



ch

st

R

st  
D J

ch

st

ch

	st	
	ch	

st

st

Python  
Python

OS

ch

ch

R

d

Hadoop

HDFS                    HDFS                    ;  
HDFS  
Hive  
HQL

Python

HTTP

HTTP

Cookie

urllib3

Requests

HTTP

Chrome

Xpath

Beautiful Soup

JSON

matplotlib pyplot

	st	st st st
	ch	ch  ch ch R ch d ch
		d drst  d st R  st



		st ?      dr                      st
		st r?      st      ft      st    rst dst    st    ch
		st r?      st      ft      st    rst dst    st    ch
		st r?      st      ft      st    rst dst    st    ch