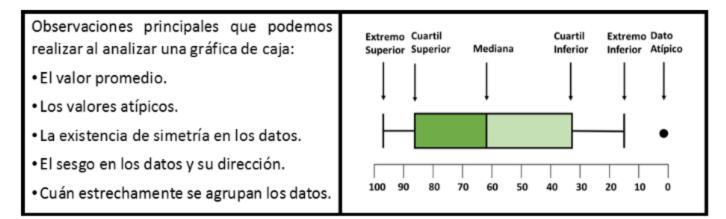
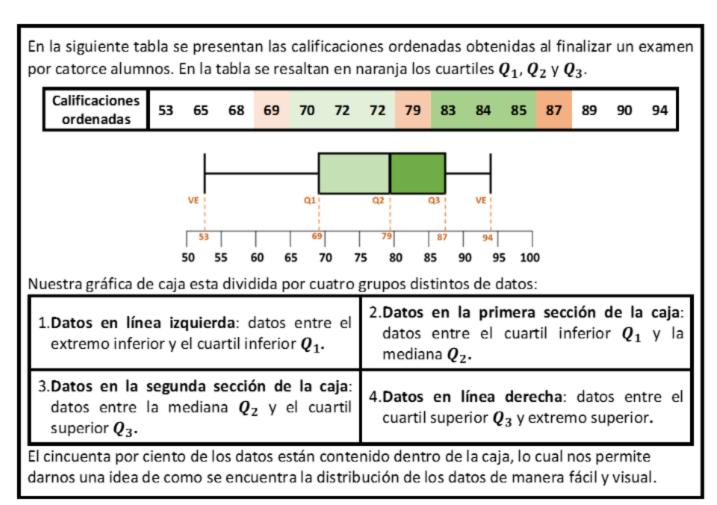
## Gráficas de Caja

Las gráficas de caja son una conveniente representación gráfica de un grupo de datos numéricos utilizando sus cuartiles como referencias. Este tipo de gráficas son utilizadas en la estadística descriptiva y son una excelente forma de analizar datos rápidamente con la ventaja de ocupar un menor espacio, lo cual es útil al comparar distribuciones entre muchos conjuntos de datos.

Las líneas que se extienden paralelas a las cajas, comúnmente llamadas bigotes, se utilizan para indicar la variabilidad fuera del intervalo formado entre el cuartil inferior y el superior. Los valores atípicos se representan como puntos individuales alineados a los bigotes.



Ahora que ya conocemos todos los detalles que conforman una gráfica de caja, mostraremos un ejemplo practico de como se representan datos utilizando este tipo de visualización.



```
In []: import pandas as pd

Heart Disease UCI Cleveland
    https://www.kaggle.com/ronitf/heart-disease-uci

In []: df = pd.read_csv('data/heart/heart.csv')
    df.head(7)

In []: import matplotlib.pyplot as plt
    %matplotlib inline
    df.boxplot('chol')
    plt.title("Gráfico de cajas Colesterol")
    plt.show()

In []: # cajas separando datos, una caja para mujeres y otra para hombres
    import matplotlib.pyplot as plt
    df.boxplot(by = 'sex',column = 'chol')

# Pandas no es muy bueno con Los títulos grandes
```

## Mejorando el gráfico con libreria Seaborn

plt.savefig('boxplot.png',dpi=100) # Se exporta la gráfica

```
In []: import seaborn as sns sns.boxplot(x = 'sex', y = 'chol', data = df)

In []: ## El ancho es mucho, se puede modificar sns.set() sns.boxplot(x = 'sex', y = 'chol', data = df, width=0.2) # El 20% del ancho

In []: # LLuvia de puntos en la caja sns.swarmplot(x = 'sex', y = 'chol', data = df, alpha=0.5) # Donde hay más puntos es donde esta la Media

In []: # LLuvia de puntos en la caja sns.swarmplot(x = 'sex', y = 'chol', data = df, alpha=0.5) sns.boxplot(x = 'sex', y = 'chol', data = df, width=0.2)

In []: # Un título para la grafica y exportarla sns.boxplot(x = 'sex', y = 'chol', data = df, width=0.2) sns.swarmplot(x = 'sex', y = 'chol', data = df, alpha=0.5) plt.title("Colesterol por sexo")
```