

3. 数据可视化

3.1 matplotlib简介与安装

matplotlib是一个用于Python的2D绘图库，它可以生成多种格式的高质量图像。matplotlib广泛应用于数据可视化领域，具有以下优点：

- 绘制多种图形：支持折线图、柱状图、散点图、饼图等多种图形类型
- 高度可定制：可以调整图形的各种属性，如颜色、线型、标签、图例等
- 易于集成：与其他数据分析和可视化库（如Pandas、seaborn、scikit-learn等）具有良好的兼容性，可以无缝集成
- 生成多种输出格式：支持生成PNG、PDF、SVG、EPS等多种图像格式，以满足不同场景的需求

3.2 使用matplotlib绘图

折线图绘制

- 使用plot()方法绘制折线图。例如，绘制一条代表学生成绩的折线图：

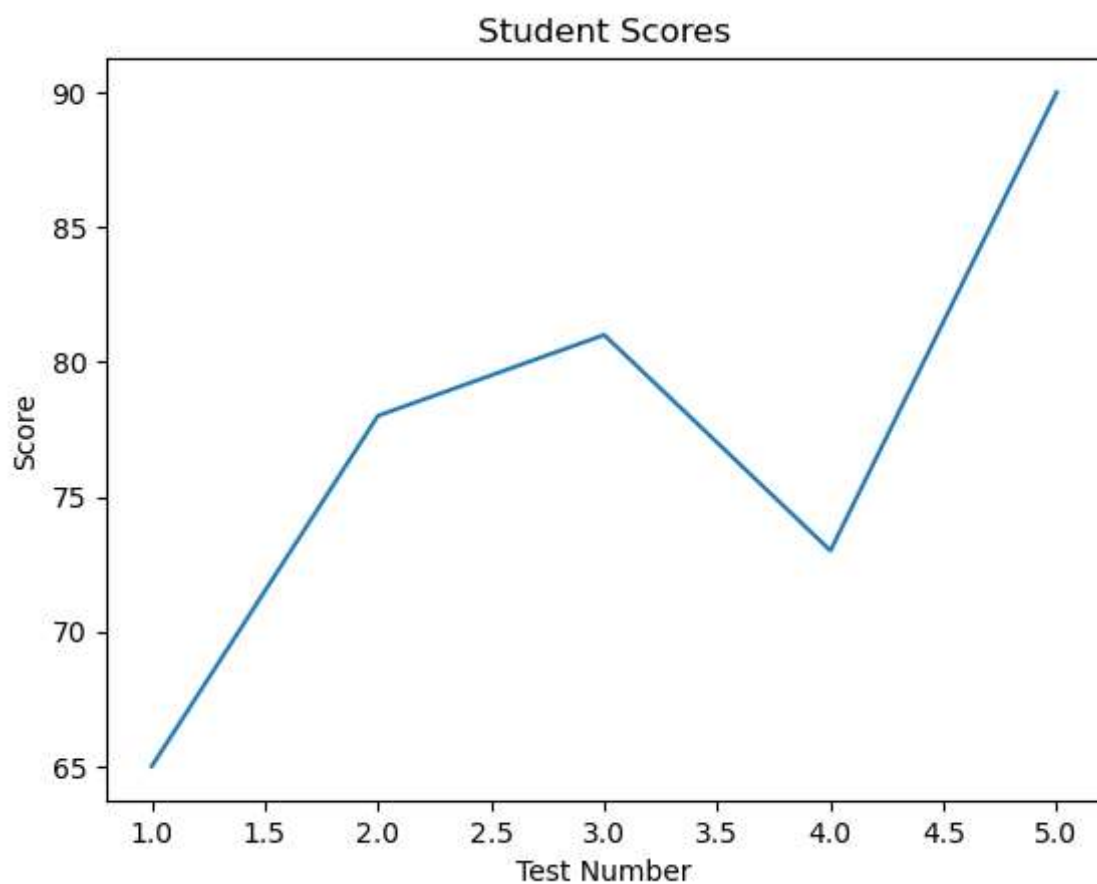
In [3]:

```
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y = [65, 78, 81, 73, 90]

plt.plot(x, y) #绘制x-y的折线图
plt.xlabel('Test Number') #x轴标签
plt.ylabel('Score') #y轴标签
plt.title('Student Scores') #图的标题

plt.show() #展示图片
```



柱状图绘制

- 使用bar()方法绘制柱状图。例如，绘制一张展示不同科目平均成绩的柱状图：

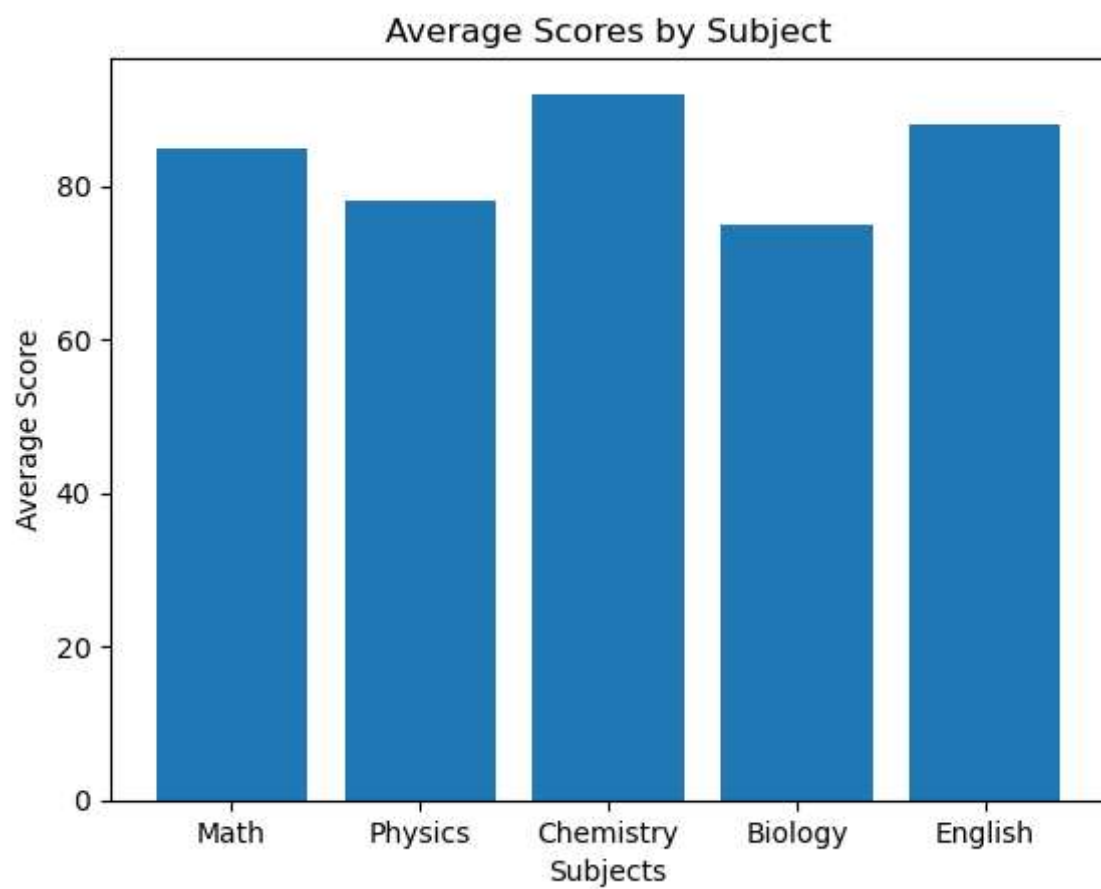
In [4]:

```
import matplotlib.pyplot as plt

# 示例数据
subjects = ['Math', 'Physics', 'Chemistry', 'Biology', 'English']
avg_scores = [85, 78, 92, 75, 88]

plt.bar(subjects, avg_scores)
plt.xlabel('Subjects')
plt.ylabel('Average Score')
plt.title('Average Scores by Subject')

plt.show()
```



散点图绘制

- 使用scatter()方法绘制散点图。例如，绘制一张展示学生年龄与成绩之间关系的散点图：

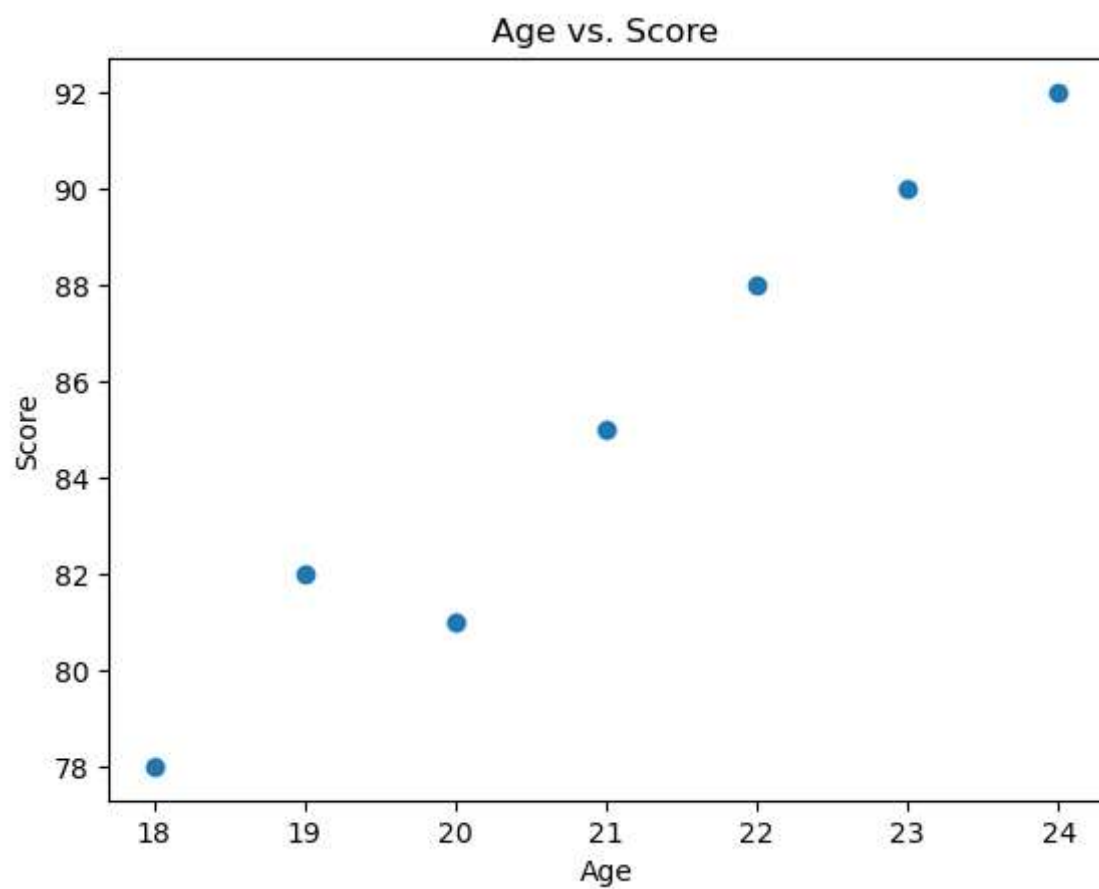
In [5]:

```
import matplotlib.pyplot as plt

# 示例数据
ages = [18, 19, 20, 21, 22, 23, 24]
scores = [78, 82, 81, 85, 88, 90, 92]

plt.scatter(ages, scores)
plt.xlabel('Age')
plt.ylabel('Score')
plt.title('Age vs. Score')

plt.show()
```



绘制饼图

- 使用pie()方法绘制饼图。例如，绘制一张展示不同科目成绩占比的饼图：

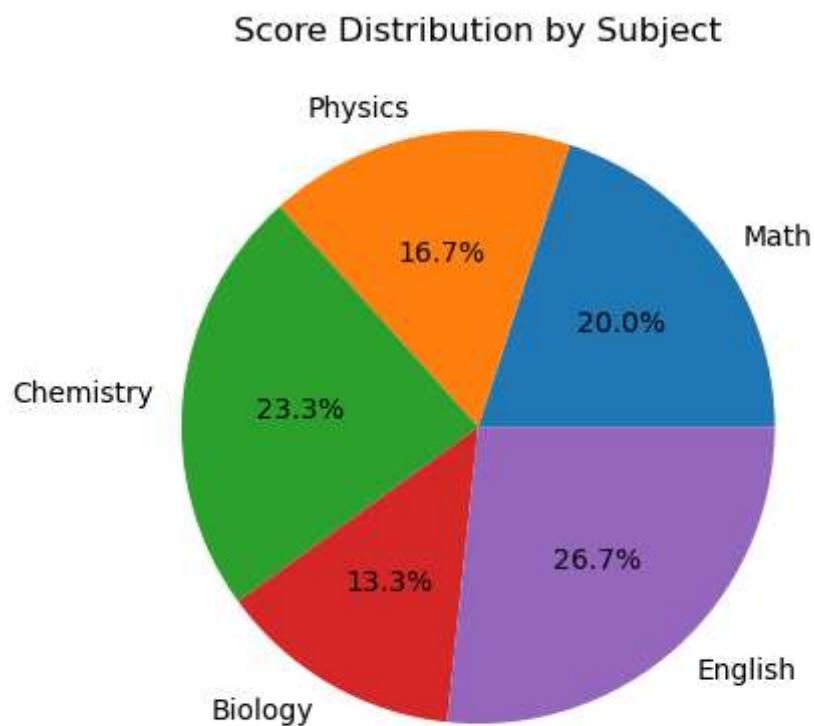
In [6]:

```
import matplotlib.pyplot as plt

# 示例数据
subjects = ['Math', 'Physics', 'Chemistry', 'Biology', 'English']
scores = [300, 250, 350, 200, 400]

plt.pie(scores, labels=subjects, autopct='%.1f%%')
plt.title('Score Distribution by Subject')

plt.show()
```



如何提升图片观感?

——使用参数, 定制化图片

In [38]:

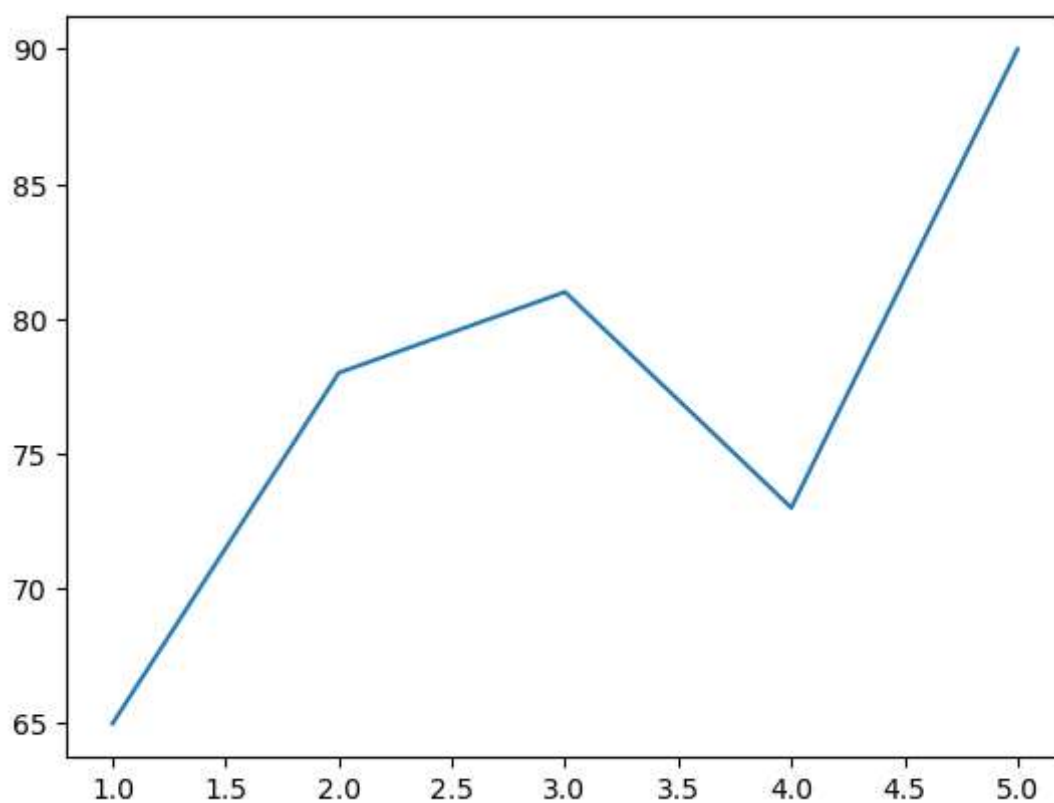
```
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]

#折线图绘制
plt.plot(x, y)
```

Out[38]:

[<matplotlib.lines.Line2D at 0x18bfe2d5e20>]



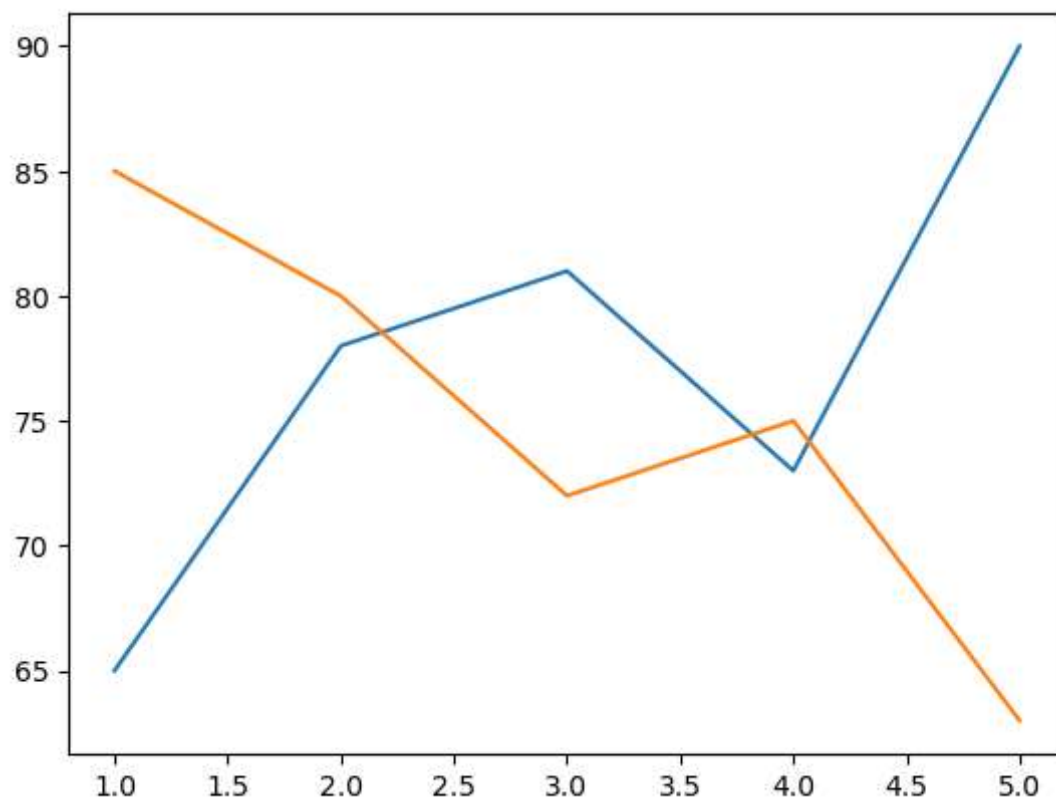
同一幅图中绘制多组数据

In [39]:

```
# 添加另一组数据
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]
y2=[85, 80, 72, 75, 63]
plt.plot(x, y1)
plt.plot(x, y2)

plt.show()
```



添加图例

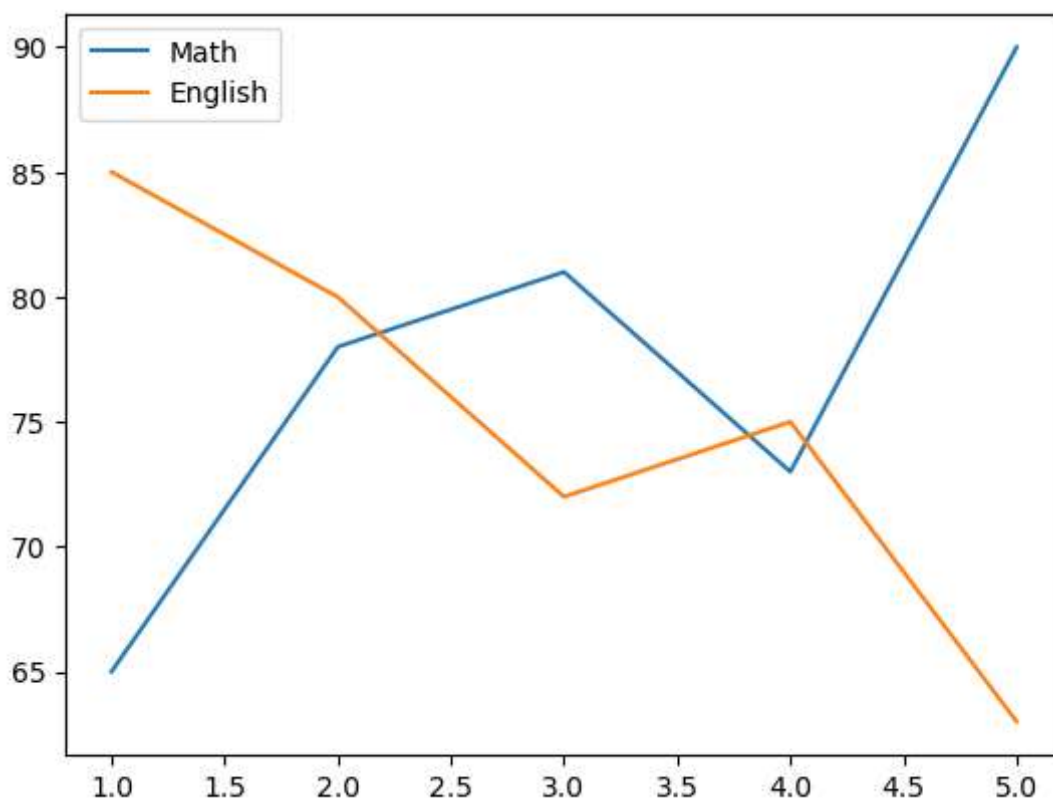
In [40]:

```
# 添加另一组数据
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]
y2=[85, 80, 72, 75, 63]
plt.plot(x, y1, label="Math")
plt.plot(x, y2, label="English")

#添加图例
plt.legend()

plt.show()
```



添加网格线

In [41]:

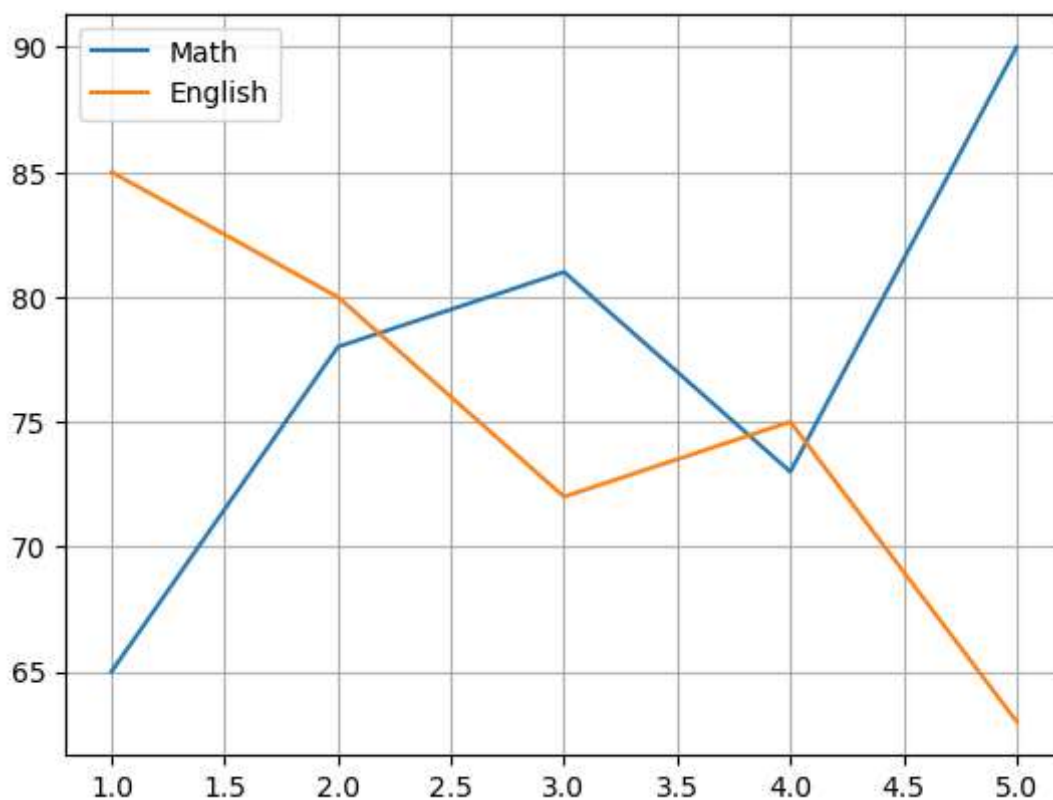
```
# 添加另一组数据
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]
y2=[85,80,72,75,63]
plt.plot(x,y1,label="Math") #中文会乱码，之后会讲到如何解决
plt.plot(x,y2,label="English")

#添加图例
plt.legend()

#添加网格线
plt.grid()

plt.show()
```



添加x轴和y轴标签

In [42]:

```
# 添加另一组数据
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]
y2=[85,80,72,75,63]
plt.plot(x,y1,label="Math") #中文会乱码，之后会讲到如何解决
plt.plot(x,y2,label="English")

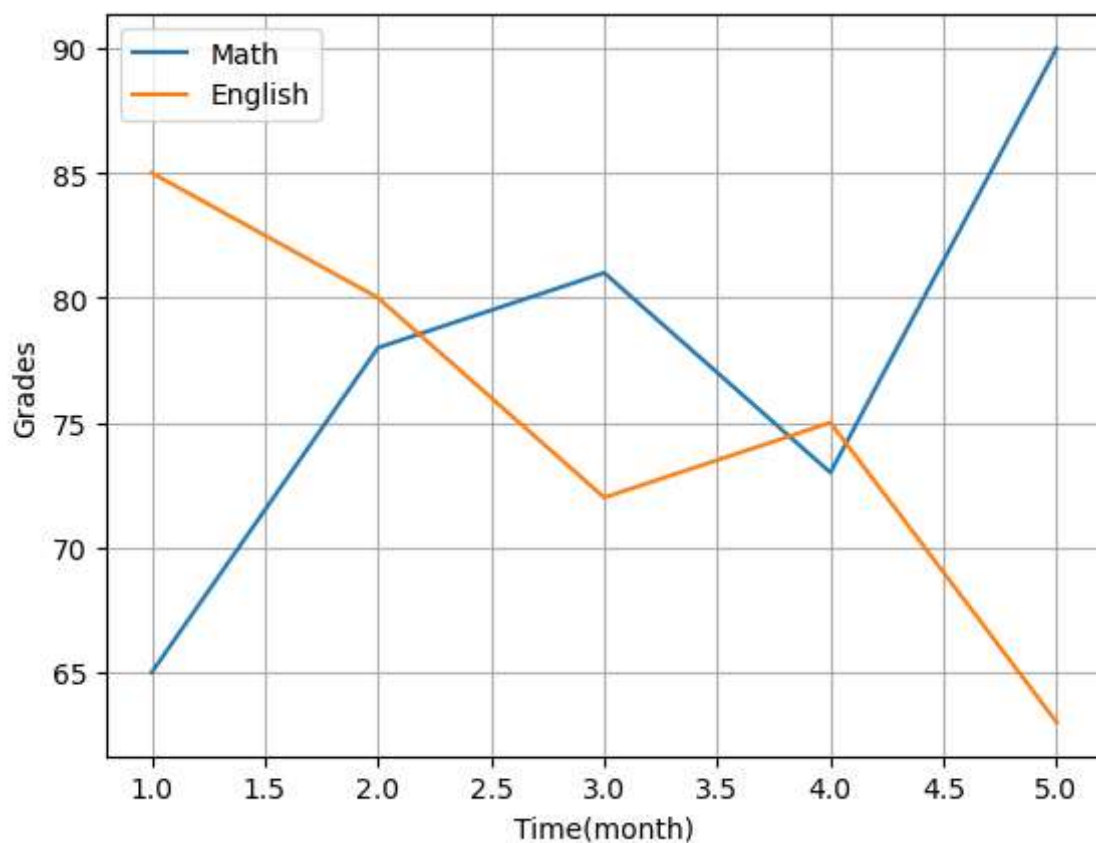
#添加图例
plt.legend()

#添加网格线
plt.grid()

#添加x轴标签
plt.xlabel("Time(month)")

#添加y轴标签
plt.ylabel("Grades")

plt.show()
```



标记数据点(Marker)

In [43]:

```
# 添加另一组数据
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]
y2=[85, 80, 72, 75, 63]
plt.plot(x, y1, label="Math", marker="o") #中文会乱码，之后会讲到如何解决
plt.plot(x, y2, label="English", marker="*")

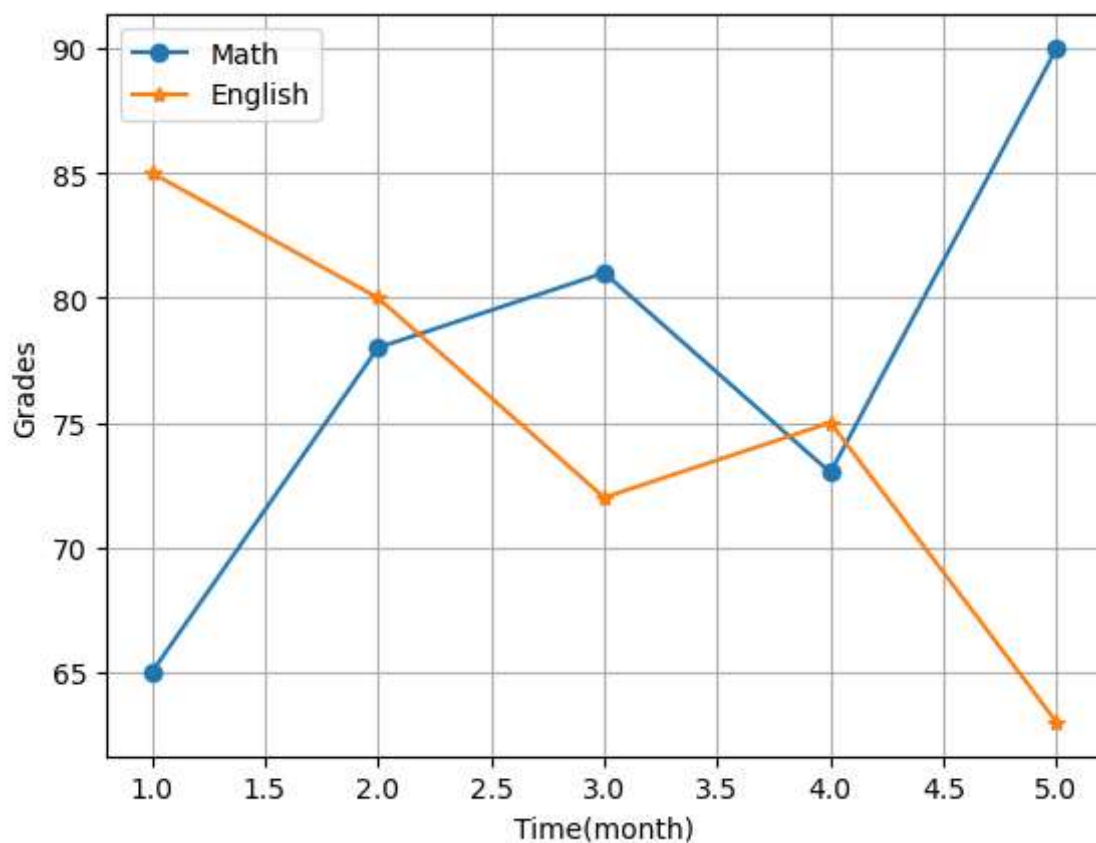
#添加图例
plt.legend()

#添加网格线
plt.grid()

#添加x轴标签
plt.xlabel("Time(month)")

#添加y轴标签
plt.ylabel("Grades")

plt.show()
```



改变线条类型

In [44]:

```
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]
y2=[85, 80, 72, 75, 63]
plt.plot(x,y1,label="Math",marker="o",linestyle="--") #中文会乱码，之后会讲到如何解决
plt.plot(x,y2,label="English",marker="*",linestyle="--")

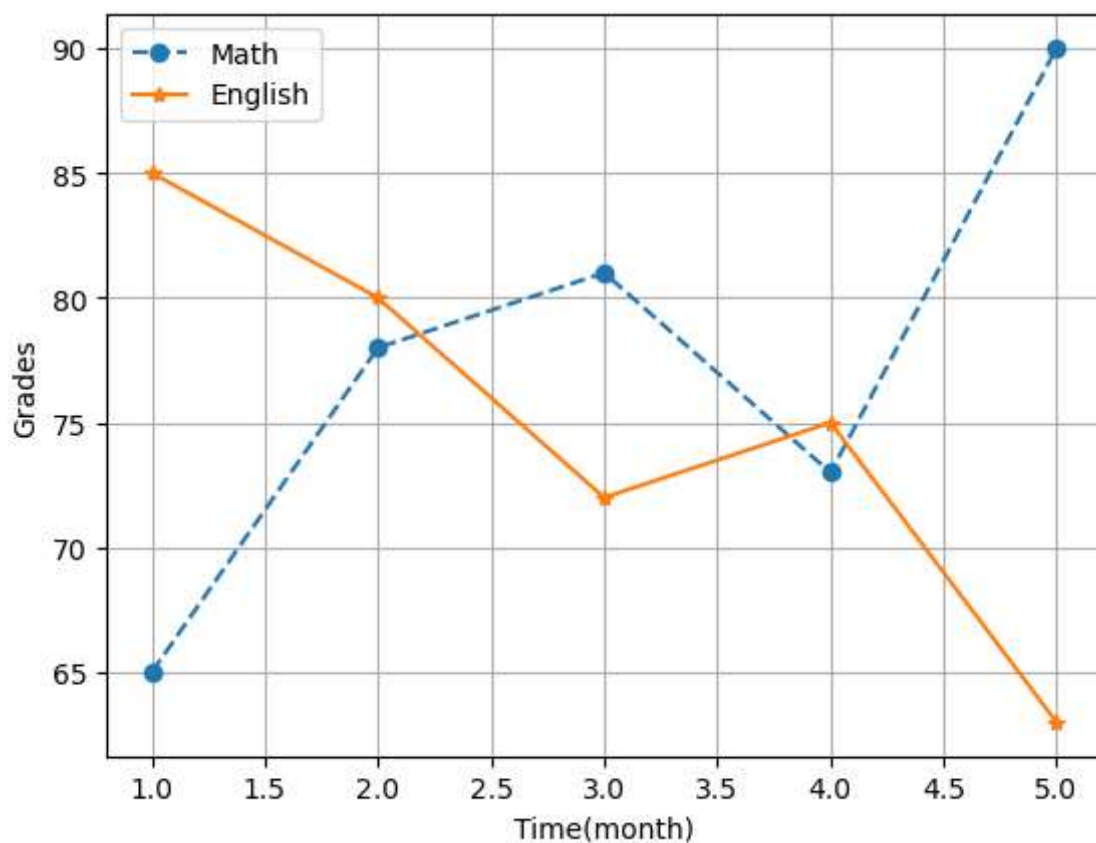
#添加图例
plt.legend()

#添加网格线
plt.grid()

#添加x轴标签
plt.xlabel("Time(month)")

#添加y轴标签
plt.ylabel("Grades")

plt.show()
```



更改线条颜色

In [52]:

```
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]
y2=[85, 80, 72, 75, 63]
plt.plot(x,y1,label="Math",marker="o",linestyle="--",color="#fdb66d") #中文会乱码，之后会讲到如
plt.plot(x,y2,label="English",marker="*",color="#0984e3")

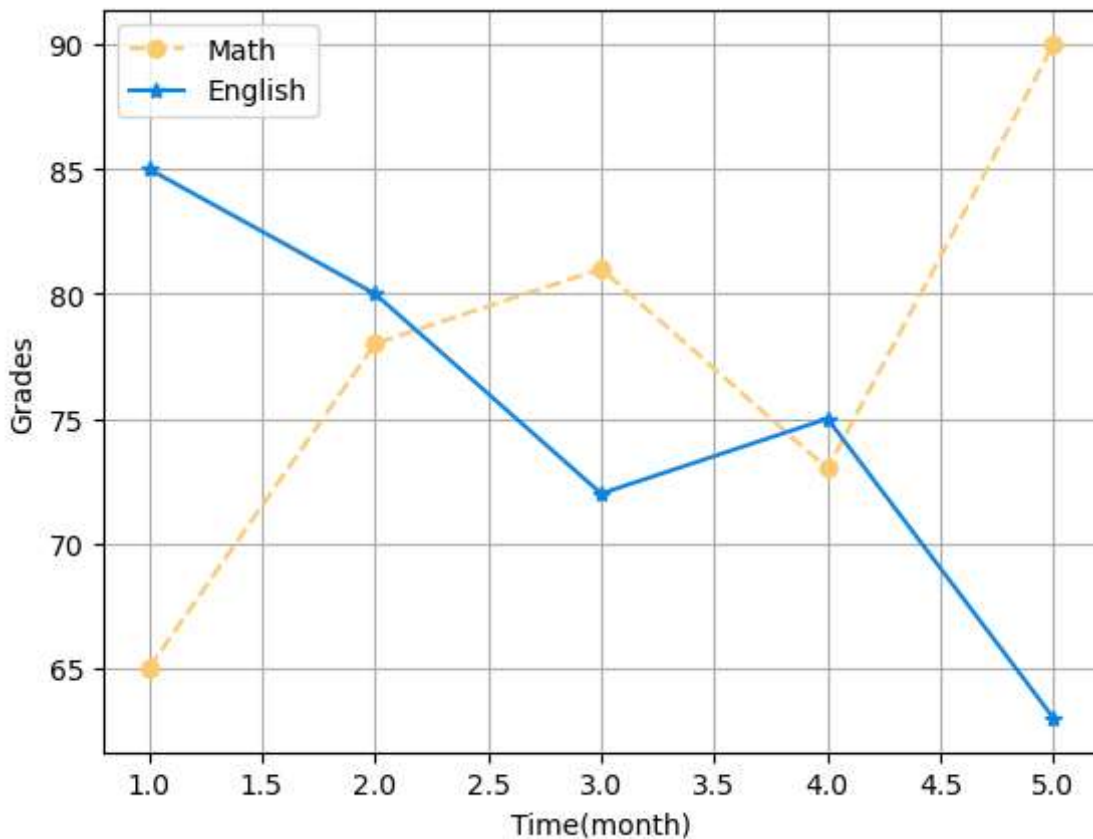
#添加图例
plt.legend()

#添加网格线
plt.grid()

#添加x轴标签
plt.xlabel("Time(month)")

#添加y轴标签
plt.ylabel("Grades")

plt.show()
```



更改字体

In [57]:

```
import matplotlib.pyplot as plt #导入matplotlib 模块中的pyplot库

# 示例数据
x = [1, 2, 3, 4, 5]
y1 = [65, 78, 81, 73, 90]
y2=[85, 80, 72, 75, 63]
plt.plot(x,y1,label="Math",marker="o",linestyle="--",color="#fdb66d") #中文会乱码，之后会讲到如
plt.plot(x,y2,label="English",marker="*",color="#0072bc")

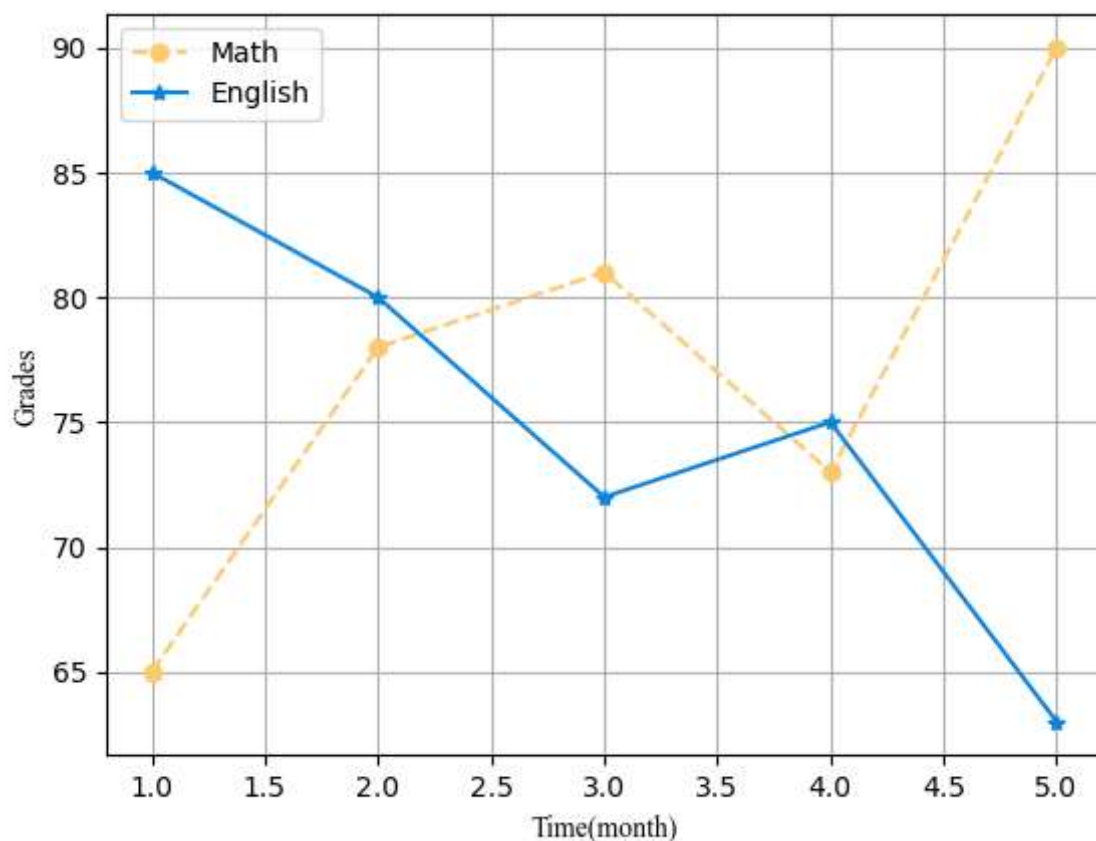
#添加图例
plt.legend()

#添加网格线
plt.grid()

#添加x轴标签
plt.xlabel("Time(month)",font="Times New Roman")

#添加y轴标签
plt.ylabel("Grades",font="Times New Roman")

plt.show()
```



In []:

