Untitled

2024-01-08

df\_conc=read.csv('Concrete\_Data1.csv')  
head(df\_conc)str(df\_conc)

## 'data.frame': 1030 obs. of 9 variables:  
## $ Cement : num 540 540 332 332 199 ...  
## $ Blast.Furnace.Slag: num 0 0 142 142 132 ...  
## $ Fly.Ash : num 0 0 0 0 0 0 0 0 0 0 ...  
## $ Water : num 162 162 228 228 192 228 228 228 228 228 ...  
## $ Superplasticizer : num 2.5 2.5 0 0 0 0 0 0 0 0 ...  
## $ Coarse.Aggregate : num 1040 1055 932 932 978 ...  
## $ Fine.Aggregate : num 676 676 594 594 826 ...  
## $ Age.days. : num 28 28 270 365 360 90 365 28 28 28 ...  
## $ Concrete.strength : num 80 61.9 40.3 41 44.3 ...

lm\_model <- lm(Concrete.strength ~ ., data = df\_conc)  
summary(lm\_model)

## Multiple R-squared: 0.6155, Adjusted R-squared: 0.6125   
## F-statistic: 204.3 on 8 and 1021 DF, p-value: < 2.2e-16

plot(lm\_model)

boxplot(df\_conc$Concrete.strength ~ df\_conc$Age.days., xlab = "Age in days", ylab = "Concrete Strength", main = "Concrete Strength vs. Age")

pairs(df\_conc)

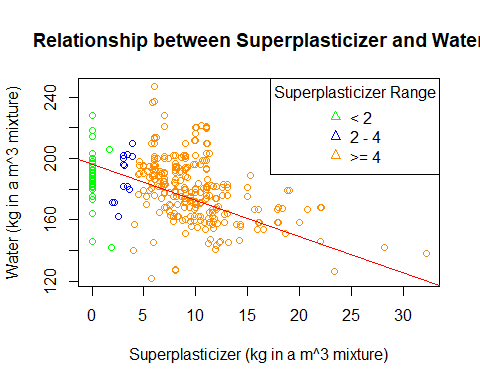
library(ggplot2)  
library(scales)  
  
# Set the range for the x-axis (Age.days.) and the break point  
range\_days <- c(0, 400)  
break\_point <- 50  
  
ggplot(df\_conc, aes(x = Age.days., y = Cement, color = Concrete.strength)) +  
 geom\_point(size = 3) + # Use point size for visibility  
 scale\_x\_continuous(limits = range\_days, breaks = seq(0, range\_days[2], by = break\_point)) +  
 scale\_color\_gradient(low = "tan", high = "darkred", guide = "legend") +  
 labs(x = "Age (days)", y = "Cement (kg in a m^3 mixture)", color = "Concrete Strength (Mpa)") +  
 theme\_minimal()

library(ggplot2)  
# Set the range for the x-axis (Age.days.) and the break point  
range\_days <- c(0, 400)  
break\_point <- 50  
ggplot(df\_conc, aes(x = Cement, y = Concrete.strength, color = Water, size = Water)) +  
 geom\_point() +  
 labs(x = "Cement (kg in a m^3 mixture)", y = "Concrete Strength (Mpa megapascals)", color = "Water", size = "Water") +  
 scale\_size\_continuous(guide = FALSE) + # Remove size legend  
 guides(color = guide\_legend(title = "Water (kg in a m^3 mixture) ")) + # Set a single legend for color  
 theme\_minimal()

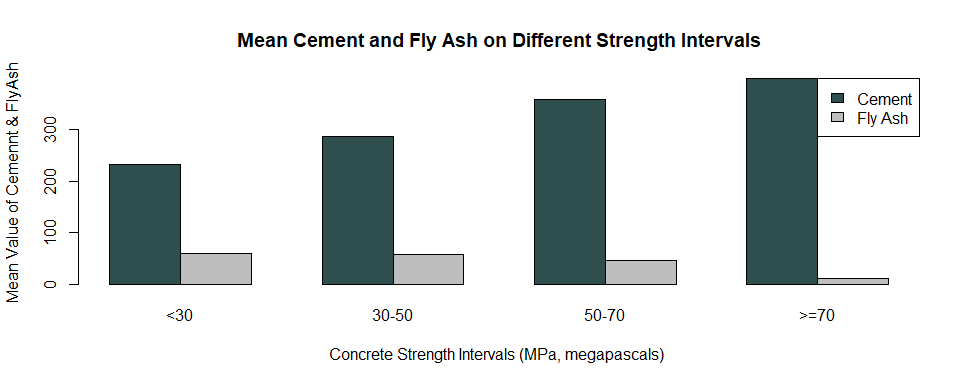
library(ggplot2)  
  
ggplot(df\_conc, aes(x = Cement, y = Concrete.strength, color = Fly.Ash)) +  
 geom\_point() +  
 scale\_color\_gradient(low = "lightgrey", high = "darkslategrey", name = "Fly Ash") +  
 scale\_size\_continuous(range = c(2, 8), name = "Fly Ash") +  
 labs(x = "Cement (kg in a m^3 mixture)", y = "Concrete Strength (MPa, megapascals) ") +  
 theme\_minimal()

# Define intervals for Concrete Strength  
df\_conc$Strength\_Group <- cut(df\_conc$Concrete.strength, breaks = c(0, 30, 50, 70, max(df\_conc$Concrete.strength)), labels = c("<30", "30-50", "50-70", ">=70"))  
  
# Calculate mean Cement and Fly Ash for each interval of Concrete Strength  
aggregate\_cement\_flyash <- aggregate(cbind(Cement, Fly.Ash) ~ Strength\_Group, data = df\_conc, FUN = mean)  
  
# Create a bar plot for Cement and Fly Ash on different intervals of Concrete Strength  
barplot(t(aggregate\_cement\_flyash[, c("Cement", "Fly.Ash")]), beside = TRUE,   
 col = c("darkslategrey", "gray"),   
 names.arg = aggregate\_cement\_flyash$Strength\_Group,   
 main = "Mean Cement and Fly Ash on Different Strength Intervals",  
 xlab = "Concrete Strength Intervals (MPa, megapascals) ", ylab = "Mean Value of Cemennt & FlyAsh")  
  
legend("topright", legend = c("Cement", "Fly Ash"), fill = c("darkslategrey", "gray"))

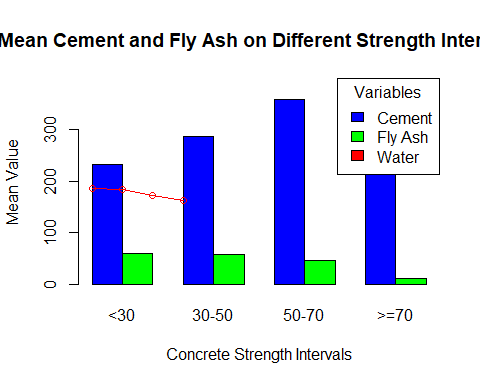
# Define colors based on Superplasticizer values  
colors <- ifelse(df\_conc$Superplasticizer < 2, "green",   
 ifelse(df\_conc$Superplasticizer >= 2 & df\_conc$Superplasticizer < 4, "blue", "darkorange"))  
  
# Scatter plot of Superplasticizer against Water with colored points  
plot(df\_conc$Superplasticizer, df\_conc$Water,   
 xlab = "Superplasticizer (kg in a m^3 mixture)", ylab = "Water (kg in a m^3 mixture)",  
 main = "Relationship between Superplasticizer and Water",  
 col = colors)  
  
# Adding a trend line  
abline(lm(df\_conc$Water ~ df\_conc$Superplasticizer), col = "red")  
  
# Adding a legend  
legend("topright", legend = c("< 2", "2 - 4", ">= 4"), col = c("green", "blue", "darkorange"), pch = 2,  
 title = "Superplasticizer Range")



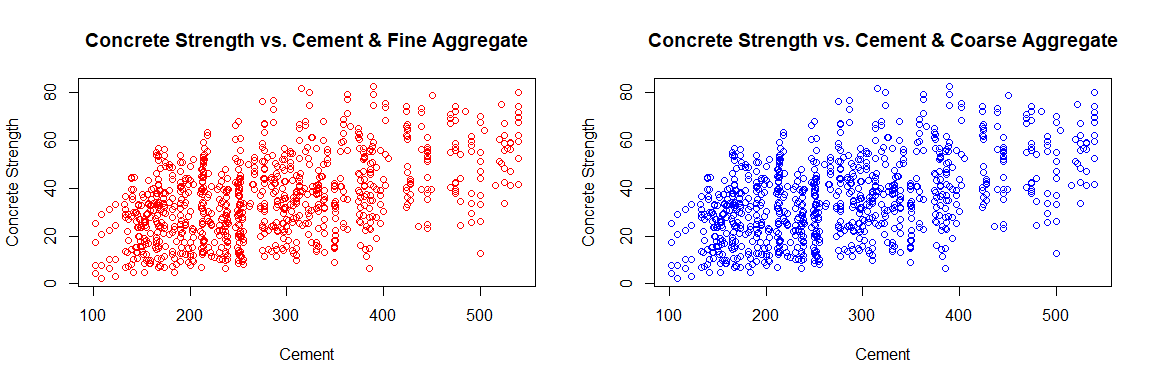
# Define intervals for Concrete Strength  
df\_conc$Strength\_Group <- cut(df\_conc$Concrete.strength, breaks = c(0, 30, 50, 70, max(df\_conc$Concrete.strength)), labels = c("<30", "30-50", "50-70", ">=70"))  
  
# Calculate mean Cement and Fly Ash for each interval of Concrete Strength  
aggregate\_cement\_flyash <- aggregate(cbind(Cement, Fly.Ash) ~ Strength\_Group, data = df\_conc, FUN = mean)  
  
# Create a bar plot for Cement and Fly Ash on different intervals of Concrete Strength  
barplot(t(aggregate\_cement\_flyash[, c("Cement", "Fly.Ash")]), beside = TRUE,   
 col = c("darkslategrey", "gray"),   
 names.arg = aggregate\_cement\_flyash$Strength\_Group,   
 main = "Mean Cement and Fly Ash on Different Strength Intervals",  
 xlab = "Concrete Strength Intervals (MPa, megapascals) ", ylab = "Mean Value of Cemennt & FlyAsh")  
  
legend("topright", legend = c("Cement", "Fly Ash"), fill = c("darkslategrey", "gray"))



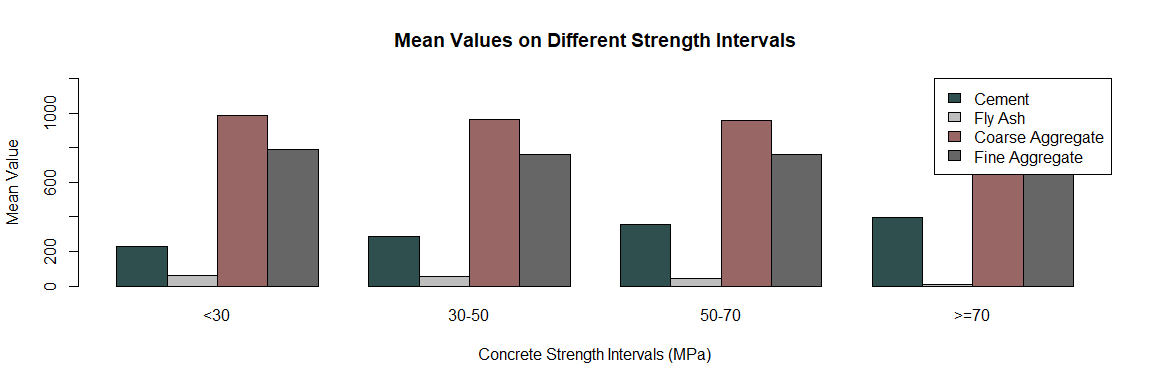
# Calculate mean Water for each interval of Concrete Strength  
aggregate\_water <- aggregate(Water ~ Strength\_Group, data = df\_conc, FUN = mean)  
  
# Create a bar plot for Cement and Fly Ash on different intervals of Concrete Strength  
barplot(t(aggregate\_cement\_flyash[, c("Cement", "Fly.Ash")]), beside = TRUE,   
 col = c("blue", "green"),   
 names.arg = aggregate\_cement\_flyash$Strength\_Group,   
 main = "Mean Cement and Fly Ash on Different Strength Intervals",  
 xlab = "Concrete Strength Intervals", ylab = "Mean Value")  
  
# Add a line for average Water  
lines(aggregate\_water$Strength\_Group, aggregate\_water$Water, type = "o", col = "red")  
legend("topright", legend = c("Cement", "Fly Ash", "Water"), fill = c("blue", "green", "red"), title = "Variables")



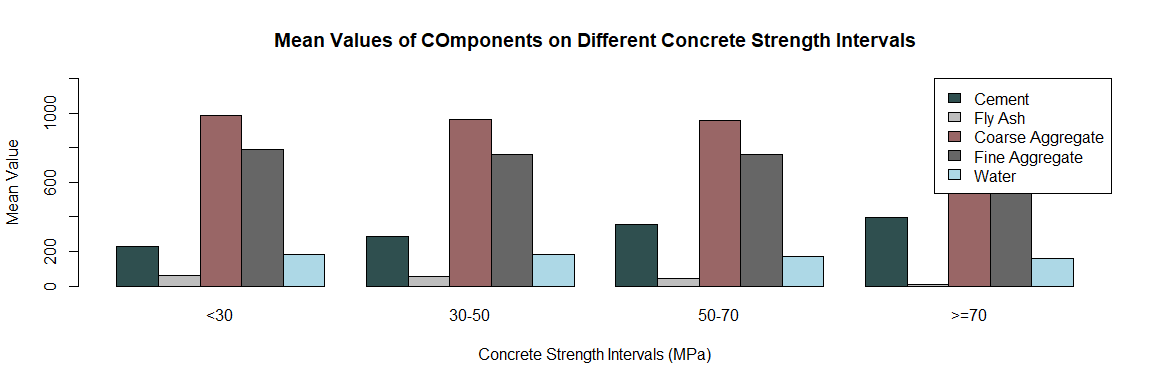
par(mfrow = c(1, 2)) # Set up a 1x2 grid for plots  
  
# Plot 1: Concrete Strength vs. Cement and Fine Aggregate  
plot(df\_conc$Cement, df\_conc$Concrete.strength, xlab = "Cement", ylab = "Concrete Strength",   
 main = "Concrete Strength vs. Cement & Fine Aggregate", col = "red")  
points(df\_conc$Fine.Aggregate,col = "red")  
  
# Plot 2: Concrete Strength vs. Cement and Coarse Aggregate  
plot(df\_conc$Cement, df\_conc$Concrete.strength, xlab = "Cement", ylab = "Concrete Strength",   
 main = "Concrete Strength vs. Cement & Coarse Aggregate", col = "blue")  
points(df\_conc$Coarse.Aggregate, df\_conc$Concrete.strength, col = "blue")



# Define intervals for Concrete Strength  
df\_conc$Strength\_Group <- cut(df\_conc$Concrete.strength, breaks = c(0, 30, 50, 70, max(df\_conc$Concrete.strength)), labels = c("<30", "30-50", "50-70", ">=70"))  
  
# Calculate mean Cement, Fly Ash, Coarse Aggregate, and Fine Aggregate for each interval of Concrete Strength  
aggregate\_values <- aggregate(cbind(Cement, Fly.Ash, Coarse.Aggregate, Fine.Aggregate) ~ Strength\_Group, data = df\_conc, FUN = mean)  
  
# Create a bar plot for Cement, Fly Ash, Coarse Aggregate, and Fine Aggregate on different intervals of Concrete Strength  
barplot(t(aggregate\_values[, c("Cement", "Fly.Ash", "Coarse.Aggregate", "Fine.Aggregate")]), beside = TRUE,   
 col = c("darkslategrey", "gray", "#996666", "#666666"),   
 names.arg = aggregate\_values$Strength\_Group,   
 main = "Mean Values on Different Strength Intervals",  
 xlab = "Concrete Strength Intervals (MPa)", ylab = "Mean Value",  
 ylim = c(0, 1200)) # Set the y-axis limits to accommodate Coarse Aggregate max value  
  
legend("topright", legend = c("Cement", "Fly Ash", "Coarse Aggregate", "Fine Aggregate"),   
 fill = c("darkslategrey", "gray", "#996666", "#666666"))



# Define intervals for Concrete Strength  
df\_conc$Strength\_Group <- cut(df\_conc$Concrete.strength, breaks = c(0, 30, 50, 70, max(df\_conc$Concrete.strength)), labels = c("<30", "30-50", "50-70", ">=70"))  
  
# Calculate mean Cement, Fly Ash, Coarse Aggregate, Fine Aggregate, and Water for each interval of Concrete Strength  
aggregate\_values <- aggregate(cbind(Cement, Fly.Ash, Coarse.Aggregate, Fine.Aggregate, Water) ~ Strength\_Group, data = df\_conc, FUN = mean)  
  
# Create a bar plot for Cement, Fly Ash, Coarse Aggregate, Fine Aggregate, and Water on different intervals of Concrete Strength  
barplot(t(aggregate\_values[, c("Cement", "Fly.Ash", "Coarse.Aggregate", "Fine.Aggregate", "Water")]), beside = TRUE,   
 col = c("darkslategrey", "gray", "#996666", "#666666", "lightblue"),   
 names.arg = aggregate\_values$Strength\_Group,   
 main = "Mean Values of COmponents on Different Concrete Strength Intervals",  
 xlab = "Concrete Strength Intervals (MPa)", ylab = "Mean Value",  
 ylim = c(0, 1200)) # Set the y-axis limits to accommodate Coarse Aggregate max value  
legend("topright", legend = c("Cement", "Fly Ash", "Coarse Aggregate", "Fine Aggregate", "Water"),   
 fill = c("darkslategrey", "gray", "#996666", "#666666", "lightblue"))



# Define colors based on Superplasticizer values  
colors <- ifelse(df\_conc$Superplasticizer < 2, "green",   
 ifelse(df\_conc$Superplasticizer >= 2 & df\_conc$Superplasticizer < 4, "blue", "darkorange"))  
par(mfrow = c(1, 2)) # Set up a 1x2 grid for plots  
# Scatter plot of Superplasticizer against Water with colored points based on Superplasticizer range  
plot(df\_conc$Superplasticizer, df\_conc$Water,   
 xlab = "Superplasticizer (kg in a m^3 mixture)", ylab = "Water (kg in a m^3 mixture)",  
 main = "Relationship between Superplasticizer and Water",  
 col = colors)  
  
# Adding a trend line for Water vs. Superplasticizer  
abline(lm(df\_conc$Water ~ df\_conc$Superplasticizer), col = "red")  
  
# Adding a legend  
legend("topright", legend = c("< 2", "2 - 4", ">= 4"), col = c("green", "blue", "darkorange"), pch = 2,  
 title = "Superplasticizer Range")  
  
# Create a second plot to show the effect on response variable (Concrete.strength)  
  
  
# Scatter plot of Superplasticizer against Concrete.strength with colored points based on Superplasticizer range  
plot(df\_conc$Superplasticizer, df\_conc$Concrete.strength,   
 xlab = "Superplasticizer (kg in a m^3 mixture)", ylab = "Concrete Strength (MPa)",  
 main = "Effect of Superplasticizer on Concrete Strength",  
 col = colors)  
  
# Adding a trend line for Concrete.strength vs. Superplasticizer  
abline(lm(df\_conc$Concrete.strength ~ df\_conc$Superplasticizer), col = "red")  
  
# Adding a legend  
legend("topright", legend = c("< 2", "2 - 4", ">= 4"), col = c("green", "blue", "darkorange"), pch = 2,  
 title = "Superplasticizer Range")  
  
  
# Adding a legend  
legend("topright", legend = c("< 2", "2 - 4", ">= 4"), col = c("green", "blue", "darkorange"), pch = 2,  
 title = "Superplasticizer Range")

