# python工程师（web开发和爬虫方向）-面试经历

### **一、这家公司主要对亚马逊商品进行数据采集，问的问题比较杂。**

1. 是否了解线程的同步和异步？
2. 是否了解网络的同步和异步？
3. 链表和顺序表储存时各自有什么优点？
4. 使用redis搭建分布式系统时如何处理网络延迟和网络异常？
5. 数据仓库是什么？
6. 假设有一个爬虫，从网络上获取数据的频率快，本地写入数据的频率慢，使用什么数据结构好？
7. 你是否了解谷歌的无头浏览器？
8. 你是否了解MySQL数据库的几种引擎？
9. redis数据库有哪几种数据结构？

### **二、这家是做网络电视应用（教育/游戏等）的后台（.APK）**

1. 是否了解django中的manage.py自定义的用法？
2. django的常用功能有哪些？
3. django有哪些优势？
4. 是否对django的admin进行定制过？
5. 在django中有使用过原生sql语句吗？（查了一下可能是涉及到ORM的性能优化）

### **三、做大数据征信业务的公司，招django工程师**

1. django有什么优点？
2. 是否了解django admin定制？
3. 描述一下你的项目。

### **四、做银行/金融业决策系统，招django工程师**

1. 详细描述一下做某个项目的过程，描述的是一个爬虫项目的过程
2. 是否有对爬虫采集结果进行数据分析，有哪些？
3. 假设现在数据已经爬取完成，需要开发一个接口给用户，你会设置哪些筛选维度？（薪资，地点，职位关键词等）
4. 筛选维度设计好后，现在需要在用户输入某个维度的条件后查询与用户查询目标最为匹配的条目，如何设计查询逻辑算法？（拿薪资举例，以500为区间建立引索，将薪资的下限和上限存入每个区间里面,不能有冗余空间。以用户查询值为基准，逆序查询指定个数的条目，前端显示最相关的数据。需要处理非500倍数的值。）   
   据实际例子去想解决思路，喜欢这种面试方式。
5. 决策系统的某个模块的数据结构如何设计？（使用向量集）