数据提取方法：截屏，微信识别图片文字

数据：

罩衣

鞋，裙子，T恤

牛仔裤，T恤

牛仔裤，鞋，T恤

牛仔裤，短裤

鞋，T恤

牛仔裤，裙子

牛仔裤，鞋，短裤，T恤

牛仔裤

牛仔裤，鞋，T恤

T恤

罩衣，牛仔裤，鞋，裙子，T恤

牛仔裤，鞋，短裤，T恤

鞋，裙子，T恤

牛仔裤，T恤

裙子，T恤

罩衣，牛仔裤，裙子

牛仔裤，鞋，短裤，T恤

牛仔裤

牛仔裤，鞋，短裤，T恤

求频繁项集，可以直接数，也可以用库函数，此处使用python库函数apriori算法走完整流程：

from apyori import apriori

import apyori

# 最小支持度为5

minimum\_support = 5 / len(data)

# 最大信度为1

maximum\_confidence = 1

# 使用Apriori算法寻找频繁项集

frequent\_itemsets = apriori(data, min\_support=minimum\_support, max\_length=maximum\_confidence)

import typing

# 打印结果

results: typing.List[apyori.RelationRecord] = list(frequent\_itemsets)

# print(results)

for item in results:

    print(list(item.items)[0])

得到频繁项集

T恤 牛仔裤 短裤 裙子 鞋