

Summary of data quality plan

Feature	Data Quality issue	Solution Strategy
Animal ID	Outliers [Primary key]	Do nothing
Name_Intake	Null values	Replace with Unique value "Unknown"
Name_Intake	Special values '1','135grams' etc	Replace with Unique value "Unknown"
Name_Intake	Name prefixed with '**'	Remove prefix '**'
Found Location	Unformatted locaion field	Prepend with unique value "Unknown"
Intake Type	Outliers	Do nothing
Intake Condition	Outliers	Do Nothing
Animal Type_Intake	Outliers	Do nothing
Breed_Intake	Too many breed categories	abstract breed categories
Color_Intake	Too many color categories	abstract color categories
Age upon Intake	Imprecision	Calculate accurate value in weeks using feature DateTime_Intake and Date of Birth
Age upon Outcome	IMprecision	Calculate accurate value in weeks using feature DateTime_Intake and Date of Birth
DateTime_Intake	Data entry errors	Intercahnge with feature DateTime_Outcome
DateTime_Outcome	Data entry errors	Intercahnge with feature DateTime_Intake
Date of Birth	Outliers	Do nothing
Sex upon Intake	Special values 'Unknown'	Replace with Unique value "Unknown Unknown"
Sex upon Outcome	Special values 'Unknown'	Replace with Unique value "Unknown Unknown"
binary_outcome	No issue	Do nothing