

## TNT 겨울방학 스터디 - 1주차 과제

1. 드라이브에 제시된 Climate Change 파일을 사용하여 다음 문제를 해결하세요.

Climate Change 파일은 1958-03-06 ~ 2016-12-06간 특정 도시의 기온과 CO2 변화량을 나타낸 데이터프레임입니다.

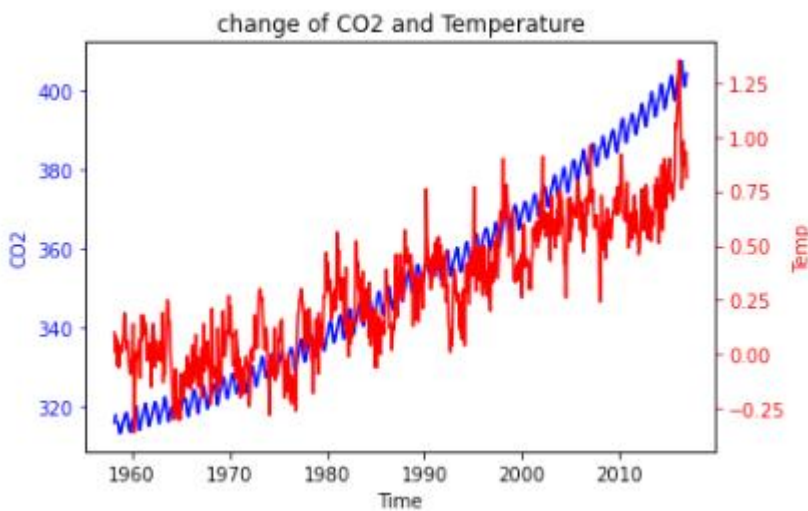
각 열이 나타내는 정보는 다음과 같습니다.

date : 해당 일자

co2 : 해당 일자의 co2 농도

relative\_temp : 해당 일자의 상대온도

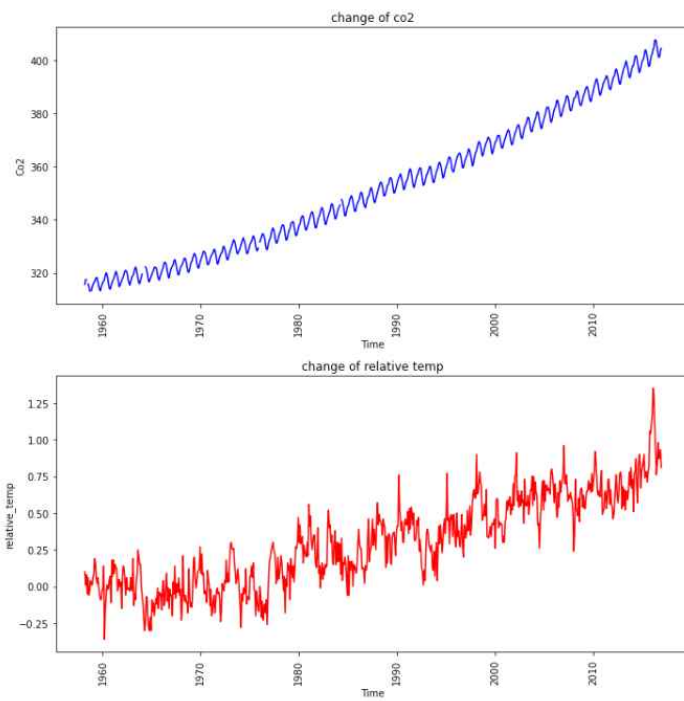
1) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.



가이드라인

- ① 파란색 그래프는 시간에 따른 co2 농도 변화, 빨간색 그래프는 시간에 따른 상대온도 변화를 나타냅니다.
- ② x축은 Time, 왼쪽 y축은 CO2, 오른쪽 y축은 Temp로 labeling 해야 합니다.
- ③ 제목은 change of CO2 and Temperature로 해주세요.
- ④ 각 y축의 label은 그래프색과 동일하게 해주세요.
- ⑤ 주어진 climate change의 전체 데이터를 사용해주세요.

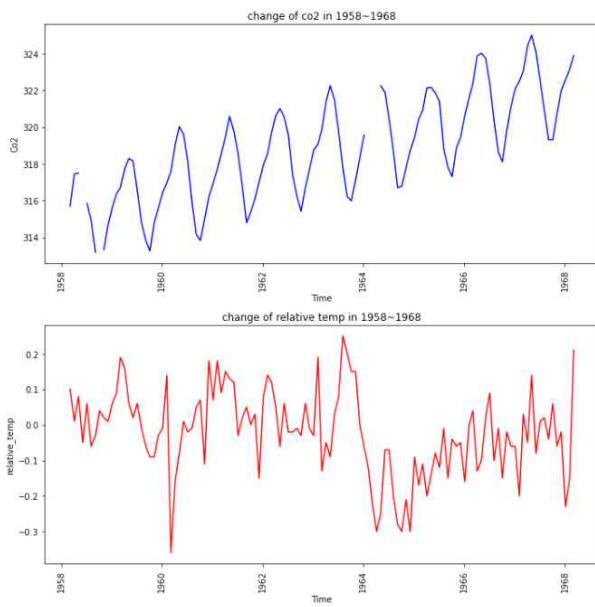
2) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.



가이드라인

- ① 첫 번째 그래프는 시간에 따른 CO2 농도의 변화, 두 번째 그래프는 시간에 따른 상대온도의 변화를 나타냅니다.
- ② 각각의 그래프에 존재하는 x축, y축 labeling, title을 제시해주세요.
- ③ x축의 labeling의 경우 90도로 회전시켜주세요.
- ④ 그래프끼리 겹치면 안 됩니다.

3) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.



가이드라인

- ① 현재 데이터는 1958-03-06에서 1968-03-06까지의 CO2와 상대온도 변화를 나타낸 그래프입니다.
- ② 각각의 그래프에 존재하는 x축, y축 labeling, title을 제시해주세요.
- ③ x축의 labeling의 경우 90도로 회전시켜주세요.
- ④ 그래프끼리 겹치면 안 됩니다.

4) 위의 세 그래프를 이용하여 추론할 수 있는 내용을 서술해주세요.

## 2. 드라이브에 제시된 mpg 데이터를 이용하여 다음 문제를 해결하세요.

mpg 데이터는 여러 차량의 mpg(연비)와 이와 관련된 여러 자료들이 있는 데이터프레임입니다.

각 열이 나타내는 정보는 다음과 같습니다.

mpg : 차량의 연비

cylinders : 차량의 엔진 분류(실린더 수)

displacement : 배기량

horsepower : 마력

weight : 차량무게

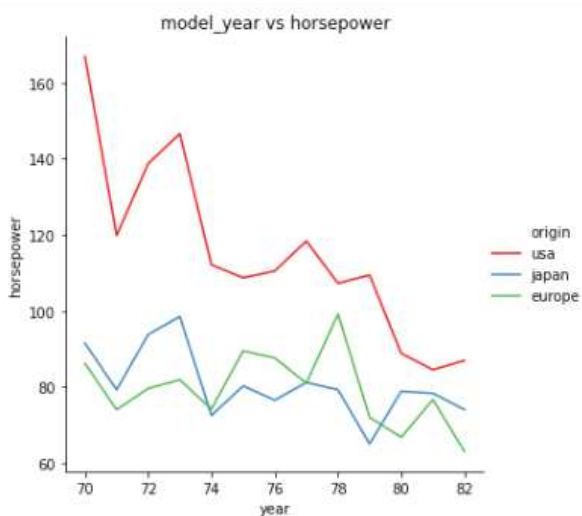
accerleration : 차량 가속

model\_year : 차량 모델 제작년도

origin : 차량 모델 제작 국가(회사가 속한 국가)

name : 차량 이름

1) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.

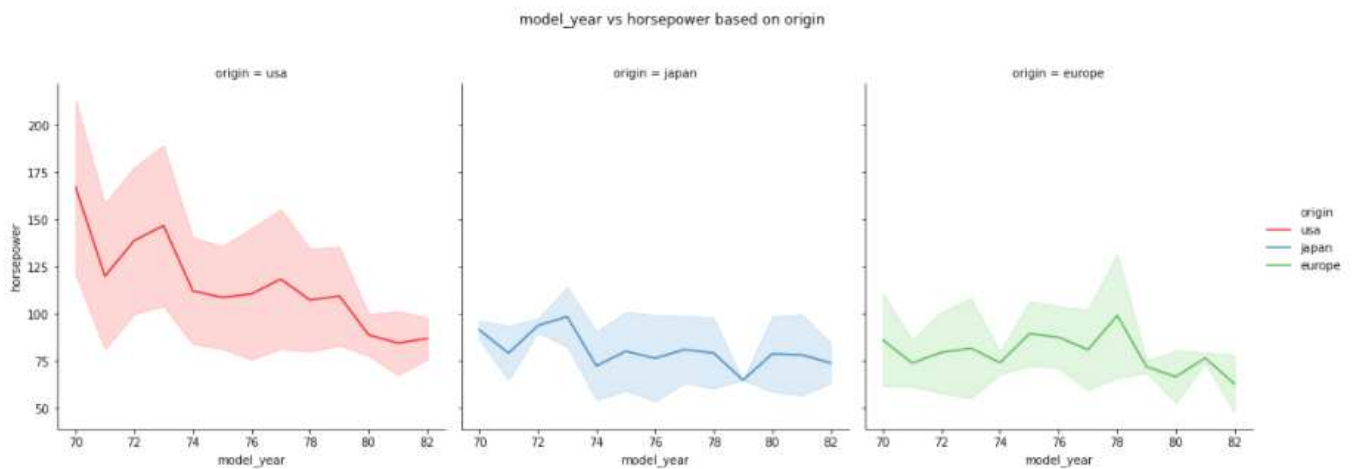


가이드라인

- ① x축은 model\_year, y축은 horsepower를 이용하여 만들어주세요.
- ② 해당 년도의 차량들의 horsepower의 평균값을 선으로 나타내주세요.
- ③ origin을 구분하여 그래프를 그려주세요.
- ④ 그래프의 색에 따라 origin이 무엇인지 표시해주세요.
- ⑤ x축, y축 label, title을 위의 그래프와 같이 작성해주세요.
- ⑥ 그래프 색은 자유롭게 제시해도 괜찮습니다.

(1) 다음 그래프에서 얻을 수 있는 정보는 무엇인가요?

2) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.

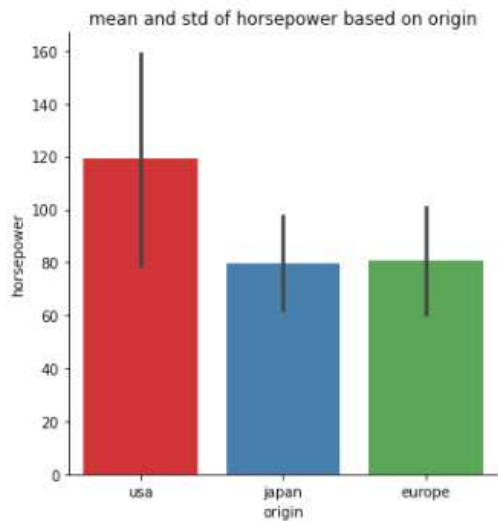


가이드라인

- ① x축은 model\_year, y축은 horsepower를 이용해주세요.
- ② 해당 년도 차량들의 horsepower의 평균값을 진한선으로 나타내주세요
- ③ 연한색으로 칠해진 구역은 해당 년도 차량의 horsepower 표준편차를 나타냅니다.
- ④ 그래프의 색에 따라 origin이 무엇인지 표시해주세요.
- ⑤ x축, y축 label, title을 위의 그래프와 같이 작성해주세요.
- ⑥ 그래프 색은 자유롭게 제시해도 괜찮습니다.

(2) 다음 그래프에서 얻을 수 있는 정보는 무엇인가요?

3) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.

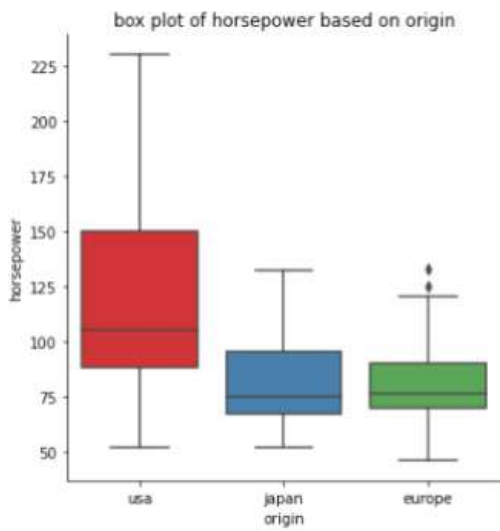


가이드라인

- ① origin에 따른 horsepower의 평균과 표준편차를 제시해주세요.(평균은 색이 있는 막대로, 표준편차는 검은 색 얇은 막대로 표시하면 됩니다.)
- ② x, y label과 title을 다음과 같이 설정해주세요.
- ③ 그래프 색은 자유롭게 제시해도 괜찮습니다.

(1) 위 그래프에서 얻을 수 있는 정보는 무엇인가요?

4) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.



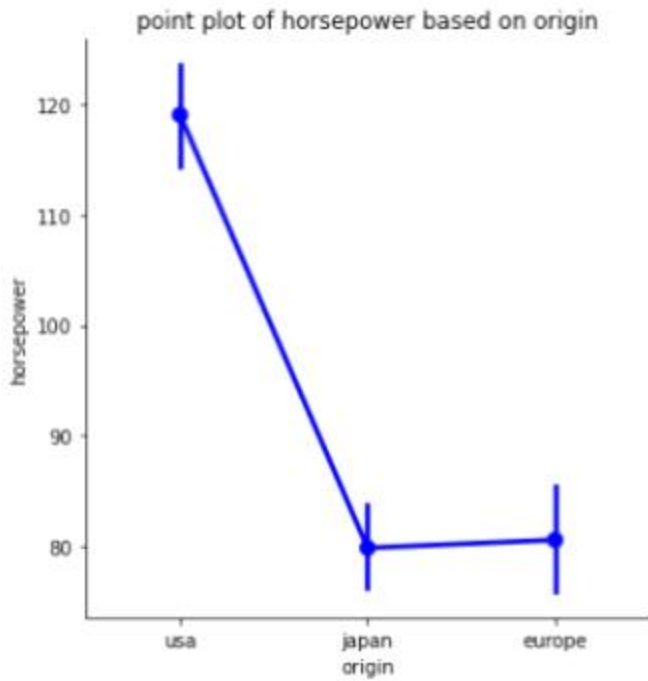
가이드라인

- ① origin에 따른 horsepower의 boxplot을 제시해주세요.
- ② x, y label과 title을 다음과 같이 설정해주세요.
- ③ 그래프 색은 자유롭게 제시해도 괜찮습니다.

(1) 위의 그래프에서 얻을 수 있는 정보는 무엇인가요?

(2) box plot에서 그래프 요소가 무엇을 의미하는지 작성해주세요.(ex 상자안 검은색 선, 상자 밖 점들 등등..)

4) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.



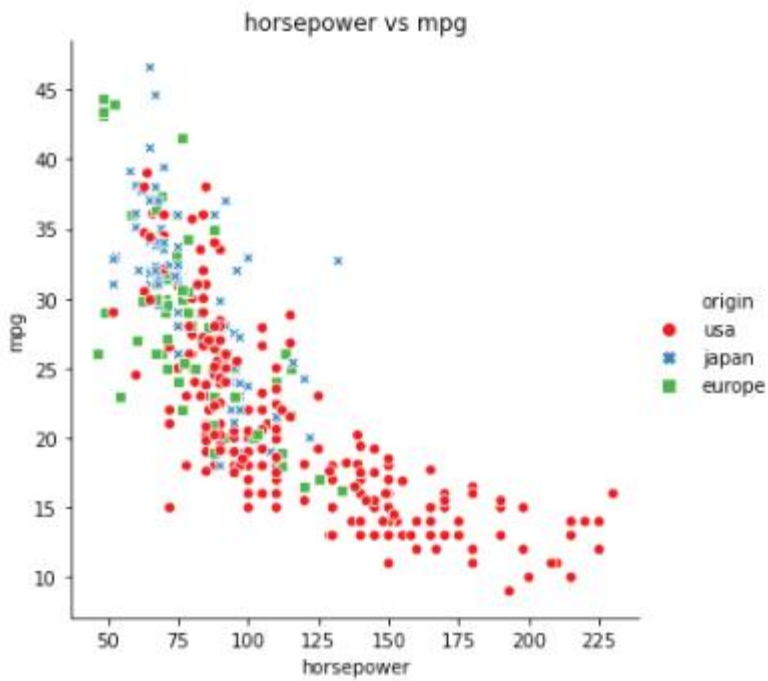
가이드라인

- ① origin에 따른 horsepower의 평균을 점으로, 평균에 대한 95% 신뢰구간을 세로 막대로 제시해주세요.
- ② x, y label과 title을 다음과 같이 설정해주세요.
- ③ 그래프 색은 자유롭게 제시해도 괜찮습니다.

(1) 위의 그래프에서 얻을 수 있는 정보는 무엇인가요?



5) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.



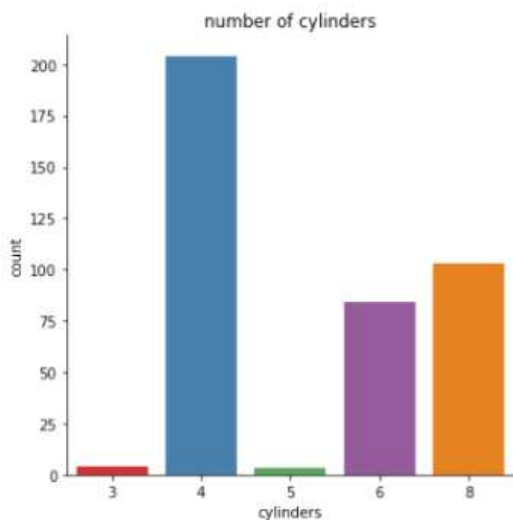
가이드라인

- ① horsepower와 mpg에 대한 산점도를 제시해주세요.
- ② origin에 따라 점의 색과 모양을 다르게 설정해주세요.
- ③ x, y label, 제목을 다음과 같이 설정해주세요
- ④ 점의 모양과 색에 따라 origin이 무엇인지 표시해주세요.

(1) 위의 그래프에서 얻을 수 있는 정보는 무엇인가요?

6) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.

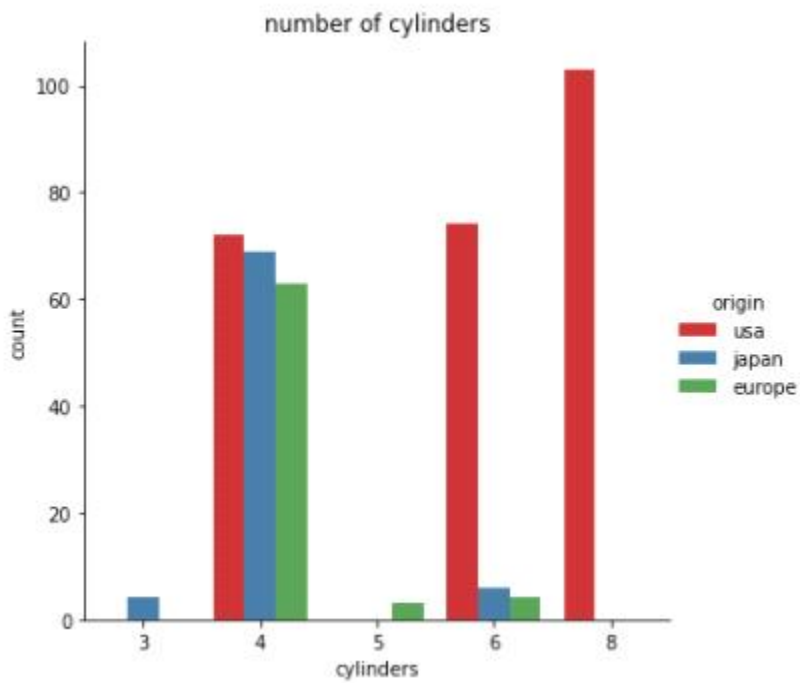
(1)



가이드라인

- ① cylinder 수에 따른 자동차 수를 나타낸 그래프를 만들어주세요.
- ② x, y label, 제목을 다음과 같이 설정해주세요.
- ③ 그래프 색은 자유롭게 제시해도 괜찮습니다.

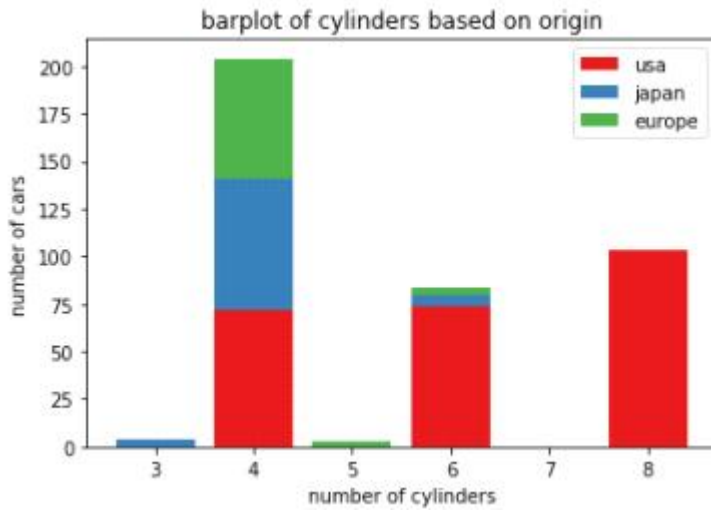
(2)



가이드라인

- ① cylinder 수에 따른 자동차 수를 나타낸 그래프를 만들어주세요.
- ② origin에 따라 자동차 수를 구분해주세요.
- ③ x, y label, 제목을 다음과 같이 설정해주세요.
- ④ 색에 따라 origin이 무엇인지 표시해주세요.
- ⑤ 그래프 색은 자유롭게 제시해도 괜찮습니다.

(3)



가이드라인

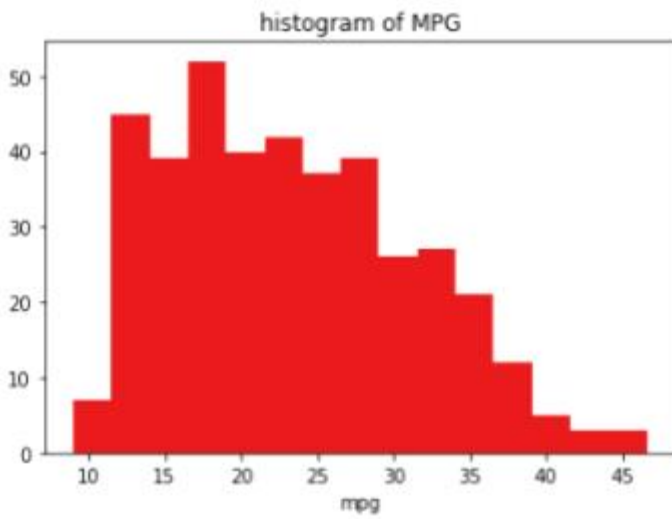
- ① cylinder 수에 따른 자동차 수를 나타낸 그래프를 만들어주세요.
- ② usa - japan - europe 순으로 cylinder 별 자동차 수를 쌓아서 그래프를 만들어주세요.
- ③ x, y label, 제목을 다음과 같이 설정해주세요.
- ④ 색에 따라 origin이 무엇인지 표시해주세요.
- ⑤ 그래프 색은 자유롭게 제시해도 괜찮습니다.

(4) 각 그래프에서 얻을 수 있는 정보는 무엇인가요?

(5) 3개의 그래프가 가지는 장단점은 무엇인가요?

7) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.

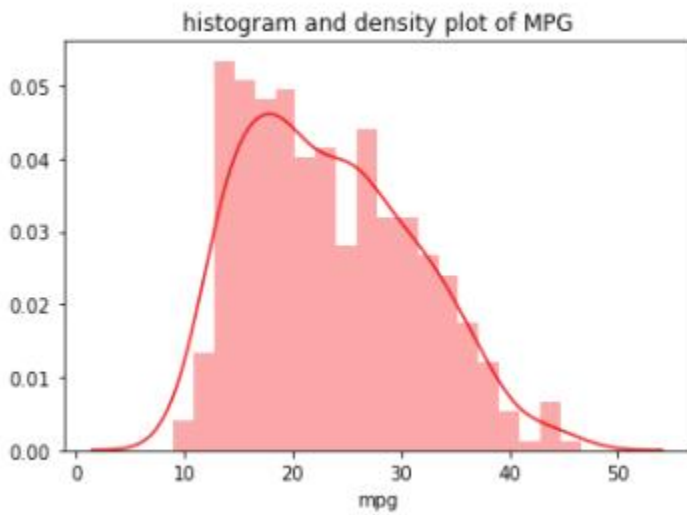
(1)



가이드라인

- ① mpg에 대한 히스토그램을 그려주세요.
- ② 제목과 x label을 다음과 같이 설정해주세요.
- ③ y축은 count로 제시해주세요.
- ④ 히스토그램 색은 자유롭게 설정해도 괜찮습니다.

(2)

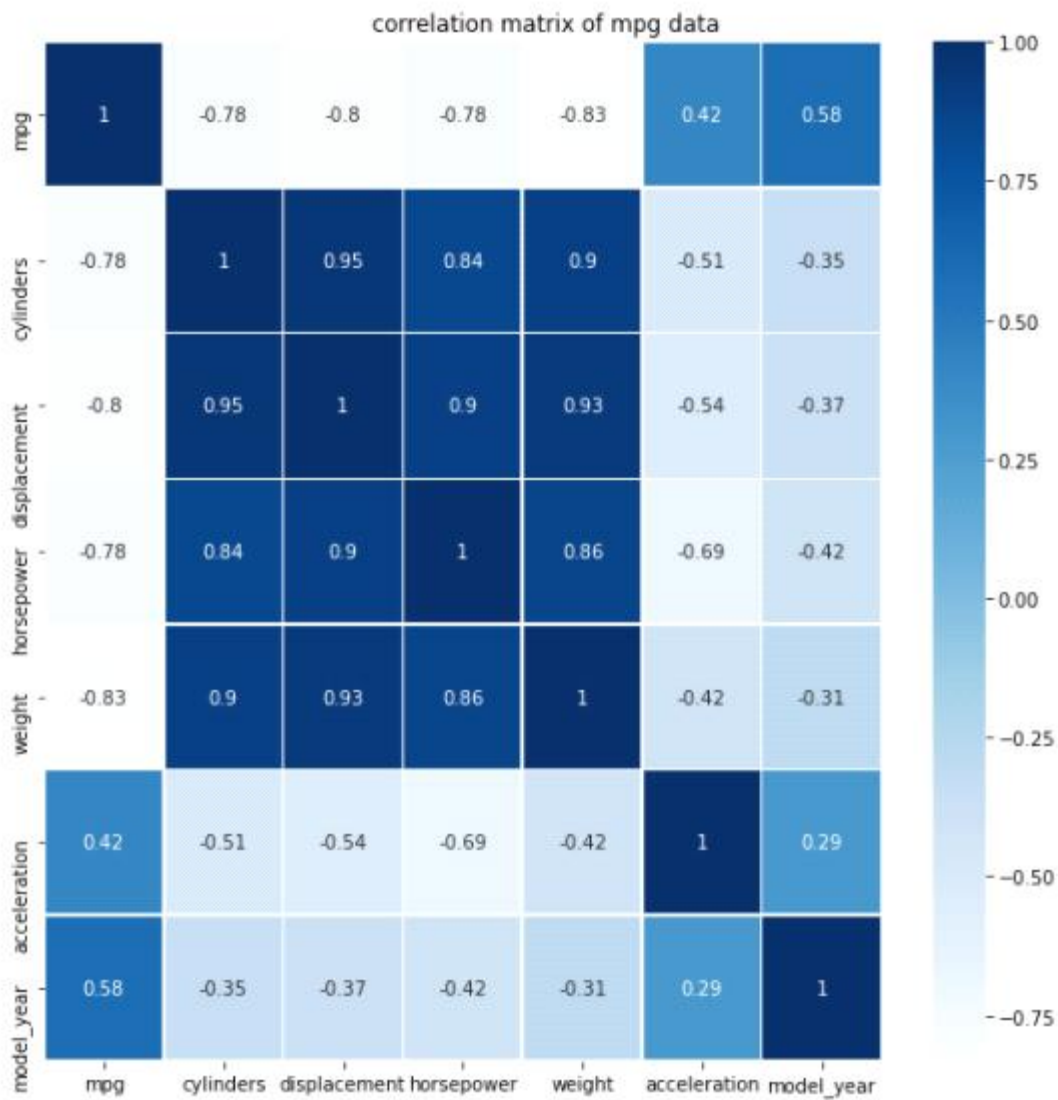


가이드라인

- ① mpg에 대한 히스토그램과 density plot을 그려주세요.
- ② 제목과 x label을 다음과 같이 설정해주세요.
- ③ y축은 frequency로 제시해주세요.
- ④ 히스토그램 색은 자유롭게 설정해도 괜찮습니다. 다만 히스토그램은 음영처리를 해주세요.

(3) 다음 그래프에서 얻을 수 있는 정보는 무엇인가요?

8) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.

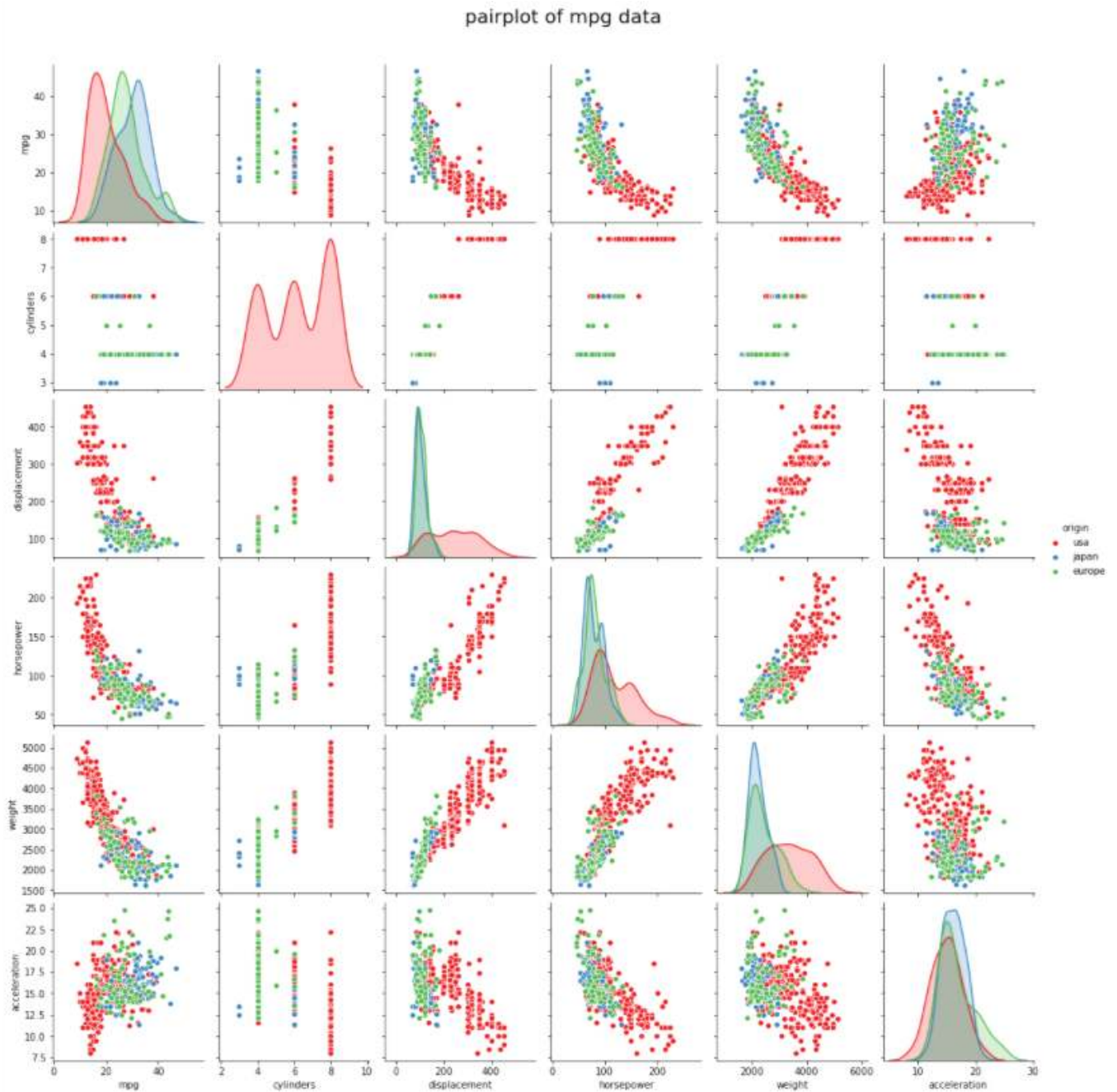


가이드라인

- ① mpg, cylinders, displacement, horsepower, weight, acceleration, model\_year에 대한 상관계수 matrix를 만들고 이를 이용하여 color heatmap을 만들어주세요.
- ② 각 칸에 상관계수 수치를 제시해주세요
- ③ 색에 따라 값이 얼마인지 알 수 있게 제시해주세요. (오른쪽에 나타낸 bar)
- ④ 색은 자유롭게 사용해도 괜찮습니다.

(1) 위의 자료에서 얻을 수 있는 정보는 무엇인가요?

9) 위의 데이터프레임을 이용하여 다음과 같은 그래프를 만드세요.



가이드라인

- ① 변수로 mpg, cylinders, displacement, horsepower, weight, acceleration을 이용해주세요.
- ② origin별로 구분하여 그래프를 그려주세요
- ③ 대각성분에 해당하는 그래프는 density plot을, 이외 성분에 해당하는 그래프는 산점도를 그려주세요.
- ④ 색에 따라 origin이 무엇인지 표시해주세요.
- ⑤ 색은 자유롭게 사용해도 괜찮습니다.