





.

46 4

kk

p k

p k

kk

p k

kk

4

4	p k	
	kk	

**M**

p k

4

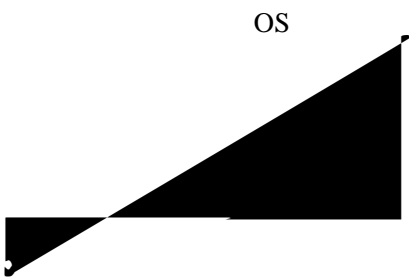
p k

•  
• M  
Python  
Python

M

M

1



**a**

kk

kk

ft

•  
• a



Hadoop

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 4 4



6

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