Here put a nice title for your report: it should say something about your question or your most important result

Your name comes here

18 Jan 2015

**Introduction**

In this section introduce the problem and the data you are going to analyze. Cite relevant literature like this (Matthews 2000). All the reference you cite should appear in the **Literature cited** section at the end of your report. If you use this template in RStudio, you can transform your report into a word file, simply pressing the button "Knit Word" in the bar above. If you need help with the Markdown language, click on the "?" button next to the "Knit Word"" button above. To make sure that your next paragraph starts on the following line, leave two spaces after the paragraph (for example after the title of each section).

**Methods**

*Data*  
Provide details about the origin of the data you are analyzing. If you collected the data on your own, please provide sufficient detail to understand how, where and in which context you collected it (was this data collected for your thesis? Did you collect it especially for this project? etc.). If the data was obtained from an other source please give details about how and from where it was obtained. If you downloaded it from the internet plese provide and cite the original source providing as well the exact link from where you downloaded it.

You should also provide a brief, but accurate, description of all the variables included in your analysis. You want to make sure that I clearly can understand what you were trying to do and why. And if I do not understand what the variables you included in your analysis are, it will be very difficoult for me to evaluate your work.

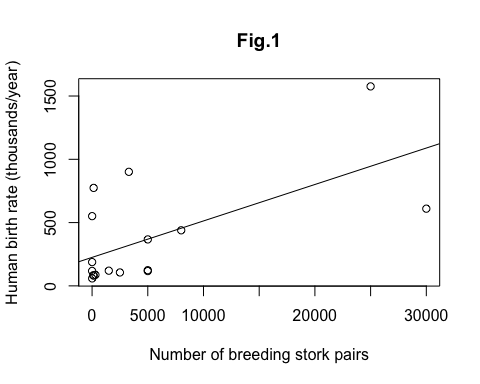
*Statistical methods* Describe briefly the statistical methods you used for your analysis. Specify what kind of method you used and for what, specifying as well the function and packages you used in R. E.g.: "I used simple linear regression to analyze the relationship between number of breeding stork pairs and the human birth rate in European countries. All analyses were conducted in the statistical environment R version 3.0.2. (R Core Team 2014). For linear regression I used the function "lm" in the {stats} package".

**Results** This is the most important part of your report. Please try to write down exactly what you did and why, providing all the code you used to perform your analysis and as well to generate your plots. Include your code in code chunks like in the example below.

E.g.

As shown also by Matthews (2000), the number of breeding pairs in European countries, is positivelly related to the human birth rate in those countries (Fig. 1)

storks.dat <- read.csv("storks.csv")   
with(storks.dat, plot(Birth~Storks, xlab="Number of breeding stork pairs", ylab="Human birth rate (thousands/year)", main="Fig.1"))  
storks.lm<-lm(Birth~Storks,data=storks.dat)  
abline(storks.lm)



plot of chunk unnamed-chunk-1

The linear regression between number of storks (Storks) and human birth rates (Birth) appears to be highly significant (Beta= ..., s.e. = ..., df= ... , p-value = ....; Fig. 2a):

summary(storks.lm)

##   
## Call:  
## lm(formula = Birth ~ Storks, data = storks.dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -479 -166 -145 -2 631   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 225.0287 93.5606 2.41 0.0295 \*   
## Storks 0.0288 0.0094 3.06 0.0079 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 332 on 15 degrees of freedom  
## Multiple R-squared: 0.385, Adjusted R-squared: 0.344   
## F-statistic: 9.38 on 1 and 15 DF, p-value: 0.0079

However, it can be shown that Birth is also significantly related to a third variable (Area; Beta= ..., s.e. = ..., df= ... , p-value = .... ):

storks2.lm<-lm(Birth~Area,data=storks.dat)  
summary(storks2.lm)

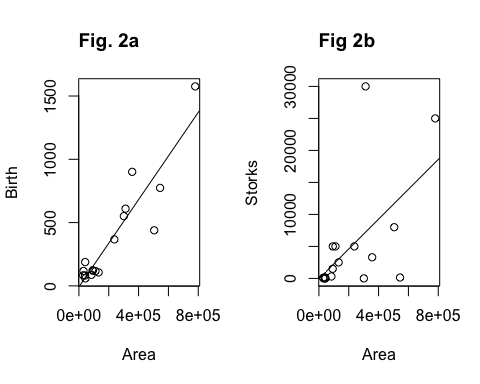
##   
## Call:  
## lm(formula = Birth ~ Area, data = storks.dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -422.9 -49.7 -7.5 73.2 293.7   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -7.775499 56.937678 -0.14 0.89   
## Area 0.001723 0.000186 9.26 1.4e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 163 on 15 degrees of freedom  
## Multiple R-squared: 0.851, Adjusted R-squared: 0.841   
## F-statistic: 85.7 on 1 and 15 DF, p-value: 1.36e-07

A significant relationship can be also shown regressing Storks vs. Area (Beta= ..., s.e. = ..., df= ... , p-value = ....; Fig. 2b):

storks3.lm<-lm(Storks~Area,data=storks.dat)  
summary(storks3.lm)

##   
## Call:  
## lm(formula = Storks ~ Area, data = storks.dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12479 -1594 -650 -475 22773   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -6.07e+01 2.59e+03 -0.02 0.982   
## Area 2.33e-02 8.47e-03 2.75 0.015 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7440 on 15 degrees of freedom  
## Multiple R-squared: 0.336, Adjusted R-squared: 0.291   
## F-statistic: 7.58 on 1 and 15 DF, p-value: 0.0148

par(mfrow = c( 1,2))  
  
with(storks.dat,plot(Birth~Area))  
title("Fig. 2a",adj=0)  
abline(storks2.lm)  
  
with(storks.dat,plot(Storks~Area))  
title("Fig 2b",adj=0)  
abline(storks3.lm)



plot of chunk unnamed-chunk-5

Country area seems indeed to be strongly correlated both to human birth rate and the number of stork pairs. I thus explored what happens when we statistically control for the effect of the confounding variable Area, including this variable in the model of the relationship between Storks and Birth:

storks4.lm <- lm(Birth ~ Area + Storks, data = storks.dat)  
summary(storks4.lm)

Call:  
lm(formula = Birth ~ Area + Storks, data = storks.dat)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-400.7 -57.5 -27.9 77.1 323.4   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) -7.411687 56.702180 -0.13 0.90   
Area 0.001583 0.000227 6.96 6.6e-06 \*\*\*  
Storks 0.005995 0.005651 1.06 0.31   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 163 on 14 degrees of freedom  
Multiple R-squared: 0.862, Adjusted R-squared: 0.842   
F-statistic: 43.8 on 2 and 14 DF, p-value: 9.45e-07

In the above analysis it appears clear that when we statistically control for Area, which is probably affecting indipendently both Storks and Birth, the effect of Storks on Birth appears to be actually not significant (Beta= ..., s.e. = ..., df= ... , p-value = ....)

**Discussion**

In the discussion you should briefly discuss your results. Pinpoint what they add to what you previously knew (from the literature or from your own hypotheses or observations) about the biological problem and scientific question you wanted to analyze and which you had presented in the introduction. It's a good idea to start your discussion with a very syntetic summary of your results. Please cite all relevant literature, and try interpret your results for their biological more than for their statistical relevance here.

**Literaure cited**

Matthews (2000). Storks Deliver Babies (p=0.008). *Teaching Statistics 22(2):36-38*

R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org/>.