



## 11주차 과제

과 목	데이터사이언스프로그래밍
담당 교수	이 두 호
학 번	201720970
학 과	소프트웨어·미디어·산업공학부
이 름	권 대 한

# 1. iris 데이터에서 각 종별로 Sepal.Length 가 5.5 보다 큰 관측치 개수를 출력하는 코드를 작성하라.

<실행 코드>

가) iris <- iris

나) iris %>% group\_by(Species) %>% filter(Sepal.Length > 5.5) %>% summarise(n = n())

코드 가)에서 Dataset의 Iris를 지역 변수 iris로 선언하였으며, 코드 나)에서 iris를 Species 필드의 내용으로 그룹화하였으며, 이후 Sepal.Length의 데이터가 5.5를 초과하는 데이터만 요약함으로, 각 필드의 데이터 개수를 출력할 수 있다.

<실행 결과>

```
> iris <- iris
> iris %>% group_by(Species) %>% filter(Sepal.Length > 5.5) %>% summarise(n = n())
# A tibble: 3 x 2
  Species      n
  <fct>    <int>
1 setosa         3
2 versicolor    39
3 virginica     49
```