

Package ‘ggResidpanel’

June 1, 2019

Type Package

Title Panels and Interactive Versions of Diagnostic Plots using
'ggplot2'

Version 0.3.0

Description An R package for creating panels of diagnostic plots for residuals from a model using ggplot2 and plotly to analyze residuals and model assumptions from a variety of viewpoints. It also allows for the creation of interactive diagnostic plots.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

URL <https://goodekat.github.io/ggResidpanel/>

Imports cowplot, ggplot2, grDevices, grid, MASS, plotly, qqplotr,
stats, stringr

RoxygenNote 6.1.1

Suggests dplyr, forcats, knitr, lme4, lmerTest, nlme, randomForest,
rmarkdown, rpart, testthat, vdiff, wesanderson

VignetteBuilder knitr

Depends R (>= 3.0.0)

NeedsCompilation no

Author Katherine Goode [aut, cre],
Kathleen Rey [aut]

Maintainer Katherine Goode <kgoode@iastate.edu>

Repository CRAN

Date/Publication 2019-05-31 23:20:04 UTC

R topics documented:

penguins	2
resid_auxpanel	2
resid_compare	4

resid_interact	7
resid_panel	9
resid_xpanel	12

Index	14
--------------	-----------

penguins	<i>Penguins Dataset</i>
----------	-------------------------

Description

A dataset that contains information from a study done on dives of emperor penguins by Jessica Meir and others. The scientists were interested in understanding how the heart rate of the penguins relates to the depth and duration of the dive. The study involved attaching a device to penguins that recorded the heart rate of the bird during a dive. The dataset contains multiple observations recorded from 9 penguins. The dataset has 125 observations and 4 variables.

Usage

penguins

Format

A data.frame.

Details

The variables in the dataset are as follows.

heartrate	Heart rate of the penguin during the dive (beats per minute)
depth	Depth of the dive (meters)
duration	Duration of the dive (minutes)
bird	The id number associated with a penguin

Source

https://dasl.datadescription.com/datafile/penguins/?_sfm_methods=Regression&_sfm_cases=4+59943&sf_paged=9

resid_auxpanel	<i>Panel of Diagnostic Residual Plots.</i>
----------------	--

Description

Creates a panel of residual diagnostic plots given inputs of residuals and fitted values.

Usage

```
resid_auxpanel(residuals, predicted, plots = "default", bins = 30,
  smoother = FALSE, qqline = TRUE, qqbands = FALSE, scale = 1,
  theme = "bw", axis.text.size = 10, title.text.size = 12,
  title.opt = TRUE, nrow = NULL)
```

Arguments

<code>residuals</code>	The residuals from the model.
<code>predicted</code>	The fitted values from the model.
<code>plots</code>	Plots chosen to include in the panel of plots. The default panel includes a residual plot, a normal quantile plot, an index plot, and a histogram of the residuals. (See details for the options available.)
<code>bins</code>	Number of bins for histogram of the residuals. Default is set to 30.
<code>smoother</code>	Indicates whether or not to include a smoother on the residual plot and/or index plot. Specify TRUE or FALSE. Default is set to FALSE.
<code>qqline</code>	Indicates whether to include a 1-1 line on the qq-plot. Specify TRUE or FALSE. Default is set to TRUE.
<code>qqbands</code>	Indicates whether to include confidence bands on the qq-plot. Specify TRUE or FALSE. Default is set to FALSE.
<code>scale</code>	Scales the size of the graphs in a panel. Takes values in (0,1].
<code>theme</code>	ggplot2 theme to be used. Options are "bw", "classic", and "grey" (or "gray"). Default is "bw".
<code>axis.text.size</code>	Specifies the size of the text for the axis labels of all plots.
<code>title.text.size</code>	Specifies the size of the text for the titles of all plots.
<code>title.opt</code>	Indicates whether or not to include a title on the plots. Specify TRUE or FALSE. Default is set to TRUE.
<code>nrow</code>	Sets the number of rows in the panel.

Details

The following grid options can be chosen for the plots argument.

- "all": This creates a panel of all plot types included in the package that are available for `resid_auxpanel`. (See plot descriptions under individual options.)
- "default": This creates a panel with a residual plot, a normal quantile plot of the residuals, an index plot of the residuals, and a histogram of the residuals.
- "SAS": This creates a panel of a residual plot, a normal quantile plot of the residuals, a histogram of the residuals, and a boxplot of the residuals. This was modeled after the `residpanel` option in `proc mixed` from SAS version 9.4.
- A vector of individual plots can also be specified. For example, one can specify `plots = c("boxplot", "hist")` or `plots = "qq"`. The individual plot options are as follows.
 - "boxplot": A boxplot of residuals

- "hist": A histogram of residuals
- "index": A plot of residuals versus observation number
- "qq": A normal quantile plot of residuals
- "resid": A plot of residuals versus predicted values

Details on the creation of the plots can be found in the details section of the help file for resid_panel.

Value

A panel of residual diagnostic plots containing plots specified.

Examples

```
# Fit a regression tree to the penguins data
penguin_tree <- rpart::rpart(heartrate ~ depth + duration, data = penguins)

# Obtain the predictions from the model on the observed data
penguin_tree_pred <- predict(penguin_tree)

# Obtain the residuals from the model
penguin_tree_resid <- penguins$heartrate - penguin_tree_pred

# Create a panel with the residual and index plot
resid_auxpanel(residuals = penguin_tree_resid,
               predicted = penguin_tree_pred,
               plots = c("resid", "index", "yvp"))
```

resid_compare

Panel of Diagnostic Residual Plots Across Multiple Models.

Description

Creates a panel of residual diagnostic plots given a list of models. Currently accepts models of type "lm", "glm", "lmerMod", "lmerModLmerTest", and "glmerMod".

Usage

```
resid_compare(models, plots = "default", type = NA, bins = 30,
              smoother = FALSE, qqline = TRUE, qqbands = FALSE, scale = 1,
              theme = "bw", axis.text.size = 10, title.text.size = 12,
              title.opt = TRUE, nrow = NULL)
```

Arguments

<code>models</code>	List of models fit using either <code>lm</code> , <code>glm</code> , <code>lmer</code> , <code>lmerTest</code> , or <code>glmer</code> .
<code>plots</code>	Plots chosen to include in the panel of plots. The default panel includes a residual plot, a normal quantile plot, an index plot, and a histogram of the residuals. (See details for the options available.)
<code>type</code>	Type of residuals to use in the plot. If not specified, the default residual type for each model type is used. (See details for the options available.)
<code>bins</code>	Number of bins to use when creating a histogram of the residuals. Default is set to 30.
<code>smoother</code>	Indicates whether or not to include a smoother on the index, residual-leverage, location-scale, and residual plots. Specify <code>TRUE</code> or <code>FALSE</code> . Default is set to <code>FALSE</code> .
<code>qqline</code>	Indicates whether to include a 1-1 line on the qq-plot. Specify <code>TRUE</code> or <code>FALSE</code> . Default is set to <code>TRUE</code> .
<code>qqbands</code>	Indicates whether to include confidence bands on the qq-plot. Specify <code>TRUE</code> or <code>FALSE</code> . Default is set to <code>FALSE</code> .
<code>scale</code>	Scales the size of the graphs in the panel. Takes values in (0,1].
<code>theme</code>	ggplot2 theme to be used. Current options are <code>"bw"</code> , <code>"classic"</code> , and <code>"grey"</code> (or <code>"gray"</code>). Default is <code>"bw"</code> .
<code>axis.text.size</code>	Specifies the size of the text for the axis labels of all plots in the panel.
<code>title.text.size</code>	Specifies the size of the text for the titles of all plots in the panel.
<code>title.opt</code>	Indicates whether or not to include a title on the plots in the panel. Specify <code>TRUE</code> or <code>FALSE</code> . Default is set to <code>TRUE</code> .
<code>nrow</code>	Sets the number of rows in the panel.

Details

The first two sections below contain information on the available input options for the plots and type arguments in `resid_compare`. The third section contains details relating to the creation of the plots.

Options for Plots

The following options can be chosen for the `plots` argument.

- `"all"`: This creates a panel of all plot types included in the package that are available for the model type input into `residpanel`. (See note below.)
- `"default"`: This creates a panel with a residual plot, a normal quantile plot of the residuals, an index plot of the residuals, and a histogram of the residuals.
- `"R"`: This creates a panel with a residual plot, a normal quantile plot of the residuals, a location-scale plot, and a leverage versus residuals plot. This was modeled after the plots shown in R if the `plot()` base function is applied to an `lm` model. This option can only be used with an `lm` or `glm` model.

- "SAS": This creates a panel with a residual plot, a normal quantile plot of the residuals, a histogram of the residuals, and a boxplot of the residuals. This was modeled after the residpanel option in proc mixed from SAS version 9.4.
- A vector of individual plots can also be specified. For example, one can specify plots = c("boxplot", "hist") or plots = "qq". The individual plot options are as follows.
 - "boxplot": A boxplot of residuals
 - "cookd": A plot of Cook's D values versus observation numbers
 - "hist": A histogram of residuals
 - "index": A plot of residuals versus observation numbers
 - "ls": A location scale plot of the residuals
 - "qq": A normal quantile plot of residuals
 - "lev": A plot of leverage values versus residuals
 - "resid": A plot of residuals versus predicted values
 - "yvp": A plot of observed response values versus predicted values

Note: "cookd", "ls", and "lev" are only available for "lm" and "glm" models.

Options for Type

Several residual types are available to be requested based on the model type that is input into resid_panel. These currently are as follows.

- lm residual options
 - "pearson": The Pearson residuals
 - "response": The raw residuals (Default for "lm")
 - "standardized": The standardized raw residuals
- glm residual options
 - "pearson": The Pearson residuals
 - "deviance": The deviance residuals (Default for "glm")
 - "response": The raw residuals
 - "stand.deviance": The standardized deviance residuals
 - "stand.pearson": The standardized Pearson residuals
- lmer, lmerTest, and lme residual options
 - "pearson": The Pearson residuals (Default for "lmer", "lmerTest", and "lme")
 - "response": The raw residuals
- glmer residual options
 - "pearson": The Pearson residuals
 - "deviance": The deviance residuals (Default for "glmer")
 - "response": The raw residuals

Note: The plots of "ls" and "lev" only accept standardized residuals.

Details on the Creation of Plots

Boxplot (boxplot) Boxplot of the residuals.

- Cook's D** (cookd) The horizontal line represents a cut-off to identify highly influential points. The horizontal line is placed at $4/n$ where n is the number of data points used in the model.
- Histogram** (hist) Plots a histogram of the residuals. The density curve overlaid has mean equal to zero and standard deviation equal to the standard deviation of the residuals.
- Index Plot** (index) Plots the residuals on the y-axis and the observation number associated with the residual on the x-axis.
- Leverage Plot** (lev) Plots the standardized residuals on the y-axis and the leverage values on the x-axis. A lowess curve is overlaid, and Cook's D contours are included for $\alpha = 0.5$ and $\alpha = 1$.
- Location-Scale Plot** (ls) Plots the square root of the absolute value of the standardized residuals on the y-axis and the predicted values on the x-axis. The predicted values are plotted on the original scale for glm and glmer models. A lowess curve is overlaid.
- QQ Plot** (qq) Makes use of the R package qqplotr for creating a normal quantile plot of the residuals.
- Residual Plot** (resid) Plots the residuals on the y-axis and the predicted values on the x-axis. The predicted values are plotted on the original scale for glm and glmer models.
- Response vs. Predicted** (yvp) Plots the response variable from the model on the y-axis and the predicted values on the x-axis. Both response variable and predicted values are plotted on the original scale for glm and glmer models.

Value

A panel of residual diagnostic plots containing plots specified for each model.

Examples

```
# Fit two models to the penguins data
penguin_model <- lme4::lmer(heartrate ~ depth + duration + (1|bird), data = penguins)
penguin_model_log2 <- lme4::lmer(log(heartrate) ~ depth + duration + I(duration^2) +
(1|bird), data = penguins)

# Compare the residuals from the model
resid_compare(list(penguin_model, penguin_model_log2))

# Adjust some options in the panel of plots
resid_compare(list(penguin_model, penguin_model_log2), plots = c("resid", "yvp"),
smoother = TRUE, theme = "grey")
```

resid_interact

Panel of Interactive Versions of Diagnostic Residual Plots.

Description

Creates a panel of interactive residual diagnostic plots given a model. Currently accepts models of type "lm", "glm", "lmerMod", "lmerModLmerTest", "lme", and "glmerMod".

Usage

```
resid_interact(model, plots = "default", type = NA, bins = 30,
  smoother = FALSE, qqline = TRUE, scale = 0.9, theme = "bw",
  axis.text.size = 10, title.text.size = 12, title.opt = TRUE,
  nrow = NULL)
```

Arguments

<code>model</code>	Model fit using either <code>lm</code> , <code>glm</code> , <code>lmer</code> , <code>lmerTest</code> , <code>lme</code> , or <code>glmer</code> .
<code>plots</code>	Plots chosen to include in the panel of plots. The default panel includes a residual plot, a normal quantile plot, an index plot, and a histogram of the residuals. (See details in the help file for <code>resid_panel</code> for the options available.)
<code>type</code>	Type of residuals to use in the plot. If not specified, the default residual type for each model type is used. (See details in the help file for <code>resid_panel</code> for the options available.)
<code>bins</code>	Number of bins to use when creating a histogram of the residuals. Default is set to 30.
<code>smoother</code>	Indicates whether or not to include a smoother on the index, residual-leverage, location-scale, and residual plots. Specify <code>TRUE</code> or <code>FALSE</code> . Default is set to <code>FALSE</code> .
<code>qqline</code>	Indicates whether to include a 1-1 line on the qq-plot. Specify <code>TRUE</code> or <code>FALSE</code> . Default is set to <code>TRUE</code> . (The option of <code>qqbands</code> has not been implemented in <code>plotly</code> , so it is not available as an option with <code>resid_interact</code> .)
<code>scale</code>	Scales the size of the graphs in the panel. Takes values in (0,1].
<code>theme</code>	<code>ggplot2</code> theme to be used. Current options are <code>"bw"</code> , <code>"classic"</code> , and <code>"grey"</code> (or <code>"gray"</code>). Default is <code>"bw"</code> .
<code>axis.text.size</code>	Specifies the size of the text for the axis labels of all plots in the panel.
<code>title.text.size</code>	Specifies the size of the text for the titles of all plots in the panel.
<code>title.opt</code>	Indicates whether or not to include a title on the plots in the panel. Specify <code>TRUE</code> or <code>FALSE</code> . Default is set to <code>TRUE</code> .
<code>nrow</code>	Sets the number of rows in the panel.

Details

Details on the creation of the plots can be found in the details section of the help file for `resid_panel`.

Value

A panel of interactive residual diagnostic plots containing plots specified.

Examples

```
# Fit a model to the penguin data
penguin_model <- lme4::lmer(heartrate ~ depth + duration + (1|bird), data = penguins)

# Create the default interactive panel
resid_interact(penguin_model)

# Select all plots to include in the panel and set the smoother option to TRUE
resid_interact(penguin_model, plots = "all", smoother = TRUE)

# Select only the residual plot and qq-plot to be included in the panel and
# set the number of rows to 2
resid_interact(penguin_model, plots = c("resid", "qq"), nrow = 2)

# Choose the SAS panel of plots, change the theme to classic, and remove the
# titles of the plots
resid_interact(penguin_model, plots = "SAS", theme = "classic", title.opt = FALSE)
```

resid_panel

*Panel of Diagnostic Residual Plots.***Description**

Creates a panel of residual diagnostic plots given a model. Currently accepts models of type "lm", "glm", "lmerMod", "lmerModLmerTest", "lme", and "glmerMod".

Usage

```
resid_panel(model, plots = "default", type = NA, bins = 30,
  smoother = FALSE, qqline = TRUE, qqbands = FALSE, scale = 1,
  theme = "bw", axis.text.size = 10, title.text.size = 12,
  title.opt = TRUE, nrow = NULL)
```

Arguments

model	Model fit using either lm, glm, lmer, lmerTest, lme, or glmer.
plots	Plots chosen to include in the panel of plots. The default panel includes a residual plot, a normal quantile plot, an index plot, and a histogram of the residuals. (See details for the options available.)
type	Type of residuals to use in the plot. If not specified, the default residual type for each model type is used. (See details for the options available.)
bins	Number of bins to use when creating a histogram of the residuals. Default is set to 30.
smoother	Indicates whether or not to include a smoother on the index, residual-leverage, location-scale, and residual plots. Specify TRUE or FALSE. Default is set to FALSE.

qqline	Indicates whether to include a 1-1 line on the qq-plot. Specify TRUE or FALSE. Default is set to TRUE.
qqbands	Indicates whether to include confidence bands on the qq-plot. Specify TRUE or FALSE. Default is set to FALSE.
scale	Scales the size of the graphs in the panel. Takes values in (0,1].
theme	ggplot2 theme to be used. Current options are "bw", "classic", and "grey" (or "gray"). Default is "bw".
axis.text.size	Specifies the size of the text for the axis labels of all plots in the panel.
title.text.size	Specifies the size of the text for the titles of all plots in the panel.
title.opt	Indicates whether or not to include a title on the plots in the panel. Specify TRUE or FALSE. Default is set to TRUE.
nrow	Sets the number of rows in the panel.

Details

The first two sections below contain information on the available input options for the plots and type arguments in `resid_panel`. The third section contains details relating to the creation of the plots.

Options for Plots

The following options can be chosen for the `plots` argument.

- "all": This creates a panel of all plot types included in the package that are available for the model type input into `resid_panel`. (See note below.)
- "default": This creates a panel with a residual plot, a normal quantile plot of the residuals, an index plot of the residuals, and a histogram of the residuals.
- "R": This creates a panel with a residual plot, a normal quantile plot of the residuals, a location-scale plot, and a residuals versus leverage plot. This was modeled after the plots shown in R if the `plot()` base function is applied to an `lm` model. This option can only be used with an `lm` or `glm` model.
- "SAS": This creates a panel with a residual plot, a normal quantile plot of the residuals, a histogram of the residuals, and a boxplot of the residuals. This was modeled after the `residualpanel` option in `proc mixed` from SAS version 9.4.
- A vector of individual plots can also be specified. For example, one can specify `plots = c("boxplot", "hist")` or `plots = "qq"`. The individual plot options are as follows.
 - "boxplot": A boxplot of residuals
 - "cookd": A plot of Cook's D values versus observation numbers
 - "hist": A histogram of residuals
 - "index": A plot of residuals versus observation numbers
 - "ls": A location scale plot of the residuals
 - "qq": A normal quantile plot of residuals
 - "lev": A plot of standardized residuals versus leverage values
 - "resid": A plot of residuals versus predicted values
 - "yvp": A plot of observed response values versus predicted values

Note: "cookd", "1s", and "lev" are only available for "lm" and "glm" models.

Options for Type

Several residual types are available to be requested based on the model type that is input into resid_panel. These currently are as follows.

- lm residual options
 - "pearson": The Pearson residuals
 - "response": The raw residuals (Default for "lm")
 - "standardized": The standardized raw residuals
- glm residual options
 - "pearson": The Pearson residuals
 - "deviance": The deviance residuals (Default for "glm")
 - "response": The raw residuals
 - "stand.deviance": The standardized deviance residuals
 - "stand.pearson": The standardized Pearson residuals
- lmer, lmerTest, and lme residual options
 - "pearson": The Pearson residuals (Default for "lmer", "lmerTest", and "lme")
 - "response": The raw residuals
- glmer residual options
 - "pearson": The Pearson residuals
 - "deviance": The deviance residuals (Default for "glmer")
 - "response": The raw residuals

Note: The plots of "1s" and "lev" only accept standardized residuals.

Details on the Creation of Plots

Boxplot (boxplot) Boxplot of the residuals.

Cook's D (cookd) The horizontal line represents a cut-off to identify highly influential points. The horizontal line is placed at $4/n$ where n is the number of data points used in the model.

Histogram (hist) Plots a histogram of the residuals. The density curve overlaid has mean equal to zero and standard deviation equal to the standard deviation of the residuals.

Index Plot (index) Plots the residuals on the y-axis and the observation number associated with the residual on the x-axis.

Residual-Leverage Plot (lev) Plots the standardized residuals on the y-axis and the leverage values on the x-axis with a loess curve is overlaid. Cook's D contour lines (which are a function of leverage and standardized residuals) are plotted as the red dashed lines for Cook's D values of 0.5 and 1.

Location-Scale Plot (1s) Plots the square root of the absolute value of the standardized residuals on the y-axis and the predicted values on the x-axis. The predicted values are plotted on the original scale for glm and glmer models. A loess curve is overlaid.

QQ Plot (qq) Makes use of the R package qqplotr for creating a normal quantile plot of the residuals.

Residual Plot (resid) Plots the residuals on the y-axis and the predicted values on the x-axis. The predicted values are plotted on the original scale for glm and glmer models.

Response vs. Predicted (yvp) Plots the response variable from the model on the y-axis and the predicted values on the x-axis. Both response variable and predicted values are plotted on the original scale for glm and glmer models.

Value

A panel of residual diagnostic plots containing plots specified.

Examples

```
# Fit a model to the penguin data
penguin_model <- lme4::lmer(heartrate ~ depth + duration + (1|bird), data = penguins)

# Create the default panel
resid_panel(penguin_model)

# Select all plots to include in the panel and set the smoother option to TRUE
resid_panel(penguin_model, plots = "all", smoother = TRUE)

# Select only the residual plot and qq-plot to be included in the panel,
# request confidence bands on the qq plot, and set the number of rows to 2
resid_panel(penguin_model, plots = c("resid", "qq"), qqbands = TRUE, nrow = 2)

# Choose the SAS panel of plots, change the theme to classic, and remove the
# titles of the plots
resid_panel(penguin_model, plots = "SAS", theme = "classic", title.opt = FALSE)
```

resid_xpanel	<i>Panel of Plots of Residuals or Response Variable versus Predictor Variables.</i>
--------------	---

Description

Creates a panel of plots of the residuals or response variable versus the predictor (x) variables in the model. Interactions between predictor variables are not included. Currently accepts models of type "lm", "glm", "lmerMod", "lmerModLmerTest", "lme", and "glmerMod".

Usage

```
resid_xpanel(model, yvar = "residual", type = NA, smoother = FALSE,
  scale = 1, theme = "bw", axis.text.size = 10,
  title.text.size = 12, title.opt = TRUE, nrow = NULL)
```

Arguments

model	Model fit using either <code>lm</code> , <code>glm</code> , <code>lmer</code> , <code>lmerTest</code> , <code>lme</code> , or <code>glmer</code> .
yvar	Specifies the variable to put on the y-axis of the plots. Options are "residual" (default) or "response".
type	Type of residuals to use in the plot. If not specified, the default residual type for each model type is used. (See details for the options available.)
smoother	Indicates whether or not to include a smoother on the plots. Specify TRUE or FALSE. Default is set to FALSE.
scale	Scales the size of the graphs in the panel. Takes values in (0,1].
theme	ggplot2 theme to be used. Current options are "bw", "classic", and "grey" (or "gray"). Default is "bw".
axis.text.size	Specifies the size of the text for the axis labels of all plots in the panel.
title.text.size	Specifies the size of the text for the titles of all plots in the panel.
title.opt	Indicates whether or not to include a title on the plots in the panel. Specify TRUE or FALSE. Default is set to TRUE.
nrow	Sets the number of rows in the panel.

Details

Note that for x variables that are factors, the levels shown on the x-axis will be in the order that the levels are ordered in the dataframe. This can be adjusted by reordering the levels of the factor before the model is fit.

Value

A panel of plots of the residuals or response variable versus the predictor variables.

Examples

```
# Fit a model to the penguin data
penguin_model <- lme4::lmer(heartrate ~ depth + duration + (1|bird), data = penguins)

# Create plots of the residuals versus the predictor variables
resid_xpanel(penguin_model, theme = "classic")

# Create plots of the response variable versus the predictor variables
resid_xpanel(penguin_model, yvar = "response", theme = "classic", smoother = TRUE)
```

Index

*Topic **datasets**

penguins, [2](#)

penguins, [2](#)

resid_auxpanel, [2](#)

resid_compare, [4](#)

resid_interact, [7](#)

resid_panel, [9](#)

resid_xpanel, [12](#)