

빅데이터 분석을 위한 파이썬 입문 기말고사 (20241519 조예성)

iterable 자료형을 주면 모아서 하나의 iterable로 만들어 주는 함수: zip()

- 푸는 방법: zip(*data)

반복가능 자료형을 묶어 반복가능 자료형을 만들 수 있는데 이를 칭하는 용어: 집적화

값과 키를 갖는 연속형 자료구조: 딕셔너리

딕셔너리 대표 메소드 3개: .keys(), .values(), .items()

- 키와 값을 튜플로 반환: .items

※딕셔너리에서 ['키'] 에 키가 없을 경우 오류

이름이 없는 함수: 람다함수

빈 집합 생성: set()

for문 작동시 정렬: sorted(set())

텍스트 데이터 다루는 도구: 정규식

문자열 분할 메소드: .split()

리스트 형태 문자열 합체 메소드: '문자'.join(문자열 리스트)

특정 문자(열) 횟수: str.count(char)

특정 문자(열)로 시작, 끝 검사: startswith(), endswith()

좌우 문자열 공백 제거: str.strip()

- 왼쪽만, 오른쪽만, lstrip(), rstrip()

부분 문자열 찾는 메소드: str.find()

- 찾고자 하는 문자열의 시작 위치를 반환한다
- 찾고자 하는 문자열이 없으면? -1반환

raw 문자열 출력: r'문자열'

문장을 문자, 단어, 어휘 집합 단위로 나누는 작업: 토큰화

string에서 워드클라우드 생성: wordcloud.generate(str)

차트 그리는 라이브러리: matplotlib

선그래프: plt.plot

막대그래프: plt.bar

산점도: plt.scatter()

파이그래프: plt.pie()

도수분포도: plt.hist()

효율적인 데이터분석 위한 라이브러리: pandas

2x2 서브플롯 만들기: fig, axs = plt.subplots(2,2, figsize = 5,5)

- axs[0,0].hist()

지도 시각화 라이브러리: folium

기본 지도 그리기: map = folium.Map(location = [x, y], zoom_start = n, control_scales = False)

마커 추가하기: folium.marker([경도, 위도]).add_to(지도)

- 마커 속성 이외 4가지: popup, tooltip, icon = folium.Icon(color..)

원형 마커 추가하기: folium.Circle()

마커 군집화 import: from folium.plugins import MarkerCluster

- 지도에 군집 적용: mc = MarkerCluster.add_to(map)
- 클러스터는 mc에 add_to