Are Chicago-area Chinese restaurants more/less likely to have online reservations than Chicagoarea Indian restaurants

Packages used for the code

```
library(httr)
library(httpuv)
library(ggplot2)
library(jsonlite)
```

Keys used to make connection(yelp api and R)

```
consumerKey = "56kWW13J51fkcAYQ2CUOnQ"
consumerSecret = "WzYimV2HyVFeBrueIKZAQDQK-Qo"
token = "PqelJy9-fZijYnearhi9Moph7gWMFSau"
token_secret = "aeBla3kHKB4jeg4E_cX4Ea65Qyk"
```

Authorization

```
myapp = oauth_app("YELP", key=consumerKey, secret=consumerSecret)
sig=sign_oauth1.0(myapp, token=token,token_secret=token_secret)
```

Limit set for obtaining data in one result

```
limit<- 20
```

Query used to get data(chinese restaurant)

Fetch data using Yelp API of chinese restaurant(20 at a time),than make a data frame of it. Total number of data is 80.

```
yelpurl <- paste0("https://api.yelp.com/v2/search/?term=reservation&location=Chicago,IL&offset=1&limit=

# To get the data
locationdata=GET(yelpurl, sig)
locationdataContent = content(locationdata)
```

```
# Convert JSON to list
locationdataList=jsonlite::fromJSON(toJSON(locationdataContent))

# Convert list to dataframe containing lists
yelpfirst<- data.frame(locationdataList)


# Second set of data
yelpurl2<- paste0("https://api.yelp.com/v2/search/?term=reservation&location=Chicago,IL&offset=21&limit=

# To get the data
locationdata=GET(yelpurl2, sig)
locationdataContent = content(locationdata)

# Convert JSON to list
locationdataList=jsonlite::fromJSON(toJSON(locationdataContent))

# Convert list to dataframe containing lists
yelpsecond<- data.frame(locationdataList)

# Third set of data
yelpurl3 <- paste0("https://api.yelp.com/v2/search/?term=reservation&location=Chicago,IL&offset=120&lim

# To get the data
locationdata=GET(yelpurl3, sig)
locationdataContent = content(locationdata)

# Convert JSON to list
locationdataList=jsonlite::fromJSON(toJSON(locationdataContent))

# Convert list to dataframe containing lists
yelpthird<- data.frame(locationdataList)

# Fourth set of data
yelpurl4 <- paste0("https://api.yelp.com/v2/search/?term=reservation&location=Chicago,IL&offset=141&lim

# To get the data
locationdata=GET(yelpurl4, sig)
locationdataContent = content(locationdata)

# Convert JSON to list
locationdataList=jsonlite::fromJSON(toJSON(locationdataContent))

# Convert list to dataframe containing lists
yelpfourth<- data.frame(locationdataList)
```

Scrubbing the data(chinese restaurant)

After getting data of chinese restaurant, unneccessary data should be deleted and keep the final one which consist of useful data

```
# Deleting unneccessary cloumns from each set of data
yelpfirst1 <- yelpfirst[-c(1:10, 12:24)]
yelpfirst20 <- yelpfirst1[-c(2:3)]

yelpsecond1 <- yelpsecond[-c(1:10, 12:24)]
yelpsecond20 <- yelpsecond1[-c(2:3)]

yelpthird1 <- yelpthird[-c(1:10, 12:24)]
yelpthird20 <- yelpthird1[-c(2:3)]

yelpfourth1 <- yelpfourth[-c(1:10, 12:26)]
yelpfourth20 <- yelpfourth1[-c(2:3)]

# Combining all the data set into single one
Yelpfinal <- rbind(yelpfirst20, yelpsecond20, yelpthird20, yelpfourth20)

# Deleting unnccessary rows from the final data set
Yelpfinal1 <- Yelpfinal[-c(1:4, 6:9, 11:14, 16:37, 39:44, 46:69, 71:91),]

## inserting new column with name index
Yelpfinal1$index <- seq.int(nrow(Yelpfinal1))

# Chnage the values of rows to same
Yelpfinal1$index <- "chinese"
```

Query used to get data(indian restaurant)

Fetch data using Yelp API of indian restaurant(20 at a time),than make a data frame of it. Total number of data is 80.

```
# First set of data and query used to get data(indian restaurant)
yelpurl5 <- paste0("https://api.yelp.com/v2/search/?term=reservation&location=Chicago,IL&offset=1&limit=

# To get the data
locationdata=GET(yelpurl5, sig)
locationdataContent = content(locationdata)

# Convert JSON to list
locationdataList=jsonlite::fromJSON(toJSON(locationdataContent))

# Convert list to dataframe containing lists
yelpfifth<- data.frame(locationdataList)


# Second set of data
yelpurl6<- paste0("https://api.yelp.com/v2/search/?term=reservation&location=Chicago,IL&offset=50&limit=

# To get the data
locationdata=GET(yelpurl6, sig)
locationdataContent = content(locationdata)

# Convert JSON to list
```

```
locationdataList=jsonlite::fromJSON(toJSON(locationdataContent))

# Convert list to dataframe containing lists
yelpsixth<- data.frame(locationdataList)

# Third set of data
yelpurl7 <- paste0("https://api.yelp.com/v2/search/?term=reservation&location=Chicago,IL&offset=70&limi

# To get the data
locationdata=GET(yelpurl7, sig)
locationdataContent = content(locationdata)

# Convert JSON to list
locationdataList=jsonlite::fromJSON(toJSON(locationdataContent))

# Convert list to dataframe containing lists
yelpseventh<- data.frame(locationdataList)

# Fourth set of data
yelpurl8 <- paste0("https://api.yelp.com/v2/search/?term=reservation&location=Chicago,IL&offset=90&limi

# To get the data
locationdata=GET(yelpurl8, sig)
locationdataContent = content(locationdata)

# Convert JSON to list
locationdataList=jsonlite::fromJSON(toJSON(locationdataContent))

# Convert list to dataframe containing lists
yelpeigth<- data.frame(locationdataList)
```

Scrubbing the data(indian restaurant)

After getting data of indian restaurant, unneccessary data should be deleted and keep the final one which consist of useful data

```
# Deleting unneccessary cloumns from each set of data
yelpfifth1<- yelpfifth[-c(1:10, 12:24)]
yelpfifth20<- yelpfifth1[-c(2:3)]

yelpsixth1 <- yelpsixth[-c(1:10, 12:25)]
yelpsixth20 <- yelpsixth1[-c(2:3)]

yelpseventh1 <- yelpseventh[-c(1:10, 12:22)]
yelpseventh20 <- yelpseventh1[-c(1, 3:7)]

yelpeigth1 <- yelpeigth[-c(1:10, 12:25)]
yelpeigth20 <- yelpeigth1[-c(2, 3, 5)]


# Combining all the data set into single one
Yelpfinal2 <- rbind(yelpfifth20, yelpsixth20, yelpseventh20, yelpeigth20)
```

```
# Deleting unnccessary rows from the final data set
Yelpfinal2 <- Yelpfinal2[-c(1:9, 11:36, 38, 39, 41:45, 47:61, 63:69, 71:78),]

# inserting new column with name index
Yelpfinal2$index <- seq.int(nrow(Yelpfinal2))

# Chnage the values of rows to same in the index column
Yelpfinal2$index <- "indian"
```

Exploring the data

str(Yelpfinal1) 'data.frame': 7 obs. of 3 variables: $ businesses.name :List of 7

str(Yelpfinal2) 'data.frame': 6 obs. of 3 variables: $ businesses.name :List of 6

class(Yelpfinal1) [1] "data.frame"

class(Yelpfinal2) [1] "data.frame"

Combining data set of chinese and indian restaurant to one

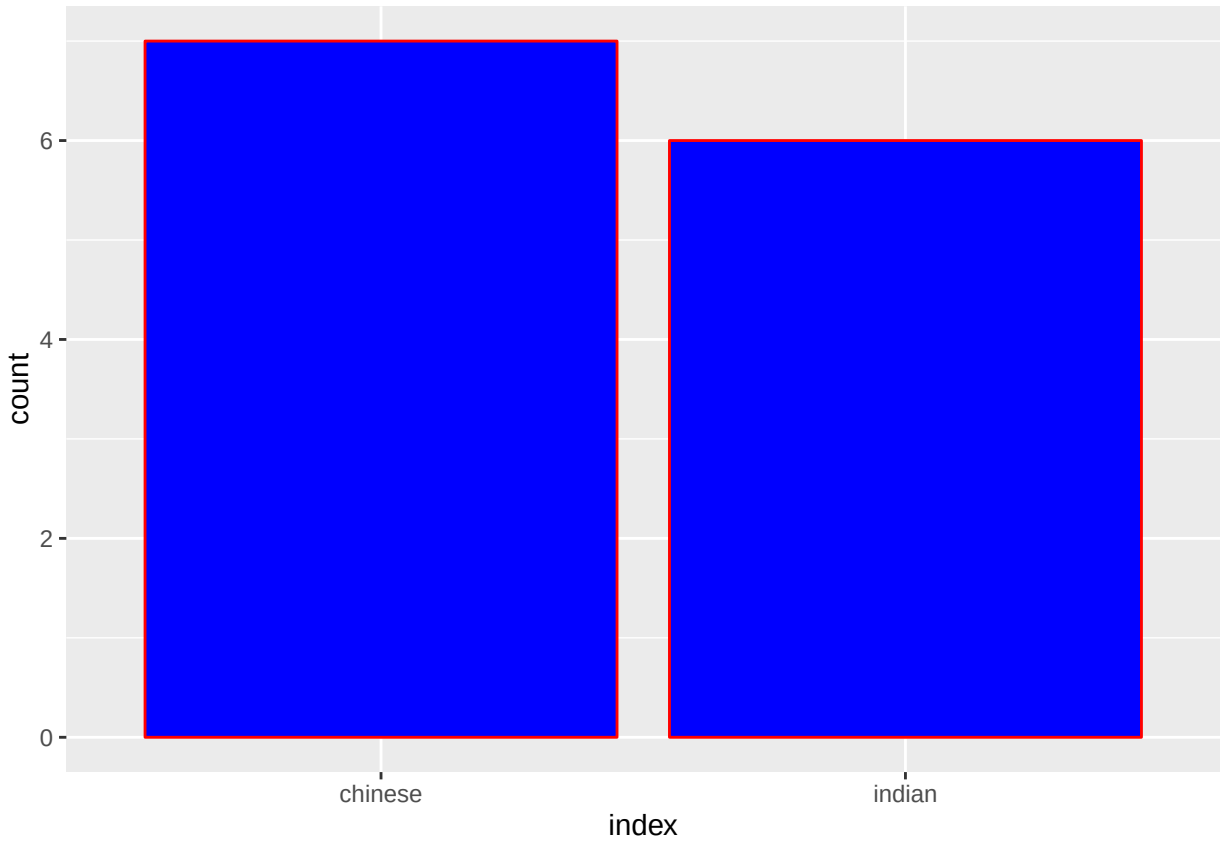
The final data from chinese restaurant and indian restaurant is combined into a variable named "Result" for ploting graphs

```
Result<- rbind(Yelpfinal1,Yelpfinal2)
```

Ploting of different graphs

Bar Graph

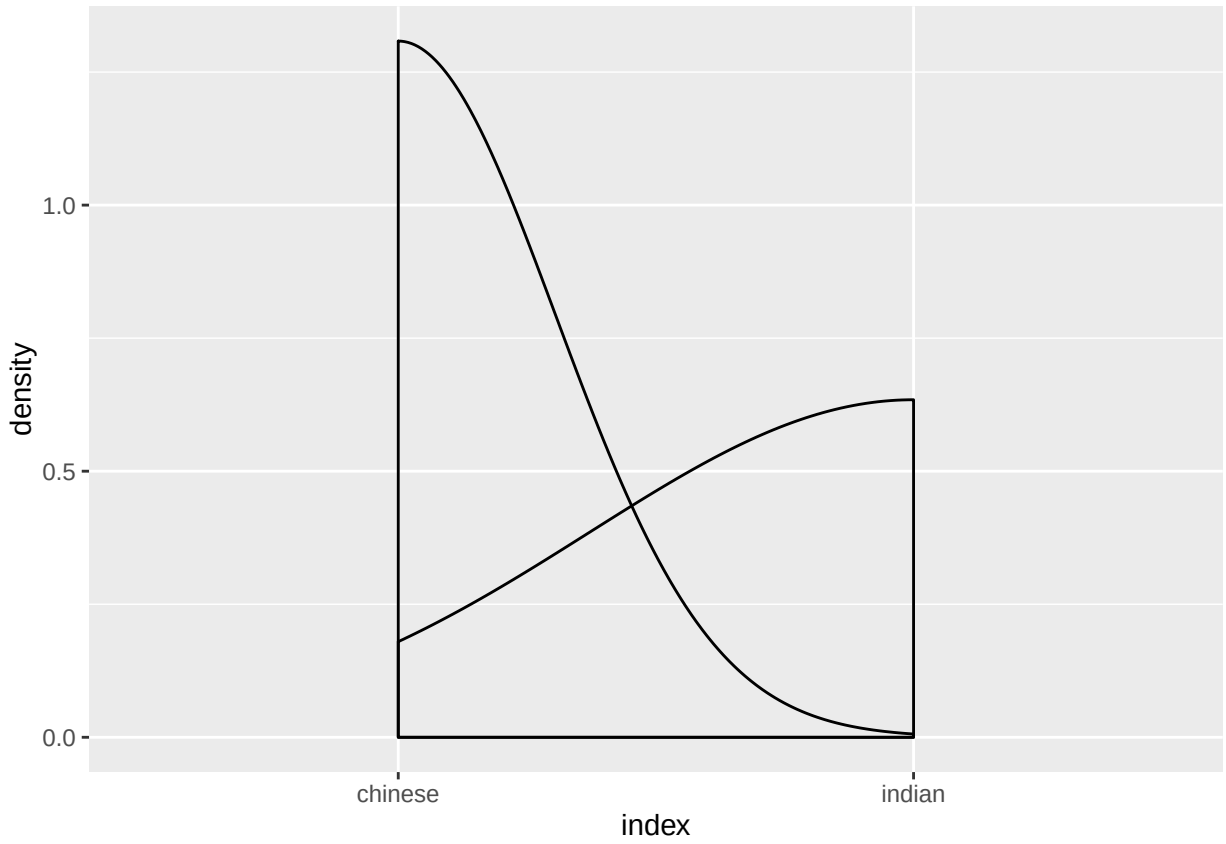
```
ggplot(Result, aes(x=index))+geom_bar(colour= "red",fill= "blue")
```



The graphshows that chinese restaurant has more reservation than the indian restaurant by 1.

Density Graph

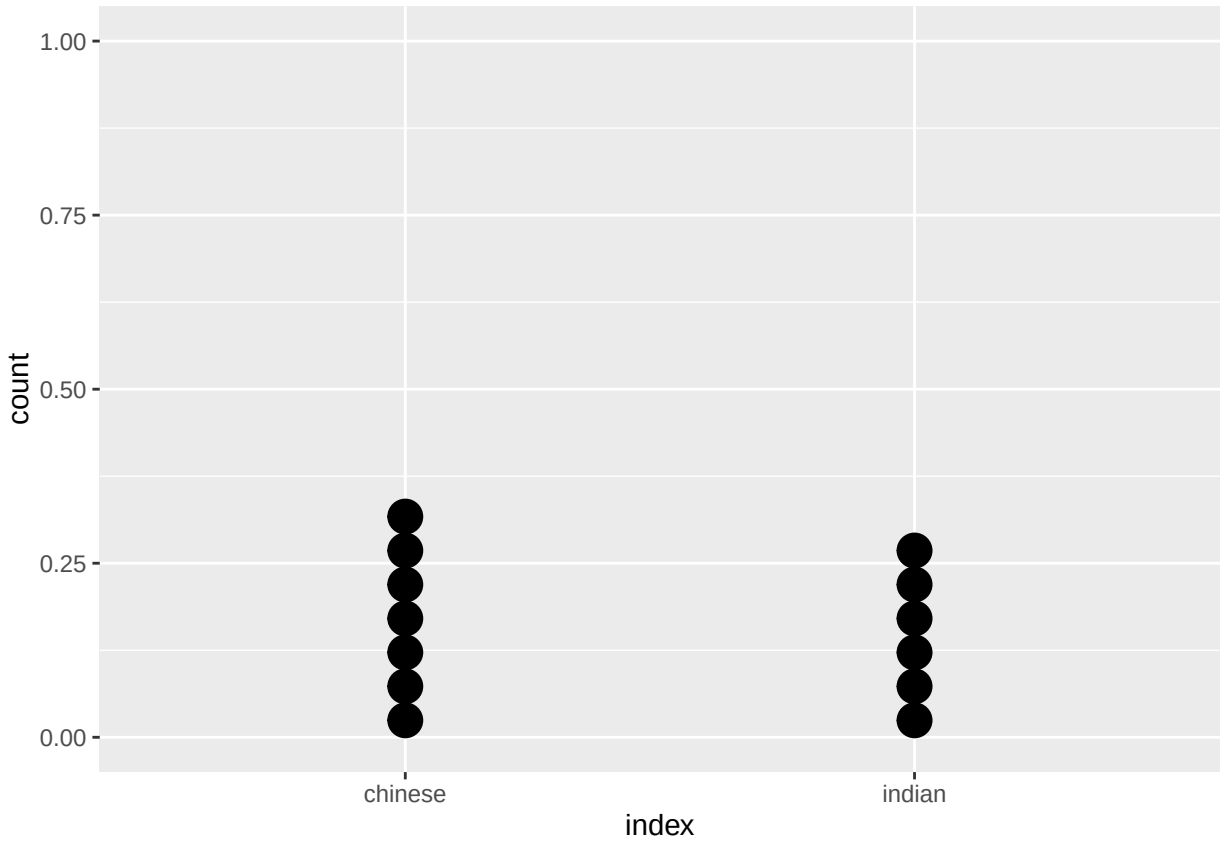
```
ggplot(Result, aes(x=index))+geom_density(adjust= 0.5)
```



Dot Graph

```
ggplot(Result, aes(x=index))+geom_dotplot(dotsize = 2.0)
```

```
## `stat_bindot()` using `bins = 30`. Pick better value with `binwidth`.
```



Chinese restaurant has value 1 more than the indian restaurant.