

HIERARCHICAL CLUSTERING

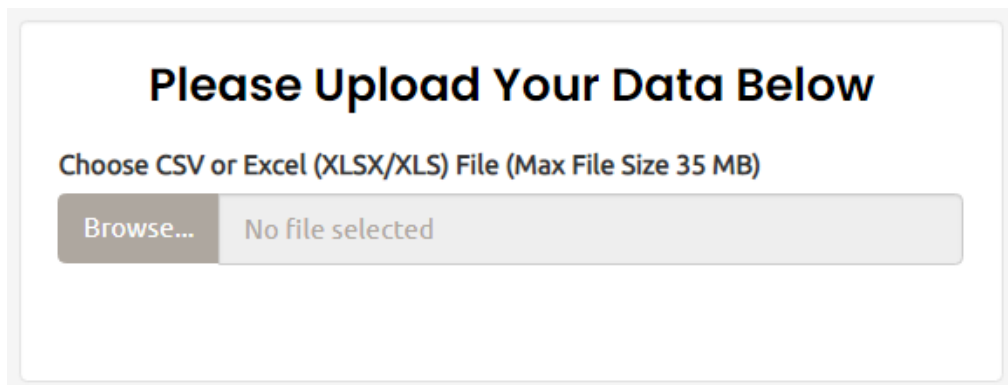
Hierarchical Clustering is an algorithm that groups similar objects into a group (cluster). There are 3 hierarchical clustering methods.

- Single Linkage: A grouping method based on the smallest distance between each object.
- Complete Linkage: A grouping method based on the greatest distance between each object.
- Average Linkage: A grouping method based on the average between each object.

ANALYZING DATA USING HIERARCHICAL CLUSTERING

STEP 1: Uploading Your Data

1. On the upload panel, click Browse and select the data you want to upload. The app will automatically show you the data in “10 Top Data” section.

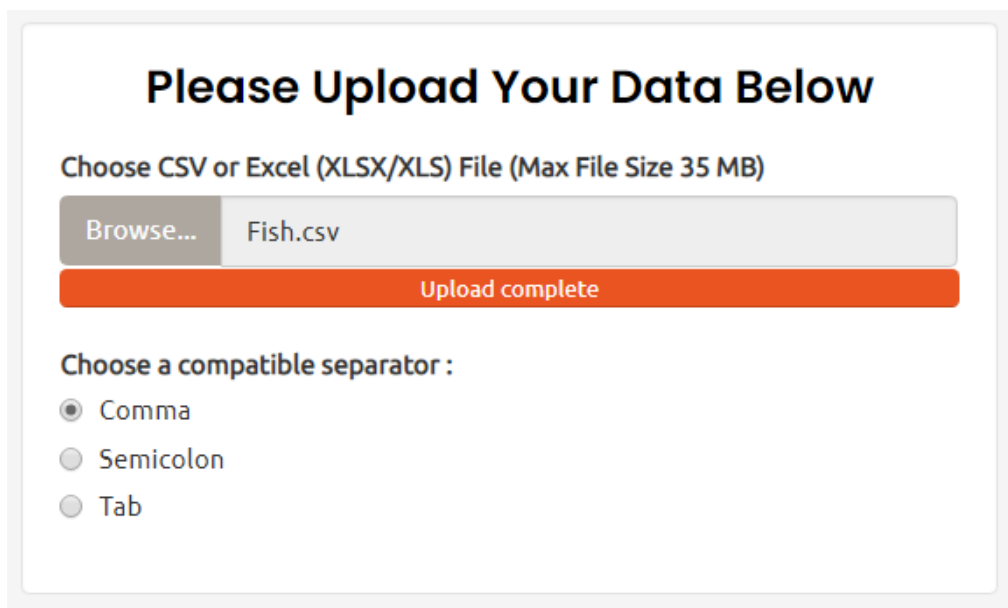


Please Upload Your Data Below

Choose CSV or Excel (XLSX/XLS) File (Max File Size 35 MB)

Browse... No file selected

Note : If you select the data with **csv** format, you need to choose one of the three separators that are available in a selection. Make sure the separator is compatible with the data.



Please Upload Your Data Below

Choose CSV or Excel (XLSX/XLS) File (Max File Size 35 MB)

Browse... Fish.csv

Upload complete

Choose a compatible separator :

☒ Comma

☐ Semicolon

☐ Tab

STEP 2: Selecting Variables

1. Choose hierarchical clustering method. It can be single, complete or average linkage.

Choose Hierarchical Cluster Method
☒ Single Linkage
☐ Complete Linkage
☐ Average Linkage

2. Select the categorical variable. The variable can be numeric, characters or string.

Categorical variable

Species ▼

3. Enter the cluster counts.

Cluster counts

6

Note : The minimum number of cluster counts is 2 with the maximum being the number of categorical variable classes.

STEP 3: Submitting Your Selections

1. After selecting the variables and entering the cluster counts, the submit button will be active and you can click it to show the results in “Summary of HAC” section.

Choose Hierarchical Cluster Method
☐ Single Linkage
☒ Complete Linkage
☐ Average Linkage

Categorical variable

Species ▼

Cluster counts

6

Submit

Introduction
10 Top Data
Summary of HAC
Graphics Result

```

[1] "Selected Method : COMPLETE LINKAGE"

Clustering Methods:
agnes

Cluster sizes:
2 3 4 5 6

Validation Measures:
      2      3      4      5      6

agnes APN  0.0000 0.0357 0.0357 0.0000 0.0238
      AD   0.9303 0.5507 0.3346 0.1076 0.0480
      ADM  0.0000 0.0679 0.0679 0.0000 0.0135
      FOM  0.4171 0.3228 0.2700 0.0985 0.0889

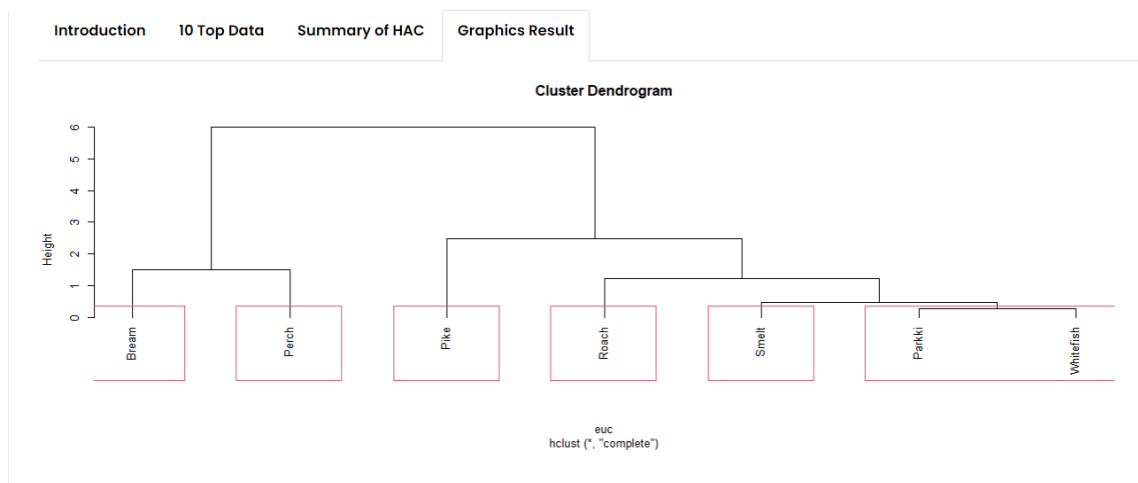
Optimal Scores:

  Score Method Clusters
APN 0.0000 agnes 2
AD  0.0480 agnes 6
ADM 0.0000 agnes 2
FOM 0.0889 agnes 6

Single Linkage Complete Linkage Average Linkage
1      0.9266245      0.9292318      0.9298349

```

- You can also see the graphic result by clicking on the “Graphics Result” tab.



STEP 4: Downloading Results

- The “Download Results” button will appear after the summary result are shown. Click it to export the results to your device in PDF format.

Submit

Download Results