# 미국의 년도별 인구수 변화

**년 도 (x)**

# 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 19

인 구 수 (y)

76.1 83.8 92.4 100.5 106.5 115.8 123.1 127.3 132.5 133.4 151.9 165.1 180.0 193.5 204.0 215.5 22



In [10]:

*#* 연습문제 *7 p343, node (15)*

*#* 데이터를 보고 물음에 답하라*! # [* 조건 *]*

*# 1.* 회귀직선식과 상관계수를 구하여라

*# 2.* 비선형 모델을 찾기 위해*, log(*인구수*)* 대 년도*(1900~)*을 그리고 회귀직선과 상관계수를 구하여라

*# 3.* 각각의 모델을 이용해 *2100*년의 인구수를 예측하고 인구수가 *300(millions)*에 이르는 시기를 예측하라

*# 4. (3)*의 정답에 대한 정확성을 논평하라*(???)*

**import** numpy **as** np

**from** sklearn.linear\_model **import** LinearRegression

X **=** np**.**array([1900, 1905, 1910, 1915, 1920, 1925, 1930, 1935, 1940, 1945, 1950, 1955, 1960, 1965, 1970, 1975, 1980, 1985, 1990])**.**resh

Y **=** np**.**array([76.1, 83.8, 92.4, 100.5, 106.5, 115.8, 123.1, 127.3, 132.5, 133.4, 151.9, 165.1, 180.0, 193.5, 204.0,215.5 ,227.2 ,237.9 ,249.

log\_Y **=** np**.**log(Y)

model **=** LinearRegression() model**.**fit(X,Y)

print('1. 회귀직선식과 상관계수를 구하여라') print("Intercept: ", round((model**.**intercept\_), 4))

print("Coe **f**icient: ", round((model**.**coef\_[0]), 4)) correlation\_coe **f**icient **=** np**.**corrcoef(X**.**reshape(1,**-**1), Y)[0][1] print("상관계수 : ", round((correlation\_coe **f**icient), 4))

population\_2100\_modelA **=** model**.**predict(np**.**array([[2100]]))[0] year\_modelA **=** (300 **-** model**.**intercept\_) **/** model**.**coef\_[0]

model**.**fit(X,log\_Y)

print('\n2. 비선형 모델을 찾기 위해, log(인구수) 대 년도(1900~)을 그리고 회귀직선과 상관계수를 구하여라') print("Intercept: ", round((model**.**intercept\_), 4))

print("Coe **f**icient: ", round((model**.**coef\_[0]), 4)) correlation\_coe **f**icient **=** np**.**corrcoef(X**.**reshape(1,**-**1), Y)[0][1] print("상관계수 : ", round((correlation\_coe **f**icient), 4))

population\_2100\_modelB **=** model**.**predict(np**.**array([[2100]]))[0] year\_modelB **=** (300 **-** model**.**intercept\_) **/** model**.**coef\_[0]

print('\n3. 각각의 모델을 이용해 2100년의 인구수를 예측하고 인구수가 300(mi**l**ions)에 이르는 시기를 예측하라') print(f"2100년의 인구수 예측 (모델A) {population\_2100\_modelA}")

print(f"300 mi**l**ian에 인구수가 도달하는 시점 (모델A) : {year\_modelA}\n")

print(f"2100년의 인구수 예측 (모델B) {population\_2100\_modelB}") print(f"300 mi**l**ian에 인구수가 도달하는 시점 (모델B) : {year\_modelB}")

1. 회귀직선식과 상관계수를 구하여라

Intercept: -3593.284

Coefficient: 1.9264

상관계수 : 0.989

1. 비선형 모델을 찾기 위해, log(인구수) 대 년도(1900~)을 그리고 회귀직선과 상관계수를 구하여라

Intercept: -20.2833

Coefficient: 0.013

상관계수 : 0.989

1. 각각의 모델을 이용해 2100년의 인구수를 예측하고 인구수가 300(millions)에 이르는 시기를 예측하라

2100년의 인구수 예측 (모델A) 452.0528070175437

300 millian에 인구수가 도달하는 시점 (모델A) : 2021.0669204568223

2100년의 인구수 예측 (모델B) 6.983727241844431

300 millian에 인구수가 도달하는 시점 (모델B) : 24666.983716788473

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js