BS3

July 16, 2024

```python
[11]: import pandas as pd
      import numpy as np
```

```python
[12]: df = pd.read_excel("C:/Users/thala/Downloads/Bird Strikes data.xlsx")
```

```python
[13]: df
```

```
[13]:        Record ID Aircraft: Type                    Airport: Name Altitude bin  \
       0         202152       Airplane                   LAGUARDIA NY     > 1000 ft
       1         208159       Airplane   DALLAS/FORT WORTH INTL ARPT     < 1000 ft
       2         207601       Airplane              LAKEFRONT AIRPORT     < 1000 ft
       3         215953       Airplane           SEATTLE-TACOMA INTL     < 1000 ft
       4         219878       Airplane                  NORFOLK INTL     < 1000 ft
       ...          ...            ...                           ...           ...
       25553     321151       Airplane              REDDING MUNICIPAL     > 1000 ft
       25554     319677       Airplane                 ORLANDO INTL     < 1000 ft
       25555     319680            NaN                          NaN           NaN
       25556     319679       Airplane   DETROIT METRO WAYNE COUNTY ARPT   < 1000 ft
       25557     319593       Airplane   ABRAHAM LINCOLN CAPITAL ARPT     < 1000 ft

              Aircraft: Make/Model Wildlife: Number struck  \
       0                 B-737-400                Over 100
       1                     MD-80                Over 100
       2                     C-500                Over 100
       3                 B-737-400                Over 100
       4              CL-RJ100/200                Over 100
       ...                     ...                     ...
       25553               EMB-120                       1
       25554                 A-321                       1
       25555                EC-135                     NaN
       25556             B-757-200                       1
       25557             B-737-400                       1

              Wildlife: Number Struck Actual Effect: Impact to flight FlightDate  \
       0                          859           Engine Shut Down 2000-11-23
       1                          424                        NaN 2001-07-25
       2                          261                        NaN 2001-09-14
       3                          806      Precautionary Landing 2002-09-05
```

4	942	NaN	2003-06-23
…	…	…	…
25553	1	NaN	2011-12-30
25554	1	NaN	2011-12-30
25555	1	NaN	NaT
25556	1	NaN	2011-12-31
25557	1	NaN	2011-12-31

	Effect: Indicated Damage	…	Remains of wildlife sent to Smithsonian \
0	Caused damage	…	False
1	Caused damage	…	False
2	No damage	…	False
3	No damage	…	False
4	No damage	…	False
…	…	… …	…
25553	No damage	…	False
25554	No damage	…	False
25555	No damage	…	False
25556	No damage	…	False
25557	Caused damage	…	False

	Remarks	Wildlife: Size \
0	FLT 753. PILOT REPTD A HUNDRED BIRDS ON UNKN T…	Medium
1	102 CARCASSES FOUND. 1 LDG LIGHT ON NOSE GEAR …	Small
2	FLEW UNDER A VERY LARGE FLOCK OF BIRDS OVER AP…	Small
3	NOTAM WARNING. 26 BIRDS HIT THE A/C, FORCING A…	Small
4	NO DMG REPTD.	Small
…	…	…
25553	DUCK? NO DMG REPTD.	Large
25554	NaN	Small
25555	STRUCK BIRD ON RT FRONT DURING T/O. BIRD REPTD…	NaN
25556	PILOTS REPORT STRIKING UNKNOWN BIRD ON RWY 21L…	Medium
25557	HIT CENTER OF RADOME, CAVING IN ABOUT 12". RAD…	Medium

	Conditions: Sky	Wildlife: Species \
0	No Cloud	Unknown bird - medium
1	Some Cloud	Rock pigeon
2	No Cloud	European starling
3	Some Cloud	European starling
4	No Cloud	European starling
…	…	…
25553	Overcast	Unknown bird - large
25554	Some Cloud	Tree swallow
25555	No Cloud	Unknown bird - small
25556	Some Cloud	Unknown bird - medium
25557	No Cloud	Red-tailed hawk

```
        Pilot warned of birds or wildlife?  Cost: Total $ Feet above ground  \
0                                         N          30736            1500.0
1                                         Y              0               0.0
2                                         N              0              50.0
3                                         Y              0              50.0
4                                         N              0              50.0
...                                     ...            ...               ...
25553                                     N              0            1500.0
25554                                     Y              0               0.0
25555                                   NaN              0               NaN
25556                                     Y              0               0.0
25557                                     N              0               0.0

        Number of people injured Is Aircraft Large?
0                              0               Yes
1                              0                No
2                              0                No
3                              0               Yes
4                              0                No
...                          ...               ...
25553                          0                No
25554                          0                No
25555                          0               NaN
25556                          0               Yes
25557                          0               Yes

[25558 rows x 26 columns]
```

```python
[14]: df['Aircraft: Type'] = df['Aircraft: Type'].fillna('Not Available')
      df['Airport: Name'] = df['Airport: Name'].fillna('Not Available')
      df['Altitude bin'] = df['Altitude bin'].fillna('Not Available')
      df['Aircraft: Make/Model'] = df['Aircraft: Make/Model'].fillna('Not Available')
      df['Wildlife: Number struck'] = df['Wildlife: Number struck'].fillna('Not␣
       ↪Available')
      df['Effect: Impact to flight'] = df['Effect: Impact to flight'].fillna('Not␣
       ↪Available')
      df['When: Phase of flight']=df['When: Phase of flight'].fillna('Not Available')
      df_cleaned = df.fillna('Unvailable')
      display(df_cleaned.head(10))
      df_cleaned.info()
```

```
   Record ID Aircraft: Type              Airport: Name Altitude bin  \
0     202152       Airplane               LAGUARDIA NY    > 1000 ft
1     208159       Airplane  DALLAS/FORT WORTH INTL ARPT   < 1000 ft
2     207601       Airplane            LAKEFRONT AIRPORT   < 1000 ft
3     215953       Airplane        SEATTLE-TACOMA INTL     < 1000 ft
4     219878       Airplane               NORFOLK INTL     < 1000 ft
5     218432       Airplane          GUAYAQUIL/S BOLIVAR   < 1000 ft
```

```
6    221697        Airplane              NEW CASTLE COUNTY   < 1000 ft
7    236635        Airplane   WASHINGTON DULLES INTL ARPT   < 1000 ft
8    207369        Airplane                   ATLANTA INTL   < 1000 ft
9    204371        Airplane   ORLANDO SANFORD INTL AIRPORT   < 1000 ft


   Aircraft: Make/Model Wildlife: Number struck  \
0            B-737-400                   Over 100
1                MD-80                   Over 100
2                C-500                   Over 100
3            B-737-400                   Over 100
4           CL-RJ100/200                 Over 100
5                A-300                   Over 100
6           LEARJET-25                   Over 100
7                A-320                   Over 100
8              DC-9-30                    2 to 10
9                A-330                    2 to 10


   Wildlife: Number Struck Actual Effect: Impact to flight  \
0                           859           Engine Shut Down
1                           424              Not Available
2                           261              Not Available
3                           806       Precautionary Landing
4                           942              Not Available
5                           537              Not Available
6                           227                      Other
7                           320                      Other
8                             9           Aborted Take-off
9                             4              Not Available


           FlightDate Effect: Indicated Damage   …  \
0 2000-11-23 00:00:00            Caused damage   …
1 2001-07-25 00:00:00            Caused damage   …
2 2001-09-14 00:00:00               No damage   …
3 2002-09-05 00:00:00               No damage   …
4 2003-06-23 00:00:00               No damage   …
5 2003-07-24 00:00:00               No damage   …
6 2003-08-17 00:00:00            Caused damage   …
7 2006-03-01 00:00:00            Caused damage   …
8 2000-01-06 00:00:00               No damage   …
9 2000-01-07 00:00:00               No damage   …


  Remains of wildlife sent to Smithsonian  \
0                                    False
1                                    False
2                                    False
3                                    False
4                                    False
5                                    False
```

```
6                                    True
7                                   False
8                                   False
9                                   False


                                            Remarks  Wildlife: Size  \
0  FLT 753. PILOT REPTD A HUNDRED BIRDS ON UNKN T…          Medium
1  102 CARCASSES FOUND. 1 LDG LIGHT ON NOSE GEAR …           Small
2  FLEW UNDER A VERY LARGE FLOCK OF BIRDS OVER AP…           Small
3  NOTAM WARNING. 26 BIRDS HIT THE A/C, FORCING A…           Small
4                                  NO DMG REPTD.            Small
5         NO DMG. BIRD REMAINS ON F/O WINDSCREEN.           Small
6                                     Unavailable           Small
7  WS ASSISTED IN CLEAN-UP OF 273 STARLINGS AND 1…          Small
8                                     Unavailable           Small
9                                        FLT 057           Small


  Conditions: Sky       Wildlife: Species  Pilot warned of birds or wildlife?  \
0       No Cloud  Unknown bird – medium                                    N
1     Some Cloud            Rock pigeon                                    Y
2       No Cloud      European starling                                    N
3     Some Cloud      European starling                                    Y
4       No Cloud      European starling                                    N
5       No Cloud   Unknown bird – small                                    N
6       No Cloud      European starling                                    N
7     Some Cloud      European starling                                    Y
8     Some Cloud            Rock pigeon                                    N
9     Some Cloud   Unknown bird – small                                    N


   Cost: Total $ Feet above ground Number of people injured Is Aircraft Large?
0          30736            1500.0                        0                Yes
1              0               0.0                        0                 No
2              0              50.0                        0                 No
3              0              50.0                        0                Yes
4              0              50.0                        0                 No
5              0               0.0                        0                 No
6        1481711             150.0                        0                 No
7        1483141             100.0                        0                 No
8              0               0.0                        0                 No
9              0               0.0                        0                 No


[10 rows x 26 columns]

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25558 entries, 0 to 25557
Data columns (total 26 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
```

```
0   Record ID                                25558 non-null  int64
1   Aircraft: Type                           25558 non-null  object
2   Airport: Name                            25558 non-null  object
3   Altitude bin                             25558 non-null  object
4   Aircraft: Make/Model                     25558 non-null  object
5   Wildlife: Number struck                  25558 non-null  object
6   Wildlife: Number Struck Actual           25558 non-null  int64
7   Effect: Impact to flight                 25558 non-null  object
8   FlightDate                               25558 non-null  object
9   Effect: Indicated Damage                 25558 non-null  object
10  Aircraft: Number of engines?             25558 non-null  object
11  Aircraft: Airline/Operator               25558 non-null  object
12  Origin State                             25558 non-null  object
13  When: Phase of flight                    25558 non-null  object
14  Conditions: Precipitation                25558 non-null  object
15  Remains of wildlife collected?           25558 non-null  bool
16  Remains of wildlife sent to Smithsonian  25558 non-null  bool
17  Remarks                                  25558 non-null  object
18  Wildlife: Size                           25558 non-null  object
19  Conditions: Sky                          25558 non-null  object
20  Wildlife: Species                        25558 non-null  object
21  Pilot warned of birds or wildlife?       25558 non-null  object
22  Cost: Total $                            25558 non-null  int64
23  Feet above ground                        25558 non-null  object
24  Number of people injured                 25558 non-null  int64
25  Is Aircraft Large?                       25558 non-null  object
dtypes: bool(2), int64(4), object(20)
memory usage: 4.7+ MB
```

[15]: `df_cleaned.describe()`

[15]:
	Record ID	Wildlife: Number Struck Actual	Cost: Total $
count	25558.000000	25558.000000	2.555800e+04
mean	253916.085609	2.691525	5.567354e+03
std	38510.453382	12.793975	1.219713e+05
min	1195.000000	1.000000	0.000000e+00
25%	225783.750000	1.000000	0.000000e+00
50%	248749.000000	1.000000	0.000000e+00
75%	269168.750000	1.000000	0.000000e+00
max	321909.000000	942.000000	1.239775e+07

	Number of people injured
count	25558.000000
mean	0.001056
std	0.050420
min	0.000000
25%	0.000000

```
50%                        0.000000
75%                        0.000000
max                        6.000000
```

[16]: `display(df_cleaned.head(10))`

```
   Record ID Aircraft: Type              Airport: Name Altitude bin  \
0     202152       Airplane                 LAGUARDIA NY     > 1000 ft
1     208159       Airplane    DALLAS/FORT WORTH INTL ARPT   < 1000 ft
2     207601       Airplane             LAKEFRONT AIRPORT    < 1000 ft
3     215953       Airplane           SEATTLE-TACOMA INTL    < 1000 ft
4     219878       Airplane                 NORFOLK INTL     < 1000 ft
5     218432       Airplane           GUAYAQUIL/S BOLIVAR    < 1000 ft
6     221697       Airplane             NEW CASTLE COUNTY    < 1000 ft
7     236635       Airplane    WASHINGTON DULLES INTL ARPT   < 1000 ft
8     207369       Airplane                 ATLANTA INTL     < 1000 ft
9     204371       Airplane    ORLANDO SANFORD INTL AIRPORT  < 1000 ft


  Aircraft: Make/Model Wildlife: Number struck  \
0            B-737-400                  Over 100
1                MD-80                  Over 100
2                C-500                  Over 100
3            B-737-400                  Over 100
4          CL-RJ100/200                 Over 100
5                A-300                  Over 100
6           LEARJET-25                  Over 100
7                A-320                  Over 100
8              DC-9-30                    2 to 10
9                A-330                    2 to 10


  Wildlife: Number Struck Actual Effect: Impact to flight  \
0                            859           Engine Shut Down
1                            424              Not Available
2                            261              Not Available
3                            806         Precautionary Landing
4                            942              Not Available
5                            537              Not Available
6                            227                      Other
7                            320                      Other
8                              9           Aborted Take-off
9                              4              Not Available


          FlightDate Effect: Indicated Damage   … \
0  2000-11-23 00:00:00           Caused damage   …
1  2001-07-25 00:00:00           Caused damage   …
2  2001-09-14 00:00:00              No damage   …
3  2002-09-05 00:00:00              No damage   …
4  2003-06-23 00:00:00              No damage   …
```

```
5  2003-07-24 00:00:00                    No damage  …
6  2003-08-17 00:00:00                Caused damage  …
7  2006-03-01 00:00:00                Caused damage  …
8  2000-01-06 00:00:00                    No damage  …
9  2000-01-07 00:00:00                    No damage  …

  Remains of wildlife sent to Smithsonian  \
0                                   False
1                                   False
2                                   False
3                                   False
4                                   False
5                                   False
6                                    True
7                                   False
8                                   False
9                                   False


                                          Remarks Wildlife: Size  \
0  FLT 753. PILOT REPTD A HUNDRED BIRDS ON UNKN T…          Medium
1  102 CARCASSES FOUND. 1 LDG LIGHT ON NOSE GEAR …           Small
2  FLEW UNDER A VERY LARGE FLOCK OF BIRDS OVER AP…           Small
3  NOTAM WARNING. 26 BIRDS HIT THE A/C, FORCING A…           Small
4                               NO DMG REPTD.               Small
5          NO DMG. BIRD REMAINS ON F/O WINDSCREEN.          Small
6                                   Unavailable              Small
7  WS ASSISTED IN CLEAN-UP OF 273 STARLINGS AND 1…           Small
8                                   Unavailable              Small
9                                       FLT 057              Small


  Conditions: Sky        Wildlife: Species  Pilot warned of birds or wildlife?  \
0        No Cloud  Unknown bird - medium                                    N
1      Some Cloud            Rock pigeon                                    Y
2        No Cloud       European starling                                    N
3      Some Cloud       European starling                                    Y
4        No Cloud       European starling                                    N
5        No Cloud    Unknown bird - small                                    N
6        No Cloud       European starling                                    N
7      Some Cloud       European starling                                    Y
8      Some Cloud            Rock pigeon                                    N
9      Some Cloud    Unknown bird - small                                    N


  Cost: Total $ Feet above ground Number of people injured Is Aircraft Large?
0          30736            1500.0                        0                Yes
1              0               0.0                        0                 No
2              0              50.0                        0                 No
3              0              50.0                        0                Yes
4              0              50.0                        0                 No
```

```
5           0           0.0                         0           No
6     1481711         150.0                         0           No
7     1483141         100.0                         0           No
8           0           0.0                         0           No
9           0           0.0                         0           No

[10 rows x 26 columns]
```

[17]: `df_cleaned.isnull().sum()`

```
[17]: Record ID                               0
      Aircraft: Type                          0
      Airport: Name                           0
      Altitude bin                            0
      Aircraft: Make/Model                    0
      Wildlife: Number struck                 0
      Wildlife: Number Struck Actual          0
      Effect: Impact to flight                0
      FlightDate                              0
      Effect: Indicated Damage                0
      Aircraft: Number of engines?            0
      Aircraft: Airline/Operator              0
      Origin State                            0
      When: Phase of flight                   0
      Conditions: Precipitation               0
      Remains of wildlife collected?          0
      Remains of wildlife sent to Smithsonian 0
      Remarks                                 0
      Wildlife: Size                          0
      Conditions: Sky                         0
      Wildlife: Species                       0
      Pilot warned of birds or wildlife?      0
      Cost: Total $                           0
      Feet above ground                       0
      Number of people injured                0
      Is Aircraft Large?                      0
      dtype: int64
```

[18]: `df.to_excel("C:/Users/thala/Downloads/BS3_Cleaned.xlsx",index= False)`

[ ]: