Spring 2019

Mathematical Analysis of the Duck Migration to Louisiana

Brandon Garcia

Mathematical Analysis of the Duck Migration to Louisiana

Brandon Garcia

May 10, 2019

Project Advisor: Dr. Katie A. Evans

Abstract

The purpose of this project is to research the relationship between duck migration and weather patterns, more specifically trying to determine if the rainfall and temperature in a given year affects the migration patterns of ducks. Duck hunters and conservationists alike have observed an overall decrease in the duck population in Louisiana over the past 70 years. Though some years have seen an increase, the population has not recovered to the level from the 1950s. These observations have led to many questions about what have happened to the ducks or where have the ducks gone. Using different forms of Regression in Excel and Minitab this project investigates recent national patterns to try and correlate a pattern between the duck population and natural occurrences. At the end of the project it was discovered that there is a correlation between the duck population in Louisiana to Louisiana Temperature, Louisiana Precipitation, Minnesota Temperature, and Ontario Precipitation. This correlation provides a better understanding of how nature has an impact on a population of a single species and allows us to better predict the duck population at the end of the year from statistics in the beginning of the year.

Keywords: Duck migration, weather patterns, Regression, Louisiana Temperature, Louisiana Precipitation, Minnesota Temperature, Ontario Precipitation

1 Introduction

Imagine this, you are walking along the riverbank with your children and they are feeding the ducks corn and you get to see the joy on their faces as the ducks quack away at them.

This is a sight that you love to see and your kids love to have the interaction with the ducks. As time goes on you keep going back to feed the ducks, but over time you notice the ducks are slowly disappearing. You recall that the ducks have been there since you were a child so why are they gone suddenly? You deduce that they must have gotten old and passed away like many creatures of Earth. This is a reasonable reaction, however, since the 1950s the duck population has been on a decline and though they have come back in recent years they are nowhere near the numbers they have been in the past [1].

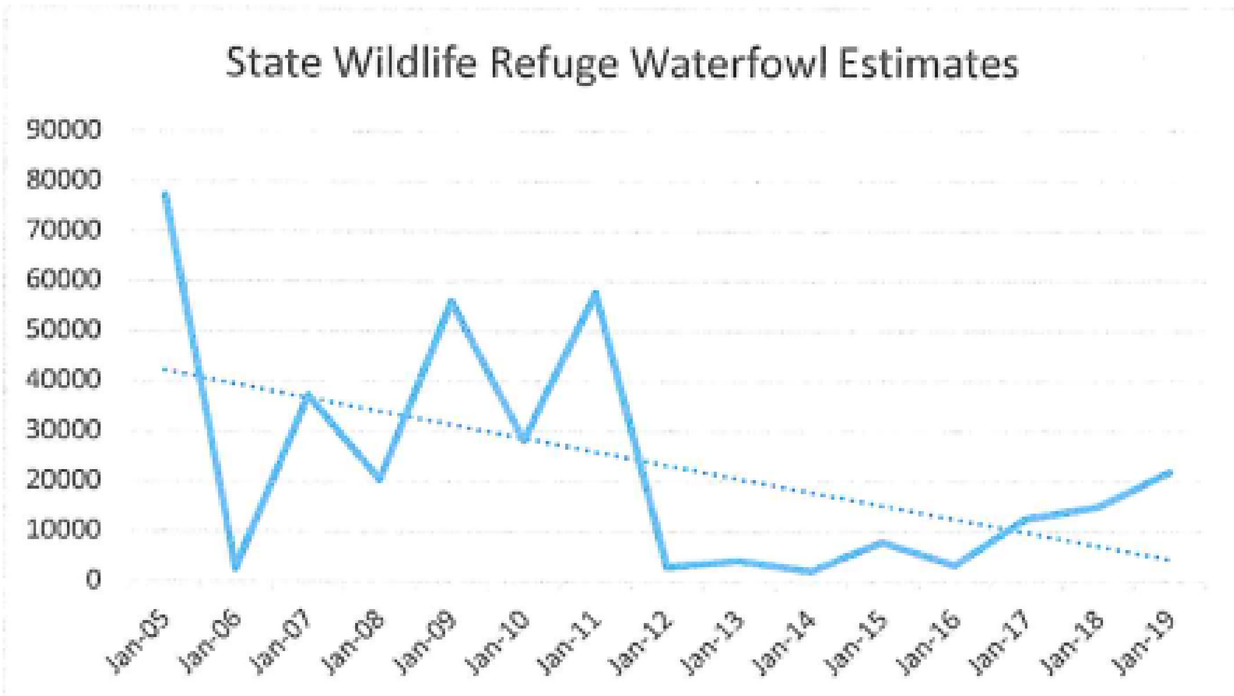

Figure 1: State Wildlife Refuge Waterfowl Estimates from Louisiana Wildlife and Fisheries [21]

Duck hunters are everywhere in the south and theres a chance that you or someone you know duck hunts. There has even probably been a time that they complain about not seeing ducks, or as many ducks, on their hunts. In Louisiana, the duck population in 2017 was 2,054,000 but in 2000 it was as high as 5,000,000 [21], over half the population of ducks disappeared in as little as 17 years. This is not just a worry for duck hunters though. Ducks are responsible for eating bugs and algae/vegetation in the water and they are responsible for fertilizing most of the soil in wetlands [20]. The predators that prey on ducks would also be affected and humans would no longer be able to hunt the ducks. Without duck hunters this would impact the economy, specifically the ones tied into the hunting community. So the purpose of this project was to try and better understand what has caused the duck population to decline and to try and find a hope of preventing further decreases.

In order to look at stopping the decline we must first look at the possible reasons for the decline in duck population in Louisiana. Two of the biggest factors would have to be the precipitation and temperature in the flyway and breeding grounds of the ducks. These are major factors because the reason a duck would leave the north is because the water is frozen and it has no food supply. It travels south to look for water that is not frozen and for an area with an abundance of food. As they travel south they will make stops along the way in states north of Louisiana until the cold hits them and they are pushed even farther south. A duck will fly ahead of a cold front a majority of the time and duck hunters who are aware of this will be sure to hunt on days leading to these cold fronts. Thus as a hypothesis, if there is a mild year with high precipitation the water in the northern states would be higher than it usually is, resulting in fewer ducks in Louisiana. If there is a cold year with low precipitation the water in the northern states would be lower than it usually is resulting in more ducks in Louisiana. This could be a huge factor on why in some years duck hunters see more ducks than other years. Another thing to look at is the weather through the states that the ducks would fly through and see if that ties into affecting the population of ducks in Louisiana as well.

2 Background Literature

Many organizations are trying to help the duck population get back to their prominent numbers. One, if not the most notable, is Ducks Unlimited. This nonprofit organization is primarily focused on conserving, restoring, and managing wetlands so that the duck population can be ever lasting [14]. With the help of Ducks Unlimited we can see the duck population over the years in certain areas throughout the United States and Canada. They use both ground and aerial surveyors to give a rough estimate of the duck population. They do this by looking at the data from the ground and aerial surveys and use visibility correction factors to get a general idea of how many ducks there are. The visibility correction factors look at areas with dense forests where they can see minimal amount of ducks in the area from the aerial surveyors but with the help of ground surveyors they know that there are ducks in these areas. This results in a rough estimate of the population in that area but it is an estimate that is viable. [16]. This is a time consuming and tedious task for the surveyors but the information gathered is very important to Ducks Unlimited and the hunters across the nation.

Other organizations determined to help the duck population are U.S. Fish and Wildlife Services and individual states Wildlife and Fisheries Departments. The U.S. Fish and Wildlife Services also does their own survey of waterfowl population and survey of their habitats to give more data on the populations of ducks in their breeding grounds. They use their own technology, with the help of state wildlife agencies and the Canadian Wildlife Service, to give the best estimate possible [17]. Similar to Ducks Unlimited they use aerial and ground surveys to conclude their estimates. The individual states Wildlife and Fisheries Departments use the funds gathered from hunters buying hunting licenses, permits, and tickets that a hunter may receive from a Game Warden in order to ensure that habitats are being conserved. They ultimately use the money that hunters spend in order to hunt legally and put that back into the wildlife. They also gather information to see how many ducks go to certain states, and they do that by collecting information from the hunters. On Wildlife Management Areas (WMA), a hunter must report what they are doing when they get to the WMA and they must report their kill when they leave the WMA. The Game Wardens will collect this data and analyze it so that they can have a better understanding of how many ducks actually come to their respected state. In this project we use the data collected from the Louisiana Wildlife and Fisheries to get the duck population in Louisiana over the past 18 years.

As we stated, our hypothesis is that when there is a lot of precipitation and a colder year there is a higher population in Louisiana, and when there is a low amount of precipitation and a hotter year there is a lower population in Louisiana. The areas that we are most curious about are in the Mississippi Flyway. The Mississippi Flyway follows the Mississippi River from the northern states all the way down to Louisiana and Mississippi and this is the main output for the duck population in Louisiana [23]. The part of the flyway that we are primarily focused on consists of the breeding grounds, Missouri, Arkansas, and Louisiana [23]. The breeding grounds are Wisconsin, Minnesota, Ontario (Canada), and Michigan. We looked at the weather characteristics in these states as we followed them down to Louisiana. In order to get correct data we had to go to several cites in order to look at weather over the months and years in each state. We began by collecting data for both the precipitation and temperature from 2000-2017 in Louisiana [7], Arkansas [8], Missouri [5], Wisconsin [19], Ontario [15], Michigan [9], and Minnesota [4]. The original data, as well as the duck population in Louisiana [21], can be found in an exert from Excel in Appendix A.

3 Methods

Once the data was collected we went through several obstacles to get to the point where we were satisfied with the results. In order to be satisfied with the results we used Multiple Linear Regression to find p-values and R-squared values so that we had some metric that shows our data is acceptable/appropriate. P-values are calculated probabilities that help you determine the significance of your results and for it to be significant to our results we used $P < 0.05$. This means that our results would have a less than 1 in 20 chance of being wrong [13]. R-Squared values are statistical measures that represents the proportion of the variance for a dependent variable that is explained by an independent variable or variables in a regression model [26]. It ranges from $0 - 1$ but is expressed in percentages usually and the closer to 1 or 100% the better the model is. For Multiple Regression, which has several independent variables, the R-Squared must be adjusted which compares the descriptive power of regression models. The Adjusted R-Squared compensates for multiple independent variables, so with more variables being added to a model the Adjusted R-Square is what is best to look at. For our models we accepted R-Squared values greater than 50% because we wanted the variance to be over half the proportion between the dependent and independent variables. Thus between the p-values and R-Squared values we felt that our data was very acceptable at the end of the analysis.

The first tool we used to look at Linear Regression was Excel. We put all the information into Excel, and then used graphs to check for a correlation between the Duck Population and the Temperature or Precipitation in certain states. We were implementing Regression models in Excel and with all of the different data it was only trying to show a relation between two variables, the Population and another variable ranging from Precipitation in Louisiana to Temperature in Ontario. What Excel was implementing was Linear Regression, which attempts to model a relationship between two variables, one which is independent and the other dependent [3]. However, though this is a good model when looking at two variables, we are trying to find a good model that shows a relationship between multiple independent variables and a single dependent variable.

Multiple Linear Regression attempts to model a relationship between two or more independent variables and a single dependent variable [6], which is exactly what we were wanting to implement with our data. Since Excels built-in statistical methods are limited in capability we thought it would be best to proceed with the regression models in Minitab. We converted the same data from Excel to Minitab and began to run models to try and see which independent variables (Temperature and Precipitation) most related to the dependent variable (Population). In the beginning, we were not getting the results we were looking for, we were getting p-values that were not accepted while getting R-Squared values that were accepted. This contradiction resulted in us trying other Regression Models to see if we could find a better fit model in Minitab, and if that did not work we even thought about moving on to other statistical based tools like Matlab. The Regressions that were implemented were Logistic Regression, Nonlinear Regression, Poisson Regression, and Orthogonal Regression. We also looked into the Cluster Obstervatins and Item Analysis which are other statistical methods that Minitab allows you to use.

Logistic Regression is used when the dependent variable is dichotomous (binary), and like multiple linear regression, it is used with one or more independent variables [18]. However, since our dependent variable was not dichotomous this regression model was not very helpful for this data analysis. Nonlinear Regression models are regression equations that do not follow the rules for a linear model meaning that they are simply nonlinear. Nonlinear Regression models tend to fit an enormous variety of curves. Which forces the user of these models to conduct research to determine which functional form will provide the best fit for the users data. One of the biggest differences between Linear and Nonlinear Regression models are that Nonlinear Regression will have invalid R-Squared values nor can statistical software calculate p-values for the terms either [24]. With no metric to show that a model is acceptable or appropriate it made Nonlinear Regression unusable. Poisson Regression is similar to both multiple linear and logistic regression models. It is similar to multiple linear regression since it shows a relationship between multiple independent variables to a single dependent variable. However, the dependent variable is an observed count that follows the Poisson distribution making it have a discrete response variable similar to the logistic regression. The difference between the logistic regression and this model is that Poisson Regression results in nonnegative integers [2]. Even though it does show a relationship between multiple independent variables and a single dependent variable it still is not suitable for our models we are looking for in this analysis so like logistic regression, it is not helpful. Orthogonal Regression is used to show the relationship between two continuous variables, one independent and one dependent. It is similar to simple linear regression except that both the independent and the dependent variables contain measurement error. In linear regression only the independent variable has error [12]. Since Orthogonal Regression only shows the relationship between a single independent variable and dependent variable it was not a viable model to use on multiple independent variables. Cluster Observations is used to join observations that share common characteristics into groups and is mainly used when the analyst does not have any initial information about how to form groups with their data [10]. Since we knew the information about our data and we already had them broken into their appropriate groups of Temperature and Precipitation it was not useful to use Cluster Observations. Item Analysis tells us how well a set of questions measure one characteristic and helps to identify questions that are problematic which was of no use with the mathematical analysis of the duck population to Louisiana [25]. Going through these other models it made it clear that for our statistical data it was best to be analyzed by using Multiple Linear Regression.

4 Data

p-values for Regression Models Tested

Model #	Louisiana Temp	Louisiana Prec	Missouri Temp	Missouri Prec	Arkansas Temp	Arkansas Prec	Wisconsin Temp	Wisconsin Prec	Minnesota Temp	Minnesota Prec	Ontario Temp	Ontario Prec	Michigan Temp	Michigan Prec	Model R-Square (%)	Adj R-Square (%)
1	0.666	0.268	0.329	0.582	0.87	0.629	0.548	0.262	0.138	0.212	0.251	0.201	0.835	0.807	82.13	0
2	0.835	0.369	0.278	0.556	0.794	0.624	0.541	0.307	0.161	0.264	0.273	0.267	0.84	0.856	76.64	0
3	0.855	0.589	0.743	0.588	0.788	0.424									31.16	0
4							0.713	0.83	0.301	0.844	0.771	0.315	0.319	0.688	30.2	0
5	0.386		0.396		0.167										14.89	0
6		0.317		0.38		0.39									29.35	14.21
7							0.244		0.132		0.768		0.346		20.92	0
8								0.635		0.847		0.246		0.36	16.7	0
9	0.8876	0.2248	0.1062	0.3296	0.8425	0.5248	0.3648	0.1862	0.0568	0.129	0.1472	0.1379			75.9	0
10	0.423	0.969	0.882	0.796	0.555	0.722	0.424	0.531	0.693	0.527	0.886	0.529	0.616	0.529	90.86	22.32
11	0.502	0.171	0.381	0.566	0.986	0.646	0.531	0.205	0.107	0.155	0.21	0.137	0.87	0.736	88.26	33.5
12		0.007	0.027					0.043	0.018	0.049	0.138	0.007			72.58	53.39
13		0.021					0.541	0.399	0.387	0.386		0.068			52.28	26.25
14	0.278		0.631	0.973	0.543	0.308	0.577	0.484	0.207	0.255	0.317	0.322	0.711	0.952	75.32	0
15	0.256		0.572		0.624		0.199		0.101		0.984		0.876		37.99	0
16		0.201		0.531		0.667		0.469		0.422		0.147		0.975	55.06	23.61
17		0.015						0.386		0.387		0.025			47.49	31.33
18	0.038	0.008							0.052		0.787	0.026			64.51	49.72
19	0.044	0.01							0.098			0.156			65.2	46.22
20	0.028	0.016					0.236	0.855	0.056			0.041			69.18	52.36
21	0.261	0.039			0.373	0.963			0.069			0.054			67.08	49.12
22	0.14	0.017	0.847	0.876					0.113			0.036			64.56	45.23
23	0.031	0.006							0.038			0.021			64.28	53.29
24		0.01	0.086					0.103	0.052	0.161		0.015			65.94	47.36
25	0.107	0.007	0.596					0.142	0.044	0.109		0.016			74.05	55.89

Acceptable p-values
Acceptable R-square
Acceptable R-square (Adj)

Figure 2: P-Values, R-Squared, and Adjusted R-Squared for Regression Models Tested

From the data pictured above, and in Appendix C, it is visible that we did a total of 25 different regressions in Minitab. Most of them were Fitted-Regression Models. Fitted Regression Models are models that describe the relationship between a set of independent variables and a continuous dependent variable and we accomplished this by choosing Stat > Regression > Regression > Fit Regression Model [11]. This would bring up a small window, pictured below, and with this window we were able to choose the response variable, or the dependent variable, and then choose the predictors, or independent variables, that we wanted to show a relationship with. After several trials and errors we discovered that in 'options' we could choose to transform the data, more specifically we could choose a Box-Cox Transformation.

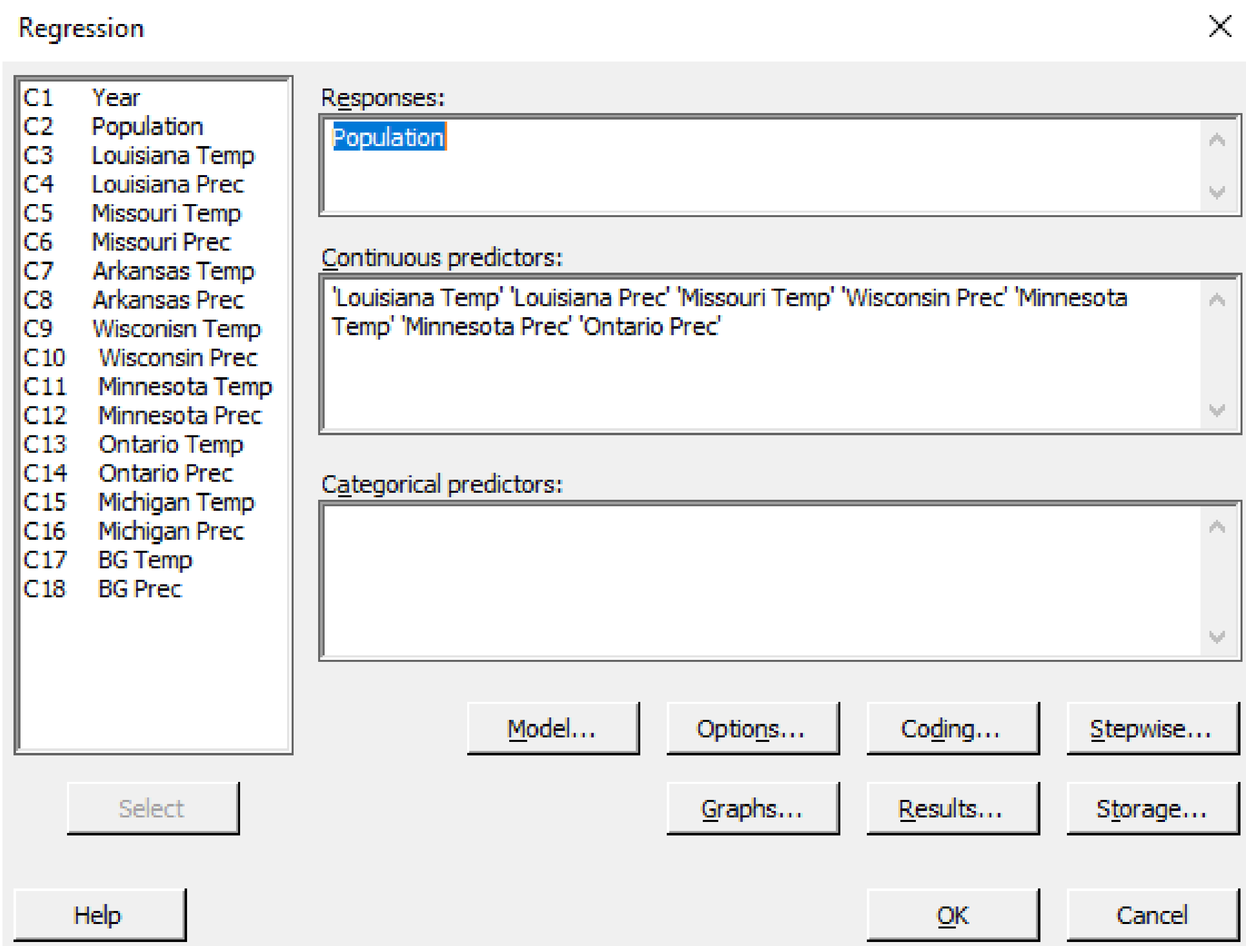

Figure 3: Steps to use Fitted Regression in Minitab

A Box-Cox Transformation would result in the best possible Fitted Regression Model. Box-Cox Transformation is a way to transform non-normal dependent variables into a normal shape resulting in an ability to run a broader number of tests [22]. In Figure 1, we have cleaned up the data so that we are focusing on the Variables, p-Values, and R-Squared values. From here we have highlighted the data in yellow to represent the accepted p-Values,

values which are lower than .05. We also highlighted the accepted R-Squared values, which are higher than 50 percent, in a light blue and the acceptable R-Squared (adj) values which are highlighted in a light red. From the picture we can see that there a lot of accepted R-Squared values but there are not many p-Values that are accepted and there are even less amount of accepted R-Squared Values (adj). In fact, Trial 10 had the highest R-Squared value but it did not have any accepted p-Values. Trial 25 had the highest R-Squared (adj) with an acceptable R-Squared but it resulted in several p-values that were not acceptable. Trial 23 had all p-Values accepted and even had an accepted R-Squared value and accepted R-Squared (adj) value. Trial 23's values were so good that it is the trial that we concluded was the best fit for the project. By looking at Trial 23's values we can see that by looking at the Louisiana Temperature, Louisiana Precipitation, Minnesota Temperature, and Ontario Precipitation we would be able to account for 53.29% of the duck population in Louisiana.

Knowing this we can now look at the Regression Equation that was given by the Multiple Linear Regression model from Minitab. The regression equation in Appendix C is $-\text{Population}^{(-1)} = (0.000003085034502) - (0.0000000051380423) \times \text{Louisiana Precipitation} - (0.0000000057818968) \times \text{Ontario Precipitation} + (0.0000000363464390) \times \text{Minnesota Temperature} - (0.0000000662295036) \times \text{Louisiana Temperature}$. With this equation we can plug in the data for a given year and have an estimate of the population of ducks in Louisiana in January. We say January because the duck population used in this data was found in January of each year.

5 Conclusion

With the world in dismay and the temperatures and precipitation getting hotter and less, it is a matter of time before mother nature takes it toll on the world we know today. Everyday it seems we hear that a day sets a new record on how hot it is or we see wildfires spreading across beautiful lands due to droughts in the summer heat. With the temperature on the rise and the precipitation on the decline, it is a wonder that humans are not seeing a decline of more and more species. Hunters and conservationists notice these things though, because they are on the lookout for the changes in a species population. It may be for hunting or just going out to bird watch, but these are the people noticing a change in an animals population. These are the people who make it known and these are maybe the people to listen to before we start losing a species all together. The duck population in Louisiana is drastically on the decline and though this may not be a necessary species for some people it should be a sign of what is happening around us. As we can see from this data analyst, 53.29% of the duck population in Louisiana can be figured out from the Louisiana Temperature, Louisiana Precipitation, Minnesota Temperature, and Ontario Precipitation. These four factors account for more than 50% of the duck population in Louisiana in a given year and these are just Temperature and Precipitation alone. With a these two factors constantly changing it will be interesting to see what the duck population in Louisiana will look like in the near future.

Waterfowl populations are adapted well to short-term swings in habitat conditions, but we must continue to guard against the long-term loss of breeding habitat" [27]. Going forward the project may need to have a wider variety of data in terms of years. Instead of looking at 2000-2017 it could be expanded into the 1900s to see how the graphs look then. For this project we focused primarily on Linear Regression, maybe looking at a bigger picture it could be another form of regression model so this is maybe something to look into as well. Someone who picks up this project could also look into the other possibilities of the loss of the duck population in Louisiana, maybe from hunters in the states above Louisiana or any other predator species that targets waterfowl.

References

[1] Breeding grounds data.

[2] Home.

[3] Linear regression.

[4] Minnesota climate trends.

[5] Missouri climate.

[6] Multiple linear regression.

[7] Noaa national centers for enviromental information state climate summaries.

[8] Noaa national centers for enviromental information state climate summaries.

[9] Noaa national centers for enviromental information state climate summaries.

[10] Overview for cluster observations.

[11] Overview for fit regression model.

[12] Overview for orthogonal regression.

[13] P values.

[14] Rating for ducks unlimited.

[15] Toronto monthly climate averages.

[16] Waterfowl breeding and habitat surveys.

[17] Waterfowl population and habitat surveys are underway!

[18] What is logistic regression?

[19] Wisconsin climate.

[20] Mallards are everywhere, and that's great for wetlands, Mar 2017.

[21] Aerial waterfowl surveys, 2018.

[22] Box cox transformation, May 2018.

[23] Du projects: Mississippi flyway, 2018.

[24] Jim Frost. The difference between linear and nonlinear regression models, Nov 2018.

[25] Meredith Griffith. Item analysis with cronbach's alpha for reliable surveys.

[26] Adam Hayes. R-squared, May 2019.

[27] Joe Macaluso. Survey shows duck breeding numbers down, Aug 2018.

6 Appendix A

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Louisiana				Missouri		Arkansas		Wisconsin		Minnesota		Ontario		Michigan		Breeding	
2	Year	Precipitation (in)	Average Temp. (F)	Population	Prec.	Ave. Temp. (F	Prec	Avg Temp	Avg Temp	Prec	Avg Temp	Prec	Avg Temp	Prec	Avg Temp	Prec	Avg Temp	Avg Prec
3	2000	46.71	67.25	5000000	37.37	56.2	40.16	63.2	41.93	32.4	41.74	27.93	51.07	45.13	47.2	40.08	45.485	36.385
4	2001	78.26	66	3150000	35.29	57.7	56.01	63	45.08	35.4	42.99	29.23	54.72	34.62	49.3	35.38	48.0225	33.6575
5	2002	74.42	66.67	2750000	40.95	57.9	51.91	62.8	43.92	36	41.97	28.51	47.28	35.4	49	22.25	45.5425	30.54
6	2003	53.22	67.25	3200000	46.06	56.5	35.37	62.2	44.42	28.44	41.44	22.17	56.99	35.59	47.2	29.7	47.5125	28.975
7	2004	80.35	67.67	1659000	42.27	57.6	55.37	63.3	43	35.4	41.21	29.79	53.1	37.65	47.7	29.91	46.2525	33.1875
8	2005	41.33	67.67	2537000	37.85	58	38.62	64	44.5	29.64	42.98	31.51	49.85	31.98	48	28.96	46.3325	30.5225
9	2006	51.97	67.67	2953000	29.93	58.5	43.05	64	45.5	30.84	44.32	22.68	76.05	42.39	48.8	38.46	53.6675	33.5925
10	2007	57.75	67.42	4286000	30.57	58.3	45.51	63.9	44.25	34.2	42.32	29.14	57.97	26.69	48.5	30.94	48.26	30.2425
11	2008	58.25	66.83	1831000	57.96	55.5	46.5	62.1	41.08	33.48	39.35	27.31	50.71	45.65	47.5	34.38	44.66	35.205
12	2009	62.71	67	1595000	50.92	56.6	78.02	62.3	42.3	30.12	40.02	26.35	45.95	66.12	46.3	35.4	43.6425	39.4975
13	2010	42.46	66.75	2699000	39.07	58	31.97	63.7	45.42	39.12	42.79	33.32	48.65	60.42	49.2	25.57	46.515	39.6075
14	2011	35.87	67.75	3284000	47.17	58.7	48.7	64.1	43.5	30.72	41.94	24.29	48.51	43.35	48.8	40.95	45.6875	34.8275
15	2012	66.97	68.83	2792000	32.3	61.2	45.69	65.1	46.42	29.16	45.12	26.15	52.25	30.77	51.6	30.85	48.8475	29.2325
16	2013	56.56	66.33	2621000	42.68	56.4	52.66	61.9	41.16	36.72	39.23	28.96	50.59	34.64	48.2	32.63	44.795	33.2375
17	2014	60	65.25	2872000	43.43	55.8	41.41	60.6	39.3	37.2	38.61	29.04	48.2	38.7	46.2	35.42	43.0775	35.09
18	2015	65.12	68.33	3385000	61.24	58.8	51.31	63.2	44.17	35.52	43.52	28.26	50	37.9	49.6	29.3	46.8225	32.745
19	2016	80.96	69	1849000	41.44	60.4	54.38	64.5	45.83	39.48	44.5	32.17	51.49	32.59	51.5	30.74	48.33	33.745
20	2017	69.32	69.17	2054000	36.65	60.2	50.8	64.3	44	37.8	42.73	27.21	49.1	42	49.2	35.71	46.2575	35.68

Figure 4: Raw Data of Collected Precipitation and Temperature from Each State in Excel

7 Appendix B

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
	Year	Population	Louisiana Temp	Louisiana Prec	Missouri Temp	Missouri Prec	Arkansas Temp	Arkansas Prec	Wisconisn Temp	Wisconsin Prec	Minnesota Temp	Minnesota Prec	Ontario Temp	Ontario Prec	Michigan Temp	Michigan Prec
1	2000	5000000	67.25	46.71	56.2	37.37	63.2	40.16	41.93	32.40	41.74	27.93	51.07	45.13	47.2	40.08
2	2001	3150000	66.00	78.26	57.7	35.29	63.0	56.01	45.08	35.40	42.99	29.23	54.72	34.62	49.3	35.38
3	2002	2750000	66.67	74.42	57.9	40.95	62.8	51.91	43.92	36.00	41.97	28.51	47.28	35.40	49.0	22.25
4	2003	3200000	67.25	53.22	56.5	46.06	62.2	35.37	44.42	28.44	41.44	22.17	56.99	35.59	47.2	29.70
5	2004	1659000	67.67	80.35	57.6	42.27	63.3	55.37	43.00	35.40	41.21	29.79	53.10	37.65	47.7	29.91
6	2005	2537000	67.67	41.33	58.0	37.85	64.0	38.62	44.50	29.64	42.98	31.51	49.85	31.98	48.0	28.96
7	2006	2953000	67.67	51.97	58.5	29.93	64.0	43.05	45.50	30.84	44.32	22.68	76.05	42.39	48.8	38.46
8	2007	4286000	67.42	57.75	58.3	30.57	63.9	45.51	44.25	34.20	42.32	29.14	57.97	26.69	48.5	30.94
9	2008	1831000	66.83	58.25	55.5	57.96	62.1	46.50	41.08	33.48	39.35	27.31	50.71	45.65	47.5	34.38
10	2009	1595000	67.00	62.71	56.6	50.92	62.3	78.02	42.30	30.12	40.02	26.35	45.95	66.12	46.3	35.40
11	2010	2699000	66.75	42.46	58.0	39.07	63.7	31.97	45.42	39.12	42.79	33.32	48.65	60.42	49.2	25.57
12	2011	3284000	67.75	35.87	58.7	47.17	64.1	48.70	43.50	30.72	41.94	24.29	48.51	43.35	48.8	40.95
13	2012	2792000	68.83	66.97	61.2	32.30	65.1	45.69	46.42	29.16	45.12	26.15	52.25	30.77	51.6	30.85
14	2013	2621000	66.33	56.56	56.4	42.68	61.9	52.66	41.16	36.72	39.23	28.96	50.59	34.64	48.2	32.63
15	2014	2872000	65.25	60.00	55.8	43.43	60.6	41.41	39.30	37.20	38.61	29.04	48.20	38.70	46.2	35.42
16	2015	3385000	68.33	65.12	58.8	61.24	63.2	51.31	44.17	35.52	43.52	28.26	50.00	37.90	49.6	29.30
17	2016	1849000	69.00	80.96	60.4	41.44	64.5	54.38	45.83	39.48	44.50	32.17	51.49	32.59	51.5	30.74
18	2017	2054000	69.17	69.32	60.2	36.65	64.3	50.80	44.00	37.80	42.73	27.21	49.10	42.00	49.2	35.71

Figure 5: Raw Data of Collected Precipitation and Temperature from Each State in Minitab

8 Appendix C

p-values for Regression Models Tested

Model #	Louisiana Temp	Louisiana Prec	Missouri Temp	Missouri Prec	Arkansas Temp	Arkansas Prec	Wisconsin Temp	Wisconsin Prec	Minnesota Temp	Minnesota Prec	Ontario Temp	Ontario Prec	Michigan Temp	Michigan Prec	Model R-Square (%)	Adj R-Square (%)
1	0.666	0.268	0.329	0.582	0.87	0.629	0.548	0.262	0.138	0.212	0.251	0.201	0.835	0.807	82.13	0
2	0.835	0.369	0.278	0.556	0.794	0.624	0.541	0.307	0.161	0.264	0.273	0.267	0.84	0.856	76.64	0
3	0.855	0.589	0.743	0.588	0.788	0.424									31.16	0
4							0.713	0.83	0.301	0.844	0.771	0.315	0.319	0.688	30.2	0
5	0.386		0.396		0.167										14.89	0
6		0.317		0.38		0.39									29.35	14.21
7							0.244		0.132		0.768		0.346		20.92	0
8								0.635		0.847		0.246		0.36	16.7	0
9	0.8876	0.2248	0.1062	0.3296	0.8425	0.5248	0.3648	0.1862	0.0568	0.129	0.1472	0.1379			75.9	0
10	0.423	0.969	0.882	0.796	0.555	0.722	0.424	0.531	0.693	0.527	0.886	0.529	0.616	0.529	90.86	22.32
11	0.502	0.171	0.381	0.566	0.986	0.646	0.531	0.205	0.107	0.155	0.21	0.137	0.87	0.736	88.26	33.5
12		0.007	0.027					0.043	0.018	0.049	0.138	0.007			72.58	53.39
13		0.021					0.541	0.399	0.387	0.386		0.068			52.28	26.25
14	0.278		0.631	0.973	0.543	0.308	0.577	0.484	0.207	0.255	0.317	0.322	0.711	0.952	75.32	0
15	0.256		0.572		0.624		0.199		0.101		0.984		0.876		37.99	0
16		0.201		0.531		0.667		0.469		0.422		0.147		0.975	55.06	23.61
17		0.015						0.386		0.387		0.025			47.49	31.33
18	0.038	0.008							0.052		0.787	0.026			64.51	49.72
19	0.044	0.01							0.098			0.156			65.2	46.22
20	0.028	0.016					0.236	0.855	0.056			0.041			69.18	52.36
21	0.261	0.039			0.373	0.963			0.069			0.054			67.08	49.12
22	0.14	0.017	0.847	0.876					0.113			0.036			64.56	45.23
23	0.031	0.006							0.038			0.021			64.28	53.29
24		0.01	0.086					0.103	0.052	0.161		0.015			65.94	47.36
25	0.107	0.007	0.596					0.142	0.044	0.109		0.016			74.05	55.89

Acceptable p-values

Acceptable R-square

Acceptable R-square (Adj)

Figure 6: P-Values R-Squared and Adjusted R-Squared for Regression Models Tested

9 Appendix D

Regression Analysis: Population versus Louisiana Prec, ... uisiana Temp

Method

Box-Cox transformation	
Rounded λ	-1
Estimated λ	-1.08976
95% CI for λ	(-1.09226, -1.08426)

Analysis of Variance for Transformed Response

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Regression	4	0.000000	64.28%	0.000000	0.000000	5.85	0.006
Louisiana Prec	1	0.000000	19.97%	0.000000	0.000000	10.70	0.006
Ontario Prec	1	0.000000	26.04%	0.000000	0.000000	6.84	0.021
Minnesota Temp	1	0.000000	2.29%	0.000000	0.000000	5.36	0.038
Louisiana Temp	1	0.000000	15.99%	0.000000	0.000000	5.82	0.031
Error	13	0.000000	35.72%	0.000000	0.000000		
Total	17	0.000000	100.00%				

Model Summary for Transformed Response

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)
0.0000001	64.28%	53.29%	0.0000000	41.35%

Coefficients for Transformed Response

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	0.000003	0.000001	(-0.000000, 0.000006)	2.09	0.057	
Louisiana Prec	-0.000000	0.000000	(-0.000000, -0.000000)	-3.27	0.006	1.11
Ontario Prec	-0.000000	0.000000	(-0.000000, -0.000000)	-2.62	0.021	1.18
Minnesota Temp	0.000000	0.000000	(0.000000, 0.000000)	2.32	0.038	2.06
Louisiana Temp	-0.000000	0.000000	(-0.000000, -0.000000)	-2.41	0.031	1.97

Figure 7: Regression Analysis Output of Model 23

Coefficients for Transformed Response

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	0.000003	0.000001	(-0.000000, 0.000006)	2.09	0.057	
Louisiana Prec	-0.000000	0.000000	(-0.000000, -0.000000)	-3.27	0.006	1.11
Ontario Prec	-0.000000	0.000000	(-0.000000, -0.000000)	-2.62	0.021	1.18
Minnesota Temp	0.000000	0.000000	(0.000000, 0.000000)	2.32	0.038	2.06
Louisiana Temp	-0.000000	0.000000	(-0.000000, -0.000000)	-2.41	0.031	1.97

Regression Equation

-Population^-1 = 0.000003 - 0.000000 Louisiana Prec - 0.000000 Ontario Prec
+ 0.000000 Minnesota Temp - 0.000000 Louisiana Temp

Fits and Diagnostics for Unusual Observations

Original Response

Obs	Population	Fit	95% CI
6	2537000	4313946	(3067653, 7265827)

Fits and Diagnostics for Unusual Observations

Transformed Response

Obs	Population'	Fit	SE Fit	95% CI	Resid	Std Resid
6	-0.000000	-0.000000	0.000000	(-0.000000, -0.000000)	-0.000000	-2.29

Obs	Del Resid	HI	Cook's D	DFITS	
6	-2.84	0.273712	0.39	-1.74410	R

Population' = transformed response
R Large residual

Figure 8: Regression Analysis Output of Model 23 Continued