

# anova\_\_example

*Zongyan Wang*

*Sunday, December 06, 2015*

**This document gives the example on how to do ANOVA and LS-D(HSD) upon different hospital**

**Load package**

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.1.3
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.1.3
```

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

```
library(agricolae)
```

```
## Warning: package 'agricolae' was built under R version 3.1.3
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

**data loading and clean**

```

phy_drugs = fread("F:/Academic/Stat 992/group project/toy_data/wi_cardi2.tab", sep="\t", header=FALSE)
hospitals = unique(phy_drugs[, 20])
setkey(phy_drugs,V1)
#phy_drugs$zip = DT[as.character(phy_drugs$V1), mult = "first"]$`Zip Code`
if(F){
  phy_drugs = phy_drugs[`V9`=="METOPROLOL SUCCINATE"]
}

#First aggregate ratios over physicians
grouped_by_physician = phy_drugs %>%
  group_by(V1) %>%
  select(drug_name = V8, generic_name = V9, total_claim_cnt = V11)

#Calculates ratios based on number of brand name vs total claims
phy_bg_ratios = summarise(grouped_by_physician,
  bg_ratio = (sum(as.vector(total_claim_cnt[as.vector(drug_name) != as.vector(g
)

phy_bg_ratios$hospital <- phy_drugs[.(phy_bg_ratios$V1), mult = "first"]$V20

```

## ANOVA

```

data <- phy_bg_ratios
data <- cbind(data, trt=factor(data$hospital, labels=1:length(unique(data$hospital))))
model <- aov(bg_ratio ~ trt, data = data)
summary(model)

```

```

##              Df Sum Sq Mean Sq F value Pr(>F)
## trt          60  1.034  0.01723   1.556 0.0109 *
## Residuals    238  2.635  0.01107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The result shows there are significant different among hospital.

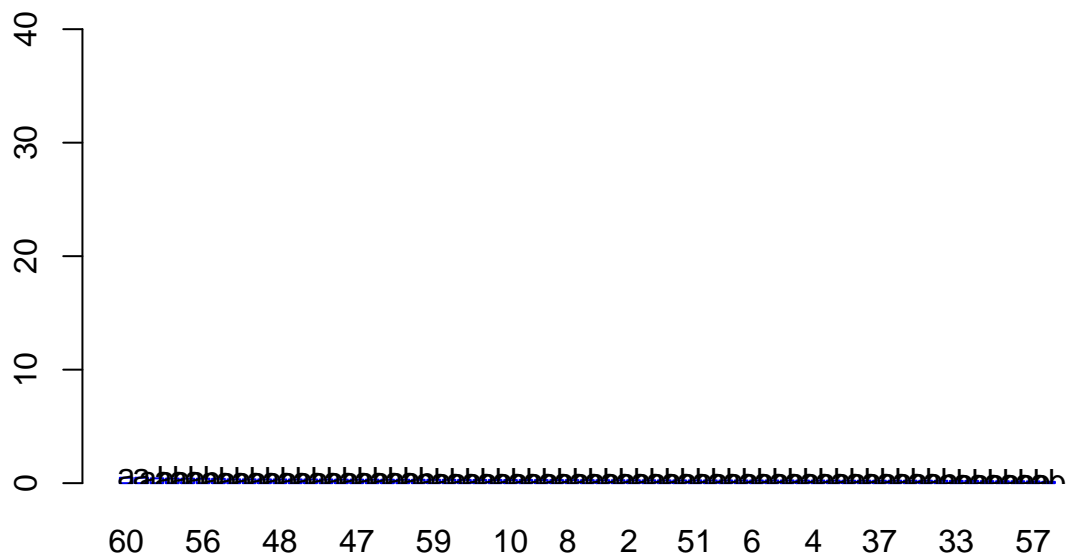
## Doing LSD and HSD test

```

lsd.out <- LSD.test(model, "trt", p.adj = "bonferroni")
hsd.out <- HSD.test(model, "trt", group = T)

bar.group(lsd.out$groups, ylim=c(0,45), density=4, border="blue")
bar.group(hsd.out$groups, ylim=c(0,45), density=4, border="blue")

```



```
#lsd.out$groups
#hsd.out$groups
```

You can match the group number with hospital with the data we use

```
id.group <- data %>%
  select(hospital, trt) %>%
  unique()
hsd.out$groups.id <- merge(id.group, hsd.out$groups, by = "trt")
hsd.out$groups.id
```

```
##      trt
## 1:  10
## 2:  11
## 3:  12
## 4:  13
## 5:  14
## 6:  15
## 7:  16
## 8:  17
## 9:  18
##10:  19
##11:  20
##12:  21
```

## 13: 22  
 ## 14: 23  
 ## 15: 24  
 ## 16: 25  
 ## 17: 26  
 ## 18: 27  
 ## 19: 28  
 ## 20: 29  
 ## 21: 30  
 ## 22: 31  
 ## 23: 32  
 ## 24: 33  
 ## 25: 34  
 ## 26: 35  
 ## 27: 36  
 ## 28: 37  
 ## 29: 38  
 ## 30: 39  
 ## 31: 40  
 ## 32: 41  
 ## 33: 42  
 ## 34: 43  
 ## 35: 44  
 ## 36: 45  
 ## 37: 46  
 ## 38: 47  
 ## 39: 48  
 ## 40: 49  
 ## 41: 50  
 ## 42: 51  
 ## 43: 52  
 ## 44: 53  
 ## 45: 54  
 ## 46: 55  
 ## 47: 56  
 ## 48: 57  
 ## 49: 58  
 ## 50: 59  
 ## 51: 60  
 ## 52: 61

## trt

##

hospital

## 1:	BAYCARE AURORA, LLC
## 2:	BELLIN MEMORIAL HOSPITAL, INC
## 3:	BELOIT HEALTH SYSTEM INC
## 4:	CARDIOLOGY ASSOCIATES OF WAUKESHA, S.C.
## 5:	COLUMBIA ST MARY'S HOSPITAL OZAUKEE, INC.
## 6:	COLUMBIA ST. MARY'S HOSPITAL MILWAUKEE, INC
## 7:	CULLINANE CARDIOVASCULAR CONSULTANTS, SC
## 8:	DEAN HEALTH SYSTEMS, INC
## 9:	EAU CLAIRE HEART INSTITUTE, SC
## 10:	EAU CLAIRE MEDICAL CLINIC SC
## 11:	EDDY D. CO, MD, SC
## 12:	FARZAD KAMRANI,MD.SC.

```

## 13: FROEDTERT AND THE MEDICAL COLLEGE OF WISCONSIN COMMUNITY PHYSICIANS IN
## 14:                                GUNDERSEN CLINIC LTD
## 15:                                GUNDERSEN LUTHERAN MEDICAL CENTER, INC.
## 16:                                HOLY FAMILY MEMORIAL INC
## 17:                                HUDSON PHYSICIANS SC
## 18:                                JOAN T. HARNEY GNADT, MD, SC
## 19:                                LAKE COUNTRY CARDIOVASCULAR ASSOCIATES LTD
## 20:                                LAKESHORE MEDICAL CLINIC LTD
## 21:                                LOUIE COULIS MD SC
## 22:                                MARSHFIELD CLINIC INC
## 23:                                MAYO CLINIC HEALTH SYSTEM-RED CEDAR, INC
## 24:                                MAYO CLINIC HEALTH SYSTEM - EAU CLAIRE CLINIC, INC
## 25:                                MAYO CLINIC HEALTH SYSTEM - FRANCISCAN MEDICAL CENTER, INC.
## 26:                                MERCY HEALTH SYSTEM CORPORATION
## 27:                                MERITER HOSPITAL, INC.
## 28:                                MERITER MEDICAL GROUP INC
## 29:                                MILE BLUFF MEDICAL CENTER INC
## 30:                                MILWAUKEE CARDIOVASCULAR CNTR S.C.
## 31:                                NETWORK HEALTH SYSTEM, INC.
## 32:                                OAKLEAF CLINICS SC
## 33:                                PRACTICE MANAGEMENT GROUP INC
## 34:                                PROHEALTH CARE MEDICAL ASSOCIATES INC
## 35:                                SIXTEENTH STREET COMMUNITY HEALTH CENTERS INC
## 36:                                ST VINCENT HOSPITAL-HOSPITAL SISTERS-THIRD ORDER OF ST FRANCIS
## 37:                                ST. JOSEPH'S ELECTROCARDIOGRAPHIC ASSOCIATES
## 38:                                ST. NICHOLAS HOSPITAL-SISTERS OF THE THIRD ORDER OF ST. FRANCIS
## 39:                                SUHAS K SHELGIKAR MD SC
## 40:                                THE MEDICAL COLLEGE OF WISCONSIN INC
## 41:                                THE MONROE CLINIC, INC.
## 42:                                THEDACARE, INCORPORATED
## 43:                                UNITED HOSPITAL SYSTEM INC
## 44:                                UNIVERSITY OF WISCONSIN MEDICAL FOUNDATION INC
## 45:                                WATERTOWN HEART INSTITUTE SC
## 46:                                WATERTOWN REGIONAL MEDICAL CENTER, INC
## 47:                                WAUKESHA HEART INSTITUTE
## 48:                                WESTFIELDS HOSPITAL, INC.
## 49:                                WHEATON FRANCISCAN MEDICAL GROUP
## 50:                                WISCONSIN HEALTH FUND
## 51:                                WISCONSIN HEART CENTER S C
## 52:                                ZENITH HEALTHCARE SC
##                                hospital
##
##          means  M
## 1: 0.2777667 ab
## 2: 0.2207583 ab
## 3: 0.2714276 ab
## 4: 0.2983064 ab
## 5: 0.3147188 ab
## 6: 0.2784896 ab
## 7: 0.4251262 ab
## 8: 0.2618529 ab
## 9: 0.2445100 ab
## 10: 0.2230500 ab
## 11: 0.2929120 ab
## 12: 0.3091187 ab

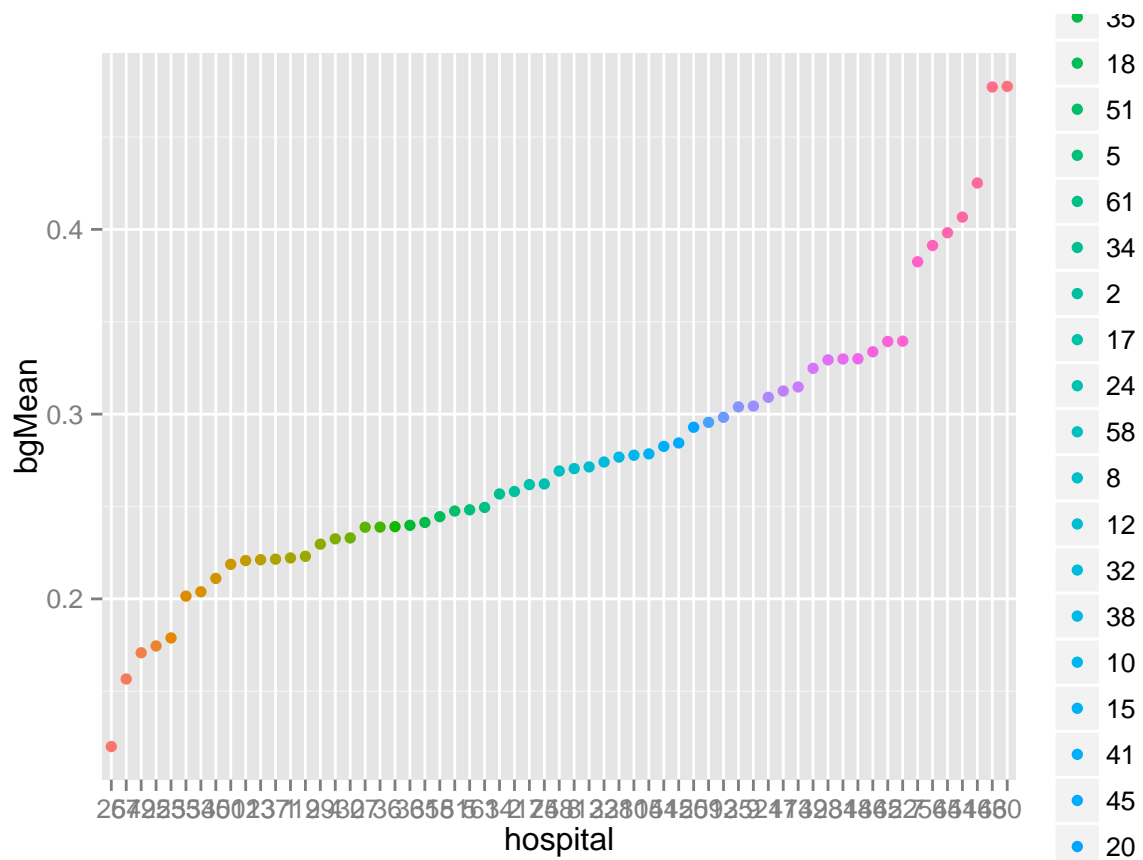
```

```
## 13: 0.1745037 ab
## 14: 0.2212075 ab
## 15: 0.2622252 ab
## 16: 0.3039344 ab
## 17: 0.1200686 ab
## 18: 0.2387944 ab
## 19: 0.3293824 ab
## 20: 0.2296248 ab
## 21: 0.2329979 ab
## 22: 0.3298661 ab
## 23: 0.2740851 ab
## 24: 0.2038654 ab
## 25: 0.2567954 ab
## 26: 0.2413616 ab
## 27: 0.2398291 ab
## 28: 0.2215102 ab
## 29: 0.2767385 ab
## 30: 0.3248027 ab
## 31: 0.2110743 ab
## 32: 0.2825382 ab
## 33: 0.3393418 ab
## 34: 0.4770871 a
## 35: 0.3981913 ab
## 36: 0.2843912 ab
## 37: 0.3337723 ab
## 38: 0.3125156 ab
## 39: 0.3299663 ab
## 40: 0.1708218 ab
## 41: 0.2187086 ab
## 42: 0.2475537 ab
## 43: 0.3394982 ab
## 44: 0.1788730 ab
## 45: 0.4066732 ab
## 46: 0.2014802 ab
## 47: 0.3912818 ab
## 48: 0.1566502 ab
## 49: 0.2692160 ab
## 50: 0.2955656 ab
## 51: 0.4773980 a
## 52: 0.2495088 ab
##           means M
```

This can also be applied to LSD

## plot

```
bgmean <- data %>%
  group_by(trt) %>%
  summarize(bgMean = mean(bg_ratio)) %>%
  mutate(hospital = reorder(x = trt, X = bgMean, min))
ggplot(bgmean, aes(x = hospital, y = bgMean)) +
  geom_point(aes(color = hospital))
```



```
# geom_text(aes(x = hospital, y = bgMean, label = hospital), size = 2.5)
ggsave(filename = "wi_cardio.png", plot = last_plot())
```

```
## Saving 6.5 x 4.5 in image
```