

analyze.R

joebrew

Tue Jun 23 16:46:10 2015

```
options(scipen=999)
```

```
library(readxl)
```

```
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'  
##  
## The following object is masked from 'package:stats':  
##  
##     filter  
##  
## The following objects are masked from 'package:base':  
##  
##     intersect, setdiff, setequal, union
```

```
library(waffle)
```

```
## Loading required package: ggplot2
```

```
library(RColorBrewer)  
df <- read_excel('../data/thegardenproducereport14-15.xlsx',  
                 skip = 1)  
names(df) <- c('item', 'number', 'month', 'uom', 'qty', 'price', 'state')  
  
# Get rid of grand total  
df <- df[which(is.na(df$item) | df$item != 'Grand Total'),]  
  
# Get right names in item / number / state (de-aggregate)  
for (i in 1:nrow(df)){  
  while(is.na(df$item[i])){  
    new_row <- i-1  
    df$item[i] <- df$item[new_row]  
  }  
  while(is.na(df$number[i])){  
    new_row <- i-1  
    df$number[i] <- df$number[new_row]  
  }  
  #print(i)  
}  
  
# Fix the funky date stuff (get rid of month headers and take previous rows)  
df$is_date <- grepl("^[[[:digit:]]", df$month)  
df$date <- ifelse(df$is_date,  
                 format(as.Date(df$month, format = '%d-%B-%y'), '%Y-%m-%d'),
```

```

      NA)
df$date <- as.Date(df$date)

# repair dates and states
for (i in 2:nrow(df)){
  while(is.na(df$date[i])){
    new_row <- i-1
    df$date[i] <- df$date[new_row]
  }
  while(is.na(df$state[i])){
    new_row <- i-1
    df$state[i] <- df$state[new_row]
  }
  #print(i)
}}

# Get rid of the aggregated month rows
df <- df[which(!is.na(df$qty)),
        c('date', 'item', 'number', 'uom', 'qty', 'price', 'state')]

# Clean up state names
df$state <- toupper(df$state)
df$state[which(df$state == 'N/A')] <- NA
df$state[which(df$state == 'FL (ORANGE)')] <- 'FLORIDA'
df$state[which(df$state == 'COSATA RICA')] <- 'COSTA RICA'

# Give florida / non-florida column
df$florida <- ifelse(df$state == 'FLORIDA', 'Florida', 'Non-Florida')

# Get rid of the 30 dollars with no state associated
df <- df[which(!is.na(df$state)),]

#####
# WHAT PERCENTAGE OF OVERALL PRODUCE IS FLORIDA PRODUCE?
#####
simple <- df %>%
  group_by(florida) %>%
  summarise(dollars = sum(price, na.rm = TRUE))
simple$color <- adjustcolor(ifelse(simple$florida == 'Florida', 'blue', 'black'), alpha.f = 0.6)
complicated <- df %>%
  group_by(state) %>%
  summarise(dollars = sum(price, na.rm = TRUE))
complicated <- complicated[order(complicated$dollars),]
complicated$color <- adjustcolor(ifelse(complicated$state == 'FLORIDA', 'blue', 'black'), alpha.f = 0.6)

#####
# VISUALIZE
#####
par(mar = c(7,6,5,4))

bp <- barplot(simple$dollars,
              names.arg = simple$florida,
              las = 1,

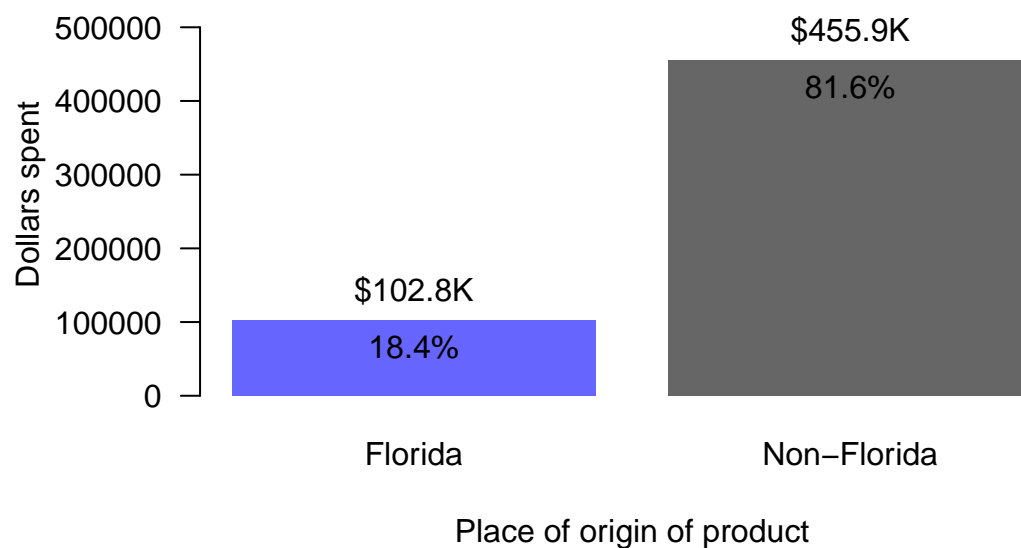
```

```

        ylab = NA,
        xlab = 'Place of origin of product',
        col = simple$color,
        border = NA,
        ylim = c(0, max(simple$dollars, na.rm = TRUE) * 1.2))
mtext(side = 2, line = 4, 'Dollars spent')
text(x = bp[,1],
     y = simple$dollars,
     labels = paste0('$', round(simple$dollars / 1000, digits = 1), 'K'),
     pos = 3)
text(x = bp[,1],
     y = simple$dollars,
     labels = paste0(round(simple$dollars / sum(simple$dollars) * 100, digits = 1), '%'),
     pos = 1)
title(main = 'Where do our food dollars go?')

```

Where do our food dollars go?

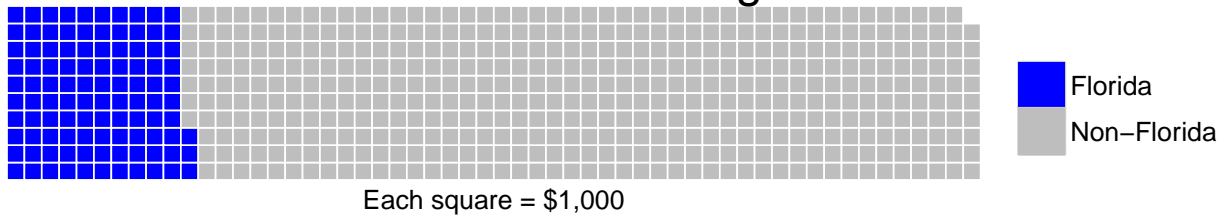


```

##
vec <- round(simple$dollars/1000)
names(vec) <- simple$florida
waffle(vec, colors = c('blue', 'grey'),
       xlab = 'Each square = $1,000',
       size = 0.3,
       title = 'Where do our food dollars go?')

```

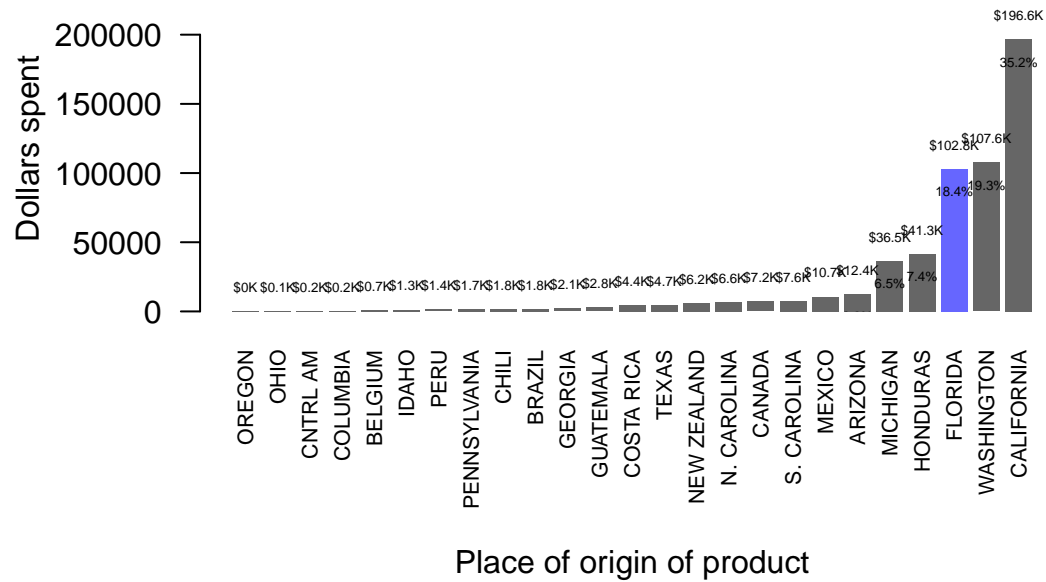
Where do our food dollars go?



```
#####
par(mar = c(9,6,5,4))
bp <- barplot(complicated$dollars,
              #names.arg = complicated$state,
              las = 1,
              ylab = NA,
              xlab = NA,
              col = complicated$color,
              border = NA,
              ylim = c(0, max(complicated$dollars, na.rm = TRUE) * 1.2))
mtext(side = 2, line = 4, 'Dollars spent')
mtext(side = 1, line = 6, 'Place of origin of product')
axis(side = 1,
     at = bp[,1],
     labels = complicated$state,
     las = 3,
     cex.axis = 0.7,
     tick = FALSE)

text(x = bp[,1],
     y = complicated$dollars,
     labels = paste0('$', round(complicated$dollars / 1000, digits = 1), 'K'),
     pos = 3,
     cex = 0.4)
text(x = bp[,1],
     y = complicated$dollars,
     labels = paste0(round(complicated$dollars / sum(complicated$dollars) * 100, digits = 1), '%'),
     pos = 1,
     cex = 0.4)
title(main = 'Where do our food dollars go?\n(detailed)')
```

Where do our food dollars go? (detailed)



```
vec <- round(complicated$dollars/1000)
names(vec) <- complicated$state
vec <- rev(vec)
cols <- colorRampPalette(brewer.pal(8, 'Set2'))(length(vec))
cols[which(names(vec) == 'FLORIDA')] <- 'blue'
waffle(vec, xlab = 'Each square = $1,000',
        colors = cols,
        size = 0.3,
        title = 'Where do our food dollars go?',
        rows = 20)
```

