# IST769 Homework 10 Submission

## Basic Information

Your Name: Srihari Busam   
Your SUID: sbusam  
Your Email: sbusam@syr.edu  
Date Due: 12/09/2021  
Homework #: 10

**NOTE:** I did setup kafa containers locally on my home computer. It took a while to make the setup running. The other point is that the queries not running without putting the post fix “EMIT CHANGES” to the queries. The output windows also very different from the lecture or the homework quick walkthrough. However, the with the postfix the queries are running as expected it seems. Just want to bring this to your attention.

## QUESTIONS:

## **Q1:**

Write KSQL to create a stream named **weblogs** from the JSON keys in the **weblogs** Kafka topic. Make sure to set the TIMESTAMP property to the timestamp from the stream.

## **Answer**

|  |
| --- |
| Code |
| CREATE STREAM weblogs\_new (TimeStamp bigint, Browser varchar, OS varchar, Uri varchar, User varchar) WITH (KAFKA\_TOPIC='weblogs', TIMESTAMP='TimeStamp', VALUE\_FORMAT='JSON');  DESCRIBE weblogs EXTENDED;  SELECT \* FROM weblogs EMIT CHANGES; |

Screen shot:

Text

Description automatically generated

## Q2:

Write a KSQL statement create a persistent stream/table called **homepage** which only displays visitors to the root of the website (/). It should display all columns from the **weblogs** stream.

## Answer

|  |
| --- |
| Code |
| CREATE STREAM homepage AS SELECT \* FROM weblogs WHERE Uri = '/';  DESCRIBE homepage;  SELECT \* FROM homepage EMIT CHANGES; |

Screenshot:

Text

Description automatically generated

## Q3:

Write a KSQL statement to count operating systems users (os) in 60 second windows. After 60 seconds, the counter should reset, and counts should begin again.

## Answer

From the screenshot it can be observed that the os counter resets to 1 as the window goes beyond 1 min.

|  |
| --- |
| Code |
| SELECT OS, COUNT(\*) FROM weblogs WINDOW TUMBLING (SIZE 1 MINUTE) GROUP BY OS EMIT CHANGES; |

Screenshot:

Graphical user interface, text

Description automatically generated with medium confidence

## Q4:

Write a KSQL persistent stream/table called **user\_activity** which will display a count of user activity on the website within 1-minute sessions.

## Answer

|  |
| --- |
| CODE |
| CREATE STREAM weblogs2 AS SELECT 1 AS Grp, \* FROM weblogs;  CREATE TABLE user\_activity\_new AS SELECT Grp, COUNT(\*) FROM weblogs2 WINDOW TUMBLING (SIZE 1 MINUTE) GROUP BY Grp;  DESCRIBE user\_activity\_new;  SELECT \* FROM user\_activity\_new EMIT CHANGES; |

Text

Description automatically generated

## Q5:

Write a KSQL statement to display users who have more than 1 pages of activity in a 1-minute session.

## Answer

|  |
| --- |
| Code |
| SELECT User, COUNT(\*) FROM weblogs WINDOW TUMBLING (SIZE 1 MINUTE) GROUP BY User HAVING COUNT(\*) > 2 EMIT CHANGES; |

Screenshot:

Text

Description automatically generated