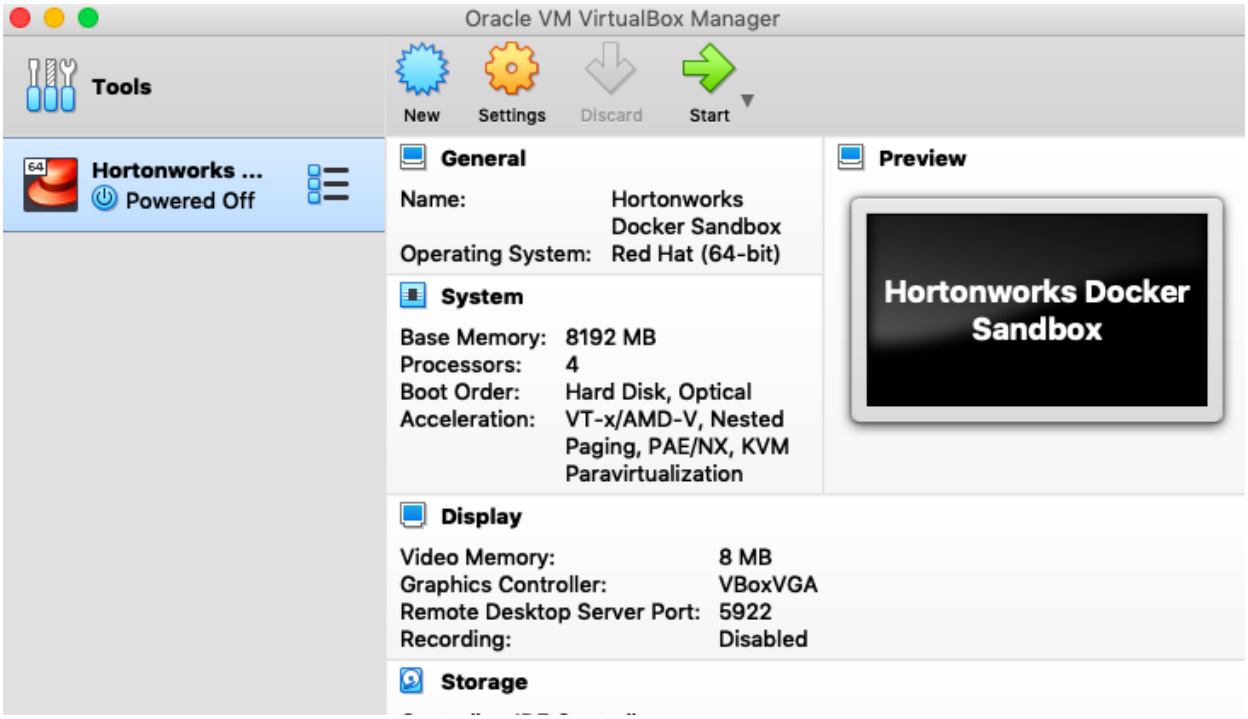
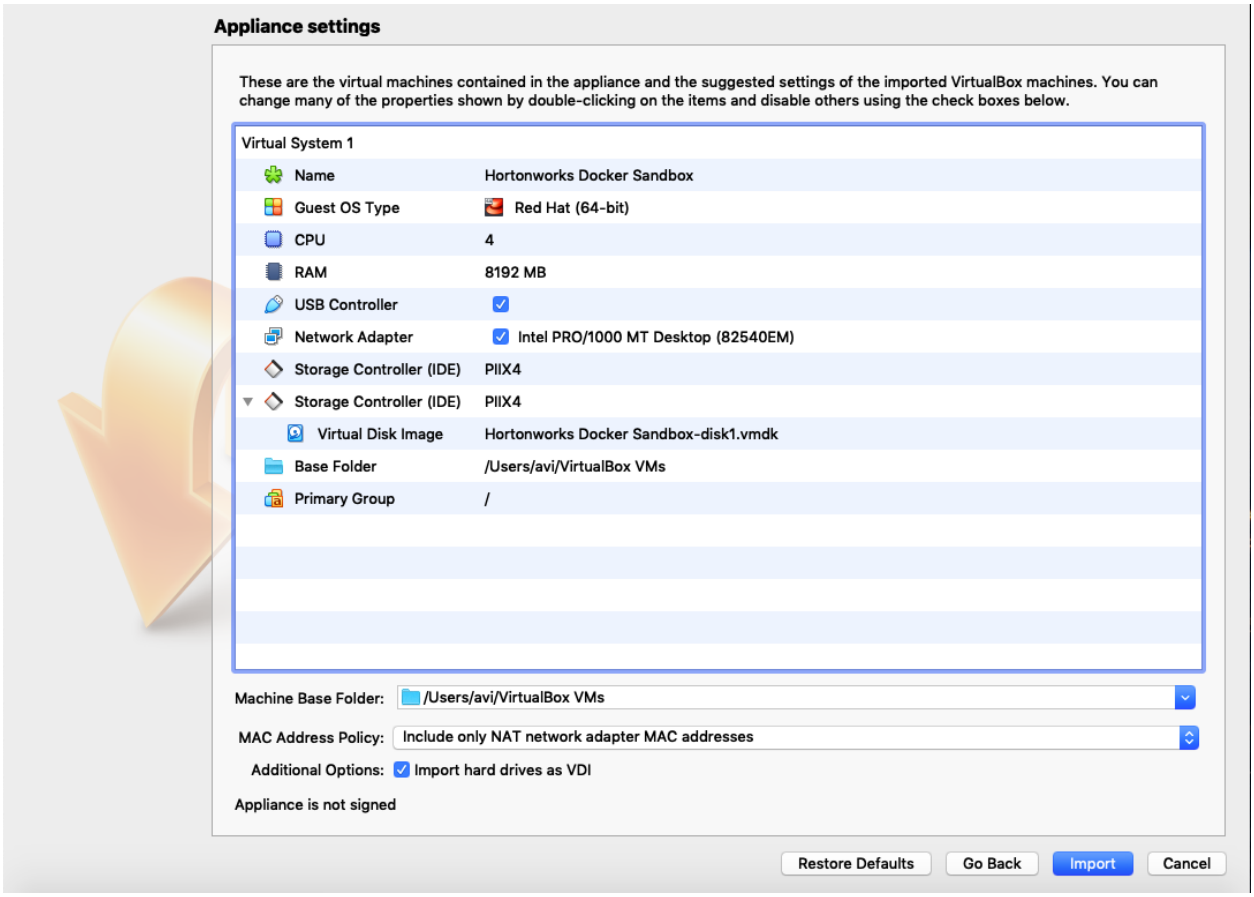
CLOUD COMPUTING
ASSIGNMENT 3

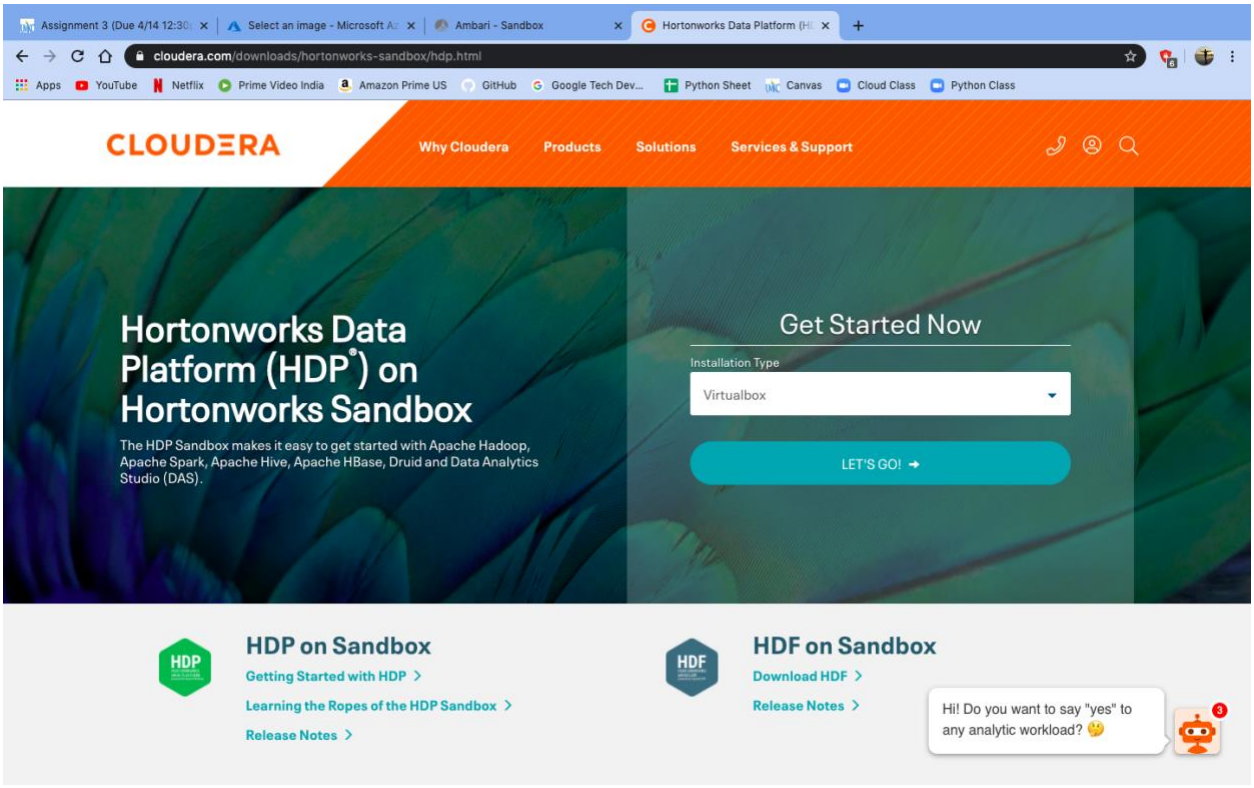
Avinash Ganguri
16293133

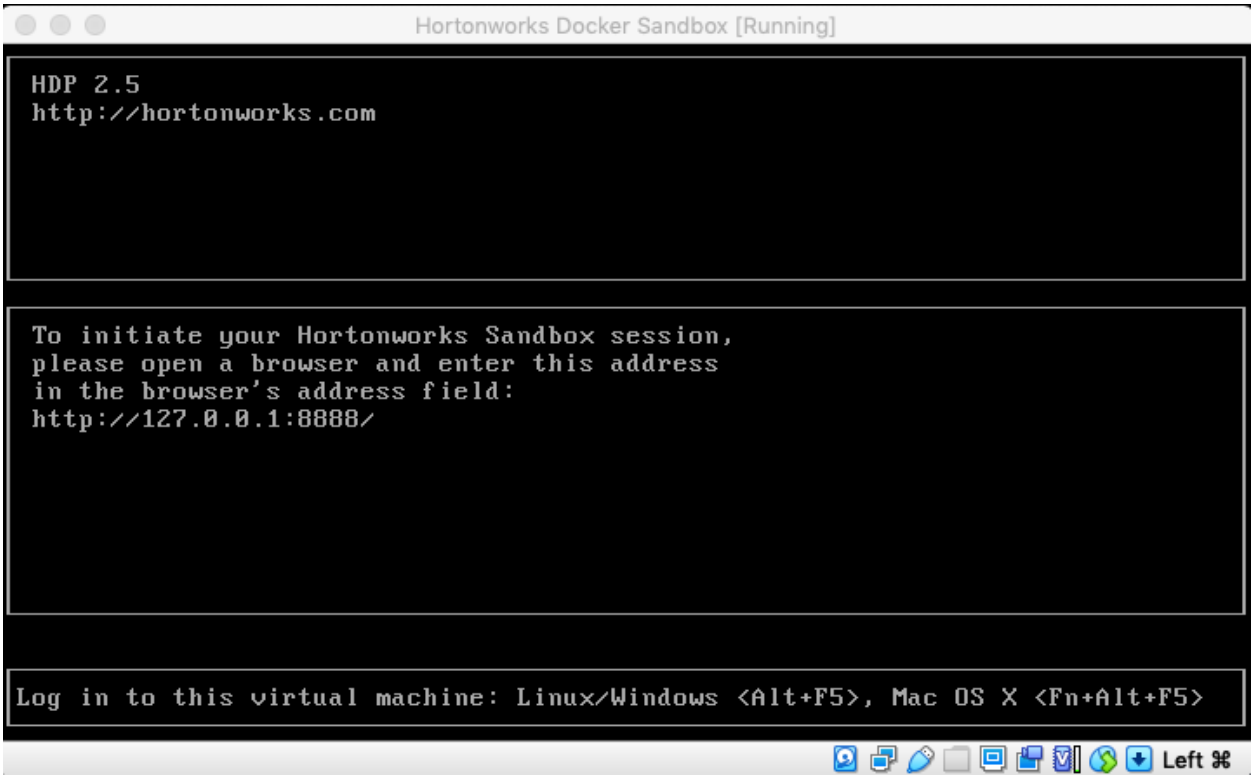
Step1: Downloaded the Virtual box hosted hypervisor for Mac





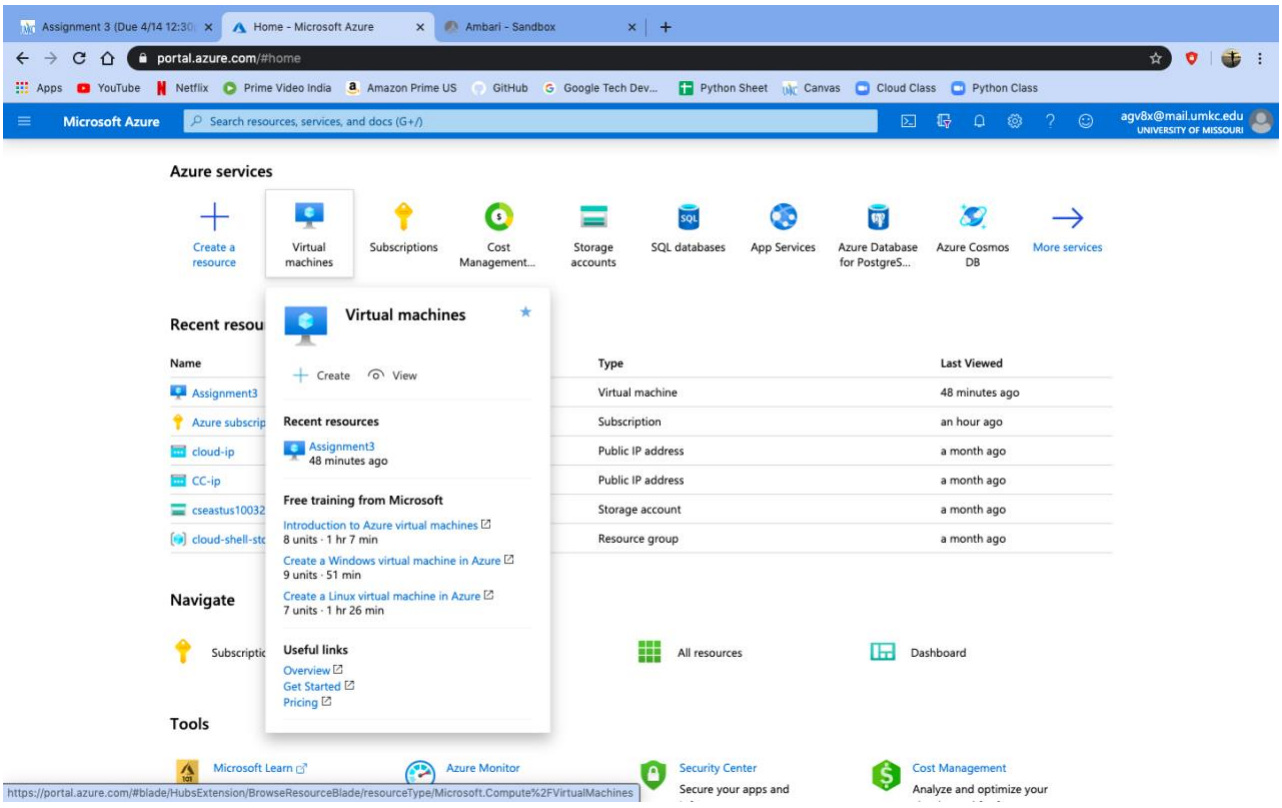
Step2: Download Hortonworks HDP Image from the site,

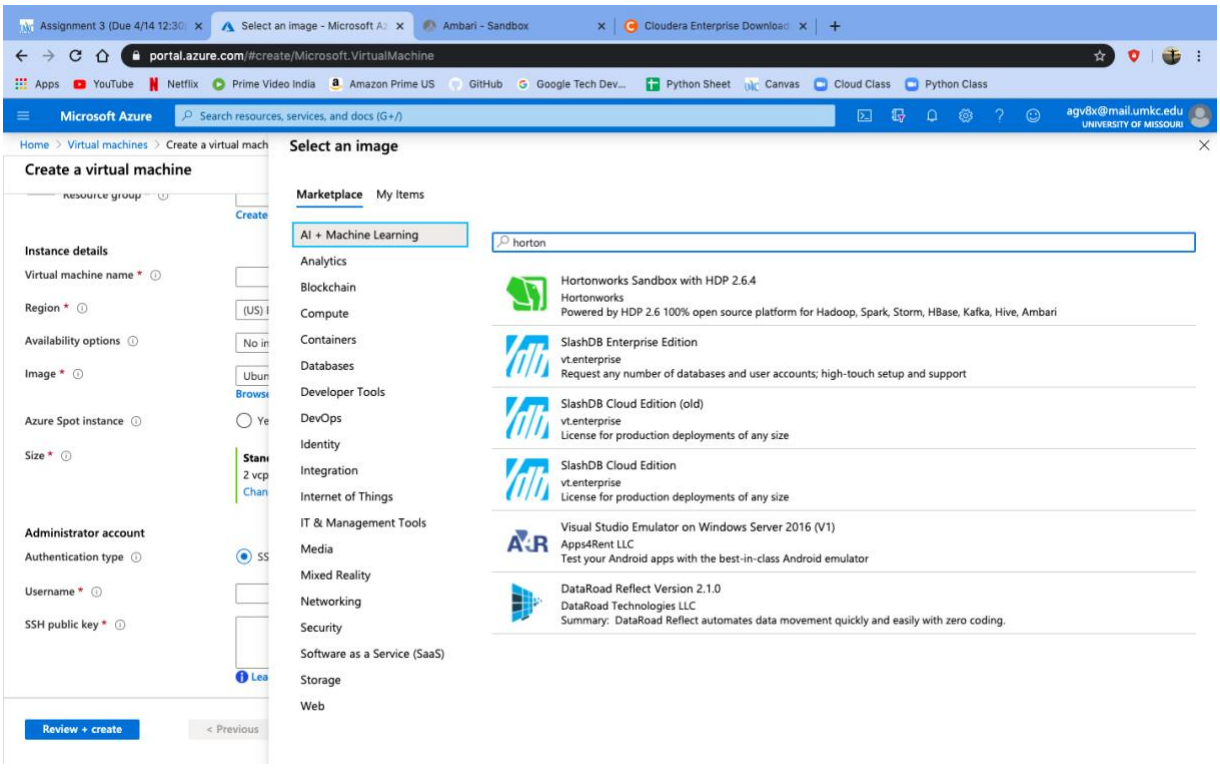


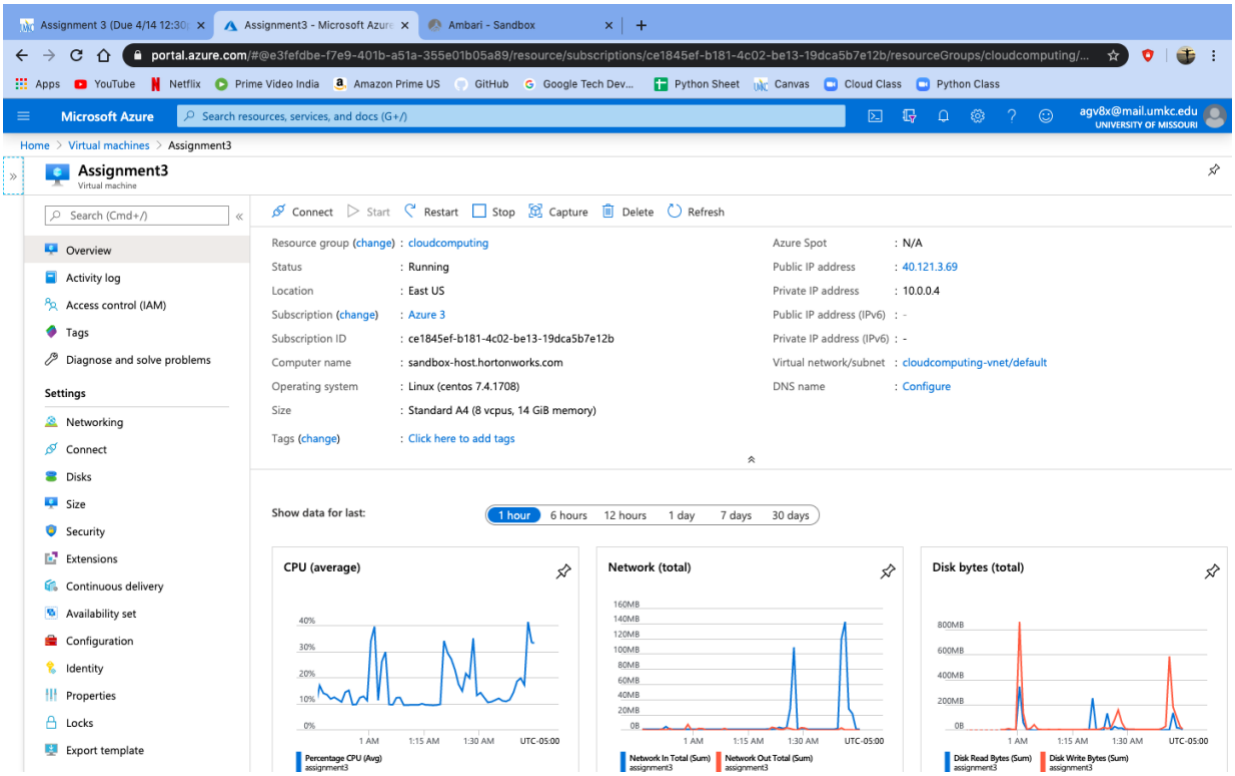


And my system configurations are,

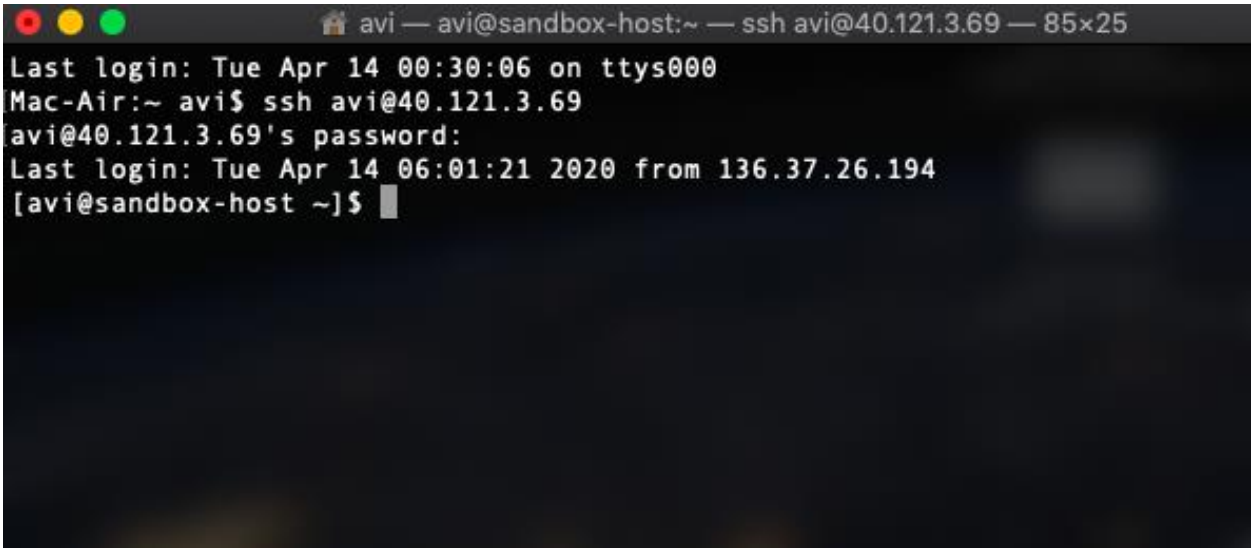
*To run the Image Hortonworks HDP in Virtual Box, we need **RAM** more than 8GB. On running the virtual box my system crashes for every 5 mins, so I have to choose another alternative for running Hortonworks HDP Image. Therefore, I used **Microsoft Azure services** to create a Virtual Machine and run the **Hortonworks HDP 2.5 Version** in it*.


Assignment 3 (Due 4/14 12:30)
Home - Microsoft Azure
Ambari - Sandbox
portal.azure.com/#home
Apps YouTube Netflix Prime Video India Amazon Prime US GitHub Google Tech Dev... Python Sheet Canvas Cloud Class Python Class
Microsoft Azure
Search resources, services, and docs (G+/)
agv8x@mail.umkc.edu
UNIVERSITY OF MISSOURI
Azure services
Create a resource
Virtual machines
Subscriptions
Cost Management...
Storage accounts
SQL databases
App Services
Azure Database for PostgreS...
Azure Cosmos DB
More services
Recent resou
Name
Type
Last Viewed
Assignment3
Virtual machine
48 minutes ago
Azure subscrip
Subscription
an hour ago
cloud-ip
Public IP address
a month ago
CC-ip
Public IP address
a month ago
cseastus10032
Storage account
a month ago
cloud-shell-sto
Resource group
a month ago
Virtual machines
Create
View
Recent resources
Assignment3
48 minutes ago
Free training from Microsoft
Introduction to Azure virtual machines
8 units · 1 hr 7 min
Create a Windows virtual machine in Azure
9 units · 51 min
Create a Linux virtual machine in Azure
7 units · 1 hr 26 min
Useful links
Overview
Get Started
Pricing
Navigate
Subscriptio
All resources
Dashboard
Tools
Microsoft Learn
Azure Monitor
Security Center
Secure your apps and
Cost Management
Analyze and optimize your
https://portal.azure.com/#blade/HubsExtension/BrowseResourceBlade/resourceType/Microsoft.Compute%2FVirtualMachines


Assignment 3 (Due 4/14 12:30)
Select an image - Microsoft Az
Ambari - Sandbox
Cloudera Enterprise Download
portal.azure.com/#create/Microsoft.VirtualMachine
Apps YouTube Netflix Prime Video India Amazon Prime US GitHub Google Tech Dev... Python Sheet Canvas Cloud Class Python Class
Microsoft Azure
Search resources, services, and docs (G+/)
agv8x@mail.umkc.edu
UNIVERSITY OF MISSOURI
Home > Virtual machines > Create a virtual mach
Create a virtual machine
Instance details
Virtual machine name *
Region *
Availability options
Image *
Azure Spot instance
Size *
Administrator account
Authentication type
Username *
SSH public key *
Review + create
< Previous
Select an image
Marketplace
My Items
AI + Machine Learning
Analytics
Blockchain
Compute
Containers
Databases
Developer Tools
DevOps
Identity
Integration
Internet of Things
IT & Management Tools
Media
Mixed Reality
Networking
Security
Software as a Service (SaaS)
Storage
Web
horton
Hortonworks Sandbox with HDP 2.6.4
Hortonworks
Powered by HDP 2.6 100% open source platform for Hadoop, Spark, Storm, HBase, Kafka, Hive, Ambari
SlashDB Enterprise Edition
vt.enterprise
Request any number of databases and user accounts; high-touch setup and support
SlashDB Cloud Edition (old)
vt.enterprise
License for production deployments of any size
SlashDB Cloud Edition
vt.enterprise
License for production deployments of any size
Visual Studio Emulator on Windows Server 2016 (V1)
Apps4Rent LLC
Test your Android apps with the best-in-class Android emulator
DataRoad Reflect Version 2.1.0
DataRoad Technologies LLC
Summary: DataRoad Reflect automates data movement quickly and easily with zero coding.

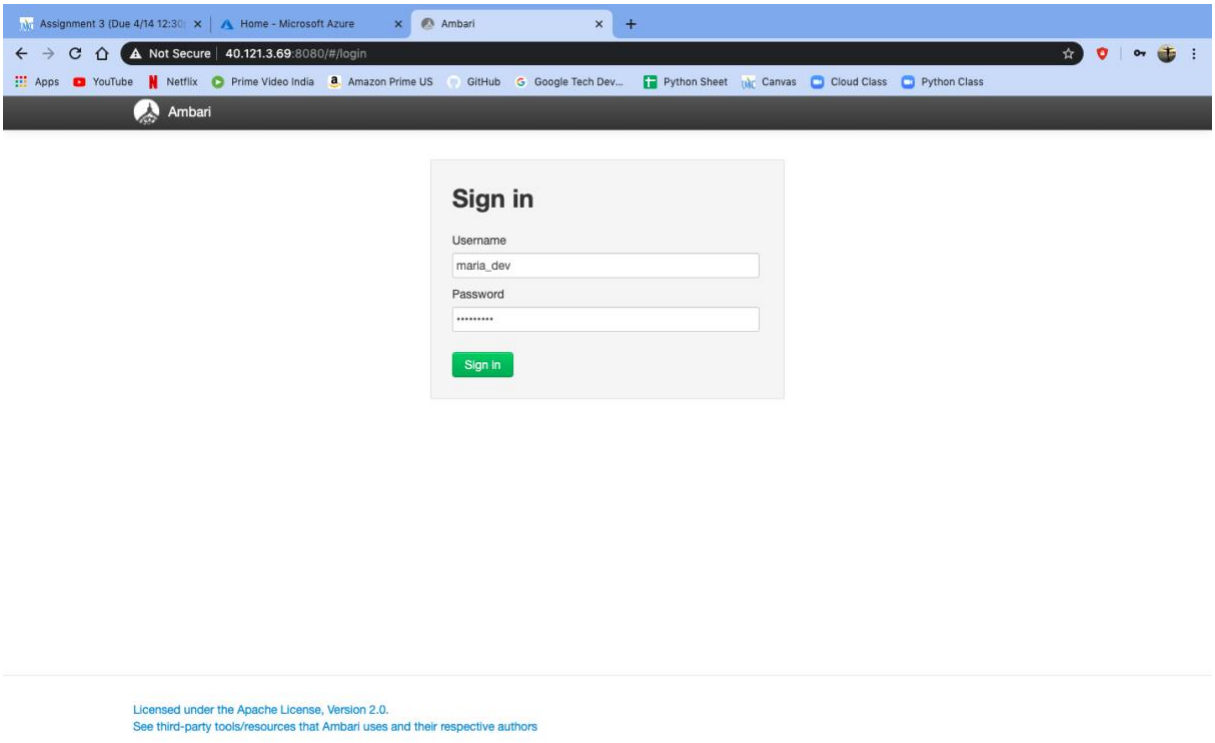


And then connecting the VM to my localhost through SSH,

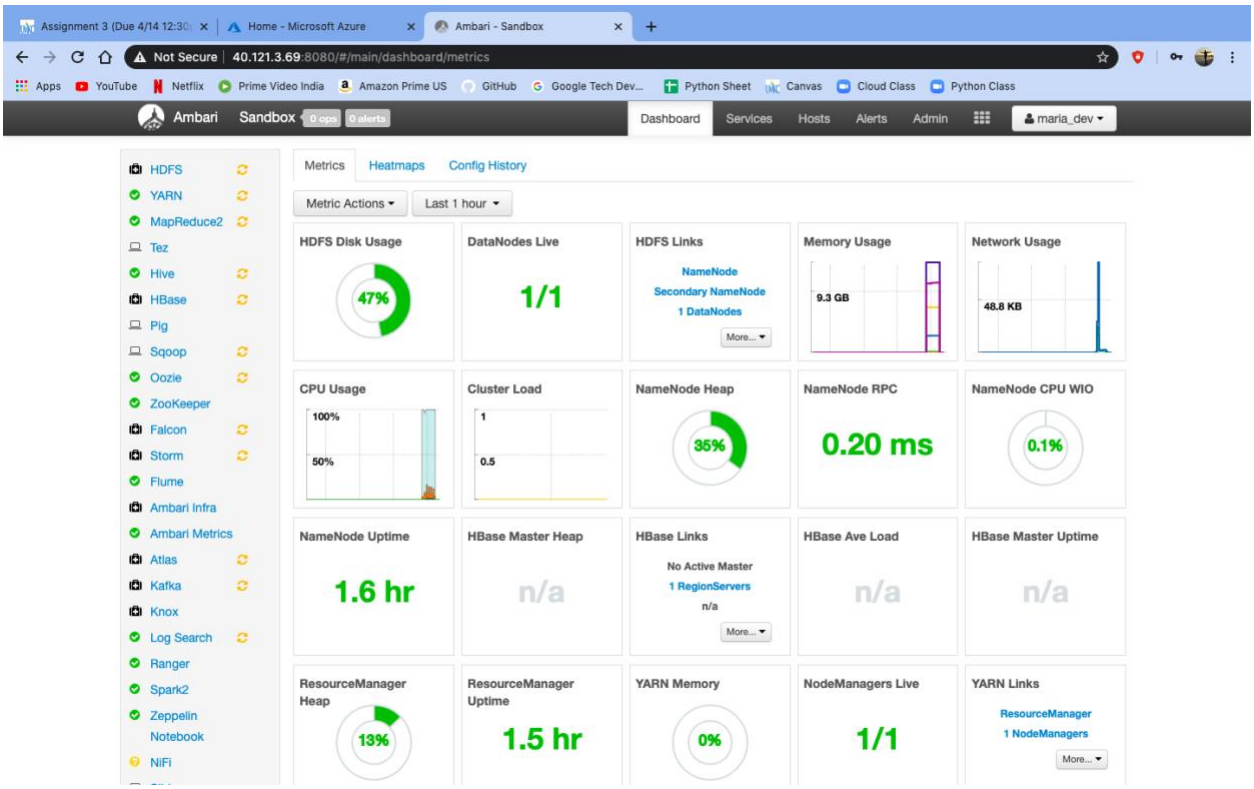


Step 4:

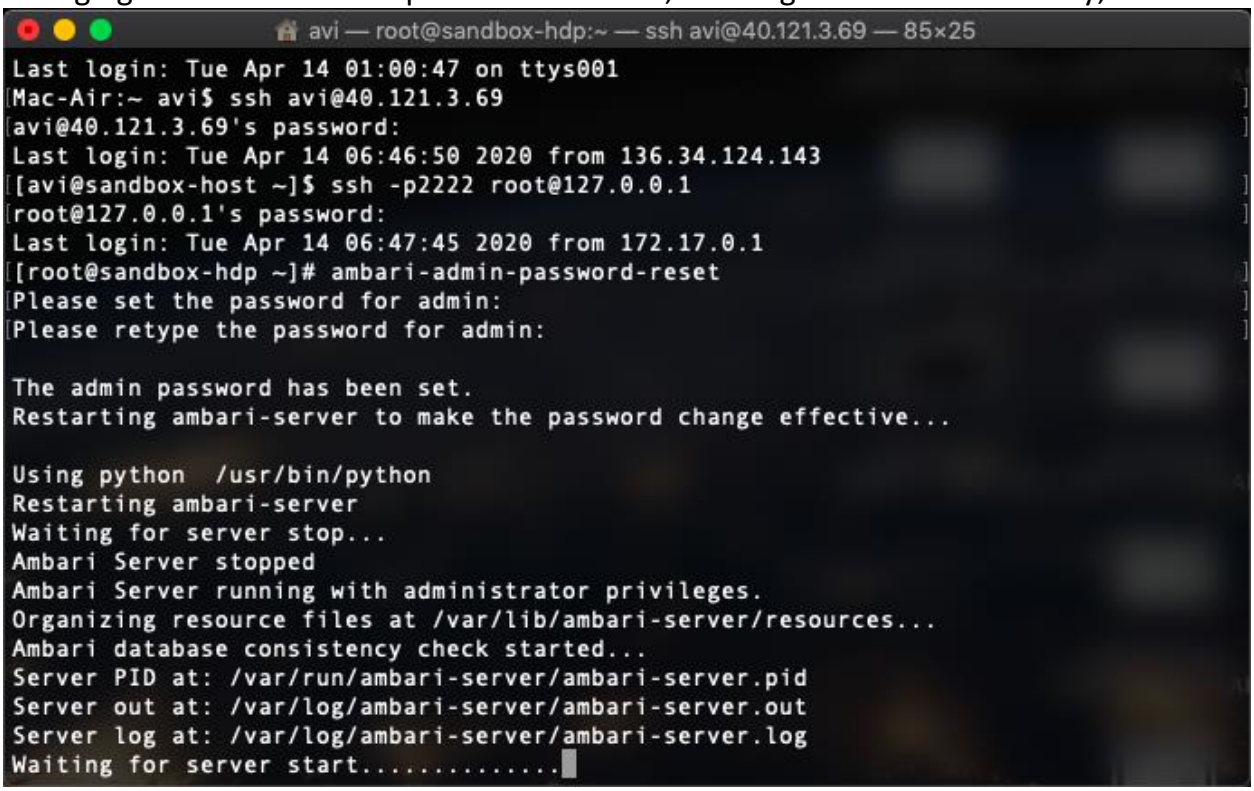
Using the public IP from Azure VM, open it in the browser,
Enter username: maria_dev, password: maria_dev

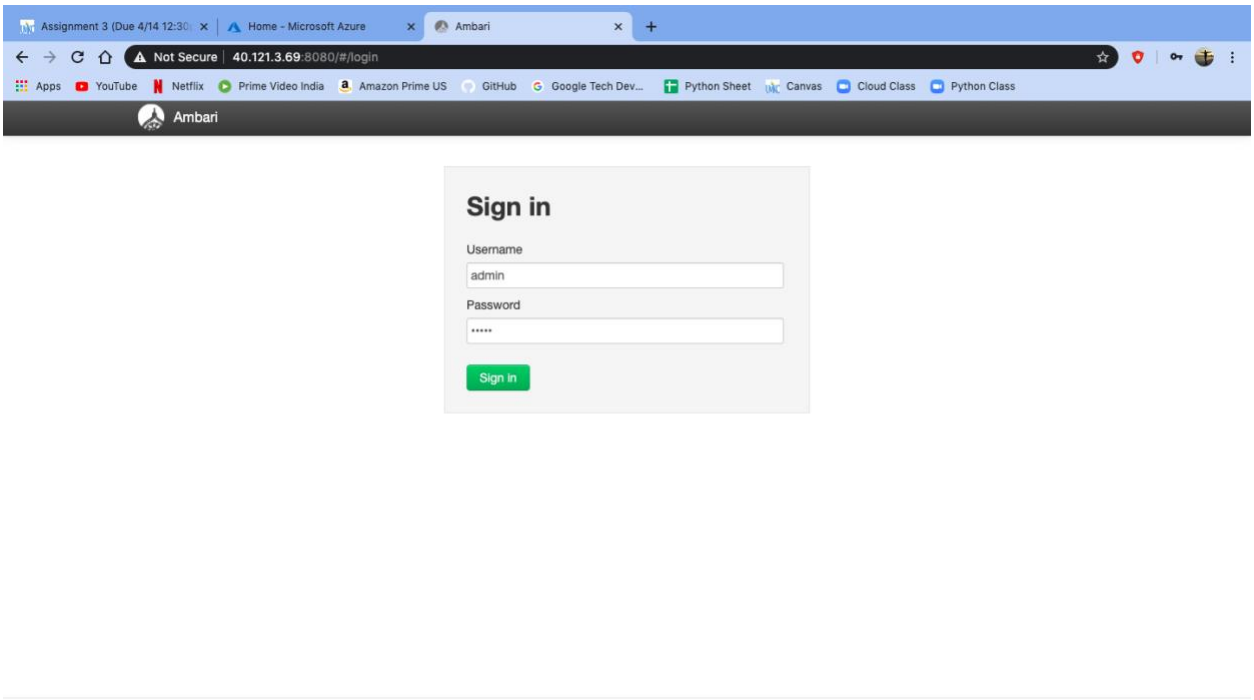


The Ambari login page shows up, after successful login you can see the dashboard,

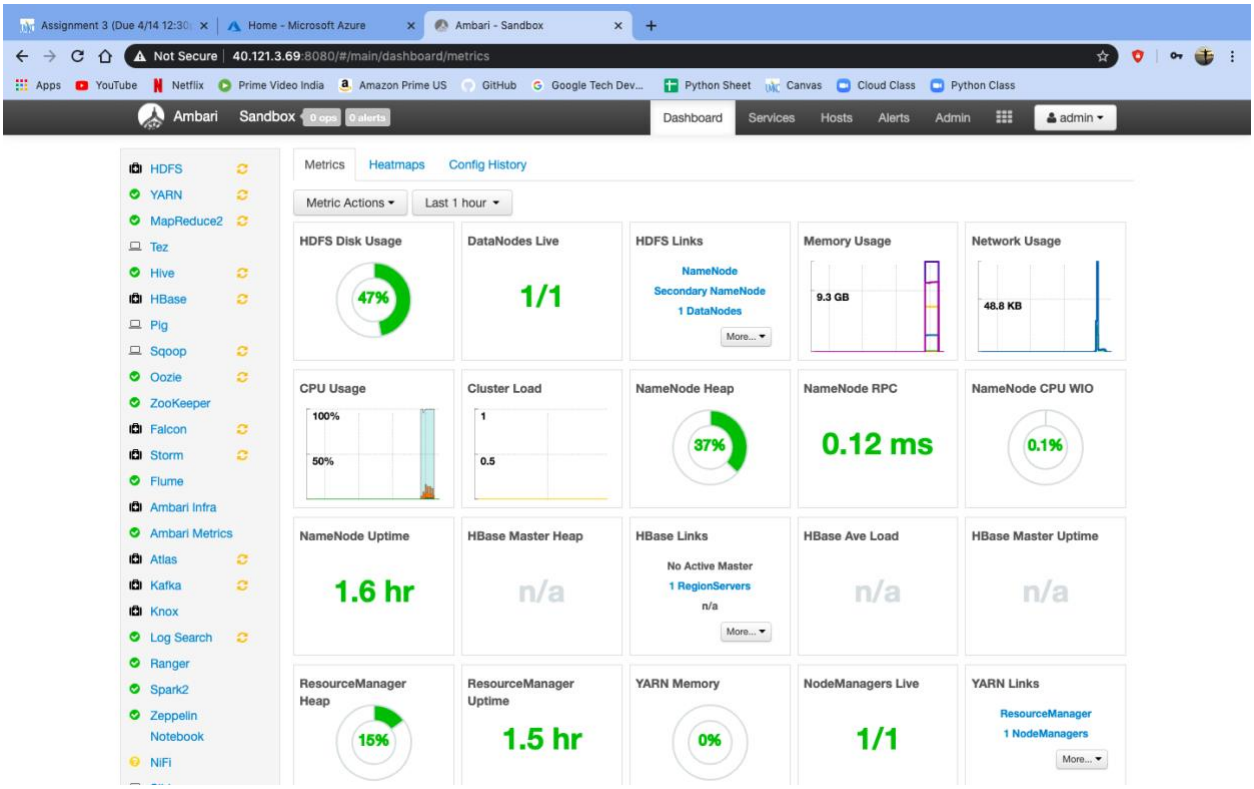


Changing the username and password and also, creating admin credentials by,



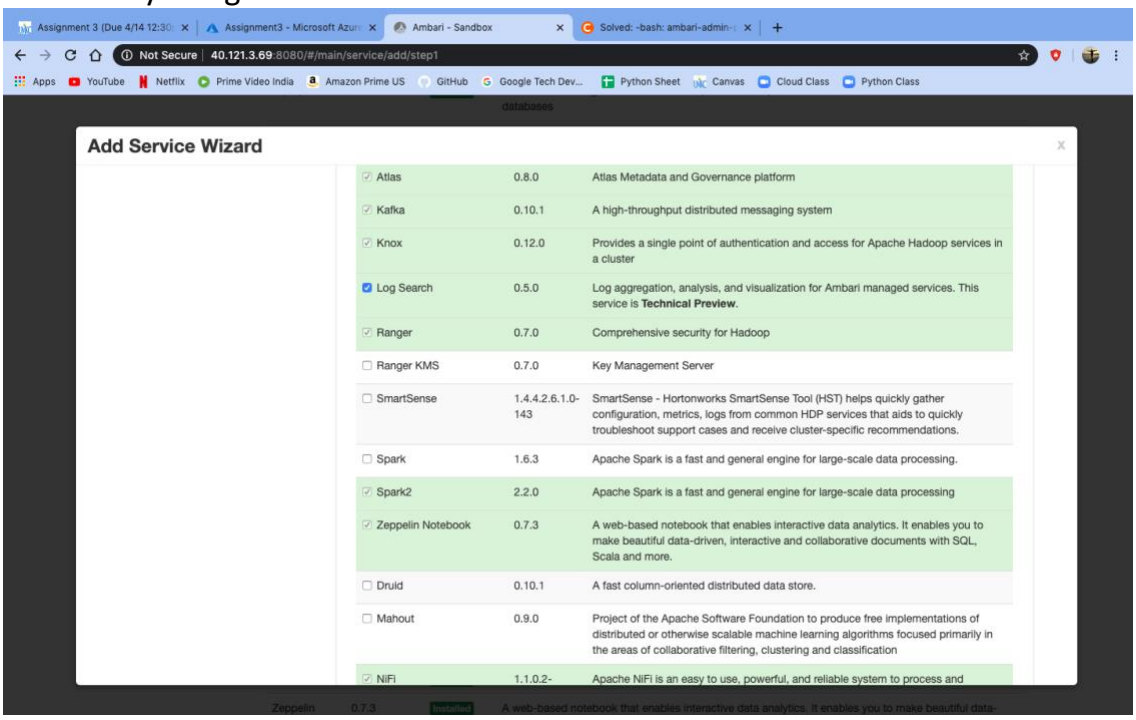


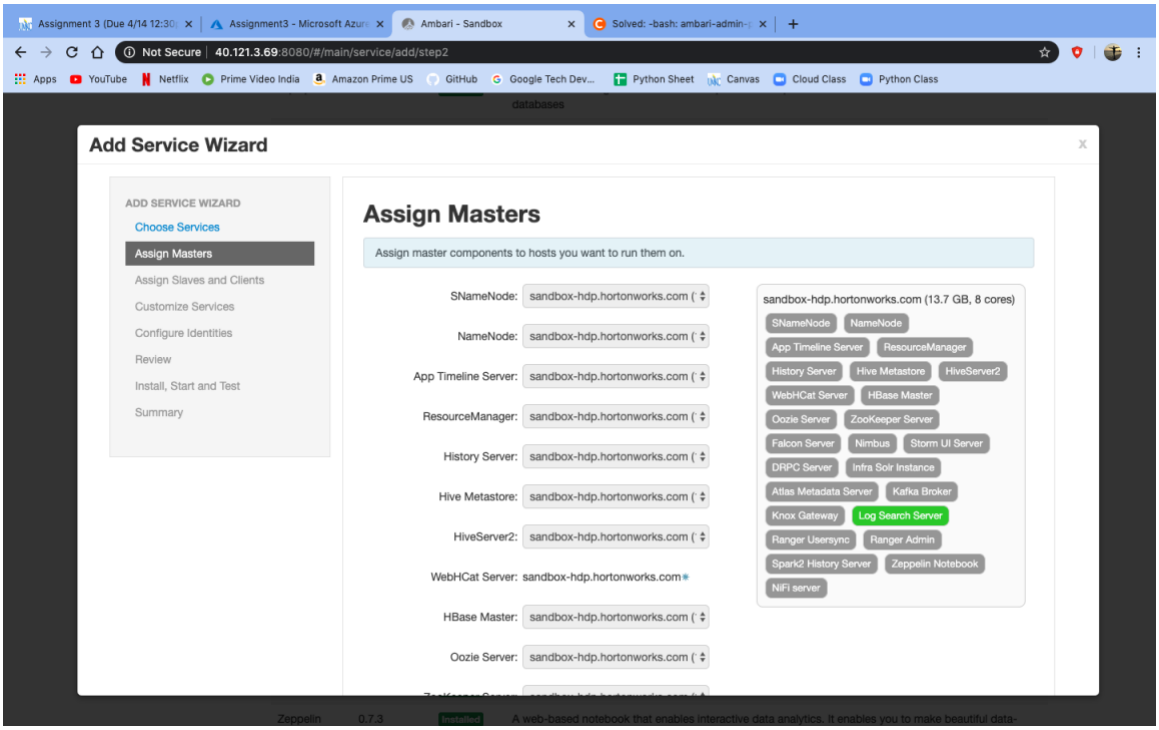
After logging with the admin details, the dashboard page



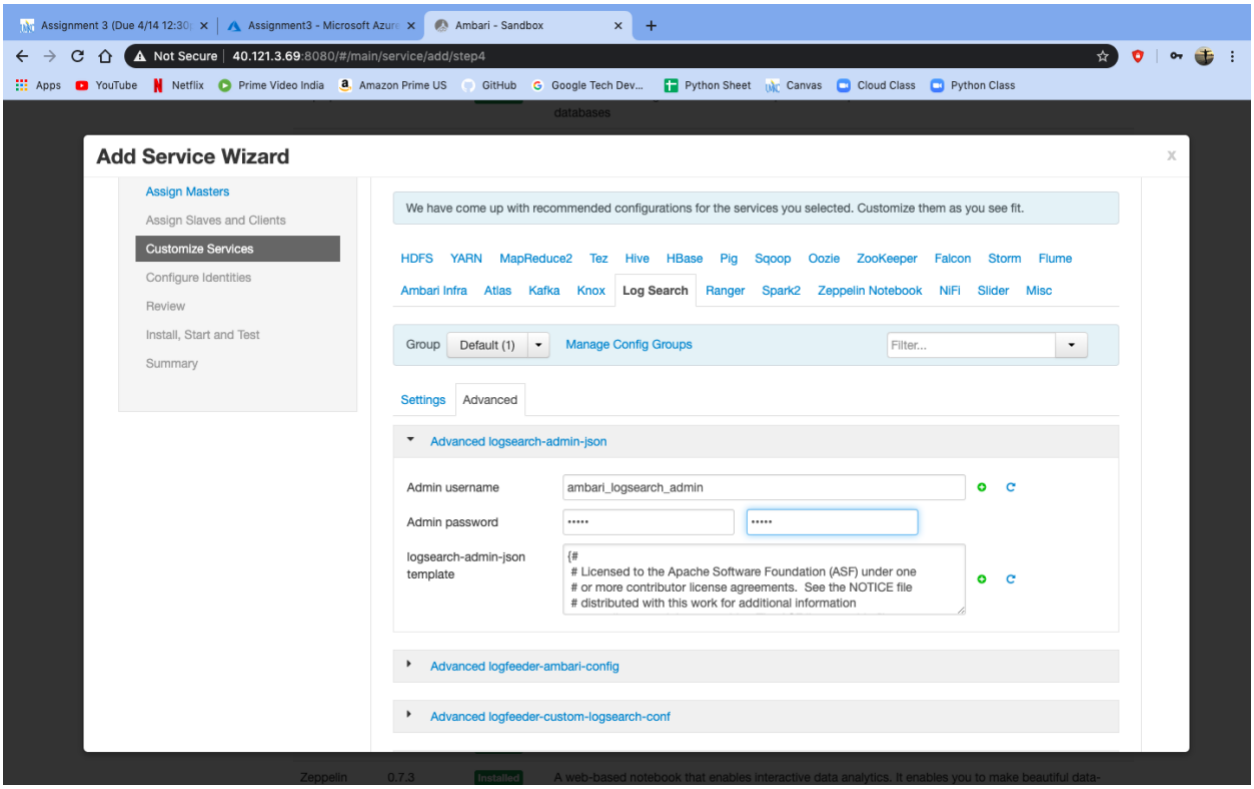
Step 5: Explore Ambari using web browser. Explore its capabilities and show at least three things with screenshots of its function (eg. installing other modules, setting system alert parameter and method, etc.)

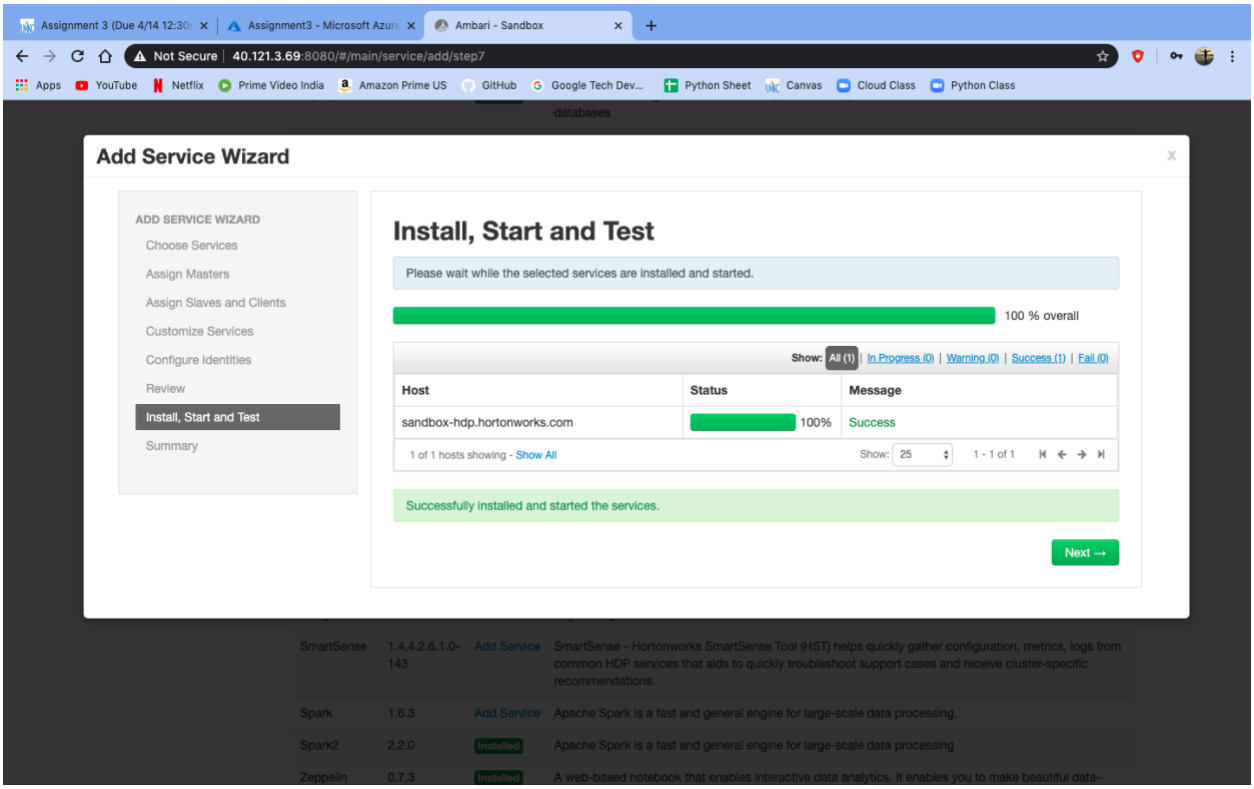
1) A Log Search Service is added, and only an Admin can add it,so we have to login only using admin details.

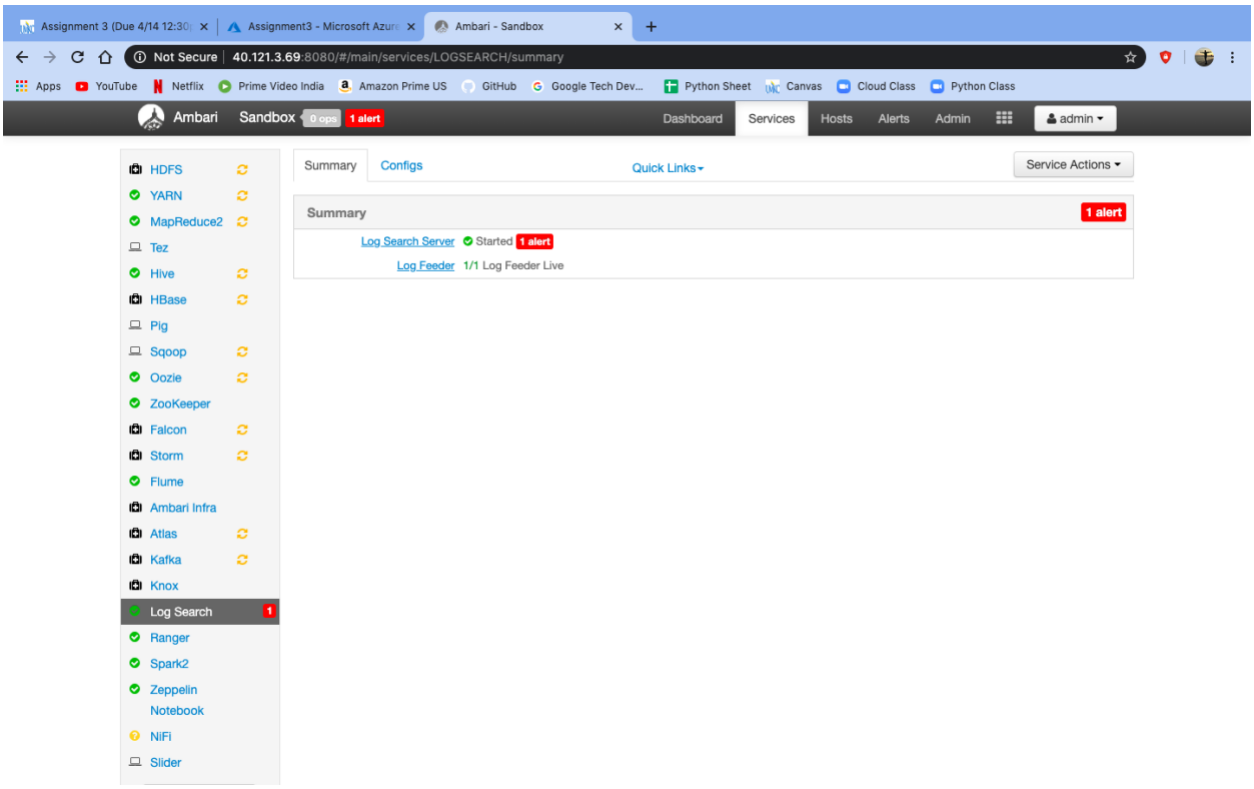




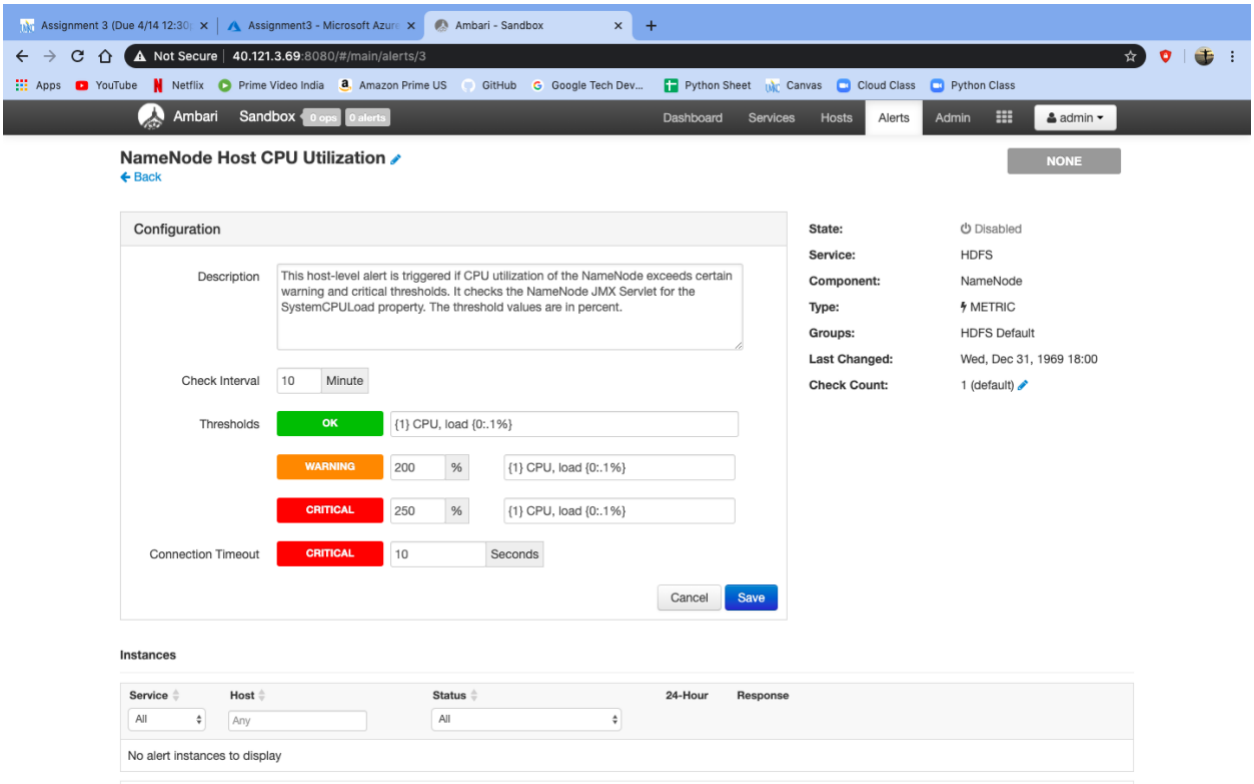
After successfully deploying,



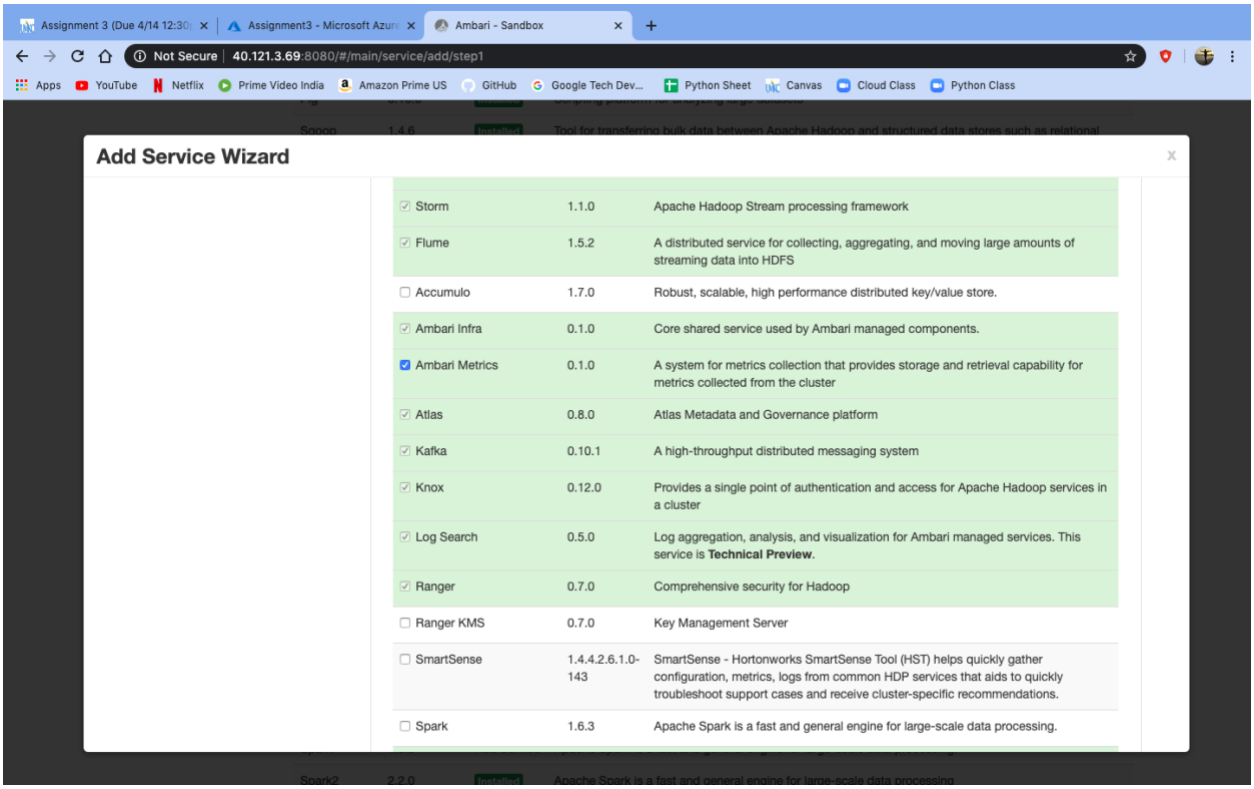


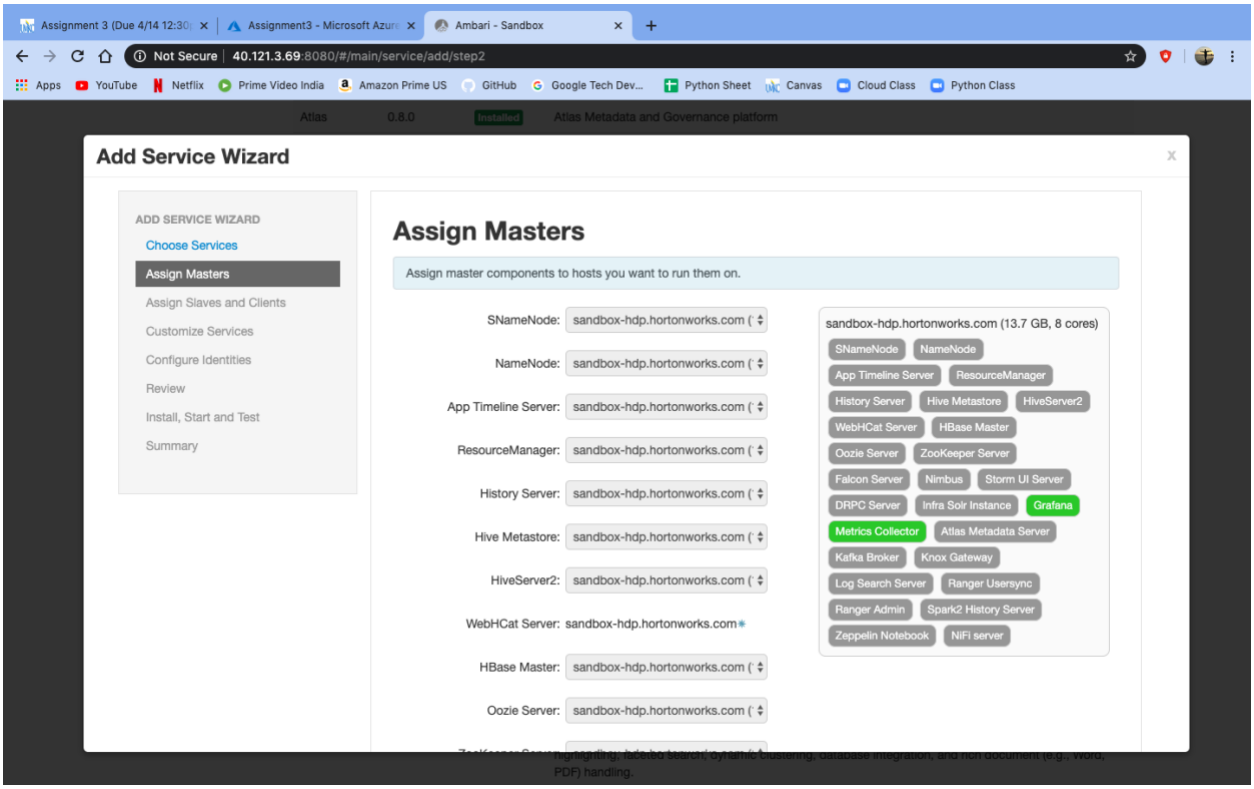


2) Adding NameNode Host CPU Utilization Alert by changing the interval from 5 minsto 10 mins and connection timeout 5 secs to 10 secs.

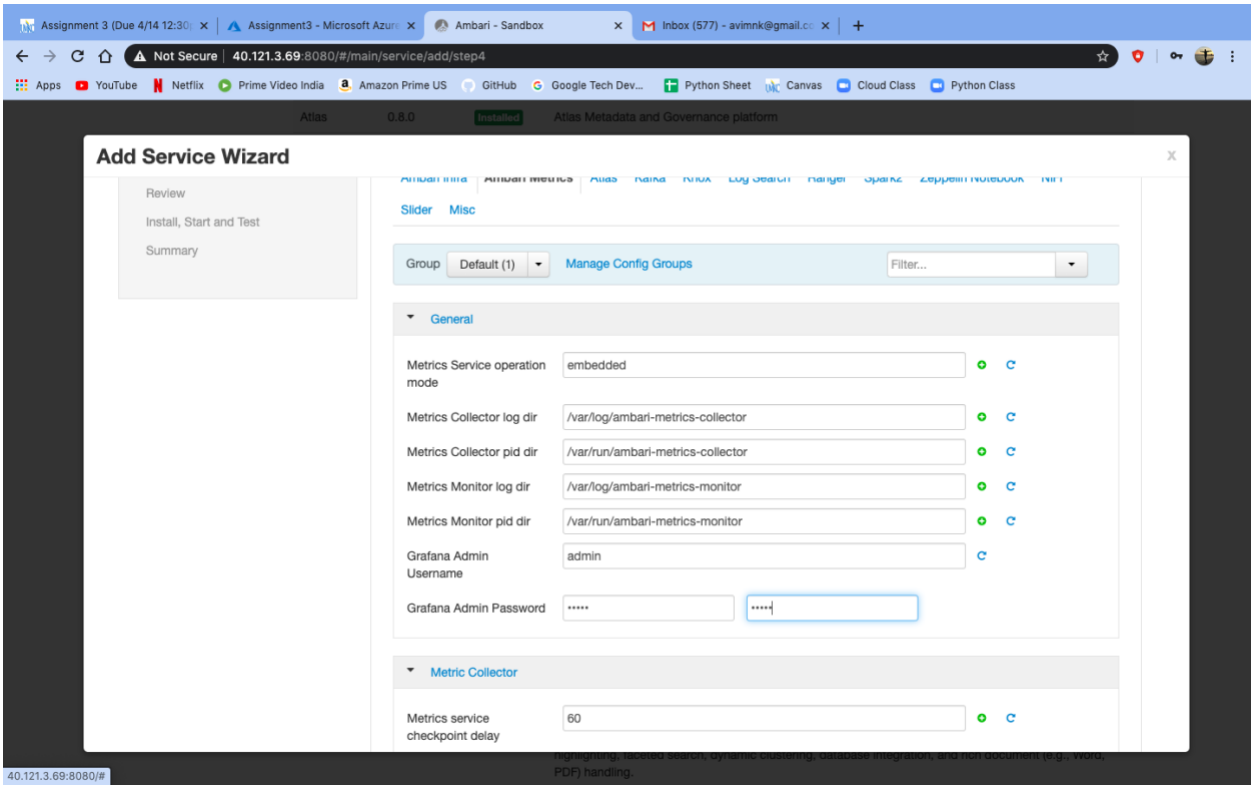


3) Adding Ambari Metrics Service,

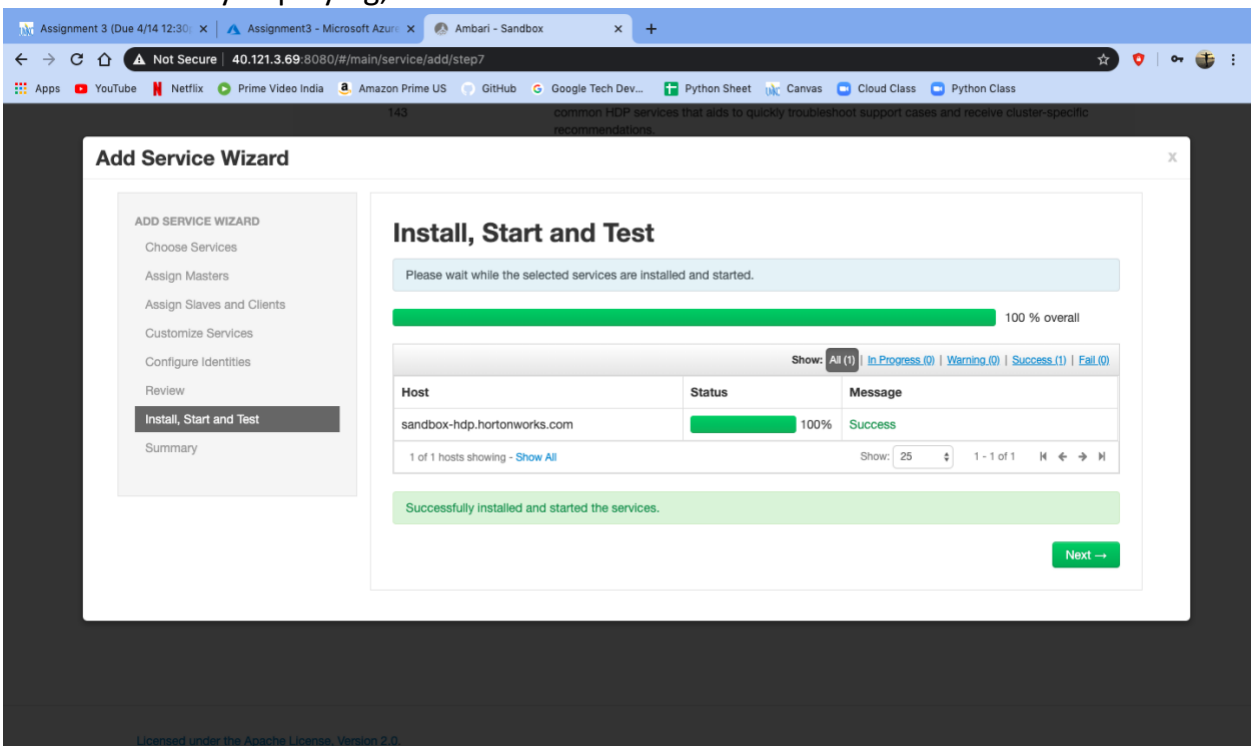


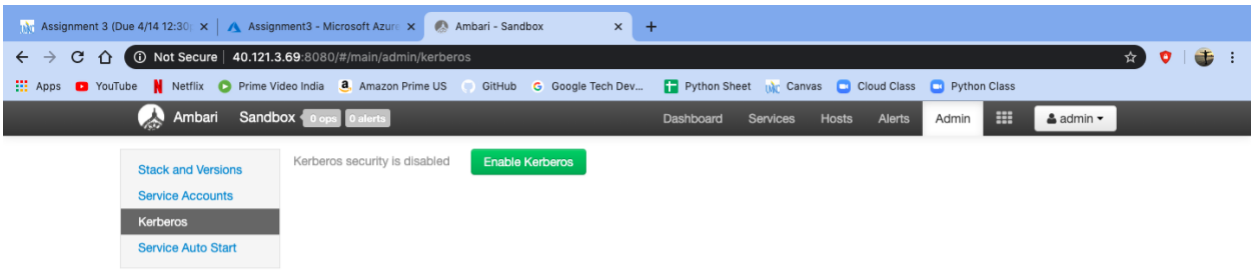


Creating admin password,



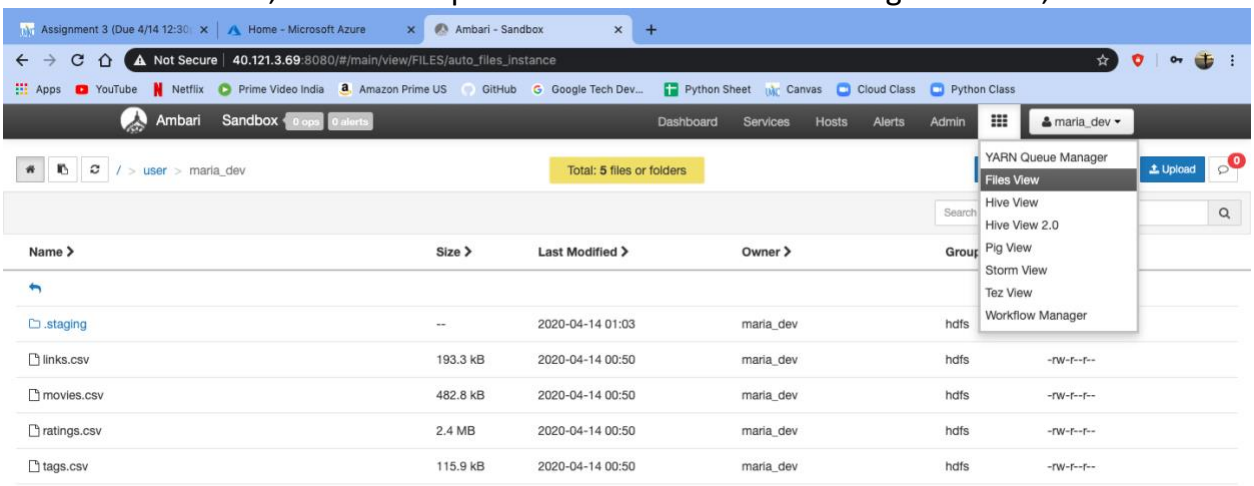
After succesfully deploying,



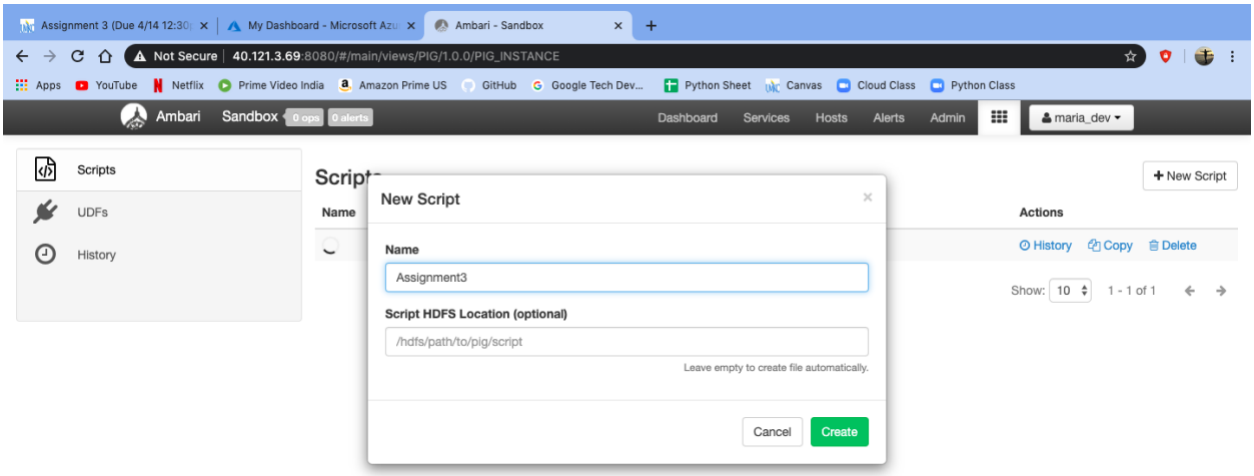


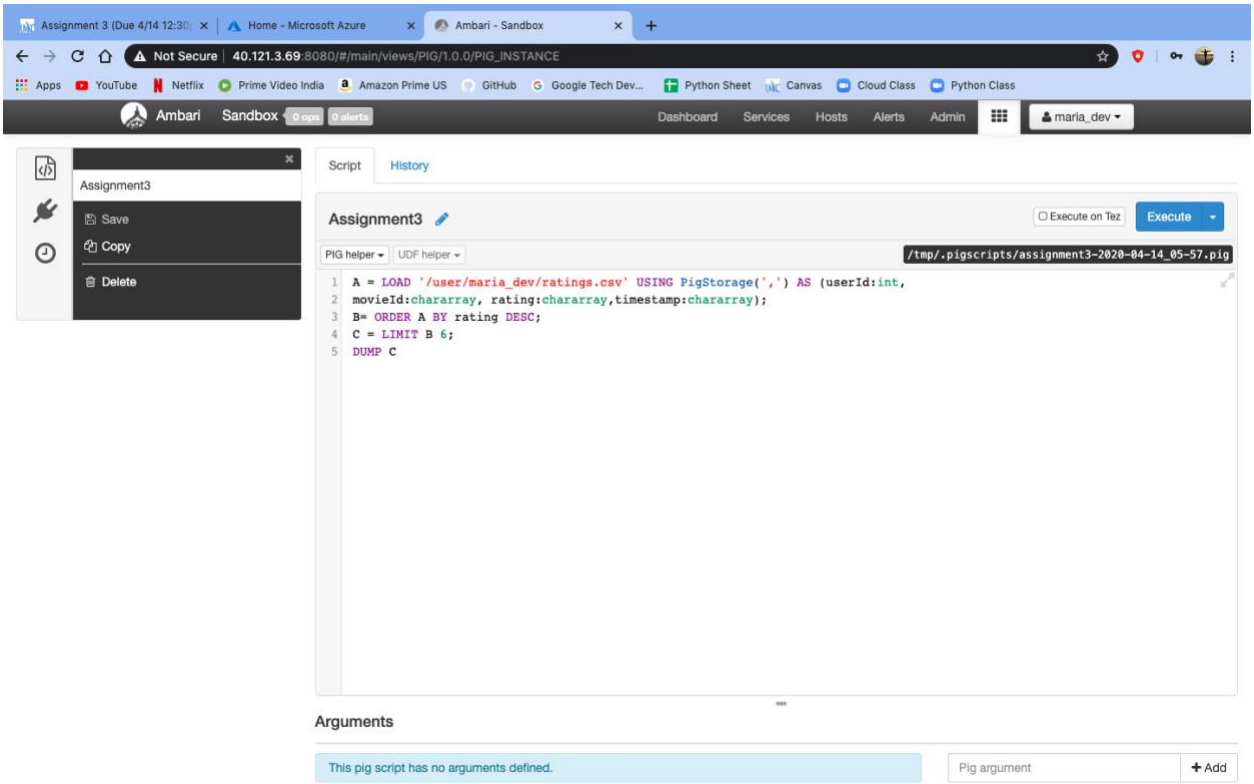
Step 6: To explore Ambari -> Pig View and find top 5 highly rated movies.

This is the file view, where we upload the csv files such as ratings.csv here,

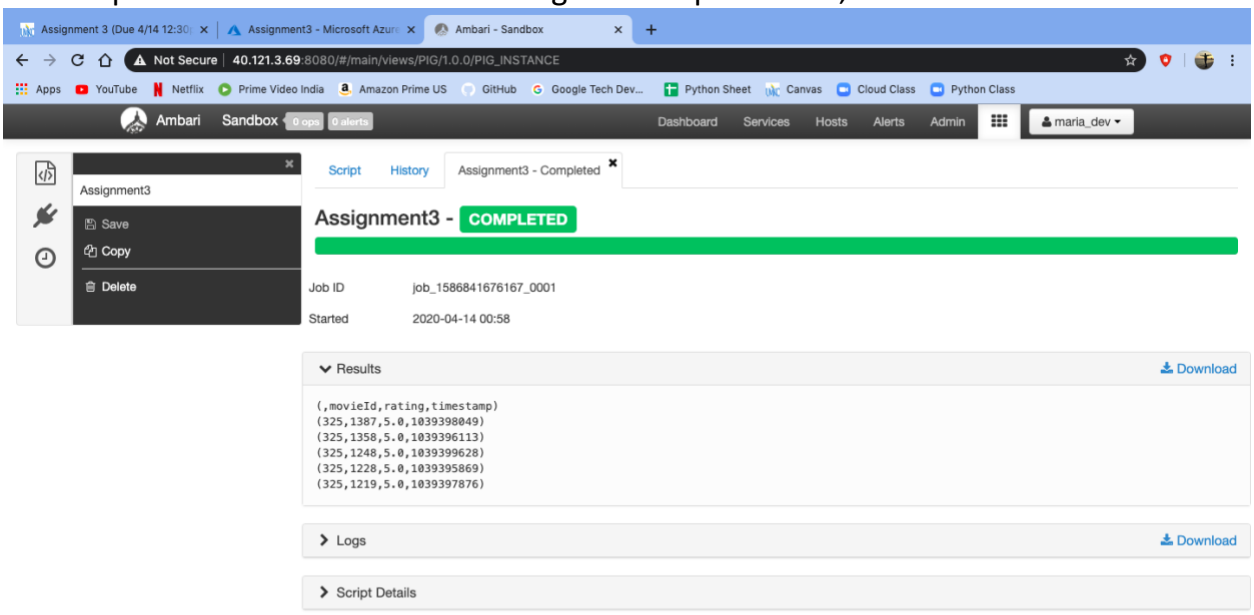


In the Pig view, a new script is created and we write our code here,

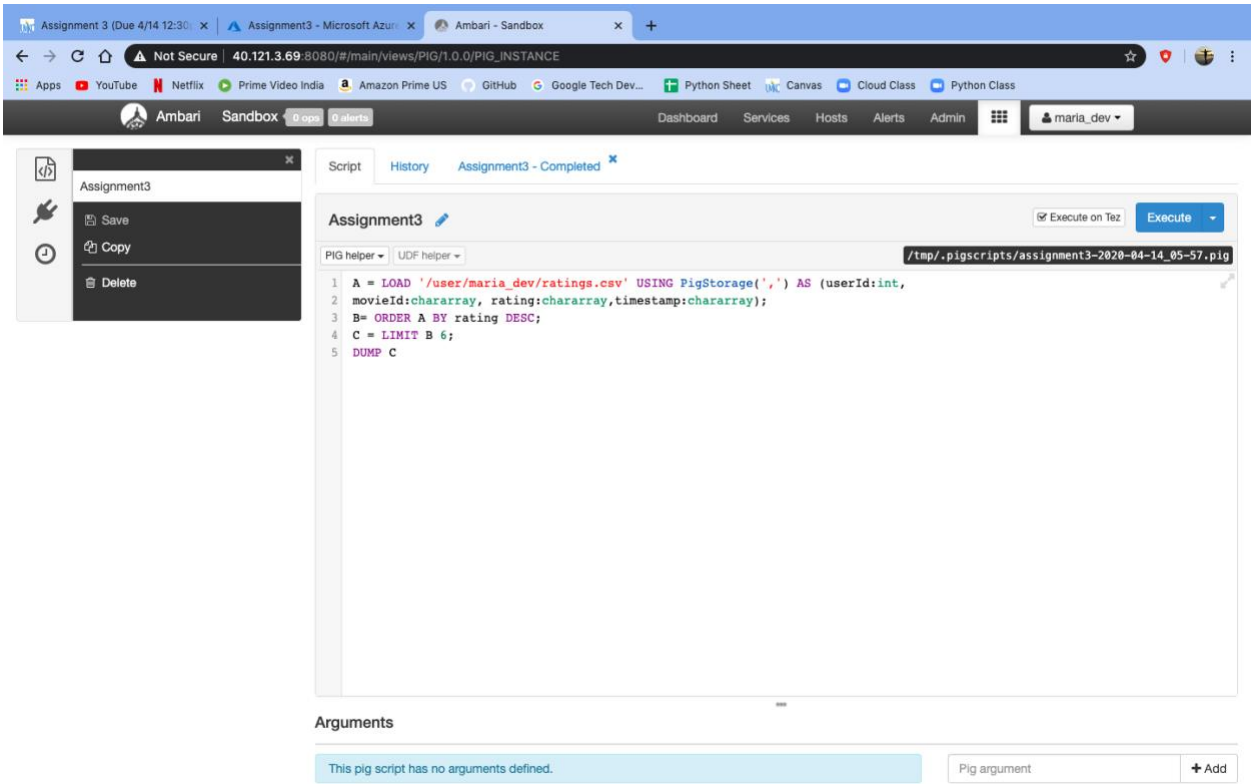


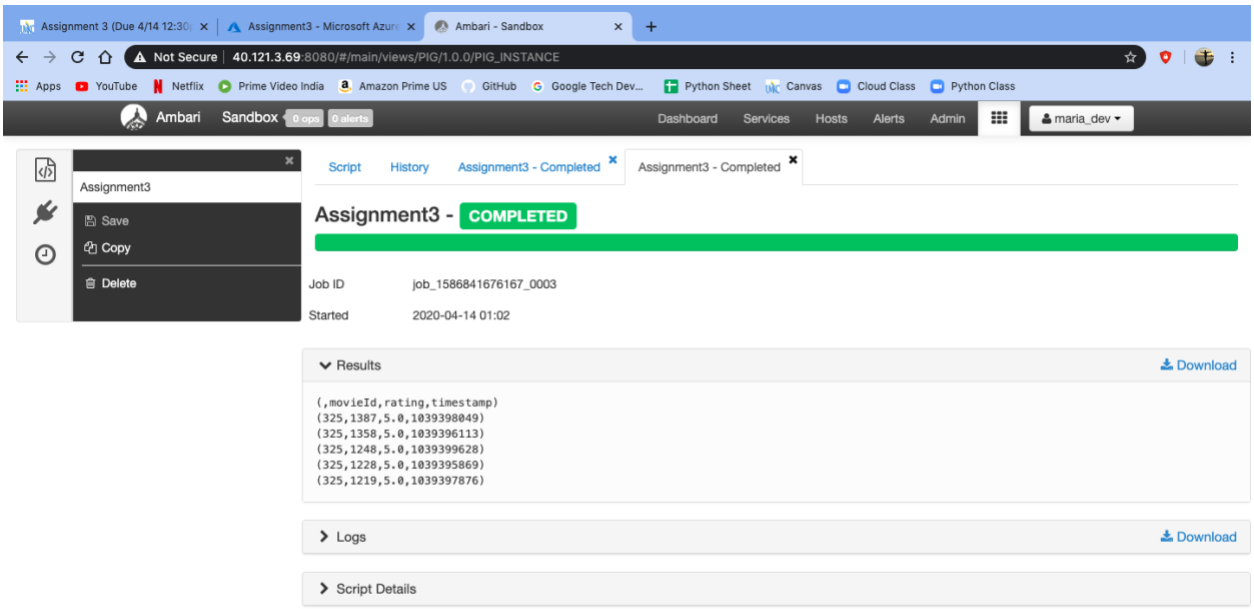


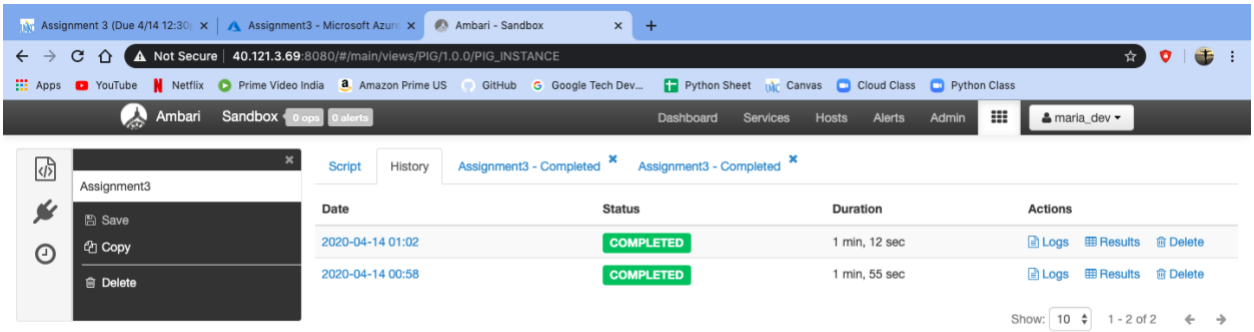
Our script file is executed here and we get the Top 5 results,



Step 7: Do the above Step 6 with Tez enabled,







With the Tez enabled, the script was executed in 1 min 12 secs and with the Mapreduce it was executed in 1 min 55 sec. Hence, *Tez runs faster than MapReduce.*