



# 14. 상관분석

## chap14\_Correlation 수업내용

- 1) 상관분석 개요
- 2) 피어슨 상관계수
- 3) 상관분석 실습
- 4) 상관분석 결과 제시



# 1) 상관분석 개요

## 상관관계 분석(Correlation Analysis)

- 변수 간 관련성 분석 방법
- 하나의 변수가 다른 변수와 관련성 분석
- 예, 광고비와 매출액 사이의 관련성 등 분석

### 【상관관계분석 중요사항】

- 회귀분석 전 변수 간 관련성 분석(가설 검정 전 수행)
- 상관계수 → **피어슨(Pearson) R계수** 이용 관련성 유무
  - ✓ 상관관계분석 척도:
  - ✓ 피어슨 상관계수(Pearson correlation coefficient : r)



## 2) 피어슨 상관계수 $r$

### 【피어슨 상관계수 $R$ 】

피어슨 상관계수 $R$	상관관계 정도
$\pm 0.9$ 이상	매우 높은 상관관계
$\pm 0.9 \sim \pm 0.7$	높은 상관관계
$\pm 0.7 \sim \pm 0.4$	다소 높은 상관관계
$\pm 0.4 \sim \pm 0.2$	낮은 상관관계
$\pm 0.2$ 미만	상관관계 없음

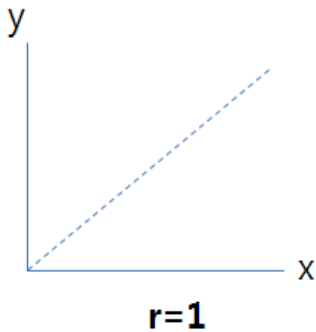
※ 상관계수  $r$ 은 -1에서 +1까지의 값을 가진다. 또한 가장 높은 완전 상관관계의 상관계수는 1이고, 두 변수간에 전혀 상관관계가 없으면 상관계수는 0이다.



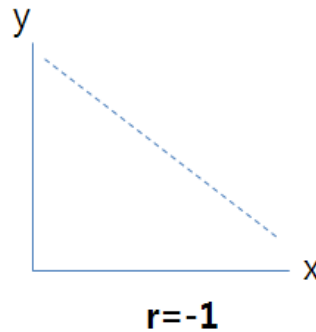
## 2) 피어슨 상관계수 $r$

- 상관계수  $r$ 과 상관관계 정도

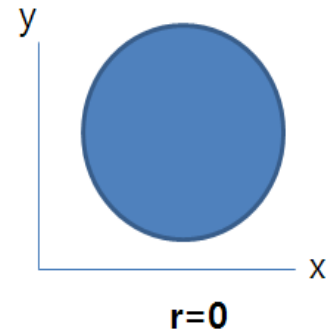
완전 정(+) 상관관계



완전 부(-) 상관관계



상관관계 없음





### 3) 상관분석 실습

# 데이터셋 가져오기

```
result <- read.csv("C:/Rwork/data/drinking_water.csv", header=T)  
head(result)
```

# 상관계수 보기

```
cor(result$친밀도, result$적절성)  
cor(result$친밀도, result$만족도)
```

# 전체 변수 간 상관계수 보기

```
cor(result, method="pearson") # 피어슨 상관계수 - default
```

```
cor(result, method="spearman") # spearman 상관계수(서열척도)
```



### 3) 상관분석 실습

# 상관관계 시각화  
library(corrplot)

```
corr <- cor(product, method="pearson")  
corrplot(corr, method="number")
```





## 4) 상관분석 결과 제시

### 【논문에서 상관관계 분석 결과 제시 방법】

- 일반적으로 상관관계 분석 결과를 논문에서 제시할 경우 해당 기술통계량(평균과 표준편차)과 피어슨 상관계수 함께 제시

분석 단위	평균 (Mean)	표준편차 (Std. Deviation)	분석 단위 간 상관관계 (Inter-Analysis Correlations)		
			1	2	3
1. 친밀도	2.928	0.9703446	1		
2. 적절성	3.133	0.8596574	.499	1	
3. 만족도	3.095	0.8287436	.467	.767	1