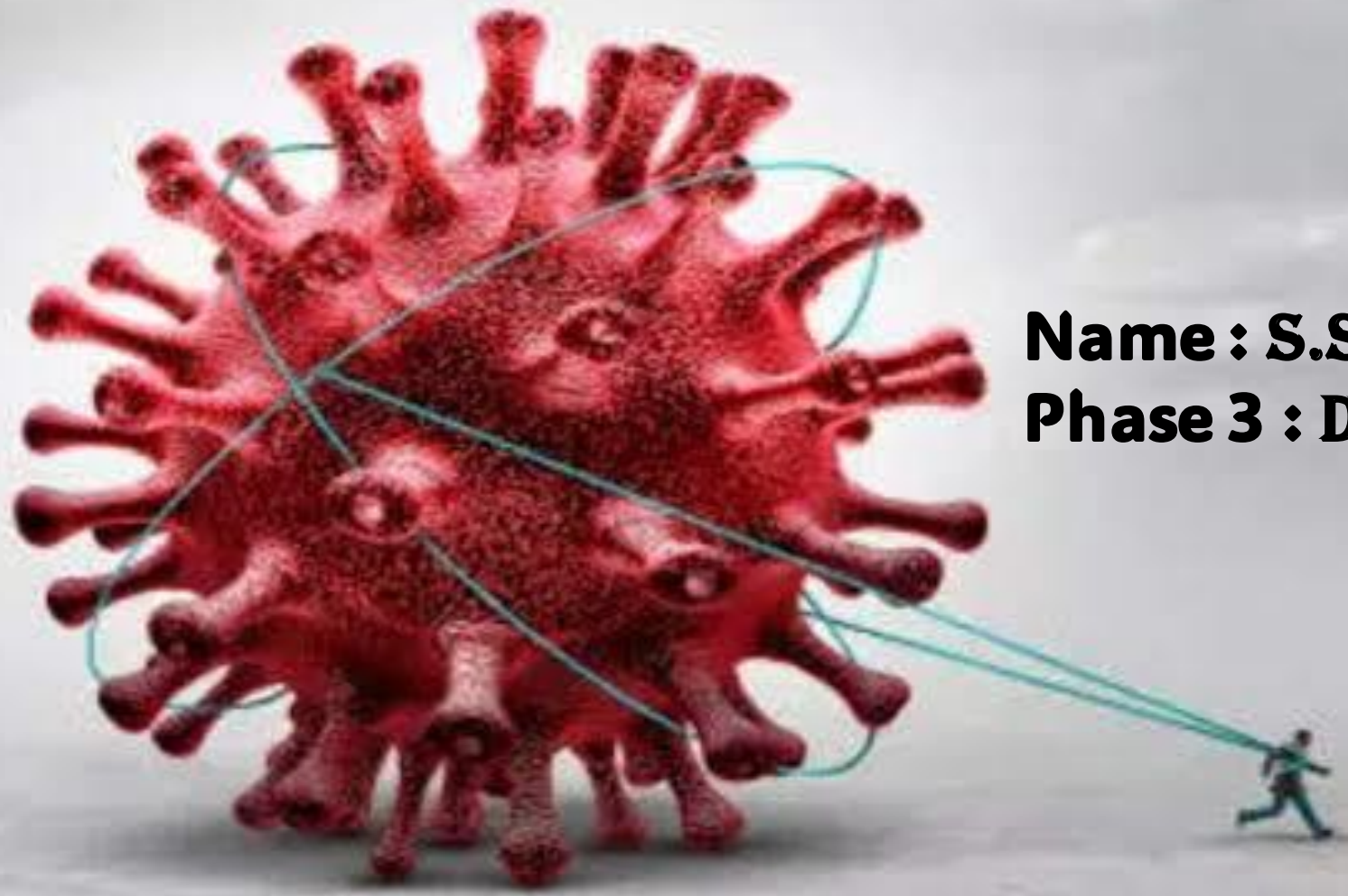
COVID-19 case analysis



Name : S.SUVENTHAN

Phase 3 : DEVELOPMENT PART 1

COVID-19 case analysis

INTRODUCTION :

Covid can be very contagious and spreads quickly. Over one million people have died from COVID-19 in the United States. COVID-19 most often causes respiratory symptoms that can feel much like a cold, the flu, or pneumonia. COVID-19 may attack more than your lungs and respiratory system.

Given Dataset :

Out[5]:

	dateRep	day	month	year	cases	deaths	countriesAndTerritories
0	31-05-2021	31	5	2021	366	5	Austria
1	30-05-2021	30	5	2021	570	6	Austria
2	29-05-2021	29	5	2021	538	11	Austria
3	28-05-2021	28	5	2021	639	4	Austria
4	27-05-2021	27	5	2021	405	19	Austria
...	...	...	...	...	...	...	...
2725	06-03-2021	6	3	2021	3455	17	Sweden
2726	05-03-2021	5	3	2021	4069	12	Sweden
2727	04-03-2021	4	3	2021	4884	14	Sweden
2728	03-03-2021	3	3	2021	4876	19	Sweden
2729	02-03-2021	2	3	2021	6191	19	Sweden

2730 rows × 7 columns

Steps to follow :

Import libraries

Start by importing necessary libraries

```
[4]: import pandas as pd
     import numpy as np
     from sklearn.model_selection import train_test_split
     from sklearn.preprocessing import StandardScaler
```

Load the dataset :

Load your dataset into Pandas data frame.You can tupicaly find covid-19 case analysis datasets in CSV format,but you can adapt this code to other formats as needed.

```
In [5]: ds=pd.read_csv("D:\\New folder (2)\\covid_19_cases4.csv")
        ds
```

Exploratory Data Analysis

This include checking for missing value, exploring the data's statics, and visualizing it to identify patterns.

```
ds.isnull().sum()
```

```
ds.describe()
```

Handling of missing values

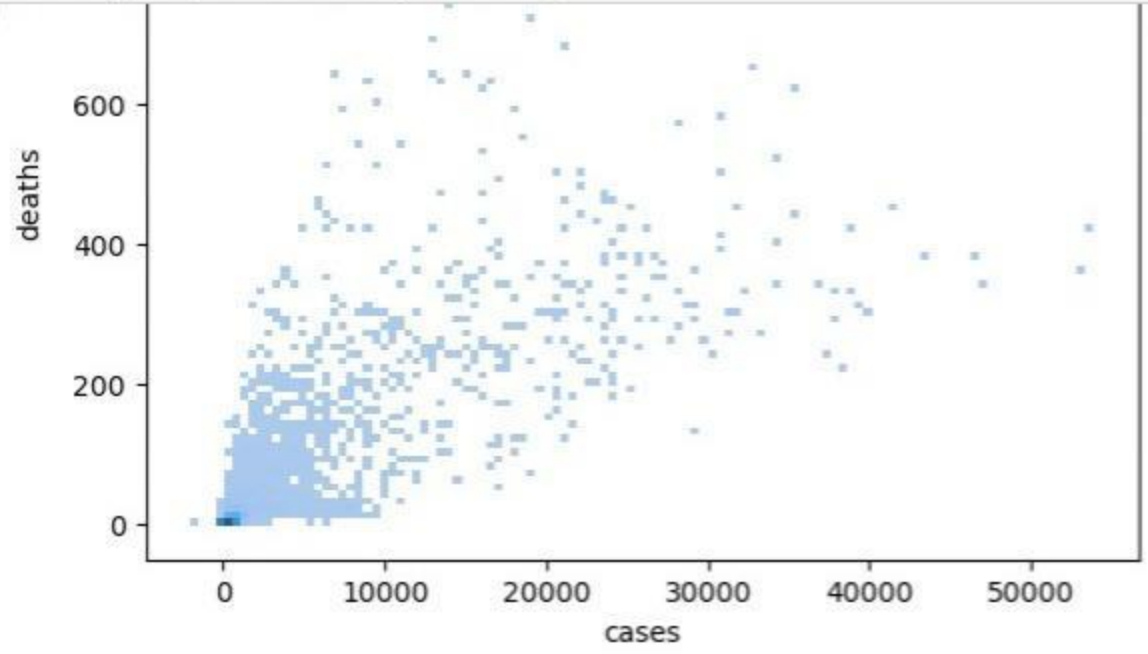
Handling missing values is the deletion of the rows or columns having null values. If any columns have more than half of the values as null then you can drop the entire column. Rows can also be dropped if having one or more columns values as null.

```
ds.isnull().sum()
```

```
dateRep                    0
day                        0
month                      0
year                       0
cases                      0
deaths                     0
countriesAndTerritories    0
dtype: int64
```

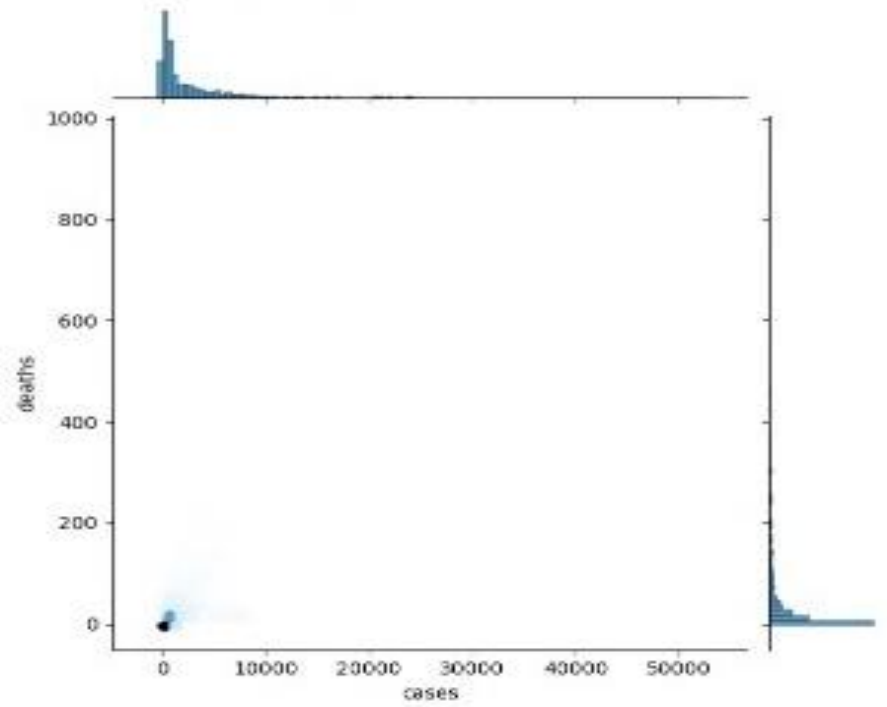
Visualization & preprocessing of data

```
In [15]: sns.histplot(ds,x='cases',y='deaths')
```



In [21]:
```
sns.jointplot(ds,x='cases',y='deaths',kind='hex')
```

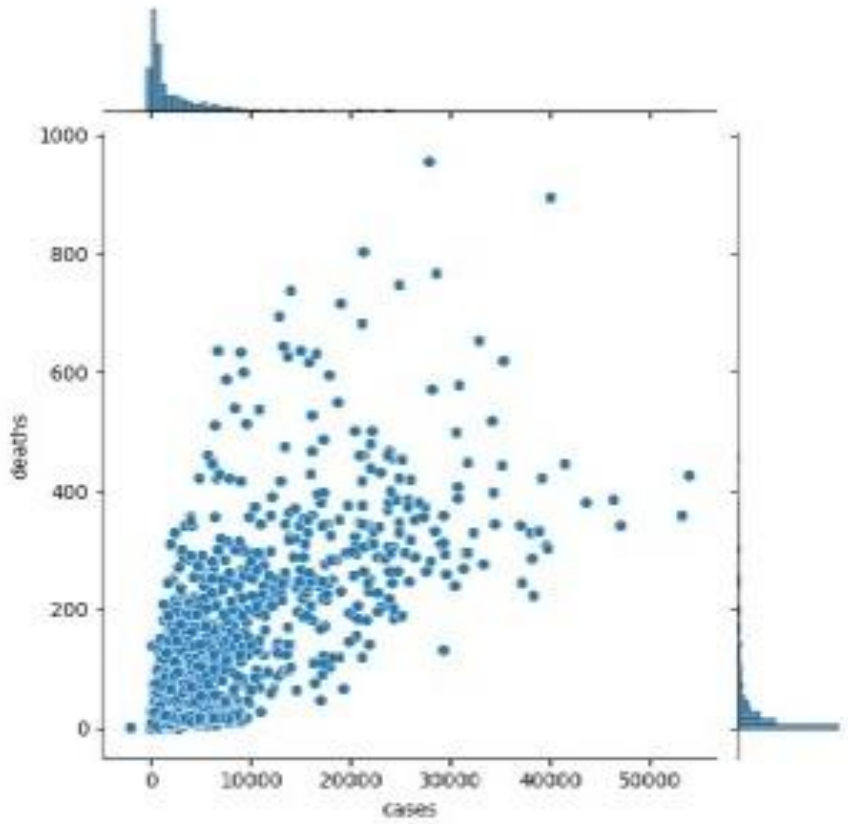
Out[21]:
```
<seaborn.axisgrid.JointGrid at 0x25521ded698>
```



In [22]:
```
sns.jointplot(ds,x='cases',y='deaths')
```

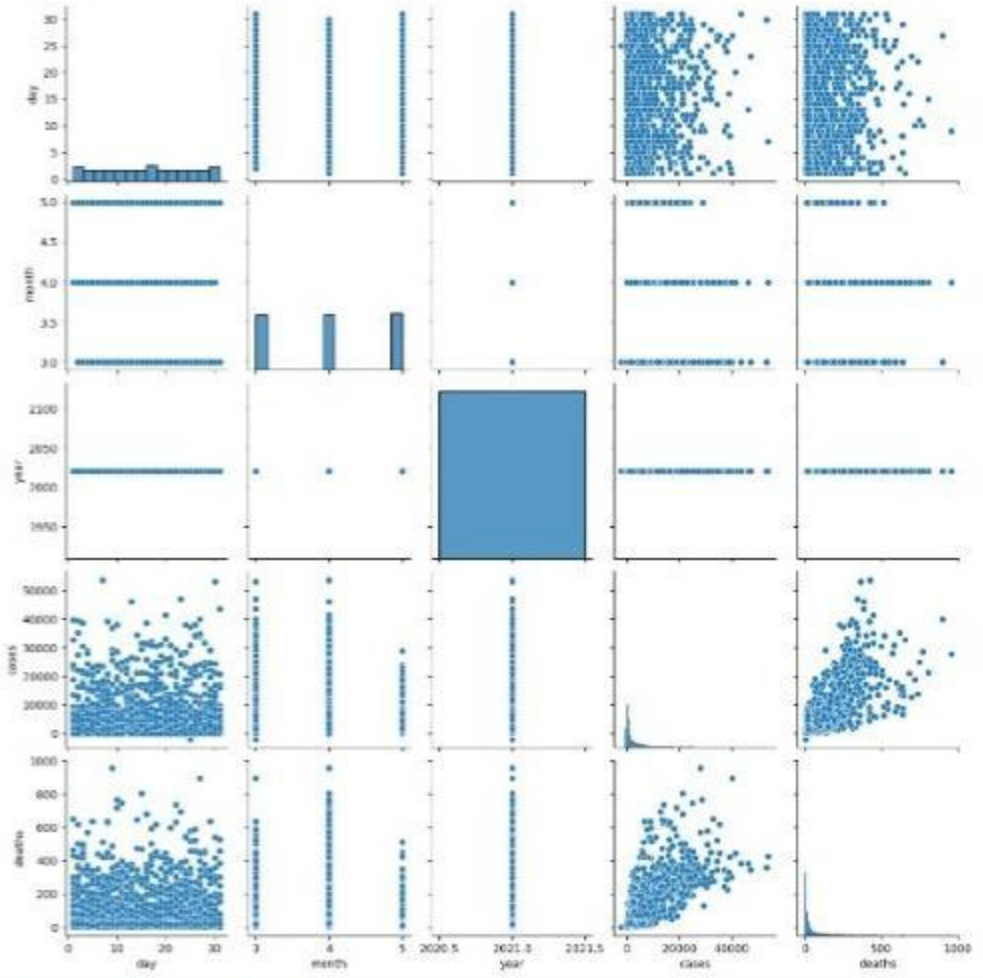
In [22]:
```
sns.jointplot(ds,x='cases',y='deaths')
```

Out[22]:
```
<seaborn.axisgrid.JointGrid at 0x25522218750>
```



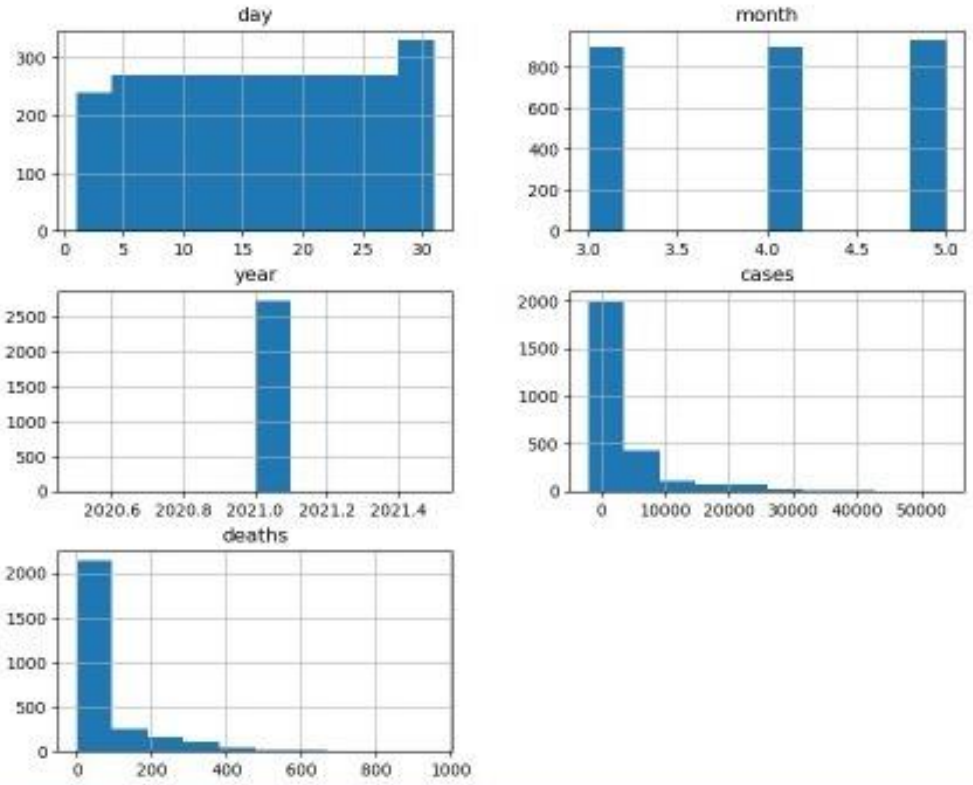
```
sns.pairplot(dataset)
```

Out[29]: <seaborn.axisgrid.PairGrid at 0x25522c03790>



In [38]:
```
ds.hist(figsize=(10,8))
```

Out[38]:
```
array([[<Axes: title={'center': 'day'}>,
        <Axes: title={'center': 'month'}>],
       [<Axes: title={'center': 'year'}>,
        <Axes: title={'center': 'cases'}>],
       [<Axes: title={'center': 'deaths'}>, <Axes: >]], dtype=object)
```



```
In [31]: ds.corr(numeric_only=True)
```

Out[31]:

	day	month	year	cases	deaths
day	1.000000	-0.022973	NaN	-0.026988	-0.038128
month	-0.022973	1.000000	NaN	-0.172412	-0.126515
year	NaN	NaN	NaN	NaN	NaN
cases	-0.026988	-0.172412	NaN	1.000000	0.766309
deaths	-0.038128	-0.126515	NaN	0.766309	1.000000

Conclusion

The COVID-19 pandemic has had a significant impact on the world. It has highlighted the importance of preparedness, collaboration, and innovation. While the future outlook is uncertain, we have learned many lessons that will help us better prepare for future pandemics.