|  |
| --- |
| **SSAC**  **파이썬을 활용한 비즈니스 빅데이터 분석가 양성 과정**  **(TEMPLATE) 자가학습 워크시트** |

|  |  |
| --- | --- |
| **이름** | 백승재 |
| **날짜** | 2021-11-09 |
| **온라인 콘텐츠(인강)**  **학습 범위** | * 직장인을 위한 Python : ch01\_04~11 |
| **복습 노트 작성 (자유 양식)** | Ch 04: 다중 그래프-fig.axes,subplots  - subplot(row,column,index), 각 옵션 사이 콤마 생략 가능  - 여러 개의 subplot 그리기(subplot set그리기) : fig,axes = plts.ubplots(행의 개수, 열의 갯수)  Ch 05: 기본 그래프 스타일링  - tiltle: plt.title(‘title name’,fontsize=)  - X축, Y축 라벨 설정 : plt.xlabel (‘label name’,fontsize)  - tick 설정 : xtick(rotation=90 or 30), 날짜 데이터인 경우  - 범례 (legend) 설정 : plt.legend()  - x,y limit 설정 : plt.xlimt(st,end)  - 스타일 세부 설정 : 마커(maker), 라인(linestyle), 컬러(color), 투명도(alpha), 그리드 설정(plt.grid)  Ch 06: 그래프 이미지를 저장 및 활용하기: plt.savefig(‘filename.png’,dpi = )  Ch 07 : Scatterplot  - scatterplot option : s: Marker size, c: Marker, cmap : c에 값을 컬러맵에 맞춰 색깔 부여, alpha : 투명도  Ch 08. Barplot,barh  - bar option : align: bar 정렬, alpha : 투명도, color: 컬러 설정  - horizontal type : barh  - 비교 그래프 설정하기 : width 설정 -> axes.bar(x-width/2,y\_1,width), axes.bar(x+width/2,y\_1,width)  Ch 09. Lineplot : plt.plot(x,y)  Ch 10.Areaplot :  - 경계선 상/하를 구분지어 시각화 할 때 사용 : plt.fill\_between () & plt.plot()  - 여러 line plot 겹치는 부분이 있는 경우 : plt.fill\_between() 여러 개 그리고 투명도를 차등하여 옵션에 적용해줌 |
| **퀴즈 제출** | Q1. 다음은 x, y에 대해 산점도로 표현한 그래프이다. 해석으로 적절한 것은?  x = 2 \* np.random.rand(n)  y = 2 \* np.pi \* np.random.rand(n)    1) subplot형태이며 각 plot은 (3,1,1), (3,1,2), (3,1,3)에 해당한다.  2) 산점도는 데이터의 분포를 시각화할 때 사용하며 시각화 시 데이터 x,y를 따로 설정하지 않아도 된다.  3) 산점도의 option 중 cmap을 사용하면 조합되어 있는 색들을 c에 값들을 지정한 cmap list에 맞춰 변형해서 색깔을 적용할 수 있다.  4) 컬러 바는 plt.legend()로 설정할 수 있다.  정답 3  1) subplot형태이며 각 plot은 (1,3,1), (1,3,2), (1,3,3)에 해당한다.  2) 산점도는 데이터의 분포를 시각화할 때 사용하며 시각화 시 데이터 x,y를 따로 설정해야함, plt.scattter(x,y)  4) 컬러바는 plt.colorbar()로 설정할 수 있다.  Q2. 다음은 수강 과목(x)에 대한 두 학생(y\_1, y\_2)의 성적을 비교한 막대 그래프에 대한 설명으로 적절하지 않은 것은?    1) 상기 그래프는 막대그래프로 넓이 지정과 넓이 설정을 해줘야 한다.  2) 막대그래프는 barh타입으로 ytick은 x로 설정해야 한다.  3) 수강 과목은 범주형 변수로 x 축 데이터를 범주형 변수의 길이를 array형태로 변경해줘야 한다.  4) 비교 막대 그래프를 통해 두 학생의 평균을 알 수 있다.  정답 : 4 box plot으로 시각화하면 두 학생의 평균을 알 수 있다.  Q3. 다음은 세 개의 곡선 함수를 영역 그래프로 시각화 것이다. 설명으로 적절하지 않은 것은?  x = np.arange(1, 10, 0.05)  y\_1 = np.cos(x)+1  y\_2 = np.sin(x)+1  y\_3 = y\_1 \* y\_2 / np.pi    1) 각각의 함수에 대해 plt.fill\_between을 사용하여 시각화 한 것이다.  2) 세 개의 그래프는 fill\_beween option 중 color를 각각 다르게 설정 하였다.  3) 세 개의 그래프를 동시에 표현하는 경우 투명도를 동등하게 설정해야 시각적으로 뚜렷하게 보인다.  4) 영역그래프를 사용할 때 x,y값이 필요하다.  정답 : 3 세 개의 그래프를 동시에 표현하는 경우 그래프가 겹치는 부분은 색이 진해지기 때문에 투명도 (=alpha)를 다르게 설정한다. |
| **교육생 사진 (교육생 모습 + 교육 화면)** |  |
| **화면 캡쳐**  **(교육 화면 + 교육일시)** |  |