

Report for 02935 Introduction to applied statistics and R for
PhD students, Winter 2025

Ida Bomholt Dyrholm Jacobsen

Summary

Table of contents

Summary	1
Introduction	3
Description of data	4
Scientific question	5
Statistical analyses	6
results	6
Discussion	6
PCA	6
Conclusion	27
Appendix	28
The client	29

Introduction

Description of data

The data process in the following report was collected in Kangerluasunnguaq (Kobbefjord, Nuuk) in Southeast Greenland in 2024. The data was collected in 100 plots placed in an area of interest of approximately 12 km² by randomised stratified sampling. Stratification was based on elevation (5 bins) and NDVI (4 bins).

Within each circular plot of 1 m² the abundance and maximum height of all vascular plant species was recorded. Abundance was assessed by means of the well known Braun-Blanquet scale (8 step version). Abundance was assessed for bryophytes and lichens collectively as well.

In all plots soil temperature, soil moisture, and general vegetation height, was measured as well. Soil temperature was measured with a generic thermometer (°C, 4 measurements, mean calculated). Soil moisture was measured with a ThetaProbe (% water content, 4 measurements, mean calculated). General vegetation height was measured with a generic ruler (cm, 4 measurements, mean calculated)

The final processed data contains data from 100 plots and 72 species of vascular plants.

```
'data.frame':  961 obs. of  11 variables:
 $ plot_name      : chr  "MP002" "MP003" "MP004" "MP005" ...
 $ taxon          : Factor w/ 77 levels "Agrostis mertensii",...: 68 64 26 6 64 6 6 26 64 65 ..
 $ height        : int   15 14 7 6 23 37 23 9 35 2 ...
 $ bb            : Factor w/ 8 levels "0.5","2.5","12.5",...: 3 4 3 4 3 5 3 4 6 3 ...
 $ bb_num        : chr   "12.5" "37.5" "12.5" "37.5" ...
 $ mean_soil_moisture : num   95.42 3.22 25.23 35.2 28.3 ...
 $ mean_soil_temp   : num    7.05 6.5 4 6.9 4.58 ...
 $ mean_veg_height  : num    9.25 2.25 10.75 3 6.25 ...
 $ rowid          : int    1 2 3 4 5 6 7 9 10 11 ...
 $ position       : chr   "taxon_1" "taxon_1" "taxon_1" "taxon_1" ...
 $ other_vegetation_type: logi   NA NA NA NA NA NA ...
```

Scientific question

- Does the most common species have different preferences (= abundance) for soil moisture and/or temperature?
 - *
- What model can describe the relationship between the abundance of the most common species (> 10 observations in plots) and soil moisture?
- Can PCA reveal any kind of structure in the collected data based the abundance of different species?
 - If so, what do these structures indicate?

Statistical analyses

The relationship between soil moisture and abundance of a given species cannot be assumed to be linear. Any given species of plants will have a preference for soil moisture (a wide or narrow interval) where frequency is highest. Abundance of said species will decrease with higher or lower soil moisture than this optimum. Hence, the relation could be considered to have somewhat a bell shape, but uncertain whether this would be symmetrical.

Thus, for the purpose of the analysis a generalized additive model (GAM) have been chosen. This choice was made with emphasis on the fact that this model does not make any assumptions about the relationship of the explanatory and response variable.

- Estimation: Which parameter values t the observations best? How certain are we of our estimates?
- Model check : Are the assumptions on the underlying model fulfilled? Logically this should come rst, but for practical reasons it comes after estimation.
- Simplifying the model (te

Model choice (expected relation ship, data types)

results

Discussion

PCA is a commonly used method in ecology and vegetation science to aid in classification of vegetation structures. This is due to the advantages of

Assumed model

assumption

evaluation

- Statistical analyses
- results

PCA

Can causes the plots to be most different? With 73 taxa and the abundance of bare ground, bryophytes and lichen, what abundances are most nessesary to describe the differences of the plots. Some occurrences might be redundant in explaining the variation between plots.

If the combination of characteristics

What causes wines to be different? • With the 13 characteristics, we can distinguish wines through differences in the charateristics. But are all 13 characteristics necessary? Some may be redundant. • If we can identify scales (linear combinations of the characteristics) where the characteristics vary the most, we can also find a scale that differentiates optimally between the wines.

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2    3.5.1      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr      1.0.2
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()     masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

Attaching package: 'janitor'
```

The following objects are masked from 'package:stats':

```
chisq.test, fisher.test
```

Warning: Values from `bb_num` are not uniquely identified; output will contain list-cols.

```
* Use `values_fn = list` to suppress this warning.
* Use `values_fn = {summary_fun}` to summarise duplicates.
* Use the following dplyr code to identify duplicates.
{data} |>
  dplyr::summarise(n = dplyr::n(), .by = c(plot_name, taxon)) |>
  dplyr::filter(n > 1L)
```

Warning: There were 74 warnings in `mutate()`.

The first warning was:

```
i In argument: `across(-1, ~as.numeric(as.character(.)))`.
```

Caused by warning:

```
! NAs introduced by coercion
```

```
i Run `dplyr::last_dplyr_warnings()` to see the 73 remaining warnings.
```

```
[1] 0
```

```
# A tibble: 100 x 78
```

	plot_name	scirpus_caespitosus	salix_glauca	empetrum_nigrum	betula_nana
	<chr>	<dbl>	<dbl>	<dbl>	<dbl>
1	MP002	12.5	0	0.5	0
2	MP003	0	37.5	37.5	0
3	MP004	0	0	12.5	37.5
4	MP005	0	0	0.5	37.5
5	MP006	0	12.5	62.5	0
6	MP007	0	0	87.5	62.5
7	MP001	0	0	87.5	12.5
8	MP008	0	0	37.5	0.1
9	MP009	0	87.5	62.5	0
10	MP010	0	0	0	0

```
# i 90 more rows
```

```
# i 73 more variables: salix_herbacea <dbl>, carex_bigelowii <dbl>,
# vaccinium_uliginosum <dbl>, oxyria_digyna <dbl>, huperzia_selago <dbl>,
# gnaphalium_supinum <dbl>, poa_alpina <dbl>, poa_glauca <dbl>,
# saxifraga_oppositifolia <dbl>, juncus_trifidus <dbl>,
# angelica_archangelica <dbl>, luzula_spicata <dbl>, carex_glareosa <dbl>,
# deschampsia_flexuosa <dbl>, polygonum_viviparum <dbl>, ...
```


tibble [100 x 78] (S3: tbl_df/tbl/data.frame)

\$ plot_name	: chr [1:100]	"MP002" "MP003" "MP004" "MP005" ...
\$ scirpus_caespitosus	: num [1:100]	12.5 0 0 0 0 0 0 0 0 0 ...
\$ salix_glauca	: num [1:100]	0 37.5 0 0 12.5 0 0 0 87.5 0 ...
\$ empetrum_nigrum	: num [1:100]	0.5 37.5 12.5 0.5 62.5 87.5 87.5 37.5 62.5 0 ...
\$ betula_nana	: num [1:100]	0 0 37.5 37.5 0 62.5 12.5 0.1 0 0 ...
\$ salix_herbacea	: num [1:100]	0 0 0 0 0 0 0 0.5 0 12.5 ...
\$ carex_bigelowii	: num [1:100]	12.5 0 2.5 0.1 0 0.1 2.5 2.5 0.5 0.1 ...
\$ vaccinium_uliginosum	: num [1:100]	37.5 0 0 12.5 37.5 0 37.5 2.5 12.5 0 ...
\$ oxyria_digyna	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ huperzia_selago	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ gnaphalium_supinum	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ poa_alpina	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ poa_glauca	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ saxifraga_oppositifolia	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ juncus_trifidus	: num [1:100]	0 0.5 0 0 0 0 0 0 0 12.5 ...
\$ angelica_archangelica	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ luzula_spicata	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ carex_glareosa	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ deschampsia_flexuosa	: num [1:100]	0.1 2.5 0 0 0 0 0.5 0.1 0 12.5 ...
\$ polygonum_viviparum	: num [1:100]	0 0 0 0 0 0 0 0 0.01 0 ...
\$ phyllodoce_coerulea	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ luzula_multiflora	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ luzula_confusa	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ silene_acaulis	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ carex_rariflora	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ loiseleuria_procumbens	: num [1:100]	0 0 0 0.1 0 0 0 0.5 0 0 ...
\$ salix_arctophila	: num [1:100]	0 0 0 0 0 0 0 0 0.1 0 ...
\$ festuca_brachyphylla	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ ledum_groenlandicum	: num [1:100]	0 0 0 0 0 0 2.5 0 0.1 0 ...
\$ lycopodium_annotinum	: num [1:100]	0 0 0 0 0 12.5 12.5 0 0 0 ...
\$ eriophorum_angustifolium	: num [1:100]	0.5 0 0 0 0 0 0 0 0 0 ...
\$ chamaenerion_latifolium	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ rhodiola_rosea	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ calamagrostis_langsdorfii	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ festuca_rubra	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ juniperus_communis	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ diphasiastrum_complanatum	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ campanula_gieseckiana	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ luzula_parviflora	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ gymnocarpium_dryopteris	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ taraxacum_croceum	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ plantago_maritima	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ linnaea_borealis_ssp_americana	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ deschampsia_alpina	: num [1:100]	0 0 0 0 0 2.5 0 0 0 0 ...
\$ cerastium_alpinum	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ sibbaldia_procumbens	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ listera_cordata	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ potentilla_tridentata	: num [1:100]	0 0 0 0 0 0 0 0.1 0 0 ...
\$ carex_brunnescens	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ tofieldia_pusilla	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ dryopteris_assimilis	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ diphasiastrum_alpinum	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ pyrola_grandiflora	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ potentilla_crantzii	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ poa_pratensis	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ carex_scirpoidea	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...
\$ bartsia_alpina	: num [1:100]	0 0 0 0 0 0 0 0 0 0 ...

```

$ coptis_trifolia      : num [1:100] 0.1 0 0 0 0 0 0 0 0 0 0 ...
$ veronica_wormskjoldii : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ agrostis_mertensii   : num [1:100] 0 0 0 0 0 0 0 0 0.1 0 0 ...
$ veronica_alpina      : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ harrimanella_hypnoides : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ diapensia_lapponica  : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ stellaria_calycantha : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ scirpis_caespitosus  : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ chamaenerion_angustifolium : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ alchemilla_alpina    : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ pedicularis_lapponica : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ viscaria_alpina      : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ pedicularis_flammea  : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ equisetum_silvaticum : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ hieracium_hyparcticum : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ poa_nemoralis        : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ carex_canescens      : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ phleum_commutatum    : num [1:100] 0 0 0 0 0 0 0 0 0 0 0 ...
$ bryophyte            : num [1:100] 87.5 12.5 87.5 87.5 37.5 12.5 62.5 0 2.5 2.5 ...
$ lichen               : num [1:100] 0 62.5 12.5 87.5 12.5 2.5 0 12.5 0 62.5 ...
$ bareground           : num [1:100] 0 37.5 0.01 0 0 0 0 12.5 12.5 62.5 ...

```

Attaching package: 'ChemometricsWithR'

The following objects are masked from 'package:stats':

loadings, screeplot

```

[1] "scores"          "loadings"        "var"             "totalvar"
[5] "centered.data"

```

PCA model of a mean-centered matrix of 100 by 76
Number of PCs to cover 90 percent of the variance: 37

	Var	Cumul. var.
PC 1	6.834400	6.83440
PC 2	6.073946	12.90835
PC 3	5.068934	17.97728
PC 4	4.840477	22.81776
PC 5	4.163365	26.98112
PC 10	2.822849	43.38922

	PC 1	PC 2	PC 3	PC 4
scirpus_caespitosus	-0.02379151	0.0002786357	-0.02915946	-0.01321870
salix_glauca	-0.13993305	-0.0066342397	0.30125279	0.03762653
empetrum_nigrum	-0.11550968	0.0642820311	-0.07727665	-0.02161595

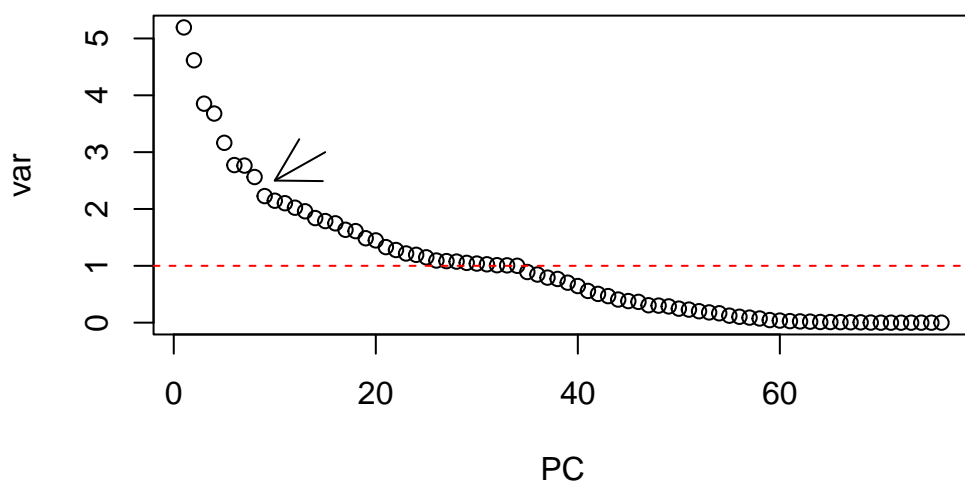
	PC 5	PC 6	PC 7	PC 8	PC 9
scirpus_caespitosus	0.04257509	-0.04575855	0.016432593	-0.03762496	0.08254694
salix_glauca	-0.02042728	-0.05851153	-0.039278744	0.02001079	0.04712893
empetrum_nigrum	0.13115038	-0.05850372	-0.001136057	-0.15122685	0.12824624

	PC 10	PC 11	PC 12	PC 13	PC 14
scirpus_caespitosus	-0.23692846	0.07849664	-0.08862788	0.05318152	-0.19908100
salix_glauca	0.07408672	0.05266168	0.05789762	-0.02524768	0.06867752
empetrum_nigrum	0.06436084	-0.01469251	-0.11630637	0.23551154	-0.03877864

	PC 15	PC 16	PC 17	PC 18	PC 19
--	-------	-------	-------	-------	-------

scirpus_caespitosus	0.37349004	-0.072123167	0.35881616	-0.1301318	-0.04458802
salix_glauca	-0.03004669	-0.003994224	0.08604947	0.1621147	-0.32050032
empetrum_nigrum	-0.06142780	0.144779168	-0.01675500	0.2780506	0.10421265
	PC 20	PC 21	PC 22	PC 23	PC 24
scirpus_caespitosus	0.22176634	-0.08660349	-0.001123991	0.05652370	0.03515630
salix_glauca	-0.08557510	0.10030238	0.007940519	0.10162646	0.16675243
empetrum_nigrum	-0.04258597	-0.22503363	-0.266103595	-0.06709857	0.02671033
	PC 25	PC 26	PC 27	PC 28	
scirpus_caespitosus	-0.02262733	-0.005142075	-0.001449962	-0.005948812	
salix_glauca	-0.07222090	0.067558690	-0.025919178	-0.043869272	
empetrum_nigrum	0.08047646	-0.028741651	0.093663947	0.010941079	
	PC 29	PC 30	PC 31	PC 32	
scirpus_caespitosus	0.001028597	0.0004828267	0.002920071	0.004146281	
salix_glauca	-0.040172805	-0.0287644746	0.046955796	0.006608961	
empetrum_nigrum	-0.015297020	-0.0376444308	0.002533122	-0.002096815	
	PC 33	PC 34	PC 35	PC 36	
scirpus_caespitosus	0.004111769	0.003558343	0.003401163	-0.030120357	
salix_glauca	0.026392798	0.021818894	-0.005471512	-0.003555740	
empetrum_nigrum	-0.009586569	0.025056143	0.015555584	-0.009012015	
	PC 37	PC 38	PC 39	PC 40	
scirpus_caespitosus	-0.004345183	-0.0210863272	0.02676775	-0.04851726	
salix_glauca	0.081860202	0.0003150324	-0.01410896	0.01194866	
empetrum_nigrum	0.195484110	0.1266890769	-0.11602912	0.04361899	
	PC 41	PC 42	PC 43	PC 44	PC 45
scirpus_caespitosus	0.10265984	0.03533248	-0.04250135	-0.22244679	0.01966824
salix_glauca	-0.01827668	0.01774871	-0.39848282	-0.09454788	-0.19566519
empetrum_nigrum	-0.34532044	-0.03100345	0.21885931	0.12377472	-0.05838907
	PC 46	PC 47	PC 48	PC 49	PC 50
scirpus_caespitosus	0.01787691	0.01933158	0.43735187	-0.2332758	0.3292791
salix_glauca	-0.21670535	0.02515220	-0.31668906	-0.1228940	0.1863840
empetrum_nigrum	0.10171339	-0.08973389	0.04898743	0.1862022	0.0885980
	PC 51	PC 52	PC 53	PC 54	PC 55
scirpus_caespitosus	0.03253064	0.2004875	-0.16108044	0.15035201	0.07016883
salix_glauca	0.08428987	0.2855922	0.04960067	0.06951756	-0.02347552
empetrum_nigrum	0.04044612	0.2986558	-0.23138054	0.08918307	-0.02218976
	PC 56	PC 57	PC 58	PC 59	
scirpus_caespitosus	-0.01430575	-0.01215051	-0.003881416	0.002995895	
salix_glauca	0.10317242	0.24815821	0.262690303	-0.001208050	
empetrum_nigrum	0.06435506	0.30438071	-0.069513559	0.012796508	
	PC 60	PC 61	PC 62	PC 63	
scirpus_caespitosus	0.006429032	0.005282158	0.02953841	-0.001755620	
salix_glauca	0.003885018	-0.028060309	0.03138839	0.002979555	
empetrum_nigrum	0.022429641	0.051423893	0.01303850	0.017818231	
	PC 64	PC 65	PC 66	PC 67	
scirpus_caespitosus	-0.01863068	-0.0005312578	0.008031640	0.005287108	
salix_glauca	-0.01900699	-0.0001926705	-0.013403964	-0.003641698	
empetrum_nigrum	0.02278277	0.0400219889	0.003070548	0.017190797	
	PC 68	PC 69	PC 70	PC 71	
scirpus_caespitosus	0.0020057887	0.007297768	0.0002603157	9.512622e-05	
salix_glauca	0.0145075311	0.017454714	-0.0001072518	-1.367580e-04	
empetrum_nigrum	0.0009543681	-0.016645871	0.0004053946	3.376961e-04	
	PC 72	PC 73	PC 74	PC 75	
scirpus_caespitosus	3.550208e-06	1.851124e-16	0.000000e+00	0.000000e+00	
salix_glauca	5.569003e-06	7.549419e-17	3.589501e-17	-1.341151e-16	
empetrum_nigrum	-2.072272e-05	-1.189144e-16	2.918691e-16	-5.831357e-17	
	PC 76				
scirpus_caespitosus	0.000000e+00				
salix_glauca	9.622918e-17				

empetrum_nigrum -3.057764e-16



```
#| label: pca-x-t-eigenvalues-1
#| echo: false
#| output: false

X<-var(scale(pca_stat))
T<-eigen(X)$vectors
sum(is.na(X))
```

```
[1] 0
```

```
head(T,n=3)
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	-0.02379151	-0.0002786357	0.02915946	-0.01321870	-0.04257509	0.04575855
[2,]	-0.13993305	0.0066342397	-0.30125279	0.03762653	0.02042728	0.05851153
[3,]	-0.11550968	-0.0642820311	0.07727665	-0.02161595	-0.13115038	0.05850372
	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]
[1,]	-0.016432593	0.03762496	-0.08254694	-0.23692846	0.07849664	0.08862788
[2,]	0.039278744	-0.02001079	-0.04712893	0.07408672	0.05266168	-0.05789762
[3,]	0.001136057	0.15122685	-0.12824624	0.06436084	-0.01469251	0.11630637
	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]
[1,]	-0.05318152	-0.19908100	0.37349004	0.072123167	0.35881616	0.1301318
[2,]	0.02524768	0.06867752	-0.03004669	0.003994224	0.08604947	-0.1621147
[3,]	-0.23551154	-0.03877864	-0.06142780	-0.144779168	-0.01675500	-0.2780506
	[,19]	[,20]	[,21]	[,22]	[,23]	[,24]
[1,]	0.04458802	-0.22176634	-0.08660349	0.001123991	0.05652370	0.03515630
[2,]	0.32050032	0.08557510	0.10030238	-0.007940519	0.10162646	0.16675243
[3,]	-0.10421265	0.04258597	-0.22503363	0.266103595	-0.06709857	0.02671033
	[,25]	[,26]	[,27]	[,28]	[,29]	
[1,]	-0.02262733	-0.005142075	0.001449962	0.005948812	-0.001028597	
[2,]	-0.07222090	0.067558690	0.025919178	0.043869272	0.040172805	
[3,]	0.08047646	-0.028741651	-0.093663947	-0.010941079	0.015297020	
	[,30]	[,31]	[,32]	[,33]	[,34]	
[1,]	0.0004828267	-0.002920071	0.004146281	-0.004111769	-0.003558343	
[2,]	-0.0287644746	-0.046955796	0.006608961	-0.026392798	-0.021818894	
[3,]	-0.0376444308	-0.002533122	-0.002096815	0.009586569	-0.025056143	
	[,35]	[,36]	[,37]	[,38]	[,39]	
[1,]	-0.003401163	-0.030120357	0.004345183	-0.0210863272	-0.02676775	
[2,]	0.005471512	-0.003555740	-0.081860202	0.0003150324	0.01410896	

```

[3,] -0.015555584 -0.009012015 -0.195484110 0.1266890769 0.11602912
      [,40]      [,41]      [,42]      [,43]      [,44]      [,45]
[1,] 0.04851726 0.10265984 0.03533248 0.04250135 -0.22244679 -0.01966824
[2,] -0.01194866 -0.01827668 0.01774871 0.39848282 -0.09454788 0.19566519
[3,] -0.04361899 -0.34532044 -0.03100345 -0.21885931 0.12377472 0.05838907
      [,46]      [,47]      [,48]      [,49]      [,50]      [,51]
[1,] 0.01787691 0.01933158 0.43735187 0.2332758 -0.3292791 -0.03253064
[2,] -0.21670535 0.02515220 -0.31668906 0.1228940 -0.1863840 -0.08428987
[3,] 0.10171339 -0.08973389 0.04898743 -0.1862022 -0.0885980 -0.04044612
      [,52]      [,53]      [,54]      [,55]      [,56]      [,57]
[1,] -0.2004875 0.16108044 0.15035201 -0.07016883 -0.01430575 -0.01215051
[2,] -0.2855922 -0.04960067 0.06951756 0.02347552 0.10317242 0.24815821
[3,] -0.2986558 0.23138054 0.08918307 0.02218976 0.06435506 0.30438071
      [,58]      [,59]      [,60]      [,61]      [,62]
[1,] 0.003881416 -0.002995895 -0.006429032 0.005282158 0.02953841
[2,] -0.262690303 0.001208050 -0.003885018 -0.028060309 0.03138839
[3,] 0.069513559 -0.012796508 -0.022429641 0.051423893 0.01303850
      [,63]      [,64]      [,65]      [,66]      [,67]
[1,] -0.001755620 0.01863068 -0.0005312578 0.008031640 -0.005287108
[2,] 0.002979555 0.01900699 -0.0001926705 -0.013403964 0.003641698
[3,] 0.017818231 -0.02278277 0.0400219889 0.003070548 -0.017190797
      [,68]      [,69]      [,70]      [,71]      [,72]
[1,] 0.0020057887 0.007297768 -0.0002603157 -9.512622e-05 3.550208e-06
[2,] 0.0145075311 0.017454714 0.0001072518 1.367580e-04 5.569003e-06
[3,] 0.0009543681 -0.016645871 -0.0004053946 -3.376961e-04 -2.072272e-05
      [,73]      [,74]      [,75]      [,76]
[1,] 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
[2,] -6.104022e-14 -4.481232e-14 3.544326e-14 5.959669e-15
[3,] 2.464478e-13 1.367634e-13 -1.067870e-13 -1.575131e-14

```

```

Lambda<-t(T)%*%X%*%T
round(Lambda, digits=3)

```

```

      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
[1,] 5.194 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[2,] 0.000 4.616 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[3,] 0.000 0.000 3.852 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[4,] 0.000 0.000 0.000 3.679 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[5,] 0.000 0.000 0.000 0.000 3.164 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[6,] 0.000 0.000 0.000 0.000 0.000 2.772 0.000 0.000 0.000 0.000 0.000 0.000
[7,] 0.000 0.000 0.000 0.000 0.000 0.000 2.763 0.000 0.000 0.000 0.000 0.000
[8,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 2.563 0.000 0.000 0.000 0.000
[9,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 2.227 0.000 0.000 0.000
[10,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 2.145 0.000 0.000
[11,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 2.102 0.000
[12,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 2.022
[13,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[14,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[15,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[16,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[17,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[18,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[19,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[20,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[21,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[22,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[23,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
[24,] 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000

```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[27,]	0	0	0	0
[28,]	0	0	0	0
[29,]	0	0	0	0
[30,]	0	0	0	0
[31,]	0	0	0	0
[32,]	0	0	0	0
[33,]	0	0	0	0
[34,]	0	0	0	0
[35,]	0	0	0	0
[36,]	0	0	0	0
[37,]	0	0	0	0
[38,]	0	0	0	0
[39,]	0	0	0	0
[40,]	0	0	0	0
[41,]	0	0	0	0
[42,]	0	0	0	0
[43,]	0	0	0	0
[44,]	0	0	0	0
[45,]	0	0	0	0
[46,]	0	0	0	0
[47,]	0	0	0	0
[48,]	0	0	0	0
[49,]	0	0	0	0
[50,]	0	0	0	0
[51,]	0	0	0	0
[52,]	0	0	0	0
[53,]	0	0	0	0
[54,]	0	0	0	0
[55,]	0	0	0	0
[56,]	0	0	0	0
[57,]	0	0	0	0
[58,]	0	0	0	0
[59,]	0	0	0	0
[60,]	0	0	0	0
[61,]	0	0	0	0
[62,]	0	0	0	0
[63,]	0	0	0	0
[64,]	0	0	0	0
[65,]	0	0	0	0
[66,]	0	0	0	0
[67,]	0	0	0	0
[68,]	0	0	0	0
[69,]	0	0	0	0
[70,]	0	0	0	0
[71,]	0	0	0	0
[72,]	0	0	0	0
[73,]	0	0	0	0
[74,]	0	0	0	0
[75,]	0	0	0	0
[76,]	0	0	0	0

T[,1]

[1]	-0.023791507	-0.139933049	-0.115509679	-0.060987867	0.176681495
[6]	-0.021593321	-0.078278012	0.236909365	0.221066129	0.194602741
[11]	0.255542557	0.324472862	0.034806896	0.013229959	0.101461065
[16]	0.275594535	0.008793231	-0.083385687	0.062401479	-0.025606073
[21]	0.062473359	0.052339366	-0.007657970	0.002480708	-0.002196198

```
[26] 0.006799823 0.040421943 -0.044839537 -0.084489919 -0.028367259
[31] 0.041308315 0.248260975 -0.032667893 0.008793231 -0.023588232
[36] -0.010568506 -0.072798308 0.018482643 -0.065422089 0.036347026
[41] 0.008793231 -0.084312228 -0.021490072 0.246518875 0.193100855
[46] 0.010656960 -0.025697801 -0.016349741 -0.012593350 -0.020375772
[51] -0.078403059 0.168610701 -0.021347349 0.171270076 -0.019498893
[56] 0.005564110 0.002168323 -0.041023953 0.006254542 0.315956123
[61] 0.053862625 -0.019461870 -0.054828988 -0.012908725 -0.083128288
[66] -0.018691661 -0.020509952 0.025204272 -0.013429377 0.008793940
[71] -0.099416596 -0.061529064 0.298144876 -0.041801973 -0.014256728
[76] 0.156155449
```

```
round(T[,1],digits=2)
```

```
[1] -0.02 -0.14 -0.12 -0.06 0.18 -0.02 -0.08 0.24 0.22 0.19 0.26 0.32
[13] 0.03 0.01 0.10 0.28 0.01 -0.08 0.06 -0.03 0.06 0.05 -0.01 0.00
[25] 0.00 0.01 0.04 -0.04 -0.08 -0.03 0.04 0.25 -0.03 0.01 -0.02 -0.01
[37] -0.07 0.02 -0.07 0.04 0.01 -0.08 -0.02 0.25 0.19 0.01 -0.03 -0.02
[49] -0.01 -0.02 -0.08 0.17 -0.02 0.17 -0.02 0.01 0.00 -0.04 0.01 0.32
[61] 0.05 -0.02 -0.05 -0.01 -0.08 -0.02 -0.02 0.03 -0.01 0.01 -0.10 -0.06
[73] 0.30 -0.04 -0.01 0.16
```

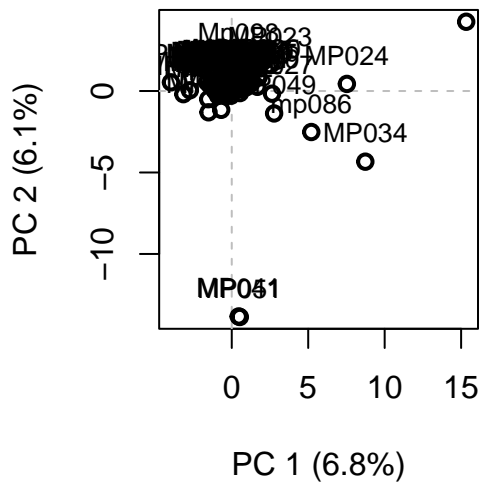
```
sum(diag(Lambda))
```

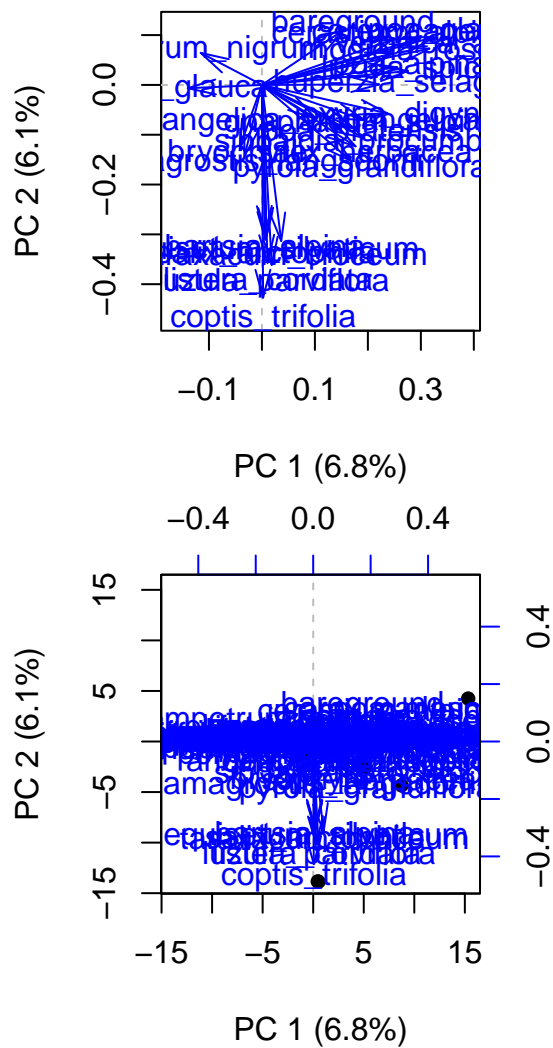
```
[1] 76
```

```
plot_names <- levels(as.factor(wide_data2$plot_name))
```

```
scoreplot(mp.PC, lwd=2)
```

```
text(mp.PC$scores[,1], mp.PC$scores[,2], labels=plot_names, pos=3, cex=0.8)
```

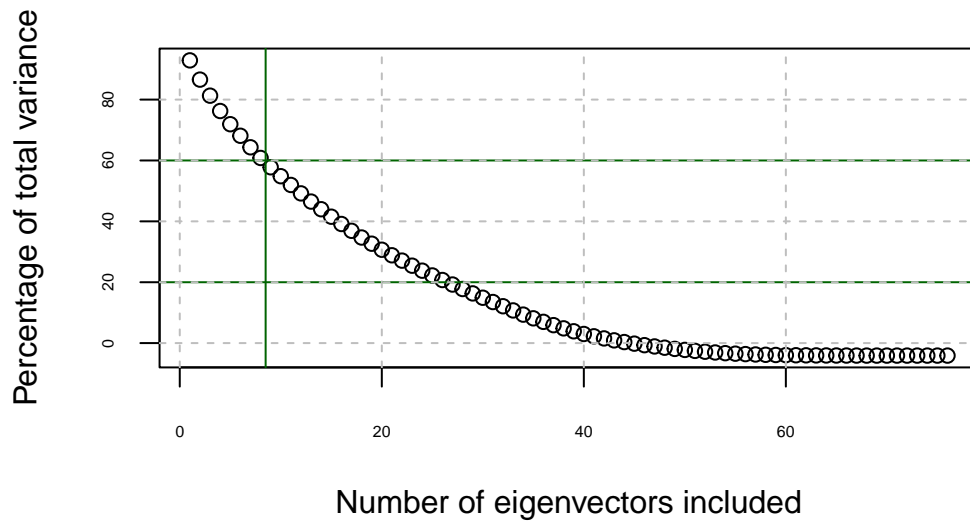




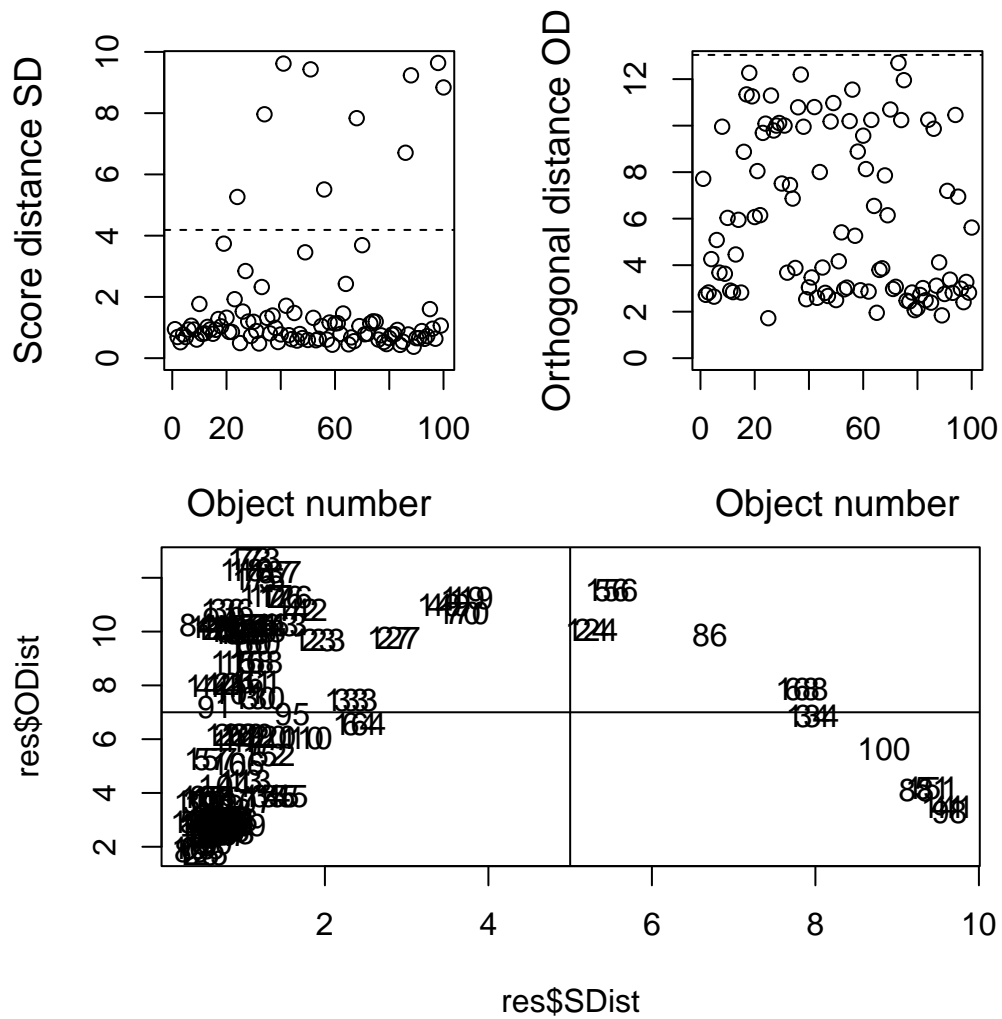
```
#| label: plot-percentage-variance-unexplained
#| echo: false
#| out-width: 100%

plot(100*(73-cumsum(diag(Lambda)))/73,type="b",
main="Percentage Variance Unexplained",
xlab='Number of eigenvectors included',
ylab='Percentage of total variance',
  cex.axis = 0.5) # Reduce size of axis text)
abline(h = 20, col = "darkgreen")
abline(h = 60, col = "darkgreen")
abline(v = 8.5, col = "darkgreen")
grid(nx = NULL, ny = NULL,
      lty = 2,          # Grid line type
      col = "gray",     # Grid line color
      lwd = 1)         # Grid line width
```


Percentage Variance Unexplained

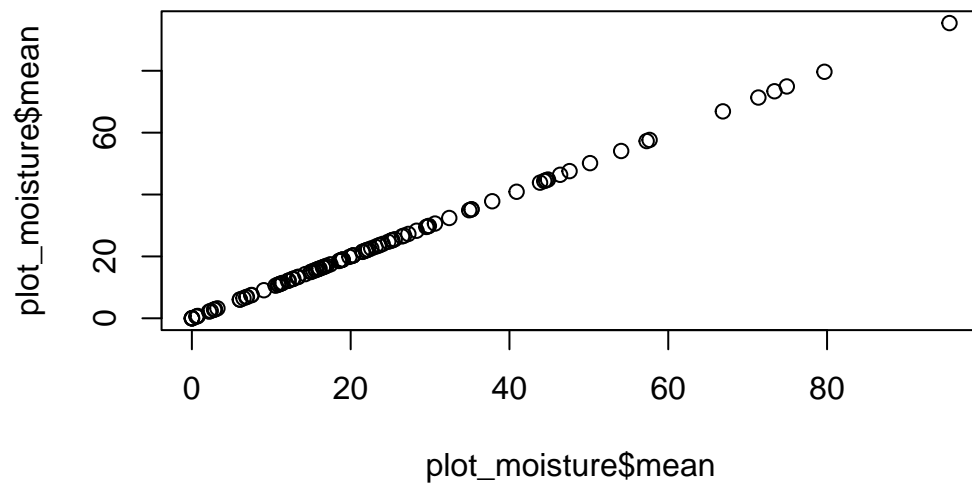


Loading required package: rpart



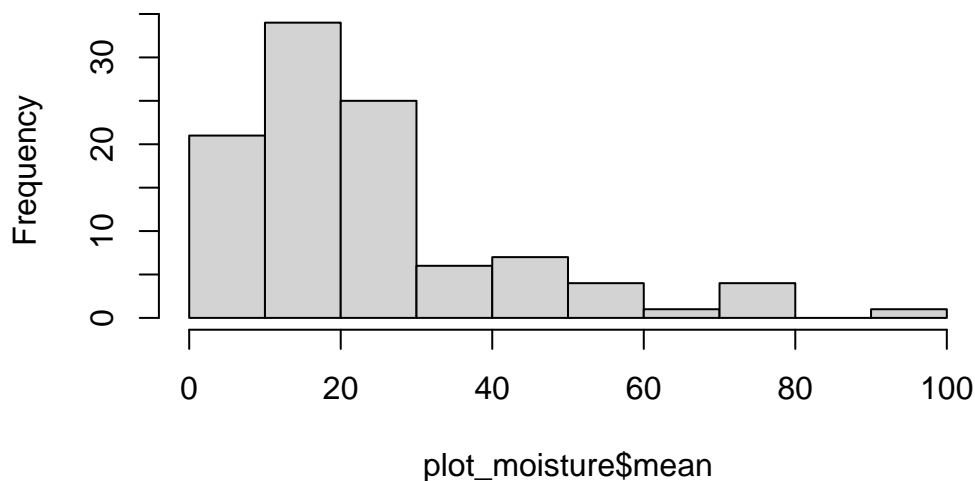
```
plot_moisture <- stat_mappingplants |>
  group_by(plot_name) |>
  summarize(mean = mean(mean_soil_moisture))
```

```
plot(x = plot_moisture$mean, y = plot_moisture$mean)
```



```
hist(plot_moisture$mean)
```

Histogram of plot_moisture\$mean



```
# # Install and load vegan package
# install.packages("vegan")
# library(vegan)
#
# # Compute the Bray-Curtis dissimilarity matrix
# bray_curtis <- vegdist(pca_stat, method = "bray")
#
# # Convert to a matrix (since vegdist returns a dist object)
# bray_curtis_matrix <- as.matrix(bray_curtis)
#
# # Perform classical MDS (similar to PCA, but using dissimilarity matrix)
# mds <- cmdscale(bray_curtis_matrix, k = 2) # k is the number of dimensions you want to keep (e.g.
#
# # Plot the MDS result
# plot(mds, main = "PCA-like plot using Bray-Curtis Dissimilarity")
```

```

# #| label: chunk-6
# # Perform k-means clustering (example: 3 clusters)
# set.seed(42)
# kmeans_result <- kmeans(wide_data2[, -1], centers = 32) # Exclude first column (plot_name)
#
# # Add the cluster labels to the dataset
# wide_data2$cluster <- as.factor(kmeans_result$cluster)
#
# # Fit the tree model with k-means cluster labels as the response
# tree_model <- tree(cluster ~ ., data = wide_data2[, -1]) # Exclude non-numeric columns (plot_name)
#
# plot(tree_model)
# text(tree_model, pretty = 0)
#
# # Print the tree model
# summary(tree_model)

```

Conclusion

Appendix

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

```
# stat_mappingplants <- readRDS("~/Library/CloudStorage/OneDrive-Aarhusuniversitet/MappingPlants/adm
#
# taxa <- pivot |>
#   group_by(taxon) |>
#   count()
#
# soil_moi <- stat_mappingplants
#
# plot(x = stat_mappingplants$mean_soil_moisture, y = stat_mappingplants$mean_veg_height)
#
# hist(stat_mappingplants$mean_soil_moisture)
#
# empnig <- stat_mappingplants[stat_mappingplants$taxon=="Empetrum nigrum",]
#
# plot(x = empnig$mean_soil_moisture, y = empnig$bb_num)
#
#
# par(mfrow = c(2,2), mgp = c(2,0.7,0), mar = c(3,3,1,1))
# model <- gam(bb_num ~ s(mean_soil_moisture), data = stat_mappingplants[stat_mappingplants$taxon ==
# plot(model)
# par(mfrow = c(1,1))
#
#
# summary(model)
#
# library(tree)
# model<-tree(bb_num ~ ., data = empnig)
# plot(model)
# text(model)
```

```
taxon_counts <- table(stat_mappingplants$taxon)

# Convert to a data frame for easier manipulation
taxon_summary <- data.frame(
  Taxon = names(taxon_counts),
  Count = as.vector(taxon_counts)
)

taxon_summary <- taxon_summary[order(-taxon_summary$Count), ]
```

```
# Count the occurrences of each taxon
taxon_counts <- table(stat_mappingplants$taxon)

# Get the names of taxa with more than 10 observations
```

```
frequent_taxons <- names(taxon_counts[taxon_counts > 10])

# Filter the dataset
filtered_df <- stat_mappingplants[stat_mappingplants$taxon %in% frequent_taxons, ]
```

The client

The client, i.e. the receiver, of this report is my future self. I have basic understanding of statistics, statistical methods, and want to further my expertise in this areas both to explore the data I collect and have available as well as to document known phenomena of this same data. I have advance knowledge in biology and ecology. I do not have extensive of intuitive understanding of statistics and this report is aimed at document the learning outcomes of the data processing with the purpose of statistical reporting.

It is my interest to gain an applied and hand on approach to statistics, answer the reserach question at hand, explore the data I have collected

- What does the client already know? (basic/advanced science on the subject, statistical methods, project circumstances)
- What does the client not know? (basic/advanced science on the subject, statistical methods, project circumstances)
- What is the interest of the Client? (research question, p-values, effect parameters, issues with data handling)
- What is NOT the interest of the Client? (R code, issues with data handling, intermediate analyses)
- Adapt the contents and structure (not the results though) to fit the knowledge and interests of the Client.