



•

xc

sg k

x

sg k  
R

xc

sg k

xc

	sg k	
	xc	

**M**

sg k

sg k

•

• **M**

Python  
Python

**M**

**M**

**1**

OS

**a**

xc

xc

E

•

• **a**

Hadoop

HDFS                    HDFS                    ;  
HDFS  
Hive  
HQL

•  
•

Python  
HTTP  
HTTP  
Cookie

urllib3    Requests    HTTP  
Chrome                    Xpath    Beautiful Soup  
JSON

•

•

matplotlib pyplot

•

	sg k	sg k sg k sg k
	xc	xc  xc xc E xc xc
		tff            tff s            RR  xtff etf tff  xsg



		gss ?                    k                    gs
	R	gss ?            s            ffk e ks            s    sx    gs            c
		gss ?            s            ffk e ks            s    sx    gs            c
		gss ?            s            ffk e ks            s    sx    gs            c