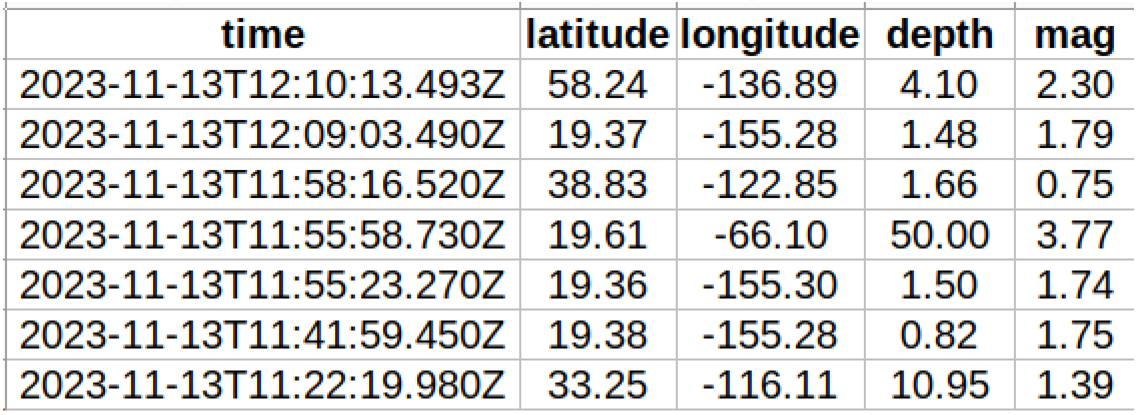
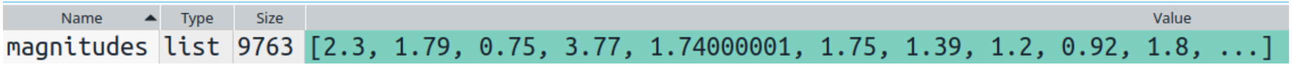
Exercise 1: Histogram, Box Plot and Violin Plot of the Earthquake Magnitudes

The file “`earthquakes_Oct_2023.csv`” (available from Moodle) contains the details of earthquakes over a period of a month. The following is an excerpt:



time	latitude	longitude	depth	mag
2023-11-13T12:10:13.493Z	58.24	-136.89	4.10	2.30
2023-11-13T12:09:03.490Z	19.37	-155.28	1.48	1.79
2023-11-13T11:58:16.520Z	38.83	-122.85	1.66	0.75
2023-11-13T11:55:58.730Z	19.61	-66.10	50.00	3.77
2023-11-13T11:55:23.270Z	19.36	-155.30	1.50	1.74
2023-11-13T11:41:59.450Z	19.38	-155.28	0.82	1.75
2023-11-13T11:22:19.980Z	33.25	-116.11	10.95	1.39

The Python program `mpl_earthquakes.py` retrieves the data from the file and returns a list of the earthquake magnitudes:



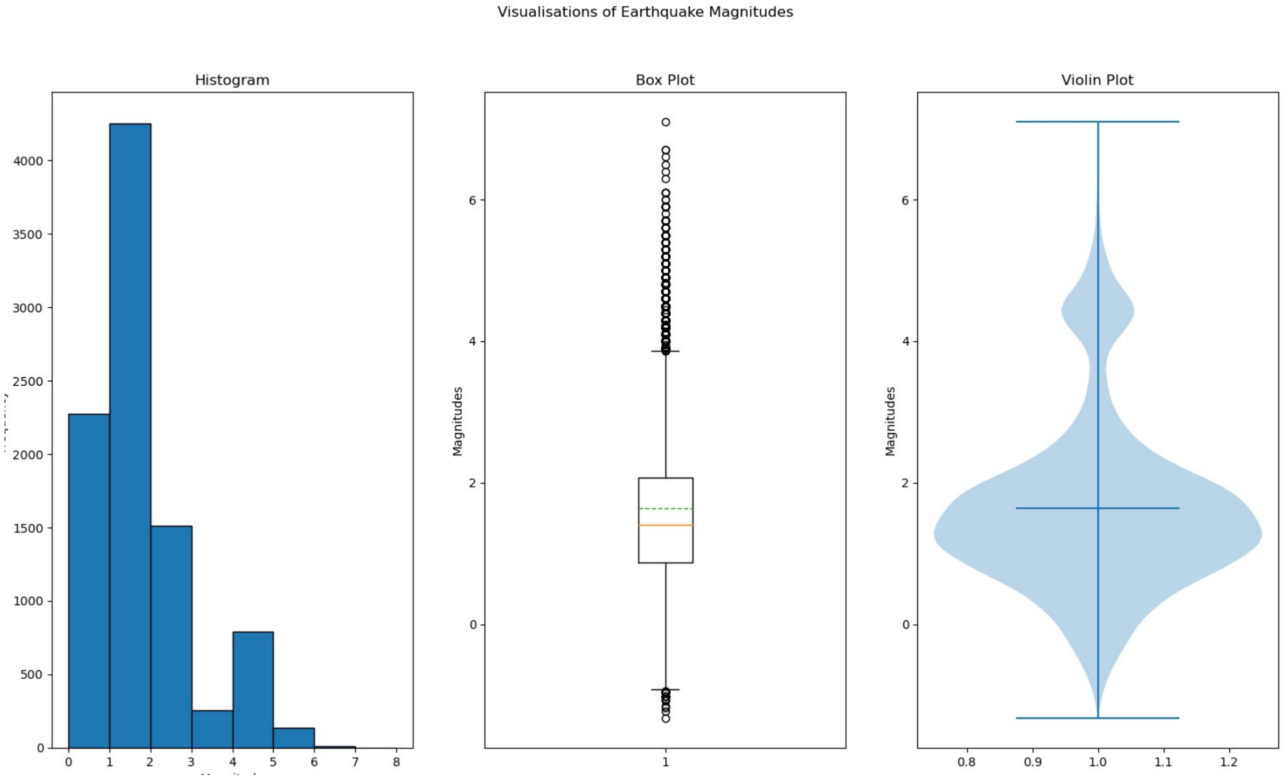
Name ▲	Type	Size	Value
magnitudes	list	9763	[2.3, 1.79, 0.75, 3.77, 1.74000001, 1.75, 1.39, 1.2, 0.92, 1.8, ...]

Modify the program to display the following:

- A histogram
- A box plot
- A violin plot

of the earthquake magnitudes.

Sample Visualisation



Guidelines

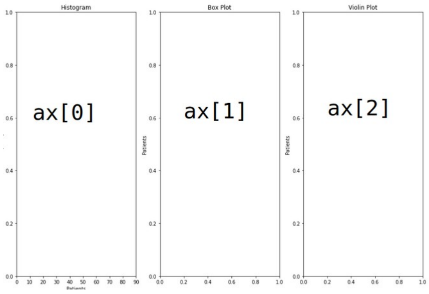
Create a Figure `fig` and Axes `axs` which can show the three plots in a row:

```
     fig, axs = plt.subplots(1,3, figsize=(15,10))
```

`(1,3)` means 1 row and 3 columns. If you want 3 rows and 1 column, use `subplots(3)`

You can set a title for the figure using `fig.suptitle(`*`text`*`)`

You'll refer to the 3 plots as `axs[0]` `axs[1]` and `axs[2]`



For the box plot, use the function `boxplot()`
with keyword arguments `showmeans` and `meanline` set to `True`

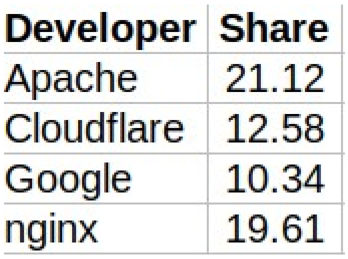
For the violin plot, use the function `violinplot()`
with keyword argument `showmeans` to `True`

For the histogram, you need to specify the groups for the bars, called *bins*:

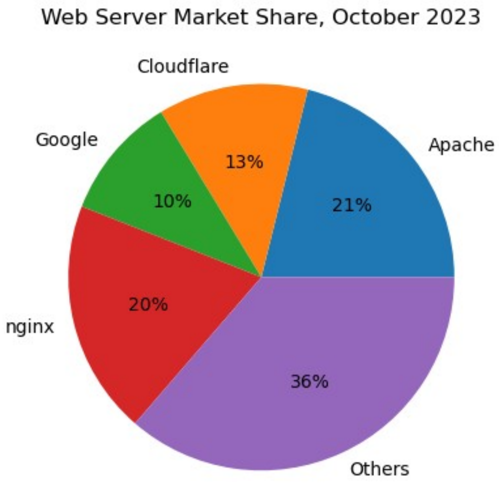
1. Create a list for the `bins` using `range(int(max(magnitudes)+2)`
2. Set the `xticks` to the `bins` list
3. Display the histogram using the function `hist` with `magnitudes` and `bins`.
   For greater visibility of the bars, use `ec="black"` (`ec` stands for edge colour).

Exercise 2: Web Server Market Share

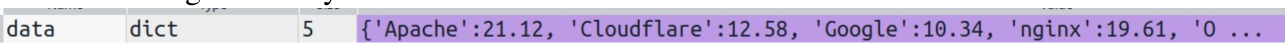
The file `webservers_202310.csv` contains information on the market share of the most popular web servers from October 2023 (https://news.netcraft.com/archives/category/web-server-survey/)



Developer	Share
Apache	21.12
Cloudflare	12.58
Google	10.34
nginx	19.61



The program `mpl_webservers.py` retrieves the information from the file and provides the data in the following dictionary:



data	dict	5	{'Apache':21.12, 'Cloudflare':12.58, 'Google':10.34, 'nginx':19.61, 'O ...

Modify the program so that it creates a Pie Chart to visualise the data, as follows:

1) Import `matplotlib.pyplot` as `plt`

2) Create a Figure `fig` and Axes `ax` using `plt.subplots()`.

3) Set the title, for the Axes `ax`.

4) Create the pie chart, including the dictionary keys as labels and percentage share on each slice:

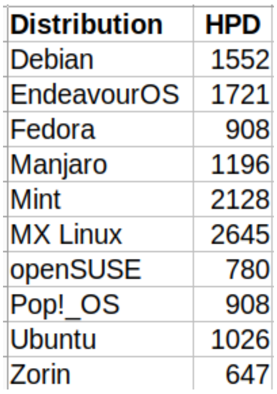
```
ax.pie(dict.values(),labels=dict.keys(),autopct="%.f%%")
```

5) Show the plot (if not using Spyder)

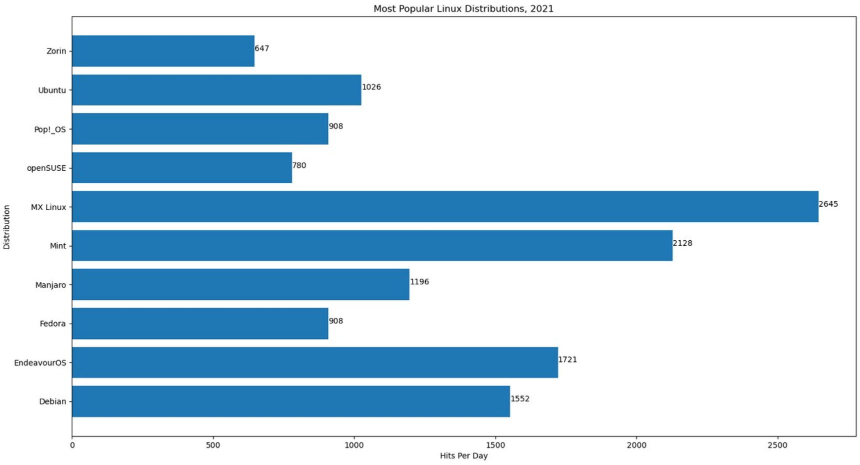
6) Save the figure and upload it to Moodle.

Exercise 3: Popularity of Linux Distributions

The file `distrowatch_Nov_2023.csv` contains information on the Top 10 most popular Linux Distributions of 2023 (https://distrowatch.com/)

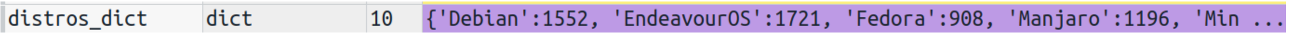


Distribution	HPD
Debian	1552
EndeavourOS	1721
Fedora	908
Manjaro	1196
Mint	2128
MX Linux	2645
openSUSE	780
Pop!_OS	908
Ubuntu	1026
Zorin	647



HPD = hits per day

The program `mpl_distrowatch.py` retrieves the information from the file and provides the data in the following dictionary:



```
distros_dict    dict    10    {'Debian':1552, 'EndeavourOS':1721, 'Fedora':908, 'Manjaro':1196, 'Min ...
```

Modify the program so that it creates a Horizontal Bar Chart to visualise the data, as follows:

1) Import `matplotlib.pyplot` as `plt`

2) Create a Figure `fig` and Axes `ax` using `plt.subplots()`.

3) Set the title, x-axis label and y-axis label for the Axes `ax`.

4) Create the bar chart, using the dictionary keys as labels:

```
    ax.barh(list(dict.keys()),dict.values())
```

5) Display the values at the end of the bars:

```
        for index, value in enumerate(dict.values()):
            ax.text(value,index,str(value))
```

6) Show the plot (if not using Spyder)

7) Save the figure and upload it to Moodle.