1. Numpy:   
    Using NumPy create random vector of size 15 having only Integers in the range 1-20.   
   

First of all, imported numpy as np, with np from numbers 1 to 20 randomly printed with size of 15 as creating random vector it is written as np.random.randint(1,20,15) and assigned it x variable printed it.

1. Reshape the array to 3 by 5

2. Print array shape.

A picture containing table

Description automatically generated

In the above code, the reshape of the array is done to 3 by 5 to existing array ‘x’ as x.reshape(3,5) and assigned it ‘y’ variable and y is printed the output will be as shown .

[[10 7 16 6 17]

[11 12 18 11 13]

[ 1 10 7 3 12]]

Then array shape is printed as print(y.shape), the output is (3,5)

3. Replace the max in each row by 0

Graphical user interface, application

Description automatically generated  
The y array, replacing the max in each row by 0. The maximum in the each row as follows:  
In the 1st row it is 17 , in the 2nd row it is 18, in the 3rd row it is 12 are replaced by 0.  
By using For loop np.where(y == [[i] for i in np.amax(y, axis = 1], 0, y) and printed it.   
  
The output is [[10 7 16 6 0]

[11 12 0 11 13]

[ 1 10 7 3 0]]

2.Pandas

1. Read the provided CSV file ‘data.csv’. <https://drive.google.com/drive/folders/1h8C3mLsso-R-sIOLsvoYwPLzy2fJ4IOF?usp=sharing>

2. Show the basic statistical description about the data.   
3. Check if the data has null values. a. Replace the null values with the mean   
4. Select at least two columns and aggregate the data using: min, max, count, mean.   
5. Filter the dataframe to select the rows with calories values between 500 and 1000.   
6. Filter the dataframe to select the rows with calories values > 500 and pulse < 100.   
7. Create a new “df\_modified” dataframe that contains all the columns from df except for “Maxpulse”. 8. Delete the “Maxpulse” column from the main df dataframe   
9. Convert the datatype of Calories column to int datatype.   
10. Using pandas create a scatter plot for the two columns (Duration and Calories).

Table

Description automatically generated with medium confidence

Chart

Description automatically generated

Imported pandas as pd. A document in data.csv format is given in the link. After downloading the data.csv file, it is uploaded in Jupiter and then run the command as pd.read\_csv("data.csv") and printed it and statistical description is printed as print(df.describe()) and printed info as print(df.info()) and checked if the data has null values as print(df.isnull().values.any(). to Replace the null values with the mean, command   
print(df.isnull().sum())

df['Calories'].fillna(value=df['Calories'].mean(), inplace=True)

print(df)

And Selected at least two columns and aggregate the data using: min, max, count, mean i.e pulse and Calories as print(df.isnull().values.any())

dff = df[["Pulse","Calories"]].min()

df.groupby(["Duration","Pulse"]).agg(['min', 'max', 'count', 'mean'])

Filter the dataframe to select the rows with calories values between 500 and 1000 and Filter the dataframe to select the rows with calories values > 500 and pulse < 100 as

df[((df["Calories"]>500) & (df["Calories"]<1000))]

df[((df["Calories"]>500) | (df["Calories"]<100))]

and Creating a new “df\_modified” dataframe that contains all the columns from df except for “Maxpulse” as df\_modified = df.loc[:, df.columns != 'Maxpulse']

df\_modified  
and deleting the “Maxpulse” column from the main df dataframe as df.drop(["Maxpulse"],axis=1,inplace=True)

and Converting the datatype of Calories column to int datatype as   
df["Calories"] = df["Calories"].astype(int)  
  
and by using pandas create a scatter plot for the two columns (Duration and Calories) as df.info()

df.plot.scatter(x="Duration",y="Calories")

3. Matplotlib  
 1. Write a Python programming to create a below chart of the popularity of programming Languages.  
 2. Sample data:   
Programming languages: Java, Python, PHP, JavaScript, C#, C++ Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

Chart

Description automatically generated

Here, in the above problem matplotlib is imported as plt , To create a chart of the popularity , the data is plotted by using languages 'Java', 'Python', 'PHP', 'JavaScript', 'C#', 'C++' and popularity as 22.2, 17.6, 8.8, 8, 7.7, 6.7.   
Colour in chart is shown as ["#1f77b4", "#ff7f0e", "#2ca02c", "#d62728", "#9467bd", "#8c564b"]. # exploding the 1st slice by using command as explode = (0.1, 0, 0, 0,0,0) and # Plot

plt.pie(popuratity, explode=explode, labels=languages, colors=colors, autopct='%1.1f%%', shadow=True, startangle=140) , plt.axis('equal') ,plt.show()

Video link:  
<https://youtu.be/WrsnqQN68wo>