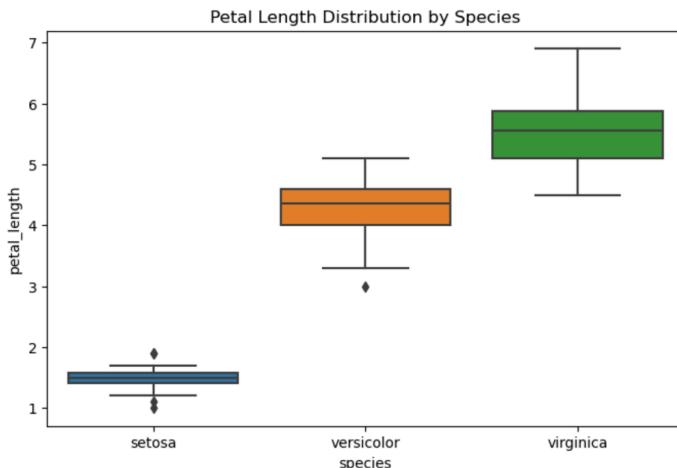


# result

## 시각화

- species 별 petal length 분포 (boxplot)



setosa는 두 종에 비해 petal length가 현저히 짧고 분산 또한 매우 작았다.

versicolor는 중간 정도의 petal length를 보였으며, virginica는 세 종 중 가장 큰 중앙값과 상대적으로 넓은 분포를 나타냈다. 이를 통해 종별 petal length에 차이가 존재함을 시각적으로 확인할 수 있었다.

## 정규성 검정 (Shapiro-Wilk)

H0: 각 species의 petal length는 정규분포를 따른다.

H1: 각 species의 petal length는 정규분포를 따르지 않는다.

setosa: p-value = 0.0548

versicolor: p-value = 0.1585

virginica: p-value = 0.1098

Shapiro-Wilk 정규성 검정을 실시한 결과, setosa의 p-value는 0.0548, versicolor는 0.1585, virginica는 0.1098로 모두 0.05 이상으로 나타났다.

따라서 세 species 모두 petal length가 정규분포를 따른다는 귀무가설을 기각할 수 없으며, 정규성 가정을 만족한다고 판단하였다.

## 등분산성 검정 (Levene)

Levene test p-value: 0.0000

Levene 등분산성 검정을 실시한 결과 p-value가 0.05 미만으로 나타나, 세 species 간 petal length의 분산이 동일하다는 귀무가설은 기각되었다.

다만 과제 지시사항에 따라 이후 분석에서는 등분산성을 만족한다고 가정하고 ANOVA를 진행하였다.

## ANOVA

- one-way anova

```
F-statistic: 1180.1612  
p-value: 2.8568e-91
```

One-way ANOVA 결과, F-statistic는 1180.1612로 나타났으며 p-value는 0.05보다 매우 작게 나타났다( $p < 0.05$ ).

이에 따라 세 species의 petal length 평균이 모두 같다는 귀무가설은 기각되었고, 종별 petal length 평균에는 통계적으로 유의미한 차이가 존재함을 확인하였다.

- turkey HSD 사후검정

```
Multiple Comparison of Means - Tukey HSD, FWER=0.05  
=====  
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 종 쌍(setosa-versicolor, setosa-virginica, versicolor-virginica) 사이에서 petal length 평균 차이가 통계적으로 유의하다.

## 결과 요약

Boxplot 시각화 결과, setosa는 다른 두 종에 비해 petal length가 현저히 짧았으며, virginica는 가장 긴 분포를 보였고 versicolor는 그 중간에 위치하였다.

One-way ANOVA 결과 세 species 간 petal length 평균 차이는 통계적으로 유의미하였으며( $p < 0.05$ ), Tukey HSD 사후검정 결과 모든 종 쌍 간 평균 차이가 유의한 것으로 나타났다.

종합하면 petal length는 setosa가 가장 짧고, virginica가 가장 길며, versicolor는 중간 수준이라고 결론지을 수 있다.

## 회귀 분석

선형 회귀 분석 결과, 모델의 MSE는 0.1300으로 낮은 값을 보였으며

$R^2$ 는 0.9603으로 입력 변수들이 petal length를 매우 잘 설명하는 것으로 나타났다.

회귀계수 분석 결과, petal\_width가 petal\_length에 가장 큰 양의 영향을 미치는 변수였으며, sepal\_length 또한 양의 영향을 보였다. 반면 sepal\_width는 음의 계수를 보여 petal\_length와 음의 관계를 가지는 것으로 나타났다.