第二次作业

Test1

%matplotlib inline

import pandas as pd

import numpy as np

import matplotlib as mpl

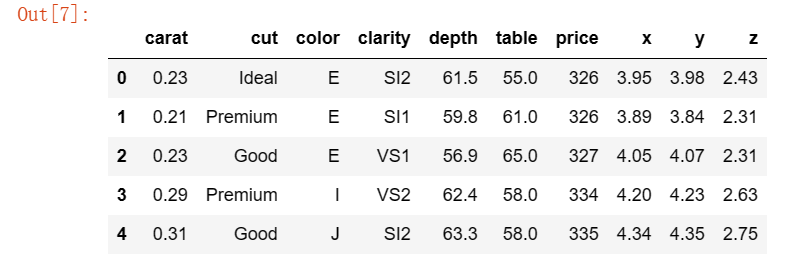
import matplotlib.pyplot as plt

import seaborn as sns

#读取数据

df=sns.load\_dataset("diamonds")

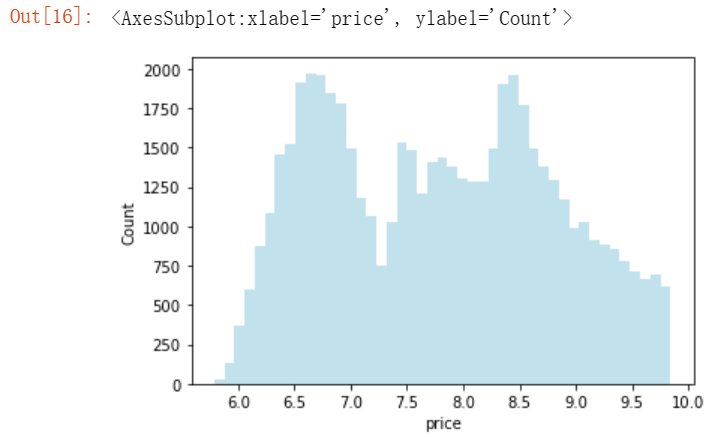
df.head()



#drawing img1

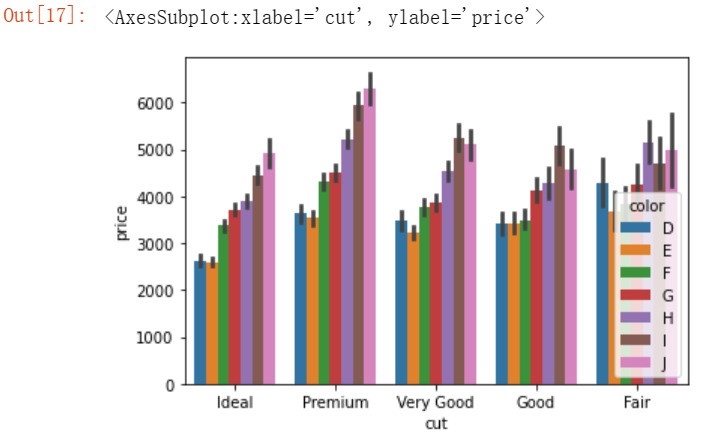
img1=np.log(df["price"])

sns.histplot(img1,edgecolor=None,color='lightblue')



#drawing img2

sns.barplot(x='cut',y='price',data=df,hue='color')



Test2

#读取数据

%matplotlib inline

import numpy as np

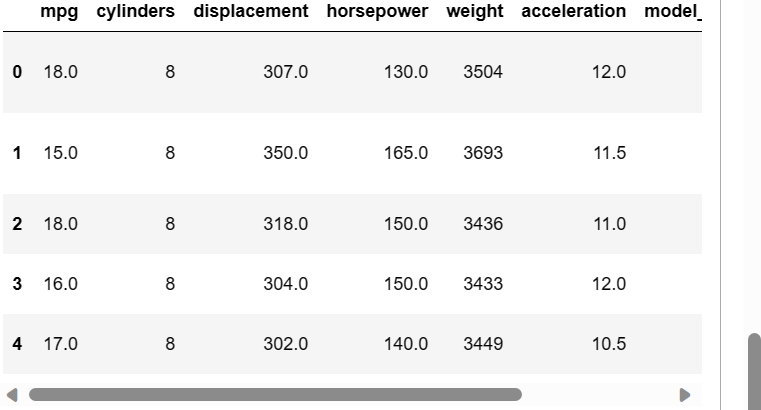
import matplotlib.pyplot as plt

import seaborn as sns

import pandas as pd

df=sns.load\_dataset("mpg")

df.head()



#绘制折线图

fig,ax=plt.subplots()

df.groupby('model\_year')['mpg'].mean()

x\_positions=range(len(mpg\_mean))

y\_values=mpg\_mean.values

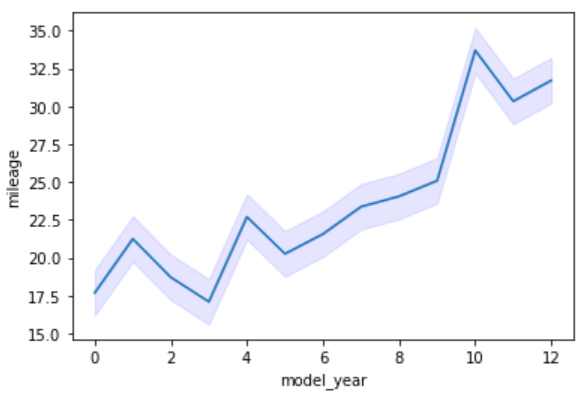
ax.plot(x\_positions,y\_values,markersize=6)

ax.fill\_between(x\_positions, y\_values, y\_values+1.5, color='blue', alpha=0.1)

ax.fill\_between(x\_positions, y\_values, y\_values-1.5, color='blue', alpha=0.1)

ax.set(xlabel='model\_year',ylabel='mileage')

plt.show()

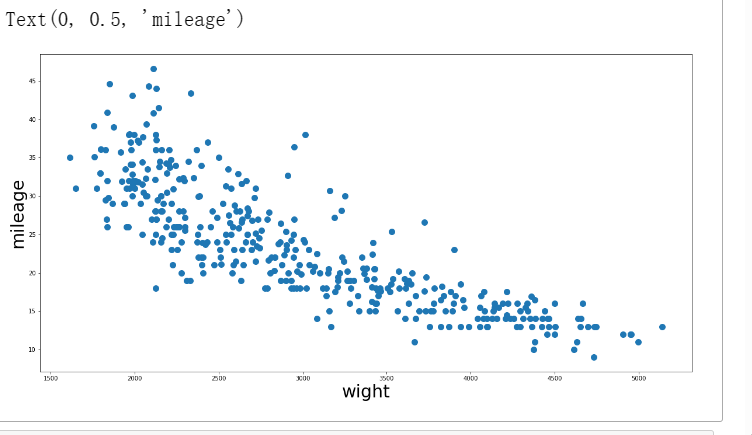


#绘制散点图

df.plot(kind='scatter',x='weight',y='mpg',figsize=(20,10),s=100)

plt.xlabel('wight',fontsize=30)

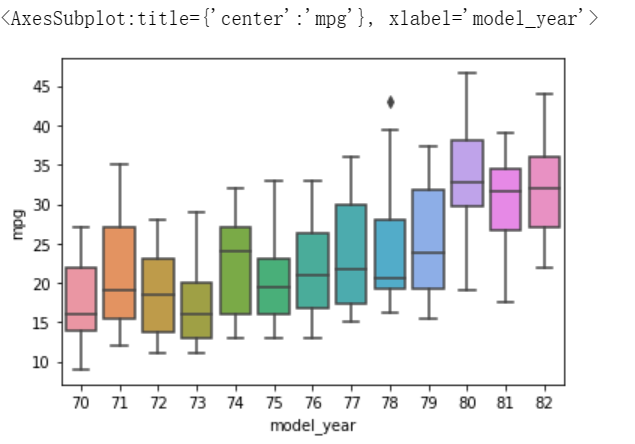
plt.ylabel('mileage',fontsize=30)



#绘制箱线图

sns.boxplot(data=df,x='model\_year',y='mileage')

df.boxplot(column='mileage',by='model\_year')



Test3

#读取microsoft\_stock文件

由于从Jupyter notebook读取总是失效，因此直接从绝对路径读取

%matplotlib inline

import numpy as np

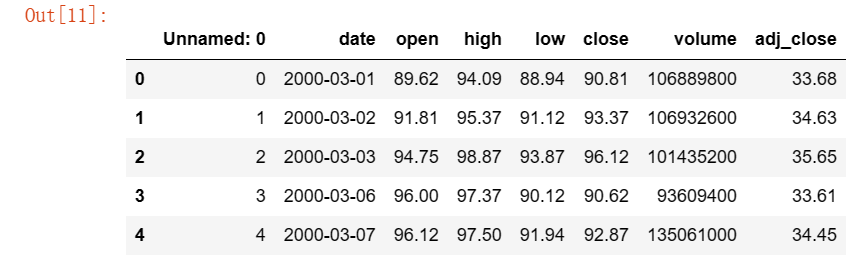
import matplotlib.pyplot as plt

import seaborn as sns

import pandas as pd

df=pd.read\_csv("D:\\学\\大数据开发课程设计\\数据开发-张志旺\\数据开发-张志旺\\microsoft\_stock.csv")

df.head()



相同方法读取google\_stock文件



#绘图

df['date']=pd.to\_datetime(df['date'])

df1['date']=pd.to\_datetime(df1['date'])

plt.figure(figsize=(12, 6))

plt.plot(df['date'],df['close'],label="microsoft\_stock")

plt.plot(df1['date'],df1['close'],label="google\_stock",color="red")

plt.xlabel("Year")

plt.ylabel("Stock Price")

plt.legend(loc=1)

plt.grid(True)

plt.show()

