|  |  |
| --- | --- |
| 교육 제목 | **python 기본** |
| 교육 일시 | 2021.11.08 |
| 교육 장소 | 영우글로벌러닝 |
| **교육 내용** | |
| 오전 | 각 열의 데이터 개수 세기  열 데이터 개수 확인  DF.객체.count()  열 데이터의 고유 값 개수  DF객체[‘열 이름’].value\_count()  모든 평균 값  DF.객체.mean()  특정 열의 평균값  DF객체[“열 이름”].mean()  같은 형식의 중간값, 최대값 최소값  mean()자리에 median(),max(),min() 사용  표준편차와 상관계수  모든 열 : DF객체.std()  특정 열 : DF객체[“열 이름”].std()  상관계수도 같은 형식에 .corr()사용  # 상관계수는 최소 두개의 열 필요  그래프  DF.객체.plot(kind=”hist, bar, barh ... 등“)으로 이용 |
| 오후 | 시각화 도구  Matplotlib = 모듈  선 그래프  선 그래프는 연속하는 데이터 값들을 직선 또는 곡선으로 연결해여 데이터 값 사이의 관계를 나타냄  그래프 꾸미기  ‘o’ : 선그래프가 아닌 점 그래프로 표현  marker= ‘o’ : 마커 모양(‘o’, ‘+’, ‘\*’, ‘.’)  markerfacecolor= ‘원하는 색’ : 마커 배경색  markersize = 값 : 마커 크기  color = ‘원하는 색’ : 선의 색  linewidth = 값 : 선의 두께  label = ‘라벨’ : 라벨 지정  면적 그래프  면적 그래프는 각 열의 데이터를 선으로 구현하는데 선 그래프와 x축 사이의 공간에 색이 입혀진다.  색의 투명도는 alpha로 변경가능  막대 그래프  막대 그래프는 데이터 값의 크기에 비례하여 높이를 갖는 직사각형 막대로 표현한다.  히스토그램  히스토그램은 변수가 하나인 단변수 데이터의 빈도수를 그래프로 표현한다.  산점도  서로 다른 두 변수 사이의 관계를 나타낸다, 각 변수는 연속되는 값을 갖는다.  파이차트  원을 파이조각처럼 나누어서 표현한다.  박스 플롯  범주형 데이터의 분포를 파악하는데 적합하다. |