## 中国大学生计算机设计大赛作品简介

### 参赛年份：\_\_2019\_\_\_\_

## 【50763】工业危险废弃物可视化追踪

### 作品分类：

大类：\_\_\_\_大数据类\_\_\_\_\_\_\_ 小类：\_\_\_\_\_\_\_数据可视化\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### 作品或演示：

* 百度云盘分享地址：https://pan.baidu.com/s/1YzU1mBe4PgLS-0-jddSRWg （提取码bymn）

【注意】1. 百度云盘分享链接形式必须类似于：**http://pan.baidu.com/s/1eQF1CVg**，否则不合要求。请使用分享功能而不是直接复制地址栏的链接取得。

2. 可以将最新更新的作品进行分享。

3. 可以分享作品，也可以分享演示视频，或同时分享作品和演示视频。

4. **所有分享内容请放在一个文件夹内。**

5. **源码等机密信息请勿放在分享文件夹内。**

### 作品简介：

|  |
| --- |
| 工业危险废弃物可视化追踪是我们团队创作的数据可视化作品，作品以多种类图表展现了全国各省份近年来工业危险废物产生、利用和贮存情况。并利用回归分析对短期趋势进行了预测，以直观的方式唤起人们对我国工业危险废弃物现状的关注。作品创作过程中使用了python语言pyecharts模块对来自国家统计局的可靠来源数据进行了处理分析。作品的设计意义是结合近期内公众对江苏省多地化工事故的关注，引导人们聚焦昆山、泰州两次爆炸事故相关原因：工业危险废弃物的处理和贮存，加强环保和安全意识，对危险隐患进行监督和反映，希望减少此类事故再次发生。作品的优势在于根据数据特质清晰展示，以简洁美观的图形指向我国工业危险废弃物现状和危机。危机 |

### 安装说明：

|  |
| --- |
| 在谷歌浏览器中打开即可 |

### 演示效果：

|  |
| --- |
| 选取表现作品典型功能或效果的截图5-8幅，  设置每幅图片宽度不超过12厘米、高度不超过10厘米。  总页数不超过3页 |

### 设计思路：

|  |
| --- |
| 2019年3月底到4月初以来，江苏省连续发生多起化工厂爆炸事故，造成了人员伤亡和财产损失，引发了社会对化工行业的广泛关注。在对事故原因进行初步调查后，据报道昆山与泰州两起爆炸事故中都存在废弃物贮存不当的情况。抱着探寻的态度，我们产生了将全国今年工业危险废弃物的生产、贮存、利用、处置情况以可视化的方式进行展示。  我们希望通过这种方式引导公众对工业危险废弃物加以重视，呼吁政府和企业合理贮存和处理，减少类似事故再次发生的概率。  **我们的设计思路如下：**   1. 首先登入国家数据平台，从中国国家统计年鉴（<http://www.stats.gov.cn/tjsj/ndsj/>）中，获取2012-2017年工业危险废弃物相关数据 2. 利用OCR工具（https://zhcn.109876543210.com/）对相关数据进行了识别，将其保存为excel文件格式 3. 下载安装python中pyecharts包，学习了pyecharts新老版本的配置文档   （https://pyecharts.org/#/zh-cn/intro）   1. 对数据类型和大小进行了初步查看，选定可视化图表 2. 利用地图热力图从空间和时间分析了2012-2017年全国不同地区工业危险废弃物贮存量的密度分布 3. 利用散点图对各省处置利用量同生产量进行综合分析，引入面积反映贮存量，完成对数据的全貌概括 4. 由于获取的数据直到2017年，我们调用了多种方法对已有数据回归分析，产生2018和2019年的预测数据 5. 利用3D柱状图将各省处置利用率进行了展示分析各省危险废弃物的处理能力 6. 利用词云图反映了2018-2019年预测各省危险废弃物贮存量的增长率 7. 添加对各张图表的视觉效果进行了反复检验，调整参数获得了较佳的视感 8. 利用凡科网平台（https://jz.fkw.com/）设计了展示页面，将数据的分析结果进行了总结   我们的作品简洁美观，数据完整，其来源的可信度具有较高保障，合理选取了不同类型的图标进行了数据展示，同时消除了大数量级数据不易只管理解的特点，消除了普通人的理解障碍。作品易于传播，有助于实现向社会大众宣传的目的。 |

### 设计重点与难点：

|  |
| --- |
| **1.重点难点**  设计的重点在于如何将不同类别的数据都选取合适的图表形式进行展示，尽力避免不同省份间数据量的巨大差异影响最终的可视化结果。我们对pyecharts包中的不同类型的图标进行了反复的尝试，尝试了利用不同类型数据构造出更能反映本质特征的数据，以及如何尽可能改善视觉感官，消除大众对数据的陌生感。  设计的难点主要在于作品刚开始的数据来源问题，从包括学研网等数据平台上获取的数据普遍存在数据陈旧和完整度太低的问题，而从私人付费平台上获取数据存在价格昂贵和可信度存疑的痛处，爬虫抓取也存在数据筛选和反爬虫制约的问题。仔细思考数据需求后，从国家统计年鉴中获取的数据利用OCR进行了识别。  还有一个难点就是pyecharts包在设计过程中进行了更新，新版的配置文档尚未完成编辑，空白率较高，对于过程中尝试某些使用频率不是很高的图表的使用造成了很大的困难。  **2.关键技术**  作品使用了python语言，pyecharts包，OCR识别工具和页面设计平台，其中最关键的是对pyecharts包中不同图标的合理配置和调用。 |