


# Tutorial 1

# Outline

Creat an AWS Account

Data preprocessing

# Create AWS account

English ▾

## Create an AWS account

Email address

Password

Confirm password

AWS account name ⓘ

Continue

[Sign in to an existing AWS account](#)

© 2019 Amazon Web Services, Inc. or its affiliates.  
All rights reserved.  
[Privacy Policy](#) | [Terms of Use](#)

### AWS Accounts Include 12 Months of Free Tier Access

Including use of Amazon EC2, Amazon S3, and Amazon DynamoDB  
Visit [aws.amazon.com/free](https://aws.amazon.com/free) for full offer terms

# Contact Information



English ▾

## Contact Information

All fields are required.

Please select the account type and complete the fields below with your contact details.

Account type ⓘ  
☐ Professional ☒ Personal

Full name

Phone number

Country/Region

Address

City

State / Province or region

Postal code

☒ Check here to indicate that you have read and agree to the terms of the AWS Customer Agreement

Create Account and Continue

# Payment Information



English ▾

## Payment Information

We use your payment information to verify your identity and only for usage in excess of the [AWS Free Tier Limits](#). We will not charge you for usage below the AWS Free Tier Limits. For more information, see the [frequently asked questions](#).

**i** When you submit your payment information, we will charge \$1 USD/EUR to your credit card as a verification charge to ensure your card is valid. The amount may show as pending in your credit card statement for 3-5 days until the verification is completed, at which time the charge will be removed. You may be redirected to your bank website to authorize the verification charge.

Credit card number

Expiration date

Cardholder's name

Billing address

☐ Use a new address

Verify and Add

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved.

[Privacy Policy](#) | [Terms of Use](#) | [Sign Out](#)

# Select a Support Plan



English ▾

## Select a Support Plan

AWS offers a selection of support plans to meet your needs. Choose the support plan that best aligns with your AWS usage. [Learn more](#)



### Basic Plan

Free

- Included with all accounts
- 24/7 self-service access to forums and resources
- Best practice checks to help improve security and performance
- Access to health status and notifications



### Developer Plan

From \$29/month

- For early adoption, testing and development
- Email access to AWS Support during business hours
- 1 primary contact can open an unlimited number of support cases
- 12-hour response time for nonproduction systems



### Business Plan

From \$100/month

- For production workloads & business-critical dependencies
- 24/7 chat, phone, and email access to AWS Support
- Unlimited contacts can open an unlimited number of support cases
- 1-hour response time for production systems

### Need Enterprise level support?

Contact your account manager for additional information on running business and mission critical-workloads on AWS (starting at \$15,000/month). [Learn more](#)

© 2019 Amazon Web Services, Inc. or its affiliates. All rights reserved.

[Privacy Policy](#) [Terms of Use](#) [Sign Out](#)

# Sign in



## Sign in

Email address of your AWS account

Or to sign in as an IAM user, enter your account ID or account alias instead.

Next

New to AWS?

Create a new AWS account



## AWS Accounts Include 12 Months of Free Tier Access

Including use of Amazon EC2,  
Amazon S3, and Amazon DynamoDB

Visit [aws.amazon.com/free](https://aws.amazon.com/free) for full offer terms

### About Amazon.com Sign in

Amazon Web Services uses information from your Amazon.com account to identify you and allow access to Amazon Web Services. Your use of this site is governed by our Terms of Use and Privacy Policy linked below. Your use of Amazon Web Services products and services is governed by the AWS Customer Agreement linked below unless you have entered into a separate agreement with Amazon Web Services or an AWS Value Added Reseller to purchase these products and services. The AWS Customer Agreement was updated on March 31, 2017. For more information about these updates, see [Recent Changes](#).

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Terms of Use](#) | [Privacy Policy](#) | [AWS Customer Agreement](#)

English ▼

# Launch a virtual machine with EC2

## AWS Management Console

### AWS services

#### Find Services

You can enter names, keywords or acronyms.

Example: Relational Database Service, database, RDS

#### Recently visited services

 CodeStar

 Billing

 IAM

 Support

#### ► All services

### Access resources on the go



Access the Management Console using the AWS Console Mobile App. [Learn more](#)

### Explore AWS

#### Amazon SageMaker

Machine learning for every developer and data scientist. [Learn more](#)

#### Amazon GuardDuty

Protect your AWS accounts and workloads with intelligent threat detection. [Learn more](#)

#### CloudEndure Migration

Re-host a large number of machines to AWS without worrying about compatibility, performance disruption, or long cutover windows. [Get started](#)

#### EC2 Spot Instances

Run fault-tolerant workloads on Spot Instances and save up to 90% on compute. [Learn more](#)

### Build a solution

Get started with simple wizards and automated workflows.

#### Launch a virtual machine

With EC2

2-3 minutes



#### Build a web app

With Elastic Beanstalk

6 minutes



#### Build using virtual servers

With Lightsail

1-2 minutes



#### Connect an IoT device

With AWS IoT

5 minutes



#### Start a development project

With CodeStar

5 minutes



#### Register a domain

With Route 53

3 minutes



#### Deploy a serverless microservice

With Lambda, API Gateway

2 minutes



#### Host a static web app


With AWS Amplify Console

5 minutes





# launch an instance (ubuntu 18.04 or ubuntu 16.04)

 Services ▾ Resource Groups ▾ ★

comp4331 ▾ Ohio ▾ Support ▾

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

## Step 1: Choose an Amazon Machine Image (AMI)

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. You can select an AMI provided by AWS, our user community, or the AWS Marketplace; or you can select one of your own AMIs.

Q Search for an AMI by entering a search term e.g. "Windows" X


Quick Start

My AMIs

AWS Marketplace

Community AMIs

☐ Free tier only ⓘ

**Amazon Linux**  
Free tier eligible


**Amazon Linux 2 AMI (HVM), SSD Volume Type** - ami-00c037f7f2ec15c3 (64-bit x86) / ami-014d175d64de0a174 (64-bit Arm)

Amazon Linux 2 comes with five years support. It provides Linux kernel 4.14 tuned for optimal performance on Amazon EC2, systemd 219, GCC 7.3, Glibc 2.26, Binutils 2.29.1, and the latest software packages through extras.

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

☒ 64-bit (x86)  
☐ 64-bit (Arm)

Select

**Amazon Linux**  
Free tier eligible


**Amazon Linux AMI 2018.03.0 (HVM), SSD Volume Type** - ami-0c64dd618a49aeec8

The Amazon Linux AMI is an EBS-backed, AWS-supported image. The default image includes AWS command line tools, Python, Ruby, Perl, and Java. The repositories include Docker, PHP, MySQL, PostgreSQL, and other packages.

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

64-bit (x86)

Select

**Red Hat**  
Free tier eligible


**Red Hat Enterprise Linux 8 (HVM), SSD Volume Type** - ami-0520e698dd500b1d1 (64-bit x86) / ami-0099847d600887c9f (64-bit Arm)

Red Hat Enterprise Linux version 8 (HVM), EBS General Purpose (SSD) Volume Type

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

☒ 64-bit (x86)  
☐ 64-bit (Arm)

Select

**SUSE Linux**  
Free tier eligible


**SUSE Linux Enterprise Server 15 SP1 (HVM), SSD Volume Type** - ami-0e0bae59dc35fe89a (64-bit x86) / ami-0b49a8f443e46ff20 (64-bit Arm)

SUSE Linux Enterprise Server 15 Service Pack 1 (HVM), EBS General Purpose (SSD) Volume Type. Public Cloud, Advanced Systems Management, Web and Scripting, and Legacy modules enabled.

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

☒ 64-bit (x86)  
☐ 64-bit (Arm)

Select

**Ubuntu**  
Free tier eligible


**Ubuntu Server 18.04 LTS (HVM), SSD Volume Type** - ami-05c1fa8df71875112 (64-bit x86) / ami-0606a0d9f566249d3 (64-bit Arm)

Ubuntu Server 18.04 LTS (HVM), EBS General Purpose (SSD) Volume Type. Support available from Canonical (<http://www.ubuntu.com/cloud/services>).

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

☒ 64-bit (x86)  
☐ 64-bit (Arm)

Select

**Amazon RDS**

**Are you launching a database instance? Try Amazon RDS.**

Amazon Relational Database Service (RDS) makes it easy to set up, operate, and scale your database on AWS by automating time-consuming database management tasks. With RDS, you can easily deploy **Amazon Aurora, MariaDB, MySQL, Oracle, PostgreSQL, and SQL Server** databases on AWS. Aurora is a MySQL- and PostgreSQL-compatible, enterprise-class database at 1/10th the cost of commercial databases. [Learn more about RDS](#)

Launch a database using RDS

Feedback

English (US)

© 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

# launch an instance

aws

Services

Resource Groups

comp4331

Ohio

Support

1. Choose AMI2. Choose Instance Type3. Configure Instance4. Add Storage5. Add Tags6. Configure Security Group7. Review

## Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. [Learn more](#) about instance types and how they can meet your computing needs.

Filter by: All Instance typesCurrent generationShow/Hide Columns

Currently selected: t2.micro (Variable ECUs, 1 vCPUs, 2.5 GHz, Intel Xeon Family, 1 GiB memory, EBS only)


	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance	IPv6 Support
<input type="checkbox"/>	General purpose	t2.nano	1	0.5	EBS only	-	Low to Moderate	Yes
<input checked="" type="checkbox"/>	General purpose	t2.micro Free tier eligible	1	1	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.small	1	2	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.medium	2	4	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.large	2	8	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.xlarge	4	16	EBS only	-	Moderate	Yes
<input type="checkbox"/>	General purpose	t2.2xlarge	8	32	EBS only	-	Moderate	Yes
<input type="checkbox"/>	General purpose	t3a.nano	2	0.5	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	General purpose	t3a.micro	2	1	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	General purpose	t3a.small	2	2	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	General purpose	t3a.medium	2	4	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	General purpose	t3a.large	2	8	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	General purpose	t3a.xlarge	4	16	EBS only	Yes	Up to 5 Gigabit	Yes

CancelPreviousReview and LaunchNext: Configure Instance Details

FeedbackEnglish (US)

© 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy PolicyTerms of Use

# launch an instance

Services ▾Resource Groups ▾★


comp4331 ▾Ohio ▾Support ▾

1. Choose AMI2. Choose Instance Type3. Configure Instance4. Add Storage5. Add Tags6. Configure Security Group7. Review

## Step 7: Review Instance Launch

Please review your instance launch details. You can go back to edit changes for each section. Click **Launch** to assign a key pair to your instance and complete the launch process.

▼ AMI Details

Free tier eligible

**Amazon Linux 2 AMI (HVM), SSD Volume Type - ami-00c03f7f7f2ec15c3**  
Amazon Linux 2 comes with five years support. It provides Linux kernel 4.14 tuned for optimal performance on Amazon EC2, systemd 219, GCC 7.3, Glibc 2.26, Binutils 2.29.1, and the latest software packages through extras.  
Root Device Type: ebsVirtualization type: hvm

[Edit AMI](#)

▼ Instance Type

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
t2.micro	Variable	1	1	EBS only	-	Low to Moderate

[Edit instance type](#)

▼ Security Groups

**Security group name**

launch-wizard-1

**Description**

launch-wizard-1 created 2019-09-16T16:52:13.989+08:00

Type ⓘ	Protocol ⓘ	Port Range ⓘ	Source ⓘ	Description ⓘ
This security group has no rules				

[Edit security groups](#)

► Instance Details

[Edit instance details](#)

► Storage

[Edit storage](#)

► Tags

[Edit tags](#)

Cancel

Previous

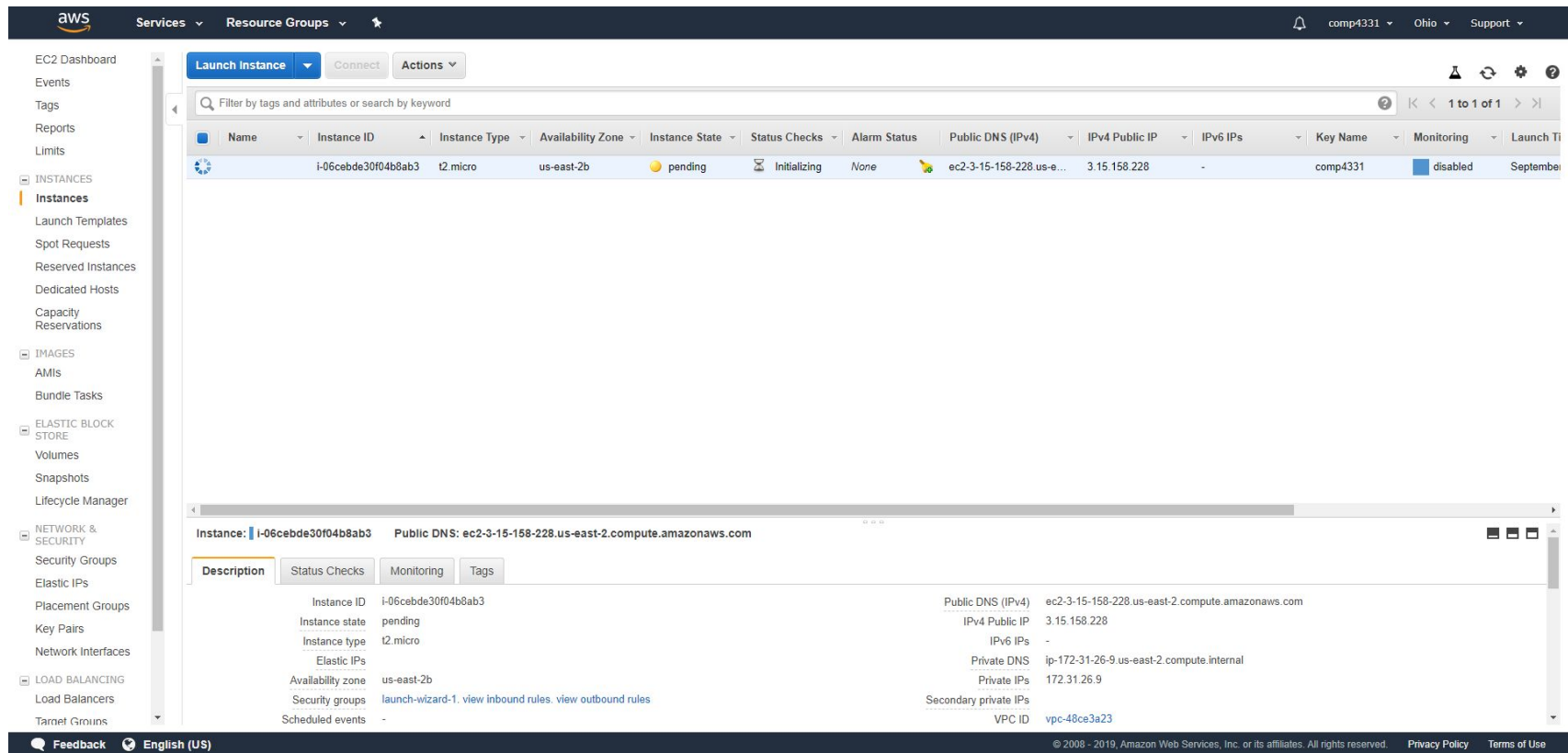
**Launch**

Feedback

English (US)

© 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

# launch an instance



The screenshot displays the AWS Management Console interface. At the top, the navigation bar includes the AWS logo, 'Services', 'Resource Groups', and a user profile section with a notification bell, account ID 'comp4331', region 'Ohio', and a 'Support' link. The left-hand navigation pane lists various AWS services, with 'INSTANCES' and 'Instances' highlighted. The main content area features a 'Launch Instance' button, a 'Connect' button, and an 'Actions' dropdown. Below these is a search bar and a table of instances. The table has columns for Name, Instance ID, Instance Type, Availability Zone, Instance State, Status Checks, Alarm Status, Public DNS (IPv4), IPv4 Public IP, IPv6 IPs, Key Name, Monitoring, and Launch Time. A single instance is listed with ID 'i-06cebbe30f04b8ab3', type 't2.micro', and state 'pending'. Below the table, a detailed view for the selected instance is shown, including tabs for Description, Status Checks, Monitoring, and Tags. The Description tab is active, displaying details such as Instance ID, Instance state, Instance type, Elastic IPs, Availability zone, Security groups, and Scheduled events. The right side of this section lists network-related information: Public DNS (IPv4), IPv4 Public IP, IPv6 IPs, Private DNS, Private IPs, Secondary private IPs, and VPC ID.

Launch Instance Connect Actions

Filter by tags and attributes or search by keyword

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)	IPv4 Public IP	IPv6 IPs	Key Name	Monitoring	Launch Time
	i-06cebbe30f04b8ab3	t2.micro	us-east-2b	pending	Initializing	None	ec2-3-15-158-228 us-e...	3.15.158.228	-	comp4331	disabled	September

Instance: i-06cebbe30f04b8ab3 Public DNS: ec2-3-15-158-228.us-east-2.compute.amazonaws.com

Description Status Checks Monitoring Tags

Property	Value
Instance ID	i-06cebbe30f04b8ab3
Instance state	pending
Instance type	t2.micro
Elastic IPs	
Availability zone	us-east-2b
Security groups	launch-wizard-1. <a href="#">view inbound rules</a> . <a href="#">view outbound rules</a>
Scheduled events	-

Property	Value
Public DNS (IPv4)	ec2-3-15-158-228.us-east-2.compute.amazonaws.com
IPv4 Public IP	3.15.158.228
IPv6 IPs	-
Private DNS	ip-172-31-26-9.us-east-2.compute.internal
Private IPs	172.31.26.9
Secondary private IPs	
VPC ID	vpc-48ce3a23

Feedback English (US)

© 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

# Connect To Your Instance

## Creat a keypair , name it and download it

**Connect To Your Instance**

I would like to connect with

- ☒ A standalone SSH client ⓘ
- ☐ EC2 Instance Connect (browser-based SSH connection) ⓘ
- ☐ A Java SSH Client directly from my browser (Java required) ⓘ

**To access your instance:**

1. Open an SSH client. (find out how to [connect using PuTTY](#) )
2. Locate your private key file (comp4331.pem). The wizard automatically detects the key you used to launch the instance.
3. Your key must not be publicly viewable for SSH to work. Use this command if needed:

```
chmod 400 comp4331.pem
```

4. Connect to your instance using its Public DNS:

```
ec2-3-15-158-228.us-east-2.compute.amazonaws.com
```

**Example:**

```
ssh -i "comp4331.pem" ec2-user@ec2-3-15-158-228.us-east-2.compute.amazonaws.com
```

Please note that in most cases the username above will be correct, however please ensure that you read your AMI usage instructions to ensure that the AMI owner has not changed the default AMI username.

If you need any assistance connecting to your instance, please see our [connection documentation](#).

**Close**

Creat a key pair!!! name it and download it

Os system: Ubuntu 18.04 or 16.04 (The fifth choice in the website) if you choose amazon linux wrongly, input `sudo yum install git` to install git tool.

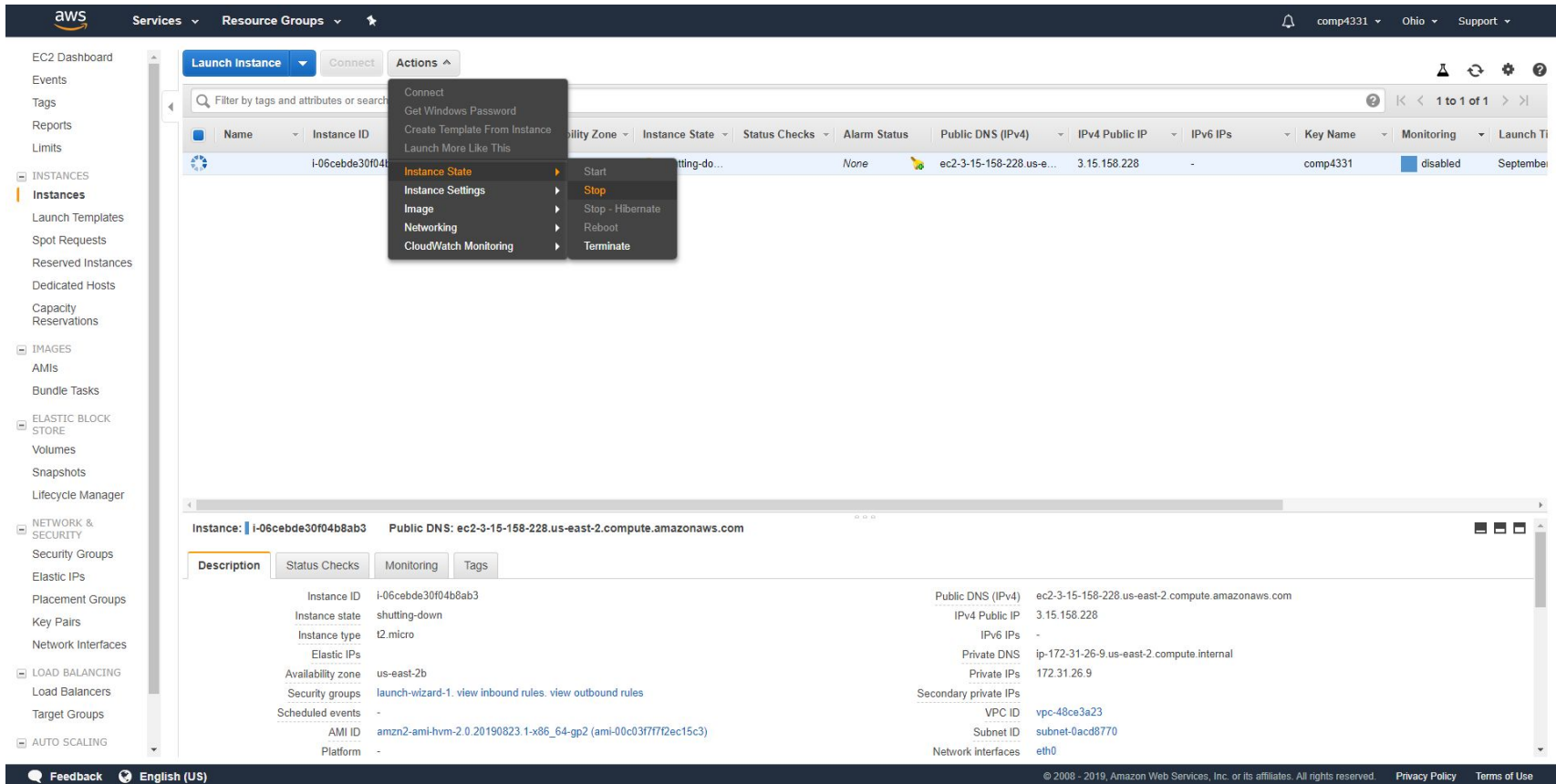
After lanching an instance, you need ssh connect to the virtual machine, since we only open port22.

For MacOS users: 1. open terminal 2. `chmod 400 yourkeypair.pem`

3. `ssh -i "yourkeypair.pem" ubuntu@ec2-3-15-42-77.us-east-2.compute.amazonaws.com`

For windows users: 1. download mobaxterm or other ssh tools 2,3. same as MacOS users

# Stop or terminal your instance



The screenshot displays the AWS Management Console interface. On the left, a navigation sidebar lists various services including EC2 Dashboard, INSTANCES, IMAGES, ELASTIC BLOCK STORE, NETWORK & SECURITY, LOAD BALANCING, and AUTO SCALING. The main content area shows a list of EC2 instances. An 'Actions' dropdown menu is open for the instance 'i-06cebd30f04b8ab3', with the 'Instance State' sub-menu expanded, highlighting the 'Stop' option. Below the instance list, the details for instance 'i-06cebd30f04b8ab3' are shown, including its state ('shutting-down'), type ('t2.micro'), and various network and security configurations.

**Actions Menu:**

- Connect
- Get Windows Password
- Create Template From Instance
- Launch More Like This
- Instance State**
  - Start
  - Stop**
  - Stop - Hibernate
- Instance Settings
  - Reboot
- Image
- Networking
- CloudWatch Monitoring
  - Terminate

**Instance Details:**

Category	Property	Value
Description	Instance ID	i-06cebd30f04b8ab3
	Instance state	shutting-down
	Instance type	t2.micro
	Elastic IPs	-
Availability zone	us-east-2b	
	Security groups	launch-wizard-1. <a href="#">view inbound rules.</a> <a href="#">view outbound rules</a>
Scheduled events	-	
	AMI ID	amzn2-ami-hvm-2.0.20190823.1-x86_64-gp2 (ami-00c037f7f72ec15c3)
Platform	-	
	Public DNS (IPv4)	ec2-3-15-158-228.us-east-2.compute.amazonaws.com
IPv4 Public IP	3.15.158.228	
	IPv6 Public IP	-
Private DNS	ip-172-31-26-9.us-east-2.compute.internal	
	Private IPs	172.31.26.9
Secondary private IPs	-	
	VPC ID	vpc-48ce3a23
Subnet ID	subnet-0acd8770	
	Network interfaces	eth0

# Data preprocessing

Example: Iris dataset

	sepal length	sepal width	petal length	petal width	target
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

iris dataset head

Python Libraries: pandas, matplotlib, sklearn

Code: git clone <https://github.com/jiaxinjie97/COMP4331.git>



git clone <https://github.com/jiaxinxie97/COMP4331.git>

# Principal component analysis (PCA)

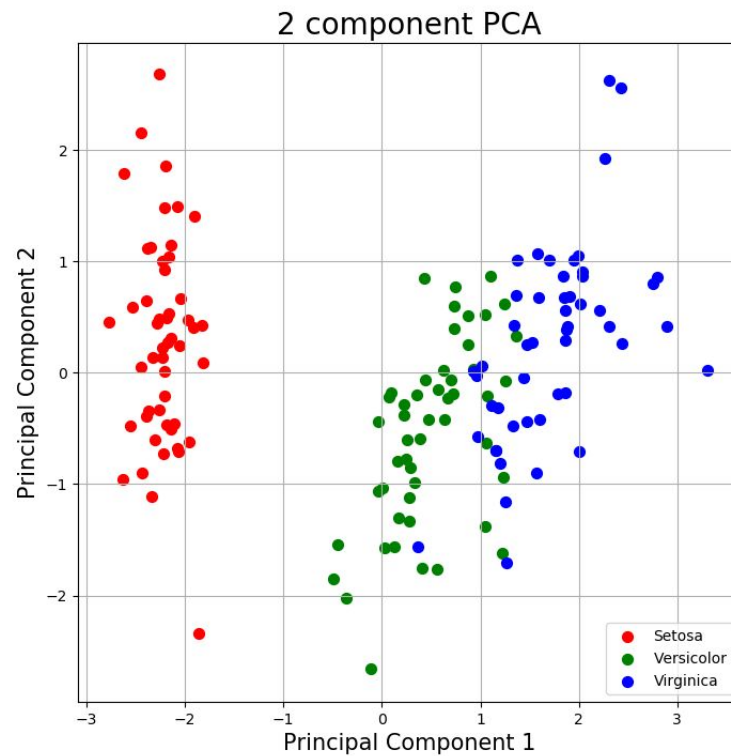
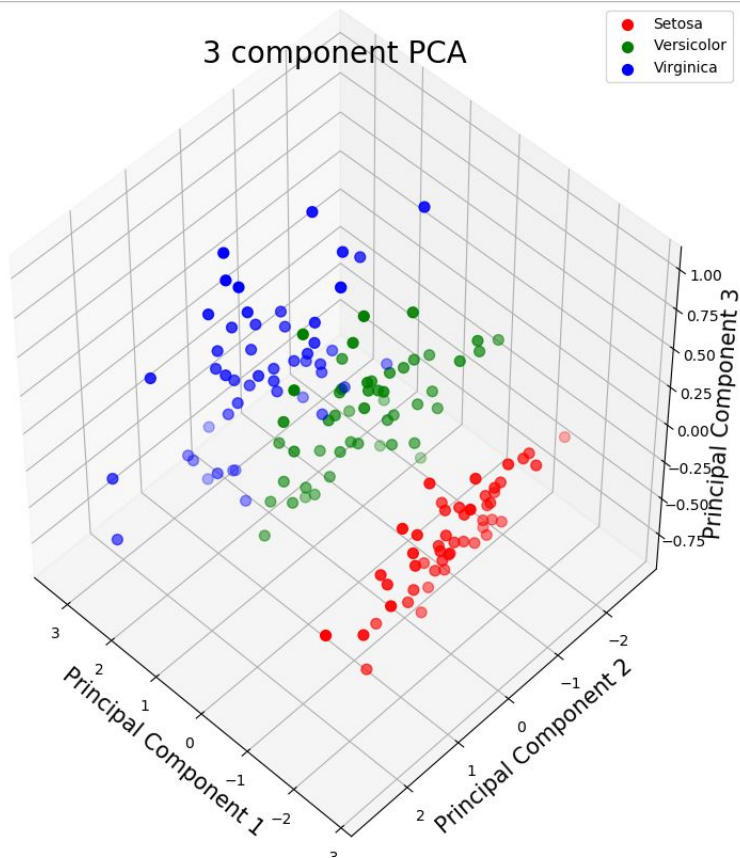
```
1  from sklearn import decomposition
2  # 3 components
3  pca = decomposition.PCA(n_components=3)
4  PA3 = pca.fit_transform(X)
5  # 2 components
6  pca = decomposition.PCA(n_components=2)
7  PA2 = pca.fit_transform(X)
8
```

Function: `sklearn.decomposition.PCA`

More details plz see <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>

Change `n_components`?

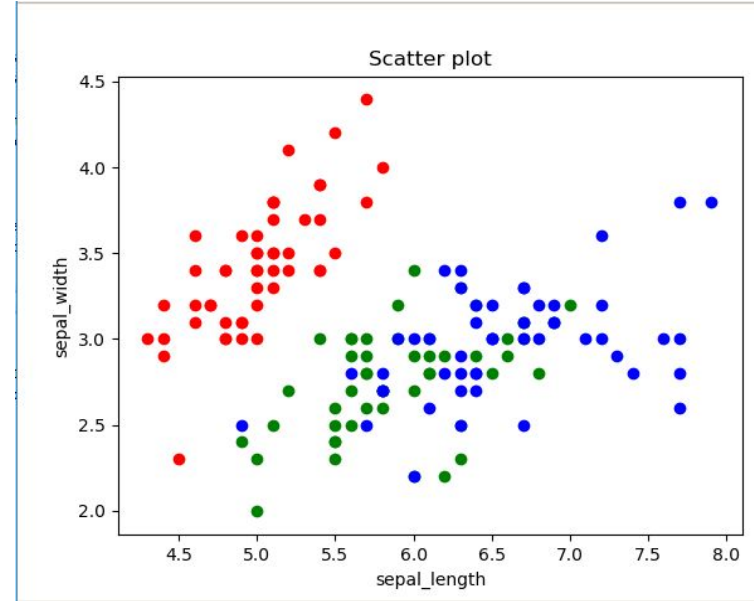
# Principal component analysis (PCA)



# Scatter plot

```
1 import matplotlib.pyplot as plt
2 #scatter plot
3 colors = {'Setosa':'r', 'Versicolor':'g', 'Virginica':'b'}
4 # create a figure and axis
5 fig, ax = plt.subplots()
6 # plot each data-point
7 for i in range(len(iris['sepal_length'])):
8     ax.scatter(iris['sepal_length'][i], iris['sepal_width'][i],\
9               color=colors[iris['class'][i]])
10 # set a title and labels
11 ax.set_title('Scatter Plot')
12 ax.set_xlabel('sepal_length')
13 ax.set_ylabel('sepal_width')
14 plt.show()
15
```

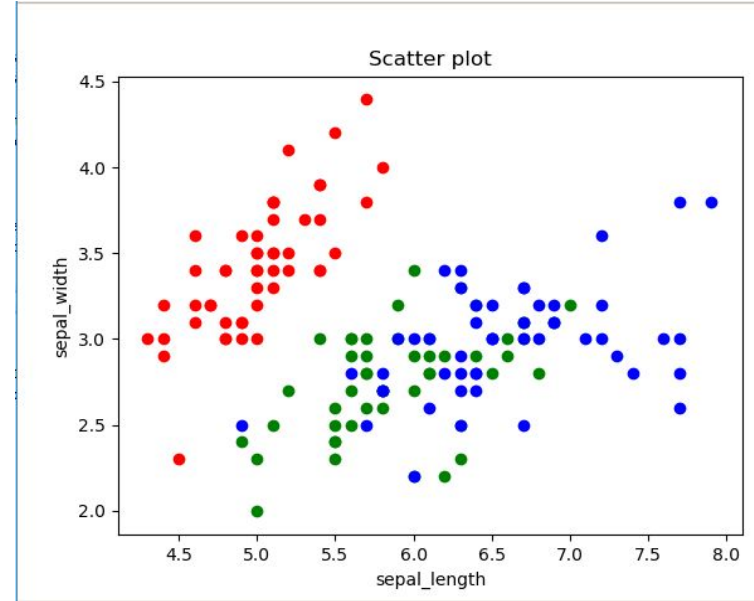
How to draw a scatter plot with sepal\_length for the x-axis and petal\_length for the y\_axis?



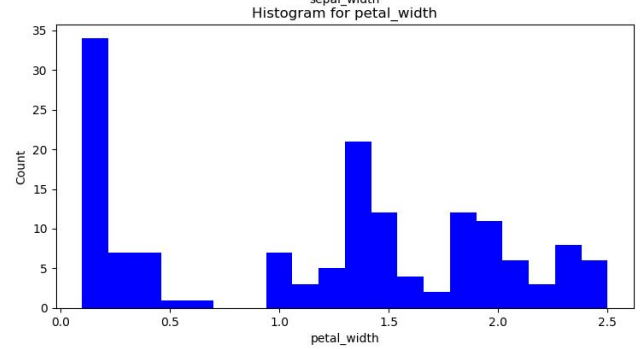
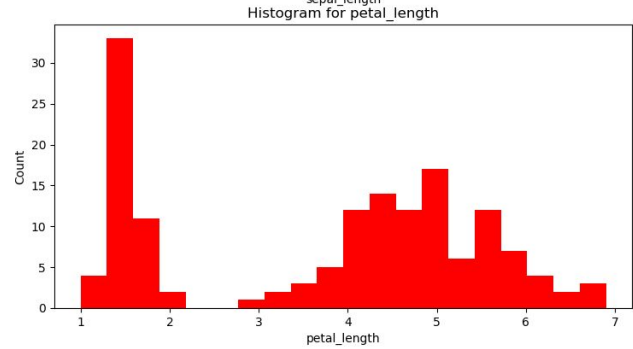
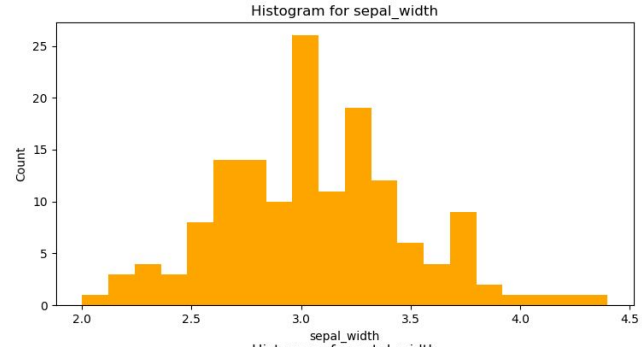
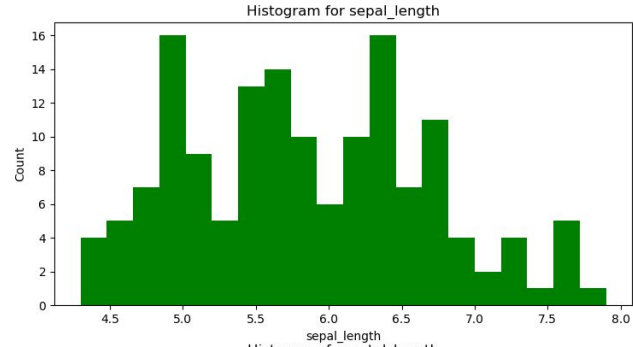
# Scatter plot

```
1 import matplotlib.pyplot as plt
2 #scatter plot
3 colors = {'Setosa':'r', 'Versicolor':'g', 'Virginica':'b'}
4 # create a figure and axis
5 fig, ax = plt.subplots()
6 # plot each data-point
7 for i in range(len(iris['sepal_length'])):
8     ax.scatter(iris['sepal_length'][i], iris['sepal_width'][i],\
9               color=colors[iris['class'][i]])
10 # set a title and labels
11 ax.set_title('Scatter Plot')
12 ax.set_xlabel('sepal_length')
13 ax.set_ylabel('sepal_width')
14 plt.show()
15
```

How to draw a scatter plot with sepal\_length for the x-axis and petal\_length for the y\_axis?



# Histogram

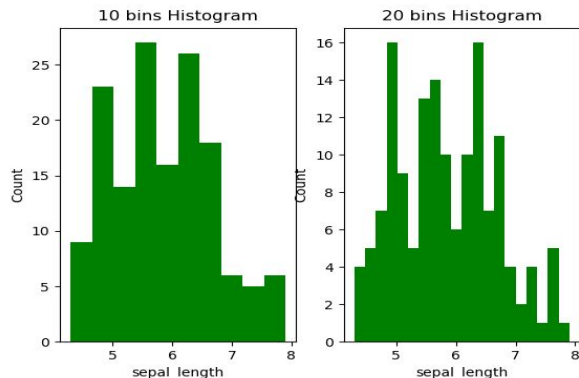


# Histogram

```
1 import matplotlib.pyplot as plt
2 plt.figure()
3 x = iris["sepal_length"]
4 plt.hist(x, bins = 20, color = "green")
5 plt.title("Histogram for sepal_length")
6 plt.xlabel("sepal_length")
7 plt.ylabel("Count")
```

Function: `matplotlib.pyplot.hist` More details plz see [https://matplotlib.org/3.1.1/api/\\_as\\_gen/matplotlib.pyplot.hist.html](https://matplotlib.org/3.1.1/api/_as_gen/matplotlib.pyplot.hist.html)

Change bins? `plt.hist(x, bins = 10, color = "green")`

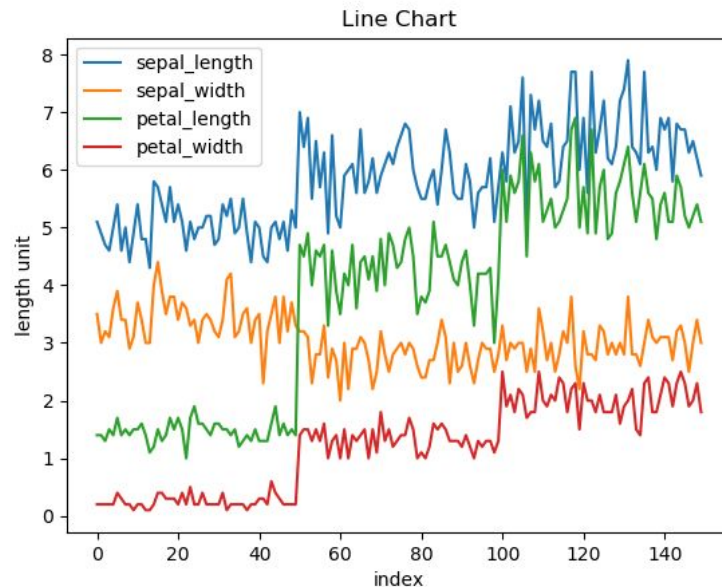


Try to change other parameters, such as range.

# Line Chart

```
1 import matplotlib.pyplot as plt
2 #Line Chart
3 columns = iris.columns.drop(['class'])
4 # create x data
5 x_data = range(0, iris.shape[0])
6 # create figure and axis
7 fig, ax = plt.subplots()
8 # plot each column
9 for column in columns:
10     ax.plot(x_data, iris[column])
11 # set title and legend
12 ax.set_title('Line Chart')
13 plt.xlabel('index')
14 plt.ylabel('length unit')
15 ax.legend(['sepal_length', 'sepal_width', 'petal_length', 'petal_width'])
16 plt.show()
```

How to draw a Line Chart if we only want to plot sepal\_length and sepal\_width?





# Box Plot

```
1 import matplotlib.pyplot as plt
2 #Box plot
3 plt.figure()
4 new_iris=iris[["sepal_length", "sepal_width",\
5 "petal_length", "petal_width"]]
6 new_iris.boxplot()
7 plt.title('Box Plot')
8 plt.show()
```

Same question as Line Chart. if we only want to include sepal\_length and sepal\_width?

