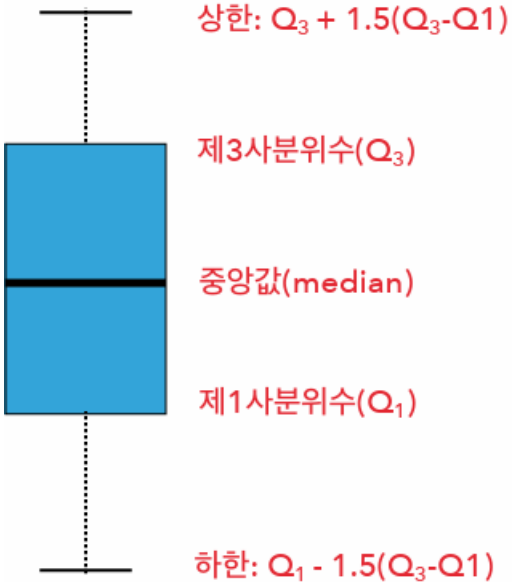


<p>Elementary Statistics</p> <p>모집단(population): 통계분석의 대상이 되는 모든 개체들의 집합 ▶  모수(parameter): 모집단의 특성을 나타내는 수치 (평균, 분산, 비율 등) ▶ 표본(sample): 모집단으로부터 일정한 규칙에 의해 수집한  모집단의 부분집합 ▶ 통계량(statistic): 모집단의 특성을 추측하기  위해 사용되는 표본의 함수(표본</p>	
<div style="text-align: center;"> <p><b>Population 모집단</b></p> <div style="border: 1px solid black; padding: 5px; display: inline-block;">True mean (parameter)</div> <p><math>\mu</math></p> <p><b>Sample 표본</b></p> <div style="border: 1px solid black; padding: 5px; display: inline-block;">Sample mean (statistic)</div> <math display="block">\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i</math> </div> <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <span>통계적 추론</span> <span>데이터 수집</span> </div>	
<pre>x &lt;- c(2,4.4,3,3,2,2.2,2,4) mean(x) median(x) x.tab &lt;- table(x)  #tan so max(x.tab) x.tab[x.tab==max(x.tab)]</pre>	
<pre>data() chickwts chickwts[1:3,] str(chickwts) tapply(chickwts\$weight, INDEX=chickwts\$feed, FUN=mean) tapply(chickwts\$weight, INDEX=chickwts\$feed, FUN=median)</pre>	<pre>mean(chickwts\$weight) median(chickwts\$weight) c.tab &lt;- table(chickwts\$feed)</pre>

	tính trung vị (median) của (weight) của gà con (chicks) cho mỗi nhóm theo loại thức ăn
<pre>table(chickwts\$feed) table(chickwts\$feed)/nrow(chickwts) round(table(chickwts\$feed)/nrow(chickwts), 3) sum(chickwts\$feed=="casein")/nrow(chickwts)</pre>	
<pre>chickwts\$feed == "casein" as.numeric(chickwts\$feed == "casein") sum(chickwts\$feed=="casein")</pre>	
<b>QUANTILES</b>	
<pre>quantile(chickwts\$weight, prob=0.25) quantile(chickwts\$weight, prob=c(0,0.25,0.5,0.75,1)) summary(chickwts\$weight)</pre>	
<pre>var(chickwts\$weight) sd(chickwts\$weight) IQR(chickwts\$weight) = Q3 - Q1 (Q3:0.75-quantile <math>\hat{\mu}_k</math>, Q1: 0.25-quantile <math>\hat{\mu}_k</math>) sd(chickwts\$weight) / mean(chickwts\$weight)</pre>	<p>(phương sai mẫu độ lệch chuẩn mẫu khoảng tứ phân vị <math>cv = s / \bar{x}</math></p>
<b>Basic Data Visualization</b>	
<pre>mtcars[1:3,] carb.tab &lt;- table(mtcars\$carb) carb.tab barplot(carb.tab)</pre>	Freq of carb columns
<pre>carb.matrix &lt;- table(mtcars\$am, mtcars\$carb) carb.matrix barplot(carb.matrix, beside=T)</pre>	
<b>HISTOGRAMS</b>	
<pre>mtcars\$hp hist(mtcars\$hp) hist(mtcars\$hp, breaks=seq(0,400,25), col=5, main="Horsepower", xlab="HP") seq(0,400,25)</pre>	<p><b>mtcars\$hp: công suất</b> bins với độ rộng là 25, bắt đầu từ 0 và kết thúc ở 400</p>

<pre>str(iris) par(mfrow=c(2,2))</pre>	Chia
<pre>for (k in 1:4) hist(iris[[k]], main=colnames(iris[k]), xlab=colnames(iris[k]), col=k+1)</pre>	
<pre>iris[k] iris[[k]]</pre>	
BOX PLOTS	
	
<pre>&gt; par(mfrow=c(2,2)) &gt; boxplot(mtcars\$mpg, col=2, main="mpg") &gt; boxplot(mtcars\$qsec, col=3, main="qsec") &gt; boxplot(mtcars\$wt, col=4, main="wt") &gt; boxplot(mtcars\$dis, col=4, main="dis")</pre>	
SCATTERPLOTS	
<pre>&gt; plot(mtcars\$wt, mtcars\$mpg, xlab="weight", ylab="mpg", cex=1.3) # plot() 함수를 이용하여 산점도 작성  points(mtcars[mtcars\$cyl==4,6],mtcars[mtcars\$cyl==4,1],pch=19,col=2) points(mtcars[mtcars\$cyl==6,6],mtcars[mtcars\$cyl==6,1],pch=19,col=3) points(mtcars[mtcars\$cyl==8,6],mtcars[mtcars\$cyl==8,1],pch=19,col=4)</pre>	

Thêm các điểm cho các xe có 4 xi-lanh, với màu sắc là màu thứ 2 trong bảng màu mặc định.	
<code>pairs(mtcars[c(1,2,4,6)], pch=19, col=4)</code>	
#Xây dựng biểu đồ phân tán giữa bốn biến: hiệu suất nhiên liệu (mpg), xi lanh (xi lanh), công suất (mã lực) và trọng lượng (wt).	