# Sample Queries for CloudTrail to Athena

**Example 1: Find all create-bucket actions**

```
SELECT * FROM <CloudTrail Trail table name>
```

where eventName = ‘CreateBucket’;

**Example 2: Find all users who signed in to the console in a set of Regions.**

**This may require signing in and signing out to have this activity show up because you just created your trail. You can do this on a different browser/incognito mode.**

**Allow a few minutes for your trail to start delivering logs before running the below code.**

**To search through JSON arrays in a column, use “.” This will then be *columnname.fieldname* as in the example below.**

```
SELECT
    eventTime,
    useridentity.arn,
    awsRegion
FROM
    <CloudTrail Trail table name>
WHERE
    awsRegion in (‘us-east-1’, ‘us-west-2’)
AND
    eventName = ‘ConsoleLogin’
```

**Example 3: Create a CloudTrail log file table with partitions for performance improvement.**

```
CREATE EXTERNAL TABLE <New CloudTrail Table name>
eventversion STRING,
useridentity STRUCT<
               type:STRING,
               principalid:STRING,
               arn:STRING,
               accountid:STRING,
               invokedby:STRING,
```

```
                accesskeyid:STRING,
                userName:STRING,
  sessioncontext:STRUCT<
    attributes:STRUCT<
                mfaauthenticated:STRING,
                creationdate:STRING>,
    sessionissuer:STRUCT<
                type:STRING,
                principalId:STRING,
                arn:STRING,
                accountId:STRING,
                userName:STRING>,
    ec2RoleDelivery:string,
    webIdFederationData:map<string,string>
  >
>,
eventtime STRING,
eventsource STRING,
eventname STRING,
awsregion STRING,
sourceipaddress STRING,
useragent STRING,
errorcode STRING,
errormessage STRING,
requestparameters STRING,
responseelements STRING,
additionaleventdata STRING,
requestid STRING,
eventid STRING,
resources ARRAY<STRUCT<
                arn:STRING,
                accountid:STRING,
                type:STRING>>,
eventtype STRING,
apiversion STRING,
```

```
readonly STRING,
recipientaccountid STRING,
serviceeventdetails STRING,
sharedeventid STRING,
vpcendpointid STRING,
tlsDetails struct<
  tlsVersion:string,
  cipherSuite:string,
  clientProvidedHostHeader:string>
)
PARTITIONED BY (region string, year string, month string, day string)
ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'
STORED AS INPUTFORMAT
'com.amazon.emr.cloudtrail.CloudTrailInputFormat'
OUTPUTFORMAT
'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://CloudTrail_bucket_name/AWSLogs/Account_ID/CloudTrail/';
```

1. Open s3.
    a. Click on the log data bucket (the bucket made when we connected CloudTrail to Athena using the AWS workflow).
    b. **Make sure its not the result bucket.** (This would have the keyword "result" in its name. *Tip:* You can download results of your Athena queries from the result bucket.)
2. Descend through this path: AWSLogs/ > your_account _id > CloudTrail/.
3. Click **Copy the S3 URI** button on the "CloudTrail/" object screen.
4. Replace the *LOCATION* in the above query with the S3 URI in step 3. *LOCATION* points to the location of your log data.
5. Run the query.
6. Note how "partitioned" appears on the tile name of the new table in Athena.
7. Try to preview data by clicking on the three vertical dots to the right of your new table. Nothing shows. This is because you need to specify the partition you wish to query.

    *Less data scanned = quicker results and reduced cost*

8. Specify the partition. (Replace all text in red with actual values as in screenshot below.)

    ALTER TABLE <*New CloudTrail Table name*>
    ADD PARTITION (region='region',

year=’yyyy ‘,
month=’mm’,
day=’dd’)
LOCATION
‘s3://cloudtrail_bucket_name/AWSLogs/Account_ID/CloudTrail/region/yyyy/mm/dd/’

9. Replace *LOCATION (up to “..CloudTrail/”)* using the S3 path in the *LOCATION* keyword in the create statement above. *LOCATION* points to the location of your log data.

10. Choose a date, for example, today.

```
ALTER TABLE cloudtrail_logs ADD
    PARTITION (region='us-east-1',
               year='2023',
               month='03',
               day='09')
    LOCATION 's3://aws-cloudtrail-logs-testev/AWSLogs/497395817824/CloudTrail/us-east-1/2023/03/09/'
```

11. To use less fields (for example, region only), modify your create table statement by removing the fields you do not want. Your alter table statement will therefore only need the region field.

12. Preview data in your new table again; now you see data in your partition.