

1. Utilizing the analytics_demo data set via HBase data browsing, using the prefix search capability - within domain 15 **which country France, Italy or US had the most visits?** (Hint: use the row key search for each country; write the result for each country)

- a) US
- b) Italy
- c) France

2. How many different domains are there in the analytics_demo data set (**Using the Hue web interface**)? *Hint:* domains are rows, not individual cells in the rows. Only count the default shown, don't select all.

- a) 10
- b) 100
- c) 15
- d) 55

3. Utilizing the analytics_demo data set via HBase data browsing - **which domain has the most (total) site visits** for the US? (hint: sort columns by descending order)

- a) domain.10
- b) domain.0
- c) domain.12
- d) domain.13

4. Which 2 domains are the least visited (smallest total number of site visits) for all three countries together?

- a) domain.16 and domain.15
- b) domain.17 and domain.12
- c) domain.16 and domain.12
- d) domain.14 and domain.0

What to Check with your TAs your queries and their results and answer to above questions.

	Queries	Results
question 1	your query for each country (3 queries)	your result for each country total: 3 results for US, Italy and France
question 2	your query or analysis pipeline to find the number of domains as requested in the question	A number showing the number of domains
question3	your query to find most (total) site visits for the US for the given domains	4 numbers for the given domain in the question3
question4	your query to find smallest total number of site visits	The 2 domain number as the result

- For more practice you can do similar analytics using the provided demo datasets in Cloudera or HBase examples.
- Access to Hue interface via <http://demo.gethue.com> easier WITHOUT NEED TO VBOX