



46 4

kk

p k

p k

kk

p k

kk

4

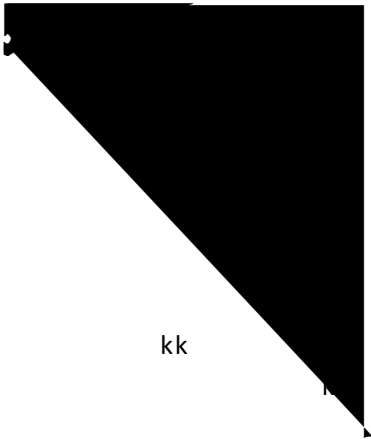
4	p k	
	kk	

p k

p k

Python
Python

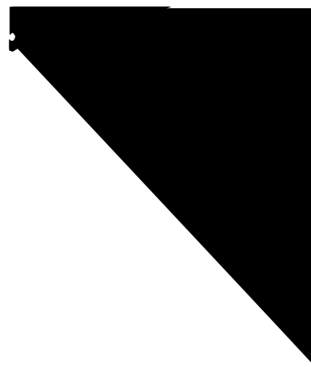
OS



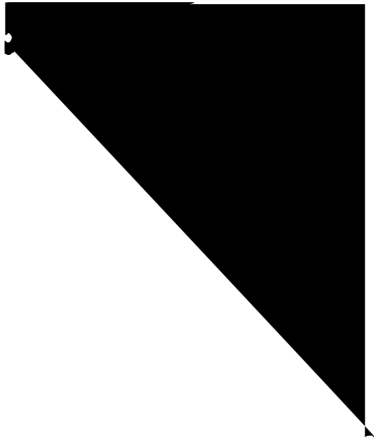
kk

ft

Hadoop

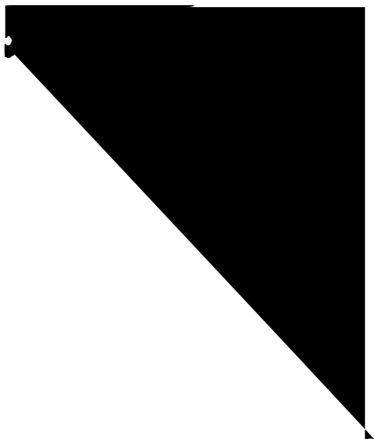


HDFS HDFS ;
HDFS ;
Hive
HQL

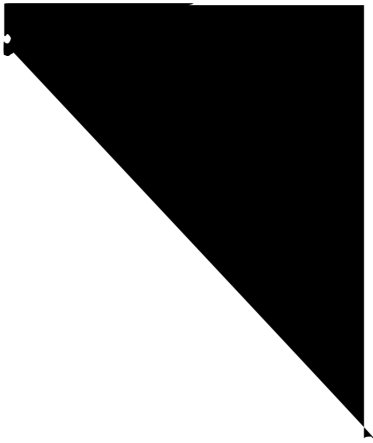


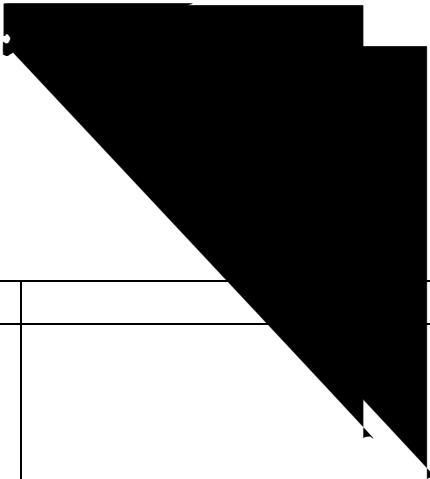
Python
HTTP
HTTP
Cookie

urllib3 Requests HTTP
Chrome Xpath Beautiful Soup
JSON



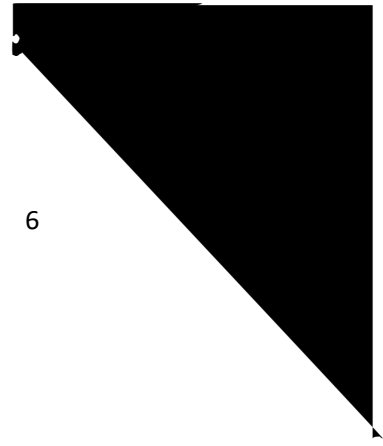
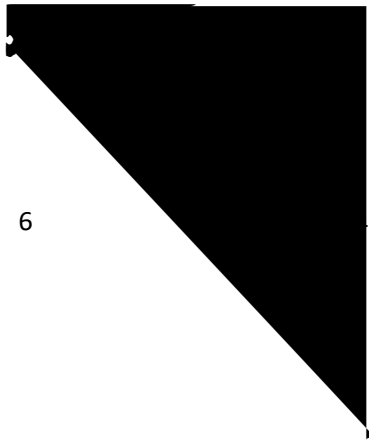
matplotlib pyplot





4		
	p k	p k p k p k
	kk	kk kk kk kk ft kk
		ft p pkk k p

4



6

4	64	pp ft 64 6 6	
	64	pp p 6 k ck	ft 64 64 44
		pp p 6 k ck	ft 64 4 446
		pp p 6 k ck	ft 64 6 644