Pandas Descriptive Statistics Assignment

Aircraft wildlife strikes data | 1990 - 2015

In this exercise, we will extract and analyze aircraft wildlife strikes data, and we will determine the probability of each part of an aircraft getting damaged by an aircraft wildlife strike

In [ ]:
```
# Import the necessary libraries

import pandas as pd
import matplotlib.pyplot as plt
import requests
import io
```

In [ ]:
```
# Read our data from Google Drive

file_id = "1TAD7Uyc9PjByt_q13uvGXGeubXnujnUi"
url = f"https://drive.google.com/uc?id={file_id}"

# Download the contents of the CSV file
download = requests.get(url).content

# Read the CSV file into a Pandas DataFrame
df = pd.read_csv(io.StringIO(download.decode("utf-8")), low_memory=False)
```

In [ ]:
```
# Explore the data

df.head()
```

Out[ ]:

	Record ID	Incident Year	Incident Month	Incident Day	Operator ID	Operator	Aircraft	Aircraft Type	Aircraf Mak
0	127128	1990	1	1	DAL	DELTA AIR LINES	B-757-200	A	14
1	129779	1990	1	1	HAL	HAWAIIAN AIR	DC-9	A	58
2	129780	1990	1	2	UNK	UNKNOWN	UNKNOWN	NaN	Nal
3	2258	1990	1	3	MIL	MILITARY	A-10A	A	34
4	2257	1990	1	3	MIL	MILITARY	F-16	A	56

5 rows × 66 columns

In [ ]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 174104 entries, 0 to 174103
Data columns (total 66 columns):
 #   Column                Non-Null Count   Dtype  
---  ------                --------------   -----  
 0   Record ID             174104 non-null  int64  
 1   Incident Year         174104 non-null  int64  
 2   Incident Month        174104 non-null  int64  
 3   Incident Day          174104 non-null  int64  
 4   Operator ID           174104 non-null  object 
 5   Operator              174104 non-null  object 
 6   Aircraft              174104 non-null  object 
 7   Aircraft Type         133074 non-null  object 
 8   Aircraft Make         131051 non-null  object 
 9   Aircraft Model        122439 non-null  object 
 10  Aircraft Mass         127320 non-null  float64
 11  Engine Make           123434 non-null  float64
 12  Engine Model          121988 non-null  object 
 13  Engines               127342 non-null  float64
 14  Engine Type           127282 non-null  object 
 15  Engine1 Position      126193 non-null  object 
 16  Engine2 Position      118715 non-null  float64
 17  Engine3 Position      11659 non-null   object 
 18  Engine4 Position      3092 non-null    float64
 19  Airport ID            174104 non-null  object 
 20  Airport               173814 non-null  object 
 21  State                 152128 non-null  object 
 22  FAA Region            155202 non-null  object 
 23  Warning Issued        76418 non-null   object 
 24  Flight Phase          118802 non-null  object 
 25  Visibility            109933 non-null  object 
 26  Precipitation         88322 non-null   object 
 27  Height                103677 non-null  float64
 28  Speed                 71258 non-null   float64
 29  Distance              99713 non-null   float64
 30  Species ID            174104 non-null  object 
 31  Species Name          174024 non-null  object 
 32  Species Quantity      169627 non-null  object 
 33  Flight Impact         99465 non-null   object 
 34  Fatalities            565 non-null     float64
 35  Injuries              229 non-null     float64
 36  Aircraft Damage       174104 non-null  int64  
 37  Radome Strike         174104 non-null  int64  
 38  Radome Damage         174104 non-null  int64  
 39  Windshield Strike     174104 non-null  int64  
 40  Windshield Damage     174104 non-null  int64  
 41  Nose Strike           174104 non-null  int64  
 42  Nose Damage           174104 non-null  int64  
 43  Engine1 Strike        174104 non-null  int64  
 44  Engine1 Damage        174104 non-null  int64  
 45  Engine2 Strike        174104 non-null  int64  
 46  Engine2 Damage        174104 non-null  int64  
 47  Engine3 Strike        174104 non-null  int64  
 48  Engine3 Damage        174104 non-null  int64  
 49  Engine4 Strike        174104 non-null  int64  
 50  Engine4 Damage        174104 non-null  int64
```

```
 51  Engine Ingested       174104 non-null  int64
 52  Propeller Strike      174104 non-null  int64
 53  Propeller Damage      174104 non-null  int64
 54  Wing or Rotor Strike  174104 non-null  int64
 55  Wing or Rotor Damage  174104 non-null  int64
 56  Fuselage Strike       174104 non-null  int64
 57  Fuselage Damage       174104 non-null  int64
 58  Landing Gear Strike   174104 non-null  int64
 59  Landing Gear Damage   174104 non-null  int64
 60  Tail Strike           174104 non-null  int64
 61  Tail Damage           174104 non-null  int64
 62  Lights Strike         174104 non-null  int64
 63  Lights Damage         174104 non-null  int64
 64  Other Strike          174104 non-null  int64
 65  Other Damage          174104 non-null  int64
dtypes: float64(10), int64(34), object(22)
memory usage: 87.7+ MB
```

In [ ]:
```
df.describe()
```

Out[ ]:

	Record ID	Incident Year	Incident Month	Incident Day	Aircraft Mass	Engin
count	174104.000000	174104.000000	174104.000000	174104.000000	127320.000000	123434.
mean	241204.036915	2006.036392	7.171840	15.712264	3.510611	21.
std	94013.682213	6.747708	2.790152	8.799405	0.873783	11.
min	1000.000000	1990.000000	1.000000	1.000000	1.000000	1.
25%	205411.750000	2001.000000	5.000000	8.000000	3.000000	10.
50%	249102.500000	2007.000000	8.000000	16.000000	4.000000	22.
75%	322592.250000	2012.000000	9.000000	23.000000	4.000000	34.
max	367445.000000	2015.000000	12.000000	31.000000	5.000000	92.

8 rows × 44 columns

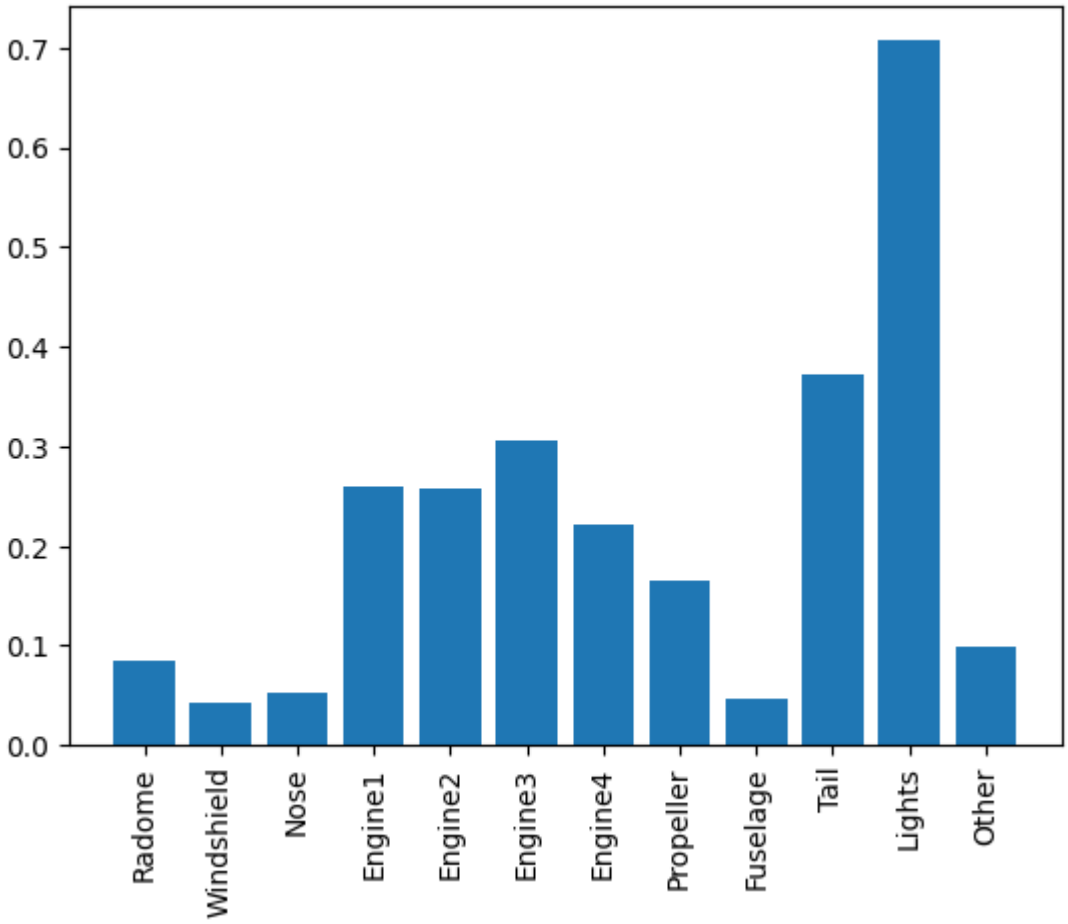
Now we are going to calculate the probability of each part of the flight getting damaged and plot these probabilities

In [ ]:
```
strikes = {}
for c in df.columns:
    column_name = c.split(" ")
    # print(len(col_sep), col_sep)
```

```
    if len(column_name) > 1 and column_name[1] == "Strike":
        strikes[column_name[0]] = df[column_name[0] + " Damage"].sum() / df[c].sum(
```

In [ ]:
```
# Calculate the probability of each part of the aircraft getting damaged and find t
plt.bar(strikes.keys(), strikes.values())
plt.xticks(rotation=90)
print(max(strikes, key=strikes.get))
```

```
Lights
```



In [ ]: