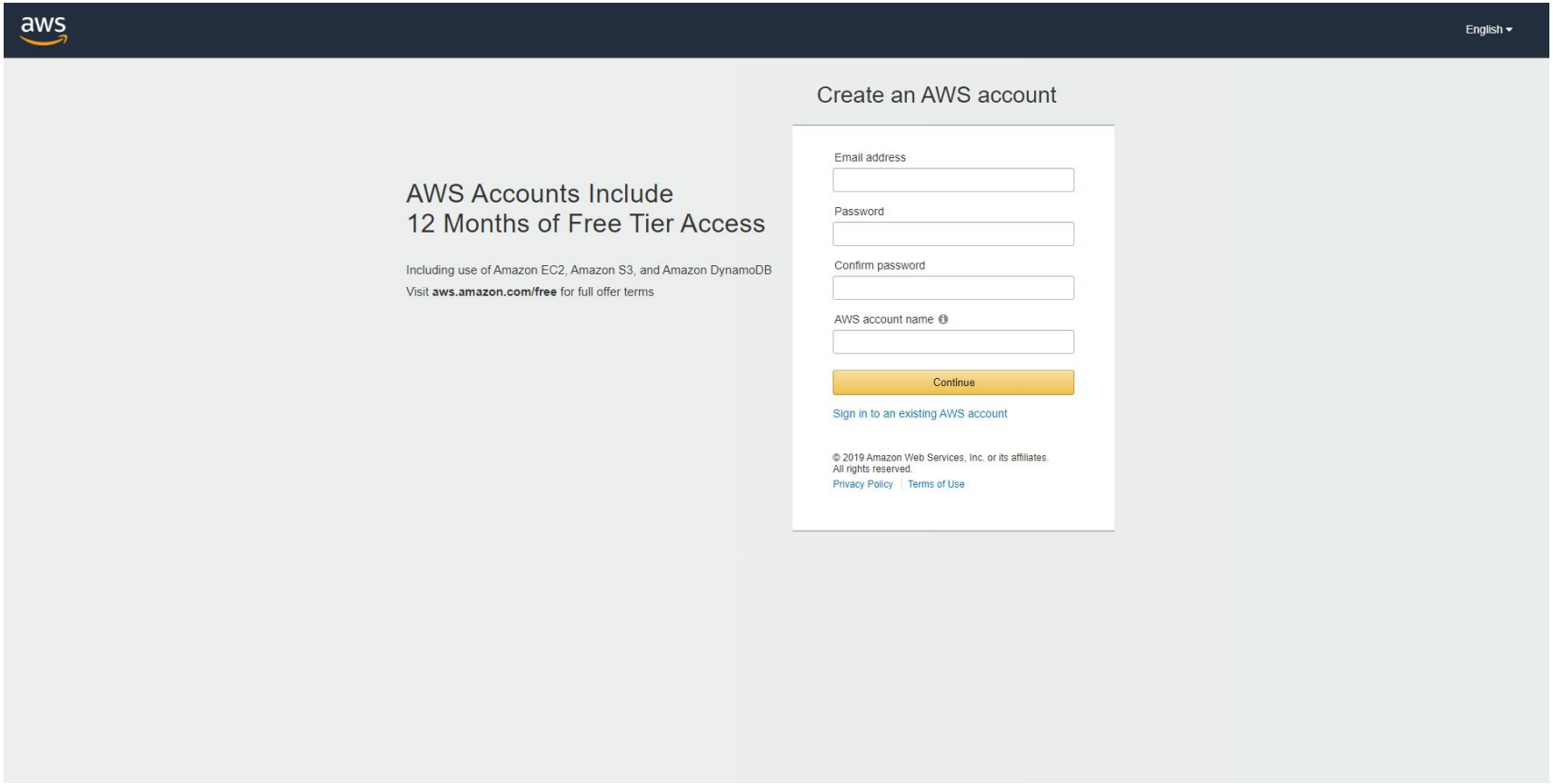
Tutorial 1

Outline

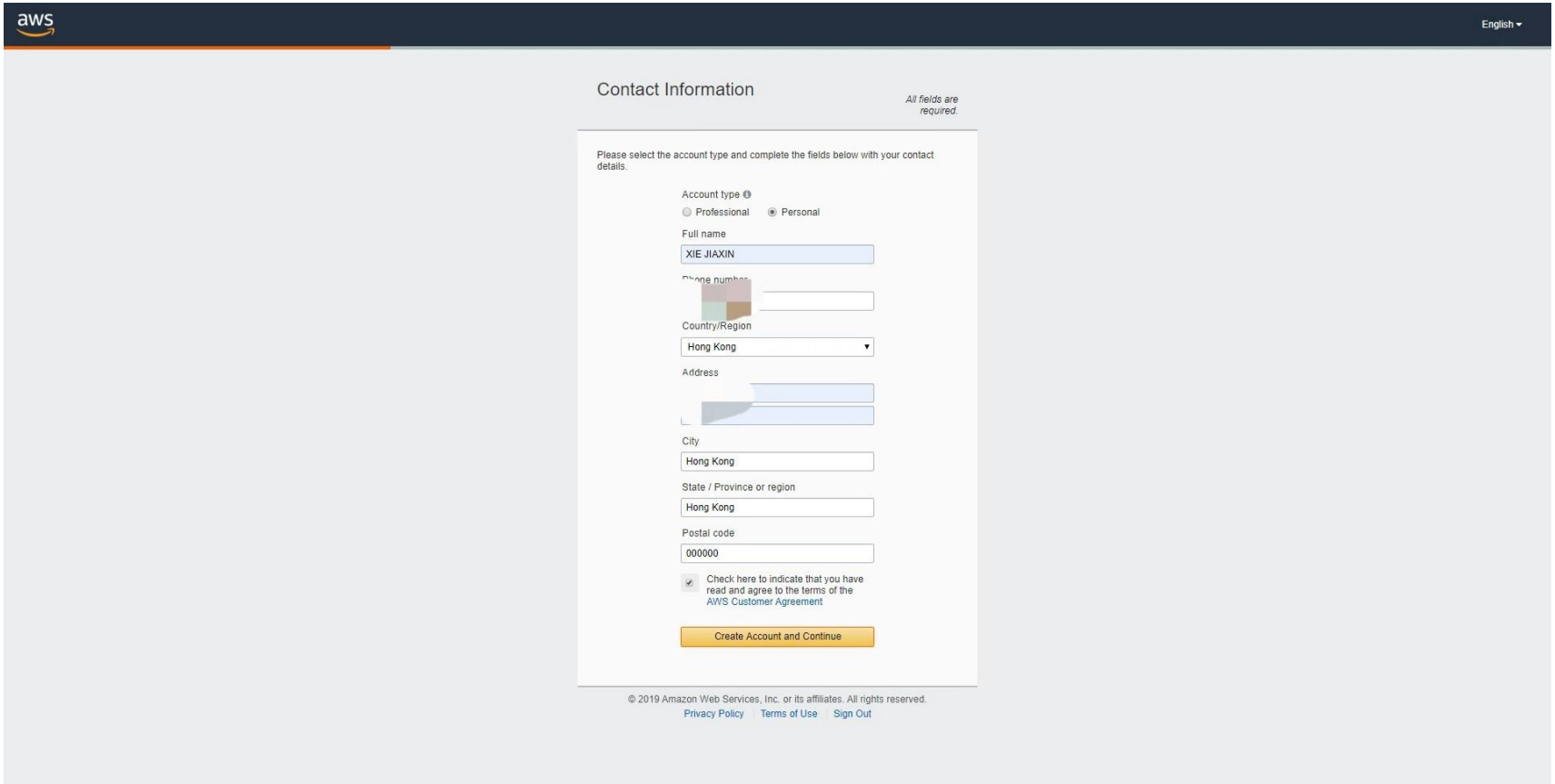
Creat an AWS Account

Data preprocessing

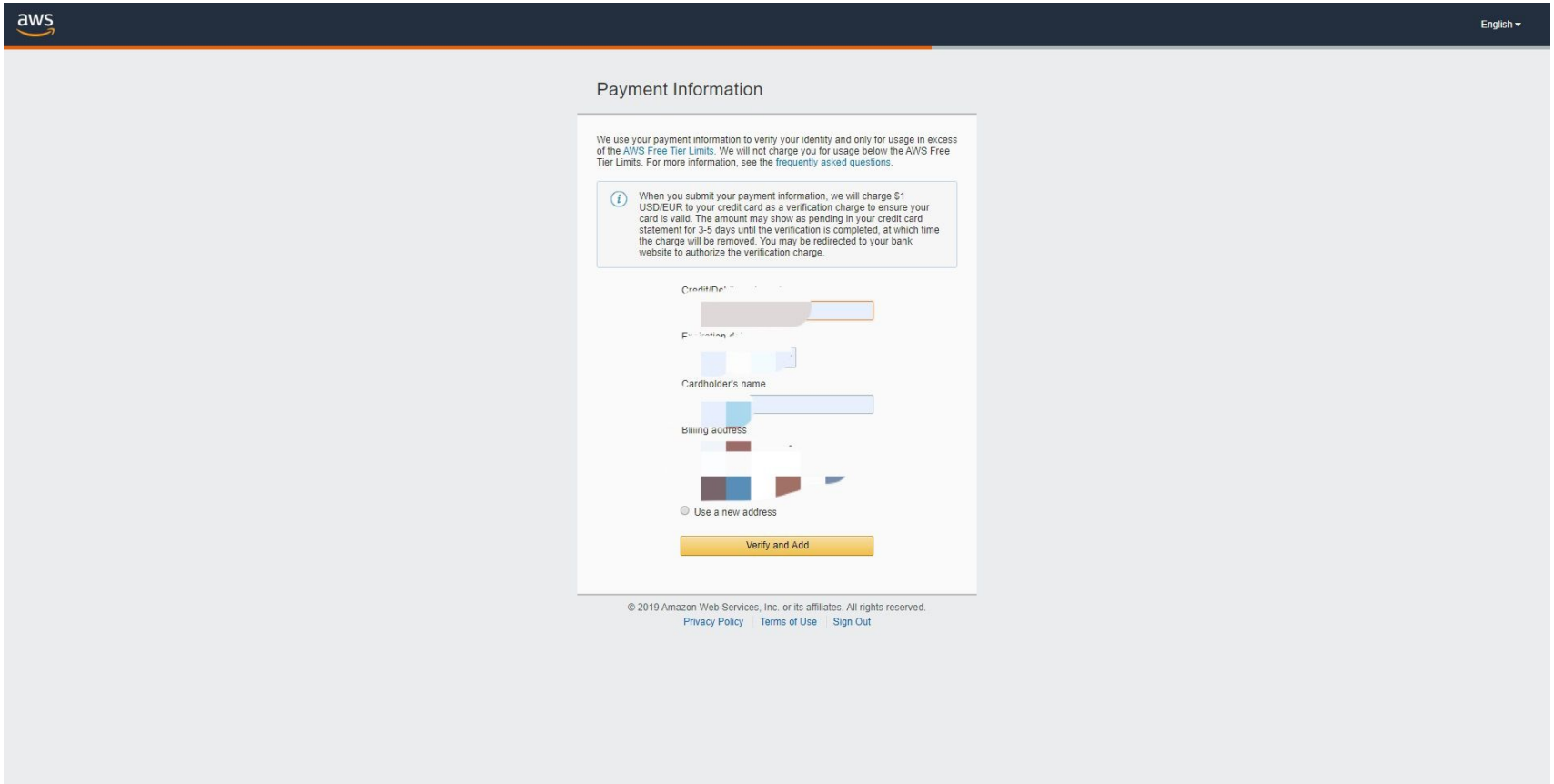
Create AWS account



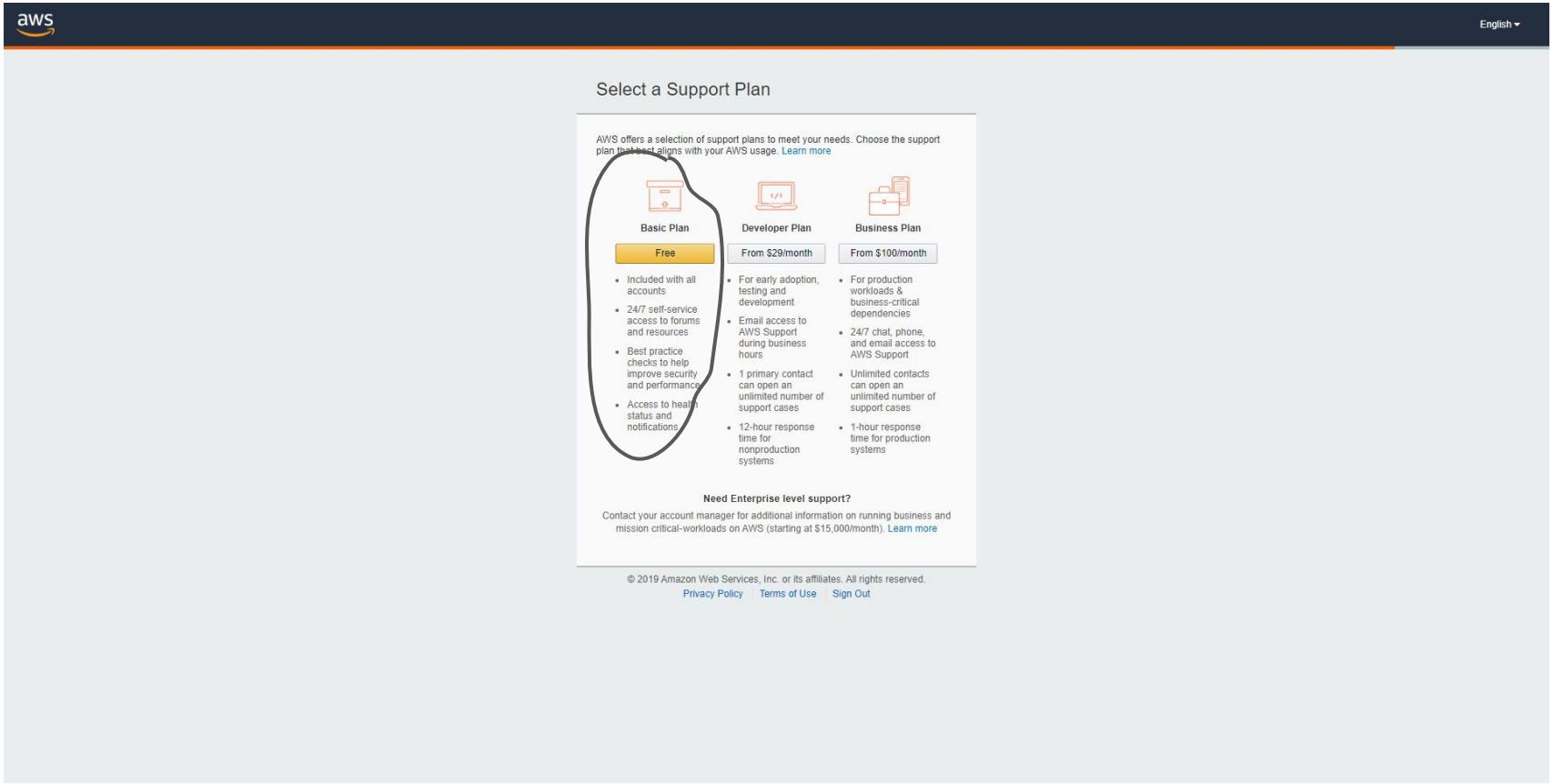
Contact Information



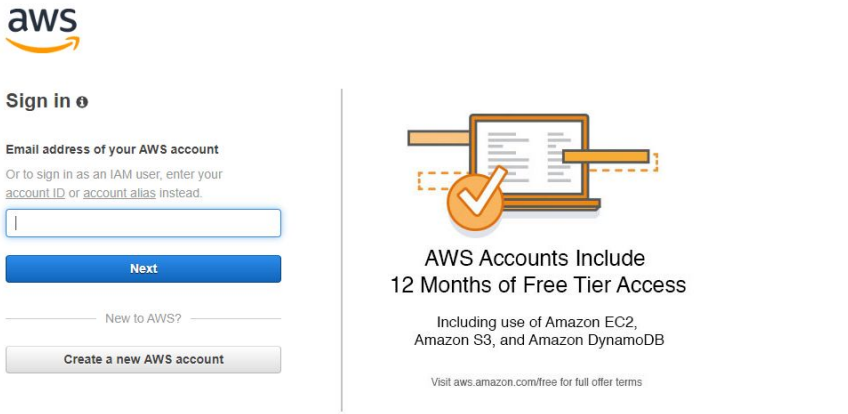
Payment Information



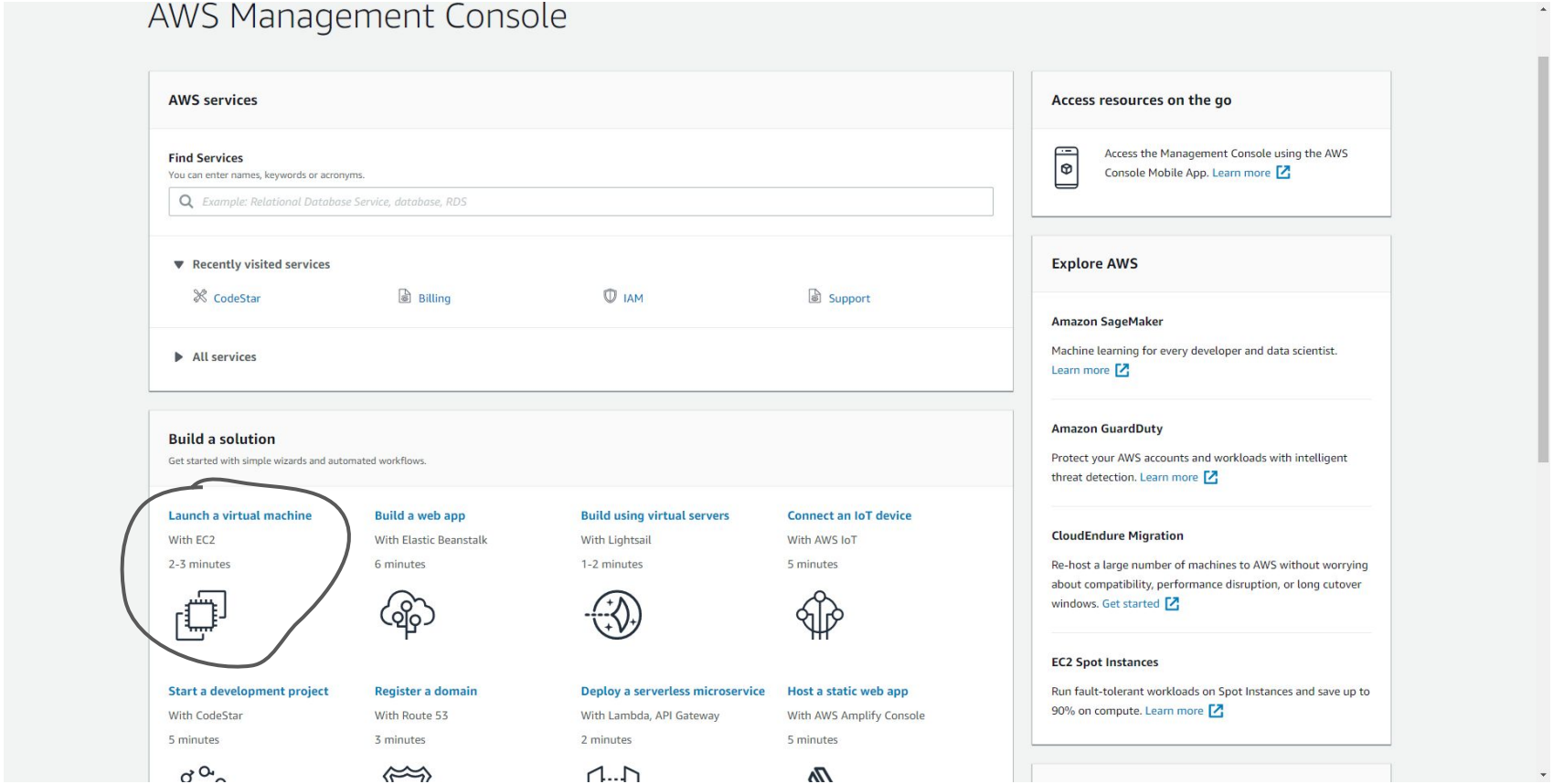
Select a Support Plan



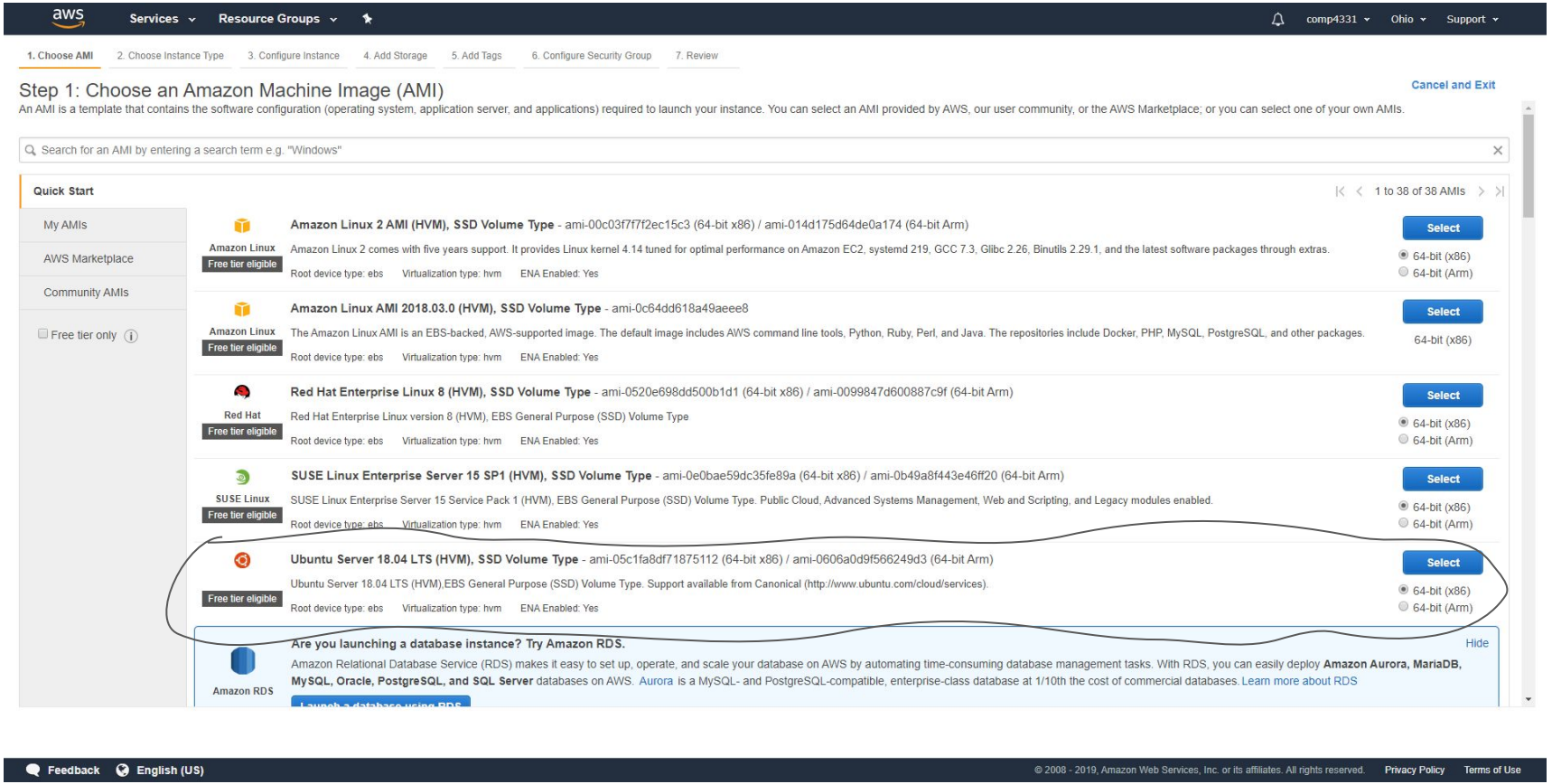
Sign in


aws
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 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

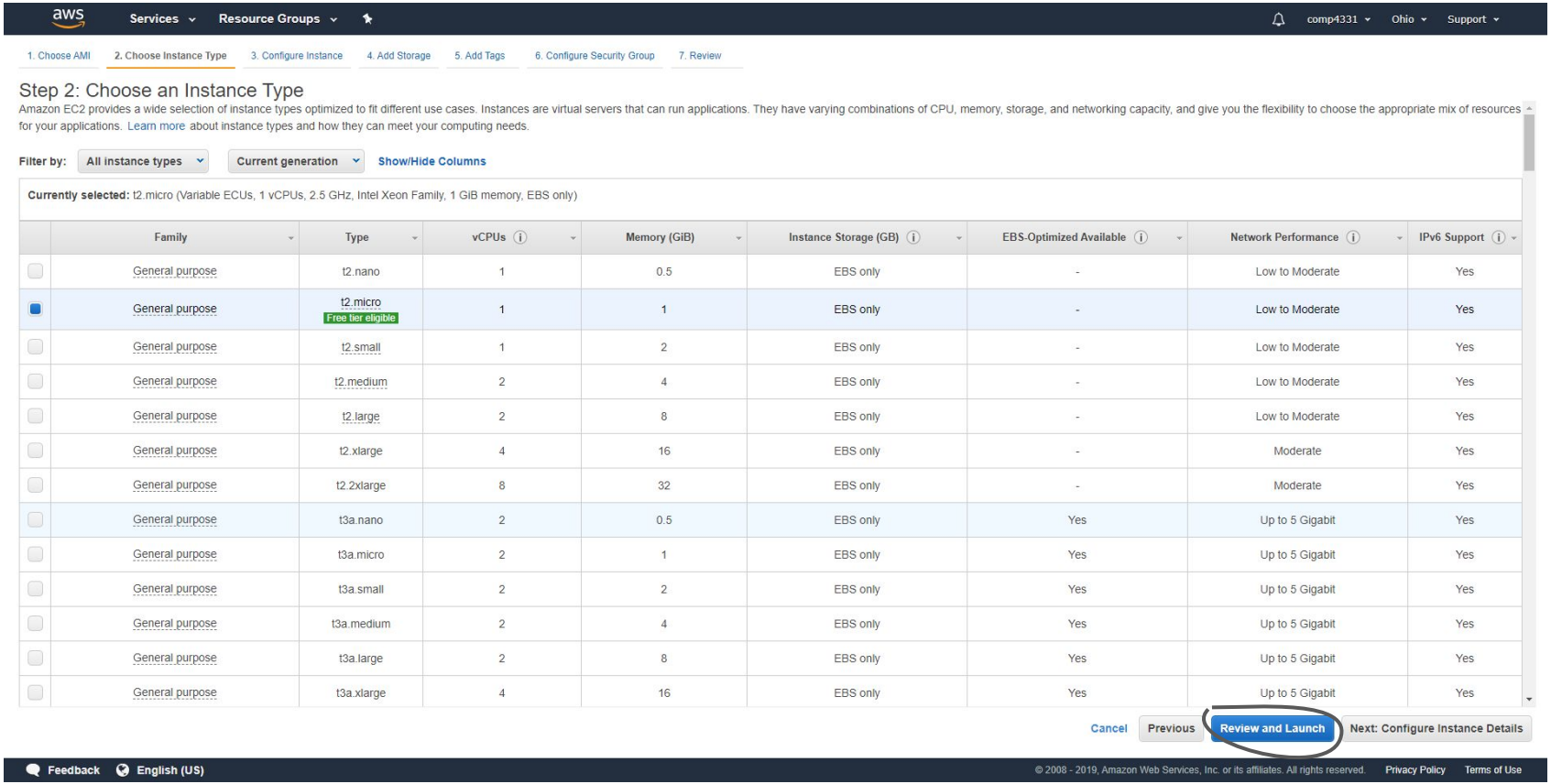
Launch a virtual machine with EC2



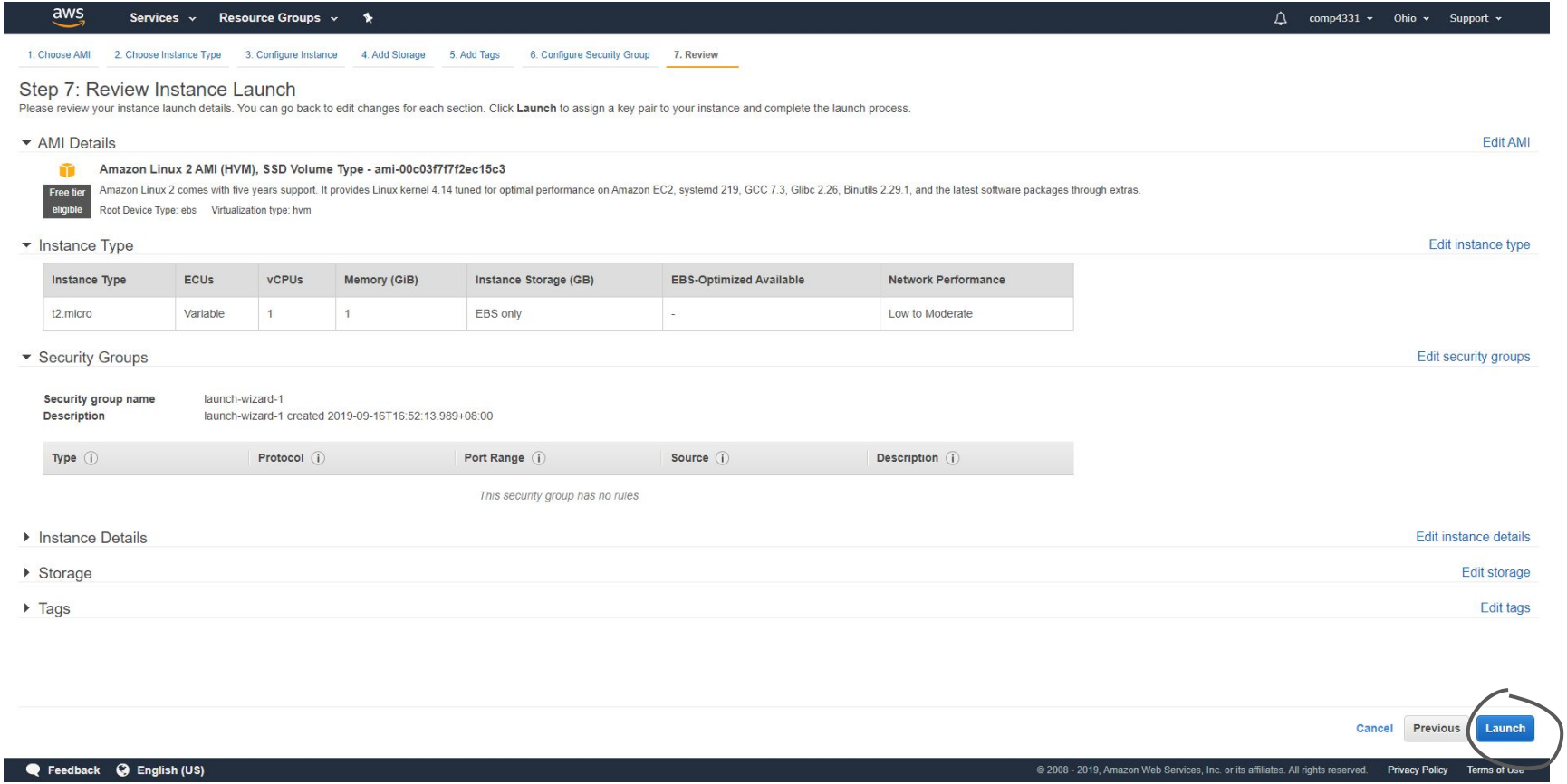
launch an instance (ubuntu 18.04 or ubuntu 16.04)



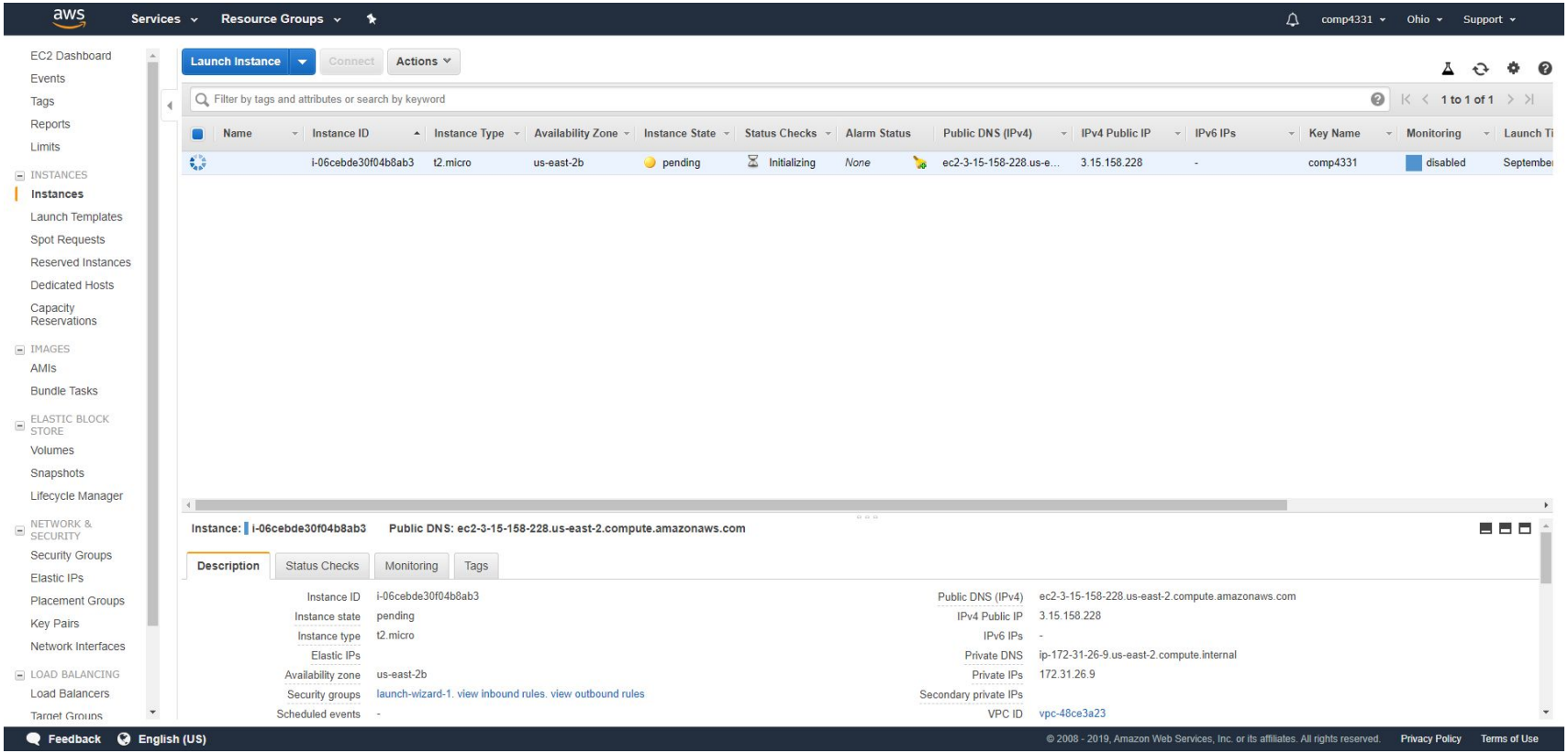
launch an instance



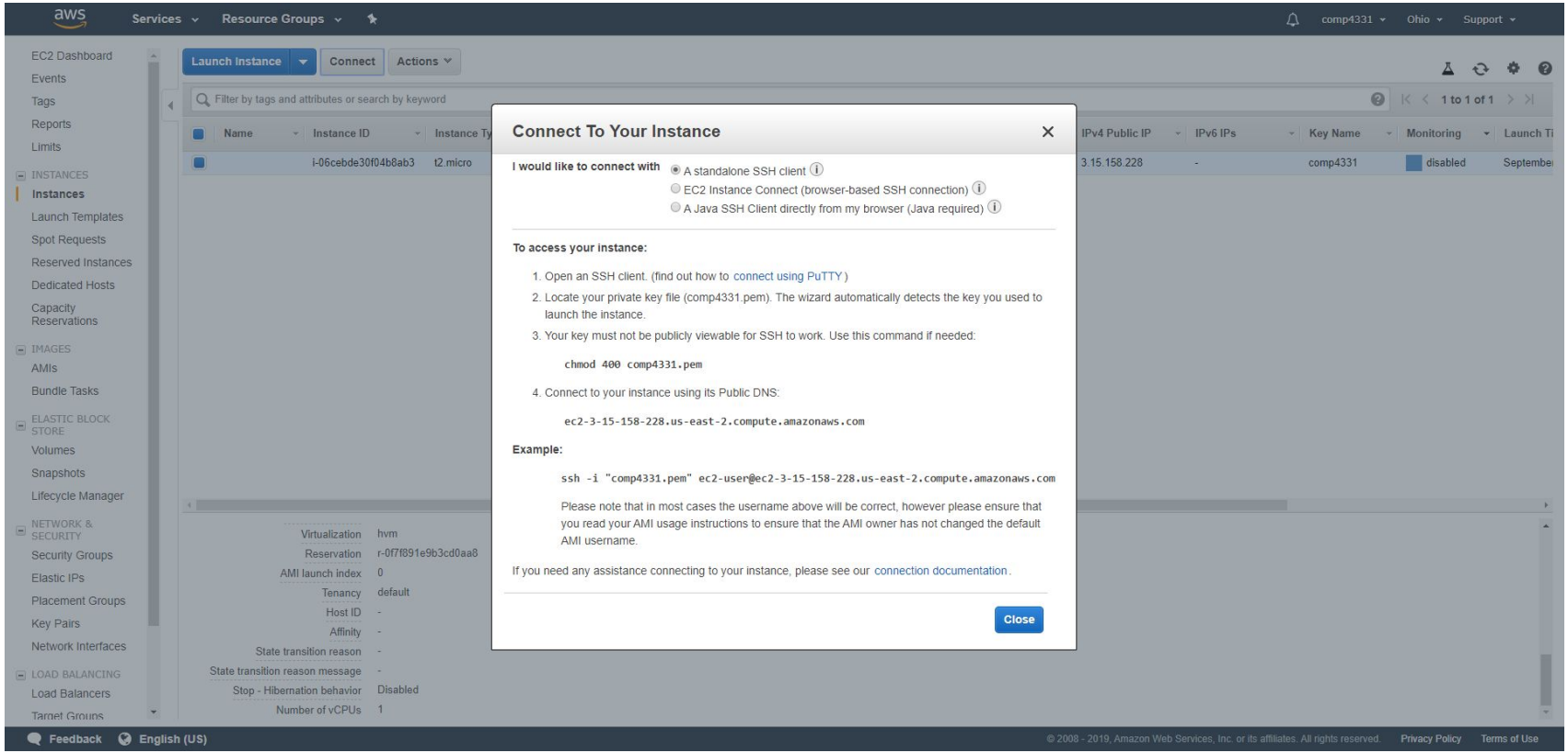
launch an instance



launch an instance



Connect To Your Instance
Creat a keypair , name it and download it



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.

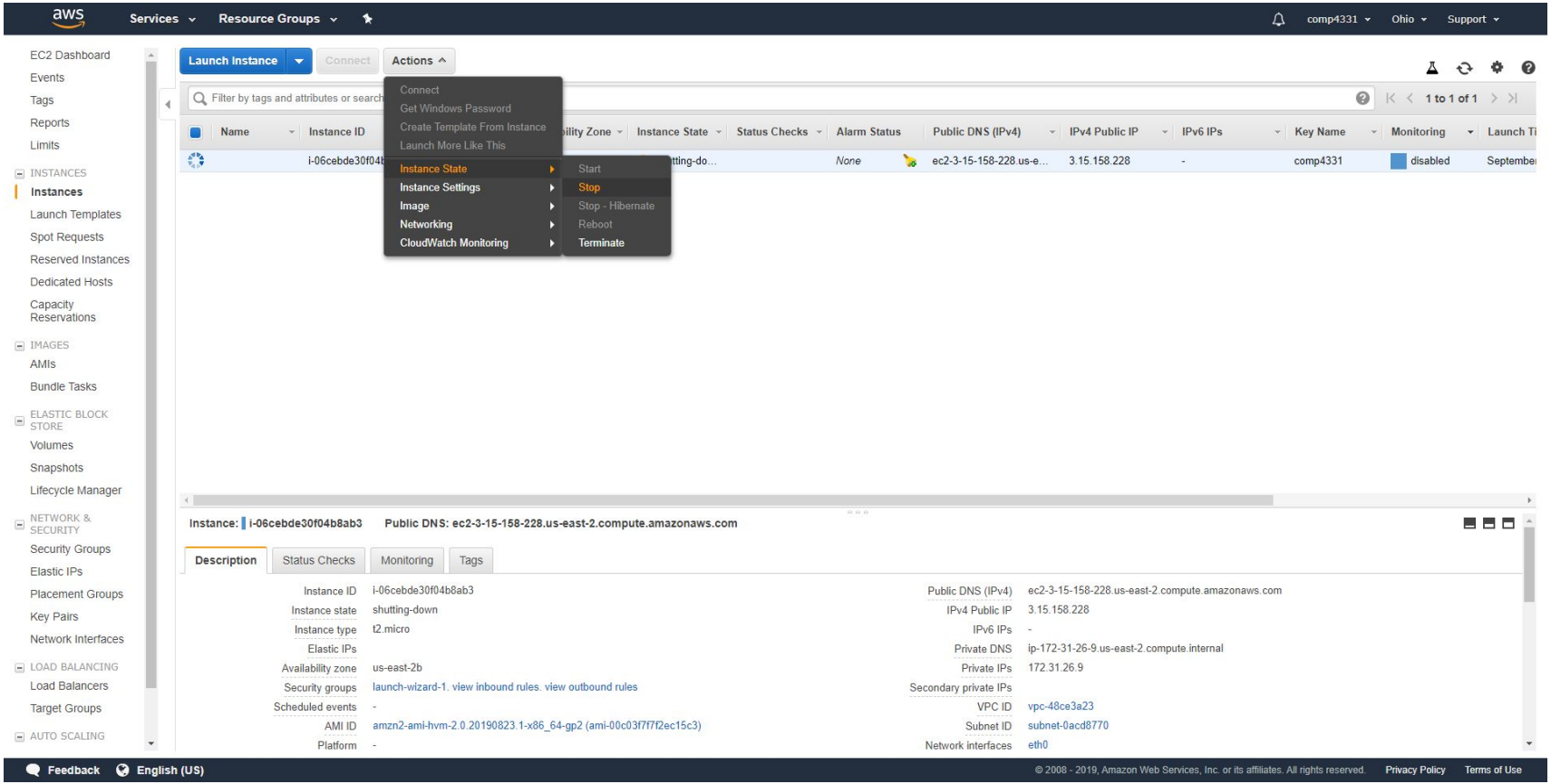
After lanching an instance, you need ssh connect to the virtual machine, since we only open port22.

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



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/jiaxinxie97/COMP4331.git

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

Principal component analysis (PCA)

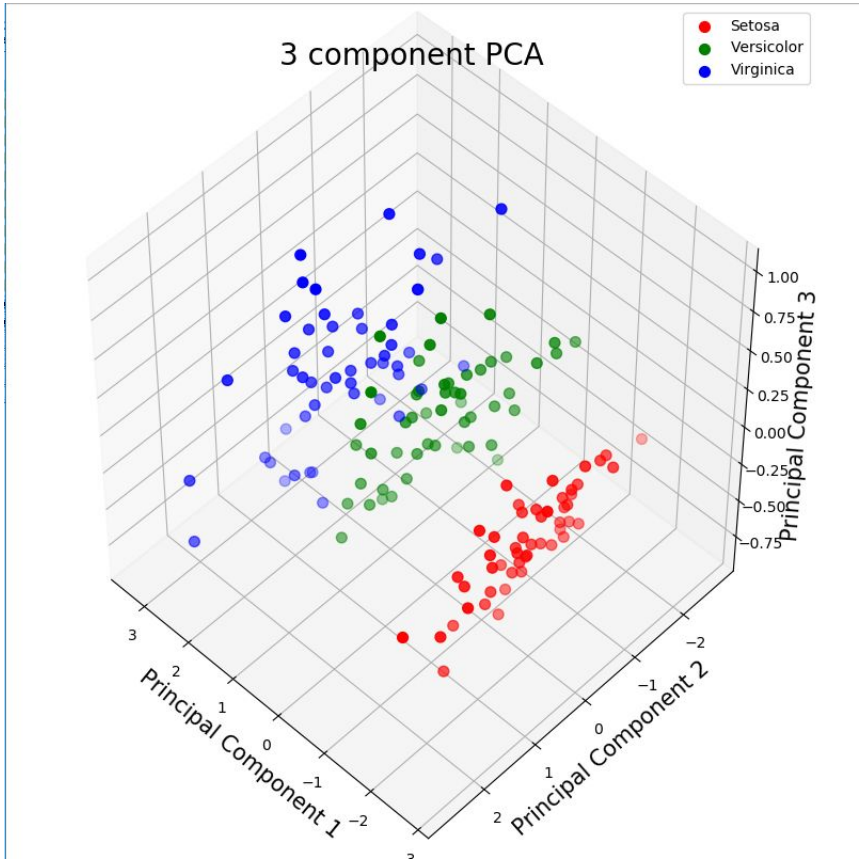
```
from sklearn import decomposition
# 3 components
pca = decomposition.PCA(n_components=3)
PA3 = pca.fit_transform(X)
# 2 components
pca = decomposition.PCA(n_components=2)
PA2 = pca.fit_transform(X)
```

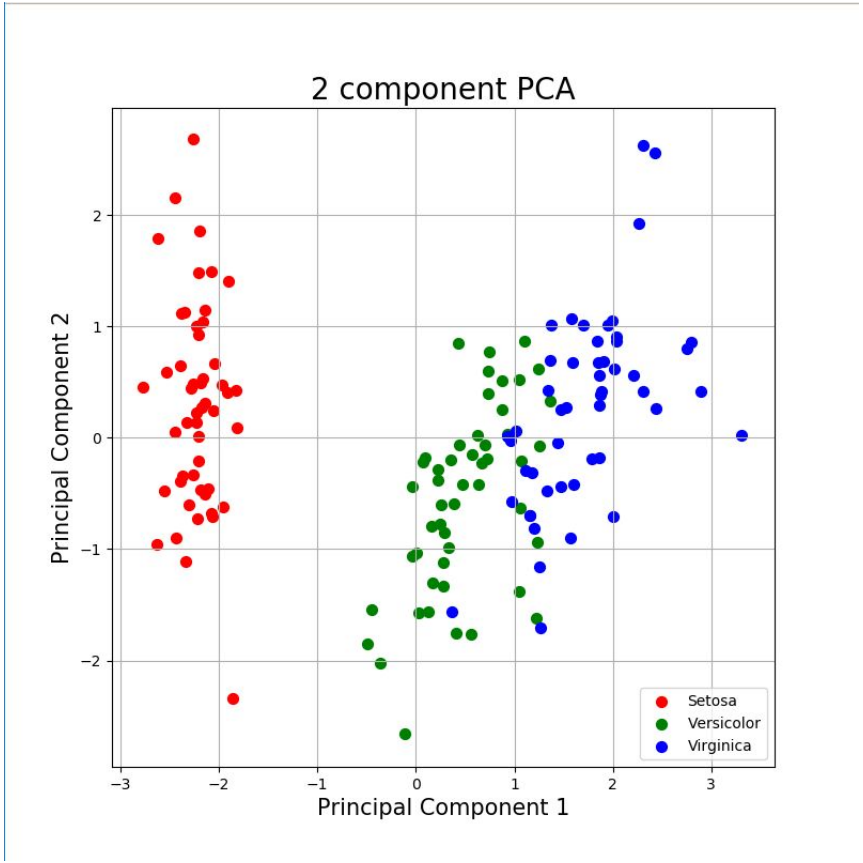
Function: sklearn.decomposition.PCA

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

Change n_componets?

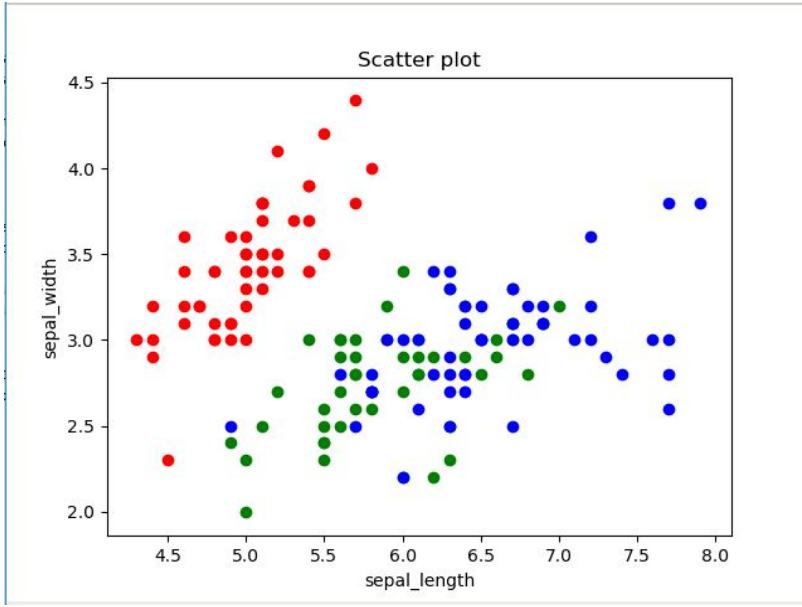
Principal component analysis (PCA)





Scatter plot

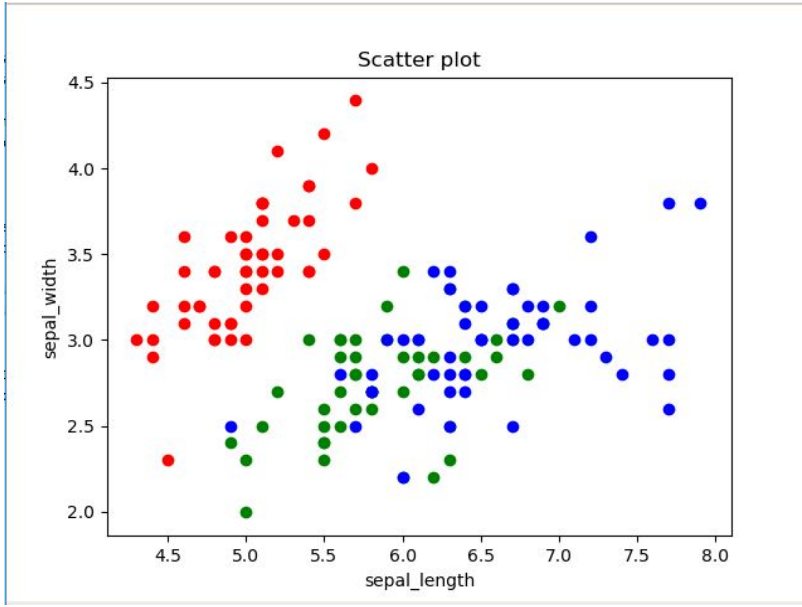
```python
import matplotlib.pyplot as plt
#scatter plot
colors = {'Setosa':'r', 'Versicolor':'g', 'Virginica':'b'}
# create a figure and axis
fig, ax = plt.subplots()
# plot each data-point
for i in range(len(iris['sepal_length'])):
    ax.scatter(iris['sepal_length'][i], iris['sepal_width'][i],\
    color=colors[iris['class'][i]])
# set a title and labels
ax.set_title('Scatter Plot')
ax.set_xlabel('sepal_length')
ax.set_ylabel('sepal_width')
plt.show()
```



How to draw a scatter plot with sepal_length for the x-axis and petal_length for the y_axis?

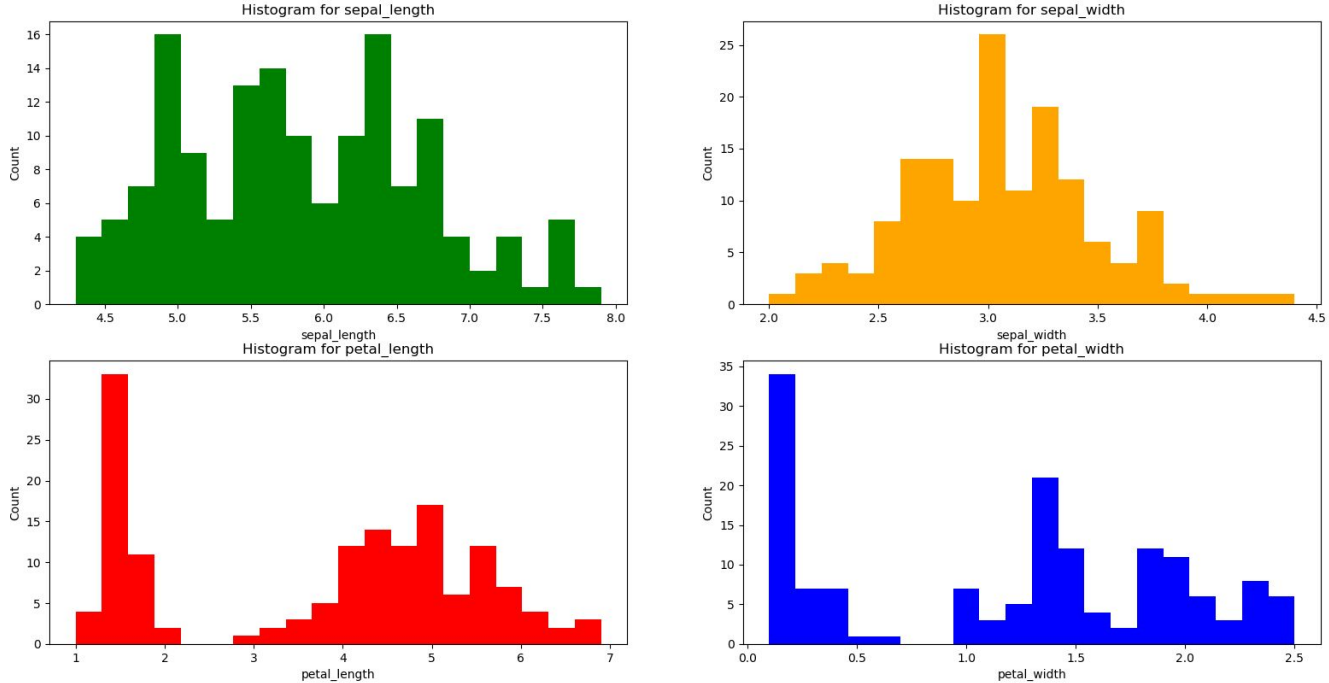
Scatter plot

```python
import matplotlib.pyplot as plt
#scatter plot
colors = {'Setosa':'r', 'Versicolor':'g', 'Virginica':'b'}
# create a figure and axis
fig, ax = plt.subplots()
# plot each data-point
for i in range(len(iris['sepal_length'])):
    ax.scatter(iris['sepal_length'][i], iris['sepal_width'][i],\
    color=colors[iris['class'][i]])
# set a title and labels
ax.set_title('Scatter Plot')
ax.set_xlabel('sepal_length')
ax.set_ylabel('sepal_width')
plt.show()
```



How to draw a scatter plot with sepal_length for the x-axis and petal_length for the y_axis?

Histogram

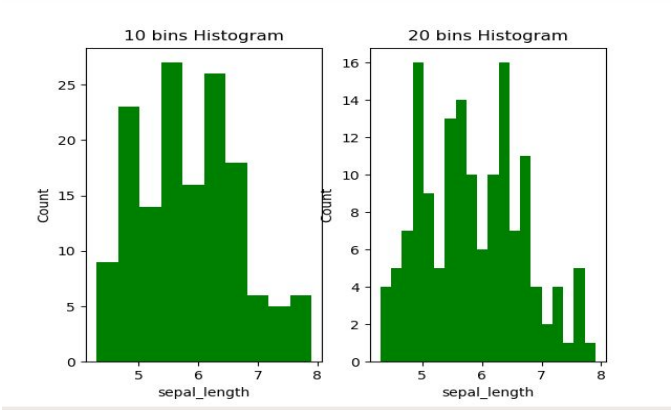

Histogram for sepal_length
Count
sepal_length
Histogram for sepal_width
Count
sepal_width
Histogram for petal_length
Count
petal_length
Histogram for petal_width
Count
petal_width

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

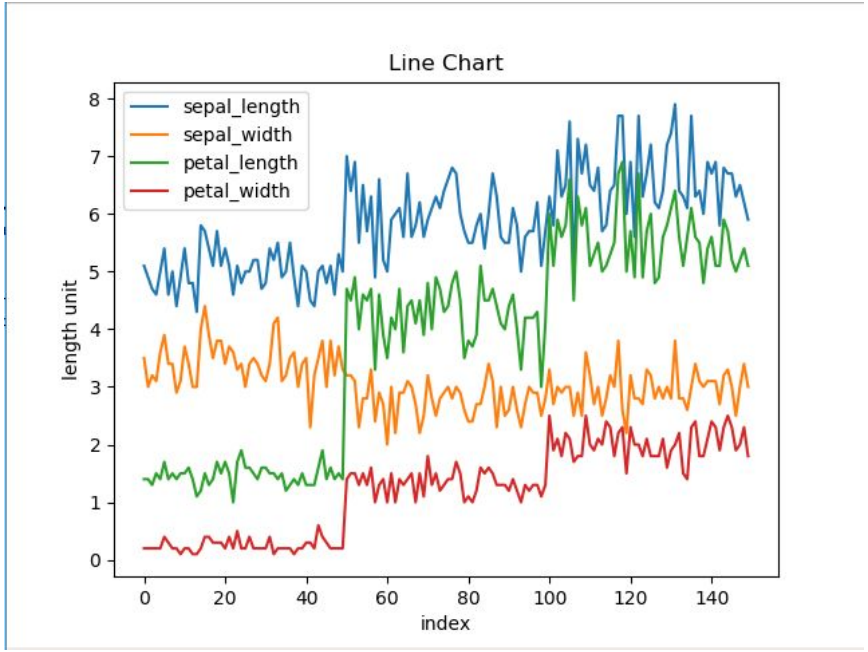
Change bins? plt.hist(x, bins = 10, color = "green")



Try to change other parameters, such as range.

Line Chart

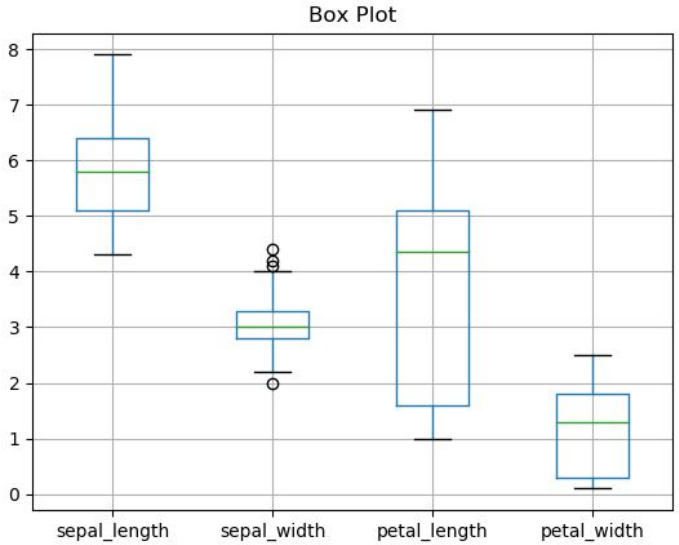
```python
import matplotlib.pyplot as plt
#Line Chart
columns = iris.columns.drop(['class'])
# create x data
x_data = range(0, iris.shape[0])
# create figure and axis
fig, ax = plt.subplots()
# plot each column
for column in columns:
    ax.plot(x_data, iris[column])
# set title and legend
ax.set_title('Line Chart')
plt.xlabel('index')
plt.ylabel('length unit')
ax.legend(['sepal_length', 'sepal_width', 'petal_length', 'petal_width'])
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?