... 

# 计算所有变量的相关矩阵

print(wine.corr())

# 从红葡萄酒和白葡萄酒的数据中取出一个“小”样本来进行绘图

def take\_sample(data\_frame, replace=False, n=200):

return data\_frame.loc[np.random.choice(data\_frame.index, replace=replace, size=n)]

reds\_sample = take\_sample(wine.loc[wine['type'] == 'red', :])

whites\_sample = take\_sample(wine.loc[wine['type'] == 'white', :])

wine\_sample = pd.concat([reds\_sample, whites\_sample])

wine['in\_sample'] = np.where(wine.index.isin(wine\_sample.index), 1., 0.)

print(pd.crosstab(wine.in\_sample, wine.type, margins=True))

# 查看成对变量之间的关系

sns.set\_style("dark")

g = sns.pairplot(wine\_sample, kind='reg', plot\_kws={"ci": False, "x\_jitter": 0.25, "y\_jitter": 0.25}, hue='type',

diag\_kind='hist', diag\_kws={"bins": 10, "alpha": 1.0}, palette=dict(red="red", white="white"),

markers=["o", "s"], vars=['quality', 'alcohol', 'residual\_sugar'])

print(g)

plt.suptitle('Histograms and Scatter Plots of Quality, Alcohol, and Residual Sugar', fontsize=14,

horizontalalignment='center', verticalalignment='top', x=0.5, y=0.999)

plt.show()

代码运行的结果如下：



