ANOVA_Example

Arif_Furkan

2024-06-30

Loading Libraries and Data

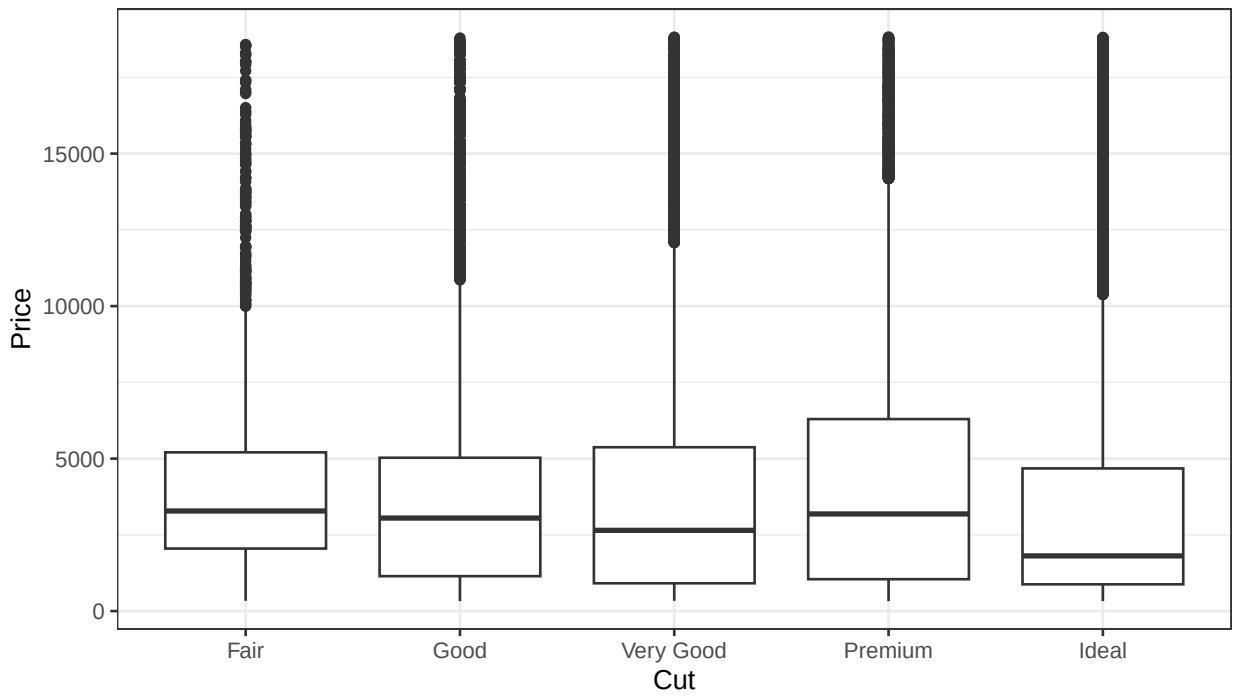
```
require(ggplot2)
```

```
## Zorunlu paket yükleniyor: ggplot2
```

```
data(diamonds)
```

Creating a Boxplot

```
ggplot(data = diamonds) + geom_boxplot(aes(cut,price))+
theme_bw() + xlab("Cut") + ylab("Price")
```



One-Way ANOVA

```
anova.1 <- aov(price~cut, data = diamonds)
summary(anova.1)
```

```
##                Df    Sum Sq   Mean Sq F value Pr(>F)
## cut             4 1.104e+10 2.760e+09   175.7 <2e-16 ***
## Residuals   53935 8.474e+11 1.571e+07
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print(anova.1)
```

```
## Call:
##    aov(formula = price ~ cut, data = diamonds)
##
## Terms:
##                         cut    Residuals
## Sum of Squares   11041745359 847431390159
## Deg. of Freedom            4        53935
##
## Residual standard error: 3963.847
## Estimated effects may be unbalanced
```

Fitting a Linear Model

```
fit <- lm(price~cut, data = diamonds)
anova.2 <- anova(fit)
anova.2
```

```
## Analysis of Variance Table
##
## Response: price
##              Df     Sum Sq    Mean Sq F value    Pr(>F)
## cut           4 1.1042e+10 2760436340  175.69 < 2.2e-16 ***
## Residuals 53935 8.4743e+11   15712087
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Another ANOVA using the car package

```
require(car)
```

```
## Zorunlu paket yükleniyor: car
```

```
## Warning: package 'car' was built under R version 4.4.1
```

```
## Zorunlu paket yükleniyor: carData
```

```
fit <- lm(price ~ cut, data = diamonds)
anova.3 <- Anova(fit)
require(car)
anova.3 <- Anova(fit)
anova.3
```

```
## Anova Table (Type II tests)
##
## Response: price
##               Sum Sq    Df F value    Pr(>F)
## cut       1.1042e+10     4  175.69 < 2.2e-16 ***
## Residuals 8.4743e+11 53935
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pairwise T-Tests

Performs pairwise t-tests to compare the means of diamond prices between different cuts.

```
pairwise.t.test(diamonds$price, diamonds$cut, p.adjust.method = "bonferroni")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  diamonds$price and diamonds$cut
##
##           Fair    Good    Very Good Premium
## Good      0.0016  -       -         -
## Very Good 0.0034  1.0000  -         -
## Premium   0.3077  < 2e-16 < 2e-16   -
## Ideal     < 2e-16 5.7e-13 < 2e-16   < 2e-16
##
## P value adjustment method: bonferroni
```

From the results, it can be seen that the Fair-Premium and Good-Very Good groups are not statistically different from each other, but there is a statistically significant difference in the others.

Tukey's Honest Significant Difference Test

```
TukeyHSD(anova.1, ordered = FALSE)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = price ~ cut, data = diamonds)
##
## $cut
##                          diff         lwr        upr     p adj
## Good-Fair          -429.89331  -740.44880  -119.3378 0.0014980
## Very Good-Fair     -376.99787  -663.86215   -90.1336 0.0031094
## Premium-Fair        225.49994   -59.26664   510.2665 0.1950425
## Ideal-Fair         -901.21579 -1180.57139  -621.8602 0.0000000
## Very Good-Good       52.89544  -130.15186   235.9427 0.9341158
## Premium-Good        655.39325   475.65120   835.1353 0.0000000
## Ideal-Good         -471.32248  -642.36268  -300.2823 0.0000000
## Premium-Very Good   602.49781   467.76249   737.2331 0.0000000
## Ideal-Very Good    -524.21792  -647.10467  -401.3312 0.0000000
## Ideal-Premium     -1126.71573 -1244.62267 -1008.8088 0.0000000
```

As can be seen, consistent with the above result, there is no statistically significant difference between Fai -Premium and Good-Very Good, but there is a difference between the other groups.

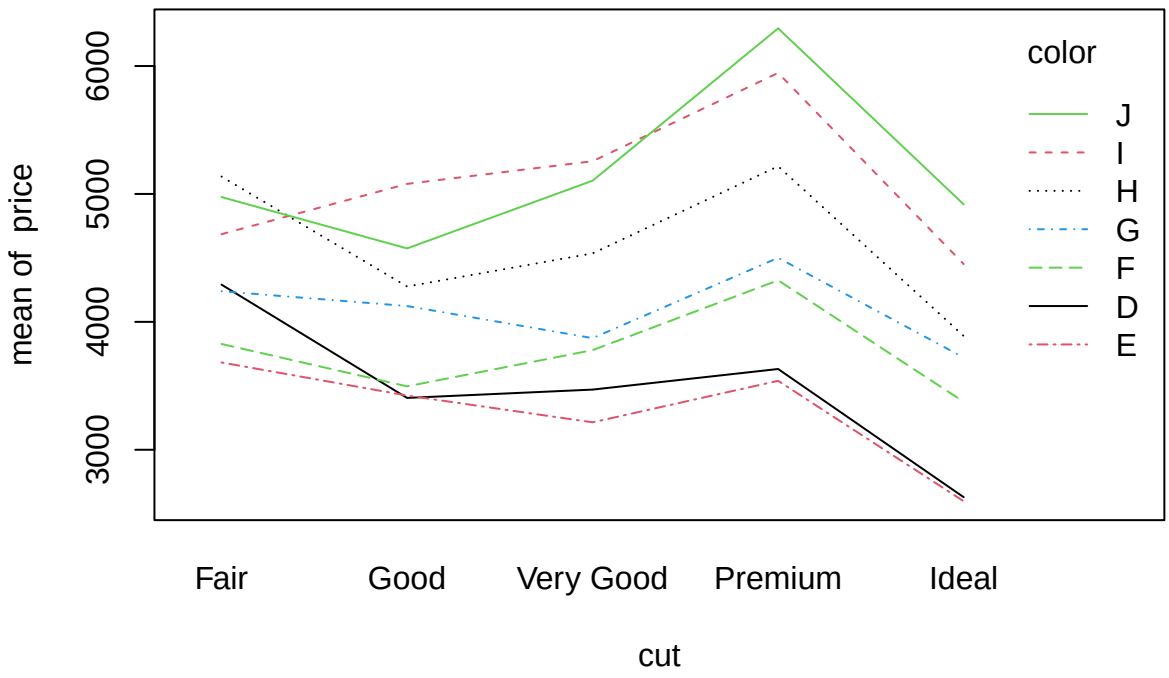
Two-Way ANOVA

```
fit.2 <- lm(price ~ cut*color, data = diamonds)
library(car)
anova.4 <- Anova(fit.2, type = "III")
anova.4
```

```
## Anova Table (Type III tests)
##
## Response: price
##                 Sum Sq    Df    F value    Pr(>F)
## (Intercept) 3.7606e+11     1 24713.0009 < 2.2e-16 ***
## cut         8.7891e+09     4   144.3969 < 2.2e-16 ***
## color       9.4602e+09     6   103.6142 < 2.2e-16 ***
## cut:color   1.6535e+09    24     4.5274 1.001e-12 ***
## Residuals   8.2027e+11 53905
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction Plot

```
with(diamonds,{
interaction.plot(cut, color, price, col = c(1:4))})
```



While the price change is the same in most colors, there is no price decrease in the good cuts in Colors I and G. Apart from this, the price change generally occurs in the same way.