mycelis            Mycelis muralis <NA>
> 6             Oxalis corniculata       Oxalis          Oxalis corniculata <NA>
>                                  G4  G1TSN  G2TSN G3TSN  G4TSN Taxa_Kingdom
> 1                                <NA>  43054  43083  <NA>  <NA>      Plantae
> 2 Fritillaria affinis var. affinis  42932 507870  <NA> 531396      Plantae
> 3                                <NA>  25108 508697  <NA>  <NA>      Plantae
> 4                                <NA>  27886  27890  <NA>  <NA>      Plantae
> 5                                <NA> 500432 503893  <NA>  <NA>      Plantae
> 6                                <NA>  29062  29067  <NA>  <NA>      Plantae
>    Taxa_Phylum    Taxa_Class  Taxa_Order   Taxa_Family
> 1 Tracheophyta Magnoliopsida   Liliales Melanthiaceae
> 2 Tracheophyta Magnoliopsida   Liliales    Liliaceae
> 3 Tracheophyta Magnoliopsida    Rosales     Rosaceae
> 4 Tracheophyta Magnoliopsida Santalales   Santalaceae
> 5 Tracheophyta Magnoliopsida  Asterales    Asteraceae
> 6 Tracheophyta Magnoliopsida Oxalidales   Oxalidaceae
>                                                               Common_Name
> 1                                                           Idaho trillium
> 2                               Checker lily, Checker lily, Checker lily
> 3                                                     Cusick's serviceberry
> 4                                                   Western dwarf mistletoe
> 5                                                              Wall-lettuce
> 6 'Ihi, Creeping oxalis, Yellow oxalis, Yellow wood sorrel, Creeping woodsorrel
```

### Resolving IFWIS TaxonIDs

It is also useful for the new vegetation data to be associated with the IFWIS Taxonomic database for Idaho. This function requires a data frame output of `rectify_taxa` since it will use the most specific ITIS taxonomic serial number to assign an IFWIS TaxonID to an observation.

```
## Resolve IFWIS Taxon IDs
fsvm_dummy_ifwis <- fsvm::ifwis_resolve(dat = fsvm_dummy_rectified)
> Loading required package: dplyr
> Warning: package 'dplyr' was built under R version 4.0.5
>
> Attaching package: 'dplyr'
> The following objects are masked from 'package:stats':
>
>     filter, lag
> The following objects are masked from 'package:base':
>
>     intersect, setdiff, setequal, union
> [1] "Loading IFWIS Look-up Table..."
> [1] "Gathering TSNs from Data..."
> [1] "Matching TaxonID by downstream ITIS TSN..."
> [1] "Cleaning up..."
head(fsvm_dummy_ifwis)
>          TranKey        PlotKey         Source  DataType SampleYear
> 1  CPNWH_872726   CPNWH_872726 CPNWH_ID_Herbarium Presence      1970
> 2  CPNWH_878330   CPNWH_878330 CPNWH_ID_Herbarium Presence      1970
```

```
> 3  CPNWH_925921  CPNWH_925921 CPNWH_ID_Herbarium Presence        1932
> 4 CPNWH_3266691 CPNWH_3266691 CPNWH_ID_Herbarium Presence        2019
> 5 CPNWH_3266690 CPNWH_3266690 CPNWH_ID_Herbarium Presence        2019
> 6 CPNWH_3266688 CPNWH_3266688 CPNWH_ID_Herbarium Presence        2019
>   PercentCover PlotArea                              SpeciesName
> 1            0        0                        Trillium petiolatum
> 2            0        0                       Fritillaria lanceolata
> 3            0        0 Amelanchier alnifolia var. cusickii
> 4            0        0                        Arceuthobium laricis
> 5            0        0                          Mycelis muralis
> 6            0        0                         Oxalis corniculata
>                         AcceptedName            G1                        G2  G3
> 1              Trillium petiolatum       Trillium       Trillium petiolatum <NA>
> 2 Fritillaria affinis var. affinis   Fritillaria        Fritillaria affinis <NA>
> 3              Amelanchier cusickii   Amelanchier        Amelanchier cusickii <NA>
> 4          Arceuthobium campylopodum Arceuthobium Arceuthobium campylopodum <NA>
> 5                  Mycelis muralis       Mycelis           Mycelis muralis <NA>
> 6               Oxalis corniculata        Oxalis         Oxalis corniculata <NA>
>                                G4  G1TSN  G2TSN G3TSN  G4TSN Taxa_Kingdom
> 1                             <NA>  43054  43083  <NA>   <NA>      Plantae
> 2 Fritillaria affinis var. affinis  42932 507870  <NA> 531396      Plantae
> 3                             <NA>  25108 508697  <NA>   <NA>      Plantae
> 4                             <NA>  27886  27890  <NA>   <NA>      Plantae
> 5                             <NA> 500432 503893  <NA>   <NA>      Plantae
> 6                             <NA>  29062  29067  <NA>   <NA>      Plantae
>    Taxa_Phylum   Taxa_Class Taxa_Order   Taxa_Family
> 1 Tracheophyta Magnoliopsida   Liliales Melanthiaceae
> 2 Tracheophyta Magnoliopsida   Liliales     Liliaceae
> 3 Tracheophyta Magnoliopsida     Rosales      Rosaceae
> 4 Tracheophyta Magnoliopsida Santalales    Santalaceae
> 5 Tracheophyta Magnoliopsida  Asterales     Asteraceae
> 6 Tracheophyta Magnoliopsida Oxalidales    Oxalidaceae
>                                                                 Common_Name
> 1                                                             Idaho trillium
> 2                                   Checker lily, Checker lily, Checker lily
> 3                                                        Cusick's serviceberry
> 4                                                     Western dwarf mistletoe
> 5                                                                 Wall-lettuce
> 6 'Ihi, Creeping oxalis, Yellow oxalis, Yellow wood sorrel, Creeping woodsorrel
>   TaxonID
> 1   49076
> 2   87798
> 3      NA
> 4   46958
> 5   61669
> 6   49055
```

### Classify Forage Groups

Now that our species observations have been rectified, we next classify their forage utility to mule deer, elk, moose, and/or sage-grouse which can be useful groupings for machine learning species distribution model training. The 'lib_loc' parameter corresponds with the file path of your machine's R library and 'sel' can be one of "all", "forage only", "mule deer", "elk", "moose", or "sage grouse". If any argument other than "all" is used (default), then only vegetation species considered to be forage will be returned in the output data frame.

8

```
## Classifying Forage Groups
fsvm_dummy_forage <- fsvm::assign_forage(fielddata = fsvm_dummy_ifwis,
                                         lib_loc = "C:/Users/rritson/Documents/R/win-library/4.0",
                                         sel = "all")
head(fsvm_dummy_forage)
>        TranKey      PlotKey              Source DataType SampleYear
> 1  CPNWH_872726   CPNWH_872726 CPNWH_ID_Herbarium Presence       1970
> 2  CPNWH_878330   CPNWH_878330 CPNWH_ID_Herbarium Presence       1970
> 3  CPNWH_925921   CPNWH_925921 CPNWH_ID_Herbarium Presence       1932
> 4 CPNWH_3266691 CPNWH_3266691 CPNWH_ID_Herbarium Presence       2019
> 5 CPNWH_3266690 CPNWH_3266690 CPNWH_ID_Herbarium Presence       2019
> 6 CPNWH_3266688 CPNWH_3266688 CPNWH_ID_Herbarium Presence       2019
>   PercentCover PlotArea                        SpeciesName
> 1            0        0             Trillium petiolatum
> 2            0        0           Fritillaria lanceolata
> 3            0        0 Amelanchier alnifolia var. cusickii
> 4            0        0              Arceuthobium laricis
> 5            0        0                  Mycelis muralis
> 6            0        0                Oxalis corniculata
>                     AcceptedName           G1                     G2   G3
> 1             Trillium petiolatum     Trillium      Trillium petiolatum <NA>
> 2 Fritillaria affinis var. affinis  Fritillaria       Fritillaria affinis <NA>
> 3            Amelanchier cusickii  Amelanchier       Amelanchier cusickii <NA>
> 4        Arceuthobium campylopodum Arceuthobium Arceuthobium campylopodum <NA>
> 5             Mycelis muralis       Mycelis          Mycelis muralis <NA>
> 6            Oxalis corniculata      Oxalis        Oxalis corniculata <NA>
>                              G4  G1TSN  G2TSN G3TSN  G4TSN Taxa_Kingdom
> 1                          <NA>  43054  43083  <NA>  <NA>     Plantae
> 2 Fritillaria affinis var. affinis  42932 507870  <NA> 531396     Plantae
> 3                          <NA>  25108 508697  <NA>  <NA>     Plantae
> 4                          <NA>  27886  27890  <NA>  <NA>     Plantae
> 5                          <NA> 500432 503893  <NA>  <NA>     Plantae
> 6                          <NA>  29062  29067  <NA>  <NA>     Plantae
>    Taxa_Phylum   Taxa_Class Taxa_Order   Taxa_Family
> 1 Tracheophyta Magnoliopsida   Liliales Melanthiaceae
> 2 Tracheophyta Magnoliopsida   Liliales    Liliaceae
> 3 Tracheophyta Magnoliopsida    Rosales     Rosaceae
> 4 Tracheophyta Magnoliopsida Santalales   Santalaceae
> 5 Tracheophyta Magnoliopsida  Asterales    Asteraceae
> 6 Tracheophyta Magnoliopsida Oxalidales   Oxalidaceae
>                                                              Common_Name
> 1                                                          Idaho trillium
> 2                                        Checker lily, Checker lily, Checker lily
> 3                                                        Cusick's serviceberry
> 4                                                      Western dwarf mistletoe
> 5                                                               Wall-lettuce
> 6 'Ihi, Creeping oxalis, Yellow oxalis, Yellow wood sorrel, Creeping woodsorrel
>   TaxonID MuleDeer SageGrouse Elk Moose
> 1   49076        N          N   N     N
> 2   87798        N          N   N     N
> 3      NA        Y          N   Y     Y
> 4   46958        N          N   N     N
> 5   61669        N          N   N     N
```

```
> 6    49055        N        N    N    N

## Classify Mule Deer Forage
fsvm_muledeer <- fsvm::assign_forage(fielddata = fsvm_dummy_ifwis,
                                     lib_loc = "C:/Users/rritson/Documents/R/win-library/4.0",
                                     sel = "mule deer")
head(fsvm_muledeer)
>        TranKey      PlotKey             Source DataType SampleYear PercentCover
> 1 CPNWH_925921 CPNWH_925921 CPNWH_ID_Herbarium Presence       1932            0
> 2 CPNWH_870815 CPNWH_870815 CPNWH_ID_Herbarium Presence       1913            0
> 3 CPNWH_868595 CPNWH_868595 CPNWH_ID_Herbarium Presence       1913            0
>   PlotArea                        SpeciesName            AcceptedName
> 1        0 Amelanchier alnifolia var. cusickii    Amelanchier cusickii
> 2        0                  Stipa occidentalis Achnatherum occidentale
> 3        0                         Poa secunda    Puccinellia rupestris
>           G1                      G2   G3   G4  G1TSN  G2TSN G3TSN G4TSN
> 1 Amelanchier    Amelanchier cusickii <NA> <NA>  25108 508697  <NA>  <NA>
> 2 Achnatherum Achnatherum occidentale <NA> <NA> 500933 507950  <NA>  <NA>
> 3 Puccinellia    Puccinellia rupestris <NA> <NA>  41176  41223  <NA>  <NA>
>   Taxa_Kingdom  Taxa_Phylum     Taxa_Class Taxa_Order Taxa_Family
> 1      Plantae Tracheophyta Magnoliopsida     Rosales    Rosaceae
> 2      Plantae Tracheophyta Magnoliopsida      Poales     Poaceae
> 3      Plantae Tracheophyta Magnoliopsida      Poales     Poaceae
>
> 1
> 2 Western needlegrass, California needlegrass, Western needlegrass, Western needle grass, Pubescent we
> 3
>   TaxonID MuleDeer SageGrouse Elk Moose
> 1      NA        Y          N   Y     Y
> 2   76940        Y          N   Y     N
> 3   42930        Y          N   Y     N
```

## Formatting

In order to accurately model fine scale vegetation distributions, we need accurate locations for each species observation in a survey. How this is accomplished depends on the protocol used to collect the data: Line Point Intercept, Macro Cover Plot, or Simple Vegetation Survey. Given the current modeling methodology, we are primarily concerned with interpolating the point intercepts of LPI surveys as generally only the starting point of the transect is recorded (`assign_Lpi`). Future methods may incorporate polygons containing observations, which can be created with the function `assign_Plot`. Point interpolations should be performed on raw data prior before merging with other data. Plot interpolation can be performed at any point as long as relevant metadata of the data sources is available (plot dimensions, azimuths, etc.). It is also helpful to classify the ecological region of the survey locations (`assign_ecoregions`) which can be used for future reference. Finally, we illustrate the entire formatting and taxonomy process for raw IDFG LPI survey data using the composite function `lpi2fsvm`.

### LPI Survey Interpolation

For Line Point Intercept (LPI) surveys, coordinates of each point intercept along a survey line transect need to be interpolated given one of the following: a starting coordinate paired with an azimuth and survey interval, an ending coordinate paired with an azimuth and survey interval, or a starting, middle, and ending coordinates paired with a survey interval. The following example uses raw data from an LPI survey with an azimuth and a start coordinate for each survey. The line is 50 meters with intercepts every 0.5 meters.

```
## Line Point Intercept Interpolation Example
#Load Data
lpi <- data.table::fread(
"A:/Fine scale vegetation analysis/understory_veg_model/data/FieldData/DataFormatting/original_data/dima
)
lpi1 <- lpi[c(1:100),c(3,5,6,12,13,15,16)]
head(lpi1)
>           LineKey PointLoc TopCanopy Easting Northing azimuth distance
> 1: 1402241645214620      0.5      none   292095  4987824       0      100
> 2: 1402241645214620      1.0    artrw8   292095  4987824       0      100
> 3: 1402241645214620      1.5     PHLO2   292095  4987824       0      100
> 4: 1402241645214620     10.0     pssp6   292095  4987824       0      100
> 5: 1402241645214620     10.5      pose   292095  4987824       0      100
> 6: 1402241645214620     11.0     PHLO2   292095  4987824       0      100
plot(lpi1[,c("Easting","Northing")],xlim=c(292090,292140),ylim=c(4987820,4987860),
     main = "LPI Start Location")
```
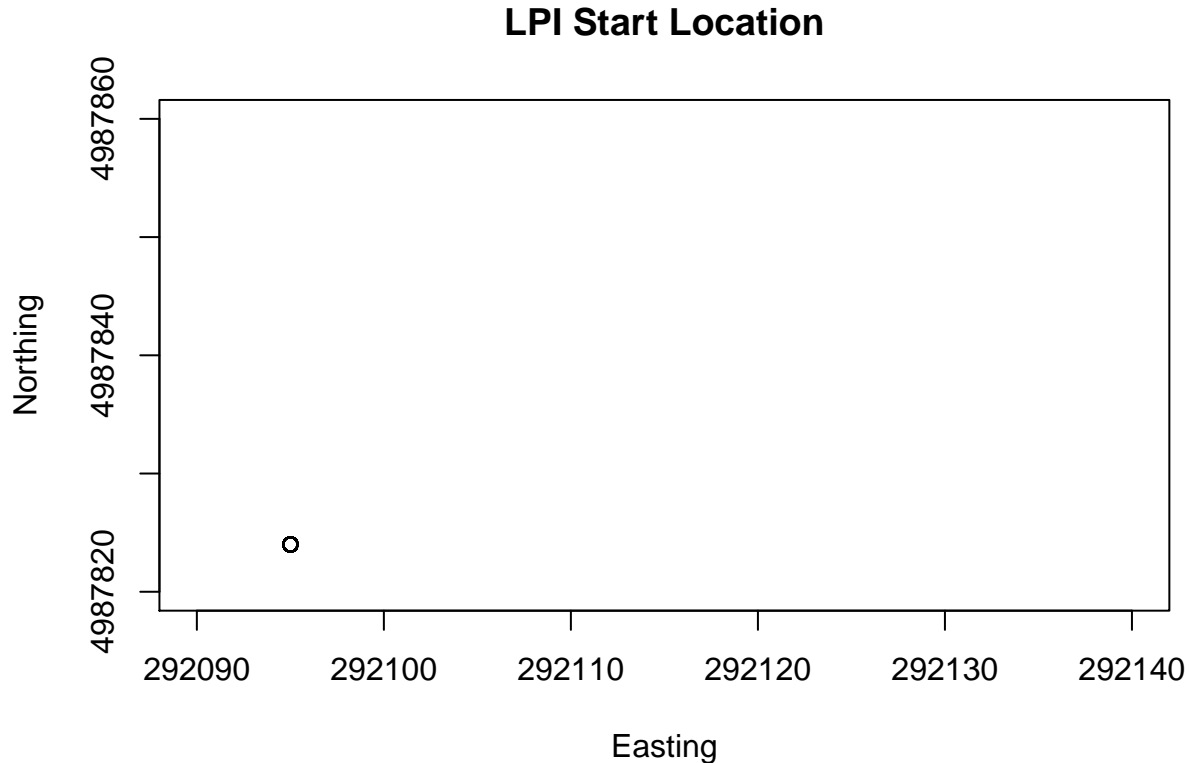


**LPI Start Location**

```
#Assign locations to points in LPI survey
lpi.locs <- fsvm::assign_Lpi(lpi = lpi1, x = "Easting", y = "Northing", ID = "LineKey",
                            interval = "PointLoc", n = "distance", units = "m",
                            azimuth = "azimuth", coord.type = "Start", datum = "IDTM")
head(lpi.locs)
> # A tibble: 6 x 4
>   ParentGlobalID MeterMrkr Easting Northing
>         <int64>      <dbl>   <dbl>    <dbl>
> 1           1e15        0.5 292095. 4987824.
```

```
> 2           1e15       1   292096. 4987825.
> 3           1e15     1.5 292096. 4987825.
> 4           1e15       2   292097. 4987825.
> 5           1e15     2.5 292097. 4987826.
> 6           1e15       3   292098. 4987826.

#Plot
plot(lpi.locs[,c("Easting","Northing")],xlim=c(292090,292140),ylim=c(4987820,4987860),
     main = "Interpolated LPI Points")
```



**Interpolated LPI Points**
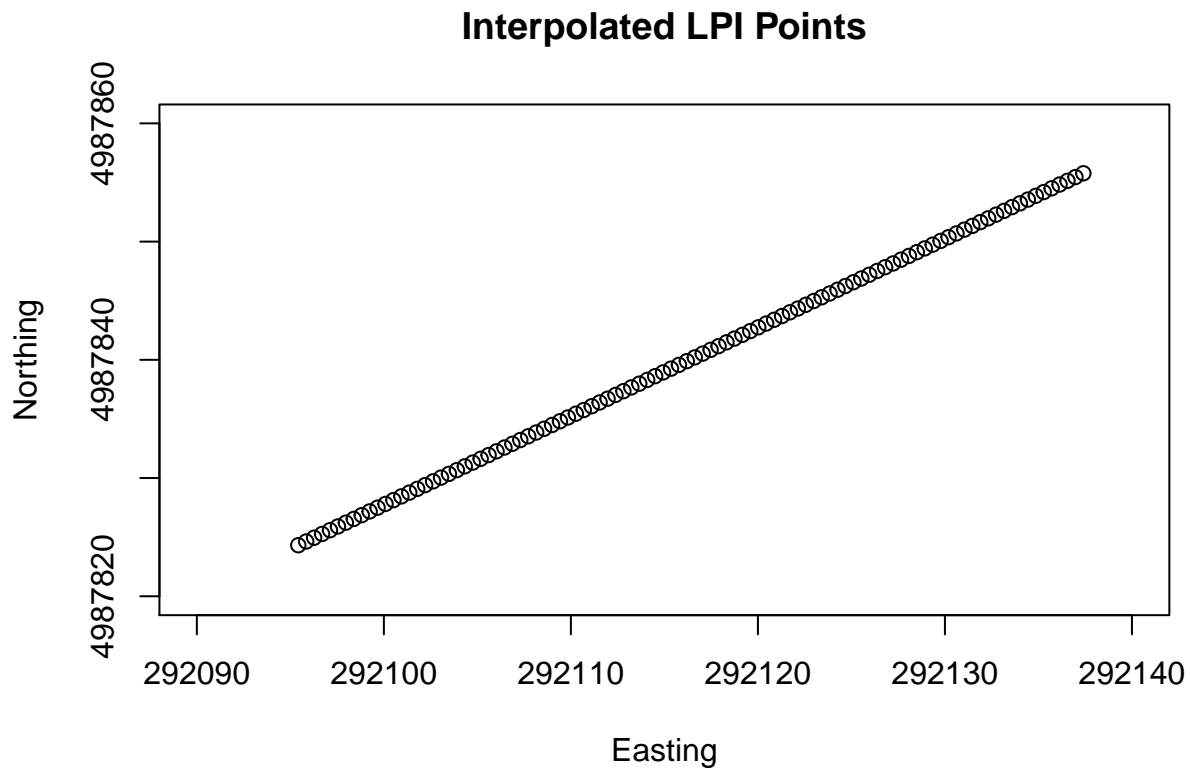
### Create Plot Polygon

The next function creates plot polygons around survey locations and calculates plot area. This includes drawing a specified circular buffer around a survey location like an IDFG Simple Veg Survey (Survey123 app), or a rectangular polygon like with Macro-cover plots. The following example uses a Simple Veg Survey type data point. **For more information on downloading Simple Veg Survey and other survey data from Survey123, please see documentation for `getSurvey123` and the 'Survey 123 Download' vignette.**

```
## Survey 123 Example
#Load Data
svs <- data.table::fread(
"A:/Fine scale vegetation analysis/understory_veg_model/data/FieldData/DataFormatting/original_data/Simp
)
svs <- svs[c(100:101),c(2,4,9,10,11)]
head(svs)
```

```
>                                GlobalID          Date Latitude Longitude
> 1: 61c91854-4ae6-4ad4-8523-a5fe048822dd 7/13/2020 18:00 43.04101 -116.2140
> 2: ba38fd6d-43f8-4387-9398-994beb099b6d 7/13/2020 18:00 43.04112 -116.2145
>    Radius
> 1:      5
> 2:      5
svs.pts <- ggplot(data = svs) +
  geom_point(aes(x = `Longitude`, y = `Latitude`), shape=3, color = "red") +
  ggtitle("Simple Veg Survey Points")
plot(svs.pts)
```

## Simple Veg Survey Points
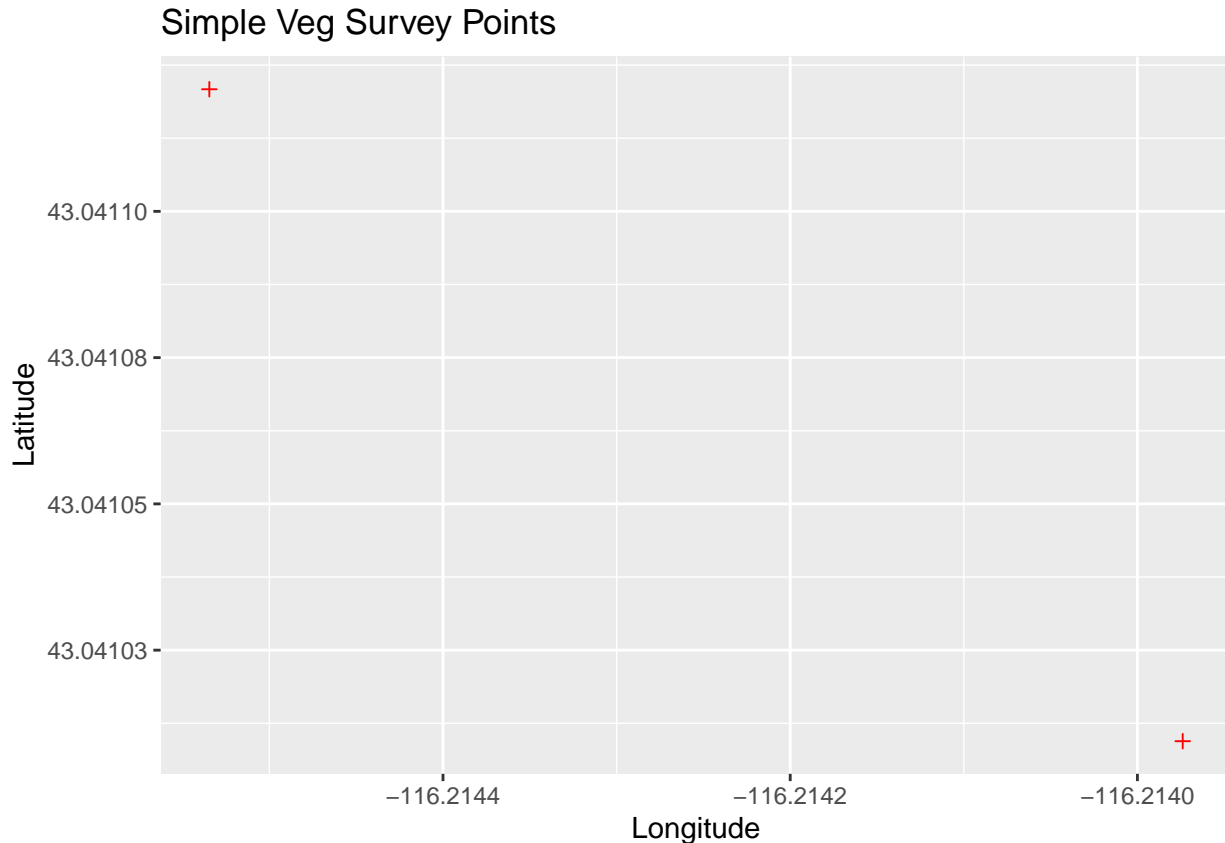


```
#Buffer Simple Veg Survey location by survey radius
svs.locs <- fsvm::assign_Plot(dat = svs, x = "Longitude", y = "Latitude",
                         ID = "GlobalID", size = "Radius", units = "m",
                         type = "SVS", proj = "WGS84")
head(svs.locs)
> Simple feature collection with 2 features and 6 fields
> Geometry type: POLYGON
> Dimension:     XY
> Bounding box:  xmin: 2319609 ymin: 1317966 xmax: 2319665 ymax: 1317990
> CRS:           +proj=tmerc +lat_0=42 +lon_0=-114 +k=0.9996 +x_0=2500000 +y_0=1200000 +datum=NAD83 +uni
>                                GlobalID          Date Radius
> 1 61c91854-4ae6-4ad4-8523-a5fe048822dd 7/13/2020 18:00      5
> 2 ba38fd6d-43f8-4387-9398-994beb099b6d 7/13/2020 18:00      5
>                       geometry Easting Northing  Plot_Ar
> 1 POLYGON ((2319665 1317971, ... 2319660  1317971 78.53982
```

```
> 2 POLYGON ((2319619 1317985, ... 2319614  1317985 78.53982

#Plot
svs.plot <- ggplot(data = svs.locs) + geom_sf(aes(geometry = geometry), fill = "gray") +
  coord_sf(datum = sf::st_crs(
    "+proj=tmerc +lat_0=42 +lon_0=-114 +k=0.9996 +x_0=2500000 +y_0=1200000 +ellps=GRS80 +units=m +no_de
  geom_point(aes(x = `Easting`, y = `Northing`), shape=3, color = "red") +
  ggtitle("Buffered Simple Veg Survey Plots")
plot(svs.plot)
```

### Buffered Simple Veg Survey Plots
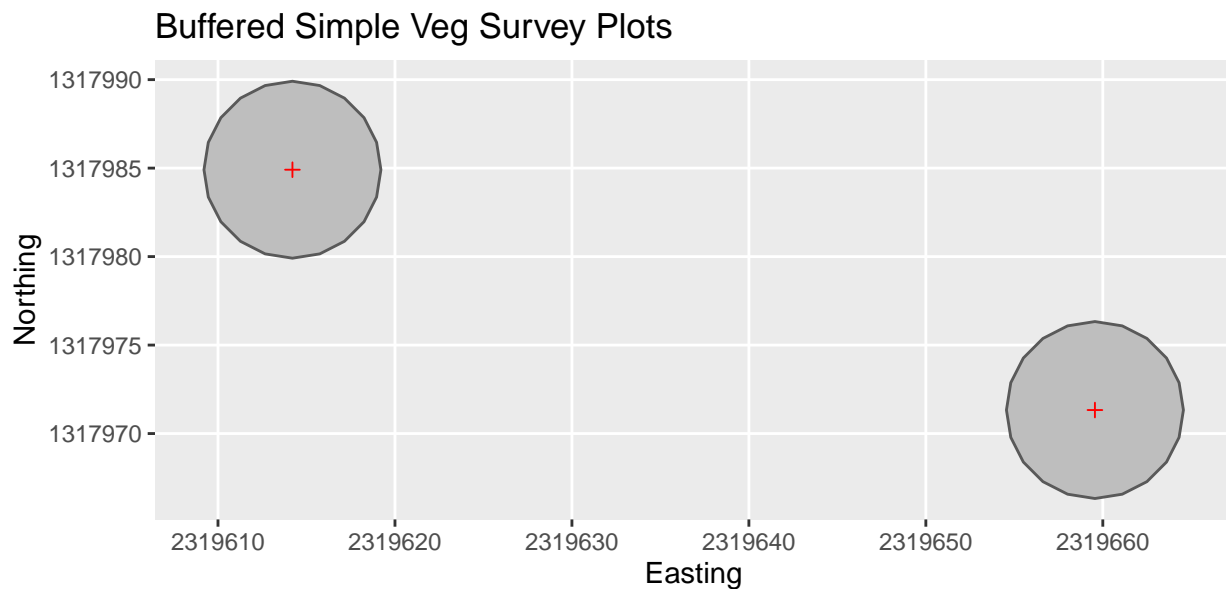


Note that `assign_Plot` automatically projects coordinates to Idaho Transverse Mercator from the original supplied datum (typically WGS84, which is a default). Other source datums can be supplied to the 'proj' argument using a valid proj4string (https://spatialreference.org/ref/).

The following example draws survey polygon for a Macro Cover Plot given the starting location of the baseline and first transect, azimuth of the baseline and first transect, and lengths of the baseline and transects (typically 60 meter baseline with 30 meter transects, defaults).

```
## Cover Plot Example
#Load Data
macro <- data.table::fread(
"A:/Fine scale vegetation analysis/understory_veg_model/data/FieldData/DataFormatting/original_data/sie
)
macro <- macro[,c(1:4,8:11)]
macro$ID <- paste0(macro$Cluster,"_",macro$Macroplot)
macro <- macro[1, c(1,2,9,5:8)]
head(macro)
```

```
>        Zone      Date  ID BaseHeadin TransectHe X_COORD Y_COORD
> 1: Sawtooth 6/23/2017 1_A        185         95  667843 4891037
macro.pts <- ggplot(data = macro) +
  geom_point(aes(`X_COORD`,`Y_COORD`), shape=8, color="red", size=7) +
  ggtitle("Macroplot locations") + xlim(667793,667863) + ylim(4890973,4891038)
plot(macro.pts)
```



Macroplot locations
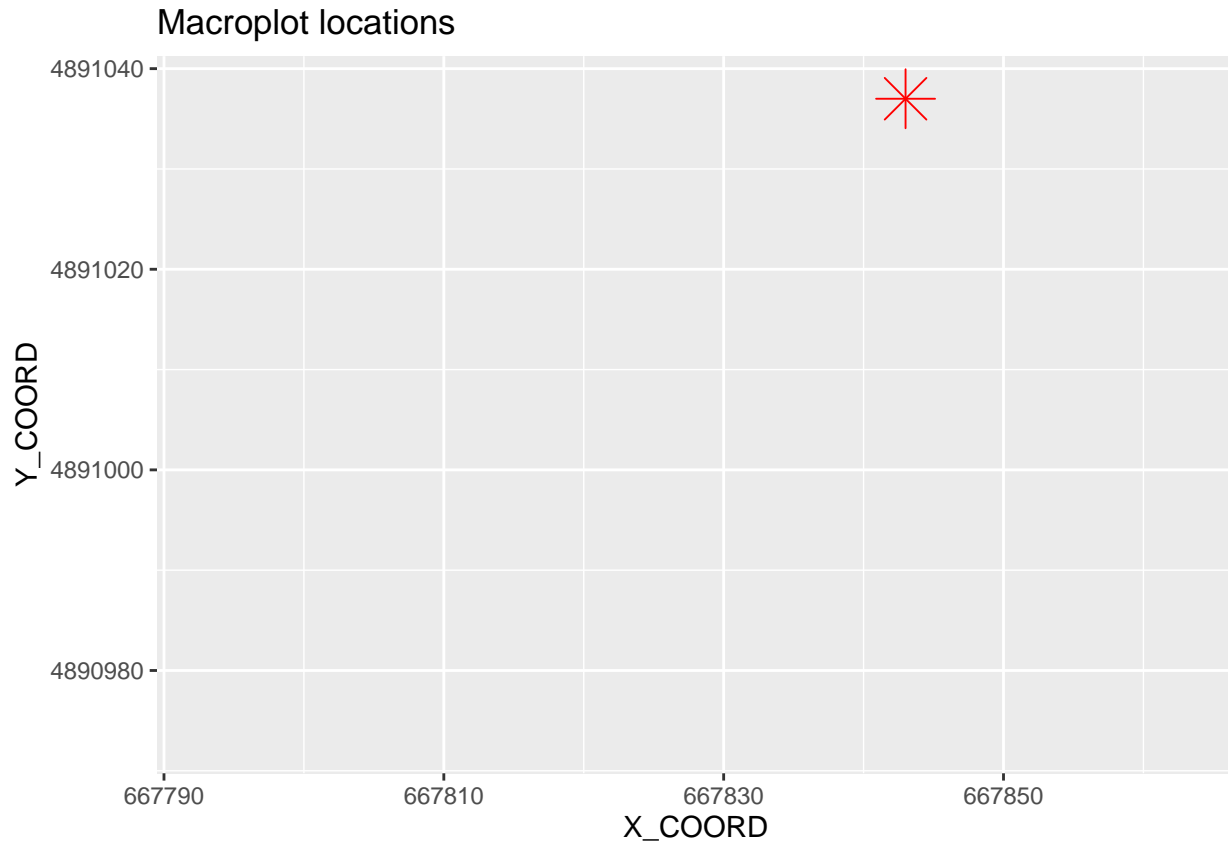
```
#Draw macroplot polygon based on azimuths, distances, and starting location
macro.poly <- fsvm::assign_Plot(dat = macro, x = "X_COORD", y = "Y_COORD", ID = "ID",
                                units = "m", type = "Macro", proj = "IDTM",
                                base_dir = "BaseHeadin", tran_dir = "TransectHe",
                                base_len = 60, tran_len = 30)
head(macro.poly)
> Simple feature collection with 1 feature and 8 fields
> Geometry type: POLYGON
> Dimension:     XY
> Bounding box:  xmin: 667794.2 ymin: 4890974 xmax: 667862.5 ymax: 4891037
> CRS:           +proj=tmerc +lat_0=42 +lon_0=-114 +k=0.9996 +x_0=2500000 +y_0=1200000 +datum=NAD83 +un:
>       Zone      Date  ID BaseHeadin TransectHe                          geometry
> 1 Sawtooth 6/23/2017 1_A        185         95 POLYGON ((667843 4891037, 6...
>   Easting Northing Plot_Ar
> 1  667843  4891037    1800

#Plot
macro.plot <- ggplot(data = macro.poly) + geom_sf(aes(geometry = geometry), fill = "gray") +
  coord_sf(datum = sf::st_crs(
```

```
    "+proj=tmerc +lat_0=42 +lon_0=-114 +k=0.9996 +x_0=2500000 +y_0=1200000 +ellps=GRS80 +units=m +no_de
  geom_point(aes(`Easting`,`Northing`), shape=8, color="red", size=7) +
  ggtitle("Macroplot polygon") + xlim(667793,667863) + ylim(4890973,4891038)
plot(macro.plot)
```



Macroplot polygon
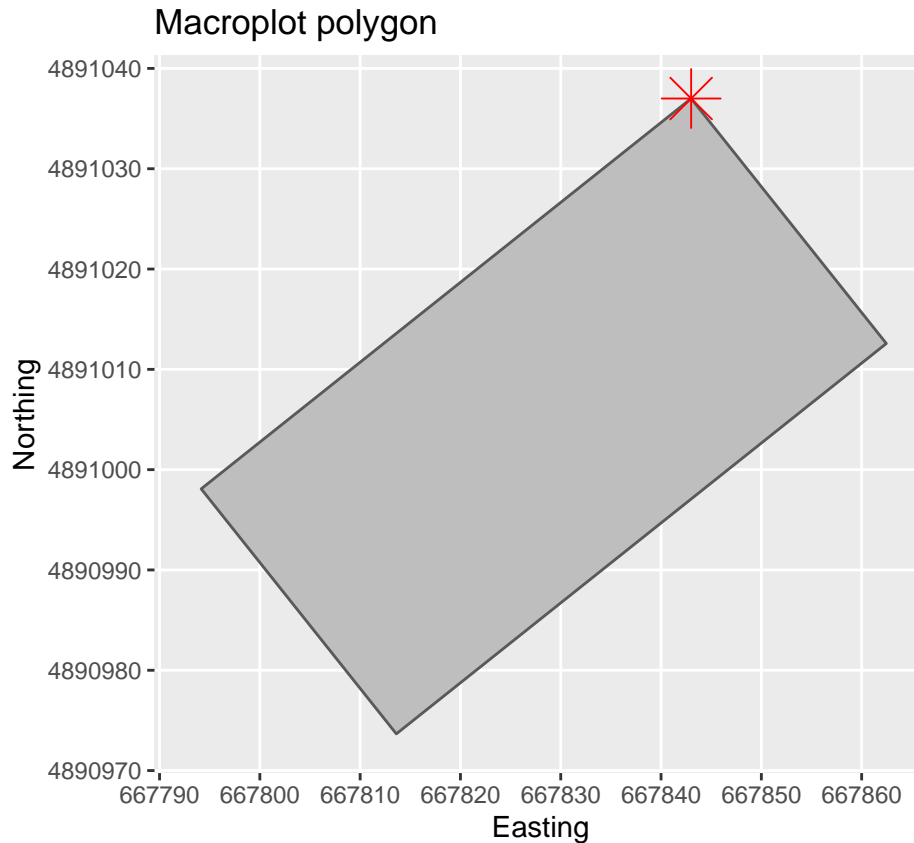
### Classify Eco-Region of field data

Knowing the regional ecological classification of a survey location could be helpful for future analyses or interpretation of results. This is easily accomplished by assigning Bailey's Eco-Regions to field survey locations using a shapefile stored withing the package using `assign_ecoregions`.

```
#Load data
fsvm_dummy <- fsvm::getSQLData()[c(1:50),c(2:9,24:26)]

## Assign Bailey's Eco-Regions to field data
dummy_eco <- fsvm::assign_ecoregions(DAT = fsvm_dummy,
                                     lib_loc = "C:/Users/rritson/Documents/R/win-library/4.0")
> Warning in OGRSpatialRef(dsn, layer, morphFromESRI = morphFromESRI, dumpSRS =
> dumpSRS, : Discarded datum Not_specified_based_on_Clarke_1866_ellipsoid in Proj4
> definition: +proj=aea +lat_0=41 +lon_0=-117 +lat_1=43 +lat_2=48 +x_0=700000
> +y_0=0 +ellps=clrk66 +units=m +no_defs
> OGR data source with driver: ESRI Shapefile
> Source: "C:\Users\rritson\Documents\R\win-library\4.0\fsvm\esri\ecoregions", layer: "Baileys_ecoregion
> with 330 features
> It has 7 fields
> Warning: attribute variables are assumed to be spatially constant throughout all
```

```
> geometries
head(dummy_eco)
>         TranKey        PlotKey              Source DataType SampleYear
> 1  CPNWH_872726  CPNWH_872726 CPNWH_ID_Herbarium Presence       1970
> 2  CPNWH_878330  CPNWH_878330 CPNWH_ID_Herbarium Presence       1970
> 3  CPNWH_925921  CPNWH_925921 CPNWH_ID_Herbarium Presence       1932
> 4 CPNWH_3266691 CPNWH_3266691 CPNWH_ID_Herbarium Presence       2019
> 5 CPNWH_3266690 CPNWH_3266690 CPNWH_ID_Herbarium Presence       2019
> 6 CPNWH_3266688 CPNWH_3266688 CPNWH_ID_Herbarium Presence       2019
>   PercentCover PlotArea                               SpeciesName    QuadPolyID
> 1            0        0                       Trillium petiolatum q47116d8_2146
> 2            0        0                     Fritillaria lanceolata q47116d8_2146
> 3            0        0 Amelanchier alnifolia var. cusickii q47116e7_1021
> 4            0        0                     Arceuthobium laricis q47116e7_5205
> 5            0        0                        Mycelis muralis q47116e7_5639
> 6            0        0                     Oxalis corniculata q47116e7_5645
>   Easting Northing EcoCode
> 1 2274771  1813965   M333A
> 2 2274771  1813965   M333A
> 3 2284256  1828336   M333A
> 4 2287516  1826851   M333A
> 5 2287463  1826671   M333A
> 6 2287575  1826634   M333A
```

**Convert a Raw IDFG Line-point intercept survey to `fsvm` format**

The only required input for this function is a file path to the folder containing the raw LPI files ('frmLPI.csv','tblLPI.csv, and 'LowerLevels.csv'). These are expected to be downloads generated from Survey123 which may or may not be used in future IDFG data collection efforts. The 'export' parameter will save the output data frame to the 'lpi_path' file path named 'LPI_IDFG_FINAL.csv'

```
## View raw IDFG LPI Data
frmLPI <- data.table::fread("A:/Fine scale vegetation analysis/understory_veg_model/data/FieldData/DataI
head(frmLPI)
>    ObjectID                              GlobalID Office      Plot ID
> 1:        1 1d29ed49-daa0-48fc-ba03-ace69ee1923d   IDFG q42112c8_5750
> 2:        2 bd9f8428-021b-43c6-8312-38176e11a25b   IDFG q42112c8_1270
> 3:        3 bb45f077-3be2-4fcb-98cf-ad54b175d96c   IDFG q42112c8_1126
> 4:        4 cb6969ad-5977-43b4-a5ef-328701f43f7f   IDFG q42112c8_5770
> 5:        5 87da484f-7363-4458-8ec1-8b566cfdcc3e   IDFG q42113c1_3152
> 6:        6 431c390e-bfa3-4514-85f1-766e4e8e8c23   IDFG q42113c1_2971
>                Plot Key Line Number Azimuth (Â°)                 Line Key
> 1: IDFG_2019_q42112c8_5750           1          10 IDFG_2019_q42112c8_5750_1
> 2: IDFG_2019_q42112c8_1270           1           2 IDFG_2019_q42112c8_1270_1
> 3: IDFG_2019_q42112c8_1126           1         332 IDFG_2019_q42112c8_1126_1
> 4: IDFG_2019_q42112c8_5770           1           3 IDFG_2019_q42112c8_5770_1
> 5: IDFG_2019_q42113c1_3152           1          68 IDFG_2019_q42113c1_3152_1
> 6: IDFG_2019_q42113c1_2971           1         270 IDFG_2019_q42113c1_2971_1
>        Observer Other Observer     Recorder Other Recorder
> 1: MathenyClayton           NA MillerJennifer           NA
> 2: MathenyClayton           NA MillerJennifer           NA
> 3: MillerJennifer           NA MathenyClayton           NA
> 4: MillerJennifer           NA MathenyClayton           NA
> 5: MathenyClayton           NA MillerJennifer           NA
```

```
> 6: MillerJennifer          NA MathenyClayton          NA
>               Date FormDate   Yr Point Interval (cm)
> 1: 8/9/2019 6:00:00 PM 20190809 2019                50
> 2: 8/9/2019 6:00:00 PM 20190809 2019                50
> 3: 8/9/2019 6:00:00 PM 20190809 2019                50
> 4: 8/9/2019 6:00:00 PM 20190809 2019                50
> 5: 8/8/2019 6:00:00 PM 20190808 2019                50
> 6: 8/8/2019 6:00:00 PM 20190808 2019                50
>    Are you Collecting Vegetation Heights? Height Interval (# of Positions)
> 1:                                      1                                5
> 2:                                      1                                5
> 3:                                      1                                5
> 4:                                      1                                5
> 5:                                      1                                5
> 6:                                      1                                5
>    shrubshape_collect Direction
> 1:                  0   Forward
> 2:                  0   Forward
> 3:                  0   Forward
> 4:                  0   Forward
> 5:                  0   Forward
> 6:                  0   Forward
>                                                                           Notes
> 1:
> 2:
> 3: MARE would have been represented in woody heights, but missed out because listed as herb species
> 4:                                                    Popcorn flower; arabis or sissymbrium?
> 5:                                                           Ermi is a woody species
> 6:                                                    Cacm001= carex; thin?
>    Warning! Before submitting the form ensure that the LPI is not on a repeat/position # that include
> 1:
> 2:
> 3:
> 4:
> 5:
> 6:
>                                          AllUnknownCodesList
> 1:            AFCM003, , , CRCM001, , AFcm004, CRcm001, GACM001
> 2:                                   LUcm001, , , , ,
> 3:                                          , , , , ,
> 4:                      , , , ARCM001, ARcm001, ,
> 5:                      , , , AFcm005, , ACcm001
> 6: CAcm001, CAcm001, CAcm001, CAcm001, Cacm001, , CAcm001, , CAcm001
>    Points Completed Max Woody Height Average Woody Height Max Herbaceous Height
> 1:               50              123             64.44444                   125
> 2:               50              127             69.60000                    77
> 3:               50               NA                   NA                    54
> 4:               50               52             36.85714                    69
> 5:               50               94             49.66667                   108
> 6:               50              134             73.57143                    94
>    Average Herbaceous Height Percent Bare Ground Percent Foliar Cover
> 1:                      49.4                    0                   96
> 2:                      58.9                    2                   88
```

18

```
> 3:                      40.9                        0                       84
> 4:                      40.2                        4                       86
> 5:                      53.7                        0                       88
> 6:                      45.7                        0                      100
>    Percent Basal Cover Percent Standing Dead Cover Percent Biotic Cover
> 1:                   6                          0                      0
> 2:                   4                          0                      0
> 3:                   6                          0                      2
> 4:                   2                          4                     24
> 5:                   4                          0                     22
> 6:                   0                          2                      2
>    Percent Litter Cover Percent Rock Cover Percent Ground Cover
> 1:                   78                 36                  84360
> 2:                   86                 22                    112
> 3:                   68                 34                    110
> 4:                   70                 36                    132
> 5:                   82                 22                    130
> 6:                   92                 12                    106
>          CreationDate Creator             EditDate Editor        x        y
> 1: 8/12/2019 3:00:04 PM        8/12/2019 3:00:04 PM        -112.9761 42.35054
> 2: 8/12/2019 3:00:08 PM        8/12/2019 3:00:08 PM        -112.9716 42.37120
> 3: 8/12/2019 3:00:11 PM        8/12/2019 3:00:11 PM        -112.9710 42.37138
> 4: 8/12/2019 3:00:13 PM        8/12/2019 3:00:13 PM        -112.9757 42.35040
> 5: 8/12/2019 3:00:16 PM        8/12/2019 3:00:16 PM        -113.0410 42.35548
> 6: 8/12/2019 3:00:19 PM        8/12/2019 3:00:19 PM        -113.0401 42.35610

tblLPI <- data.table::fread("A:/Fine scale vegetation analysis/understory_veg_model/data/FieldData/Datal
head(tblLPI)
>    ObjectID                              GlobalID                 RecKey
> 1:        1 68d7bd49-eb21-4320-a5a5-f52f42b59a9e IDFG_2019_q42112c8_5750_1
> 2:        2 de29ce29-a093-4b4e-b405-070a7e134116 IDFG_2019_q42112c8_5750_1
> 3:        3 1164ea4c-0147-43a0-ae36-0200c7f0c31d IDFG_2019_q42112c8_5750_1
> 4:        4 9cf4c018-0304-4994-abc2-91a48e254eed IDFG_2019_q42112c8_5750_1
> 5:        5 996d9211-e18e-4e3e-8a83-f59cbca4930c IDFG_2019_q42112c8_5750_1
> 6:        6 4d4e22a6-adcf-471e-beda-563d4ca3501c IDFG_2019_q42112c8_5750_1
>    PointNbr MeterMrkr PointLabel Canopy (>1.5m) Present Directly Over Point?
> 1:        1       0.5         NA                                          0
> 2:        2       1.0         NA                                          0
> 3:        3       1.5         NA                                          0
> 4:        4       2.0         NA                                          0
> 5:        5       2.5         NA                                          0
> 6:        6       3.0         NA                                          0
>    Top Layer Unknown Code for Top Layer Top Layer Live/Dead? Soil Surface
> 1:     LECI4                                              L            S
> 2:     LECI4                                              L            P
> 3:     LECI4                                              L           GR
> 4:     SYOR2                                              L           GR
> 5:     SYOR2                                              L            S
> 6:     SYOR2                                              L           GR
>    Basal Plant Unknown Code for Basal Basal Live/Dead? Shrub Shape
> 1:                                                            NA
> 2:       ARTR2                               D            NA
> 3:                                                            NA
```

19

```
> 4:                                                          NA
> 5:                                                          NA
> 6:                                                          NA
>    Woody Species Woody Live/Dead? Unknown Code for Woody Woody Height (cm)
> 1:                                                      NA                NA
> 2:                                                      NA                NA
> 3:                                                      NA                NA
> 4:                                                      NA                NA
> 5:         SYOR2               L                        NA                61
> 6:                                                      NA                NA
>    Woody Species 2 Woody 2 Live/Dead? Unknown Code for Woody 2
> 1:
> 2:
> 3:
> 4:
> 5:
> 6:
>    Woody Height 2 (cm) Woody Species 3 Woody 3 Live/Dead?
> 1:                  NA
> 2:                  NA
> 3:                  NA
> 4:                  NA
> 5:                  NA
> 6:                  NA
>    Unknown Code for Woody 3 Woody Height 3 (cm) Herb Species
> 1:                                           NA
> 2:                                           NA
> 3:                                           NA
> 4:                                           NA
> 5:                                           NA        PSSP6
> 6:                                           NA
>    Unknown Code for Herbaceous Herbaceous Live/Dead? Herb Height (cm)
> 1:                                                                  NA
> 2:                                                                  NA
> 3:                                                                  NA
> 4:                                                                  NA
> 5:                                                 L                47
> 6:                                                                  NA
>    Herb Species 2 Unknown Code for Herbaceous 2 Herbaceous 2 Live/Dead?
> 1:
> 2:
> 3:
> 4:
> 5:         AFXXXX                        AFcm004                       L
> 6:
>    Herb Height 2 (cm) Herb Species 3 Unknown Code for Herbaceous 3
> 1:                  NA
> 2:                  NA
> 3:                  NA
> 4:                  NA
> 5:                  24          COLI2
> 6:                  NA
>    Herbaceous 3 Live/Dead? Herb Height 3 (cm) bg fc bc nonzerowoodyhgt
```

```
> 1:                                                 NA   0 1 0               0
> 2:                                                 NA   0 1 1               0
> 3:                                                 NA   0 1 0               0
> 4:                                                 NA   0 1 0               0
> 5:                                 L                17   0 1 0               1
> 6:                                                 NA   0 1 0               0
>    nonzeroherbhgt nonzeroherbhgt2 littercoverc rockcoverc bioticcoverc
> 1:              0               0            0          0            0
> 2:              0               0            0          0            0
> 3:              0               0            1          1            0
> 4:              0               0            1          1            0
> 5:              1               1            1          0            0
> 6:              0               0            1          1            0
>    deadcoverc                          ParentGlobalID        CreationDate Creator
> 1:          0 1d29ed49-daa0-48fc-ba03-ace69ee1923d 8/12/2019 3:00:05 PM
> 2:          0 1d29ed49-daa0-48fc-ba03-ace69ee1923d 8/12/2019 3:00:05 PM
> 3:          0 1d29ed49-daa0-48fc-ba03-ace69ee1923d 8/12/2019 3:00:05 PM
> 4:          0 1d29ed49-daa0-48fc-ba03-ace69ee1923d 8/12/2019 3:00:05 PM
> 5:          0 1d29ed49-daa0-48fc-ba03-ace69ee1923d 8/12/2019 3:00:05 PM
> 6:          0 1d29ed49-daa0-48fc-ba03-ace69ee1923d 8/12/2019 3:00:05 PM
>                 EditDate Editor
> 1: 8/12/2019 3:00:05 PM
> 2: 8/12/2019 3:00:05 PM
> 3: 8/12/2019 3:00:05 PM
> 4: 8/12/2019 3:00:05 PM
> 5: 8/12/2019 3:00:05 PM
> 6: 8/12/2019 3:00:05 PM


LowerLevels <- data.table::fread("A:/Fine scale vegetation analysis/understory_veg_model/data/FieldData,
head(LowerLevels)
>    ObjectID                             GlobalID level_nbr  Lower
> 1:        1 aa15b867-137b-4b46-b041-fbffdb032776         1 CRXXU3
> 2:        2 a8c1a23c-371d-4975-a2cf-740ec34a91cd         2     HL
> 3:        3 3b89c39d-960f-453b-ba9a-263661c93db0         1
> 4:        4 2c69175a-74f5-4aa5-9918-5e39d877cc47         1 SYOR2
> 5:        5 ea473baf-571d-41af-8f8d-51a062414c03         2     HL
> 6:        6 d5113850-47da-4bc6-bb6a-52ddbba5351b         1 PSSP6
>    Unknown Code for Lower ChkboxLower                          ParentGlobalID
> 1:             CRcm001           L 68d7bd49-eb21-4320-a5a5-f52f42b59a9e
> 2:                                 68d7bd49-eb21-4320-a5a5-f52f42b59a9e
> 3:                                 de29ce29-a093-4b4e-b405-070a7e134116
> 4:                               L 1164ea4c-0147-43a0-ae36-0200c7f0c31d
> 5:                                 1164ea4c-0147-43a0-ae36-0200c7f0c31d
> 6:                               L 9cf4c018-0304-4994-abc2-91a48e254eed
>            CreationDate Creator             EditDate Editor
> 1: 8/12/2019 3:00:05 PM         8/12/2019 3:00:05 PM
> 2: 8/12/2019 3:00:05 PM         8/12/2019 3:00:05 PM
> 3: 8/12/2019 3:00:05 PM         8/12/2019 3:00:05 PM
> 4: 8/12/2019 3:00:05 PM         8/12/2019 3:00:05 PM
> 5: 8/12/2019 3:00:05 PM         8/12/2019 3:00:05 PM
> 6: 8/12/2019 3:00:05 PM         8/12/2019 3:00:05 PM


## Using lpi2fsvm function
```

```
lpi<-"A:/Fine scale vegetation analysis/understory_veg_model/data/FieldData/DataFormatting/original_data
IDFG_LPI <- fsvm::lpi2fsvm(lpi_path = lpi,
                           lib_loc = "C:/Users/rritson/Documents/R/win-library/4.0",
                           export = F)
> [1] "Loading Required Files..."
> [1] "Cleaning Files..."
> [1] "Assigning LPI Points..."
> Joining, by = c("TranKey", "PlotKey")
> [1] "Translate Codes..."
> [1] "Rectify Taxa..."
> [1] "Begin Rectifying Names..."
> [1] "Iteration 1 of 1"
> [1] "Assigning Accepted Name..."
> [1] "Assigning Model Groups by Accepted Name..."
> [1] "Getting G1 TSN..."
> [1] "Getting G2 TSN..."
> [1] "Getting G3 TSN..."
> Warning: Unknown or uninitialised column: `current_taxon_id`.
> [1] "Getting G4 TSN..."
> Warning: Unknown or uninitialised column: `current_taxon_id`.
> [1] "Getting Hierarchy..."
> [1] "Joining Names and TSNs to Data..."
> [1] "Assign IFWIS TaxonID..."
> [1] "Loading IFWIS Look-up Table..."
> [1] "Gathering TSNs from Data..."
> [1] "Matching TaxonID by downstream ITIS TSN..."
> [1] "Cleaning up..."
> [1] "Assigning Forage..."
> [1] "Assigning EcoCodes..."
> Warning in OGRSpatialRef(dsn, layer, morphFromESRI = morphFromESRI, dumpSRS =
> dumpSRS, : Discarded datum Not_specified_based_on_Clarke_1866_ellipsoid in Proj4
> definition: +proj=aea +lat_0=41 +lon_0=-117 +lat_1=43 +lat_2=48 +x_0=700000
> +y_0=0 +ellps=clrk66 +units=m +no_defs
> OGR data source with driver: ESRI Shapefile
> Source: "C:\Users\rritson\Documents\R\win-library\4.0\fsvm\esri\ecoregions", layer: "Baileys_ecoregion
> with 330 features
> It has 7 fields
> Warning: attribute variables are assumed to be spatially constant throughout all
> geometries
> [1] "COMPLETE"
head(IDFG_LPI)
> # A tibble: 6 x 30
>   TranKey  PlotKey Source DataType SampleYear Easting Northing Code  SpeciesName
>   <chr>    <chr>   <chr>  <chr>         <int>   <dbl>    <dbl> <chr> <chr>
> 1 IDFG_20~ IDFG_2~ IDFG_~ LPI            2019 2.40e6 1372570. N     NO VEG
> 2 IDFG_20~ IDFG_2~ IDFG_~ LPI            2019 2.40e6 1372570. N     NO VEG
> 3 IDFG_20~ IDFG_2~ IDFG_~ LPI            2019 2.40e6 1372570. SPBE2 Spiraea be~
> 4 IDFG_20~ IDFG_2~ IDFG_~ LPI            2019 2.40e6 1372570. HL    NO VEG
> 5 IDFG_20~ IDFG_2~ IDFG_~ LPI            2019 2.40e6 1372570. AMAL2 Amelanchie~
> 6 IDFG_20~ IDFG_2~ IDFG_~ LPI            2019 2.40e6 1372570. PHMA5 Physocarpu~
> # ... with 21 more variables: AcceptedName <chr>, G1 <chr>, G2 <chr>, G3 <chr>,
> #   G4 <chr>, G1TSN <chr>, G2TSN <chr>, G3TSN <chr>, G4TSN <chr>,
> #   Taxa_Kingdom <chr>, Taxa_Phylum <chr>, Taxa_Class <chr>, Taxa_Order <chr>,
```

```
> #   Taxa_Family <chr>, Common_Name <list>, TaxonID <int>, MuleDeer <chr>,
> #   SageGrouse <chr>, Elk <chr>, Moose <chr>, EcoCode <chr>
```

See 'Extracting QuadPolyIDs' vignette to continue formatting your data.

For details on training or predicting from species distribution models with 'fsvm', please see appropriate vignette.