



Data Visualization Exercises

12 July 2024

Licence: CC-by-sa

Before you start:

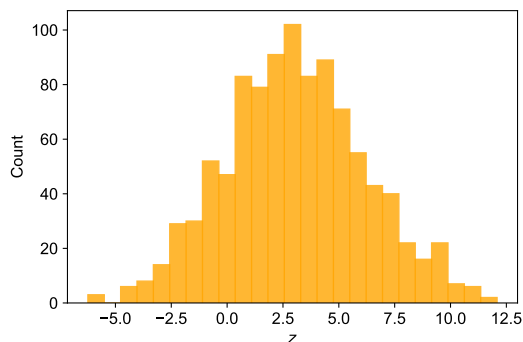
- create a suitable directory for this exercise
- download the lecture material from http://www.physik.uzh.ch/~python/python/lecture_visual/ for reference.

For the first four exercises you find in the material notebooks where the steps to load the data are already included, so that you can focus on creating the charts.

Exercise 1: Replicating plots – Simple

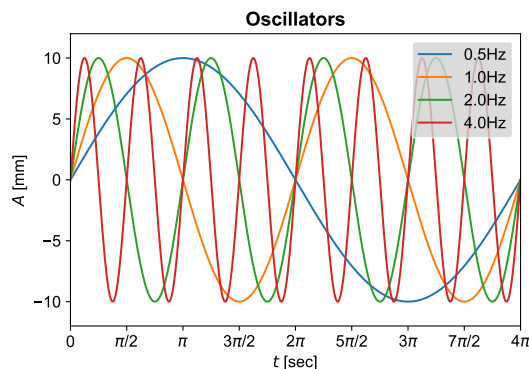
Try to replicate the following plots. The data required is stored in the file `Example_1_data.csv`.

Histogram of the variable z

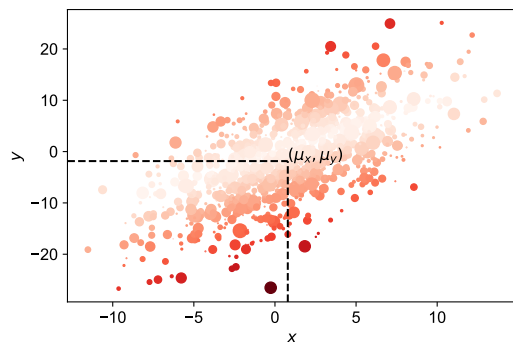


- Color is orange
- There are 25 bins.

Line chart of oscillations



- The lines represent the functions $A(t) = A_0 \sin(2\pi\nu t)$ with $A_0 = 10$ for $\nu = 0.5, 1, 2, 4$.
- Create a legend with a grey background.
- Put ticks at every $\pi/2$ -step.

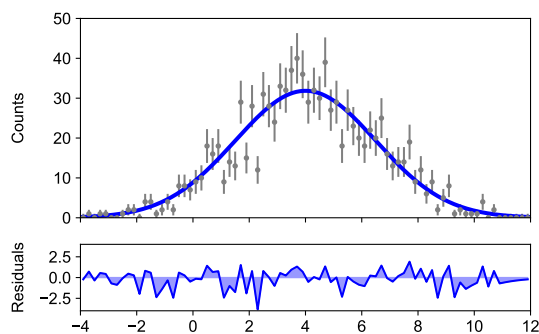


Scatter plot of x and y

- The size is defined by the variable z .
- The value representing the color is $|y - x|$.
- The color map is `Reds`.
- Dashed black lines show the mean of x and y , μ_x and μ_y .
- There is an annotation of the point (μ_x, μ_y) .

Exercise 2: Replicating plots – Medium

Try to replicate the following plot. The data required is stored in the file `Example_2_data.csv` with `x`, `y`, and `yerr` of the data points (*i.e.* no uncertainty for x). The mean of the distribution is $\mu = 4.0$ and the standard deviation is $\sigma = 2.5$.

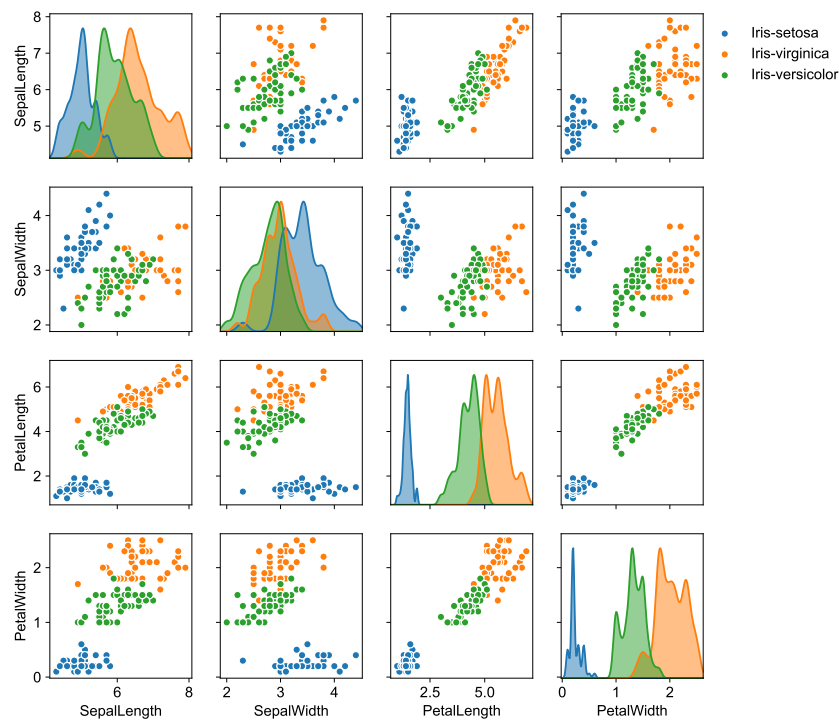


Data points with a line chart and secondary plot for residuals

- The top chart is three times higher than the bottom one.
- The data points are shown as gray error-bars.
- The line is showing a normal distribution with the aforementioned parameters.
- The residuals are shown as an area using `ax.fill_between`.

Exercise 3: Replicating plots – Complex

Try to replicate the pairplot shown in the lecture, but without `seaborn`. Just use `matplotlib`. The data is stored in `iris.csv`.



Exercise 4: Geo-spatial Visualisation

There is a shape file of the Boroughs of London at `statistical-gis-boundaries-london/ESRI/London_Borough_Excluding_MHW.shp`.

You can join it with the file `london-borough-profiles.csv`. Create the following visualisations:

- A Choropleth map showing the happiness of the different Boroughs (Variable `Happiness_score_2011-14_(out_of_10)`)
- B Centroids of the Boroughs (You can get the coordinates of the centroids of elements in a geopandas data frame `gdf` via `gdf.geometry.centroid.x` and `gdf.geometry.centroid.y`.)
- C Choropleth map showing the population of the different Boroughs (Use the natural-breaks scheme to select the different colors.)