模块三：业务分析与可视化

## 任务一：数据可视化（20分）

### 子任务一：基于 Echarts 的数据可视化分析（5分）

气候变化正在迅速地改变地球。随着全球气温不断升高、海平面上升、极端天气事件频繁发生，人们对于地球的未来更加担忧。为了更好地了解气候变化的趋势、预测未来天气趋势，我们创建了“天气数据库”，用于收集、记录、可视化展示来自全球各地的气象数据和天气预报信息。请在“index.html”文件中编写代码实现功能,数据文件名为“chengdu.js”:

1、文件内记录了四川省成都市2021年的天气数据，统计每个月的平均最高气温和每个月的平均最低气温，将统计得到的数据格式转换为Echarts所需的数据格式；

2、将第1步中构建的数据作为输入，通过Echarts绘制柱状图；

3、使用浏览器打开“index.html”文件，然后将渲染结果截图并粘贴到客户端桌面**【result/模块3答题报告.docx】**中对应的任务序号下。

### 子任务二：基于Python进行灾害损失可视化分析（5分）

现有一份关于灾害的数据集“disaster.csv”，列名包含年份、国家、灾害类型、灾害子类型、区域、灾害频次、总死亡人数、总经济损失（千美元）。绘图过程中如有使用中文字体的地方，请统一使用“SimSun”字体。

需要分析并绘制中国各年份死亡人数折线图和中国各年份总经济损失折线图，将两张图作为子图绘制于同一个画布中，画布大小为15\*10，子图布局为2行1列。将**两张折线图**截图并粘贴到客户端桌面**【result/模块3答题报告.docx】**中对应的任务序号下。

具体要求如下：

**（1）分析并绘制中国各年份死亡人数折线图：**

①使用matplotlib库绘图。

②计算中国各年的总死亡人数，单位为“万人”。

③将图形绘制于第一张子图中。

④图像标题为“死亡人数”,字体大小20。

⑤x轴标签为“年份”，字体大小16，y轴标签为“总死亡人数（万人）”，字体大小16。

⑥折线图颜色为“red”，折线图上每个点需要显示出来，点型为实心圆

**（2）分析并绘制中国各年份总经济损失折线图：**

①使用matplotlib库绘图。

②计算中国各年的总经济损失，单位为亿元，假设1美元=7元。

③将图形绘制于第二张子图中。

④图像标题为“经济损失”，字体大小20。

⑤x轴标签为“年份”，字体大小16，y轴标签为“总经济损失（亿元）”，字体大小16。

⑥折线图颜色为“violet”，折线图上每个点需要显示出来，点型为实心圆。

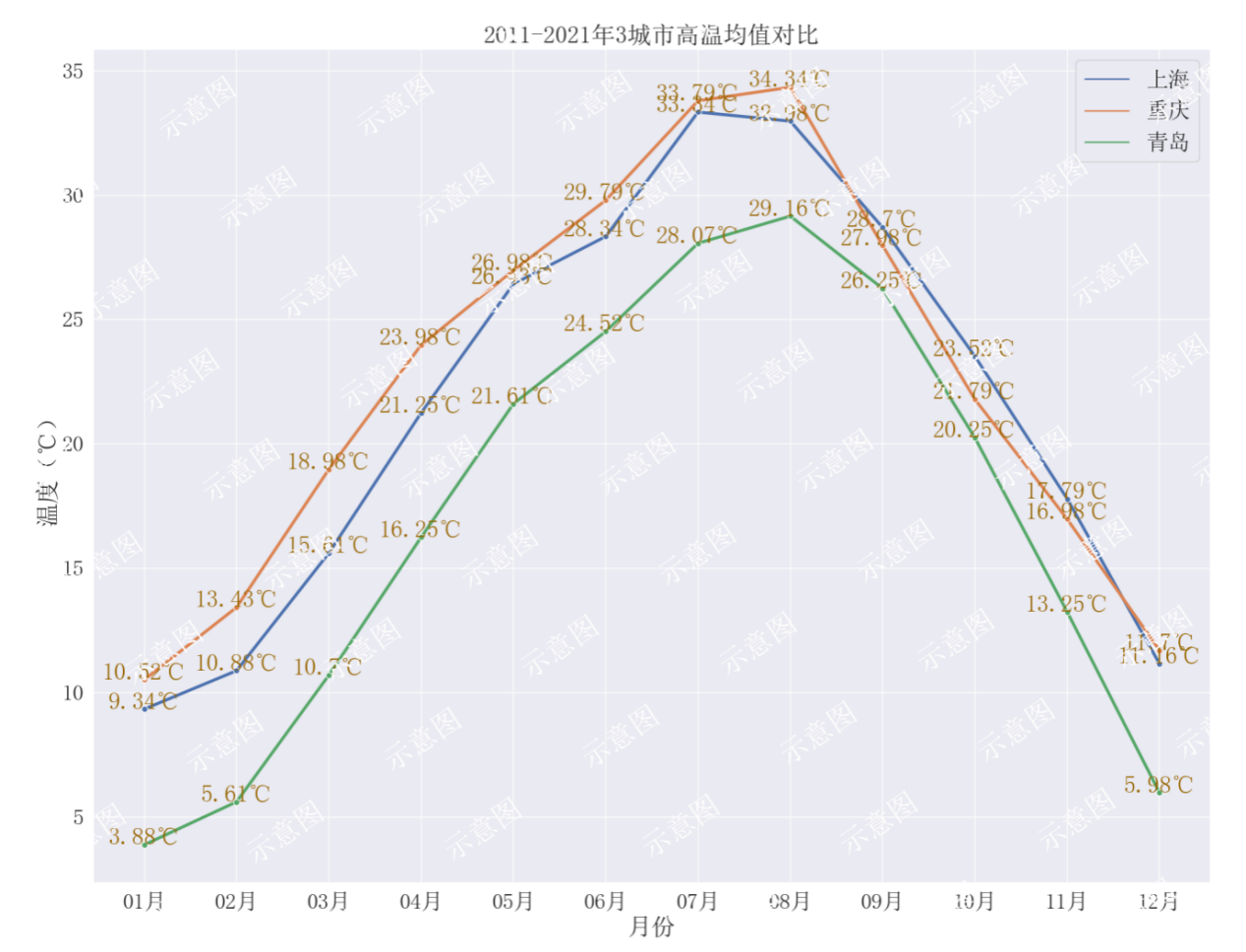
### 子任务三：基于Python进行城市全年温度对比分析（5分）

现有一份关于2011-2022年全国各城市的每月天气数据集，字段说明如下表：

|  |  |
| --- | --- |
| **列名** | **字段说明** |
| city | 城市 |
| month | 年-月 |
| avg\_high\_tem | 平均高温 |
| avg\_low\_tem | 平均低温 |
| extreme\_high\_tem | 极端高温 |
| extreme\_low\_tem | 极端低温 |
| avg\_air\_quality | 平均空气质量指数 |
| best\_air | 最好空气指数 |
| best\_air\_date | 最好空气指数日期 |
| worst\_air | 最差空气指数 |
| worst\_air\_date | 最差空气指数日期 |

请编写代码，从数据集“clean\_month.csv”中取出重庆、上海、青岛三市在2011-2021年的数据；计算出三市在这10年中所有月份平均高温的平均值，并使用该指标绘制出三个城市一年中的高温变化图来分析三市的气温情况，具体绘图要求如下：

1. 使用Seaborn绘制以上指标的折线图，三个城市的折线图绘制在同一张图中，主题设置为“darkgrid”，字体为“SimSun”，字体缩放因子设置为2；
2. 折线图中需要添加数据标记，形状为圆点，线宽为3，并通过对城市分组给三条曲线设置不同的颜色，采用分组后Seaborn的默认配色；
3. 折线图需显示数据标签，颜色为“#996600”，数据标签中需要带有单位（℃）并且保留两位小数；
4. 图像标题为“2011-2021年3城市高温均值对比”，横轴标签为“月份”，纵轴标签为“温度（℃）”；
5. 横轴刻度为月份，刻度标签为“01月 02月 ... 12月”;
6. 图像右上角显示图例；
7. 绘图过程中如有使用中文字体的地方，请统一使用“SimSun”字体。
8. 绘制完成后将图片粘贴到客户端桌面**【result/模块3答题报告.docx】**中对应的任务序号下。



结果参考示意图（图中数据请以实际计算结果为准）

### 子任务四：基于Python实现历史最高温城市排名分析（5分）

现有一份关于2011-2022年全国各城市的每日天气数据集，字段说明如下表：

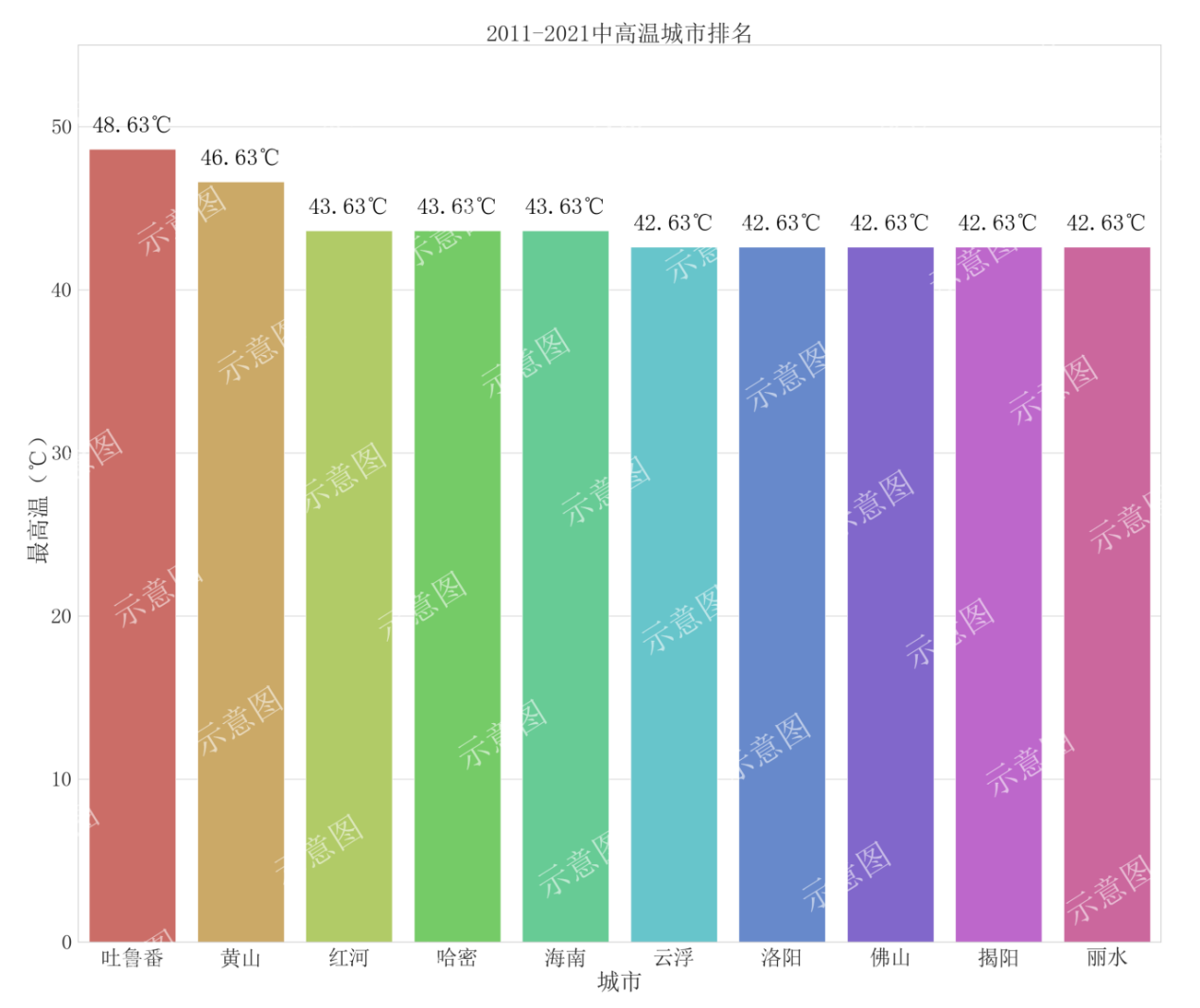
|  |  |
| --- | --- |
| **列名** | **字段说明** |
| city | 城市 |
| hightest\_tem | 最高气温 |
| lowest\_tem | 最低气温 |
| weather | 天气 |
| date | 日期 |
| wind\_direction | 风向 |
| wind\_level | 风力等级 |
| month\_day | 月-日 |
| weekday | 星期 |

请编写代码实现功能，数据集为“clean\_day.csv”。

绘图过程中如有使用中文字体的地方，请统一使用“SimSun”字体。

从数据集中取出2011-2021年的数据，计算出每个城市在这10年中出现的最高温度，然后将该指标排名前10的城市数据取出并绘制出这些城市最高温统计图进行分析，具体绘图要求如下：

1. 使用Seaborn绘制出以上指标的柱状图，主题设置为“whitegrid”，字体为“SimSun”，字体缩放因子设置为3；
2. 柱状图颜色设置为Seaborn中调色板“hls”的默认颜色；
3. 柱状图上需要显示数据标签，数据标签中需要带有单位（℃），颜色为黑色；
4. 设置图像标题为“2011-2021中高温城市排名”；
5. 横轴标签为“城市”，纵轴标签为“最高温（℃）”;
6. 横轴的刻度标签为各城市的名字；
7. 纵轴的刻度范围是（0,55）。
8. 绘制完成后将图片粘贴到客户端桌面**【result/模块3答题报告.docx】**中对应的任务序号下。

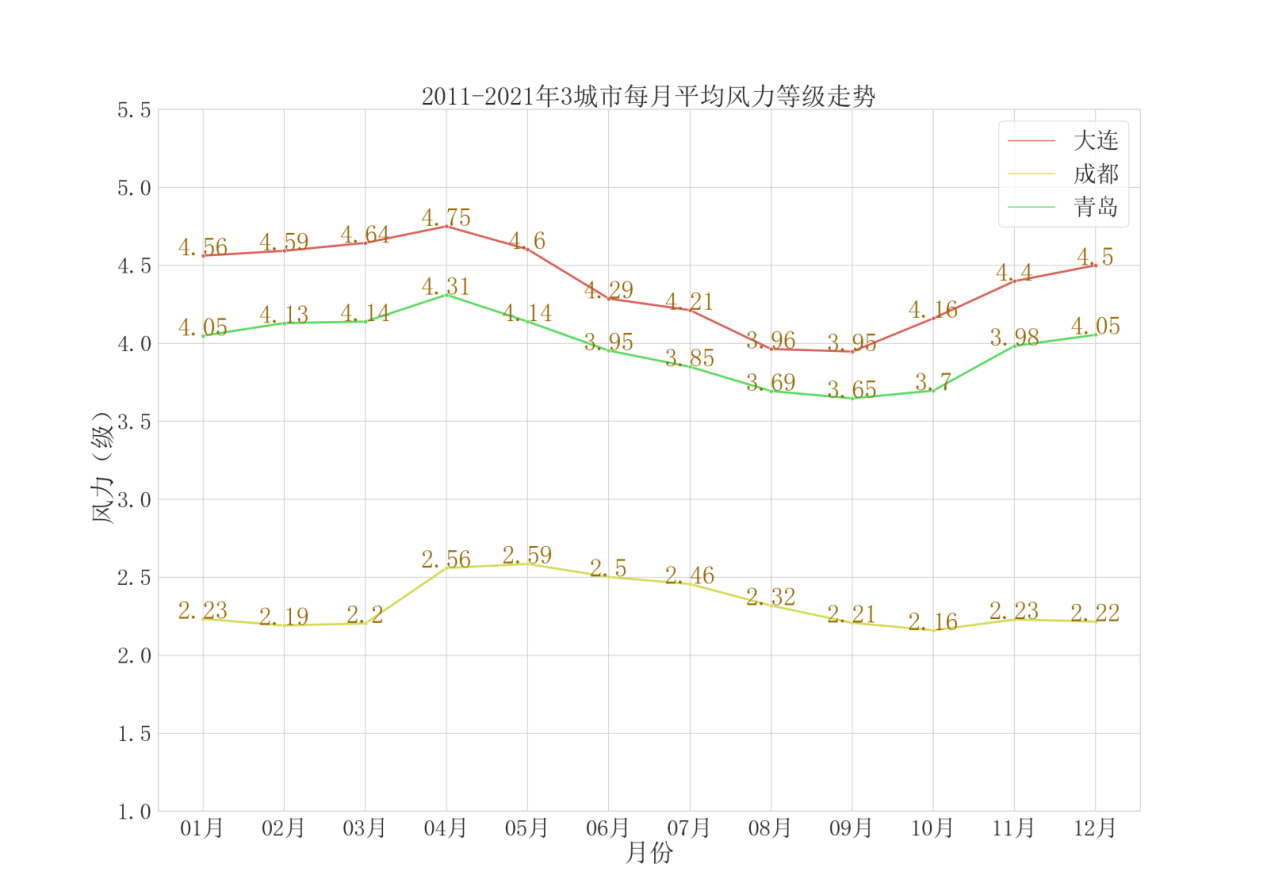


结果参考示意图（图中数据请以实际计算结果为准）

## 任务二：业务分析（10分）

现有一份关于天气的可视化分析结果，请对结果进行业务分析，并给出可解释型结论。

下图是3个城市10年来每月平均风力等级走势图，请对该图以及图表中的数据进行分析并给出合理解释。



风力等级走势图

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模块三：  业务分析与可视化 | 任务一：  数据可视化 | 使用Python可视化库或Web前端框架对数据进行呈现，包括但不限于柱状图、折线图、玫瑰图、气泡图、饼状图、条形图、雷达图、散点图等效果。 |  | 20% |
| 任务二：  业务分析 | 报表分析，对大数据项目的业务场景和数据进行分析，撰写报告。 | 10% |
| 职业素养 | | 团队分工明确合理、操作规范、文明竞赛。 | 5% |

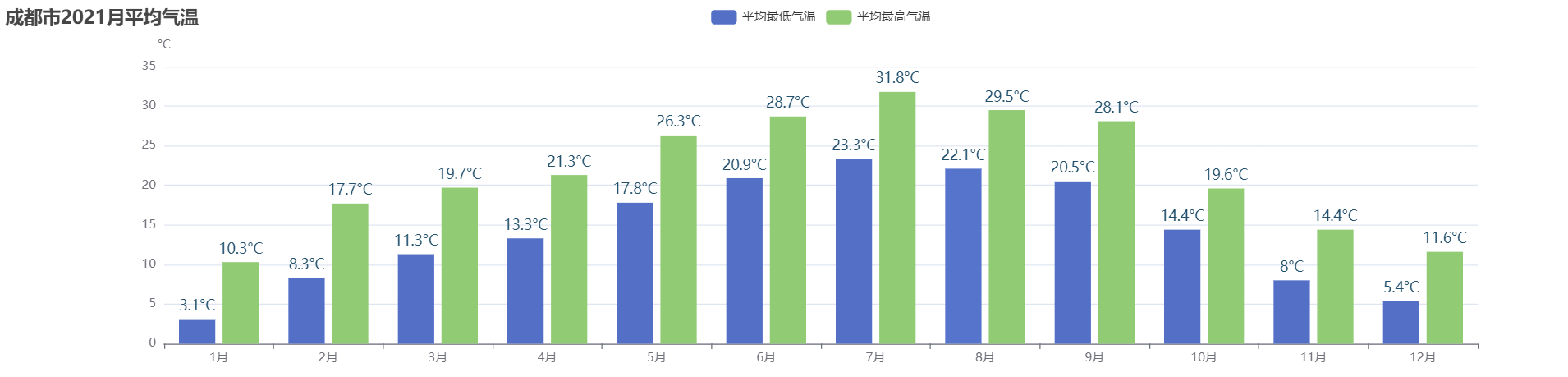
# **大数据应用与服务赛项-标准答案（A卷）**

## 模块三：业务数据分析与可视化（30分）

### 任务一：数据可视化（20分）

#### 子任务一：基于 Echarts 的数据可视化分析（5分）

Echarts柱状图渲染完成后截图粘贴到此处。结果如下图：5分

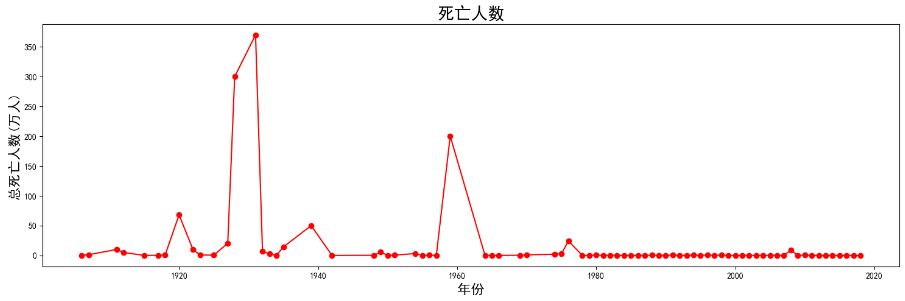


得分点：

1. 显示为柱状图得1分；
2. 统计“平均最高气温”每个数据标签正确得2分；
3. 统计“平均最低气温”每个数据标签正确得2分；

#### 子任务二：基于Pythont进行灾害损失可视化分析（5分）

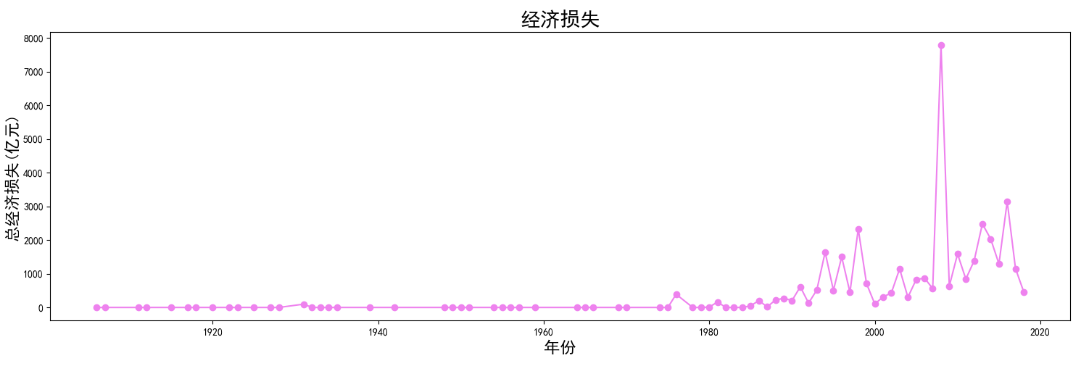
1）中国各年份死亡人数折线图截图（2.5分）



得分点：

1. 标题为“死亡人数”，x轴标签为“年份”,y轴标签为“总死亡人数（万人）”，得1分。
2. 折线图颜色、趋势一致得1分。
3. 折线图上的点是实心圆得0.5分。

2）中国各年份总经济损失折线图截图（2.5分）

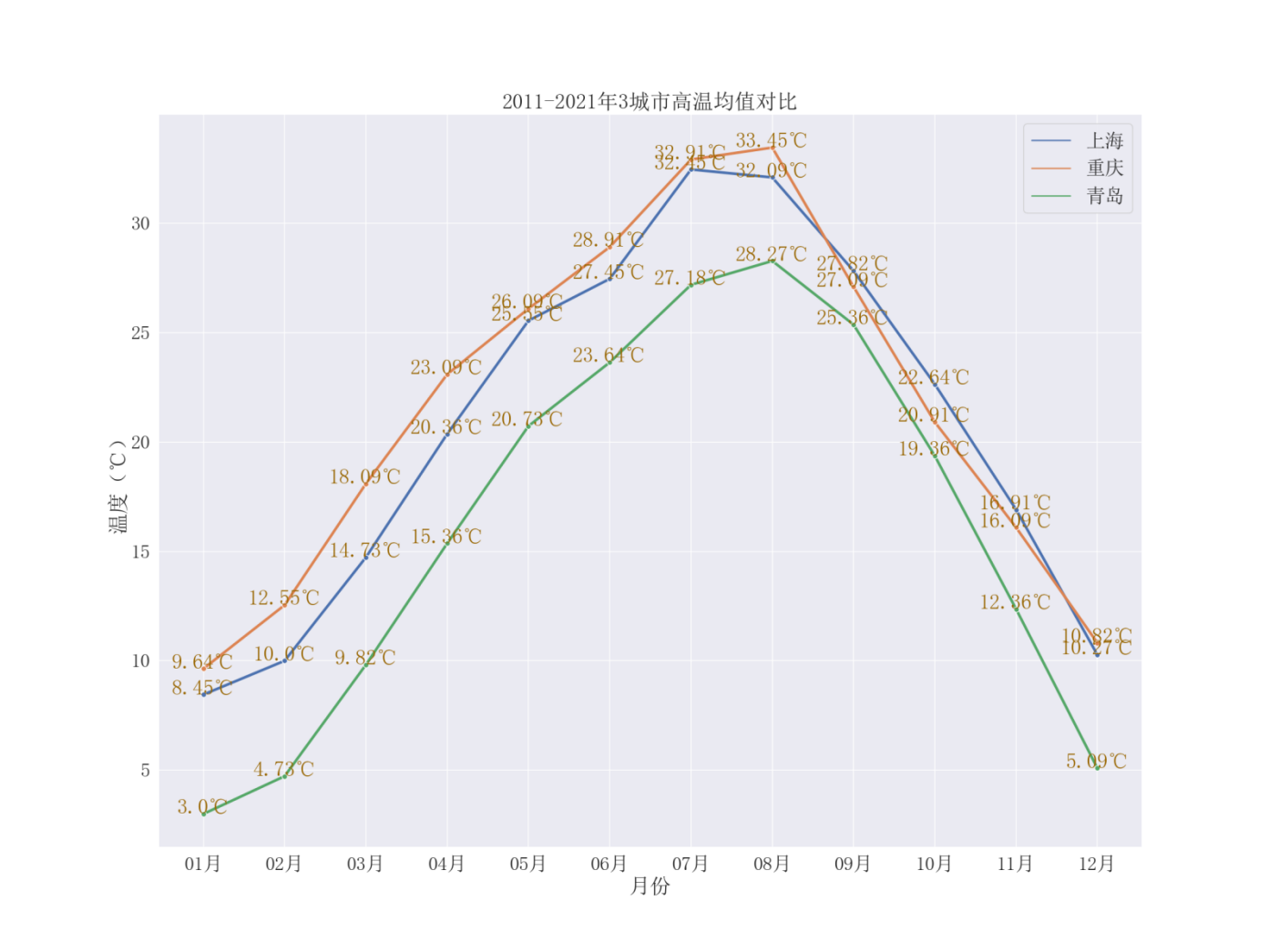


得分点：

1. 标题为“经济损失”，x轴标签为“年份”,y轴标签为“总经济损失（亿元）”得1分。
2. 折线图颜色、趋势一致得1分。
3. 折线图上的点是实心圆得0.5分。

#### 子任务三：基于Python进行城市全年温度对比分析（5分）

结果如下图：5分

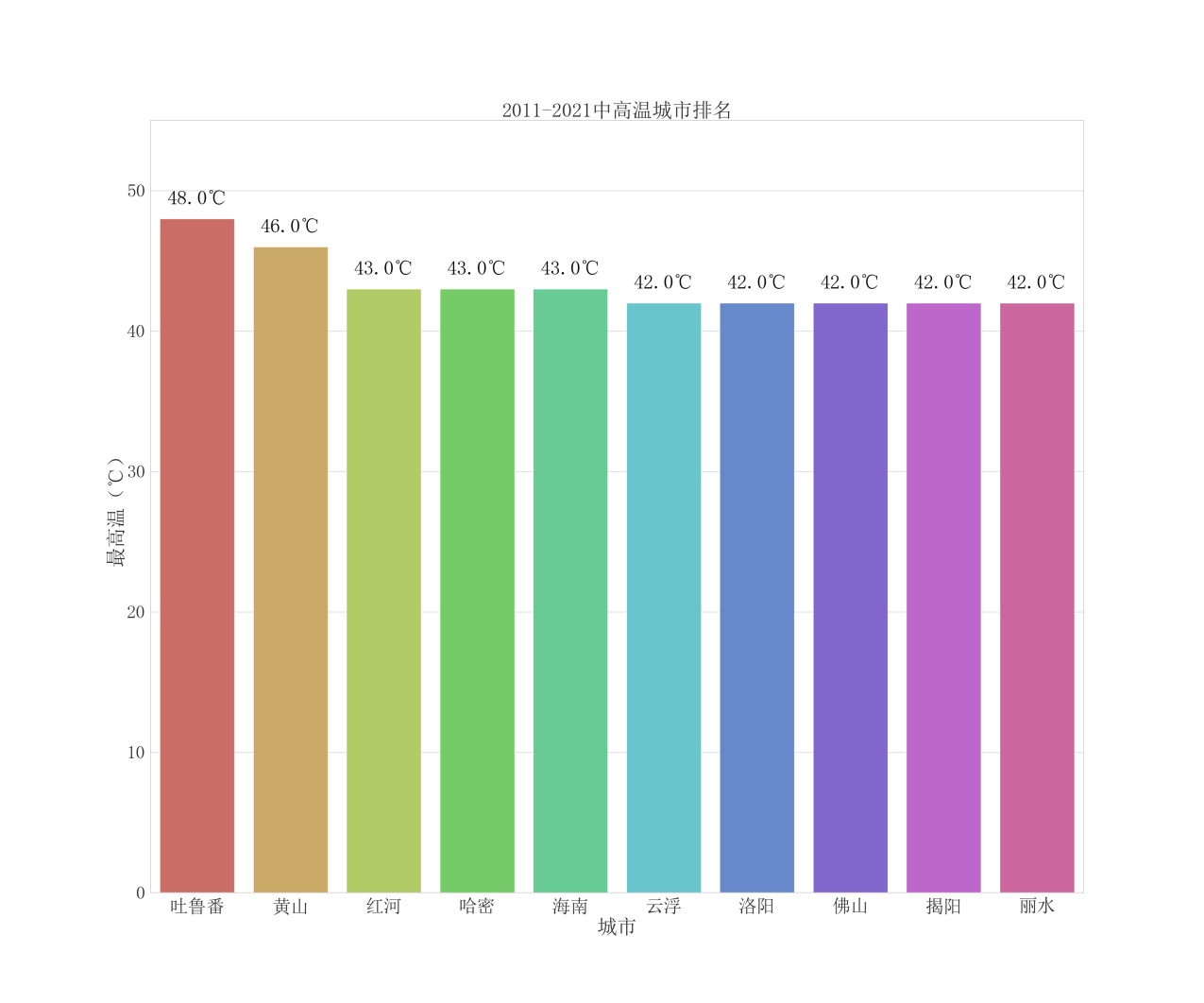


得分点：

1. 标题一致得0.5分；
2. 折线图显示“青岛”、“上海”、“重庆”三种不同颜色得1.5分；
3. 横轴标题显示“月份1-12月”得0.5分；
4. 纵轴标题显示“温度°C”得0.5分；
5. 数据一致得1分；
6. 图例背景颜色显示一致得1分

#### 子任务四：基于Python实现历史最高温城市排名分析（5分）

结果如下图：5分



得分点：

1. 图像标题显示“2011-2021中高温城市排名”得0.5分；
2. 每个数据标签单位带有°C得1分；
3. 柱状图颜色一致得0.5分；
4. 横轴标题显示“城市”且不重复得1分；
5. 纵轴标题显示“最高温（°C）”且刻度区间在0-55以内得1分；
6. 每个数据标签一致得0.5分；
7. 主题颜色一致得0.5分；

### 任务二：业务分析（10分）

从图可知，成都的风力明显比青岛和大连小，是因为成都处理四川盆地内，

周围的高山抵挡了大部分的气流。大连和青岛的风力走势一致，且呈现冬

季风大、夏季风小的规律。因为夏季风力主要来自东南海洋，而冬季风力

主要来自西伯利亚。（此题为主观答案，描述合理即可）

得分点：

1. 出现成都、大连、青岛三个城市描述各得2分；共6分；
2. 描述成都风力走势小得2分；
3. 描述大连和青岛风力走势一致得2分；

## 职业素养（5分）

团队分工明确合理、操作规范、文明竞赛。