Plot 2 Way ANOVA in R

Tidy Data

Notes:

(1) This PDF is part of YouTube tutorial (https://youtu.be/HYUOF0oSwKc). This PDF is for individual, personal usage only.
(2) The author accepts no responsibility for the topicality, correctness, completeness or quality of the information provided.

```
# download the data from Github
df<- read.csv("https://raw.githubusercontent.com/TidyPython/interactions/main/city_brand_sales.csv")

# print out the data
print(df)
```

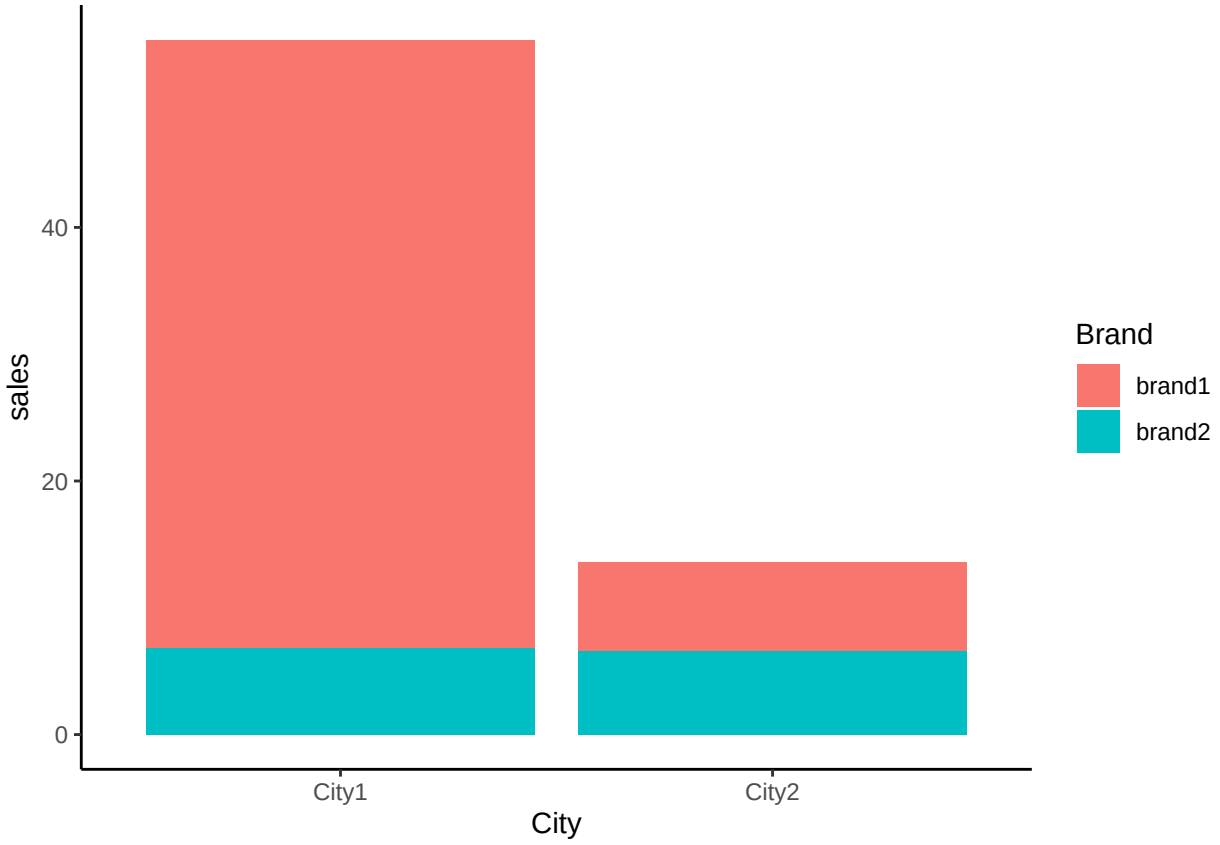
```
##     City  Brand sales
## 1  City1 brand1    70
## 2  City1 brand2    10
## 3  City1 brand1   100
## 4  City1 brand2     2
## 5  City1 brand1    30
## 6  City1 brand2     2
## 7  City1 brand1    20
## 8  City1 brand2    10
## 9  City1 brand1    20
## 10 City1 brand2    10
## 11 City2 brand1     9
## 12 City2 brand2    10
## 13 City2 brand1     5
## 14 City2 brand2     4
## 15 City2 brand1     4
## 16 City2 brand2     4
## 17 City2 brand1     5
## 18 City2 brand2     4
## 19 City2 brand1    12
## 20 City2 brand2    11
```

```
car::Anova(lm(sales ~ City*Brand, data = df),type=3)
```

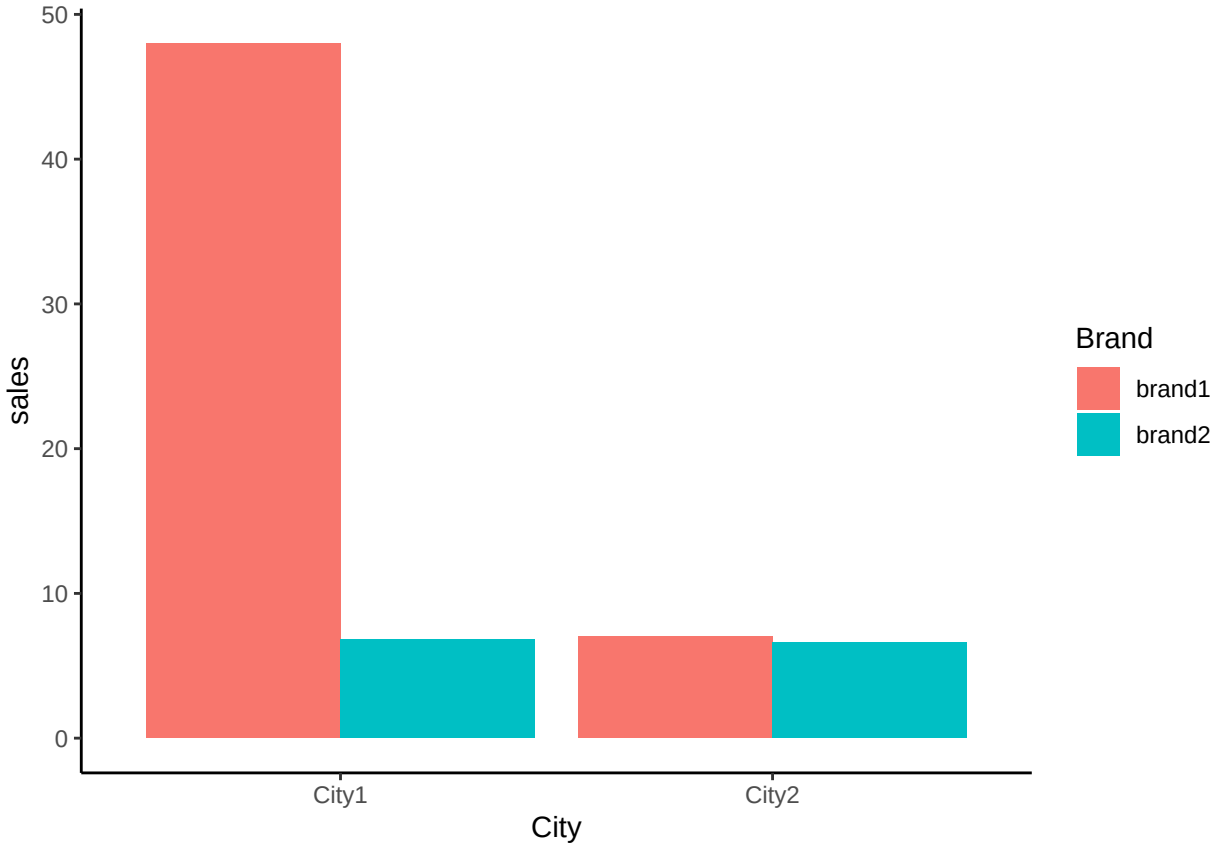
```
## Anova Table (Type III tests)
##
## Response: sales
##              Sum Sq Df F value    Pr(>F)
## (Intercept) 11520.0  1 35.0818 2.143e-05 ***
```

```
## City         4202.5  1 12.7979  0.002516 **
## Brand        4243.6  1 12.9230  0.002425 **
## City:Brand   2080.8  1  6.3367  0.022865 *
## Residuals    5254.0 16
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

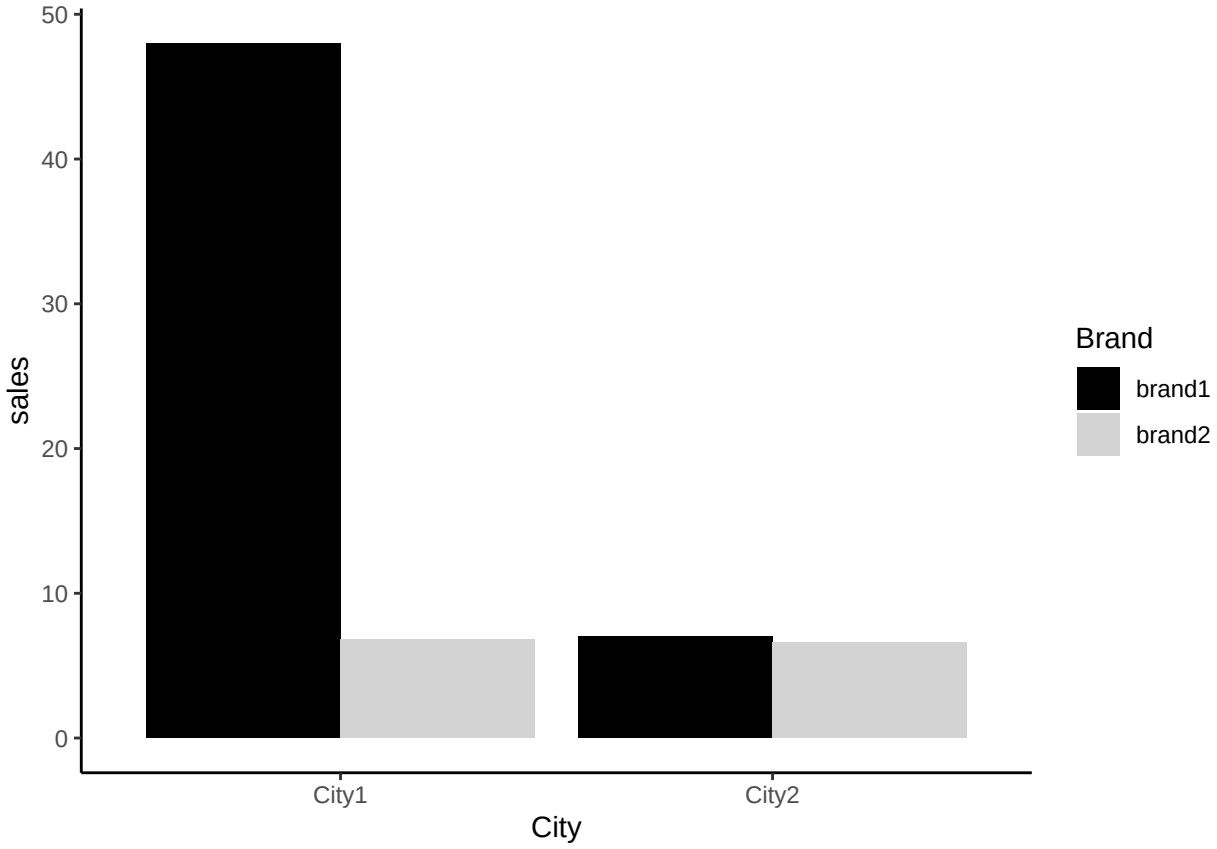
```
library(ggplot2)
# bar chart
ggplot(data=df,aes(x = City, y = sales, fill =Brand)) +
  geom_bar(stat = "summary", fun = mean)+ theme_classic()
```



```r
# Change bar position
ggplot(data=df,aes(x = City, y = sales, fill =Brand)) +
  geom_bar(stat = "summary",fun = mean, position = "dodge")+ theme_classic()
```



```r
# Change color
ggplot(data=df,aes(x = City, y = sales, fill =Brand)) +
  geom_bar(stat = "summary", fun = mean, position = "dodge")+
  theme_classic()+ scale_fill_manual(values=c('black','lightgray'))
```



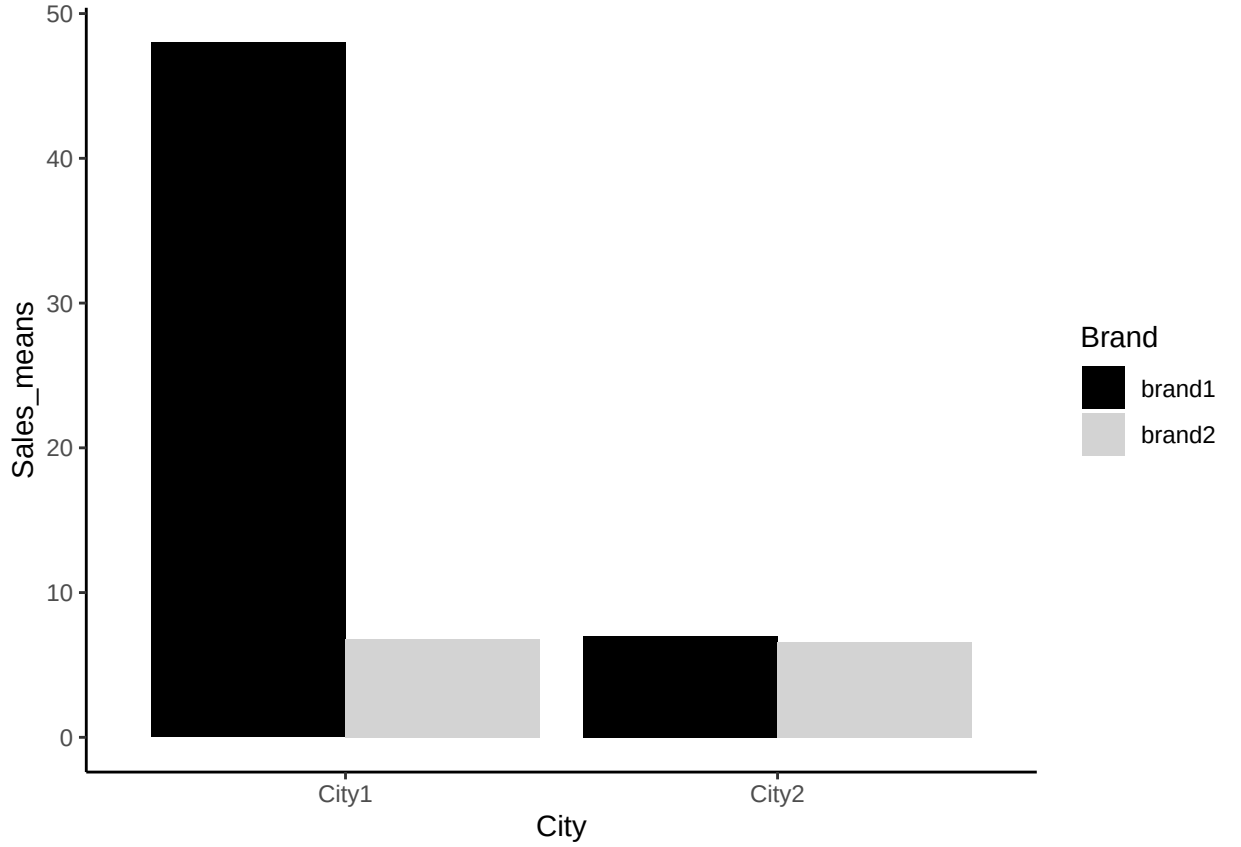
```
library(dplyr)

# use dplyr to calculate means grouped by City and Brand
means=df%>% group_by(City,Brand) %>% summarise_at(vars(sales), list(Sales_means = mean))

# print out the table of means
print(means)
```

```
## # A tibble: 4 x 3
## # Groups:   City [2]
##   City  Brand  Sales_means
##   <chr> <chr>        <dbl>
## 1 City1 brand1        48
## 2 City1 brand2         6.8
## 3 City2 brand1         7
## 4 City2 brand2         6.6
```

```
ggplot(data=means,aes(x = City, y = Sales_means, fill =Brand)) +
  geom_bar(stat = "identity", position = "dodge")+
  theme_classic()+ scale_fill_manual(values=c('black','lightgray'))
```



```r
# line chart
ggplot(means,aes(x = City, y = Sales_means, colour = Brand, group = Brand)) +
  geom_point(size = 4) + geom_line()+theme_classic()
```