NKC - Animal Shelter and Rescue statistics: Data Wrangling

Data wrangling and Cleaning process performed by Santhosha J

Shelter And Rescue Statistics Data - 2016

For this project, I used both Excel and Python (Jupyter Notebook). Before importing the data into Python, I performed initial cleaning in Excel, like normalizing column names and removing unwanted rows or columns. After these basic adjustments, I imported the data into Python for further wrangling and analysis.

In [1]:
```
# Importing all the required libraries
import numpy as np        # For numerical operations and array manipulation
import pandas as pd       # For data manipulation and analysis
```

In [2]:
```
# Loading the data
data_2016 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [3]:
```
data_2016.head()
```

Out[3]:

	PACFA License Number	Facility Name	Zip Code	County	Adult Dogs: Beginning Count	Adult Dogs: Beginning Foster Count	Adult Dogs: Stray	Adult Dogs: Owner Relinquished	Adult Dogs: Owner Requested Euthanasia upon intake	Adult Dogs: Transfer in from another Colorado organization	...	Other: Other Live Outcomes (ie: tnr / snr)	Other: Died	C M S
0	PL0010TH	German Shepherd Rescue of Central Colorado	80449	Park County	1.0	NaN	NaN	6.0	NaN	NaN	...	NaN	NaN	
1	PL002RLS	Gritty Pittie Rescue	80501	Boulder County	NaN	NaN	NaN	1.0	NaN	NaN	...	NaN	NaN	
2	PL001V1M	Mountian High collie & sheltie Rescue	80232	Jefferson County	NaN	NaN	NaN	2.0	NaN	1.0	...	NaN	NaN	
3	PL002TR8	Doggy Dog World Rescue	80125	Douglas County	25.0	25.0	NaN	7.0	NaN	33.0	...	NaN	NaN	
4	PL001A7K	Surface Creek Shelter	81413	Delta County	6.0	NaN	131.0	26.0	NaN	12.0	...	NaN	NaN	

5 rows × 194 columns

In [4]:
```
data_2016.shape   # total rows and columns
```

Out[4]: (257, 194)

Since the data was originally in wide format, I transformed it into long format using unpivot to standardize and simplify the process.

In [5]:
```
# Define ID columns
id_vars = ["PACFA License Number", "Facility Name", "Zip Code", "County", "Location"]

# Melt the rest of the columns into two: "Metric" and "Value"
data_2016_unpivoted = data_2016.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2016_unpivoted = data_2016_unpivoted.dropna(subset=["Value"])
```

In [6]:
```
data_2016_unpivoted.head()
```

Out[6]:

	PACFA License Number	Facility Name	Zip Code	County	Location	Metric	Value
0	PL0010TH	German Shepherd Rescue of Central Colorado	80449	Park County	Hartsel Colorado (35.500801, -117.9478)	Adult Dogs: Beginning Count	1.0
3	PL002TR8	Doggy Dog World Rescue	80125	Douglas County	Littleton (39.612653, -105.016198)	Adult Dogs: Beginning Count	25.0
4	PL001A7K	Surface Creek Shelter	81413	Delta County	Cedaredge (38.900738, -107.923767)	Adult Dogs: Beginning Count	6.0
5	PL001H6U	Delta County Citizens for Animal Welfare and S...	CO 81416	Delta County	Delta (38.741684, -108.070175)	Adult Dogs: Beginning Count	7.0
6	PL0006XU	Dalmatian Rescue of Colorado	80526	Larimer County	Fort Collins (40.588972, -105.082459)	Adult Dogs: Beginning Count	3.0

In [7]:
```
data_2016_unpivoted.shape
```

Out[7]:
```
(6374, 7)
```

After converting the data into long format, I created a separate column for `Animal Type` by splitting the metric column.

In [8]:
```
# Split Metric column into two at the first colon (:)
data_2016_unpivoted[["Animal Type", "Metric Type"]] = data_2016_unpivoted["Metric"].str.split(":", n=1, expand=True)

# Clean up whitespace
data_2016_unpivoted["Animal Type"] = data_2016_unpivoted["Animal Type"].str.strip()
data_2016_unpivoted["Metric Type"] = data_2016_unpivoted["Metric Type"].str.strip()

# Removing old 'Metric' column
data_2016_unpivoted.drop(["Metric"], axis=1, inplace=True)
```

In [9]:
```
data_2016_unpivoted.head()
```

Out[9]:

	PACFA License Number	Facility Name	Zip Code	County	Location	Value	Animal Type	Metric Type
0	PL0010TH	German Shepherd Rescue of Central Colorado	80449	Park County	Hartsel Colorado (35.500801, -117.9478)	1.0	Adult Dogs	Beginning Count
3	PL002TR8	Doggy Dog World Rescue	80125	Douglas County	Littleton (39.612653, -105.016198)	25.0	Adult Dogs	Beginning Count
4	PL001A7K	Surface Creek Shelter	81413	Delta County	Cedaredge (38.900738, -107.923767)	6.0	Adult Dogs	Beginning Count
5	PL001H6U	Delta County Citizens for Animal Welfare and S...	CO 81416	Delta County	Delta (38.741684, -108.070175)	7.0	Adult Dogs	Beginning Count
6	PL0006XU	Dalmatian Rescue of Colorado	80526	Larimer County	Fort Collins (40.588972, -105.082459)	3.0	Adult Dogs	Beginning Count

In the Metric Type column, I removed unnecessary categories to ensure consistency with other data files and renamed certain categories for improved clarity and easier understanding.

In [10]:
```
data_2016_unpivoted["Metric Type"].unique()
```

Out[10]:
```
array(['Beginning Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished', 'Owner Requested Euthanasia upon intake',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transfer or Rescue out to another Colorado organization',
       'Transfer or Rescue out to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died', 'Missing / Stolen',
       'Shelter Euthanasia', 'Owner Requested Euthanasia', 'Ending Count',
       'Foster Count', 'AVG LOS', 'Notes'], dtype=object)
```

In [11]:
```
# Drop rows where metric type is 'AVG LOS' or 'Notes'
data_2016_unpivoted = data_2016_unpivoted[~data_2016_unpivoted['Metric Type'].isin(['AVG LOS', 'Notes'])]
# (df['metric type'].isin([...]) → checks if the value is in the given list and ~ (tilde) → negates it (means "NOT").)
```

In [12]:
```
data_2016_unpivoted.shape
```

Out[12]:
```
(5496, 8)
```

In [13]:
```
# Renaming categories inside Metric Type column
data_2016_unpivoted["Metric Type"] = data_2016_unpivoted["Metric Type"].replace(
    {"Beginning Count": "Beginning Shelter Count",
     "Beginning Foster Count": "Beginning Foster Count",
     "Ending Count": "Ending Shelter Count",
     "Foster Count": "Ending Foster Count",
     "Transfer or Rescue out to another Colorado organization": "Transferred out to another Colorado organization",
     "Transfer or Rescue out to an Out of State organization": "Transferred to an Out of State organization"
    })

# Check the unique values after renaming
print(data_2016_unpivoted['Metric Type'].unique())
```

```
['Beginning Shelter Count' 'Beginning Foster Count' 'Stray'
 'Owner Relinquished' 'Owner Requested Euthanasia upon intake'
 'Transfer in from another Colorado organization'
 'Transfer in from Out of State organization'
 'Other: TNR / Protective Custody / Returns / Disaster Relief' 'Adoption'
 'Returned To Owner (RTO)'
 'Transferred out to another Colorado organization'
 'Transferred to an Out of State organization'
 'Other Live Outcomes (ie: tnr / snr)' 'Died' 'Missing / Stolen'
 'Shelter Euthanasia' 'Owner Requested Euthanasia' 'Ending Shelter Count'
 'Ending Foster Count']
```

Added a new column `Flow Type` to map and categorize records into Start of Year Count, End of Year Count, Intake, and Outcome, facilitating structured analysis and aggregation in reports.

In [14]:
```
# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",

    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Owner Requested Euthanasia upon intake": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
```

```
    "Missing / Stolen": "Outcome (Negative)",
    "Shelter Euthanasia": "Outcome (Negative)",
    "Owner Requested Euthanasia": "Outcome (Negative)"
}

# Apply mapping to create new column
data_2016_unpivoted["Flow Type"] = data_2016_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2016_unpivoted = data_2016_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [15]:
```
# Convert to integer
data_2016_unpivoted["Animal Count"] = data_2016_unpivoted["Animal Count"].astype("int64")
```

In [16]:
```
data_2016_unpivoted.head()
```

Out[16]:

	PACFA License Number	Facility Name	Zip Code	County	Location	Animal Count	Animal Type	Event Type	Flow Type
0	PL0010TH	German Shepherd Rescue of Central Colorado	80449	Park County	Hartsel Colorado (35.500801, -117.9478)	1	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	PL002TR8	Doggy Dog World Rescue	80125	Douglas County	Littleton (39.612653, -105.016198)	25	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	PL001A7K	Surface Creek Shelter	81413	Delta County	Cedaredge (38.900738, -107.923767)	6	Adult Dogs	Beginning Shelter Count	Start of Year Count
5	PL001H6U	Delta County Citizens for Animal Welfare and S...	CO 81416	Delta County	Delta (38.741684, -108.070175)	7	Adult Dogs	Beginning Shelter Count	Start of Year Count
6	PL0006XU	Dalmatian Rescue of Colorado	80526	Larimer County	Fort Collins (40.588972, -105.082459)	3	Adult Dogs	Beginning Shelter Count	Start of Year Count

Added a `Reporting Year` column to the entire table to accommodate subsequent years of data and maintain a unified dataset.

In [17]:
```
# Adding Reporting year column to the whole table
```

```
data_2016_unpivoted["Reporting Year"] = 2016
```

Extracted two separate columns from the 'Animal Type' field - `Species` and `Age Group`, to enable more granular analysis and reporting, and dropped the original 'Animal Type' column to avoid redundancy.

In [18]:
```
# Adding Species column
data_2016_unpivoted["Species"] = (data_2016_unpivoted["Animal Type"]
                                  .str.replace("Adult", "", case=False)
                                  .str.replace("Juvenile", "", case=False)
                                  .str.strip()
                                 )
```

In [19]:
```
data_2016_unpivoted["Species"].unique()
```

Out[19]:
```
array(['Dogs', 'Cats', 'Birds', 'Small Mammal', 'Reptiles', 'Rabbits',
       'Other'], dtype=object)
```

In [20]:
```
# Adding Age group column
data_2016_unpivoted["Age Group"] = np.where(
    data_2016_unpivoted["Animal Type"].str.contains("Juvenile", case=False, na=False), "Juvenile",
    np.where(data_2016_unpivoted["Animal Type"].str.contains("Adult", case=False, na=False), "Adult",
             "Unknown")
)

# (case=False → Ignore uppercase/lowercase and na=False → Treat missing values as "does not match")
```

In [21]:
```
data_2016_unpivoted["Age Group"].unique()
```

Out[21]:
```
array(['Adult', 'Juvenile', 'Unknown'], dtype=object)
```

In [22]:
```
# Dropped Animal Type column
data_2016_unpivoted = data_2016_unpivoted.drop(columns=["Animal Type"], errors="ignore")
```

Extracted structured information from the 'Location' column, creating separate `City`, `Longitude`, and `Latitude` fields for improved geospatial analysis and reporting.

In [23]:
```
# Split into "city" and "(lat, long)"
data_2016_unpivoted[["City", "Coordinates"]] = data_2016_unpivoted["Location"].str.split("(", n=1, expand=True)
```

```
# (splits each string at the first "(" -- n=1 → ensures we split only once -- expand=True → puts results into two new columns.

# Clean up:
# removes extra spaces at the beginning or end.
data_2016_unpivoted["City"] = data_2016_unpivoted["City"].str.strip()

# removes the closing parenthesis from Coordinates
data_2016_unpivoted["Coordinates"] = data_2016_unpivoted["Coordinates"].str.replace(")", "").str.strip()

# Split coordinates into lat & long
data_2016_unpivoted[["Latitude", "Longitude"]] = data_2016_unpivoted["Coordinates"].str.split(",", n=1, expand=True)

# Convert to numeric
data_2016_unpivoted["Latitude"] = data_2016_unpivoted["Latitude"].astype(float)
data_2016_unpivoted["Longitude"] = data_2016_unpivoted["Longitude"].astype(float)
```

In [24]:
```
# Removing unnecessary columns
data_2016_unpivoted = data_2016_unpivoted.drop(columns=["Location", "Coordinates"], errors="ignore")
```

In [25]:
```
data_2016_unpivoted.head()
```

Out[25]:

	PACFA License Number	Facility Name	Zip Code	County	Animal Count	Event Type	Flow Type	Reporting Year	Species	Age Group	City	Latitude	Longitude
0	PL0010TH	German Shepherd Rescue of Central Colorado	80449	Park County	1	Beginning Shelter Count	Start of Year Count	2016	Dogs	Adult	Hartsel Colorado	35.500801	-117.947800
3	PL002TR8	Doggy Dog World Rescue	80125	Douglas County	25	Beginning Shelter Count	Start of Year Count	2016	Dogs	Adult	Littleton	39.612653	-105.016198
4	PL001A7K	Surface Creek Shelter	81413	Delta County	6	Beginning Shelter Count	Start of Year Count	2016	Dogs	Adult	Cedaredge	38.900738	-107.923767
5	PL001H6U	Delta County Citizens for Animal Welfare and S...	CO 81416	Delta County	7	Beginning Shelter Count	Start of Year Count	2016	Dogs	Adult	Delta	38.741684	-108.070175
6	PL0006XU	Dalmatian Rescue of Colorado	80526	Larimer County	3	Beginning Shelter Count	Start of Year Count	2016	Dogs	Adult	Fort Collins	40.588972	-105.082459

Verified that the columns contain no null values, ensuring the data is in a consistent and reliable structure for analysis.

In [26]:
```
data_2016_unpivoted.isnull().sum()
```

Out[26]:
```
PACFA License Number    0
Facility Name           0
Zip Code                0
County                  0
Animal Count            0
Event Type              0
Flow Type               0
Reporting Year          0
Species                 0
Age Group               0
City                    0
Latitude                0
Longitude               0
dtype: int64
```

In [27]:
```
data_2016_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 5496 entries, 0 to 47769
Data columns (total 13 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   PACFA License Number  5496 non-null   object
 1   Facility Name         5496 non-null   object
 2   Zip Code              5496 non-null   object
 3   County                5496 non-null   object
 4   Animal Count          5496 non-null   int64
 5   Event Type            5496 non-null   object
 6   Flow Type             5496 non-null   object
 7   Reporting Year        5496 non-null   int64
 8   Species               5496 non-null   object
 9   Age Group             5496 non-null   object
 10  City                  5496 non-null   object
 11  Latitude              5496 non-null   float64
 12  Longitude             5496 non-null   float64
dtypes: float64(2), int64(2), object(9)
memory usage: 601.1+ KB
```

Verified all event types and their respective counts against the summary document, ensuring that the data matches and is accurate.

In [29]:
```
# Sum of Animal Count per Event Type
data_2016_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[29]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	3411
1	End of Year Count	Ending Shelter Count	10080
2	Intake	Other: TNR / Protective Custody / Returns / Di...	13638
3	Intake	Owner Relinquished	37478
4	Intake	Owner Requested Euthanasia upon intake	6015
5	Intake	Stray	64945
6	Intake	Transfer in from Out of State organization	34888
7	Intake	Transfer in from another Colorado organization	15493
8	Outcome (Negative)	Died	2147
9	Outcome (Negative)	Missing / Stolen	185
10	Outcome (Negative)	Owner Requested Euthanasia	5977
11	Outcome (Negative)	Shelter Euthanasia	12130
12	Outcome (Positive)	Adoption	105722
13	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	6162
14	Outcome (Positive)	Returned To Owner (RTO)	27710
15	Outcome (Positive)	Transferred out to another Colorado organization	12826
16	Outcome (Positive)	Transferred to an Out of State organization	337
17	Start of Year Count	Beginning Foster Count	3599
18	Start of Year Count	Beginning Shelter Count	9678

In [30]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2016 = data_2016_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in

summary_2016.to_excel("summary_2016.xlsx", index=False)
```

In [31]:
```
# Rearrange the columns
Cleaned_2016_Shelter_And_Rescue_Statistics_final = data_2016_unpivoted[["PACFA License Number", "Facility Name", "County",
                                                                        "City", "Zip Code", "Latitude", "Longitude", "Reporting Y
                                                                        "Species", "Age Group", "Flow Type", "Event Type",
                                                                        "Animal Count"]]
```

In [32]:
```
Cleaned_2016_Shelter_And_Rescue_Statistics_final.head()
```

Out[32]:

	PACFA License Number	Facility Name	County	City	Zip Code	Latitude	Longitude	Reporting Year	Species	Age Group	Flow Type	Event Type	Animal Count
0	PL0010TH	German Shepherd Rescue of Central Colorado	Park County	Hartsel Colorado	80449	35.500801	-117.947800	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	1
3	PL002TR8	Doggy Dog World Rescue	Douglas County	Littleton	80125	39.612653	-105.016198	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	25
4	PL001A7K	Surface Creek Shelter	Delta County	Cedaredge	81413	38.900738	-107.923767	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	6
5	PL001H6U	Delta County Citizens for Animal Welfare and S...	Delta County	Delta	CO 81416	38.741684	-108.070175	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	7
6	PL0006XU	Dalmatian Rescue of Colorado	Larimer County	Fort Collins	80526	40.588972	-105.082459	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	3

In [33]:
```
Cleaned_2016_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2016_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2016_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2016_Shelter_And_Rescue_Statistics_final.xlsx
```

In [34]:
```
from IPython.display import FileLink
FileLink('Cleaned_2016_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[34]: Cleaned_2016_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:

Shelter And Rescue Statistics Data - 2017

In [1]:
```
# Importing all the required libraries
import numpy as np        # For numerical operations and array manipulation
import pandas as pd       # For data manipulation and analysis
```

In [2]:
```
# Loading the data
data_2017 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [3]:
```
data_2017.head()
```

Out[3]:

	Facility Name	Zip Code	County	Adult Dogs: In Shelter Count as of 1/1/2017	Adult Dogs: In Foster Care Count 1/1/2017	Adult Dogs: Stray	Adult Dogs: Owner Relinquished	Adult Dogs: Owner Requested Euthanasia upon intake	Adult Dogs: Transfer in from another Colorado organization	Adult Dogs: Transfer in from Out of State organization	...	Other: Transferred out to another Colorado organization	Trai to a orga
0	Peaceful Animal Adoption Shelter (PAAS)	74301	-	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	
1	FOUND NATION ANIMAL RESCUE	80003	Jefferson County	0.0	20.0	0.0	7.0	0.0	5.0	190.0	...	NaN	
2	COLORADO WYOMING VIZSLA RESCUE GROUP INC	80007	Jefferson County	3.0	18.0	0.0	16.0	0.0	2.0	4.0	...	NaN	
3	Golden Retriever Rescue of The Rockies	80007	Jefferson County	6.0	3.0	2.0	117.0	0.0	18.0	83.0	...	NaN	
4	City of Aurora Animal Shelter	80011	Adams County	50.0	0.0	1618.0	191.0	73.0	27.0	0.0	...	0.0	

5 rows × 183 columns

In [4]:
```
data_2017.shape   # total rows and columns
```

Out[4]: (275, 183)

In [5]:
```
# Unpivoting the DataFrame from wide to long format

# Define ID (identifier variables) columns
id_vars = ["Facility Name", "Zip Code", "County", "Location"]

# Unpivot the rest of the columns into two: "Metric" and "Value"
data_2017_unpivoted = data_2017.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2017_unpivoted = data_2017_unpivoted.dropna(subset=["Value"])
```

In [6]:
```
data_2017_unpivoted.head()
```

Out[6]:

	Facility Name	Zip Code	County	Location	Metric	Value
1	FOUND NATION ANIMAL RESCUE	80003	Jefferson County	ARVADA 80003 (39.82682, -105.06527)	Adult Dogs: In Shelter Count as of 1/1/2017	0.0
2	COLORADO WYOMING VIZSLA RESCUE GROUP INC	80007	Jefferson County	Arvada 80007 (39.839939, -105.186131)	Adult Dogs: In Shelter Count as of 1/1/2017	3.0
3	Golden Retriever Rescue of The Rockies	80007	Jefferson County	Arvada 80007 (39.839939, -105.186131)	Adult Dogs: In Shelter Count as of 1/1/2017	6.0
4	City of Aurora Animal Shelter	80011	Adams County	Aurora 80011 (39.74187, -104.799113)	Adult Dogs: In Shelter Count as of 1/1/2017	50.0
5	Colorado St. Bernard Rescue, Inc.	80011	Arapahoe County	Aurora 80011 (39.74187, -104.799113)	Adult Dogs: In Shelter Count as of 1/1/2017	0.0

In [7]:
```
data_2017_unpivoted.shape
```

Out[7]: (17455, 6)

In [8]:
```
# Split Metric column into two at the first colon (:)
data_2017_unpivoted[["Animal Type", "Metric Type"]] = data_2017_unpivoted["Metric"].str.split(":", n=1, expand=True)

# Clean up whitespace
data_2017_unpivoted["Animal Type"] = data_2017_unpivoted["Animal Type"].str.strip()
data_2017_unpivoted["Metric Type"] = data_2017_unpivoted["Metric Type"].str.strip()

# Removing old 'Metric' column
data_2017_unpivoted.drop(["Metric"], axis=1, inplace=True)
```

In [9]:
```
data_2017_unpivoted.head()
```

Out[9]:

	Facility Name	Zip Code	County	Location	Value	Animal Type	Metric Type
1	FOUND NATION ANIMAL RESCUE	80003	Jefferson County	ARVADA 80003 (39.82682, -105.06527)	0.0	Adult Dogs	In Shelter Count as of 1/1/2017
2	COLORADO WYOMING VIZSLA RESCUE GROUP INC	80007	Jefferson County	Arvada 80007 (39.839939, -105.186131)	3.0	Adult Dogs	In Shelter Count as of 1/1/2017
3	Golden Retriever Rescue of The Rockies	80007	Jefferson County	Arvada 80007 (39.839939, -105.186131)	6.0	Adult Dogs	In Shelter Count as of 1/1/2017
4	City of Aurora Animal Shelter	80011	Adams County	Aurora 80011 (39.74187, -104.799113)	50.0	Adult Dogs	In Shelter Count as of 1/1/2017
5	Colorado St. Bernard Rescue, Inc.	80011	Arapahoe County	Aurora 80011 (39.74187, -104.799113)	0.0	Adult Dogs	In Shelter Count as of 1/1/2017

In [10]:
```
data_2017_unpivoted["Metric Type"].unique()
```

Out[10]:
```
array(['In Shelter Count as of  1/1/2017',
       'In Foster Care Count 1/1/2017', 'Stray', 'Owner Relinquished',
       'Owner Requested Euthanasia upon intake',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR/Protective Custody/Returns/Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an out of state organization',
       'Other live outcomes (ie: tnr/snr)', 'Died', 'Missing/Stolen',
       'Shelter Euthanasia', 'Owner Requested Euthanasia',
       'In Shelter Count as of 12/31/2017',
       'In Foster Care Count 12/31/2017',
       'In Shelter Count as of 1/1/2017',
       'Other: TNR/Protective Custody/Returns/Disaster relief',
       'Other live outcomes (ie tnr/snr)', 'Missing / Stolen',
       'Other Live Outcomes (ie:tnr / snr)',
       'Other Live Outcomes(ie: tnr/snr)',
       'Other Live Outcomes (ie: tnr/snr)',
       'In Foster Care Count as of 12/31/2017',
       'Other Live Outcomes ( ie: tnr/snr)',
       'TNR/Protective Custody/Returns/Disaster Relief'], dtype=object)
```

In [11]:
```
# Renaming categories inside Metric Type column
replace_map = {
    # Beginning Counts
    "In Shelter Count as of  1/1/2017": "Beginning Shelter Count",
    "In Shelter Count as of 1/1/2017": "Beginning Shelter Count",
    "In Foster Care Count 1/1/2017": "Beginning Foster Count",

    # Ending Counts
    "In Shelter Count as of 12/31/2017": "Ending Shelter Count",
    "In Foster Care Count 12/31/2017": "Ending Foster Count",
    "In Foster Care Count as of 12/31/2017": "Ending Foster Count",

    # TNR / Protective Custody variations
    "Other: TNR/Protective Custody/Returns/Disaster Relief": "Other: TNR / Protective Custody / Returns / Disaster Relief",
    "Other: TNR/Protective Custody/Returns/Disaster relief": "Other: TNR / Protective Custody / Returns / Disaster Relief",
    "TNR/Protective Custody/Returns/Disaster Relief": "Other: TNR / Protective Custody / Returns / Disaster Relief",
```

```
    # Transfer
    "Transferred to an out of state organization": "Transferred to an Out of State organization",

    # Other Live Outcomes variations
    "Other live outcomes (ie: tnr/snr)": "Other Live Outcomes (ie: tnr / snr)",
    "Other live outcomes (ie tnr/snr)": "Other Live Outcomes (ie: tnr / snr)",
    "Other Live Outcomes (ie:tnr / snr)": "Other Live Outcomes (ie: tnr / snr)",
    "Other Live Outcomes(ie: tnr/snr)": "Other Live Outcomes (ie: tnr / snr)",
    "Other Live Outcomes (ie: tnr/snr)": "Other Live Outcomes (ie: tnr / snr)",
    "Other Live Outcomes ( ie: tnr/snr)": "Other Live Outcomes (ie: tnr / snr)",

    # Missing variations
    "Missing/Stolen": "Missing / Stolen"
}

data_2017_unpivoted["Metric Type"] = data_2017_unpivoted["Metric Type"].replace(replace_map)
```

In [12]:
```
# Checking the unique values after normalizing the category names
data_2017_unpivoted["Metric Type"].unique()
```

Out[12]:
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished', 'Owner Requested Euthanasia upon intake',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died', 'Missing / Stolen',
       'Shelter Euthanasia', 'Owner Requested Euthanasia',
       'Ending Shelter Count', 'Ending Foster Count'], dtype=object)
```

In [13]:
```
# Added a new column Flow Type to map and categorize records

# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",
```

```
    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Owner Requested Euthanasia upon intake": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
    "Missing / Stolen": "Outcome (Negative)",
    "Shelter Euthanasia": "Outcome (Negative)",
    "Owner Requested Euthanasia": "Outcome (Negative)"
}

# Apply mapping to create new column
data_2017_unpivoted["Flow Type"] = data_2017_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2017_unpivoted = data_2017_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [14]:
```
# Convert to integer numeric column
data_2017_unpivoted["Animal Count"] = data_2017_unpivoted["Animal Count"].astype("int64")
```

In [15]:
```
data_2017_unpivoted.head()
```

Out[15]:

	Facility Name	Zip Code	County	Location	Animal Count	Animal Type	Event Type	Flow Type
1	FOUND NATION ANIMAL RESCUE	80003	Jefferson County	ARVADA 80003 (39.82682, -105.06527)	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	COLORADO WYOMING VIZSLA RESCUE GROUP INC	80007	Jefferson County	Arvada 80007 (39.839939, -105.186131)	3	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	Golden Retriever Rescue of The Rockies	80007	Jefferson County	Arvada 80007 (39.839939, -105.186131)	6	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	City of Aurora Animal Shelter	80011	Adams County	Aurora 80011 (39.74187, -104.799113)	50	Adult Dogs	Beginning Shelter Count	Start of Year Count
5	Colorado St. Bernard Rescue, Inc.	80011	Arapahoe County	Aurora 80011 (39.74187, -104.799113)	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [16]:
```
# Adding Reporting year column to the whole table
data_2017_unpivoted["Reporting Year"] = 2017
```

In [17]:
```
# Adding Species column by using Animal Type column
data_2017_unpivoted["Species"] = (data_2017_unpivoted["Animal Type"]
                                  .str.replace("Adult", "", case=False)
                                  .str.replace("Juvenile", "", case=False)
                                  .str.strip()
                                 )
# Checking unique values
data_2017_unpivoted["Species"].unique()
```

Out[17]:
```
array(['Dogs', 'Cats', 'Birds', 'Small Mammal', 'Sm Mammal',
       'Reptiles & Amphibians', 'Rabbits', 'Other'], dtype=object)
```

In [18]:
```
# Standardize similar species names
species_replace = {"Sm Mammal": "Small Mammal"}

# Apply replacement
data_2017_unpivoted["Species"] = data_2017_unpivoted["Species"].replace(species_replace)
```

```
# Now, checking unique values
data_2017_unpivoted["Species"].unique()
```

Out[18]:
```
array(['Dogs', 'Cats', 'Birds', 'Small Mammal', 'Reptiles & Amphibians',
       'Rabbits', 'Other'], dtype=object)
```

In [19]:
```
# Adding Age group column by using Animal Type column
data_2017_unpivoted["Age Group"] = np.where(
    data_2017_unpivoted["Animal Type"].str.contains("Juvenile", case=False, na=False), "Juvenile",
    np.where(data_2017_unpivoted["Animal Type"].str.contains("Adult", case=False, na=False), "Adult",
             "Unknown")
)

# checking unique values
data_2017_unpivoted["Age Group"].unique()
```

Out[19]:
```
array(['Adult', 'Juvenile', 'Unknown'], dtype=object)
```

In [20]:
```
# Removed the 'Animal Type' column to avoid redundancy after creating separate 'Species' and 'Age Group' columns.
data_2017_unpivoted = data_2017_unpivoted.drop(columns=["Animal Type"], errors="ignore")
```

In [21]:
```
data_2017_unpivoted.head()
```

Out[21]:

	Facility Name	Zip Code	County	Location	Animal Count	Event Type	Flow Type	Reporting Year	Species	Age Group
1	FOUND NATION ANIMAL RESCUE	80003	Jefferson County	ARVADA 80003 (39.82682, -105.06527)	0	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult
2	COLORADO WYOMING VIZSLA RESCUE GROUP INC	80007	Jefferson County	Arvada 80007 (39.839939, -105.186131)	3	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult
3	Golden Retriever Rescue of The Rockies	80007	Jefferson County	Arvada 80007 (39.839939, -105.186131)	6	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult
4	City of Aurora Animal Shelter	80011	Adams County	Aurora 80011 (39.74187, -104.799113)	50	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult
5	Colorado St. Bernard Rescue, Inc.	80011	Arapahoe County	Aurora 80011 (39.74187, -104.799113)	0	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult

In [22]:

```
# Extracting useful information from 'Location' column

# Splitting into 'PlaceZip' and 'Coordinates'
data_2017_unpivoted[["PlaceZip", "Coordinates"]] = data_2017_unpivoted["Location"].str.split("(", n=1, expand=True)

# Removing any extra Spaces from extracted columns
data_2017_unpivoted["PlaceZip"] = data_2017_unpivoted["PlaceZip"].str.strip()
data_2017_unpivoted["Coordinates"] = data_2017_unpivoted["Coordinates"].str.replace(")", "").str.strip()

# Splitting coordinates into Latitude & Longitude
data_2017_unpivoted[["Latitude", "Longitude"]] = data_2017_unpivoted["Coordinates"].str.split(",", n=1, expand=True)

# Converting to numeric
data_2017_unpivoted["Latitude"] = data_2017_unpivoted["Latitude"].astype(float)
data_2017_unpivoted["Longitude"] = data_2017_unpivoted["Longitude"].astype(float)

# Splitting PlaceZip into City and Zip Code
```

```
data_2017_unpivoted[["City", "Zip Code Extracted"]] = data_2017_unpivoted["PlaceZip"].str.rsplit(" ", n=1, expand=True)

# Drop helper columns
data_2017_unpivoted = data_2017_unpivoted.drop(columns=["PlaceZip", "Coordinates", "Location"])
```

In [23]:
```
data_2017_unpivoted.head()
```

Out[23]:

	Facility Name	Zip Code	County	Animal Count	Event Type	Flow Type	Reporting Year	Species	Age Group	Latitude	Longitude	City	Zip Code Extracted
1	FOUND NATION ANIMAL RESCUE	80003	Jefferson County	0	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult	39.826820	-105.065270	ARVADA	80003
2	COLORADO WYOMING VIZSLA RESCUE GROUP INC	80007	Jefferson County	3	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult	39.839939	-105.186131	Arvada	80007
3	Golden Retriever Rescue of The Rockies	80007	Jefferson County	6	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult	39.839939	-105.186131	Arvada	80007
4	City of Aurora Animal Shelter	80011	Adams County	50	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult	39.741870	-104.799113	Aurora	80011
5	Colorado St. Bernard Rescue, Inc.	80011	Arapahoe County	0	Beginning Shelter Count	Start of Year Count	2017	Dogs	Adult	39.741870	-104.799113	Aurora	80011

In [24]:
```
# Convert both to string and strip spaces
data_2017_unpivoted["Zip Code"] = data_2017_unpivoted["Zip Code"].astype(str).str.strip()
data_2017_unpivoted["Zip Code Extracted"] = data_2017_unpivoted["Zip Code Extracted"].astype(str).str.strip()
```

```
# Creating a similarity check column
data_2017_unpivoted["Zip_Match"] = data_2017_unpivoted["Zip Code"] == data_2017_unpivoted["Zip Code Extracted"]

# Check mismatches
mismatches = data_2017_unpivoted[data_2017_unpivoted["Zip_Match"] == False]
print(mismatches[["Zip Code", "Zip Code Extracted", "City"]])

# Calculate match rate
match_rate = data_2017_unpivoted["Zip_Match"].mean() * 100
print(f"Zip Code Match Rate: {match_rate:.2f}%")
```

```
Empty DataFrame
Columns: [Zip Code, Zip Code Extracted, City]
Index: []
Zip Code Match Rate: 100.00%
```

In [25]:
```
# Since both ZIP Code columns contained identical values, one column was removed to eliminate duplication.
data_2017_unpivoted = data_2017_unpivoted.drop(columns=["Zip Code Extracted", "Zip_Match"])
```

In [26]:
```
# Verified all event types and their respective counts against the summary document
data_2017_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[26]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	3172
1	End of Year Count	Ending Shelter Count	6272
2	Intake	Other: TNR / Protective Custody / Returns / Di...	13637
3	Intake	Owner Relinquished	37412
4	Intake	Owner Requested Euthanasia upon intake	5964
5	Intake	Stray	59827
6	Intake	Transfer in from Out of State organization	36789
7	Intake	Transfer in from another Colorado organization	13856
8	Outcome (Negative)	Died	1991
9	Outcome (Negative)	Missing / Stolen	121
10	Outcome (Negative)	Owner Requested Euthanasia	5911
11	Outcome (Negative)	Shelter Euthanasia	11078
12	Outcome (Positive)	Adoption	103969
13	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	4232
14	Outcome (Positive)	Returned To Owner (RTO)	26701
15	Outcome (Positive)	Transferred out to another Colorado organization	12072
16	Outcome (Positive)	Transferred to an Out of State organization	567
17	Start of Year Count	Beginning Foster Count	2794
18	Start of Year Count	Beginning Shelter Count	6407

In [27]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2017 = data_2017_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in
```

```
summary_2017.to_excel("summary_2017.xlsx", index=False)
```

In [28]:
```
# Verified that all the columns contain no null values
data_2017_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 17455 entries, 1 to 49222
Data columns (total 12 columns):
 #   Column          Non-Null Count  Dtype  
---  ------          --------------  -----  
 0   Facility Name   17455 non-null  object 
 1   Zip Code        17455 non-null  object 
 2   County          17455 non-null  object 
 3   Animal Count    17455 non-null  int64  
 4   Event Type      17455 non-null  object 
 5   Flow Type       17455 non-null  object 
 6   Reporting Year  17455 non-null  int64  
 7   Species         17455 non-null  object 
 8   Age Group       17455 non-null  object 
 9   Latitude        17455 non-null  float64
 10  Longitude       17455 non-null  float64
 11  City            17455 non-null  object 
dtypes: float64(2), int64(2), object(8)
memory usage: 1.7+ MB
```

In [29]:
```
data_2017_unpivoted.isnull().sum()
```

Out[29]:
```
Facility Name     0
Zip Code          0
County            0
Animal Count      0
Event Type        0
Flow Type         0
Reporting Year    0
Species           0
Age Group         0
Latitude          0
Longitude         0
City              0
dtype: int64
```

In [30]:
```
# Rearranging the columns
Cleaned_2017_Shelter_And_Rescue_Statistics_final = data_2017_unpivoted[["Facility Name", "County", "City", "Zip Code",
                                                                        "Latitude", "Longitude", "Reporting Year", "Species",
                                                                        "Age Group", "Flow Type", "Event Type", "Animal Count"
```

In [31]:
```
Cleaned_2017_Shelter_And_Rescue_Statistics_final.head()
```

Out[31]:

	Facility Name	County	City	Zip Code	Latitude	Longitude	Reporting Year	Species	Age Group	Flow Type	Event Type	Animal Count
1	FOUND NATION ANIMAL RESCUE	Jefferson County	ARVADA	80003	39.826820	-105.065270	2017	Dogs	Adult	Start of Year Count	Beginning Shelter Count	0
2	COLORADO WYOMING VIZSLA RESCUE GROUP INC	Jefferson County	Arvada	80007	39.839939	-105.186131	2017	Dogs	Adult	Start of Year Count	Beginning Shelter Count	3
3	Golden Retriever Rescue of The Rockies	Jefferson County	Arvada	80007	39.839939	-105.186131	2017	Dogs	Adult	Start of Year Count	Beginning Shelter Count	6
4	City of Aurora Animal Shelter	Adams County	Aurora	80011	39.741870	-104.799113	2017	Dogs	Adult	Start of Year Count	Beginning Shelter Count	50
5	Colorado St. Bernard Rescue, Inc.	Arapahoe County	Aurora	80011	39.741870	-104.799113	2017	Dogs	Adult	Start of Year Count	Beginning Shelter Count	0

In [32]:
```
Cleaned_2017_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2017_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2017_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2017_Shelter_And_Rescue_Statistics_final.xlsx
```

In [34]:
```
from IPython.display import FileLink
FileLink('Cleaned_2017_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[34]: Cleaned_2017_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:

Shelter And Rescue Statistics Data - 2018

In [1]:
```
# Importing all the required libraries
import numpy as np        # For numerical operations and array manipulation
import pandas as pd       # For data manipulation and analysis
```

In [2]:
```
# Loading the data
data_2018 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [3]:
```
data_2018.head()
```

Out[3]:

	Facility Name	County	Adult Dogs: In Shelter Count as of 1/1/2018	Adult Dogs: In Foster Care Count 1/1/2018	Adult Dogs: Stray	Adult Dogs: Owner Relinquished	Adult Dogs: Owner Requested Euthanasia upon intake	Adult Dogs: Transfer in from another Colorado organization	Adult Dogs: Transfer in from Out of State organization	Adult Dogs: Other: TNR/Protective Custody/Returns/Disaster Relief	...	R
0	2nd Chance Vizsla Rescue, Inc.	Larimer County	0.0	4.0	0.0	2.0	0.0	2.0	0.0	0.0	...	
1	4 Paws 4 Life Rescue	Douglas County	24.0	0.0	0.0	41.0	0.0	0.0	337.0	0.0	...	
2	9 Lives Rescue	El Paso County	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	
3	Acadiana Animal Aid	Huerfano County	82.0	10.0	100.0	110.0	0.0	0.0	565.0	9.0	...	
4	Adams County Animal Shelter & Adoption Center	Adams County	55.0	15.0	2482.0	740.0	206.0	44.0	2.0	251.0	...	

5 rows × 173 columns

In [4]:
```
data_2018.shape  # Total rows and columns (wide format)
```

Out[4]: (328, 173)

In [5]:
```
# Unpivoting the DataFrame from wide to long format

# Define ID (identifier variables) columns
id_vars = ["Facility Name", "County"]

# Unpivot the rest of the columns into two: "Metric" and "Value"
data_2018_unpivoted = data_2018.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2018_unpivoted = data_2018_unpivoted.dropna(subset=["Value"])
```

In [6]:
```
data_2018_unpivoted.head()
```

Out[6]:

	Facility Name	County	Metric	Value
0	2nd Chance Vizsla Rescue, Inc.	Larimer County	Adult Dogs: In Shelter Count as of 1/1/2018	0.0
1	4 Paws 4 Life Rescue	Douglas County	Adult Dogs: In Shelter Count as of 1/1/2018	24.0
3	Acadiana Animal Aid	Huerfano County	Adult Dogs: In Shelter Count as of 1/1/2018	82.0
4	Adams County Animal Shelter & Adoption Center	Adams County	Adult Dogs: In Shelter Count as of 1/1/2018	55.0
5	Adoptable Animal Rescue Force	Teller County	Adult Dogs: In Shelter Count as of 1/1/2018	0.0

In [7]:
```
data_2018_unpivoted.shape   # total rows and columns after unpivoting
```

Out[7]:
```
(20506, 4)
```

In [8]:
```
# Split Metric column into two at the first colon (:)
data_2018_unpivoted[["Animal Type", "Metric Type"]] = data_2018_unpivoted["Metric"].str.split(":", n=1, expand=True)

# Clean up whitespace
data_2018_unpivoted["Animal Type"] = data_2018_unpivoted["Animal Type"].str.strip()
data_2018_unpivoted["Metric Type"] = data_2018_unpivoted["Metric Type"].str.strip()
```

```
# Removing old 'Metric' column
data_2018_unpivoted.drop(["Metric"], axis=1, inplace=True)
```

In [9]:
```
data_2018_unpivoted.head()
```

Out[9]:

	Facility Name	County	Value	Animal Type	Metric Type
0	2nd Chance Vizsla Rescue, Inc.	Larimer County	0.0	Adult Dogs	In Shelter Count as of 1/1/2018
1	4 Paws 4 Life Rescue	Douglas County	24.0	Adult Dogs	In Shelter Count as of 1/1/2018
3	Acadiana Animal Aid	Huerfano County	82.0	Adult Dogs	In Shelter Count as of 1/1/2018
4	Adams County Animal Shelter & Adoption Center	Adams County	55.0	Adult Dogs	In Shelter Count as of 1/1/2018
5	Adoptable Animal Rescue Force	Teller County	0.0	Adult Dogs	In Shelter Count as of 1/1/2018

In [10]:
```
data_2018_unpivoted["Metric Type"].unique()
```

Out[10]:
```
array(['In Shelter Count as of  1/1/2018',
       'In Foster Care Count 1/1/2018', 'Stray', 'Owner Relinquished',
       'Owner Requested Euthanasia upon intake',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR/Protective Custody/Returns/Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other live outcomes (ie: TNR/SNR)', 'Died', 'Missing/Stolen',
       'Shelter Euthanasia', 'Owner Requested Euthanasia',
       'In Shelter Count as of 12/31/2018',
       'In Foster Care Count 12/31/2018',
       'In Shelter Count as of 1/1/2018',
       'Other: TNR/Protective Custody/Returns/Disaster relief',
       'Other live outcomes (ie TNR/SNR)', 'Missing / Stolen',
       'Other Live Outcomes (ie: TNR/SNR)',
       'In Foster Care Count as of 12/31/2018',
       'TNR/Protective Custody/Returns/Disaster Relief'], dtype=object)
```

In [11]:
```
# Normalizing category names inside Metric Type column

replace_map = {
    # Beginning Counts
    "In Shelter Count as of  1/1/2018": "Beginning Shelter Count",
    "In Shelter Count as of 1/1/2018": "Beginning Shelter Count",
    "In Foster Care Count 1/1/2018": "Beginning Foster Count",

    # Ending Counts
    "In Shelter Count as of 12/31/2018": "Ending Shelter Count",
    "In Foster Care Count 12/31/2018": "Ending Foster Count",
    "In Foster Care Count as of 12/31/2018": "Ending Foster Count",

    # TNR / Protective Custody variations
    "Other: TNR/Protective Custody/Returns/Disaster Relief": "Other: TNR / Protective Custody / Returns / Disaster Relief",
    "Other: TNR/Protective Custody/Returns/Disaster relief": "Other: TNR / Protective Custody / Returns / Disaster Relief",
    "TNR/Protective Custody/Returns/Disaster Relief": "Other: TNR / Protective Custody / Returns / Disaster Relief",

    # Other Live Outcomes variations
    "Other live outcomes (ie: TNR/SNR)": "Other Live Outcomes (ie: tnr / snr)",
    "Other live outcomes (ie TNR/SNR)": "Other Live Outcomes (ie: tnr / snr)",
    "Other Live Outcomes (ie: TNR/SNR)": "Other Live Outcomes (ie: tnr / snr)",

    # Missing variations
    "Missing/Stolen": "Missing / Stolen"
}

data_2018_unpivoted["Metric Type"] = data_2018_unpivoted["Metric Type"].replace(replace_map)
```

In [12]:
```
data_2018_unpivoted["Metric Type"].unique()  # after normalizing category names
```

Out[12]:
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished', 'Owner Requested Euthanasia upon intake',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died', 'Missing / Stolen',
       'Shelter Euthanasia', 'Owner Requested Euthanasia',
       'Ending Shelter Count', 'Ending Foster Count'], dtype=object)
```

In [13]:
```python
# Added a new column Flow Type to map and categorize records

# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",

    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Owner Requested Euthanasia upon intake": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
    "Missing / Stolen": "Outcome (Negative)",
```

```
    "Shelter Euthanasia": "Outcome (Negative)",
    "Owner Requested Euthanasia": "Outcome (Negative)"
}

# Apply mapping to create new column
data_2018_unpivoted["Flow Type"] = data_2018_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2018_unpivoted = data_2018_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [14]:
```
# Convert to integer numeric column
data_2018_unpivoted["Animal Count"] = data_2018_unpivoted["Animal Count"].astype("int64")
```

In [15]:
```
data_2018_unpivoted.head()
```

Out[15]:

	Facility Name	County	Animal Count	Animal Type	Event Type	Flow Type
0	2nd Chance Vizsla Rescue, Inc.	Larimer County	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	4 Paws 4 Life Rescue	Douglas County	24	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	Acadiana Animal Aid	Huerfano County	82	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	Adams County Animal Shelter & Adoption Center	Adams County	55	Adult Dogs	Beginning Shelter Count	Start of Year Count
5	Adoptable Animal Rescue Force	Teller County	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [16]:
```
# Adding Reporting year column to the whole table
data_2018_unpivoted["Reporting Year"] = 2018
```

In [17]:
```
# Adding Species column by using Animal Type column
data_2018_unpivoted["Species"] = (data_2018_unpivoted["Animal Type"]
                                  .str.replace("Adult", "", case=False)
```

```
                                    .str.replace("Juvenile", "", case=False)
                                    .str.strip()
                                   )
# Checking unique values
data_2018_unpivoted["Species"].unique()
```

Out[17]:
```
array(['Dogs', 'Cats', 'Birds', 'Small Mammal', 'Sm Mammal',
       'Reptiles & Amphibians', 'Rabbits', 'Other'], dtype=object)
```

In [18]:
```
# Standardize similar species names
species_replace = {"Sm Mammal": "Small Mammal"}

# Apply replacement
data_2018_unpivoted["Species"] = data_2018_unpivoted["Species"].replace(species_replace)

# Now, checking unique values
data_2018_unpivoted["Species"].unique()
```

Out[18]:
```
array(['Dogs', 'Cats', 'Birds', 'Small Mammal', 'Reptiles & Amphibians',
       'Rabbits', 'Other'], dtype=object)
```

In [19]:
```
# Adding Age group column by using Animal Type column
data_2018_unpivoted["Age Group"] = np.where(
    data_2018_unpivoted["Animal Type"].str.contains("Juvenile", case=False, na=False), "Juvenile",
    np.where(data_2018_unpivoted["Animal Type"].str.contains("Adult", case=False, na=False), "Adult",
             "Unknown")
)

# checking unique values
data_2018_unpivoted["Age Group"].unique()
```

Out[19]:
```
array(['Adult', 'Juvenile', 'Unknown'], dtype=object)
```

In [20]:
```
# Removed the 'Animal Type' column to avoid redundancy after creating separate 'Species' and 'Age Group' columns.
data_2018_unpivoted = data_2018_unpivoted.drop(columns=["Animal Type"], errors="ignore")
```

In [21]:
```
data_2018_unpivoted.head()
```

Out[21]:

	Facility Name	County	Animal Count	Event Type	Flow Type	Reporting Year	Species	Age Group
0	2nd Chance Vizsla Rescue, Inc.	Larimer County	0	Beginning Shelter Count	Start of Year Count	2018	Dogs	Adult
1	4 Paws 4 Life Rescue	Douglas County	24	Beginning Shelter Count	Start of Year Count	2018	Dogs	Adult
3	Acadiana Animal Aid	Huerfano County	82	Beginning Shelter Count	Start of Year Count	2018	Dogs	Adult
4	Adams County Animal Shelter & Adoption Center	Adams County	55	Beginning Shelter Count	Start of Year Count	2018	Dogs	Adult
5	Adoptable Animal Rescue Force	Teller County	0	Beginning Shelter Count	Start of Year Count	2018	Dogs	Adult

In [22]:
```
# Verified all event types and their respective counts against the summary document
data_2018_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[22]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	4283
1	End of Year Count	Ending Shelter Count	8152
2	Intake	Other: TNR / Protective Custody / Returns / Di...	15542
3	Intake	Owner Relinquished	40882
4	Intake	Owner Requested Euthanasia upon intake	6183
5	Intake	Stray	64774
6	Intake	Transfer in from Out of State organization	43032
7	Intake	Transfer in from another Colorado organization	13880
8	Outcome (Negative)	Died	2323
9	Outcome (Negative)	Missing / Stolen	100
10	Outcome (Negative)	Owner Requested Euthanasia	6096
11	Outcome (Negative)	Shelter Euthanasia	10380
12	Outcome (Positive)	Adoption	115019
13	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	5414
14	Outcome (Positive)	Returned To Owner (RTO)	28435
15	Outcome (Positive)	Transferred out to another Colorado organization	12461
16	Outcome (Positive)	Transferred to an Out of State organization	2095
17	Start of Year Count	Beginning Foster Count	3291
18	Start of Year Count	Beginning Shelter Count	7076

Note: In the 2018 summary document, the animal counts for 'Transferred out to another Colorado organization' and 'Owner Requested Euthanasia' do not match the result values above. However, verification against the dataset confirms that both values are correct.

In [23]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2018 = data_2018_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in

summary_2018.to_excel("summary_2018.xlsx", index=False)
```

In [24]:
```
# Verified that all the columns contain no null values
data_2018_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 20506 entries, 0 to 56085
Data columns (total 8 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Facility Name   20506 non-null  object
 1   County          20506 non-null  object
 2   Animal Count    20506 non-null  int64 
 3   Event Type      20506 non-null  object
 4   Flow Type       20506 non-null  object
 5   Reporting Year  20506 non-null  int64 
 6   Species         20506 non-null  object
 7   Age Group       20506 non-null  object
dtypes: int64(2), object(6)
memory usage: 1.4+ MB
```

In [25]:
```
data_2018_unpivoted.isnull().sum()
```

Out[25]:
```
Facility Name     0
County            0
Animal Count      0
Event Type        0
Flow Type         0
Reporting Year    0
Species           0
Age Group         0
dtype: int64
```

In [26]:
```
# Rearranging the columns
Cleaned_2018_Shelter_And_Rescue_Statistics_final = data_2018_unpivoted[["Facility Name", "County", "Reporting Year", "Species"
                                                                        "Age Group", "Flow Type", "Event Type", "Animal Count"
```

In [27]:
```
Cleaned_2018_Shelter_And_Rescue_Statistics_final.head()
```

Out[27]:

	Facility Name	County	Reporting Year	Species	Age Group	Flow Type	Event Type	Animal Count
0	2nd Chance Vizsla Rescue, Inc.	Larimer County	2018	Dogs	Adult	Start of Year Count	Beginning Shelter Count	0
1	4 Paws 4 Life Rescue	Douglas County	2018	Dogs	Adult	Start of Year Count	Beginning Shelter Count	24
3	Acadiana Animal Aid	Huerfano County	2018	Dogs	Adult	Start of Year Count	Beginning Shelter Count	82
4	Adams County Animal Shelter & Adoption Center	Adams County	2018	Dogs	Adult	Start of Year Count	Beginning Shelter Count	55
5	Adoptable Animal Rescue Force	Teller County	2018	Dogs	Adult	Start of Year Count	Beginning Shelter Count	0

In [28]:
```
Cleaned_2018_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2018_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2018_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2018_Shelter_And_Rescue_Statistics_final.xlsx
```

In [29]:
```
from IPython.display import FileLink
FileLink('Cleaned_2018_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[29]: Cleaned_2018_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:
```

```

Shelter And Rescue Statistics Data - 2019

In [33]:
```
# Importing all the required libraries
import numpy as np
import pandas as pd
```

In [34]:
```
# Loading the data
data_2019 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [35]:
```
data_2019.head()
```

Out[35]:

	Facility Name	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Dogs-Adult	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Dogs-Juvenile	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Cats-Adult	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Cats-Juvenile	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Birds	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Small Mammals	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Reptiles & Amphibians	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Rabbits	Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Other, Fish, Livestock, etc.	...	Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Other, Fish, Livestock, etc.	Endi Anim Statistic In Fost Count as 12/31/20 - Dog Adu
0	2 Blondes All Breed Rescue	59	30	0	0	0	0	0	0	0	...	0	
1	2nd Chance Vizsla Rescue, Inc.	0	0	0	0	0	0	0	0	0	...	0	
2	4 Paws 4 Life Rescue	26	16	6	3	0	0	0	0	0	...	0	
3	9 Lives Rescue	0	0	0	0	0	0	0	0	0	...	0	
4	Acadiana Animal Aid	17	44	76	9	0	0	0	0	0	...	0	

5 rows × 154 columns

In [36]:
```
data_2019.shape  # total rows and columns (in wide format)
```

Out[36]: (349, 154)

In [37]:
```
# Unpivoting the DataFrame from wide to long format

# Define ID (identifier variables) columns
id_vars = ["Facility Name"]

# Unpivot the rest of the columns into two: "Metric" and "Value"
data_2019_unpivoted = data_2019.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2019_unpivoted = data_2019_unpivoted.dropna(subset=["Value"])
```

In [38]:
```
data_2019_unpivoted.head()
```

Out[38]:

	Facility Name	Metric	Value
0	2 Blondes All Breed Rescue	Starting Animal Statistics - In Shelter Count ...	59
1	2nd Chance Vizsla Rescue, Inc.	Starting Animal Statistics - In Shelter Count ...	0
2	4 Paws 4 Life Rescue	Starting Animal Statistics - In Shelter Count ...	26
3	9 Lives Rescue	Starting Animal Statistics - In Shelter Count ...	0
4	Acadiana Animal Aid	Starting Animal Statistics - In Shelter Count ...	17

In [39]:
```
data_2019_unpivoted.shape   # total rows and columns after unpivoting (long format)
```

Out[39]: (53397, 3)

In [40]:
```
data_2019_unpivoted["Metric"].unique()
```

Out[40]:
```
array(['Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Dogs-Adult',
       'Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Dogs-Juvenile',
       'Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Cats-Adult',
       'Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Cats-Juvenile',
       'Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Birds',
       'Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Small Mammals',
       'Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Reptiles & Amphibians',
       'Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Rabbits',
       'Starting Animal Statistics - In Shelter Count as of 1/1/2019 - Other, Fish, Livestock, etc.',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Dogs-Adult',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Dogs-Juvenile',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Cats-Adult',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Cats-Juvenile',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Birds',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Small Mammals',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Reptiles & Amphibians',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Rabbits',
       'Starting Animal Statistics - In Foster Care Count as 1/1/2019 - Other, Fish, Livestock, etc.',
       'Animal Intake Statistics - Stray - Dogs-Adult',
       'Animal Intake Statistics - Stray - Dogs-Juvenile',
       'Animal Intake Statistics - Stray - Cats-Adult',
       'Animal Intake Statistics - Stray - Cats-Juvenile',
       'Animal Intake Statistics - Stray - Birds',
       'Animal Intake Statistics - Stray - Small Mammals',
       'Animal Intake Statistics - Stray - Reptiles & Amphibians',
       'Animal Intake Statistics - Stray - Rabbits',
       'Animal Intake Statistics - Stray - Other, Fish, Livestock, etc.',
       'Animal Intake Statistics - Owner Relinquished - Dogs-Adult',
       'Animal Intake Statistics - Owner Relinquished - Dogs-Juvenile',
       'Animal Intake Statistics - Owner Relinquished - Cats-Adult',
       'Animal Intake Statistics - Owner Relinquished - Cats-Juvenile',
       'Animal Intake Statistics - Owner Relinquished - Birds',
       'Animal Intake Statistics - Owner Relinquished - Small Mammals',
       'Animal Intake Statistics - Owner Relinquished - Reptiles & Amphibians',
       'Animal Intake Statistics - Owner Relinquished - Rabbits',
       'Animal Intake Statistics - Owner Relinquished - Other, Fish, Livestock, etc.',
       'Animal Intake Statistics - Transfer In from a Colorado Organization - Dogs-Adult',
       'Animal Intake Statistics - Transfer In from a Colorado Organization - Dogs-Juvenile',
       'Animal Intake Statistics - Transfer In from a Colorado Organization - Cats-Adult',
       'Animal Intake Statistics - Transfer In from a Colorado Organization - Cats-Juvenile',
```

```
            'Animal Intake Statistics - Transfer In from a Colorado Organization - Birds',
            'Animal Intake Statistics - Transfer In from a Colorado Organization - Small Mammals',
            'Animal Intake Statistics - Transfer In from a Colorado Organization - Reptiles & Amphibians',
            'Animal Intake Statistics - Transfer In from a Colorado Organization - Rabbits',
            'Animal Intake Statistics - Transfer In from a Colorado Organization - Other, Fish, Livestock, etc.',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Dogs-Adult',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Dogs-Juvenile',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Cats-Adult',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Cats-Juvenile',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Birds',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Small Mammals',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Reptiles & Amphibians',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Rabbits',
            'Animal Intake Statistics - Transfer In from an Out of State Organization - Other, Fish, Livestock, etc.',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Dogs-Adult',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Dogs-Juvenile',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Cats-Adult',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Cats-Juvenile',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Birds',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Small Mammals',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Reptiles & Amphibians',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Rabbits',
            'Animal Intake Statistics - Other; TNR/Protective Custody/Returns/Disaster Relief, etc. - Other, Fish, Livestock, et
c.',
            'Animal Outcome Statistics - Adoption - Dogs-Adult',
            'Animal Outcome Statistics - Adoption - Dogs-Juvenile',
            'Animal Outcome Statistics - Adoption - Cats-Adult',
            'Animal Outcome Statistics - Adoption - Cats-Juvenile',
            'Animal Outcome Statistics - Adoption - Birds',
            'Animal Outcome Statistics - Adoption - Small Mammals',
            'Animal Outcome Statistics - Adoption - Reptiles & Amphibians',
            'Animal Outcome Statistics - Adoption - Rabbits',
            'Animal Outcome Statistics - Adoption - Other, Fish, Livestock, etc.',
            'Animal Outcome Statistics - Return to Owner - Dogs-Adult',
            'Animal Outcome Statistics - Return to Owner - Dogs-Juvenile',
            'Animal Outcome Statistics - Return to Owner - Cats-Adult',
            'Animal Outcome Statistics - Return to Owner - Cats-Juvenile',
            'Animal Outcome Statistics - Return to Owner - Birds',
            'Animal Outcome Statistics - Return to Owner - Small Mammals',
            'Animal Outcome Statistics - Return to Owner - Reptiles & Amphibians',
            'Animal Outcome Statistics - Return to Owner - Rabbits',
```

```
'Animal Outcome Statistics - Return to Owner - Other, Fish, Livestock, etc.',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Dogs-Adult',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Dogs-Juvenile',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Cats-Adult',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Cats-Juvenile',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Birds',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Small Mammals',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Reptiles & Amphibians',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Rabbits',
'Animal Outcome Statistics - Transfer Out to a Colorado Organization - Other, Fish, Livestock, etc.',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Dogs-Adult',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Dogs-Juvenile',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Cats-Adult',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Cats-Juvenile',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Birds',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Small Mammals',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Reptiles & Amphibians',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Rabbits',
'Animal Outcome Statistics - Transfer Out to an Out of State Organization - Other, Fish, Livestock, etc.',
'Animal Outcome Statistics - Other Live Outcomes - Dogs-Adult',
'Animal Outcome Statistics - Other Live Outcomes - Dogs-Juvenile',
'Animal Outcome Statistics - Other Live Outcomes - Cats-Adult',
'Animal Outcome Statistics - Other Live Outcomes - Cats-Juvenile',
'Animal Outcome Statistics - Other Live Outcomes - Birds',
'Animal Outcome Statistics - Other Live Outcomes - Small Mammals',
'Animal Outcome Statistics - Other Live Outcomes - Reptiles & Amphibians',
'Animal Outcome Statistics - Other Live Outcomes - Rabbits',
'Animal Outcome Statistics - Other Live Outcomes - Other, Fish, Livestock, etc.',
'Animal Outcome Statistics - Deaths - Dogs-Adult',
'Animal Outcome Statistics - Deaths - Dogs-Juvenile',
'Animal Outcome Statistics - Deaths - Cats-Adult',
'Animal Outcome Statistics - Deaths - Cats-Juvenile',
'Animal Outcome Statistics - Deaths - Birds',
'Animal Outcome Statistics - Deaths - Small Mammals',
'Animal Outcome Statistics - Deaths - Reptiles & Amphibians',
'Animal Outcome Statistics - Deaths - Rabbits',
'Animal Outcome Statistics - Deaths - Other, Fish, Livestock, etc.',
'Animal Outcome Statistics - Missing/Stolen - Dogs-Adult',
'Animal Outcome Statistics - Missing/Stolen - Dogs-Juvenile',
'Animal Outcome Statistics - Missing/Stolen - Cats-Adult',
'Animal Outcome Statistics - Missing/Stolen - Cats-Juvenile',
```

```
       'Animal Outcome Statistics - Missing/Stolen - Birds',
       'Animal Outcome Statistics - Missing/Stolen - Small Mammals',
       'Animal Outcome Statistics - Missing/Stolen - Reptiles & Amphibians',
       'Animal Outcome Statistics - Missing/Stolen - Rabbits',
       'Animal Outcome Statistics - Missing/Stolen - Other, Fish, Livestock, etc.',
       'Animal Outcome Statistics - Euthanasia - Dogs-Adult',
       'Animal Outcome Statistics - Euthanasia - Dogs-Juvenile',
       'Animal Outcome Statistics - Euthanasia - Cats-Adult',
       'Animal Outcome Statistics - Euthanasia - Cats-Juvenile',
       'Animal Outcome Statistics - Euthanasia - Birds',
       'Animal Outcome Statistics - Euthanasia - Small Mammals',
       'Animal Outcome Statistics - Euthanasia - Reptiles & Amphibians',
       'Animal Outcome Statistics - Euthanasia - Rabbits',
       'Animal Outcome Statistics - Euthanasia - Other, Fish, Livestock, etc.',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Dogs-Adult',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Dogs-Juvenile',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Cats-Adult',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Cats-Juvenile',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Birds',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Small Mammals',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Reptiles & Amphibians',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Rabbits',
       'Ending Animal Statistics - In Shelter Count as of 12/31/2019 - Other, Fish, Livestock, etc.',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Dogs-Adult',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Dogs-Juvenile',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Cats-Adult',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Cats-Juvenile',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Birds',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Small Mammals',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Reptiles & Amphibians',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Rabbits',
       'Ending Animal Statistics - In Foster Count as of 12/31/2019 - Other, Fish, Livestock, etc.'],
      dtype=object)
```

In [41]:

```python
# Splitting Metric column into 3 parts
split_cols = data_2019_unpivoted["Metric"].str.split(" - ", n=2, expand=True)

# split the text at " - ", but only into 3 pieces (n=2 means max 2 splits)

# Assigning meaningful column names
data_2019_unpivoted["Section"] = split_cols[0]
```

```
data_2019_unpivoted["Metric Type"] = split_cols[1]
data_2019_unpivoted["Animal Raw"] = split_cols[2]

# Clean up whitespace
data_2019_unpivoted["Section"] = data_2019_unpivoted["Section"].str.strip()
data_2019_unpivoted["Metric Type"] = data_2019_unpivoted["Metric Type"].str.strip()
data_2019_unpivoted["Animal Raw"] = data_2019_unpivoted["Animal Raw"].str.strip()
```

In [42]:
```
data_2019_unpivoted.head()
```

Out[42]:

	Facility Name	Metric	Value	Section	Metric Type	Animal Raw
0	2 Blondes All Breed Rescue	Starting Animal Statistics - In Shelter Count ...	59	Starting Animal Statistics	In Shelter Count as of 1/1/2019	Dogs-Adult
1	2nd Chance Vizsla Rescue, Inc.	Starting Animal Statistics - In Shelter Count ...	0	Starting Animal Statistics	In Shelter Count as of 1/1/2019	Dogs-Adult
2	4 Paws 4 Life Rescue	Starting Animal Statistics - In Shelter Count ...	26	Starting Animal Statistics	In Shelter Count as of 1/1/2019	Dogs-Adult
3	9 Lives Rescue	Starting Animal Statistics - In Shelter Count ...	0	Starting Animal Statistics	In Shelter Count as of 1/1/2019	Dogs-Adult
4	Acadiana Animal Aid	Starting Animal Statistics - In Shelter Count ...	17	Starting Animal Statistics	In Shelter Count as of 1/1/2019	Dogs-Adult

In [43]:
```
data_2019_unpivoted["Metric Type"].unique()
```

Out[43]:
```
array(['In Shelter Count as of 1/1/2019',
       'In Foster Care Count as 1/1/2019', 'Stray', 'Owner Relinquished',
       'Transfer In from a Colorado Organization',
       'Transfer In from an Out of State Organization',
       'Other; TNR/Protective Custody/Returns/Disaster Relief, etc.',
       'Adoption', 'Return to Owner',
       'Transfer Out to a Colorado Organization',
       'Transfer Out to an Out of State Organization',
       'Other Live Outcomes', 'Deaths', 'Missing/Stolen', 'Euthanasia',
       'In Shelter Count as of 12/31/2019',
       'In Foster Count as of 12/31/2019'], dtype=object)
```

In [44]:
```
# Standardizing category names inside Metric Type column
metric_map = {
    "In Shelter Count as of 1/1/2019": "Beginning Shelter Count",
    "In Foster Care Count as 1/1/2019": "Beginning Foster Count",
    "Stray": "Stray",
    "Owner Relinquished": "Owner Relinquished",
    "Transfer In from a Colorado Organization": "Transfer in from another Colorado organization",
    "Transfer In from an Out of State Organization": "Transfer in from Out of State organization",
    "Other; TNR/Protective Custody/Returns/Disaster Relief, etc.": "Other: TNR / Protective Custody / Returns / Disaster Relie
    "Adoption": "Adoption",
    "Return to Owner": "Returned To Owner (RTO)",
    "Transfer Out to a Colorado Organization": "Transferred out to another Colorado organization",
    "Transfer Out to an Out of State Organization": "Transferred to an Out of State organization",
    "Other Live Outcomes": "Other Live Outcomes (ie: tnr / snr)",
    "Deaths": "Died",
    "Missing/Stolen": "Missing / Stolen",
    "Euthanasia": "Shelter Euthanasia",
    "In Shelter Count as of 12/31/2019": "Ending Shelter Count",
    "In Foster Count as of 12/31/2019": "Ending Foster Count"
}

# Apply mapping
data_2019_unpivoted["Metric Type"] = data_2019_unpivoted["Metric Type"].replace(metric_map)
```

In [45]:
```
data_2019_unpivoted["Metric Type"].unique()
```

Out[45]:
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died', 'Missing / Stolen',
       'Shelter Euthanasia', 'Ending Shelter Count',
       'Ending Foster Count'], dtype=object)
```

In [46]:
```
# Added a new column Flow Type to map and categorize records

# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",

    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
    "Missing / Stolen": "Outcome (Negative)",
    "Shelter Euthanasia": "Outcome (Negative)"
```

```
}

# Apply mapping to create new column
data_2019_unpivoted["Flow Type"] = data_2019_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2019_unpivoted = data_2019_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [47]:
```
# Convert to integer numeric column
data_2019_unpivoted["Animal Count"] = data_2019_unpivoted["Animal Count"].astype("int64")
```

In [48]:
```
data_2019_unpivoted.head()
```

Out[48]:

	Facility Name	Metric	Animal Count	Section	Event Type	Animal Raw	Flow Type
0	2 Blondes All Breed Rescue	Starting Animal Statistics - In Shelter Count ...	59	Starting Animal Statistics	Beginning Shelter Count	Dogs-Adult	Start of Year Count
1	2nd Chance Vizsla Rescue, Inc.	Starting Animal Statistics - In Shelter Count ...	0	Starting Animal Statistics	Beginning Shelter Count	Dogs-Adult	Start of Year Count
2	4 Paws 4 Life Rescue	Starting Animal Statistics - In Shelter Count ...	26	Starting Animal Statistics	Beginning Shelter Count	Dogs-Adult	Start of Year Count
3	9 Lives Rescue	Starting Animal Statistics - In Shelter Count ...	0	Starting Animal Statistics	Beginning Shelter Count	Dogs-Adult	Start of Year Count
4	Acadiana Animal Aid	Starting Animal Statistics - In Shelter Count ...	17	Starting Animal Statistics	Beginning Shelter Count	Dogs-Adult	Start of Year Count

In [49]:
```
# Removed the 'Metric' & 'Section' columns to avoid redundancy.
data_2019_unpivoted = data_2019_unpivoted.drop(columns=["Metric", "Section"], errors="ignore")
```

In [50]:
```
data_2019_unpivoted.head()
```

Out[50]:

	Facility Name	Animal Count	Event Type	Animal Raw	Flow Type
0	2 Blondes All Breed Rescue	59	Beginning Shelter Count	Dogs-Adult	Start of Year Count
1	2nd Chance Vizsla Rescue, Inc.	0	Beginning Shelter Count	Dogs-Adult	Start of Year Count
2	4 Paws 4 Life Rescue	26	Beginning Shelter Count	Dogs-Adult	Start of Year Count
3	9 Lives Rescue	0	Beginning Shelter Count	Dogs-Adult	Start of Year Count
4	Acadiana Animal Aid	17	Beginning Shelter Count	Dogs-Adult	Start of Year Count

In [51]:
```
# Splitting Animal Raw into Species & Age Group
animal_split = data_2019_unpivoted["Animal Raw"].str.split("-", n=1, expand=True)

# Assigning meaningful column names and Clean up extra whitespace
data_2019_unpivoted["Species"] = animal_split[0].str.strip()
data_2019_unpivoted["Age Group"] = animal_split[1].fillna("Unknown").str.strip()

# Checking unique values
print(data_2019_unpivoted["Species"].unique())
print(data_2019_unpivoted["Age Group"].unique())
```

```
['Dogs' 'Cats' 'Birds' 'Small Mammals' 'Reptiles & Amphibians' 'Rabbits'
 'Other, Fish, Livestock, etc.']
['Adult' 'Juvenile' 'Unknown']
```

In [52]:
```
# Standardize similar species names
species_replace = {"Other, Fish, Livestock, etc.": "Other"}

# Apply replacement
data_2019_unpivoted["Species"] = data_2019_unpivoted["Species"].replace(species_replace)

# Now, checking unique values
data_2019_unpivoted["Species"].unique()
```

Out[52]:
```
array(['Dogs', 'Cats', 'Birds', 'Small Mammals', 'Reptiles & Amphibians',
       'Rabbits', 'Other'], dtype=object)
```

In [53]:
```
# Removed the 'Animal Raw' column to avoid redundancy after creating separate 'Species' and 'Age Group' columns.
```

```
data_2019_unpivoted = data_2019_unpivoted.drop(columns=["Animal Raw"], errors="ignore")
```

In [54]:
```
# Adding Reporting year column to the whole table
data_2019_unpivoted["Reporting Year"] = 2019
```

In [55]:
```
data_2019_unpivoted.head()
```

Out[55]:

	Facility Name	Animal Count	Event Type	Flow Type	Species	Age Group	Reporting Year
0	2 Blondes All Breed Rescue	59	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2019
1	2nd Chance Vizsla Rescue, Inc.	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2019
2	4 Paws 4 Life Rescue	26	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2019
3	9 Lives Rescue	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2019
4	Acadiana Animal Aid	17	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2019

In [56]:
```
data_2019_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[56]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	5008
1	End of Year Count	Ending Shelter Count	8787
2	Intake	Other: TNR / Protective Custody / Returns / Di...	19124
3	Intake	Owner Relinquished	43594
4	Intake	Stray	58990
5	Intake	Transfer in from Out of State organization	45680
6	Intake	Transfer in from another Colorado organization	12306
7	Outcome (Negative)	Died	2317
8	Outcome (Negative)	Missing / Stolen	134
9	Outcome (Negative)	Shelter Euthanasia	10823
10	Outcome (Positive)	Adoption	118446
11	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	5598
12	Outcome (Positive)	Returned To Owner (RTO)	26418
13	Outcome (Positive)	Transferred out to another Colorado organization	12167
14	Outcome (Positive)	Transferred to an Out of State organization	2671
15	Start of Year Count	Beginning Foster Count	4513
16	Start of Year Count	Beginning Shelter Count	9485

Note: In the 2019 summary document, the animal counts also do not match the result values above. However, verification against the dataset confirms that these result values are correct. The discrepancy is due to some numbers in the data file being formatted as text.

In [57]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2019 = data_2019_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in
```

```
summary_2019.to_excel("summary_2019.xlsx", index=False)
```

In [58]:
```
# Verified that all the columns contain no null values
data_2019_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 53397 entries, 0 to 53396
Data columns (total 7 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Facility Name   53397 non-null  object
 1   Animal Count    53397 non-null  int64 
 2   Event Type      53397 non-null  object
 3   Flow Type       53397 non-null  object
 4   Species         53397 non-null  object
 5   Age Group       53397 non-null  object
 6   Reporting Year  53397 non-null  int64 
dtypes: int64(2), object(5)
memory usage: 2.9+ MB
```

In [59]:
```
data_2019_unpivoted.isnull().sum()
```

Out[59]:
```
Facility Name     0
Animal Count      0
Event Type        0
Flow Type         0
Species           0
Age Group         0
Reporting Year    0
dtype: int64
```

In [60]:
```
# Rearranging the columns
Cleaned_2019_Shelter_And_Rescue_Statistics_final = data_2019_unpivoted[["Facility Name", "Reporting Year", "Species",
                                                                        "Age Group", "Flow Type", "Event Type", "Animal Count"
```

In [61]:
```
Cleaned_2019_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2019_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2019_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2019_Shelter_And_Rescue_Statistics_final.xlsx
```

In [62]:
```
from IPython.display import FileLink
FileLink('Cleaned_2019_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[62]: Cleaned_2019_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:
```

```

Shelter And Rescue Statistics Data - 2020

In [1]:
```
# Importing all the required libraries
import numpy as np
import pandas as pd
```

In [2]:
```
# Loading the data
data_2020 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [3]:
```
data_2020.head()
```

Out[3]:

	Facility Name	1/1/2020\n Adult Dogs\n In Shelter	1/1/2020\n Adult Dogs\n In Foster Care	2020\n Adult Dogs\n Stray	2020\n Adult Dogs\n Owner Relinquished	2020\n Adult Dogs \n Transfer from another Colorado Organization	2020\n Adult Dogs\n Transfer from Out of State	2020\n Adult Dogs\n Other	2020\n Adult Dogs\n Adoption	2020\n Adult Dogs\n Returned to Owner (RTO)	...	2020\n Other\n Adoption	2020\n Other\n Returned to Owner (RTO)
0	2 Blondes All Breed Rescue, Inc.	78	47	0	0	0	379	40	498	0	...	0	0
1	2nd Chance Vizsla Rescue, Inc.	0	0	0	0	0	0	0	0	0	...	0	0
2	4 Paws 4 Life Rescue	5	0	0	11	0	454	0	459	0	...	0	0
3	9 Lives Rescue	0	0	0	0	0	0	0	0	0	...	0	0
4	A Friend of Jack Rescue	0	0	0	1	0	181	0	175	0	...	0	0

5 rows × 154 columns

In [4]:
```
data_2020.shape
```

Out[4]: (356, 154)

In [5]:
```
# Unpivoting the DataFrame from wide to long format

# Define ID (identifier variables) columns
id_vars = ["Facility Name"]

# Unpivot the rest of the columns into two: "Metric" and "Value"
data_2020_unpivoted = data_2020.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2020_unpivoted = data_2020_unpivoted.dropna(subset=["Value"])
```

In [6]:
```
data_2020_unpivoted.head()
```

Out[6]:

	Facility Name	Metric	Value
0	2 Blondes All Breed Rescue, Inc.	1/1/2020\n Adult Dogs\n In Shelter	78
1	2nd Chance Vizsla Rescue, Inc.	1/1/2020\n Adult Dogs\n In Shelter	0
2	4 Paws 4 Life Rescue	1/1/2020\n Adult Dogs\n In Shelter	5
3	9 Lives Rescue	1/1/2020\n Adult Dogs\n In Shelter	0
4	A Friend of Jack Rescue	1/1/2020\n Adult Dogs\n In Shelter	0

In [7]:
```
data_2020_unpivoted.shape   # total rows and columns after unpivoting
```

Out[7]:
```
(54468, 3)
```

In [8]:
```
data_2020_unpivoted["Metric"].unique()
```

Out[8]:
```
array(['1/1/2020\n Adult Dogs\n In Shelter',
       '1/1/2020\n Adult Dogs\n In Foster Care',
       '2020\n Adult Dogs\n Stray',
       '2020\n Adult Dogs\n Owner Relinquished',
       '2020\n Adult Dogs \n Transfer from another Colorado Organization',
       '2020\n Adult Dogs\n Transfer from Out of State',
       '2020\n Adult Dogs\n Other', '2020\n Adult Dogs\n Adoption',
       '2020\n Adult Dogs\n Returned to Owner (RTO)',
       '2020\n Adult Dogs\n Transfer to another Colorado Organization',
       '2020\n Adult Dogs\n Transfer to Out of State',
       '2020\n Adult Dogs\n Other.1', '2020\n Adult Dogs\n Deaths',
       '2020\n Adult Dogs\n Euthanasia',
       '2020\n Adult Dogs\n Missing/Stolen',
       '12/31/2020\n Adult Dogs\n In Shelter',
       '12/31/2020\n Adult Dogs\n In Foster Care',
       '1/1/2020\n Juvenile Dogs\n In Shelter',
       '1/1/2020\n Juvenile Dogs\n In Foster Care',
       '2020\n Juvenile Dogs\n Stray',
       '2020\n Juvenile Dogs\n Owner Relinquished',
       '2020\n Juvenile Dogs\n Transfer from another Colorado Organization',
       '2020\n Juvenile Dogs\n Transfer from Out of State',
       '2020\n Juvenile Dogs\n Other', '2020\n Juvenile Dogs\n Adoption',
       '2020\n Juvenile Dogs\n Returned to Owner (RTO)',
       '2020\n Juvenile Dogs\n Transfer to another Colorado Organization',
       '2020\n Juvenile Dogs\n Transfer to Out of State',
       '2020\n Juvenile Dogs\n Other.1', '2020\n Juvenile Dogs\n Deaths',
       '2020\n Juvenile Dogs\n Euthanasia',
       '2020\n Juvenile Dogs\n Missing/Stolen',
       '12/31/2020\n Juvenile Dogs\n In Shelter',
       '12/31/2020\n Juvenile Dogs\n In Foster Care',
       '1/1/2020\n Adult Cats\n In Shelter',
       '1/1/2020\n Adult Cats\n In Foster Care',
       '2020\n Adult Cats\n Stray',
       '2020\n Adult Cats\n Owner Relinquished',
       '2020\n Adult Cats\n Transfer from another Colorado Organization',
       '2020\n Adult Cats\n Transfer from Out of State',
       '2020\n Adult Cats\n Other', '2020\n Adult Cats\n Adoption',
       '2020\n Adult Cats\n Returned to Owner (RTO)',
       '2020\n Adult Cats\n Transfer to another Colorado Organization',
       '2020\n Adult Cats\n Transfer to Out of State',
```

```
'2020\n Adult Cats\n Other.1', '2020\n Adult Cats\n Deaths',
'2020\n Adult Cats\n Euthanasia',
'2020\n Adult Cats\n Missing/Stolen',
'12/31/2020\n Adult Cats\n In Shelter',
'12/31/2020\n Adult Cats\n In Foster Care',
'1/1/2020\n Juvenile Cats\n In Shelter',
'1/1/2020\n Juvenile Cats\n In Foster Care',
'2020\n Juvenile Cats\n Stray',
'2020\n Juvenile Cats\n Owner Relinquished',
'2020\n Juvenile Cats\n Transfer from another Colorado Organization',
'2020\n Juvenile Cats\n Transfer from Out of State',
'2020\n Juvenile Cats\n Other', '2020\n Juvenile Cats\n Adoption',
'2020\n Juvenile Cats\n Returned to Owner (RTO)',
'2020\n Juvenile Cats\n Transfer to another Colorado Organization',
'2020\n Juvenile Cats\n Transfer to Out of State',
'2020\n Juvenile Cats\n Other.1', '2020\n Juvenile Cats\n Deaths',
'2020\n Juvenile Cats\n Euthanasia',
'2020\n Juvenile Cats\n Missing/Stolen',
'12/31/2020\n Juvenile Cats\n In Shelter',
'12/31/2020\n Juvenile Cats\n In Foster Care',
'1/1/2020\n Birds\n In Shelter',
'1/1/2020\n Birds\n In Foster Care', '2020\n Birds\n Stray',
'2020\n Birds\n Owner Relinquished',
'2020\n Birds\n Transfer from another Colorado Organization',
'2020\n Birds\n Transfer from Out of State',
'2020\n Birds\n Other', '2020\n Birds\n Adoption',
'2020\n Birds\n Returned to Owner (RTO)',
'2020\n Birds\n Transfer to another Colorado Organization',
'2020\n Birds\n Transfer to Out of State',
'2020\n Birds\n Other.1', '2020\n Birds\n Deaths',
'2020\n Birds\n Euthanasia', '2020\n Birds\n Missing/Stolen',
'12/31/2020\n Birds\n In Shelter',
'12/31/2020\n Birds\n In Foster Care',
'1/1/2020\n Small Mammals\n In Shelter',
'1/1/2020\n Small Mammals\n In Foster Care',
'2020\n Small Mammals\n Stray',
'2020\n Small Mammals\n Owner Relinquished',
'2020\n Small Mammals\n Transfer from another Colorado Organization',
'2020\n Small Mammals\n Transfer from Out of State',
'2020\n Small Mammals\n Other', '2020\n Small Mammals\n Adoption',
'2020\n Small Mammals\n Returned to Owner (RTO)',
```

'2020\n Small Mammals\n Transfer to another Colorado Organization',
'2020\n Small Mammals\n Transfer to Out of State',
'2020\n Small Mammals\n Other.1', '2020\n Small Mammals\n Deaths',
'2020\n Small Mammals\n Euthanasia',
'2020\n Small Mammals\n Missing/Stolen',
'12/31/2020\n Small Mammals\n In Shelter',
'12/31/2020\n Small Mammals\n In Foster Care',
'1/1/2020\n Reptiles & Amphibians\n In Shelter',
'1/1/2020\n Reptiles & Amphibians\n In Foster Care',
'2020\n Reptiles & Amphibians\n Stray',
'2020\n Reptiles & Amphibians\n Owner Relinquished',
'2020\n Reptiles & Amphibians\n Transfer from another Colorado Organization',
'2020\n Reptiles & Amphibians\n Transfer from Out of State',
'2020\n Reptiles & Amphibians\n Other',
'2020\n Reptiles & Amphibians\n Adoption',
'2020\n Reptiles & Amphibians\n Returned to Owner (RTO)',
'2020\n Reptiles & Amphibians\n Transfer to another Colorado Organization',
'2020\n Reptiles & Amphibians\n Transfer to Out of State',
'2020\n Reptiles & Amphibians\n Other.1',
'2020\n Reptiles & Amphibians\n Deaths',
'2020\n Reptiles & Amphibians\n Euthanasia',
'2020\n Reptiles & Amphibians\n Missing/Stolen',
'12/31/2020\n Reptiles & Amphibians\n In Shelter',
'12/31/2020\n Reptiles & Amphibians\n In Foster Care',
'1/1/2020\n Rabbits\n In Shelter',
'1/1/2020\n Rabbits\n In Foster Care', '2020\n Rabbits\n Stray',
'2020\n Rabbits\n Owner Relinquished',
'2020\n Rabbits\n Transfer from another Colorado Organization',
'2020\n Rabbits\n Transfer from Out of State',
'2020\n Rabbits\n Other', '2020\n Rabbits\n Adoption',
'2020\n Rabbits\n Returned to Owner (RTO)',
'2020\n Rabbits\n Transfer to another Colorado Organization',
'2020\n Rabbits\n Transfer to Out of State',
'2020\n Rabbits\n Other.1', '2020\n Rabbits\n Deaths',
'2020\n Rabbits\n Euthanasia', '2020\n Rabbits\n Missing/Stolen',
'12/31/2020\n Rabbits\n In Shelter',
'12/31/2020\n Rabbits\n In Foster Care',
'1/1/2020\n Other\n In Shelter',
'1/1/2020\n Other\n In Foster Care', '2020\n Other\n Stray',
'2020\n Other\n Owner Relinquished',
'2020\n Other\n Transfer from another Colorado Organization',

```
       '2020\n Other\n Transfer from Out of State',
       '2020\n Other\n Other', '2020\n Other\n Adoption',
       '2020\n Other\n Returned to Owner (RTO)',
       '2020\n Other\n Transfer to another Colorado Organization',
       '2020\n Other\n Transfer to Out of State',
       '2020\n Other\n Other.1', '2020\n Other\n Deaths',
       '2020\n Other\n Euthanasia', '2020\n Other\n Missing/Stolen',
       '12/31/2020\n Other\n In Shelter',
       '12/31/2020\n Other\n In Foster Care'], dtype=object)
```

In [9]:
```
# Splitting 'Metric' column by newline character to create two columns
split_cols = data_2020_unpivoted["Metric"].str.split("\n", expand=True)

# Animal Type will always be the second part (index 1)
data_2020_unpivoted["Animal Type"] = split_cols[1].str.strip()

# Metric Type will be first part (date) + third part (metric)
data_2020_unpivoted["Metric Type"] = (split_cols[0].str.strip() + " " + split_cols[2].str.strip())
```

In [10]:
```
data_2020_unpivoted.head()
```

Out[10]:

	Facility Name	Metric	Value	Animal Type	Metric Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2020\n Adult Dogs\n In Shelter	78	Adult Dogs	1/1/2020 In Shelter
1	2nd Chance Vizsla Rescue, Inc.	1/1/2020\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2020 In Shelter
2	4 Paws 4 Life Rescue	1/1/2020\n Adult Dogs\n In Shelter	5	Adult Dogs	1/1/2020 In Shelter
3	9 Lives Rescue	1/1/2020\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2020 In Shelter
4	A Friend of Jack Rescue	1/1/2020\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2020 In Shelter

In [11]:
```
# checking unique values after split
data_2020_unpivoted["Metric Type"].unique()
```

Out[11]:
```
array(['1/1/2020 In Shelter', '1/1/2020 In Foster Care', '2020 Stray',
       '2020 Owner Relinquished',
       '2020 Transfer from another Colorado Organization',
       '2020 Transfer from Out of State', '2020 Other', '2020 Adoption',
       '2020 Returned to Owner (RTO)',
       '2020 Transfer to another Colorado Organization',
       '2020 Transfer to Out of State', '2020 Other.1', '2020 Deaths',
       '2020 Euthanasia', '2020 Missing/Stolen', '12/31/2020 In Shelter',
       '12/31/2020 In Foster Care'], dtype=object)
```

In [12]:
```
# Standardizing category names inside Metric Type column

metric_map = {
    # Start-of-year counts
    "1/1/2020 In Shelter": "Beginning Shelter Count",
    "1/1/2020 In Foster Care": "Beginning Foster Count",

    # Intakes
    "2020 Stray": "Stray",
    "2020 Owner Relinquished": "Owner Relinquished",
    "2020 Transfer from another Colorado Organization": "Transfer in from another Colorado organization",
    "2020 Transfer from Out of State": "Transfer in from Out of State organization",
    "2020 Other": "Other: TNR / Protective Custody / Returns / Disaster Relief",

    # Outcomes
    "2020 Adoption": "Adoption",
    "2020 Returned to Owner (RTO)": "Returned To Owner (RTO)",
    "2020 Transfer to another Colorado Organization": "Transferred out to another Colorado organization",
    "2020 Transfer to Out of State": "Transferred to an Out of State organization",
    "2020 Other.1": "Other Live Outcomes (ie: tnr / snr)",
    "2020 Deaths": "Died",
    "2020 Euthanasia": "Shelter Euthanasia",
    "2020 Missing/Stolen": "Missing / Stolen",

    # End-of-year counts
    "12/31/2020 In Shelter": "Ending Shelter Count",
    "12/31/2020 In Foster Care": "Ending Foster Count"
}
```

```
# Applying mapping
data_2020_unpivoted["Metric Type"] = data_2020_unpivoted["Metric Type"].replace(metric_map)
```

In [13]:
```
# Checking unique values after normalizing
data_2020_unpivoted["Metric Type"].unique()
```

Out[13]:
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died',
       'Shelter Euthanasia', 'Missing / Stolen', 'Ending Shelter Count',
       'Ending Foster Count'], dtype=object)
```

In [14]:
```
# Added a new column Flow Type to map and categorize records

# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",

    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
```

```
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
    "Missing / Stolen": "Outcome (Negative)",
    "Shelter Euthanasia": "Outcome (Negative)"
}

# Apply mapping to create new column
data_2020_unpivoted["Flow Type"] = data_2020_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2020_unpivoted = data_2020_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [15]:
```
data_2020_unpivoted.head()
```

Out[15]:

	Facility Name	Metric	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2020\n Adult Dogs\n In Shelter	78	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	2nd Chance Vizsla Rescue, Inc.	1/1/2020\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	4 Paws 4 Life Rescue	1/1/2020\n Adult Dogs\n In Shelter	5	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	9 Lives Rescue	1/1/2020\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	A Friend of Jack Rescue	1/1/2020\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [16]:
```
# Removing 'Metric' column to avoid redundancy.
data_2020_unpivoted = data_2020_unpivoted.drop(columns=["Metric"], errors="ignore")
```

In [17]:
```
data_2020_unpivoted.head()
```

Out[17]:

	Facility Name	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	78	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	2nd Chance Vizsla Rescue, Inc.	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	4 Paws 4 Life Rescue	5	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	9 Lives Rescue	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	A Friend of Jack Rescue	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [18]:
```
# Checking unique values of Animal Type column
data_2020_unpivoted["Animal Type"].unique()
```

Out[18]:
```
array(['Adult Dogs', 'Juvenile Dogs', 'Adult Cats', 'Juvenile Cats',
       'Birds', 'Small Mammals', 'Reptiles & Amphibians', 'Rabbits',
       'Other'], dtype=object)
```

In [19]:
```
# Creating two columns by splitting the 'Animal Type' column

# Species column
data_2020_unpivoted["Species"] = (data_2020_unpivoted["Animal Type"]
                                  .str.replace("Adult", "", case=False)
                                  .str.replace("Juvenile", "", case=False)
                                  .str.strip())

# Age Group column
data_2020_unpivoted["Age Group"] = np.where(
    data_2020_unpivoted["Animal Type"].str.contains("Adult", case=False, na=False), "Adult",
    np.where(data_2020_unpivoted["Animal Type"].str.contains("Juvenile", case=False, na=False), "Juvenile", "Unknown")
)

# Checking unique values
print(data_2020_unpivoted["Species"].unique())
print(data_2020_unpivoted["Age Group"].unique())
```

```
['Dogs' 'Cats' 'Birds' 'Small Mammals' 'Reptiles & Amphibians' 'Rabbits'
 'Other']
['Adult' 'Juvenile' 'Unknown']
```

In [20]:
```
# Removing 'Animal Type' column to avoid redundancy.
data_2020_unpivoted = data_2020_unpivoted.drop(columns=["Animal Type"], errors="ignore")
```

In [21]:
```
# Adding Reporting year column to the whole table
data_2020_unpivoted["Reporting Year"] = 2020
```

In [22]:
```
data_2020_unpivoted.head()
```

Out[22]:

	Facility Name	Animal Count	Event Type	Flow Type	Species	Age Group	Reporting Year
0	2 Blondes All Breed Rescue, Inc.	78	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2020
1	2nd Chance Vizsla Rescue, Inc.	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2020
2	4 Paws 4 Life Rescue	5	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2020
3	9 Lives Rescue	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2020
4	A Friend of Jack Rescue	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2020

In [23]:
```
# Checking Animal Count by Flow Type & Event Type
data_2020_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[23]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	5723
1	End of Year Count	Ending Shelter Count	7716
2	Intake	Other: TNR / Protective Custody / Returns / Di...	15165
3	Intake	Owner Relinquished	37459
4	Intake	Stray	50584
5	Intake	Transfer in from Out of State organization	50359
6	Intake	Transfer in from another Colorado organization	12188
7	Outcome (Negative)	Died	1808
8	Outcome (Negative)	Missing / Stolen	80
9	Outcome (Negative)	Shelter Euthanasia	7927
10	Outcome (Positive)	Adoption	117655
11	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	6127
12	Outcome (Positive)	Returned To Owner (RTO)	21196
13	Outcome (Positive)	Transferred out to another Colorado organization	10557
14	Outcome (Positive)	Transferred to an Out of State organization	1008
15	Start of Year Count	Beginning Foster Count	5357
16	Start of Year Count	Beginning Shelter Count	8685

Note: In the 2020 summary document, the animal counts do not match the result values above. However, verification against the dataset confirms that these result values are correct. The discrepancy is due to some numbers in the original data file being formatted as text.

In [24]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2020 = data_2020_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in
```

```
summary_2020.to_excel("summary_2020.xlsx", index=False)
```

In [25]:
```
# Verified that all the columns contain no null values
data_2020_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 54468 entries, 0 to 54467
Data columns (total 7 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Facility Name   54468 non-null  object
 1   Animal Count    54468 non-null  int64 
 2   Event Type      54468 non-null  object
 3   Flow Type       54468 non-null  object
 4   Species         54468 non-null  object
 5   Age Group       54468 non-null  object
 6   Reporting Year  54468 non-null  int64 
dtypes: int64(2), object(5)
memory usage: 2.9+ MB
```

In [26]:
```
data_2020_unpivoted.isnull().sum()
```

Out[26]:
```
Facility Name     0
Animal Count      0
Event Type        0
Flow Type         0
Species           0
Age Group         0
Reporting Year    0
dtype: int64
```

In [27]:
```
# Rearranging the columns
Cleaned_2020_Shelter_And_Rescue_Statistics_final = data_2020_unpivoted[["Facility Name", "Reporting Year", "Species",
                                                                        "Age Group", "Flow Type", "Event Type", "Animal Count"
```

In [28]:
```
Cleaned_2020_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2020_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2020_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2020_Shelter_And_Rescue_Statistics_final.xlsx
```

In [29]:
```
from IPython.display import FileLink
FileLink('Cleaned_2020_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[29]: Cleaned_2020_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:
```
```

Shelter And Rescue Statistics Data - 2021

In [1]:
```
# Importing all the required libraries
import numpy as np
import pandas as pd
```

In [2]:
```
# Loading the data
data_2021 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [3]:
```
data_2021.head()
```

Out[3]:

	Facility Name	1/1/2021\n Adult Dogs\n In Shelter	1/1/2021\n Adult Dogs\n In Foster Care	2021\n Adult Dogs\n Stray	2021\n Adult Dogs\n Owner Relinquished	2021\n Adult Dogs\n Transfer from another Colorado Organization	2021\n Adult Dogs\n Transfer from Out of State	2021\n Adult Dogs\n Other	2021\n Adult Dogs\n Adoption	2021\n Adult Dogs\n Returned to Owner (RTO)	...	2021\n Other\n Adoption	2021\n Other\n Returned to Owner (RTO)
0	2 Blondes All Breed Rescue, Inc.	26	14	0	5	4	481	72	530	0	...	0	0
1	2nd Chance Vizsla Rescue, Inc	0	7	0	0	0	0	0	0	0	...	0	0
2	4 Paws 4 Life Rescue	91	9	0	10	13	215	0	292	0	...	0	0
3	9 Lives Rescue	0	0	0	0	0	0	0	0	0	...	0	0
4	A Cat Rescue Out in the Sticks	0	0	0	0	0	0	0	0	0	...	0	0

5 rows × 154 columns

In [4]:
```
data_2021.shape
```

Out[4]: (354, 154)

In [5]:
```
# Unpivoting the DataFrame from wide to long format

# Define ID (identifier variables) columns
id_vars = ["Facility Name"]

# Unpivot the rest of the columns into two: "Metric" and "Value"
data_2021_unpivoted = data_2021.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2021_unpivoted = data_2021_unpivoted.dropna(subset=["Value"])
```

In [6]:
```
data_2021_unpivoted.head()
```

Out[6]:

	Facility Name	Metric	Value
0	2 Blondes All Breed Rescue, Inc.	1/1/2021\n Adult Dogs\n In Shelter	26
1	2nd Chance Vizsla Rescue, Inc	1/1/2021\n Adult Dogs\n In Shelter	0
2	4 Paws 4 Life Rescue	1/1/2021\n Adult Dogs\n In Shelter	91
3	9 Lives Rescue	1/1/2021\n Adult Dogs\n In Shelter	0
4	A Cat Rescue Out in the Sticks	1/1/2021\n Adult Dogs\n In Shelter	0

In [7]:
```
data_2021_unpivoted.shape
```

Out[7]: (54162, 3)

In [8]:
```
data_2021_unpivoted["Metric"].unique()
```

Out[8]:
```
array(['1/1/2021\n Adult Dogs\n In Shelter',
       '1/1/2021\n Adult Dogs\n In Foster Care',
       '2021\n Adult Dogs\n Stray',
       '2021\n Adult Dogs\n Owner Relinquished',
       '2021\n Adult Dogs\n Transfer from another Colorado Organization',
       '2021\n Adult Dogs\n Transfer from Out of State',
       '2021\n Adult Dogs\n Other', '2021\n Adult Dogs\n Adoption',
       '2021\n Adult Dogs\n Returned to Owner (RTO)',
       '2021\n Adult Dogs\n Transfer to another Colorado Organization',
       '2021\n Adult Dogs\n Transfer to Out of State',
       '2021\n Adult Dogs\n Other.1', '2021\n Adult Dogs\n Deaths',
       '2021\n Adult Dogs\n Euthanasia',
       '2021\n Adult Dogs\n Missing/Stolen',
       '12/31/2021\n Adult Dogs\n In Shelter',
       '12/31/2021\n Adult Dogs\n In Foster Care',
       '1/1/2021\n Juvenile Dogs\n In Shelter',
       '1/1/2021\n Juvenile Dogs\n In Foster Care',
       '2021\n Juvenile Dogs\n Stray',
       '2021\n Juvenile Dogs\n Owner Relinquished',
       '2021\n Juvenile Dogs\n Transfer from another Colorado Organization',
       '2021\n Juvenile Dogs\n Transfer from Out of State',
       '2021\n Juvenile Dogs\n Other', '2021\n Juvenile Dogs\n Adoption',
       '2021\n Juvenile Dogs\n Returned to Owner (RTO)',
       '2021\n Juvenile Dogs\n Transfer to another Colorado Organization',
       '2021\n Juvenile Dogs\n Transfer to Out of State',
       '2021\n Juvenile Dogs\n Other.1', '2021\n Juvenile Dogs\n Deaths',
       '2021\n Juvenile Dogs\n Euthanasia',
       '2021\n Juvenile Dogs\n Missing/Stolen',
       '12/31/2021\n Juvenile Dogs\n In Shelter',
       '12/31/2021\n Juvenile Dogs\n In Foster Care',
       '1/1/2021\n Adult Cats\n In Shelter',
       '1/1/2021\n Adult Cats\n In Foster Care',
       '2021\n Adult Cats\n Stray',
       '2021\n Adult Cats\n Owner Relinquished',
       '2021\n Adult Cats\n Transfer from another Colorado Organization',
       '2021\n Adult Cats\n Transfer from Out of State',
       '2021\n Adult Cats\n Other', '2021\n Adult Cats\n Adoption',
       '2021\n Adult Cats\n Returned to Owner (RTO)',
       '2021\n Adult Cats\n Transfer to another Colorado Organization',
       '2021\n Adult Cats\n Transfer to Out of State',
```

'2021\n Adult Cats\n Other.1', '2021\n Adult Cats\n Deaths',
'2021\n Adult Cats\n Euthanasia',
'2021\n Adult Cats\n Missing/Stolen',
'12/31/2021\n Adult Cats\n In Shelter',
'12/31/2021\n Adult Cats\n In Foster Care',
'1/1/2021\n Juvenile Cats\n In Shelter',
'1/1/2021\n Juvenile Cats\n In Foster Care',
'2021\n Juvenile Cats\n Stray',
'2021\n Juvenile Cats\n Owner Relinquished',
'2021\n Juvenile Cats\n Transfer from another Colorado Organization',
'2021\n Juvenile Cats\n Transfer from Out of State',
'2021\n Juvenile Cats\n Other', '2021\n Juvenile Cats\n Adoption',
'2021\n Juvenile Cats\n Returned to Owner (RTO)',
'2021\n Juvenile Cats\n Transfer to another Colorado Organization',
'2021\n Juvenile Cats\n Transfer to Out of State',
'2021\n Juvenile Cats\n Other.1', '2021\n Juvenile Cats\n Deaths',
'2021\n Juvenile Cats\n Euthanasia',
'2021\n Juvenile Cats\n Missing/Stolen',
'12/31/2021\n Juvenile Cats\n In Shelter',
'12/31/2021\n Juvenile Cats\n In Foster Care',
'1/1/2021\n Birds\n In Shelter',
'1/1/2021\n Birds\n In Foster Care', '2021\n Birds\n Stray',
'2021\n Birds\n Owner Relinquished',
'2021\n Birds\n Transfer from another Colorado Organization',
'2021\n Birds\n Transfer from Out of State',
'2021\n Birds\n Other', '2021\n Birds\n Adoption',
'2021\n Birds\n Returned to Owner (RTO)',
'2021\n Birds\n Transfer to another Colorado Organization',
'2021\n Birds\n Transfer to Out of State',
'2021\n Birds\n Other.1', '2021\n Birds\n Deaths',
'2021\n Birds\n Euthanasia', '2021\n Birds\n Missing/Stolen',
'12/31/2021\n Birds\n In Shelter',
'12/31/2021\n Birds\n In Foster Care',
'1/1/2021\n Small Mammals\n In Shelter',
'1/1/2021\n Small Mammals\n In Foster Care',
'2021\n Small Mammals\n Stray',
'2021\n Small Mammals\n Owner Relinquished',
'2021\n Small Mammals\n Transfer from another Colorado Organization',
'2021\n Small Mammals\n Transfer from Out of State',
'2021\n Small Mammals\n Other', '2021\n Small Mammals\n Adoption',
'2021\n Small Mammals\n Returned to Owner (RTO)',

```
'2021\n Small Mammals\n Transfer to another Colorado Organization',
'2021\n Small Mammals\n Transfer to Out of State',
'2021\n Small Mammals\n Other.1', '2021\n Small Mammals\n Deaths',
'2021\n Small Mammals\n Euthanasia',
'2021\n Small Mammals\n Missing/Stolen',
'12/31/2021\n Small Mammals\n In Shelter',
'12/31/2021\n Small Mammals\n In Foster Care',
'1/1/2021\n Reptiles & Amphibians\n In Shelter',
'1/1/2021\n Reptiles & Amphibians\n In Foster Care',
'2021\n Reptiles & Amphibians\n Stray',
'2021\n Reptiles & Amphibians\n Owner Relinquished',
'2021\n Reptiles & Amphibians\n Transfer from another Colorado Organization',
'2021\n Reptiles & Amphibians\n Transfer from Out of State',
'2021\n Reptiles & Amphibians\n Other',
'2021\n Reptiles & Amphibians\n Adoption',
'2021\n Reptiles & Amphibians\n Returned to Owner (RTO)',
'2021\n Reptiles & Amphibians\n Transfer to another Colorado Organization',
'2021\n Reptiles & Amphibians\n Transfer to Out of State',
'2021\n Reptiles & Amphibians\n Other.1',
'2021\n Reptiles & Amphibians\n Deaths',
'2021\n Reptiles & Amphibians\n Euthanasia',
'2021\n Reptiles & Amphibians\n Missing/Stolen',
'12/31/2021\n Reptiles & Amphibians\n In Shelter',
'12/31/2021\n Reptiles & Amphibians\n In Foster Care',
'1/1/2021\n Rabbits\n In Shelter',
'1/1/2021\n Rabbits\n In Foster Care', '2021\n Rabbits\n Stray',
'2021\n Rabbits\n Owner Relinquished',
'2021\n Rabbits\n Transfer from another Colorado Organization',
'2021\n Rabbits\n Transfer from Out of State',
'2021\n Rabbits\n Other', '2021\n Rabbits\n Adoption',
'2021\n Rabbits\n Returned to Owner (RTO)',
'2021\n Rabbits\n Transfer to another Colorado Organization',
'2021\n Rabbits\n Transfer to Out of State',
'2021\n Rabbits\n Other.1', '2021\n Rabbits\n Deaths',
'2021\n Rabbits\n Euthanasia', '2021\n Rabbits\n Missing/Stolen',
'12/31/2021\n Rabbits\n In Shelter',
'12/31/2021\n Rabbits\n In Foster Care',
'1/1/2021\n Other\n In Shelter',
'1/1/2021\n Other\n In Foster Care', '2021\n Other\n Stray',
'2021\n Other\n Owner Relinquished',
'2021\n Other\n Transfer from another Colorado Organization',
```

```
       '2021\n Other\n Transfer from Out of State',
       '2021\n Other\n Other', '2021\n Other\n Adoption',
       '2021\n Other\n Returned to Owner (RTO)',
       '2021\n Other\n Transfer to another Colorado Organization',
       '2021\n Other\n Transfer to Out of State',
       '2021\n Other\n Other.1', '2021\n Other\n Deaths',
       '2021\n Other\n Euthanasia', '2021\n Other\n Missing/Stolen',
       '12/31/2021\n Other\n In Shelter',
       '12/31/2021\n Other\n In Foster Care'], dtype=object)
```

In [9]:
```
# Splitting 'Metric' column by newline character to create two columns
split_cols = data_2021_unpivoted["Metric"].str.split("\n", expand=True)

# Animal Type will always be the second part (index 1)
data_2021_unpivoted["Animal Type"] = split_cols[1].str.strip()

# Metric Type will be first part (date) + third part (metric)
data_2021_unpivoted["Metric Type"] = (split_cols[0].str.strip() + " " + split_cols[2].str.strip())
```

In [10]:
```
data_2021_unpivoted.head()
```

Out[10]:

	Facility Name	Metric	Value	Animal Type	Metric Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2021\n Adult Dogs\n In Shelter	26	Adult Dogs	1/1/2021 In Shelter
1	2nd Chance Vizsla Rescue, Inc	1/1/2021\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2021 In Shelter
2	4 Paws 4 Life Rescue	1/1/2021\n Adult Dogs\n In Shelter	91	Adult Dogs	1/1/2021 In Shelter
3	9 Lives Rescue	1/1/2021\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2021 In Shelter
4	A Cat Rescue Out in the Sticks	1/1/2021\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2021 In Shelter

In [11]:
```
# Checking Metric Type unique values after splitting
data_2021_unpivoted["Metric Type"].unique()
```

Out[11]:
```
array(['1/1/2021 In Shelter', '1/1/2021 In Foster Care', '2021 Stray',
       '2021 Owner Relinquished',
       '2021 Transfer from another Colorado Organization',
       '2021 Transfer from Out of State', '2021 Other', '2021 Adoption',
       '2021 Returned to Owner (RTO)',
       '2021 Transfer to another Colorado Organization',
       '2021 Transfer to Out of State', '2021 Other.1', '2021 Deaths',
       '2021 Euthanasia', '2021 Missing/Stolen', '12/31/2021 In Shelter',
       '12/31/2021 In Foster Care'], dtype=object)
```

In [12]:
```
# Standardizing category names inside Metric Type column

metric_map = {
    # Start-of-year counts
    "1/1/2021 In Shelter": "Beginning Shelter Count",
    "1/1/2021 In Foster Care": "Beginning Foster Count",

    # Intakes
    "2021 Stray": "Stray",
    "2021 Owner Relinquished": "Owner Relinquished",
    "2021 Transfer from another Colorado Organization": "Transfer in from another Colorado organization",
    "2021 Transfer from Out of State": "Transfer in from Out of State organization",
    "2021 Other": "Other: TNR / Protective Custody / Returns / Disaster Relief",

    # Outcomes
    "2021 Adoption": "Adoption",
    "2021 Returned to Owner (RTO)": "Returned To Owner (RTO)",
    "2021 Transfer to another Colorado Organization": "Transferred out to another Colorado organization",
    "2021 Transfer to Out of State": "Transferred to an Out of State organization",
    "2021 Other.1": "Other Live Outcomes (ie: tnr / snr)",
    "2021 Deaths": "Died",
    "2021 Euthanasia": "Shelter Euthanasia",
    "2021 Missing/Stolen": "Missing / Stolen",

    # End-of-year counts
    "12/31/2021 In Shelter": "Ending Shelter Count",
    "12/31/2021 In Foster Care": "Ending Foster Count"
}
```

```
# Applying mapping
data_2021_unpivoted["Metric Type"] = data_2021_unpivoted["Metric Type"].replace(metric_map)
```

In [13]:
```
# Checking Metric Type unique values after standardizing categories
data_2021_unpivoted["Metric Type"].unique()
```

Out[13]:
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died',
       'Shelter Euthanasia', 'Missing / Stolen', 'Ending Shelter Count',
       'Ending Foster Count'], dtype=object)
```

In [14]:
```
# Added a new column Flow Type to map and categorize records

# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",

    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
```

```
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
    "Missing / Stolen": "Outcome (Negative)",
    "Shelter Euthanasia": "Outcome (Negative)"
}

# Apply mapping to create new column
data_2021_unpivoted["Flow Type"] = data_2021_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2021_unpivoted = data_2021_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [15]:
```
data_2021_unpivoted.head()
```

Out[15]:

	Facility Name	Metric	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2021\n Adult Dogs\n In Shelter	26	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	2nd Chance Vizsla Rescue, Inc	1/1/2021\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	4 Paws 4 Life Rescue	1/1/2021\n Adult Dogs\n In Shelter	91	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	9 Lives Rescue	1/1/2021\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	A Cat Rescue Out in the Sticks	1/1/2021\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [16]:
```
# Removing 'Metric' column to avoid redundancy.
data_2021_unpivoted = data_2021_unpivoted.drop(columns=["Metric"], errors="ignore")
```

In [17]:
```
data_2021_unpivoted.head()
```

Out[17]:

	Facility Name	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	26	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	2nd Chance Vizsla Rescue, Inc	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	4 Paws 4 Life Rescue	91	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	9 Lives Rescue	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	A Cat Rescue Out in the Sticks	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [18]:
```
# Checking unique values of Animal Type column
data_2021_unpivoted["Animal Type"].unique()
```

Out[18]:
```
array(['Adult Dogs', 'Juvenile Dogs', 'Adult Cats', 'Juvenile Cats',
       'Birds', 'Small Mammals', 'Reptiles & Amphibians', 'Rabbits',
       'Other'], dtype=object)
```

In [19]:
```
# Creating two columns by splitting the 'Animal Type' column

# Species column
data_2021_unpivoted["Species"] = (data_2021_unpivoted["Animal Type"]
                                  .str.replace("Adult", "", case=False)
                                  .str.replace("Juvenile", "", case=False)
                                  .str.strip())

# Age Group column
data_2021_unpivoted["Age Group"] = np.where(
    data_2021_unpivoted["Animal Type"].str.contains("Adult", case=False, na=False), "Adult",
    np.where(data_2021_unpivoted["Animal Type"].str.contains("Juvenile", case=False, na=False), "Juvenile", "Unknown")
)

# Checking unique values
print(data_2021_unpivoted["Species"].unique())
print(data_2021_unpivoted["Age Group"].unique())
```

```
['Dogs' 'Cats' 'Birds' 'Small Mammals' 'Reptiles & Amphibians' 'Rabbits'
 'Other']
['Adult' 'Juvenile' 'Unknown']
```

In [20]:
```
# Removing 'Animal Type' column to avoid redundancy.
data_2021_unpivoted = data_2021_unpivoted.drop(columns=["Animal Type"], errors="ignore")
```

In [21]:
```
# Adding Reporting year column to the whole table
data_2021_unpivoted["Reporting Year"] = 2021
```

In [22]:
```
data_2021_unpivoted.head()
```

Out[22]:

	Facility Name	Animal Count	Event Type	Flow Type	Species	Age Group	Reporting Year
0	2 Blondes All Breed Rescue, Inc.	26	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2021
1	2nd Chance Vizsla Rescue, Inc	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2021
2	4 Paws 4 Life Rescue	91	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2021
3	9 Lives Rescue	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2021
4	A Cat Rescue Out in the Sticks	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2021

In [23]:
```
# Checking Animal Count by Flow Type & Event Type
data_2021_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[23]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	6851
1	End of Year Count	Ending Shelter Count	8511
2	Intake	Other: TNR / Protective Custody / Returns / Di...	13861
3	Intake	Owner Relinquished	41671
4	Intake	Stray	55204
5	Intake	Transfer in from Out of State organization	44521
6	Intake	Transfer in from another Colorado organization	12118
7	Outcome (Negative)	Died	1878
8	Outcome (Negative)	Missing / Stolen	103
9	Outcome (Negative)	Shelter Euthanasia	7977
10	Outcome (Positive)	Adoption	115980
11	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	6326
12	Outcome (Positive)	Returned To Owner (RTO)	23436
13	Outcome (Positive)	Transferred out to another Colorado organization	9824
14	Outcome (Positive)	Transferred to an Out of State organization	277
15	Start of Year Count	Beginning Foster Count	5643
16	Start of Year Count	Beginning Shelter Count	8145

Note: In the 2021 summary document, the animal counts do not match the result values above. However, verification against the dataset confirms that these result values are correct.

In [24]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2021 = data_2021_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in
```

```
summary_2021.to_excel("summary_2021.xlsx", index=False)
```

In [25]:
```
# Verified that all the columns contain no null values
data_2021_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 54162 entries, 0 to 54161
Data columns (total 7 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Facility Name   54162 non-null  object
 1   Animal Count    54162 non-null  int64 
 2   Event Type      54162 non-null  object
 3   Flow Type       54162 non-null  object
 4   Species         54162 non-null  object
 5   Age Group       54162 non-null  object
 6   Reporting Year  54162 non-null  int64 
dtypes: int64(2), object(5)
memory usage: 2.9+ MB
```

In [26]:
```
data_2021_unpivoted.isnull().sum()
```

Out[26]:
```
Facility Name     0
Animal Count      0
Event Type        0
Flow Type         0
Species           0
Age Group         0
Reporting Year    0
dtype: int64
```

In [27]:
```
# Rearranging the columns
Cleaned_2021_Shelter_And_Rescue_Statistics_final = data_2021_unpivoted[["Facility Name", "Reporting Year", "Species",
                                                                        "Age Group", "Flow Type", "Event Type", "Animal Count"
```

In [28]:
```
Cleaned_2021_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2021_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2021_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2021_Shelter_And_Rescue_Statistics_final.xlsx
```

In [29]:
```
from IPython.display import FileLink
FileLink('Cleaned_2021_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[29]: Cleaned_2021_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:
```

```

Shelter And Rescue Statistics Data - 2022

In [1]:
```
# Importing all the required libraries
import numpy as np
import pandas as pd
```

In [2]:
```
# Loading the data
data_2022 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [3]:
```
data_2022.head()
```

Out[3]:

	Facility Name	1/1/2022\n Adult Dogs\n In Shelter	1/1/2022\n Adult Dogs\n In Foster Care	2022\n Adult Dogs\n Stray	2022\n Adult Dogs\n Owner Relinquished	2022\n Adult Dogs\n Transfer from another Colorado Organization	2022\n Adult Dogs\n Transfer from Out of State	2022\n Adult Dogs\n Other	2022\n Adult Dogs\n Adoption	2022\n Adult Dogs\n Returned to Owner (RTO)	...	2022\n Other\n Adoption	2022\n Other\n Returned to Owner (RTO)
0	2 Blondes All Breed Rescue, Inc.	38	65	0	4	7	329	106	426	0	...	0	0
1	2nd Chance Vizsla Rescue Inc	0	9	0	3	0	0	0	5	0	...	0	0
2	4 Paws 4 Life Rescue	0	14	66	0	59	193	0	317	0	...	0	0
3	7 Paws Rescue Ranch	0	0	0	0	0	3	0	0	0	...	0	0
4	9 Lives Rescue	0	0	0	0	0	0	0	0	0	...	0	0

5 rows × 154 columns

In [4]:
```
data_2022.shape
```

Out[4]: (369, 154)

In [5]:
```
# Unpivoting the DataFrame from wide to long format

# Define ID (identifier variables) columns
id_vars = ["Facility Name"]

# Unpivot the rest of the columns into two: "Metric" and "Value"
data_2022_unpivoted = data_2022.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2022_unpivoted = data_2022_unpivoted.dropna(subset=["Value"])
```

In [6]:
```
data_2022_unpivoted.head()
```

Out[6]:

	Facility Name	Metric	Value
0	2 Blondes All Breed Rescue, Inc.	1/1/2022\n Adult Dogs\n In Shelter	38
1	2nd Chance Vizsla Rescue Inc	1/1/2022\n Adult Dogs\n In Shelter	0
2	4 Paws 4 Life Rescue	1/1/2022\n Adult Dogs\n In Shelter	0
3	7 Paws Rescue Ranch	1/1/2022\n Adult Dogs\n In Shelter	0
4	9 Lives Rescue	1/1/2022\n Adult Dogs\n In Shelter	0

In [7]:
```
data_2022_unpivoted.shape
```

Out[7]:
```
(56457, 3)
```

In [8]:
```
data_2022_unpivoted["Metric"].unique()
```

Out[8]:
```
array(['1/1/2022\n Adult Dogs\n In Shelter',
       '1/1/2022\n Adult Dogs\n In Foster Care',
       '2022\n Adult Dogs\n Stray',
       '2022\n Adult Dogs\n Owner Relinquished',
       '2022\n Adult Dogs\n Transfer from another Colorado Organization',
       '2022\n Adult Dogs\n Transfer from Out of State',
       '2022\n Adult Dogs\n Other', '2022\n Adult Dogs\n Adoption',
       '2022\n Adult Dogs\n Returned to Owner (RTO)',
       '2022\n Adult Dogs\n Transfer to another Colorado Organization',
       '2022\n Adult Dogs\n Transfer to Out of State',
       '2022\n Adult Dogs\n Other.1', '2022\n Adult Dogs\n Deaths',
       '2022\n Adult Dogs\n Euthanasia',
       '2022\n Adult Dogs\n Missing/Stolen',
       '12/31/2022\n Adult Dogs\n In Shelter',
       '12/31/2022\n Adult Dogs\n In Foster Care',
       '1/1/2022\n Juvenile Dogs\n In Shelter',
       '1/1/2022\n Juvenile Dogs\n In Foster Care',
       '2022\n Juvenile Dogs\n Stray',
       '2022\n Juvenile Dogs\n Owner Relinquished',
       '2022\n Juvenile Dogs\n Transfer from another Colorado Organization',
       '2022\n Juvenile Dogs\n Transfer from Out of State',
       '2022\n Juvenile Dogs\n Other', '2022\n Juvenile Dogs\n Adoption',
       '2022\n Juvenile Dogs\n Returned to Owner (RTO)',
       '2022\n Juvenile Dogs\n Transfer to another Colorado Organization',
       '2022\n Juvenile Dogs\n Transfer to Out of State',
       '2022\n Juvenile Dogs\n Other.1', '2022\n Juvenile Dogs\n Deaths',
       '2022\n Juvenile Dogs\n Euthanasia',
       '2022\n Juvenile Dogs\n Missing/Stolen',
       '12/31/2022\n Juvenile Dogs\n In Shelter',
       '12/31/2022\n Juvenile Dogs\n In Foster Care',
       '1/1/2022\n Adult Cats\n In Shelter',
       '1/1/2022\n Adult Cats\n In Foster Care',
       '2022\n Adult Cats\n Stray',
       '2022\n Adult Cats\n Owner Relinquished',
       '2022\n Adult Cats\n Transfer from another Colorado Organization',
       '2022\n Adult Cats\n Transfer from Out of State',
       '2022\n Adult Cats\n Other', '2022\n Adult Cats\n Adoption',
       '2022\n Adult Cats\n Returned to Owner (RTO)',
       '2022\n Adult Cats\n Transfer to another Colorado Organization',
       '2022\n Adult Cats\n Transfer to Out of State',
```

```
'2022\n Adult Cats\n Other.1', '2022\n Adult Cats\n Deaths',
'2022\n Adult Cats\n Euthanasia',
'2022\n Adult Cats\n Missing/Stolen',
'12/31/2022\n Adult Cats\n In Shelter',
'12/31/2022\n Adult Cats\n In Foster Care',
'1/1/2022\n Juvenile Cats\n In Shelter',
'1/1/2022\n Juvenile Cats\n In Foster Care',
'2022\n Juvenile Cats\n Stray',
'2022\n Juvenile Cats\n Owner Relinquished',
'2022\n Juvenile Cats\n Transfer from another Colorado Organization',
'2022\n Juvenile Cats\n Transfer from Out of State',
'2022\n Juvenile Cats\n Other', '2022\n Juvenile Cats\n Adoption',
'2022\n Juvenile Cats\n Returned to Owner (RTO)',
'2022\n Juvenile Cats\n Transfer to another Colorado Organization',
'2022\n Juvenile Cats\n Transfer to Out of State',
'2022\n Juvenile Cats\n Other.1', '2022\n Juvenile Cats\n Deaths',
'2022\n Juvenile Cats\n Euthanasia',
'2022\n Juvenile Cats\n Missing/Stolen',
'12/31/2022\n Juvenile Cats\n In Shelter',
'12/31/2022\n Juvenile Cats\n In Foster Care',
'1/1/2022\n Birds\n In Shelter',
'1/1/2022\n Birds\n In Foster Care', '2022\n Birds\n Stray',
'2022\n Birds\n Owner Relinquished',
'2022\n Birds\n Transfer from another Colorado Organization',
'2022\n Birds\n Transfer from Out of State',
'2022\n Birds\n Other', '2022\n Birds\n Adoption',
'2022\n Birds\n Returned to Owner (RTO)',
'2022\n Birds\n Transfer to another Colorado Organization',
'2022\n Birds\n Transfer to Out of State',
'2022\n Birds\n Other.1', '2022\n Birds\n Deaths',
'2022\n Birds\n Euthanasia', '2022\n Birds\n Missing/Stolen',
'12/31/2022\n Birds\n In Shelter',
'12/31/2022\n Birds\n In Foster Care',
'1/1/2022\n Small Mammals\n In Shelter',
'1/1/2022\n Small Mammals\n In Foster Care',
'2022\n Small Mammals\n Stray',
'2022\n Small Mammals\n Owner Relinquished',
'2022\n Small Mammals\n Transfer from another Colorado Organization',
'2022\n Small Mammals\n Transfer from Out of State',
'2022\n Small Mammals\n Other', '2022\n Small Mammals\n Adoption',
'2022\n Small Mammals\n Returned to Owner (RTO)',
```

```
'2022\n Small Mammals\n Transfer to another Colorado Organization',
'2022\n Small Mammals\n Transfer to Out of State',
'2022\n Small Mammals\n Other.1', '2022\n Small Mammals\n Deaths',
'2022\n Small Mammals\n Euthanasia',
'2022\n Small Mammals\n Missing/Stolen',
'12/31/2022\n Small Mammals\n In Shelter',
'12/31/2022\n Small Mammals\n In Foster Care',
'1/1/2022\n Reptiles & Amphibians\n In Shelter',
'1/1/2022\n Reptiles & Amphibians\n In Foster Care',
'2022\n Reptiles & Amphibians\n Stray',
'2022\n Reptiles & Amphibians\n Owner Relinquished',
'2022\n Reptiles & Amphibians\n Transfer from another Colorado Organization',
'2022\n Reptiles & Amphibians\n Transfer from Out of State',
'2022\n Reptiles & Amphibians\n Other',
'2022\n Reptiles & Amphibians\n Adoption',
'2022\n Reptiles & Amphibians\n Returned to Owner (RTO)',
'2022\n Reptiles & Amphibians\n Transfer to another Colorado Organization',
'2022\n Reptiles & Amphibians\n Transfer to Out of State',
'2022\n Reptiles & Amphibians\n Other.1',
'2022\n Reptiles & Amphibians\n Deaths',
'2022\n Reptiles & Amphibians\n Euthanasia',
'2022\n Reptiles & Amphibians\n Missing/Stolen',
'12/31/2022\n Reptiles & Amphibians\n In Shelter',
'12/31/2022\n Reptiles & Amphibians\n In Foster Care',
'1/1/2022\n Rabbits\n In Shelter',
'1/1/2022\n Rabbits\n In Foster Care', '2022\n Rabbits\n Stray',
'2022\n Rabbits\n Owner Relinquished',
'2022\n Rabbits\n Transfer from another Colorado Organization',
'2022\n Rabbits\n Transfer from Out of State',
'2022\n Rabbits\n Other', '2022\n Rabbits\n Adoption',
'2022\n Rabbits\n Returned to Owner (RTO)',
'2022\n Rabbits\n Transfer to another Colorado Organization',
'2022\n Rabbits\n Transfer to Out of State',
'2022\n Rabbits\n Other.1', '2022\n Rabbits\n Deaths',
'2022\n Rabbits\n Euthanasia', '2022\n Rabbits\n Missing/Stolen',
'12/31/2022\n Rabbits\n In Shelter',
'12/31/2022\n Rabbits\n In Foster Care',
'1/1/2022\n Other\n In Shelter',
'1/1/2022\n Other\n In Foster Care', '2022\n Other\n Stray',
'2022\n Other\n Owner Relinquished',
'2022\n Other\n Transfer from another Colorado Organization',
```

```
       '2022\n Other\n Transfer from Out of State',
       '2022\n Other\n Other', '2022\n Other\n Adoption',
       '2022\n Other\n Returned to Owner (RTO)',
       '2022\n Other\n Transfer to another Colorado Organization',
       '2022\n Other\n Transfer to Out of State',
       '2022\n Other\n Other.1', '2022\n Other\n Deaths',
       '2022\n Other\n Euthanasia', '2022\n Other\n Missing/Stolen',
       '12/31/2022\n Other\n In Shelter',
       '12/31/2022\n Other\n In Foster Care'], dtype=object)
```

In [9]:
```
# Splitting 'Metric' column by newline character to create two columns
split_cols = data_2022_unpivoted["Metric"].str.split("\n", expand=True)

# Animal Type will always be the second part (index 1)
data_2022_unpivoted["Animal Type"] = split_cols[1].str.strip()

# Metric Type will be first part (date) + third part (metric)
data_2022_unpivoted["Metric Type"] = (split_cols[0].str.strip() + " " + split_cols[2].str.strip())
```

In [10]:
```
data_2022_unpivoted.head()
```

Out[10]:

	Facility Name	Metric	Value	Animal Type	Metric Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2022\n Adult Dogs\n In Shelter	38	Adult Dogs	1/1/2022 In Shelter
1	2nd Chance Vizsla Rescue Inc	1/1/2022\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2022 In Shelter
2	4 Paws 4 Life Rescue	1/1/2022\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2022 In Shelter
3	7 Paws Rescue Ranch	1/1/2022\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2022 In Shelter
4	9 Lives Rescue	1/1/2022\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2022 In Shelter

In [11]:
```
# Checking Metric Type unique values after splitting
data_2022_unpivoted["Metric Type"].unique()
```

Out[11]:
```
array(['1/1/2022 In Shelter', '1/1/2022 In Foster Care', '2022 Stray',
       '2022 Owner Relinquished',
       '2022 Transfer from another Colorado Organization',
       '2022 Transfer from Out of State', '2022 Other', '2022 Adoption',
       '2022 Returned to Owner (RTO)',
       '2022 Transfer to another Colorado Organization',
       '2022 Transfer to Out of State', '2022 Other.1', '2022 Deaths',
       '2022 Euthanasia', '2022 Missing/Stolen', '12/31/2022 In Shelter',
       '12/31/2022 In Foster Care'], dtype=object)
```

In [12]:
```
# Standardizing category names inside Metric Type column

metric_map = {
    # Start-of-year counts
    "1/1/2022 In Shelter": "Beginning Shelter Count",
    "1/1/2022 In Foster Care": "Beginning Foster Count",

    # Intakes
    "2022 Stray": "Stray",
    "2022 Owner Relinquished": "Owner Relinquished",
    "2022 Transfer from another Colorado Organization": "Transfer in from another Colorado organization",
    "2022 Transfer from Out of State": "Transfer in from Out of State organization",
    "2022 Other": "Other: TNR / Protective Custody / Returns / Disaster Relief",

    # Outcomes
    "2022 Adoption": "Adoption",
    "2022 Returned to Owner (RTO)": "Returned To Owner (RTO)",
    "2022 Transfer to another Colorado Organization": "Transferred out to another Colorado organization",
    "2022 Transfer to Out of State": "Transferred to an Out of State organization",
    "2022 Other.1": "Other Live Outcomes (ie: tnr / snr)",
    "2022 Deaths": "Died",
    "2022 Euthanasia": "Shelter Euthanasia",
    "2022 Missing/Stolen": "Missing / Stolen",

    # End-of-year counts
    "12/31/2022 In Shelter": "Ending Shelter Count",
    "12/31/2022 In Foster Care": "Ending Foster Count"
}
```

```
# Applying mapping
data_2022_unpivoted["Metric Type"] = data_2022_unpivoted["Metric Type"].replace(metric_map)
```

In [13]:
```
# Checking Metric Type unique values after standardizing categories
data_2022_unpivoted["Metric Type"].unique()
```

Out[13]:
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died',
       'Shelter Euthanasia', 'Missing / Stolen', 'Ending Shelter Count',
       'Ending Foster Count'], dtype=object)
```

In [14]:
```
# Added a new column Flow Type to map and categorize records

# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",

    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
```

```
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
    "Missing / Stolen": "Outcome (Negative)",
    "Shelter Euthanasia": "Outcome (Negative)"
}

# Apply mapping to create new column
data_2022_unpivoted["Flow Type"] = data_2022_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2022_unpivoted = data_2022_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [15]:
```
# Removing 'Metric' column to avoid redundancy.
data_2022_unpivoted = data_2022_unpivoted.drop(columns=["Metric"], errors="ignore")
```

In [16]:
```
data_2022_unpivoted.head()
```

Out[16]:

	Facility Name	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	38	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	2nd Chance Vizsla Rescue Inc	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	4 Paws 4 Life Rescue	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	7 Paws Rescue Ranch	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	9 Lives Rescue	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [17]:
```
# Checking unique values of Animal Type column
data_2022_unpivoted["Animal Type"].unique()
```

Out[17]:
```
array(['Adult Dogs', 'Juvenile Dogs', 'Adult Cats', 'Juvenile Cats',
       'Birds', 'Small Mammals', 'Reptiles & Amphibians', 'Rabbits',
       'Other'], dtype=object)
```

In [18]:
```
# Creating two columns by splitting the 'Animal Type' column

# Species column
data_2022_unpivoted["Species"] = (data_2022_unpivoted["Animal Type"]
                                  .str.replace("Adult", "", case=False)
                                  .str.replace("Juvenile", "", case=False)
                                  .str.strip())

# Age Group column
data_2022_unpivoted["Age Group"] = np.where(
    data_2022_unpivoted["Animal Type"].str.contains("Adult", case=False, na=False), "Adult",
    np.where(data_2022_unpivoted["Animal Type"].str.contains("Juvenile", case=False, na=False), "Juvenile", "Unknown")
)

# Checking unique values
print(data_2022_unpivoted["Species"].unique())
print(data_2022_unpivoted["Age Group"].unique())
```

```
['Dogs' 'Cats' 'Birds' 'Small Mammals' 'Reptiles & Amphibians' 'Rabbits'
 'Other']
['Adult' 'Juvenile' 'Unknown']
```

In [19]:
```
# Removing 'Animal Type' column to avoid redundancy.
data_2022_unpivoted = data_2022_unpivoted.drop(columns=["Animal Type"], errors="ignore")
```

In [20]:
```
# Adding Reporting year column to the whole table
data_2022_unpivoted["Reporting Year"] = 2022
```

In [21]:
```
data_2022_unpivoted.head()
```

Out[21]:

	Facility Name	Animal Count	Event Type	Flow Type	Species	Age Group	Reporting Year
0	2 Blondes All Breed Rescue, Inc.	38	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2022
1	2nd Chance Vizsla Rescue Inc	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2022
2	4 Paws 4 Life Rescue	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2022
3	7 Paws Rescue Ranch	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2022
4	9 Lives Rescue	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2022

In [22]:
```
# Checking Animal Count by Flow Type & Event Type
data_2022_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[22]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	6091
1	End of Year Count	Ending Shelter Count	9757
2	Intake	Other: TNR / Protective Custody / Returns / Di...	18781
3	Intake	Owner Relinquished	48408
4	Intake	Stray	64921
5	Intake	Transfer in from Out of State organization	36411
6	Intake	Transfer in from another Colorado organization	14352
7	Outcome (Negative)	Died	2052
8	Outcome (Negative)	Missing / Stolen	75
9	Outcome (Negative)	Shelter Euthanasia	10475
10	Outcome (Positive)	Adoption	124924
11	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	6892
12	Outcome (Positive)	Returned To Owner (RTO)	25721
13	Outcome (Positive)	Transferred out to another Colorado organization	10757
14	Outcome (Positive)	Transferred to an Out of State organization	1294
15	Start of Year Count	Beginning Foster Count	6449
16	Start of Year Count	Beginning Shelter Count	8716

Note: In the 2022 summary document, the animal counts do not match the result values above. However, verification against the dataset confirms that these result values are correct.

In [23]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2022 = data_2022_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in
```

```
summary_2022.to_excel("summary_2022.xlsx", index=False)
```

In [24]:
```
# Verified that all the columns contain no null values
data_2022_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 56457 entries, 0 to 56456
Data columns (total 7 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Facility Name   56457 non-null  object
 1   Animal Count    56457 non-null  int64 
 2   Event Type      56457 non-null  object
 3   Flow Type       56457 non-null  object
 4   Species         56457 non-null  object
 5   Age Group       56457 non-null  object
 6   Reporting Year  56457 non-null  int64 
dtypes: int64(2), object(5)
memory usage: 3.0+ MB
```

In [25]:
```
data_2022_unpivoted.isnull().sum()
```

Out[25]:
```
Facility Name     0
Animal Count      0
Event Type        0
Flow Type         0
Species           0
Age Group         0
Reporting Year    0
dtype: int64
```

In [26]:
```
# Rearranging the columns
Cleaned_2022_Shelter_And_Rescue_Statistics_final = data_2022_unpivoted[["Facility Name", "Reporting Year", "Species",
                                                                        "Age Group", "Flow Type", "Event Type", "Animal Count"
```

In [27]:
```
Cleaned_2022_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2022_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2022_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2022_Shelter_And_Rescue_Statistics_final.xlsx
```

In [28]:
```
from IPython.display import FileLink
FileLink('Cleaned_2022_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[28]: Cleaned_2022_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:
```

```

Shelter And Rescue Statistics Data - 2023

In [1]:
```
# Importing all the required libraries
import numpy as np
import pandas as pd
```

In [2]:
```
# Loading the data
data_2023 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [3]:
```
data_2023.head()
```

Out[3]:

	Facility Name	1/1/2023\n Adult Dogs\n In Shelter	1/1/2023\n Adult Dogs\n In Foster Care	2023\n Adult Dogs\n Stray	2023\n Adult Dogs\n Owner Relinquished	2023\n Adult Dogs\n Transfer from another Colorado Organization	2023\n Adult Dogs\n Transfer from Out of State	2023\n Adult Dogs\n Other	2023\n Adult Dogs\n Adoption	2023\n Adult Dogs\n Returned to Owner (RTO)	...	2023\n Other\n Adoption	2023\n Other\n Returned to Owner (RTO)
0	2 Blondes All Breed Rescue, Inc.	32	86	0	14	4	326	60	428	0	...	0	0
1	4 Paws 4 Life Rescue	0	0	0	19	0	253	0	257	0	...	0	0
2	5280 Reptile Room North	0	0	0	0	0	0	0	0	0	...	0	0
3	5280 Reptile Room west	0	0	0	0	0	0	0	0	0	...	0	0
4	7 Paws Rescue Ranch	2	0	0	0	0	2	0	1	0	...	0	0

5 rows × 154 columns

In [4]:
```
data_2023.shape  # total rows and columns (original dataset)
```

Out[4]: (347, 154)

In [5]:
```
# Unpivoting the DataFrame from wide to long format

# Define ID (identifier variables) columns
id_vars = ["Facility Name"]

# Unpivot the rest of the columns into two: "Metric" and "Value"
data_2023_unpivoted = data_2023.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2023_unpivoted = data_2023_unpivoted.dropna(subset=["Value"])
```

In [6]:
```
data_2023_unpivoted.head()
```

Out[6]:

	Facility Name	Metric	Value
0	2 Blondes All Breed Rescue, Inc.	1/1/2023\n Adult Dogs\n In Shelter	32
1	4 Paws 4 Life Rescue	1/1/2023\n Adult Dogs\n In Shelter	0
2	5280 Reptile Room North	1/1/2023\n Adult Dogs\n In Shelter	0
3	5280 Reptile Room west	1/1/2023\n Adult Dogs\n In Shelter	0
4	7 Paws Rescue Ranch	1/1/2023\n Adult Dogs\n In Shelter	2

In [7]:
```
data_2023_unpivoted.shape  # total rows and columns after unpivoting
```

Out[7]: (53091, 3)

In [8]:
```
# checking unique values of the Metric column
data_2023_unpivoted["Metric"].unique()
```

Out[8]:
```
array(['1/1/2023\n Adult Dogs\n In Shelter',
       '1/1/2023\n Adult Dogs\n In Foster Care',
       '2023\n Adult Dogs\n Stray',
       '2023\n Adult Dogs\n Owner Relinquished',
       '2023\n Adult Dogs\n Transfer from another Colorado Organization',
       '2023\n Adult Dogs\n Transfer from Out of State',
       '2023\n Adult Dogs\n Other', '2023\n Adult Dogs\n Adoption',
       '2023\n Adult Dogs\n Returned to Owner (RTO)',
       '2023\n Adult Dogs\n Transfer to another Colorado Organization',
       '2023\n Adult Dogs\n Transfer to Out of State',
       '2023\n Adult Dogs\n Other.1', '2023\n Adult Dogs\n Deaths',
       '2023\n Adult Dogs\n Euthanasia',
       '2023\n Adult Dogs\n Missing/Stolen',
       '12/31/2023\n Adult Dogs\n In Shelter',
       '12/31/2023\n Adult Dogs\n In Foster Care',
       '1/1/2023\n Juvenile Dogs\n In Shelter',
       '1/1/2023\n Juvenile Dogs\n In Foster Care',
       '2023\n Juvenile Dogs\n Stray',
       '2023\n Juvenile Dogs\n Owner Relinquished',
       '2023\n Juvenile Dogs\n Transfer from another Colorado Organization',
       '2023\n Juvenile Dogs\n Transfer from Out of State',
       '2023\n Juvenile Dogs\n Other', '2023\n Juvenile Dogs\n Adoption',
       '2023\n Juvenile Dogs\n Returned to Owner (RTO)',
       '2023\n Juvenile Dogs\n Transfer to another Colorado Organization',
       '2023\n Juvenile Dogs\n Transfer to Out of State',
       '2023\n Juvenile Dogs\n Other.1', '2023\n Juvenile Dogs\n Deaths',
       '2023\n Juvenile Dogs\n Euthanasia',
       '2023\n Juvenile Dogs\n Missing/Stolen',
       '12/31/2023\n Juvenile Dogs\n In Shelter',
       '12/31/2023\n Juvenile Dogs\n In Foster Care',
       '1/1/2023\n Adult Cats\n In Shelter',
       '1/1/2023\n Adult Cats\n In Foster Care',
       '2023\n Adult Cats\n Stray',
       '2023\n Adult Cats\n Owner Relinquished',
       '2023\n Adult Cats\n Transfer from another Colorado Organization',
       '2023\n Adult Cats\n Transfer from Out of State',
       '2023\n Adult Cats\n Other', '2023\n Adult Cats\n Adoption',
       '2023\n Adult Cats\n Returned to Owner (RTO)',
       '2023\n Adult Cats\n Transfer to another Colorado Organization',
       '2023\n Adult Cats\n Transfer to Out of State',
```

'2023\n Adult Cats\n Other.1', '2023\n Adult Cats\n Deaths',
'2023\n Adult Cats\n Euthanasia',
'2023\n Adult Cats\n Missing/Stolen',
'12/31/2023\n Adult Cats\n In Shelter',
'12/31/2023\n Adult Cats\n In Foster Care',
'1/1/2023\n Juvenile Cats\n In Shelter',
'1/1/2023\n Juvenile Cats\n In Foster Care',
'2023\n Juvenile Cats\n Stray',
'2023\n Juvenile Cats\n Owner Relinquished',
'2023\n Juvenile Cats\n Transfer from another Colorado Organization',
'2023\n Juvenile Cats\n Transfer from Out of State',
'2023\n Juvenile Cats\n Other', '2023\n Juvenile Cats\n Adoption',
'2023\n Juvenile Cats\n Returned to Owner (RTO)',
'2023\n Juvenile Cats\n Transfer to another Colorado Organization',
'2023\n Juvenile Cats\n Transfer to Out of State',
'2023\n Juvenile Cats\n Other.1', '2023\n Juvenile Cats\n Deaths',
'2023\n Juvenile Cats\n Euthanasia',
'2023\n Juvenile Cats\n Missing/Stolen',
'12/31/2023\n Juvenile Cats\n In Shelter',
'12/31/2023\n Juvenile Cats\n In Foster Care',
'1/1/2023\n Birds\n In Shelter',
'1/1/2023\n Birds\n In Foster Care', '2023\n Birds\n Stray',
'2023\n Birds\n Owner Relinquished',
'2023\n Birds\n Transfer from another Colorado Organization',
'2023\n Birds\n Transfer from Out of State',
'2023\n Birds\n Other', '2023\n Birds\n Adoption',
'2023\n Birds\n Returned to Owner (RTO)',
'2023\n Birds\n Transfer to another Colorado Organization',
'2023\n Birds\n Transfer to Out of State',
'2023\n Birds\n Other.1', '2023\n Birds\n Deaths',
'2023\n Birds\n Euthanasia', '2023\n Birds\n Missing/Stolen',
'12/31/2023\n Birds\n In Shelter',
'12/31/2023\n Birds\n In Foster Care',
'1/1/2023\n Small Mammals\n In Shelter',
'1/1/2023\n Small Mammals\n In Foster Care',
'2023\n Small Mammals\n Stray',
'2023\n Small Mammals\n Owner Relinquished',
'2023\n Small Mammals\n Transfer from another Colorado Organization',
'2023\n Small Mammals\n Transfer from Out of State',
'2023\n Small Mammals\n Other', '2023\n Small Mammals\n Adoption',
'2023\n Small Mammals\n Returned to Owner (RTO)',

```
'2023\n Small Mammals\n Transfer to another Colorado Organization',
'2023\n Small Mammals\n Transfer to Out of State',
'2023\n Small Mammals\n Other.1', '2023\n Small Mammals\n Deaths',
'2023\n Small Mammals\n Euthanasia',
'2023\n Small Mammals\n Missing/Stolen',
'12/31/2023\n Small Mammals\n In Shelter',
'12/31/2023\n Small Mammals\n In Foster Care',
'1/1/2023\n Reptiles & Amphibians\n In Shelter',
'1/1/2023\n Reptiles & Amphibians\n In Foster Care',
'2023\n Reptiles & Amphibians\n Stray',
'2023\n Reptiles & Amphibians\n Owner Relinquished',
'2023\n Reptiles & Amphibians\n Transfer from another Colorado Organization',
'2023\n Reptiles & Amphibians\n Transfer from Out of State',
'2023\n Reptiles & Amphibians\n Other',
'2023\n Reptiles & Amphibians\n Adoption',
'2023\n Reptiles & Amphibians\n Returned to Owner (RTO)',
'2023\n Reptiles & Amphibians\n Transfer to another Colorado Organization',
'2023\n Reptiles & Amphibians\n Transfer to Out of State',
'2023\n Reptiles & Amphibians\n Other.1',
'2023\n Reptiles & Amphibians\n Deaths',
'2023\n Reptiles & Amphibians\n Euthanasia',
'2023\n Reptiles & Amphibians\n Missing/Stolen',
'12/31/2023\n Reptiles & Amphibians\n In Shelter',
'12/31/2023\n Reptiles & Amphibians\n In Foster Care',
'1/1/2023\n Rabbits\n In Shelter',
'1/1/2023\n Rabbits\n In Foster Care', '2023\n Rabbits\n Stray',
'2023\n Rabbits\n Owner Relinquished',
'2023\n Rabbits\n Transfer from another Colorado Organization',
'2023\n Rabbits\n Transfer from Out of State',
'2023\n Rabbits\n Other', '2023\n Rabbits\n Adoption',
'2023\n Rabbits\n Returned to Owner (RTO)',
'2023\n Rabbits\n Transfer to another Colorado Organization',
'2023\n Rabbits\n Transfer to Out of State',
'2023\n Rabbits\n Other.1', '2023\n Rabbits\n Deaths',
'2023\n Rabbits\n Euthanasia', '2023\n Rabbits\n Missing/Stolen',
'12/31/2023\n Rabbits\n In Shelter',
'12/31/2023\n Rabbits\n In Foster Care',
'1/1/2023\n Other\n In Shelter',
'1/1/2023\n Other\n In Foster Care', '2023\n Other\n Stray',
'2023\n Other\n Owner Relinquished',
'2023\n Other\n Transfer from another Colorado Organization',
```

```
'2023\n Other\n Transfer from Out of State',
'2023\n Other\n Other', '2023\n Other\n Adoption',
'2023\n Other\n Returned to Owner (RTO)',
'2023\n Other\n Transfer to another Colorado Organization',
'2023\n Other\n Transfer to Out of State',
'2023\n Other\n Other.1', '2023\n Other\n Deaths',
'2023\n Other\n Euthanasia', '2023\n Other\n Missing/Stolen',
'12/31/2023\n Other\n In Shelter',
'12/31/2023\n Other\n In Foster Care'], dtype=object)
```

In [9]:
```
# Splitting 'Metric' column by newline character to create two columns
split_cols = data_2023_unpivoted["Metric"].str.split("\n", expand=True)

# Animal Type will always be the second part (index 1)
data_2023_unpivoted["Animal Type"] = split_cols[1].str.strip()

# Metric Type will be first part (date) + third part (metric)
data_2023_unpivoted["Metric Type"] = (split_cols[0].str.strip() + " " + split_cols[2].str.strip())
```

In [10]:
```
data_2023_unpivoted.head()
```

Out[10]:

	Facility Name	Metric	Value	Animal Type	Metric Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2023\n Adult Dogs\n In Shelter	32	Adult Dogs	1/1/2023 In Shelter
1	4 Paws 4 Life Rescue	1/1/2023\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2023 In Shelter
2	5280 Reptile Room North	1/1/2023\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2023 In Shelter
3	5280 Reptile Room west	1/1/2023\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2023 In Shelter
4	7 Paws Rescue Ranch	1/1/2023\n Adult Dogs\n In Shelter	2	Adult Dogs	1/1/2023 In Shelter

In [11]:
```
# Checking Metric Type unique values after splitting
data_2023_unpivoted["Metric Type"].unique()
```

Out[11]:
```
array(['1/1/2023 In Shelter', '1/1/2023 In Foster Care', '2023 Stray',
       '2023 Owner Relinquished',
       '2023 Transfer from another Colorado Organization',
       '2023 Transfer from Out of State', '2023 Other', '2023 Adoption',
       '2023 Returned to Owner (RTO)',
       '2023 Transfer to another Colorado Organization',
       '2023 Transfer to Out of State', '2023 Other.1', '2023 Deaths',
       '2023 Euthanasia', '2023 Missing/Stolen', '12/31/2023 In Shelter',
       '12/31/2023 In Foster Care'], dtype=object)
```

In [12]:
```
# Standardizing category names inside Metric Type column

metric_map = {
    # Start-of-year counts
    "1/1/2023 In Shelter": "Beginning Shelter Count",
    "1/1/2023 In Foster Care": "Beginning Foster Count",

    # Intakes
    "2023 Stray": "Stray",
    "2023 Owner Relinquished": "Owner Relinquished",
    "2023 Transfer from another Colorado Organization": "Transfer in from another Colorado organization",
    "2023 Transfer from Out of State": "Transfer in from Out of State organization",
    "2023 Other": "Other: TNR / Protective Custody / Returns / Disaster Relief",

    # Outcomes
    "2023 Adoption": "Adoption",
    "2023 Returned to Owner (RTO)": "Returned To Owner (RTO)",
    "2023 Transfer to another Colorado Organization": "Transferred out to another Colorado organization",
    "2023 Transfer to Out of State": "Transferred to an Out of State organization",
    "2023 Other.1": "Other Live Outcomes (ie: tnr / snr)",
    "2023 Deaths": "Died",
    "2023 Euthanasia": "Shelter Euthanasia",
    "2023 Missing/Stolen": "Missing / Stolen",

    # End-of-year counts
    "12/31/2023 In Shelter": "Ending Shelter Count",
    "12/31/2023 In Foster Care": "Ending Foster Count"
}
```

```
# Applying mapping
data_2023_unpivoted["Metric Type"] = data_2023_unpivoted["Metric Type"].replace(metric_map)
```

In [13]:
```
# Checking Metric Type unique values after standardizing categories
data_2023_unpivoted["Metric Type"].unique()
```

Out[13]:
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died',
       'Shelter Euthanasia', 'Missing / Stolen', 'Ending Shelter Count',
       'Ending Foster Count'], dtype=object)
```

In [14]:
```
# Added a new column Flow Type to map and categorize records

# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",

    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
```

```
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
    "Missing / Stolen": "Outcome (Negative)",
    "Shelter Euthanasia": "Outcome (Negative)"
}

# Apply mapping to create new column
data_2023_unpivoted["Flow Type"] = data_2023_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2023_unpivoted = data_2023_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [15]:
```
data_2023_unpivoted.head()
```

Out[15]:

	Facility Name	Metric	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2023\n Adult Dogs\n In Shelter	32	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	4 Paws 4 Life Rescue	1/1/2023\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	5280 Reptile Room North	1/1/2023\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	5280 Reptile Room west	1/1/2023\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	7 Paws Rescue Ranch	1/1/2023\n Adult Dogs\n In Shelter	2	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [16]:
```
# Removing 'Metric' column to avoid redundancy.
data_2023_unpivoted = data_2023_unpivoted.drop(columns=["Metric"], errors="ignore")
```

In [17]:
```
data_2023_unpivoted.head()
```

Out[17]:

	Facility Name	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	32	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	4 Paws 4 Life Rescue	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	5280 Reptile Room North	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	5280 Reptile Room west	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	7 Paws Rescue Ranch	2	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [18]:
```
# Checking unique values of Animal Type column
data_2023_unpivoted["Animal Type"].unique()
```

Out[18]:
```
array(['Adult Dogs', 'Juvenile Dogs', 'Adult Cats', 'Juvenile Cats',
       'Birds', 'Small Mammals', 'Reptiles & Amphibians', 'Rabbits',
       'Other'], dtype=object)
```

In [19]:
```
# Creating two columns by splitting the 'Animal Type' column

# Species column
data_2023_unpivoted["Species"] = (data_2023_unpivoted["Animal Type"]
                                  .str.replace("Adult", "", case=False)
                                  .str.replace("Juvenile", "", case=False)
                                  .str.strip())

# Age Group column
data_2023_unpivoted["Age Group"] = np.where(
    data_2023_unpivoted["Animal Type"].str.contains("Adult", case=False, na=False), "Adult",
    np.where(data_2023_unpivoted["Animal Type"].str.contains("Juvenile", case=False, na=False), "Juvenile", "Unknown")
)

# Checking unique values
print(data_2023_unpivoted["Species"].unique())
print(data_2023_unpivoted["Age Group"].unique())
```

```
['Dogs' 'Cats' 'Birds' 'Small Mammals' 'Reptiles & Amphibians' 'Rabbits'
 'Other']
['Adult' 'Juvenile' 'Unknown']
```

In [20]:
```
# Removing 'Animal Type' column to avoid redundancy.
data_2023_unpivoted = data_2023_unpivoted.drop(columns=["Animal Type"], errors="ignore")
```

In [21]:
```
# Adding Reporting year column to the whole table
data_2023_unpivoted["Reporting Year"] = 2023
```

In [22]:
```
data_2023_unpivoted.head()
```

Out[22]:

	Facility Name	Animal Count	Event Type	Flow Type	Species	Age Group	Reporting Year
0	2 Blondes All Breed Rescue, Inc.	32	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2023
1	4 Paws 4 Life Rescue	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2023
2	5280 Reptile Room North	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2023
3	5280 Reptile Room west	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2023
4	7 Paws Rescue Ranch	2	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2023

In [23]:
```
# Checking Animal Count by Flow Type & Event Type
data_2023_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[23]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	6074
1	End of Year Count	Ending Shelter Count	10344
2	Intake	Other: TNR / Protective Custody / Returns / Di...	16052
3	Intake	Owner Relinquished	48994
4	Intake	Stray	60590
5	Intake	Transfer in from Out of State organization	32284
6	Intake	Transfer in from another Colorado organization	15142
7	Outcome (Negative)	Died	2056
8	Outcome (Negative)	Missing / Stolen	122
9	Outcome (Negative)	Shelter Euthanasia	11644
10	Outcome (Positive)	Adoption	115032
11	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	7219
12	Outcome (Positive)	Returned To Owner (RTO)	24165
13	Outcome (Positive)	Transferred out to another Colorado organization	12022
14	Outcome (Positive)	Transferred to an Out of State organization	406
15	Start of Year Count	Beginning Foster Count	6653
16	Start of Year Count	Beginning Shelter Count	9369

Note: In the 2023 summary document, some of the animal counts do not match the result values above. However, verification against the dataset confirms that these result values are correct.

In [24]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2023 = data_2023_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in
```

```
summary_2023.to_excel("summary_2023.xlsx", index=False)
```

In [25]:
```
# Verified that all the columns contain no null values
data_2023_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 53091 entries, 0 to 53090
Data columns (total 7 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Facility Name   53091 non-null  object
 1   Animal Count    53091 non-null  int64 
 2   Event Type      53091 non-null  object
 3   Flow Type       53091 non-null  object
 4   Species         53091 non-null  object
 5   Age Group       53091 non-null  object
 6   Reporting Year  53091 non-null  int64 
dtypes: int64(2), object(5)
memory usage: 2.8+ MB
```

In [26]:
```
data_2023_unpivoted.isnull().sum()
```

Out[26]:
```
Facility Name     0
Animal Count      0
Event Type        0
Flow Type         0
Species           0
Age Group         0
Reporting Year    0
dtype: int64
```

In [27]:
```
# Rearranging the columns
Cleaned_2023_Shelter_And_Rescue_Statistics_final = data_2023_unpivoted[["Facility Name", "Reporting Year", "Species",
                                                                        "Age Group", "Flow Type", "Event Type", "Animal Count"
```

In [28]:
```
Cleaned_2023_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2023_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2023_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2023_Shelter_And_Rescue_Statistics_final.xlsx
```

In [29]:
```
from IPython.display import FileLink
FileLink('Cleaned_2023_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[29]: Cleaned_2023_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:
```

```

Shelter And Rescue Statistics Data - 2024

In [1]:
```
# Importing all the required libraries
import numpy as np
import pandas as pd
```

In [2]:
```
# Loading the data
data_2024 = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Data
```

In [3]:
```
data_2024.head()
```

Out[3]:

	Facility Name	1/1/2024\n Adult Dogs\n In Shelter	1/1/2024\n Adult Dogs\n In Foster Care	2024\n Adult Dogs\n Stray	2024\n Adult Dogs\n Owner Relinquished	2024\n Adult Dogs\n Transfer from another Colorado Organization	2024\n Adult Dogs\n Transfer from Out of State	2024\n Adult Dogs\n Other	2024\n Adult Dogs\n Adoption	2024\n Adult Dogs\n Returned to Owner (RTO)	...	2024\n Other\n Adoption	2024\n Other\n Returned to Owner (RTO)
0	2 Blondes All Breed Rescue, Inc.	33	57	0	8	17	221	20	274	0	...	0	0
1	4 Paws 4 Life Rescue	7	3	0	15	7	179	0	192	0	...	0	0
2	5280 Reptile Room North	0	0	0	0	0	0	0	0	0	...	0	0
3	5280 Reptile Room West	0	0	0	0	0	0	0	0	0	...	0	0
4	9 Lives Rescue	0	0	0	0	0	0	0	0	0	...	0	0

5 rows × 154 columns

In [4]:
```
data_2024.shape  # total rows and columns count of raw data
```

Out[4]: (360, 154)

In [5]:
```
# Unpivoting the DataFrame from wide to long format

# Define ID (identifier variables) columns
id_vars = ["Facility Name"]

# Unpivot the rest of the columns into two: "Metric" and "Value"
data_2024_unpivoted = data_2024.melt(id_vars = id_vars,
                                     var_name = "Metric",     # This will contain the old column names
                                     value_name = "Value"     # This will contain the numbers
)

# Drop rows where Metric Value is null
data_2024_unpivoted = data_2024_unpivoted.dropna(subset=["Value"])
```

In [6]:
```
data_2024_unpivoted.head()
```

Out[6]:

	Facility Name	Metric	Value
0	2 Blondes All Breed Rescue, Inc.	1/1/2024\n Adult Dogs\n In Shelter	33
1	4 Paws 4 Life Rescue	1/1/2024\n Adult Dogs\n In Shelter	7
2	5280 Reptile Room North	1/1/2024\n Adult Dogs\n In Shelter	0
3	5280 Reptile Room West	1/1/2024\n Adult Dogs\n In Shelter	0
4	9 Lives Rescue	1/1/2024\n Adult Dogs\n In Shelter	0

In [7]:
```
data_2024_unpivoted.shape   # total rows and columns count after unpivoting
```

Out[7]:
```
(55080, 3)
```

In [8]:
```
# checking unique values of the Metric column
data_2024_unpivoted["Metric"].unique()
```

Out[8]:
```
array(['1/1/2024\n Adult Dogs\n In Shelter',
       '1/1/2024\n Adult Dogs\n In Foster Care',
       '2024\n Adult Dogs\n Stray',
       '2024\n Adult Dogs\n Owner Relinquished',
       '2024\n Adult Dogs\n Transfer from another Colorado Organization',
       '2024\n Adult Dogs\n Transfer from Out of State',
       '2024\n Adult Dogs\n Other', '2024\n Adult Dogs\n Adoption',
       '2024\n Adult Dogs\n Returned to Owner (RTO)',
       '2024\n Adult Dogs\n Transfer to another Colorado Organization',
       '2024\n Adult Dogs\n Transfer to Out of State',
       '2024\n Adult Dogs\n Other.1', '2024\n Adult Dogs\n Deaths',
       '2024\n Adult Dogs\n Euthanasia',
       '2024\n Adult Dogs\n Missing/Stolen',
       '12/31/2024\n Adult Dogs\n In Shelter',
       '12/31/2024\n Adult Dogs\n In Foster Care',
       '1/1/2024\n Juvenile Dogs\n In Shelter',
       '1/1/2024\n Juvenile Dogs\n In Foster Care',
       '2024\n Juvenile Dogs\n Stray',
       '2024\n Juvenile Dogs\n Owner Relinquished',
       '2024\n Juvenile Dogs\n Transfer from another Colorado Organization',
       '2024\n Juvenile Dogs\n Transfer from Out of State',
       '2024\n Juvenile Dogs\n Other', '2024\n Juvenile Dogs\n Adoption',
       '2024\n Juvenile Dogs\n Returned to Owner (RTO)',
       '2024\n Juvenile Dogs\n Transfer to another Colorado Organization',
       '2024\n Juvenile Dogs\n Transfer to Out of State',
       '2024\n Juvenile Dogs\n Other.1', '2024\n Juvenile Dogs\n Deaths',
       '2024\n Juvenile Dogs\n Euthanasia',
       '2024\n Juvenile Dogs\n Missing/Stolen',
       '12/31/2024\n Juvenile Dogs\n In Shelter',
       '12/31/2024\n Juvenile Dogs\n In Foster Care',
       '1/1/2024\n Adult Cats\n In Shelter',
       '1/1/2024\n Adult Cats\n In Foster Care',
       '2024\n Adult Cats\n Stray',
       '2024\n Adult Cats\n Owner Relinquished',
       '2024\n Adult Cats\n Transfer from another Colorado Organization',
       '2024\n Adult Cats\n Transfer from Out of State',
       '2024\n Adult Cats\n Other', '2024\n Adult Cats\n Adoption',
       '2024\n Adult Cats\n Returned to Owner (RTO)',
       '2024\n Adult Cats\n Transfer to another Colorado Organization',
       '2024\n Adult Cats\n Transfer to Out of State',
```

```
'2024\n Adult Cats\n Other.1', '2024\n Adult Cats\n Deaths',
'2024\n Adult Cats\n Euthanasia',
'2024\n Adult Cats\n Missing/Stolen',
'12/31/2024\n Adult Cats\n In Shelter',
'12/31/2024\n Adult Cats\n In Foster Care',
'1/1/2024\n Juvenile Cats\n In Shelter',
'1/1/2024\n Juvenile Cats\n In Foster Care',
'2024\n Juvenile Cats\n Stray',
'2024\n Juvenile Cats\n Owner Relinquished',
'2024\n Juvenile Cats\n Transfer from another Colorado Organization',
'2024\n Juvenile Cats\n Transfer from Out of State',
'2024\n Juvenile Cats\n Other', '2024\n Juvenile Cats\n Adoption',
'2024\n Juvenile Cats\n Returned to Owner (RTO)',
'2024\n Juvenile Cats\n Transfer to another Colorado Organization',
'2024\n Juvenile Cats\n Transfer to Out of State',
'2024\n Juvenile Cats\n Other.1', '2024\n Juvenile Cats\n Deaths',
'2024\n Juvenile Cats\n Euthanasia',
'2024\n Juvenile Cats\n Missing/Stolen',
'12/31/2024\n Juvenile Cats\n In Shelter',
'12/31/2024\n Juvenile Cats\n In Foster Care',
'1/1/2024\n Birds\n In Shelter',
'1/1/2024\n Birds\n In Foster Care', '2024\n Birds\n Stray',
'2024\n Birds\n Owner Relinquished',
'2024\n Birds\n Transfer from another Colorado Organization',
'2024\n Birds\n Transfer from Out of State',
'2024\n Birds\n Other', '2024\n Birds\n Adoption',
'2024\n Birds\n Returned to Owner (RTO)',
'2024\n Birds\n Transfer to another Colorado Organization',
'2024\n Birds\n Transfer to Out of State',
'2024\n Birds\n Other.1', '2024\n Birds\n Deaths',
'2024\n Birds\n Euthanasia', '2024\n Birds\n Missing/Stolen',
'12/31/2024\n Birds\n In Shelter',
'12/31/2024\n Birds\n In Foster Care',
'1/1/2024\n Small Mammals\n In Shelter',
'1/1/2024\n Small Mammals\n In Foster Care',
'2024\n Small Mammals\n Stray',
'2024\n Small Mammals\n Owner Relinquished',
'2024\n Small Mammals\n Transfer from another Colorado Organization',
'2024\n Small Mammals\n Transfer from Out of State',
'2024\n Small Mammals\n Other', '2024\n Small Mammals\n Adoption',
'2024\n Small Mammals\n Returned to Owner (RTO)',
```

'2024\n Small Mammals\n Transfer to another Colorado Organization',
'2024\n Small Mammals\n Transfer to Out of State',
'2024\n Small Mammals\n Other.1', '2024\n Small Mammals\n Deaths',
'2024\n Small Mammals\n Euthanasia',
'2024\n Small Mammals\n Missing/Stolen',
'12/31/2024\n Small Mammals\n In Shelter',
'12/31/2024\n Small Mammals\n In Foster Care',
'1/1/2024\n Reptiles & Amphibians\n In Shelter',
'1/1/2024\n Reptiles & Amphibians\n In Foster Care',
'2024\n Reptiles & Amphibians\n Stray',
'2024\n Reptiles & Amphibians\n Owner Relinquished',
'2024\n Reptiles & Amphibians\n Transfer from another Colorado Organization',
'2024\n Reptiles & Amphibians\n Transfer from Out of State',
'2024\n Reptiles & Amphibians\n Other',
'2024\n Reptiles & Amphibians\n Adoption',
'2024\n Reptiles & Amphibians\n Returned to Owner (RTO)',
'2024\n Reptiles & Amphibians\n Transfer to another Colorado Organization',
'2024\n Reptiles & Amphibians\n Transfer to Out of State',
'2024\n Reptiles & Amphibians\n Other.1',
'2024\n Reptiles & Amphibians\n Deaths',
'2024\n Reptiles & Amphibians\n Euthanasia',
'2024\n Reptiles & Amphibians\n Missing/Stolen',
'12/31/2024\n Reptiles & Amphibians\n In Shelter',
'12/31/2024\n Reptiles & Amphibians\n In Foster Care',
'1/1/2024\n Rabbits\n In Shelter',
'1/1/2024\n Rabbits\n In Foster Care', '2024\n Rabbits\n Stray',
'2024\n Rabbits\n Owner Relinquished',
'2024\n Rabbits\n Transfer from another Colorado Organization',
'2024\n Rabbits\n Transfer from Out of State',
'2024\n Rabbits\n Other', '2024\n Rabbits\n Adoption',
'2024\n Rabbits\n Returned to Owner (RTO)',
'2024\n Rabbits\n Transfer to another Colorado Organization',
'2024\n Rabbits\n Transfer to Out of State',
'2024\n Rabbits\n Other.1', '2024\n Rabbits\n Deaths',
'2024\n Rabbits\n Euthanasia', '2024\n Rabbits\n Missing/Stolen',
'12/31/2024\n Rabbits\n In Shelter',
'12/31/2024\n Rabbits\n In Foster Care',
'1/1/2024\n Other\n In Shelter',
'1/1/2024\n Other\n In Foster Care', '2024\n Other\n Stray',
'2024\n Other\n Owner Relinquished',
'2024\n Other\n Transfer from another Colorado Organization',

```
       '2024\n Other\n Transfer from Out of State',
       '2024\n Other\n Other', '2024\n Other\n Adoption',
       '2024\n Other\n Returned to Owner (RTO)',
       '2024\n Other\n Transfer to another Colorado Organization',
       '2024\n Other\n Transfer to Out of State',
       '2024\n Other\n Other.1', '2024\n Other\n Deaths',
       '2024\n Other\n Euthanasia', '2024\n Other\n Missing/Stolen',
       '12/31/2024\n Other\n In Shelter',
       '12/31/2024\n Other\n In Foster Care'], dtype=object)
```

In [9]:
```
# Splitting 'Metric' column by newline character to create two columns
split_cols = data_2024_unpivoted["Metric"].str.split("\n", expand=True)

# Animal Type will always be the second part (index 1)
data_2024_unpivoted["Animal Type"] = split_cols[1].str.strip()

# Metric Type will be first part (date) + third part (metric)
data_2024_unpivoted["Metric Type"] = (split_cols[0].str.strip() + " " + split_cols[2].str.strip())
```

In [10]:
```
data_2024_unpivoted.head()
```

Out[10]:

	Facility Name	Metric	Value	Animal Type	Metric Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2024\n Adult Dogs\n In Shelter	33	Adult Dogs	1/1/2024 In Shelter
1	4 Paws 4 Life Rescue	1/1/2024\n Adult Dogs\n In Shelter	7	Adult Dogs	1/1/2024 In Shelter
2	5280 Reptile Room North	1/1/2024\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2024 In Shelter
3	5280 Reptile Room West	1/1/2024\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2024 In Shelter
4	9 Lives Rescue	1/1/2024\n Adult Dogs\n In Shelter	0	Adult Dogs	1/1/2024 In Shelter

In [11]:
```
# Checking Metric Type unique values after splitting
data_2024_unpivoted["Metric Type"].unique()
```

Out[11]:
```
array(['1/1/2024 In Shelter', '1/1/2024 In Foster Care', '2024 Stray',
       '2024 Owner Relinquished',
       '2024 Transfer from another Colorado Organization',
       '2024 Transfer from Out of State', '2024 Other', '2024 Adoption',
       '2024 Returned to Owner (RTO)',
       '2024 Transfer to another Colorado Organization',
       '2024 Transfer to Out of State', '2024 Other.1', '2024 Deaths',
       '2024 Euthanasia', '2024 Missing/Stolen', '12/31/2024 In Shelter',
       '12/31/2024 In Foster Care'], dtype=object)
```

In [12]:
```
# Standardizing category names inside Metric Type column

metric_map = {
    # Start-of-year counts
    "1/1/2024 In Shelter": "Beginning Shelter Count",
    "1/1/2024 In Foster Care": "Beginning Foster Count",

    # Intakes
    "2024 Stray": "Stray",
    "2024 Owner Relinquished": "Owner Relinquished",
    "2024 Transfer from another Colorado Organization": "Transfer in from another Colorado organization",
    "2024 Transfer from Out of State": "Transfer in from Out of State organization",
    "2024 Other": "Other: TNR / Protective Custody / Returns / Disaster Relief",

    # Outcomes
    "2024 Adoption": "Adoption",
    "2024 Returned to Owner (RTO)": "Returned To Owner (RTO)",
    "2024 Transfer to another Colorado Organization": "Transferred out to another Colorado organization",
    "2024 Transfer to Out of State": "Transferred to an Out of State organization",
    "2024 Other.1": "Other Live Outcomes (ie: tnr / snr)",
    "2024 Deaths": "Died",
    "2024 Euthanasia": "Shelter Euthanasia",
    "2024 Missing/Stolen": "Missing / Stolen",

    # End-of-year counts
    "12/31/2024 In Shelter": "Ending Shelter Count",
    "12/31/2024 In Foster Care": "Ending Foster Count"
}
```

```
# Applying mapping
data_2024_unpivoted["Metric Type"] = data_2024_unpivoted["Metric Type"].replace(metric_map)
```

In [13]:
```
# Checking Metric Type unique values after standardizing categories
data_2024_unpivoted["Metric Type"].unique()
```

Out[13]:
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died',
       'Shelter Euthanasia', 'Missing / Stolen', 'Ending Shelter Count',
       'Ending Foster Count'], dtype=object)
```

In [14]:
```
# Added a new column Flow Type to map and categorize records

# Define mapping
flow_map = {
    # Start Count
    "Beginning Shelter Count": "Start of Year Count",
    "Beginning Foster Count": "Start of Year Count",

    # Ending
    "Ending Shelter Count": "End of Year Count",
    "Ending Foster Count": "End of Year Count",

    # Intake
    "Stray": "Intake",
    "Owner Relinquished": "Intake",
    "Transfer in from another Colorado organization": "Intake",
    "Transfer in from Out of State organization": "Intake",
    "Other: TNR / Protective Custody / Returns / Disaster Relief": "Intake",

    # Outcome
    "Adoption": "Outcome (Positive)",
    "Returned To Owner (RTO)": "Outcome (Positive)",
```

```
    "Transferred out to another Colorado organization": "Outcome (Positive)",
    "Transferred to an Out of State organization": "Outcome (Positive)",
    "Other Live Outcomes (ie: tnr / snr)": "Outcome (Positive)",
    "Died": "Outcome (Negative)",
    "Missing / Stolen": "Outcome (Negative)",
    "Shelter Euthanasia": "Outcome (Negative)"
}

# Apply mapping to create new column
data_2024_unpivoted["Flow Type"] = data_2024_unpivoted["Metric Type"].map(flow_map)

# Renaming the columns
data_2024_unpivoted = data_2024_unpivoted.rename(columns={"Metric Type": "Event Type",
                                                          "Value": "Animal Count"})
```

In [15]:
```
data_2024_unpivoted.head()
```

Out[15]:

	Facility Name	Metric	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	1/1/2024\n Adult Dogs\n In Shelter	33	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	4 Paws 4 Life Rescue	1/1/2024\n Adult Dogs\n In Shelter	7	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	5280 Reptile Room North	1/1/2024\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	5280 Reptile Room West	1/1/2024\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	9 Lives Rescue	1/1/2024\n Adult Dogs\n In Shelter	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [16]:
```
# Removing 'Metric' column to avoid redundancy.
data_2024_unpivoted = data_2024_unpivoted.drop(columns=["Metric"], errors="ignore")
```

In [17]:
```
data_2024_unpivoted.head()
```

Out[17]:

	Facility Name	Animal Count	Animal Type	Event Type	Flow Type
0	2 Blondes All Breed Rescue, Inc.	33	Adult Dogs	Beginning Shelter Count	Start of Year Count
1	4 Paws 4 Life Rescue	7	Adult Dogs	Beginning Shelter Count	Start of Year Count
2	5280 Reptile Room North	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
3	5280 Reptile Room West	0	Adult Dogs	Beginning Shelter Count	Start of Year Count
4	9 Lives Rescue	0	Adult Dogs	Beginning Shelter Count	Start of Year Count

In [18]:
```
# Checking unique values of Animal Type column
data_2024_unpivoted["Animal Type"].unique()
```

Out[18]:
```
array(['Adult Dogs', 'Juvenile Dogs', 'Adult Cats', 'Juvenile Cats',
       'Birds', 'Small Mammals', 'Reptiles & Amphibians', 'Rabbits',
       'Other'], dtype=object)
```

In [19]:
```
# Creating two columns by splitting the 'Animal Type' column

# Species column
data_2024_unpivoted["Species"] = (data_2024_unpivoted["Animal Type"]
                                  .str.replace("Adult", "", case=False)
                                  .str.replace("Juvenile", "", case=False)
                                  .str.strip())

# Age Group column
data_2024_unpivoted["Age Group"] = np.where(
    data_2024_unpivoted["Animal Type"].str.contains("Adult", case=False, na=False), "Adult",
    np.where(data_2024_unpivoted["Animal Type"].str.contains("Juvenile", case=False, na=False), "Juvenile", "Unknown")
)

# Checking unique values
print(data_2024_unpivoted["Species"].unique())
print(data_2024_unpivoted["Age Group"].unique())
```

```
['Dogs' 'Cats' 'Birds' 'Small Mammals' 'Reptiles & Amphibians' 'Rabbits'
 'Other']
['Adult' 'Juvenile' 'Unknown']
```

In [20]:
```
# Removing 'Animal Type' column to avoid redundancy.
data_2024_unpivoted = data_2024_unpivoted.drop(columns=["Animal Type"], errors="ignore")
```

In [21]:
```
# Adding Reporting year column to the whole table
data_2024_unpivoted["Reporting Year"] = 2024
```

In [22]:
```
data_2024_unpivoted.head()
```

Out[22]:

	Facility Name	Animal Count	Event Type	Flow Type	Species	Age Group	Reporting Year
0	2 Blondes All Breed Rescue, Inc.	33	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2024
1	4 Paws 4 Life Rescue	7	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2024
2	5280 Reptile Room North	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2024
3	5280 Reptile Room West	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2024
4	9 Lives Rescue	0	Beginning Shelter Count	Start of Year Count	Dogs	Adult	2024

In [23]:
```
# Checking Animal Count by Flow Type & Event Type
data_2024_unpivoted.groupby(["Flow Type", "Event Type"])["Animal Count"].sum().reset_index()
```

Out[23]:

	Flow Type	Event Type	Animal Count
0	End of Year Count	Ending Foster Count	6088
1	End of Year Count	Ending Shelter Count	8835
2	Intake	Other: TNR / Protective Custody / Returns / Di...	16259
3	Intake	Owner Relinquished	47867
4	Intake	Stray	65886
5	Intake	Transfer in from Out of State organization	27485
6	Intake	Transfer in from another Colorado organization	18396
7	Outcome (Negative)	Died	1891
8	Outcome (Negative)	Missing / Stolen	72
9	Outcome (Negative)	Shelter Euthanasia	11138
10	Outcome (Positive)	Adoption	113917
11	Outcome (Positive)	Other Live Outcomes (ie: tnr / snr)	9097
12	Outcome (Positive)	Returned To Owner (RTO)	25525
13	Outcome (Positive)	Transferred out to another Colorado organization	15214
14	Outcome (Positive)	Transferred to an Out of State organization	799
15	Start of Year Count	Beginning Foster Count	6538
16	Start of Year Count	Beginning Shelter Count	10145

Note: In the 2024 summary document, some of the animal counts do not match the result values above. However, verification against the dataset confirms that these result values are correct.

In [24]:
```
# Sum of Animal Count per Event Type, Species, and Age Group
summary_2024 = data_2024_unpivoted.groupby(["Flow Type", "Event Type", "Species", "Age Group"])["Animal Count"].sum().reset_in
```

```
summary_2024.to_excel("summary_2024.xlsx", index=False)
```

In [25]:
```
# Verified that all the columns contain no null values
data_2024_unpivoted.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 55080 entries, 0 to 55079
Data columns (total 7 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Facility Name   55080 non-null  object
 1   Animal Count    55080 non-null  int64 
 2   Event Type      55080 non-null  object
 3   Flow Type       55080 non-null  object
 4   Species         55080 non-null  object
 5   Age Group       55080 non-null  object
 6   Reporting Year  55080 non-null  int64 
dtypes: int64(2), object(5)
memory usage: 2.9+ MB
```

In [26]:
```
data_2024_unpivoted.isnull().sum()
```

Out[26]:
```
Facility Name     0
Animal Count      0
Event Type        0
Flow Type         0
Species           0
Age Group         0
Reporting Year    0
dtype: int64
```

In [27]:
```
# Rearranging the columns
Cleaned_2024_Shelter_And_Rescue_Statistics_final = data_2024_unpivoted[["Facility Name", "Reporting Year", "Species",
                                                                        "Age Group", "Flow Type", "Event Type", "Animal Count"
```

In [28]:
```
Cleaned_2024_Shelter_And_Rescue_Statistics_final.to_excel("Cleaned_2024_Shelter_And_Rescue_Statistics_final.xlsx", index=False
print("Column info saved to Cleaned_2024_Shelter_And_Rescue_Statistics_final.xlsx")
```

```
Column info saved to Cleaned_2024_Shelter_And_Rescue_Statistics_final.xlsx
```

In [29]:
```
from IPython.display import FileLink
FileLink('Cleaned_2024_Shelter_And_Rescue_Statistics_final.xlsx')
```

Out[29]: Cleaned_2024_Shelter_And_Rescue_Statistics_final.xlsx

In [ ]:

Combining Clean Data Across All Years (2016-2024)

In [136…
```
import pandas as pd   # for data handling
import glob # helps grab all Excel files in a folder at once
```

In [138…
```
# Loading all year data files at once

folder_path = "D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Datasets/Cleaned
files = glob.glob(folder_path + "/*.xlsx")

print("Files found:", files)   # check all files detected
```

```
Files found: ['D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Datasets/Cleaned
Final\\Cleaned_2016_Shelter_And_Rescue_Statistics_final.xlsx', 'D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers pr
ojects/No Kill Colarado/Cleaned Datasets/Cleaned Final\\Cleaned_2017_Shelter_And_Rescue_Statistics_final.xlsx', 'D:/DATA ANALYT
ICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Datasets/Cleaned Final\\Cleaned_2018_Shelter_A
nd_Rescue_Statistics_final.xlsx', 'D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Clean
ed Datasets/Cleaned Final\\Cleaned_2019_Shelter_And_Rescue_Statistics_final.xlsx', 'D:/DATA ANALYTICS/Real world data projects/
Data ChangeMakers projects/No Kill Colarado/Cleaned Datasets/Cleaned Final\\Cleaned_2020_Shelter_And_Rescue_Statistics_final.xl
sx', 'D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Datasets/Cleaned Final\\Cl
eaned_2021_Shelter_And_Rescue_Statistics_final.xlsx', 'D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No
Kill Colarado/Cleaned Datasets/Cleaned Final\\Cleaned_2022_Shelter_And_Rescue_Statistics_final.xlsx', 'D:/DATA ANALYTICS/Real w
orld data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Datasets/Cleaned Final\\Cleaned_2023_Shelter_And_Rescue_
Statistics_final.xlsx', 'D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Cleaned Dataset
s/Cleaned Final\\Cleaned_2024_Shelter_And_Rescue_Statistics_final.xlsx']
```

The below code loops through each file, loads it as a DataFrame, saves it into a list, and finally merges everything into one big DataFrame.

In [139…
```
# creating an empty Python list named 'dfs' to store one DataFrame for each file (each year's data).
dfs = []
```

```
# loop through a list called 'files' to go through each file one by one.
for file in files:
    df = pd.read_excel(file)
# Adds the cleaned DataFrame to the list dfs one by one
    dfs.append(df)   # append instead of overwrite

# combine all: stacks all DataFrames together row-wise
all_data_2016_2024 = pd.concat(dfs, ignore_index=True)
```

In [140…
```
all_data_2016_2024.shape   # total rows and columns of combined data
```

Out[140… (370112, 13)

In [141…
```
all_data_2016_2024.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 370112 entries, 0 to 370111
Data columns (total 13 columns):
 #   Column                Non-Null Count   Dtype  
---  ------                --------------   -----  
 0   PACFA License Number  5496 non-null    object 
 1   Facility Name         370112 non-null  object 
 2   County                43457 non-null   object 
 3   City                  22951 non-null   object 
 4   Zip Code              22951 non-null   object 
 5   Latitude              22951 non-null   float64
 6   Longitude             22951 non-null   float64
 7   Reporting Year        370112 non-null  int64  
 8   Species               370112 non-null  object 
 9   Age Group             370112 non-null  object 
 10  Flow Type             370112 non-null  object 
 11  Event Type            370112 non-null  object 
 12  Animal Count          370112 non-null  int64  
dtypes: float64(2), int64(2), object(9)
memory usage: 36.7+ MB
```

In [143…
```
# Removing unwanted columns
all_data_2016_2024 = all_data_2016_2024.drop(columns=["PACFA License Number"], errors="ignore")
```

In [145…
```
# Checking rows count after dropping columns
all_data_2016_2024.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 370112 entries, 0 to 370111
Data columns (total 12 columns):
 #   Column          Non-Null Count   Dtype  
---  ------          --------------   -----  
 0   Facility Name   370112 non-null  object 
 1   County          43457 non-null   object 
 2   City            22951 non-null   object 
 3   Zip Code        22951 non-null   object 
 4   Latitude        22951 non-null   float64
 5   Longitude       22951 non-null   float64
 6   Reporting Year  370112 non-null  int64  
 7   Species         370112 non-null  object 
 8   Age Group       370112 non-null  object 
 9   Flow Type       370112 non-null  object 
 10  Event Type      370112 non-null  object 
 11  Animal Count    370112 non-null  int64  
dtypes: float64(2), int64(2), object(8)
memory usage: 33.9+ MB
```

In [146…
```
all_data_2016_2024.head()
```

Out[146...

	Facility Name	County	City	Zip Code	Latitude	Longitude	Reporting Year	Species	Age Group	Flow Type	Event Type	Animal Count
0	German Shepherd Rescue of Central Colorado	Park County	Hartsel Colorado	80449	35.500801	-117.947800	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	1
1	Doggy Dog World Rescue	Douglas County	Littleton	80125	39.612653	-105.016198	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	25
2	Surface Creek Shelter	Delta County	Cedaredge	81413	38.900738	-107.923767	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	6
3	Delta County Citizens for Animal Welfare and S...	Delta County	Delta	CO 81416	38.741684	-108.070175	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	7
4	Dalmatian Rescue of Colorado	Larimer County	Fort Collins	80526	40.588972	-105.082459	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	3

Ensuring that data from all years is fetched correctly to maintain completeness and accuracy in the consolidated dataset.

In [147...
```
# checking years
all_data_2016_2024["Reporting Year"].unique()
```

Out[147...
```
array([2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024], dtype=int64)
```

In [149...
```
# checking distinct species
all_data_2016_2024["Species"].unique()
```

Out[149...
```
array(['Dogs', 'Cats', 'Birds', 'Small Mammal', 'Reptiles', 'Rabbits',
       'Other', 'Reptiles & Amphibians', 'Small Mammals'], dtype=object)
```

In [150…
```
# Standardized Species names across the dataset
all_data_2016_2024["Species"] = all_data_2016_2024["Species"].replace({"Small Mammal": "Small Mammals"})

all_data_2016_2024["Species"].unique()
```

Out[150…
```
array(['Dogs', 'Cats', 'Birds', 'Small Mammals', 'Reptiles', 'Rabbits',
       'Other', 'Reptiles & Amphibians'], dtype=object)
```

In [151…
```
# checking distinct age group
all_data_2016_2024["Age Group"].unique()
```

Out[151…
```
array(['Adult', 'Juvenile', 'Unknown'], dtype=object)
```

In [152…
```
# checking distinct flow type
all_data_2016_2024["Flow Type"].unique()
```

Out[152…
```
array(['Start of Year Count', 'Intake', 'Outcome (Positive)',
       'Outcome (Negative)', 'End of Year Count'], dtype=object)
```

In [153…
```
# checking distinct event types
all_data_2016_2024["Event Type"].unique()
```

Out[153…
```
array(['Beginning Shelter Count', 'Beginning Foster Count', 'Stray',
       'Owner Relinquished', 'Owner Requested Euthanasia upon intake',
       'Transfer in from another Colorado organization',
       'Transfer in from Out of State organization',
       'Other: TNR / Protective Custody / Returns / Disaster Relief',
       'Adoption', 'Returned To Owner (RTO)',
       'Transferred out to another Colorado organization',
       'Transferred to an Out of State organization',
       'Other Live Outcomes (ie: tnr / snr)', 'Died', 'Missing / Stolen',
       'Shelter Euthanasia', 'Owner Requested Euthanasia',
       'Ending Shelter Count', 'Ending Foster Count'], dtype=object)
```

In [154…
```
# year wise records count
all_data_2016_2024["Reporting Year"].value_counts().sort_index()
```

Out[154…
```
Reporting Year
2016     5496
2017    17455
2018    20506
2019    53397
2020    54468
2021    54162
2022    56457
2023    53091
2024    55080
Name: count, dtype: int64
```

In [155…
```
# total number of distinct facility names across the dataset
all_data_2016_2024["Facility Name"].nunique()
```

Out[155…
```
1137
```

In [156…
```
# Saved for reference
all_data_2016_2024.to_excel("Combined_2016_to_2024_Shelter_And_Rescue_Statistics.xlsx", index=False)
```

Extracted unique facilities along with their respective location details from all years of data to create a standardized facility location reference document.

In [157…
```
# Select only facility + location columns (ignore missing ones safely)
location_cols = ["Facility Name", "County", "City", "Zip Code", "Latitude", "Longitude"]

# Extract from all year data
facility_locations = all_data_2016_2024[location_cols].copy()

# Drop duplicates
facility_locations = facility_locations.drop_duplicates()

# Keep only one record per facility (keeping the first non-null values)
facility_locations = facility_locations.groupby("Facility Name", as_index=False).first()
```

In [158…
```
facility_locations.shape
```

Out[158…
```
(1137, 6)
```

In [159…
```
facility_locations.head()
```

Out[159…

	Facility Name	County	City	Zip Code	Latitude	Longitude
0	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198
1	2 Blondes All Breed Rescue, Inc.	None	None	None	NaN	NaN
2	2nd Chance Vizsla Rescue Inc	None	None	None	NaN	NaN
3	2nd Chance Vizsla Rescue, Inc	None	None	None	NaN	NaN
4	2nd Chance Vizsla Rescue, Inc.	Larimer County	None	None	NaN	NaN

Validated that the total number of distinct facility names extracted matches the overall unique facility count across all years of data, confirming consistency in the standardized facility list.

In [160…
```
print(all_data_2016_2024["Facility Name"].nunique()) # from all year data
print(facility_locations["Facility Name"].nunique()) # from extracted data
```

```
1137
1137
```

In [161…
```
facility_locations.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1137 entries, 0 to 1136
Data columns (total 6 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   Facility Name  1137 non-null   object 
 1   County         546 non-null    object 
 2   City           384 non-null    object 
 3   Zip Code       384 non-null    object 
 4   Latitude       384 non-null    float64
 5   Longitude      384 non-null    float64
dtypes: float64(2), object(4)
memory usage: 53.4+ KB
```

In the dataset, some facility names contain ambiguities such as punctuation differences, inconsistent casing, common suffixes, and other irregularities. Due to these variations, it is not possible to reliably assign associated details like ZIP code, longitude, latitude, city, and county information to the correct facility names. To resolve this, facility names must first be standardized using defined cleaning rules, followed by fuzzy matching to align and consolidate the cleaned names.

In [162…

```
# handle punctuation, case, and common suffixes with cleaning rules

import re           # re is the regex module
import pandas as pd

def standardize_name(name):
    if pd.isna(name):            # if the facility name is blank (NaN), just return it
        return name
    name = name.lower().strip()  # makes everything lowercase and removes leading/trailing spaces.

    # Replace punctuation with space
    name = re.sub(r"[.,/*&()]", " ", name) # re.sub() is used for cleaning text (like replacing punctuations, suffixes, multip

    # Remove corporate suffixes
    name = re.sub(r"\b(inc|incorporated|llc|corp|corporation|shelter,?\s*inc)\b", "", name)

    # Normalize "and" / "&" / "/"
    name = name.replace("&", " and ").replace("/", " ")

    # Collapse multiple spaces
    name = re.sub(r"\s+", " ", name)   # \s+ finds multiple spaces and replaces them with a single space.

    return name.title().strip()       # .title() converts to title case (first letter capitalized).

# Storing cleaned names into a new column
facility_locations["Facility Name Std"] = facility_locations["Facility Name"].apply(standardize_name)
```

In [164…

```
facility_locations.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1137 entries, 0 to 1136
Data columns (total 7 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   Facility Name      1137 non-null   object 
 1   County             546 non-null    object 
 2   City               384 non-null    object 
 3   Zip Code           384 non-null    object 
 4   Latitude           384 non-null    float64
 5   Longitude          384 non-null    float64
 6   Facility Name Std  1137 non-null   object 
dtypes: float64(2), object(5)
memory usage: 62.3+ KB
```

In [165…

```python
# Handle close duplicates with fuzzy matching

from rapidfuzz import process

# gets all unique standardized facility names
unique_names = facility_locations["Facility Name Std"].unique()

master_map = {}

for name in unique_names:
    if master_map:  # only try matching if master_map has something
        result = process.extractOne(name, master_map.keys())
    else:
        result = None

    if result:
        match, score, _ = result
        if score > 90:  # strong match threshold
            master_map[name] = match
        else:
            master_map[name] = name  # new "master" entry
    else:
        master_map[name] = name  # first entry goes as is

# 'process.extractOne' compares name against all already-seen names in master_map.keys()
# If the similarity score (score) is > 90 (very close), we assume it’s the same facility.
```

```
# Otherwise, we keep it as a new entry.

# apply mapping to new column
facility_locations["Facility Name Final"] = facility_locations["Facility Name Std"].map(master_map)
```

In [166…
```
facility_locations.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1137 entries, 0 to 1136
Data columns (total 8 columns):
 #   Column              Non-Null Count  Dtype  
---  ------              --------------  -----  
 0   Facility Name       1137 non-null   object 
 1   County              546 non-null    object 
 2   City                384 non-null    object 
 3   Zip Code            384 non-null    object 
 4   Latitude            384 non-null    float64
 5   Longitude           384 non-null    float64
 6   Facility Name Std   1137 non-null   object 
 7   Facility Name Final  1137 non-null   object 
dtypes: float64(2), object(6)
memory usage: 71.2+ KB
```

In [167…
```
print(facility_locations["Facility Name"].nunique())  # total unique facility counts in original data
print(facility_locations["Facility Name Std"].nunique()) # total unique facility counts after manual (rules) cleaning
print(facility_locations["Facility Name Final"].nunique()) # total unique facility counts after fuzzy cleaning
```

```
1137
878
721
```

Grouped records by Facility Name Final and retained the first available values for County, City, Zip Code, Latitude, and Longitude to standardize the location information and consistently extend location details across identical facility names.

In [168…
```
# Choose the first non-null value for each standardized facility
location_info = facility_locations.groupby("Facility Name Final").agg({
    "County": "first",
    "City": "first",
    "Zip Code": "first",
    "Latitude": "first",
```

```
    "Longitude": "first"
}).reset_index()

# groupby("Facility Name Final") → groups all rows with the same standardized name.
# .agg({"County": "first", ...}) → picks the first non-null value for each column.

# Merge standardized location info
facility_locations = facility_locations.merge(location_info, on="Facility Name Final", how="left")
```

In [169…

```
facility_locations.head()
```

Out[169…

	Facility Name	County_x	City_x	Zip Code_x	Latitude_x	Longitude_x	Facility Name Std	Facility Name Final	County_y	City_y	Zip Code_y	Latitude_y	Longitude_y
0	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198
1	2 Blondes All Breed Rescue, Inc.	None	None	None	NaN	NaN	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198
2	2nd Chance Vizsla Rescue Inc	None	None	None	NaN	NaN	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN
3	2nd Chance Vizsla Rescue, Inc	None	None	None	NaN	NaN	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN
4	2nd Chance Vizsla Rescue, Inc.	Larimer County	None	None	NaN	NaN	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN

In [170…

```
facility_locations.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1137 entries, 0 to 1136
Data columns (total 13 columns):
 #   Column               Non-Null Count  Dtype  
---  ------               --------------  -----  
 0   Facility Name        1137 non-null   object 
 1   County_x             546 non-null    object 
 2   City_x               384 non-null    object 
 3   Zip Code_x           384 non-null    object 
 4   Latitude_x           384 non-null    float64
 5   Longitude_x          384 non-null    float64
 6   Facility Name Std    1137 non-null   object 
 7   Facility Name Final  1137 non-null   object 
 8   County_y             737 non-null    object 
 9   City_y               614 non-null    object 
 10  Zip Code_y           614 non-null    object 
 11  Latitude_y           614 non-null    float64
 12  Longitude_y          614 non-null    float64
dtypes: float64(4), object(9)
memory usage: 115.6+ KB
```

In [171…

```python
# Drop old columns to avoid duplication
facility_locations = facility_locations.drop(
    columns=["County_x", "City_x", "Zip Code_x", "Latitude_x", "Longitude_x"], errors="ignore")

# Renaming columns
facility_locations = facility_locations.rename(columns={
    "Facility Name": "Facility Name Original",
    "County_y": "County",
    "City_y": "City",
    "Zip Code_y": "Zip Code",
    "Latitude_y": "Latitude",
    "Longitude_y": "Longitude"
})

facility_locations.head()
```

Out[171…

	Facility Name Original	Facility Name Std	Facility Name Final	County	City	Zip Code	Latitude	Longitude
0	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198
1	2 Blondes All Breed Rescue, Inc.	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198
2	2nd Chance Vizsla Rescue Inc	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN
3	2nd Chance Vizsla Rescue, Inc	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN
4	2nd Chance Vizsla Rescue, Inc.	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN

In [172…

```
facility_locations.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1137 entries, 0 to 1136
Data columns (total 8 columns):
 #   Column                 Non-Null Count  Dtype  
---  ------                 --------------  -----  
 0   Facility Name Original  1137 non-null   object 
 1   Facility Name Std       1137 non-null   object 
 2   Facility Name Final     1137 non-null   object 
 3   County                  737 non-null    object 
 4   City                    614 non-null    object 
 5   Zip Code                614 non-null    object 
 6   Latitude                614 non-null    float64
 7   Longitude               614 non-null    float64
dtypes: float64(2), object(6)
memory usage: 71.2+ KB
```

Further standardized the `facility_locations` table by enriching it with additional information from the Active PACFA Facilities master file available on the website.

In [173…
```
# Loading Active PACFA master file
pacfa_master = pd.read_excel("D:/DATA ANALYTICS/Real world data projects/Data ChangeMakers projects/No Kill Colarado/Raw datas
```

In [174…
```
# checking available columns from both data sets
print(pacfa_master.columns)
print(facility_locations.columns)
```

```
Index(['Account Name', 'DBA', 'City', 'State', 'County',
       'Business License App Category Name', 'Expire Date'],
      dtype='object')
Index(['Facility Name Original', 'Facility Name Std', 'Facility Name Final',
       'County', 'City', 'Zip Code', 'Latitude', 'Longitude'],
      dtype='object')
```

In [175…
```
pacfa_master.head()
```

Out[175…

	Account Name	DBA	City	State	County	Business License App Category Name	Expire Date
0	101 Clean Dogs "LLC"	NaN	Erie	CO	BOULDER	Pet Grooming Facility - Mobile Groomer	3/2/2026
1	1480 Cafe, LLC	Pet Supplies Plus	Denver	CO	DENVER	Retail / Wholesale of Pet Animals	3/2/2026
2	1480 Cafe, LLC	Pet Supplies Plus	Denver	CO	DENVER	Pet Grooming Facility - Primary Facility Owner	3/2/2026
3	17th and Pawlished LLC	NaN	Greeley	CO	WELD	Pet Grooming Facility - Primary Facility Owner	3/2/2026
4	2 Blondes All Breed Rescue, Inc.	NaN	Lakewood	CO	JEFFERSON	Pet Animal Large Rescue	3/2/2026

In [176…
```
pacfa_master.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2958 entries, 0 to 2957
Data columns (total 7 columns):
 #   Column                              Non-Null Count  Dtype 
---  ------                              --------------  ----- 
 0   Account Name                        2958 non-null   object
 1   DBA                                 1656 non-null   object
 2   City                                2958 non-null   object
 3   State                               2958 non-null   object
 4   County                              2958 non-null   object
 5   Business License App Category Name  2958 non-null   object
 6   Expire Date                         2958 non-null   object
dtypes: object(7)
memory usage: 161.9+ KB
```

Some facility names may not match between the `pacfa_master` and `facility_locations` tables due to naming inconsistencies, even if they refer to the same facility. Since `facility_locations` has already been standardized through cleaning, we need to apply the same standardization rules to the `pacfa_master` table to maximize the number of facilities for which location information can be obtained.

In [177…

```
# Standardize pacfa_master facility names
# to make matching easier, so we can apply a similar cleaning function

def standardize_master_name(name):
    if pd.isna(name):
        return name
    name = name.lower().strip()
    name = re.sub(r"[.,/*&()]", " ", name)
    name = re.sub(r"\b(inc|incorporated|llc|corp|corporation|shelter,?\s*inc)\b", "", name)
    name = name.replace("&", " and ").replace("/", " ")
    name = re.sub(r"\s+", " ", name)
    return name.title().strip()

pacfa_master["Facility Name Std"] = pacfa_master["Account Name"].apply(standardize_master_name)
```

In [178…

```
pacfa_master.head()
```

Out[178…

	Account Name	DBA	City	State	County	Business License App Category Name	Expire Date	Facility Name Std
0	101 Clean Dogs "LLC"	NaN	Erie	CO	BOULDER	Pet Grooming Facility - Mobile Groomer	3/2/2026	101 Clean Dogs ""
1	1480 Cafe, LLC	Pet Supplies Plus	Denver	CO	DENVER	Retail / Wholesale of Pet Animals	3/2/2026	1480 Cafe
2	1480 Cafe, LLC	Pet Supplies Plus	Denver	CO	DENVER	Pet Grooming Facility - Primary Facility Owner	3/2/2026	1480 Cafe
3	17th and Pawlished LLC	NaN	Greeley	CO	WELD	Pet Grooming Facility - Primary Facility Owner	3/2/2026	17Th And Pawlished
4	2 Blondes All Breed Rescue, Inc.	NaN	Lakewood	CO	JEFFERSON	Pet Animal Large Rescue	3/2/2026	2 Blondes All Breed Rescue

After standardizing facility names in the `pacfa_master` table, extracted the available facility location information by merging it with the `facility_locations` table to enrich the dataset.

In [179…

```
# Merge on standardized names
facility_locations_final = facility_locations.merge(
    pacfa_master[["Facility Name Std", "DBA", "County", "City", "State", "Business License App Category Name", "Expire Date"]],
    left_on="Facility Name Final",
    right_on="Facility Name Std",
    how="left",
    suffixes=('_fac', '_pacfa')
)
```

In [180…

```
facility_locations_final.head()
```

Out[180...

	Facility Name Original	Facility Name Std_fac	Facility Name Final	County_fac	City_fac	Zip Code	Latitude	Longitude	Facility Name Std_pacfa	DBA	County_pacfa	City_pacfa	State	Busi Lic Cate N
0	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198	2 Blondes All Breed Rescue	NaN	JEFFERSON	Lakewood	CO	Ar L Re
1	2 Blondes All Breed Rescue, Inc.	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198	2 Blondes All Breed Rescue	NaN	JEFFERSON	Lakewood	CO	Ar L Re
2	2nd Chance Vizsla Rescue Inc	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
3	2nd Chance Vizsla Rescue, Inc	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
4	2nd Chance Vizsla Rescue, Inc.	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

After merging, city and county names come from both datasets—some match, some do not, and some are missing in the facility_locations table but can be filled from pacfa_master if available.

To standardize this City and County information for each facility, the code creates City_final and County_final columns, retaining `facility_locations` values when they match, filling missing values from `pacfa_master`, and resolving mismatches by preferring `pacfa_master` data.

In [181…

```python
# Function to decide final city/county
def decide_city_county(row):
    city_fac = row["City_fac"]
    county_fac = row["County_fac"]
    city_pacfa = row["City_pacfa"]
    county_pacfa = row["County_pacfa"]

    # Start with facility values
    final_city = city_fac
    final_county = county_fac

    # Case 1: missing -> fill from pacfa
    if pd.isna(final_city) and pd.notna(city_pacfa):
        final_city = city_pacfa
    if pd.isna(final_county) and pd.notna(county_pacfa):
        final_county = county_pacfa

    # Case 2: mismatch -> prefer pacfa
    if (pd.notna(city_pacfa) and final_city != city_pacfa):
        final_city = city_pacfa
    if (pd.notna(county_pacfa) and final_county != county_pacfa):
        final_county = county_pacfa

    # Store clean versions
    row["City_final"] = str(final_city).title() if pd.notna(final_city) else pd.NA
    row["County_final"] = str(final_county).title() if pd.notna(final_county) else pd.NA

    return row

# Apply row-wise updates
facility_locations_final = facility_locations_final.apply(decide_city_county, axis=1)

# facility_locations_final now has:
# - City_fac, County_fac (original facility info)
# - City_pacfa, County_pacfa (from PACFA master)
```

```
# - City_final, County_final (standardized)
# - Zip, Lat, Long remain unchanged
```

In [182…

```
facility_locations_final.head()
```

Out[182...

	Facility Name Original	Facility Name Std_fac	Facility Name Final	County_fac	City_fac	Zip Code	Latitude	Longitude	Facility Name Std_pacfa	DBA	County_pacfa	City_pacfa	State	Busi Lic Cate N
0	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198	2 Blondes All Breed Rescue	NaN	JEFFERSON	Lakewood	CO	Ar L Re
1	2 Blondes All Breed Rescue, Inc.	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Douglas County	Littleton	80126	39.612653	-105.016198	2 Blondes All Breed Rescue	NaN	JEFFERSON	Lakewood	CO	Ar L Re
2	2nd Chance Vizsla Rescue Inc	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
3	2nd Chance Vizsla Rescue, Inc	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
4	2nd Chance Vizsla Rescue, Inc.	2Nd Chance Vizsla Rescue	2Nd Chance Vizsla Rescue	Larimer County	None	None	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

In [183...

```
facility_locations_final.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1308 entries, 0 to 1307
Data columns (total 17 columns):
 #   Column                              Non-Null Count  Dtype
---  ------                              --------------  -----
 0   Facility Name Original              1308 non-null   object
 1   Facility Name Std_fac               1308 non-null   object
 2   Facility Name Final                 1308 non-null   object
 3   County_fac                          818 non-null    object
 4   City_fac                            675 non-null    object
 5   Zip Code                            675 non-null    object
 6   Latitude                            675 non-null    float64
 7   Longitude                           675 non-null    float64
 8   Facility Name Std_pacfa             628 non-null    object
 9   DBA                                 175 non-null    object
 10  County_pacfa                        628 non-null    object
 11  City_pacfa                          628 non-null    object
 12  State                               628 non-null    object
 13  Business License App Category Name  628 non-null    object
 14  Expire Date                         628 non-null    object
 15  City_final                          976 non-null    object
 16  County_final                        1054 non-null   object
dtypes: float64(2), object(15)
memory usage: 173.8+ KB
```

After creating the `City_final` and `County_final` columns, the PACFA master values are used to fill in any missing or mismatched entries in the `facility_locations` table. As a result, some city and county values originally present in `facility_locations` have been replaced with PACFA values to ensure consistency and accuracy. However, because the previous Zip Code, Longitude, and Latitude values were associated with the original city and county entries, these geographic fields may now be incorrect for the updated records. Therefore, it is necessary to clear these column values so that they can be recalculated or updated based on the standardized city and county information.

To ensure data consistency and accuracy, we first need to standardize city and county names across all relevant columns—`County_fac` and `City_fac`, which originate from the `facility_locations` table, and `County_final` and `City_final`, which are the consolidated final values after merging with the PACFA master data. Standardizing these names across both the source and final columns will allow accurate matching of records, ensure consistency in the final city and county values, and prevent incorrect associations with Zip Code, Longitude, and Latitude values when updates or corrections are applied.

In [184…
```
# checking County_final unique values
print(facility_locations_final["County_final"].unique())
```

```
['Jefferson' 'Larimer County' 'Douglas' <NA> 'El Paso' 'Yuma' 'Denver'
 'La Plata County' 'Pitkin County' 'Pueblo County' 'Huerfano County'
 'Adams' 'Teller County' 'Boulder' 'El Paso County' 'Morgan'
 'Arapahoe County' 'Larimer' 'Weld' 'Sedgwick County' 'Boulder County'
 'Mesa' 'Park' 'Pueblo' 'Chaffee County' 'Fremont County' 'Adams County'
 'Chaffee' 'Out Of State' 'Moffat County' 'Denver County' 'Delta County'
 'Montezuma' 'Las Animas' 'Douglas County' 'Park County' 'Otero County'
 'Bent County' 'Garfield County' 'Weld County' 'Arapahoe' 'La Plata'
 'Alamosa County' 'Mesa County' 'Archuleta' 'Jefferson County'
 'Kit Carson County' 'Otero' 'Bent' 'Lake' 'Montrose' 'Montrose County'
 'Huerfano' 'Yuma County' 'Garfield' 'Elbert County' 'Broomfield County'
 'Gunnison' 'Clear Creek' 'Delta' 'Montezuma County' 'Alamosa' 'Eagle'
 'Kit Carson' 'Elbert' 'Morgan County' 'Grand' 'Gunnison County'
 'Eagle County' 'Fremont' 'Prowers' 'Lake County' 'Logan' 'Pitkin'
 'Broomfield' 'Las Animas County' 'Gilpin' 'My County Is Not Listed'
 'Lincoln County' 'Rio Blanco County' 'Routt' 'Costilla' 'Lincoln'
 'Ouray County' 'Summit' 'Teller' 'San Miguel' 'Washington' 'Cheyenne'
 'Crowley' 'Custer' 'Dolores County' 'Rio Blanco' 'Baca'
 'San Miguel County' 'Jackson County' 'Custer County' 'Rio Grande'
 'Baca County']
```

In [186…
```
# checking County_fac unique values
print(facility_locations_final["County_fac"].unique())
```

```
['Douglas County' 'Larimer County' None 'El Paso County' 'La Plata County'
 'Pitkin County' 'Pueblo County' 'Huerfano County' 'Adams County'
 'Teller County' 'Arapahoe County' 'Jefferson County' 'Sedgwick County'
 'Boulder County' 'Summit County' 'Chaffee County' 'Fremont County'
 'Moffat County' 'Denver County' 'Weld County' 'Delta County'
 'Las Animas County' 'Park County' 'Otero County' 'Bent County'
 'Yuma County' 'Garfield County' 'Alamosa County' 'Mesa County'
 'Montezuma County' 'Morgan County' 'Kit Carson County' 'Lake County'
 'Montrose County' 'Elbert County' 'Broomfield County'
 'Clear Creek County' 'Eagle County' 'Grand County' 'Gunnison County'
 'Archuleta County' 'Prowers County' 'Logan County'
 'My County is not listed' 'Lincoln County' 'Rio Blanco County'
 'Routt County' 'Costilla County' 'Ouray County' 'Washington County'
 'Cheyenne County' 'Crowley County' 'Custer County' 'Dolores County'
 'Baca County' 'San Miguel County' 'Jackson County' 'Rio Grande County']
```

In [187…

```python
# Standardising both county columns

def standardize_county(county):
    if pd.isna(county):
        return pd.NA

    county = str(county).strip()

    # Handle special cases
    if county.lower() in ["out of state", "outofstate"]:
        return "Out of State"
    if "not listed" in county.lower():
        return "Unknown"

    # Remove duplicate casing issues
    county = county.title()

    # Ensure "County" suffix for normal values
    if not county.endswith("County") and county not in ["Out of State", "Unknown"]:
        county = county + " County"

    return county

# Apply to both columns
```

```
facility_locations_final["County_final"] = facility_locations_final["County_final"].apply(standardize_county)
facility_locations_final["County_fac"] = facility_locations_final["County_fac"].apply(standardize_county)
```

In [189…

```
# Check cleaned unique values
print("County_final unique:", facility_locations_final["County_final"].unique())
print("County_fac unique:", facility_locations_final["County_fac"].unique())
```

```
County_final unique: ['Jefferson County' 'Larimer County' 'Douglas County' <NA>
 'El Paso County' 'Yuma County' 'Denver County' 'La Plata County'
 'Pitkin County' 'Pueblo County' 'Huerfano County' 'Adams County'
 'Teller County' 'Boulder County' 'Morgan County' 'Arapahoe County'
 'Weld County' 'Sedgwick County' 'Mesa County' 'Park County'
 'Chaffee County' 'Fremont County' 'Out of State' 'Moffat County'
 'Delta County' 'Montezuma County' 'Las Animas County' 'Otero County'
 'Bent County' 'Garfield County' 'Alamosa County' 'Archuleta County'
 'Kit Carson County' 'Lake County' 'Montrose County' 'Elbert County'
 'Broomfield County' 'Gunnison County' 'Clear Creek County' 'Eagle County'
 'Grand County' 'Prowers County' 'Logan County' 'Gilpin County' 'Unknown'
 'Lincoln County' 'Rio Blanco County' 'Routt County' 'Costilla County'
 'Ouray County' 'Summit County' 'San Miguel County' 'Washington County'
 'Cheyenne County' 'Crowley County' 'Custer County' 'Dolores County'
 'Baca County' 'Jackson County' 'Rio Grande County']
County_fac unique: ['Douglas County' 'Larimer County' <NA> 'El Paso County' 'La Plata County'
 'Pitkin County' 'Pueblo County' 'Huerfano County' 'Adams County'
 'Teller County' 'Arapahoe County' 'Jefferson County' 'Sedgwick County'
 'Boulder County' 'Summit County' 'Chaffee County' 'Fremont County'
 'Moffat County' 'Denver County' 'Weld County' 'Delta County'
 'Las Animas County' 'Park County' 'Otero County' 'Bent County'
 'Yuma County' 'Garfield County' 'Alamosa County' 'Mesa County'
 'Montezuma County' 'Morgan County' 'Kit Carson County' 'Lake County'
 'Montrose County' 'Elbert County' 'Broomfield County'
 'Clear Creek County' 'Eagle County' 'Grand County' 'Gunnison County'
 'Archuleta County' 'Prowers County' 'Logan County' 'Unknown'
 'Lincoln County' 'Rio Blanco County' 'Routt County' 'Costilla County'
 'Ouray County' 'Washington County' 'Cheyenne County' 'Crowley County'
 'Custer County' 'Dolores County' 'Baca County' 'San Miguel County'
 'Jackson County' 'Rio Grande County']
```

In [191…

```
# checking City_final unique values
print(facility_locations_final["City_final"].unique())
```

```
['Lakewood' <NA> 'Sedalia' 'Colorado Springs' 'Eckley' 'Denver'
 'Fort Collins ' 'Durango ' 'Aspen ' 'Walsenburg ' 'Brighton'
 'Woodland Park' 'Northglenn' 'Boulder' 'Fort Collins' 'Colorado Springs '
 'Fort Morgan' 'Englewood ' 'Westminster' 'Littleton' 'Berthoud' 'Greeley'
 'Julesburg' 'Longmont' 'Fruita' 'Fairplay' 'Pueblo' 'Durango'
 'Buena Vista ' 'Poncha Springs' 'Canon City' 'Aurora' 'Buena Vista'
 'Aztec' 'Craig' 'Denver ' 'Wellington' 'Larkspur' 'Crawford'
 'Black Forest ' 'Cortez' 'Yoder' 'Boncarbo' 'Superior' 'Franktown '
 'La Junta ' 'Las Animas' 'Yuma' 'Thornton' 'Loma' 'Bailey' 'Eaton '
 'Alamosa' 'Grand Junction' 'Cheyenne' 'Pagosa Springs' 'Evergreen'
 'Evergreen ' 'Cripple Creek' 'Walsenburg' 'Brush' 'Burlington' 'La Junta'
 'Leadville' 'Montrose' 'Rocky Ford' 'Trrinidad' 'Wray' 'Lone Tree'
 'Glenwood Springs' 'Elizabeth' 'Rye' 'Englewood' 'Henderson' 'Morrison'
 'Johnstown' 'Broomfield' 'Loveland' 'Arvada' 'Gunnison' 'Parker'
 'Thornton ' 'Centenial ' 'Golden' 'Dumont' 'Aurora ' 'Delta' 'Matheson'
 'Commerce City' 'Calhan' 'Broomfield ' 'Silt' 'Castle Rock' 'Loveland '
 'Eagle' 'Eaton' 'Elbert' 'Estes Park' 'Fort Lupton' 'Arvada ' 'Peyton'
 'Cedaredge ' 'Fort Morgan ' 'Rifle ' 'Rifle' 'Hartsel Colorado' 'Granby'
 'Colorado City' 'Monument' 'Gunnison ' 'Salida ' 'Alamosa '
 'Manitou Springs' 'Byers ' 'Empire' 'Minturn' 'Ault' 'Naturita' 'Evans'
 'Lake George' 'Pine' 'Pueblo West ' 'Paonia' 'Windsor ' 'Elizabeth '
 'Littleton ' 'Lamar' 'Fountain' 'Fort Colllins' 'Leadville ' 'Erie'
 'Log Lane Village' 'Sterling' 'Aspen' 'Monument ' 'Carr' 'Whitewater'
 'Miami' 'Centennial' 'Mead' 'Lakewood ' 'Lafayette' 'Laporte' 'Trinidad '
 'Crested Butte' 'Spring' 'Highlands Ranch' 'Pampa' 'Fowler' 'Pueblo West'
 'Wheat Ridge ' 'Severance' 'Ault ' 'Rush' 'Westminster ' 'Rangely'
 'Carbondale ' 'Houston' 'Steamboat Springs' 'Dolores ' 'San Luis' 'Genoa'
 'Ridgway ' 'Bennett' 'Oak Creek' 'Black Forest' 'Longmont ' 'De Beque'
 'Frisco' 'Cedaredge' 'Divide' 'Naturita ' 'Avon' 'Ophir' 'Delta ' 'Lewis'
 'Penrose' 'Poncha Springs ' 'Northglenn ' 'Akron' 'Cheyenne Wells'
 'Ordway' 'Westcliffe' 'Dove Creek' 'La Veta ' 'La Veta' 'Limon' 'Meeker'
 'Platteville' 'Springfield' 'Telluride' 'Walden ' 'Walsh' 'Moab'
 'Boulder ' 'Monte Vista' 'Livermore' 'Wescliffe' 'Kersey' 'Ignacio'
 'Ignacio ' 'Dolores']
```

In [193…

```
# checking City_final unique values
print(facility_locations_final["City_fac"].unique())
```

```
['Littleton' None 'Sedalia' 'Colorado Springs' 'Fort Collins ' 'DURANGO '
 'Aspen ' 'Walsenburg ' 'Brighton ' 'Woodland Park' 'Fort Collins'
 'Colorado Springs ' 'englewood ' 'Westminster' 'Lakewood ' 'Julesburg'
 'Longmont' 'Breckenridge' 'Pueblo' 'Durango' 'Buena Vista '
 'Poncha Springs' 'Canon City' 'Aurora' 'Craig' 'Denver ' 'Windsor'
 'Larkspur' 'Crawford' 'Black Forest ' 'Yoder' 'Boncargo' 'Boulder'
 'Franktown ' 'La Junta ' 'Las Animas' 'Yuma ' 'Arvada ' 'Bailey' 'Eaton '
 'Lakewood' 'alamosa' 'Grand Junction' 'Littleton ' 'Evergreen'
 'Evergreen ' 'Cortez' 'Cripple Creek' 'Walsenburg' 'Westminster '
 'Brighton' 'Brush' 'Burlington' 'Montrose' 'Rocky Ford' 'Wray'
 'Glenwood Springs' 'Elizabeth' 'Englewood' 'Broomfield' 'Loveland '
 'Arvada' 'Strasburg' 'Aurora ' 'Thornton ' 'Centenial ' 'Loveland'
 'Dumont' 'Delta' 'Denver' 'matheson' 'Keenesburg' 'Rye' 'Broomfield '
 'Silt' 'Castle Rock' 'Eagle ' 'Eagle' 'Eaton' 'Calhan' 'Estes Park'
 'Fort Lupton' 'ARVADA ' 'Golden' 'Cortez ' 'Fort Morgan' 'Cedaredge '
 'Fort Morgan ' 'Rifle ' 'Rifle' 'Hartsel Colorado' 'Granby' 'Loma'
 'Peyton' 'Colorado City' 'MONUMENT ' 'Gunnison ' 'Gunnison' 'Salida '
 'Alamosa ' 'Byers ' 'Minturn' 'Berthoud ' 'Ault' 'Nucla' 'Pagosa Springs'
 'Evans' 'Lake George' 'Pine' 'Henderson' 'Pueblo West ' 'Paonia'
 'Windsor ' 'Durango ' 'Elizabeth ' 'Fountain' 'Fort Colllins' 'Leadville'
 'Leadville ' 'Sterling' 'Aspen' 'Monument ' 'Carr' 'Whitewater'
 'Engewood ' 'Centennial' 'Commerce City' 'Lafayette ' 'Parker'
 'Trinidad ' 'Monument' 'Crested Butte' 'CENTENNIAL ' 'Estes Park '
 'Golden ' 'Fowler' 'pueblo west' 'Wheat Ridge ' 'Ault ' 'Rush'
 'westminster ' 'Thornton' 'Parker ' 'Rangely' 'Carbondale ' 'Pueblo West'
 'Greeley' 'Steamboat Springs' 'Dolores ' 'Manitou Springs' 'San Luis'
 'Severance' 'Lamar' 'Ridgway ' 'Black Forest' 'Longmont ' 'Frisco '
 'Cedaredge' 'Divide' 'Wellington' 'NATURITA ' 'Delta ' 'Poncha Springs '
 'Northglenn ' 'Akron ' 'Cheyenne Wells' 'Westcliffe' 'Dove Creek'
 'La Veta ' 'La Veta' 'Limon' 'Meeker ' 'Platteville' 'Rangely '
 'Springfield' 'Telluride' 'Walden ' 'Boulder ' 'Monte Vista' 'Alamosa'
 'Erie' 'Ignacio ' 'Centennial ' 'Dolores']
```

In [194…

```python
# Standardising both city columns

# Define a cleaning function for city names
def clean_city(name):
    if pd.isna(name):
        return None
    name = str(name).strip()                # remove leading/trailing spaces
    name = " ".join(name.split())           # collapse multiple spaces into one
```

```
    name = name.title()                    # title case (Denver, Colorado Springs)

    # Fix common typos and variants
    corrections = {
        "Ft Collins": "Fort Collins",
        "Fort Colllins": "Fort Collins",
        "Centenial": "Centennial",
        "Centenial ": "Centennial",
        "Centenial": "Centennial",
        "Centennial ": "Centennial",
        "Centennal": "Centennial",
        "Centenniel": "Centennial",
        "Englewood ": "Englewood",
        "Engewood": "Englewood",
        "Englewood ": "Englewood",
        "Arvada ": "Arvada",
        "Arvada  ": "Arvada",
        "Arvada  ": "Arvada",
        "Arvada": "Arvada",
        "Boncargo": "Boncarbo",
        "Durango ": "Durango",
        "Durango  ": "Durango",
        "Pueblo West ": "Pueblo West",
        "Pueblo West  ": "Pueblo West",
        "Pueblo West": "Pueblo West",
        "Gunnison ": "Gunnison",
        "Gunnison  ": "Gunnison",
        "Naturita ": "Naturita",
        "Naturita  ": "Naturita",
        "La Veta ": "La Veta",
        "La Veta  ": "La Veta",
        "Ft Morgan": "Fort Morgan",
        "Ft Lupton": "Fort Lupton",
        "Trinidad ": "Trinidad",
        "Trinidad  ": "Trinidad"
    }
    return corrections.get(name, name)

# Apply cleaning to both columns
```

```
facility_locations_final["City_fac"] = facility_locations_final["City_fac"].apply(clean_city)
facility_locations_final["City_final"] = facility_locations_final["City_final"].apply(clean_city)
```

In [196…

```
# Check cleaned unique values
print("City_fac unique:", facility_locations_final["City_fac"].unique())
print("City_final unique:", facility_locations_final["City_final"].unique())
```

```
City_fac unique: ['Littleton' None 'Sedalia' 'Colorado Springs' 'Fort Collins' 'Durango'
 'Aspen' 'Walsenburg' 'Brighton' 'Woodland Park' 'Englewood' 'Westminster'
 'Lakewood' 'Julesburg' 'Longmont' 'Breckenridge' 'Pueblo' 'Buena Vista'
 'Poncha Springs' 'Canon City' 'Aurora' 'Craig' 'Denver' 'Windsor'
 'Larkspur' 'Crawford' 'Black Forest' 'Yoder' 'Boncarbo' 'Boulder'
 'Franktown' 'La Junta' 'Las Animas' 'Yuma' 'Arvada' 'Bailey' 'Eaton'
 'Alamosa' 'Grand Junction' 'Evergreen' 'Cortez' 'Cripple Creek' 'Brush'
 'Burlington' 'Montrose' 'Rocky Ford' 'Wray' 'Glenwood Springs'
 'Elizabeth' 'Broomfield' 'Loveland' 'Strasburg' 'Thornton' 'Centennial'
 'Dumont' 'Delta' 'Matheson' 'Keenesburg' 'Rye' 'Silt' 'Castle Rock'
 'Eagle' 'Calhan' 'Estes Park' 'Fort Lupton' 'Golden' 'Fort Morgan'
 'Cedaredge' 'Rifle' 'Hartsel Colorado' 'Granby' 'Loma' 'Peyton'
 'Colorado City' 'Monument' 'Gunnison' 'Salida' 'Byers' 'Minturn'
 'Berthoud' 'Ault' 'Nucla' 'Pagosa Springs' 'Evans' 'Lake George' 'Pine'
 'Henderson' 'Pueblo West' 'Paonia' 'Fountain' 'Leadville' 'Sterling'
 'Carr' 'Whitewater' 'Commerce City' 'Lafayette' 'Parker' 'Trinidad'
 'Crested Butte' 'Fowler' 'Wheat Ridge' 'Rush' 'Rangely' 'Carbondale'
 'Greeley' 'Steamboat Springs' 'Dolores' 'Manitou Springs' 'San Luis'
 'Severance' 'Lamar' 'Ridgway' 'Frisco' 'Divide' 'Wellington' 'Naturita'
 'Northglenn' 'Akron' 'Cheyenne Wells' 'Westcliffe' 'Dove Creek' 'La Veta'
 'Limon' 'Meeker' 'Platteville' 'Springfield' 'Telluride' 'Walden'
 'Monte Vista' 'Erie' 'Ignacio']
City_final unique: ['Lakewood' None 'Sedalia' 'Colorado Springs' 'Eckley' 'Denver'
 'Fort Collins' 'Durango' 'Aspen' 'Walsenburg' 'Brighton' 'Woodland Park'
 'Northglenn' 'Boulder' 'Fort Morgan' 'Englewood' 'Westminster'
 'Littleton' 'Berthoud' 'Greeley' 'Julesburg' 'Longmont' 'Fruita'
 'Fairplay' 'Pueblo' 'Buena Vista' 'Poncha Springs' 'Canon City' 'Aurora'
 'Aztec' 'Craig' 'Wellington' 'Larkspur' 'Crawford' 'Black Forest'
 'Cortez' 'Yoder' 'Boncarbo' 'Superior' 'Franktown' 'La Junta'
 'Las Animas' 'Yuma' 'Thornton' 'Loma' 'Bailey' 'Eaton' 'Alamosa'
 'Grand Junction' 'Cheyenne' 'Pagosa Springs' 'Evergreen' 'Cripple Creek'
 'Brush' 'Burlington' 'Leadville' 'Montrose' 'Rocky Ford' 'Trrinidad'
 'Wray' 'Lone Tree' 'Glenwood Springs' 'Elizabeth' 'Rye' 'Henderson'
 'Morrison' 'Johnstown' 'Broomfield' 'Loveland' 'Arvada' 'Gunnison'
 'Parker' 'Centennial' 'Golden' 'Dumont' 'Delta' 'Matheson'
 'Commerce City' 'Calhan' 'Silt' 'Castle Rock' 'Eagle' 'Elbert'
 'Estes Park' 'Fort Lupton' 'Peyton' 'Cedaredge' 'Rifle'
 'Hartsel Colorado' 'Granby' 'Colorado City' 'Monument' 'Salida'
 'Manitou Springs' 'Byers' 'Empire' 'Minturn' 'Ault' 'Naturita' 'Evans'
 'Lake George' 'Pine' 'Pueblo West' 'Paonia' 'Windsor' 'Lamar' 'Fountain'
 'Erie' 'Log Lane Village' 'Sterling' 'Carr' 'Whitewater' 'Miami' 'Mead'
```

```
'Lafayette' 'Laporte' 'Trinidad' 'Crested Butte' 'Spring'
'Highlands Ranch' 'Pampa' 'Fowler' 'Wheat Ridge' 'Severance' 'Rush'
'Rangely' 'Carbondale' 'Houston' 'Steamboat Springs' 'Dolores' 'San Luis'
'Genoa' 'Ridgway' 'Bennett' 'Oak Creek' 'De Beque' 'Frisco' 'Divide'
'Avon' 'Ophir' 'Lewis' 'Penrose' 'Akron' 'Cheyenne Wells' 'Ordway'
'Westcliffe' 'Dove Creek' 'La Veta' 'Limon' 'Meeker' 'Platteville'
'Springfield' 'Telluride' 'Walden' 'Walsh' 'Moab' 'Monte Vista'
'Livermore' 'Wescliffe' 'Kersey' 'Ignacio']
```

Now, we will retain all the original data, but for rows where both city and county do not match between the `_fac` and `_final` columns, the location-specific columns (Zip Code, Latitude, and Longitude) will be cleared. This ensures that facility_locations_final contains accurate location information only for rows where the city and county either match or have been correctly updated with valid `_final` values.

In [197…

```
# Define the location columns to clear
location_cols = ["Zip Code", "Latitude", "Longitude"]

# Create a boolean mask where both city and county do NOT match
mask = ~(
    (facility_locations_final["City_fac"] == facility_locations_final["City_final"]) &
    (facility_locations_final["County_fac"] == facility_locations_final["County_final"])
)

# Clear location info where mask is True
facility_locations_final.loc[mask, location_cols] = None
```

In [198…

```
facility_locations_final.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1308 entries, 0 to 1307
Data columns (total 17 columns):
 #   Column                              Non-Null Count  Dtype
---  ------                              --------------  -----
 0   Facility Name Original              1308 non-null   object
 1   Facility Name Std_fac               1308 non-null   object
 2   Facility Name Final                 1308 non-null   object
 3   County_fac                          818 non-null    object
 4   City_fac                            675 non-null    object
 5   Zip Code                            580 non-null    object
 6   Latitude                            580 non-null    float64
 7   Longitude                           580 non-null    float64
 8   Facility Name Std_pacfa             628 non-null    object
 9   DBA                                 175 non-null    object
 10  County_pacfa                        628 non-null    object
 11  City_pacfa                          628 non-null    object
 12  State                               628 non-null    object
 13  Business License App Category Name  628 non-null    object
 14  Expire Date                         628 non-null    object
 15  City_final                          976 non-null    object
 16  County_final                        1054 non-null   object
dtypes: float64(2), object(15)
memory usage: 173.8+ KB
```

In [199…

```
# Drop old columns (not required) to avoid duplication
facility_locations_final = facility_locations_final.drop(
    columns=["Facility Name Std_fac", "Facility Name Std_pacfa",
             "County_fac", "City_fac", "County_pacfa", "City_pacfa", "DBA",
             "Business License App Category Name", "Expire Date"], errors="ignore")
```

In [200…

```
facility_locations_final.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1308 entries, 0 to 1307
Data columns (total 8 columns):
 #   Column                 Non-Null Count  Dtype  
---  ------                 --------------  -----  
 0   Facility Name Original  1308 non-null   object 
 1   Facility Name Final     1308 non-null   object 
 2   Zip Code                580 non-null    object 
 3   Latitude                580 non-null    float64
 4   Longitude               580 non-null    float64
 5   State                   628 non-null    object 
 6   City_final              976 non-null    object 
 7   County_final            1054 non-null   object 
dtypes: float64(2), object(6)
memory usage: 81.9+ KB
```

In [201…

```
# Removing duplicate facility_locations_final by Facility Name Original
facility_locations_final = facility_locations_final.drop_duplicates(
    subset=["Facility Name Original"])
```

In [203…

```
# Ensured that only unique values are present.
facility_locations_final.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 1137 entries, 0 to 1307
Data columns (total 8 columns):
 #   Column                 Non-Null Count  Dtype  
---  ------                 --------------  -----  
 0   Facility Name Original  1137 non-null   object 
 1   Facility Name Final     1137 non-null   object 
 2   Zip Code                533 non-null    object 
 3   Latitude                533 non-null    float64
 4   Longitude               533 non-null    float64
 5   State                   457 non-null    object 
 6   City_final              805 non-null    object 
 7   County_final            883 non-null    object 
dtypes: float64(2), object(6)
memory usage: 79.9+ KB
```

Replaced all missing values (None or np.nan) in the applicable columns with pd.NA to ensure consistent handling of null values across the dataset.

In [204…
```
facility_locations_final["City_final"] = facility_locations_final["City_final"].fillna(pd.NA)
facility_locations_final["County_final"] = facility_locations_final["County_final"].fillna(pd.NA)
facility_locations_final["Zip Code"] = facility_locations_final["Zip Code"].fillna(pd.NA)
facility_locations_final["Latitude"] = facility_locations_final["Latitude"].fillna(pd.NA)
facility_locations_final["Longitude"] = facility_locations_final["Longitude"].fillna(pd.NA)
facility_locations_final["State"] = facility_locations_final["State"].fillna(pd.NA)
```

In [205…
```
facility_locations_final.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 1137 entries, 0 to 1307
Data columns (total 8 columns):
 #   Column                  Non-Null Count  Dtype  
---  ------                  --------------  -----  
 0   Facility Name Original  1137 non-null   object 
 1   Facility Name Final     1137 non-null   object 
 2   Zip Code                533 non-null    object 
 3   Latitude                533 non-null    float64
 4   Longitude               533 non-null    float64
 5   State                   457 non-null    object 
 6   City_final              805 non-null    object 
 7   County_final            883 non-null    object 
dtypes: float64(2), object(6)
memory usage: 79.9+ KB
```

In [207…
```
# Rename columns
facility_locations_final = facility_locations_final.rename(columns={
    "City_final": "City",
    "County_final": "County"
})

# Reorder columns
facility_locations_final = facility_locations_final[[
    "Facility Name Original",
    "Facility Name Final",
    "City",
```

```
    "County",
    "State",
    "Zip Code",
    "Latitude",
    "Longitude"
]]
```

In [208…
```
facility_locations_final.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 1137 entries, 0 to 1307
Data columns (total 8 columns):
 #   Column                  Non-Null Count  Dtype  
---  ------                  --------------  -----  
 0   Facility Name Original  1137 non-null   object 
 1   Facility Name Final     1137 non-null   object 
 2   City                    805 non-null    object 
 3   County                  883 non-null    object 
 4   State                   457 non-null    object 
 5   Zip Code                533 non-null    object 
 6   Latitude                533 non-null    float64
 7   Longitude               533 non-null    float64
dtypes: float64(2), object(6)
memory usage: 79.9+ KB
```

In [209…
```
facility_locations_final.head()
```

Out[209…

	Facility Name Original	Facility Name Final	City	County	State	Zip Code	Latitude	Longitude
0	2 Blondes All Breed Rescue	2 Blondes All Breed Rescue	Lakewood	Jefferson County	CO	<NA>	NaN	NaN
1	2 Blondes All Breed Rescue, Inc.	2 Blondes All Breed Rescue	Lakewood	Jefferson County	CO	<NA>	NaN	NaN
2	2nd Chance Vizsla Rescue Inc	2Nd Chance Vizsla Rescue	<NA>	Larimer County	<NA>	<NA>	NaN	NaN
3	2nd Chance Vizsla Rescue, Inc	2Nd Chance Vizsla Rescue	<NA>	Larimer County	<NA>	<NA>	NaN	NaN
4	2nd Chance Vizsla Rescue, Inc.	2Nd Chance Vizsla Rescue	<NA>	Larimer County	<NA>	<NA>	NaN	NaN

In [210…
```
# Saved for reference
```

```
facility_locations_final.to_excel("Unique_Facility_Locations_Final.xlsx", index=False)
```

In [252…

We now have the final datasets for all years and the unique facility location data. The next step is to combine both datasets to create a standardized final file, consolidating facility and location information along with shelter and rescue statistics.

In [211…

```
# Combined all year shelter and rescue statistics data
all_data_2016_2024.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 370112 entries, 0 to 370111
Data columns (total 12 columns):
 #   Column          Non-Null Count   Dtype  
---  ------          --------------   -----  
 0   Facility Name   370112 non-null  object 
 1   County          43457 non-null   object 
 2   City            22951 non-null   object 
 3   Zip Code        22951 non-null   object 
 4   Latitude        22951 non-null   float64
 5   Longitude       22951 non-null   float64
 6   Reporting Year  370112 non-null  int64  
 7   Species         370112 non-null  object 
 8   Age Group       370112 non-null  object 
 9   Flow Type       370112 non-null  object 
 10  Event Type      370112 non-null  object 
 11  Animal Count    370112 non-null  int64  
dtypes: float64(2), int64(2), object(8)
memory usage: 33.9+ MB
```

In [212…

```
# Unique facility and location data
facility_locations_final.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 1137 entries, 0 to 1307
Data columns (total 8 columns):
 #   Column                  Non-Null Count  Dtype  
---  ------                  --------------  -----  
 0   Facility Name Original  1137 non-null   object 
 1   Facility Name Final     1137 non-null   object 
 2   City                    805 non-null    object 
 3   County                  883 non-null    object 
 4   State                   457 non-null    object 
 5   Zip Code                533 non-null    object 
 6   Latitude                533 non-null    float64
 7   Longitude               533 non-null    float64
dtypes: float64(2), object(6)
memory usage: 79.9+ KB
```

Merged `all_data_2016_2024` with `facility_locations_final` using Facility Name Original as the key to consolidate facility, location, and historical data into a single dataset.

In [213…

```
# Merge all_data_2016_2024 with facility_locations_final on Facility Name Original
merged = all_data_2016_2024.merge(
    facility_locations_final,
    left_on="Facility Name",
    right_on="Facility Name Original",
    how="left",
    suffixes=("", "_fac")
)

# Columns to update
update_cols = ["City", "County", "Zip Code", "Latitude", "Longitude"]

# Overwrite original columns with facility_locations_final values where available
for col in update_cols:
    merged[col].update(merged[col + "_fac"])

# Final updated table
all_data_2016_2024_updated = merged
```

In [214…

```
all_data_2016_2024_updated.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 370112 entries, 0 to 370111
Data columns (total 20 columns):
 #   Column                  Non-Null Count   Dtype  
---  ------                  --------------   -----  
 0   Facility Name           370112 non-null  object 
 1   County                  311360 non-null  object 
 2   City                    295502 non-null  object 
 3   Zip Code                193334 non-null  object 
 4   Latitude                193334 non-null  float64
 5   Longitude               193334 non-null  float64
 6   Reporting Year          370112 non-null  int64  
 7   Species                 370112 non-null  object 
 8   Age Group               370112 non-null  object 
 9   Flow Type               370112 non-null  object 
 10  Event Type              370112 non-null  object 
 11  Animal Count            370112 non-null  int64  
 12  Facility Name Original  370112 non-null  object 
 13  Facility Name Final     370112 non-null  object 
 14  City_fac                295502 non-null  object 
 15  County_fac              311360 non-null  object 
 16  State                   204191 non-null  object 
 17  Zip Code_fac            191137 non-null  object 
 18  Latitude_fac            191137 non-null  float64
 19  Longitude_fac           191137 non-null  float64
dtypes: float64(4), int64(2), object(14)
memory usage: 56.5+ MB
```

In [215…

```
# Removing unnecessary columns
all_data_2016_2024_updated = all_data_2016_2024_updated.drop(
    columns=["Facility Name Original", "Facility Name Final", "City_fac",
             "County_fac", "Zip Code_fac", "Latitude_fac", "Longitude_fac"], errors="ignore")
```

The final consolidated dataset combines all-year data with standardized facility location information (where available), providing a comprehensive view of facility operations and associated metrics.

In [217…

```
all_data_2016_2024_updated.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 370112 entries, 0 to 370111
Data columns (total 13 columns):
 #   Column          Non-Null Count   Dtype
---  ------          --------------   -----
 0   Facility Name   370112 non-null  object
 1   County          311360 non-null  object
 2   City            295502 non-null  object
 3   Zip Code        193334 non-null  object
 4   Latitude        193334 non-null  float64
 5   Longitude       193334 non-null  float64
 6   Reporting Year  370112 non-null  int64
 7   Species         370112 non-null  object
 8   Age Group       370112 non-null  object
 9   Flow Type       370112 non-null  object
 10  Event Type      370112 non-null  object
 11  Animal Count    370112 non-null  int64
 12  State           204191 non-null  object
dtypes: float64(2), int64(2), object(9)
memory usage: 36.7+ MB
```

In [221…

```
# Reorder columns
all_data_2016_2024_updated = all_data_2016_2024_updated[[
    "Facility Name", "City", "County", "State", "Zip Code", "Latitude", "Longitude",
    "Reporting Year", "Species", "Age Group", "Flow Type", "Event Type", "Animal Count"
]]
```

Final cleaned output

In [222…

```
all_data_2016_2024_updated.head()
```

Out[222...

	Facility Name	City	County	State	Zip Code	Latitude	Longitude	Reporting Year	Species	Age Group	Flow Type	Event Type	Animal Count
0	German Shepherd Rescue of Central Colorado	Hartsel Colorado	Park County	<NA>	80449	35.500801	-117.947800	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	1
1	Doggy Dog World Rescue	Littleton	Jefferson County	CO	80125	39.612653	-105.016198	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	25
2	Surface Creek Shelter	Cedaredge	Delta County	CO	81413	38.900738	-107.923767	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	6
3	Delta County Citizens for Animal Welfare and S...	Delta	Delta County	<NA>	CO 81416	38.741684	-108.070175	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	7
4	Dalmatian Rescue of Colorado	Colorado Springs	El Paso County	CO	80526	40.588972	-105.082459	2016	Dogs	Adult	Start of Year Count	Beginning Shelter Count	3

In [223...

```
# Saved for reference
all_data_2016_2024_updated.to_excel("Cleaned_2016_to_2024_Shelter_And_Rescue_Statistics_final.xlsx", index=False)
```

In [ ]: