

DSC640_ASumbaraju_Wk1_2_Charts

SumbarajuAditya

6/18/2021

Read data from csv file

```
hotdog <- read.csv('C:/BU/DSC640/ex1-2/hotdog-places.csv', header = TRUE)
```

```
head(hotdog)
```

| ## | LOCATION | X2000 | X2001 | X2002 | X2003 | X2004 | X2005 | X2006 | X2007 | X2008 | X2009 | X2010 |
|------|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| ## 1 | EAST | 25 | 50.0 | 50.5 | 44.5 | 53.5 | 49 | 54 | 66 | 59 | 68.0 | 54 |
| ## 2 | MID-WEST | 24 | 31.0 | 26.0 | 30.5 | 38.0 | 37 | 52 | 63 | 59 | 64.5 | 43 |
| ## 3 | WEST | 22 | 23.5 | 25.5 | 29.5 | 32.0 | 32 | 37 | 49 | 42 | 55.0 | 37 |

Data preperation

```
library(tidyr)
```

```
hotdog_df <- gather(hotdog, year, quantity, X2000:X2010)
```

```
hotdog_df$year <- gsub("[a-zA-Z ]", "", hotdog_df$year)
```

```
hotdog_df$year <- as.factor(hotdog_df$year)
```

```
hotdog_df$quantity <- as.numeric(hotdog_df$quantity)
```

```
hotdog_df
```

| ## | LOCATION | year | quantity |
|-------|----------|------|----------|
| ## 1 | EAST | 2000 | 25.0 |
| ## 2 | MID-WEST | 2000 | 24.0 |
| ## 3 | WEST | 2000 | 22.0 |
| ## 4 | EAST | 2001 | 50.0 |
| ## 5 | MID-WEST | 2001 | 31.0 |
| ## 6 | WEST | 2001 | 23.5 |
| ## 7 | EAST | 2002 | 50.5 |
| ## 8 | MID-WEST | 2002 | 26.0 |
| ## 9 | WEST | 2002 | 25.5 |
| ## 10 | EAST | 2003 | 44.5 |
| ## 11 | MID-WEST | 2003 | 30.5 |
| ## 12 | WEST | 2003 | 29.5 |
| ## 13 | EAST | 2004 | 53.5 |
| ## 14 | MID-WEST | 2004 | 38.0 |
| ## 15 | WEST | 2004 | 32.0 |
| ## 16 | EAST | 2005 | 49.0 |

```
## 17 MID-WEST 2005      37.0
## 18      WEST 2005      32.0
## 19      EAST 2006      54.0
## 20 MID-WEST 2006      52.0
## 21      WEST 2006      37.0
## 22      EAST 2007      66.0
## 23 MID-WEST 2007      63.0
## 24      WEST 2007      49.0
## 25      EAST 2008      59.0
## 26 MID-WEST 2008      59.0
## 27      WEST 2008      42.0
## 28      EAST 2009      68.0
## 29 MID-WEST 2009      64.5
## 30      WEST 2009      55.0
## 31      EAST 2010      54.0
## 32 MID-WEST 2010      43.0
## 33      WEST 2010      37.0
```

```
library(dplyr)
```

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

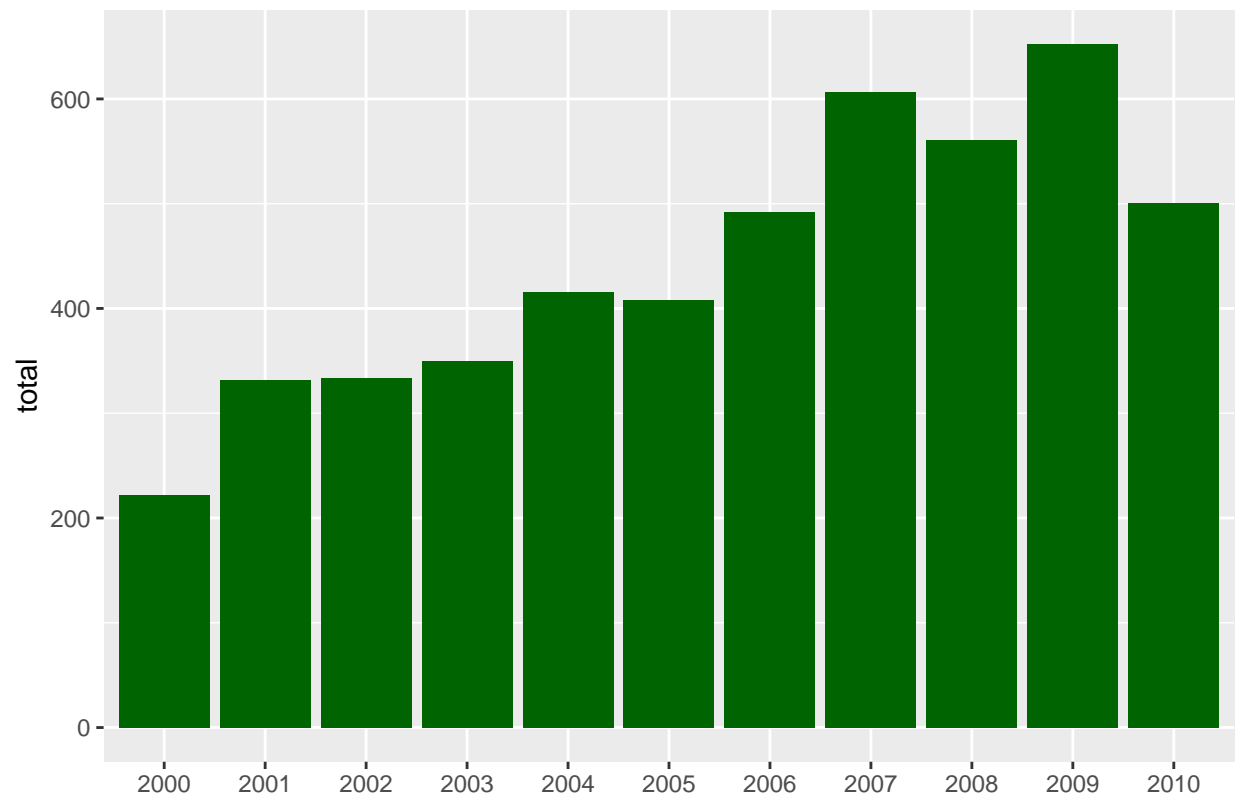
```
library(ggplot2)
library(reshape)
```

```
##
## Attaching package: 'reshape'
## The following object is masked from 'package:dplyr':
##
##      rename
## The following objects are masked from 'package:tidyr':
##
##      expand, smiths
```

```
##barplot
hotdog_by_year <- hotdog_df%>% group_by(year) %>% mutate(total = sum(quantity, year))
```

```
ggplot(hotdog_by_year, aes(x=year, y=total)) +
  geom_bar(stat = 'identity', fill="dark GREEN") +
  labs(title = 'Hot Dogs - Year based consumption Chart', ylab('Count')) +
  theme(axis.title.x = element_blank())
```

Hot Dogs – Year based consumption Chart

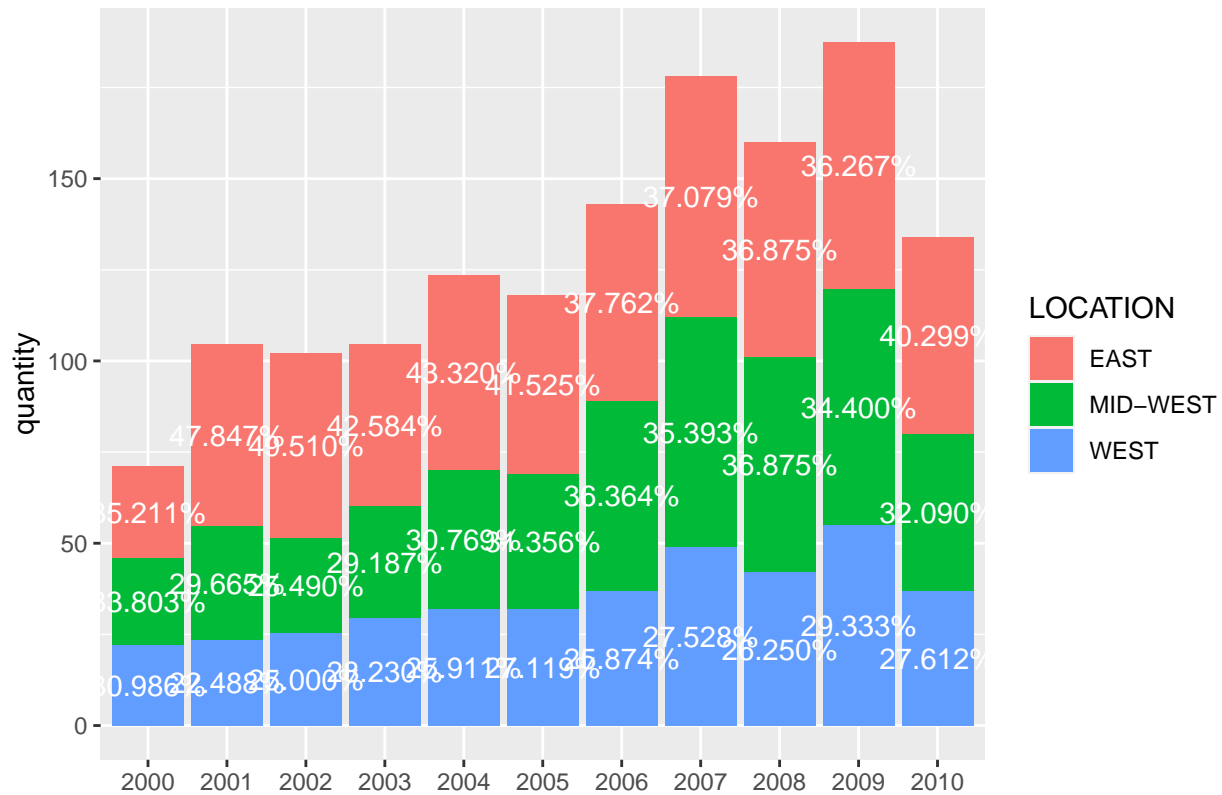


Stacked chart

```
hotdog_Location_year <- hotdog_df %>% group_by(year) %>% mutate(pct= quantity / sum(quantity))

ggplot(hotdog_Location_year, aes(fill=LOCATION, x=year, y=quantity)) +
  geom_bar(position="stack", stat = 'identity') +
  geom_text(aes(label = scales::percent(pct)), position = position_stack(vjust= .5), color="white") +
  labs(title = 'Hot Dogs - Location based consumption stacked Chart per year', ylab('Count')) +
  theme(axis.title.x = element_blank())
```

Hot Dogs – Location based consumption stacked Chart per year



piechart

```
library(plotrix)
```

```
#Pie Chart
```

```
#Distribution of first second and third place hotdogs consumed
```

```
hotdog_Location <- hotdog_df %>% group_by(LOCATION) %>% summarise(total = sum(quantity))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

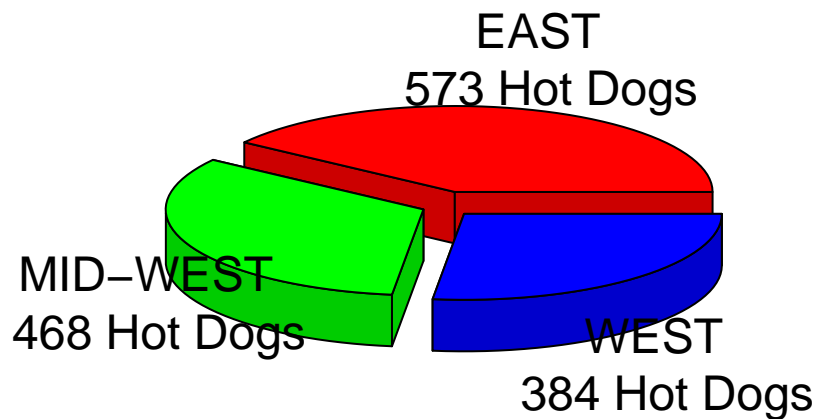
```
hotdog_Location$total <- as.integer(hotdog_Location$total)
```

```
slices <- hotdog_Location$total
```

```
labels <- paste0(hotdog_Location$LOCATION, " \n ", hotdog_Location$total, ' Hot Dogs')
```

```
pie3D(slices, main="Hot Dogs Pie chart - % of Hotdogs consumed per Location", labels = labels,explode=0)
```

Hot Dogs Pie chart – % of Hotdogs consumed per Location



```
## donut chart
# Compute percentages
hotdog_Location$fraction = hotdog_Location$total / sum(hotdog_Location$total)

# Compute the cumulative percentages (top of each rectangle)
hotdog_Location$ymax = cumsum(hotdog_Location$fraction)

# Compute the bottom of each rectangle
hotdog_Location$ymin = c(0, head(hotdog_Location$ymax, n=-1))

# Compute label position
hotdog_Location$labelPosition <- (hotdog_Location$ymax + hotdog_Location$ymin) / 2

# Compute a good label
hotdog_Location$label <- paste0(hotdog_Location$LOCATION, " \n ", hotdog_Location$total, ' Hot Dogs')

# Make the plot
ggplot(hotdog_Location, aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=LOCATION)) +
  geom_rect() +
  geom_label(x=2, aes(y=labelPosition, label=label), size=4, color='white') +
  coord_polar(theta="y") + xlim(c(-1, 4)) + # without this parameter it looks like a pie chart
  labs(title='Distribution of Hot Dogs Consumed by Place') +
  theme_void() +
  theme(legend.position = "none", plot.title = element_text(hjust = 0.5))
```

Distribution of Hot Dogs Consumed by Place

