

ReDiag Tutorial: Illustrated Examples

Introduction

We showcase the versatile capabilities of ReDiag through two examples using real datasets. The steps, which are also available in the *Manual tab* of the app, can be summarized as follows:

1. Data Input
2. Define the Model
3. Model Diagnostics
4. Data Transformation
5. Reassemble the Model
6. Generate a Report

Example 1: Behavioural Test Data

This example uses data from a study involving behavioural outcomes in C57BL/6 mice subjected to a bone marrow transplantation procedure. The dataset includes variables for locomotion and anxiety-related behaviour. We assess the impact of treatment and anxiety on velocity, highlighting issues with linearity and how to correct them.

Steps for Example 1

1. Data Input: Select the '*Open field test*' example dataset in the *Data Input* tab.
2. Define the Model:

Fit the initial model:

$$velocity = \beta_0 + \beta_1 treatment + \beta_2 anxiety +$$

3. Model Diagnostics: Residual analysis reveals violations of linearity (Figure 1a).

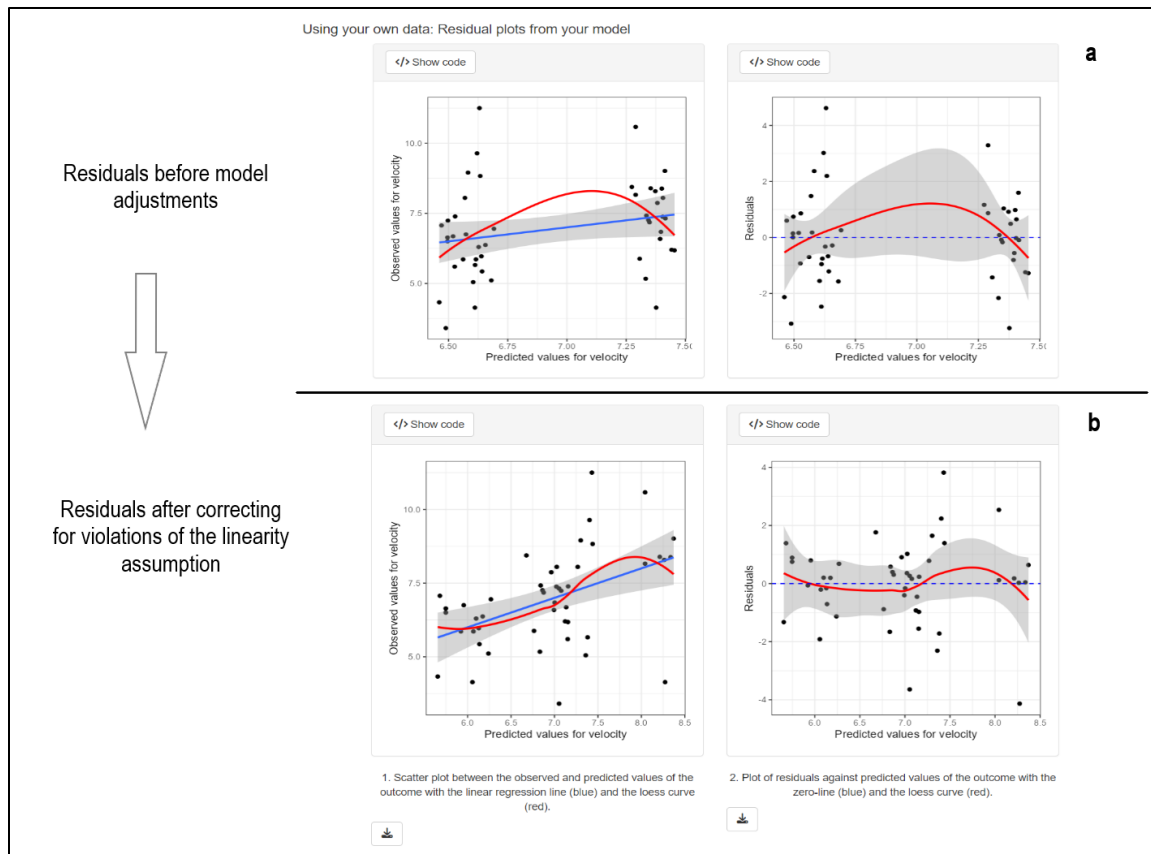


Figure 1. Residual analysis of the fitted model. Screenshot of residual plots to assess the linearity assumption: (a) plots before adjustments and (b) after correcting for violations.

4. Remedies: Address omitted variable bias by including 'Age' in the model.

5. Updated Model:

Fit the updated model:

$$velocity = \beta_0 + \beta_1 treatment + \beta_2 age + \beta_3 anxiety + \varepsilon$$

6. Results: Residuals now meet the linearity assumption (Figure 1b), consistent with the study's findings.

Example 2: Cage Occupancy Data

This example uses data from a study on the impact of cage occupancy and bedding volume on ammonia build-up in ventilated cages. We demonstrate how ReDiag can assess normality and homoscedasticity assumptions and apply a Box-Cox transformation to address violations.

Steps for Example 2

1. Data Input: Load the 'Occupancy' example dataset on ammonia levels in ventilated cages.

2. Define the Model:

Fit the initial model:

$$\text{ammonia} = \beta_0 + \beta_1 \text{occupancy} + \beta_2 \text{cage} + \beta_3 \text{breeding} + \beta_4 \text{occupancy} * \text{cage} + \varepsilon$$

3. Model Diagnostics: Identify violations of normality and homoscedasticity using residual plots (Figures 2a, 3a).

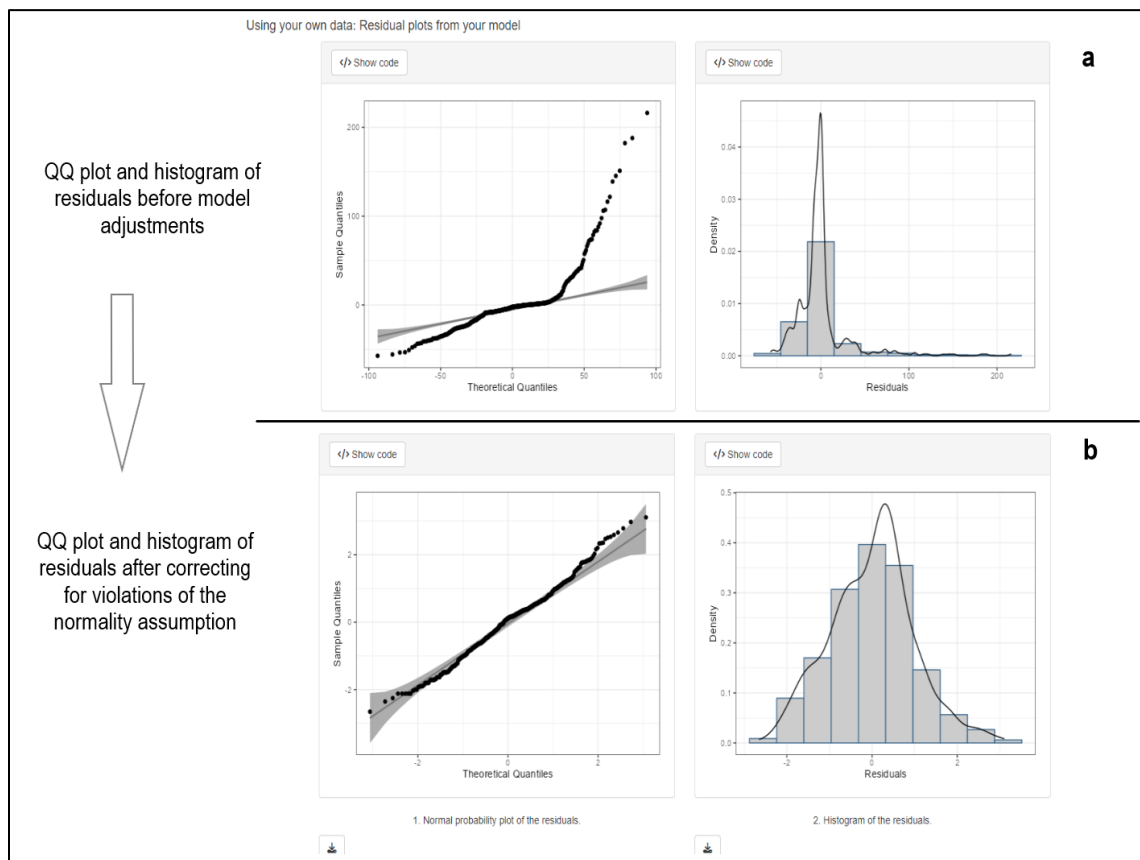


Figure 2. Residual analysis of the fitted model. Screenshot of plots to assess the normality assumption: (a) plots before adjustments and (b) after correcting for violations.

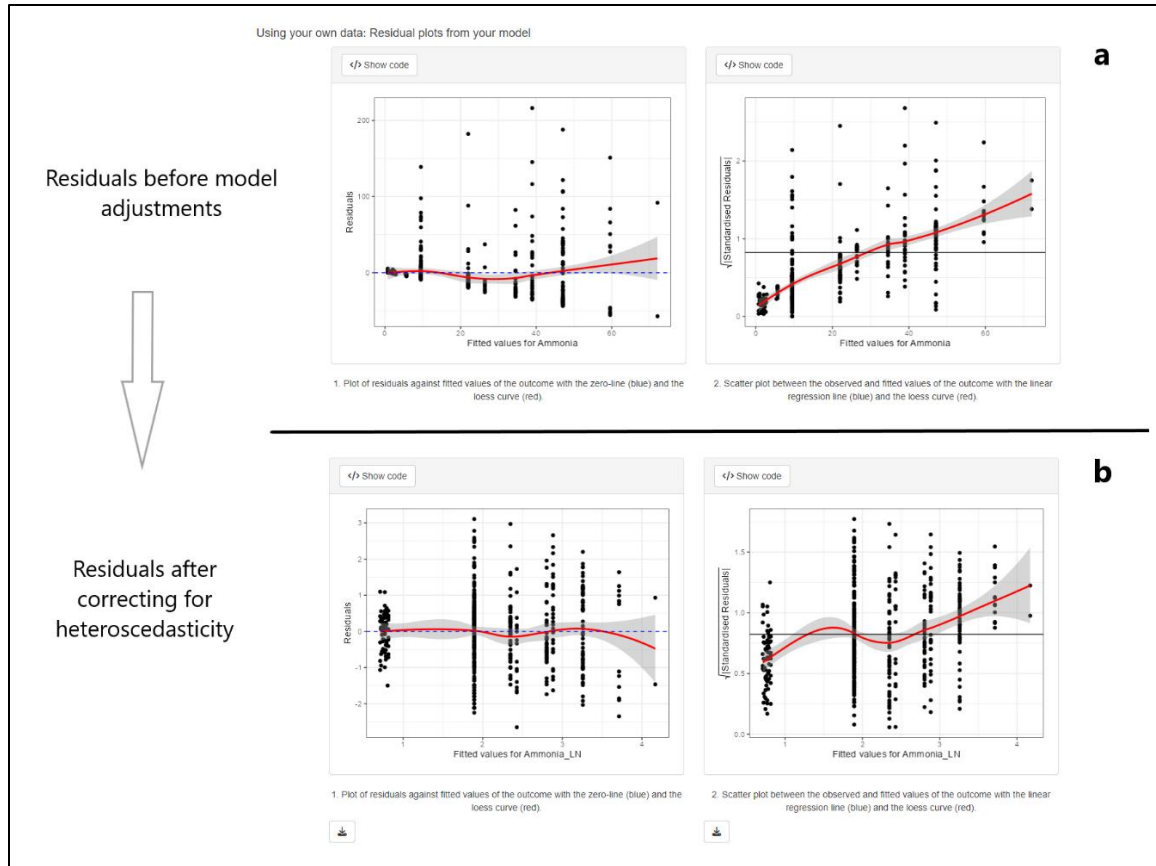


Figure 3. Residual analysis of the fitted model. Screenshot of plots to assess the homoscedasticity assumption: (a) plots before adjustments and (b) after correcting for violations.

4. Remedies: Apply a Box-Cox transformation (Figure 4) to correct violations.

5. Updated Model:

Reassemble the model using a log-transformed outcome:

$$\log(\text{ammonia}) = \beta_0 + \beta_1 \text{occupancy} + \beta_2 \text{cage} + \beta_3 \text{breeding} + \beta_4 \text{occupancy} * \text{cage} + \varepsilon$$

6. Results: Residuals meet normality and homoscedasticity assumptions (Figures 2b, 3b).

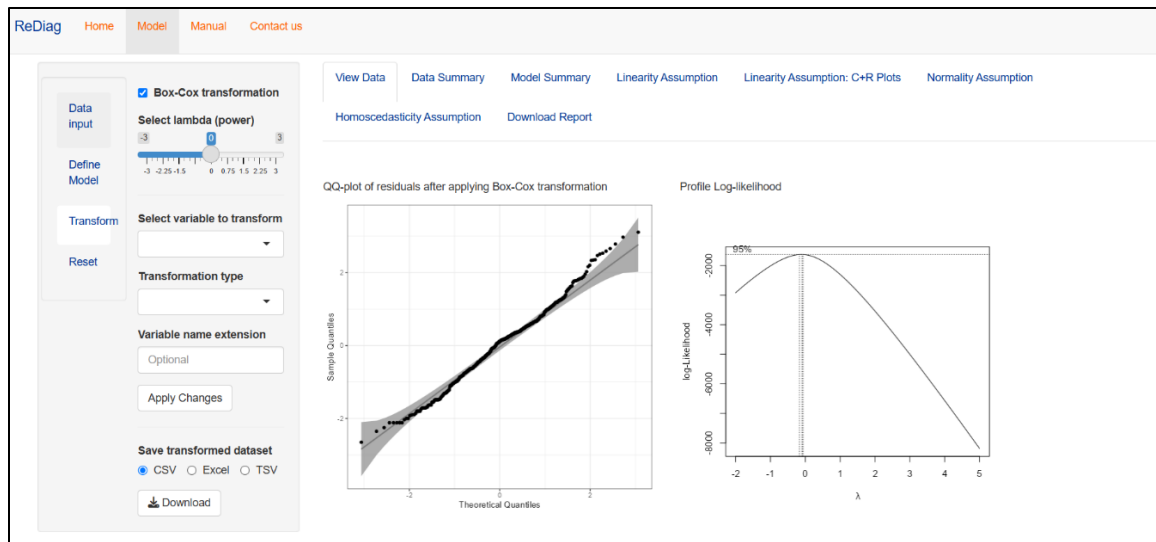


Figure 4. Box-Cox transformations. A screenshot of the interactive plot where the slider was utilised to search for the correct power transformation.

Conclusion

ReDiag effectively diagnoses and addresses violations in regression assumptions. By leveraging interactive visualisations and transformations, it ensures robust model fitting and valid statistical inferences, making it an invaluable tool for researchers.