

# 数据可视化报告

南开第一深情

陈星烨 2212518 周末 2211349 高鸿浩 2213411

## 1 作品简介

### 1.1 作品主题与问题场景

地震，是地球内部发生的急剧破裂产生的震波，在一定范围内引起地面振动的现象。地震是地球上常见的自然灾害之一，对人类社会造成了极大的破坏和伤害。地震有主要几个参数，包括地震发生时刻、震中位置、震源深度、震级、最大震度等。研究这些参数之间的关系，可以帮助我们更好地了解地震的规律和特点，从而提高地震预测和防灾减灾工作的效果。

### 1.2 基本概念

#### 1.2.1 基本参数

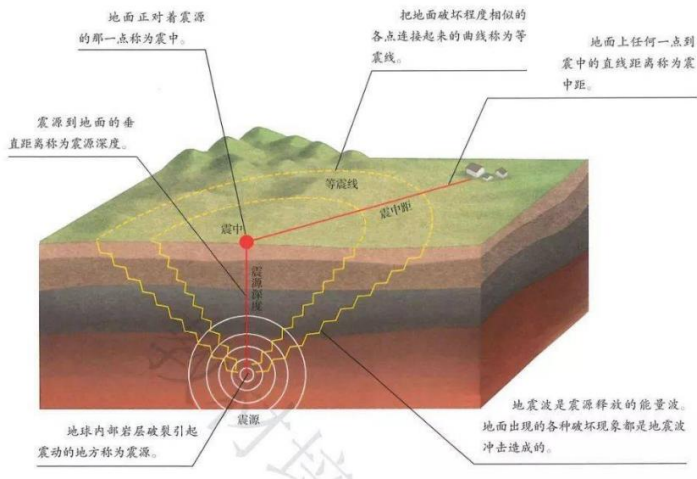


图 1 地震示意图

震源:地球内部发生地震的地方叫震源，也称震源区或震源域。它是一个区域，但研究地震时，常把它看成一个点。

震源深度:如果把震源看成一个点，那么这个点到地面的垂直距离就称为震源深度。

震中:指地面上正对着震源的那一点。

震中距:在地面上，从震中到任一点的距离叫做震中距。

#### 1.2.2 烈度或震度

烈度，在日本称作震度，是用来描述地震时地面受到的影响或者破坏程度的量，可以用 I (Intensity) 来表示。在中国分为 I 度至 XII 度一共十二个等级，在日本分为 1、

2、3、4、5 弱 (5-)、5 强 (5+)、6 弱 (6-)、6 强 (6+)、7 一共九个等级。日本气象厅规定了烈度数值的计算方法，即 JMA 计测烈度算法，并规定

$$I = 2 \log(V_a) + 0.94$$

其中 $I$ 为计测烈度值， $V_a$ 为对三分向加速度时程带通滤波后再组合而成的合成时程中持时大于等于 0.3s 时的幅值。根据 $I$ 的不同值，对震度的划分如下表：

表 计测震度与震度转化表

计测震度 $I$	震度
$I < 0.5$	0
$0.5 \leq I < 1.5$	1
$1.5 \leq I < 2.5$	2
$2.5 \leq I < 3.5$	3
$3.5 \leq I < 4.5$	4
$4.5 \leq I < 5.0$	5-
$5.0 \leq I < 5.5$	5+
$5.5 \leq I < 6.0$	6-
$6.0 \leq I < 6.5$	6+
$I \geq 6.5$	7

中国的烈度制与日本的震度制之间可以互相转换，例如，山东平原地震极震区最大烈度为 VIII 度，换算成震度约为 5+；甘肃积石山地震震中最大烈度 9.5 度，换算成震度约为 6+。

1.2.3 震级

震级，是表示震源释放能量多少的级别，一般记作  $M$  (Magnitude)。一次地震只有一个震级。震源所释放的能量越多，震级就越大。震级可以有矩震级 ( $M_w$ ) 与里氏震级 ( $M_L$ )、面波震级 ( $M_s$ ) 及体波震级 ( $M_b$ ) 等表示方法。目前国际通用的是里氏震级划分方法，定义在距离震中 100 千米处的观测点地震仪记录到的最大水平位移为 1 微米的地震作为 0 级地震。2008 年, Yamamoto 等人根据 JMA 计测烈度算法提出了烈度震级 ( $M_l$ )，其计算方法如下：

$$M_l = \frac{I}{2} + \log(r) + \frac{\pi f T}{2.3Q} + b - \log C_i$$

其中， $r$ 为震中距， $f$ 为卓越频率， $T$ 为 S 波走时， $Q$ 为品质因子， $b$ 为系统校正值， $C_i$ 为台站校正值， $I$ 为由台站记录计算得到的 JMA 计测烈度值。

1.3 应用价值

震源深度、震中距与烈度的关系反映了地震发生的位置，对地震的强烈程度、破坏范围和地震波传播等有重要影响。通过对震源深度、震中距与烈度的关系的研究，可以深入了解地震的发生机制和地壳运动规律，为地震预测、减少地震危害具有重要意义。

通过对震中纬度、经度和地震发生时刻的研究，了解不同地区地震活动的时空分布规律。这有助于预测地震的可能发生时间，提高地震预警系统的准确性，确定地震活动带和地震危险区域，指导城市规划和土地利用，减少地震灾害风险。

## 1.4 分析目标

- 1、分析地震各个参数之间的关系，以及地震各个参数与地面破坏程度（烈度）之间的关系；
- 2、分析各个地震之间的时空关联性，为地震的防御措施提供理论依据；
- 3、分析几个大型地震的特点，并为大型地震的防御总结经验。

# 2 数据介绍

## 2.1 数据来源与基本内容

### 2.1.1 数据来源

地震数据主要来源于日本气象厅官网：<https://www.jma.go.jp>；日本地图的 geojson 数据来源于老师的提供。

### 2.1.2 数据的基本内容

从网站上得到的数据被整理成了三个数据集。

第一个数据集包括了从 2000 年 1 月 1 日至 2023 年 11 月 23 日日本本土及附近海域发生地震的详细信息，包括地震发生日、地震发生时刻、震中地名、纬度、经度、震源深度、震级、最大烈度，一共 8 个维度。

第二个数据集包含了日本自 2000 年以来经历过 6 次的大地震的所有观测点以及其实测烈度。

第三个数据集包含了日本本土所有观测点的名字及其经纬度数据。

## 2.2 数据预处理

结合可视分析需求，考虑到庞大的数据处理会严重影响作品的渲染效率，因此本作品对数据进行了不同方面的预处理。最终用于可视分析系统输入的数据均被转换为 csv 格式，具体的数据预处理过程如下：

### 2.2.1 数据清洗

将数据中带有“不明”的无用数据全部删除，利用 python 程序将有用的几栏数据单独制成一个数据集

### 2.2.2 数据处理

根据图的不同需求，将数据处理成可以被读取的形式。例如在地震数量条形图中，会利用 python 程序提前将不同烈度地震的数量统计出来，并新制成一个数据集。有关方位颜色映射的数据处理方式较为复杂，具体见下文分析。

### 2.2.3 方位颜色映射

方位颜色映射的处理步骤较为复杂，经过地名哈希对应、方位角计算和正规化、颜色方位角对应三个步骤实现。

地名哈希对应，是指将上述第三个数据集作为一个哈希表，并在其中查找观测点所对应经纬度数据，再将其返回给程序。

方位角指的是以震源为参考系，观测点所在的方位角 $\varphi$ ，范围为 $(-\pi, \pi]$ ，其示意图如下：

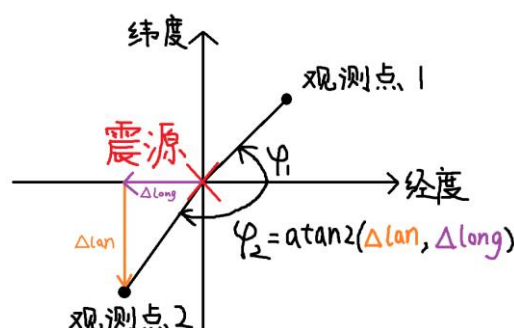


图 2 方位角示意图

在计算方位角时，由于观测范围只在日本附近，所以可以把地球的球面近似成平面，用 atan2 函数来近似其实际方位角，atan2 的函数表达式如下：

$$atan2(y, x) = \begin{cases} \arctan\left(\frac{y}{x}\right), & x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi, & y \geq 0, x < 0 \\ \arctan\left(\frac{y}{x}\right) - \pi, & y < 0, x < 0 \\ +\frac{\pi}{2}, & y > 0, x = 0 \\ -\frac{\pi}{2}, & y < 0, x = 0 \\ undefined, & y = 0, x = 0 \end{cases}$$

其中的 $y$ 在实际情况下表示的是纬度的差值， $x$ 在实际情况下表示的是经度的差值。  
设分别用 $\Delta long$ 和 $\Delta lan$ 表示经纬度的差值，则有：

$$\Delta long = long - Long$$

$$\Delta lan = lan - Lan$$

其中 $Long$ 、 $Lan$ 分别表示震源的经度和纬度， $long$ 和 $lan$ 分别表示观测点的经度和纬度，则其相对于震源的方位角为：

$$\varphi = atan2(\Delta long, \Delta lan)$$

为了计算方便，我们将计算出的方位角统一到了 $(-3, 3]$ 的区间内，记为 $\varphi_n$ ，有如下公式：

$$\varphi_n = \frac{3}{\pi} \varphi$$

方位对应我们采用了颜色对应的方式，利用了 $(R, G, B)$ 循环渐变的方法，设计出了如下对应函数：

$$(R(\varphi_n), G(\varphi_n), B(\varphi_n)) = \begin{cases} (255, (3 + \varphi_n) \times 255, 0), & -3 \leq \varphi_n < -2 \\ (-(1 + \varphi_n) \times 255, 255, 0), & -2 \leq \varphi_n < -1 \\ (0, 255, (1 + \varphi_n) \times 255), & -1 \leq \varphi_n < 0 \\ (0, (1 - \varphi_n) \times 255, 255), & 0 \leq \varphi_n < 1 \\ ((\varphi_n - 1) \times 255, 0, 255), & 1 \leq \varphi_n < 2 \\ (255, 0, (3 - \varphi_n) \times 255), & 2 \leq \varphi_n \leq 3 \\ (0, 0, 0), & \varphi_n = undefined \end{cases}$$

函数图像与颜色效果如下图所示：

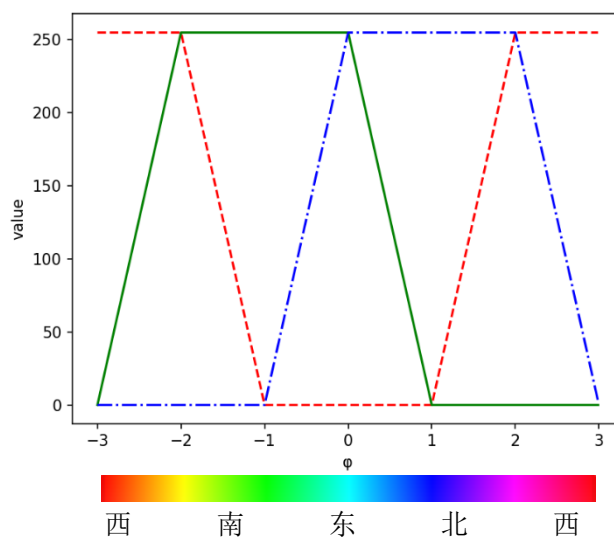


图 3 方位角颜色映射图像

### 3 可视化与交互设计

#### 3.1 总体介绍

本作品以多个图的方式呈现。

第一张图把从 2000 到 2023 年的所有地震数据作为输入，画出了显示不同地震的震级、最大烈度和震源深度之间关系的气泡图。图片支持用户选择不同烈度的点进行显示，以便比较不同烈度点的分布情况。

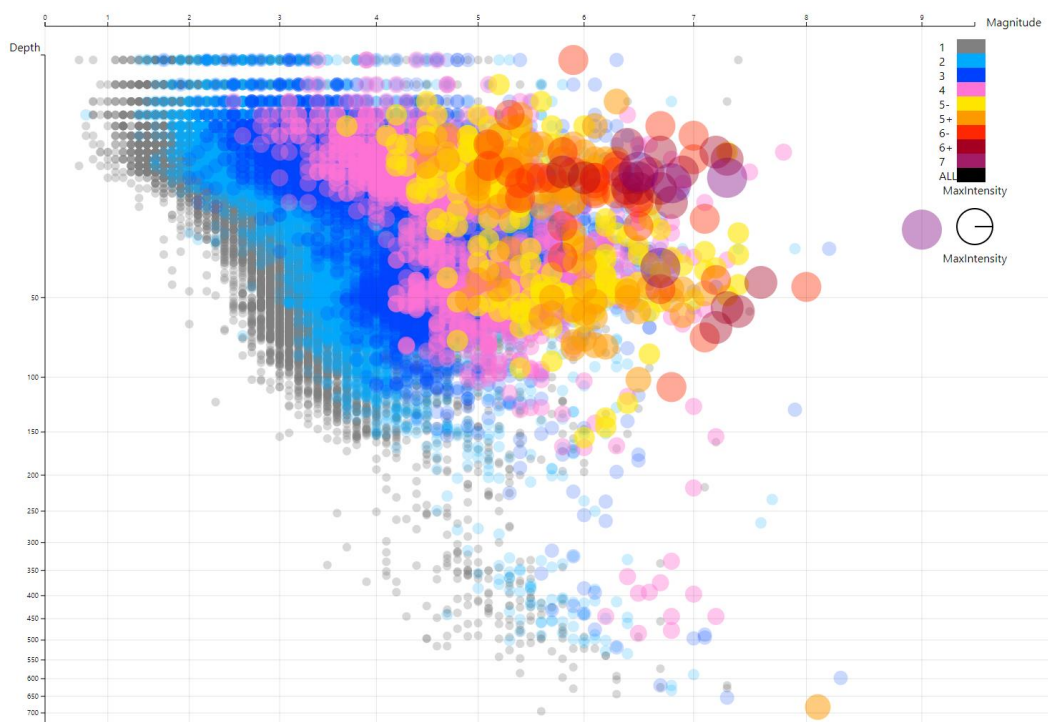
第二张图将震源数据按照不同烈度统计数量，绘制出了显示不同年份、不同烈度的地震数量的堆叠条形图。这张图支持用户选择同一烈度的地震的统计数量进行显示。

第三张图在日本地图上播放从 2000 年至 2023 年地震按时间发生的动画，直观的展示了地震发生的时间分布情况和空间分布情况。

第四张图包含两个子图，其中左图显示的是这一次地震的所有站点实际测得的最大烈度数据，右图展示的是这一次地震中最大烈度、震中距以及震中方位之间的关系。

作品设计实现了多种视图，辅以合理的布局，同时支持灵活直观的交互设计，可以展示对应日期，不同地区的地震发生的位置、震级等，方便用户开展对地震的时空分布规律、地震各个参数之间的关系等分析任务。

#### 3.2 震源气泡图



### 3.2.1 视图设计

地震气泡图采用坐标系展示 2000 至 2023 年所有地震的震源深度、最大烈度和震级的关系。纵轴表示深度，横轴表示震级，颜色和气泡的大小表示最大烈度。因数据最大烈度主要集中在 3 到 6+ 之间，数据 Depth 主要集中于 0 至 100km 之间，所以坐标轴采用指数和对数坐标系，可以更加清晰地表现出两者的关系。

### 3.2.2 交互设计

用户通过切换最大烈度，可以实现观察单个最大烈度与其他数据的关系，也可以实现观察多个最大烈度数据之间的关系。

### 3.2.3 代码实现

用 svg 标签将每一个气泡画出来，再用 on(“click”，...)函数，实现每当点击所选的烈度的图例时，可以使其他烈度的气泡隐藏，只显示当前所选烈度的气泡。此时再点击其他烈度的图例时，会在原气泡不变的基础上，追加显示所选烈度的气泡。当点击 ALL 的图例时，会恢复原来的图像。

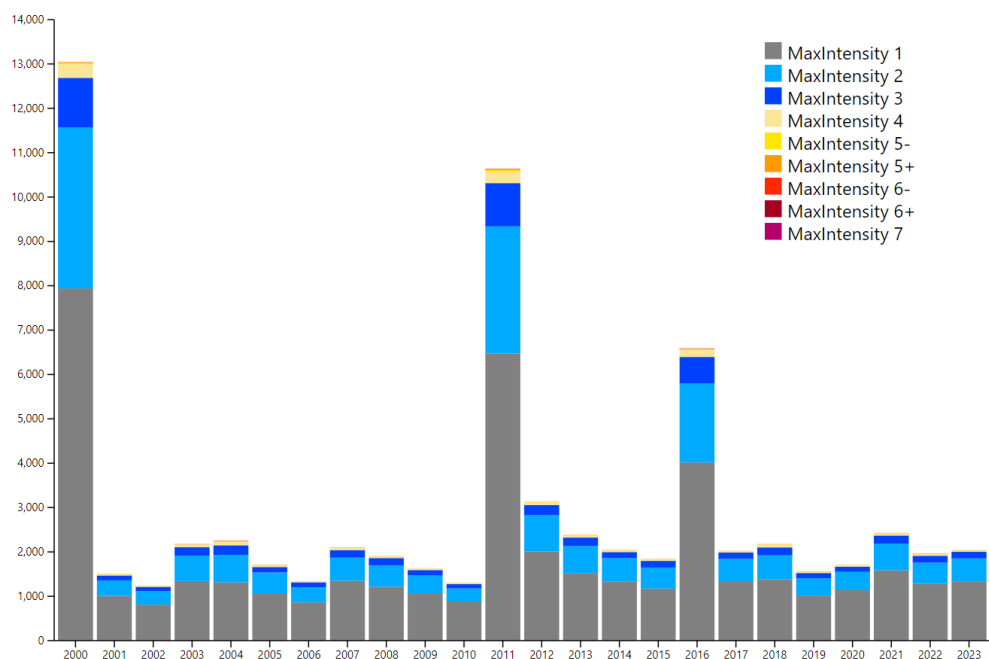


### 3.3.4 图像分析

结合图例，可以通过不同的观察方式得到以下的结论：

- 1、整体上，随着深度增加，地震造成的破坏强度在下降，即最大烈度在下降；
- 2、在震级较低时，震源深度较深的地震在地面上基本无感，因此图像左下方有一大片空白；
- 3、整体上，随着震级增加，地震造成的破坏在增强，即最大烈度在升高；
- 4、整体上，震级越大，地震发生的次数越少，即气泡的数量在逐渐减少；

### 3.3 地震数量条形图



#### 3.3.1 视图设计

地震数量条形图采用坐标系展示 2000 至 2023 年之间每年地震总数和最大烈度的关系。其中纵轴表示地震总数，横轴表示年份，颜色的大小表示最大烈度。

#### 3.3.2 交互设计

用户通过切换最大烈度可以观察单个最大烈度地震数量的年份分布。由于烈度较大的地震发生次数较少，所以实现了可以选择不同烈度数据的交互。

#### 3.3.3 代码实现

根据最大烈度划分不同的 key，并以数量的多少进行堆叠；在实现选择烈度单独呈现上，找到这一烈度是最大数量，并以此重新调整 y 轴比例尺，绘制图像。

#### 3.3.4 图像分析

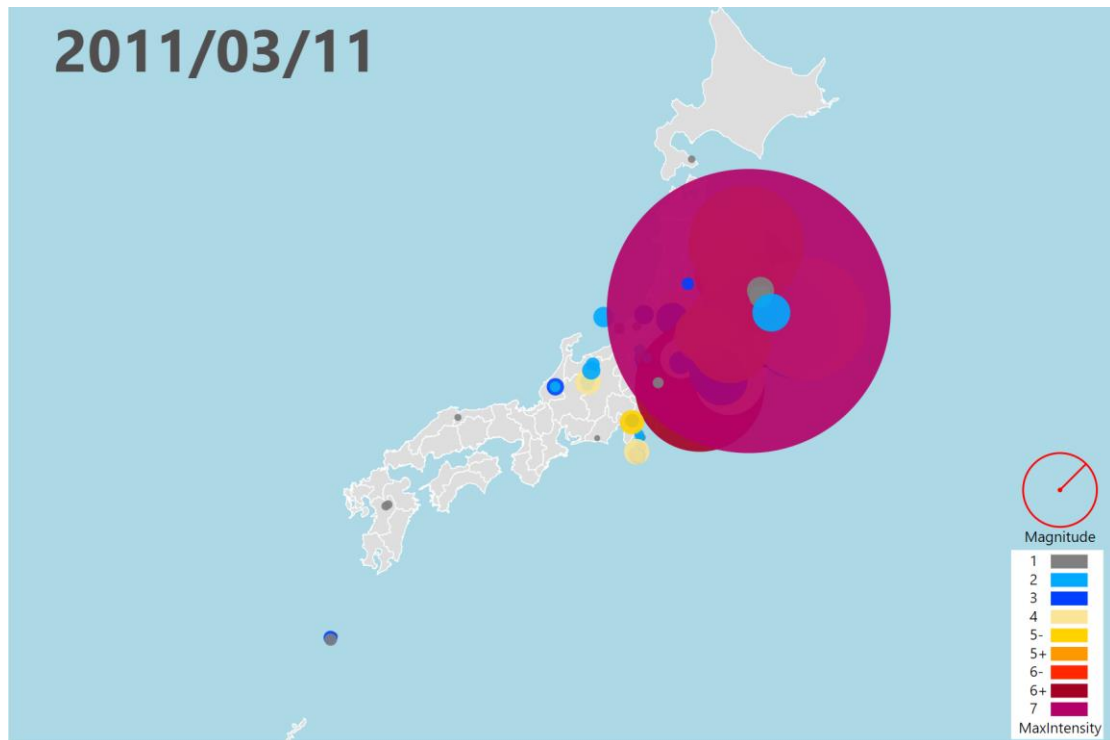
在图像中，2000 年、2011 年和 2016 年的地震数量明显偏多，这是因为在这些发生了一些大的地质事件。2000 年 6 月至 8 月，三宅岛火山喷发，火山爆发伴随的有感地震近万次；2011 年 3 月 11 日，在三陆冲发生了里氏 9.0 级的东日本大地震，其余震过程持续数年，所以 2011 年及其随后几年内地震发生的数量整体有增加，但是随着时间推移，余震发生次数降低，所以地震数量整体呈下降趋势；2016 年 4 月 14 日和 4 月 16 日，在熊本发生接连发生了两次强震，里氏震级分别为 M6.7 和 M7.3，这两次地震也引发了数千次余震。

除去这几年以外，2004 年、2018 年、2021 年、2022 年、2023 年等地震发生次数也较多，这是因为这几年也发生过几次强震过程，这一结论可以切换不同烈度显示观察得到。

而且观察烈度为 6-、6+、7 的强震数量，会发现日本每年都会有至少一次强震过程。

### 3.4 地震时间动画图





#### 3.4.1 视图设计

地震时间动画图以日本地图为背景，圆圈的中心位置表示地震的震源位置，圆圈的大小表示震级的大小，圆圈的颜色表示地震的最大烈度，图例如图右下角所示。

#### 3.4.2 交互设计

有三个交互按钮，`restart` 按钮可以重制动画并停在第一帧，`stop` 按钮可以暂停动画播放，`continue` 按钮可以继续播放。注意按完 `restart` 后要再按一次 `continue` 按钮。

#### 3.4.3 代码实现

首先将数据整理到 `sequential` 数组中，并将数组元素按照日期从早到晚排序。这个数组中的每一个元素是一个小数组，每个小数组存放了同一天中所有的地震信息。播放动画是只需要利用 `for` 循环控制输入每一天的数据即可。

初始化函数 `renderinit` 中初始化了地图和第一帧的画面，切换帧画面函数 `renderupdate` 中实现了切换上下帧的功能，具体实现是利用 `d3` 提供的 `transition` 中的 `duration` 函数控制动画的播放速度，每经过一帧，将上一针添加的 `id` 为 `earthquake` 的 `g` 标签移除，然后根据下一帧的数据添加新的 `id` 为 `earthquake` 的 `g` 标签。利用 `ansync` 和 `await` 确保每一帧加载完之后再加载下一帧。

动画函数 `anime` 中利用 `for` 循环将每一帧数据输入 `renderupdate` 函数中，实现动画。

按钮 `restart`、`stop`、和 `continue` 通过控制全局变量 `frame`（当前播放帧）、`temp_start`（动画临时开始帧）和 `temp_end`（动画临时结束帧），进而控制 `for` 循环结束条件“`frame < temp_end`”，实现动画的重制、暂停和继续功能。

#### 3.4.4 图像分析

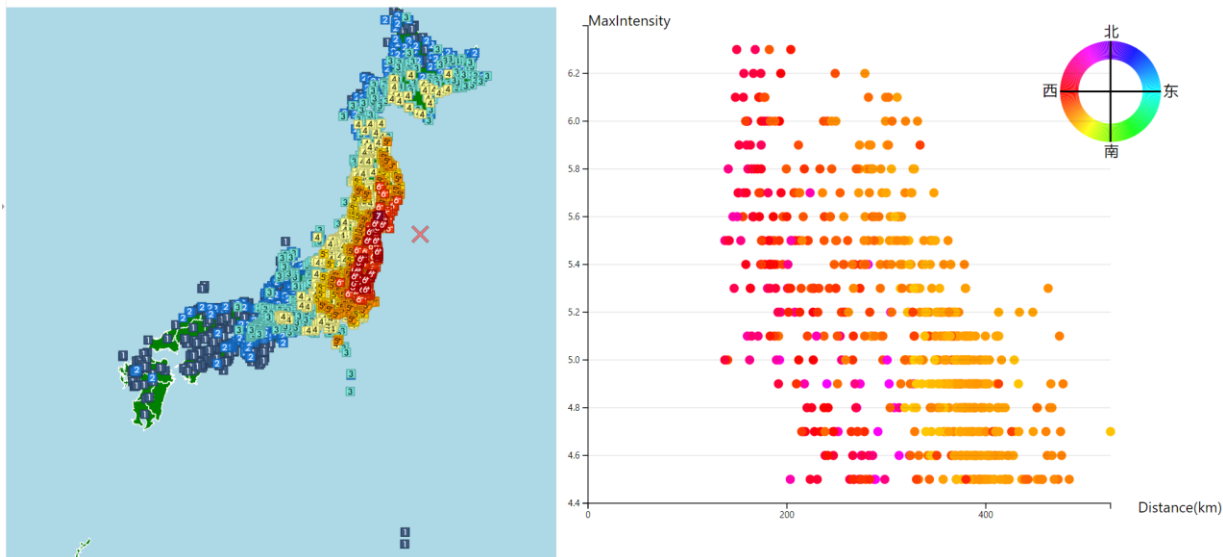
从总体上看，日本全境都遭受着地震威胁。实际上，日本地处欧亚板块、太平洋板块与菲律宾板块的交界处，全国遍布断层，地震易发，平均每天都会发生三四次地震。这些地震多为小地震，六级以上的大地震每年也至少会发生一次。

尽管如此，地震发生的位置还是有几个集中的位置：冲绳群岛，九州中西部及其近海地区，横贯九州东部、四国、本州东南部的南海断裂带，本州中国地方中东部和近畿地方，长野县、岐阜县、新潟县，石川县能登半岛，伊纪群岛和小笠原群岛，本州岛东邻近海，北海道四周沿海等。这些都是断层非常活跃的地方，也经常发生大地震。

还能从图中观察到地震与地震之间的时间关联性。例如主震和余震的关系，前震和主震的关系以及大地震与诱发地震的关系（参照东日本大地震中震源不在本州岛东部的大地震）。还能观察到一些有意思的地震活动过程，例如火山爆发过程中地震活动过程（2000 年 6 月至 8 月三宅岛的火山爆发过程）以及群震的活动过程（参照吐喀喇列岛 2023 年 5 月和 9 月的地震过程）。

### 3.5 烈度数据分析对比图

发生于2011年3月11日的东日本大地震



#### 3.5.1 视图设计

烈度数据分析对比图采用两个图像对比的方式，综合分析震中距、最大烈度和观测点方位的关系。

#### 3.5.2 代码实现

左图导入 `geojson` 绘制出地图，根据观测点的经纬度，并通过观测点最大震度加载相应的震度贴图和震源位置，最后通过 `zoom` 实现地图的缩放功能。

右图根据传入的 `csv` 文件，`x` 轴对应震中距，`y` 轴对应最大烈度，颜色对应方位角，用 `svg` 画出气泡图。再用 `d3.hsl` 使图例上方位与颜色对应起来，在图例标上文字即可。

#### 3.5.3 图像分析

上图展示的是东日本大地震的烈度数据分析对比图。我们可以得到以下结论：

1、在右图中，左侧有空隙，说明地震发生在海中，即震源离陆地之间有一段距离，这一点从左图中也可以得到验证；

2、在右图中，沿着某一条横线从左到右看，结合图例可知，观测点所在方位从西北逐渐转向西南，即在同一烈度下，西北的观测点距离西南近，说明震源的西北方向更靠近陆地，西南方向相对远离陆地；

3、在右图中，上方的点既有位于西北方向的点，也有位于西南方向的点，经过上文分析，西南方向距离震源较远，受到的破坏却不比西北方向距离近的点少，证明地震的造成的破坏是有方向性的。事实上，此类构造型地震一般是由断层破裂产生的，我们把断层破裂的方向称作震源，断层破裂时会释放地震波，地震波会对地表造成破坏。但是断层破裂是有方向的，在断层破裂方向上的地方遭到的破坏程度也会更大。具体到这次地震上来说，结合左图，我们可以很清晰地看出，断层是自东北向西南破裂的，所以位于震源西南方向上的很多距离较远的地方也遭受了严重破坏。所以断层附近的居民需要提高防震意识，在断层附近的建筑等抗震等级应该高于其他地区。

4、通过对比不同地震，可以发现，相同深度的情况下，震级越大，影响范围越大，破坏程度越强，最大烈度越大；相同震级的情况下，深度越深，影响范围就越大，但是破坏程度越弱，最大烈度越小。

## 4 遇到的困难

1、在获得数据时，得到的原始数据不是 csv 文件，需要手动处理成可以利用的 csv 文件，有时候还需要将数据重组。在处理数据时，需要有一定的专业知识，对数据进行相应的数学处理。观测点信息处理任务较为繁琐，需要对信息分类划分成观测点名称，经纬度坐标，观测震度等多个参数进行处理。

2、画条形图时，查看每个震度单独情况，需要重新绘制坐标轴，用此震度最大数量作为 y 轴最大范围。

3、画条形图时，后续在添加重置原图表功能时，在一个函数中实现比较麻烦，需要先保存原始数据，在重置时引用原始数据绘制原图表。

4、画气泡图时，由于数据是按照时间顺序排列，而不是按照最大烈度，因此气泡图层出现了错误，所以为了确保它们在正确的位置上，我们先绘制了不同最大烈度的气泡，再在此基础上进行叠加。

5、绘制地图时用 d3 的 json 接口导入 geojson 文件出现问题，最终通过直接 import 的方式解决。

6、在加载观测点图片时遇到困难，发现 svg 并不能直接使用 url 路径，这样的话在 html 上加载不出来，需要先用 import 获取路径才能使用。

## 5 讨论与总结

本作品使用从 2000 到 2023 年的地震数据集，围绕分析日本地震最大烈度、震级、震源深度、位置的关系，揭示不同地区地震活动的时空规律，对数据进行多方面处理，并最终设计实现了日本地震可视分析系统，深入了解地震的发生机制和地壳运动规律，为地震预测、减少地震危害具有重要意义。