Azure Databricks:

Pre-Requisites:

1. Outlook User account or an email ID with Microsoft access (School or office)
2. Active Credit Card (without any Azure account link)
3. Active Mobile Number
4. Address

Portals:

1. www.azure.com
2. [portal.azure.com](http://portal.azure.com/)

Azure Services:

1. Resource Group
2. Virtual Machine
3. Virtual Network
4. Firewall
5. Storage Account
6. Container
7. Blob Storage
8. Azure SQL
9. Azure Databricks
10. IP Address
11. Location

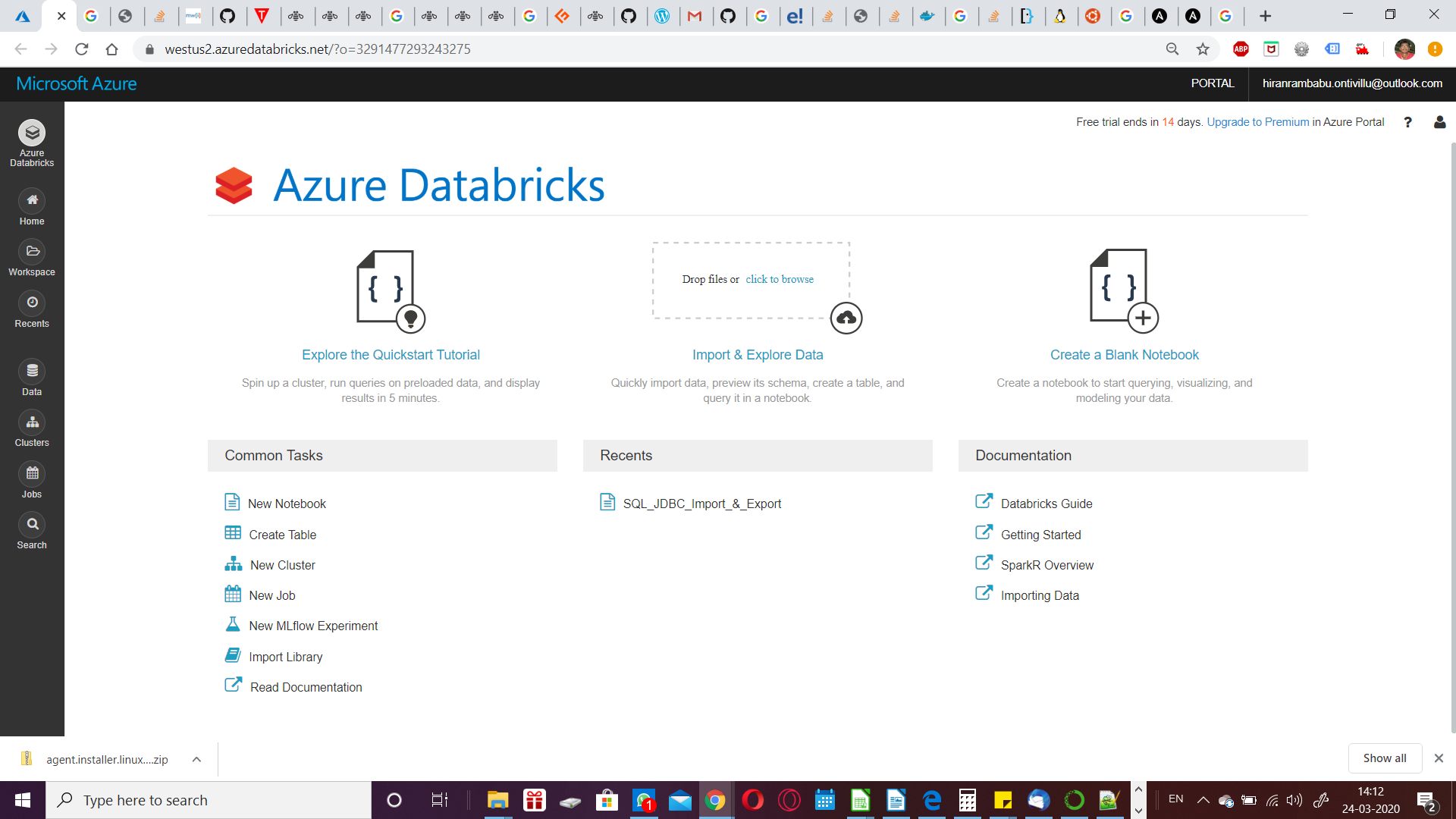
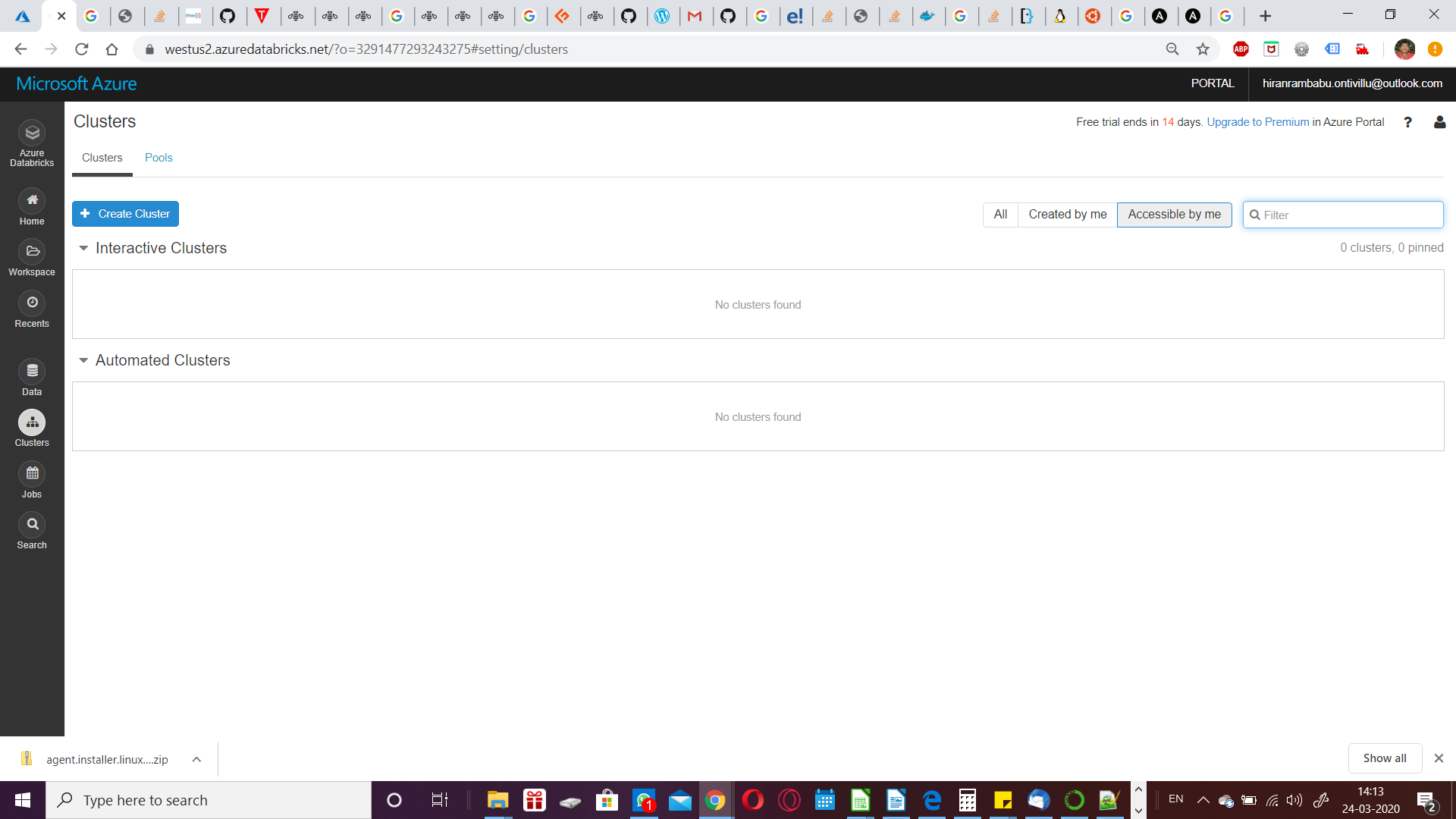
Azure Keywords:

1. Virtual Machine
2. Subscription
3. Resource Group
4. Compute
5. Public Endpoint, Private Endpoint
6. Tags
7. Services
8. Resources
9. Cost Management
10. Location
11. Region
12. Redundant
13. Access Keys
14. Geo Replication
15. Blade
16. Icon
17. Dashboard

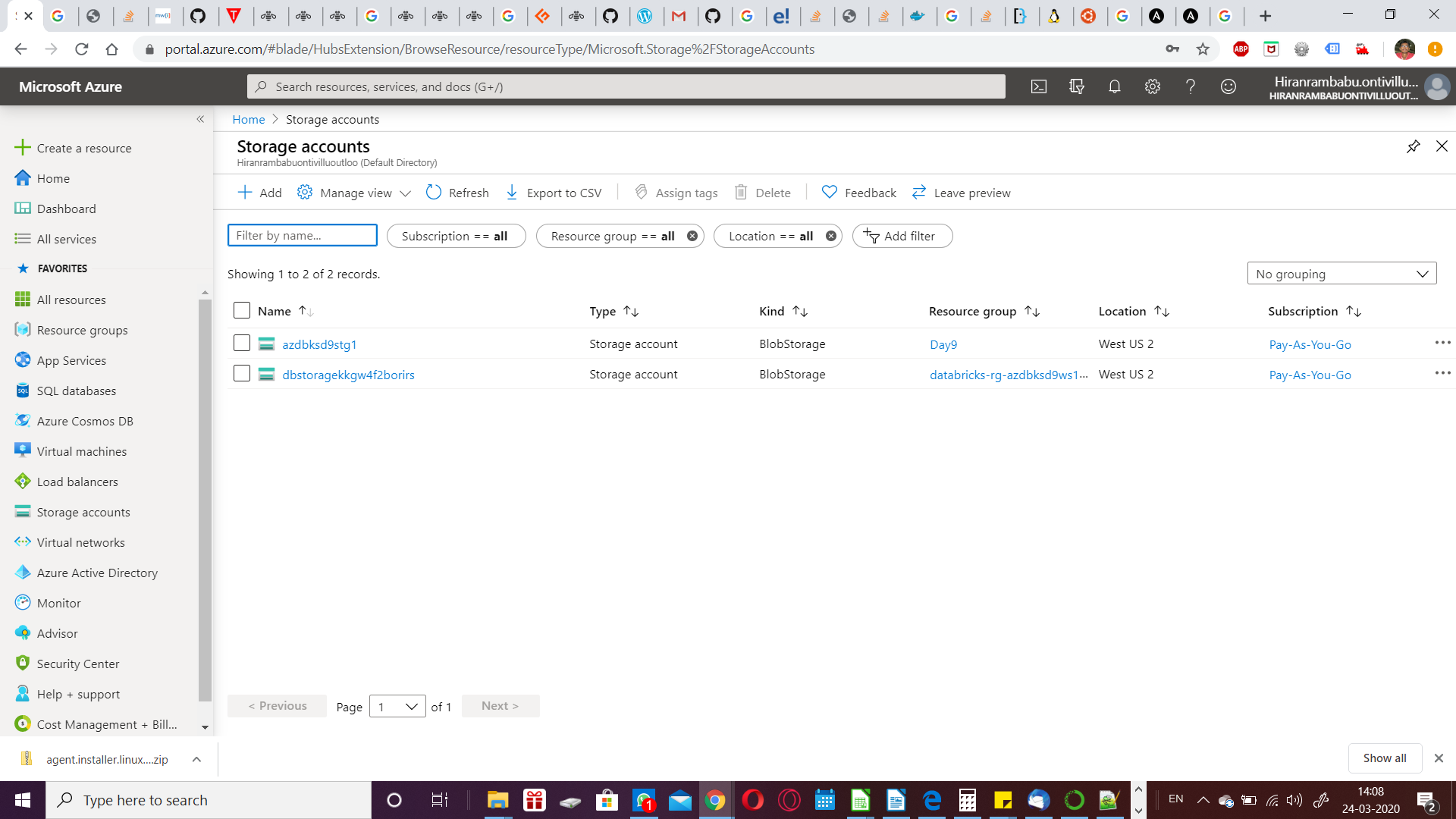
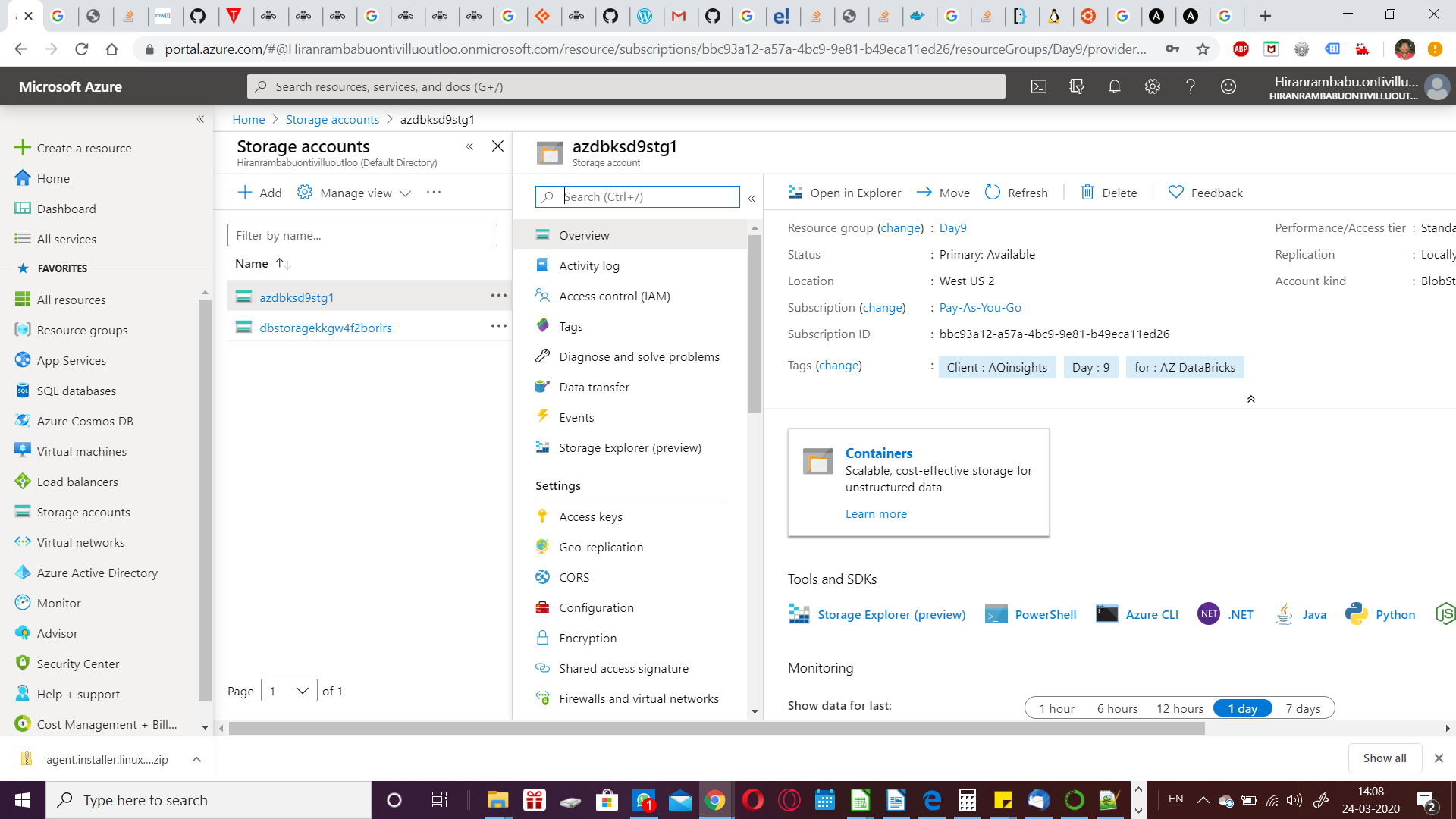
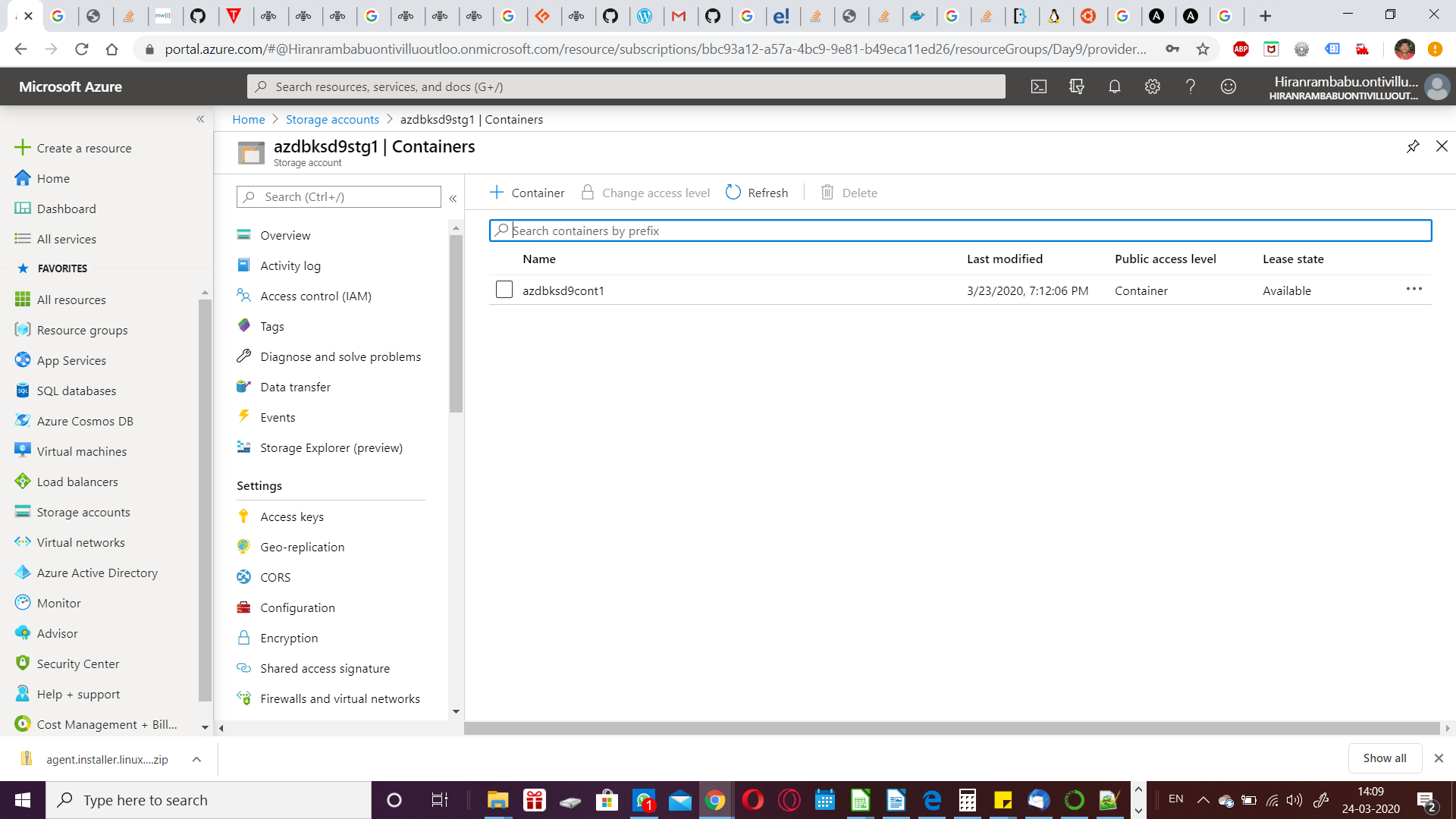
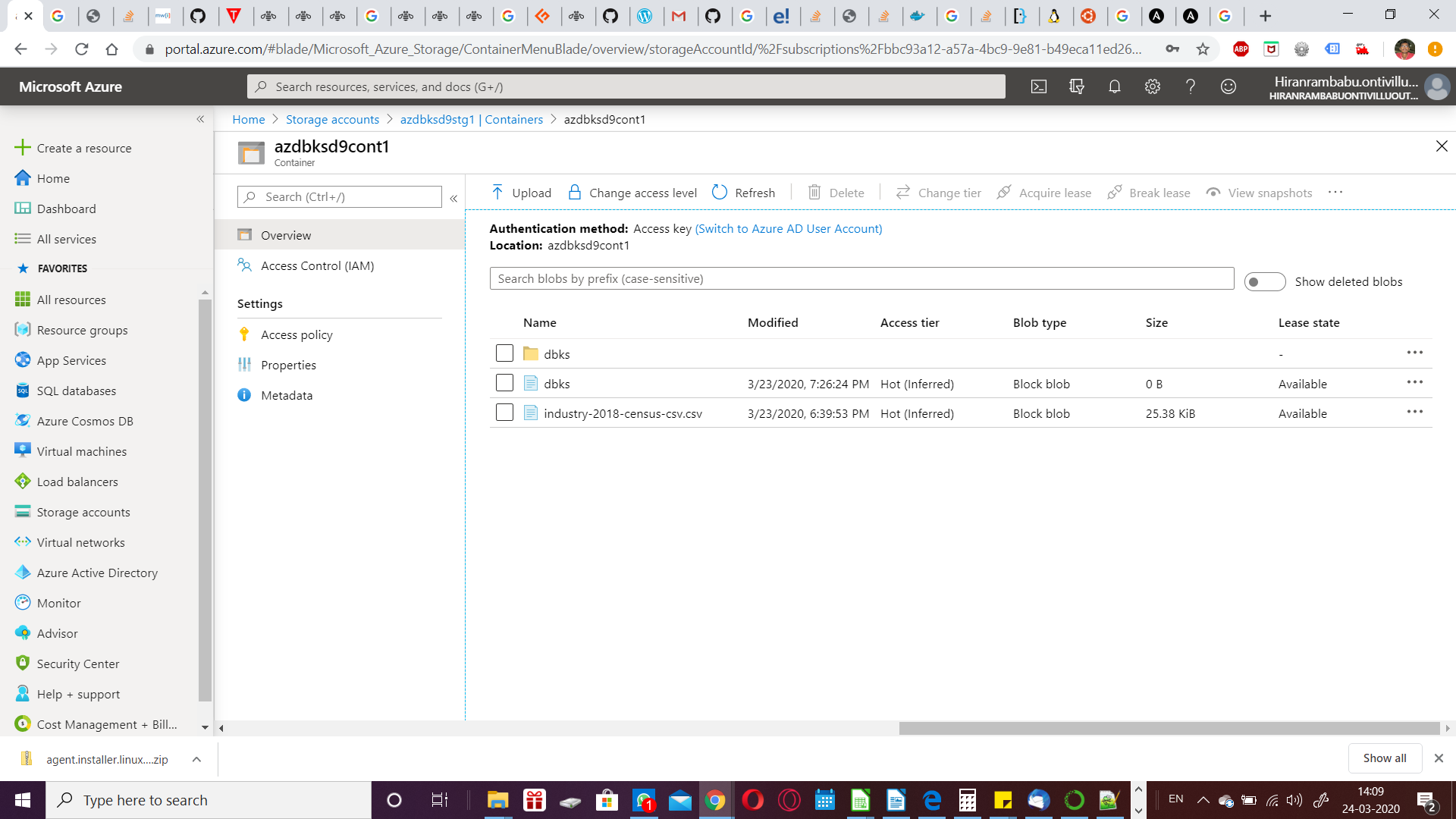
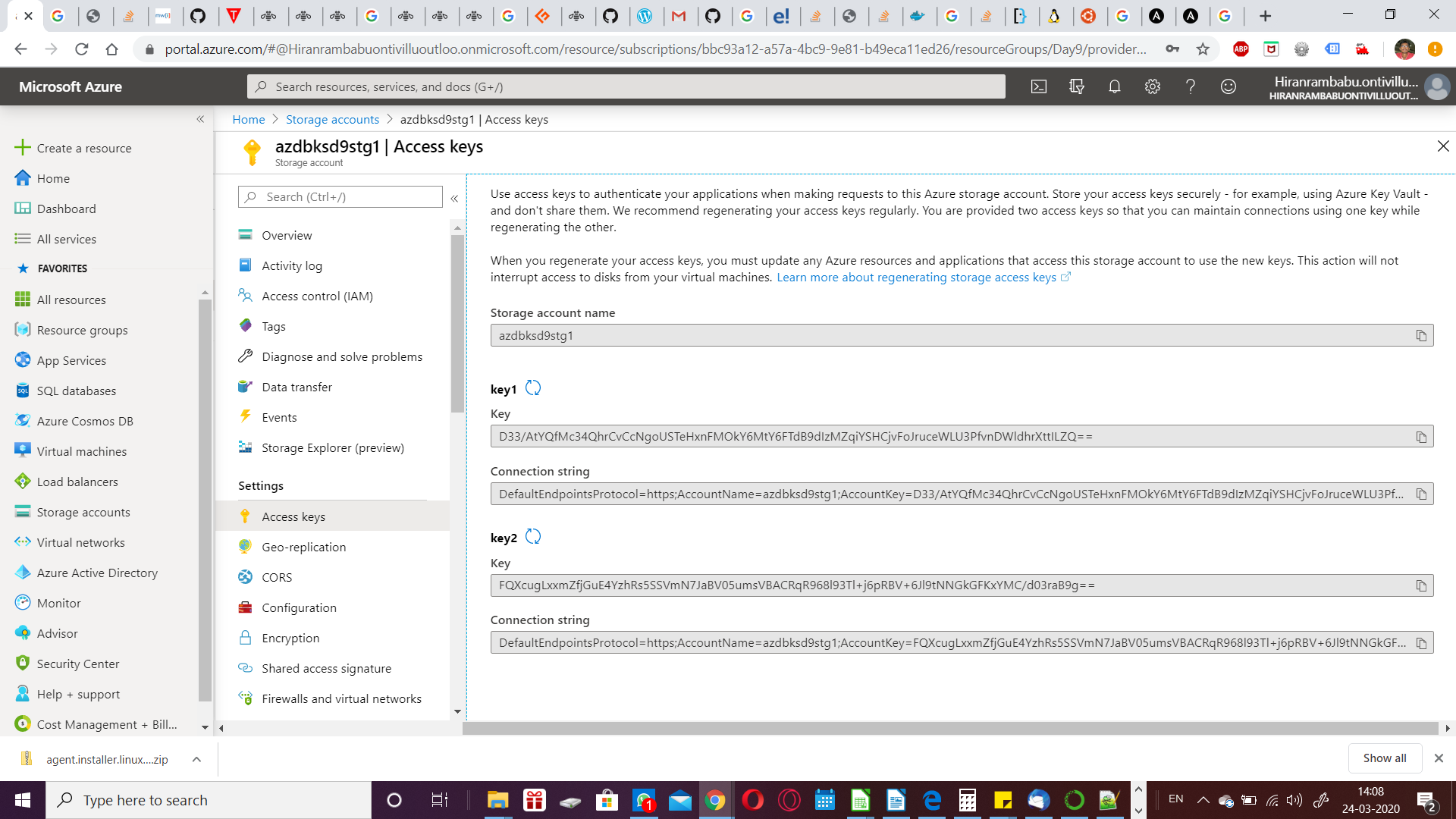
Azure Databricks:

1. Workspace
2. Cluster
3. Notebook
4. Python
5. R
6. SQL
7. SCALA
8. API
9. JOB
10. Schedule
11. Cell
12. Jupyter
13. Data
14. Datasets
15. Database
16. Data Frame
17. CSV
18. JDBC
19. Data lake
20. Git
21. Table
22. In-Memory
23. Pool
24. Attach/Detach
25. Interactive Cluster
26. Automated Cluster

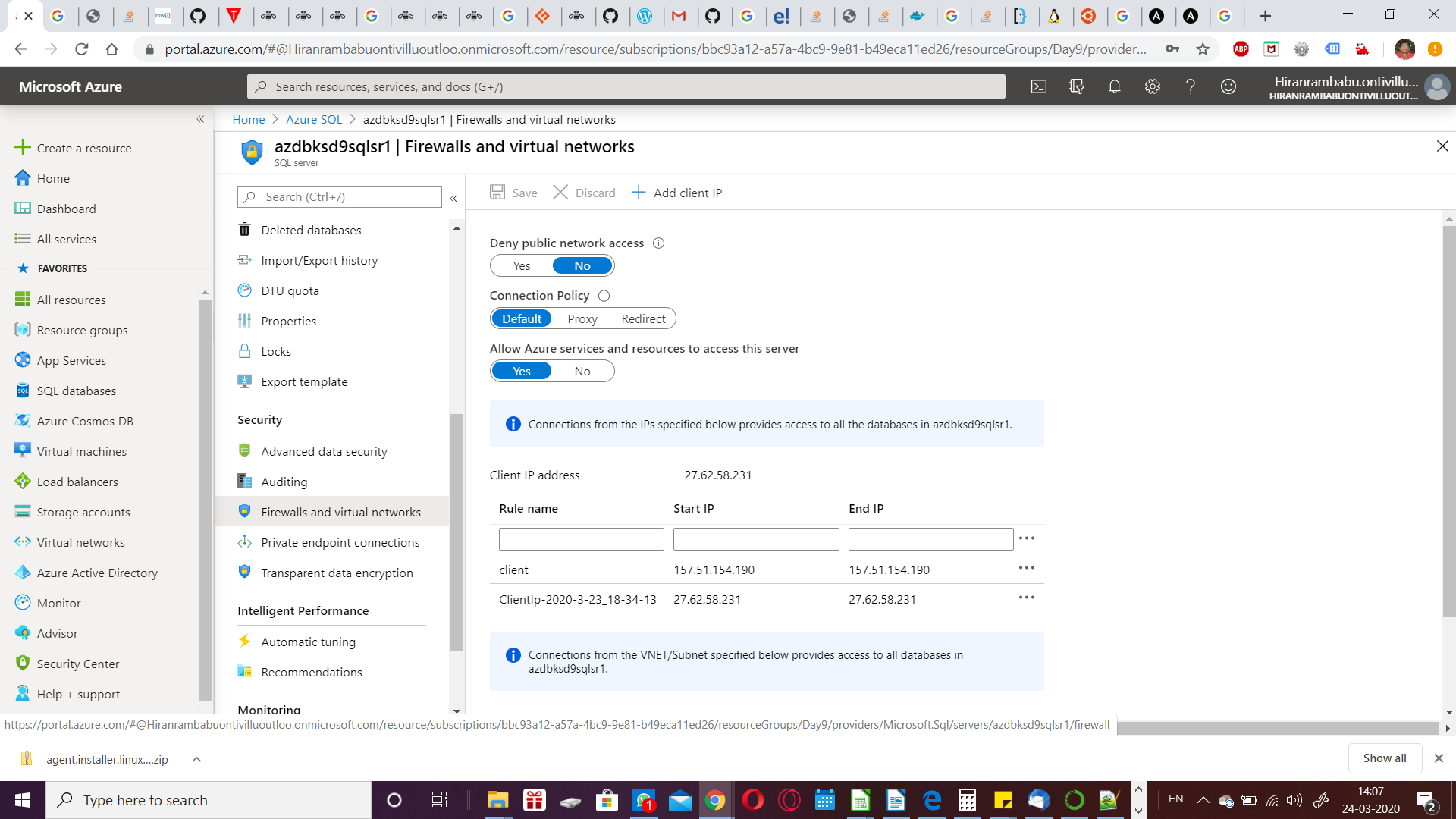
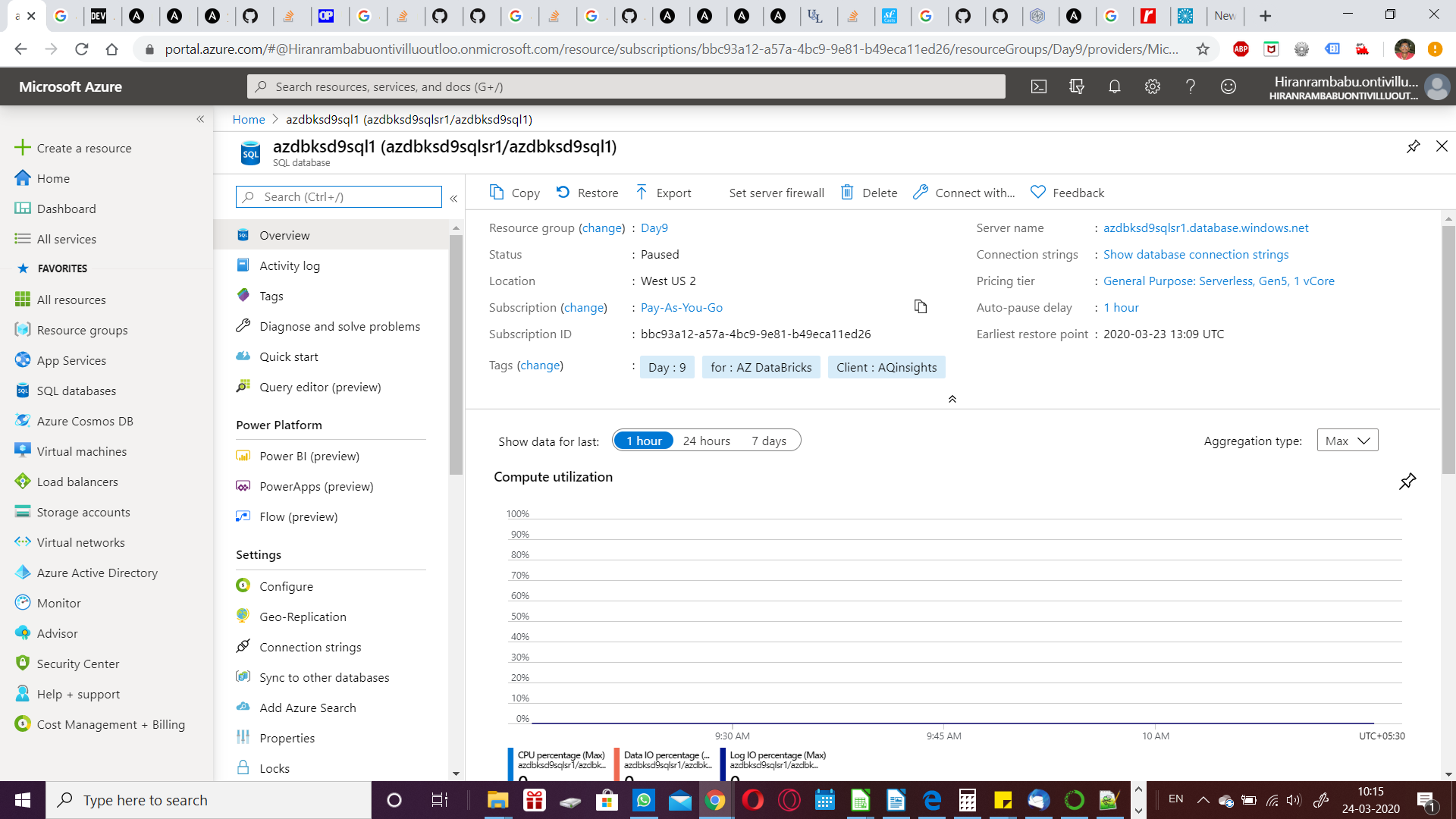
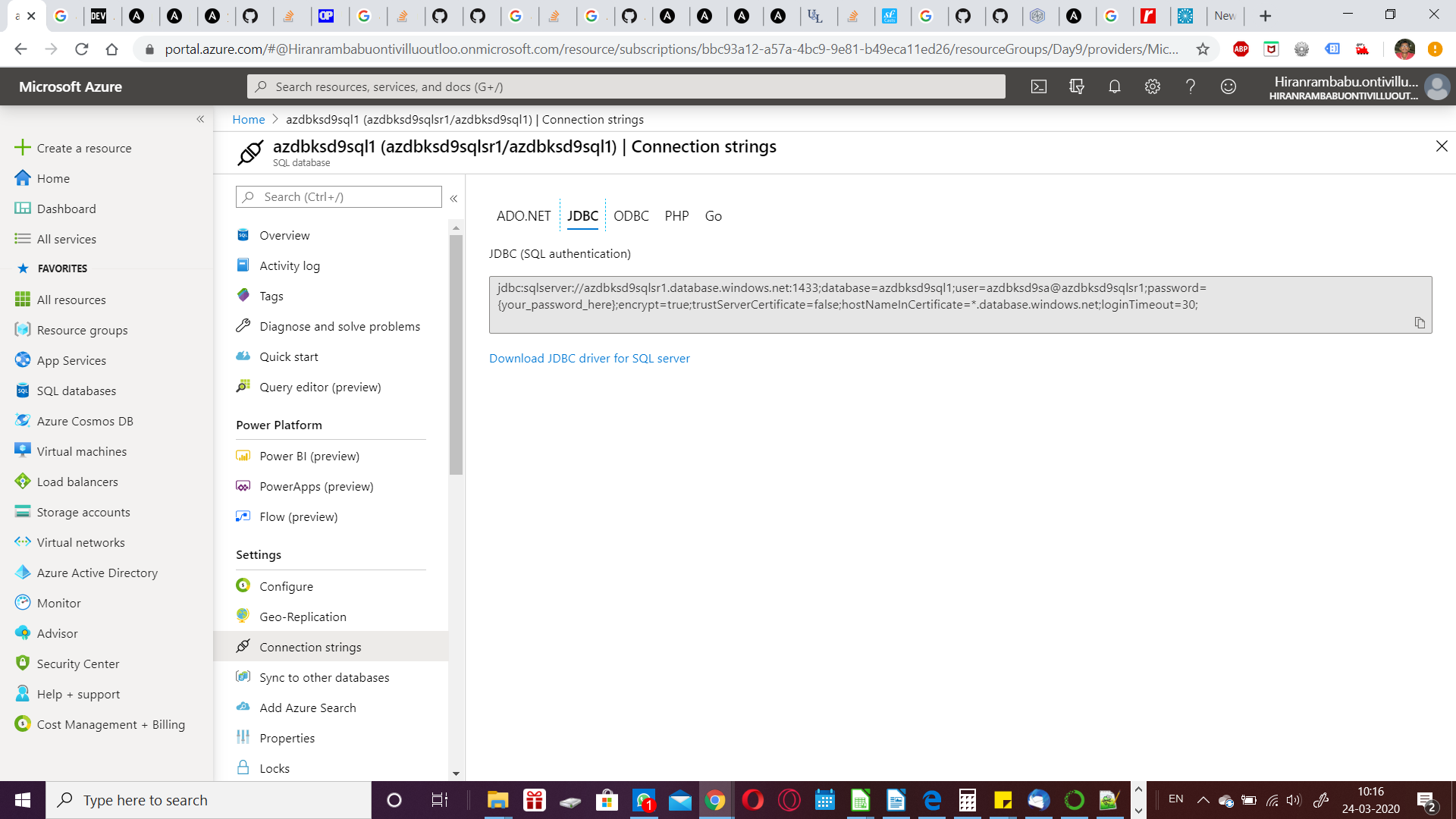
Steps:

1. open www.azure.com
2. click on “Start free”
3. Please enter all the details
4. validate
5. create account
6. Open [portal.azure.com](http://portal.azure.com/)
7. Login with the newly created azure account
8. type “Cost Management” in Search bar
9. select the active subscription (Free Subscription or Azure Subscription)
10. click on orange banner which says $200 credit or ₹13000
11. Upgrade to Pay as you go with “No support”
12. Click on “All Services” from the left hanging menu
13. select “resource group”
    1. Click +Add
    2. Select Subscription
    3. Type a name for Resource Group
    4. Select Region
    5. Click Tags
       1. provide Key and Value for tags
    6. Select Review + Create
    7. upon successful validation, please select Create
14. Upon creation of Resource Group, Type “Azure Databricks” in Search bar
    1. Click +Add
    2. Select Subscription
    3. Select the earlier created Resource Group by its name
    4. Enter a name for the Workspace Name
    5. Select Location
    6. Select Pricing Tier as “Trial (Premium – 14-days Free DBUs)”
    7. Click Networking
    8. Click Tags
       1. provide Key and Value for tags
    9. Select Review + Create
    10. upon successful validation, please select Create
15. Upon creation of Azure Databricks resource, select “go to resource”
    1. Click “Launch Workspace” - This will open a new tab
    2. Click login with azure credentials
    3. Click Cluster from left menu
    4. Click +Create Cluster
       1. Enter Cluster Name
       2. Select Standard for Cluster Mode
          1. Standard – Single job and Single user at a time – Costs low
          2. High Concurrency – Multiple user \* Multiple jobs at a time – Costs high
       3. Select None for pool
       4. Select the appropriate Databricks Runtime Version
       5. Enable or disable auto scaling as per the requirement
          1. Enable – Cluster Nodes will Expand or Shrink as per the load; Min & Max node needs to be set
          2. Disable – Cluster nodes will be static; number of nodes only to be set
       6. Enable or Disable Terminate after [x] minutes of inactivity
          1. Enable – automatically terminates the cluster nodes upon non-usage; saves cost
          2. Disable – Does not terminates the cluster nodes at any time; saves time
       7. Select the appropriate Worker Type as per the requirement
          1. This is the Compute selection
          2. Enter the Min and Max nodes if Autoscaling selected
          3. Enter the Workers Count, if Autoscaling is unselected
       8. select the appropriate Drive Type as per the requirement
          1. This is the Compute selection
       9. Click “Create Cluster”
    5. Click Azure Databricks from left menu
    6. Click “New Notebook”
       1. Provide a name
       2. Select the default language
       3. Select respective Cluster
    7. Start writing the respective code

Steps for creation of Storage:

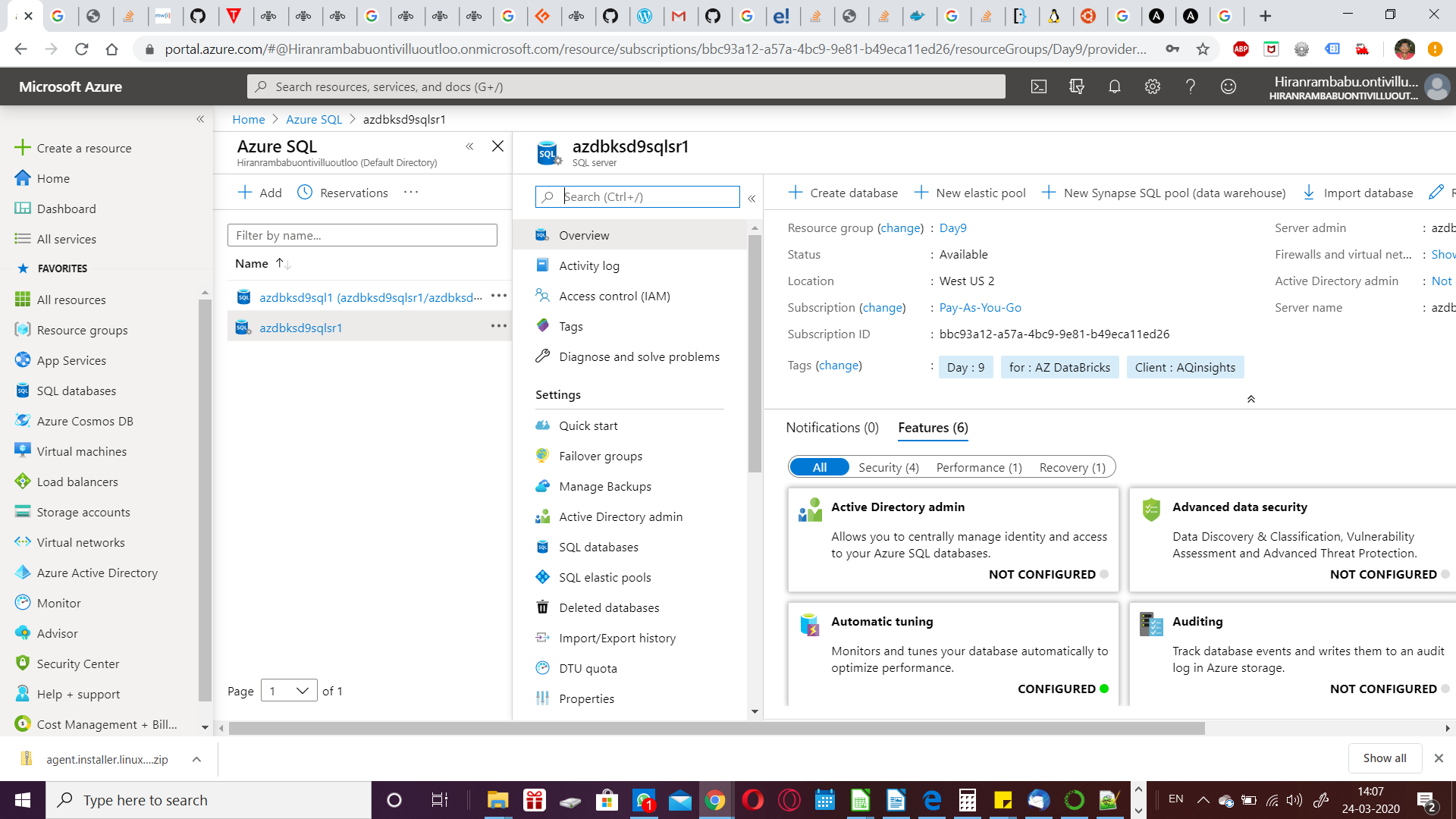
1. Open and Login [portal.azure.com](http://portal.azure.com/)
2. Type “Storage accounts” in Search bar
   1. Click +Add
   2. Select Subscription
   3. Type a name for Storage account
   4. Select Location
   5. Select Standard for Performance
   6. Select Blobstorage as Account kind
   7. Select “Locally-redundant storage (LRS) for Replication
   8. Select Hot for Access Tier
   9. Click Networking
   10. Select Public Endpoint (all networks)
   11. leave remaining all as default
   12. Click Tags
       1. provide Key and Value for tags
   13. Select Review + Create
3. Upon successful creation of Storage account, Select Goto Resource
   1. Click Containers
   2. Click + Container
      1. Provide a Name
      2. Select Container (anonymous read access for containers and blobs) as Public access level
      3. Click Create
   3. Click Upload
      1. Select a data file (.csv) for upload
   4. Click on Storage accounts and select the appropriate Storage account name
   5. Click on Access Keys from left menu
      1. Note the Storage Account Name
      2. Note the Key 1
      3. Note the Connection String

Steps for creation of Storage:

1. Type “Azure SQL” in Search bar
   1. Click +Add
      1. Under SQL databases, select single database as Resource Type
      2. Click Create
         1. Select Subscription
         2. Select appropriate Resource Group
         3. Type a name for Database
         4. Select Server or click Create New
            1. Provide a name for Server
            2. Provide an admin username
            3. Provide a password for admin username
            4. Re-type the provided password
            5. Select appropriate location
         5. Select no for SQL elastic pool
         6. Click Configure database for Compute + Storage
            1. Select Serverless for compute tier
            2. Click Apply
         7. Select Networking
         8. Select Public Endpoint
         9. Click Additional Setting
         10. Select Not Now for Enable advanced data security and leave remaining as it is
         11. Click Tags
             1. provide Key and Value for tags
         12. Select Review + Create
      3. Upon successful creation of Azure SQL, Type “Azure SQL” in Search bar
      4. Select the newly created SQL database server from the list (select the one with Resource type as SQL Server)
         1. Select Firewalls and virtual networks
            1. Confirm “No” for deny Public network access
            2. Confirm “Default” for Connection policy
            3. Confirm “Yes” for Allow Azure services and resources to access this server
            4. Type a name to rule
            5. Type Start IP and End IP as the above mentioned Client IP address
            6. Click Save
      5. Upon successful save, Type “Azure SQL” in Search bar
      6. Select the newly created SQL database from the list (select the one with Resource type as SQL database)
      7. Select Connection strings from left menu
         1. select JDBC
         2. Copy the code from the box under “JDBC (SQL Authentication)”
      8. Select query editor from left menu
         1. Provide the appropriate Login and password for SQL Authentication
         2. type the sql statement to create a new table and insert some values into the table
         3. click Run
         4. note down the table name, column name, schema

<https://github.com/hrb1989/AZdbksaqinsghts>

SQL:



Azure Databricks:

