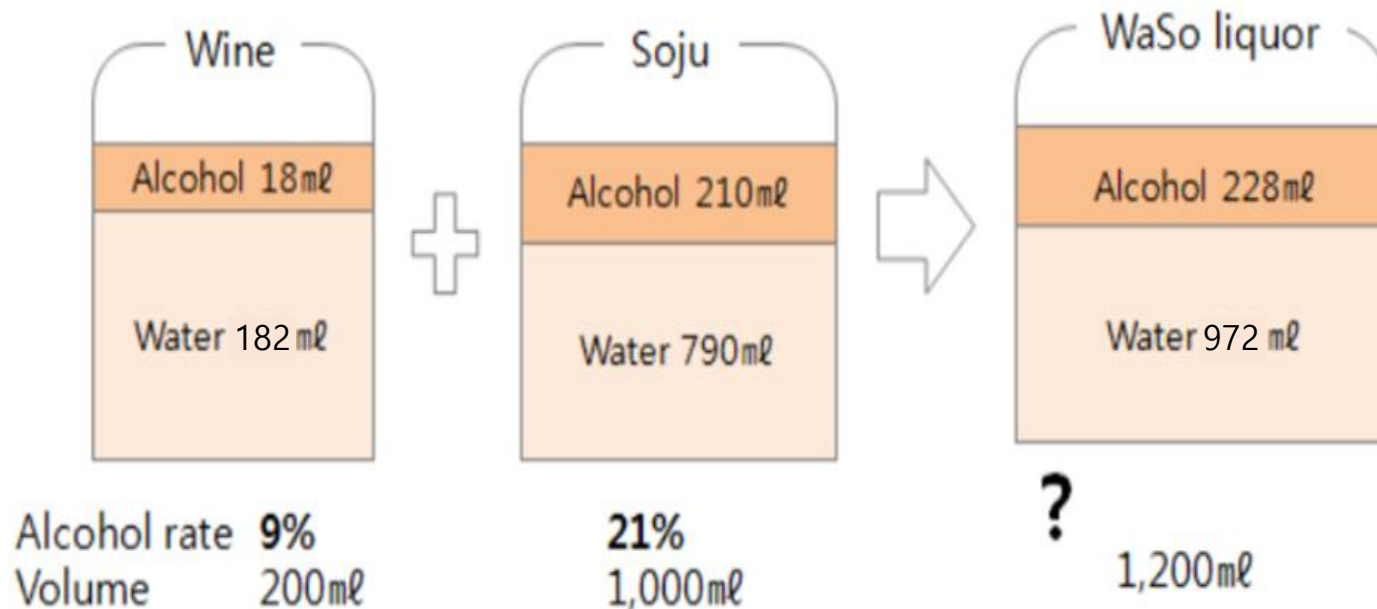


# 대푯값 - 가중평균 : 문제1)

- 가중평균(weighted average)

[문제] 알코올 도수 9%인 와인 200ml와 알코올 도수 21%인 소주 1000ml를 섞어서 만든 와소 폭탄주의 평균 알코올 도수는?



# 대포깁 - 가중평균 : 문제1)

- 가중평균(weighted average)

$$\text{Weighted Average} = \frac{W_1 \times x_1 + W_2 \times x_2 + \dots + W_n \times x_n}{W_1 + W_2 + \dots + W_n}$$

- ✓ 알코올 도수 9%인 와인 200ml
- ✓ 알코올 도수 21%인 소주 1000ml
- ✓ 와소 폭탄주의 평균 알코올 도수 = 19%

```
# weighted mean
weighted_mean1 <- (200*0.09 + 1000*0.21)/(200+1000)
weighted_mean1
#[1] 0.19

# weighted mean
alcohol <- data.frame(volume=c(200, 1000), alcohol_rate=c(0.09, 0.21))
weighted_mean2 <- weighted.mean(alcohol$alcohol_rate, alcohol$volume)
weighted_mean2
#[1] 0.19
```

# 상관분석 - 상관계수

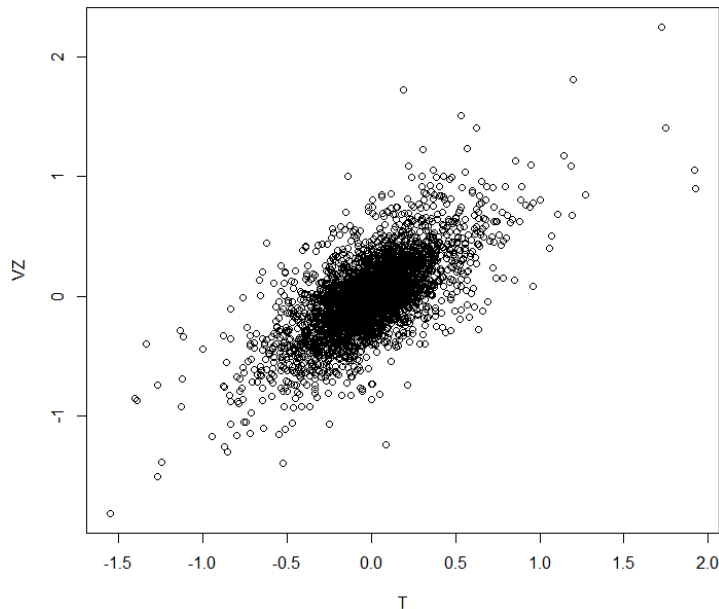
- 상관분석(Correlation Analysis)은 두 확률 변수 사이의 관련성을 파악하는 방법
- 상관계수(Correlation Coefficient)는 두 변수 간 관련성의 정도를 의미하며, 이를 계산하는 방법에는 피어슨 상관계수, 스피어만 상관 계수, 켄달의 순위 상관 계수 등이 있음
- 흔히 상관 계수라고 하면 피어슨 상관 계수를 뜻함
- 상관계수 관계의 강도
  - ①  $0 \leq | \text{상관계수} | \leq 0.1$  : 관계 없음
  - ②  $0.2 \leq | \text{상관계수} | \leq 0.4$  : 약한 관계
  - ③  $0.5 \leq | \text{상관계수} | \leq 0.7$  : 중간 관계
  - ④  $0.8 \leq | \text{상관계수} | \leq 1$  : 강한 관계

# 상관분석 - 상관계수

- 피어슨 상관계수(Pearson correlation coefficient)

- 다음은 AT&T(T)와 버라이즌(VZ) 주식의 일간 수익 산점도(상관도) 이다.

```
plot(telecom$T, telecom$VZ, xlab="T", ylab="VZ")
```



- ✓ 좌측의 산포도에서 점들이 오른쪽 위로 증가하는 것처럼 보이는 경향의 정도는 파악할 수 있음
- ✓ 산포도는 모든 사람이 동일한 해석 및 결론에 도달할 수 없음
- ✓ 이 경우 AT&T(T)와 버라이즌(VZ) 주가의 관련 정도를 숫자로 표현한 것이 상관계수

# 상관분석 - 상관계수

- 피어슨 상관계수(Pearson correlation coefficient)

- 다음은 AT&T(T)와 버라이즌(VZ) 주식의 일간 수익 **상관계수** 이다.

```
cor(telecom$T, telecom$VZ)  
#[1] 0.6805438
```

- ✓ 피어슨 상관계수는 약 0.681 으로 두 값 사이에 **중간 정도의 상관 관계**가 있고,
- ✓ AT&T(T) 주가가 상승하면 버라이즌(VZ) 주식도 상승하는 경향이 있음을 알 수 있음