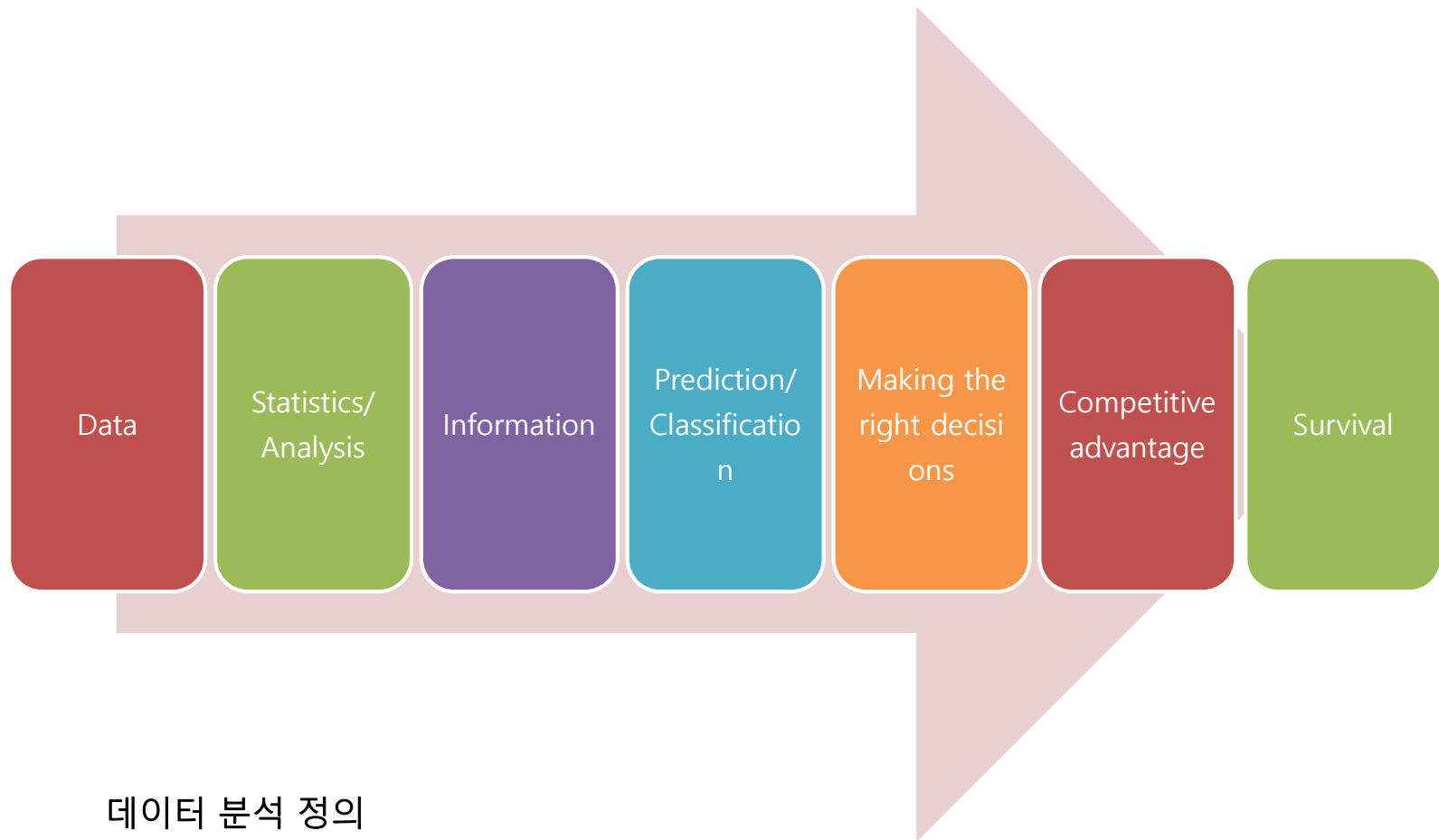


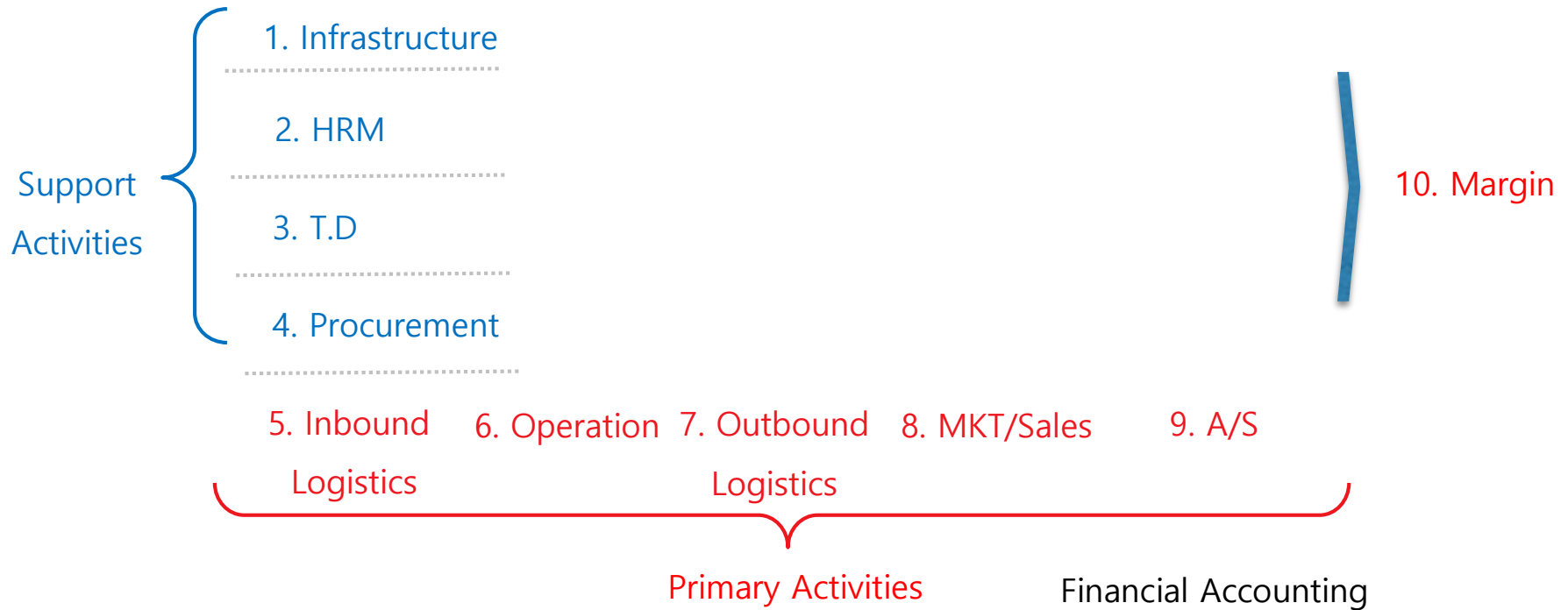
2019년 3월 정화민 교수



# 데이터 분석은 4차산업시대 생존을 위한 필수 기술



# 기업경영에서의 데이터 분석 (예: Value chain)



ERP?

CRM?



가설 :

가설이란 모집단의 특성, 특히 모수에 대한 가정 혹은 잠정적인 결론이다. 이러한 가설을 검정하기 위해 알고 싶은 내용을 기술한 가설의 종류에는 귀무가설과 대립가설이 있다.

## 귀무가설과 대립가설

대립가설(alternative hypothesis):  
표본을 통해 확실한 근거를 가지고 입증하고자 하는  
가설로서 연구가설(research hypothesis)이라고도 한다.  
(새로운 주장 또는 실제로 입증하고자 하는 가설)

대립가설과 반대되는 가설을 귀무가설(null hypothesis)이라 부르며  
현재까지 주장되어 온 것이거나 기존과 비교하여 변화 혹은 차이가  
없음을 나타내는 가설이다.

가설검정: 모집단에 대한 통계적 가설을 세우고 표본을 추출한 다음, 그 표본을 통해 얻은 정보를 이용하여 통계적 가설의 진위를 판단하는 과정 즉, 표본을 활용하여 모집단에 대입해보았을 때 새롭게 제시된 대립 가설이 옳다고 판단할 수 있는지를 평가하는 과정이다.

-증명된 바 없는 주장/가설을 귀무가설이라는 틀 속에 넣고 일부 표본을 추출하여 얻은 표본 통계량(검정 통계량)에 입각하여, 가설의 진위를 판단/증명/검정하는 통계적 추론방법

### (1) 제1종 오류

귀무가설이 참인데 잘못하여 이를 기각하게 되는 오류를 제1종 오류라 한다.

예) 결함 없는 제품을 불량품으로 판정하는 생산자 위험 오류

### (2) 제2종 오류

귀무가설이 참이 아닌데 잘못하여 이를 채택하게 되는 오류를 제2종 오류라 한다.

예) 결함 있는 제품을 정상품으로 판정하는 소비자 위험 오류

## 유의수준 5% , p값 0.05 미만

유의수준 0.05라 함은 두개 집단의 모평균은 실제 같은데도 잘못해서 귀무가설을 기각하게 될 확률  
(모평균이 같지않은 것으로 판단할 확률 . 1종 오류)

즉, 5%의 유의성이란 가설검정 결과가 “사실이 아닐 확률이 5%라는 뜻이고 사실일 확률이 95%라는 뜻.

그러므로 유의확률 p 값은 귀무가설을 기각할 수 있는 최소한의 확률의 의미함.

Ex) 다이어트 효과 검정에서 약을 먹기 전과 후의 체중 비교를 통하여 유의확률 값이 0.05 미만이고 t값이 양수면 체중이 통계적으로 차이날 정도로 줄었다고 해석함.



## 예제 R Studio로 분석하기

	A	B	C
1	Subject	Before	After
2	1	57.9	58.31
3	2	64.68	61.27
4	3	66.3	66.27
5	4	59.97	58.39
6	5	74.12	73.71
7	6	72.71	71.72
8	7	72.5	70.59
9	8	68.36	66.83
10	9	78.86	76.84
11	10	49.24	48.17
12			

다이어트 약 복용 후 체중  
변화 가설검정

귀무가설?  
(가설을 써 보세요)  
대립가설?  
(가설을 써 보세요)

분석방법: 엑셀에서 데이터  
입력 후 CSV 파일로 저장

R studio에서 분석  
통계분석 방법은 무엇입니  
까?  
(통계분석 방법 써 보세요)

A series of four horizontal bars in green, orange, dark blue, and red, stacked vertically on the left side of the slide.

## 예제 R Studio로 분석하기

---

분석과정은 수업시간에 진행합니다.