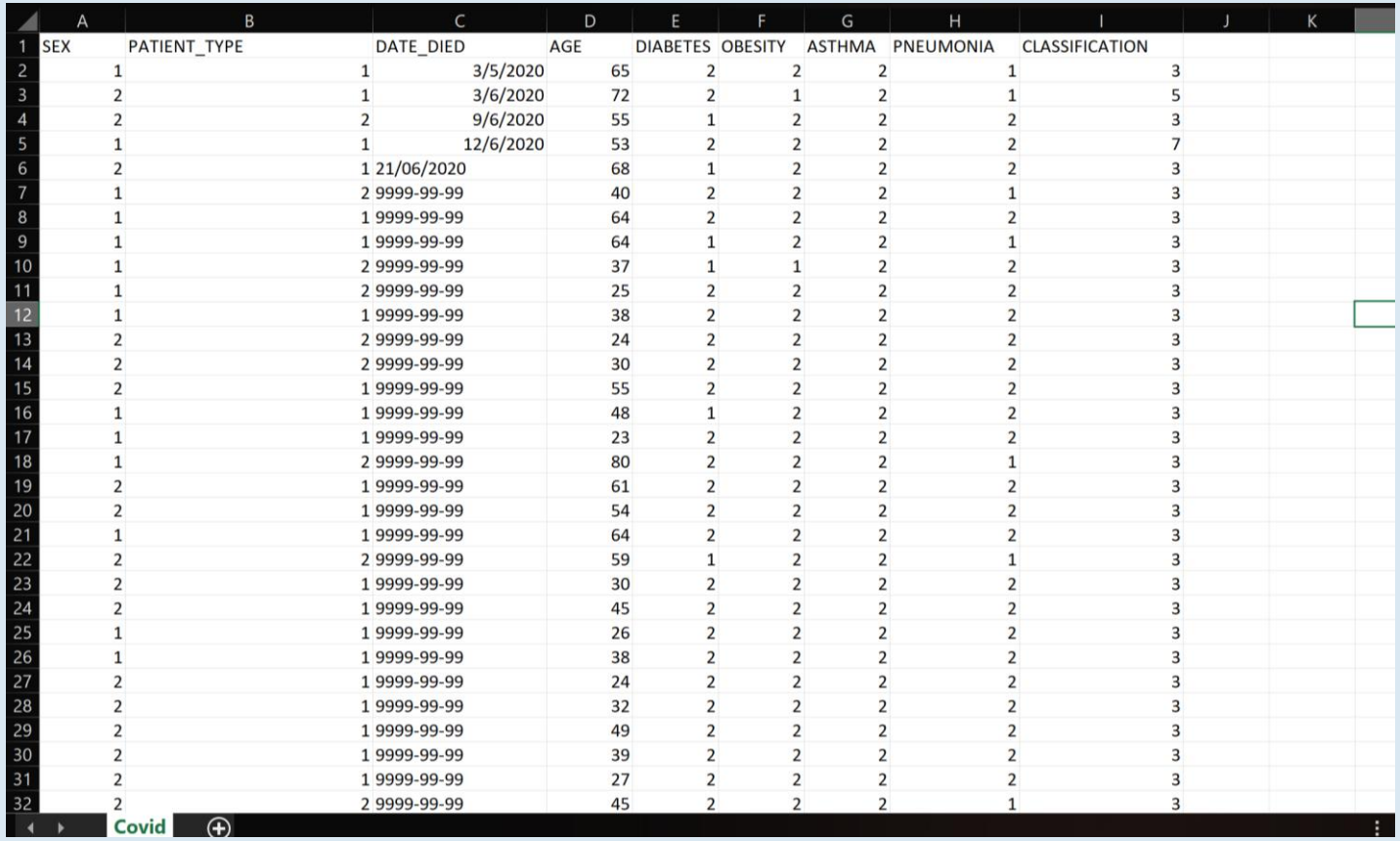
Characteristics of Covid-19 Patients

Ellie Goodman

Purpose

- Project the overall likelihood of death by Covid-19 and predict survival rates based on patient characteristics like sex, age etc.

Our Data



	A	B	C	D	E	F	G	H	I
1	SEX	PATIENT_TYPE	DATE_DIED	AGE	DIABETES	OBESITY	ASTHMA	PNEUMONIA	CLASSIFICATION
2	1	1	3/5/2020	65	2	2	2	1	3
3	2	1	3/6/2020	72	2	1	2	1	5
4	2	2	9/6/2020	55	1	2	2	2	3
5	1	1	12/6/2020	53	2	2	2	2	7
6	2	1	21/06/2020	68	1	2	2	2	3
7	1	2	9999-99-99	40	2	2	2	1	3
8	1	1	9999-99-99	64	2	2	2	2	3
9	1	1	9999-99-99	64	1	2	2	1	3
10	1	2	9999-99-99	37	1	1	2	2	3
11	1	2	9999-99-99	25	2	2	2	2	3
12	1	1	9999-99-99	38	2	2	2	2	3
13	2	2	9999-99-99	24	2	2	2	2	3
14	2	2	9999-99-99	30	2	2	2	2	3
15	2	1	9999-99-99	55	2	2	2	2	3
16	1	1	9999-99-99	48	1	2	2	2	3
17	1	1	9999-99-99	23	2	2	2	2	3
18	1	2	9999-99-99	80	2	2	2	1	3
19	2	1	9999-99-99	61	2	2	2	2	3
20	2	1	9999-99-99	54	2	2	2	2	3
21	1	1	9999-99-99	64	2	2	2	2	3
22	2	2	9999-99-99	59	1	2	2	1	3
23	2	1	9999-99-99	30	2	2	2	2	3
24	2	1	9999-99-99	45	2	2	2	2	3
25	1	1	9999-99-99	26	2	2	2	2	3
26	1	1	9999-99-99	38	2	2	2	2	3
27	2	1	9999-99-99	24	2	2	2	2	3
28	2	1	9999-99-99	32	2	2	2	2	3
29	2	1	9999-99-99	49	2	2	2	2	3
30	2	1	9999-99-99	39	2	2	2	2	3
31	2	1	9999-99-99	27	2	2	2	2	3
32	2	2	9999-99-99	45	2	2	2	1	3

Covid

- Data set from Kaggle of 1,048,575 patient deaths in Mexico of 2020
- Including patient characteristics
  - Sex, Patient Type, Age, Diabetes, Obesity, Asthma, Pneumonia and Classification

Plan

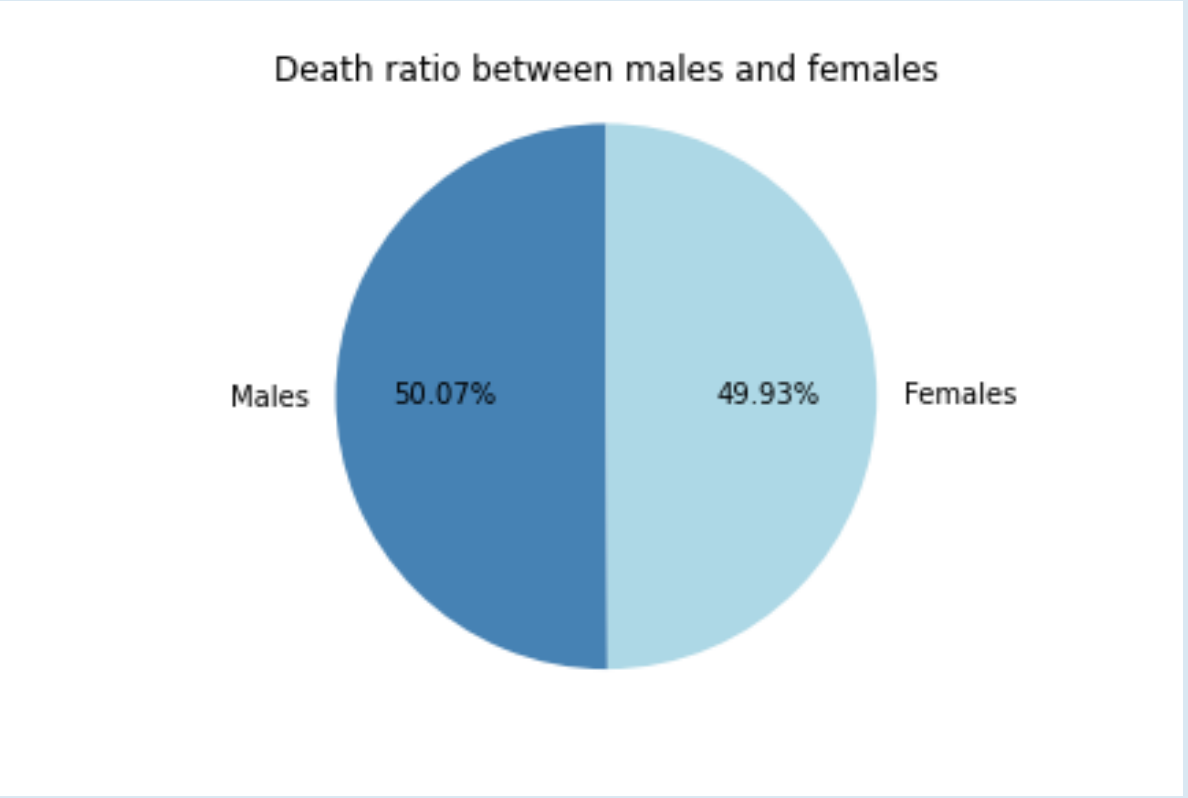
- Create 8 functions using the 8 characteristics
  - •Functions Created: getSex, getAge, getPatient, getDiabetic, getObese, getAsthma, getPneumonia, getClassification
- Import raw data and calculate the totals of each section and find the percentage and likelihood of death based off each characteristic
- Use those calculations to also either create a bar chart or pie chart to visualize results
- Imported 3 libraries
  - •Pandas, matplotlib, numpy

Sex

- In our data sex was identified by male= 1 and female= 2
- There was no relationship found between sex and the death ratio among patients

Code Output

```
-Sex % is used to distinguish covid deaths betweeen males and females.
Male deaths: 50.07%
Female deaths: 49.93%
```

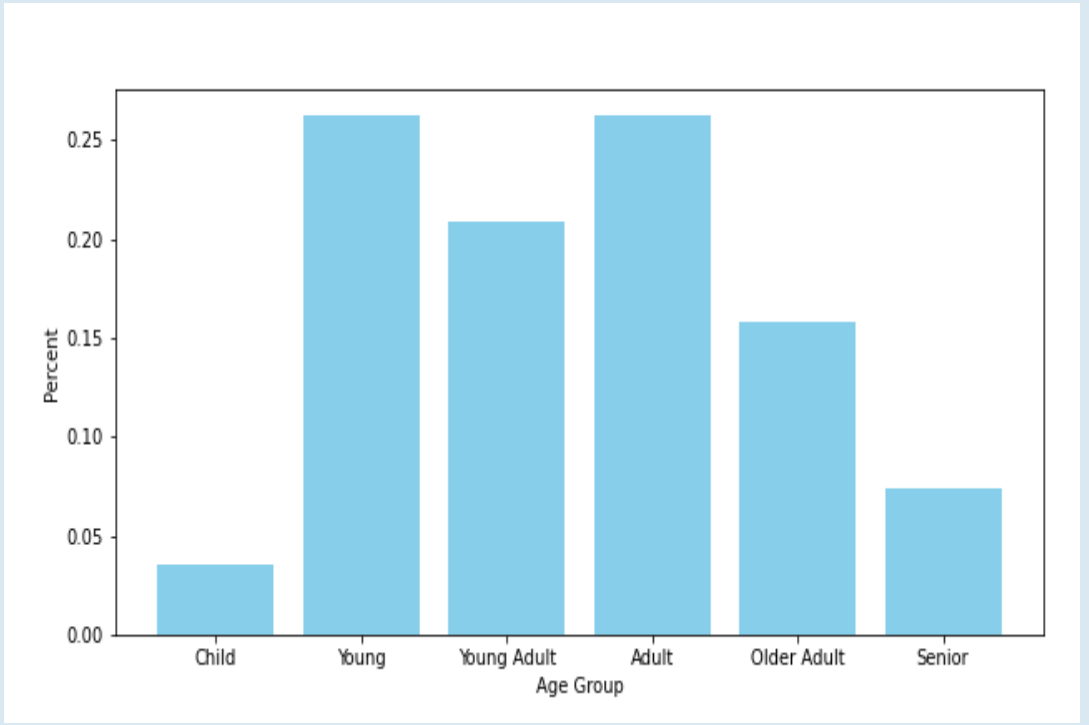


Age

- In our data we were given a range of patients from the ages of 0-100
- We broke up those ages in age groups to help identify which ages led to the most deaths
- We found that the age groups of the young, young adult and adult had the highest % of death because we believe that they had more contact with people that were infected with Covid-19

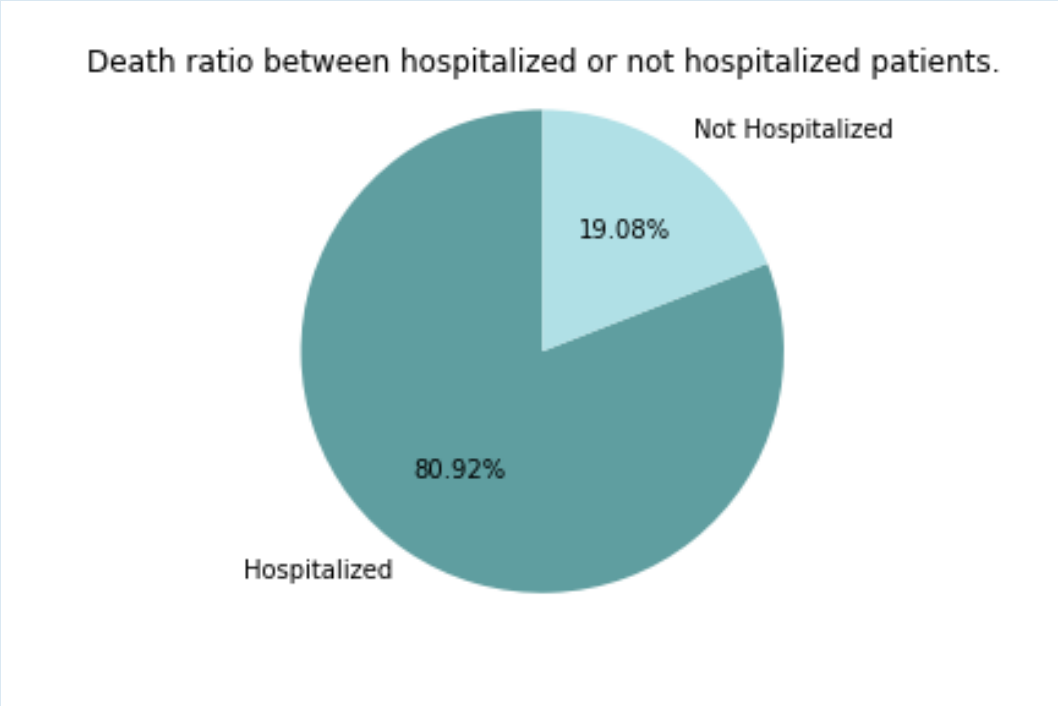
Code Output

```
-Age % is used to distinguish covid deaths between different ages.
Age Group
   0-15: Child
   16-25: Young
   26-35: Young Adult
   36-50: Adult
   51-65: Older Adult
   65-100: Senior
Age group Child has this % of death: 3.54%
Age group Young has this % of death: 26.22%
Age group Young Adult has this % of death: 20.84%
Age group Adult has this % of death: 26.22%
Age group Older Adult has this % of death: 15.79%
Age group Senior has this % of death: 7.39%
```



Hospitalized

- In our data we identified patients who were hospitalized = 1 and patients not hospitalized = 2
- We found that patients who were admitted into a hospital had a strong relationship to dying from covid

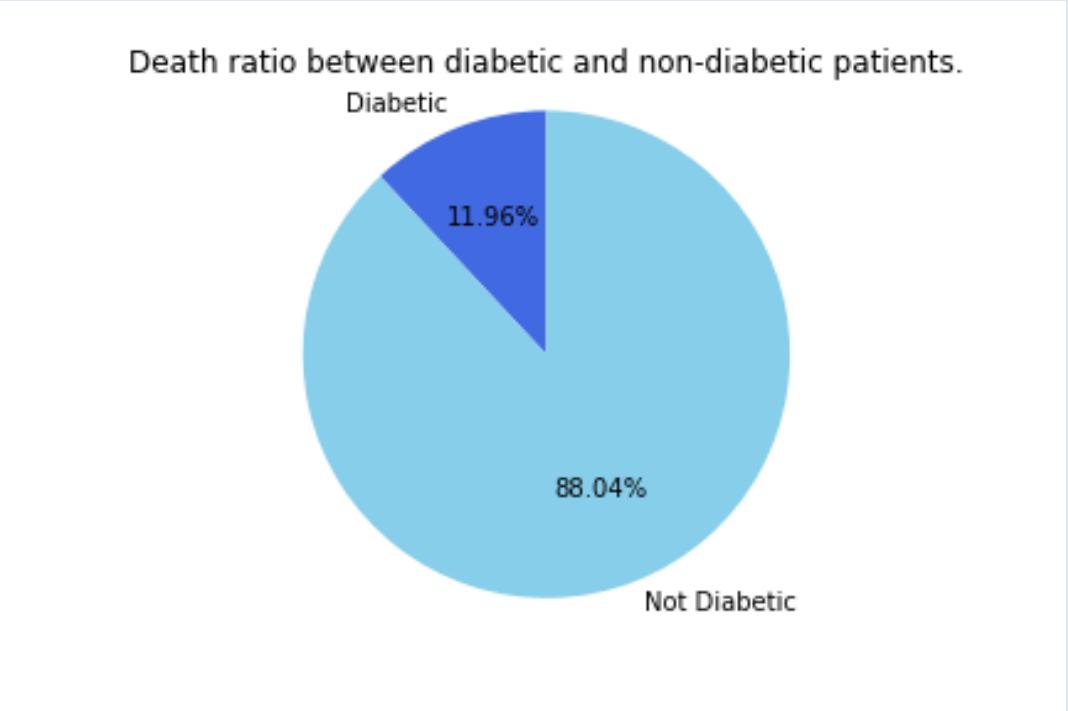


Code Output

```
-Hospitalized % is used to distinguish covid deaths between patients hospitalized or not.
Hospitalized patients: 80.92%
Not hospitalized patients: 19.08%
```

Diabetes

- In our data we identified patients who were diabetic = 1 and patients who were not =2
- We found that there was not a relationship between patients with diabetes and death from Covid

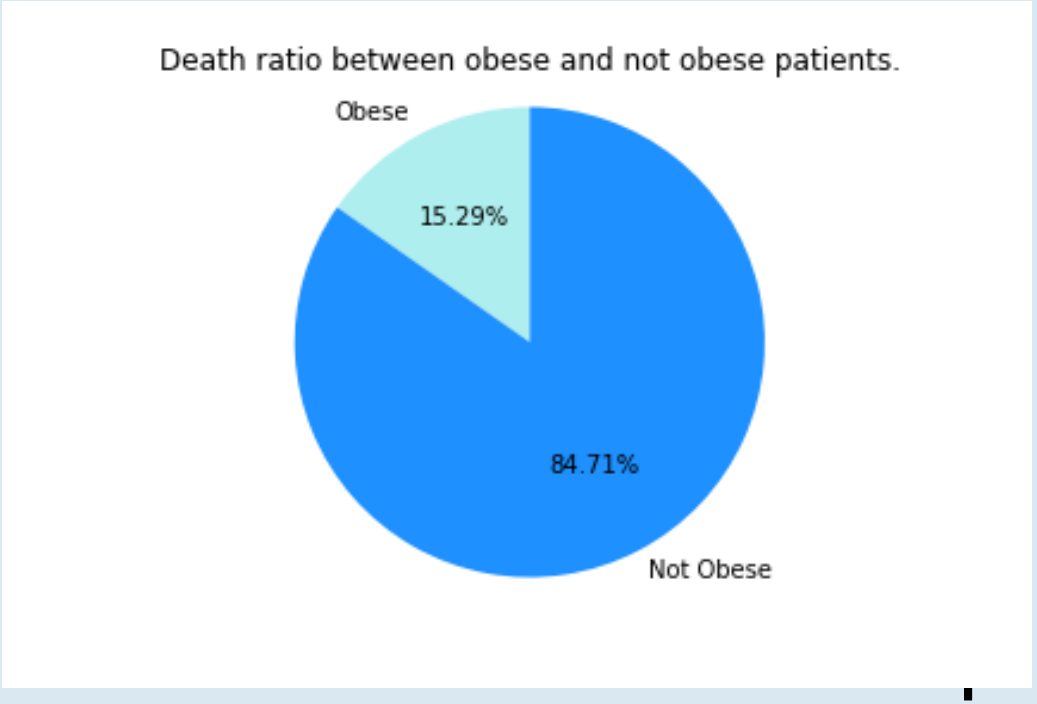


Code Output

```
-Diabetic % is used to distinguish covid deaths between patients with diabetes or not.
Diabetic patients: 11.96%
Non-diabetic patients: 88.04%
```

Obesity

- In our data obesity was identified by patients who were identified as obese = 1 and those who were not = 2
- Throughout the data, we found that there was not a significant relationship between patients who died from covid and being obese although it was little bit higher than other characteristics

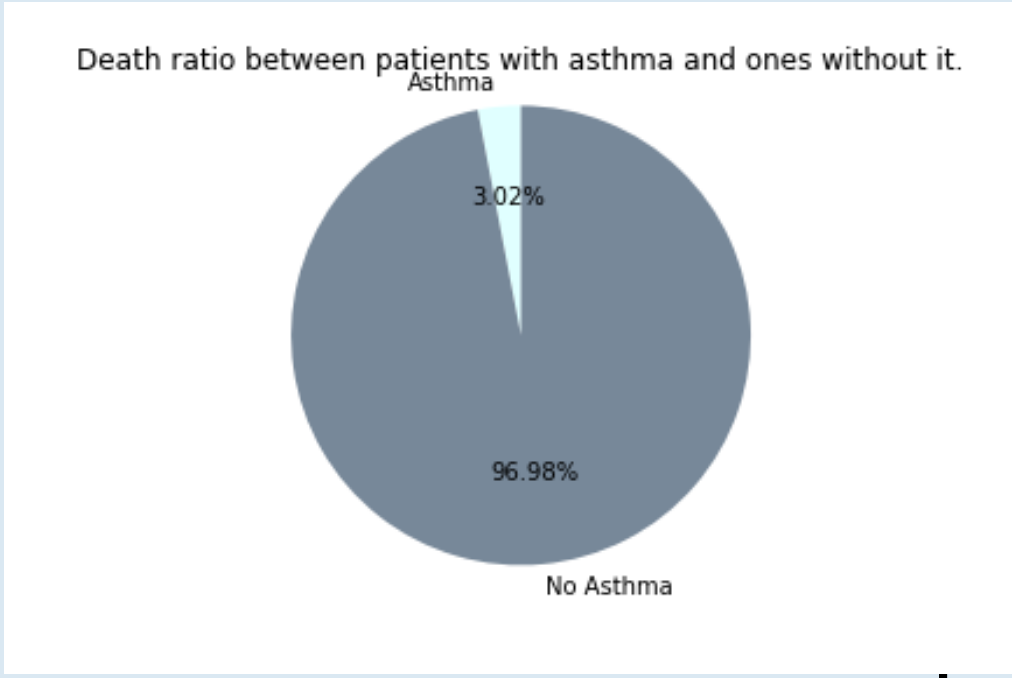


Code Output

```
-Obesity % is used to distinguish covid deaths between patients who were obese and those who were not.
Obese patients: 15.29%
Non-obese patients: 84.71%
```

Asthma

- In Our data asthma was identified by patients with asthma = 1 and patients with no asthma = 2
- There was no relationship found between covid deaths and having asthma

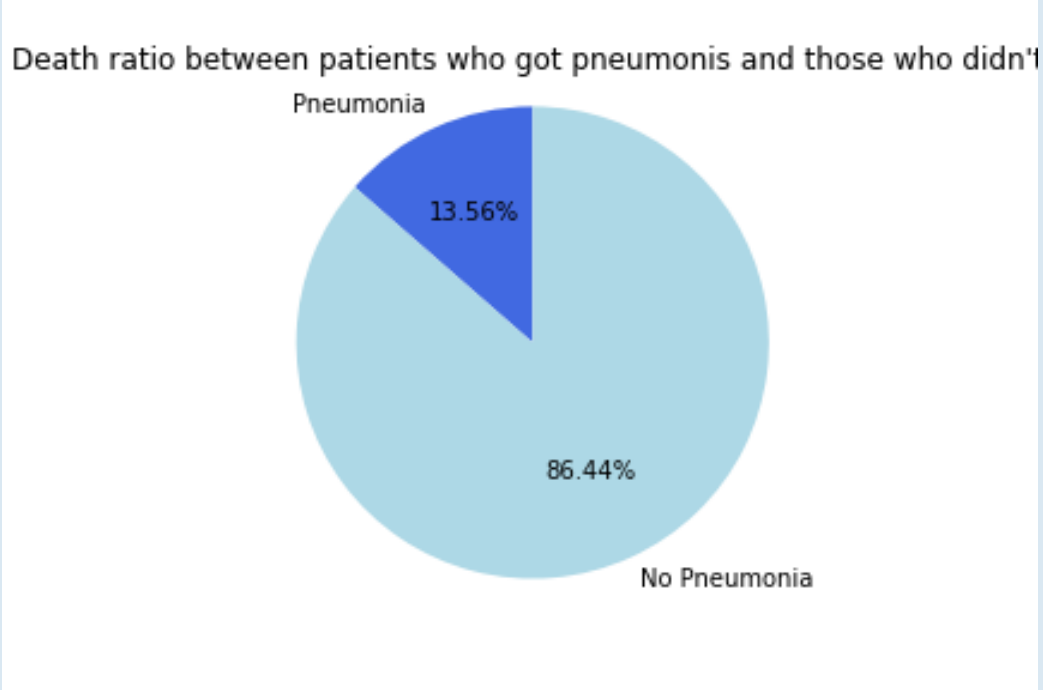


Code Output

```
-Asthma % is used to distinguish covid deaths to patient with asthma or without it.
Asthma patients: 3.02%
Without Asthma: 96.98%
```

Pneumonia

- In Our data pneumonia was identified by patients who got pneumonia= 1 and patients who did not get pneumonia= 2
- There was no relationship found between covid deaths and getting pneumonia when having covid



Code Output

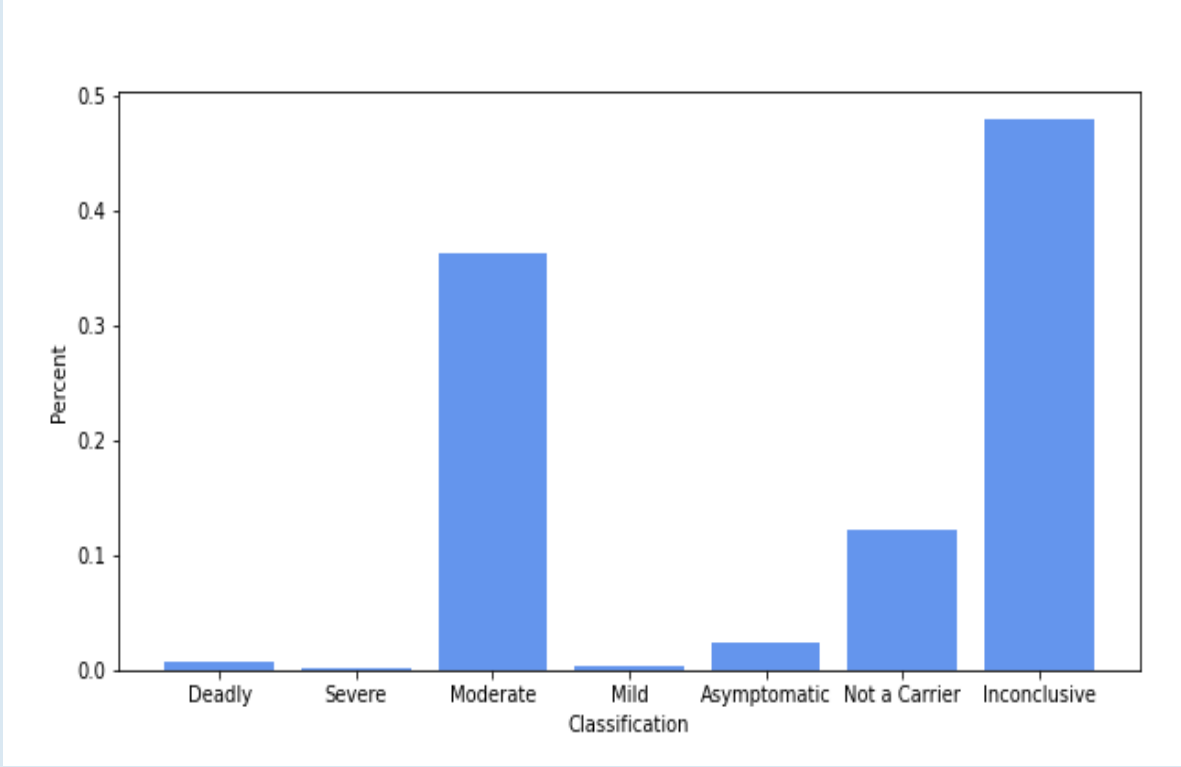
```
-Pneumonia % is used to distinguish covid deaths between patients who got pneumonia and those who did not.
Patients who got pneumonia: 13.56%
Patients who did no get pneumonia: 86.44%
```

Classification

- In Our data we were given Patient Classifications that ranged from 1-7
- We found that the patient classification that was inclusive had the highest death ratio of 48%
- We believe this is because there were not as accurate Covid- 19 tests has there are today

Code Output

```
-There are 7 types of classifications for patients this was the ratio in which each classification led to death.
Classfication Type:
   1-Deadly
   2-Severe
   3-Moderate
   4-Mild
   5-Asymptomatic
   6-Not a Carrier
   7-Inconclusive
Classification 1 had this % of death: 0.82%
Classification 2 had this % of death: 0.18%
Classification 3 had this % of death: 36.39%
Classification 4 had this % of death: 0.30%
Classification 5 had this % of death: 2.49%
Classification 6 had this % of death: 12.22%
Classification 7 had this % of death: 48.00%
```



Conclusion

Overall, we found 3 strong relationships with death from Covid-19 with characteristics such as Age, Hospitalization, Classification.

As well as a smaller relationship with patients who were obese

Works Cited

Nizri, Meir. “Covid-19 Dataset.” *Kaggle*, 13 Nov. 2022, https://www.kaggle.com/datasets/meirnizri/covid19-dataset?resource=download.