WEEK9 Assignment-1 Data Visualization

Data Visualization Assignment Problem using Matplotlib and seaborn

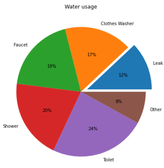
Data Folder Link:https://github.com/svkarthik86/Advanced-python/tree/main/WEEK-9%20Assignment/Data

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
%matplotlib inline
```

Question 1

Creating a Pie Chart for Water Usage Use pandas to read the data located in the subfolder data. # Load dataset data = pd.read_csv(./Data/water_usage.csv') Use a pie chart to visualize the water usage. Highlight the water Leak part percentages using the explode parameter. Show the percentages for each slice and add a title. **Sample Output**

In [2]:
```
water_leakage=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/water_usage.csv")
plt.figure(figsize=(7,7))
plt.pie(data=water_leakage,x="Percentage",labels="Usage",explode=(0.1,0,0,0,0,0),autopct="%1.f%%")
plt.title("Water usage")
plt.show()
```



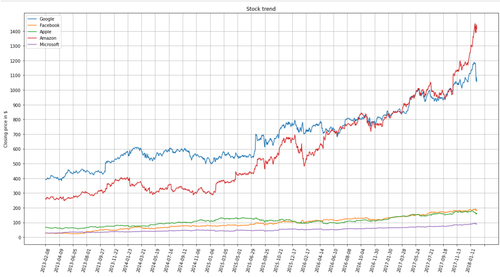
Question 2

Visualizing Stock Trends by Using a Line Plot create a line plot to show stock trends. Let's look at the following scenario: You are interested in investing in stocks. You downloaded the stock prices for the "big five": Amazon,Google, Apple, Facebook, and Microsoft.In this activity,we will create a line plot to show stock trends. Let's look at the following scenario: You are interested in investing in stocks. You downloaded the stock prices for the "big five": Amazon, Google, Apple, Facebook, and Microsoft.

Use pandas to read the data located in the subfolder data.

Use Matplotlib to create a line chart visualizing the closing prices for the past five years (whole data sequence) for all five companies. Add labels, titles, and a legend to make the visualization self-explanatory. Use plt.grid() to add a grid to your plot. **Sample Output**

In [3]:
```
google=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/GOOGL_data.csv")
amazon=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/AMZN_data.csv")
facebook=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/FB_data.csv")
apple=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/AAPL_data.csv")
microsoft=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/MSFT_data.csv")

data=pd.concat([google.date,google.close,facebook.close,apple.close,amazon.close,microsoft.close],axis=1)
data.columns=["date","Google","Facebook","Apple","Amazon","Microsoft"]
data.plot(x="date",figsize=(20,10),xticks=(np.arange(0,1300,40)),yticks=(np.arange(0,1450,100)))
plt.xticks(rotation=75)
plt.xlabel(None)
plt.title("Stock trend")
plt.ylabel("Closing price in $")
plt.grid()
plt.show()
```



Question 3

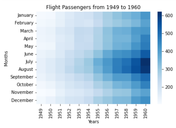
Using Heatmaps to Find Patterns in Flight Passengers' Data use a heatmap to find patterns in flight passenger data. The goal of this activity is to apply your knowledge about color palettes to choose a suitable color palette for this data.

Use pandas to read the dataset flight_details.csv located in the Data folder. The given dataset contains the monthly figures for flight passengers from the years 1949 to 1960. use the pivot() function to transform the data to a format which is suitable for heatmaps.

Use a heatmap to visualize the given data. The given dataset contains the monthly figures for flight passengers for multiple years. Use your own appropriate color map. *9Make sure that the lowest value is the brightest and the highest the darkest color.*

Sample Output

In [4]:
```
flight_data=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/flight_details.csv")

flight_data_pivot=flight_data.pivot(index="Months",columns="Years",values="Passengers")

flight_data_pivot=flight_data_pivot.loc[['January', 'February','March','April', 'May','June', 'July',  'August', 'September', 'October', 'November','December']]

plt.figure(figsize=(6,4))

sns.heatmap(flight_data_pivot,cmap="Blues")

plt.title("Flight Passengers from 1949 to 1960")
plt.xticks(rotation=90)
plt.show()
```



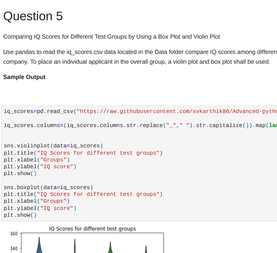
Question 4

Movie Comparison use a bar plot to compare movie scores. You are given five movies with scores from Rotten Tomatoes. The Tomatometer is the percentage of approved Tomatometer critics who have given a positive review for the movie. The Audience Score is the percentage of users who have given a score of 3.5 or higher out of 5. Compare these two scores among the five movies.

Use pandas to read the movie_scores.csv data located in the Data folder and transform the data into a useable format for Seaborn's barplot function

Sample Output

In [5]:
```
movie_scores=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/movie_scores.csv")

movie_scores=movie_scores.melt(id_vars="MovieTitle",value_vars=["Tomatometer","AudienceScore"],var_name="Type")

plt.figure(figsize=(10,5))
sns.barplot(data=movie_scores,x="MovieTitle",y="value",hue="Type",hue_order=["AudienceScore","Tomatometer"])
plt.xlabel("Movies")
plt.ylabel("Scores")
plt.title("Movies Scores Comparision")
plt.xticks(rotation=15)
plt.show()
```



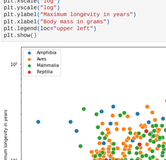
Question 5

Comparing IQ Scores for Different Test Groups by Using a Box Plot and Violin Plot

Use pandas to read the iq_scores.csv data located in the Data folder compare IQ scores among different test groups by using the violin plot that's provided by Seaborn's library. 100 people have come for an interview in a company. To place an individual applicant in the overall group, a violin plot and box plot shall be used.

Sample Output

In [6]:
```
iq_scores=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/iq_scores.csv")

iq_scores.columns=(iq_scores.columns.str.replace("_"," ").str.capitalize()).map(lambda x:" ".join([x.split(" ")[0],x.split(" ")[1].upper()]))


sns.violinplot(data=iq_scores)
plt.title("IQ Scores for different test groups")
plt.xlabel("Groups")
plt.ylabel("IQ score")
plt.show()

sns.boxplot(data=iq_scores)
plt.title("IQ Scores for different test groups")
plt.xlabel("Groups")
plt.ylabel("IQ score")
plt.show()
```



Question 6

Using a Scatter Plot to Visualize Correlation between Various Animals uses a scatter plot to show correlation within a dataset. Let's look at the following scenario: You are given a dataset containing information about various animals. Visualize the correlation between animal attributes.

The given dataset is not complete. Filter the data so you end up with samples containing a body mass and a maximum longevity. Sort the data according to the animal class.

Create a scatter plot visualizing the correlation between the body mass and the maximum longevity. Use different colors for grouping data samples according to their class. Add a legend, labels and a title. Use a log scale for both the x-axis and y-axis.

Sample Output

In [7]:
```
animal_data=pd.read_csv("https://raw.githubusercontent.com/svkarthik86/Advanced-python/main/WEEK-9%20Assignment/Data/anage_data.csv")

animal_data_new=animal_data[["Class","Maximum longevity (yrs)","Body mass (g)"]]

animal_filter_data=animal_data_new.query("(Class=='Amphibia')or(Class=='Aves')or(Class=='Reptilia')or(Class=='Mammalia')")

plt.figure(figsize=(16,8))
sns.scatterplot(data=animal_filter_data,x="Body mass (g)",y="Maximum longevity (yrs)",hue="Class",s=100)
plt.xscale("log")
plt.yscale("log")
plt.ylabel("Maximum longevity in years")
plt.xlabel("Body mass in grams")
plt.legend(loc="upper left")
plt.show()
```


In [ ]:
```

```

In [ ]:
```

```