Spark SQL运行机制

1. 解析SQL之前，会创建sparksession，把元数据保存在SessionCatalog中，涉及到表名，字段名称和字段类型。创建临时表或者视图，其实就会往SessionCatalog注册
2. 词法和语法解析parse：使用ANTLR生成未绑定的逻辑计划，共两步
    1. 词法分析：Lexical Analysis，负责将token分组成符号类
    2. 构建一个分析树或者语法树AST
3. 绑定Bind：用分析器Analyzer绑定逻辑计划，在该阶段，Analyzer会使用Analyzer Rules，并结合SessionCatalog，对未绑定的逻辑计划进行解析，生成已绑定的逻辑计划
4. 优化Optimize：生成最优执行计划，优化器也是会定义一套Rules，利用这些Rule对逻辑计划和Exepression进行迭代处理，从而使得树的节点进行和并和优化
5. 使用SparkPlanner生成物理计划：SparkSpanner使用Planning Strategies，对优化后的逻辑计划进行转换，生成可以执行的物理计划SparkPlan.
6. 使用QueryExecution执行物理计划:此时调用SparkPlan的execute方法，底层其实已经再触发JOB了，然后返回RDD

流程图