SWING 해커톤 대본

#2

다음은 데이터 분석 부분입니다.

The following is the data analysis section.

적조 데이터와 전력 데이터를 분석하기 수월하게 전처리를 해줍니다.

Pre-processing red tide data and power data for easy analysis.

#3

Red\_tide\_data.csv에 있는 컬럼 값을 확인하고, 결측치를 처리해줍니다.

Check the value of the column in Red\_tide\_data.csv and process the missing values.

우선 컬럼은 발생날짜, 조사해역, 원인생물, 생물밀도의 max,min, 수온의 max,min, 속보코드가 존재합니다.

First of all, the column contains the occurrence date, irradiation sea area, causative organism, max, min of biological density, max, min of water temperature, and breaking news code.

원인생물의 종류를 파악하기 위해서 결측값을 확인했습니다.

To determine the type of causative organism, we checked the missing value.

오직 하나만 결측값이 존재해서, 삭제처리를 해주었습니다.

There was only one missing value, so they deleted it.

#4

원인생물 중에서 가장 많이 발견되는 종을 파악하기 위해서 바차트를 사용해서 집계를 했습니다.

In order to identify the most common causative species, we used a bar chart to aggregate them.

그 결과 코클로디니움이라는 종이 압도적으로 많이 발견되었습니다.

As a result, a species called coclodinium was found overwhelmingly.

그래서 해당 종을 집중해서 분석해보기로 결정했습니다.

So we decided to focus on analyzing that species.

#5

코클로디니움 데이터프레임을 구성하고, 관측해안 컬럼값을 활용해서 시군구 컬럼을 만들어주었습니다.

The Cochlodinium data frame was constructed, and the city, county, and district columns were created using the observed coastal column values.

그리고 생물밀도와 수온의 max,min값을 활용해서 mean값의 컬럼도 만들어주었습니다.

And I made a column of mean values using the max and min values of biological density and water temperature.

그 후 앞에서 한것과 같은 방법으로 결측치를 확인하고, 삭제하는 방법으로 처리했습니다.

After that, it was processed by checking and deleting missing values in the same way as previously done.

#6

시간별 일일 전력 전력량이 있는 데이터입니다.

Data with daily power by hour.

시간당 사용량의 합계를 구하여 일별 전력 사용량으로 변경하여 분석합니다.

Obtain the sum of hourly usage and analyze it by changing it to daily power usage.

그리고 집계된 데이터는 day\_power 열에 저장되어 tableau에 사용되었습니다.

The aggregated data was stored in the day\_power column and used for tableau.

#7

2015년부터 2018년 달마다 일별 전력 사용량을 시각화 하여 살펴보았습니다.

From 2015 to 2018, we visualized the daily power usage every month.

많은 전력을 사용하는 달과 적조가 발견된 시기가 겹친다는 사실을 알아냈습니다.

We found that the moon, which uses a lot of electricity, overlaps with the time when red tide was found.

그래서 전력과 적조의 연관성을 찾기 위해 분석을 진행했습니다.

So we conducted an analysis to find the connection between power and red tide.

#8

다음은 적조와 수온에 대한 상관관계 분석입니다.

The following is a correlation analysis of red tide and water temperature.

논문에 의한 내용을 직접 데이터 분석을 통해 관계성을 밝혀내봅니다.

The contents of the paper are directly analyzed to reveal the relationship through data analysis.

#9

수온과 생물학적 밀도의 상관관계를 알아보겠습니다.

Let's look at the correlation between water temperature and biological density.

25도 근처에서 가장 많이 발견된다는 고려하여 24~26도로 구간을 지정하고, 구간 내외로 구분하여 생물학적 밀도의 상관관계를 확인합니다.

Considering that it is most commonly found near 25 degrees, we designate an interval between 24 and 26 degrees, and check the correlation of biological density by dividing it into and out of the interval.

위쪽 그림에서 초록색은 구간 외의 데이터이고, 파란색은 구간 내의 데이터입니다.

In the upper figure, green is the data outside the interval, and blue is the data within the interval.

자세히 살펴보면 24~26도 사이의 데이터들에 의하면 수온과 적조의 생물밀도는 양의 상관관계를 지니고 있음을 알 수 있습니다.

If you look closely, the data between 24 and 26 degrees show that there is a positive correlation between water temperature and red tide density.

즉, 특정 온도구간 내에서 수온이 올라가면 적조의 생물밀도가 증가함을 확인했습니다.

In other words, we found that the biological density of red tides increases when the water temperature rises within a certain temperature range.

생물밀도가 높다는 뜻은 적조가 많이 발생했다는 뜻입니다.

High biological density means a lot of red tides.

그러므로 적조가 발생하는 온도구간에서는 온도가 높아짐에 따라 적조가 많이 발생함을 증명합니다.

Therefore, we prove that in the temperature zone where red tide occurs, red tide occurs more as the temperature increases.

의의 : (Significance)

적조가 발견된 날에 대해 전력 사용이 많을 것을 예측할 수 있다.

 It can be predicted that there will be a lot of electricity use on the day the red tide is found.

그리고 추가적인 연구나 개발을 통해 적조 발생 시 전력사용량과 관련된 적절한 경고 및 제어 시스템을 구현하는 것에도 도움을 줄 수 있습니다.

Further research or development can also help implement appropriate warning and control systems related to power usage in the event of red tide.

이를 통해 블랙아웃 현상과 같은 과한 자원사용으로 인한 문제를 다루는 것에 기여할 수 있습니다.

This can contribute to dealing with problems caused by excessive resource use, such as blackouts.

한계(Limitations) :

공개된 최신 적조데이터가 부족했다.

There was a lack of the latest red tide data released.

또한 적조 발생 지점이 해안이라는 점에서, 경계가 명확하지 않아 지도 시각화가 어려웠다.

In addition, in that the point of occurrence of red tide is the coast, it was difficult to visualize the map because the boundary was not clear.

아두이노 키트를 통해 구현한 간단한 시스템이기에 관측을 통한 제어시스템까지의 작동을 구현하기 어려웠다.

Since it is a simple system implemented through the Arduino kit, it was difficult to implement the operation of the control system through observation.

더 많은 데이터나, 개발 환경이 주어진다면 나아가 예측시스템까지의 구현을 해보고 싶었다.

Given more data or development environment, I wanted to implement a prediction system.