Reproducible Research Demo

OHSU Center for Health Systems Effectiveness

Wednesday, July 22, 2015

Last update by Stephanie Renfro ([renfrst@ohsu.edu](mailto:renfrst@ohsu.edu)) on 2015-07-22 11:39:10 using R version 3.1.2 (2014-10-31).

## Purpose

This work was inspired by the following email from Farmer Ben.

From: Ben Chan   
Sent: Thursday, June 11, 2015 4:04 PM  
To: Stephanie Renfro  
Subject: What to feed chicks  
  
Hello,  
  
I'm receiving 20 baby chicks next month. Can you help me decide what to feed them?  
  
I'm choosing between the following four diets:  
  
1. Grower diet  
2. Layer diet  
3. Breeder diet  
4. High cluckage diet  
  
Thanks,  
Ben  
  
-----------------------------------------------------------------  
Ben Chan, Farmer and Research Associate  
OHSU Center for Health Systems Effectiveness  
Office: 3030 SW Moody | Mail Code: MDYCHSE  
www.ohsu.edu/chse

## Preliminaries

Start clock to calculate total runtime.

start\_program <- proc.time()

Load needed packages:

* *data.table* - for faster processing
* *knitr* - for better table display ("kable" function)
* *ggplot2* - for pretty plots

packages <- c("data.table", "knitr", "ggplot2")  
sapply(packages, require, character.only=TRUE, quietly=TRUE)

## data.table knitr ggplot2   
## TRUE TRUE TRUE

Define the CHSE color palette function.

colorPalette <- function () {  
 c(rgb( 1, 67, 134, maxColorValue=255),  
 rgb(119, 120, 123, maxColorValue=255),  
 rgb(139, 184, 234, maxColorValue=255),  
 rgb(188, 190, 192, maxColorValue=255),  
 rgb( 94, 122, 162, maxColorValue=255),  
 rgb(223, 122, 28, maxColorValue=255))  
}

## Prepare Data

This demo uses [data from an experiment on the effect of diet on early growth of chicks](http://www.inside-r.org/r-doc/datasets/ChickWeight), ChickWeight, which comes pre-loaded in any R session.

Let's take a look at the first few rows:

head(ChickWeight)

## weight Time Chick Diet  
## 1 42 0 1 1  
## 2 51 2 1 1  
## 3 59 4 1 1  
## 4 64 6 1 1  
## 5 76 8 1 1  
## 6 93 10 1 1

Let's also print a summary of the data.

*Note, by specifying the option "echo = FALSE", the resulting output will display, but not the code that generated it.*

## weight Time Chick Diet   
## Min. : 35.0 Min. : 0.00 13 : 12 1:220   
## 1st Qu.: 63.0 1st Qu.: 4.00 9 : 12 2:120   
## Median :103.0 Median :10.00 20 : 12 3:120   
## Mean :121.8 Mean :10.72 10 : 12 4:118   
## 3rd Qu.:163.8 3rd Qu.:16.00 17 : 12   
## Max. :373.0 Max. :21.00 19 : 12   
## (Other):506

Convert to data.table for faster processing.

ChickWeight <- data.table(ChickWeight)

Create a table showing mean weight at times 0, 10, and 21 days, for each of the four diet types.

mean\_ChickWeight <- ChickWeight[Time %in% c(0,10,21),  
 list(mean\_weight = round(mean(weight), digits=1)),  
 by = list(Diet,Time)]  
  
kable(mean\_ChickWeight)

|  |  |  |
| --- | --- | --- |
| Diet | Time | mean\_weight |
| 1 | 0 | 41.4 |
| 1 | 10 | 93.1 |
| 1 | 21 | 177.8 |
| 2 | 0 | 40.7 |
| 2 | 10 | 108.5 |
| 2 | 21 | 214.7 |
| 3 | 0 | 40.8 |
| 3 | 10 | 117.1 |
| 3 | 21 | 270.3 |
| 4 | 0 | 41.0 |
| 4 | 10 | 126.0 |
| 4 | 21 | 238.6 |

Create a character variable for diet. Use this variable for plotting small multiples.

ChickWeight[, dietChr := sprintf("Diet %d", Diet)]

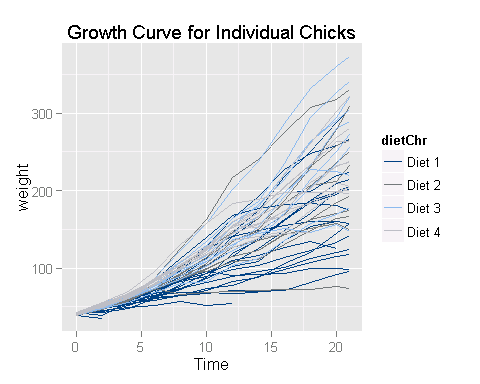
## weight Time Chick Diet dietChr  
## 1: 42 0 1 1 Diet 1  
## 2: 51 2 1 1 Diet 1  
## 3: 59 4 1 1 Diet 1  
## 4: 64 6 1 1 Diet 1  
## 5: 76 8 1 1 Diet 1  
## ---   
## 574: 175 14 50 4 Diet 4  
## 575: 205 16 50 4 Diet 4  
## 576: 234 18 50 4 Diet 4  
## 577: 264 20 50 4 Diet 4  
## 578: 264 21 50 4 Diet 4

## Growth for Individual Chicks

The following plot illustrates the growth curve for individual chicks from 0 to 21 days.

Colors represent the four diets.

ggplot() +  
 geom\_line(data=ChickWeight, aes(x=Time, y=weight, color=dietChr, group=Chick)) +  
 scale\_color\_manual(values=colorPalette()) +  
 ggtitle("Growth Curve for Individual Chicks")

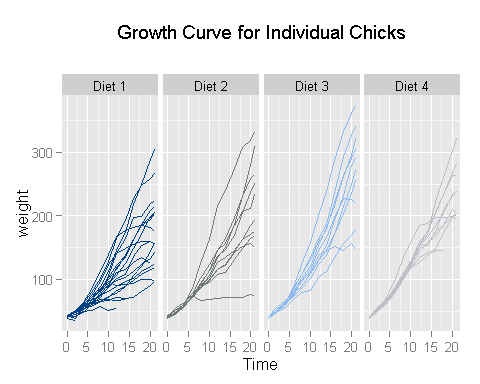


**From this plot, it is difficult to distinguish between the performance of the four diets.**

### Individual growth curves

Plot individual chick growth curves using small multiples.

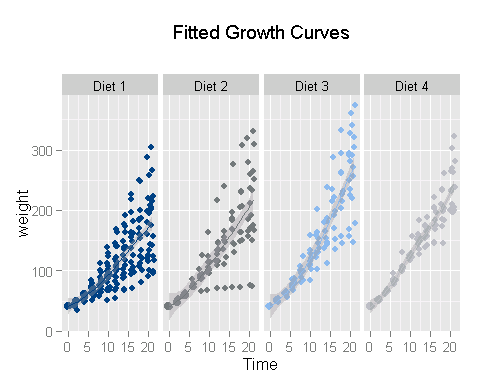
ggplot() +  
 geom\_line(data=ChickWeight, aes(x=Time, y=weight, color=dietChr, group=Chick)) +  
 facet\_wrap(~ dietChr, nrow=1) +   
 scale\_color\_manual(values=colorPalette()) +  
 theme(legend.position="none") +  
 ggtitle(bquote(atop("Growth Curve for Individual Chicks")))



### Fitted growth curves

Plot fitted growth curves using small multiples. Data points are jittered around time value.

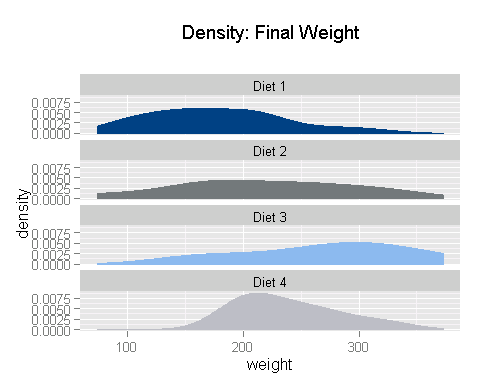
ggplot() +  
 geom\_jitter(data=ChickWeight, aes(x=Time, y=weight, colour=dietChr)) +  
 geom\_smooth(data=ChickWeight, aes(x=Time, y=weight, colour=dietChr)) +  
 facet\_wrap(~ dietChr, nrow=1) +   
 scale\_color\_manual(values=colorPalette()) +  
 theme(legend.position="none") +  
 ggtitle(bquote(atop("Fitted Growth Curves")))



### Final weight density

Plot densities by diet for chicks' final weights (day 21) using small multiples.

ggplot(ChickWeight[Time==21], aes(x=weight, colour=dietChr, fill=dietChr)) +  
 geom\_density() +  
 facet\_wrap(~ dietChr, nrow=4) +   
 scale\_color\_manual(values=colorPalette()) +  
 scale\_fill\_manual(values=colorPalette()) +  
 theme(legend.position="none") +  
 ggtitle(bquote(atop("Density: Final Weight")))



## Wrap Up

Calculate total runtime and clear memory.

time\_program <- proc.time()-start\_program  
print(paste("Total runtime:", format(time\_program[3]/60,digits=3), "minutes"))

## [1] "Total runtime: 0.0913 minutes"

rm(list=ls())  
invisible(gc())