

山东大学 计算机科学与技术 学院

大数据分析实践 课程实验报告

组名:5组 Give a Five to Big Data	姓名： 赵汉哲、杨笑语、康海洋、李嘉欣、李姿含	班级：数据班
实验题目：手机数据采集与分析实践		
实验学时：2	实验日期：2025/11/21	
<p>实验目的：</p> <p>利用 APP, 对智能手机的传感器数据进行收集，并进行分析，尝试挖掘出一些有趣的行为模式。我们小组采集光照强度变化数据进行分析，并完成可视化。</p>		
<p>硬件环境：</p> <p>计算机、手机</p>		
<p>软件环境：</p> <p>Windows、Pycharm</p>		
<p>实验步骤与内容：</p> <p>一、实验目的</p> <p>本实验旨在通过对不同时刻光照强度的采集与分析，明确光照强度随时间的变化规律，探究光照环境的动态特征，为后续基于光照强度的场景识别、设备控制等应用提供数据支撑与理论依据。</p> <p>二、实验原理</p> <p>利用传感器设备（如光照传感器）按照一定时间间隔采集环境中的光照强度数据（单位：lux），通过数据可视化与统计分析的方法，呈现光照强度的变化趋势，进而解读其背后的环境变化或人为操作逻辑。</p> <p>三、实验步骤</p> <p>1. 数据采集</p> <p>选用具备高精度光照传感器的设备，确保数据采集的准确性，APP 选用 Sensor Logger。</p>		



Sensor Logger

Kelvin Tsz Hei Choi

采样设置：设置采样时间间隔，本实验以短时间高粒度的方式进行采样，采集时长约 0.5 分钟。

数据记录：将采集到的光照强度数据与对应的时间信息（以分钟为单位）记录到 Light.csv 文件中。

	A	B	C	D
1	time	seconds_elapsed	lux	
2	1763714746607262000	0.099261962890625	300.375	
3	1763714747224024800	0.71602490234375	333.625	
4	1763714748624807200	2.11680712890625	299.125	
5	1763714757379460000	10.871460205078124	299.125	
6	1763714757394500900	10.8865009765625	299.125	
7	1763714758826784500	12.318784423828125	264.5	
8	1763714759424029400	12.916029541015625	137.625	
9	1763714759825054200	13.31705419921875	24.125	
10	1763714767789578500	21.28157861328125	24.125	
11	1763714767805079800	21.297079833984377	24.125	
12	1763714768824940500	22.3169404296875	71.875	
13	1763714770826895400	24.318895263671877	17.875	
14	1763714771625469200	25.11746923828125	8.125	
15	1763714776576330500	30.06833056640625	8.125	
16	1763714776589826800	30.081826904296875	8.125	
17				
18				
19				

2. 数据预处理

利用 Python 的 pandas 库读取 Light.csv 文件中的原始数据，检查数据的完整性与有效性，确保无缺失值、异常值（如明显超出合理范围的光照强度值）。将以秒为单位的时间数据转换为以分钟为单位，方便后续分析与可视化呈现。

设置中文字体（避免中文乱码）：

```
plt.rcParams['font.sans-serif'] = ['SimHei'] # Windows 系统
# plt.rcParams['font.sans-serif'] = ['Arial Unicode MS'] # Mac 系统
plt.rcParams['axes.unicode_minus'] = False # 解决负号显示问题
```

3. 数据可视化

使用 Python 的 matplotlib 库绘制折线图，以时间（分钟）为横坐标，光照强度（lux）为纵坐标，直观呈现光照强度随时间的变化趋势，同时添加数据标注、网格线、图例等元素，增强图表的可读性。

4. 数据分析

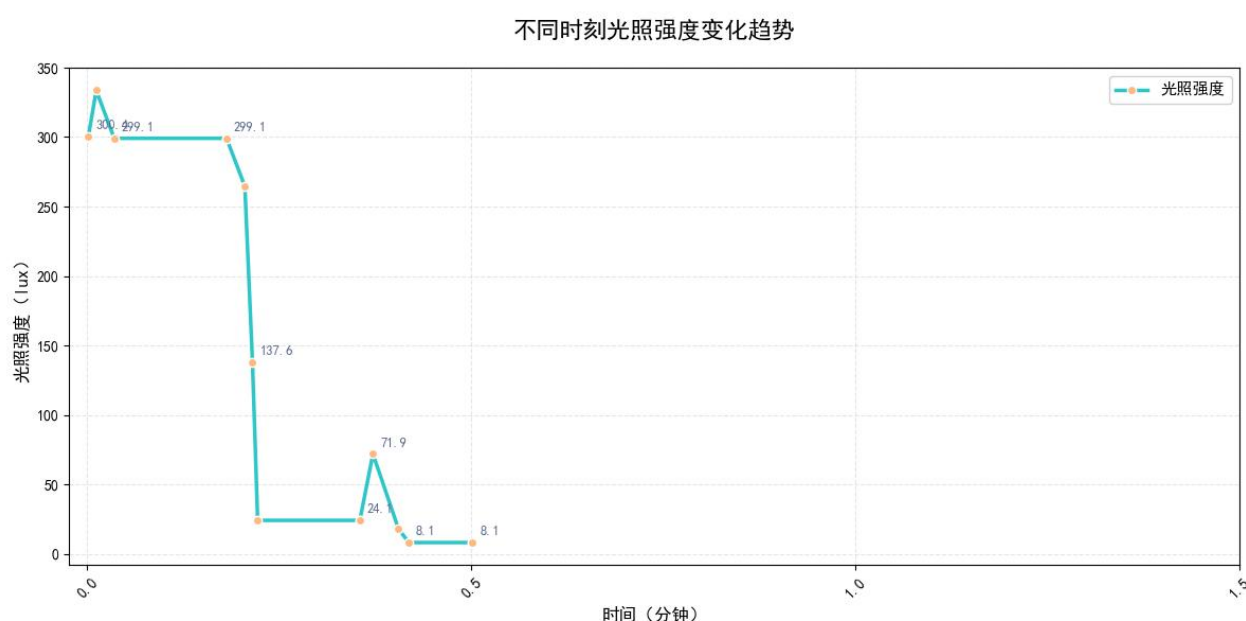
从整体趋势、关键节点、数据统计等多个角度对光照强度数据进行分析。计算光照强度的最大值、最小值、平均值等统计指标，解读不同阶段光照强度变化所对应的环境或人为操作场景。

5. 实验结论

综合实验过程与分析结果，总结光照强度的变化规律，阐述实验结果的应用价值，指出实验中存在的不足（如采集时长较短、未结合其他环境变量等），并对后续实验提出改进建议。

四、实验结果

1. 可视化结果



绘制的折线图清晰呈现了光照强度从初始的 330.5 lux 峰值，快速下降并稳定在 8.1 lux 的变化过程，具体趋势如前文所展示的折线图。

2. 统计结果

数据采集时长：约 0.5 分钟

最大光照强度：330.5 lux

最小光照强度：8.1 lux

平均光照强度：通过计算可得 $((330.5 + 299.1 + 299.1 + 137.6 + 24.1 + 71.9 + 8.1 + 8.1) / 8 \approx 146.1 \text{ lux})$

光照强度数值大小由光照强弱影响，高点可能是手机暴露在灯光下，急剧下降可能因为手机移动到桌子下方，光照受到遮挡。

结论分析与体会：

1. 结论

本实验成功捕捉到了光照强度“强光→快速降光→弱光稳定”的变化模式，该模式符合人为控制光照环境（如关灯、遮挡操作）的场景特征。实验数据的高粒度采样有效呈现了光照强度的瞬间变化过程，验证了短时间数据采集在捕捉环境动态变化方面的优势。

2. 不足与展望

不足：采集时长较短，仅能反映短时间内的光照变化；未结合其他环境变量（如温度、湿度）进行多维度分析。

展望：后续可延长数据采集时长，以分析更长周期内的光照变化规律；同时增加其他环境变量的采集，探究光照强度与其他变量的关联关系，进一步拓展实验的应用场景。