

SPLUNK IMPLEMENTATION (STUDENT DATA) CERTIFICATION PROJECT

Neeraj Agarwal
MINDTREE LTD

Problem Statement

Objective: Implement a Splunk project for the Student details. Under this project, you will have to work with the log files of Student data. Following parts should be covered

1. Log Files Creation
2. Data Inputs
3. Fields Extraction
4. Lookups
5. Alerts
6. Report
7. Dashboard

The following are the fields that are to be added in the log files:

1. Date
2. Time
3. Student_ID
4. Student_First_Name
5. Student_Last_Name
6. Student_Father_Name
7. Class

Abstract

This Document includes step by step procedure for monitoring set up of Student data. This includes

- **Creating an index**
- **Adding data sets**
- **Adding lookup tables**
- **Creating reports and alerts**
- **Creating dashboards**

Contents

Creating an index.....	4
Adding data inputs.....	5
Basic searches (Fields Extraction)	6
Lookups	11
Adding an alert.....	14
Creating Reports.....	16
Creating A Dashboard.....	20
Conclusion	21

Creating an index

First I have created an index specifically for monitoring the Student data.

Name	Actions	Type	App	Current Size	Max Size	Event Count	Earliest Event	Latest Event	Home Path	Frozen Path	Status
_audit	Edit Delete Disable	Events	system	7 MB	488.28 GB	52.9K	10 days ago	a few seconds ago	\$SPLUNK_DB/audit/db	N/A	✓ Enabled
_internal	Edit Delete Disable	Events	system	109 MB	488.28 GB	1.05M	10 days ago	4 days ago	\$SPLUNK_DB/_internal/db/db	N/A	✓ Enabled
_introspection	Edit Delete Disable	Events	system	183 MB	488.28 GB	158K	10 days ago	4 days ago	\$SPLUNK_DB/_introspection/db	N/A	✓ Enabled
_telemetry	Edit Delete Disable	Events	system	1MB	488.28 GB	2	6 days ago	6 days ago	\$SPLUNK_DB/_telemetry/db	N/A	✓ Enabled
_thefishbuckete	Edit Delete Disable	Events	system	1MB	488.28 GB	0			\$SPLUNK_DB/fishbucket/db	N/A	✓ Enabled
emp_data	Edit Delete Disable	Events	search	1MB	500 GB	300	4 days ago	4 days ago	\$SPLUNK_DB/emp_data/db	N/A	✓ Enabled
game_web	Edit Delete Disable	Events	search	53 MB	10 GB	659K	a month ago	a month ago	\$SPLUNK_DB/game_web/db	N/A	✓ Enabled
history	Edit Delete Disable	Events	system	1MB	488.28 GB	0			\$SPLUNK_DB/history/db	N/A	✓ Enabled
main	Edit Delete Disable	Events	system	1MB	488.28 GB	0			\$SPLUNK_DB/default/db	N/A	✓ Enabled
splunklogger	Edit Delete Enable	Events	system	0 B	488.28 GB	0			\$SPLUNK_DB/splunklogger/db	N/A	✗ Disabled
student_data	Edit Delete Disable	Events	search	1MB	500 GB	100	4 days ago	4 days ago	\$SPLUNK_DB/student_data/db	N/A	✓ Enabled
summary	Edit Delete Disable	Events	system	1MB	488.28 GB	0			\$SPLUNK_DB/summary/db	N/A	✓ Enabled

Adding data inputs

I have created a static csv file which contains details of 100 Students to add as a data set. I will add this file to monitor under the index created previously.

Configure this instance to monitor files and directories for data. To monitor all objects in a directory, select the directory. The Splunk platform monitors and assigns a single source type to all objects within the directory. This might cause problems if there are different object types or data sources in the directory. To assign multiple source types to objects in the same directory, configure individual data inputs for those objects. [Learn More](#)

File or Directory: /opt/splunk/Student_Data.csv

Continuously Monitor | Index Once

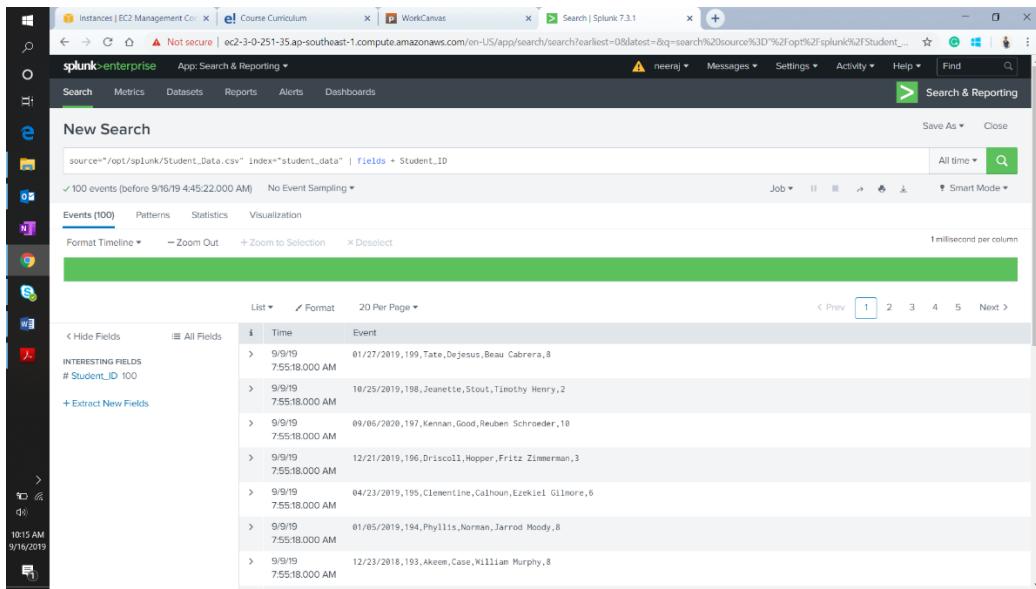
FAQ

- What kinds of files can the Splunk platform index?
- I can't access the file that I want to index. Why?
- How do I get remote data onto my Splunk platform instance?
- Can I monitor changes to files in addition to their content?
- What is a source type?
- How do I specify a whitelist or blacklist for a directory?

Full path to your data	Set host	Source type	Index	Number of files	App	Status	Actions
\$/SPLUNK_HOME/etc/splunk/version	Constant Value	splunk_version	_internal		system	Disabled Enable	
\$/SPLUNK_HOME/var/log/introspection	Constant Value	Automatic	_introspection		introspection_generator_addon	Disabled Enable	
\$/SPLUNK_HOME/var/log/splunk	Constant Value	Automatic	_internal		system	Disabled Enable	
\$/SPLUNK_HOME/var/log/splunk/license_usage_summary.log	Constant Value	Automatic	_telemetry		system	Disabled Enable	
\$/SPLUNK_HOME/var/log/watchdog/watchdog.log	Constant Value	Automatic	_internal		system	Disabled Enable	
\$/SPLUNK_HOME/var/run/splunk/search_telemetry/search_telemetry.json	Constant Value	search_telemetry	_introspection		system	Disabled Enable	
\$/SPLUNK_HOME/var/spool/splunk	Constant Value	Automatic	default		system	Disabled Enable	
\$/SPLUNK_HOME/var/spool/splunk/_stash_new	Constant Value	stash_new	default		system	Disabled Enable	
/opt/splunk/emp_data.csv	Segment	csv	emp_data	1	search	Enabled Disable Delete	
/opt/splunk/Student_Data.csv	Segment	csv	student_data	1	search	Enabled Disable Delete	
/opt/splunk/tutorialdata.zip	Segment	Automatic	game_web	1	search	Enabled Disable Delete	

Basic searches (Fields Extraction)

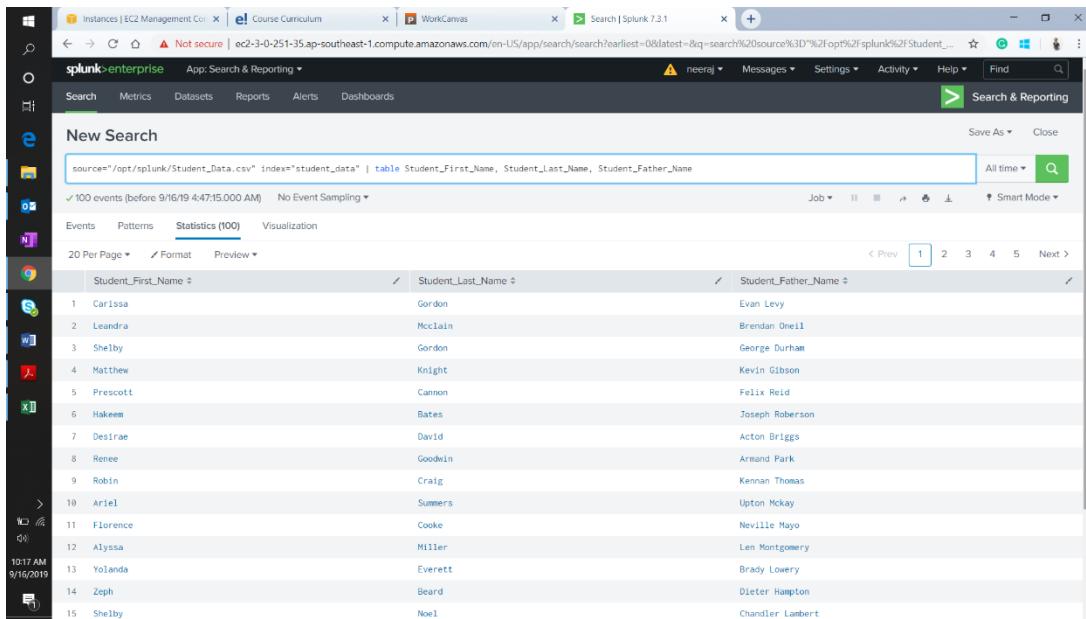
In this section, I have attached a few screenshots of the searches using various fields and commands.



The screenshot shows the Splunk 7.3.1 interface with a search titled "New Search". The search bar contains the command: `source="/opt/splunk/Student_Data.csv" index="student_data" | fields + Student_ID`. The results table shows 100 events, all from 9/9/19 at 7:55:18 AM. The table has three columns: Time, Event, and Student_ID. The data is as follows:

Time	Event	Student_ID
9/9/19 7:55:18 AM	01/27/2019,199,Tate,Dejesus,Beau Cabrera,8	
9/9/19 7:55:18 AM	10/25/2019,198,Jeannette,Stout,Timothy Henry,2	
9/9/19 7:55:18 AM	09/06/2020,197,Kennan,Good,Reuben Schroeder,10	
9/9/19 7:55:18 AM	12/21/2019,196,Driscoll,Hopper,Fritz Zimmerman,3	
9/9/19 7:55:18 AM	04/23/2019,195,Clementine,Calhoun,Ezekiel Gilmore,6	
9/9/19 7:55:18 AM	01/05/2019,194,Phyllis,Norman,Jarrod Moody,8	
9/9/19 7:55:18 AM	12/23/2018,193,Akeem,Case,William Murphy,8	

Field Extraction



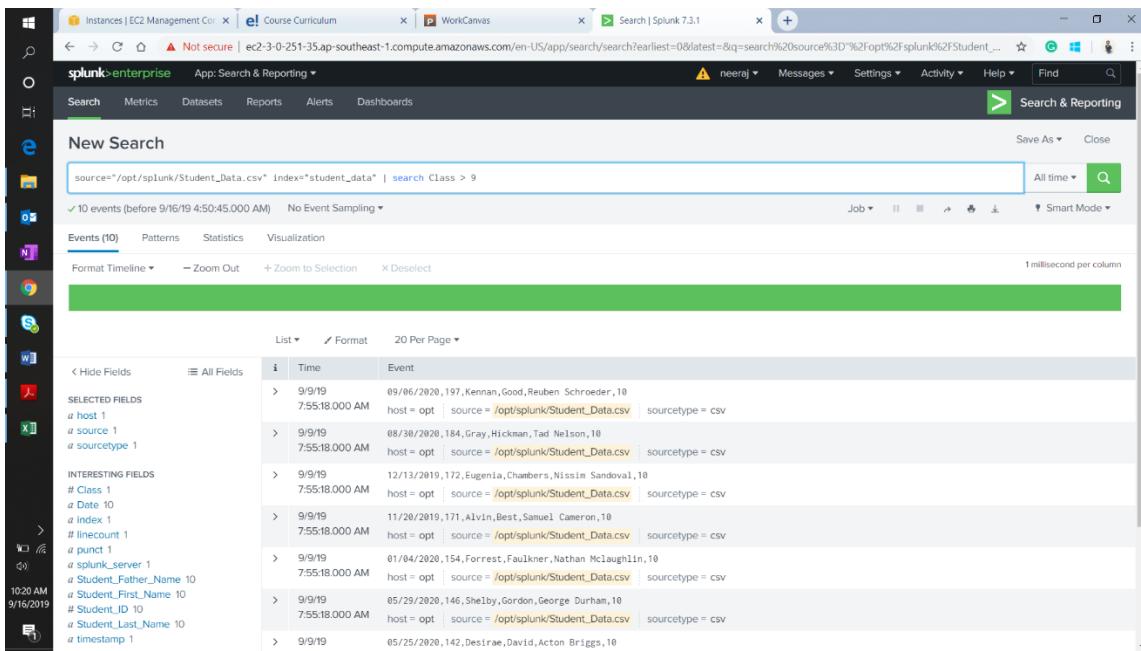
The screenshot shows the Splunk 7.3.1 interface with a search titled "New Search". The search bar contains the command: `source="/opt/splunk/Student_Data.csv" index="student_data" | table Student_First_Name, Student_Last_Name, Student_Father_Name`. The results table shows 15 events, all from 9/9/19 at 7:55:18 AM. The table has three columns: Student_First_Name, Student_Last_Name, and Student_Father_Name. The data is as follows:

Student_First_Name	Student_Last_Name	Student_Father_Name
Carissa	Gordon	Evan Levy
Leandra	McClain	Brendan O'Neill
Shelby	Gordon	George Durham
Matthew	Knight	Kevin Gibson
Prescott	Cannon	Felix Reid
Hakeem	Bates	Joseph Roberson
Desirae	David	Acton Briggs
Renee	Goodwin	Armand Park
Robin	Craig	Kennan Thomas
Ariel	Summers	Upton McKay
Florence	Cooke	Neville Mayo
Alyssa	Miller	Len Montgomery
Yolanda	Everett	Brady Lowery
Zeph	Beard	Dieter Hampton
Shelby	Noel	Chandler Lambert

Example of “Table” command

Example of “Sort” command

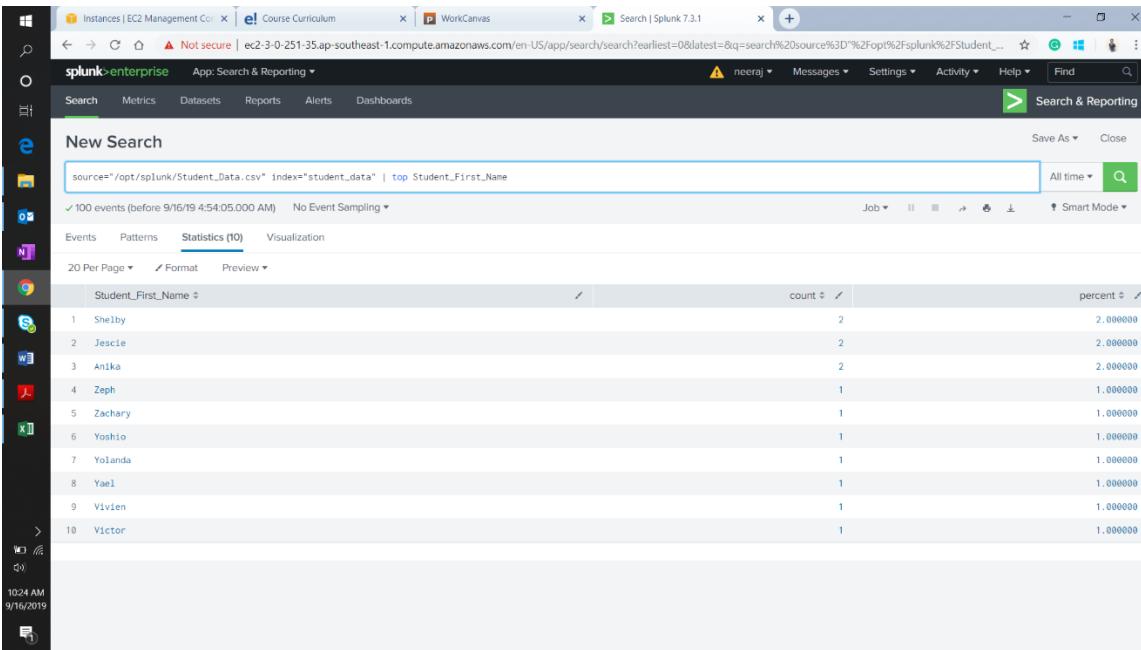
Example of “Rename” command



The screenshot shows the Splunk 7.3.1 interface with a search bar containing the command: `source="/opt/splunk/Student_Data.csv" index="student_data" | search Class > 9`. The results table displays 10 events, each with a timestamp, source, and sourcetype. The table includes columns for Time, Event, and sourcetype. The sourcetype is consistently listed as 'csv'.

Time	Event	sourcetype
09/06/2020, 197, Kennan, Good, Reuben Schroeder, 10	host = opt source = /opt/splunk/Student_Data.csv sourcetype = csv	csv
09/09/2020, 184, Gray, Hickman, Tad Nelson, 10	host = opt source = /opt/splunk/Student_Data.csv sourcetype = csv	csv
12/13/2019, 172, Eugenia, Chambers, Nissim Sandoval, 10	host = opt source = /opt/splunk/Student_Data.csv sourcetype = csv	csv
11/20/2019, 171, Alvin, Best, Samuel Cameron, 10	host = opt source = /opt/splunk/Student_Data.csv sourcetype = csv	csv
01/04/2020, 154, Forrest, Faulkner, Nathan McLaughlin, 10	host = opt source = /opt/splunk/Student_Data.csv sourcetype = csv	csv
05/29/2020, 146, Shelby, Gordon, George Durham, 10	host = opt source = /opt/splunk/Student_Data.csv sourcetype = csv	csv
05/23/2020, 142, Desiree, David, Acton Briggs, 10	host = opt source = /opt/splunk/Student_Data.csv sourcetype = csv	csv
05/25/2020, 142, Desiree, David, Acton Briggs, 10	host = opt source = /opt/splunk/Student_Data.csv sourcetype = csv	csv
09/09/2019	05/25/2020, 142, Desiree, David, Acton Briggs, 10	csv

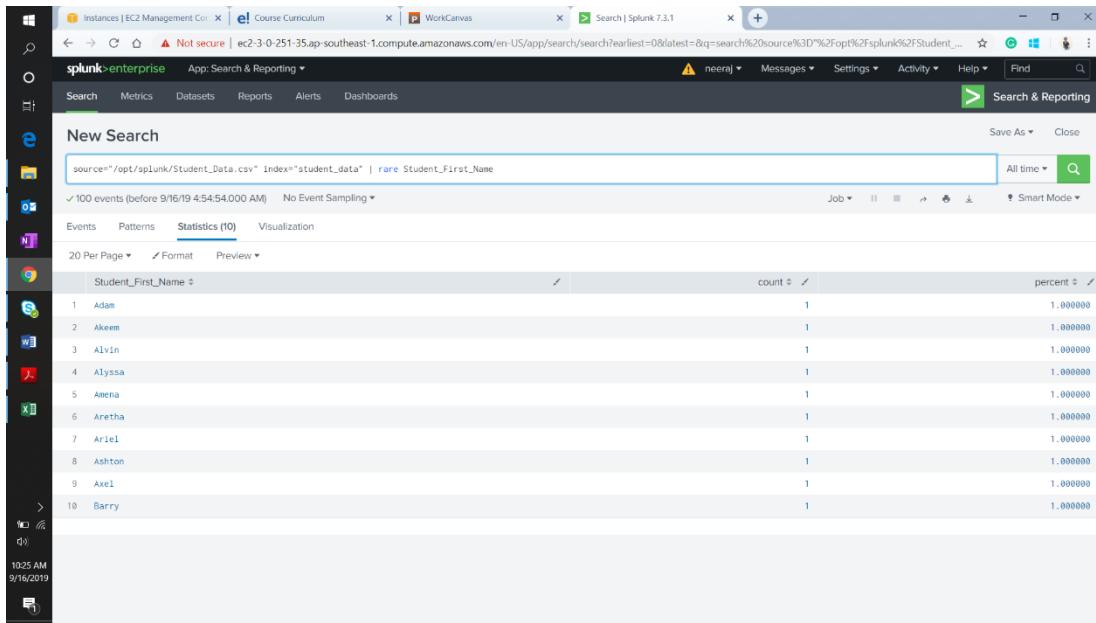
Example of “search” command



The screenshot shows the Splunk 7.3.1 interface with a search bar containing the command: `source="/opt/splunk/Student_Data.csv" index="student_data" | top Student_First_Name`. The results table displays 10 events, each with a count and percent. The table includes columns for Student_First_Name, count, and percent.

Student_First_Name	count	percent
Shelby	2	2.000000
Jescie	2	2.000000
Anika	2	2.000000
Zeph	1	1.000000
Zachary	1	1.000000
Yoshio	1	1.000000
Yolanda	1	1.000000
Yael	1	1.000000
Vivien	1	1.000000
Victor	1	1.000000

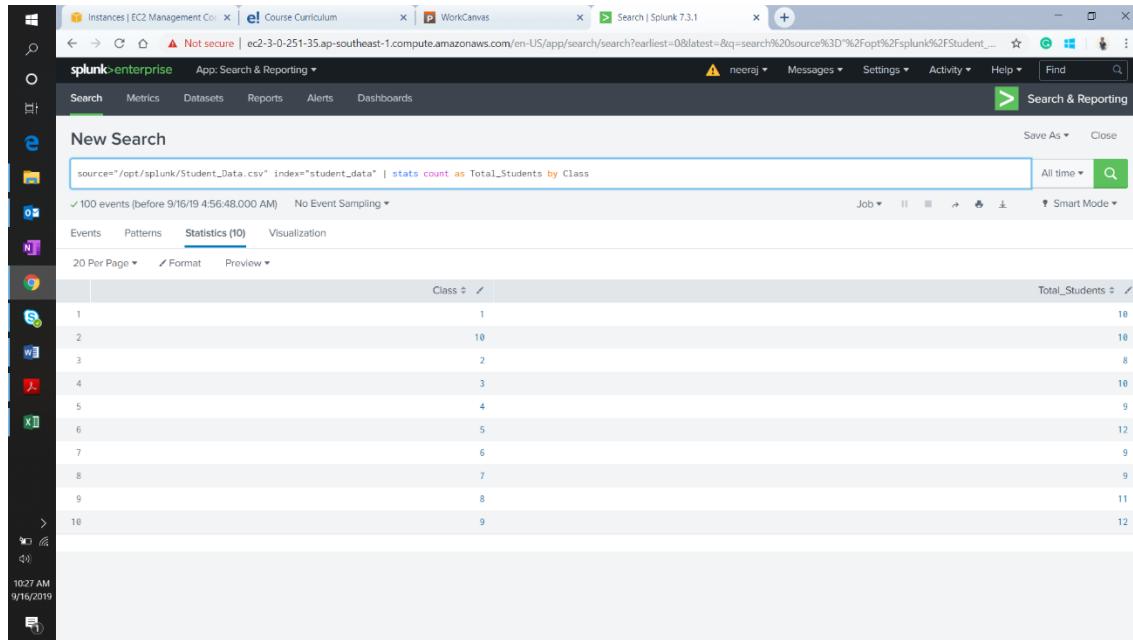
Example of “Top” command



The screenshot shows the Splunk 7.3.1 interface with a search bar containing the command: `source="/opt/splunk/Student_Data.csv" index="student_data" | rare Student_First_Name`. The results table displays 100 events, with the first 10 rows listed as follows:

Student_First_Name	count	percent
Adam	1	1.000000
Akeem	1	1.000000
Alvin	1	1.000000
Alyssa	1	1.000000
Amena	1	1.000000
Aretha	1	1.000000
Ariel	1	1.000000
Ashton	1	1.000000
Axel	1	1.000000
Barry	1	1.000000

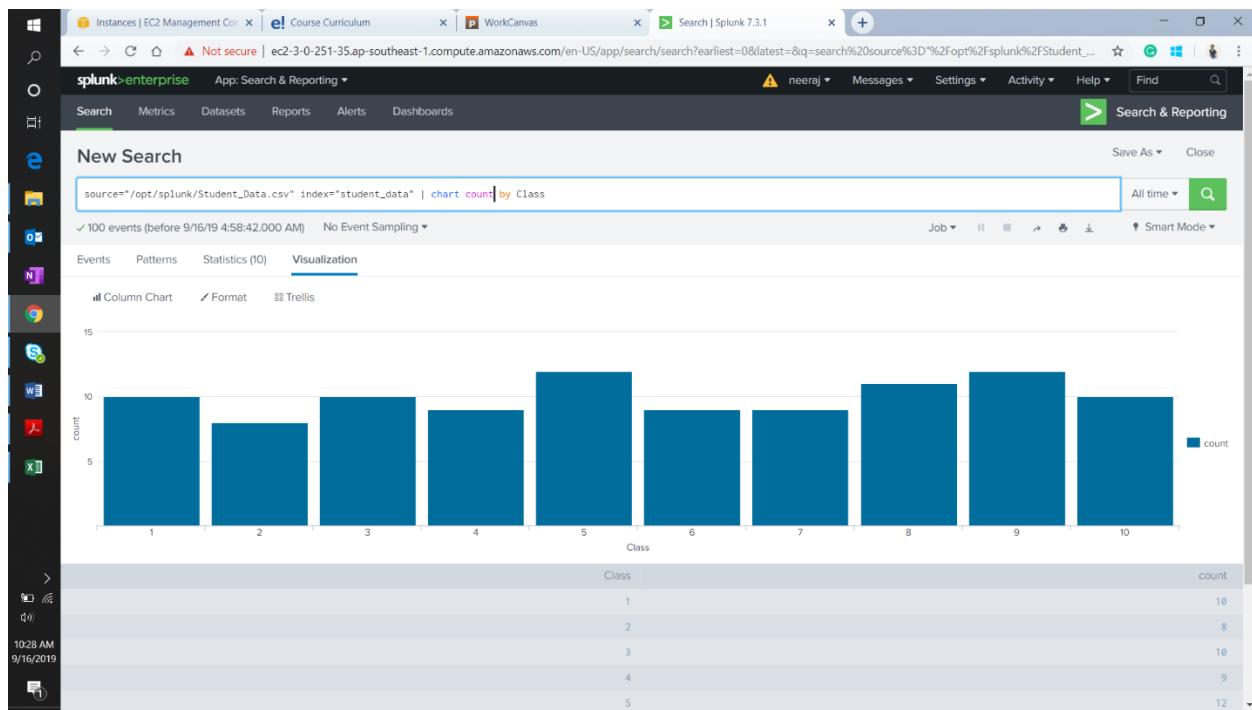
Example of “Rare” command



The screenshot shows the Splunk 7.3.1 interface with a search bar containing the command: `source="/opt/splunk/Student_Data.csv" index="student_data" | stats count as Total_Students by Class`. The results table displays 100 events, with the first 10 rows listed as follows:

Class	Total_Students
1	10
2	10
3	8
4	10
5	9
6	12
7	9
8	9
9	11
10	12

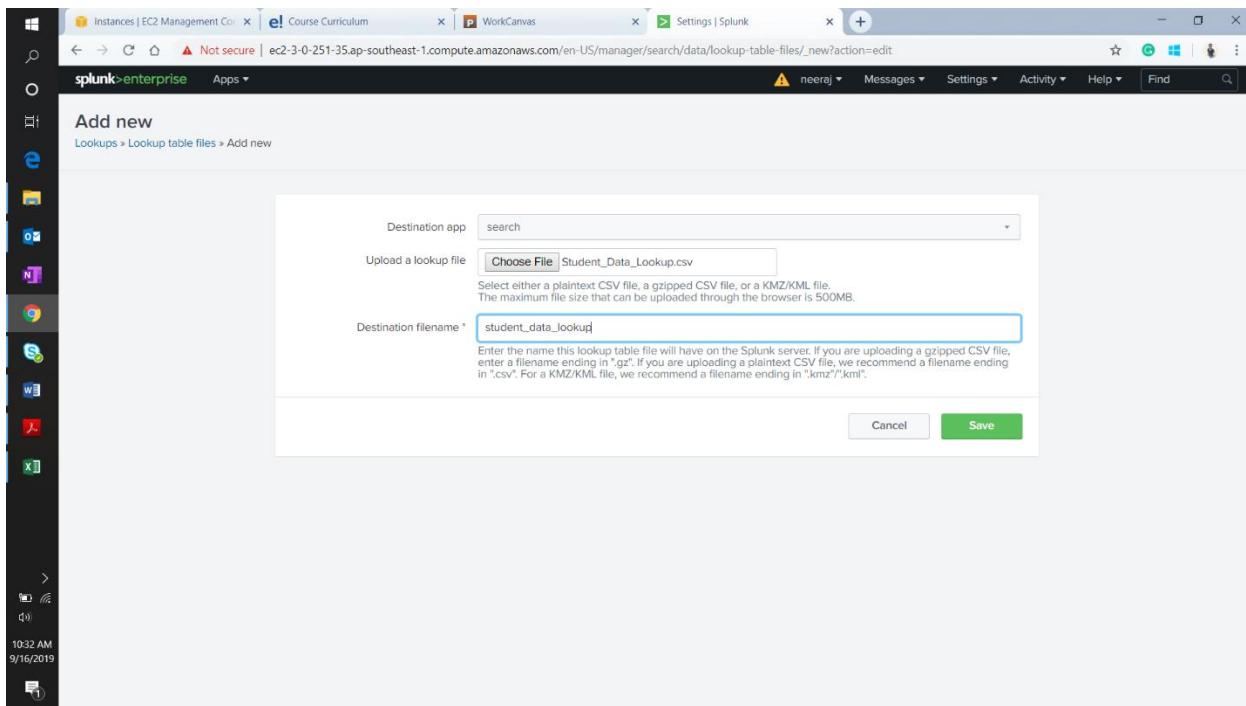
Example of “Stats” command



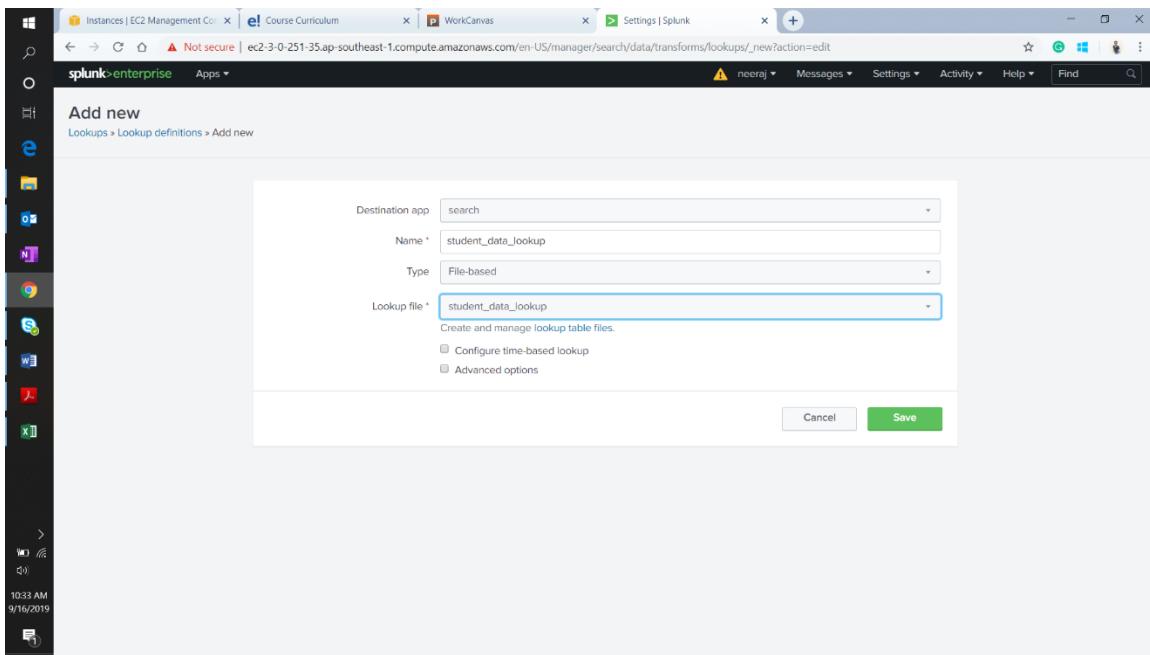
Example of “Chart” command

Lookups

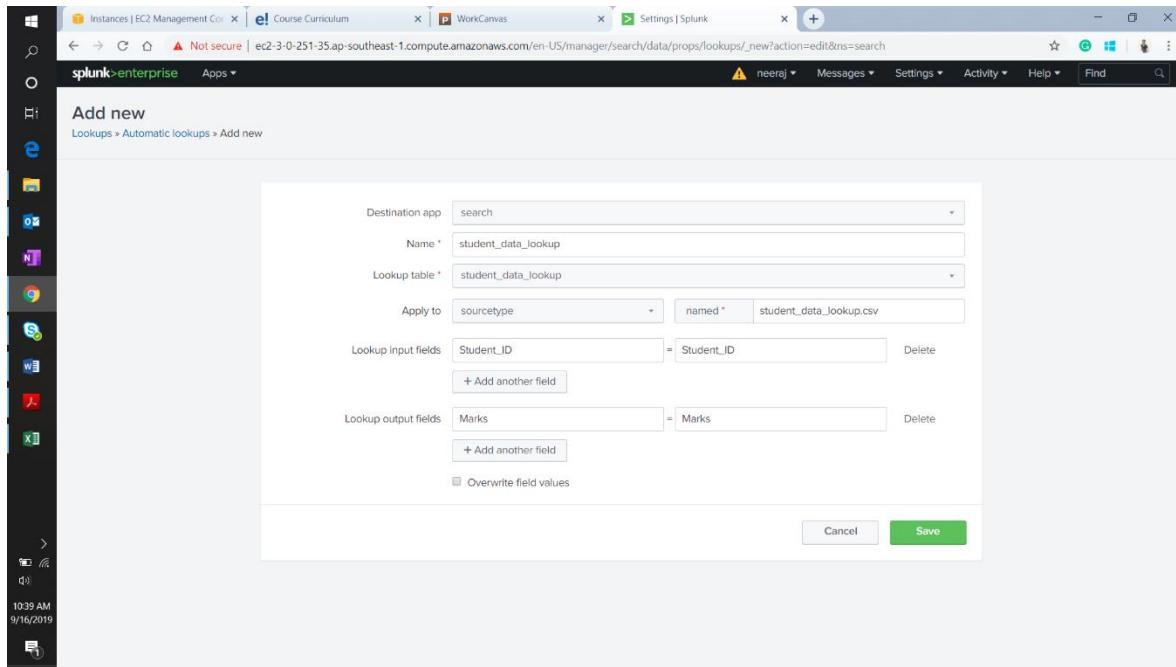
In this section, I will add a lookup table to map the Student's job title to the company they are associated with. I have created a csv file with above-mentioned data.



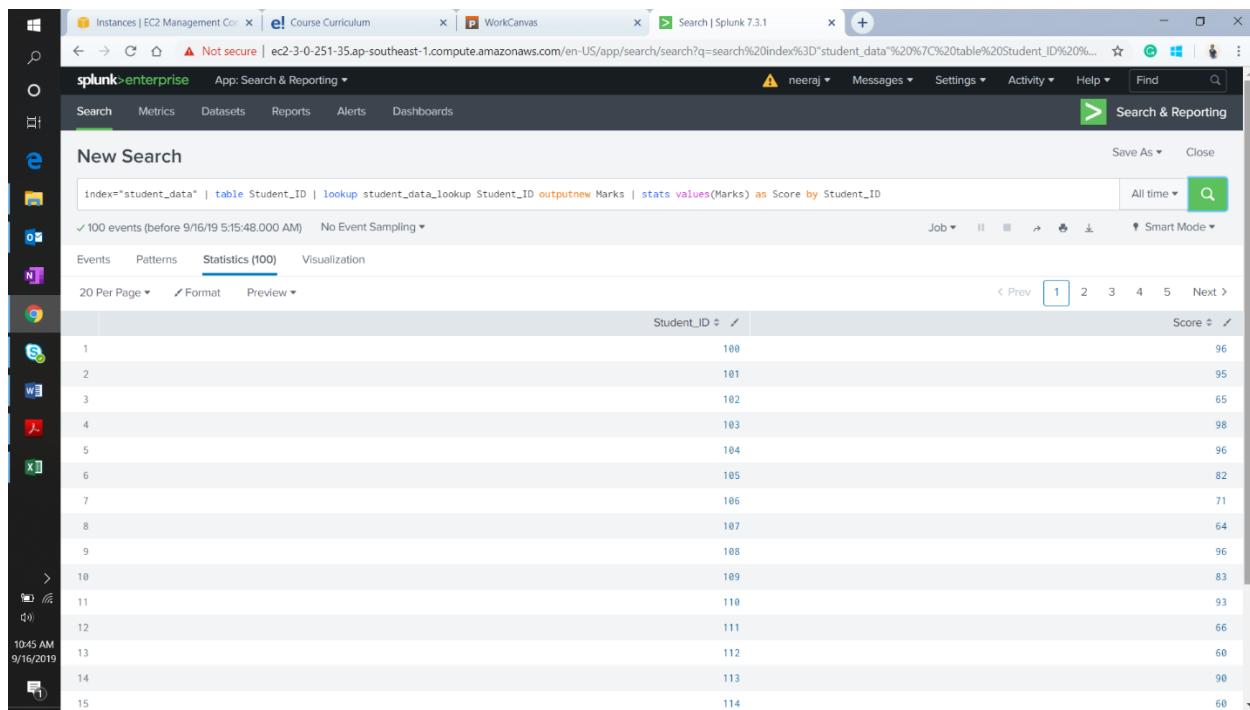
Adding the lookup file



Adding a new lookup definition



Adding automatic lookup



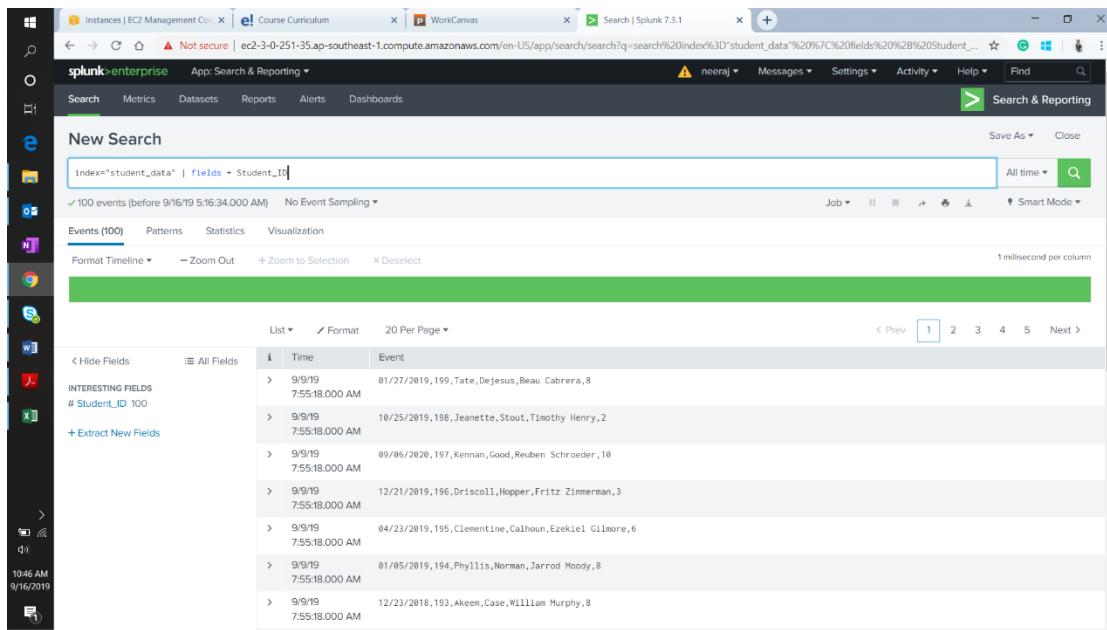
The screenshot shows the Splunk 7.3.1 interface with a search results table. The search bar contains the following command: `index="student_data" | table Student_ID | lookup student_data_lookup Student_ID outputnew Marks | stats values(Marks) as Score by Student_ID`. The results table has two columns: `Student_ID` and `Score`. The data is as follows:

Student_ID	Score
100	96
101	95
102	65
103	98
104	96
105	82
106	71
107	64
108	96
109	83
110	93
111	66
112	60
113	90
114	60

Example of a search which utilizes lookup

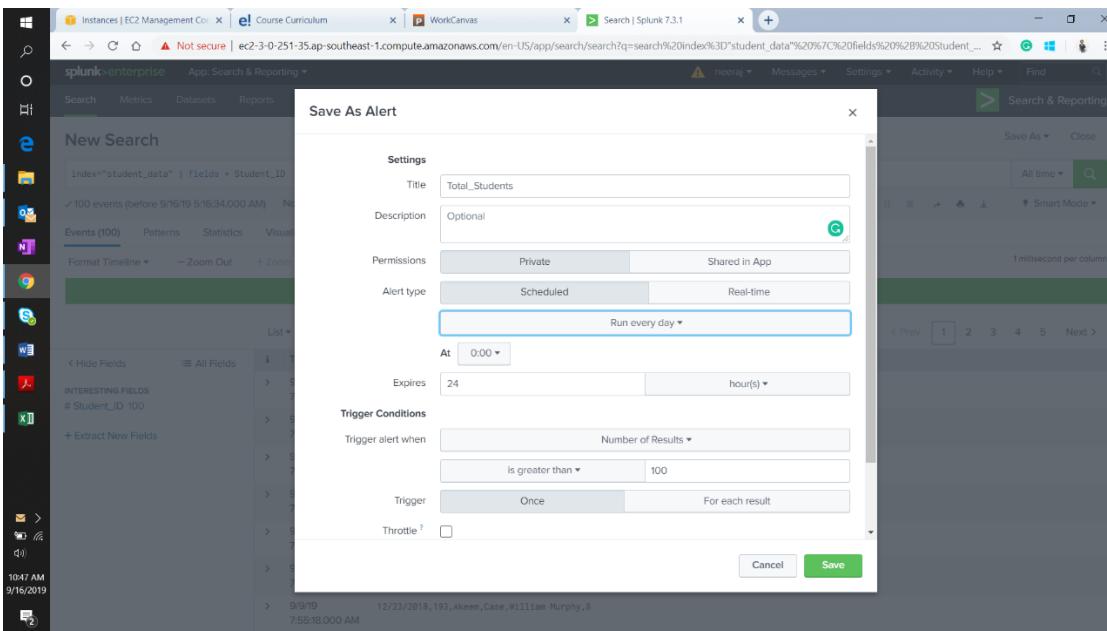
Adding an alert

Here I am configuring an alert where if the search returns showing more than 100 Students in the organization data. This search runs every day at midnight and raises an alert, as well as outputs, result in lookup.



The screenshot shows the Splunk interface with a search bar containing the query `index="student_data" | fields + Student_ID`. The results table shows 100 events, all of which have the `# Student_ID` field set to 100. The table includes columns for Time and Event. The interface includes a sidebar with various icons and a timestamp of 10:46 AM on 9/16/2019.

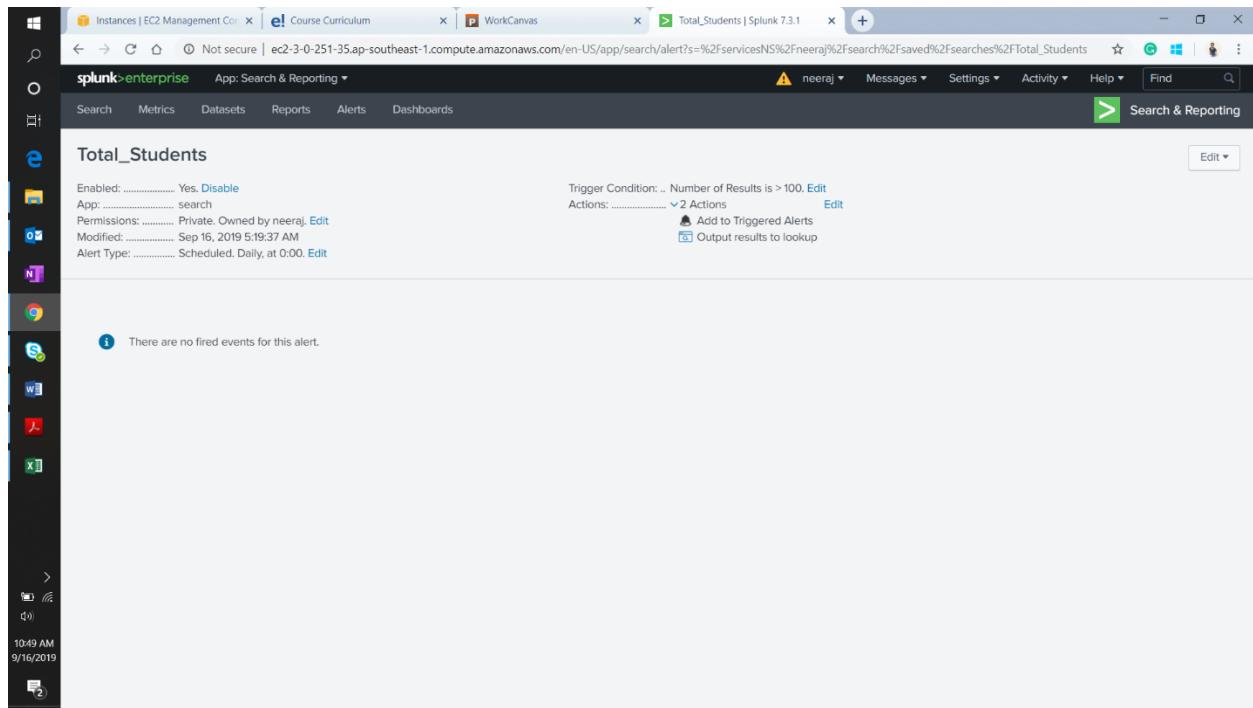
The search string to be configured as alert



The screenshot shows the Splunk interface with the 'Save As Alert' dialog box open. The dialog box is titled 'Save As Alert' and contains the following settings:

- Settings**:
 - Title: Total_Students
 - Description: Optional
 - Permissions: Private
 - Alert type: Scheduled
 - Run every day
 - At: 0:00
 - Expires: 24 hour(s)
- Trigger Conditions**:
 - Trigger alert when: Number of Results
 - is greater than: 100
 - Trigger: Once
 - For each result
 - Throttle:

At the bottom of the dialog box are 'Cancel' and 'Save' buttons.



The screenshot shows the Splunk interface with the 'Total_Students' alert configuration. The alert is enabled and set to search daily at 0:00. It triggers when the number of results is greater than 100. Two actions are defined: 'Add to Triggered Alerts' and 'Output results to lookup'. A note indicates there are no fired events for this alert. The Splunk sidebar on the left shows various monitoring and search icons.

The alert configured

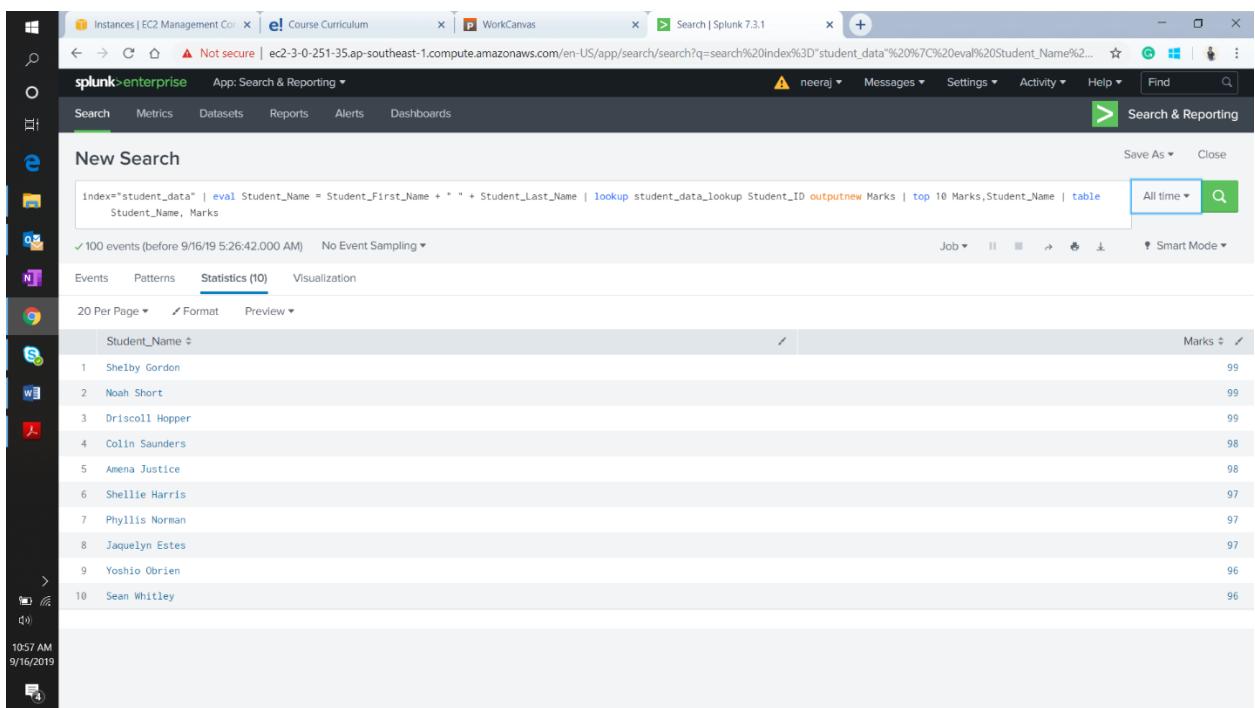
Creating Reports

For this project, I have created four reports which will be used to create a dashboard in the next section. Below is the list of the reports

1. Top 10 scoring Students

Search string :

```
index="student_data" | eval Student_Name =  
Student_First_Name + " " + Student_Last_Name | lookup  
student_data_lookup Student_ID outputnew Marks | top 10  
Marks,Student_Name | table Student_Name, Marks
```

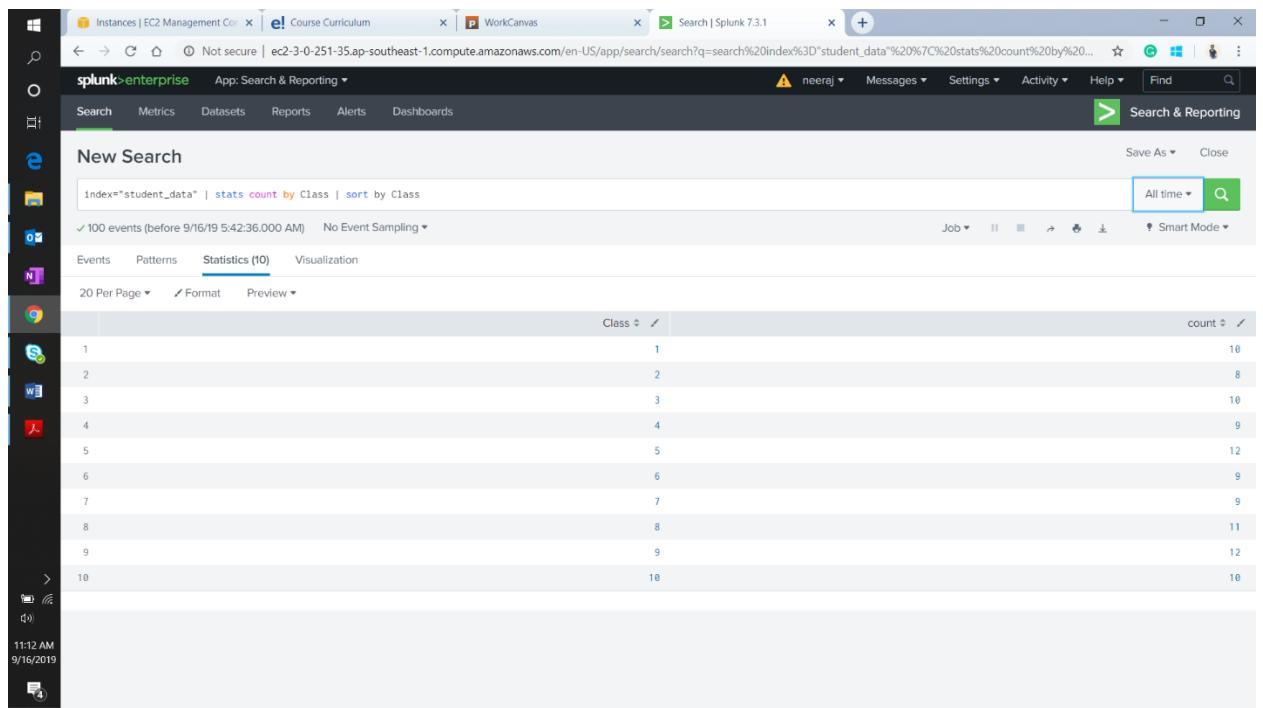


Student_Name	Marks
Shelby Gordon	99
Noah Short	99
Driscoll Hopper	99
Colin Saunders	98
Amena Justice	98
Shellie Harris	97
Phyllis Norman	97
Jaquelyn Estes	97
Yoshio Obrien	96
Sean Whitley	96

2. Total Students in each Class

Search string :

```
index="student_data" | stats count by Class | sort by Class
```



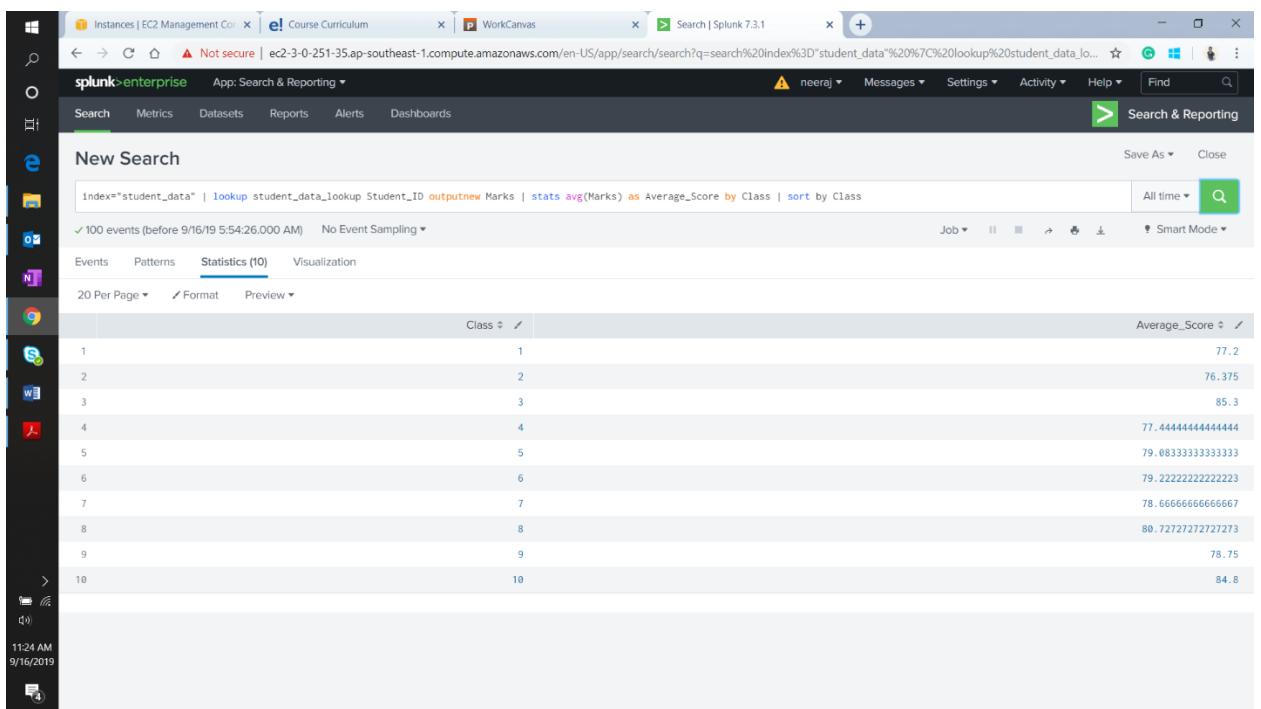
The screenshot shows the Splunk Enterprise search interface. The search bar contains the query: `index="student_data" | stats count by Class | sort by Class`. The results table shows the following data:

Class	count
1	10
2	8
3	10
4	9
5	12
6	9
7	9
8	11
9	12
10	10

3. Average Marks in each Class

Search string :

```
index="student_data" | lookup student_data_lookup Student_ID  
outputnew Marks | stats avg(Marks) as Average_Score by Class  
| sort by Class
```



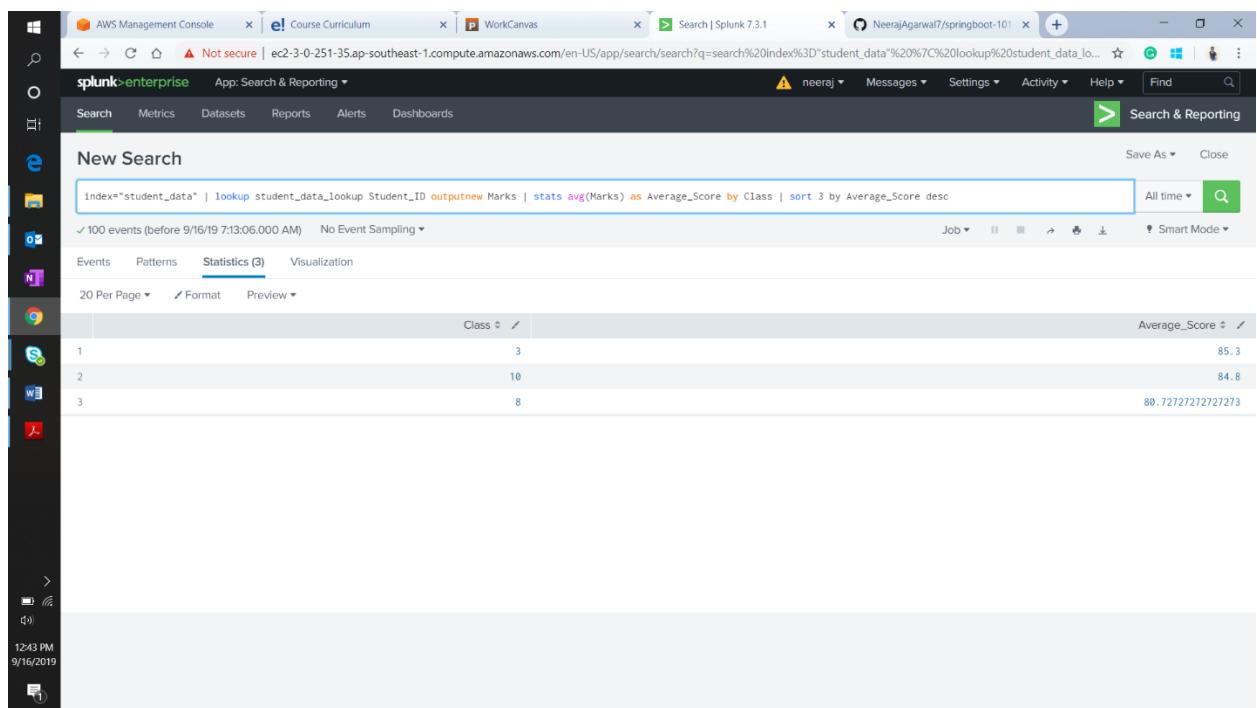
The screenshot shows the Splunk interface with the search bar containing the command: `index="student_data" | lookup student_data_lookup Student_ID outputnew Marks | stats avg(Marks) as Average_Score by Class | sort by Class`. The results table displays 10 rows of data, each representing a class and its average score:

Class	Average_Score
1	77.2
2	76.375
3	85.3
4	77.444444444444
5	79.083333333333
6	79.222222222223
7	78.66666666666667
8	80.727272727273
9	78.75
10	84.8

4. Top 3 performing class

Search String :

```
index="student_data" | lookup student_data_lookup Student_ID  
outputnew Marks | stats avg(Marks) as Average_Score by Class  
| sort 3 by Average_Score desc
```

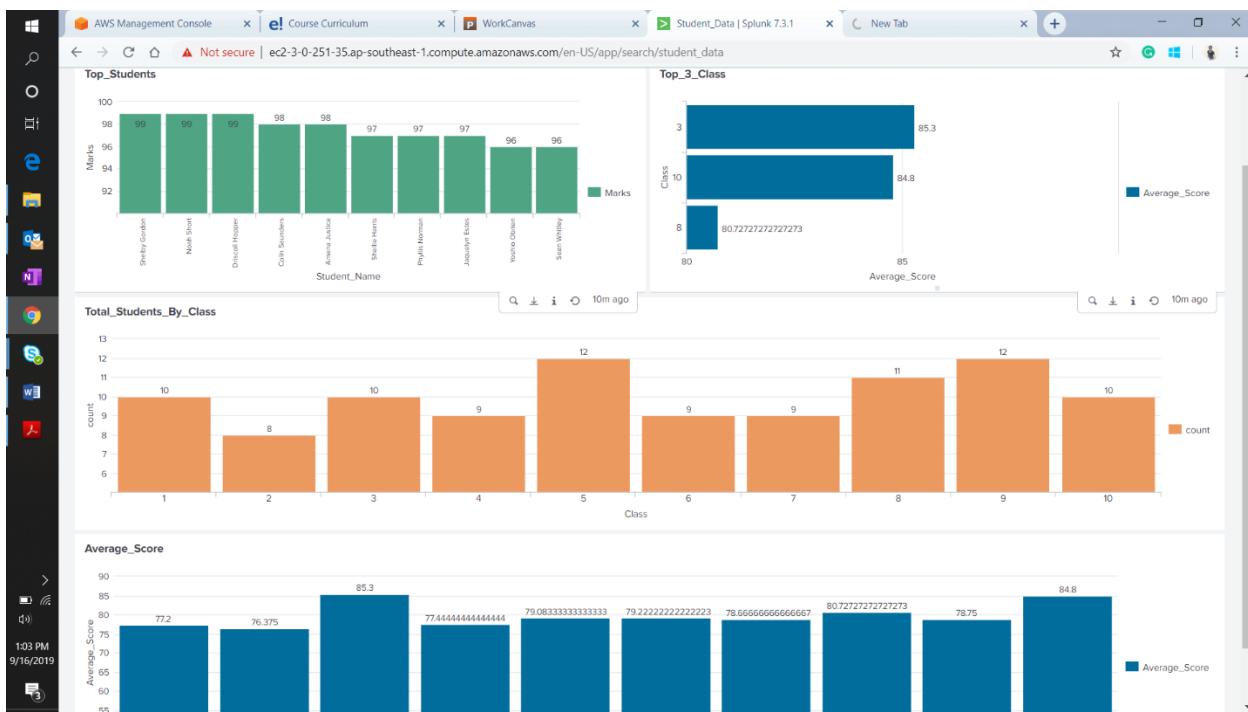


Class	Average_Score
1	85.3
2	84.8
3	80.72727272727273

Creating A Dashboard

Using the reports mentioned in the previous section I will create a simple dashboard consisting of bar charts to visualize the results of these searches instead of just a tabular form. I do this by creating a blank dashboard and selecting **add panel → new from reports**.

This is what we get after adding all 4 reports.



We can export it as pdf and share with other users as well.

Conclusion

This concludes the technical documentation for the solution to the project statement given.

As required I have,

- Created a data file with all the required data
- Created an index and added data to it
- Performed various searches
- Added lookup table
- Configured alert
- Created reports
- Created Dashboard

All the required files including the original data file will be present along with this document of reference.