



Busan science high school

2023 Ocean ICT Festival

2023 BOIF

B
32

Youtube 영상 QR

회귀분석을 이용 대한민국 각 도시별 연도별 어획량 예측 프로그램

해파리 사냥꾼: 2415 천승원 2416 최윤건

탐구 동기&목적

최근 수산업에 있어 인력 증가, 기술 발달 등 수산업 발전이 이루어졌다. 때문에 과거와 비교해 현재의 수산업 발달 정도를 알아보기 위해 연도별 지역별 어획량 데이터를 분석하고, 추가로 **미래의 어획량을 예측하는 프로그램**을 만들어보았다.

★ 예측 방식으로 추세선을 선택한 이유 : 현재 수산업 기술이 발달하고 있는 중이므로 미래에는 지금과 같이 어획량이 점차 증가할 것이라 예측, 그러나 무한히 증가하지는 않으므로 보다 가까운 미래에서만 유효한 결과를 얻을 수 있음

프로그램 구현(회귀분석)

```
In [19]: import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
import pandas as pd
f = pd.read_csv("어업별_품종별_통계_지역별.csv", encoding = 'cp949') #csv파일의 인코딩이라는 게 있는데 이 파일의 경우 cp949원

yearlist = f.iloc[0, 4:57].tolist()
#년도 데이터 행, 열 일괄해서 리스트로 만들
fishlist = f.iloc[2, 4:57].tolist()
busanfishlist = f.iloc[4, 4:57].tolist()
gyeongdofishlist = f.iloc[11, 4:57].tolist()
gangwandofishlist = f.iloc[12, 4:57].tolist()
chungcheongbukdofishlist = f.iloc[13, 4:57].tolist()
chungcheongnandofishlist = f.iloc[14, 4:57].tolist()
jeollabukdofishlist = f.iloc[15, 4:57].tolist()
jeollanadofishlist = f.iloc[16, 4:57].tolist()
gyeongsangbukdofishlist = f.iloc[17, 4:57].tolist()
gyeongsangnandofishlist = f.iloc[18, 4:57].tolist()
jejudofishlist = f.iloc[19, 4:57].tolist()
#각 지역에 맞는 데이터 행, 열 일괄해서 리스트로 만들

yearlist = [float(item) for item in yearlist]
fishlist = [float(item) for item in fishlist]
busanfishlist = [float(item) for item in busanfishlist]
gyeongdofishlist = [float(item) for item in gyeongdofishlist]
gangwandofishlist = [float(item) for item in gangwandofishlist]
chungcheongbukdofishlist = [float(item) for item in chungcheongbukdofishlist]
chungcheongnandofishlist = [float(item) for item in chungcheongnandofishlist]
jeollabukdofishlist = [float(item) for item in jeollabukdofishlist]
jeollanadofishlist = [float(item) for item in jeollanadofishlist]
gyeongsangbukdofishlist = [float(item) for item in gyeongsangbukdofishlist]
gyeongsangnandofishlist = [float(item) for item in gyeongsangnandofishlist]
jejudofishlist = [float(item) for item in jejudofishlist]
#모든 리스트 안에 요소 실수형으로 바꿈
```

	A	B	C	D	E	F	G	H	I	J	K
1	111										
2	어업별	품종별	행정구역별 판매형태별		1970	1971	1972	1973	1974	1975	1976
3	어업별	품종별	행정구역별 판매형태별 생산량-계	생산량-계	생산량-계	생산량-계	생산량-계	생산량-계	생산량-계	생산량-계	생산량-계
4	계	계	국내계	계	845841	914426	1119434	1325849	1607841	1569386	1682636
5	계	계	서울특별시계		-	-	-	-	-	-	-
6	계	계	부산광역시계	125767	152956	184606	204986	227642	221154	272322	
7	계	계	대구광역시계	-	-	-	-	-	-	-	-
8	계	계	인천광역시계	-	-	-	-	-	-	-	-
9	계	계	광주광역시계	-	-	-	-	-	-	-	-
10	계	계	대전광역시계	-	-	-	-	-	-	-	-
11	계	계	울산광역시계	-	-	-	-	-	-	-	-
12	계	계	세종특별자치계	-	-	-	-	-	-	-	-
13	계	계	강기도	계	82608	86267	89522	102343	117034	123591	113181
14	계	계	강원도	계	101038	74271	92641	105690	116749	101663	130833
15	계	계	충청북도	계	2	12	27	22	31	45	301
16	계	계	충청남도	계	32793	36097	38858	45397	61854	89565	99118
17	계	계	전라북도	계	36086	39450	45778	52967	53857	65275	69988
18	계	계	전라남도	계	170226	190921	202974	272516	414136	350036	389217
19	계	계	경상북도	계	72316	69303	88194	93747	93209	89749	141556
20	계	계	경상남도	계	198547	235781	342955	409818	478611	483807	419537
21	계	계	제주도	계	26458	29369	33680	38364	44719	44501	46584

국가통계포털(KOSIS)에서 어업 별, 지역별 어획량 데이터 다운받음

각 지역별로 데이터 분리해서 실수 형으로 전환, 리스트에 저장

```
print('원하는 지역을 입력하세요.한지역 : 전국, 부산광역시, 경기도, 강원도, 충청북도, 충청남도, 전라북도, 전라남도, 경상북도, 경상남도, 제주
지역 = str(input('지역 : '))
fish = []
if 지역 == '전국':
    fish.append(fishlist)
elif 지역 == '부산광역시':
    fish.append(busanfishlist)
elif 지역 == '경기도':
    fish.append(gyeongdofishlist)
elif 지역 == '강원도':
    fish.append(gangwandofishlist)
elif 지역 == '충청북도':
    fish.append(chungcheongbukdofishlist)
elif 지역 == '충청남도':
    fish.append(chungcheongnandofishlist)
elif 지역 == '전라북도':
    fish.append(jeollabukdofishlist)
elif 지역 == '전라남도':
    fish.append(jeollanadofishlist)
elif 지역 == '경상북도':
    fish.append(gyeongsangbukdofishlist)
elif 지역 == '경상남도':
    fish.append(gyeongsangnandofishlist)
elif 지역 == '제주도':
    fish.append(jejudofishlist)
else:
    print('지역 입력이 잘못되었습니다. 올바른 철자로 입력했는지 확인해주세요.')
```

원하는 지역 입력시 그에 따른 지역 어획량 데이터 불러와서 fish 리스트에 저장, 사용

```
x = []
for i in range(len(yearlist)):
    x.append(yearlist[i]-1970.0) #1970-2022년을 0-52로 표현, 이렇게 해야 회귀분석 볼

x_year = np.array(x).reshape(-1,1) #행렬 만들 때 1열로 설정하면 행 개수는 자동설정 되는 거
y_fish = np.array(fish).reshape(-1,1)

yearfish = LinearRegression().fit(x_year,y_fish)
a=yearfish.coef_[0]
b=yearfish.intercept_

x_year2 = np.array(yearlist)
plt.grid()
plt.scatter(x_year2, y_fish)
plt.plot([1970, 2023], [b, 53*a+b], c='r')
plt.show()

년도 = float(input('년도:'))
estimatedfish = (년도-1970.0)*a+b
print('예상 생산량 : %.2f톤'%estimatedfish)
```

년도를 x축, 어획량을 y축으로 설정해 회귀분석 진행, 그래프 출력

그래프를 통해 미래의 어획량 계산

기대효과

1.미래 어획량을 예측 함으로써 수산업 발전에 기여할 수 있을 것이라고 보인다.

2. 지역별 어획량과 그 지역의 환경을 분석함으로써 그 지역의 생태계에 맞는 방법을 택하여 생태계를 효과적으로 보존할 수 있을 것이다.



원하는 지역을 입력하세요.
지역 : 전국, 부산광역시, 경기도, 강원도, 충청북도, 충청남도, 전라북도, 전라남도, 경상북도, 경상남도, 제주도
서울특별시, 대구광역시, 인천광역시, 광주광역시, 대전광역시, 울산광역시, 세종특별자치시는 데이터의 양이 적어 유의미한 결과를 도출할 수 없음
지역 : 충청북도

