

递交MapReduce作业

```
graph TD; Start([递交MapReduce作业]) --> MapBox; Start --> ReduceBox; subgraph MapBox [Map函数]; MapStep1[程序从HDFS获取数据] --> MapStep2{数据合法性验证}; MapStep2 -- Y --> MapStep3[设置Key/Value输出]; end; subgraph ReduceBox [Reduce函数]; ReduceStep1[接收Map结果] --> ReduceStep2[计算PV]; ReduceStep2 --> ReduceStep3[将输入结果写入HDFS]; end; MapStep3 --> ReduceStep1;
```

程序从HDFS获取数据

数据合法性
验证

Y

设置Key/Value输出

Map函数

接收Map结果

计算PV

将输入结果写入HDFS

Reduce函数