



nn

v nl

P

v nl

nn

v nl

nn

	v nl	
	nn	

v nl

v nl

Python  
Python

OS

nn

nn

P

ck

Hadoop

HDFS                    HDFS                    ;  
HDFS  
Hive  
HQL

Python  
HTTP  
HTTP  
Cookie

urllib3      Requests                    HTTP  
Chrome                                    Xpath      BeautifulSoup

JSON

matplotlib pyplot

	v nl	v nl v nl v nl
	nn	nn  nn nn P nn           ck nn
		ck   ck p  ck   Pn



		! d p ! ! !
		p ! !n ! ! nl np ck !
		p ! !n ! ! nl np ck !
		p ! !n ! ! nl np ck !