

Package ‘aquats’

October 21, 2026

Type Package

Title Robust Ecological and Fisheries Data Analysis and Visualization

Version 1.2.2

Description A comprehensive statistical and visualization toolkit tailored for observational data in fisheries science and aquatic ecology. Provides streamlined workflows for univariate and factorial analysis of variance (ANOVA), analysis of covariance (ANCOVA), multivariate analysis of variance (MANOVA), permutational multivariate analysis of variance (PERMANOVA), similarity percentages (SIMPER), generalized linear models, non-linear morphometric regressions, two-group hypothesis testing, and chi-square tests. Features multivariate ordination methods including non-metric multidimensional scaling (NMDS), principal component analysis (PCA), redundancy analysis (RDA), distance-based redundancy analysis (db-RDA), and hierarchical cluster dendrograms. Computes alpha diversity metrics, species accumulation curves, water quality index (WQI) values, heavy metal ecotoxicological contamination indices, correlation heatmaps, and Mantel test network linkages. Offers specialized environmental visualizations including Sankey diagrams, Likert plots, circular bar charts, and longitudinal time-series with uncertainty ribbons. Methodological formulations for multivariate permutation tests follow Anderson (2001) <[doi:10.1111/j.1442-9993.2001.01070.pp.x](https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x)>. Ecological community dissimilarity transformations follow Legendre and Gallagher (2001) <[doi:10.1007/s004420100716](https://doi.org/10.1007/s004420100716)>, and ecotoxicological contamination index models follow Tomlinson et al. (1980) <[doi:10.1007/BF02414780](https://doi.org/10.1007/BF02414780)>.

License MIT + file LICENSE

URL <https://github.com/Ataher76/aquats>,
<https://doi.org/10.5281/zenodo.22885025>

BugReports <https://github.com/Ataher76/aquats/issues>

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

Depends R (>= 3.5.0)

Imports dplyr, emmeans, ggplot2, ggpubr, ggrepel, MASS, multcomp,
RColorBrewer, rlang, tidyr, ggalluvial, vegan

Suggests cowplot, knitr, multcompView, patchwork, readxl, rmarkdown,
writexl

VignetteBuilder knitr

NeedsCompilation no

Author Ataher Ali [aut, cre] (ORCID: <<https://orcid.org/0009-0008-8780-0903>>),
Mohammed Shahidul Alam [aut] (ORCID:
<<https://orcid.org/0000-0003-2740-5908>>)

Maintainer Ataher Ali <ataher.cu.ms@gmail.com>

Repository CRAN

Date/Publication 2026-10-21 15:00:12 UTC

Contents

ancova_analysis	3
calc_diversity	4
calc_metal_indices	5
calc_wqi	6
cluster_dendrogram	7
community_master	8
compare_two_groups	9
correlation_heatmap	10
dbrda_analysis	11
glm_analysis	13
hilsa_master	14
manova_analysis	15
mantel_heatmap_analysis	16
nmds_analysis	17
one_way_anova	18
pca_community_biplot	20
pca_correlation_circle	21
permanova_analysis	22
plot_bar	23
plot_circular_bar	25
plot_diversity	26
plot_donut	27
plot_histogram	28
plot_likert	29
plot_metal_indices	30
plot_pie3d	31
plot_sankey	32
plot_timeseries	33
plot_wqi	34
rda_analysis	35
regression_analysis	37

simper_analysis	38
specaccum_plot	39
summarize_obs	40
test_chisq	41
three_way_anova	43
two_way_anova	44
Index	46

ancova_analysis	<i>Analysis of Covariance (ANCOVA) with Adjusted Means Visualization</i>
-----------------	--

Description

Performs an Analysis of Covariance (ANCOVA) to test group differences in a continuous response variable while controlling for a continuous covariate. Extracts the ANCOVA summary table, estimated marginal means (EMMs), and generates a publication-ready regression scatter plot with parallel trend lines matching the additive model.

Usage

```
ancova_analysis(  
  data,  
  response_var,  
  factor_var,  
  covariate_var,  
  factor_levels = NULL,  
  facet = FALSE,  
  color_palette = "Set1"  
)
```

Arguments

data	A data.frame or matrix containing the variables for analysis.
response_var	Character string; column name in data representing the continuous numeric response variable.
factor_var	Character string; column name in data representing the categorical grouping factor.
covariate_var	Character string; column name in data representing the continuous numeric covariate.
factor_levels	Optional character vector; custom level ordering for factor_var. Default is NULL.
facet	Logical; if TRUE, facets panels by factor level using ggplot2::facet_wrap. Default is FALSE.
color_palette	Character string; a valid RColorBrewer palette name. Default is "Set1".

Value

A list containing:

ANCOVA_Table	Analysis of variance table (anova.lm) evaluating factor and covariate effects.
EMMs	Estimated marginal means summary object from emmeans.
Plot	The publication-ready ggplot2 scatter and regression plot.

calc_diversity	<i>Calculate Comprehensive Alpha Diversity Indices</i>
----------------	--

Description

Computes species richness, total abundance, Shannon-Wiener (H'), Gini-Simpson ($1 - D$), Margalef (d), Menhinick (R), and Pielou's evenness (J') indices simultaneously from community abundance matrices. Automatically isolates non-numeric metadata columns, guards against mathematical domain errors (e.g., division by zero), and removes empty sample observations.

Usage

```
calc_diversity(data, group_col = NULL)
```

Arguments

data	A data.frame, matrix, or single-element list containing community abundance counts and optional metadata columns.
group_col	Optional character or integer vector; column name(s) or index/indices in data specifying metadata variables to retain alongside calculated diversity indices. Default is NULL.

Value

A tidy data.frame containing sample identifiers, retained metadata columns, and the following computed diversity metrics:

Sample	Sample row names or integer index identifiers.
Abundance	Total observed individual count/abundance per sample (N).
Richness	Observed species richness count (S).
Shannon	Shannon-Wiener diversity index (H' , natural log base e).
Simpson	Gini-Simpson diversity index ($1 - D$).
Margalef	Margalef richness index ($d = (S - 1) / \ln(N)$).
Menhinick	Menhinick richness index ($R = S / \sqrt{N}$).
Pielou	Pielou's evenness index ($J' = H' / \ln(S)$).

calc_metal_indices *Heavy Metal Pollution Indices with Confidence Intervals*

Description

Computes five heavy metal pollution indices (HPI, HEI, Cd, PLI, and Nemerow PN) per observation or replicate. If replicated sampling is detected, it calculates station-level summary statistics including Mean, SD, SE, and 95% confidence intervals.

Usage

```
calc_metal_indices(
  data,
  id_col = "Station",
  standards = c(As = 10, Cd = 3, Cr = 50, Cu = 2000, Pb = 10, Zn = 3000),
  ci = 0.95,
  file = NULL
)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing metal concentration data.
id_col	Character string; column name in data identifying sampling stations or monitoring sites. Default is "Station".
standards	Named numeric vector of permissible guideline limits (S_i) in matching concentration units ($\mu\text{g/L}$ or mg/L).
ci	Numeric; confidence level for interval estimation (e.g., 0.95 for 95% confidence intervals). Default is 0.95.
file	Optional character string; file path to export tabular results (.xlsx or .csv). Default is NULL.

Value

A list containing:

Replicate_Data	Data frame containing computed indices for each individual sample observation.
Summary_Data	Data frame of station-level aggregated statistics (Mean, SD, SE, 95% CI, and pollution status).

calc_wqi	<i>Weighted Arithmetic Water Quality Index (WAWQI) with Confidence Intervals</i>
----------	--

Description

Computes the Weighted Arithmetic Water Quality Index (WAWQI) per observation or replicate. When replicated station sampling is detected, it calculates station-level summary statistics including Mean, SD, SE, 95% confidence intervals, and overall ecological water quality rating.

Usage

```
calc_wqi(
  data,
  id_col = "Station",
  standards = c(pH = 8.5, DO = 5, BOD = 5, TDS = 500, Turbidity = 5),
  ideal_values = NULL,
  ci = 0.95,
  file = NULL
)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing physicochemical water quality parameters.
id_col	Character string; column name in data identifying sampling stations or monitoring sites. Default is "Station".
standards	Optional named numeric vector of standard regulatory limits (S_i) in matching concentration units. If NULL, default benchmark guidelines (e.g., WHO/BIS) are applied. Default is NULL.
ideal_values	Optional named numeric vector of ideal zero-effect reference values (V_0) (e.g., 7.0 for pH, 14.6 mg/L for dissolved oxygen). Default is NULL.
ci	Numeric; confidence level for interval estimation (e.g., 0.95 for 95% confidence intervals). Default is 0.95.
file	Optional character string; export file path for tabular results (.xlsx or .csv). Default is NULL.

Value

A list containing:

Replicate_Data	Data frame containing computed sub-indices, relative weights, and overall WAWQI for each individual sample observation.
Summary_Data	Data frame of station-level aggregated statistics (Mean, SD, SE, 95% CI, and water quality classification rating).

Description

Computes ecological dissimilarity matrices (Bray-Curtis, Jaccard, Euclidean, etc.) with optional multivariate transformations (Hellinger, log, presence/absence, standardization) and performs agglomerative hierarchical clustering (UPGMA, Ward.D2, Complete). Generates a publication-ready ggplot2 dendrogram with color-coded group tips, a cluster cut-off threshold line, and a sample-to-cluster assignment table.

Usage

```
cluster_dendrogram(
  data,
  group_col = 1,
  dist_method = c("bray", "jaccard", "euclidean", "horn", "kulczynski", "gower",
    "manhattan", "canberra"),
  transform = c("none", "hellinger", "log", "pa", "standardize", "sqrt"),
  cluster_method = c("average", "ward.D2", "ward.D", "complete", "single", "centroid"),
  k = NULL,
  h = NULL,
  palette = "Set2",
  orientation = c("vertical", "horizontal"),
  title = NULL
)
```

Arguments

data	A data.frame, matrix, or list containing community abundance counts or environmental parameters with optional grouping metadata.
group_col	Character string or integer; column name or index in data specifying grouping metadata. Default is 1. If NULL, clustering runs without group categorization.
dist_method	Character string; dissimilarity metric passed to <code>vegan::vegdist</code> : "bray" (default), "jaccard", "euclidean", "horn", "kulczynski", "gower", "manhattan", or "canberra". Default is "bray".
transform	Character string; transformation applied via <code>vegan::decostand</code> : "none" (default), "hellinger" (recommended for abundance data), "log" ($\ln(x + 1)$), "pa" (presence/absence binary), "standardize" (z-score, recommended for water chemistry), or "sqrt". Default is "none".
cluster_method	Character string; agglomeration method passed to <code>stats::hclust</code> : "average" (UPGMA, default), "ward.D2", "ward.D", "complete", "single", or "centroid". Default is "average".
k	Optional integer; number of clusters to partition the tree into. If NULL, defaults to the number of group levels. Default is NULL.

h	Optional numeric; specific dissimilarity height threshold at which to cut the dendrogram. Default is NULL.
palette	Character string; a valid RColorBrewer palette name for group leaf tips. Default is "Set2".
orientation	Character string; tree layout: "vertical" (default) or "horizontal". Default is "vertical".
title	Optional character string; custom plot title. Default is NULL.

Value

A list containing:	
Plot	A publication-ready ggplot2 dendrogram object.
Cluster_Assignments	Tidy data frame showing sample cluster memberships.
Distance_Matrix	The computed dist object.
HClust_Object	The underlying hclust object.
Transformed_Data	The processed numeric matrix used for distance computation.

community_master	<i>Multi-Trophic Community Abundance Master List</i>
------------------	--

Description

A curated list containing species abundance matrices for phytoplankton, zooplankton, and benthic macroinvertebrates, designed for community ecology analyses, PERMANOVA, SIMPER, and multi-matrix Mantel network visualizations.

Usage

community_master

Format

- A named list with 3 data frames:
- Phytoplankton** Data frame of diatom, cyanobacteria, and green algae counts.
 - Zooplankton** Data frame of copepod, cladoceran, and rotifer counts.
 - Benthos** Data frame of chironomid, oligochaete, and mollusc counts.

compare_two_groups	<i>Two-Sample Comparison with Automated F-Test, t-Test, and Visualization</i>
--------------------	---

Description

Evaluates equality of variances using Fisher's F -test, checks group-level normality via the Shapiro-Wilk test, and performs Student's, Welch's, or paired t -test to compare means between two groups. Automatically computes effect sizes (Cohen's d), compiles group descriptive statistics, and generates a publication-ready ggplot2 graphic with significance annotations.

Usage

```
compare_two_groups(
  data,
  response_var,
  group_var,
  var_equal = "auto",
  alternative = "two.sided",
  paired = FALSE,
  conf_level = 0.95,
  plot_type = "box",
  palette = "Dark2",
  title = NULL
)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing the observations.
response_var	Character string; column name in data representing the continuous numeric response variable.
group_var	Character string; column name in data representing the categorical grouping factor (must contain exactly two distinct levels).
var_equal	Character string or logical; variance assumption for independent t -tests: "auto" (default; runs Fisher's F -test and adopts Welch's t -test if $p < 0.05$), TRUE (forces Student's t -test), or FALSE (forces Welch's t -test). Ignored if paired = TRUE. Default is "auto".
alternative	Character string; alternative hypothesis: "two.sided" (default), "greater", or "less". Default is "two.sided".
paired	Logical; if TRUE, performs a paired t -test. When TRUE, observations must be matched row-wise across groups. Default is FALSE.
conf_level	Numeric; confidence level for difference intervals and summary bounds (e.g., 0.95 for 95% confidence intervals). Default is 0.95.

plot_type	Character string; visual geometry: "box" (default), "violin", or "jitter". Default is "box".
palette	Character string; a valid RColorBrewer palette name. Default is "Dark2".
title	Optional character string; custom plot title. Default is NULL.

Value

A list containing:

Summary_Table	Tidy data frame containing sample sizes (N), Mean, SD, SE, 95% CI per group, mean difference, and Cohen's d .
T_Test	Tidy summary data frame of the t -test results (method, test statistic t , degrees of freedom, p -value, and confidence interval).
F_Test	Tidy summary data frame of Fisher's F -test for variance equality (variance ratio, degrees of freedom, and p -value).
Normality_Test	Tidy data frame of Shapiro-Wilk normality test results for each group.
Plot	A publication-ready ggplot2 object annotated with test results.
Model_Objects	A list containing raw base R test outputs (t_test, f_test, and shapiro_tests).

correlation_heatmap	<i>Flexible Correlation Matrix and Heatmap Generator for Environmental Data</i>
---------------------	---

Description

Computes correlation matrices (Pearson, Spearman, or Kendall) with exact p-values and generates publication-ready heatmaps. Supports customizable shapes (squares, circles), triangle views (lower, upper, full), optional diagonal display, significance asterisks, and automated contrast text adjustment.

Usage

```
correlation_heatmap(
  data,
  numeric_vars = NULL,
  method = c("pearson", "spearman", "kendall"),
  shape = c("square", "circle"),
  view = c("lower", "upper", "full"),
  diag = FALSE,
  label_type = NULL,
  sig_level = NULL,
  color_palette = "RdBu",
  title = NULL
)
```

Arguments

<code>data</code>	A <code>data.frame</code> or matrix containing numeric variables.
<code>numeric_vars</code>	Optional character vector; specific numeric column names in data to include. If NULL, all numeric columns are automatically detected. Default is NULL.
<code>method</code>	Character string; correlation method: "pearson" (default), "spearman", or "kendall". Default is "pearson".
<code>shape</code>	Character string; heatmap marker shape: "square" (default) or "circle". Default is "square".
<code>view</code>	Character string; matrix layout view: "lower" (default), "upper", or "full". Default is "lower".
<code>diag</code>	Logical; whether to show diagonal self-correlations. Default is FALSE.
<code>label_type</code>	Character string; content displayed inside cells: "both" (numeric value and asterisks), "stars" (only asterisks), "values" (only numeric values), or "none". Defaults to "both" for squares and "stars" for circles.
<code>sig_level</code>	Optional numeric; significance threshold (e.g., 0.05). If specified, non-significant correlations are rendered transparent or blanked out. Default is NULL.
<code>color_palette</code>	Character string; a valid diverging palette name for <code>scale_fill_distiller</code> (e.g., "RdBu", "RdYlBu", "Spectral", "BrBG"). Default is "RdBu".
<code>title</code>	Optional character string; custom plot title. Default is NULL.

Value

A list containing:

<code>Correlation_Matrix</code>	Symmetric matrix of correlation coefficients.
<code>P_Value_Matrix</code>	Matrix of exact <i>p</i> -values.
<code>Summary_Table</code>	Tidy data frame of pairwise correlations, <i>p</i> -values, and significance tiers.
<code>Plot</code>	A publication-ready ggplot2 object.

<code>dbrda_analysis</code>	<i>Distance-based Redundancy Analysis (db-RDA) Triplot</i>
-----------------------------	--

Description

Performs constrained Distance-based Redundancy Analysis (db-RDA) using ecological dissimilarity metrics (e.g., Bray-Curtis) combined with environmental variables. Automatically cleans empty samples, runs permutation significance tests, and generates a publication-ready ggplot2 Triplot with ggrepel text positioning.

Usage

```
dbrda_analysis(
  comm_data,
  env_data,
  group_col = 1,
  dist_method = "bray",
  transform = c("none", "hellinger", "log", "pa", "wisconsin"),
  color_palette = "Dark2",
  species_arrow_mult = 1,
  env_arrow_mult = 1
)
```

Arguments

<code>comm_data</code>	A data frame or matrix where grouping columns are at the start, followed by numeric species counts.
<code>env_data</code>	A data frame or matrix containing environmental predictor variables. Must have the exact same number of rows as <code>comm_data</code> .
<code>group_col</code>	Character string or integer; column name or index in <code>comm_data</code> specifying the categorical grouping factor. Default is 1.
<code>dist_method</code>	Character string; dissimilarity metric for community abundance data (e.g., "bray", "jaccard", "euclidean"). Default is "bray".
<code>transform</code>	Character string; pre-transformation applied to community abundance data via <code>vegan::decostand</code> : "none" (default), "hellinger", "log", "pa" (presence/absence), or "wisconsin". Default is "none".
<code>color_palette</code>	Character string; a valid RColorBrewer palette name. Default is "Dark2".
<code>species_arrow_mult</code>	Numeric; scaling multiplier for species vector arrows. Default is 1.
<code>env_arrow_mult</code>	Numeric; scaling multiplier for environmental vector arrows. Default is 1.

Value

A list containing:

<code>dbRDA_Object</code>	The underlying dbrda ordination object from <code>vegan</code> .
<code>Model_Significance</code>	Permutation ANOVA test for overall model significance.
<code>Variable_Significance</code>	Permutation ANOVA test for marginal term significance.
<code>Plot</code>	The publication-ready <code>ggplot2</code> triplot.

Description

Fits generalized linear models (GLM) across Gaussian, Poisson, Binomial, and Gamma families with publication-ready diagnostic and prediction graphics.

Usage

```
glm_analysis(  
  data,  
  response_var,  
  predictor_vars,  
  family_type = "poisson",  
  color_palette = "Set1"  
)
```

Arguments

data	A data.frame or matrix containing the variables for modeling.
response_var	Character string; column name in data representing the continuous or discrete dependent response variable.
predictor_vars	Character vector; column name(s) in data representing continuous or categorical predictor variables.
family_type	Character string; error distribution family: "gaussian" (default), "poisson", "binomial", or "gamma". Default is "gaussian".
color_palette	Character string; a valid RColorBrewer palette name. Default is "Dark2".

Value

A list containing:

Model	The fitted glm model object from stats.
Summary	Summary specification of the fitted model (summary.glm).
Anova_Table	Analysis of deviance table (anova.glm).
Plot	A publication-ready ggplot2 diagnostic or prediction object.

hilsa_master

*Comprehensive Master Dataset for Hilsa Shad Ecology and Statistics***Description**

A comprehensive, unified dataset containing morphological measurements, environmental parameters, catch records, and multi-factor experimental groupings for Hilsa shad (*Tenualosa ilisha*) across various habitats, seasons, and size classes. Designed to power all parametric, non-parametric, regression, and multivariate functions in the package.

Usage

```
hilsa_master
```

Format

A data frame with 120 rows and 14 variables:

Habitat Factor; sampling environment (Marine, Estuary, River).

Season Factor; fishing season (Monsoon, Dry).

Size_Class Factor; growth stage (Juvenile, Adult).

Weight_g Numeric; body weight in grams.

Total_Length_cm Numeric; total length in centimeters.

Body_Depth_cm Numeric; maximum body depth in centimeters.

Head_Length_cm Numeric; head length in centimeters.

Fin_Length_cm Numeric; pectoral fin length in centimeters.

Fishing_Hours Numeric; fishing effort in hours.

Catch_Count Integer; number of fish caught (count response).

SST_C Numeric; Sea Surface Temperature in degrees Celsius.

Salinity_ppt Numeric; water salinity in parts per thousand.

Depth_m Numeric; water depth in meters.

Catch_kg Numeric; total fish catch in kilograms.

manova_analysis	<i>Multivariate Analysis of Variance (MANOVA) with Alternative Visualizations</i>
-----------------	---

Description

Performs a one-way MANOVA across multiple continuous response variables grouped by a categorical factor, extracts multivariate test statistics (Pillai's Trace), provides univariate ANOVA breakdowns, and generates either faceted trait boxplots or a Canonical Discriminant Analysis (LDA) multivariate ordination scatter plot.

Usage

```
manova_analysis(
  data,
  response_vars,
  factor_var,
  factor_levels = NULL,
  plot_type = "boxplot",
  color_palette = "Set1"
)
```

Arguments

data	A data.frame or matrix containing the variables for analysis.
response_vars	Character vector; column names in data representing continuous numeric response variables.
factor_var	Character string; column name in data representing the categorical grouping factor.
factor_levels	Optional character vector; custom level ordering for factor_var. Default is NULL.
plot_type	Character string; visualization layout: "boxplot" (faceted response boxplots) or "lda" (linear discriminant ordination plot). Default is "boxplot".
color_palette	Character string; a valid RColorBrewer palette name. Default is "Set1".

Value

A list containing:

MANOVA_Summary	Multivariate test summary table based on Pillai's trace.
Univariate_ANOVAs	Univariate ANOVA summary tables for each response variable.
Plot	The publication-ready ggplot2 object.

mantel_heatmap_analysis

Integrated Environmental Correlation and Multi-Community Mantel Network

Description

Computes pairwise environmental correlations while performing Mantel tests across single or multiple biological communities (e.g., Phytoplankton, Zooplankton, Benthos, Fish). Generates publication-ready network diagrams with customizable straight or curved linkages and anchor pins.

Usage

```
mantel_heatmap_analysis(
  comm_data,
  env_data,
  plot_type = c("network", "bars"),
  line_style = c("curve", "straight"),
  method = c("pearson", "spearman"),
  spec_dist = "bray",
  env_dist = "euclidean",
  transform = c("none", "hellinger", "log", "pa", "sqrt"),
  only_significant = FALSE,
  line_color_by = c("significance", "community"),
  permutations = 999,
  seed = 42,
  show_cor_text = TRUE,
  color_palette = "RdBu",
  title = NULL
)
```

Arguments

comm_data	A data.frame, matrix, or a named list of data frames representing distinct biological communities.
env_data	A data.frame or matrix of numeric environmental variables.
plot_type	Character string; visualization layout: "network" (default curved/straight diagram) or "bars" (synchronized side-by-side barplot). Default is "network".
line_style	Character string; linkage geometry: "curve" (default) or "straight". Default is "curve".
method	Character string; correlation method: "pearson" (default) or "spearman". Default is "pearson".
spec_dist	Character string; dissimilarity metric for community abundance data (e.g., "bray", "jaccard"). Default is "bray".
env_dist	Character string; distance metric for environmental parameters. Default is "euclidean".

transform	Character string; transformation applied to community counts via <code>vegan::decostand</code> : "none" (default), "hellinger" (recommended for abundance), "log", "pa", or "sqrt". Default is "none".
only_significant	Logical; if TRUE, displays only statistically significant Mantel linkages ($p < 0.05$). Default is FALSE.
line_color_by	Character string; variable mapping for linkage line colors: "significance" (default) or "community". Default is "significance".
permutations	Numeric integer; number of Monte Carlo permutations for Mantel significance testing. Default is 999.
seed	Optional integer; random seed for reproducible permutations. Default is 42.
show_cor_text	Logical; if TRUE, displays correlation coefficients inside tiles. Default is TRUE.
color_palette	Character string; a valid diverging RColorBrewer palette name for the correlation heatmap. Default is "RdBu".
title	Optional character string; master figure title. Default is NULL.

Value

A list containing:

Plot	The primary publication-ready ggplot2 object.
Correlation_Matrix	Symmetric matrix of environmental correlation coefficients.
Mantel_Results	Tidy data frame containing community-specific Mantel r values, p -values, and significance tiers.

nmds_analysis	<i>Non-metric Multidimensional Scaling (NMDS) Analysis for Aquatic Communities</i>
---------------	--

Description

Performs NMDS ordination with automated data transformations (Hellinger, log, etc.), stress quality checks, and generates a publication-ready ggplot2 scatterplot with 95

Usage

```
nmds_analysis(
  data,
  group_col = 1,
  env_data = NULL,
  dist_method = "bray",
  transform = c("none", "hellinger", "log", "pa", "wisconsin"),
  k = 2,
  trymax = 100,
  palette = "Dark2",
  title = NULL
)
```

Arguments

<code>data</code>	A <code>data.frame</code> or <code>matrix</code> containing community abundance counts and grouping metadata.
<code>group_col</code>	Character string or integer; column name or index in <code>data</code> specifying the categorical grouping factor. Default is 1.
<code>env_data</code>	Optional <code>data.frame</code> or <code>matrix</code> of numeric environmental variables to fit onto the ordination space via <code>vegan::envfit</code> . Default is <code>NULL</code> .
<code>dist_method</code>	Character string; dissimilarity metric passed to <code>vegan::vegdist</code> (e.g., "bray", "jaccard", "euclidean"). Default is "bray".
<code>transform</code>	Character string; pre-transformation applied to community counts via <code>vegan::decostand</code> : "none" (default), "hellinger" (recommended for abundance data), "log", "pa" (presence/absence), or "wisconsin". Default is "none".
<code>k</code>	Numeric integer; number of ordination dimensions. Default is 2.
<code>trymax</code>	Numeric integer; maximum number of random starts for <code>vegan::metaMDS</code> . Default is 100.
<code>palette</code>	Character string; a valid <code>RColorBrewer</code> palette name for group ellipses and points. Default is "Dark2".
<code>title</code>	Optional character string; custom plot title. Default is <code>NULL</code> .

Value

A list containing:

<code>Plot</code>	A publication-ready <code>ggplot2</code> ordination object.
<code>Plot_Data</code>	Tidy data frame of NMDS site coordinates and grouping factors.
<code>Stress</code>	The final stress value of the ordination.
<code>NMDS_Object</code>	The underlying <code>metaMDS</code> object from <code>vegan</code> .
<code>Envfit_Results</code>	Vector fit results from <code>vegan::envfit</code> if environmental variables were supplied.

one_way_anova

One-Way Analysis of Variance (ANOVA) with Faceting & Visualization

Description

This function performs a one-way ANOVA, calculates Tukey's HSD post-hoc test, extracts Estimated Marginal Means (EMMs), and generates a publication-ready plot with optional faceting by a secondary categorical variable.

Usage

```

one_way_anova(
  data,
  factor_var,
  numeric_var,
  facet_var = NULL,
  factor_levels = NULL,
  plot_type = "boxplot",
  error_type = "se",
  sig_display = "letters",
  y_limits = NULL,
  add_jitter = TRUE,
  show_mean = TRUE,
  mean_color = "darkred",
  color_palette = "Set1"
)

```

Arguments

<code>data</code>	A data frame containing the variables for analysis.
<code>factor_var</code>	Character string; column name in data representing the primary categorical grouping factor (x-axis).
<code>numeric_var</code>	Character string; column name in data representing the continuous numeric response variable.
<code>facet_var</code>	Optional character string; column name in data representing a secondary categorical factor to facet panels by. Default is NULL.
<code>factor_levels</code>	Optional character vector; custom level ordering for <code>factor_var</code> . Default is NULL.
<code>plot_type</code>	Character string; display geometry: "boxplot", "barplot", or "pointrange". Default is "boxplot".
<code>error_type</code>	Character string; error bar dispersion metric for bar and point-range plots: "se" (Standard Error) or "sd" (Standard Deviation). Default is "se".
<code>sig_display</code>	Character string; method to annotate statistical significance: "letters" (compact letter display from Tukey's HSD) or "stars" (pairwise significance brackets). Default is "letters".
<code>y_limits</code>	Optional numeric vector of length 2; lower and upper limits for the y-axis (e.g., <code>c(0, 100)</code>). Default is NULL.
<code>add_jitter</code>	Logical; if TRUE, overlays jittered raw observation points on boxplots. Default is TRUE.
<code>show_mean</code>	Logical; if TRUE, displays group mean indicator diamonds inside boxplots. Default is TRUE.
<code>mean_color</code>	Character string; hex code or color name for group mean indicators. Default is "darkred".
<code>color_palette</code>	Character string; a valid RColorBrewer palette name. Default is "Set1".

Value

A list containing:

ANOVA_Summary	Global ANOVA summary table (F-statistic and p-value).
TukeyHSD	Tukey HSD post-hoc pairwise comparisons.
EMMs	Estimated Marginal Means object from emmeans.
Summary_Table	Aggregated sample sizes, means, SDs, SEs, and confidence limits.
Plot	The publication-ready ggplot2 object.

pca_community_biplot *Principal Component Analysis (PCA) Community Biplot*

Description

Performs Principal Component Analysis (PCA / RDA) on aquatic community data. Generates a publication-ready ggplot2 biplot showing site coordinates with 95

Usage

```
pca_community_biplot(
  data,
  group_col = 1,
  scale = FALSE,
  transform = c("none", "hellinger", "log", "pa", "wisconsin"),
  top_n_species = NULL,
  color_palette = "Dark2",
  arrow_multiplier = 1,
  title = NULL
)
```

Arguments

data	A data.frame or matrix containing grouping metadata and numeric species abundance counts.
group_col	Character string or integer; column name or index in data specifying the categorical grouping factor. Default is 1.
scale	Logical; whether to standardize species variables to unit variance. Setting to TRUE is recommended when variables span different units or orders of magnitude. Default is FALSE.
transform	Character string; pre-transformation applied to community counts via <code>vegan::decostand</code> : "none" (default), "hellinger" (recommended for abundance data), "log", "pa" (presence/absence), or "wisconsin". Default is "none".
top_n_species	Optional integer; number of top species to display based on vector loading length to reduce text clutter. If NULL, all species arrows are plotted. Default is NULL.

color_palette	Character string; a valid RColorBrewer palette name for site groups. Default is "Dark2".
arrow_multiplier	Numeric; scaling multiplier for species loading arrows. Default is 1.
title	Optional character string; custom plot title. Default is NULL.

Value

A list containing:

Plot	A publication-ready ggplot2 biplot object.
Plot_Data	Tidy data frame of site PC scores and grouping metadata.
Species_Scores	Tidy data frame of species loadings and vector coordinates.
PCA_Object	The underlying rda ordination object from vegan.

pca_correlation_circle

PCA Correlation Circle with cos² Gradient for Environmental Parameters

Description

Generates a publication-grade PCA correlation circle (variables factor map) using base stats and ggplot2. Colors environmental parameters by their representation quality (cos²) on the selected principal components.

Usage

```
pca_correlation_circle(
  data,
  exclude_cols = 1,
  dim1 = 1,
  dim2 = 2,
  palette = c("#2A9D8F", "#E9C46A", "#F4A261", "#E76F51"),
  title = NULL
)
```

Arguments

data	A data.frame or matrix containing metadata columns and numeric environmental variables.
exclude_cols	Optional character or integer vector; column names or indices in data to exclude from PCA (e.g., non-numeric metadata columns). Default is 1. If NULL, all columns are analyzed.
dim1	Numeric integer; index of the first principal component to plot on the x-axis. Default is 1.

dim2	Numeric integer; index of the second principal component to plot on the y-axis. Default is 2.
palette	Character vector; hex color codes defining the color gradient for representation quality ($\cos^2$). Default is <code>c("#2A9D8F", "#E9C46A", "#F4A261", "#E76F51")</code> .
title	Optional character string; custom plot title. Default is NULL.

Value

A list containing:

Plot	A publication-ready ggplot2 correlation circle object.
Variable_Stats	Data frame of variable coordinates, correlations, and $\cos^2$ values.
Variance_Summary	Data frame detailing eigenvalues and percentage variance explained per PC.
Sample_Scores	Data frame of sample/station scores along the principal components.
PCA_Object	The underlying prcomp object from stats.

permanova_analysis	<i>Permutational Multivariate Analysis of Variance (PERMANOVA)</i>
--------------------	--

Description

Performs PERMANOVA (using `vegan::adonis2`) to test for statistical differences in community composition between groups. Automatically cleans empty samples, validates numeric matrices, supports ecological transformations, and handles flexible data ingestion.

Usage

```
permanova_analysis(
  data,
  group_col = 1,
  dist_method = "bray",
  transform = c("none", "hellinger", "log", "pa", "wisconsin"),
  permutations = 999,
  seed = 42
)
```

Arguments

data	A data.frame or matrix containing grouping metadata and numeric species abundance counts.
group_col	Character string or integer; column name or index in data specifying the categorical grouping factor. Default is 1.
dist_method	Character string; dissimilarity or distance metric passed to <code>vegan::vegdist</code> (e.g., "bray", "jaccard", "euclidean"). Default is "bray".

transform	Character string; pre-transformation applied to community counts via <code>vegan::decostand</code> : "none" (default), "hellinger" (recommended for abundance data), "log", "pa" (presence/absence), or "wisconsin". Default is "none".
permutations	Numeric integer; number of Monte Carlo permutations for significance testing. Default is 999.
seed	Optional integer; random seed for reproducible permutation testing. Default is 42.

Value

A list containing:

PERMANOVA_Table

The ANOVA-style permutation results table (`anova.cca`) from `vegan::adonis2`.

Cleaned_Data

A data frame combining the filtered grouping factor and community matrix.

Distance_Method

Character string indicating the dissimilarity metric applied.

plot_bar

Flexible and Batch Publication-Ready Bar and Lollipop Plots

Description

Generates publication-ready bar and lollipop plots from a data frame, CSV, or Excel file. Supports frequency counting, summary metric aggregation (mean, sum, identity), grouped/stacked arrangements, automated data labelling, coordinate flipping, and multi-column batch plotting.

Usage

```
plot_bar(
  data,
  x,
  y = NULL,
  group = NULL,
  type = c("bar", "lollipop"),
  stat = c("mean", "sum", "identity"),
  position = c("dodge", "stack", "fill"),
  show_labels = FALSE,
  horizontal = FALSE,
  palette = "Set2",
  bar_width = 0.7,
  point_size = 3.5,
  line_size = 0.8,
  xlab = NULL,
  ylab = NULL,
  title_prefix = ""
)

plot_lollipop(..., point_size = 3.5, line_size = 0.8)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv, .xlsx, or .xls file containing the survey data.
x	Character vector; column name(s) in data representing the primary categorical grouping variable(s). If multiple variables are passed without y, a list of plots is returned.
y	Optional character vector; column name(s) in data representing numeric response variable(s) to aggregate or plot directly. If multiple variables are passed, a list of plots is returned. Default is NULL.
group	Optional character string; column name in data representing a secondary categorical factor for grouped or stacked layouts. Default is NULL.
type	Character string; visual geometry style: "bar" (default) or "lollipop". Default is "bar".
stat	Character string; aggregation statistic when y is provided: "mean" (default), "sum", or "identity" (values plotted as-is). Default is "mean".
position	Character string; arrangement geometry: "dodge" (default for comparisons), "stack", or "fill" (proportional 100% stacked). For lollipop plots, "dodge" is always used. Default is "dodge".
show_labels	Logical; if TRUE, displays numeric value labels on or above the geometries. Default is FALSE.
horizontal	Logical; if TRUE, flips Cartesian coordinates horizontally for legible category labels. Default is FALSE.
palette	Character string; a valid RColorBrewer palette name (e.g., "Set2", "Dark2"). Default is "Set2".
bar_width	Numeric; relative width of bars or spacing factor for dodging. Default is 0.7.
point_size	Numeric; size of the marker heads when type = "lollipop". Default is 3.5.
line_size	Numeric; width of the stem lines when type = "lollipop". Default is 0.8.
xlab	Optional character string; custom x-axis title. Default is NULL.
ylab	Optional character string; custom y-axis title. Default is NULL.
title_prefix	Optional character string; text prefix prepended to the plot title. Default is "".
...	Additional arguments passed to plot_bar() when calling plot_lollipop().

Value

A single ggplot2 object, or a named list of ggplot2 objects if batch plotting across multiple columns.

plot_circular_bar	<i>Publication-Grade Circular or Radial Bar Plot</i>
-------------------	--

Description

Builds circular bar charts with concentric reference gridlines, scale labels, and category labels. Best suited for data with six or more categories.

Usage

```
plot_circular_bar(
  data,
  x,
  y = NULL,
  group = NULL,
  stat = c("mean", "sum", "identity"),
  inner_radius = 0.35,
  show_grid = TRUE,
  palette = "Set2",
  title = NULL
)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing the category observations.
x	Character string; column name in data representing the primary categorical grouping factor positioned radially around the circle.
y	Optional character string; column name in data representing the numeric variable to aggregate. If NULL, observation frequencies (counts) are plotted. Default is NULL.
group	Optional character string; column name in data representing a secondary categorical factor for bar fill coloring. Default is NULL.
stat	Character string; aggregation statistic when y is provided: "mean" (default), "sum", or "identity" (values plotted as-is). Default is "mean".
inner_radius	Numeric; fractional proportion of the plot radius reserved for the central inner void (between 0 and 1). Default is 0.35.
show_grid	Logical; if TRUE, renders concentric reference grid rings and numeric scale values. Default is TRUE.
palette	Character string; a valid RColorBrewer palette name. Default is "Set2".
title	Optional character string; custom plot title. Default is NULL.

Value

A publication-ready ggplot2 object.

Description

Generates publication-ready boxplots for alpha diversity metrics across single or two-factor environmental groupings. Provides optional parametric (ANOVA) or non-parametric (Kruskal-Wallis) hypothesis testing, with optional faceting.

Usage

```
plot_diversity(
  data,
  group_col = 1,
  facet_var = NULL,
  index = c("shannon", "simpson", "richness", "abundance", "margalef", "menhinick",
    "pielou"),
  test = c("anova", "kruskal", "none"),
  color_palette = "Dark2",
  permutations = 1000,
  file = NULL
)
```

Arguments

data	A data.frame or matrix containing community abundance counts and grouping metadata columns.
group_col	Character string or integer vector; column name(s) or index/indices in data specifying one or two categorical grouping factors.
facet_var	Optional character string; column name in data to facet panels by. Default is NULL.
index	Character string; alpha diversity metric to compute and plot: "shannon" (default), "simpson", "richness", "abundance", "margalef", "menhinick", or "pielou". Default is "shannon".
test	Character string; statistical hypothesis test to perform across groups: "anova" (default), "kruskal", or "none". Default is "anova".
color_palette	Character string; a valid RColorBrewer palette name. Default is "Dark2".
permutations	Numeric integer; number of bootstrap iterations for confidence intervals. Default is 1000.
file	Optional character string; export file path for summary tables (.xlsx or .csv). Default is NULL.

Value

A list containing:

Diversity_Data	Data frame of individual sample diversity metric values.
Group_Summary	Data frame of aggregated group-level summary statistics.
Test_Result	Statistical test summary object (aov or kruskal.test).
Plot	A publication-ready ggplot2 boxplot object.

plot_donut	<i>Publication-Ready Single, Double, and Triple-Layered Donut Plots</i>
------------	---

Description

Generates publication-ready single or nested multi-layered (sunburst) donut charts. Supports up to three concentric hierarchical tiers with automatic alignment and percentage threshold labelling.

Usage

```
plot_donut(
  data,
  levels,
  value = NULL,
  show_labels = TRUE,
  min_percent = 4,
  palette = "Set2",
  title = NULL
)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing categorical composition data.
levels	Character vector; column names in data representing 1 to 3 categorical grouping variables, ordered from the innermost ring to the outermost ring.
value	Optional character string; column name in data representing a numeric variable to aggregate (sum). If NULL, observation frequencies (counts) are plotted. Default is NULL.
show_labels	Logical; if TRUE, displays percentage labels directly on segments. Default is TRUE.
min_percent	Numeric; minimum percentage (0 to 100) required to display a text label on a segment to prevent visual clutter. Default is 4.
palette	Character string; a valid RColorBrewer palette name. Default is "Set2".
title	Optional character string; custom plot title. Default is NULL.

Value

A publication-ready ggplot2 object.

plot_histogram

*Publication-Ready Histogram and Density Plot***Description**

Generates publication-ready frequency histograms for continuous fisheries and environmental variables. Supports optional kernel density curves, theoretical normal distribution overlays, and mean/median reference lines.

Usage

```
plot_histogram(
  data,
  x,
  group = NULL,
  bins = 30,
  binwidth = NULL,
  add_density = FALSE,
  add_normal = FALSE,
  show_stats = TRUE,
  facet = FALSE,
  palette = "Blues",
  xlab = NULL,
  ylab = "Frequency",
  title = NULL
)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing the observations.
x	Character string; column name in data representing the continuous numeric variable to plot.
group	Optional character string; column name in data representing a categorical grouping factor for faceting or fill coloring. Default is NULL.
bins	Numeric integer; number of histogram bins. Default is 30.
binwidth	Optional numeric; explicit width of the histogram bins (overrides bins). Default is NULL.
add_density	Logical; if TRUE, overlays an empirical kernel density curve. Default is FALSE.
add_normal	Logical; if TRUE, overlays a theoretical normal distribution curve. Default is FALSE.
show_stats	Logical; if TRUE, adds dashed reference lines indicating the mean and median. Default is FALSE.
facet	Logical; if TRUE and group is provided, facets panels by group. Default is FALSE.

palette	Character string; a valid RColorBrewer palette name. Default is "Blues".
xlab	Optional character string; custom x-axis label. Default is NULL.
ylab	Optional character string; custom y-axis label. Default is "Frequency".
title	Optional character string; custom plot title. Default is NULL.

Value

A publication-ready ggplot2 object.

plot_likert	<i>Publication-Ready Likert Scale Visualization</i>
-------------	---

Description

Generates modern diverging or 100% stacked horizontal bar charts for survey items and perception data.

Usage

```
plot_likert(
  data,
  items,
  levels = NULL,
  type = c("diverging", "stacked"),
  palette = "RdYlBu",
  show_labels = TRUE,
  clean_labels = TRUE,
  title = NULL
)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing survey responses.
items	Character vector; column names in data representing Likert question items.
levels	Optional character vector; custom ordered levels of the Likert response scale (e.g., from strongly disagree to strongly agree). Default is NULL.
type	Character string; plot display style: "diverging" (default) or "stacked" (100% stacked). Default is "diverging".
palette	Character string; a valid RColorBrewer palette name. Default is "RdYlBu".
show_labels	Logical; if TRUE, displays percentage labels on segments. Default is TRUE.
clean_labels	Logical; if TRUE, replaces underscores with spaces in item label names. Default is TRUE.
title	Optional character string; custom plot title. Default is NULL.

Value

A publication-ready ggplot2 object.

plot_metal_indices	<i>Visualize Heavy Metal Pollution Indices with Confidence Intervals</i>
--------------------	--

Description

Generates publication-ready plots for heavy metal pollution metrics (HPI, HEI, Cd, PLI, or Nemerow PN) across monitoring stations, featuring 95% confidence interval whiskers, regulatory benchmark lines, and intuitive status fills.

Usage

```
plot_metal_indices(  
  data,  
  metric = c("HPI", "HEI", "Cd", "PLI", "Nemerow_PN"),  
  id_col = "Station",  
  type = c("bar", "lollipop"),  
  show_ci = TRUE,  
  show_threshold = TRUE,  
  palette = "auto",  
  title = NULL,  
  ...  
)
```

Arguments

data	A list returned by calc_metal_indices() or a raw data.frame containing water quality or sediment heavy metal concentrations.
metric	Character string; index to visualize: "HPI" (Heavy Metal Pollution Index, default), "HEI" (Heavy Metal Evaluation Index), "Cd" (Degree of Contamination), "PLI" (Pollution Load Index), or "Nemerow_PN" (Nemerow Pollution Index). Default is "HPI".
id_col	Character string; column name in data identifying sampling stations or monitoring sites. Default is "Station".
type	Character string; plot display geometry: "bar" (default) or "lollipop". Default is "bar".
show_ci	Logical; if TRUE, displays 95% confidence interval whiskers when replicate observations exist per station. Default is TRUE.
show_threshold	Logical; if TRUE, overlays horizontal dashed reference lines representing regulatory benchmark thresholds. Default is TRUE.
palette	Character string; status color theme: "auto" for standard ecological condition grading (green = clean, red = polluted) or a valid RColorBrewer palette name. Default is "auto".

title	Optional character string; custom plot title. Default is NULL.
...	Additional arguments passed to <code>calc_metal_indices()</code> when raw data is provided.

Value

A publication-ready ggplot2 object.

plot_pie3d	<i>Publication-Ready 3D Depth Pie Chart</i>
------------	---

Description

Generates an isometric 3D pie chart with perspective tilt, extruded depth ribbons, and realistic shading natively in ggplot2.

Usage

```
plot_pie3d(
  data,
  x,
  y = NULL,
  stat = c("sum", "mean", "identity"),
  tilt = 0.5,
  depth = 0.25,
  show_labels = TRUE,
  min_percent = 4,
  palette = "Set2",
  title = NULL
)
```

Arguments

data	A <code>data.frame</code> , <code>matrix</code> , or character string file path to a <code>.csv</code> or <code>.xlsx</code> file containing the category data.
x	Character string; column name in data representing the categorical factor for slice partitioning.
y	Optional character string; column name in data representing the numeric variable to aggregate. If NULL, observation frequencies (counts) are plotted. Default is NULL.
stat	Character string; aggregation statistic when y is provided: "sum" (default), "mean", or "identity" (values plotted as-is). Default is "sum".
tilt	Numeric; perspective vertical compression factor between 0.2 and 0.8. Default is 0.5.
depth	Numeric; extrusion depth thickness for the 3D base rim. Default is 0.25.

show_labels	Logical; if TRUE, displays percentage labels directly on the slices. Default is TRUE.
min_percent	Numeric; minimum threshold percentage required to display slice label text. Default is 4.
palette	Character string; a valid RColorBrewer palette name. Default is "Set2".
title	Optional character string; custom plot title. Default is NULL.

Value

A publication-ready ggplot2 object.

plot_sankey	<i>Publication-Ready Sankey and Alluvial Flow Diagram</i>
-------------	---

Description

Generates publication-ready multi-tier alluvial and Sankey flow diagrams to visualize categorical transitions, trophic pathways, diet shifts, or spatial routing in ecological and environmental data. Automatically pre-aggregates frequencies, guards against missing strata, dynamically adapts color palettes, and returns a fully customizable ggplot2 object.

Usage

```
plot_sankey(  
  data,  
  axes,  
  value = NULL,  
  fill_axis = 1,  
  alpha = 0.65,  
  width = 0.25,  
  stratum_fill = "grey92",  
  stratum_color = "grey40",  
  label_size = 3.5,  
  palette = "Set2",  
  title = NULL  
)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing the observations.
axes	Character vector; column names in data specifying two or more categorical variables to serve as consecutive vertical axes (strata) from left to right.
value	Optional character string; column name in data representing the flow magnitude, frequency, or biomass. If NULL, observation frequencies (row counts) are calculated automatically. Default is NULL.

fill_axis	Character string or integer; column name or index in axes used to color the alluvial ribbons. Default is 1.
alpha	Numeric; transparency of the alluvial flow ribbons (between 0 and 1). Default is 0.65.
width	Numeric; relative horizontal width of stratum bars (between 0 and 1). Default is 0.25.
stratum_fill	Character string; fill color of the stratum bounding boxes. Default is "grey92".
stratum_color	Character string; border color of the stratum bounding boxes. Default is "grey40".
label_size	Numeric; font size for stratum category labels. Default is 3.5.
palette	Character string; a valid RColorBrewer palette name. Default is "Set2".
title	Optional character string; custom plot title. Default is NULL.

Value

A publication-ready ggplot2 alluvial diagram object.

plot_timeseries	<i>Publication-Ready Temporal Tracking and Factorial Time-Series Plots</i>
-----------------	--

Description

Generates publication-ready time-series and line graphs for aquatic, limnological, and fisheries datasets. Supports single-factor trajectories, two-factor comparisons, and three-factor nested layouts with automated replicate aggregation (Mean $\pm$ SE, SD, or 95

Usage

```
plot_timeseries(
  data,
  time_var,
  y_var,
  color_var = NULL,
  linetype_var = NULL,
  facet_var = NULL,
  ribbon = c("ci", "se", "sd", "none"),
  show_points = TRUE,
  point_size = 2.5,
  line_width = 1,
  ribbon_alpha = 0.2,
  smooth = FALSE,
  color_palette = "Dark2",
  xlab = NULL,
  ylab = NULL,
  title = NULL
)
```

Arguments

data	A <code>data.frame</code> , <code>matrix</code> , or file path to a <code>.csv</code> or <code>.xlsx</code> file containing the time-series data.
time_var	Character string; column name in data representing time steps (Date, POSIXct, numeric year, or calendar factor).
y_var	Character string; column name in data representing the continuous numeric response variable.
color_var	Optional character string; column name in data mapped to line and point colors (secondary factor). Default is <code>NULL</code> .
linetype_var	Optional character string; column name in data mapped to line styles (tertiary factor). Default is <code>NULL</code> .
facet_var	Optional character string; column name in data used to facet plots into sub-panels. Default is <code>NULL</code> .
ribbon	Character string; uncertainty band style when replicates exist per time point: <code>"ci"</code> (95 "sd" (± 1 Standard Deviation), or <code>"none"</code> . Default is <code>"ci"</code> .
show_points	Logical; if <code>TRUE</code> (default), renders points at each observation or aggregated mean time step.
point_size	Numeric; size of data points. Default is <code>2.5</code> .
line_width	Numeric; thickness of trajectory lines. Default is <code>1.0</code> .
ribbon_alpha	Numeric; opacity of uncertainty ribbons (0 to 1). Default is <code>0.20</code> .
smooth	Logical; if <code>TRUE</code> , adds a smoothing trendline. Default is <code>FALSE</code> .
color_palette	Character string; a valid <code>RColorBrewer</code> palette name. Default is <code>"Dark2"</code> .
xlab	Optional character string; custom x-axis label. Default is <code>NULL</code> .
ylab	Optional character string; custom y-axis label. Default is <code>NULL</code> .
title	Optional character string; custom plot title. Default is <code>NULL</code> .

Value

A list containing:

Plot	The publication-ready <code>ggplot2</code> object.
Summary_Data	A tidy data frame of aggregated Means, SDs, SEs, and Confidence Intervals.

plot_wqi

Visualize Water Quality Index (WQI) with Confidence Intervals

Description

Generates publication-ready WQI plots across monitoring stations with 95% confidence interval whiskers, status fills, and background benchmark bands for WAWQI classification tiers.

Usage

```
plot_wqi(
  data,
  id_col = "Station",
  type = c("bar", "lollipop"),
  show_ci = TRUE,
  show_bands = TRUE,
  title = NULL,
  ...
)
```

Arguments

<code>data</code>	A list returned by <code>calc_wqi()</code> or a raw <code>data.frame</code> containing physicochemical water quality observations.
<code>id_col</code>	Character string; column name in <code>data</code> identifying sampling stations or monitoring sites. Default is "Station".
<code>type</code>	Character string; plot display geometry: "bar" (default) or "lollipop". Default is "bar".
<code>show_ci</code>	Logical; if TRUE, displays 95% confidence interval whiskers when replicate observations exist per station. Default is TRUE.
<code>show_bands</code>	Logical; if TRUE, overlays shaded horizontal background bands representing WAWQI rating classification tiers (e.g., Excellent, Good, Poor). Default is TRUE.
<code>title</code>	Optional character string; custom plot title. Default is NULL.
<code>...</code>	Additional arguments passed to <code>calc_wqi()</code> when raw data is provided.

Value

A publication-ready `ggplot2` object.

rda_analysis

Redundancy Analysis (RDA) Triplot

Description

Performs constrained Redundancy Analysis (RDA) combining community and environmental data. Automatically cleans empty samples from both datasets, runs permutation significance tests, and generates a publication-ready `ggplot2` Triplot with `ggrepel` label adjustments.

Usage

```
rda_analysis(
  comm_data,
  env_data,
  group_col = 1,
  transform = c("none", "hellinger", "log", "pa", "wisconsin"),
  color_palette = "Dark2",
  species_arrow_mult = 1,
  env_arrow_mult = 1
)
```

Arguments

<code>comm_data</code>	A data.frame or matrix where grouping metadata is at the start, followed by numeric species abundance counts.
<code>env_data</code>	A data.frame or matrix containing environmental predictor variables. Must have the exact same number of rows as <code>comm_data</code> .
<code>group_col</code>	Character string or integer; column name or index in <code>comm_data</code> specifying the categorical grouping factor. Default is 1.
<code>transform</code>	Character string; pre-transformation applied to community counts via <code>vegan::decostand</code> : "none" (default), "hellinger", "log", "pa" (presence/absence), or "wisconsin". Default is "none".
<code>color_palette</code>	Character string; a valid <code>RColorBrewer</code> palette name. Default is "Dark2".
<code>species_arrow_mult</code>	Numeric; scaling multiplier for species vector arrows. Default is 1.
<code>env_arrow_mult</code>	Numeric; scaling multiplier for environmental vector arrows. Default is 1.

Value

A list containing:

<code>RDA_Object</code>	The underlying rda ordination object from <code>vegan</code> .
<code>Model_Significance</code>	Permutation ANOVA test for overall model significance.
<code>Variable_Significance</code>	Permutation ANOVA test for marginal term significance.
<code>Plot</code>	The publication-ready <code>ggplot2</code> triplot.

Description

Fits linear, allometric power, exponential, logarithmic, quadratic polynomial, or cubic regression models. Automatically formats algebraic mathematical formulas, R-squared values, and p-values, displays confidence or prediction intervals, and renders publication-ready ggplot2 graphics.

Usage

```
regression_analysis(
  data,
  x_var,
  y_var,
  fit_type = c("linear", "power", "exponential", "logarithmic", "polynomial", "cubic"),
  group_var = NULL,
  facet = FALSE,
  interval = c("confidence", "prediction", "none"),
  show_equation = TRUE,
  color_palette = "Set1",
  point_size = 2.2,
  point_alpha = 0.75,
  title = NULL
)
```

Arguments

data	A data.frame, matrix, or tibble containing the variables for analysis.
x_var	Character string; column name in data representing the independent numeric predictor variable.
y_var	Character string; column name in data representing the dependent numeric response variable.
fit_type	Character string; regression model functional form: "linear" (default), "power" (allometric), "exponential", "logarithmic", "polynomial" (quadratic), or "cubic".
group_var	Optional character string; column name in data representing a categorical grouping factor for stratified analysis. Default is NULL.
facet	Logical; if TRUE and group_var is provided, facets panels by group. Default is FALSE.
interval	Character string; uncertainty band type: "confidence" (default), "prediction", or "none".
show_equation	Logical; if TRUE (default), prints the fitted algebraic formula, R^2 , and p-value on the canvas.
color_palette	Character string; a valid RColorBrewer palette name. Default is "Set1".

point_size	Numeric; size of scatter plot points. Default is 2.2.
point_alpha	Numeric; opacity of scatter plot points (0 to 1). Default is 0.75.
title	Optional character string; custom plot title. Default is NULL.

Value

A list containing:

Plot	Publication-ready ggplot2 object.
Models	Named list of fitted lm model objects per group.
Equation_Table	Data frame of algebraic formulas, R^2 , and p-values.
Predictions	Data frame of calculated predictions with confidence/prediction limits.

simper_analysis	<i>Similarity Percentages (SIMPER) Analysis</i>
-----------------	---

Description

Identifies which species primarily contribute to the differences between groups using Bray-Curtis dissimilarity decomposition (vegan::simper). Automatically cleans empty samples, validates missing values, supports ecological transformations, and generates a publication-ready ggplot2 bar chart.

Usage

```
simper_analysis(
  data,
  group_col = 1,
  top_n = 5,
  transform = c("none", "hellinger", "log", "pa", "wisconsin"),
  permutations = 999,
  seed = 42
)
```

Arguments

data	A data.frame or matrix where grouping metadata is at the start, followed by numeric species abundance counts.
group_col	Character string or integer; column name or index in data specifying the categorical grouping factor. Default is 1.
top_n	Numeric integer; number of top contributing species to display per pairwise comparison. Default is 5.
transform	Character string; pre-transformation applied to community counts via vegan::decostand: "none" (default), "hellinger" (recommended for abundance data), "log", "pa" (presence/absence), or "wisconsin". Default is "none".
permutations	Numeric integer; number of permutations for p -value estimation. Default is 999.
seed	Optional integer; random seed for reproducible permutation testing. Default is 42.

Value

A list containing:

SIMPER_Object	The raw simper ordination/decomposition object from vegan.
Summary_Table	Tidy data frame of top contributing species, average contributions, cumulative percentages, and significance per group comparison.
Plot	A publication-ready ggplot2 bar chart of top contributors.

specaccum_plot	<i>Publication-Ready Species Accumulation Curves</i>
----------------	--

Description

Computes and plots species accumulation curves using `vegan::specaccum`. Supports pooled curves or comparative grouped curves across habitats, handles methods without standard deviations defensively, and formats outputs with publication-ready styling.

Usage

```
specaccum_plot(
  data,
  group_col = 1,
  by_group = FALSE,
  method = "exact",
  palette = "Set2",
  line_color = "#2A9D8F",
  ci = 1.96,
  show_ci = TRUE,
  title = NULL
)
```

Arguments

data	A <code>data.frame</code> or matrix containing community abundance counts with optional grouping metadata column(s).
group_col	Character string or integer vector; column name(s) or index/indices in data specifying grouping metadata. Default is 1.
by_group	Logical; if TRUE, generates comparative accumulation curves for each level of the grouping factor. Default is FALSE.
method	Character string; accumulation method passed to <code>vegan::specaccum</code> (e.g., "exact", "random", "collector"). Default is "exact".
palette	Character string; a valid <code>RColorBrewer</code> palette name for grouped curves. Default is "Set2".
line_color	Character string; hex code or color name for single pooled curves. Default is "#2A9D8F".

ci	Numeric; multiplier for confidence intervals (e.g., 1.96 for 95% confidence intervals). Default is 1.96.
show_ci	Logical; if TRUE, displays confidence ribbons when standard deviations are available. Default is TRUE.
title	Optional character string; custom plot title. Default is NULL.

Value

A list containing:	
Specaccum_Object	The raw specaccum object (or list of objects if grouped) from vegan.
Curve_Data	A tidy data frame with sampling effort, estimated species richness, and confidence bounds.
Plot	The publication-ready ggplot2 object.

summarize_obs	<i>Comprehensive Observation Summary with Excel Export</i>
---------------	--

Description

Computes descriptive statistics (N, Mean, SD, SE, Median, Min, Max, and Mean $\pm$ SD) for numeric variables across optional grouping factors. Supports direct file ingestion and exporting publication-ready tables directly to Excel (.xlsx) or CSV.

Usage

```
summarize_obs(data, vars = NULL, group = NULL, digits = 2, file = NULL)
```

Arguments

data	A data.frame, matrix, or character string file path to a .csv or .xlsx file containing the observations.
vars	Optional character vector; column names in data representing continuous numeric variables to summarize. If NULL, all numeric columns are automatically evaluated. Default is NULL.
group	Optional character vector; column name(s) in data representing categorical grouping factor(s) (e.g., "Habitat" or c("Habitat", "Season")). Default is NULL.
digits	Numeric integer; number of decimal places for rounding reported statistics. Default is 2.
file	Optional character string; export file path for the resulting summary table (must end in .xlsx or .csv). Default is NULL.

Value

A tidy data.frame containing computed descriptive metrics.

Examples

```
# Generate self-contained sample survey data
fish_survey <- data.frame(
  River = rep(c("River A", "River B"), each = 15),
  Length = c(rnorm(15, 20, 2), rnorm(15, 24, 3)),
  Weight = c(rnorm(15, 120, 15), rnorm(15, 180, 20))
)

summary_tbl <- summarize_obs(
  data = fish_survey,
  vars = c("Length", "Weight"),
  group = "River",
  digits = 2
)
head(summary_tbl)
```

test_chisq	<i>Chi-Square Test of Independence and Goodness-of-Fit with Visualization</i>
------------	---

Description

Performs Pearson's Chi-Square (χ^2) test of independence or goodness-of-fit on categorical survey, ecological, or environmental frequency data. Automatically evaluates expected cell frequencies, applies Yates' continuity correction or Monte Carlo simulation when sample sizes are small, computes effect sizes (Cramér's V and Phi coefficient ϕ), extracts standardized Pearson residuals, and generates a publication-ready ggplot2 stacked or dodged bar chart.

Usage

```
test_chisq(
  data,
  x,
  y = NULL,
  p = NULL,
  correct = TRUE,
  simulate_p = TRUE,
  B = 2000,
  plot_type = "stacked",
  palette = "Set2",
  title = NULL
)
```

Arguments

data	A data.frame, matrix, table, or character string file path to a .csv or .xlsx file containing the observations.
x	Character string; column name in data representing the primary categorical factor (rows of the contingency table).
y	Optional character string; column name in data representing the secondary categorical factor (columns of the contingency table). If NULL, a Chi-Square goodness-of-fit test is performed on x. Default is NULL.
p	Optional numeric vector; theoretical probabilities for the goodness-of-fit test when y = NULL. Must sum to 1 and match the number of levels in x. Default is NULL (equal probabilities across categories).
correct	Logical; if TRUE, applies Yates' continuity correction for 2×2 contingency tables. Default is TRUE.
simulate_p	Logical; if TRUE, computes <i>p</i> -values via Monte Carlo simulation when expected cell counts fall below 5. Default is TRUE.
B	Numeric integer; number of Monte Carlo replicates used when simulate_p = TRUE. Default is 2000.
plot_type	Character string; visual geometry: "stacked" (100% stacked proportion bar chart; default) or "dodged" (grouped absolute count bar chart). Default is "stacked".
palette	Character string; a valid RColorBrewer palette name. Default is "Set2".
title	Optional character string; custom plot title. Default is NULL.

Value

A list containing:

Chi_Square_Test

Tidy data frame containing the χ^2 statistic, degrees of freedom, *p*-value, test method, and effect sizes (Cramér's *V* and Phi ϕ).

Observed_Counts

Cross-tabulated data frame of observed cell frequencies with marginal totals.

Expected_Counts

Data frame of theoretical expected cell frequencies under the null hypothesis of independence.

Standardized_Residuals

Data frame of standardized Pearson residuals ($|r| > 2$ indicates significant deviation).

Plot

A publication-ready ggplot2 bar chart annotated with test statistics.

Model_Object

The underlying htest object returned by stats::chisq.test.

three_way_anova	<i>Three-Way Analysis of Variance (ANOVA) with Faceted Interaction Visualization</i>
-----------------	--

Description

Performs a three-way factorial ANOVA with full interaction effects, calculates Tukey's HSD post-hoc tests, extracts Estimated Marginal Means (EMMs), and generates faceted publication-ready interaction plots with compact letter displays. Automatically handles NA values safely, converts matrix inputs, and coerces categorical variables.

Usage

```
three_way_anova(
  data,
  factor1_var,
  factor2_var,
  factor3_var,
  numeric_var,
  factor1_levels = NULL,
  factor2_levels = NULL,
  factor3_levels = NULL,
  plot_type = "boxplot",
  error_type = "se",
  y_limits = NULL,
  add_jitter = TRUE,
  show_mean = TRUE,
  mean_color = "darkred",
  color_palette = "Set1"
)
```

Arguments

data	A data frame or matrix in long format containing the variables for analysis.
factor1_var	Character string; column name in data representing the first categorical factor (x-axis).
factor2_var	Character string; column name in data representing the second categorical factor (fill/color/groups).
factor3_var	Character string; column name in data representing the third categorical factor (facet panels).
numeric_var	Character string; column name in data representing the continuous numeric response variable.
factor1_levels	Optional character vector; custom level ordering for factor 1. Default is NULL.
factor2_levels	Optional character vector; custom level ordering for factor 2. Default is NULL.
factor3_levels	Optional character vector; custom level ordering for factor 3. Default is NULL.

plot_type	Character string; display geometry: "boxplot", "barplot", or "pointrange". Default is "boxplot".
error_type	Character string; error bar dispersion metric: "se" (Standard Error) or "sd" (Standard Deviation). Default is "se".
y_limits	Optional numeric vector of length 2; lower and upper limits for the y-axis (e.g., c(0, 100)). Default is NULL.
add_jitter	Logical; if TRUE, overlays jittered raw observation points. Default is TRUE.
show_mean	Logical; if TRUE, displays group mean indicator diamonds. Default is TRUE.
mean_color	Character string; hex code or color name for group mean indicators. Default is "darkred".
color_palette	Character string; a valid RColorBrewer palette name. Default is "Set1".

Value

A list containing:

ANOVA_Summary	Global ANOVA summary table with 3-way interaction terms.
TukeyHSD	Tukey HSD pairwise comparison results across interaction levels.
EMMs	Estimated Marginal Means object from emmeans.
Summary_Table	Aggregated sample sizes, means, SDs, SEs, and confidence limits.
Plot	The publication-ready faceted ggplot2 object.

two_way_anova	<i>Two-Way Analysis of Variance (ANOVA) with Interaction & Faceting Visualization</i>
---------------	---

Description

Performs a two-way factorial ANOVA with interaction effects, calculates Tukey’s HSD post-hoc tests, extracts Estimated Marginal Means (EMMs), and generates publication-ready interaction plots with optional facetting.

Usage

```
two_way_anova(  
  data,  
  factor1_var,  
  factor2_var,  
  numeric_var,  
  facet_var = NULL,  
  factor1_levels = NULL,  
  factor2_levels = NULL,  
  plot_type = "boxplot",  
  error_type = "se",
```

```

    y_limits = NULL,
    add_jitter = TRUE,
    show_mean = TRUE,
    mean_color = "darkred",
    color_palette = "Set1"
  )

```

Arguments

<code>data</code>	A data frame containing the variables for analysis.
<code>factor1_var</code>	Character string; column name in data representing the primary categorical factor (x-axis).
<code>factor2_var</code>	Character string; column name in data representing the secondary categorical factor (fill/color/groups).
<code>numeric_var</code>	Character string; column name in data representing the continuous numeric response variable.
<code>facet_var</code>	Optional character string; column name in data to facet panels by (e.g., <code>factor2_var</code>). Default is <code>NULL</code> .
<code>factor1_levels</code>	Optional character vector; custom level ordering for factor 1. Default is <code>NULL</code> .
<code>factor2_levels</code>	Optional character vector; custom level ordering for factor 2. Default is <code>NULL</code> .
<code>plot_type</code>	Character string; display geometry: "boxplot", "barplot", or "pointrange". Default is "boxplot".
<code>error_type</code>	Character string; error bar dispersion metric: "se" (Standard Error) or "sd" (Standard Deviation). Default is "se".
<code>y_limits</code>	Optional numeric vector of length 2; lower and upper limits for the y-axis (e.g., <code>c(0, 100)</code>). Default is <code>NULL</code> .
<code>add_jitter</code>	Logical; if <code>TRUE</code> , overlays jittered raw observation points. Default is <code>TRUE</code> .
<code>show_mean</code>	Logical; if <code>TRUE</code> , displays group mean indicator diamonds. Default is <code>TRUE</code> .
<code>mean_color</code>	Character string; hex code or color name for group mean indicators. Default is "darkred".
<code>color_palette</code>	Character string; a valid <code>RColorBrewer</code> palette name. Default is "Set1".

Value

A list containing:

<code>ANOVA_Summary</code>	Global ANOVA summary table with interaction terms.
<code>TukeyHSD</code>	Tukey HSD pairwise comparison results across interaction levels.
<code>EMMs</code>	Estimated Marginal Means object from <code>emmeans</code> .
<code>Summary_Table</code>	Aggregated sample sizes, means, SDs, SEs, and confidence limits.
<code>Plot</code>	The publication-ready <code>ggplot2</code> object.

Index

* datasets

- community_master, 8
- hilsa_master, 14

ancova_analysis, 3

calc_diversity, 4
calc_metal_indices, 5
calc_wqi, 6
cluster_dendrogram, 7
community_master, 8
compare_two_groups, 9
correlation_heatmap, 10

dbrda_analysis, 11

glm_analysis, 13

hilsa_master, 14

manova_analysis, 15

mantel_heatmap_analysis, 16

nmds_analysis, 17

one_way_anova, 18

pca_community_biplot, 20
pca_correlation_circle, 21
permanova_analysis, 22
plot_bar, 23
plot_circular_bar, 25
plot_diversity, 26
plot_donut, 27
plot_histogram, 28
plot_likert, 29
plot_lollipop(plot_bar), 23
plot_metal_indices, 30
plot_pie3d, 31
plot_sankey, 32
plot_timeseries, 33

plot_wqi, 34

rda_analysis, 35

regression_analysis, 37

simper_analysis, 38

specaccum_plot, 39

summarize_obs, 40

test_chisq, 41

three_way_anova, 43

two_way_anova, 44