



北京航空航天大学  
BEIHANG UNIVERSITY

# 大学计算机基础

北京航空航天大学  
计算机学院



# 课程目录

第1章 计算与计算思维 (2学时)

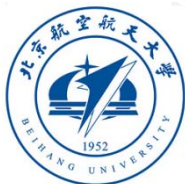
第2章 算法与问题求解 (2学时)

第3章 算法的程序实现 (12学时)

第4章 数据的组织 (12学时)

第5章 数据处理与应用 (8学时)

第6章 计算机科学发展前沿 (4学时)





## 第5章 数据处理与应用

5.1 对文件的访问

5.2 数据可视化方法

5.3 使用Excel绘图（自学）

5.4 使用Matplotlib绘图





## 5.2 数据可视化方法

- ◆ 5.2.1 数据可视化
- ◆ 5.2.2 典型图形展示方式及对比
- ◆ 5.2.3 运算结果可视化工具





## 让我们来做个对比

- **神舟飞天瞬间，民航为其“让”出一条巨大通道**
- **在神舟飞船向太空冲刺的20分钟里，如果你能够查看空管雷达，就会发现，从内蒙古西部到山东江苏交界的黄海之滨，我国民航客机规整地“让”出了一条巨大通道，而通道中央，就是神舟十一号拔地而起，造访天宫的轨迹。**





## 让我们来做对比









## 5.2.1 数据可视化

- 数据可视化技术的基本思想是将每一个数据项作为单个图元元素表示，大量的数据集构成数据图像，同时将数据的各个属性值以多维数据的形式表示，可以从不同的维度观察数据，从而对数据进行更深入的观察和分析，应用领域包括：
  - 数据采集
  - 数据分析
  - 数据管理
  - 数据挖掘
  - 科学计算结果展示与科技论文撰写

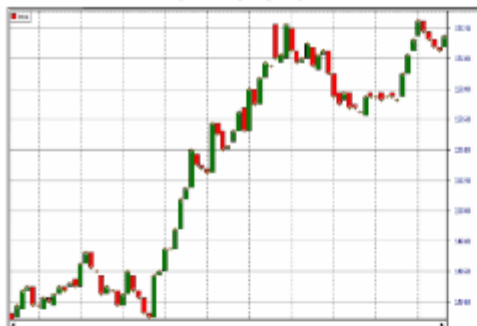




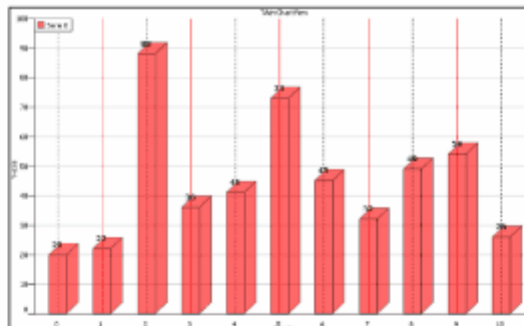


# 典型图形展示方式及对比

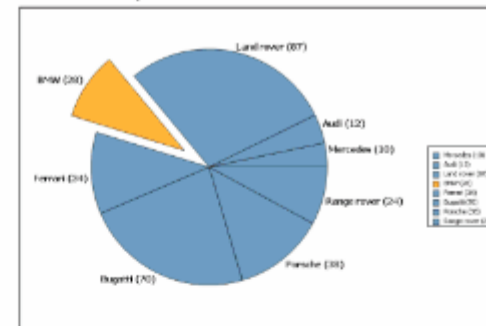
OHLC Chart



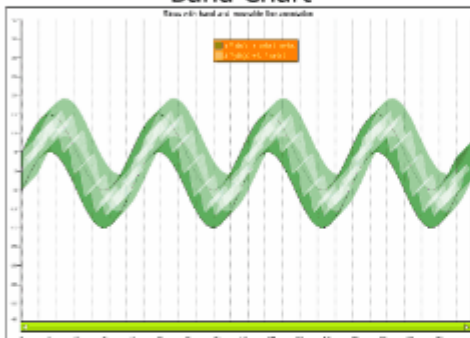
3D Bar Chart



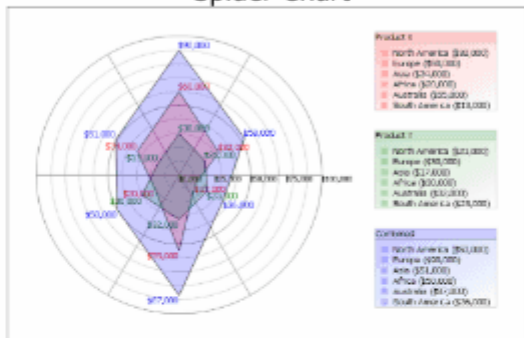
Exploded Pie Slice Chart



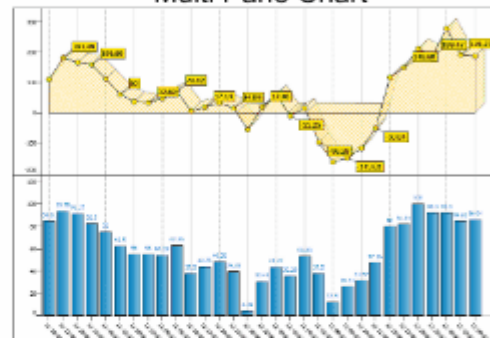
Band Chart



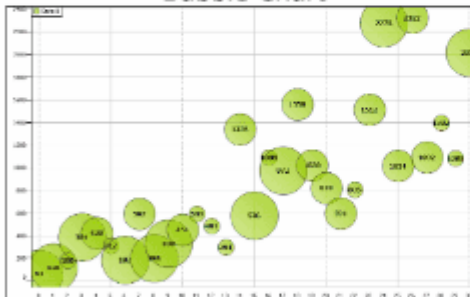
Spider Chart



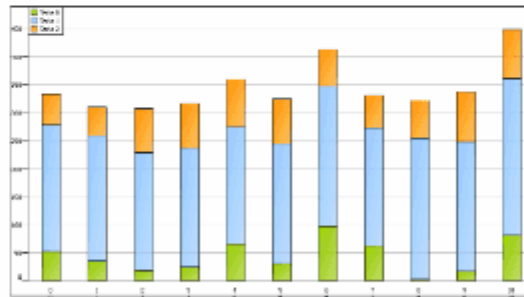
Multi Pane Chart



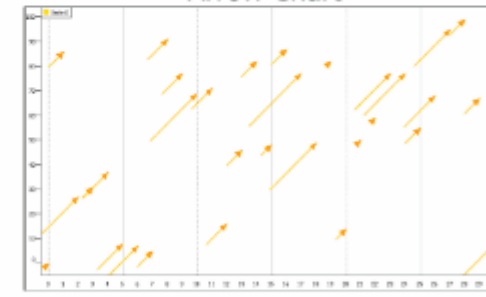
Bubble Chart



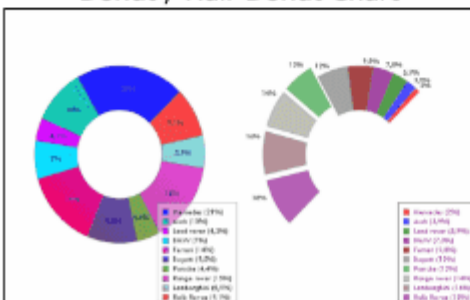
Stacked Bar Chart



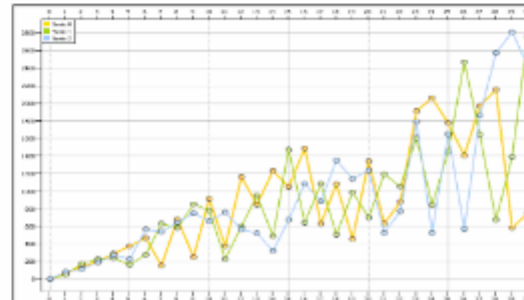
Arrow Chart



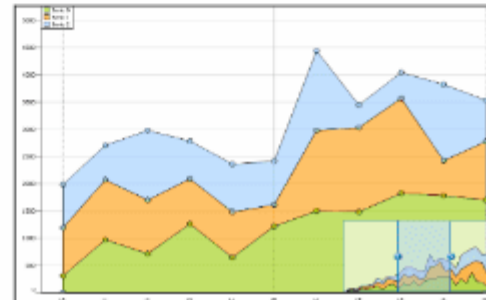
Donut / Half-Donut Chart



Line Chart



Area Chart with Zoomwindow



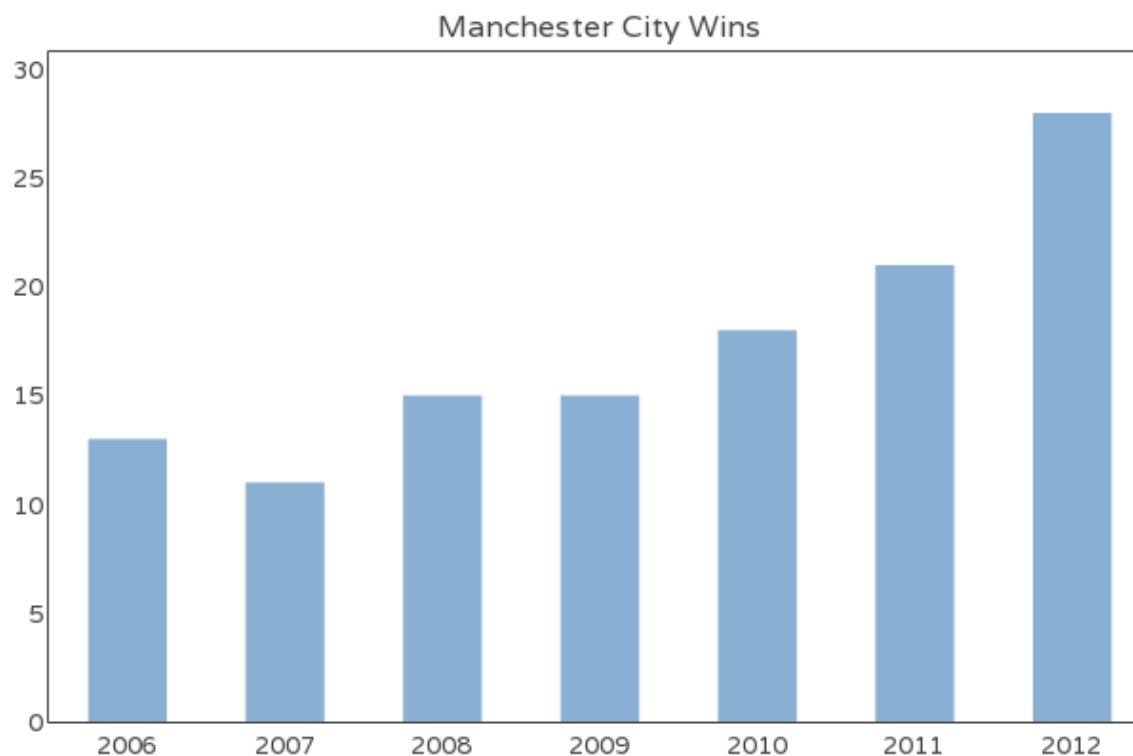


## 5.2.2 典型图形展示方式及对比

### ■ 柱状图

- 适用场合是二维数据集（每个数据点包括两个值 $x$ 和 $y$ ），但只有一个维度需要比较

年销售额示例



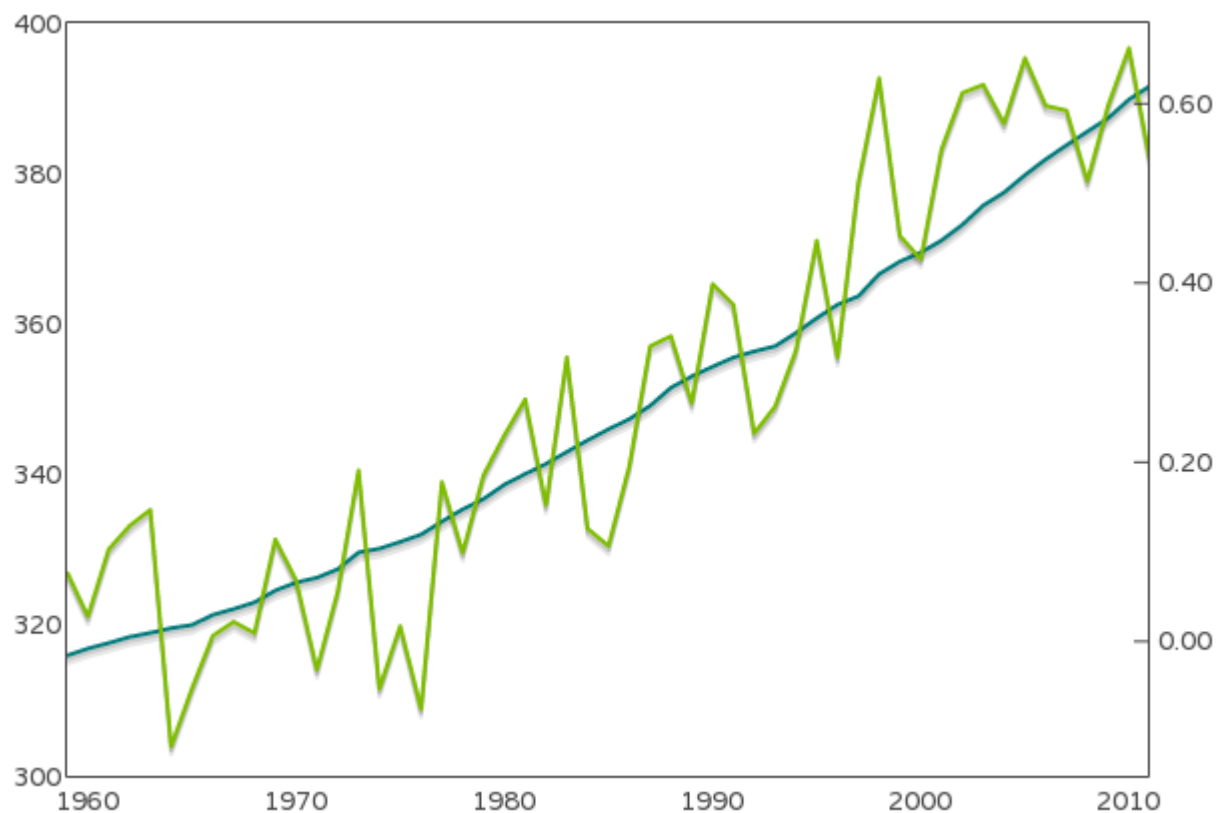


## 5.2.2 典型图形展示方式及对比

### ■ 折线图

- 适合二维的大数据集，尤其是那些趋势比单个数据点更重要的场合

尤其适用于  
两个数据  
趋势比较



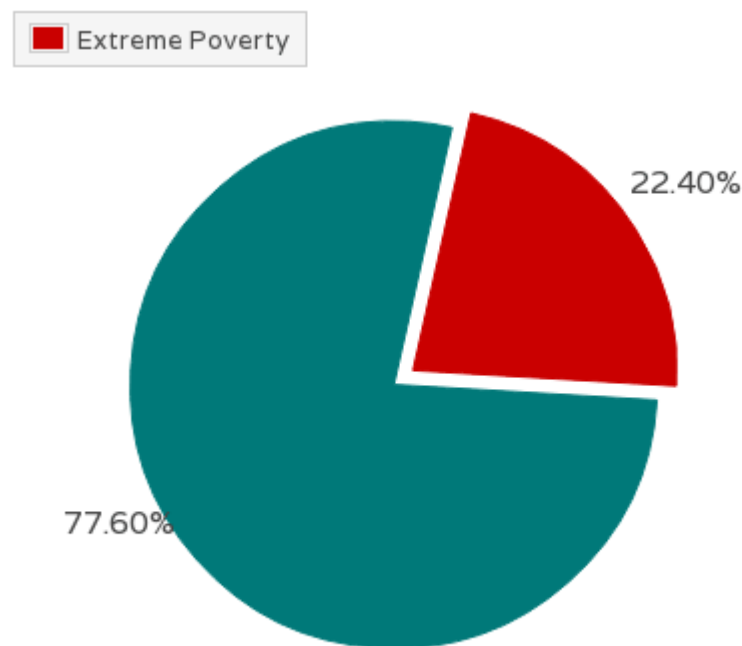


## 5.2.2 典型图形展示方式及对比

### ■ 饼图

- 反映某个部分占整体的比重

贫穷人口占总体  
人口比例



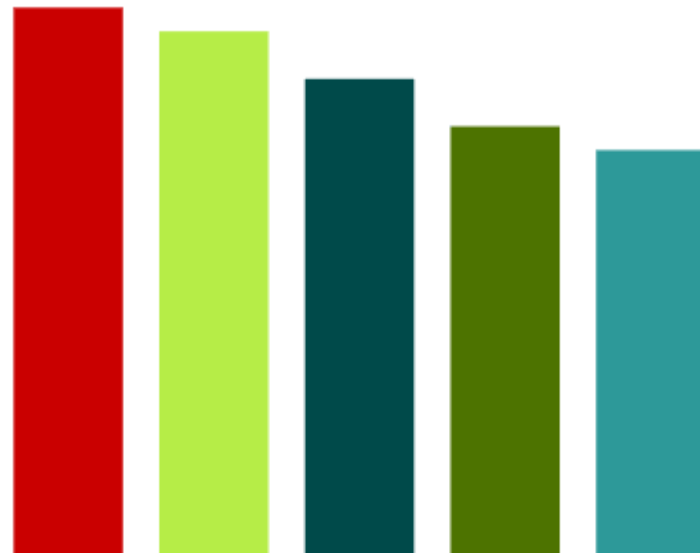


## 5.2.2 典型图形展示方式及对比

### ■ 饼图

- 肉眼对面积大小不敏感，所以，一般情况常用柱状图替代饼图

左侧饼图的五个色块的面积排序，不容易看出来。换成柱状图，更加直观



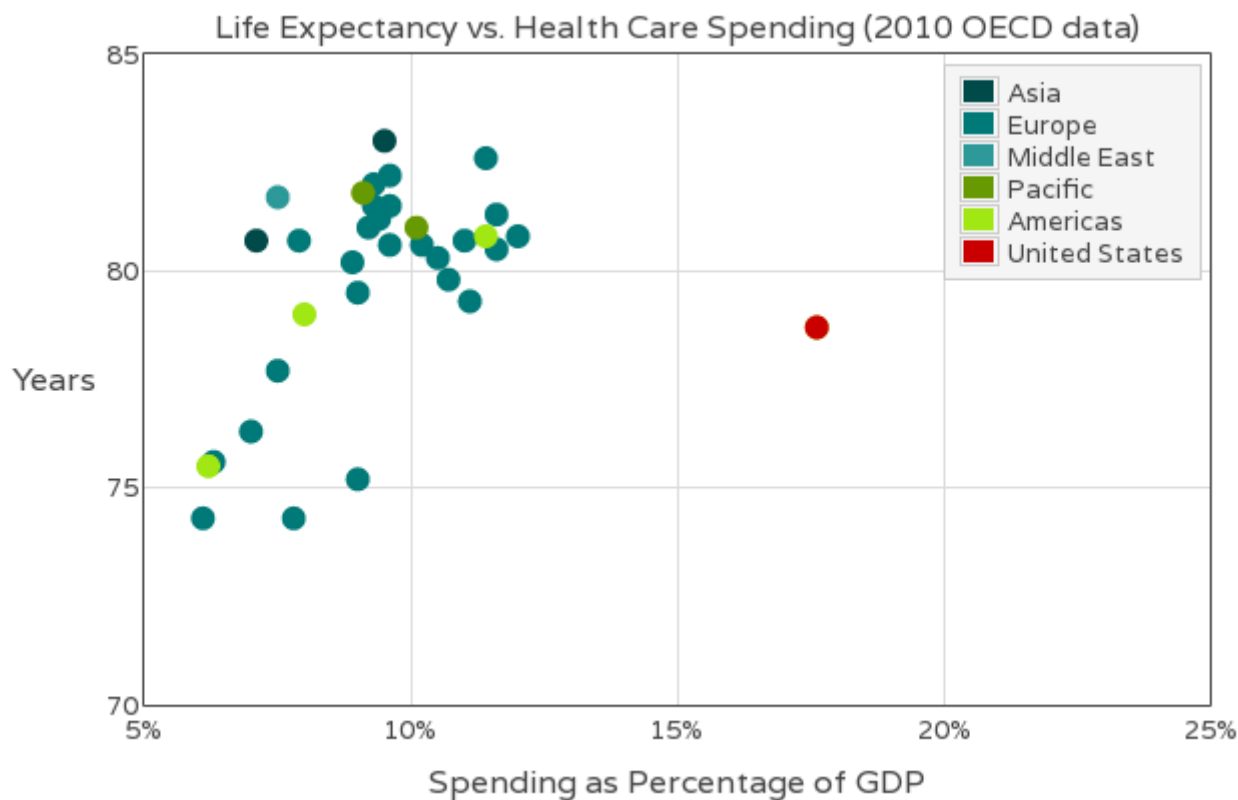


## 5.2.2 典型图形展示方式及对比

### ■ 散点图

- 散点图适用于三维数据集，但其中只有两维需要比较

各国的医疗支出与预期寿命，三个维度分别为国家、医疗支出、预期寿命，只有后两个维度需要比较



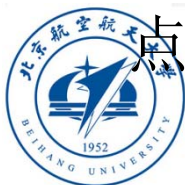
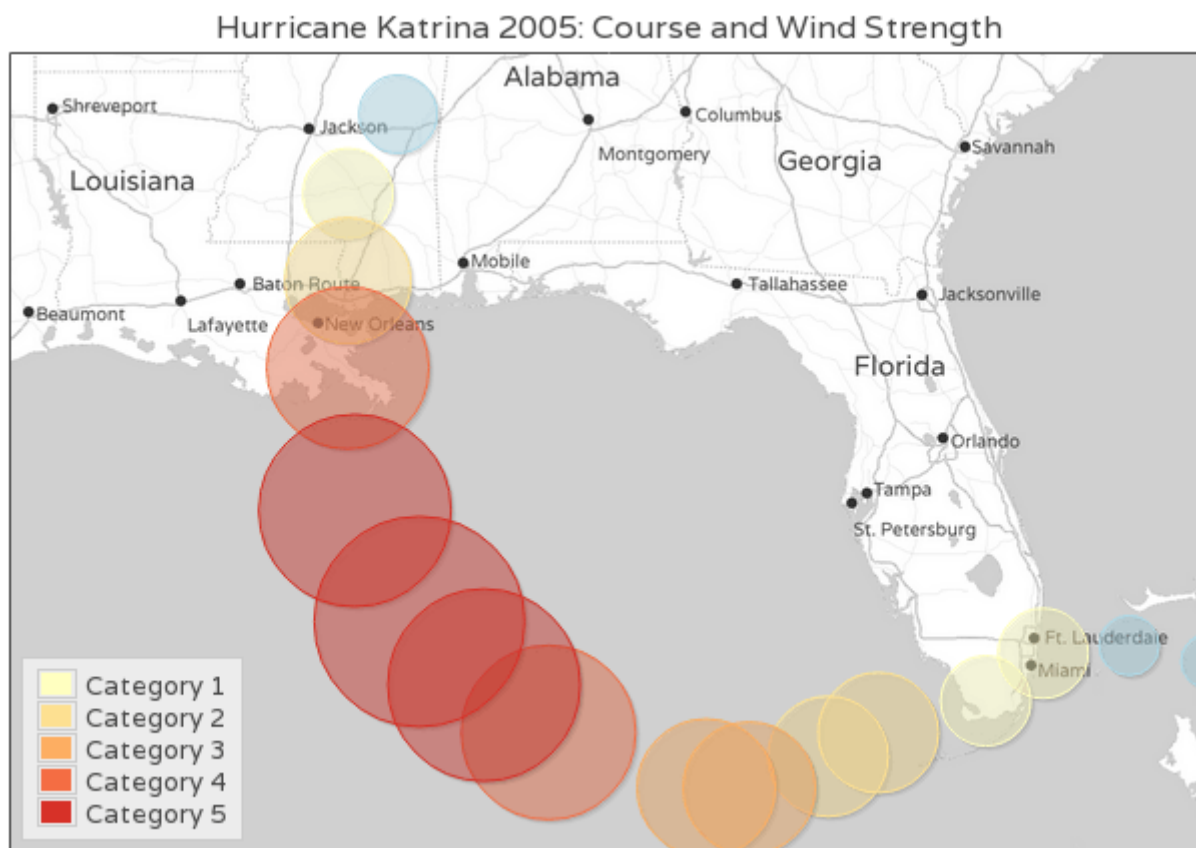


## 5.2.2 典型图形展示方式及对比

### ■ 气泡图

■ 散点图的一种变体，通过每个点的面积大小，反映第三维

卡特里娜飓风的路径，三个维度分别为经度、纬度、强度。点的面积越大，就代表强度越大；颜色，表示每个点的风力等级





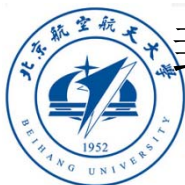
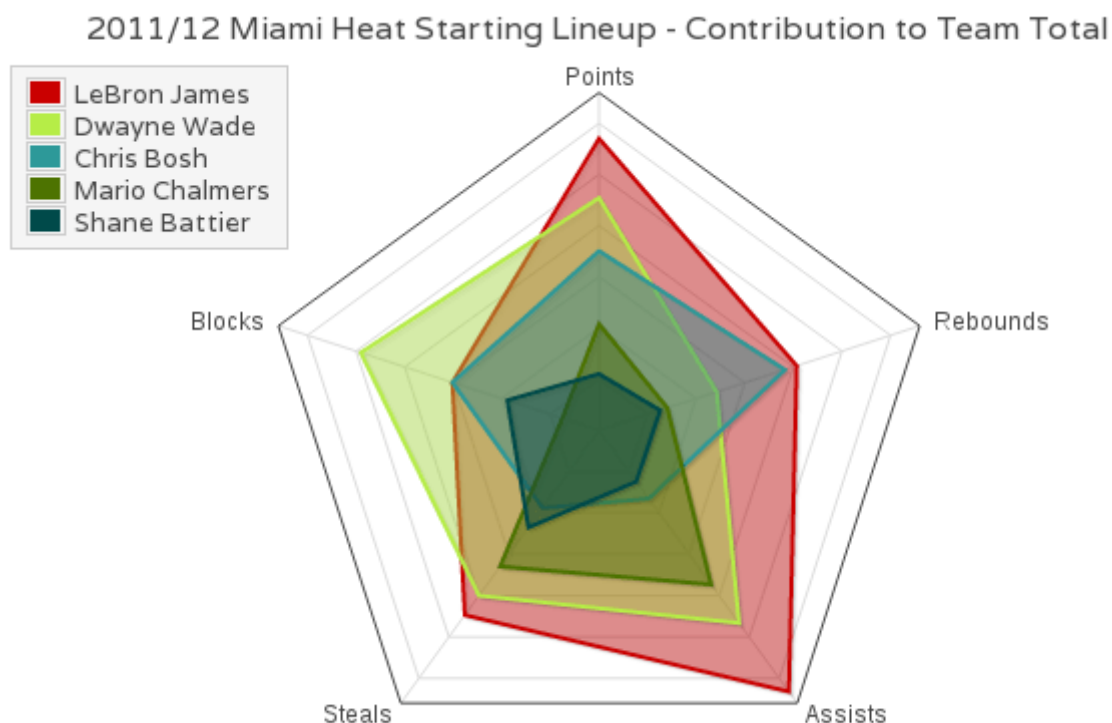


## 5.2.2 典型图形展示方式及对比

### ■ 雷达图

- 雷达图适用于多维数据（四维以上），且每个维度必须可以排序，它有一个局限，就是数据点最多6个，否则无法辨别，因此适用场合有限

迈阿密热火队  
首发的五名篮球  
球选手的数据。  
除了姓名，每  
个数据点有五  
个维度，分别  
是得分、篮板、  
助攻、抢断、  
封盖

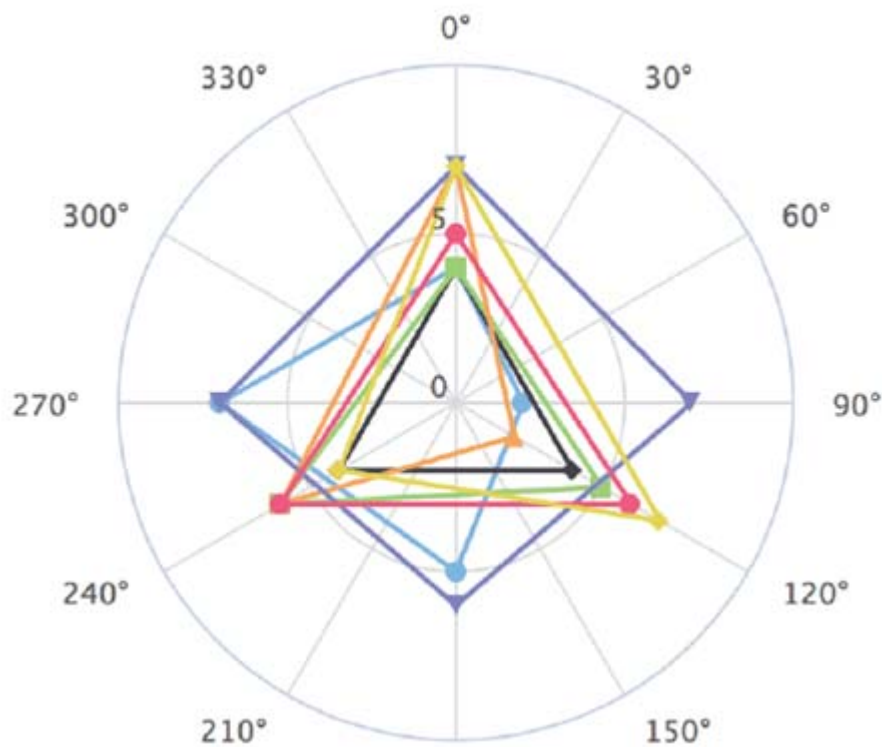




## 5.2.2 典型图形展示方式及对比

### ■ 雷达图

- 雷达图适用于多维数据（四维以上），且每个维度必须可以排序，它有一个局限，就是数据点最多6个，否则无法辨别，因此适用场合有限





## 5.2.2 典型图形展示方式及对比

图表	维度	注意点
柱状图	二维	只需比较其中一维
折线图	二维	适用于较大的数据集
饼图	二维	只适用反映部分与整体的关系
散点图	二维或三维	有两个维度需要比较
气泡图	三维或四维	只有两个维度可以精确辨识
雷达图	四维以上	数据点不超过6个





## 5.2.3 运算结果可视化工具

### ■ 常用的科学计算绘图工具

- ◆ **Excel** : 常用Office套件
- ◆ **Matlab** : MATLAB是美国MathWorks公司出品的商业数学软件，用于算法开发、数据可视化、数据分析以及数值计算的高级技术计算语言和交互式环境
- ◆ **matplotlib** : matplotlib是基于Python语言的开源项目，旨在为Python提供一个数据绘图包





## 5.2.3 运算结果可视化工具

- 使用Excel绘制柱状图和饼图
- 请结合课件和 “学生.xlsx” 自行学习



# 案例：制作三维簇状柱形图

- ▶ 打开 “学生” 工作簿的 “成绩” 的工作表
- ▶ 选择 “姓名” 列、“高等数学” 列、“外语” 列、“平均分” 列制作 “三维簇状柱形图”
- ▶ 关注：
  - 选择数据、修改系列、修改分类轴、图表名称与坐标轴名



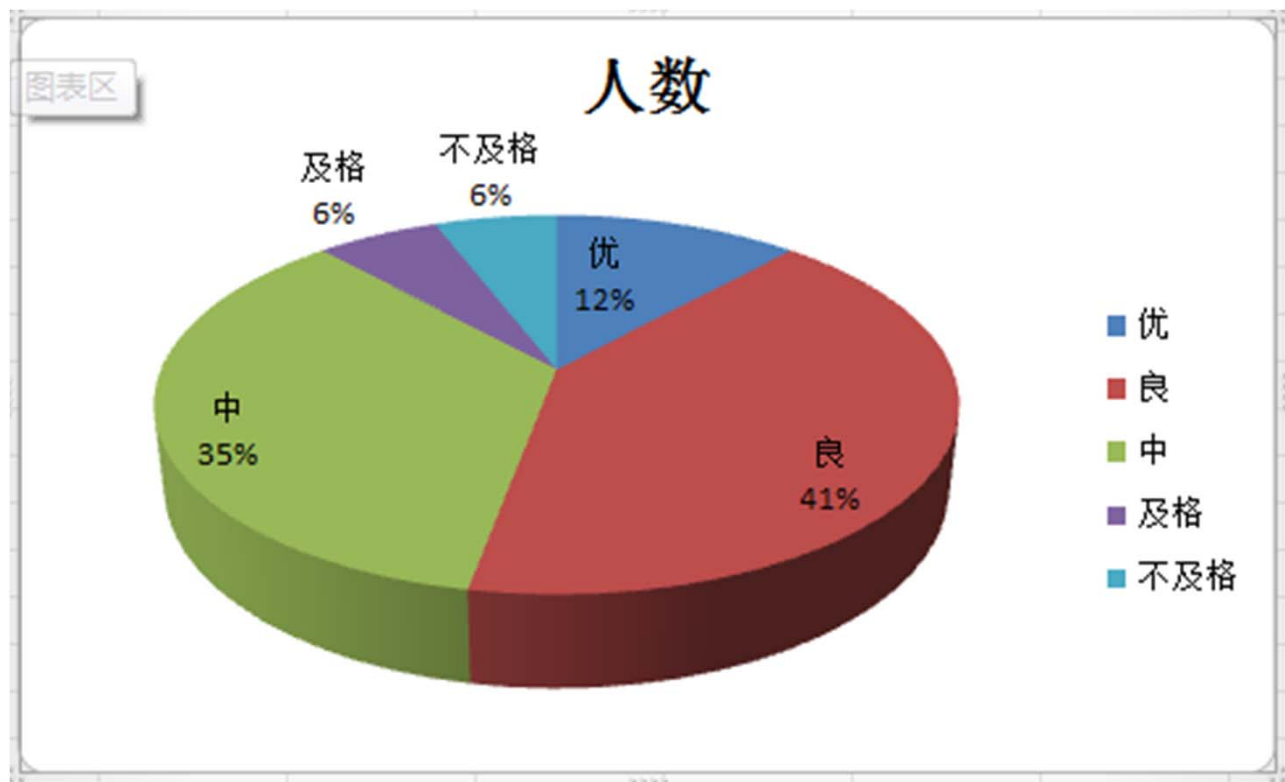
## 步骤提示：成绩簇状柱形图

- ▶ 先选中“学号”等字段名与数据区域，单击【插入】|【柱形图】中的簇状柱形图。
- ▶ 调整图表的大小后，在【布局】选项卡添加【坐标轴标题】和【坐标轴标题】。
- ▶ 单击选中图表上代表平均分的柱形，单击右键【添加数据标签】



## 案例：制作三维饼图

- 根据平均分，统计五级评分结果，制作三维饼图



# 步骤提示：饼图

- ▶ 选中“评级”和“人数”字段名及其列数据，单击【插入】|【饼图】中的三维饼图
- ▶ 【布局】|【数据标签】|【其他数据标签选项】，设置标签包括【类别名称】和【百分比】

评级	人数
优	2
良	7
中	6
及格	1
不及格	1



## 5.3 matplotlib

- ◆ 5.3.1 matplotlib安装
- ◆ 5.3.2 绘制柱状图
- ◆ 5.3.3 绘制折线图
- ◆ 5.3.4 其它图形的绘制





## 5.3.1 matplotlib安装

### ■ 安装配置

- ◆ **matplotlib** : <http://matplotlib.org/>
- ◆ **numPy** : matplotlib依赖包 ,  
<http://www.lfd.uci.edu/~gohlke/pythonlibs/#numpy>
- ◆ 安装指令 : `pip install matplotlib`





## 5.3.1 matplotlib 内容纵览

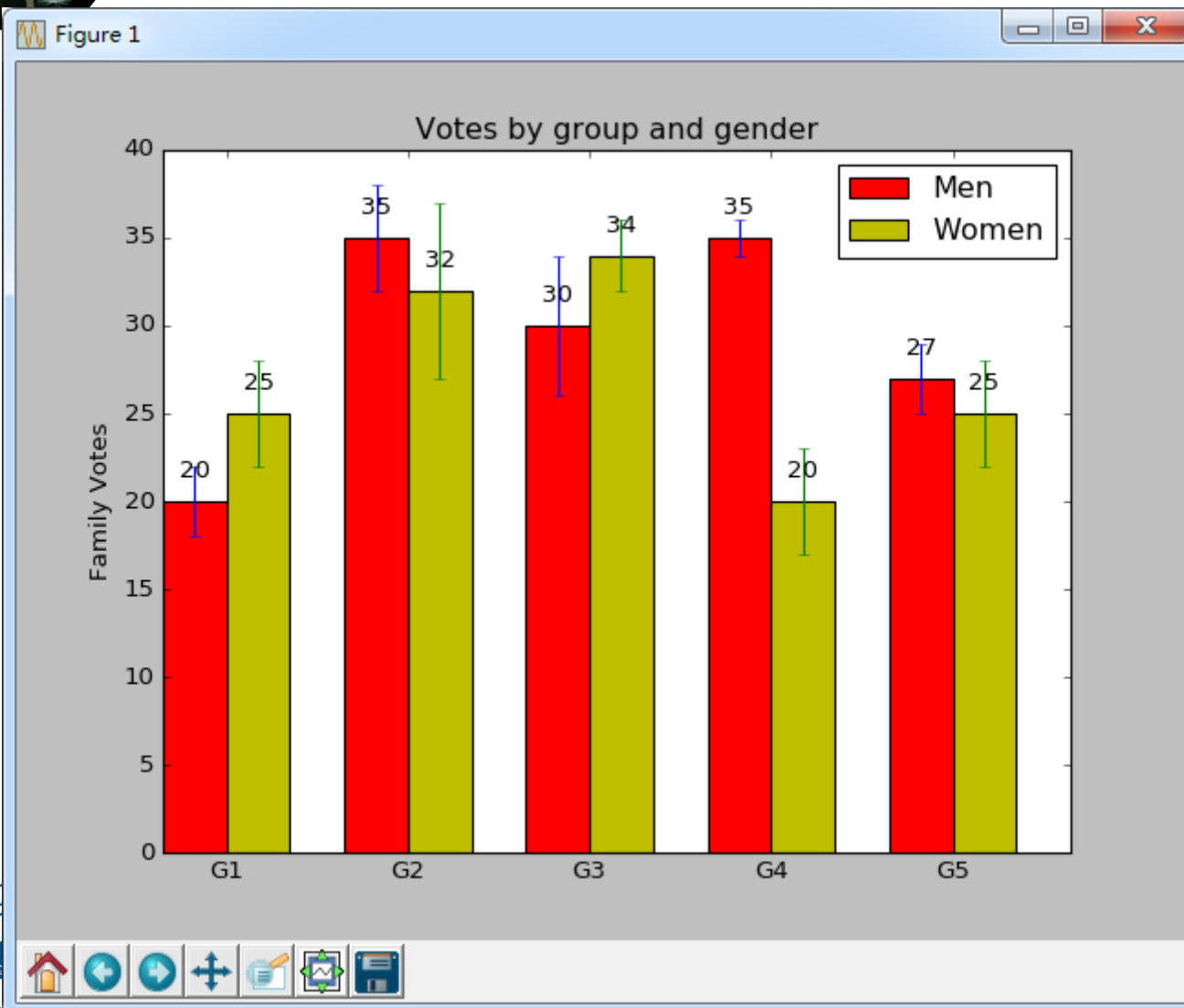
- ◆ matplotlib的关键绘制功能，均集中在matplotlib.pyplot中，例如：
  - ◆ matplotlib.pyplot.bar 用于绘制柱状图
  - ◆ matplotlib.pyplot.plot 用于绘制折线图
  - ◆ matplotlib.pyplot.xlabel 用户设置坐标系X轴的说明
- ◆ matplotlib的官网，可以查阅所有支持的绘制函数

[http://matplotlib.org/api/pyplot\\_summary.html](http://matplotlib.org/api/pyplot_summary.html)

<http://docs.scipy.org/doc/numpy/genindex.html#A>



## 5.3.2 绘制柱状图



### ◆ 要素

- ◆ 数据
- ◆ 误差
- ◆ 图例
- ◆ 坐标轴



## 5.3.2 绘制柱状图

### ◆ 数据

```
#!/usr/bin/env python
# a bar plot with errorbars
import numpy as np
import matplotlib.pyplot as plt
N = 5
menMeans = (20, 35, 30, 35, 27)
ind = np.arange(N) # the x locations for the groups
width = 0.35 # the width of the bars
fig, ax = plt.subplots()
rects1 = ax.bar(ind, menMeans, width)
plt.show()
```

导入matplotlib库

待绘制的数据

横坐标, 5个元素

每个柱的宽度

绘制

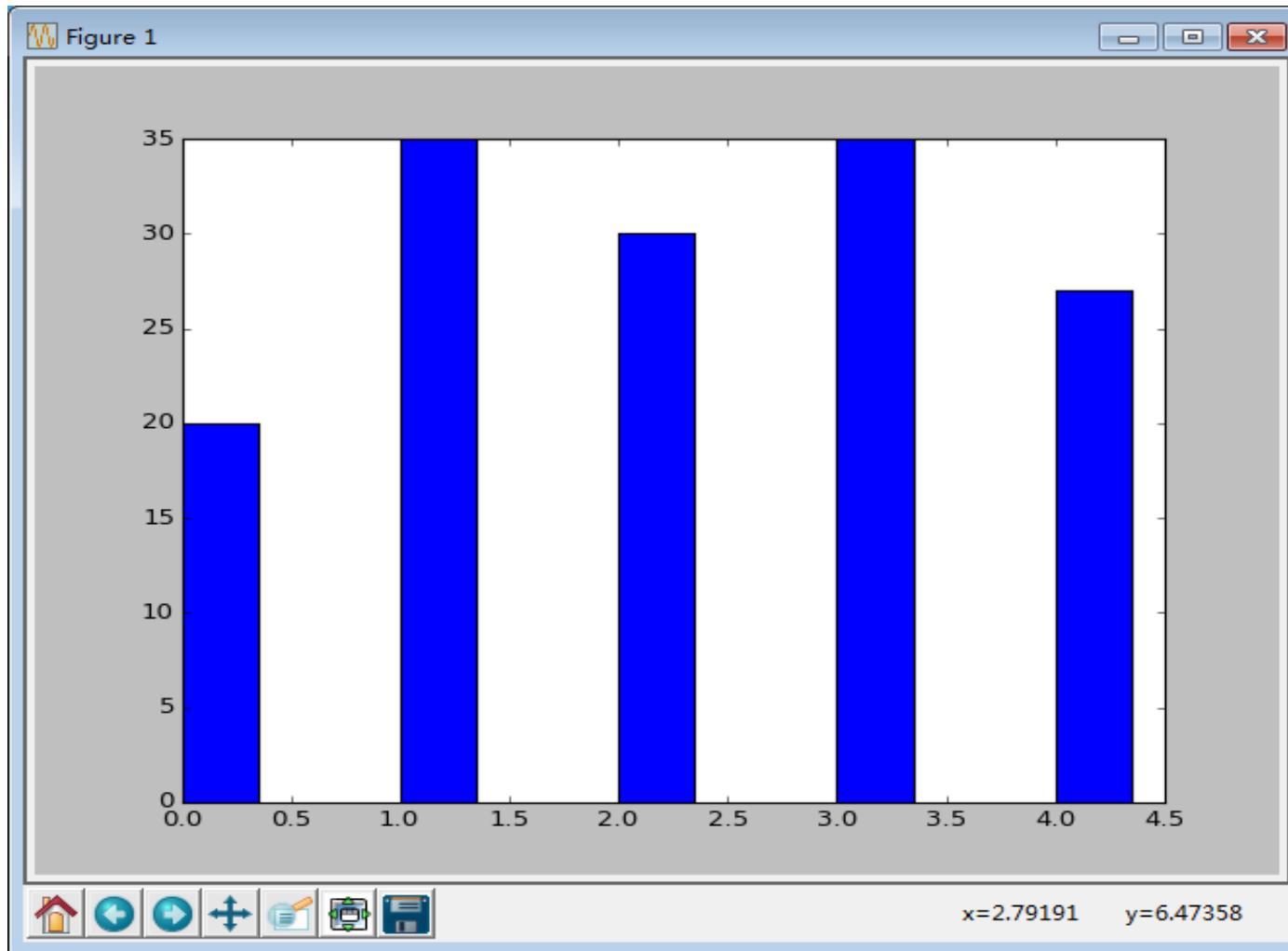
展示出来







## 5.3.2 绘制柱状图





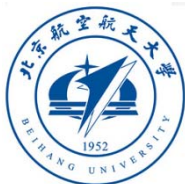
## 5.3.2 绘制基本要素

**matplotlib** 是python最著名的绘图库，它提供了一整套和**matlab**相似的命令API，十分适合交互式地进行制图。而且也可以方便地将它作为绘图控件，嵌入GUI应用程序中。

它的文档相当完备，并且 **Gallery**页面 中有上百幅缩略图，打开之后都有源程序。因此如果你需要绘制某种类型的图，只需要在这个页面中浏览/复制/粘贴一下，基本上都能搞定。

matplotlib中的快速绘图的函数库可以通过如下语句载入：

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





## 5.3.2 绘制基本要素

pylab模块

matplotlib还提供了名为pylab的模块，其中包括了许多numpy和pyplot中常用的函数，方便用户快速进行计算和绘图，

figure创建一个绘图对象，并且使它成为当前的绘图对象。

plot函数的调用方式很灵活，用关键字参数指定各种属性：

**label:** 给所绘制的曲线一个名字，此名字在图例(legend)中显示。

只要在字符串前后添加"\$"符号，matplotlib就会使用其内嵌的latex引擎绘制的数学公式。

**color:** 指定曲线的颜色

**linewidth:** 指定曲线的宽度





## 5.3.2 绘制柱状图

### ◆ 误差

为了更加清楚的表示数据的统计规律，我们习惯为每一个统计量加上误差

```
#!/usr/bin/env python
# a bar plot with errorbars
import numpy as np
import matplotlib.pyplot as plt
N = 5
menMeans = (20, 35, 30, 35, 27)
menStd = (2, 3, 4, 1, 2)
ind = np.arange(N) # the x locations for the groups
width = 0.35      # the width of the bars
fig, ax = plt.subplots()
rects1 = ax.bar(ind, menMeans, width, color = 'r',
yerr=menStd)
plt.show()
```

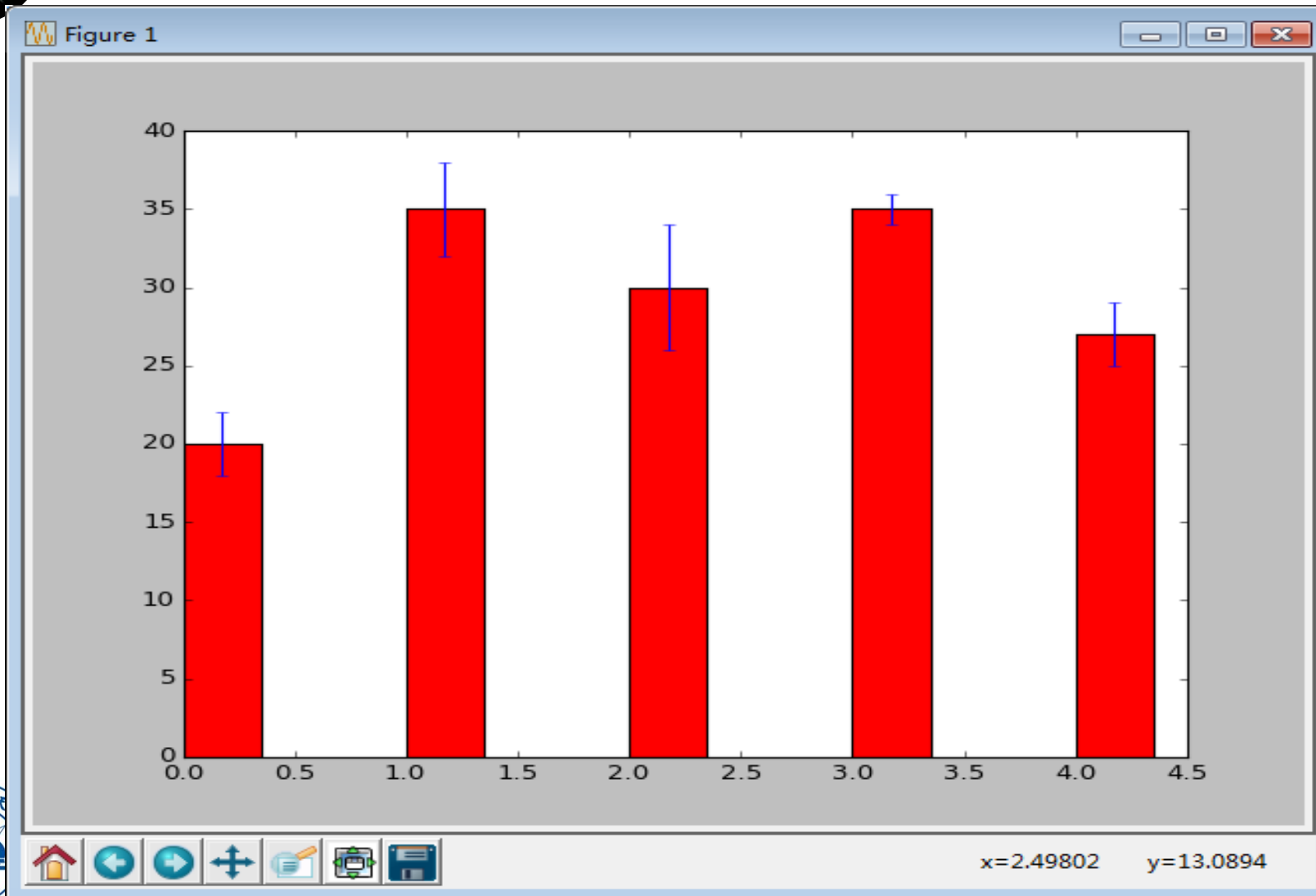
误差数据

颜色变化，红色





## 5.3.2 绘制柱状图





## 5.3.2 绘制柱状图

### ◆ 图例

如果一幅图中有多种数据，为了增强图的可读性，可以加入图例

```
#!/usr/bin/env python
# a bar plot with errorbars
import numpy as np
import matplotlib.pyplot as plt
N = 5
menMeans = (20, 35, 30, 35, 27)
womenMeans = (25, 32, 34, 20, 25)
menStd = (2, 3, 4, 1, 2)
womenStd = (3, 5, 2, 3, 3)
ind = np.arange(N) # the x locations for the groups
width = 0.35 # the width of the bars
fig, ax = plt.subplots()
rects1 = ax.bar(ind, menMeans, width, color = 'r', yerr=menStd)
rects2 = ax.bar(ind+width, womenMeans, width, color='y',
yerr=womenStd)
ax.legend( (rects1[0], rects2[0]), ('Men', 'Women') )
plt.show()
```

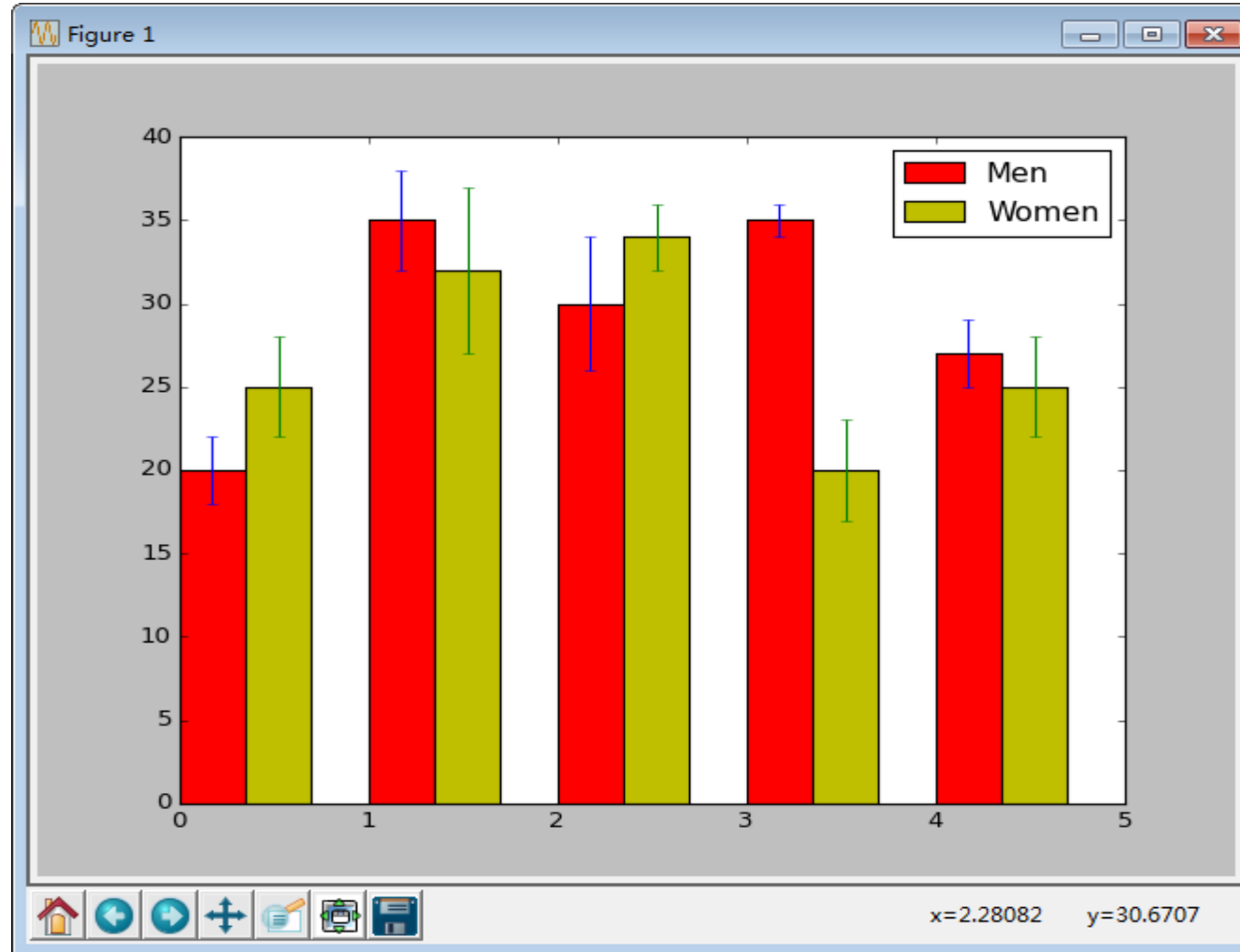
两类数据

图例





## 5.3.2 绘制柱状图







## 5.3.2 绘制柱状图

### ◆ 坐标轴

```
# add some text for labels, title and axes ticks
ax.set_ylabel('Family votes')
ax.set_title('Votes by group and gender')
ax.set_xticks(ind+width)
ax.set_xticklabels( ('G1', 'G2', 'G3', 'G4', 'G5') )
```

纵坐标标题

题目

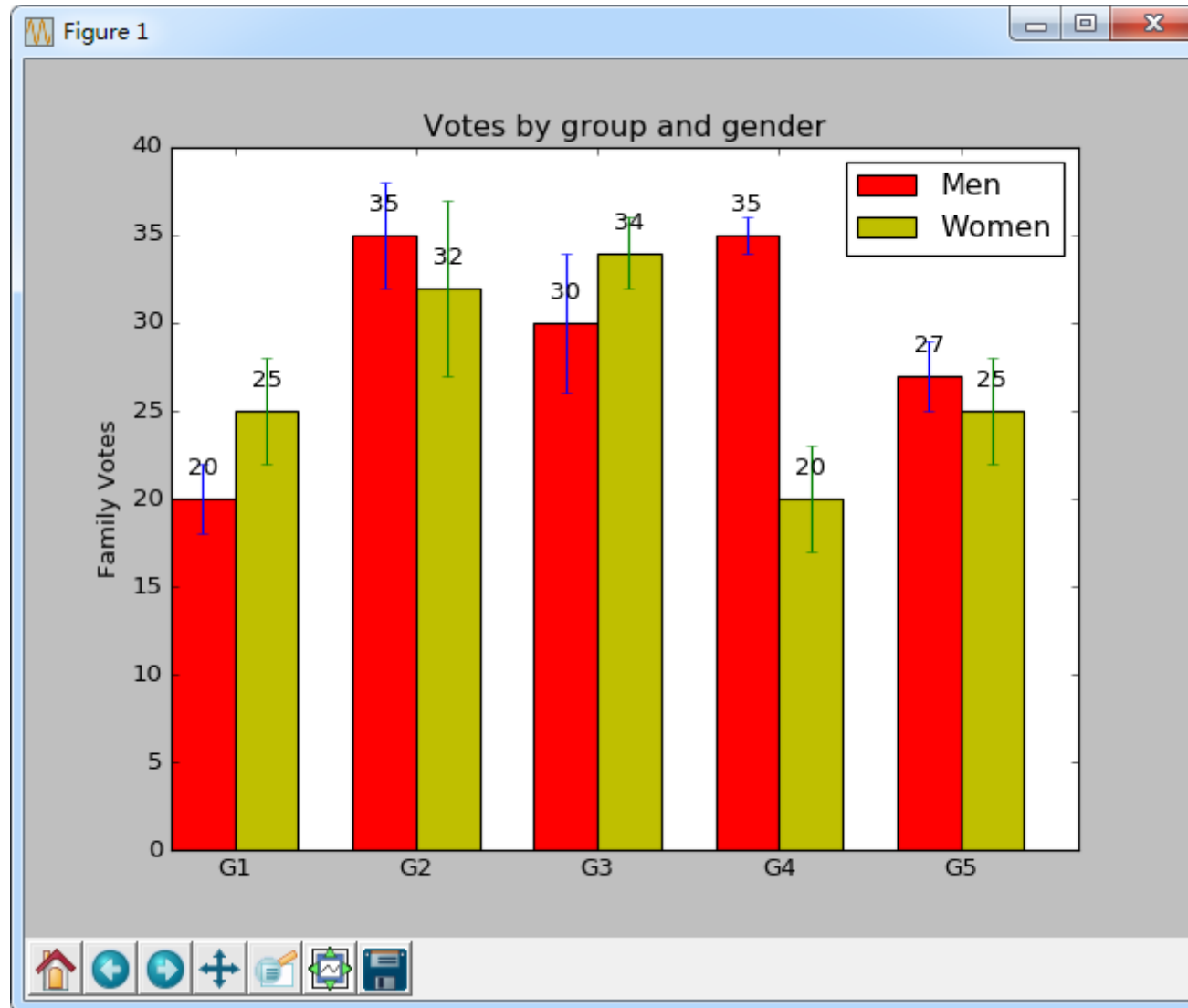
横坐标标题

```
# attach some text labels
for rect in rects:
    height = rect.get_height()
    ax.text(rect.get_x()+rect.get_width()/2., 1.05*height,
            height, ha='center', va='bottom')
```





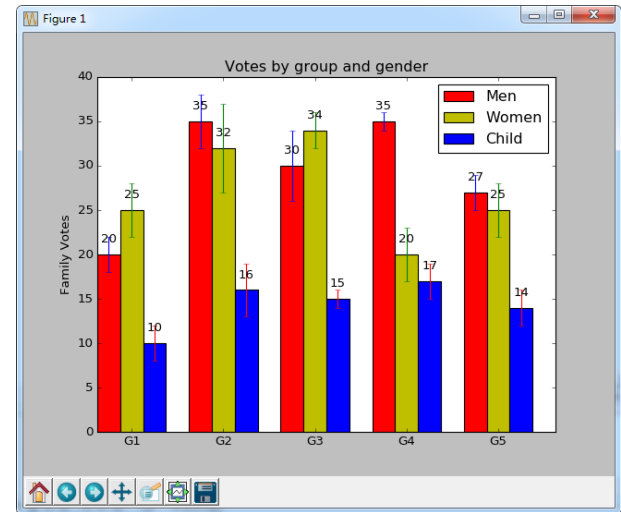
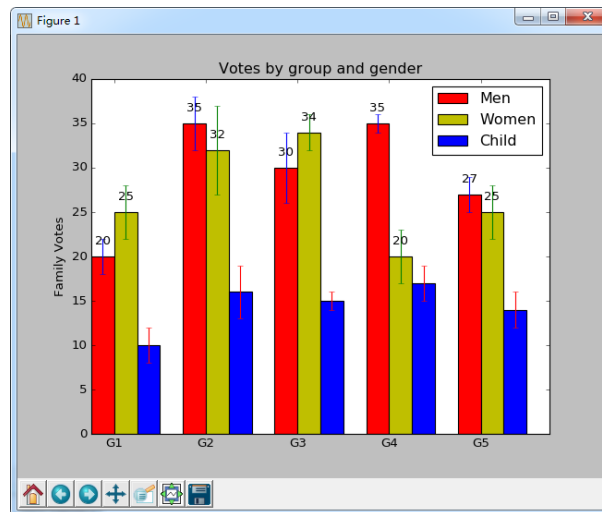
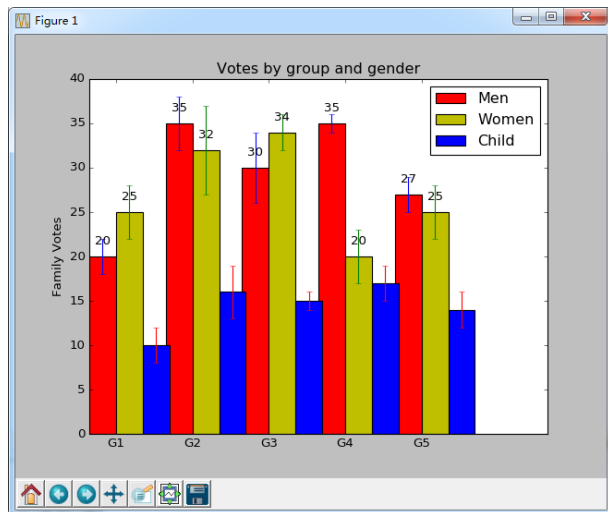
## 5.3.2 绘制柱状图





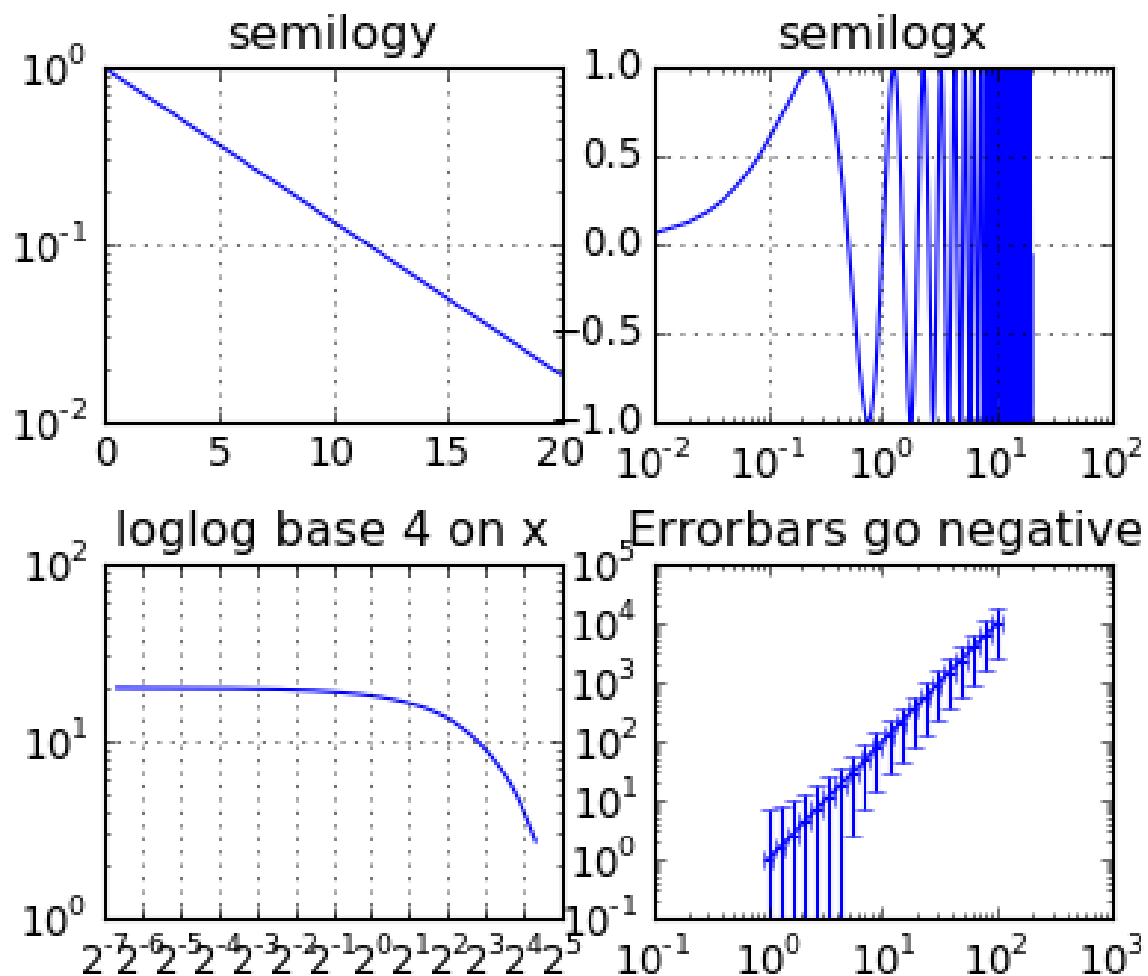
## 课内练习1

- 请为上面的柱状图增加一组"柱子"，代表Child，取值为 ( 10, 16, 15, 17, 14 )，误差为 ( 2, 3, 1, 2, 2 )
- 第一步：立出第三组"柱子"
- 第二步、第三步：一直美下去.....





## 5.3.3 绘制折线图



### 要素

- ◆ 折线
- ◆ 颜色、数据点与线型
- ◆ 子图



## 5.3.3 绘制折线图

### ◆ 折线/曲线

```
import numpy as np  
import matplotlib.pyplot as plt
```

```
x1 = np.linspace(0.0, 5.0)  
y1 = np.cos(2 * np.pi * x1) * np.exp(-x1)
```

```
plt.plot(x1, y1, 'ko-')
```

```
plt.title('A tale of 2 subplots')  
plt.ylabel('Damped oscillation')  
plt.show()
```

数据准备

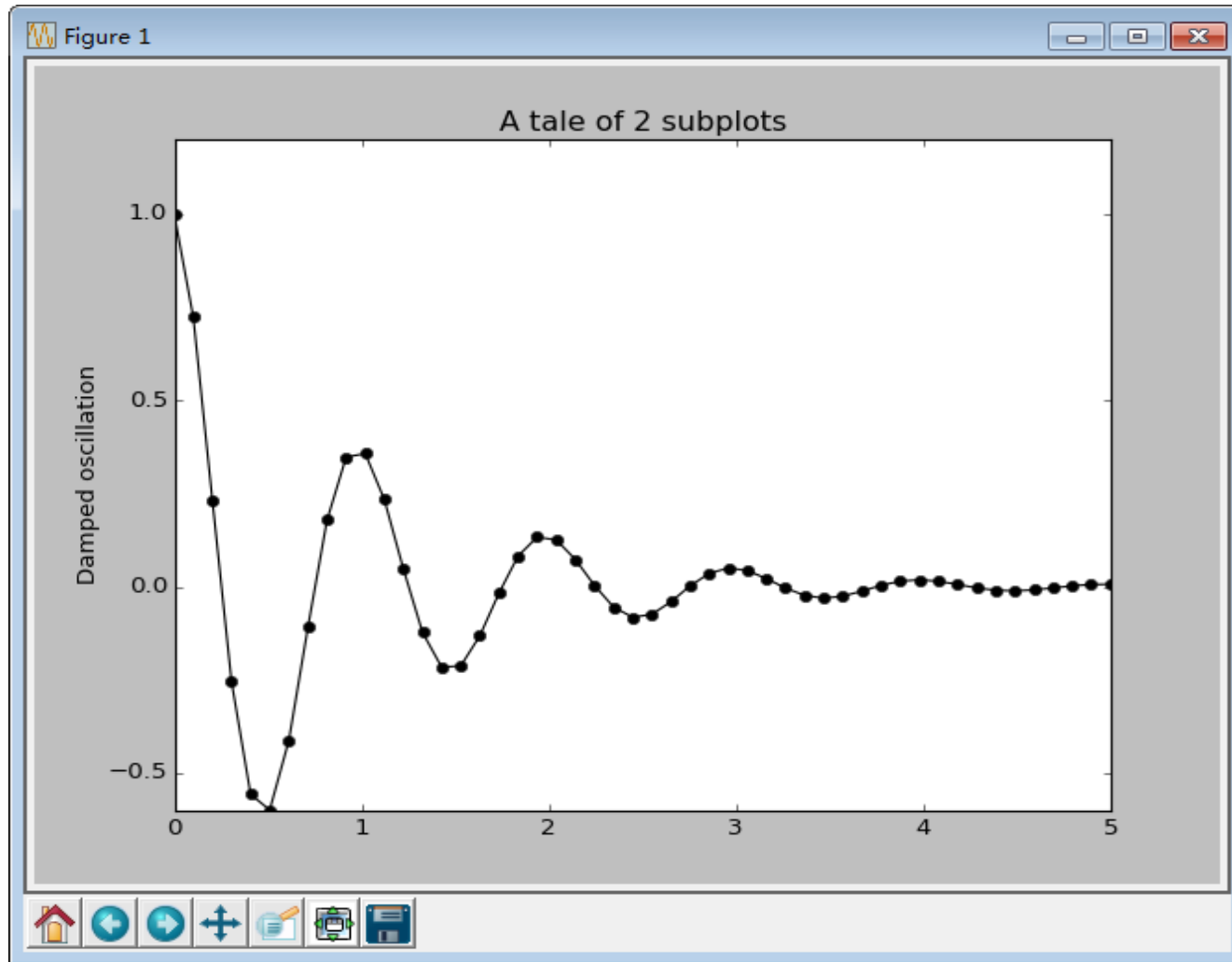
绘图

坐标轴





## 5.3.3 绘制折线图





## 5.3.3 绘制折线图

### ◆ 颜色、数据点与线型

```
import numpy as np
import matplotlib.pyplot as plt
x1 = np.linspace(0.0, 5.0)
x2 = np.linspace(0.0, 5.0)
y1 = np.cos(2 * np.pi * x1) *
np.exp(-x1)
y2 = np.cos(2 * np.pi * x2)
plt.plot(x1, y1, 'ko-')
plt.plot(x2, y2, 'ro-')
plt.title('A tale of 2 subplots')
plt.ylabel('Damped oscillation')
plt.show()
```

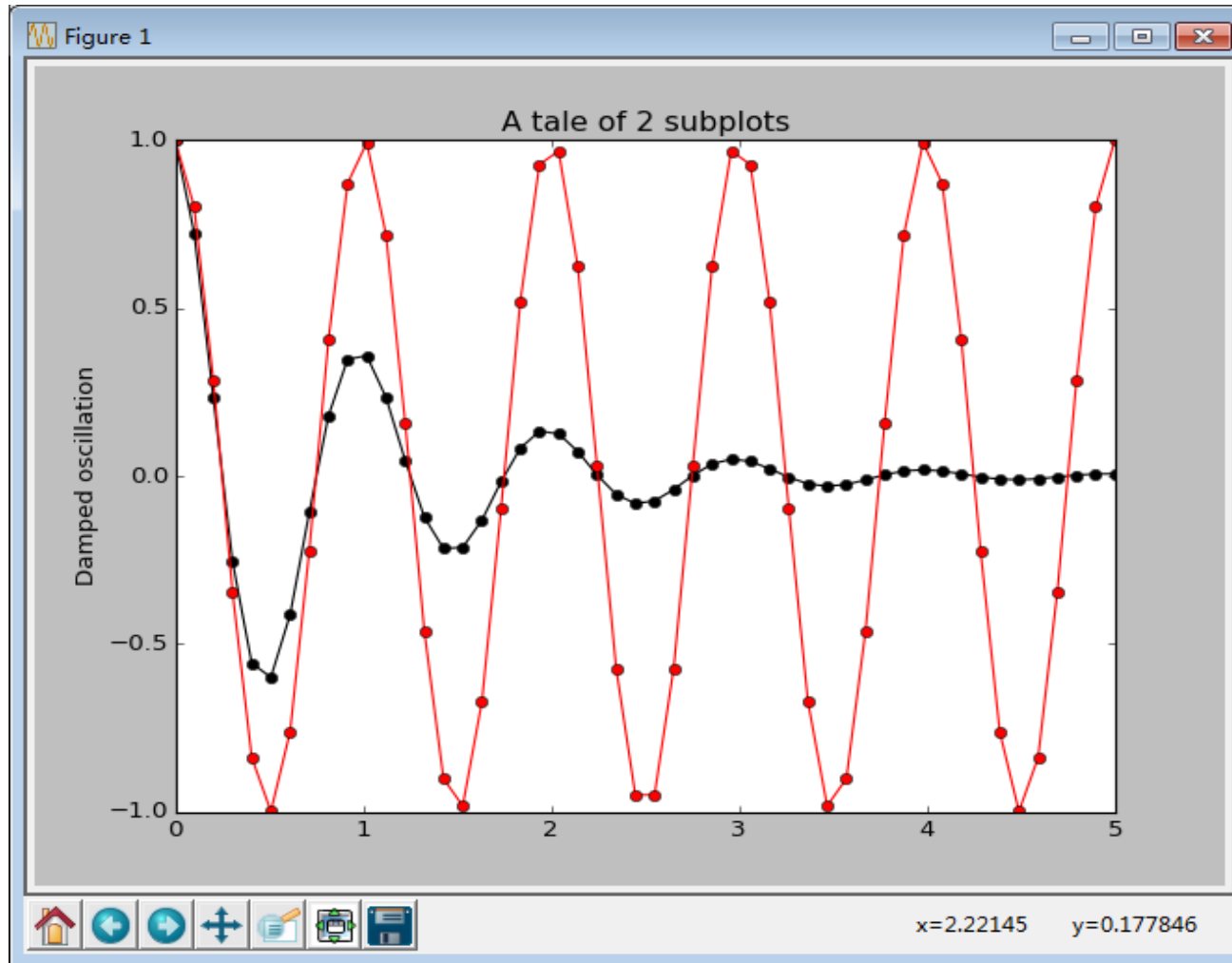
(颜色、数据点标记、线型)

两条曲线在一幅图

- 颜色: r-red, y-yellow, b-blue, k-black, g-green...
- 数据点: o-实心圆点, d-菱形, \*-\*, +-+, x-x...
- 线型: -实线, --虚线, -.点虚线



## 5.3.3 绘制折线图







## 5.3.3 绘制折线图

### ◆ 子图

```
import numpy as np
import matplotlib.pyplot as plt
x1 = np.linspace(0.0, 5.0)
x2 = np.linspace(0.0, 2.0)
y1 = np.cos(2 * np.pi * x1) * np.exp(-x1)
y2 = np.cos(2 * np.pi * x2)
plt.subplot(2, 1, 1)
plt.plot(x1, y1, 'ko-')
plt.title('A tale of 2 subplots')
plt.ylabel('Damped oscillation')
plt.subplot(2, 1, 2)
plt.plot(x2, y2, 'r.-')
plt.xlabel('time (s)')
plt.ylabel('Undamped')
plt.show()
```

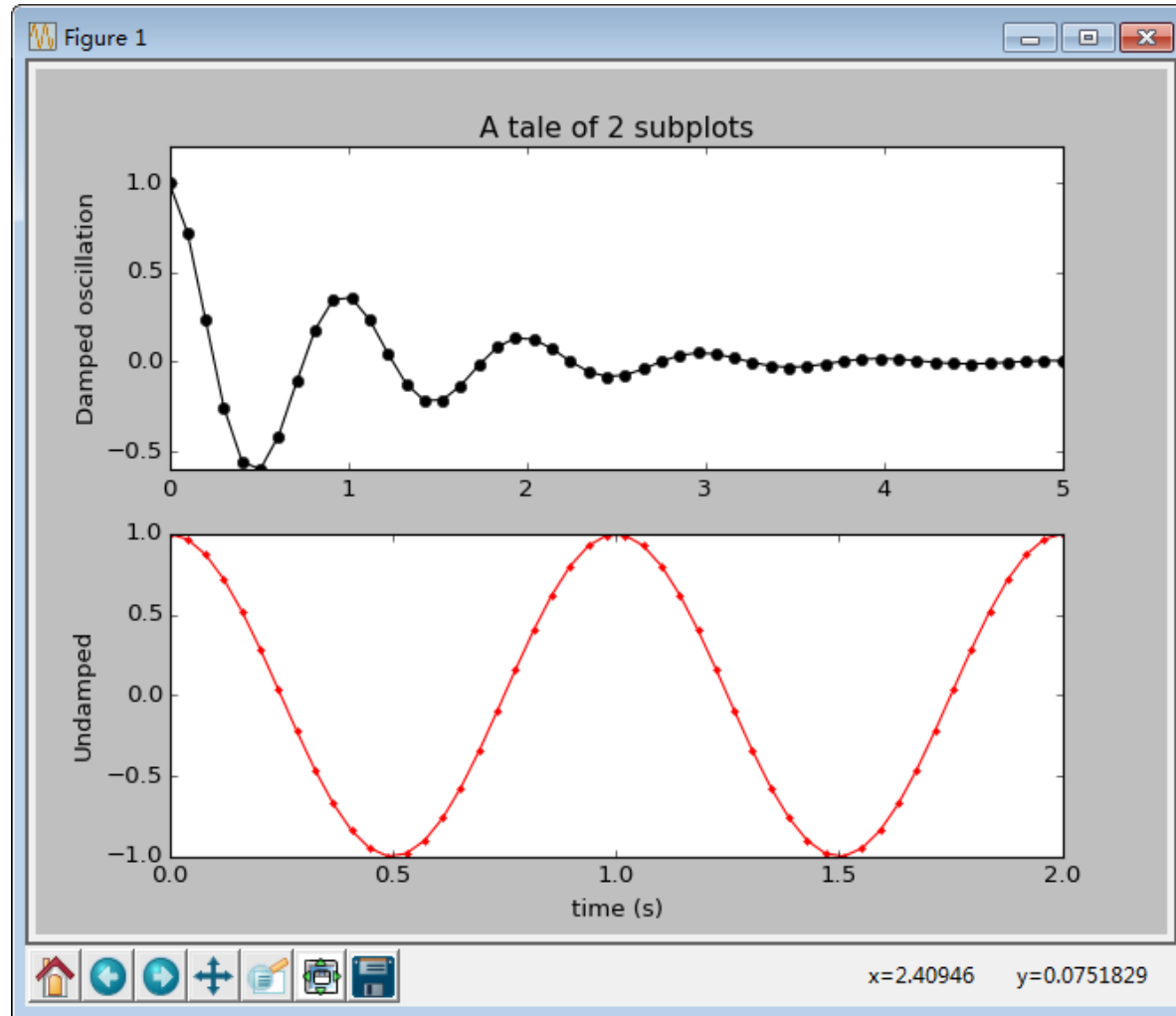
子图1（2行，1列，  
第1副图）

子图2（2行，1列，  
第2副图）





## 5.3.3 绘制折线图





## 5.3.3 绘制基本要素

一个绘图对象(`figure`)可以包含多个轴(`axis`), 在Matplotlib中用轴表示一个绘图区域, 可以将其理解为子图。

我们可以使用`subplot`函数快速绘制有多个轴的图表。`subplot`函数的调用形式如下:

**`subplot(numRows, numCols, plotNum)`**

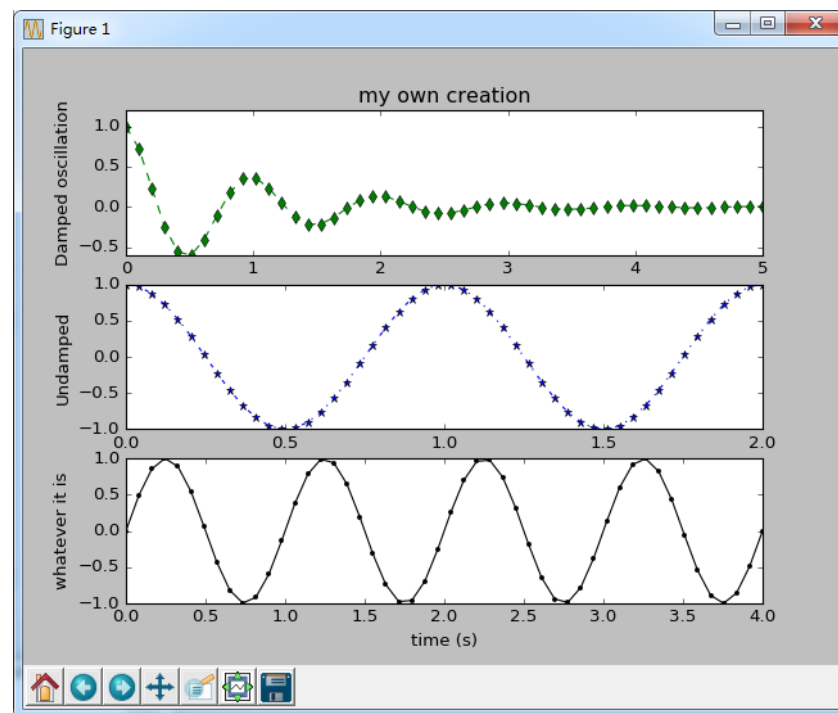
`subplot`将整个绘图区域等分为`numRows`行 $\times$ `numCols`列个子区域, 然后按照从左到右, 从上到下的顺序对每个子区域进行编号, 左上的子区域的编号为1。如果`numRows`, `numCols`和`plotNum`这三个数都小于10的话, 可以把它们缩写为一个整数, 例如`subplot(323)`和`subplot(3,2,3)`是相同的。`subplot`在`plotNum`指定的区域中创建一个轴对象。如果新创建的轴和之前创建的轴重叠的话, 之前的轴将被删除。





## 课堂练习2

- Step1：将刚才的折线图替换成不同的颜色、数据点和线型，并用注释解释说明具体的选择
- Step2：再增加一个子图，自行定义曲线函数，并设置相应的x和y坐标轴，修改图标题





## 5.3.4 其它图形绘制

- ◆ 其它图形绘制
  - ◆ Matplotlib官网 , <http://Matplotlib.org>
  - ◆ Examples页面

**matplotlib**  
Version 2.2.2

home | examples | tutorials | pyplot | docs » modules | index

### Gallery

This gallery contains examples of the many things you can do with Matplotlib. Click on any image to see the full image and source code.

For longer tutorials, see our [tutorials page](#). You can also find [external resources](#) and a [FAQ](#) in our [user guide](#).

### Matplotlib API

These examples use the Matplotlib api rather than the pylab/pyplot procedural state machine. For robust, production level scripts, or for applications or web application servers, we recommend you use the Matplotlib API directly as it gives you the maximum control over your figures, axes and plotting commands.

The example `agg_oo.py` is the simplest example of using the Agg backend which is readily ported to other output formats. This example is a good starting point if you are a web application developer. Many of the other examples in this directory use `matplotlib.pyplot` just to create the figure and show calls, and use the API for everything else. This is a good solution for production quality scripts. For full fledged GUI applications, see the `user_interfaces` examples.

**Table Of Contents**

- Gallery
- Matplotlib API
- Pyplot
- Subplots, axes and figures
- Color
- Statistics
- Lines, bars and markers
- Images, contours and fields
- Shapes and collections
- Text, labels and annotations
- Pie and polar charts
- Style sheets
- Showcase
- Animation
- Axes Grid
- Axis Artist
- Event Handling
- Front Page
- Miscellaneous

51

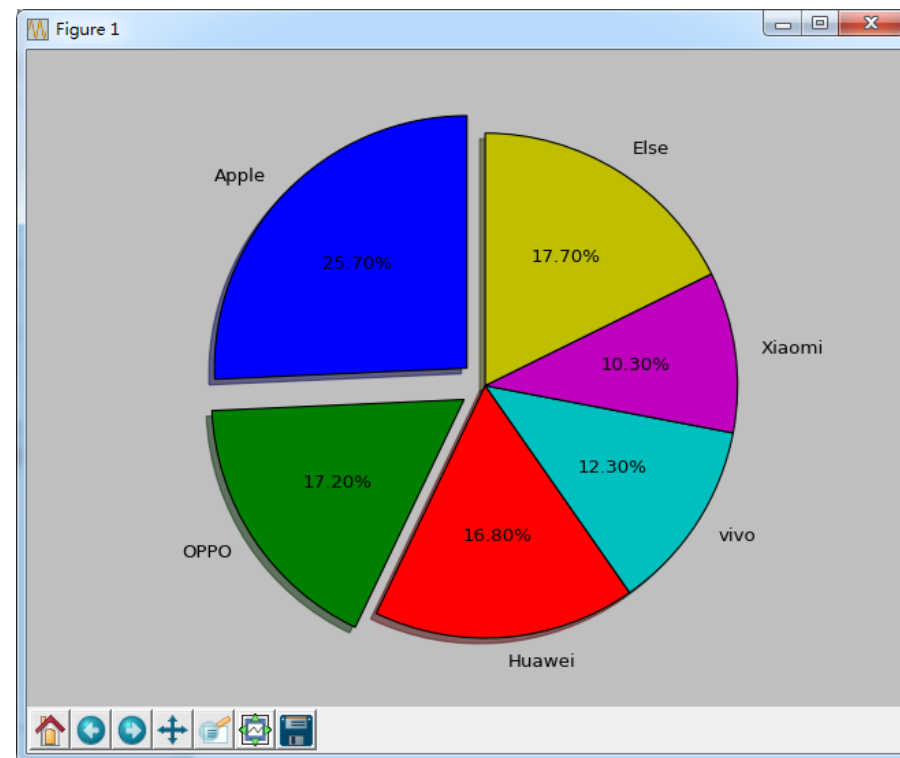




## 热身题

- 请根据matplotlib官网上的example，根据下表绘制一张饼图

Type	Percentage
Apple	25.7
OPPO	17.2
Huawei	16.8
vivo	12.3
Xiaomi	10.3
Else	17.7



数据来源：2018年第一季度  
中国手机市场份额调查

