import findspark  
import os

os.environ["SPARK\_HOME"] = "/home/ser/Dev/Spark/spark-1.6.0-bin-hadoop2.6"  
os.environ["PYSPARK\_PYTHON"] = "/home/ser/anaconda3/bin/python"  
os.environ['PYSPARK\_SUBMIT\_ARGS'] = "--master local[4] pyspark-shell"  
findspark.init()  
  
import numpy as np  
import sys  
import sep

from operator import add

import pyspark

from pyspark import SparkContext

if \_\_name\_\_ == "\_\_main\_\_":

sc = SparkContext(appName="SourceExtractor")

#sc = pyspark.SparkContext("local", "Simple App")

#rdd = sc.fitsData("/home/ser/Dev/Kira/Kira-master/scratch/spark/data/")

#rdd = sc.fitsData("/Users/zhaozhang/projects/Kira/scratch/spark-ec2/data/")

#catalog = rdd.map(lambda (key, fits): (key, extract(np.copy(fits))))

#catalog.saveAsTextFile("temp-output")

sc.stop()

#logData = sc.textFile("/home/ser/test-start.py").cache()

#numAs = logData.filter(lambda s: 'a' in s).count()

#numBs = logData.filter(lambda s: 'b' in s).count()

#print("Lines with a: %i, with b: %i" % (numAs, numBs))

#lst = sc.parallelize(range(1000))

#lst.collect()