Spark提供的分区方式有3种一种是HashPartitioner，另一种是RangePartitioner。

Hash分区有可能会出现数据倾斜，Rang分区无法保证分区内有序。

自定义分区需要实现Partitioner类重写numPartitions和getPartitioner方法。

注意：重新分区是一件非常消耗性能的事情所以慎重使用。像频繁建立连接的这种可以通过分区来改善，重新分区的使用一定要仔细权衡利弊。

