Multiple regression

[brouwern@gmail.com](mailto:brouwern@gmail.com)

December 5, 2016

Table of Contents

# Sample Paper Using Lion data

A an example of how to write and format your final project, using lion data from Whitlock & Shulter chapte 17, which is originally from Whitmas et al 2004 in Science.

# Preliminaries

#The following sets up the data fro the analysis  
#Set working directory  
  
setwd("C:/Users/lisanjie2/Desktop/TEACHING/1\_STATS\_CalU/1\_STAT\_CalU\_2016\_by\_NLB/Lecture/Unit3\_regression/last\_week")

Load data

dat <- read.csv("lion\_age\_by\_pop.csv")

# Content of a scientific paper

## Does the relationship between nose pigmentation and age vary between different populations of African lion (*Panthera leo*)?

## ABSTRACT

For information on the format of "Structured abstacts", see <https://www.nlm.nih.gov/bsd/policy/structured_abstracts.html>

**INTRODUCTION**: Being able to accurately understand the age structure of population

**OBJECTIVES**: The primary of objective of this study was to determine if the age of lions can be predicted from the amount of pigementation on their noses, and whether this relationship varies between populations.

**METHODS**:(3-4 sentneces on methods; likely to be longest or 2nd longest part of abstract)

**RESULTS**:(2-5 sentences on results; likely to be longest or 2nd longest part of abstract)

**CONCLUSION**(2-3 sentences stating the biological/ecological/scientific conclusions that can be drawn from the study).

## INTRODUCTION

* Please include 3 to 4 citations from **peer reviewed journal articles** here.
* (There needs to be 4 references in the papers; at least 3 have to be in the introduction)
* References should be properly formated, ie "Brouwer et al (2015)."
* Use your introductory material from the previous assignment
* Please end your intro with a sentence or two stating what the overall goal/hypothesis /question of the study is.
* Please update your intro based on the suggestions
* I do not expect you to write a perfect introduction, but I expect to see improvement from your first draft

## METHODS

### *Study species*

* If you study is focused on a particular species include a few sentences about its basic biology in case your reader is not an expert on that species.
* All statements of fact should be cited
* ie, "The African lion (*Panthera leo*) is the largest carnivore in the world and can live to ages of ... (citation). While endangered in much of its range, trophy hunting is legal in parts of Africa (cite) ...""

### *Study site*

* If your study occurred at a particular location that someone would have to re-visit if they wanted to replicate or follow up on your, provide information about where it is located (State, County, etc) and what the physical conditions are like.
* ie, "This study was conduct on the Serengeti of Tanzania and teh Ngorogoro crater. The Serengeti is ... (cite). Ngorogo crater is ... (cite)."

## *Field/Lab methods*

* How data was collected

## *Statistical methods*

I tested whether there was a significant relationship between nose pigmentation and lion age using linear regression. To determine if this relationship varied between lion populations I tested for an effect of population and a pigmenation\*popualtion interaction. Data were logged transformed to meet the assumptions of linear regression. All analyses were carried out in R 3.3.1 (R Core Team 2016).

## RESULTS

### *Descriptive statistics*

* You should start you results with a few sentences about the general characteristics of the data you collected

Data on nose pigmentation and age was collected for 32 lions; 22 lions were from the Serengeti population and 10 from Ngorogoro. (Raw data are availabe in Table A1 in Appendix 1). The mean age of lions in The Serengeti sample popualtion was 3 years (SE = 0.3; Figure 1) and the mean in the Ngorogoro sample was 7.13 (SE = 0.85). The distribution of ages for each population is shown in Figure A1.

### *Statistical results*

There was a significant positive relationship between nose pigmentation and lion age (F = 44.364, p < 0.0001). The Serengeti and Ngorogoro populations had different intercepts (F = 23.17, p < 0.0001) but there was no evidence of a significant population by pigmentation interaction (F = 0.72, p = 0.40). The best fitting model therefore had seperate intercepts for each popualtion but a single slope (Figure 2; slope = 1.74, SE = 0.29).

## DISCUSSION

* A discussion section is optional
* Up to 5 points will be given as extra credit for a reasonably well-written discussion. THe discussion should include references

## ACKNOWLEDGEMETNS

* Recognize anyone who helped you on the project, such as professor who you consulted with on the project, roommates who helped you collect data, or classmates who helped you with stats or code.

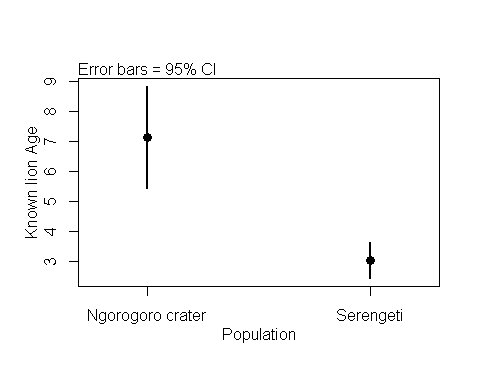
## REFERENCES

* You should have 4 references from peer reviewed journals
* Your 5th reference should be for R, as shown below.

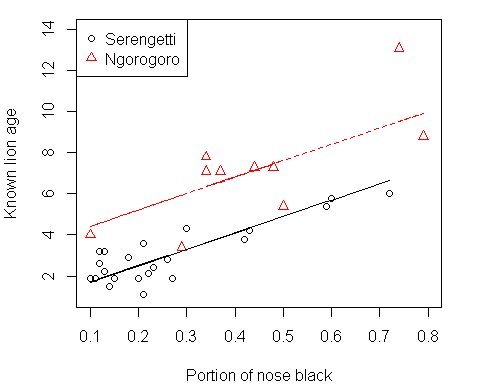
R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

## FIGURES

**Figure 1:** Mean ages in the two samples of lions. Error bars are approximate 95% confidence intervals.



**Figure 2:** Relationship between nose pigmentation and age of male African lions (*Leo panthera*) in the Serengeti (n=22) and Ngorogoro populations (n = 10), Tanzania.



## APPENDIX 1: Raw data table

**Table A1:** Raw data from 32 African lions (*Panthera leo*) from Tanzania. Originally published ... [add reference]

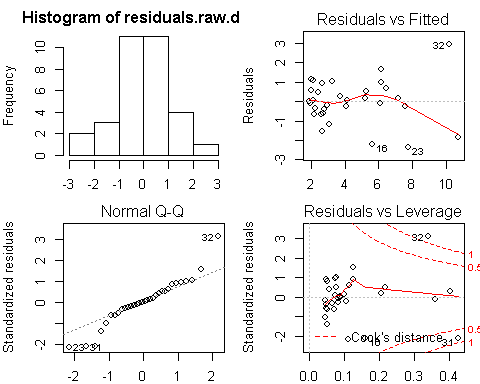
[Add table]

## APPENDIX 2: Raw data plots

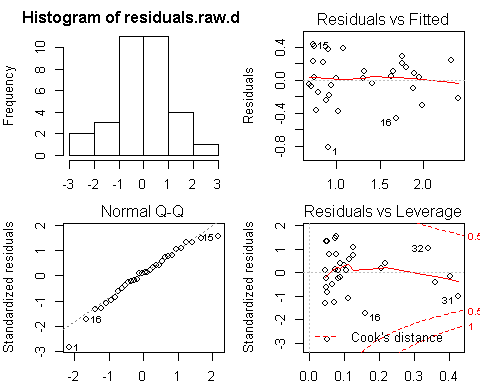
**Table A2:** Distribution of ages in the samples used for regression analysis.

## APPENDIX 3: Model diagnostics

**Table A1-1: Diagnostic plots for un-tranformed data**



**Table A1-2: Diagnostic plots for log-tranformed data**



# Notes

## Title Notes

* Species names should be italized in titles and abstracts

## Abstract

* Abstracts almost never have citations and never have quotes from authors.
* Don't say you transformed that data in abstract, just the methods

## Introduction

## Methods

### Statistical methods

* In your methods need to include a statement like this. "All analyses were carried out in R 3.3.1 (R Core Team 2016)."

In your references include:

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

* This should be your 5th (or greater) reference. \* You should have 4 other references from peer-reviewed scientific papers.
* Unless you use a very special package or technique you do not need to mention the function or code you used.
* That is, you DO NOT need to say "I used the t.test function", or "I used the plot2means function."

### Results

* If you transformed that data, just report those results (ie F values and P values)