

场景一

对应知识点：第三章-各类图形

在实际的数据分析过程中，我们往往需要对一个表格中的多个项目进行比较分析，这个时候，组合柱状图就派上了用场。通过第三章所学的知识，你能否将下表数据用组合柱状图的形式呈现？其中，横坐标为水果种类，纵坐标为价格与数量。提示：进行两次画图操作，注意计算好柱形图的x轴位置，保持图表的美观噢~

水果	价格	数量
香蕉	2	5
苹果	4	3
草莓	10	6

In [2]:

#参考答案:

```
import pandas as pd
```

```
import numpy as np
```

```
import matplotlib.pyplot as plt
```

```
raw_data = {'水果': ['香蕉', '苹果', '草莓'],  
            '价格': [2, 4, 10],  
            '数量': [5, 3, 6],}
```

```
df = pd.DataFrame(raw_data)
```

```
pos = list(range(len(df['水果'])))
```

```
width = 0.25
```

```
plt.bar([p - width/2 for p in pos], df['价格'], width, color='#FFCCCC', label=df['水果'][0])
```

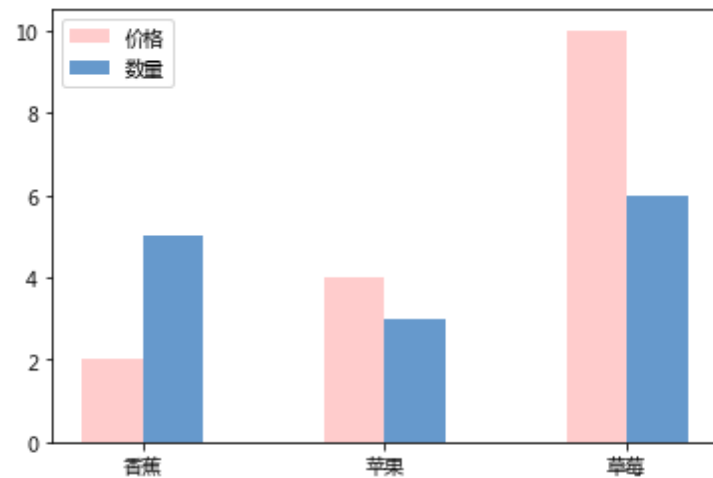
```
plt.bar([p + width/2 for p in pos], df['数量'], width, color='#6699CC', label=df['水果'][1])
```

```
plt.xticks(np.linspace(0, 2, 3), df['水果'])
```

```
plt.legend(['价格', '数量'], loc='upper left')
```

Out[2]:

<matplotlib.legend.Legend at 0x7fb46bafc940>



场景二

对应知识点：第三章-各类图形

针对KLab提供的“人力资源分析数据集” (/home/kesci/input/human_one/HR_comma_sep.csv) ， 数据集中的字段包括：

对环境的满意度(satisfaction_level)

每月加班平均时间(average_montly_hours) 对现有的项目的满意度(last_evaluation)

请你利用散点图挖掘出员工对环境满意度和其每月加班平均时间的关系

In [3]:

```
#参考答案:
HR_data = pd.read_csv('/home/kesci/input/human_one/HR_comma_sep.csv')
HR_data.shape
```

Out[3]:

(14999, 10)

In [4]:

```
HR_data.head()
```

Out[4]:

	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	left	promotion_la
0	0.38	0.53	2	157	3	0	1	0
1	0.80	0.86	5	262	6	0	1	0
2	0.11	0.88	7	272	4	0	1	0
3	0.72	0.87	5	223	5	0	1	0
4	0.37	0.52	2	159	3	0	1	0

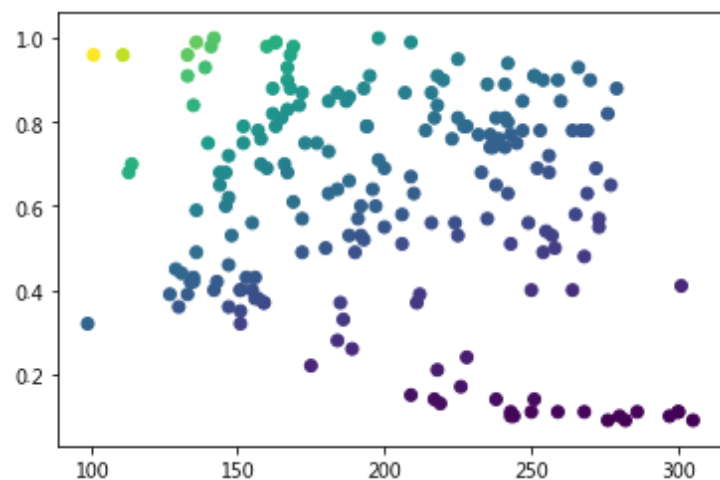


In [5]:

```
from sklearn.utils import shuffle
data = shuffle(HR_data).iloc[:200,:]
X = data['average_monthly_hours'].tolist()
Y = data['satisfaction_level'].tolist()
T = np.arctan2(Y,X)
plt.scatter(X,Y,c=T)
```

Out[5]:

<matplotlib.collections.PathCollection at 0x7fb46636f9e8>



场景三

对应知识点：第四章-多图合并显示，第三章-各类图形

针对“人力资源分析数据集”，挖掘对现有的项目的满意度和月平均加班时间之间的关系；并将此图和上图合并显示

In [6]:

#参考答案:

```
data = shuffle(HR_data).iloc[:200,:]  
X = data['average_monthly_hours'].tolist()  
Y1 = data['satisfaction_level'].tolist()  
Y2 = data['last_evaluation'].tolist()  
T1 = np.arctan2(Y1, X)  
T2 = np.arctan2(Y2, X)
```

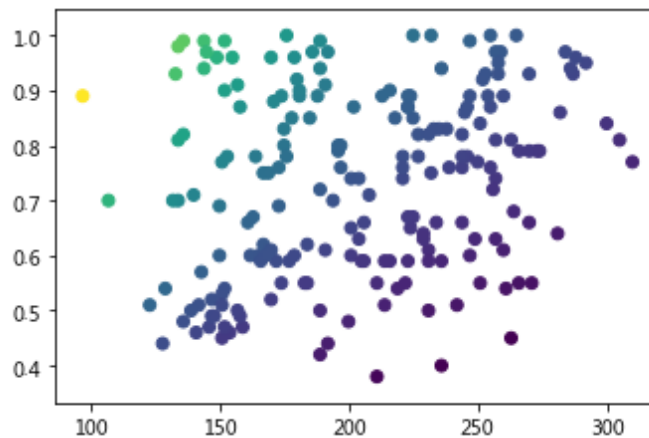
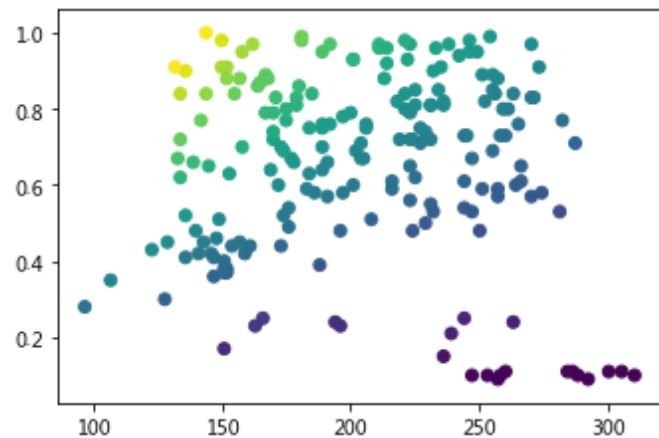
```
plt.figure(figsize=(12, 8))
```

```
plt.subplot(2, 2, 1)  
plt.scatter(X, Y1, c=T1)
```

```
plt.subplot(2, 2, 2)  
plt.scatter(X, Y2, c=T2)
```

Out[6]:

<matplotlib.collections.PathCollection at 0x7fb4663581d0>



场景四

对应知识点：第五章-3D作图与动画，第三章-各类图形

请在3D坐标系下画出函数 $Z=3(1-X)^2 \exp(-(X^2) - (Y+1)^2) - 10(X/5 - X^3 - Y^5) \exp(-X^2-Y^2) - 1/3 \exp(-(X+1)^2 - Y^2)$;并将其投影至 $x=-4$ 和 $z=-7$ 平面上;并添加 colorbar

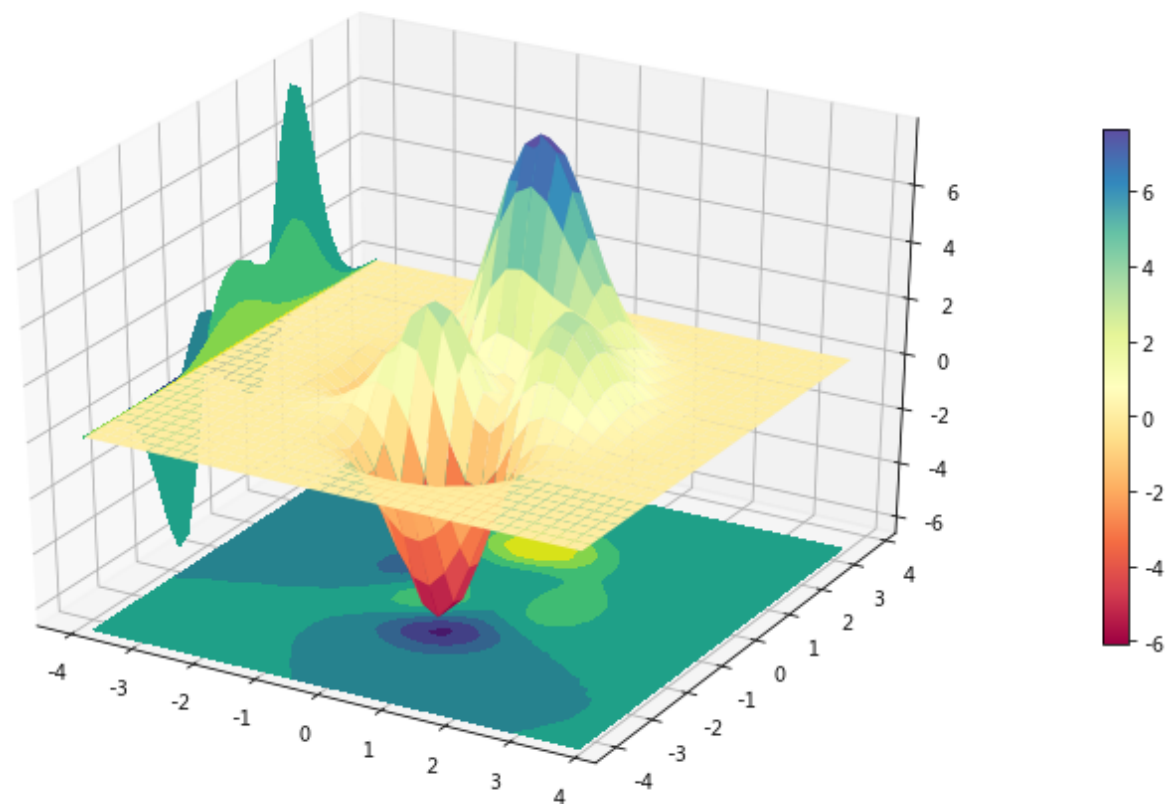
In [7]:

#参考答案:

```
from mpl_toolkits.mplot3d import Axes3D
X = np.arange(-4, 4, 0.25)
Y = np.arange(-4, 4, 0.25)
X, Y = np.meshgrid(X, Y)
Z = 3 * (1-X)**2 * np.e**(-(X**2) - (Y+1)**2) - 10*(X/5 - X**3 - Y**5)*np.e**(-X**2-Y**2) - 1/3*np.e**(-(X+1)**2 - Y**2)
fig = plt.figure(figsize=(10,6))
ax = Axes3D(fig)
cmap = 'Spectral'
m = ax.plot_surface(X, Y, Z, rstride=1, cstride=1, cmap=plt.get_cmap(cmap))
ax.contourf(X, Y, Z, zdir='x', offset=-4)
ax.contourf(X, Y, Z, zdir='z', offset=-7)
plt.colorbar(m, shrink=0.6)
```

Out[7]:

<matplotlib.colorbar.Colorbar at 0x7fb466235978>



场景五（拓展）

实际上，画图的工具有很多，Matplotlib 只是其中一种，至此小伙伴们学习到的知识已经可以完成实际工程应用中的大部分画图任务。但如果想要追求更加完美的画图效果，大家可以参考一个工具：Seaborn，它在 Matplotlib 基础上进行了更高级的封装，利用它我们可以更加方便的作出更有吸引力的图。请点击[本链接](#)查看 Seaborn 中 distplot 函数的使用方法，画出“人力资源分析数据集”中“last_evaluation”字段的分布图。

In [8]:

#参考答案:

```
import seaborn as sns
sns.set_style('darkgrid')
sns.distplot(HR_data['last_evaluation'])
```

/opt/conda/lib/python3.5/site-packages/matplotlib/axes/_axes.py:6462: UserWarning: The 'normed' kwarg is deprecated, and has been replaced by the 'd' warnings.warn("The 'normed' kwarg is deprecated, and has been ")

Out[8]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fb45dalc6a0>

