

题目十六：1824-2018 年北极涛动指数 NAO 数据分析

【背景】

北极涛动北纬 55°地区之间的地面大气压力的周期性变化。位于北纬国的阿拉斯加州南部、苏格兰的格拉斯哥和俄罗斯的莫斯科。北纬 55°就是低压，反之亦然。当北极的气压为低值时，北极涛动指数为正。

北极涛动指数为正时，高压区向南扩展，风暴路径北移，穿过斯堪的纳维亚和美国的阿拉斯加州，地中海地区和美国加利福尼亚州天气干燥，而欧洲和亚洲天气温暖。

当北极涛动指数为负时，中纬度地区的气压相对较低，地中海地区和美国加利福尼亚州天气变得潮湿，出现更多风暴和龙卷风，而欧洲和亚洲的内陆地区比较寒冷。

一些气候学家认为北大西洋波动/北半球环行模式(NAO/NAM)是影响范围更为广泛的北极涛动(AO)的一部分。据此，这种现象有时被总称为北极涛动/大西洋波动(AO/NAO)。

【数据说明】

数据集为 1824-2018 年北极涛动指数 NAO 数据，如下图所示。

Year	January	February	March	April	May	June	July	August	September	October	November	December
1824	-0.16	0.25	-1.44	1.46	1.34	-3.94	-2.75	-0.08	0.19	-99.99	-0.7	-0.01
1825	-0.23	0.21	0.33	-0.28	0.13	0.41	-0.92	1.43	-0.95	1.98	1.06	-1.31
1826	-3.05	4.87	-0.97	1.78	-1.2	0.83	1.89	2.72	-0.76	0.18	-2.41	-0.59
1827	-0.45	-3.72	1.83	-0.83	1.2	-0.07	2.02	-3.56	-0.07	-3.02	-1.42	2.7
1828	1.27	0.37	-0.18	0.04	-1.59	-1.33	-4.4	-2.54	-2.78	0.1	-2.57	3.04
1829	-2.48	0.32	-2.54	0.12	1.8	-0.1	0.33	0.77	0.78	0.71	-0.33	-0.43
1830	-2.33	1.2	3.58	3.08	-0.05	-0.85	3.19	-0.35	2.04	2.04	2.19	-3.13
1831	-2.91	1.4	1.48	-3.15	-2.47	-1.36	2.71	-3.04	-1.53	0.85	0.26	0.36
1832	-0.04	0.83	2.12	-1.51	-1.96	-3.62	-2.57	0.92	1.45	2.25	0.62	3.32
1833	-0.36	2.52	-2.89	2.02	0.69	-1.52	0.13	-1.74	-0.93	-1.75	1.4	4.17
1834	3.07	2.66	1.37	-2.38	-1.03	0.27	-0.73	-0.86	-0.62	0.3	-2.28	0.11
1835	0.37	3.37	1.54	-1.02	0.58	0.1	0.57	2.35	0.29	-0.3	-1.31	-1.46
1836	1.47	0.06	2.28	0.87	-1.5	2.95	4.33	2.8	-1.59	-1.17	2.04	-1.41

- 字段说明：
- (1) Year 列，表示年份
 - (2) 列 January~December 分别表示对应月份的 NAO 值。

【任务】

- 用 pandas 库读取 “nao.long.data.csv” 文件，将所有时间抽取为单独的列 Date（形式为 YYYY-MM-01），所有异常平均值抽取为一个单独的列 SOI，将所有缺失值丢弃处理，异常值-99.99 全部替换为-1，并导出到新的 txt 文件 “nao_dropnan.txt”，第一行为表头，列名分别为 Date 和 NAO，且表头和数据行中的不同字段信息都是用逗号分割，如下图所示。

```
Date,NAO
1824-01-01,-0.16
1824-02-01,-0.25
1824-03-01,-1.44
1824-04-01,-1.46
1824-05-01,-1.34
1824-06-01,-3.94
1824-07-01,-2.75
1824-08-01,-0.08
1824-09-01,0.19
1824-10-01,-1.0
1824-11-01,-0.7
1824-12-01,-0.01
```

- 2、重新读取新的数据集 “nao_dropnan.txt”，选择 NAO 字段，统计最大值、最小值、平均值。
- 3、重新读取文件 “nao_dropnan.txt”，利用第三步统计结果最大值 `maxValue`、最小值 `minValue`，利用 `category = [minValue, 0,maxValue]`和 `labels = ['ColdRelate', 'WarmRelate']`将 NAO 进行离散化；并将离散化结果作为一个新的列 Label 添加到原始数据集，并保存为 “nao_dropnan_result.csv”，从左到右三个列名分别为 Date、NAO、Label；根据离散化结果画出饼状图，保存为 “nao_pie.png”，要求分辨率不低于 300dpi。
- 4、重新读取文件 “nao_dropnan_result.csv”，利用 `matplotlib` 库，可视化显示 NAO 值，要求包括图例、图标题，x 轴刻度以年显示且间隔为 10，y 轴显示刻度值，曲线颜色为蓝色。

【要求】

- 1、根据以上数据处理任务，设计并编程实现“数据分析与可视化系统”，要求
 - ① 各个任务选择用菜单实现（菜单可用字符串输出模拟，或者 Tkinter 形式实现）。
 - ② 各个任务名称自己定义，须由独立的函数实现，且每个任务执行成功与否须给出必要的文字提示。
 - ③ 数据输入和结果输出的文件名须由人工输入，且输出结果都要以文件形式保存。
 - ④ 为保持程序的健壮性，各个任务执行过程中需要进行必要的判断（如文件是否存在、输入是否合法等）、程序异常控制等。
- 2、根据以上统计结果，书写不少于 300 字的结果分析。