

1021

D m 4 ffk

m ffk D m

4 m ffk D m

1

1	ffk	
2	D m	
3		
4		

ffk

ffk

Python
Python

OS

2

D m D m D D

Hadoop

2

HDFS HDFS ;
HDFS
Hive
HQL

2

Python
HTTP
 HTTP
Cookie

 urllib3 Requests HTTP
 Chrome Xpath Beautiful Soup

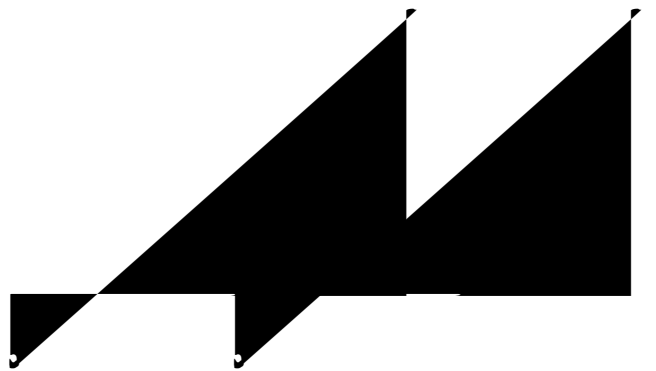
JSON

2

3

matplotlib pyplot

2




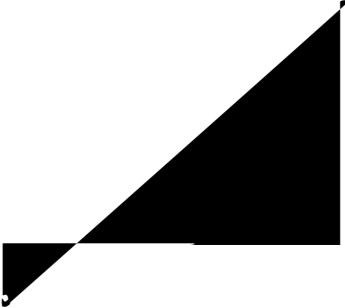
100

1

4

2

2

1		
2	ffk	ffk ffk ffk
3	D m	D m F D m D m D D m D D m
4		3 D m ffk
		

1

20

1 1 1

20

1	201	ffkm. . .201 03.3 0 ffk
2	. 3 2 3 201	ffkm .. 2 0 . .m ffk 201 0124211 1
3		ffkm .. 2 0 . .m ffk 201 1221102 31
4		ffkm .. 2 0 . .m ffk 201 02011 3320