Big Data Tools for Managers

Dept. of MBA, Siddaganga Institute of Technology-Tumkur

Working with Advertising dataset

Assume, We are the statistical consultant hired by a client to provide advice on improving their sales of particular product. We have Advertisement dataset consist of the sales of that product in 200 different markets , along with advertising budgets for each market like : TV ,Radio and Newspaper.

Is it possible for our client, to directly increase sales of their product on any of the market ?; On other hand , Is there any way to control the advertising expenditure for media ?

Download Advertising data

Note: This link take you in another page and then click on File menu -> Save Page As -> Open file window and click on Save

- **X**— Sr. Number
- **TV**— Budget for the Advertising in TV channel
- **Radio**— Budget for the Advertising in Radio channel
- **Newspaper**— Budget for the Advertising in Newspaper
- **Sales**— Dependent varible Expected Sales from the Market by using Independent Variables(TV, Radio, Newspaper)

Write R Code for below questions.

1. Read Advertising dataset in R
2. Get the dimension of Advertising dataset
3. Display column names of dataset
4. View data in Excel like screen
5. Get the quick summary of dataset
6. Look the Correlation Sales with all the variables (TV, Newspaper, Radio)
7. Prepare Simple Linear Regression Model & Display Summary of Model
8. Compute Sum of Square Error for Simple Regression Model
9. Prepare Multiple Linear Regression Model(Adding indepent variable into Model)
10. Compute Sum of Square Error for Multiple Regression Model
11. Prediction on new data
12. Prepare Formula

```r
# Read Advertising dataset in R
data <- read.csv("Advertising.csv")
```

```r
# Get the dimension of Advertising dataset
dim(data)
```

```r
# Display column names of dataset
colnames(data)
```

```r
# View data in Excel like screen
View(data)
```

```r
# Get the quick summary of dataset
summary(data)
```

```r
# Look the Correlation Sales with all the variables (TV, Newspaper, Radio)
cor(data$Radio, data$Sales)
cor(data$TV, data$Sales)
```

```r
plot(data$TV,data$Sales)
```

```r
# Prepare Simple Linear Regression Model & Display Summary of Model
model_1 <- lm(Sales ~ Radio, data)
summary(model_1)
```

```r
# Compute Sum of Square Error for Simple Regression Model
SSE_model_1<- sum(model_1$residuals^2)
print(SSE_model_1)
```

```r
# Prepare Multiple Linear Regression Model(Adding indepent variable into Model)
model_2 <- lm(Sales~Radio+TV, data)
summary(model_2)
```

```r
#Compute SSE for model2
SSE_model_2 <- sum(model_2$residuals^2)
print(SSE_model_2)
```

```r
# Prediction on new data
Radio <- c(1000)
TV <- c(1000)
new_data <- data.frame(Radio,TV)

predict(model_2, new_data)
```

```r
#Prepare Formula
b0 <- model_2$coefficients['(Intercept)']
b1 <- model_2$coefficients['Radio']
b2 <- model_2$coefficients['TV']
```

```r
Radio <- 200
TV <- 200

b0+(b1*Radio)+(b2*TV)
```