

Let's practice - 7

Import the mouse data from the file `course_dataset/mice_data_mod.csv`.
This file contains the same data as `mice_data.csv` and in addition, two more columns.

1. Run **str()** to check your data frame: did it load correctly?
2. Convert genotype and diet to factor variables.
3. Make a **scatter plot** of respiratory rate against mouse weights using the function `plot()`.
 - use solid circles as plotting symbol
 - add a title
 - customize the axis labels ("Weight [g]", "Respiratory Rate [bpm]")
 - color the points **by genotype**.
4. Fit a trend line using the function **abline()**
5. Add a legend for the genotype

Let's practice - 8

This is a continuation of the previous practice: we will continue to plot the mice data

1. Plot a **histogram** of mouse weight and customize it with title, labels, colors. Represent the density line on top.
2. Make **boxplots** of weights from WT and KO mice. Customize with title, labels, colors.
3. *Optional:* Repeat 2 with diet instead of genotype.

Let's practice - 9

This is a continuation of the previous practice: we will continue to plot the mice data

1. Make a multi-panel figure with the **four graphics (from the previous exercises) on one page**
2. Change the code to export the figure to a **pdf** file with paper size A4. Set width and height arguments in the call to pdf() to make it look nice.
3. **Optional:** Export an histogram (from previous exercise) to a **png** file. Set width and height arguments in the call to png() to make it look nice.

Let's practice - 10

Come back to the mice data-set stored in the "mice_data" data frame.

- 1) Considering WT mice weight and KO mice weight separately, check the assumption of normality graphically.
- 2) Make an appropriate plot to visualize the mouse weights grouped by genotype.
- 3) Perform a test to see whether the mouse weight is different between the two genotypes.
- 4) *Repeat steps 1 to 3 for the diet variable.*

Let's practice - 11

The data set "Pima" comes from a study on diabetes in women of Pima Indian heritage.

We are using a subset (Pima.tr).

1. Load the package MASS using `library()`. Load the dataset Pima.tr using `data(Pima.tr)`. Use `?Pima.tr` to get an idea which variables it contains.
2. Hypothesis: Blood glucose level (glu) is associated with diastolic blood pressure (bp). Run a linear model to test the hypothesis.
3. Visualize the fit with a scatter plot and a trend line.
4. Check assumptions of the model (homoscedasticity, mean of residual at 0, normality of the residuals) graphically.