# Midterm 1 W24

**Maya Ziebarth**

**2024-02-06**

# Instructions

Answer the following questions and complete the exercises in RMarkdown. Please embed all of your code and push your final work to your repository. Your code must be organized, clean, and run free from errors. Remember, you must remove the `#` for any included code chunks to run. Be sure to add your name to the author header above.

Your code must knit in order to be considered. If you are stuck and cannot answer a question, then comment out your code and knit the document. You may use your notes, labs, and homework to help you complete this exam. Do not use any other resources- including AI assistance.

Don't forget to answer any questions that are asked in the prompt!

Be sure to push your completed midterm to your repository. This exam is worth 30 points.

# Background

In the data folder, you will find data related to a study on wolf mortality collected by the National Park Service. You should start by reading the `README_NPSwolfdata.pdf` file. This will provide an abstract of the study and an explanation of variables.

The data are from: Cassidy, Kira et al. (2022). Gray wolf packs and human-caused wolf mortality. Dryad (https://doi.org/10.5061/dryad.mkkwh713f).

# Load the libraries

```
library("tidyverse")
library("janitor")
```

# Load the wolves data

In these data, the authors used `NULL` to represent missing values. I am correcting this for you below and using `janitor` to clean the column names.

```
wolves <- read.csv("data/NPS_wolfmortalitydata.csv", na = c("NULL")) %>% clean_name
s()
```

# Questions

Problem 1. (1 point) Let's start with some data exploration. What are the variable (column) names?

```
names(wolves)
```

```
##  [1] "park"         "biolyr"       "pack"         "packcode"     "packsize_aug"
##  [6] "mort_yn"      "mort_all"     "mort_lead"    "mort_nonlead" "reprody1"
## [11] "persisty1"
```

Problem 2. (1 point) Use the function of your choice to summarize the data and get an idea of its structure.

```
glimpse(wolves)
```

```
## Rows: 864
## Columns: 11
## $ park         <chr> "DENA", "DENA", "DENA", "DENA", "DENA", "DENA", "DENA", "…
## $ biolyr       <int> 1996, 1991, 2017, 1996, 1992, 1994, 2007, 2007, 1995, 200…
## $ pack         <chr> "McKinley River1", "Birch Creek N", "Eagle Gorge", "East …
## $ packcode     <int> 89, 58, 71, 72, 74, 77, 101, 108, 109, 53, 63, 66, 70, 72…
## $ packsize_aug <dbl> 12, 5, 8, 13, 7, 6, 10, NA, 9, 8, 7, 11, 0, 19, 15, 12, 1…
## $ mort_yn      <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, …
## $ mort_all     <int> 4, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, …
## $ mort_lead    <int> 2, 2, 0, 0, 0, 0, 1, 2, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, …
## $ mort_nonlead <int> 2, 0, 2, 2, 2, 2, 1, 0, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, …
## $ reprody1     <int> 0, 0, NA, 1, NA, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1…
## $ persisty1    <int> 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, …
```

```
str(wolves)
```

```
## 'data.frame':    864 obs. of  11 variables:
##  $ park        : chr  "DENA" "DENA" "DENA" "DENA" ...
##  $ biolyr      : int  1996 1991 2017 1996 1992 1994 2007 2007 1995 2003 ...
##  $ pack        : chr  "McKinley River1" "Birch Creek N" "Eagle Gorge" "East Fork"
...
##  $ packcode    : int  89 58 71 72 74 77 101 108 109 53 ...
##  $ packsize_aug: num  12 5 8 13 7 6 10 NA 9 8 ...
##  $ mort_yn     : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ mort_all    : int  4 2 2 2 2 2 2 2 2 1 ...
##  $ mort_lead   : int  2 2 0 0 0 0 1 2 1 1 ...
##  $ mort_nonlead: int  2 0 2 2 2 2 1 0 1 0 ...
##  $ reprody1    : int  0 0 NA 1 NA 0 0 1 0 1 ...
##  $ persisty1   : int  0 0 1 1 1 1 0 1 0 1 ...
```

Problem 3. (3 points) Which parks/ reserves are represented in the data? Don't just use the abstract, pull this information from the data.

```
wolves %>%
  group_by(park) %>%
  summarize(n_distinct=n())
```

```
## # A tibble: 5 × 2
##   park  n_distinct
##   <chr>      <int>
## 1 DENA         340
## 2 GNTP          77
## 3 VNP           48
## 4 YNP          248
## 5 YUCH         151
```

Problem 4. (4 points) Which park has the largest number of wolf packs?

```
wolves %>%
  select(park, packsize_aug) %>%
  group_by(park) %>%
  summarize(pack_total=sum(packsize_aug, na.rm=T)) %>%
  arrange(desc(pack_total))
```

```
## # A tibble: 5 × 2
##   park  pack_total
##   <chr>      <dbl>
## 1 YNP         2731
## 2 DENA        2500
## 3 YUCH        1048
## 4 GNTP         781.
## 5 VNP           50
```

Problem 5. (4 points) Which park has the highest total number of human-caused mortalities `mort_all` ?

```
wolves %>%
  select(park, mort_all) %>%
  group_by(park) %>%
  summarize(mort_all_total=sum(mort_all, na.rm=T)) %>%
  arrange(desc(mort_all_total))
```

```
## # A tibble: 5 × 2
##   park  mort_all_total
##   <chr>          <int>
## 1 YUCH             136
## 2 YNP               72
## 3 DENA              64
## 4 GNTP              38
## 5 VNP               11
```

The wolves in Yellowstone National Park (https://www.nps.gov/yell/learn/nature/wolf-restoration.htm) are an incredible conservation success story. Let's focus our attention on this park.

Problem 6. (2 points) Create a new object "ynp" that only includes the data from Yellowstone National Park.

```
ynp <- filter(wolves, park=="YNP")
```

Problem 7. (3 points) Among the Yellowstone wolf packs, the Druid Peak Pack (https://www.pbs.org/wnet/nature/in-the-valley-of-the-wolves-the-druid-wolf-pack-story/209/) is one of most famous. What was the average pack size of this pack for the years represented in the data?

```
ynp %>%
  filter(pack=="druid") %>%
  group_by(biolyr) %>%
  summarize(mean_pack_druid=mean(packsize_aug)) %>%
  arrange(desc(mean_pack_druid))
```

```
## # A tibble: 15 × 2
##    biolyr mean_pack_druid
##     <int>           <dbl>
##  1   2001              37
##  2   2000              27
##  3   2008              21
##  4   2003              18
##  5   2007              18
##  6   2002              16
##  7   2006              15
##  8   2004              13
##  9   2009              12
## 10   1999               9
## 11   1998               8
## 12   1996               5
## 13   1997               5
## 14   2005               5
## 15   2010               0
```

Problem 8. (4 points) Pack dynamics can be hard to predict- even for strong packs like the Druid Peak pack. At which year did the Druid Peak pack have the largest pack size? What do you think happened in 2010?

```
ynp %>%
  filter(pack=="druid") %>%
  group_by(biolyr) %>%
  summarize(largest_pack_druid=max(packsize_aug)) %>%
  arrange(desc(largest_pack_druid))
```

```
## # A tibble: 15 × 2
##    biolyr largest_pack_druid
##     <int>              <dbl>
##  1   2001                 37
##  2   2000                 27
##  3   2008                 21
##  4   2003                 18
##  5   2007                 18
##  6   2002                 16
##  7   2006                 15
##  8   2004                 13
##  9   2009                 12
## 10   1999                  9
## 11   1998                  8
## 12   1996                  5
## 13   1997                  5
## 14   2005                  5
## 15   2010                  0
```

Problem 9. (5 points) Among the YNP wolf packs, which one has had the highest overall persistence `persisty1` for the years represented in the data? Look this pack up online and tell me what is unique about its behavior- specifically, what prey animals does this pack specialize on?

```
ynp %>%
  mutate_if(is.character, factor) %>%
  count(pack, persisty1, sort = T)
```

```
##               pack persisty1  n
## 1          mollies         1 26
## 2           cougar         1 20
## 3        yelldelta         1 18
## 4            druid         1 13
## 5          leopold         1 12
## 6            agate         1 10
## 7            8mile         1  9
## 8           canyon         1  9
```

```
## 9      gibbon/mary         1  9
## 10        nezperce         1  9
## 11        junction         1  8
## 12           lamar         1  8
## 13            rose         1  8
## 14            swan         1  6
## 15          wapiti         1  6
## 16       blacktail         1  5
## 17          slough         1  5
## 18      geode/hell         1  4
## 19        quadrant         1  4
## 20          hayden         1  3
## 21         phantom         1  3
## 22        prospect         1  3
## 23         bechler         1  2
## 24           heart         1  2
## 25            lava         1  2
## 26           oxbow         1  2
## 27           snake         1  2
## 28      1118Fgroup         0  1
## 29      1155Mgroup         0  1
## 30       642Fgroup         0  1
## 31       682Mgroup         0  1
## 32       682Mgroup         1  1
## 33       694Fgroup         0  1
## 34       755Mgroup         0  1
## 35       963Fgroup         0  1
## 36           agate         0  1
## 37         biscuit         1  1
## 38     buffalofork         0  1
## 39     buffalofork         1  1
## 40          canyon         0  1
## 41       carnelian         0  1
## 42        cinnabar         1  1
## 43      cottonwood         0  1
## 44      cottonwood         1  1
## 45         crevice         1  1
## 46           druid         0  1
## 47           druid        NA  1
## 48          everts         0  1
## 49          everts         1  1
## 50     gibbon/mary         0  1
## 51        grayling         0  1
## 52        grayling         1  1
## 53          hayden         0  1
## 54            lava         0  1
## 55         leopold         0  1
```

```
## 56        lonestar         0  1
## 57        nezperce         0  1
## 58        nezperce        NA  1
## 59           oxbow         0  1
## 60        prospect         0  1
## 61 specimen/silver         0  1
## 62 specimen/silver         1  1
## 63       thorofare         0  1
## 64       thorofare         1  1
## 65       thorofare        NA  1
## 66           tower         0  1
## 67           tower         1  1
```

## The Mollie wolf pack hunts bison, which provide a lot of food for them and other scavengers. They regularly interact with bears because of their bison hunting which no other wolf pack does.

Problem 10. (3 points) Perform one analysis or exploration of your choice on the `wolves` data. Your answer needs to include at least two lines of code and not be a summary function.

```
wolves %>%
  select(park, biolyr, pack, mort_all) %>%
  filter(park=="DENA" & biolyr>=2000) %>%
  filter(pack!="East Fork") %>%
  arrange(desc(mort_all))
```

```
##     park biolyr             pack mort_all
## 1   DENA   2017      Eagle Gorge        2
## 2   DENA   2007            Pinto        2
## 3   DENA   2007           Savage        2
## 4   DENA   2003         100 Mile        1
## 5   DENA   2003    Castle Rocks2        1
## 6   DENA   2007          Chitsia        1
## 7   DENA   2003     Death Valley        1
## 8   DENA   2007      Grant Creek        1
## 9   DENA   2012      Grant Creek        1
## 10  DENA   2013      Grant Creek        1
## 11  DENA   2008       Hot Slough        1
## 12  DENA   2017  Iron Creek West        1
## 13  DENA   2014      John Hansen        1
## 14  DENA   2006  Kantishna River        1
## 15  DENA   2007          McLeod2        1
## 16  DENA   2003      Mt Margaret        1
## 17  DENA   2008      Mt Margaret        1
## 18  DENA   2009      Mt Margaret        1
## 19  DENA   2003      Muddy River        1
```

```
## 20  DENA   2016          Myrtle        1
## 21  DENA   2013    Nenana River        1
## 22  DENA   2017     Riley Creek        1
## 23  DENA   2007          Somber        1
## 24  DENA   2006      Starr Lake        1
## 25  DENA   2003    Straightaway        1
## 26  DENA   2009         Tonzona        1
## 27  DENA   2007     Hauke Creek        1
## 28  DENA   2000       Sanctuary        1
## 29  DENA   2000        100 Mile        0
## 30  DENA   2001        100 Mile        0
## 31  DENA   2002        100 Mile        0
## 32  DENA   2004        100 Mile        0
## 33  DENA   2004         Bearpaw        0
## 34  DENA   2005         Bearpaw        0
## 35  DENA   2006         Bearpaw        0
## 36  DENA   2007         Bearpaw        0
## 37  DENA   2008         Bearpaw        0
## 38  DENA   2009         Bearpaw        0
## 39  DENA   2010         Bearpaw        0
## 40  DENA   2011         Bearpaw        0
## 41  DENA   2012         Bearpaw        0
## 42  DENA   2013         Bearpaw        0
## 43  DENA   2014         Bearpaw        0
## 44  DENA   2015         Bearpaw        0
## 45  DENA   2016         Bearpaw        0
## 46  DENA   2017         Bearpaw        0
## 47  DENA   2000     Birch Hills        0
## 48  DENA   2008       Boot Lake        0
## 49  DENA   2009       Boot Lake        0
## 50  DENA   2001         Brooker        0
## 51  DENA   2002         Brooker        0
## 52  DENA   2004   Castle Rocks2        0
## 53  DENA   2006   Castle Rocks3        0
## 54  DENA   2007   Castle Rocks3        0
## 55  DENA   2008   Castle Rocks3        0
## 56  DENA   2004         Chitsia        0
## 57  DENA   2005         Chitsia        0
## 58  DENA   2006         Chitsia        0
## 59  DENA   2008         Chitsia        0
## 60  DENA   2000    Death Valley        0
## 61  DENA   2001    Death Valley        0
## 62  DENA   2002    Death Valley        0
## 63  DENA   2016     Eagle Gorge        0
## 64  DENA   2000         Foraker        0
## 65  DENA   2001         Foraker        0
## 66  DENA   2002     Grant Creek        0
```

```
## 67  DENA   2003      Grant Creek        0
## 68  DENA   2004      Grant Creek        0
## 69  DENA   2005      Grant Creek        0
## 70  DENA   2006      Grant Creek        0
## 71  DENA   2008      Grant Creek        0
## 72  DENA   2009      Grant Creek        0
## 73  DENA   2010      Grant Creek        0
## 74  DENA   2011      Grant Creek        0
## 75  DENA   2014      Grant Creek        0
## 76  DENA   2015      Grant Creek        0
## 77  DENA   2016      Grant Creek        0
## 78  DENA   2017      Grant Creek        0
## 79  DENA   2008      Hauke Creek        0
## 80  DENA   2002           Herron        0
## 81  DENA   2003           Herron        0
## 82  DENA   2004           Herron        0
## 83  DENA   2007       Hot Slough        0
## 84  DENA   2009       Hot Slough        0
## 85  DENA   2010       Hot Slough        0
## 86  DENA   2011       Hot Slough        0
## 87  DENA   2012       Hot Slough        0
## 88  DENA   2013       Hot Slough        0
## 89  DENA   2014       Hot Slough        0
## 90  DENA   2015       Hot Slough        0
## 91  DENA   2010       Iron Creek        0
## 92  DENA   2011       Iron Creek        0
## 93  DENA   2013  Iron Creek East        0
## 94  DENA   2014  Iron Creek East        0
## 95  DENA   2013  Iron Creek West        0
## 96  DENA   2014  Iron Creek West        0
## 97  DENA   2015  Iron Creek West        0
## 98  DENA   2016  Iron Creek West        0
## 99  DENA   2013      John Hansen        0
## 100 DENA   2015      John Hansen        0
## 101 DENA   2016      John Hansen        0
## 102 DENA   2017      John Hansen        0
## 103 DENA   2000  Kantishna River        0
## 104 DENA   2001  Kantishna River        0
## 105 DENA   2002  Kantishna River        0
## 106 DENA   2003  Kantishna River        0
## 107 DENA   2004  Kantishna River        0
## 108 DENA   2005  Kantishna River        0
## 109 DENA   2007  Kantishna River        0
## 110 DENA   2008  Kantishna River        0
## 111 DENA   2009  Kantishna River        0
## 112 DENA   2010  Kantishna River        0
## 113 DENA   2004   McKinley River        0
```

```
## 114 DENA   2005   McKinley River        0
## 115 DENA   2006   McKinley River        0
## 116 DENA   2007   McKinley River        0
## 117 DENA   2000  McKinley Slough        0
## 118 DENA   2001  McKinley Slough        0
## 119 DENA   2002  McKinley Slough        0
## 120 DENA   2003  McKinley Slough        0
## 121 DENA   2004  McKinley Slough        0
## 122 DENA   2005  McKinley Slough        0
## 123 DENA   2006  McKinley Slough        0
## 124 DENA   2007  McKinley Slough        0
## 125 DENA   2009  McKinley Slough        0
## 126 DENA   2010  McKinley Slough        0
## 127 DENA   2011  McKinley Slough        0
## 128 DENA   2012  McKinley Slough        0
## 129 DENA   2013  McKinley Slough        0
## 130 DENA   2014  McKinley Slough        0
## 131 DENA   2015  McKinley Slough        0
## 132 DENA   2016  McKinley Slough        0
## 133 DENA   2009      Moose Creek        0
## 134 DENA   2010      Moose Creek        0
## 135 DENA   2000      Mt Margaret        0
## 136 DENA   2001      Mt Margaret        0
## 137 DENA   2002      Mt Margaret        0
## 138 DENA   2004      Mt Margaret        0
## 139 DENA   2005      Mt Margaret        0
## 140 DENA   2006      Mt Margaret        0
## 141 DENA   2007      Mt Margaret        0
## 142 DENA   2000      Muddy River        0
## 143 DENA   2001      Muddy River        0
## 144 DENA   2002      Muddy River        0
## 145 DENA   2015           Myrtle        0
## 146 DENA   2009     Nenana River        0
## 147 DENA   2010     Nenana River        0
## 148 DENA   2011     Nenana River        0
## 149 DENA   2012     Nenana River        0
## 150 DENA   2014     Nenana River        0
## 151 DENA   2000       North Fork        0
## 152 DENA   2001       North Fork        0
## 153 DENA   2009       Otter Lake        0
## 154 DENA   2006            Pinto        0
## 155 DENA   2000      Pinto Creek        0
## 156 DENA   2001      Pinto Creek        0
## 157 DENA   2002      Pinto Creek        0
## 158 DENA   2014      Riley Creek        0
## 159 DENA   2015      Riley Creek        0
## 160 DENA   2016      Riley Creek        0
```

```
## 161 DENA   2018 Riley Creek West        0
## 162 DENA   2008           Savage        0
## 163 DENA   2008           Somber        0
## 164 DENA   2009           Somber        0
## 165 DENA   2010           Somber        0
## 166 DENA   2012           Somber        0
## 167 DENA   2013           Somber        0
## 168 DENA   2014           Somber        0
## 169 DENA   2015           Somber        0
## 170 DENA   2000         Stampede        0
## 171 DENA   2001         Stampede        0
## 172 DENA   2001       Starr Lake        0
## 173 DENA   2002       Starr Lake        0
## 174 DENA   2003       Starr Lake        0
## 175 DENA   2004       Starr Lake        0
## 176 DENA   2005       Starr Lake        0
## 177 DENA   2007       Starr Lake        0
## 178 DENA   2008       Starr Lake        0
## 179 DENA   2009       Starr Lake        0
## 180 DENA   2010       Starr Lake        0
## 181 DENA   2000     Straightaway        0
## 182 DENA   2001     Straightaway        0
## 183 DENA   2002     Straightaway        0
## 184 DENA   2004     Straightaway        0
## 185 DENA   2005     Turtle Hill         0
## 186 DENA   2006     Turtle Hill         0
```