Dallas Animal Shelter - Data Analysis – Monika Kulkarni

Objective

The purpose of this project is to analyze the **Dallas Animal Shelter Data** to:

- Normalize the dataset and create an optimized Entity-Relationship Diagram (ERD).
- Design and implement database tables using PostgreSQL.
- Load, clean, and process the data into structured formats.
- Generate insights on:
    - **Breeds with the best survival rate** over multiple years.
    - **Impact of intake type on animal outcomes**.

Data Sources

The Dallas Animal Shelter data has been extracted from dallasopendata.com. I have merged the following two datasets.

https://www.dallasopendata.com/Services/Dallas-Animal-Shelter-Data-Fiscal-Year-2022-2023/f77p-sgrc/about_data

https://www.dallasopendata.com/Services/Dallas-Animal-Shelter-Data-Fiscal-Year-2023-2025/uyte-zi7f/about_data

Approach

1. Data Preprocessing

- The dataset was sourced from the **Dallas Open Data Portal**.
- Only **cats and dogs** were considered for analysis.

2. Database Design & Normalization

A fully normalized Entity-Relationship Diagram (ERD) was created to remove redundancies and improve data integrity.

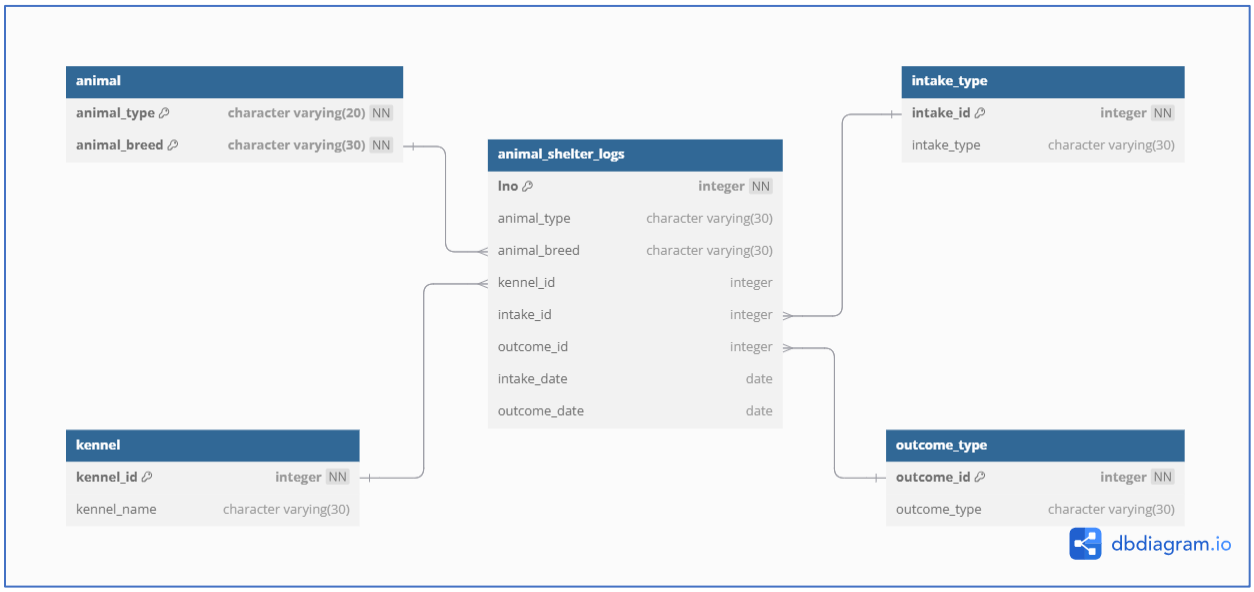
Animal - Stores unique Animal_Type and Animal_Breed.

Kennel - Stores Kennel_ID and Kennel_Name.

Intake_Type - Stores Intake_ID and Intake_Type.

Outcome_Type - Stores Outcome_ID and Outcome_Type.

Animal_Shelter_Logs - Core table linking all relationships with foreign keys.



3. SQL Implementation

- **Table Creation:**
  - Each normalized table was created using CREATE TABLE SQL statements.
- **Data Insertion:**
  - Used INSERT INTO ... SELECT DISTINCT to populate tables while avoiding duplicates.
  - Implemented sequences for primary key generation (e.g., seq_kennel, seq_intaketype).
- **Constraints & Foreign Keys:**
  - Ensured referential integrity between Animal, Kennel, IntakeType, OutcomeType, and AnimalShelterLogs.

Analysis & Reports

Report 1: Breeds Survival Rate Analysis

Approach:

Identified breeds with the best survival rates, considering only outcomes where the animal survived.

Considered only breeds with 200+ admissions per year.

Surviving Outcomes:

- FOSTER

- RETURNED TO OWNER
- TREATMENT
- SNR (Shelter-Neuter-Return)
- ADOPTION
- WILDLIFE
- TRANSFER
- TNR (Trap-Neuter-Return)

Survival Rate Report for CATS:

```sql
-- Survival Rate Report for CATS
WITH SurvivalData AS (
    SELECT
        l.animal_type,
        l.animal_breed,
        EXTRACT(YEAR FROM l.intake_date) AS intake_year,
        COUNT(*) AS total_intakes,
        SUM(CASE WHEN o.outcome_type IN ('FOSTER', 'RETURNED TO OWNER', 'TREATMENT', 'SNR', 'ADOPTION', 'WILDLIFE', 'TRANSFER', 'TNR') THEN 1 ELSE 0 END) AS survived_count
    FROM animal_shelter_logs l
    JOIN outcome_type o ON l.outcome_id = o.outcome_id
    GROUP BY l.animal_type, l.animal_breed, intake_year
)
SELECT
    animal_type,
    animal_breed,
    intake_year,
    total_intakes,
    survived_count,
    ROUND((survived_count * 100.0 / total_intakes), 2) AS survival_rate
FROM SurvivalData
WHERE total_intakes > 200 and animal_type = 'CAT'
ORDER BY animal_type, intake_year, survival_rate DESC;
```

	animal_type character varying (30)	animal_breed character varying (30)	intake_year numeric	total_intakes bigint	survived_count bigint	survival_rate numeric
1	CAT	DOMESTIC SH	2022	1217	978	80.36
2	CAT	DOMESTIC SH	2023	6917	5460	78.94
3	CAT	DOMESTIC MH	2023	457	318	69.58
4	CAT	AMER SH	2023	235	20	8.51
5	CAT	DOMESTIC SH	2024	10782	8646	80.19
6	CAT	DOMESTIC MH	2024	876	663	75.68
7	CAT	DOMESTIC LH	2024	305	205	67.21
8	CAT	AMER SH	2024	279	85	30.47
9	CAT	DOMESTIC SH	2025	557	505	90.66

Insight:

- Cat breeds, such as domestic shorthairs, show the highest survival rate, likely due to their popularity and ease of adoption.
- Breeds with lower survival rates may require more rescue efforts or specialized care programs.
- Survival rates tend to be higher in recent years, possibly indicating better adoption programs or increased awareness.

Survival Rate Report for DOGS:

Dallas Animal Shelter - Data Analysis – Monika Kulkarni

```sql
-- Survival Rate Report for DOGS
WITH SurvivalData AS (
    SELECT
        l.animal_type,
        l.animal_breed,
        EXTRACT(YEAR FROM l.intake_date) AS intake_year,
        COUNT(*) AS total_intakes,
        SUM(CASE WHEN o.outcome_type IN ('FOSTER', 'RETURNED TO OWNER', 'TREATMENT', 'SNR', 'ADOPTION', 'WILDLIFE', 'TRANSFER', 'TNR') THEN 1 ELSE 0 END) AS survived_count
    FROM animal_shelter_logs l
    JOIN outcome_type o ON l.outcome_id = o.outcome_id
    GROUP BY l.animal_type, l.animal_breed, intake_year
)
SELECT
    animal_type,
    animal_breed,
    intake_year,
    total_intakes,
    survived_count,
    ROUND((survived_count * 100.0 / total_intakes), 2) AS survival_rate
FROM SurvivalData
WHERE total_intakes > 200 and animal_type = 'DOG'
ORDER BY animal_type, intake_year, survival_rate DESC;
```

	animal_type character varying (30)	animal_breed character varying (30)	intake_year numeric	total_intakes bigint	survived_count bigint	survival_rate numeric
1	DOG	MIXED BREED	2022	3241	2548	78.62
2	DOG	PIT BULL	2022	452	237	52.43
3	DOG	GERM SHEPHERD	2022	317	152	47.95
4	DOG	MIXED BREED	2023	12937	10974	84.83
5	DOG	GERM SHEPHERD	2023	1447	916	63.3
6	DOG	LABRADOR RETR	2023	926	550	59.4
7	DOG	PIT BULL	2023	2403	1418	59.01
8	DOG	GREAT PYRENEES	2023	223	129	57.85
9	DOG	ROTTWEILER	2023	226	129	57.08
10	DOG	ALASKAN HUSKY	2023	391	218	55.75
11	DOG	SIBERIAN HUSKY	2023	255	132	51.76
12	DOG	CHIHUAHUA SH	2023	773	386	49.94
13	DOG	AM PIT BULL TER	2023	332	62	18.67
14	DOG	MIXED BREED	2024	19704	17452	88.57
15	DOG	ALASKAN HUSKY	2024	345	251	72.75
16	DOG	GERM SHEPHERD	2024	1132	823	72.7
17	DOG	PIT BULL	2024	1597	1101	68.94
18	DOG	LABRADOR RETR	2024	552	357	64.67
19	DOG	CHIHUAHUA SH	2024	561	310	55.26
20	DOG	AM PIT BULL TER	2024	234	68	29.06
21	DOG	MIXED BREED	2025	1526	1404	92.01

Insights:

- mixed breeds generally show higher survival rates.
- Breeds classified as **AM PIT BULL TER (American Pit Bull Terrier)** tend to have **lower survival rates**, possibly due to adoption hesitancy. Many cities have breed-specific legislation (BSL) restricting Pit Bull adoptions.

Report 2: Breeds Non-Survival Report

Identified breeds with the highest non-survival rates, considering only outcomes where the animal did not survive.

Non-Surviving Outcomes:

- EUTHANIZED

- DIED
- FOUND EXP (Found Expired)
- DISPOSAL

Non-Survival Rates for CATS:

```
-- Non-Survival Rates for CATS
WITH NonSurvivalData AS (
    SELECT
        l.animal_type,
        l.animal_breed,
        EXTRACT(YEAR FROM l.intake_date) AS intake_year,
        COUNT(*) AS total_intakes,
        SUM(CASE WHEN o.outcome_type IN ('EUTHANIZED', 'DIED', 'FOUND EXP', 'DISPOSAL') THEN 1 ELSE 0 END) AS non_survived_count
    FROM animal_shelter_logs l
    JOIN outcome_type o ON l.outcome_id = o.outcome_id
    GROUP BY l.animal_type, l.animal_breed, intake_year
)
SELECT
    animal_type,
    animal_breed,
    intake_year,
    total_intakes,
    non_survived_count,
    ROUND((non_survived_count * 100.0 / total_intakes), 2) AS non_survival_rate
FROM NonSurvivalData
WHERE total_intakes > 200 and animal_type = 'CAT'
ORDER BY animal_type, intake_year, non_survival_rate DESC;
```

	animal_type character varying (30)	animal_breed character varying (30)	intake_year numeric	total_intakes bigint	non_survived_count bigint	non_survival_rate numeric
1	CAT	DOMESTIC SH	2022	1217	146	12
2	CAT	AMER SH	2023	235	47	20
3	CAT	DOMESTIC SH	2023	6917	1109	16.03
4	CAT	DOMESTIC MH	2023	457	65	14.22
5	CAT	AMER SH	2024	279	50	17.92
6	CAT	DOMESTIC LH	2024	305	52	17.05
7	CAT	DOMESTIC SH	2024	10782	1612	14.95
8	CAT	DOMESTIC MH	2024	876	129	14.73
9	CAT	DOMESTIC SH	2025	557	51	9.16

Insights:

- Non-survival rates have decreased over time for DOMESTIC SH, AMERICAN SH possibly due to better medical intervention and higher adoption rates.

Non-Survival Rates for DOGS:

```sql
-- Non-Survival Rates for DOGS
WITH NonSurvivalData AS (
    SELECT
        l.animal_type,
        l.animal_breed,
        EXTRACT(YEAR FROM l.intake_date) AS intake_year,
        COUNT(*) AS total_intakes,
        SUM(CASE WHEN o.outcome_type IN ('EUTHANIZED', 'DIED', 'FOUND EXP', 'DISPOSAL') THEN 1 ELSE 0 END) AS non_survived_count
    FROM animal_shelter_logs l
    JOIN outcome_type o ON l.outcome_id = o.outcome_id
    GROUP BY l.animal_type, l.animal_breed, intake_year
)
SELECT
    animal_type,
    animal_breed,
    intake_year,
    total_intakes,
    non_survived_count,
    ROUND((non_survived_count * 100.0 / total_intakes), 2) AS non_survival_rate
FROM NonSurvivalData
WHERE total_intakes > 200 and animal_type = 'DOG'
ORDER BY animal_type, intake_year, non_survival_rate DESC;
```

	animal_type character varying (30)	animal_breed character varying (30)	intake_year numeric	total_intakes bigint	non_survived_count bigint	non_survival_rate numeric
1	DOG	PIT BULL	2022	452	151	33.41
2	DOG	GERM SHEPHERD	2022	317	86	27.13
3	DOG	MIXED BREED	2022	3241	692	21.35
4	DOG	AM PIT BULL TER	2023	332	123	37.05
5	DOG	PIT BULL	2023	2403	811	33.75
6	DOG	ROTTWEILER	2023	226	63	27.88
7	DOG	ALASKAN HUSKY	2023	391	105	26.85
8	DOG	LABRADOR RETR	2023	926	244	26.35
9	DOG	GERM SHEPHERD	2023	1447	359	24.81
10	DOG	SIBERIAN HUSKY	2023	255	62	24.31
11	DOG	GREAT PYRENEES	2023	223	50	22.42
12	DOG	CHIHUAHUA SH	2023	773	130	16.82
13	DOG	MIXED BREED	2023	12937	1938	14.98
14	DOG	AM PIT BULL TER	2024	234	65	27.78
15	DOG	PIT BULL	2024	1597	369	23.11
16	DOG	LABRADOR RETR	2024	552	110	19.93
17	DOG	ALASKAN HUSKY	2024	345	61	17.68
18	DOG	GERM SHEPHERD	2024	1132	186	16.43
19	DOG	CHIHUAHUA SH	2024	561	87	15.51
20	DOG	MIXED BREED	2024	19704	1844	9.36
21	DOG	MIXED BREED	2025	1526	120	7.86

Insights:

- Breeds classified as **AM PIT BULL TER (American Pit Bull Terrier)** tend to have **lower survival rates**, possibly due to adoption hesitancy. Many cities have breed-specific legislation (BSL) restricting Pit Bull adoptions.

Report 2: Intake Type vs Outcome Analysis

Examined how intake type influences an animal's outcome.

Dallas Animal Shelter - Data Analysis – Monika Kulkarni

Calculated the percentage of each outcome (Adoption, Euthanized, Transferred, Returned to Owner) for each intake type (Stray, Owner Surrender, Confiscated).

```
-- examine how intake type influences an animal's outcome
WITH IntakeOutcomeCounts AS (
    SELECT
        i.intake_type,
        o.outcome_type,
        COUNT(*) AS count_records
    FROM animal_shelter_logs l
    JOIN intake_type i ON l.intake_id = i.intake_id
    JOIN outcome_type o ON l.outcome_id = o.outcome_id
    GROUP BY i.intake_type, o.outcome_type
)
SELECT
    intake_type,
    outcome_type,
    count_records,
    ROUND((count_records * 100.0 / SUM(count_records) OVER (PARTITION BY intake_type)), 2) AS percentage_outcomes
FROM IntakeOutcomeCounts
ORDER BY intake_type, percentage_outcomes DESC;
```

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
1	CONFISCATED	RETURNED TO OWNER	1113	54.06
2	CONFISCATED	EUTHANIZED	377	18.31
3	CONFISCATED	ADOPTION	190	9.23
4	CONFISCATED	TRANSFER	143	6.95
5	CONFISCATED	DISPOSAL	115	5.59
6	CONFISCATED	FOSTER	112	5.44
7	CONFISCATED	DIED	6	0.29
8	CONFISCATED	SNR	3	0.15

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
9	DISPOS REQ	DISPOSAL	146	100

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
10	FOSTER	ADOPTION	7844	61.65
11	FOSTER	FOSTER	2039	16.03
12	FOSTER	TRANSFER	1875	14.74
13	FOSTER	DIED	463	3.64
14	FOSTER	EUTHANIZED	306	2.41
15	FOSTER	MISSING	127	1
16	FOSTER	RETURNED TO OWNER	62	0.49
17	FOSTER	TREATMENT	3	0.02
18	FOSTER	SNR	3	0.02
19	FOSTER	LOST EXP	1	0.01

Dallas Animal Shelter - Data Analysis – Monika Kulkarni

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
20	KEEPSAFE	RETURNED TO OWNER	551	46.69
21	KEEPSAFE	ADOPTION	291	24.66
22	KEEPSAFE	TRANSFER	126	10.68
23	KEEPSAFE	FOSTER	110	9.32
24	KEEPSAFE	EUTHANIZED	76	6.44
25	KEEPSAFE	DIED	13	1.1
26	KEEPSAFE	DISPOSAL	12	1.02
27	KEEPSAFE	MISSING	1	0.08

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
28	OWNER SURRENDER	ADOPTION	4694	38.89
29	OWNER SURRENDER	EUTHANIZED	2354	19.5
30	OWNER SURRENDER	FOSTER	2233	18.5
31	OWNER SURRENDER	TRANSFER	1833	15.19
32	OWNER SURRENDER	RETURNED TO OWNER	714	5.92
33	OWNER SURRENDER	DISPOSAL	125	1.04
34	OWNER SURRENDER	DIED	104	0.86
35	OWNER SURRENDER	MISSING	11	0.09
36	OWNER SURRENDER	SNR	1	0.01

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
37	RESOURCE	CLOSED	796	100

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
38	STRAY	ADOPTION	10815	23.16
39	STRAY	FOSTER	8853	18.96
40	STRAY	TRANSFER	6466	13.85
41	STRAY	LOST EXP	6276	13.44
42	STRAY	EUTHANIZED	5534	11.85
43	STRAY	RETURNED TO OWNER	5141	11.01
44	STRAY	FOUND EXP	2587	5.54
45	STRAY	DIED	591	1.27
46	STRAY	SNR	287	0.61
47	STRAY	DISPOSAL	120	0.26
48	STRAY	MISSING	23	0.05
49	STRAY	TREATMENT	2	0

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
50	TNR	TNR	1304	81.65
51	TNR	ADOPTION	124	7.76
52	TNR	FOSTER	122	7.64
53	TNR	TRANSFER	18	1.13
54	TNR	EUTHANIZED	15	0.94
55	TNR	DIED	6	0.38
56	TNR	SNR	4	0.25
57	TNR	RETURNED TO OWNER	3	0.19
58	TNR	MISSING	1	0.06

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
59	TRANSFER	ADOPTION	192	47.41
60	TRANSFER	FOSTER	78	19.26
61	TRANSFER	TRANSFER	50	12.35
62	TRANSFER	EUTHANIZED	43	10.62
63	TRANSFER	RETURNED TO OWNER	41	10.12
64	TRANSFER	DIED	1	0.25

	intake_type character varying (30)	outcome_type character varying (30)	count_records bigint	percentage_outcomes numeric
65	TREATMENT	TREATMENT	3335	100

Insights:

- Stray animals make up the largest percentage of shelter intakes but have a mixed outcome distribution (some returned to owners, some adopted, some euthanized).
- Owner-surrendered pets have a higher adoption rate but also a noticeable euthanasia percentage,