

HEALTH CARE COST ANALYSIS

CODES:-

```
rm(list=ls())

hops = read.csv("HospitalCosts.csv")

head(hops)

colSums(is.na(hops))

hops = na.omit(hops)

summary(hops)

str(hops)

hops$RACE = as.factor(hops$RACE)

hops$FEMALE = as.factor(hops$FEMALE)

hist(hops$AGE)

summary(as.factor(hops$AGE))

library(dplyr)

df = summarise(group_by(hops,AGE), TOTCHG = sum(TOTCHG))

df

arrange(df, desc(TOTCHG))

arrange(df, desc(TOTCHG))[1,]

df1 <- hops %>% group_by(AGE) %>% summarise(TOTCHG = sum(TOTCHG)) %>%
arrange(desc(TOTCHG))

df1[1,]

hist (hops$APRDRG)

summary(hops$APRDRG)

summary(as.factor(hops$APRDRG))

df = summarise(group_by(hops,APRDRG),TOTCHG = sum(TOTCHG))

df

arrange(df,desc(TOTCHG))

arrange(df, desc(TOTCHG))[1]

df1 = hops %>% group_by(APRDRG) %>% summarise(TOTCHG = sum(TOTCHG)) %>%
arrange(desc(TOTCHG))
```

```
df1[1,]  
str(hops$RACE)  
str(hops$TOTCHG)  
model = aov(TOTCHG ~ RACE, data = hops)  
summary(model)  
alpha = 0.05  
pvalue = summary(av)[[1]][,"Pr(>F)"][1]  
pvalue = 0.943
```

```
pvalue < alpha
```

```
av = aov(TOTCHG ~ FEMALE + AGE, data = hops)  
alpha = 0.05  
pvalue = summary(av)[[1]][,"Pr(>F)"][1]  
pvalue = 0.943  
pvalue < alpha  
Model = lm(LOS ~ AGE+ FEMALE + AGE + RACE, data = hops)  
summary(Model)  
summary(model)$coefficients[,4] < 0.05
```

```
model = lm(TOTCHG ~ ., hops)  
summary(model)
```

```
summary(model)$coefficients[,4] < 0.05
```

