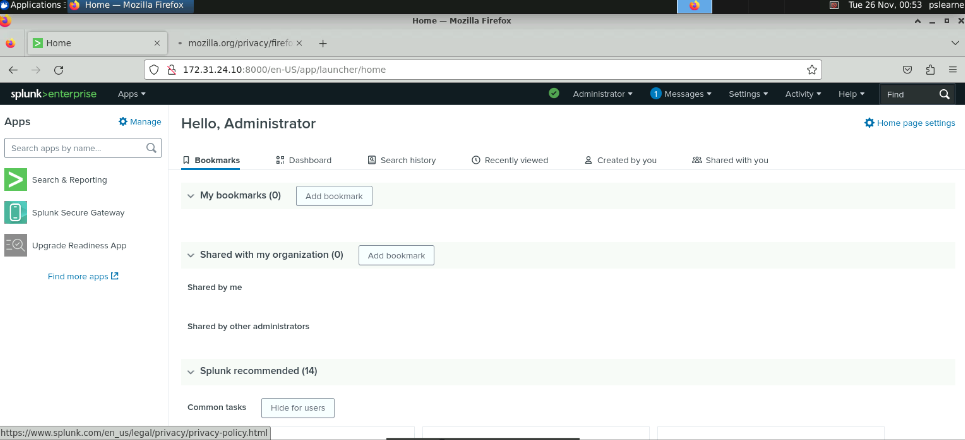
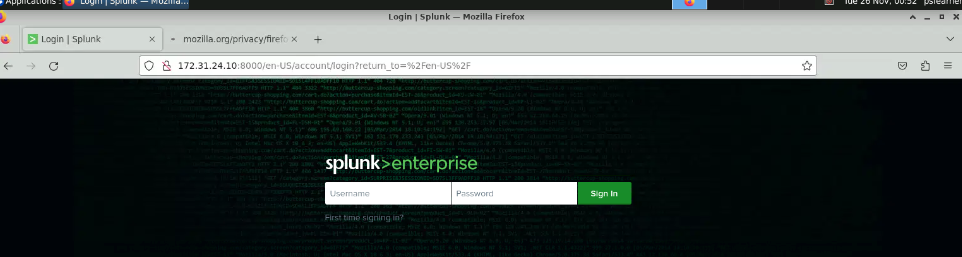
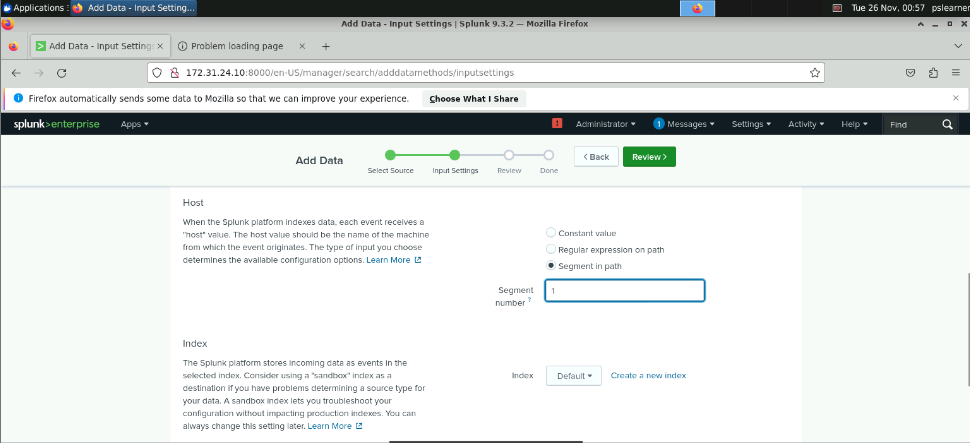
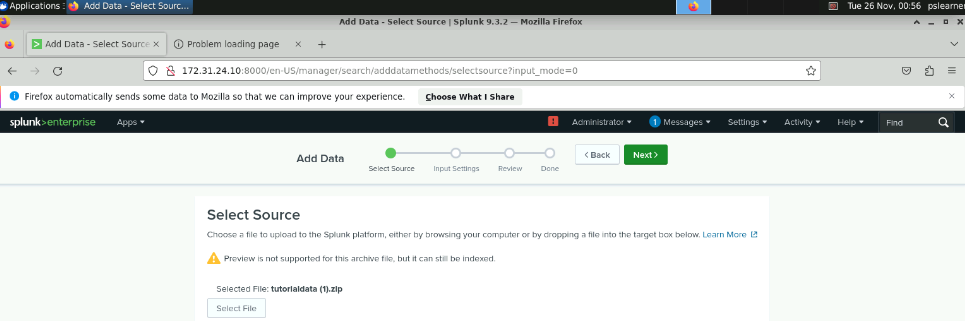
1. **Import Data into Splunk**

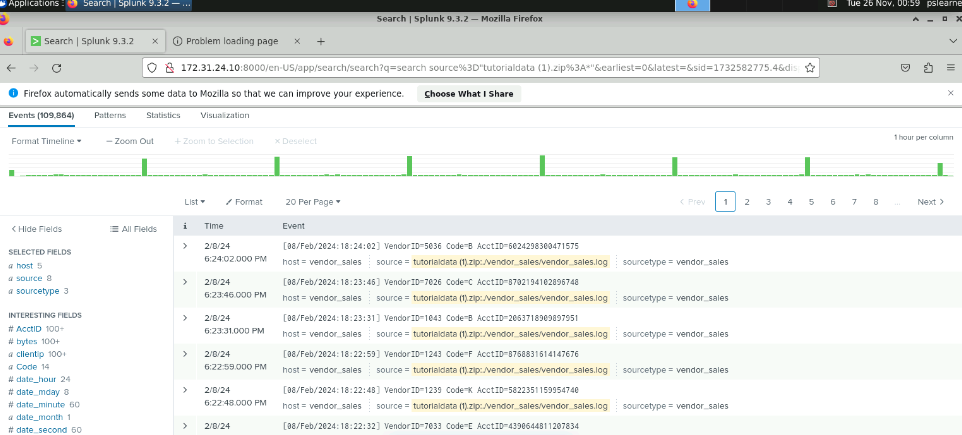
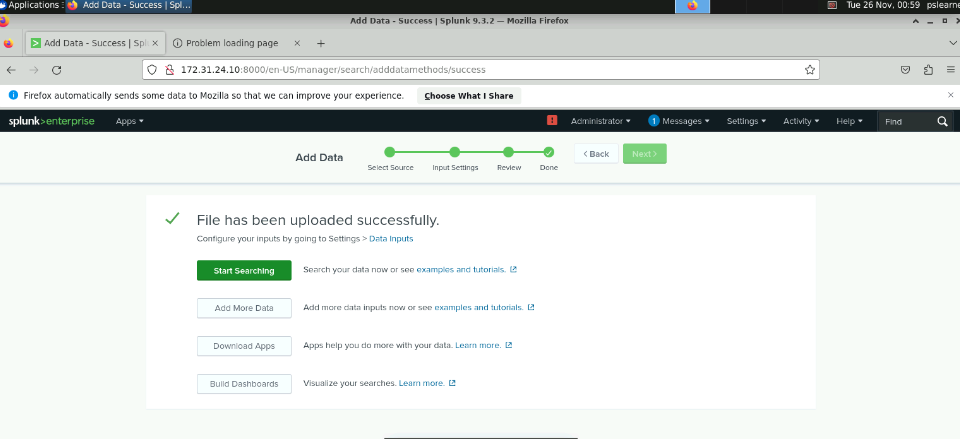
In this lab, I worked with Splunk, a leading security information and event management (SIEM) platform, to analyze log data from a company's web application, Buttercup Games. Acting as a security analyst, my first task was to connect to the Splunk server through the Firefox browser using the IP address and port 172.31.24.10:8000. After logging in with administrator credentials, I accessed the Splunk dashboard to begin the data analysis process. I uploaded the log files related to the web application and prepared the data for further investigation. This hands-on exercise provided me with practical experience in using Splunk to navigate its interface, import data, and prepare for security log analysis, simulating real-world cybersecurity operations.



After logging into Splunk, the next step was to upload log data for analysis. Using the "Add Data" option on the Splunk home page, I selected "Upload from your computer" to import the log files. I navigated to the pslearner directory, opened the lab folder, and selected the tutorialdata.zip file for upload. Once the file was selected, I proceeded to the "Input Settings" page, where I configured the Host settings. Specifically, I changed the Host value to "Segment in path" and set the segment value to 1. These configurations ensured that the data was properly indexed and ready for analysis within Splunk's environment. This process provided valuable experience in importing and configuring data sources for security analysis.



After configuring the Host settings, I proceeded by clicking "Review" to ensure the uploaded data was correctly set up for import. Once confirmed, I clicked "Submit" to complete the upload process. Splunk displayed a confirmation message indicating the file had been successfully uploaded. To verify, I navigated to the "Start Searching" option, which brought me to Splunk's search interface populated with data from the imported tutorialdata.zip file. This allowed me to confirm that the log data was properly indexed and ready for analysis. During this step, I skipped any introductory pop-ups to focus on the imported data and familiarize myself with Splunk's search capabilities.

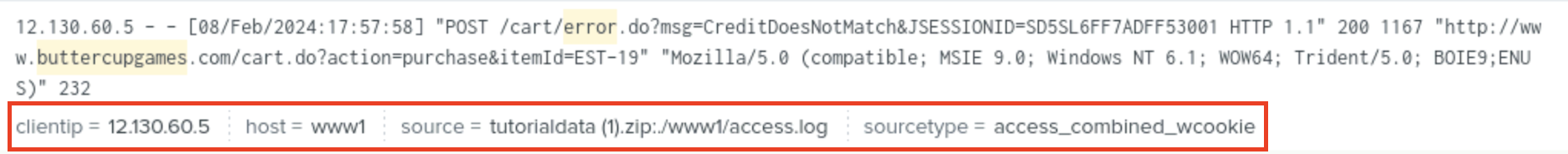
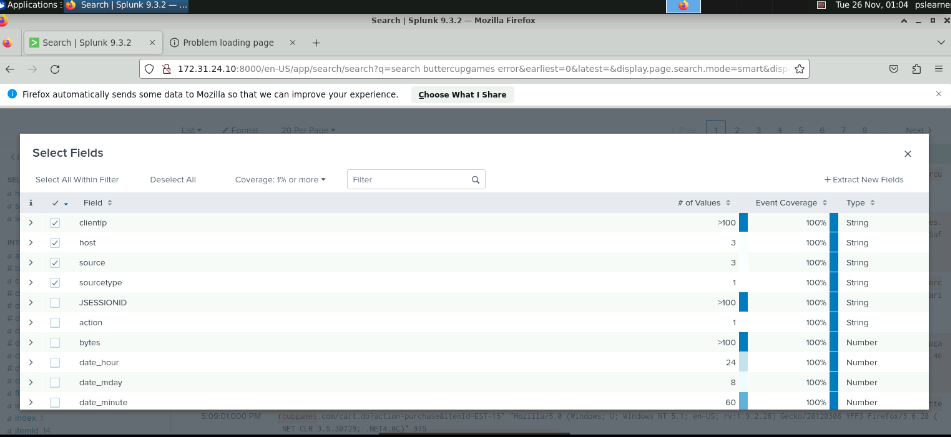


**2. Performing Searches in Splunk**

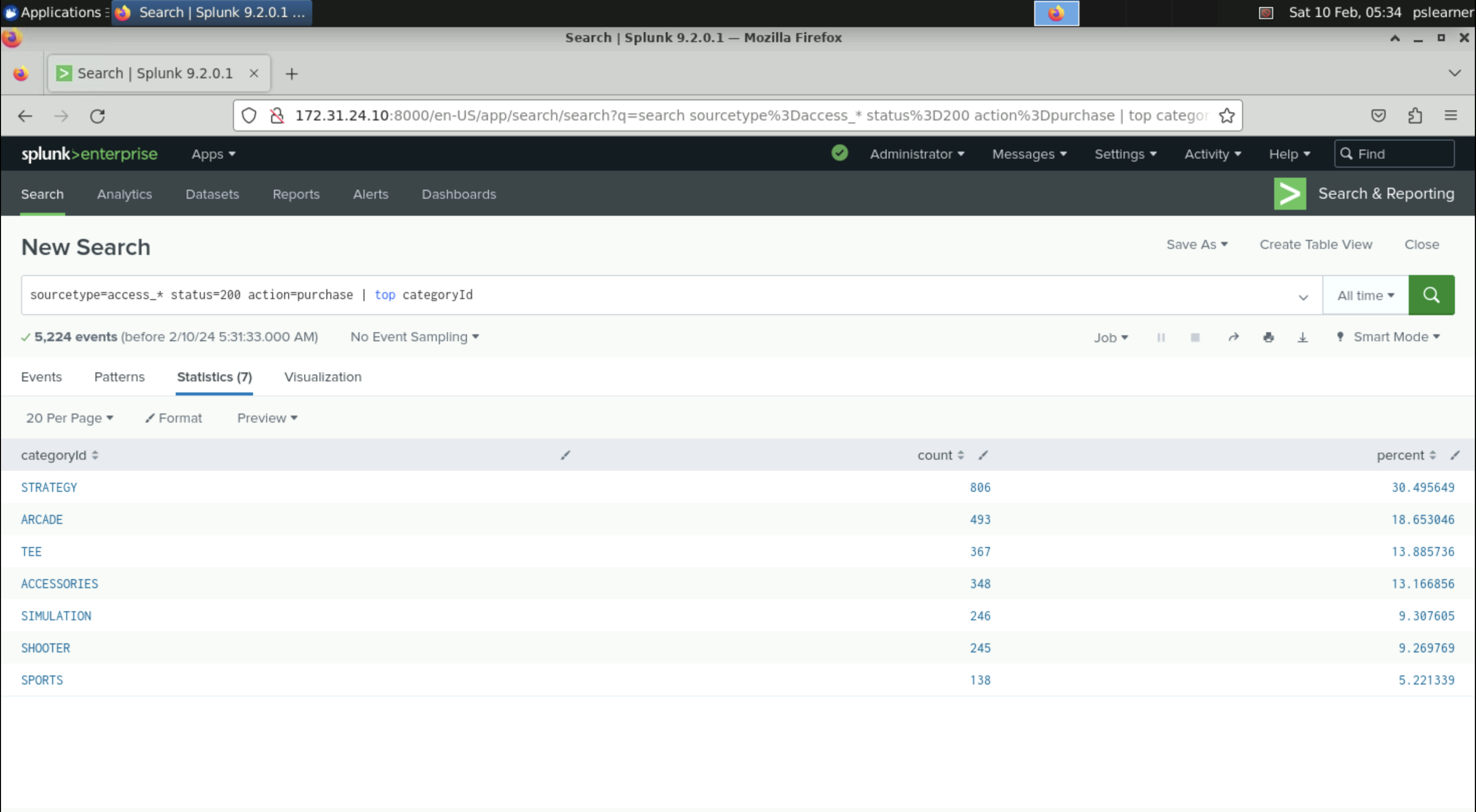
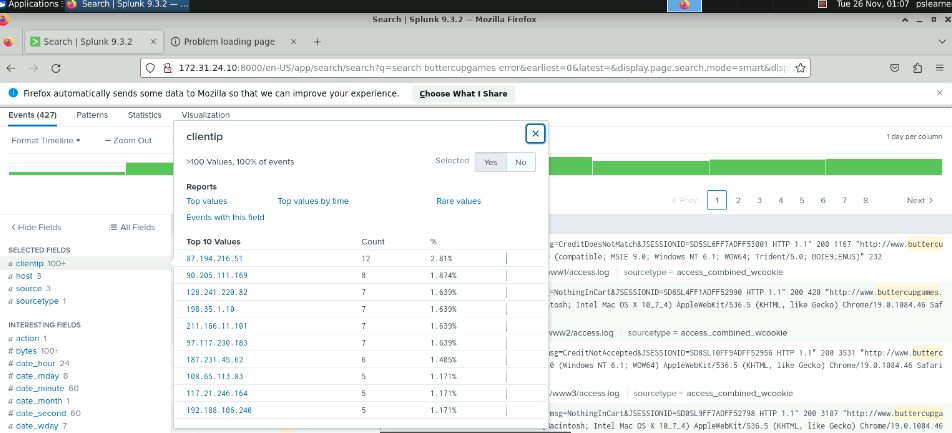
Next, I focused on performing searches within Splunk to extract specific information from the imported dataset. I began by executing a keyword search to identify any error messages related to the "buttercupgames" domain. Using the search term buttercupgames error, I retrieved results matching both keywords, leveraging Splunk's implied AND operator to streamline the query. This allowed me to identify relevant log entries that contained both "buttercupgames" and "error."



To refine my analysis, I expanded my search by incorporating field-based searches. I enabled the clientip field by selecting "All Fields" on the left panel, typing clientip into the search bar, and checking the corresponding box. This action added a new field to the search results, allowing me to view specific client IP addresses associated with the logs. Using field-based searches provided more granular control over the dataset, enabling the identification of precise details within the log data compared to keyword-based searches. This exercise demonstrated the versatility of Splunk's search capabilities, from basic keyword queries to advanced field-based analysis.

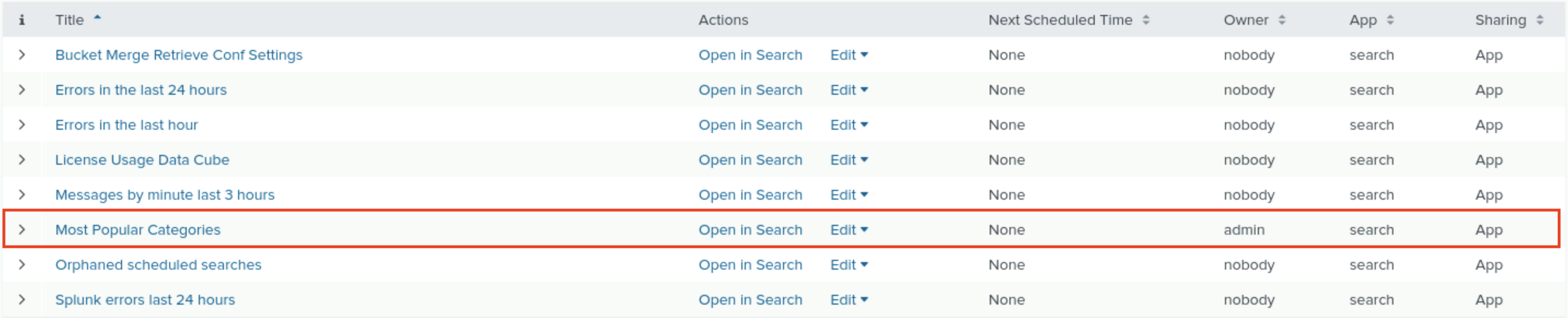
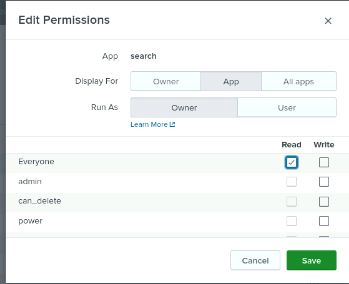
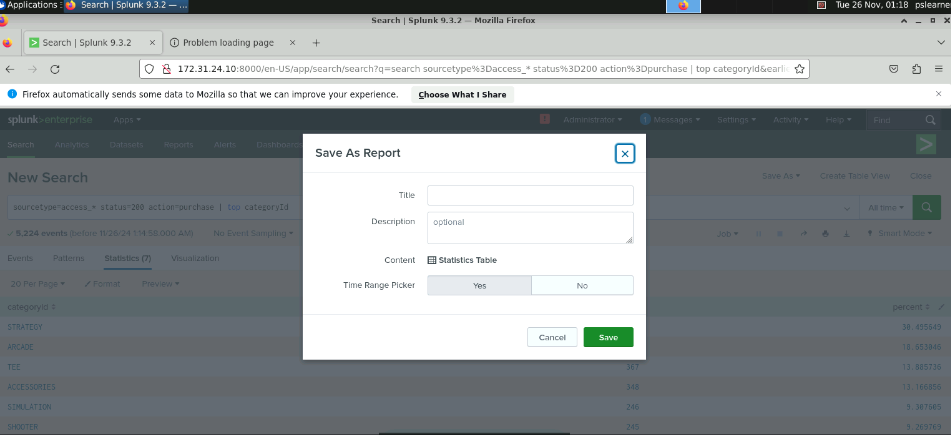


To deepen my analysis, I performed my first field-based search in Splunk. By selecting the clientip field on the left panel, I accessed a breakdown of client IPs associated with errors for the "buttercupgames" domain. This allowed me to identify the frequency of errors generated by each client IP. From the analysis, I observed that the IP address 87.194.216.51 was responsible for the most errors, making it a key focus for further investigation. This exercise highlighted the value of field-based searches in narrowing down specific data points, providing actionable insights into potential issues within the system.



**3. Create and share report**

To optimize data analysis in Splunk, I learned how to create and share reports as well as design data visualizations. Using the results from my previous transformational search, I saved a report by selecting **Save as** > **Report** and adding the title "Most popular categories" with the description "Highest Ranked CategoryIds." After saving the report, I updated its permissions by changing visibility from "Owner" to "App" and enabling "Read" access for Everyone, effectively sharing it across the application. Next, I practiced creating data visualizations by running a new search query: sourcetype=access\_\* status=200 action=purchase | top limit=10 clientip. I adjusted the time range to "All time" and explored Splunk's Visualization tab to generate a bar chart from the output. To refine the visualization, I converted the bar chart into a pie chart, showcasing Splunk's versatility in presenting data insights effectively. This exercise emphasized the importance of saving and visualizing data for streamlined analysis and efficient reporting.



In this step, I practiced creating and saving data visualizations in Splunk to better analyze and present data insights. I started by navigating back to the Search page and running a query to identify the top 10 client IPs associated with successful purchases: sourcetype=access\_\* status=200 action=purchase | top limit=10 clientip. To ensure a comprehensive analysis, I adjusted the search window to "All time" from the default 24-hour range. After generating the output, I accessed the Visualization tab directly beneath the search bar to create a visual representation of the data. Initially presented as a bar chart, I customized the visualization by switching to a pie chart via the Column Chart icon. This step demonstrated the flexibility of Splunk's visualization tools, allowing me to transform raw data into easily interpretable formats for efficient reporting and decision-making.

