↳ RDD

↳ Instance has got multiple methods

1. Map.   ↳ of Spark Context (Class)
                        ↳ .getOrCreate (conf =
                                    Spark Conf().
→ An instance in Spark   setAppName(
        lingo is Rdd.   "blah")
Rdd = {Spark Context . get ....   )

        ↳→ This creates pipelineline of work
        → This extracts data + apply pipeline.

    ↳ Map is a defined method

            ↳ Take → Function as input
                    → This creates and instantiates a
→ rdd = Rdd. textfile ('PATH')   new
                            data.   class.
        ↳ pivot pointer to the source ^path
    rdd2 =   → Applied line by line.
    → rdd. map(lamba α : α.split(' '))
            ↳ Method of rdd → Returns new
                            instance with
                            new fn added to
    ↳ rdd2.collect()   pipeline.
            ↓   ↳ like linked
        ↳ Render or   list
        Caller. → Run the
                pipeline on the src data

Read abt class method.

○ pseudo & spark
    ↳ collect → Run all fn saved in the list

↳ FIFO

Base data_storer
    ↳ numpy array based.

pop & append ↲
    `At diff ends.

↳ Map → For each mapper create new instance
    ↳ Class method.

↳ Collect → Returns result.

——————o——————x——————o——————

Flat - map.

  ↳ Explodes the data before outputting result.

  ↳ Runs. collect and then reshape it
    to (-1,)
      ↳ One list.