

Jiangwei Wu (jiw136@pitt.edu)

Song Tang (sot13@pitt.edu)

1. How many hits were made to the website item “/assets/img/homelogo.png”?

Part1.java ×		part-r-00000 ×	
1	/assets/img/home-logo.png	98744	
2			

2. How many hits were made from the IP: 10.153.239.5

Part2.java ×		part-r-00000 ×	
1	10. 153. 239. 5	547	
2			

3. Which path in the website has been hit most? How many hits were made to the path?

Part3.java ×		part-r-00000 ×	
1	"GET	26	
2	/	99303	
3	/assets/css/combined.css	117348	
4			

4. Which IP accesses the website most? How many accesses were made by it?

Part4.java ×		part-r-00000 ×	
1	10. 1. 1. 113	1	
2	10. 1. 1. 125	12	
3	10. 1. 100. 183	14	
4	10. 1. 100. 199	35	
5	10. 1. 109. 54	44	
6	10. 1. 11. 81	69	
7	10. 1. 12. 172	93	
8	10. 1. 122. 130	109	
9	10. 1. 146. 91	321	
10	10. 1. 181. 142	10686	
11	10. 102. 126. 202	12298	
12	10. 104. 62. 79	14807	
13	10. 118. 250. 30	18282	
14	10. 169. 128. 121	22516	
15	10. 173. 141. 213	47503	
16	10. 216. 113. 172	158614	
17			

5. Ngram result:

NGram.java ×		part-r-00000 ×	
1	as	6	
2	bb	1	
3	bf	4	
4	cd	1	
5	cf	1	
6	db	5	
7	dc	1	
8	ds	6	
9	dv	2	
10	fd	5	
11	sa	6	
12	sc	1	
13	sv	5	
14	vd	7	
15	vv	1	
16			