EDA ON WHEAT DATA CODE

```r
#Importing the libraries for the analysis
library(dplyr)
library(lattice)
library(markdown)

#Importing the dataset
data=read.csv("Wheat_2023 DATA.csv")

#Internal structure of the data
str(data)

#Statistical summary of the data
summary(data)

#Finding the NA values in the dataset
which(is.na(data$min_price))
which(is.na(data$max_price))
which(is.na(data$modal_price))

#Removing the NA values
data=na.omit(data)
summary(data)

#date type conversion
data["arrival_date"]=as.Date(data$arrival_date, format="%d/%m/%Y")
str(data)

#getting the unique data of the categorical attributes
unique(data["state"])

#Filtering first 9 days of data 1to 9
subset1=subset(data,arrival_date < "2023-01-10")

#second 9 days 10 to 18
subset2=subset(data, arrival_date > "2023-01-09" & arrival_date < "2023-01-19")

#Third 9 days 19 to 27
subset3=subset(data,arrival_date>"2023-01-18")
summary(subset3)
#Histogram for minimum price in subset 1
histogram(~min_price,main="MINIMUM PRICE IN SUBSET 1",data=subset1,breaks=415)
#Histogram for maximum price in subset 1
histogram(~max_price,main="MAXIMUM PRICE IN SUBSET 1",data=subset1,breaks=542)

#Histogram for minimum price in subset 2
histogram(~min_price,main="MINIMUM PRICE IN SUBSET 2",data=subset2,breaks=418)
#Histogram for maximum price in subset 2
histogram(~max_price,main="MAXIMUM PRICE IN SUBSET 2",data=subset2,breaks=540)

#Histogram for minimum price in subset 3
histogram(~min_price,main="MINIMUM PRICE IN SUBSET 3",data=subset3,breaks=381)
#Histogram for maximum price in subset 3
histogram(~max_price,main="MAXIMUM PRICE IN SUBSET 3",data=subset3,breaks=460)

#BoxPlot for min_price on the subset 1
bwplot(arrival_date~min_price, data = subset1, main = "Box Plot of min_price by arrival_date",
     xlab = "min_price", ylab = "arrival_date")
#BoxPlot for max_price on the subset 1
```

```r
bwplot(arrival_date~max_price, data = subset1, main = "Box Plot of max_price by arrival_date",
    xlab = "max_price", ylab = "arrival_date")

#BoxPlot for min_price on the subset 2
bwplot(arrival_date~min_price, data = subset2, main = "Box Plot of min_price by arrival_date",
    xlab = "min_price", ylab = "arrival_date")
#BoxPlot for max_price on the subset 2
bwplot(arrival_date~max_price, data = subset2, main = "Box Plot of max_price by arrival_date",
    xlab = "max_price", ylab = "arrival_date")

#BoxPlot for min_price on the subset 3
bwplot(arrival_date~min_price, data = subset3, main = "Box Plot of min_price by arrival_date",
    xlab = "min_price", ylab = "arrival_date")
#BoxPlot for max_price on the subset 3
bwplot(arrival_date~max_price, data = subset3, main = "Box Plot of max_price by arrival_date",
    xlab = "max_price", ylab = "arrival_date")


#scatter plot to find the correlation  between min_price and max_price with modal_price in the subset1
cor(subset1$min_price,subset1$modal_price)
xyplot(min_price ~ modal_price, data = subset1,main="correlation between min_price and modal_price")

cor(subset1$max_price,subset1$modal_price)
xyplot(max_price ~ modal_price, data = subset1,main="correlation between max_price and modal_price")


#scatter plot to find the correlation  between min_price and max_price with modal_price in the subset2
cor(subset2$min_price,subset2$modal_price)
xyplot(min_price ~ modal_price, data = subset2,main="correlation between min_price and modal_price")

cor(subset2$max_price,subset2$modal_price)
xyplot(max_price ~ modal_price, data = subset2,main="correlation between max_price and modal_price")


#scatter plot to find the correlation  between min_price and max_price with modal_price in the subset3
cor(subset3$min_price,subset3$modal_price)
xyplot(min_price ~ modal_price, data = subset3,main="correlation between min_price and modal_price")

cor(subset3$max_price,subset3$modal_price)
xyplot(max_price ~ modal_price, data = subset3,main="correlation between max_price and modal_price")

# Barchart to find the highest frequent state in the dataset
barchart(data["state"],main = "STATES WITH MAXIMUM FREQUENCY",
    xlab = "Occurance",
    ylab = "States")

#Most frequent state
mp=filter(data,state=="Madhya Pradesh")

#scatter plot for correlation between min_price and max_price with modal_price on the MadhyaPradesh
cor(mp$min_price,mp$modal_price)
xyplot(min_price ~ modal_price, data = mp,main="correlation between min_price and modal_price")

cor(mp$max_price,mp$modal_price)
xyplot(max_price ~ modal_price, data = mp,main="correlation between max_price and modal_price")


#scatter plot for whole country correlation analysis of min_price and max_price with modal_price
cor(data$min_price,data$modal_price)
xyplot(min_price ~ modal_price, data = data,main="correlation between min_price and modal_price")

cor(data$max_price,data$modal_price)
```

```
xyplot(max_price ~ modal_price, data = data,main="correlation between max_price and modal_price")
```