Difference between Spark Map vs FlatMap Operation

A **flatMap** is a transformation operation. It applies to each element of RDD and it returns the result as new RDD. It is similar to Map, but FlatMap allows returning 0, 1 or more elements from map function. In the FlatMap operation, a developer can define his own custom business logic. The same logic will be applied to all the elements of the RDD.

A FlatMap function takes one element as input process it according to custom code (specified by the developer) and returns 0 or more element at a time. flatMap() transforms an RDD of length N into another RDD of length M.

**Map** is transformation operation It applies to each element of **RDD** and it returns the result as new RDD. In the Map, operation developer can define his own custom business logic. The same logic will be applied to all the elements of RDD.

Spark Map function takes one element as input process it according to custom code (specified by the developer) and returns one element at a time. Map transforms an **RDD** of length N into another RDD of length N. The input and output RDDs will typically have the same number of records.