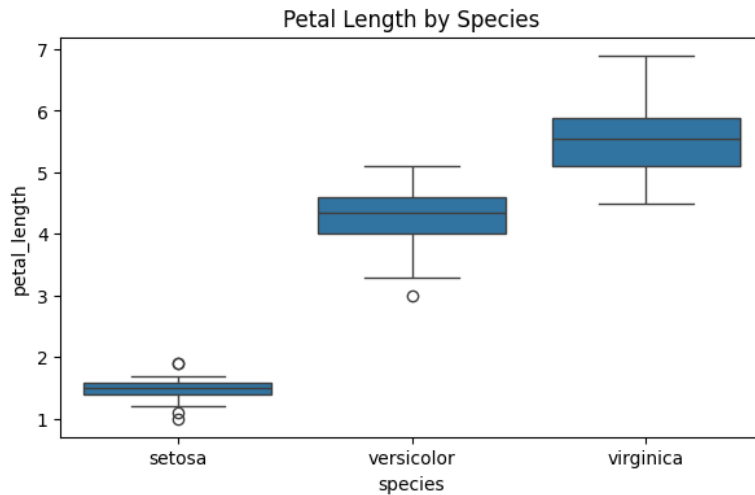


## 0. Introduction

본 과제에서는 Iris 데이터를 활용하여 <sup>1)</sup>세가지 품종별 petal length 평균에 통계적 차이가 있는지 확인하고, <sup>2)</sup>회귀분석을 통해 petal length를 예측하는 모델을 시험합니다.

## 1. Descriptive Stats & Visualization



기술통계량을 비교한 결과, setosa는 전반적으로 값이 작고, versicolor는 중간, virginica는 큰 경향을 볼 수 있습니다. Boxplot 시각화에서도 이러한 경향이 명확히 나타나 종별 평균 차이가 존재할 가능성이 있다고 판단했습니다.

## 2. Shapiro & Levene Test

Shapiro 테스트 결과, Setosa, versicolor, virginica에 대한 p값은 각각 0.0548, 0.1585, 0.109로, 0.05를 기준으로 해석했을 때, setosa만이 정규분포를 따른다고 추정할 수 있습니다.

Levene 테스트 결과 p값은 0.0000000313로 0.05보다 현저히 작기 때문에 species간 등분산성이 없다고 볼 수 있습니다.

## 3. One-way ANOVA & TUKEY HSD

One-way ANOVA를 통해 종별 petal length 평균이 동일하다는 귀무가설을 검정하였다.

H0: 세 species 간 petal\_length 평균은 동일하다.

H1: 적어도 한 species의 petal\_length 평균은 다르다.

ANOVA 결과 p값은 0에 수렴하는 값으로, 0.05보다 작게 나타나 귀무가설을 기각하였습니다. 따라서 species에 따라 petal length 평균이 통계적으로 유의미하게 다르다고 결론지을 수 있습니다.

이에 추가로 Tukey HSD 수행하여 어떤 종 사이에서 평균 차이가 유의미한지 확인한 결과, 다음과 같은 값을 얻었습니다.

```
=====
```

group1	group2	meandiff	p-adj	lower	upper	reject
-----						
setosa	versicolor	2.798	0.0	2.5942	3.0018	True
setosa	virginica	4.09	0.0	3.8862	4.2938	True
versicolor	virginica	1.292	0.0	1.0882	1.4958	True
-----						

모든 쌍에 대해 reject값이 True로, 어떤 경우에도 두 종간의 평균의 차이가 통계적으로 유의미하다는 것을 확인할 수 있습니다.

#### 4. Conclusion

지금까지의 결과를 종합하면, 결과를 종합하면 먼저 Boxplot에서 setosa < versicolor < virginica 순으로 petal\_length 분포가 증가하는 경향 확인했습니다. ANOVA 결과 p-value는 0에 가까운 값으로 종별 평균이 모두 동일하다는 가설은 기각됩니다. 또한 Tukey HSD 결과를 통해 모든 종 사이의 차이가 유의미함을 확인했으므로 species별 petal length에는 통계적으로 유의미한 차이가 존재하며, 전반적으로 virginica가 가장 길고 setosa가 가장 짧은 경향을 가진다고 결론지을 수 있습니다.

#### 5. Linear Regression

회귀분석 결과 다음과 같은 R2, MSE값을 얻었습니다.

MSE: 0.13001626031382682

R2: 0.9603293155857664

높은 R2값과 낮은 MSE값을 통해 petal length는 sepal length, sepal width, petal width 변수로 상당히 잘 예측 가능함을 확인할 수 있습니다.

또한 각 변수에 대한 회귀 계수는 다음과 같습니다.

0	sepal_length	0.722815
1	sepal_width	-0.635816
2	petal_width	1.467524

따라서 petal length는 petal width와의 연관이 가장 크며 예측에 가장 중요한 변수로 볼 수 있습니다. 반면 sepal width는 음의 계수를 보여 폭이 증가할수록 petal length가 줄어드는 관계가 나타났습니다.