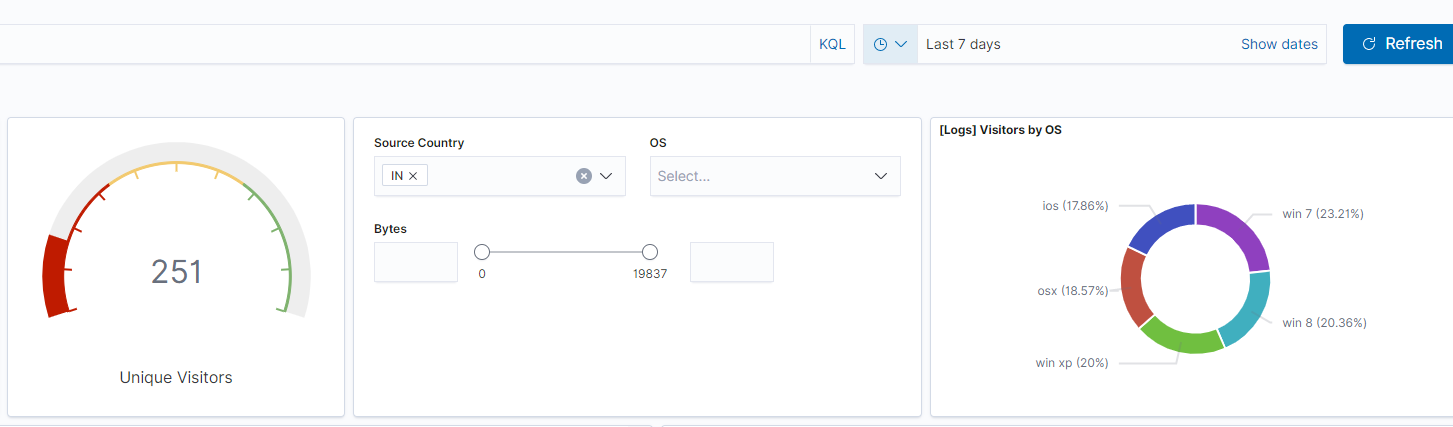
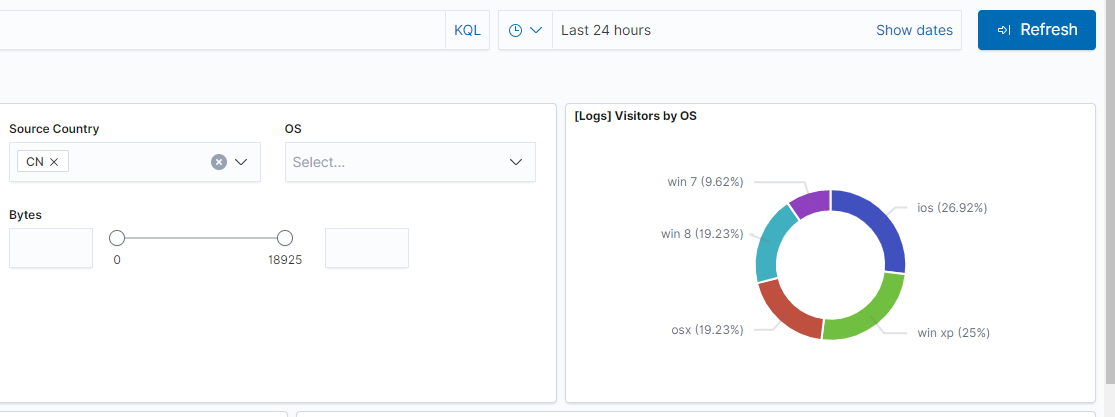
1. Add the sample web log data to Kibana.
2. Answer the following questions:
   * In the last 7 days, how many unique visitors were located in India? **251**

country (IN represents India) Time (last 7 days)

number of visitors (251)

* + In the last 24 hours, of the visitors from China, how many were using Mac OSX? 19.23 %

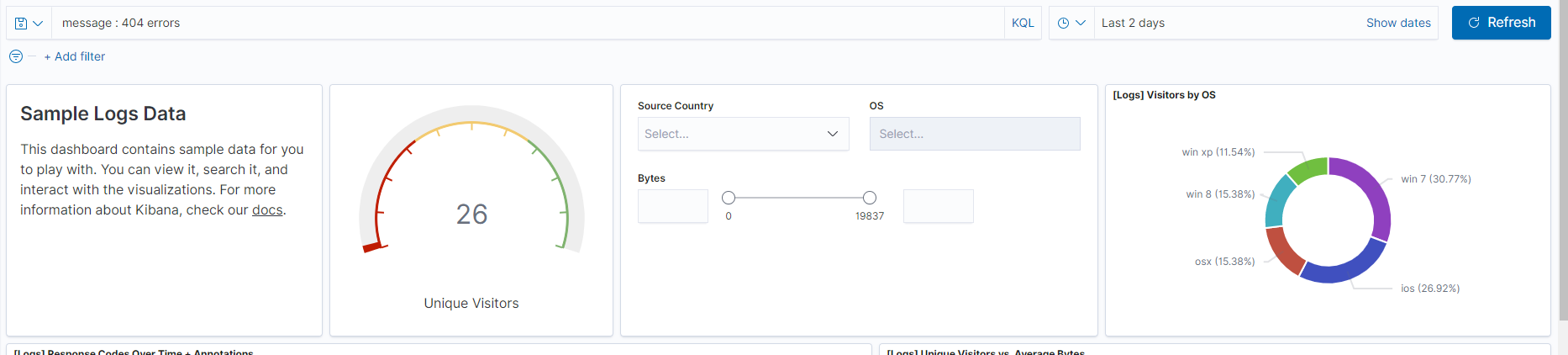
country (CN represent China) Time (last 24 hours)



OSX (19.23%)

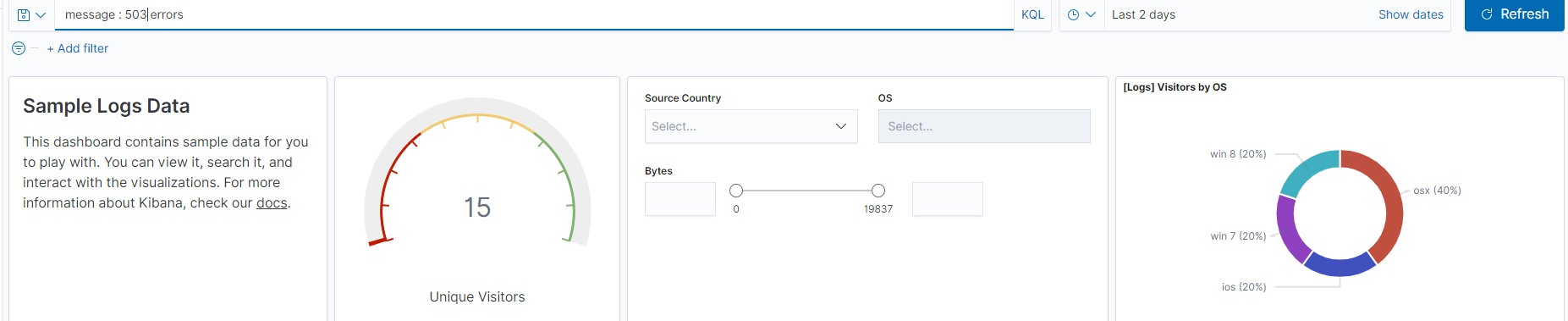
* + In the last 2 days, what percentage of visitors received 404 errors? **26**

filter (message : 404 errors) number of visitors (26)



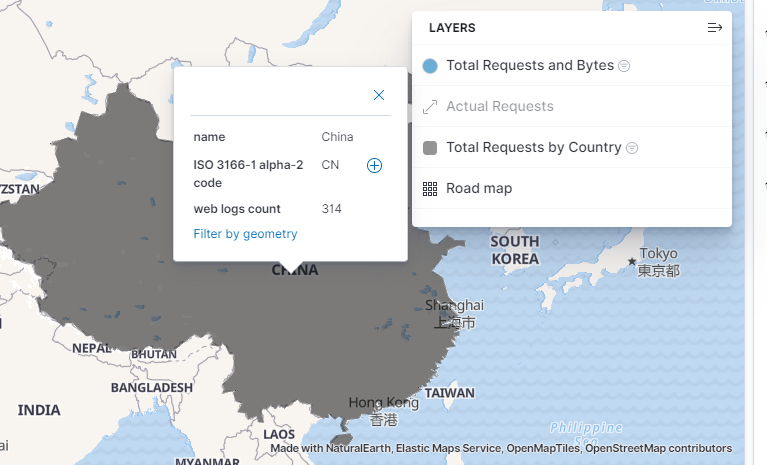
How about 503 errors? **15**

filter (message : 503 errors)number of visitors (15)

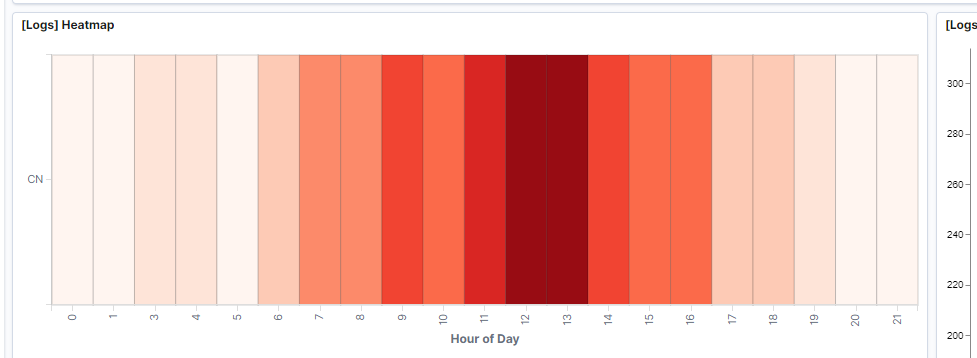


* + In the last 7 days, what country produced the majority of the traffic on the website? **China**

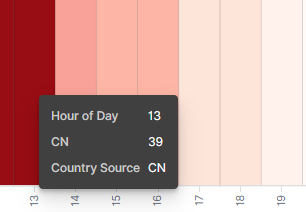
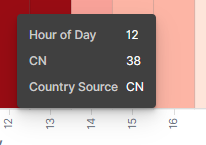
total requests

* + 
  + Of the traffic that's coming from that country, what time of day had the highest amount of activity? **1:00 PM** (note: 13 represents 1:00 PM in military time. Both 12 and 13 are the darkest shades, but hovering the mouse over each time reveals that 13 had one more request than 12. See image below)

country (CN represents China) Time



 one more visitor than 12



* + List all the types of downloaded files that have been identified for the last 7 days, along with a short description of each file type (use Google if you aren't sure about a particular file type).

gz - gz or GZ files are compressed files using gzip algorithm or a related application. gz is natively created for Linux and Unix but currently supports all major operating systems like Windows, MacOS, BSD etc

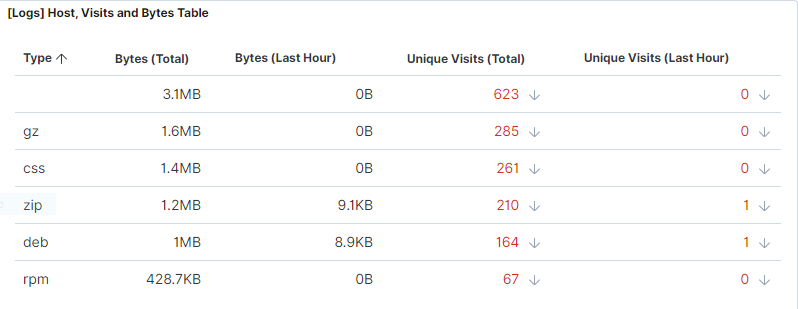
css - The CS file extension is a data file type that is associated with CHROMAom ColorSchemer Studio software. Websites, home decoration, magazines, and other target mediums use this software for developing color palettes

zip - ZIP is an archive file format that supports lossless data compression. A ZIP file may contain one or more files or directories that may have been compressed. The ZIP file format permits a number of compression algorithms, though DEFLATE is the most common

deb - deb is the format, as well as extension of the software package format for the Linux distribution Debian and its derivatives

rpm - An RPM file is an installation package originally developed for the Red Hat Linux operating system, but now used by many other Linux distributions as well. RPM files are commonly used for installing programs on Linux systems.

 types of files (gz, css, zip, deb, and rpm)

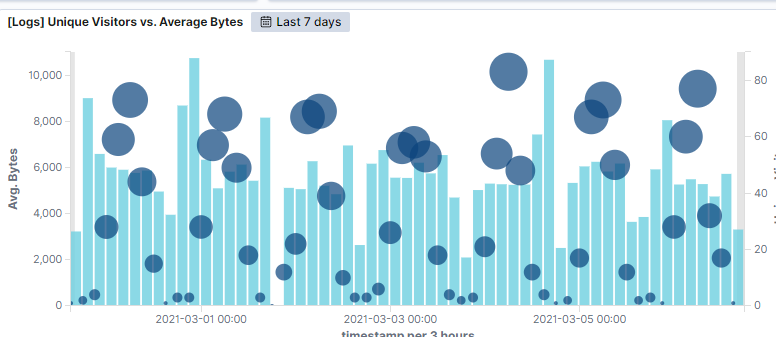


1. Now that you have a feel for the data, Let's dive a bit deeper. Look at the chart that shows Unique Visitors Vs. Average Bytes.
   * Locate the time frame in the last 7 days with the most amount of bytes (activity).

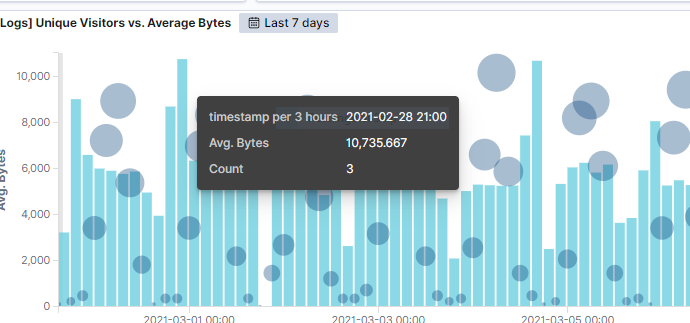
**Around 09:00 PM (21:00) on 2/28/21**

(Note: The timestamps are recorded every three hours, so this spike happened around 11:00 PM on 2/28/21 and the average amount of bytes recorded is 10,735.667 bytes.)

 filter for the last 7 days



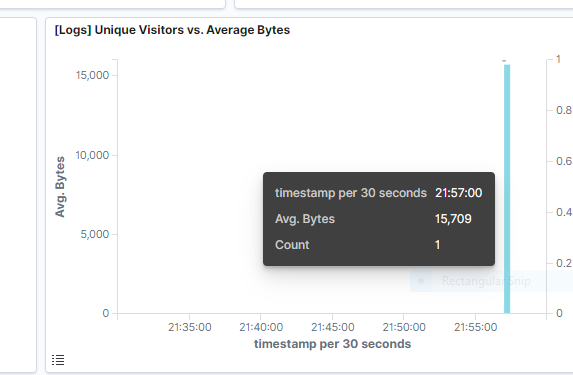
biggest spike on the graph at 21:00 on 2/28/21



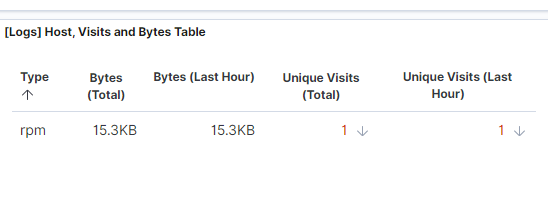
* + In your own words, is there anything that seems potentially strange about this activity? During this time frame only three visitors were present at this site. This means that these three visitors generated the most amount of bytes in a seven day span, which is a lot of bytes for such few people.

1. Filter the data by this event.
   * What is the timestamp for this event? **9:57** or **21:57**

When getting a closer look we see that highest surge occurred on 2/28/21 at 09:57 PM with the largest spike of 15,709 average bytes. Since the previous photo recorded timestamps every 3 hours, this hour was left out of the equation until we zoomed in on it by adjusting the timestamp recording.



* + What kind of file was downloaded? **rpm**

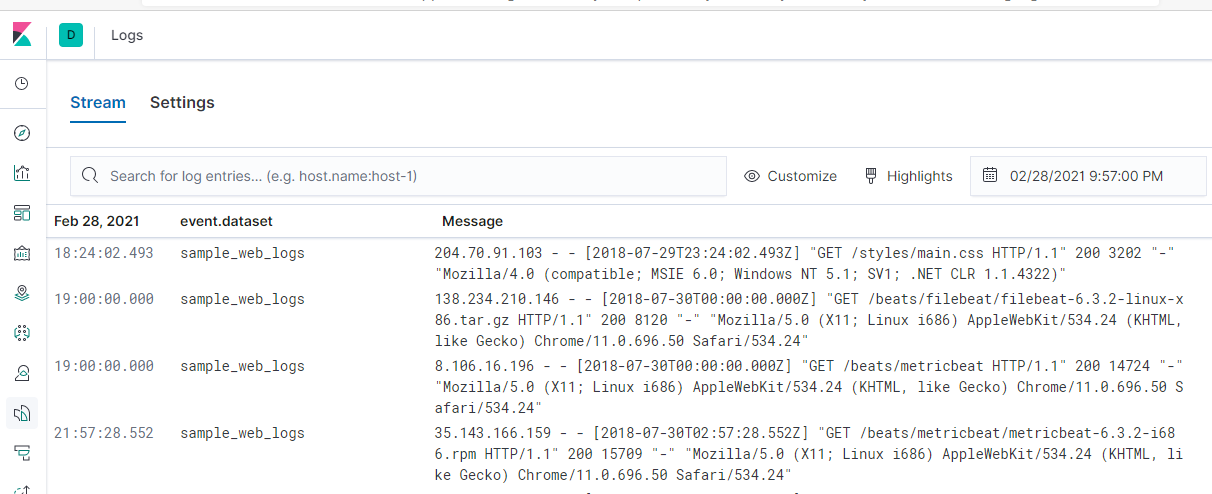


* + From what country did this activity originate? **India**



* + What HTTP response codes were encountered by this visitor?

**200; success**

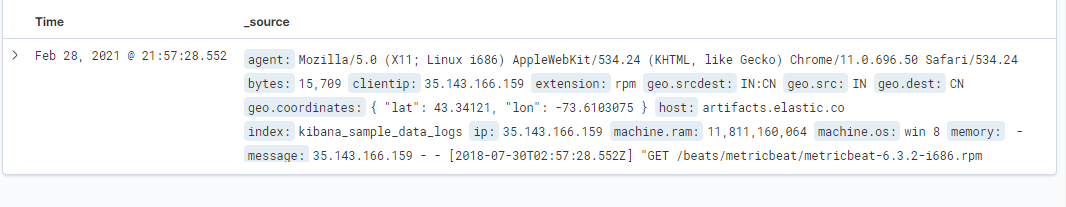


The last row has our desired packet.



Zooming, in we see the response code 200, which indicates success. 

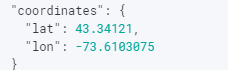
1. Switch to the Kibana Discover page to see more details about this activity. Reference this log to investigate further.



* + What is the source IP address of this activity? **35.143.166.159**



* + What are the geo coordinates of this activity? **latitude: 43.34121, longitude: -73.6103075**



* + What OS was the source machine running? **Windows 8**



* + What is the full URL that was accessed?

**https://artifacts.elastic.co/downloads/beats/metricbeat/metricbeat-6.3.2-i686.rmp**

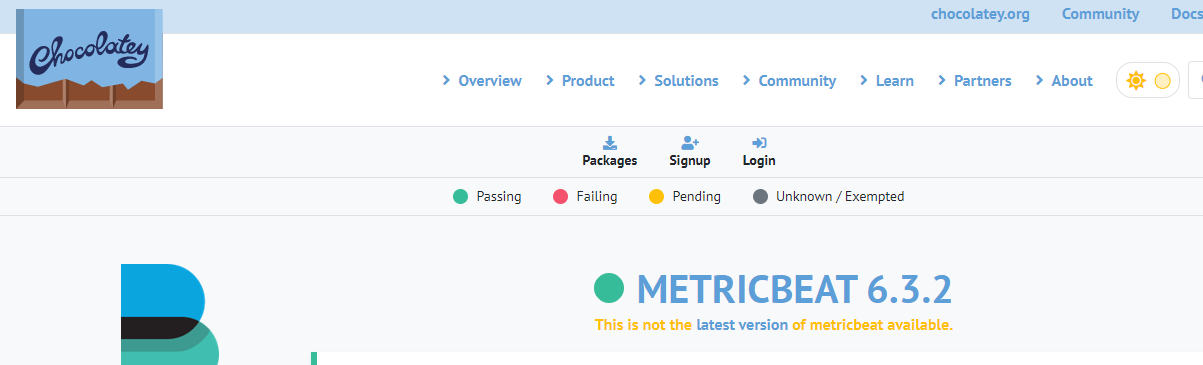
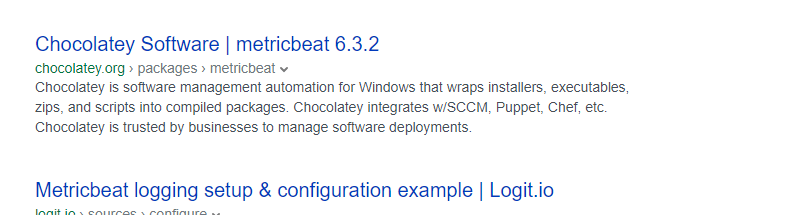


* + From what website did the visitor's traffic originate? **http://facebook.com/success/jay-c-buckey**

To find this answer navigate to Inspect > Request. Next, scroll down to the referrer header. This contains the address of the previous webpage from which the currently requested page was linked. This header allows servers to identify where people are visiting from, and may use that data for analytics, logging, or optimized caching.



1. Finish your investigation with a short overview of your insights.
   * What do you think the user was doing? Investigating the URL that the user requested and investigating, the user visited Chocolately, a software management automation for Windows that wraps installers, executables, zips, and scripts into compiled packages. The user then downloaded an rpm package manager file to install Metricbeat. This explains the reason for the large amount of activity that was observed on 2/28/21.

. 

* + Was the file they downloaded malicious? If not, what is the file used for? The rpm file that was downloaded is not malicious. A file with the RPM file extension is a Red Hat Package Manager file that's used to store installation packages on Linux operating systems. These files provide an easy way for software to be distributed, installed, upgraded, and removed since they're "packaged" in one place. The user downloaded a package manager to install Metricbeat.
  + Is there anything that seems suspicious about this activity? The direction this traffic ran across nations, from to India to Canada.



* + Is any of the traffic you inspected potentially outside of compliance guidelines? Despite the distance that the traffic ran to download an rpm file, this traffic is legitimate and does not fall outside of compliance guidelines. This demonstrates that even though Kibana displays the results, they must be analyzed for human interpretation.