

Proyecto 2

ByteMiners

06/02/2020

Descripción

Esta base de datos reúne información sobre centros educativos públicos de la ciudad de Boston, Massachusetts. Su objetivo es permitir el análisis continuo para guiar la toma de decisiones relacionadas con las inversiones en la construcción de escuelas de ese sitio. Los datos fueron recolectados por un equipo de evaluadores, conformado por arquitectos y planificadores educativos, que visitaron todos los edificios de las Escuelas Públicas de Boston, recolectando y organizando información sobre la condición física de cada edificio y sus idoneidad educativa para los programas ofrecidos dentro.

Análisis exploratorio

La base de datos se conforma de 141 observaciones y cada una cuenta con 251 variables, de las cuales cerca del 40% se encuentran vacías. Debido a lo anterior, decidimos enfocar nuestro análisis en las variables pertenecientes a las siguientes fuentes:

*Departamento de instalaciones de BPS

*SMMA (Symmes Maini & McKee Associates)

*SMMA_EA (Información recopilada como parte de la evaluación educativa)

*SMMA_FA (Información recopilada como parte de la evaluación de las instalaciones)

*Información obtenida del Departamento de Educación Primaria y Secundaria de Massachusetts (DESE / DOE)

```
library(dplyr)
library(ggplot2)
library(corrgram)
library(tidyr)
library(faraway)
setwd("C:/Users/emman/Downloads")
df <- read.csv("Buildbps.csv", sep = ",")

df$BPS_Electric_Bill = as.numeric(gsub("[\\$,]", "", df$BPS_Electric_Bill))
df$BPS_Water_Bill = as.numeric(gsub("[\\$,]", "", df$BPS_Water_Bill))
df$BPS_Gas_Bill = as.numeric(gsub("[\\$,]", "", df$BPS_Gas_Bill))
df$BPS_Total_Energy_Cost = as.numeric(gsub("[\\$,]", "", df$BPS_Total_Energy_Cost))

high <- filter(df, SMMA_Typology == "High School")
special <- filter(df, SMMA_Typology == "Special")
elementary <- filter(df, SMMA_Typology == "Elementary School")
early <- filter(df, SMMA_Typology == "Early Learning")
k8 <- filter(df, SMMA_Typology == "K-8")
middle <- filter(df, SMMA_Typology == "Middle School")

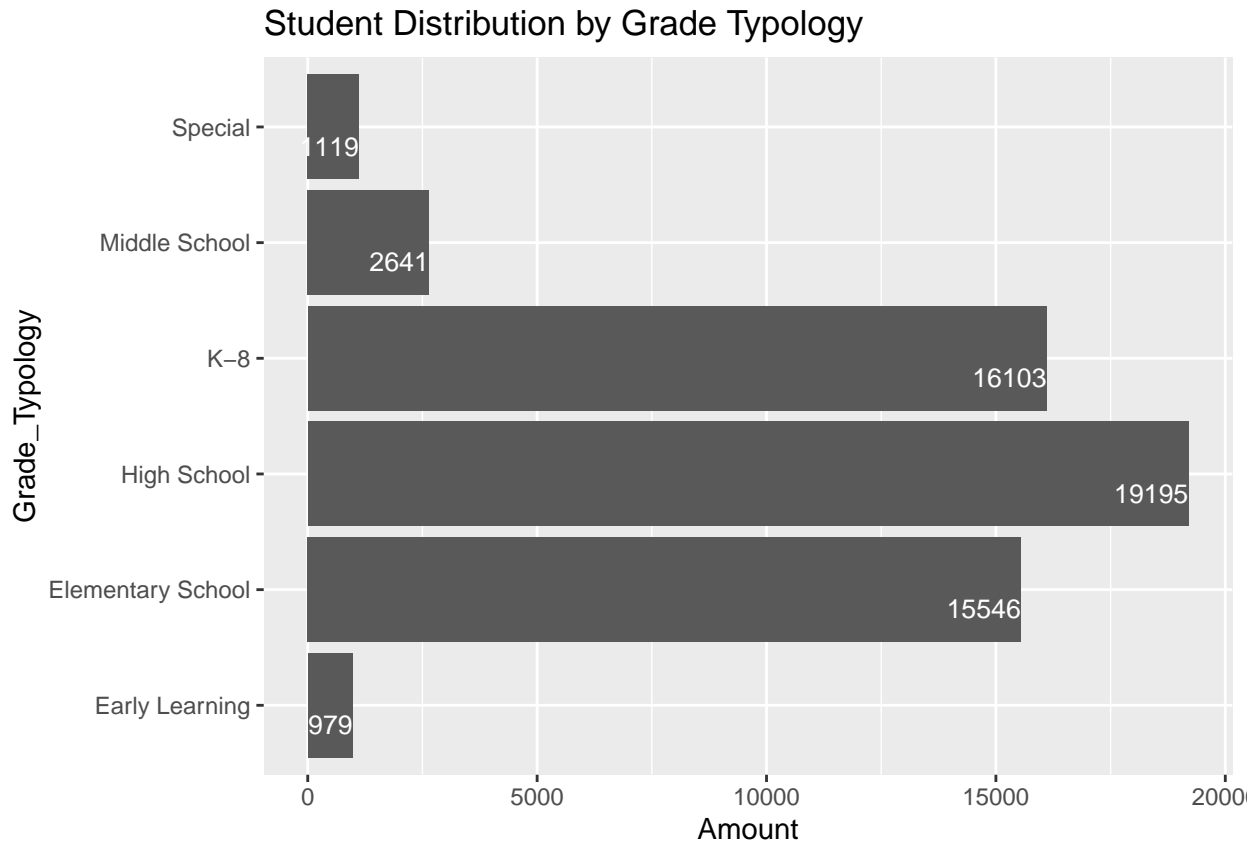
cantidad_estudiantes <- c(sum(high$DOE_Total), sum(middle$DOE_Total), sum(special$DOE_Total),
```

```

sum(k8$DOE_Total), sum(elementary$DOE_Total), sum(early$DOE_Total))

matriculados <- data.frame(Grade_Typology=c("High School", "Middle School", "Special",
                                             "K-8", "Elementary School", "Early Learning"),
                           Amount=cantidad_estudiantes)
ggplot(matriculados, aes(x=Grade_Typology, y=Amount)) + geom_bar(stat = "identity") +
  coord_flip() + labs(title="Student Distribution by Grade Typology")+
  geom_text(aes(y=Amount, label=cantidad_estudiantes),vjust=1.5,color="white",size=3.5,
            hjust="right")

```



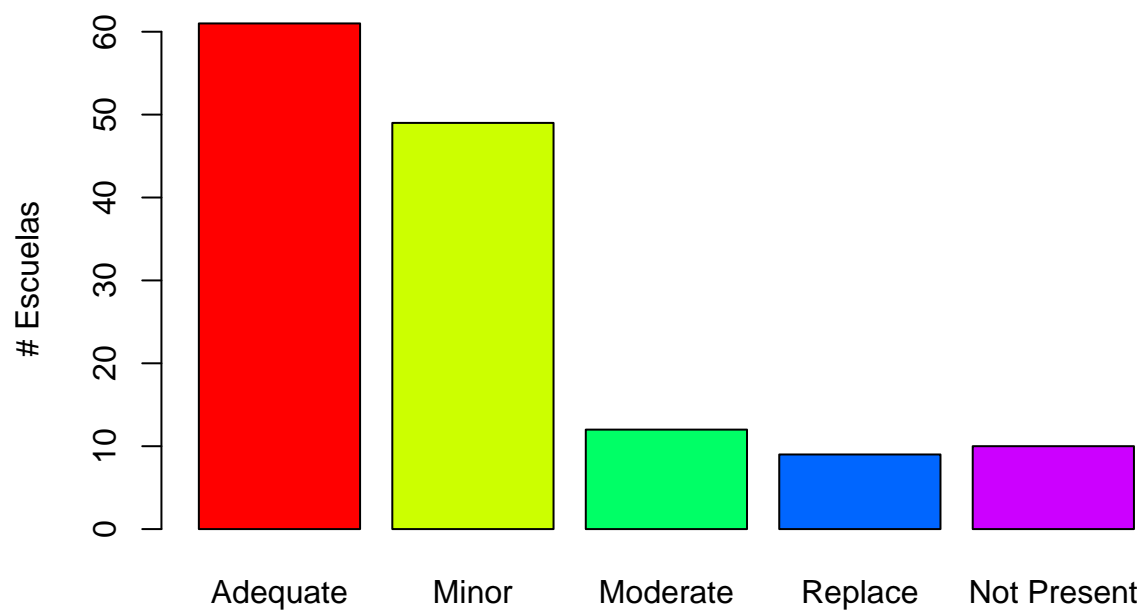
De acuerdo a las inscripciones registradas en el periodo de 2015 a 2016, podemos observar que hay una mayor población de estudiantes a nivel bachillerato (High School) en la ciudad de Boston. Este dato puede servir para determinar si es necesario fundar más escuelas de este nivel educativo o ampliar la extensión de los edificios con la finalidad de que exista el espacio adecuado para la cantidad de estudiantes necesarios.

```

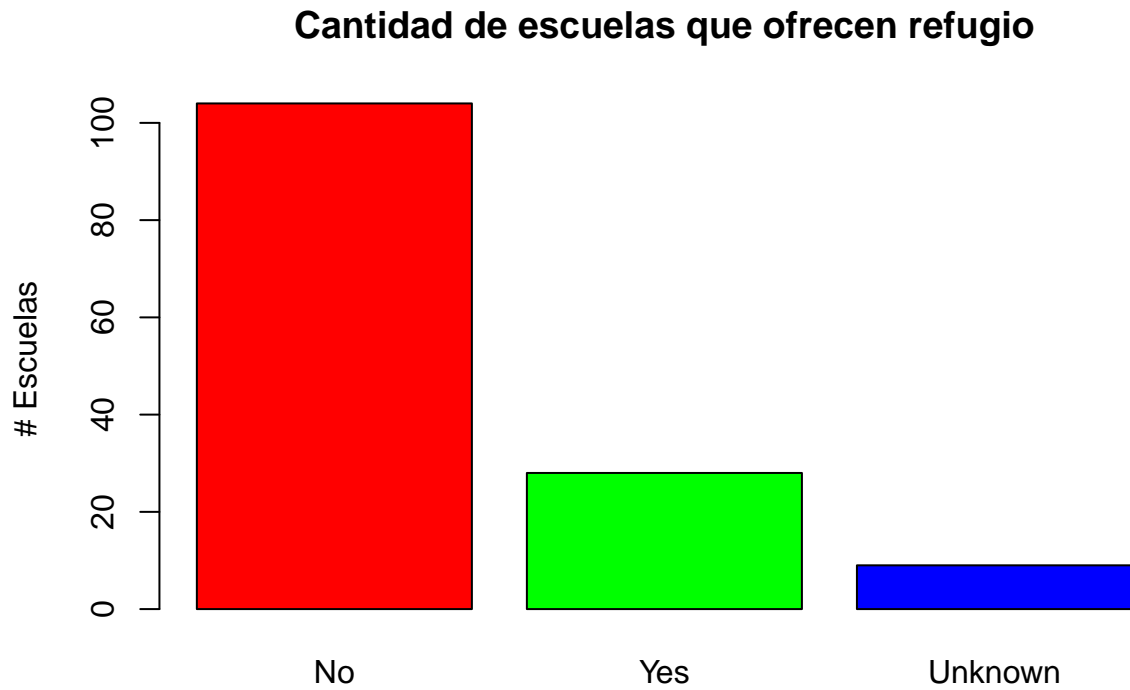
levels(df$SMMA_FA_Life_Safety) <- c(levels(df$SMMA_FA_Life_Safety), "Not Present")
df$SMMA_FA_Life_Safety[df$SMMA_FA_Life_Safety == ''] <- "Not Present"
df$SMMA_FA_Life_Safety <- factor(df$SMMA_FA_Life_Safety)
levels(df$SMMA_FA_Emergency_Shelter) <- c(levels(df$SMMA_FA_Emergency_Shelter), "Unknown")
df$SMMA_FA_Emergency_Shelter[df$SMMA_FA_Emergency_Shelter == ''] <- "Unknown"
df$SMMA_FA_Emergency_Shelter <- factor(df$SMMA_FA_Emergency_Shelter)
plot(df$SMMA_FA_Life_Safety, main = "Nivel de seguridad ofrecido por las escuelas de Boston",
     ylab = "# Escuelas", col=rainbow(5))

```

Nivel de seguridad ofrecido por las escuelas de Boston



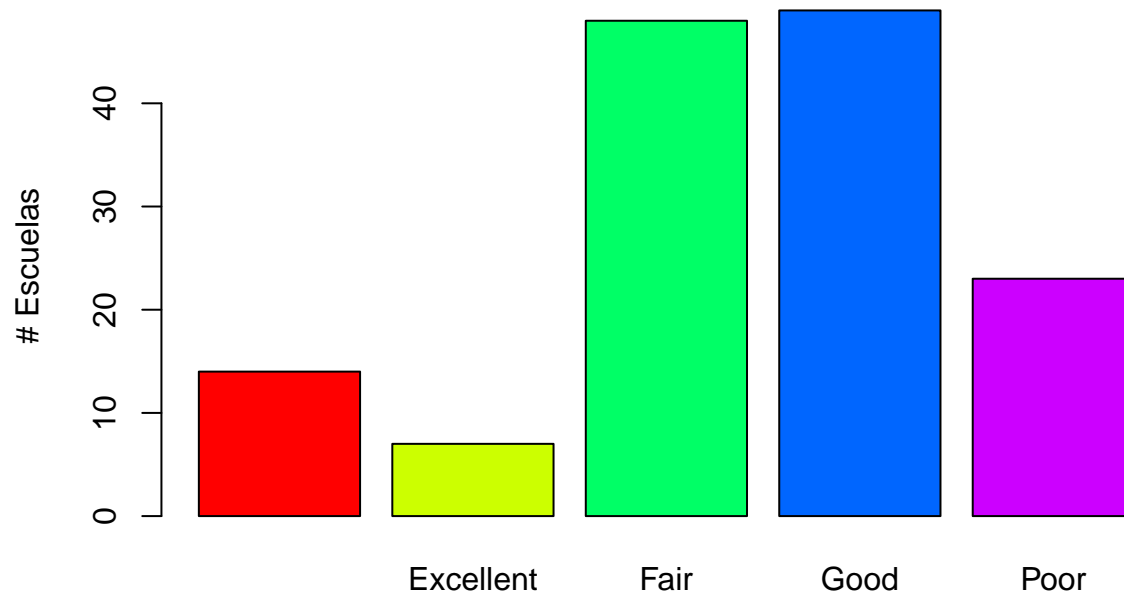
```
plot(df$SMMA_FA_Emergency_Shelter, main = "Cantidad de escuelas que ofrecen refugio",  
     ylab = "# Escuelas", col=rainbow(3))
```



Por otro lado, el cambio climático y el aumento del nivel del mar es un tema que igualmente destaca en esta exploración. Si tomamos en cuenta que Boston es una de las ciudades de Estados Unidos con mayor riesgo de inundaciones y amenazas meteorológicas, es imprescindible identificar cuántos centros educativos fungen como refugios. A través de la exploración se logró identificar que tan solo el 20% de ellos son usados como refugios de emergencia en caso de catástrofes.

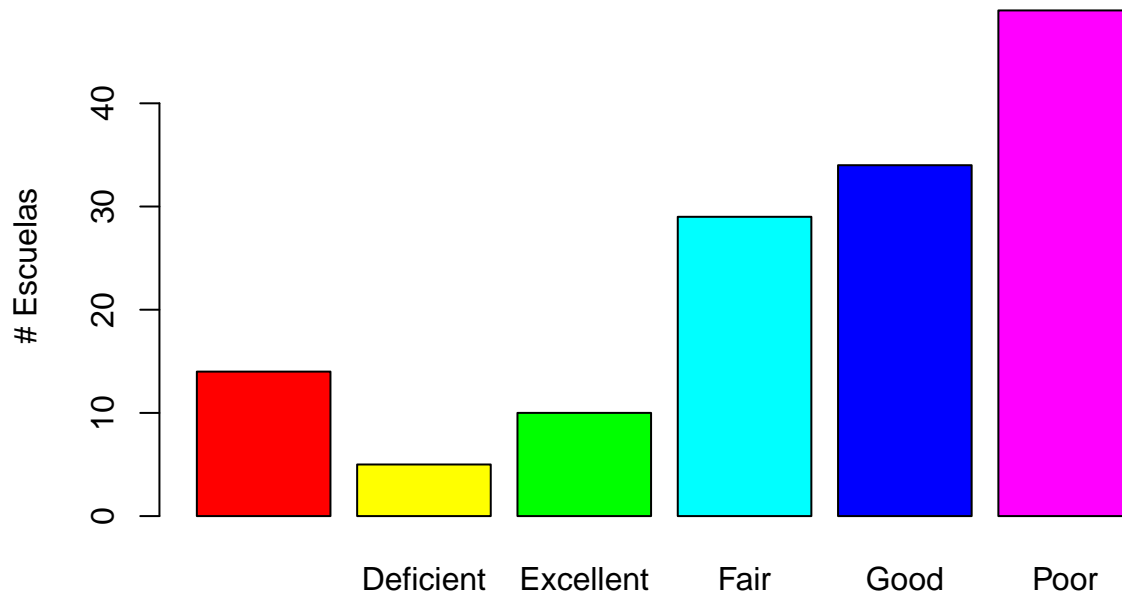
```
plot(df$SMMA_EA_Building_Technology_Wireless, main = "Calidad de instalación inalámbrica",  
      ylab = "# Escuelas", col=rainbow(5))
```

Calidad de instalación inalámbrica



```
plot(df$SMMA_EA_Building_Technology_Interactive, main = "Calidad de las TIC para la educación",  
      ylab = "# Escuelas", col=rainbow(6))
```

Calidad de las TIC para la educación



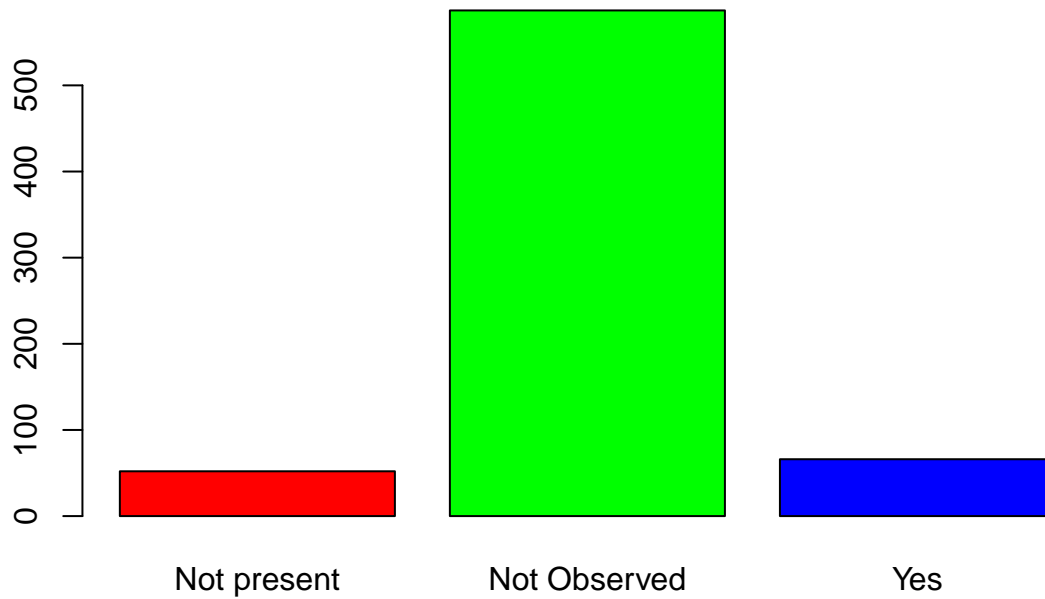
La calidad de los entornos físicos tiene un impacto directo en los resultados educativos. Es por eso que consideramos importante analizar la calidad de las Tecnologías de la Información y Comunicación empleadas en los centros educativos evaluados (por ejemplo: el uso de pizarras interactivas o cámaras), así como la calidad del cableado inalámbrico. Estos factores cualitativos influyen en la comodidad y la capacidad de los estudiantes para concentrarse en las tareas; facilita la capacidad de enseñanza dinámica de los maestros y la retención de información, lo que finalmente tiene un impacto en el rendimiento del estudiante en general.

Como se puede apreciar en la gráfica, desgraciadamente cerca del 35% de las escuelas presentan una calidad “Pobre” en cuanto al uso de TIC para la educación.

Por otro lado, si hablamos sobre la calidad de la instalación inalámbrica, la mayoría de los centros educativos presentan una buena instalación de este tipo.

```
counts <- cbind(df$SMMA_FA_Deterioration_Facade,df$SMMA_FA_Deterioration_Floor,
               df$SMMA_FA_Deterioration_Foundations,df$SMMA_FA_Deterioration_Roof,
               df$SMMA_FA_Deterioration_Walls_Columns)
counts <- table(counts)
names(counts)=c("Not present","Not Observed","Yes")
barplot(counts, main="Deterioramiento de las instalaciones en las escuelas de Boston",
        col = rainbow(3))
```

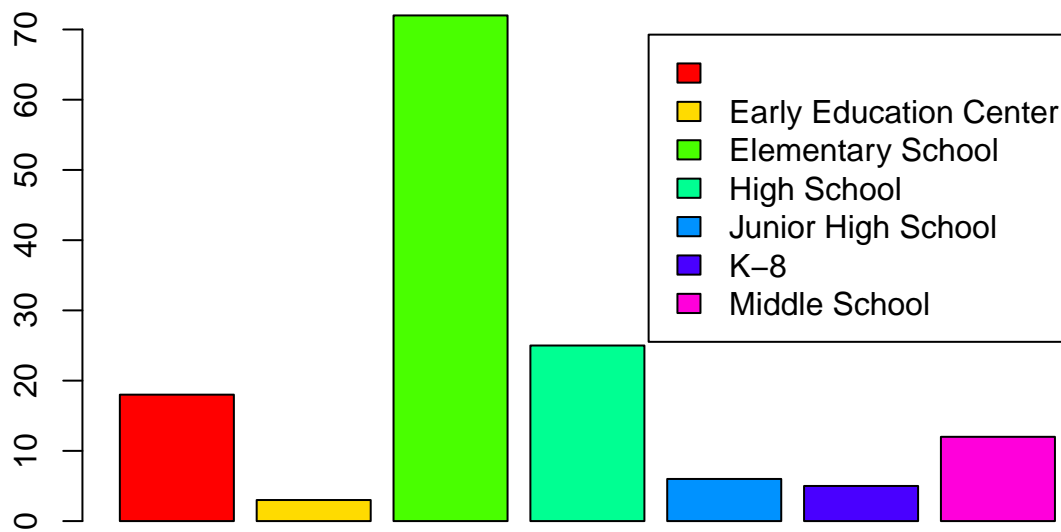
Deterioramiento de las instalaciones en las escuelas de Boston



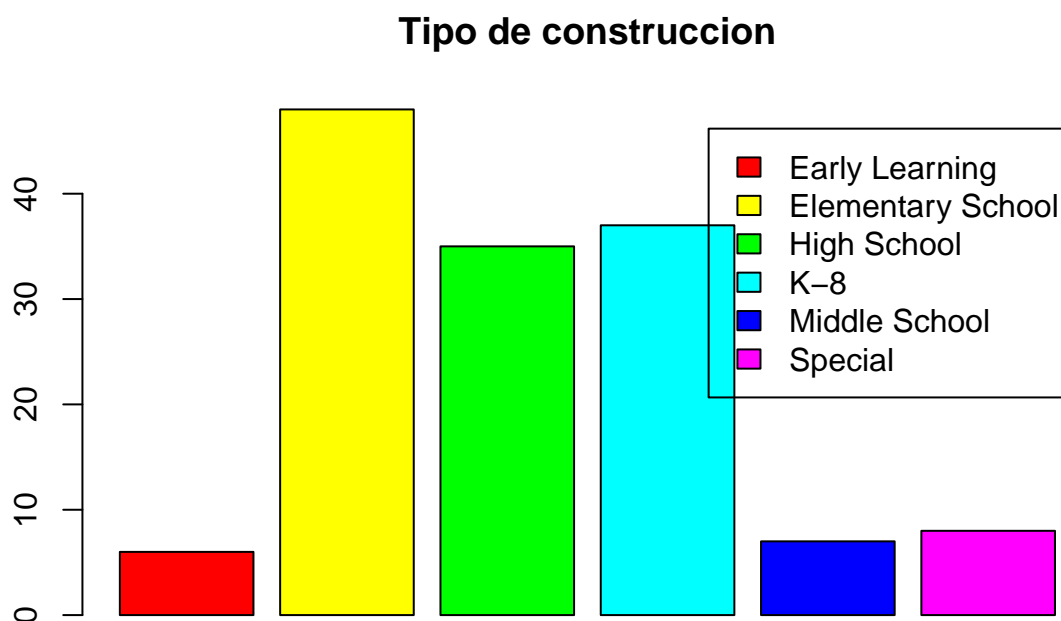
Como se puede apreciar en el grafico la mayoría de las instalaciones de las escuelas de Boston se encuentran en buen estado, no encontraron y/o observaron indicios de deterioro, son pocas las escuelas donde hay deterioramiento.

```
ori<-table(df$SMMA_EA_Building_originally_designed)
barplot(ori, main="Construccion disenada originalmente",col = rainbow(7),legend = rownames(ori),
        names.arg=FALSE)
```

Construccion disenada originalmente



```
tipo<-table(df$SMA_Typology)
barplot(tipo, main="Tipo de construccion",col = rainbow(6),legend = rownames(tipo),names.arg=FALSE)
```

Se pude notar que al comienzo, algunas de las infraestructuras fueron creadas con otros fines y con el pasar de los años, al hacer una remodelación cambio su tipolgia, un claro ejemplo es que algunas escuelas empezaron siendo Junior High School, y otras fueron creadas sin un objetivo principal.

Pero en la actualidad desaparecieron las Junior High School y se crearon escuelas especiales, de igual manera aumetó el número de K-8 hasta el punto que supero a las High Schools.