

# Actividad 3

Stefania Astudillo Bello

## Problema

Con base en los datos de ofertas de vivienda descargadas del portal Fincaraiz para apartamento de estrato 4 con área construida menor a 200 m<sup>2</sup> (vivienda4.RDS) la inmobiliaria A&C requiere el apoyo en la construcción de un modelo que lo oriente sobre los precios de inmuebles.

Con este propósito el equipo de asesores a diseñado los siguientes pasos para obtener un modelo y así poder a futuro determinar los precios de los inmuebles a negociar

```
viviendas <- readRDS(file = "vivienda4.RDS")
head(viviendas)
```

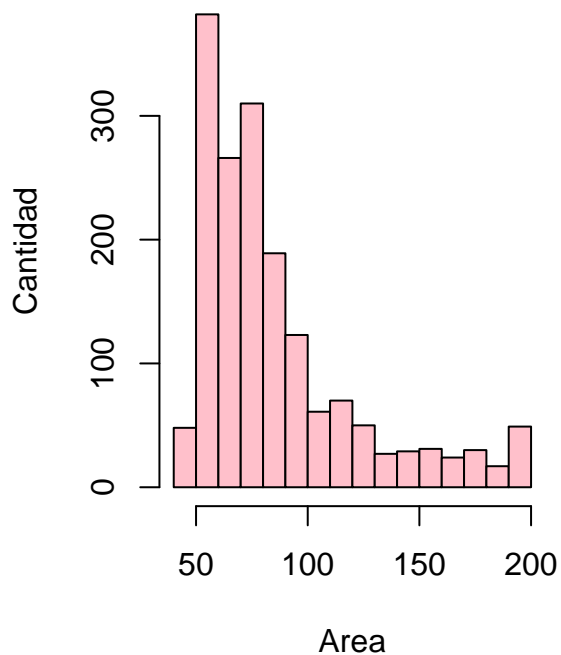
##	zona	estrato	preciom	areaconst	tipo
## 1	Zona Norte	4	220	52	Apartamento
## 2	Zona Norte	4	600	160	Casa
## 3	Zona Norte	4	320	108	Apartamento
## 4	Zona Sur	4	290	96	Apartamento
## 5	Zona Norte	4	220	82	Apartamento
## 6	Zona Norte	4	305	117	Casa

1. Realice un análisis exploratorio de las variables precio de vivienda (millones de pesos COP) y área de la vivienda (metros cuadrados) - incluir gráficos e indicadores apropiados interpretados.

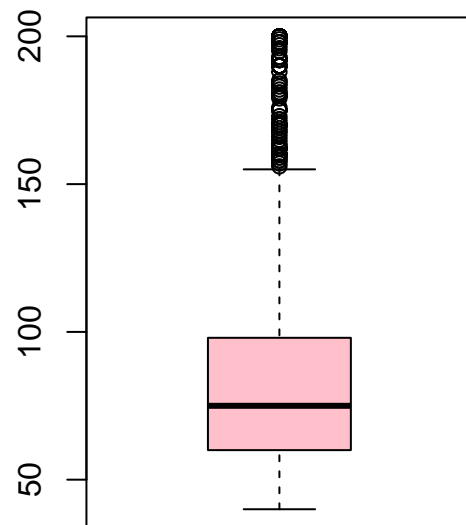
```
par(mfrow=c(1,2))
hist(viviendas$areaconst, xlab = "Area", ylab = "Cantidad", col="pink",
     main="Histograma de la variable areaconst")

boxplot(viviendas$areaconst, col="pink",
        main="Caja de la variable areaconst")
```

### Histograma de la variable areacor



### Caja de la variable areaconst



```
datosArea <- summary(viviendas$areaconst)
datosArea
```

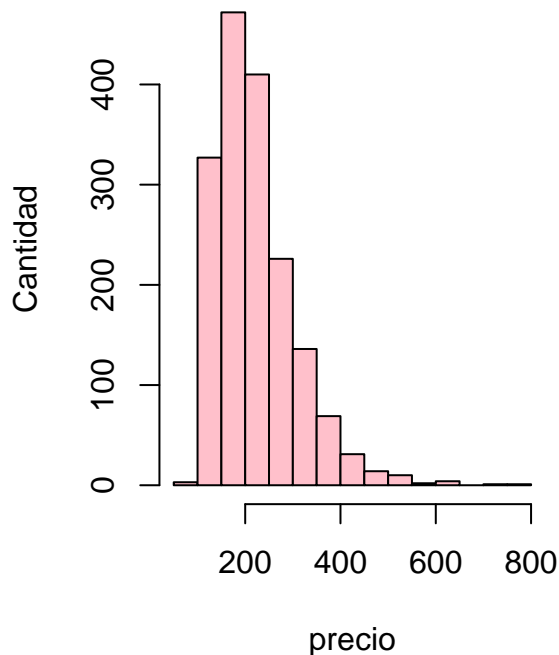
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    40.00  60.00   75.00   87.63  98.00  200.00
```

el promedio del área construida de la oferta de viviendas es de 87.63 m<sup>2</sup> para áreas que oscilan entre los 40 m<sup>2</sup> (mínimo) y los 200 m<sup>2</sup> (máximo). La distribución del área de las viviendas si muestra un sesgo bastante pronunciado a la derecha; podemos observar que el 50% de las viviendas tienen entre 40 m<sup>2</sup> y 75 m<sup>2</sup> mientras que el otro 50% tiene entre 75m<sup>2</sup> y 200 m<sup>2</sup>.

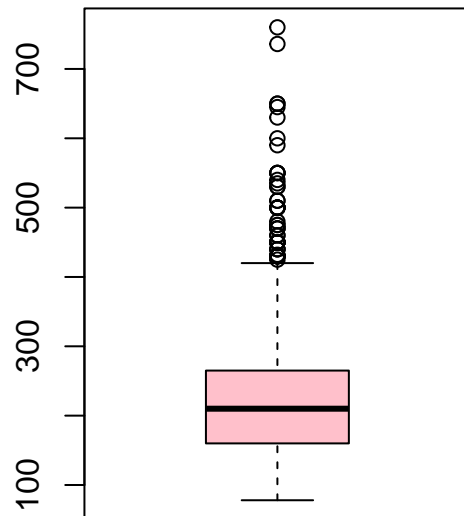
```
par(mfrow=c(1,2))
hist(viviendas$preciom, xlab = "precio", ylab = "Cantidad", col="pink",
     main="Histograma de la variable preciom")

boxplot(viviendas$preciom, col="pink",
       main="Caja de la variable preciom")
```

### Histograma de la variable precio



### Caja de la variable preciom



```
datosPrecio <- summary(viviendas$preciom)
datosPrecio
```

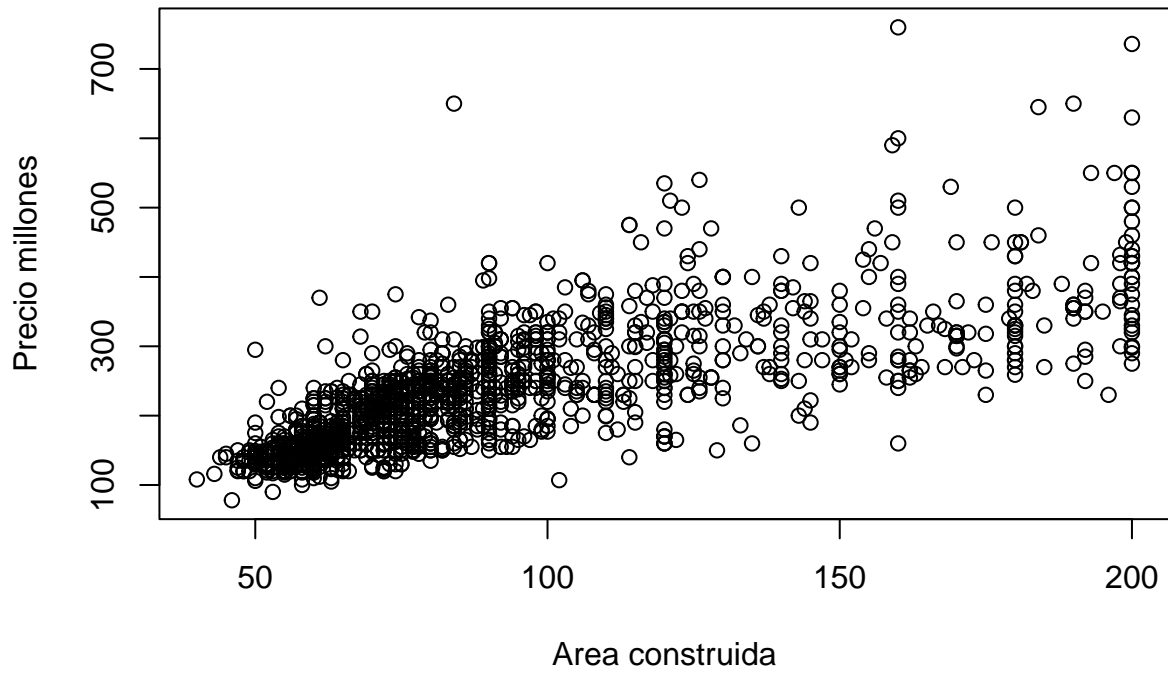
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      78.0  160.0   210.0   225.4  265.0   760.0
```

Los datos nos revelan que el promedio del precio de la oferta de viviendas es de 225.4 millones, para valores que oscilan entre los 78 millones (mínimo) y los 760 millones (máximo). podemos precisar que la mitad de las viviendas se oferta a un precio menor o igual a los 210 millones y sólo un 25% están en un rango de costosas superando los 265 millones de pesos colombianos.

2. Realice un análisis exploratorio bivariado de datos enfocado en la relación entre la variable respuesta (precio) en función de la variable predictora (area construida) - incluir gráficos e indicadores apropiados interpretados.

```
plot(viviendas$areaconst, viviendas$preciom,
      xlab = "Area construida", ylab="Precio millones", main="Análisis Exploratorio bivariado Area/Precio")
```

## Analisis Exploratorio bivariado Area/Precio



```
cor(viviendas$areaconst,viviendas$preciom)
```

```
## [1] 0.7630166
```

relacion positiva debil