Contents

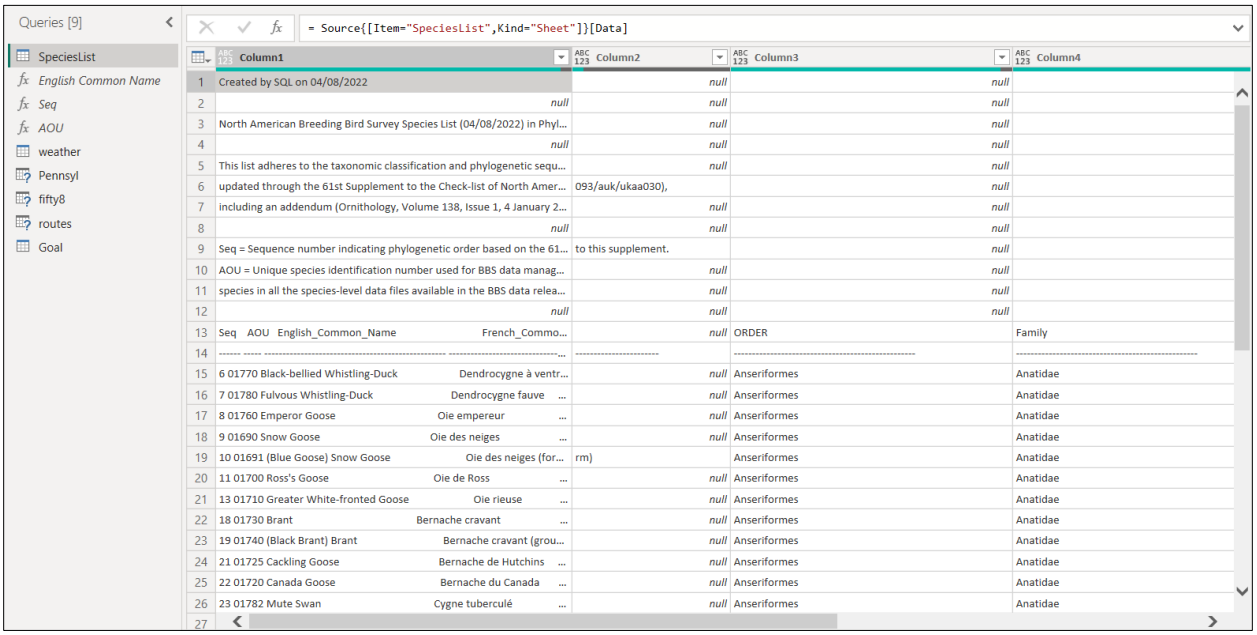
Intro

In this document, we will review the steps taken to clean and present the bird population data in a Power BI dashboard.

We will begin with importing the datasets, which consist of csv and xlsx formats.

Clean and organize data in Power Query

First is SpeciesList.xlsx, which has multiple inconsistencies to correct before analyzing:



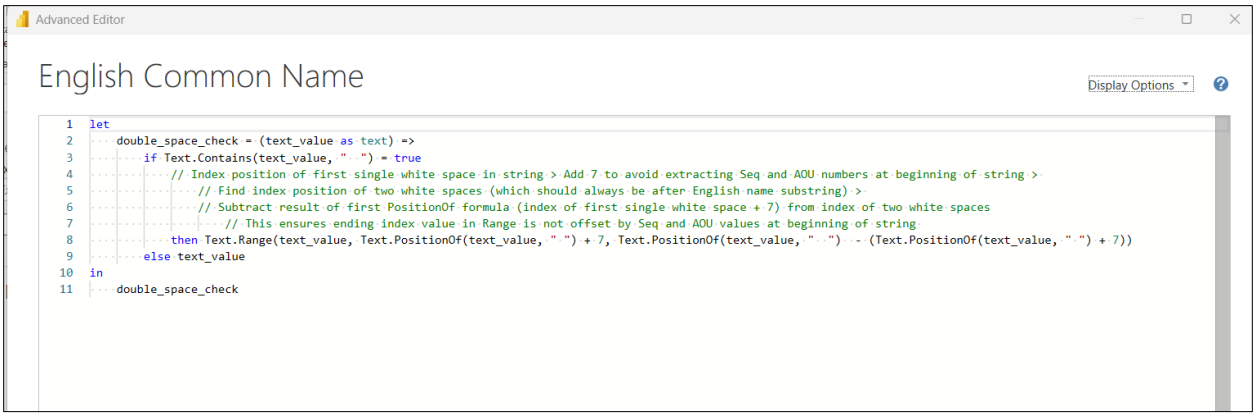
	Column1	Column2	Column3	Column4
1	Created by SQL on 04/08/2022	null	null	
2	null	null	null	
3	North American Breeding Bird Survey Species List (04/08/2022) in Phyl...	null	null	
4	null	null	null	
5	This list adheres to the taxonomic classification and phylogenetic sequ...	null	null	
6	updated through the 61st Supplement to the Check-list of North Amer...	093/auk/ukaa030),	null	
7	including an addendum (Ornithology, Volume 138, Issue 1, 4 January 2...	null	null	
8	null	null	null	
9	Seq = Sequence number indicating phylogenetic order based on the 61...	to this supplement.	null	
10	AOU = Unique species identification number used for BBS data manag...	null	null	
11	species in all the species-level data files available in the BBS data relea...	null	null	
12	null	null	null	
13	Seq AOU English_Common_Name French_Commo...	null	ORDER	Family
14	----- ----- ---------- ----------	----------	----------	----------
15	6 01770 Black-bellied Whistling-Duck Dendrocygne à ventr...	null	Anseriformes	Anatidae
16	7 01780 Fulvous Whistling-Duck Dendrocygne fauve ...	null	Anseriformes	Anatidae
17	8 01760 Emperor Goose Oie empereur ...	null	Anseriformes	Anatidae
18	9 01690 Snow Goose Oie des neiges ...	null	Anseriformes	Anatidae
19	10 01691 (Blue Goose) Snow Goose Oie des neiges (for...	rm)	Anseriformes	Anatidae
20	11 01700 Ross's Goose Oie de Ross ...	null	Anseriformes	Anatidae
21	13 01710 Greater White-fronted Goose Oie rieuse ...	null	Anseriformes	Anatidae
22	18 01730 Brant Bernache cravant ...	null	Anseriformes	Anatidae
23	19 01740 (Black Brant) Brant Bernache cravant (grou...	null	Anseriformes	Anatidae
24	21 01725 Cackling Goose Bernache de Hutchins ...	null	Anseriformes	Anatidae
25	22 01720 Canada Goose Bernache du Canada ...	null	Anseriformes	Anatidae
26	23 01782 Mute Swan Cygne tuberculé ...	null	Anseriformes	Anatidae
27				

The first 14 rows are irrelevant descriptions and a column has no data, so we will remove them.

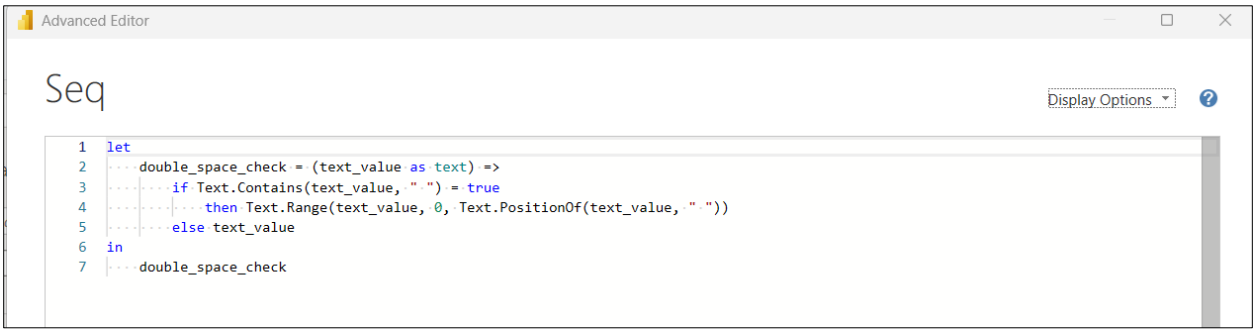
The next issue is how the first column contains independent data values (unique identification numbers and the names in English, Spanish and French). Additionally, they are separated inconsistently with different lengths of whitespace between each value.

This will be complicated to split to columns, and the standard Power Query function is unable to do it consistently, so we will write custom functions with the M language to dynamically split the column.

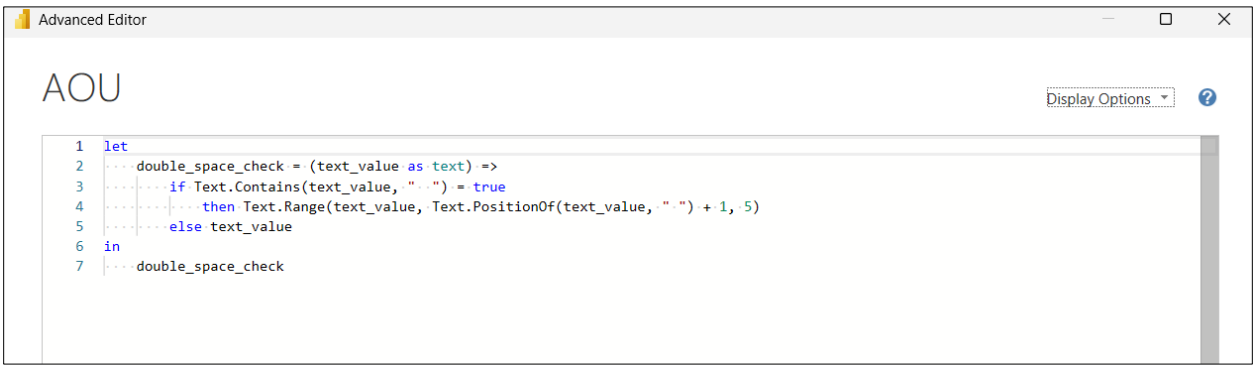
Split names:



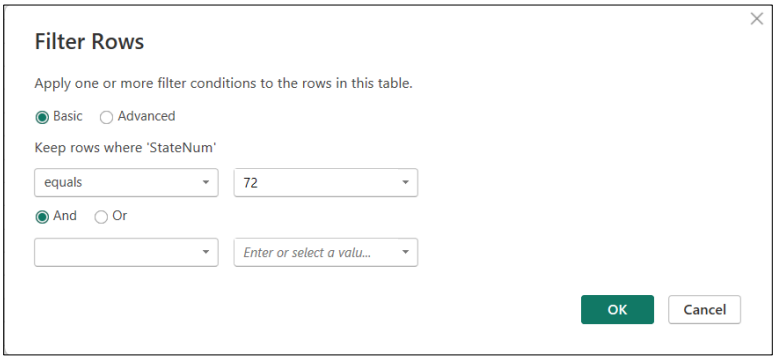
Split Seq Number:



AOU:

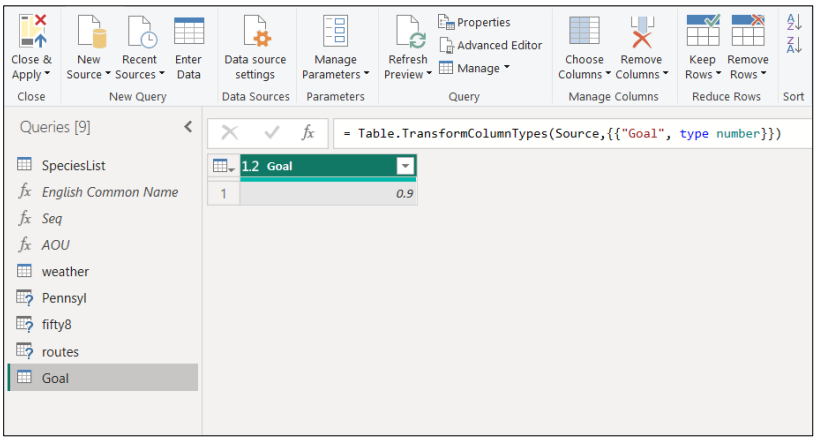


Next is weather.csv, where we will filter to only show data from Pennsylvania (i.e. StateNum = 72):



The remaining datasets (Pennsyl.csv, fifty8.csv, and routes.csv) are imported with minimal formatting changes.

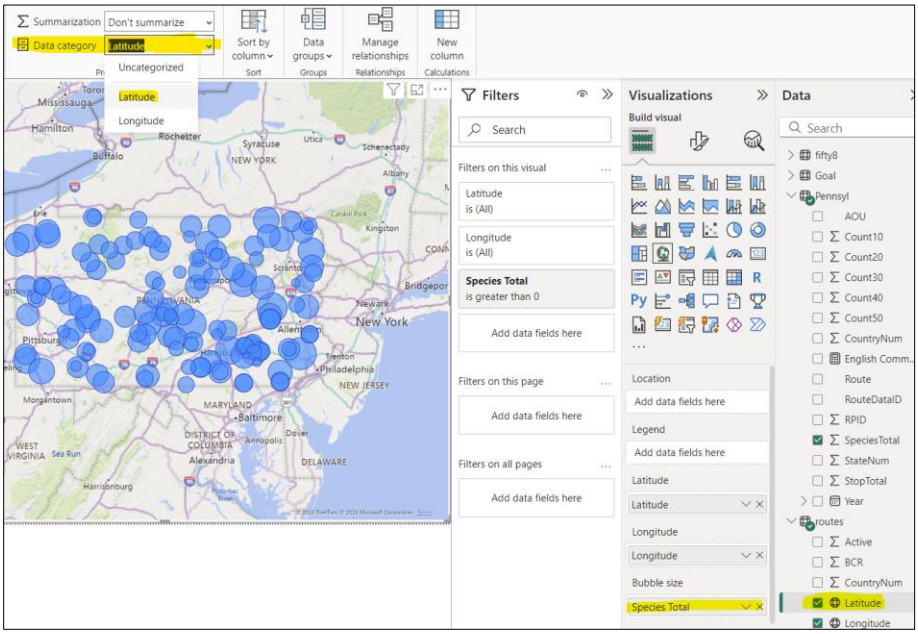
Last, we create a custom table within Power Query (Goal) to be used later for our KPI goal threshold:



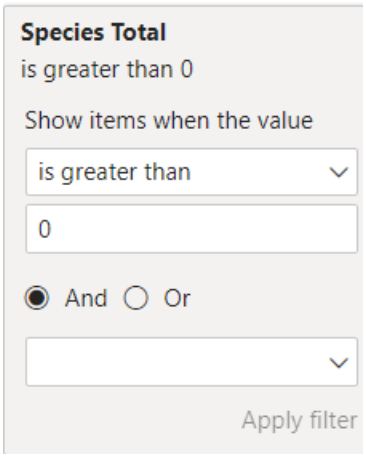
Now we are ready to import the sources to Power BI!

Species Population Map

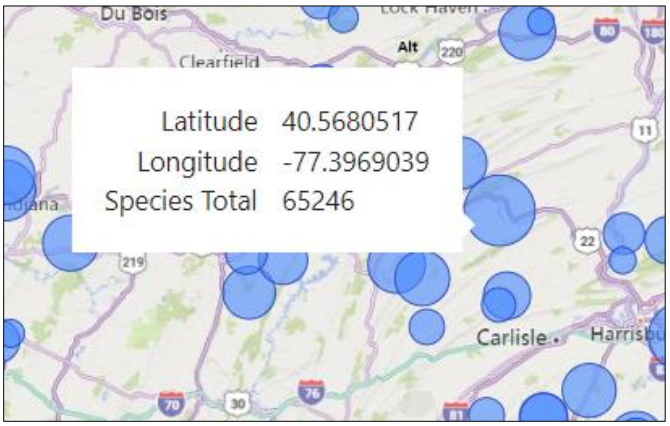
Our first page will show the number of species observed by location in a map format with bubble size representing the count of species observed. We first change the Data Category of the Latitude and Longitude fields to their respective data types to be used by the map visual. We will also input the Species Total field as the values for bubble size.



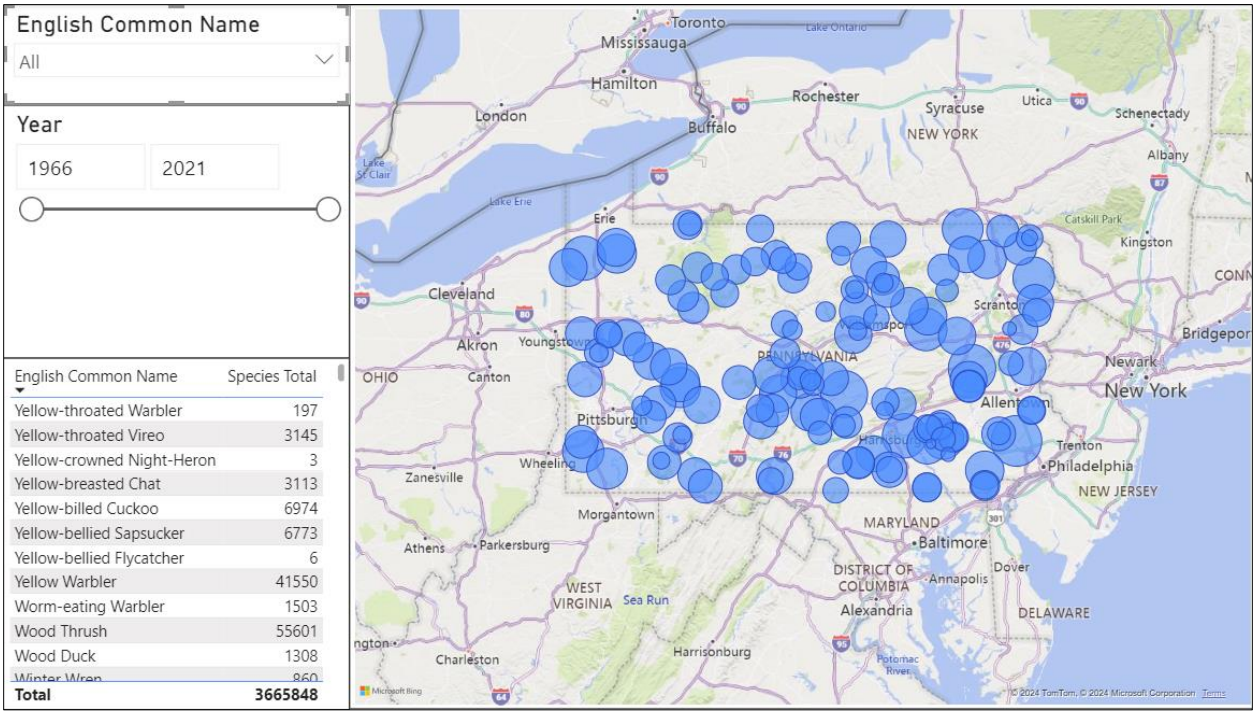
Finally, we filter out any species with a count of zero to only show relevant results in the Filters side pane:



Now we have our map visual, which can show the population count by hovering over a bubble!

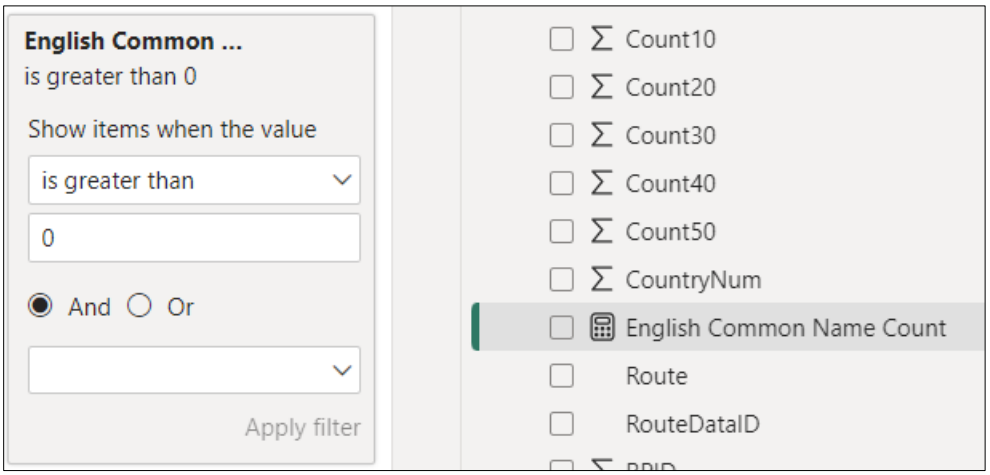


Now we will add some filters for year observed, English Common Name (species) and a table listing the Species Total, which in turn can have its rows clicked on for species filtering as well.



For the English Common Name filter, it lists all species from the SpeciesList table, even if they are not observed in any location. To fix this, we will create a measure with DAX and add it to the Filters side pane for the English Common Name dropdown menu. This will be used for all pages that use the English Common Name filter.

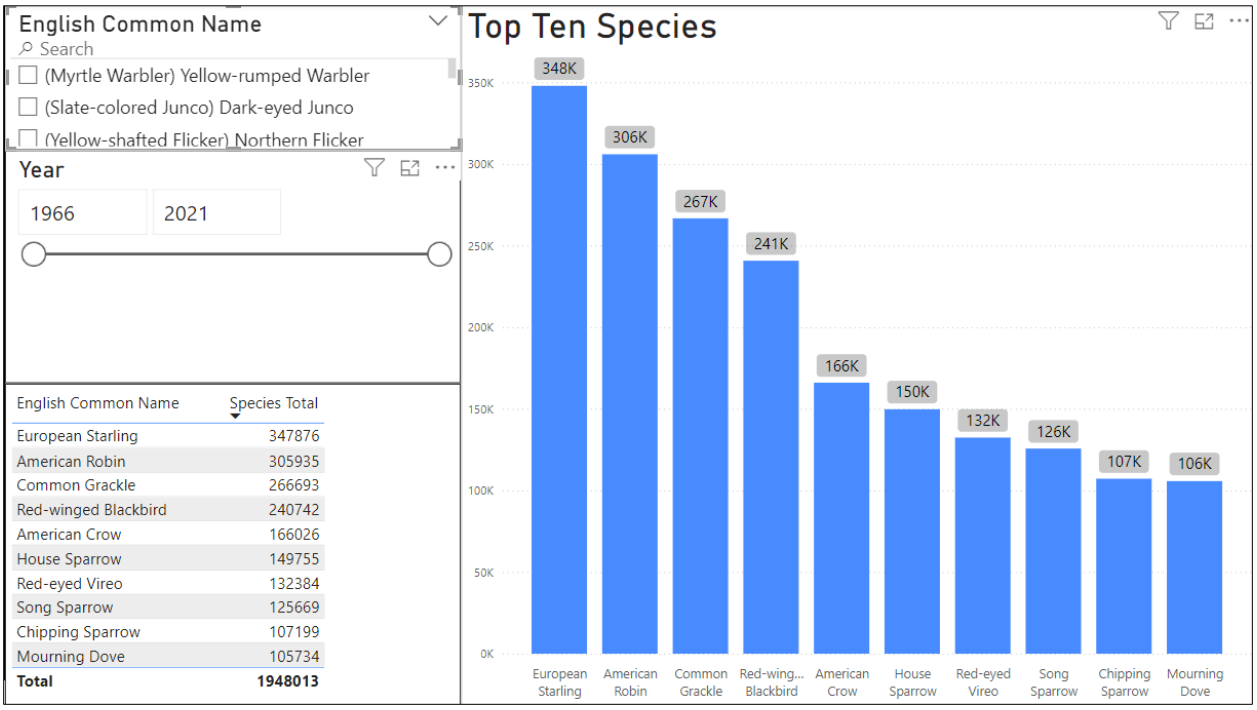
```
1 English Common Name Count = COUNTA(Pennsyl[SpeciesTotal])
```



Top Ten Species Population

The second page will illustrate the population size of species observed, with filters for year, species name.

We will add a bar graph and a table to show the count per species. They will be connected to English Common Name and Year filters on this page as well to see specific population counts by species name and year.



English Common Name	Species Total
European Starling	347876
American Robin	305935
Common Grackle	266693
Red-winged Blackbird	240742
American Crow	166026
House Sparrow	149755
Red-eyed Vireo	132384
Song Sparrow	125669
Chipping Sparrow	107199
Mourning Dove	105734
Total	**1948013**

Population Over Time

The third page will visualize the population over time in years with a line graph. Additionally, we will create groups for common bird categories like owls, hummingbirds, etc.

To create the groups, we will create a calculated column called “Species Group” in data view with a DAX formula. The logic will look for certain words in each English Common Name value for key words to properly sort them into different categories (i.e. if the name contains “owl”, label it with the “Owl” category , so “Barn Owl” would be categorized as “Owl”)

```
Species Group =
IF(CONTAINSSTRING(SpeciesList[English Common Name], "Owl"), "Owl",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Hawk"), "Hawk",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Falcon"), "Falcon",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Hummingbird"), "Hummingbird",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Duck"), "Duck",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Eagle"), "Eagle",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Woodpecker"), "Woodpecker",
    "Other"
)))))))
```

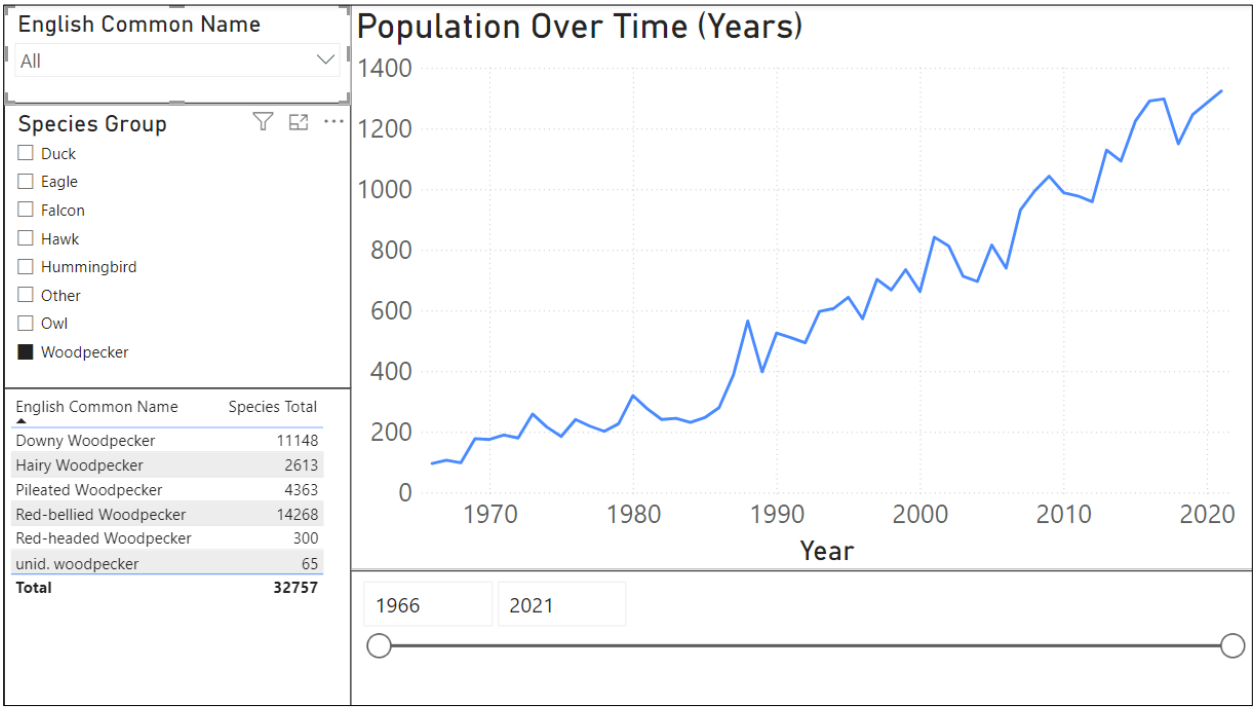
ORDER	Family	Genus	Species	English Common Name	Seq	AOU	Species Group
Strigiformes	Strigidae	Aegolius	funereus	Boreal Owl	695	3710	Owl
Strigiformes	Strigidae	Asio	flammeus	Short-eared Owl	693	3670	Owl
Strigiformes	Strigidae	Asio	otus	Long-eared Owl	691	3660	Owl
Strigiformes	Strigidae	Strix	nebulosa	Great Gray Owl	690	3700	Owl
Strigiformes	Strigidae	Strix	varia	Barred Owl	688	3680	Owl
Strigiformes	Strigidae	Strix	occidentalis	Spotted Owl	687	3690	Owl
Strigiformes	Strigidae	Athene	cunicularia	Burrowing Owl	684	3780	Owl
Strigiformes	Strigidae	Micrathene	whitneyi	Elf Owl	683	3810	Owl
Strigiformes	Strigidae	Glaucidium	brasilianum	Ferruginous Pygmy-Owl	682	3800	Owl
Strigiformes	Strigidae	Glaucidium	gnoma	Northern Pygmy-Owl	678	3790	Owl
Strigiformes	Strigidae	Surnia	ulula	Northern Hawk Owl	677	3770	Owl
Strigiformes	Strigidae	Bubo	scandiacus	Snowy Owl	676	3760	Owl
Strigiformes	Strigidae	Bubo	virginianus	Great Horned Owl	675	3750	Owl
Strigiformes	Strigidae	Megascops	asio	Eastern Screech-Owl	670	3730	Owl
Strigiformes	Strigidae	Megascops	kennicottii	Western Screech-Owl	669	3732	Owl
Strigiformes	Strigidae	Megascops	trichopsis	Whiskered Screech-Owl	666	3731	Owl
Strigiformes	Strigidae	Psiloscops	flammeolus	Flammulated Owl	665	3740	Owl
Strigiformes	Tytonidae	Tyto	alba	Barn Owl	663	3650	Owl
Galliformes	Phasianidae	Pavo	cristatus	Indian Peafowl	118	3095	Owl
Charadriiform	Scolopacid	Tringa	melanoleuca	Greater Yellowlegs	389	2540	Owl
Charadriiform	Scolopacid	Tringa	melanoleuca / flavipes	unid. Greater Yellowlegs / Lesser Yellowlegs	385	2555	Owl
Charadriiform	Scolopacid	Tringa	flavipes	Lesser Yellowlegs	384	2550	Owl
Passeriformes	Icteridae	Sturnella	magna / neglecta	unid. Eastern Meadowlark / Western Meadowlark	1349	5012	Owl
Passeriformes	Icteridae	Sturnella	neglecta	Western Meadowlark	1348	5011	Owl
Passeriformes	Icteridae	Sturnella	magna	Eastern Meadowlark	1347	5010	Owl

However, we have run into a bug in the DAX logic. The term “Owl” can be part of words not related to the species such as “Mead**owl**ark”, “F**owl**” and “Yell**owl**egs”. To fix this, we will revise the formula to include conditions excluding them.

```
Species Group =
IF(CONTAINSSTRING(SpeciesList[English Common Name], "Owl") && NOT(CONTAINSSTRING(SpeciesList[English Common Name], "Meadow")) && NOT(CONTAINSSTRING(SpeciesList
[English Common Name], "yellow")) && NOT(CONTAINSSTRING(SpeciesList[English Common Name], "fowl")), "Owl",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Hawk"), "Hawk",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Falcon"), "Falcon",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Hummingbird"), "Hummingbird",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Duck"), "Duck",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Eagle"), "Eagle",
    IF(CONTAINSSTRING(SpeciesList[English Common Name], "Woodpecker"), "Woodpecker",
    "Other"
)))))))
```

ORDER	Family	Genus	Species	English Common Name	Seq	AOU	Species Group
Strigiformes	Strigidae	Aegolius	acadicus	Northern Saw-whet Owl	696	3720	Owl
Strigiformes	Strigidae	Aegolius	funereus	Boreal Owl	695	3710	Owl
Strigiformes	Strigidae	Asio	flammeus	Short-eared Owl	693	3670	Owl
Strigiformes	Strigidae	Asio	otus	Long-eared Owl	691	3660	Owl
Strigiformes	Strigidae	Strix	nebulosa	Great Gray Owl	690	3700	Owl
Strigiformes	Strigidae	Strix	varia	Barred Owl	688	3680	Owl
Strigiformes	Strigidae	Strix	occidentalis	Spotted Owl	687	3690	Owl
Strigiformes	Strigidae	Athene	cunicularia	Burrowing Owl	684	3780	Owl
Strigiformes	Strigidae	Micrathene	whitneyi	Elf Owl	683	3810	Owl
Strigiformes	Strigidae	Glaucidium	brasilianum	Ferruginous Pygmy-Owl	682	3800	Owl
Strigiformes	Strigidae	Glaucidium	gnoma	Northern Pygmy-Owl	678	3790	Owl
Strigiformes	Strigidae	Surnia	ulula	Northern Hawk Owl	677	3770	Owl
Strigiformes	Strigidae	Bubo	scandiacus	Snowy Owl	676	3760	Owl
Strigiformes	Strigidae	Bubo	virginianus	Great Horned Owl	675	3750	Owl
Strigiformes	Strigidae	Megascops	asio	Eastern Screech-Owl	670	3730	Owl
Strigiformes	Strigidae	Megascops	kennicottii	Western Screech-Owl	669	3732	Owl
Strigiformes	Strigidae	Megascops	trichopsis	Whiskered Screech-Owl	666	3731	Owl
Strigiformes	Strigidae	Psiloscops	flammeolus	Flammulated Owl	665	3740	Owl
Strigiformes	Tytonidae	Tyto	alba	Barn Owl	663	3650	Owl

Now we can create a filter for these Species groups to add an extra dimension to this page’s visuals! As an example, here is the page with “Woodpecker” selected:



Data Collection KPI

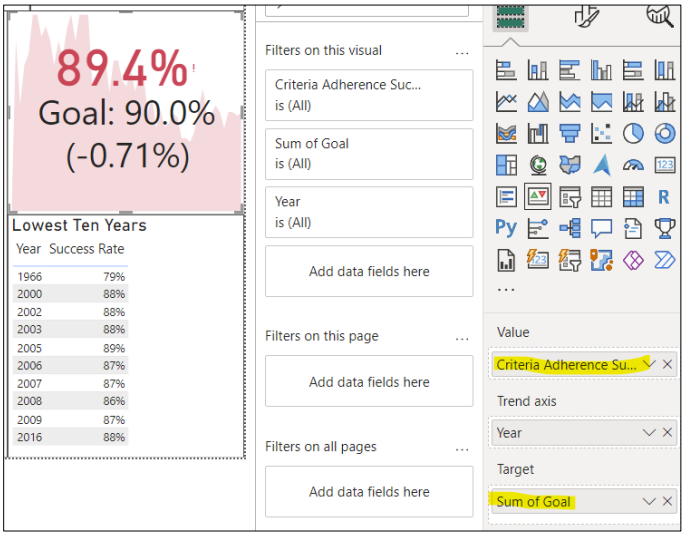
This final page will create and utilize a Key Performance Indicator visual to measure the Data Collection Success Rate of the most recent year for ensuring the accuracy and integrity of the data.

To calculate the KPI, we will first need to define what counts as a sample being "successful". If a sample does not meet all the survey's criteria to be complete and accurate, it is labelled as a "0" in the RunType column of the weather table. If it meets the criteria, it is labelled with a "1". The criteria includes the weather conditions, time, and location.

We will create a measure for the KPI visual to calculate the success rate named "Criteria Adherence Success Rate". The formula logic will be the percentage of runs that meet all data collection criteria for the most recent year:

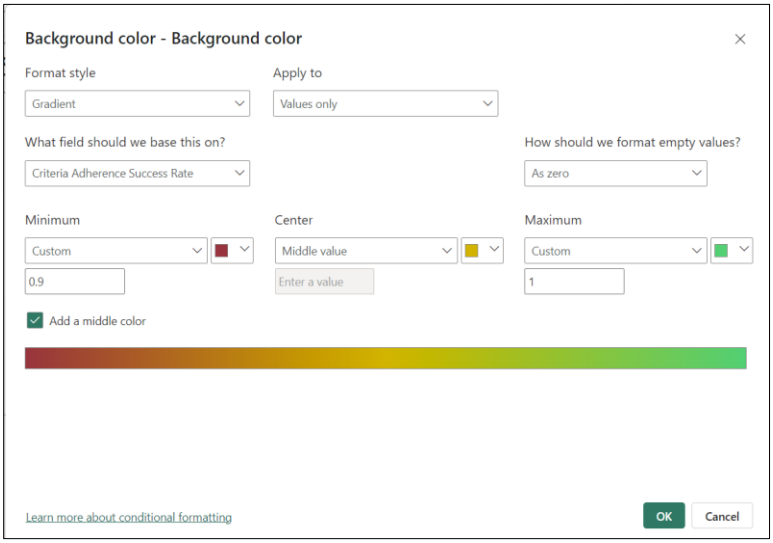
```
1 Criteria Adherence Success Rate = COUNTAX(FILTER(weather, weather[RunType] = 1), weather[RunType]) / COUNTROWS(weather)
```

Next, we will use the Criteria Adherence Success Rate" as the Value in the KPI visual, along with the Year value. Last, we add the Goal value from the Goal table to define the threshold:



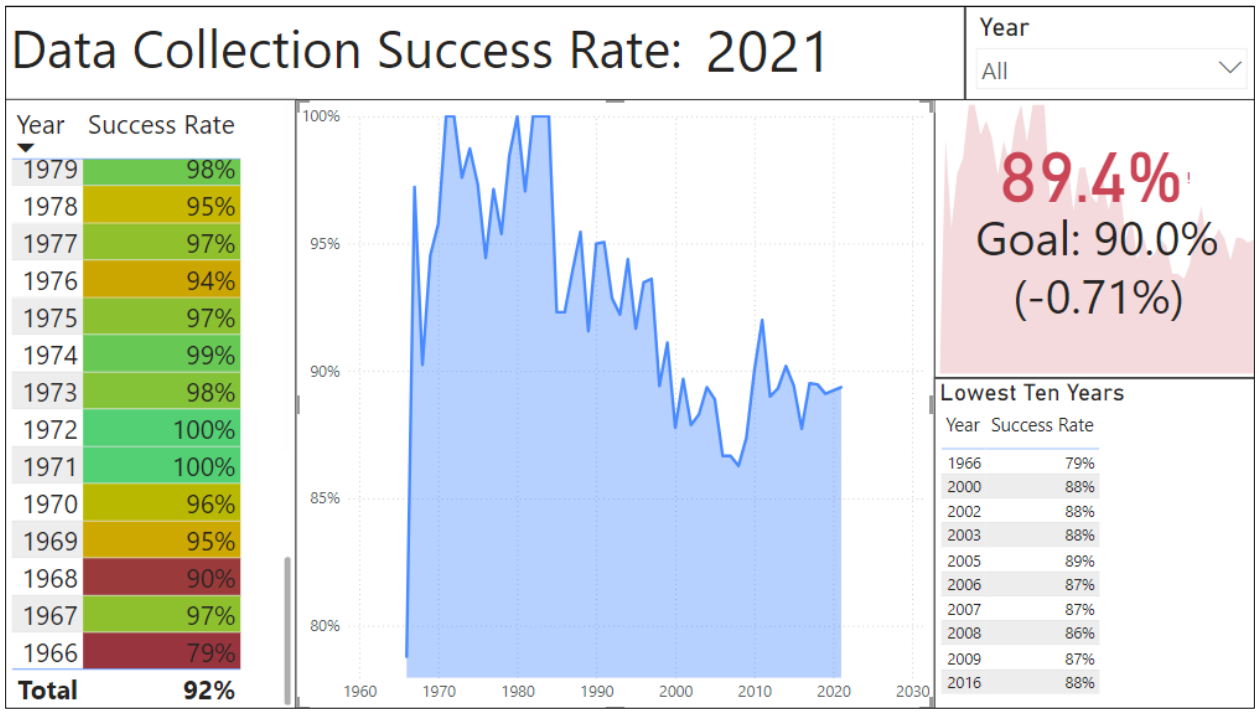
A few more visuals are added to add more perspective. One is the "Lowest Ten Years" to highlight the ten lowest years in terms of the success rate. We also add a stacked area chart to show the trend of success rates over time.

Last, we add a general table listing all years and their success rates. However, we will add an additional feature where it highlights each cell based on the value of the success rate. Green is highest, with the color gradient turning yellow the closer it is to the KPI goal/threshold of 90%, and progressively darker shades of red thereafter.



Year	Success Rate
1979	98%
1978	95%
1977	97%
1976	94%
1975	97%
1974	99%
1973	98%
1972	100%
1971	100%
1970	96%
1969	95%
1968	90%
1967	97%
1966	79%
Total	**92%**

In the end, we have our final page complete!



Citations

1. (Keith_Pardieck@usgs.gov), Keith Pardieck. BBS - USGS Patuxent Wildlife Research Center, www.pwrc.usgs.gov/bbs/. Accessed 24 June 2024.
2. “David Ziolkowski.” ScienceBase, www.sciencebase.gov/catalog/item/52b1dfa8e4b0d9b325230cd9. Accessed 24 June 2024.