

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

```
import sys
from matplotlib import pyplot as plt
from pandas import read_excel
import seaborn as sb
import numpy as np
```

2. 데이터 가져오기

```
mpg_df = read_excel("https://data.hossam.kr/D01/mpg.xlsx", index_col="id")
mpg_df
```

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	
id										
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	
id										
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	
...	
230	volkswagen	passat	2.0	2008	4	auto(s6)	f	19	28	
231	volkswagen	passat	2.0	2008	4	manual(m6)	f	21	29	
232	volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	
233	volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	
234	volkswagen	passat	3.6	2008	6	auto(s6)	f	17	26	

234 rows × 11 columns

#03. 데이터 전처리

```
df = mpg_df.astype({'manufacturer': 'category', 'model': 'category', 'year': 'category', 'cyl': 'category', 'trans': 'category', 'drv': 'category'})
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

```
'fl': 'category', 'class': 'category'})
```

df.dtypes

```
manufacturer    category
model            category
displ            float64
year            category
cyl              category
trans            category
drv              category
cty              int64
hwy              int64
fl               category
class            category
dtype: object
```

4. 그래프 전역 설정

변수 = 참인경우_할당될_값 if 조건식 else 거짓인경우_할당될_값

```
plt.rcParams['font.family'] = "AppleGothic" if sys.platform == "darwin"
plt.rcParams["font.size"] = 10
plt.rcParams["figure.figsize"] = (15, 5)
plt.rcParams['axes.unicode_minus'] = False
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

#02. 문제 풀이

문제1

자동차 제조사별로 출시한 자동차 수를 하나의 그래프로 시각화 하시오.

catplot() 함수 사용

```
g = sb.catplot(data=df, x='manufacturer', kind='count')

g.fig.set_figwidth(15)
g.fig.set_figheight(5)

plt.title("제조사별 자동차 모델 수")
plt.show()
plt.close()
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

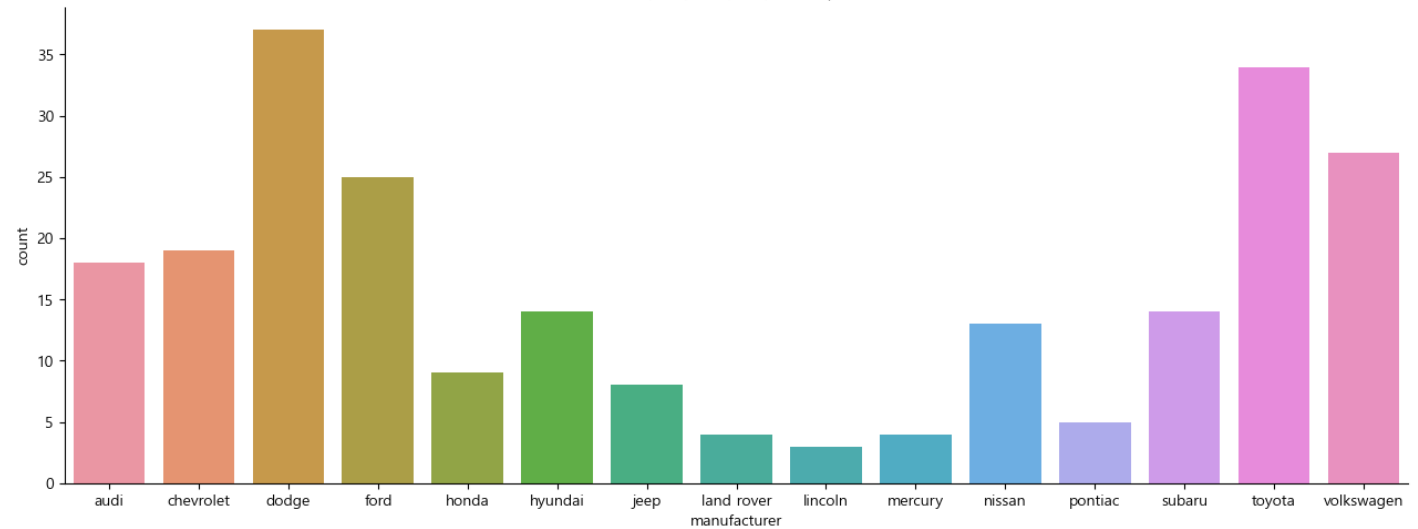
문제4

barplot() 함수 사용

문제5

연습문제(3)_풀이.ipynb

제조사별 자동차 모델 수



countplot() 함수 사용

```
sb.countplot(data=df, x="manufacturer")
plt.title("제조사별 자동차 모델 수")
plt.show()
plt.close()
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조
2. 데이터 가져오기
- #03. 데이터 전처리
4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

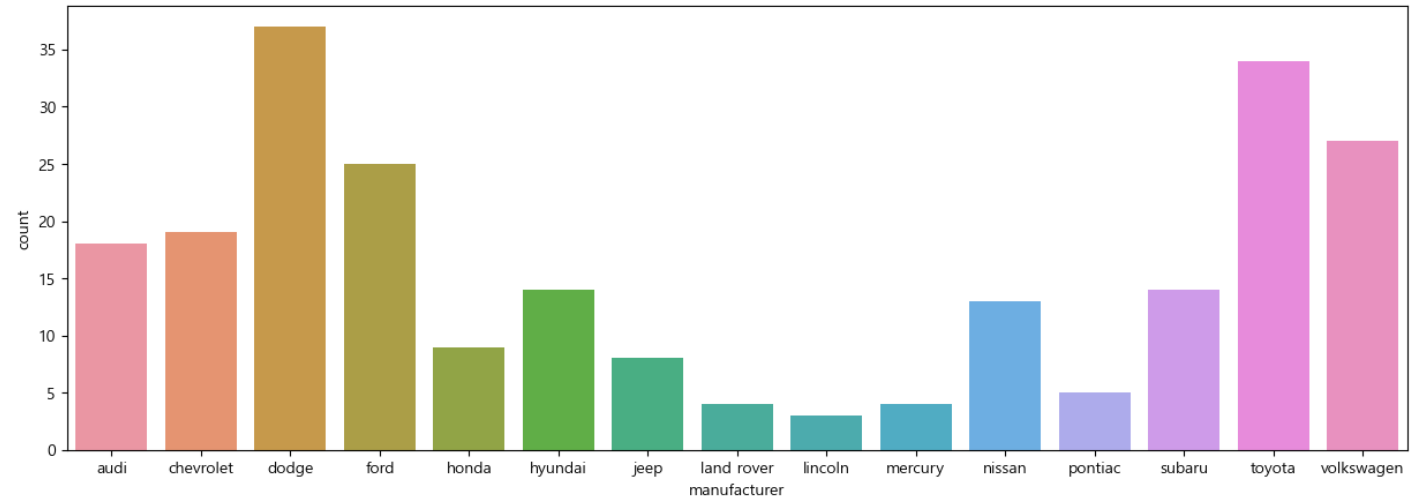
countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

제조사별 자동차 모델 수



문제2

자동차 제조사에 따른 자동차 수를 제조년도를 범주로 하여 하나의 그래프로 시각화 하시오.

catplot() 함수 사용

```
g = sb.catplot(data=df, x='manufacturer', kind='count', hue='year')

g.fig.set_figwidth(15)
g.fig.set_figheight(5)

plt.title("제조사별 자동차 모델 수")
plt.show()
plt.close()
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

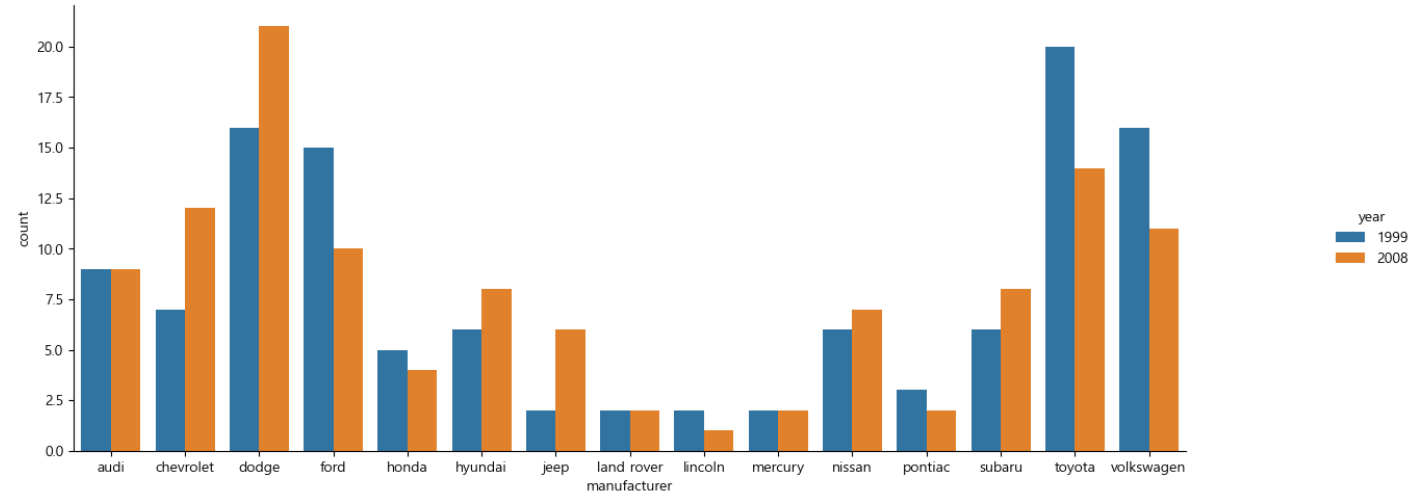
문제4

barplot() 함수 사용

문제5

연습문제(3)_풀이.ipynb

제조사별 자동차 모델 수



countplot() 함수 사용

```
sb.countplot(data=df, x="manufacturer", hue='year')
plt.title("제조사별 자동차 모델 수")
plt.show()
plt.close()
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

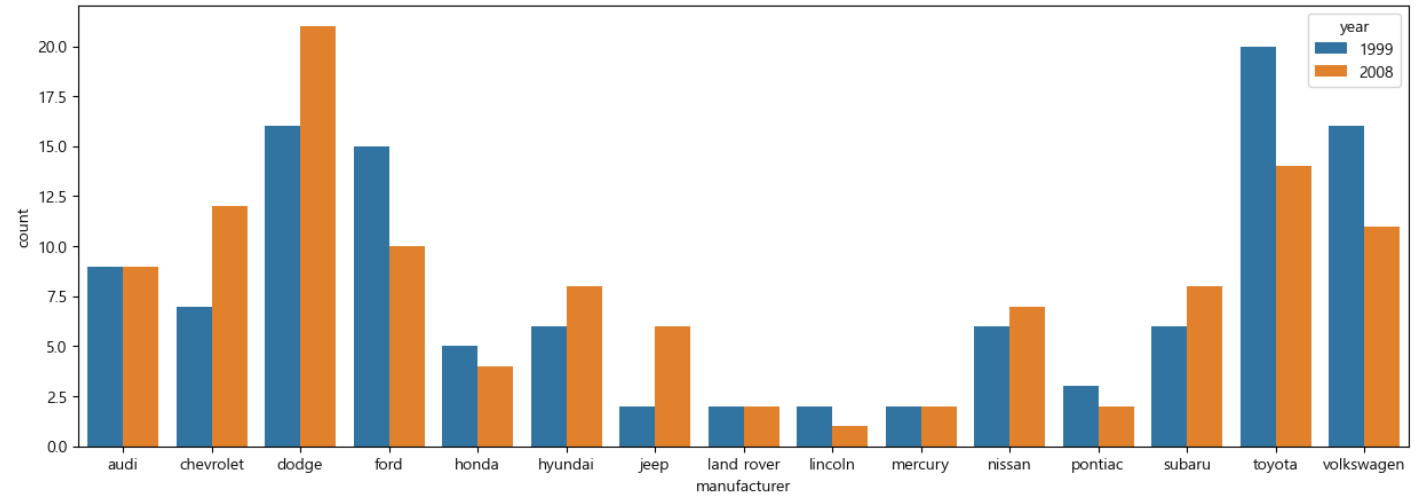
countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

제조사별 자동차 모델 수



문제3

구동방식에 따른 자동차 수를 하나의 그래프로 시각화 하시오.

catplot() 함수 사용

```
g = sb.catplot(data=df, x='drv', kind='count')

g.fig.set_figwidth(5)
g.fig.set_figheight(3)

plt.title("구동방식에 따른 자동차 수")
plt.show()
plt.close()
```


데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

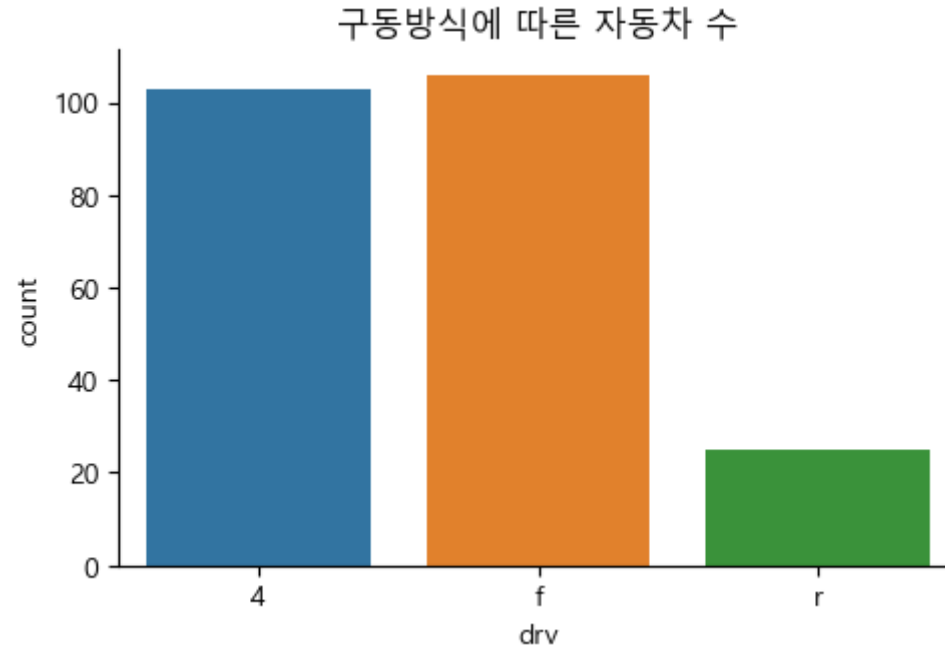
catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5



countplot() 함수 사용

```
sb.countplot(data=df, x='drv')
plt.title("구동방식에 따른 자동차 수")
plt.show()
plt.close()
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조
2. 데이터 가져오기
- #03. 데이터 전처리
4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

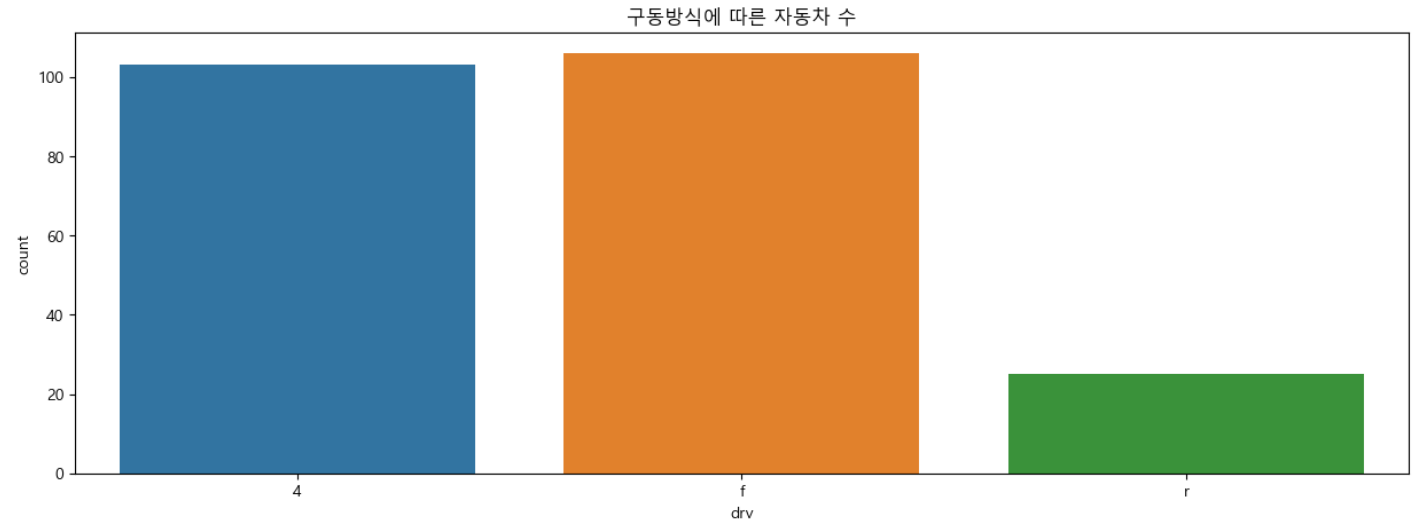
catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5



문제4

자동차 구동 방식에 따른 도심 연비 평균을 하나의 그래프로 시각화 하시오.

barplot() 함수 사용

```
plt.rcParams["figure.figsize"] = (10, 4)

sb.barplot(data=df, x='drv', y='cty', estimator='mean')
plt.title("구동방식별 평균 도심 연비")
plt.xticks([0, 1, 2], ['4률회전', '전륜구동', '후륜구동'])
plt.show()
plt.close()
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

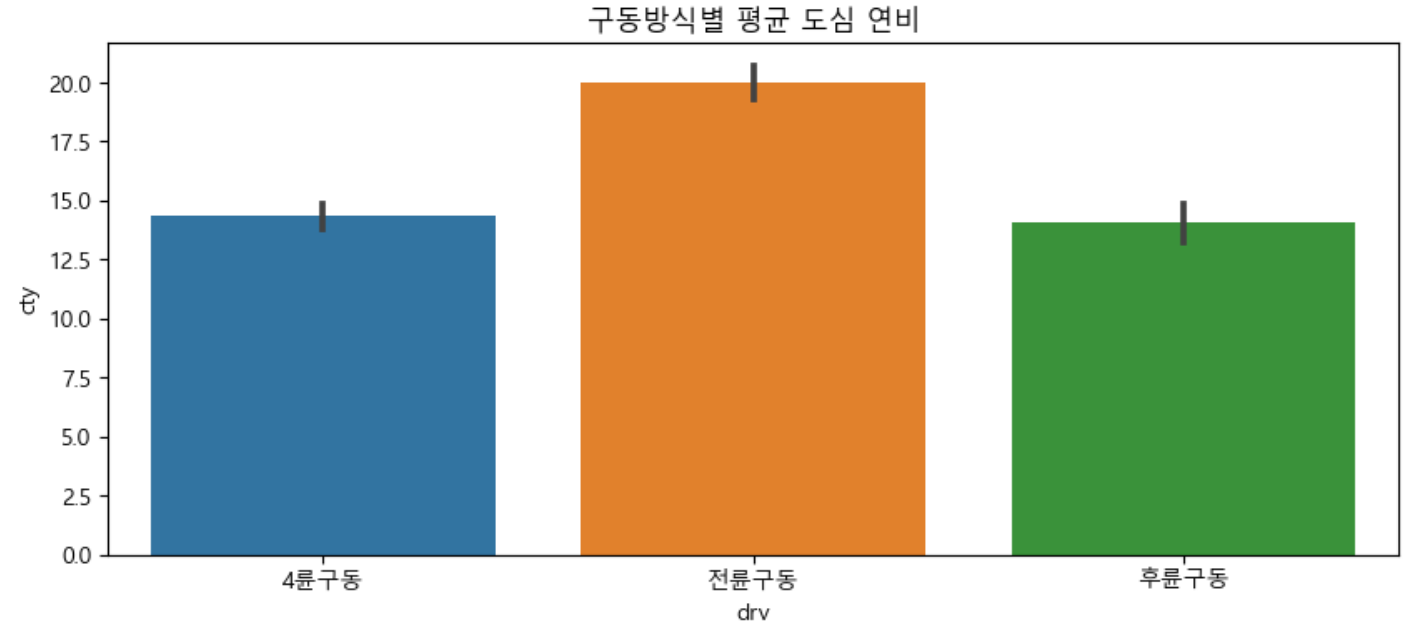
catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5



문제5

자동차 제조사에 따른 도심 연비 평균과 고속도로 연비 평균을 서브플롯 형태로 시각화 하시오.

```
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(25, 7))
```

```
fig.subplots_adjust(wspace=0.2)
```

```
fig.suptitle('제조사별 연비 비교', fontsize=30)
```

```
sb.barplot(data=df, x='manufacturer', y='cty', estimator='mean', ax=ax1)
```

```
sb.barplot(data=df, x='manufacturer', y='hwy', estimator='mean', ax=ax2)
```

```
ax1.title.set_text("평균 도심 연비")
```

```
ax2.title.set_text("평균 고속도로 연비")
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

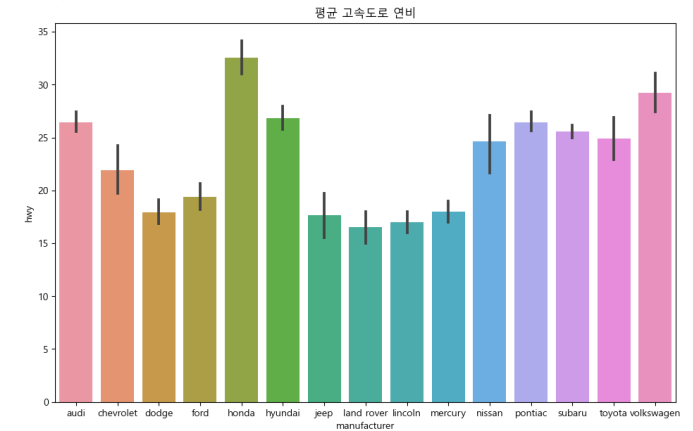
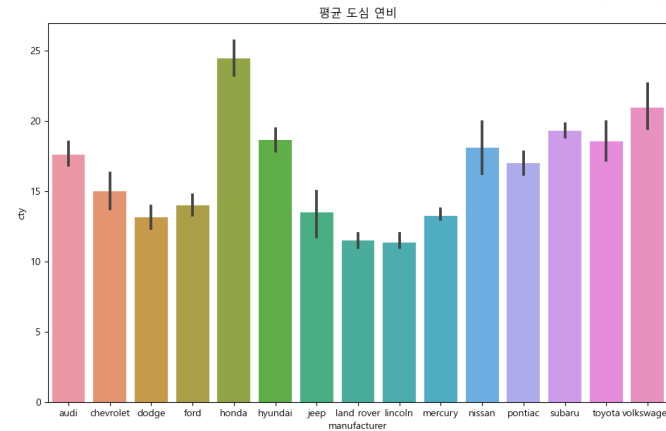
문제4

barplot() 함수 사용

문제5

```
plt.show()
plt.close()
```

제조사별 연비 비교



문제6

배기량과 도심연비, 배기량과 고속도로 연비의 상관관계를 확인하고자 한다.

서브플롯에 각 그래프를 시각화 하고 추세선을 포함시키시오.

```
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(20, 5))
```

```
fig.subplots_adjust(wspace=0.1)
```

```
fig.suptitle('배기량과 연비의 상관관계', fontsize=25)
```

```
sb.regplot(data=df, x='displ', y='cty', ax=ax1)
```

```
sb.regplot(data=df, x='displ', y='hwy', ax=ax2)
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조
2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

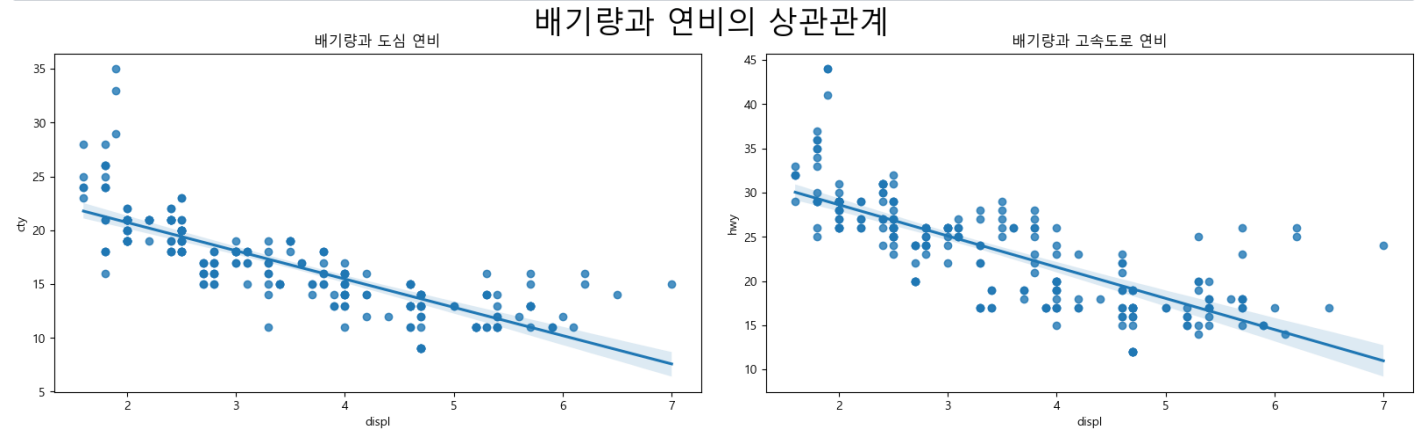
문제4

barplot() 함수 사용

문제5

```
ax1.title.set_text("배기량과 도심 연비")
ax2.title.set_text("배기량과 고속도로 연비")
```

```
plt.show()
plt.close()
```



문제 7

```
df['평균연비'] = df.loc[:, 'cty':'hwy'].mean(axis=1)
df
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	
id										
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	
...	
230	volkswagen	passat	2.0	2008	4	auto(s6)	f	19	28	
231	volkswagen	passat	2.0	2008	4	manual(m6)	f	21	29	
232	volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	
233	volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	
234	volkswagen	passat	3.6	2008	6	auto(s6)	f	17	26	

234 rows × 12 columns

```
평균값 = df['평균연비'].mean()
평균값
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

20.14957264957265

```
df['평가'] = np.where(df['평균연비'] > 평균값, 'Pass', 'Fail')
df
```

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	
id										
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	
...	
230	volkswagen	passat	2.0	2008	4	auto(s6)	f	19	28	
231	volkswagen	passat	2.0	2008	4	manual(m6)	f	21	29	
232	volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	
233	volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

연습문제(3)_풀이.ipynb

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	
id										
234	volkswagen	passat	3.6	2008	6	auto(s6)	f	17	26	

234 rows × 13 columns

```
평균합 = df['평균연비'].sum()
평균합
```

4715.0

```
df['연비효율'] = df['평균연비'] / 평균값 * 100
df
```

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	
id										
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	
id										
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	
...	
230	volkswagen	passat	2.0	2008	4	auto(s6)	f	19	28	
231	volkswagen	passat	2.0	2008	4	manual(m6)	f	21	29	
232	volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	
233	volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	
234	volkswagen	passat	3.6	2008	6	auto(s6)	f	17	26	

234 rows × 14 columns

```
조건 = [
    (df['연비효율'] ≥ 90),
    (df['연비효율'] ≥ 80),
    (df['연비효율'] ≥ 70)
]
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

```
값 = ['A', 'B', 'C']
```

```
df['등급'] = np.select(조건, 값, default='D')
df
```

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	
id										
1	audi	a4	1.8	1999	4	auto(l5)	f	18	29	
2	audi	a4	1.8	1999	4	manual(m5)	f	21	29	
3	audi	a4	2.0	2008	4	manual(m6)	f	20	31	
4	audi	a4	2.0	2008	4	auto(av)	f	21	30	
5	audi	a4	2.8	1999	6	auto(l5)	f	16	26	
...	
230	volkswagen	passat	2.0	2008	4	auto(s6)	f	19	28	
231	volkswagen	passat	2.0	2008	4	manual(m6)	f	21	29	
232	volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	
233	volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	
234	volkswagen	passat	3.6	2008	6	auto(s6)	f	17	26	

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

234 rows × 15 columns

```
fig, ((ax11, ax12), (ax21, ax22), (ax31, ax32)) = plt.subplots(3, 2, fig

# 자동차 구동방식별 평가에 대한 빈도
sb.countplot(data=df, x='drv', hue='평가', ax=ax11)

# 자동차 구동방식별 등급에 대한 빈도
sb.countplot(data=df, x='drv', hue='등급', ax=ax12)

# 자동차 연료 종류 평가에 대한 빈도
sb.countplot(data=df, x='fl', hue='평가', ax=ax21)

# 자동차 연료 종류 등급에 대한 빈도
sb.countplot(data=df, x='fl', hue='등급', ax=ax22)

# 자동차 분류별 평가에 대한 빈도
sb.countplot(data=df, x='class', hue='평가', ax=ax31)

# 자동차 분류별 등급에 대한 빈도
sb.countplot(data=df, x='class', hue='등급', ax=ax32)

plt.show()
plt.close()
```

데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조
2. 데이터 가져오기
- #03. 데이터 전처리
4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

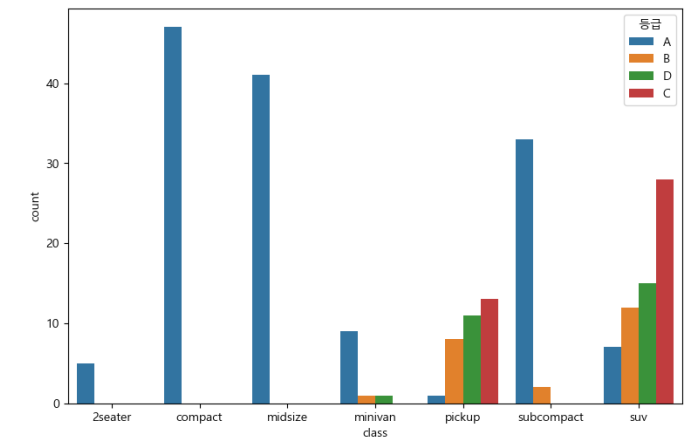
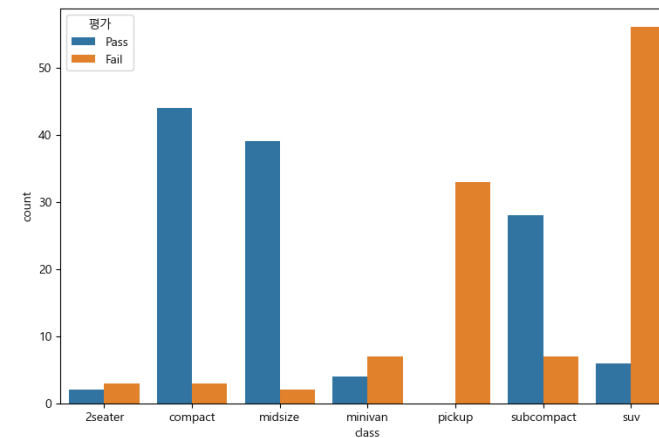
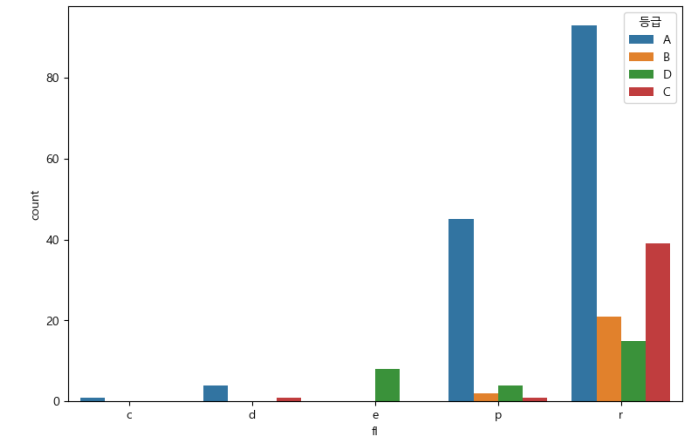
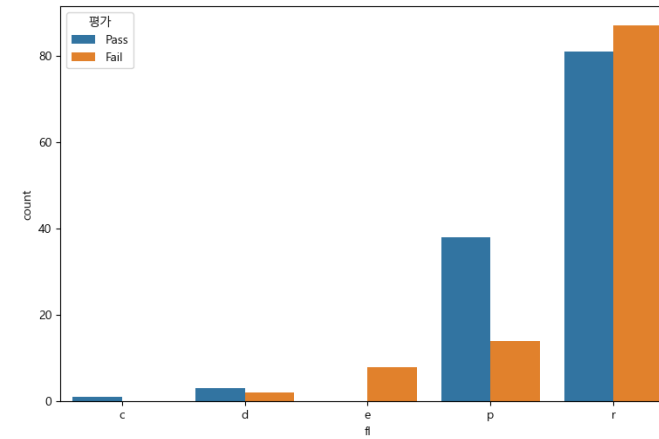
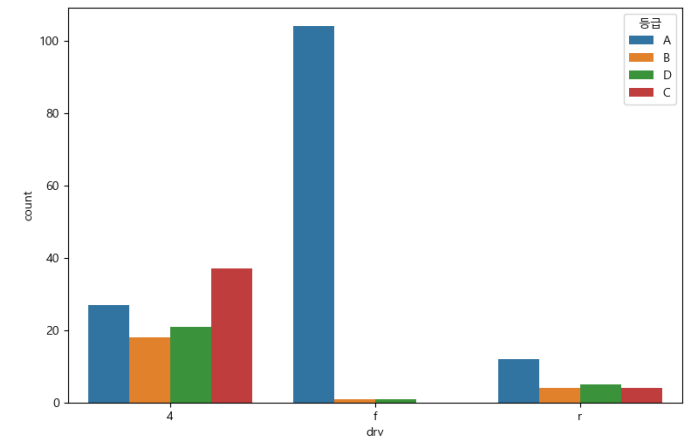
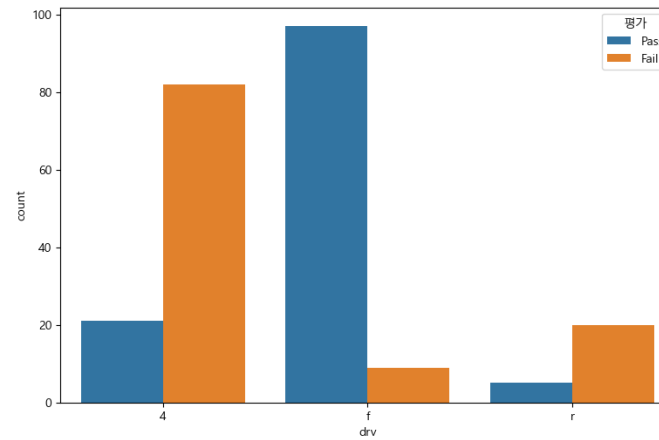
countplot() 함수 사용

문제4

barplot() 함수 사용

문제5

연습문제(3)_풀이.ipynb



데이터 시각화 연습문제 (3) - seaborn 막대, 빈도 그래프 풀이

01. 작업준비

1. 패키지 참조

2. 데이터 가져오기

#03. 데이터 전처리

4. 그래프 전역 설정

#02. 문제 풀이

문제1

catplot() 함수 사용

countplot() 함수 사용

문제2

catplot() 함수 사용

countplot() 함수 사용

문제3

catplot() 함수 사용

countplot() 함수 사용

문제4

barplot() 함수 사용

문제5