

美食数据分析

需求背景介绍

中国是一个餐饮文化大国，长期以来在某一地区由于地理环境、气候物产、文化传统以及民族习俗等因素的影响，形成有一定亲缘承袭关系、菜点风味相近，知名度较高，并为部分群众喜爱的地方风味著名流派称作菜系。其中，鲁菜、川菜、粤菜、闽菜、苏菜、浙菜、湘菜、徽菜享称为"八大菜系"。

在日常生活人们多热爱于哪些食物，哪些菜系？

数据说明

数据样例

dishes_system	dish_name	private_int_fcount	craft	taste	time	difficulty	main_ingredient	ingredient
山西菜	糖醋鲤鱼	16673	炸	酸甜味	15分钟	初中水平	鲤鱼	番茄酱,淀粉,料酒,生抽
山西菜	糖醋鲤鱼	15482	烧	酸甜味	15分钟	初中水平	鲤鱼,蔬菜粒,西兰花	面粉,水,细砂糖,陈醋,生
山西菜	平遥蜜汁长山药	10874	炸	甜味	15分钟	初中水平	山药	白糖,水,蜂蜜,植物油
山西菜	油泼面	4394	煮	香辣味	30分钟	初级入门	面粉,鸡蛋	花生油,醋,蒜,生抽,细香
山西菜	黄酒神仙鸡	3907	焖	咸鲜味	数小时	未知	整鸡,猪蹄,五花肉	姜片,葱结,冰糖,生抽,老
山西菜	过油肉	3118	炒	咸鲜味	5分钟	初级入门	里脊肉,青蒜,葱白	鸡蛋清,生抽,湿淀粉,料
山西菜	糖醋鲤鱼	2700	炸	酸甜味	30分钟	初中水平	鲤鱼,泡椒辣椒	料酒,食盐,酱油,葱,姜,大
山西菜	栗子黄焖鸡	2358	焖	咸鲜味	60分钟	初中水平	鸡(土鸡, 家养),鲜栗子	食盐,冰糖,葱,姜,蒜,料酒
山西菜	面皮	1995	拌	香辣味	30分钟	初级入门	面粉	玉米淀粉,黄瓜,胡萝卜
山西菜	虾酱豆腐	1844	炒	咸鲜味	15分钟	初中水平	北豆腐,虾酱	韭菜
山西菜	糖醋鲤鱼	1313	炸	酸甜味	15分钟	初中水平	鲤鱼	面粉
山西菜	步步高升万字扣	1188	蒸	咸鲜味	90分钟	中级掌勺	五花肉, 香辣盐菜,萝卜	生抽,蚝油,老抽,生姜,大
山西菜	油柿子	1083	炸	甜味	15分钟	中级掌勺	面粉,糯米粉	红糖
山西菜	高粱面搓长鱼鱼	975	拌	家常味	15分钟	中级掌勺	高粱面	黄瓜
山西菜	过油肉土豆片	910	爆	家常味	30分钟	新手尝试	猪里脊肉,土豆,青椒,木	蒜,小葱,生抽,鸡精,香醋
山西菜	糖醋鲤鱼	905	炸	酸甜味	60分钟	初中水平	鲤鱼,盐	面粉,淀粉,鸡蛋,大蒜,米

数据说明

dishes_system	Dish_name	Private_int_fcount	Main_infredient	infredient
菜系	菜名	收藏数	主料	辅料

实施计划

一、技术选择

拿取数据：

说明：爬取[美食杰](#)，以下所作的分析的数据全部来自于美食杰网站

技术选择：

requests 库
xpath 库分析
re 库清洗
使用 mysql 和 csv 存放

分析数据：

说明：使用 sql 分析出可视化要用的图表数据

技术选择：

sql 语句

可视化数据：

说明：拿到分析到的数据使用 echarts 制成图表

技术选择：

echarts:分析图表
jinja2: 传输数据
pyecharts:[官网](#)，将 python 和 echarts 结合一下，可以快速生成 echarts 图表
flask：使用 flask 框架生成
jieba：分词
wordcount:用来生成词云

前端展现：

说明：通过 html、css、js 制作前端页面

技术选择：

swiper:[官网](#)，生成轮播图，通过引入 swiper.js 文件编写配置即可使用
jquery:[官网](#)，实现一些卡片功能

环境准备

开发工具：pycharm, vscode

大数据分析平台：sql

分析过程

1、求出热门菜系 TOP10

代码：

```
SELECT
    dish_name,
    private_int_fcount
FROM
    foot_menu
ORDER BY
    private_int_fcount DESC
LIMIT 0,
10;
```

结果：

dish_name	private_int_fcount
水煮肉片	85142
酱爆鸡丁	58819
宫保鸡丁	47208
三蔬小炒肉丁	43974
鱼香肉丝	43747
西湖牛肉羹	42862
鱼香茄子	42618
农家小炒肉	41934
鱼香鸡蛋	38745
糖醋排骨	38494

2、分析数据中各菜系的占比情况

代码：

```
SELECT DISTINCT
    dishes_system,
    COUNT( dishes_system ) AS count
FROM
    foot_menu
GROUP BY
    dishes_system
ORDER BY
    count DESC;
```

效果：

dishes_system	count
山西菜	84
粤菜	84
闽菜	84
浙菜	84
沪菜	84
豫菜	84
湖北菜	84
西北菜	84
云贵菜	84
广西菜	84
其它菜	84
苏菜	83
清真菜	71
川菜	63
鲁菜	63
湘菜	63
徽菜	63

3.分析主料和辅料的使用率

代码：

```
SELECT
    main_ingredient
FROM
    foot_menu;

SELECT
    ingredient
FROM
    foot_menu;
```

效果：

main_ingredient	ingredient
猪里脊肉,鸡蛋,蒜薹,木	番茄酱,淀粉,料酒,生抽
挂面	面粉,水,细砂糖,陈醋,生
猪里脊肉,鸡蛋	白糖,水,蜂蜜,植物油
鲤鱼	花生油,醋,蒜,生抽,细香
草鱼	姜片,葱结,冰糖,生抽,老
丸子,辣椒,香葱	鸡蛋清,生抽,湿淀粉,料
山药,白糖	料酒,食盐,酱油,葱,姜,大
山药,白糖	食盐,冰糖,葱,姜,蒜,料酒
鸡翅中,鸡腿,色拉油,姜	玉米淀粉,黄瓜,胡萝卜
槐花,面粉	韭菜
鲤鱼,糖,醋	面粉
胡萝卜,木耳,尖椒,猪肉	生抽,蚝油,老抽,生姜,大
鲤鱼	红糖
土豆,面粉,淀粉	黄瓜
鸡腿肉,香草药20余种	蒜,小葱,生抽,鸡精,香醋
面粉,赤小豆	面粉,淀粉,鸡蛋,大蒜,米
山药,糖,油	食盐,醋,老抽,淀粉,鸡蛋

前端借鉴：

借鉴于

B 站地址：https://space.bilibili.com/32683063?spm_id_from=333.788.b_765f7570696e666f.1

源码地址：<https://github.com/local-host-8080/demo-html-css>

首页：

导航栏：23-MenuEffect

卡片：31-Card-Effect

百科：

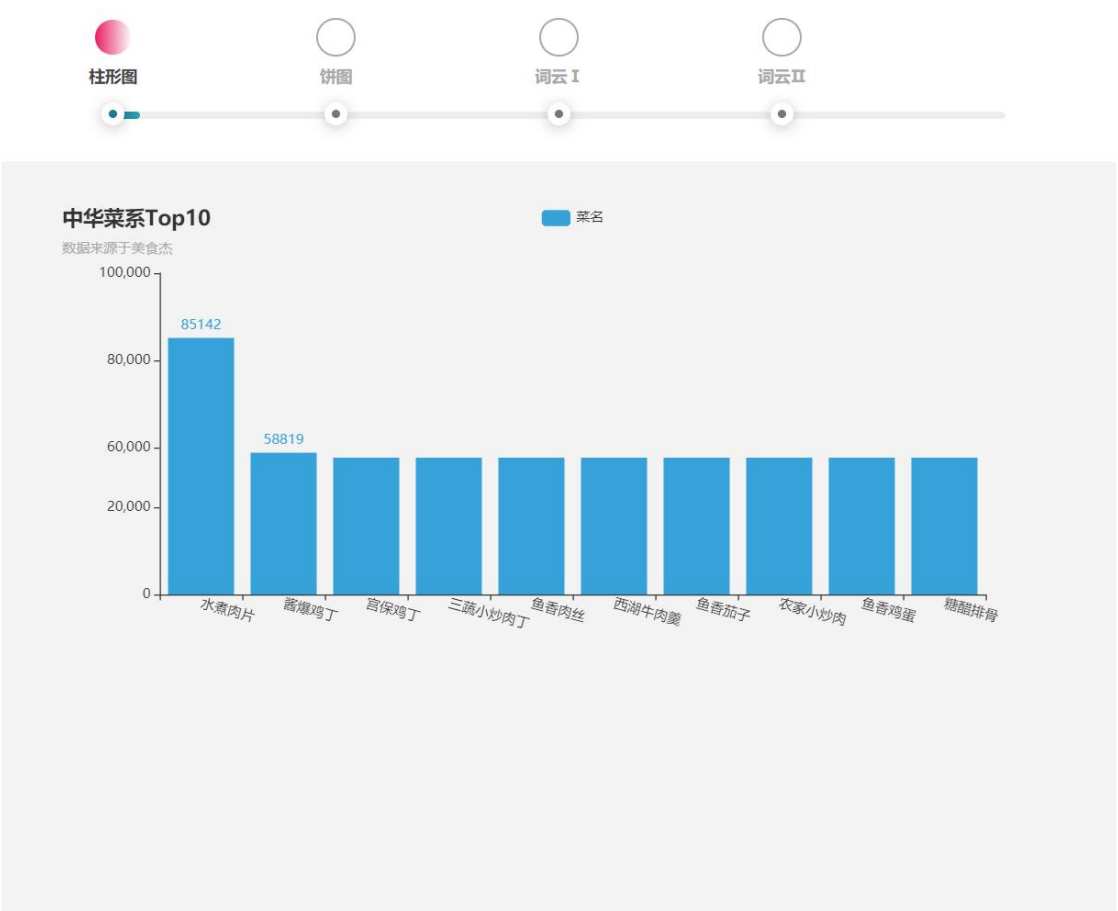
卡片：45-FullCard

展示：

图表展示：43-ProgressBar

可视化效果

可视化展示





辅料分析



结论与建议

1. 在美食杰网站上，通过热门菜系分析可以看出，前者多为肉类，相对于还是有许多人爱吃肉。
2. 前十菜系多为川菜，川菜喜欢的人多。
3. 通过各菜系占比情况可以看出江西菜、云南菜、贵州菜在同样爬取相同数量的情况之下还是居少数
4. 通过词云分析，主料一般使用五花肉、猪肉、鸡蛋、面粉、土豆居多，荤素、营养搭配居多
5. 辅料一般使用盐、料酒、生抽、老抽、白糖、姜、葱，可见盐是必不可少的。不过料酒烹调菜肴时不要放得过多，以免料酒味太重而影响菜肴本身的滋味。