JOURNAL OF INFORMATION SYSTEMS APPLIED RESEARCH

Volume 15, Issue 3
October 2022
ISSN: 1946-1836

Special Issue: Data and Business Analytics

In this issue:

iscap
Information Systems & Computing Academic Professionals

15 (3)
October 2022

The **Journal of Information Systems Applied Research** (JISAR) is a double-blind peer reviewed academic journal published by ISCAP, Information Systems and Computing Academic Professionals. Publishing frequency is three to four issues a year. The first date of publication was December 1, 2008.

JISAR is published online (https://jisar.org) in connection with CONISAR, the Conference on Information Systems Applied Research, which is also double-blind peer reviewed. Our sister publication, the Proceedings of CONISAR, features all papers, panels, workshops, and presentations from the conference. (https://conisar.org)

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the conference. At that point papers are divided into award papers (top 15%), other journal papers (top 30%), unsettled papers, and non-journal papers. The unsettled papers are subjected to a second round of blind peer review to establish whether they will be accepted to the journal or not. Those papers that are deemed of sufficient quality are accepted for publication in the JISAR journal. Currently the target acceptance rate for the journal is under 38%.

Questions should be addressed to the editor at editor@jisar.org or the publisher at publisher@jisar.org. Special thanks to members of ISCAP who perform the editorial and review processes for JISAR.

Journal of Information Systems Applied Research

Classification of Hunting-Stressed Wolf Populations Using Machine Learning

John C. Stewart
stewartj@rmu.edu

G. Alan Davis
davis@rmu.edu

Diane Igoche
igoche@rmu.edu

Department of Computer Information Systems
Robert Morris University
Moon Township, PA 15108 USA

Abstract

The preservation of Wolf populations in North America has been controversial for hundreds of years. The preservation of ecosystems or the reintroduction of wolf populations in areas to redress the ecological balance has taken place in recent decades. In other areas, wolves are hunted in an effort to manage them. Previous studies have identified physiological characteristics as an indicator of higher stress levels in individual wolf subjects in heavily hunted populations. This stress impacts reproduction, social structure and pack dynamics. The current study supports a prior study that used statistics to show elevated stress levels in hunted wolf populations. Using machine learning (k-nearest neighbor), we were able to classify individual wolf subjects as belonging to hunting-based stressed populations based on physiological data with high accuracy.

Keywords: Machine Learning, data mining, k-NN, physiological indicators, classification

1. INTRODUCTION/LITERATURE REVIEW

Human originated mortality of predator populations (i.e., hunting) has been documented as having a myriad of additional negative impacts on the affected population (Coltman, 2003, Darimont 2009). Hunting traditionally has the goal of selecting the strongest and fittest, thus impacting reproduction by reducing breeding of the healthiest members of the populations. Studies of trophy hunting of rams determined the effects of selection, placed an emphasis on harvesting of trophy rams of heavier weight and larger horn size (Festa-Bianchet, et al. 2004, Coltman, et al., 2003). Rams with higher value in terms of breeding were found to be shot at a lower age, eliminating their reproductive value to the populations (Coltman, et al., 2002).

The complicated social structure of wolf populations makes them extremely vulnerable to elevated mortality and a disruption of behavior dynamics that would occur from human intervention (Haber, 1996). While wolves can recover from a moderate decrease in population, ongoing pressures can affect behavior, the components of social structures, and genetic factors. This combination of factors can have potential long-term impacts on group and pack recovery (Rausch 1967; Haber 1996;

Jezdrzejewski et al. 2005; Sidorovich et al. 2007; Rutledge et al.2010, 2012).
Wolf populations that are heavily impacted by hunting predictably produce more female offspring (Sidorovich, et al, 2007). In addition, genetic diversity in wolf populations is affected by intense hunting (Jezdrzejewski et al. 2005). As an example, researchers have found that harvesting of wolves outside protected areas can impact the social dynamics of neighboring populations (Rutledge, et al.,2010). Further, and not surprisingly, wolf pup mortality is a critical factor in the rate of population growth (Rausch, 1967).

While changes in population numbers are easily measured, physiological impacts of hunting have only been documented in a very limited number of studies. Elevated levels of hormones like cortisol are an indicator of increased stress in hunted individual subjects (Bateson and Bradshaw 2007). Additionally, stress can negatively affect the social behavior of the target species population (Gobish, et al. 2008).

Testosterone is vital to male reproduction capability but is also an indicator of behavior. Within the social structure of the population, testosterone may be found to increase when there is an imbalance in that component (Oliveria, 2004).

Several studies have found elevated levels of the hormones cortisol, testosterone, and progesterone in pregnant females, giving an indication of the reproductive activity in the population (Foley, et al. 2001). A few studies have proposed a relationship between female testosterone levels and the social structure of the populations (Albert, et al. 1991 and Bryan, et al., 2013).

All of the negative consequences of hunting leads to the following research question: How does human caused mortality affect wolf populations on the physiological level? Only one study has evaluated hormone levels in wolf populations to determine how human-caused mortality may impact group behavior, reproduction, and social dynamics. (Bryan, et al., 2015). Additional research is needed to accurately assess the effects of hunting on wolf physiology.

2. RESEARCH METHODOLOGY

The current research seeks to determine whether individual wolves can be classified as belonging to a heavily-stressed population due to hunting, or as a member of a population with lower hunting pressure. The criteria for measuring stress will be via the measurement of hormones and reproductive steroids in the wolf's fur. More specifically, this study evaluates the hormone levels of two separate wolf populations in Northern Canada that were originally studied by Bryan, et al. in 2015.

The distinction between these two wolf populations is marked by differences in the level of hunting and the percentage reduction of the population. Wolves in the tundra-taiga area were heavily hunted using snowmobiles and firearms. Taiga is characterized by dense conifers, like spruce and pine. Conversely, tundra regions lack any tree cover. Wolves in the second area (i.e., boreal forests) had a lower level of mortality and were killed predominately by trapping. Boreal forests consist of deciduous and conifer trees, and experience wide-ranging temperatures from lows in winter to highs in summer (Musiani, M. & Paquet, P.C., 2004).

Bryan, et al., (2015), predicted that there would be elevated levels of stress and signs of increased reproduction activity in the heavily hunted tundra-taiga wolves, as evidenced by high rates of hormone production (testosterone, progesterone, and cortisol). The researchers in the 2015 study compared the tundra-taiga wolves to wolves in areas of lower hunting pressure, such as those in the boreal forest (Packard & Mech 1980, 1983; Packard, Mech & Seal 1983; Haber 1996, Bryan, et al., 2015).

Sampling Method
The samples (n=152) were collected in a prior study in Nunavut, Northwest Territories and Alberta, Canada (Musiani, et al., 2007). The samples (See Appendix, populations 1 and 2) consisted of wolf hair samples collected during the winter months. The process of extracting the hormones from the wolf hair, including quality control methodologies, is outlined in the Bryan, et al. study (2015).

Bryan, et al., (2015) used predominantly statistical analysis methods in attempting to differentiate the tundra-taiga wolves from the boreal forest wolves. The researchers used ANOVA and Welch's t-tests to compare the two wolf groups, concluding that wolves from the more heavily hunted populations had increased levels of reproduction and stress related hormones. They also determined that these physiological characteristics are in response to environmental factors, including human-induced mortality (Bryan, et al., 2015). The researchers did list confounding factors, such as ecological and genetic-based differences that could explain hormonal discrepancies. Also, the higher levels of cortisol in the tundra-taiga wolves could be attributed to extended low levels in the food supply in summer, when wolves must travel farther to catch up with migrating caribou. Finally, the massing of tundra-taiga wolf populations near caribou in summer may cause a mingling of wolves and the inevitable interactions among members of different groups (which could also explain the elevated levels of testosterone). The boreal wolves, conversely, have more traditional territories and stability, leading to fewer intergroup interactions (Walton, et al., 2001, Musiani, et al., 2007).

In order to mitigate the impact of confounding factors, the researchers used a control group of wolves (n=30) from a heavily-hunted population in a boreal forest region (See Appendix, population 3). The hormone samples in the control group showed higher levels of cortisol than in boreal forest populations. The wolves in the control group also had similar levels of cortisol as wolves in the heavily hunted northern tundra-taiga region. Therefore, the study concluded that higher cortisol levels are the result of increased mortality rates, possibly coupled with some habitat related factors (Bryan, et al., 2015).

There are several implications revealed by the differences in hormone levels in the Bryan, et al. study. First, reproduction rates are altered (and the social structure, along with the reproduction rates) when there is no longer a dominant pair (i.e., pack hierarchy), and other pack members are not prevented from breeding. The stability of the social group, characterized by a single litter per pack each year, is threatened (Haber, 1996). Second, physiological effects of the disruption in the social framework, like increased cortisol levels, can enhance wolf musculature and release stored energy (Saplosky, 1993). Lastly, high levels of testosterone aid in any challenges an individual wolf may have within the social structure, where strength and dominance of the situation are necessary (Wingfield, et al., 2001).

The current research centers upon the following research questions: Are human-exploited wolf populations more heavily impacted physiologically? Are hormone levels affected to a larger extent in exploited wolf populations, as opposed to those in less stressed populations? And finally, can the type of population an individual wolf may inhabit be identified based upon the measurement of hormone levels as indicators of stress?

Hypotheses Tested

The research hypotheses to be tested in this study are as follows:

H1: Individual wolves can be classified as belonging to a heavily exploited population based on hormone levels.

H2: Machine learning classification can be used to support the results obtained by Bryan, et al. (2015) that human-caused mortality may impact group behavior, reproduction, and social dynamics, and populations as determined by the hormone levels in affected wolves. That is, wolves can accurately be classified into one of two groups: those with high levels of hunting-induced stress, and those with less stress.

The objective of the current study is to determine whether the physiological consequences of hunting (as determined by levels of stress and reproductive hormones in hair, an indicator of elevated endocrine activity), can be used to classify wolves as belonging to a highly-stressed group or a less-stressed group.

To test these hypotheses the current study used data previously analyzed by Bryan, et al. (2015) and k-Nearest Neighbor as the classification methodology to determine wolf membership in heavily stressed versus low stressed populations, based on hormone levels. The 2015 dataset included subject wolves from two separate areas and environments. The dataset contained 45 wolves from a lightly-hunted group in a northern boreal forest, and 103 wolves from a heavily-hunted Tundra-taiga forest area.

All samples were taken as part of a prior study (Musiani, et al., 2007). The samples consisted of hair from the wolf subjects. Cortisol, testosterone, and progesterone (females) levels were measured in each hair sample. The data, listing area, gender, and levels of the three hormones can be found in the Appendix.

Machine Learning Algorithm Used

k-Nearest Neighbor (k-NN) was used to compare cortisol and testosterone levels in the different populations and to determine the accuracy in predicting each population, based on its hormone levels. Bryan et al., (2015) determined that higher levels of cortisol and testosterone were found in the tundra-taiga wolves and concluded that this higher level may be an indicator of social instability. The current study also used k-NN to compare progesterone levels in the female wolves in the two populations.

Due to the lower numbers of northern boreal forest wolves, stratified sampling was used. In addition, the data were partitioned into 70% for training and 30% for testing. A k-NN algorithm was applied to the data using the knn() function from the class package in R and RStudio. The confusionMatrix() function from the caret package was used to determine accuracy of the classification and sensitivity, specificity, Kappa, and the No Information Rate.

3. RESULTS AND DISCUSSION

The highest accuracy in predicting group membership of the wolves was 86.96% with k=3 (as shown in Table 1). The true positive rate was 100% and the false positive rate was 83%, which supports the validity of the model. The No Information Rate is higher than desired, at an elevated 78%. The Kappa statistic (a measurement of the agreement between accuracy and random chance) was 68%, which indicates moderate agreement.

Table 1. Results of Classification of Wolf Subjects based on Cortisol and Testosterone Levels (k=3)

Measurement	Value
Accuracy	.8697
Sensitivity	1.00
Specificity	0.833
Kappa	0.6849
No Information Rate	0.7826

The current study also measured the difference in progesterone levels between the female wolves in the taiga-tundra and the northern boreal forest and classified them using k-NN (Table 2). Along with k-NN, a similar sampling and data partitioning method was used to preprocess the data. After preprocessing, it was determined that k=14 had the highest accuracy in predicting classification at 0.8333, and with a sensitivity and specificity at 0.7143 and 0.8824, respectively. The No Information Rate was 0.7083, indicating the model has some validity in classification. The Kappa was 0.5966 showing, on the low end, moderate agreement between random and model accuracy.

Table 2. Results of Classification of Female Wolves Based on Progesterone Levels (k=14)

Measurement	Value
Accuracy	.8333
Sensitivity	0.7143
Specificity	0.8824
Kappa	0.5966
No Information Rate	0.7083

4. CONCLUSIONS

Past research on this topic has proposed that elevated levels of the hormones cortisol, testosterone, and progesterone in taiga-tundra wolves are explained by the synergistic effects of hunting pressures, the habitat, or sampling (Bryan, et al., 2015). In the Bryan, et al., study, the researchers compared cortisol levels in the taiga-tundra wolves to those of a control group of 30 wolf subjects (i.e., Little Smokey wolves) in a heavily hunted boreal forest area in an effort to explain the differences in habitat and ecosystem characteristics. The results of this study showed statistically higher cortisol levels in both the Little Smokey and taiga-tundra wolves, compared to the northern boreal forest wolves.

The current study used the k-NN classification algorithm to show that individual wolves can be classified as belonging to heavily hunting-pressured groups based on cortisol and testosterone levels. This classification was also shown to be at a highly-accurate level. The current study also concluded that classification of female wolves (using the k-NN classifier) is possible with a favorable accuracy, based on the females' levels of progesterone. Our results support the findings of Bryan, et al., (2015) that showed statistically-significant differences in hormone levels between taiga-tundra and boreal forest wolf populations (i.e., heavily hunted vs. lightly trapped populations). Our findings support our hypothesis that individual wolves can be classified as belonging to a heavily exploited population based on hormone levels. Additionally, k-nn, a machine learning methodology can be used as a classification mechanism for this purpose.

Prior studies have concluded that the potential ramifications of heavy human-caused mortality in wolves are substantive chronic stress, and negative alterations in reproduction and breeding practices. These negative effects on breeding, compared with non-distressed populations are not known. However, predictable genetic outcomes like in-breeding, lack of diversity, increased disease, as well as an elevated danger of population extinction are potential long-term effects of heavy hunting (Leonard, et al., 2005).

If a link exists between stress levels in wolf populations and human-based hunting, then aside from the impact on wolf populations, the effects on entire ecosystems can be influenced. Wolves are recognized as a keystone species in their natural habitat (Boyce, 2018; Ripple and Beschta, 2012). Therefore, their absence or minimization can have far reaching impacts on entire ecosystems.

Limitations of Study

In this study, we did not account for differences in male and female subjects in the analysis of cortisol and testosterone levels. This difference in levels between wolf sexes can be evaluated in a later study. It should also be noted that the sample size in this study was relatively small, particularly with the northern boreal forest wolves (i.e., n = 45). However, the research was unfortunately limited by the amount of available data. Additionally, only one machine learning algorithm for classification (i.e., the k-NN classifier). Various machine-learning techniques and models could be employed in future studies. These additional techniques could be used to determine whether wolves can be more accurately classified based hormone levels as indicators of human-caused stress.

5. REFERENCES

Albert, D., Jonik, R. & Walsh, M. (1992) Hormone-dependent aggression in male and female rats: experiential, hormonal, and neural foundations. Neuroscience & Biobehavioral Reviews, 16, 177–192.

Bateson, P. & Bradshaw, E.L. (2007) Physiological effects of hunting red deer (Cervus elaphus). Proceedings of the Royal Society of London. Series B: Biological Sciences, 264, 1707–1714.

Boyce, Mark S. (2018) Wolves for Yellowstone: dynamics in time and space, Journal of Mammalogy, 99, 5, 1021–1031.

Bryan, Heather M. et al. (2015), Heavily hunted wolves have higher stress and reproductive steroids than wolves with lower hunting pressure, Functional Ecology, Article-journal, https://doi.org/10.1111/1365-2435.12354

Bryan, H.M., Darimont, C.T., Paquet, P.C., Wynne-Edwards, K.E. & Smits, J.E. (2013b) Stress and reproductive hormones in grizzly bears reflect nutritional benefits and social consequences of a salmon foraging niche. PLoS ONE, 8, e80537.

Coltman, D.W., O'Donoghue, P., Jorgenson, J.T., Hogg, J.T., Strobeck, C. & Festa-Bianchet, M. (2003) Undesirable evolutionary consequences of trophy hunting. Nature, 426, 655–658.

Coltman, D. W., Festa-Bianchet, M., Jorgenson, J. T. & Strobeck, C. (2002) Age-dependent sexual selection in bighorn rams. Proc. R. Soc. Lond. B 269, 165–172.

Darimont, C.T., Carlson, S.M., Kinnison, M.T., Paquet, P.C., Reimchen, T.E. & Wilmers, C.C. (2009) Human predators outpace other agents of trait change in the wild. Proceedings of the National Academy of Sciences, 106, 952–954.

Festa-Bianchet, M., Coltman, D. W., Turelli, L. & Jorgenson, J. T. Relative allocation to horn and body growth in bighorn rams varies with resource availability. Behav. Ecol. (in the press)

Foley, C.A.H., Papageorge, S. & Wasser, S.K. (2001) Noninvasive stress and reproductive measures of social and ecological pressures in free ranging African elephants. Conservation Biology. 15.1134-1142.

Gobush, K.S., Mutayoba, B.M. & Wasser, S.K. (2008) Long term impacts of poaching on relatedness, stress physiology, and reproductive output of adult female African elephants. Conservation Biology, 22, 1590–1599.

Haber, G.C. (1996) Biological, conservation, and ethical implications of exploiting and controlling wolves. Conservation Biology, 10, 1068–1081.

Jezdrzejewski, W., Branicki, W., Veit, C., Me€ Augorac, I., Pilot, M.g., Bunevich, A. et al. (2005) Genetic diversity and relatedness within packs in an intensely hunted population of wolves Canis lupus. Acta Theriologica, 50, 3–22.

Leonard, J.A., Vilà, C. & Wayne, R.K. (2005) Legacy lost: genetic variability and population size of extirpated US grey wolves (Canis lupus). Molecular Ecology, 14, 9– 17.

Musiani, M. & Paquet, P.C. (2004) The practices of wolf persecution, protection, and restoration in Canada and the United States. BioScience, 54, 50–60.

Musiani, M., Leonard, J.A., Cluff, H.D., Gates, C.C., Mariani, S., Paquet, P.C. et al. (2007) Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat colour and association with migratory caribou. Molecular Ecology, 16, 4149–4170.

Oliveira, R.F. (2004) Social modulation of androgens in vertebrates: mechanisms and function. Advances in the Study of Behavior, 34, 165–239.

Packard, J.M. & Mech, L.D. (1980) Population regulation in wolves. Biosocial Mechanisms of Population Regulation (eds M.N. Cohen, R.S. Malpass & H.G. Klein), pp. 135–150. Yale University Press, New Haven, Connecticut, USA.

Packard, J.M. & Mech, L.D. (1983) Population regulation in wolves. Symposiumon Natural Regulation of Wildlife Populations. Proceedings 14 (eds F.L. Bunnell, D.S. Eastman & J.M. Peek), pp. 151–174. University of Idaho, Moscow, Idaho, USA.

Packard, J.M., Mech, L.D. & Seal, U.S. (1983) Social influences on reproduction in wolves. Wolves of Canada and Alaska (ed. L.N. Carbyn), pp. 78–85. Canadian Wildlife Service Report Series 45, Ottawa, Ontario, Canada.

R Development Core Team (2011) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rausch, R.A. (1967) Some aspects of the population ecology of wolves, Alaska. American Zoologist, 7, 253–265.

Ripple, William & Beschta, Robert. (2012). Trophic cascades in Yellowstone: The first 15 years after wolf reintroduction. Biological Conservation. 145. 205–213.

Russell, E., Koren, G., Rieder, M. & Van Uum, S.H. (2014) The detection of cortisol in human sweat: implications for measurement of cortisol in hair. Therapeutic Drug Monitoring, 36, 30–34.

Rutledge, L.Y., Patterson, B.R., Mills, K.J., Loveless, K.M., Murray, D.L. & White, B.N. (2010) Protection from harvesting restores the natural social structure of eastern wolf packs. Biological Conservation, 143, 332–339.

Rutledge, L.Y., White, B.N., Row, J.R. & Patterson, B.R. (2012) Intense harvesting of eastern wolves facilitated hybridization with coyotes. Ecology and Evolution, 2, 19–33.

Sapolsky, R.M. (1993) The physiology of dominance in stable versus unstable social hierarchies. *Primate Social Conflict* (eds W.A. Mason & S.P. Mendoza), pp. 171– 204. State University of New York Press, Albany, New York, USA.

Sidorovich, V., Stolyarov, V., Vorobei, N., Ivanova, N. & Jedrzejewska, B. (2007) Litter size, sex ratio, and age structure of gray wolves, Canis lupus, in relation to population fluctuations in northern Belarus. Canadian Journal of Zoology, 85, 295–300.

Walton, L.R., Cluff, H.D., Paquet, P.C. & Ramsay, M.A. (2001) Movement patterns of barren-ground wolves in the central Canadian Arctic. Journal of Mammalogy, 82, 867–8

Wingfield, J.C., Lynn, S. & Soma, K.K. (2001) Avoiding the ‘costs’ of testosterone: ecological bases of hormone-behavior interactions. *Brain Behavior and Evolution*, **57**, 239– 251.

APPENDICES

Appendix A: Wolf Hair Sample Data Collected during Musiani, et al. Study (2007)

Individual	Sex	Population	Colour	Cpgmg	Tpgmg	Ppgmg
1	M	2	W	15.86	5.32	NA
2	F	1	D	20.02	3.71	14.37622
3	F	2	W	9.95	5.3	21.65902
4	F	1	D	25.22	3.71	13.42507
5	M	2	D	21.13	5.34	NA
6	M	2	W	12.48	4.6	NA
7	M	1	W	26.78	4.58	NA
8	M	1	D	15.41	9.27	NA
9	F	1	D	33.87	4.81	19.9127
10	F	2	W	17.29	5.07	34.59806
11	F	1	W	9.43	4.47	25.88548
12	F	1	W	8.84	3.75	15.86882
13	F	1	D	34	4.76	33.08362
14	F	1	D	14.3	6.06	24.82876
15	M	1	D	12.16	5.75	NA
16	M	1	D	22.43	6.15	NA
17	F	2	W	26.26	4.93	25.00037
18	M	2	W	15.8	5.24	NA
19	M	1	W	7.93	4.14	NA
20	M	1	D	4.75	3.34	NA
21	M	2	W	9.17	4.02	NA
22	M	2	W	21.52	4.91	NA
23	M	1	W	10.79	3.91	NA
24	F	2	W	22.69	6.47	21.50033
25	F	2	W	22.17	4.28	31.8274
26	F	2	W	15.34	5.53	34.0765
27	F	1	W	20.48	5.06	20.21606
28	F	1	W	16.19	4.79	18.29115
29	F	1	W	24.05	3.7	21.29735
30	M	2	W	16.45	6.09	NA
31	F	2	W	21.91	4.19	36.40797
32	F	2	W	32.24	6.94	40.92793
33	F	2	W	23.99	5.97	45.9136
34	F	2	W	27.82	7.76	47.2674
35	F	2	W	19.83	6.55	40.93838
36	F	2	W	12.16	4.34	26.65583
37	F	2	W	19.05	6.34	23.90413
38	F	2	D	13.91	4.72	26.36326
39	F	2	D	17.16	9.25	34.64966
40	F	1	W	30.16	6.8	19.61885
41	F	2	W	24.38	5.49	28.12497

42	F	2	D	10.14	3.81	NA
43	M	2	W	18.4	4.98	NA
44	M	2	W	15.21	7.17	NA
45	M	2	W	24.64	15.13	NA
46	M	2	W	22.49	14.45	NA
47	M	2	W	17.42	5.36	NA
48	M	2	W	29.51	9.12	NA
49	M	2	W	27.3	10.75	NA
50	M	2	W	14.04	7.19	NA
51	M	2	W	11.77	5.17	NA
52	M	2	W	23.6	6.97	NA
53	M	2	W	18.14	5.7	NA
54	M	2	W	11.25	4.4	NA
55	F	1	W	14.82	10.81	NA
56	F	2	W	26.39	6.47	24.46521
57	M	2	W	15.15	4.52	NA
58	M	2	W	14.04	6.01	NA
59	M	2	W	21.39	7.36	NA
60	F	2	W	20.02	5.19	31.40929
61	M	2	W	24.64	14.08	NA
62	M	2	W	13.46	4.09	NA
63	M	2	W	18.79	9.74	NA
64	F	2	W	11.77	4.95	21.01472
65	F	2	W	19.96	7.62	28.06955
66	F	2	W	12.68	3.82	27.90797
67	F	2	W	19.76	5.26	27.37918
68	M	2	D	20.35	14.98	NA
69	F	2	W	17.68	5.97	53.28191
70	F	2	W	23.66	6.13	48.53432
71	F	2	W	17.23	7.24	NA
72	F	2	W	25.74	4.88	37.65696
73	F	2	W	19.89	6.35	31.90467
74	F	1	D	14.24	3.95	28.87637
75	M	2	W	17.55	5.02	NA
76	M	2	W	16.32	5.86	NA
77	M	2	W	15.34	5.78	NA
78	F	2	W	11.64	4.87	22.87393
79	M	2	W	13.65	5.04	NA
80	M	2	W	11.57	5.24	NA
81	M	2	W	20.35	5.98	NA
82	M	2	W	8.91	4.58	NA
83	M	2	W	9.1	4.4	NA
84	M	2	D	21.65	7.81	NA
85	M	1	D	10.6	3.65	NA
86	M	1	D	12.35	9.57	NA
87	F	1	D	7.93	3.83	16.77475

88	F	1	D	8	4.26	19.49892
89	F	1	D	7.61	4.24	22.56011
90	M	1	W	11.96	5.62	NA
91	M	1	D	14.82	5.35	NA
92	F	1	W	14.43	5.08	34.81566
93	F	1	D	19.57	6.81	16.67624
94	F	1	W	12.55	3.25	13.19328
95	F	1	D	12.61	3.54	13.62372
96	F	1	D	10.21	4.49	18.52082
97	M	1	D	15.99	5.82	NA
98	F	1	D	32.24	4.8	25.20981
99	M	1	D	15.41	5.68	NA
100	M	1	D	13.98	5.45	NA
101	M	1	D	16.32	6.65	NA
102	M	1	D	6.37	3.31	NA
103	M	1	W	8.19	3.81	NA
104	M	1	W	12.29	3.95	NA
105	F	2	W	12.16	4.37	13.17322
106	F	2	W	16.19	4.43	26.32807
107	F	2	W	11.83	3.48	16.40101
108	F	2	W	10.47	3.9	17.56024
109	F	2	W	21.13	5.09	29.29508
110	F	2	W	18.59	4.49	21.51784
111	F	2	W	12.09	3.96	28.49073
112	F	2	W	13	3.83	30.98607
113	F	2	W	12.09	4.65	28.62749
114	F	2	W	13.26	4.48	25.66584
115	F	2	W	12.03	4.32	19.28812
116	F	2	W	17.36	5.01	30.00925
117	F	2	W	18.14	3.56	12.7591
118	F	2	W	15.93	4.65	22.72246
119	F	2	W	12.29	5.01	23.24402
120	F	2	W	17.42	4.38	18.35924
121	F	2	W	13.2	5.3	18.88097
122	F	2	W	14.5	5.01	21.06504
123	F	2	D	11.44	4.04	16.154
124	M	2	D	11.57	5.68	NA
125	M	2	W	15.28	3.9	NA
126	M	2	W	13.46	5.1	NA
127	M	2	W	13.2	4.76	NA
128	M	2	W	11.25	4.89	NA
129	M	2	W	16.58	7.54	NA
130	M	2	W	13.2	5.07	NA
131	M	2	W	14.04	5.65	NA
132	M	2	W	17.03	5.81	NA
133	M	2	W	17.81	4.88	NA

134	M	2	W	12.48	4.86	NA
135	M	2	W	11.44	4.34	NA
136	M	2	W	40.43	9.13	NA
137	M	2	D	14.3	4.53	NA
138	M	2	W	14.89	4.32	NA
139	M	2	W	16.77	4.4	NA
140	M	2	D	9.95	4.31	NA
141	M	2	W	10.34	4.36	NA
142	M	2	W	20.54	8.06	NA
143	F	1	W	12.81	6.25	26.73429
144	F	1	W	16.51	4.62	28.10653
145	M	1	D	11.12	6.71	NA
146	M	1	D	11.64	4.51	NA
147	M	1	W	18.92	7.57	NA
148	M	2	W	19.89	5.35	NA
149	U	3		9.69	4.23	NA
150	U	3		19.37	4.26	NA
151	U	3		19.76	4.56	NA
152	U	3		11.31	7.73	NA
153	U	3		11.25	3.81	NA
154	U	3		13.85	4.28	NA
155	U	3		17.62	4.54	NA
156	U	3		22.82	4.34	NA
157	U	3		18.14	10.33	NA
158	U	3		13.52	8.12	NA
159	U	3		21.58	5.79	NA
160	U	3		8.91	29.74	NA
161	U	3		9.17	3.14	NA
162	U	3		14.17	10.32	NA
163	U	3		12.09	6.7	NA
164	U	3		54.47	61.79	NA
165	U	3		10.4	4.2	NA
166	U	3		50.31	5.48	NA
167	U	3		33.74	9.61	NA
168	U	3		14.76	8.94	NA
169	U	3		22.3	6.16	NA
170	U	3		23.21	10.59	NA
171	U	3		19.24	5.66	NA
172	U	3		13.07	4.4	NA
173	U	3		49.14	6.21	NA
174	U	3		73.19	6.41	NA
175	U	3		37.05	4.75	NA
176	U	3		16.45	7.29	NA
177	U	3		43.81	6.09	NA
178	U	3		14.89	3.53	NA

Appendix B: k-NN Algorithm and Resulting Confusion Matrix Coded in R

```
set.seed(123)
index <- initial_split(WolfData, prop = 0.7, strata =
"Population")

# index <- sample(2, nrow(WolfData), replace=TRUE,
prob=c(0.90, 0.10))

index

trainData <- WolfData[index==1,]

testData <- WolfData[index==2,]

trainData1 <- trainData[-1]
testData1 <- testData[-1]

trainDataLabels <- trainData [, 1]
testDataLabels <- testData [ ,1]


install.packages ("class")    # if necessary
library(class)

set.seed(13876)
WolfDataPred<- knn(train = trainData1, test = testData1, cl =
trainDataLabels, k=3)

##  Evaluating model performance ----

# load the "gmodels" library
install.packages('gmodels')
library(gmodels)

# Create the cross tabulation of predicted vs. actual

CrossTable(x = testDataLabels, y = WolfDataPred,
         prop.chisq=FALSE)

dim(testDataLabels)
dim(WolfDataPred)

install.packages('caret')

#Import required library
library(caret)

confusionMatrix(testDataLabels,WolfDataPred)


i=1                    # declaration to initiate for loop
k.optm=1                # declaration to initiate for loop
for (i in 1:28){
   knn.mod <-  knn(train=trainData1, test=testData1,
cl=trainDataLabels, k=i)
   k.optm[i] <- 100 * sum(testDataLabels ==
knn.mod)/NROW(testDataLabels)
   k=i
   cat(k,'=',k.optm[i],'\n')      # to print % accuracy
}
```