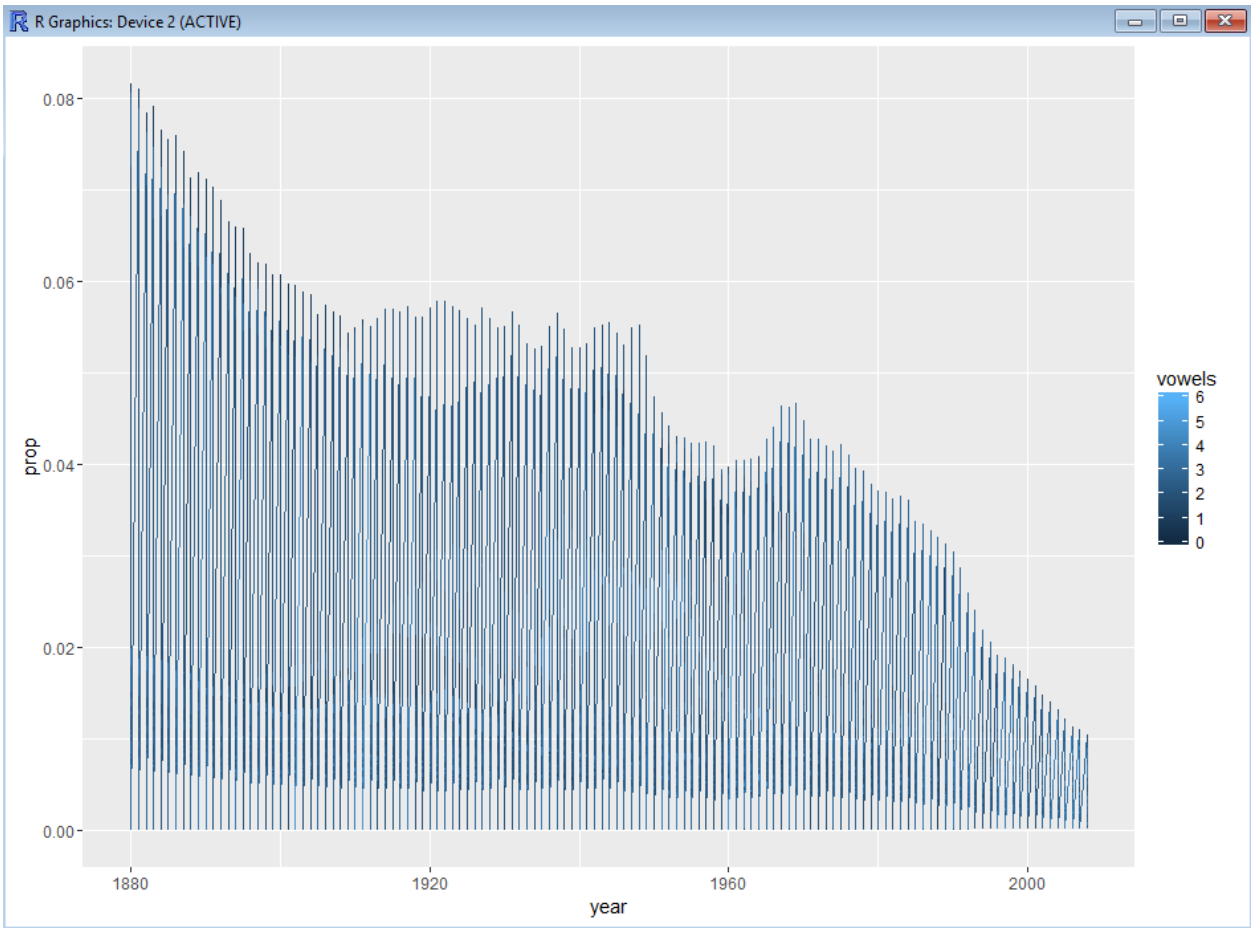
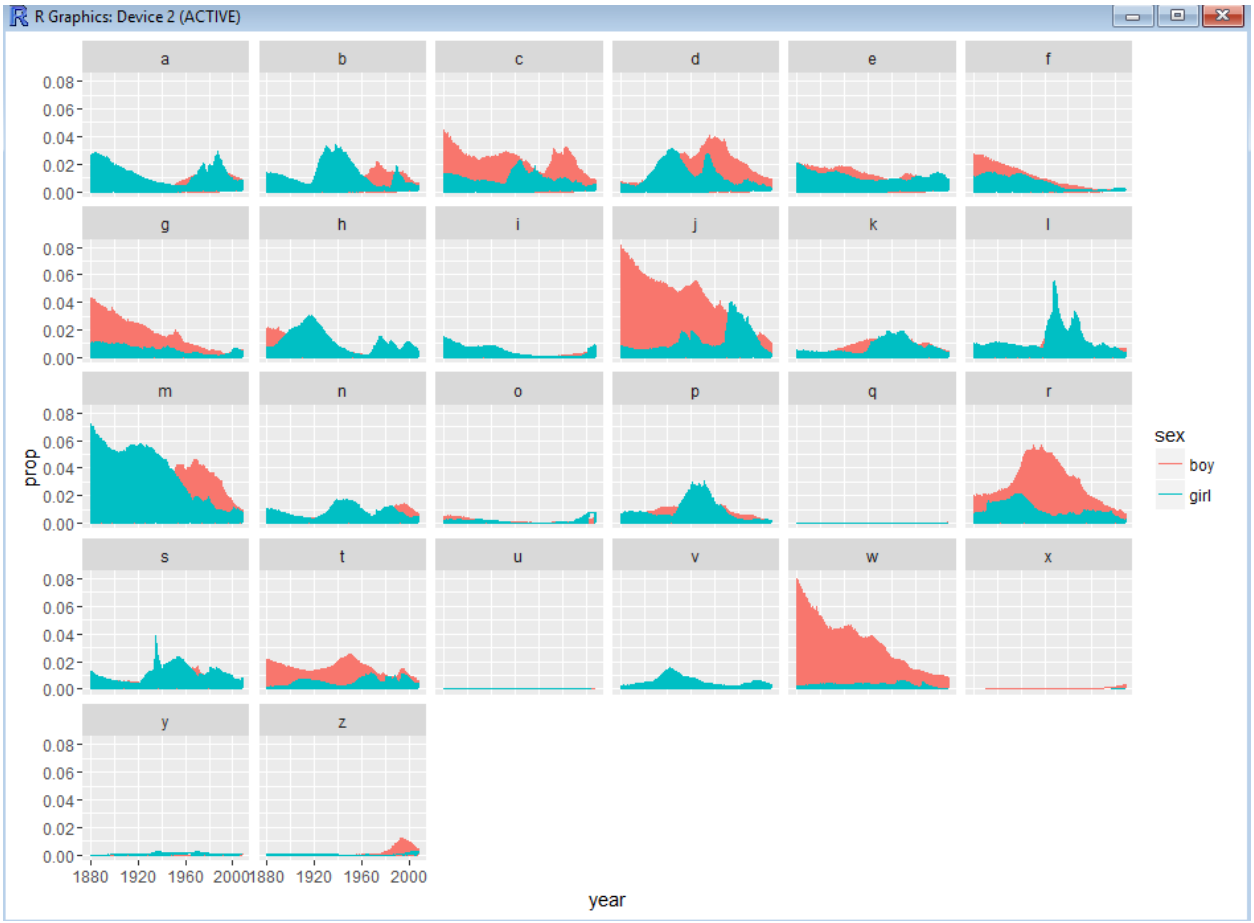
Daniel Hanks Jr
IST719 Exercise 8: Transformations

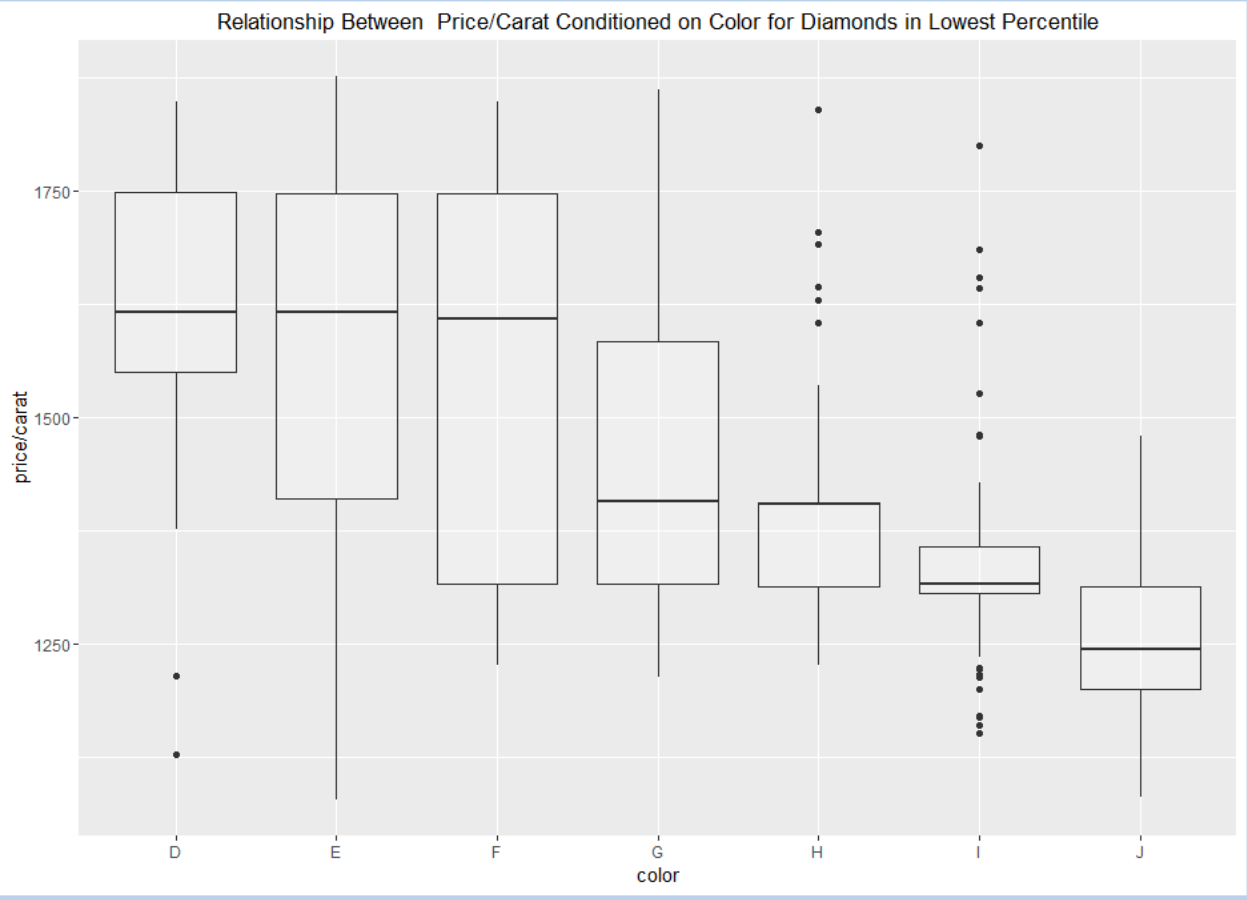
GRAPH A



Graph B


R Graphics: Device 2 (ACTIVE)
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x
y
z
0.08
0.06
0.04
0.02
0.00
1880
1920
1960
2000
prop
year
sex
boy
girl

Graph C



##

R Code for Graph A + B

```
library(plyr)

library(ggplot2)

library(stringr)


options(stringsAsFactors=FALSE)

bnames <- read.csv("baby-names2.csv.bz2")

births  <-  read.csv("births.csv")

bnames <-join(bnames,births,by=c("year","sex"))
```

```
bnames <-mutate(bnames,n=round(prop*births))


vowels <- function(x){str_length(str_replace_all(tolower(x),"[^aeiouy]",""))}


bnames <- transform(bnames,first=tolower(str_sub(name, 1,1)),last=tolower(str_sub(name,-1,-
1)),vowels=vowels(name),length=nchar(name),per10000=10000*prop,one_per=1/prop)


#Graph A

qplot(year,prop,data=bnames,colour=vowels,geom="line")


#Graph B

qplot(year,prop,data=bnames,colour=sex,geom="line")+facet_wrap(~first)

#####################################################################

R code for Graph C

library(ggplot2)
set.seed(1410) #Make the sample reproducible

#create price variable from price column in dsmall data
price <- diamonds$price

#compute the bottom 1% percentile of price (429 is the result)
bottomPrice <- quantile(price, .01)

#create a variable that contains a subset of the data that is less than 429
cheap <- subset(diamonds, price < bottomPrice)

#Investigate distribution of price per carat, condtional on color
qplot(color,price/carat,data=cheap, geom="jitter",alpha=I(1/5))

qplot(color,price/carat,data=cheap, geom="boxplot")
```