农产品价格数据分析

一 数据描述

1）数据参数 该数据每日进行采集汇总。数据范围涵盖全国主要省份（港澳台西藏海南暂无数据）的 180+ 的大型农产品批发市场，380+的农产品品类（由于季节性和地域性等特点，每日的数据中

不一定会涵盖全部的农产品品类）。

2）数据类型

|  |  |  |  |
| --- | --- | --- | --- |
| **农产品批发市场价格数据** | **中文名称** | **英文名称** | **数据类型** |
| 农产品名称 | name | STRING |
| 批发价格 | price | FLOAT |
| 采集时间 | crawl\_time | STRING |
| 批发市场名称 | market | STRING |
| 省份 | province | STRING |
| 城市 | city | STRING |

3）数据清理说明

该数据暂无敏感数据项

数据样例：该数据样例提供了 2022年 1月份的农产品批发价格的数据，每五天汇总一个表格；

所用数据：本测试只用 2022年 1.1-1.5 的数据

二、功能需求：

1. 农产品市场个数统计

（1）统计每个省份的农产品市场总数

（2）统计山东省售卖蛤蜊的农产品市场占全省农产品市场的比例

2. 农产品种类统计

（3）统计每个省农产品种类总数，找出排名前3的省份

（4）统计排名前 3 的省份共同拥有的农产品类型

3. 价格区间统计

（5）求每天价格排名前三的菜品

结果：

20221-01-01 黄鳝 120

20221-01-01 樱桃 80

20221-01-01 牛肉 58

20221-01-02 黄鳝 130

20221-01-02 人参果 70

20221-01-02 带鱼 56

………

（6）计算山西省的每种农产品的价格波动趋势，即计算每天价格均值。

结果：

2022-01-01 山西省 大豆 2.6

2022-01-02 山西省 大豆 3.1

2022-01-03 山西省 大豆 2.9

2022-01-04 山西省 大豆 2.5

2022-01-05 山西省 大豆 3.5

………

某种农产品的价格均值计算公式： PAVG =(PM1+PM2+...+PMn-max(P)-min(P))/(N-2) 其中，P 表示价格，Mn 表示 market，即农产品市场。PM1 表示 M1 农产品市场的该产品 价格，max(P)表示价格最大值，min(P)价格最小值