

EXPT-1

AIM: Install, configure and run python, numPy and Pandas.

DESCRIPTION:

NumPy:

This code demonstrates basic NumPy operations. It first creates and prints a 1D array and a 2D matrix, then performs basic arithmetic operations and element-wise operations on them. Additionally, it shows how to access specific elements and create arrays filled with zeros, ones, and an identity matrix.

CODE:

```
import numpy as np

# Create a NumPy array from a Python list
array = np.array([1, 2, 3, 4, 5])
print("Array:", array)

# Create a 2D array (matrix) from a nested list
matrix = np.array([[1, 2, 3], [4, 5, 6]])
print("Matrix:\n", matrix)

# Basic array operations
array_sum = np.sum(array)
array_mean = np.mean(array)
array_max = np.max(array)

print("Sum of array elements:", array_sum)
print("Mean of array elements:", array_mean)
print("Max of array elements:", array_max)

# Element-wise operations
array_squared = array ** 2
print("Array squared:", array_squared)

# Matrix operations
matrix_transposed = np.transpose(matrix)
print("Transposed Matrix:\n", matrix_transposed)

# Accessing elements
element = matrix[1, 2] # Access the element at row 1, column 2
print("Element at position (1, 2):", element)
```

```

# Creating arrays with specific values
zeros_array = np.zeros((2, 3)) # Create a 2x3 array of zeros
ones_array = np.ones((3, 2)) # Create a 3x2 array of ones
identity_matrix = np.eye(3) # Create a 3x3 identity matrix

print("Zeros Array:\n", zeros_array)
print("Ones Array:\n", ones_array)
print("Identity Matrix:\n", identity_matrix)

```

OUTPUT:

The screenshot shows a Linux desktop environment with a dark theme. In the top left, there's an 'Activities' button and a 'Terminal' icon. The terminal window is open and shows the command-line session above. To the right of the terminal is a file manager window showing a directory structure with various files and folders like 'ai', 'an', 'cbit', 'cruise_ship_info.csv', 'dcn', etc.

```

Activities Terminal Aug 13 10:58
Terminal
cselab7-08@cselab708-OptiPlex-3050:~$ python3 -m venv env
cselab7-08@cselab708-OptiPlex-3050:~$ env/bin/activate
bash: env/bin/activate: Permission denied
cselab7-08@cselab708-OptiPlex-3050:~$ source env/bin/activate
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 pip install numpy
python3: can't open file '/home/cselab7-08/pip': [Errno 2] No such file or directory
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 pip install numpy
python3: can't open file '/home/cselab7-08/install': [Errno 2] No such file or directory
(env) cselab7-08@cselab708-OptiPlex-3050:~$ pip3 --version
pip 22.0.2 from /home/cselab7-08/env/lib/python3.10/site-packages/pip (python 3.10)
(env) cselab7-08@cselab708-OptiPlex-3050:~$ pip3 install numpy
Collecting numpy
  Downloading numpy-2.0.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (19.5 MB)
    19.5/19.5 MB 124.0 kB/s eta 0:00:00
Installing collected packages: numpy
Successfully installed numpy-2.0.1
(env) cselab7-08@cselab708-OptiPlex-3050:~$ vi num.py # installation and execution
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 num.py
9.png
ai_lab.pdf
ai.py
an.py
an.txt
a.out
.bash_history
.bashrc
.bashrc.save
.cache/
cbit/
cbit file access.py
cbit open close program.py
cbit.txt
cnscode/
.config/
con.sh
copy.txt
cruise_ship_info.csv
dcn/
dcn/

```

PANDAS:

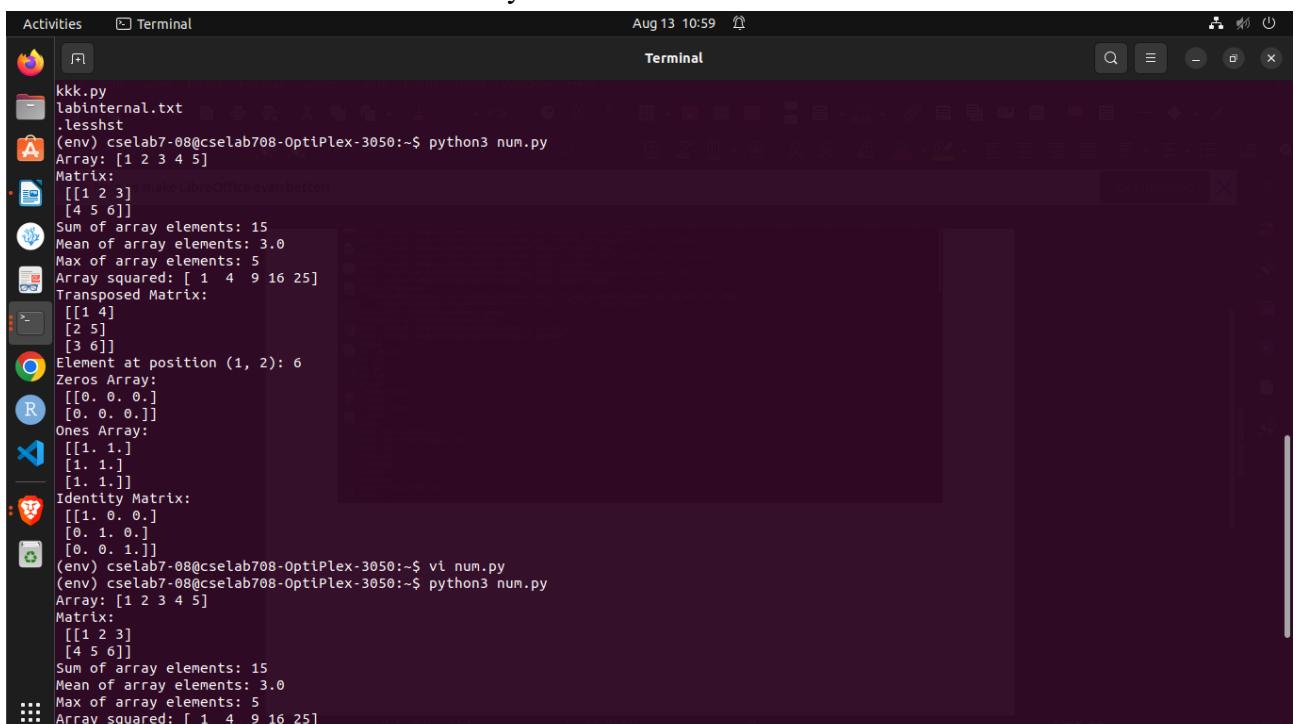
Description:

This code demonstrates basic operations using the Pandas library. It starts by creating a DataFrame from a dictionary and performs operations such as displaying rows, checking data types, and adding a new column. It also shows how to access rows and columns, filter data, compute descriptive statistics, and handle missing values by filling them with the mean of a column.

CODE

```
import pandas as pd
```

```
# Create a DataFrame from a dictionary
```



```
Activities Terminal Aug 13 10:59
kkk.py labinternal.txt .lessht
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 num.py
Array: [1 2 3 4 5]
Matrix:
[[1 2 3]
 [4 5 6]]
Sum of array elements: 15
Mean of array elements: 3.0
Max of array elements: 5
Array squared: [ 1  4  9 16 25]
Transposed Matrix:
[[1 4]
 [2 5]
 [3 6]]
Element at position (1, 2): 6
Zeros Array:
[[0. 0. 0.]
 [0. 0. 0.]]
Ones Array:
[[1. 1.]
 [1. 1.]
 [1. 1.]]
Identity Matrix:
[[1. 0. 0.]
 [0. 1. 0.]
 [0. 0. 1.]]
(env) cselab7-08@cselab708-OptiPlex-3050:~$ vi num.py
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 num.py
Array: [1 2 3 4 5]
Matrix:
[[1 2 3]
 [4 5 6]]
Sum of array elements: 15
Mean of array elements: 3.0
Max of array elements: 5
Array squared: [ 1  4  9 16 25]
data = {
    'Name': ['Alice', 'Bob', 'Charlie', 'David'],
    'Age': [24, 27, 22, 32],
    'City': ['New York', 'Los Angeles', 'Chicago', 'Houston']
}
```

```

df = pd.DataFrame(data)

print("DataFrame:\n", df)

# Basic DataFrame operations
print("\nFirst 2 rows:\n", df.head(2)) # Display the first 2 rows
print("\nData types:\n", df.dtypes) # Display the data types of each column

# Adding a new column
df['Salary'] = [70000, 80000, 60000, 90000]
print("\nDataFrame with Salary:\n", df)

# Accessing a column
print("\nAges:\n", df['Age'])

# Accessing a row by index
print("\nRow at index 1:\n", df.iloc[1])

# Filtering data
filtered_df = df[df['Age'] > 25]
print("\nFiltered DataFrame (Age > 25):\n", filtered_df)

# Descriptive statistics
print("\nDescriptive statistics:\n", df.describe())

# Handling missing values
df_with_na = df.copy()
df_with_na.loc[1, 'Salary'] = None # Introduce a missing value
print("\nDataFrame with missing values:\n", df_with_na)

# Fill missing values
df_filled = df_with_na.fillna(df['Salary'].mean()) # Fill NaNs with the mean of the Salary column
print("\nDataFrame after filling missing values:\n", df_filled)

```

OUTPUT:

Activities Terminal Aug 13 11:01

```
[0. 0. 1.]
(env) cselab7-08@cselab708-OptiPlex-3050:~$ pip3 install pandas
Collecting pandas
  Downloading pandas-2.2.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (13.0 MB)
    13.0/13.0 MB 27.0 MB/s eta 0:00:00
Collecting pytz>=2020.1
  Downloading pytz-2024.1-py2.py3-none-any.whl (505 kB)
    505.5/505.5 KB 10.0 MB/s eta 0:00:00
Requirement already satisfied: numpy>=1.22.4 in ./env/lib/python3.10/site-packages (from pandas) (2.0.1)
Collecting python-dateutil>=2.8.2
  Downloading python_dateutil-2.9.0.post0-py3-none-any.whl (229 kB)
    229.9/229.9 KB 4.9 MB/s eta 0:00:00
Collecting tzdata>=2022.7
  Downloading tzdata-2024.1-py2.py3-none-any.whl (345 kB)
    345.4/345.4 KB 6.2 MB/s eta 0:00:00
Collecting six>=1.5
  Downloading six-1.16.0-py2.py3-none-any.whl (11 kB)
Installing collected packages: pytz, tzdata, six, python-dateutil, pandas
Successfully installed pandas-2.2.2 python-dateutil-2.9.0.post0 pytz-2024.1 six-1.16.0 tzdata-2024.1
(env) cselab7-08@cselab708-OptiPlex-3050:~$ vi panda.py
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 panda.py
DataFrame:
   Name  Age      City
0  Alice  24  New York
1    Bob  27  Los Angeles
2  Charlie  22    Chicago
3   David  32    Houston
First 2 rows:
   Name  Age      City
0  Alice  24  New York
1    Bob  27  Los Angeles
Data types:
   Name    object
   Age     int64
   City   object
  dtype: object
```

EXPT-2:

AIM: Install, configure and run Hadoop and HDFS.

DESCRIPTION:

Hadoop is an open-source framework that enables distributed storage and processing of large data sets across clusters of computers. The Hadoop Distributed File System (HDFS) is a core component that provides scalable and fault-tolerant storage by dividing large files into blocks and distributing them across multiple nodes. HDFS ensures high availability and reliability through data replication and automatic recovery.

COMMANDS:

Reference: <https://www.baeldung.com/linux/hadoop-install-configure>

```
# 1. Update System Packages
sudo apt update && sudo apt upgrade -y

# 2. Install Java Development Kit (JDK)
sudo apt install default-jdk -y
java -version

# 3. Create hadoop User
sudo adduser hadoop
sudo usermod -aG sudo hadoop
sudo su - hadoop

# 4. Configure Password-less SSH Access
sudo apt install openssh-server openssh-client -y
systemctl status ssh

# Generate SSH Key Pair
ssh-keygen -t rsa -P "" -f ~/.ssh/id_rsa

# Copy Public Key
cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys

# Configure SSH Permissions
chmod 700 ~/.ssh
chmod 600 ~/.ssh/authorized_keys

# Configure Firewall for SSH Access
```

```

sudo ufw allow ssh

# Test SSH Connection
ssh localhost

# 5. Install Apache Hadoop
sudo su - hadoop

# Download and Extract Hadoop
wget https://downloads.apache.org/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz
tar -xvzf hadoop-3.3.6.tar.gz

# Move Hadoop to Installation Directory
sudo mv hadoop-3.3.6 /usr/local/hadoop
sudo mkdir /usr/local/hadoop/logs
sudo chown -R hadoop:hadoop /usr/local/hadoop

# Configure Hadoop Environment Variables
nano ~/.bashrc
# Add the following lines:
export HADOOP_HOME=/usr/local/hadoop
export HADOOP_INSTALL=$HADOOP_HOME
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib/native"
# Apply changes
source ~/.bashrc

# Configure Java Environment Variables
which javac
readlink -f /usr/bin/javac
sudo nano $HADOOP_HOME/etc/hadoop/hadoop-env.sh
# Add the following lines:
export JAVA_HOME=/usr/lib/jvm/jdk-21-oracle-x64
export HADOOP_CLASSPATH+=" $HADOOP_HOME/lib/*.jar"

# Configure Hadoop Environment
cd /usr/local/hadoop/lib
sudo wget https://jcenter.bintray.com/javax/activation/javax.activation-api/1.2.0/javax.activation-api-1.2.0.jar
hadoop version

# Edit core-site.xml Configuration

```

```

sudo nano $HADOOP_HOME/etc/hadoop/core-site.xml
# Add the following lines:
<configuration>
  <property>
    <name>fs.default.name</name>
    <value>hdfs://0.0.0.0:9000</value>
    <description>The default file system URI</description>
  </property>
</configuration>

# Create Directories for HDFS
sudo mkdir -p /home/hadoop/hdfs/{namenode,datanode}
sudo chown -R hadoop:hadoop /home/hadoop/hdfs

# Edit hdfs-site.xml Configuration
sudo nano $HADOOP_HOME/etc/hadoop/hdfs-site.xml
# Add the following lines:
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
  </property>

  <property>
    <name>dfs.name.dir</name>
    <value>file:///home/hadoop/hdfs/namenode</value>
  </property>

  <property>
    <name>dfs.data.dir</name>
    <value>file:///home/hadoop/hdfs/datanode</value>
  </property>
</configuration>

# Edit mapred-site.xml Configuration
sudo nano $HADOOP_HOME/etc/hadoop/mapred-site.xml
# Add the following lines:
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>

# Edit yarn-site.xml Configuration
sudo nano $HADOOP_HOME/etc/hadoop/yarn-site.xml
# Add the following lines:

```

```
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
</configuration>

# Validate Hadoop Configuration and Format NameNode
hdfs namenode -format

# 8. Start the Apache Hadoop Cluster
# Start NameNode and DataNode
start-dfs.sh

# Start YARN Resource and Node Managers
start-yarn.sh

# Verify All Running Components
jps
```

EXPT-3

AIM: Visualize data using basic plotting techniques in Python.

Description:

This code demonstrates data visualization using the Iris dataset. It includes loading the dataset, displaying the first few rows, and creating various plots: a pairplot to show feature relationships, a boxplot of sepal length by species, a histogram of sepal lengths, a scatter plot of sepal length vs. width, and a heatmap of feature correlations. Each plot provides insights into different aspects of the dataset.

CODE:

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load the Iris dataset from seaborn
iris = sns.load_dataset('iris')

# Display the first few rows of the dataset
print("Iris Dataset:\n", iris.head())

# Plot pairplot
sns.pairplot(iris, hue='species')
plt.title('Pairplot of Iris Dataset')
plt.show()

# Plot a boxplot for each feature
plt.figure(figsize=(10, 6))
sns.boxplot(x='species', y='sepal_length', data=iris)
plt.title('Boxplot of Sepal Length by Species')
plt.show()

# Plot a histogram of sepal lengths
plt.figure(figsize=(10, 6))
sns.histplot(iris['sepal_length'], bins=20, kde=True)
plt.title('Histogram of Sepal Length')
plt.xlabel('Sepal Length')
plt.show()

# Plot a scatter plot of sepal length vs. sepal width
plt.figure(figsize=(10, 6))
sns.scatterplot(x='sepal_length', y='sepal_width', hue='species', data=iris)
plt.title('Scatter Plot of Sepal Length vs. Sepal Width')
```

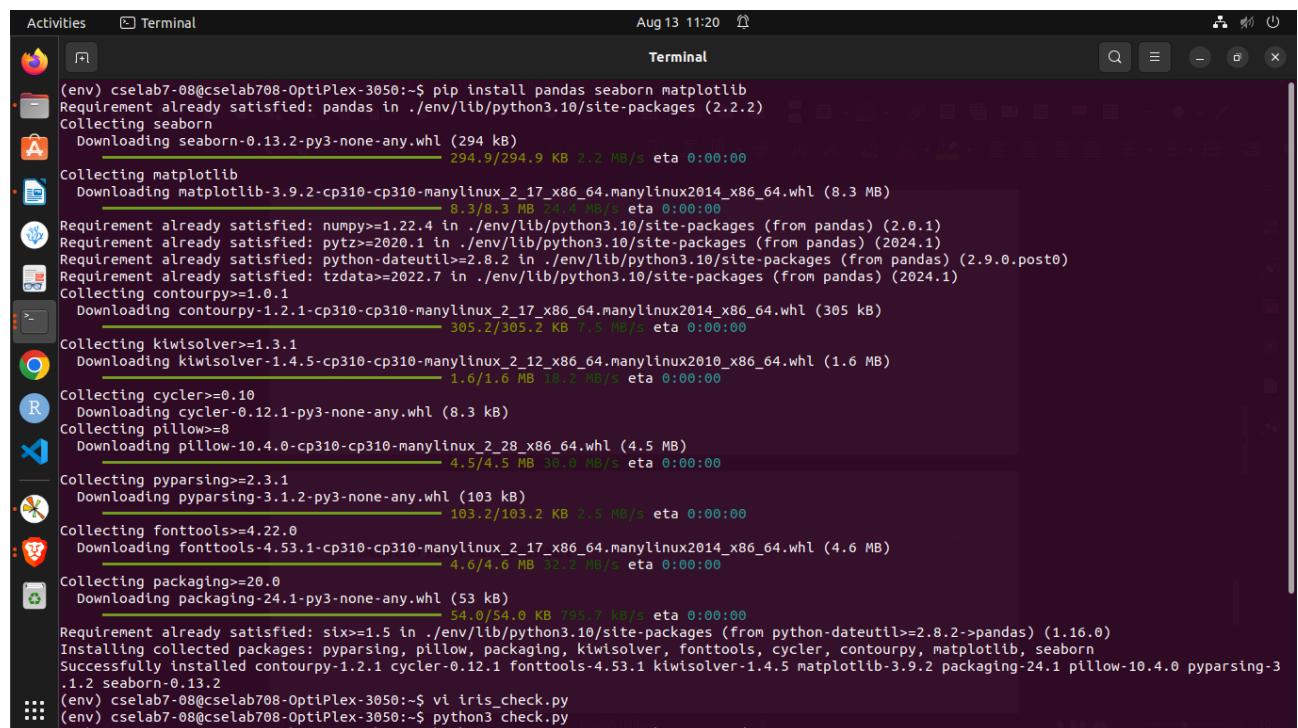
```

plt.xlabel('Sepal Length')
plt.ylabel('Sepal Width')
plt.show()

# Plot a heatmap of correlations
plt.figure(figsize=(10, 8))
correlation_matrix = iris.drop('species', axis=1).corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f')
plt.title('Heatmap of Feature Correlations')
plt.show()

```

OUTPUT:



The screenshot shows a terminal window titled "Terminal" running on a Linux desktop. The window displays the output of a pip command to install several Python packages. The packages listed include pandas, seaborn, matplotlib, numpy, pytz, python-dateutil, tzdata, contourpy, kiwisolver, cycler, pillow, pyparsing, fonttools, packaging, and six. The terminal shows the download progress for each package, indicating speeds like 2.2 MB/s and estimated times like 0:00:00. The background of the desktop shows icons for various applications like a browser, file manager, and terminal.

```

Activities Terminal Aug 13 11:20
Terminal
(env) cselab7-08@cselab708-OptiPlex-3050:~$ pip install pandas seaborn matplotlib
Requirement already satisfied: pandas in ./env/lib/python3.10/site-packages (2.2.2)
Collecting seaborn
  Downloading seaborn-0.13.2-py3-none-any.whl (294 kB)
    294.9/294.9 KB 2.2 MB/s eta 0:00:00
Collecting matplotlib
  Downloading matplotlib-3.9.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (8.3 MB)
    8.3/8.3 MB 24.4 MB/s eta 0:00:00
Requirement already satisfied: numpy>=1.22.4 in ./env/lib/python3.10/site-packages (from pandas) (2.0.1)
Requirement already satisfied: pytz>=2020.1 in ./env/lib/python3.10/site-packages (from pandas) (2024.1)
Requirement already satisfied: python-dateutil>=2.8.2 in ./env/lib/python3.10/site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: tzdata>=2022.7 in ./env/lib/python3.10/site-packages (from pandas) (2024.1)
Collecting contourpy>=1.0.1
  Downloading contourpy-1.2.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (305 kB)
    305.2/305.2 KB 0.5 MB/s eta 0:00:00
Collecting kiwisolver>=1.3.1
  Downloading kiwisolver-1.4.5-cp310-cp310-manylinux_2_12_x86_64.manylinux2010_x86_64.whl (1.6 MB)
    1.6/1.6 MB 10.2 MB/s eta 0:00:00
Collecting cycler>=0.10
  Downloading cycler-0.12.1-py3-none-any.whl (8.3 kB)
Collecting pillow>=8
  Downloading pillow-10.4.0-cp310-cp310-manylinux_2_28_x86_64.whl (4.5 MB)
    4.5/4.5 MB 30.0 MB/s eta 0:00:00
Collecting pyparsing>=2.3.1
  Downloading pyparsing-3.1.2-py3-none-any.whl (103 kB)
    103.2/103.2 KB 2.5 MB/s eta 0:00:00
Collecting fonttools>=4.22.0
  Downloading fonttools-4.53.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (4.6 MB)
    4.6/4.6 MB 32.0 MB/s eta 0:00:00
Collecting packaging>=20.0
  Downloading packaging-24.1-py3-none-any.whl (53 kB)
    54.0/54.0 KB 705.7 kB/s eta 0:00:00
Requirement already satisfied: six>=1.5 in ./env/lib/python3.10/site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Installing collected packages: pyparsing, pillow, packaging, kiwisolver, fonttools, cycler, contourpy, matplotlib, seaborn
Successfully installed contourpy-1.2.1 cycler-0.12.1 fonttools-4.53.1 kiwisolver-1.4.5 matplotlib-3.9.2 packaging-24.1 pillow-10.4.0 pyparsing-3.1.2 seaborn-0.13.2
(env) cselab7-08@cselab708-OptiPlex-3050:~$ vi iris_check.py
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 check.py

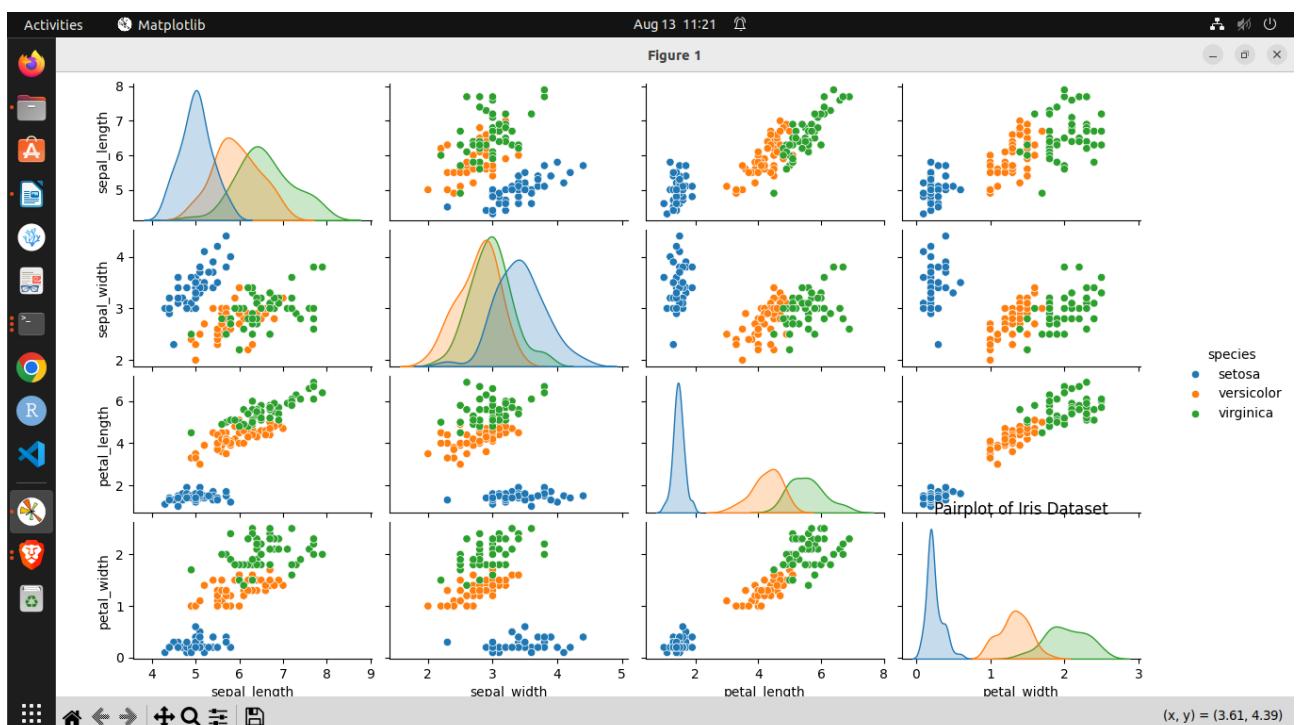
```

Activities Terminal Aug 13 11:21

```

Requirement already satisfied: python-dateutil>=2.8.2 in ./env/lib/python3.10/site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: tzdata>=2022.7 in ./env/lib/python3.10/site-packages (from pandas) (2024.1)
Collecting contourpy>=1.0.1
  Downloading contourpy-1.2.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (305 kB)
    305.2/305.2 KB 7.5 MB/s eta 0:00:00
Collecting kiwisolver>=1.3.1
  Downloading kiwisolver-1.4.5-cp310-cp310-manylinux_2_12_x86_64.manylinux2016_x86_64.whl (1.6 MB)
    1.6/1.6 MB 10.2 MB/s eta 0:00:00
Collecting cypher>=0.10
  Downloading cypher-0.12.1-py3-none-any.whl (8.3 kB)
Collecting pillow>=8
  Downloading pillow-10.4.0-cp310-cp310-manylinux_2_28_x86_64.whl (4.5 MB)
    4.5/4.5 MB 30.0 MB/s eta 0:00:00
Collecting pyparsing>=2.3.1
  Downloading pyparsing-3.1.2-py3-none-any.whl (103 kB)
    103.2/103.2 KB 2.5 MB/s eta 0:00:00
Collecting fonttools>=4.22.0
  Downloading fonttools-4.53.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (4.6 MB)
    4.6/4.6 MB 32.2 MB/s eta 0:00:00
Collecting packaging>=20.0
  Downloading packaging-24.1-py3-none-any.whl (53 kB)
    54.0/54.0 KB 795.7 kB/s eta 0:00:00
Requirement already satisfied: six>=1.5 in ./env/lib/python3.10/site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Installing collected packages: pyparsing, pillow, packaging, kiwisolver, fonttools, cypher, contourpy, matplotlib, seaborn
Successfully installed contourpy-1.2.1 cypher-0.12.1 fonttools-4.53.1 kiwisolver-1.4.5 matplotlib-3.9.2 packaging-24.1 pillow-10.4.0 pyparsing-3
.1.2 seaborn-0.13.2
(env) cselab7-08@cselab708-OptiPlex-3050:~$ vi iris_check.py
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 check.py
python3: can't open file '/home/cselab7-08/check.py': [Errno 2] No such file or directory
(env) cselab7-08@cselab708-OptiPlex-3050:~$ python3 iris_check.py
Iris Dataset:
   sepal_length  sepal_width  petal_length  petal_width species
0      5.1          3.5         1.4         0.2   setosa
1      4.9          3.0         1.4         0.2   setosa
2      4.7          3.2         1.3         0.2   setosa
3      4.6          3.1         1.5         0.2   setosa
4      5.0          3.6         1.4         0.2   setosa

```



EXPT-4

AIM: Implement NoSQL Database Operations: CRUD operations, Arrays using MongoDB.

DESCRIPTION:

To implement NoSQL database operations using MongoDB:

1. CRUD Operations:

- Create: Insert documents into a collection using `insert_one()` or `insert_many()`.
 - Read: Query documents using `find()` or `find_one()` to retrieve data based on specific criteria.
 - Update: Modify existing documents with `update_one()` or `update_many()` to change document fields.
 - Delete: Remove documents using `delete_one()` or `delete_many()` based on given conditions.

2. Arrays:

- Store arrays within documents and manipulate them using MongoDB's array operators. For example, use `'\$push` to add elements, `'\$pull` to remove elements, and `'\$addToSet` to ensure unique values in an array.

INSTALLING MONGODB

MONGODB COMMUNITY VERSION

MONGODB COMPASS

Activities Terminal Aug 20 10:58

```
Terminal
Terminal
Terminal

cselab7-08@cselab708-OptiPlex-3050:~$ sudo apt-get install gnupg curl
[sudo] password for cselab7-08:
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
curl is already the newest version (7.81.0-1ubuntu1.17).
gnupg is already the newest version (2.2.27-3ubuntu2.1).
0 upgraded, 0 newly installed, 0 to remove and 576 not upgraded.
cselab7-08@cselab708-OptiPlex-3050:~$ curl -fsSL https://www.mongodb.org/static/pgp/server-7.0.asc | \
  sudo gpg -o /usr/share/keyrings/mongodb-server-7.0.gpg \
  --dearmor
File '/usr/share/keyrings/mongodb-server-7.0.gpg' exists. Overwrite? (y/N) y
cselab7-08@cselab708-OptiPlex-3050:~$ echo [ arch=amd64,arm64 signed-by=/usr/share/keyrings/mongodb-server-7.0.gpg ] https://repo.mongodb.org/apt/ubuntu jammy/mongodb-org/7.0 multiverse" | sudo tee /etc/apt/sources.list.d/mongodb-org-7.0.list
deb [ arch=amd64,arm64 signed-by=/usr/share/keyrings/mongodb-server-7.0.gpg ] https://repo.mongodb.org/apt/ubuntu jammy/mongodb-org/7.0 multiverse
cselab7-08@cselab708-OptiPlex-3050:~$ sudo apt-get update
Ign:1 https://repo.mongodb.org/apt/ubuntu jammy/mongodb-org/7.0 InRelease
Hit:2 https://brave-browser-apt-release.s3.brave.com stable InRelease
Hit:3 https://packages.microsoft.com/repos/code stable InRelease
Hit:4 https://download.vscodium.com/debs vscodium InRelease
Get:5 https://dl.google.com/linux/chrome/deb stable InRelease [1,825 B]
Hit:6 http://in.archive.ubuntu.com/ubuntu jammy InRelease
Hit:7 https://repo.mongodb.org/apt/ubuntu jammy/mongodb-org/7.0 Release
Err:5 https://dl.google.com/linux/chrome/deb stable InRelease
The following signatures couldn't be verified because the public key is not available: NO_PUBKEY E88979FB9B30ACF2
Hit:8 http://in.archive.ubuntu.com/ubuntu jammy-updates InRelease
Hit:10 https://ppa.launchpadcontent.net/deadsnakes/ppa/ubuntu jammy InRelease
Hit:11 https://ppa.launchpadcontent.net/gns3/ppa/ubuntu jammy InRelease
Fetched 1,825 B in 1s (1,413 B/s)
Reading package lists... Done
N: Skipping acquire of configured file 'main/binary-i386/Packages' as repository 'https://brave-browser-apt-release.s3.brave.com stable InRelease' doesn't support architecture 'i386'
W: An error occurred during the signature verification. The repository is not updated and the previous index files will be used. GPG error: http://dl.google.com/linux/chrome/deb stable InRelease: The following signatures couldn't be verified because the public key is not available: NO_PUBKEY E88979FB9B30ACF2
```

Activities Terminal Aug 20 10:59

```

UBKEY E88979FB9B30ACFZ
W: Failed to fetch https://dl.google.com/linux/chrome/deb/dists/stable/InRelease  The following signatures couldn't be verified because the public key is not available: NO_PUBKEY E88979FB9B30ACFZ
W: Some index files failed to download. They have been ignored, or old ones used instead.
cselab7-08@cselab708-OptiPlex-3050:~$ sudo apt-get install -y mongodb-org
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
mongodb-org is already the newest version (7.0.12).
0 upgraded, 0 newly installed, 0 to remove and 576 not upgraded.
cselab7-08@cselab708-OptiPlex-3050:~$ echo "mongodb-org hold" | sudo dpkg --set-selections
echo "mongodb-org-database hold" | sudo dpkg --set-selections
echo "mongodb-org-server hold" | sudo dpkg --set-selections
echo "mongodb-mongosh hold" | sudo dpkg --set-selections
echo "mongodb-org-mongos hold" | sudo dpkg --set-selections
echo "mongodb-org-tools hold" | sudo dpkg --set-selections
cselab7-08@cselab708-OptiPlex-3050:~$ ps -o comm 1
systemd
cselab7-08@cselab708-OptiPlex-3050:~$ sudo systemctl start mongod
cselab7-08@cselab708-OptiPlex-3050:~$ sudo systemctl status mongod
● mongod.service - MongoDB Database Server
   Loaded: loaded (/lib/systemd/system/mongod.service; enabled; vendor preset: enabled)
   Active: active (running) since Tue 2024-08-20 10:22:31 IST; 35min ago
     Docs: https://docs.mongodb.org/manual
     Main PID: 1434 (mongod)
       Memory: 256.7M
          CPU: 17.730s
        CGroup: /system.slice/mongod.service
                 └─1434 /usr/bin/mongod -c /etc/mongod.conf

Aug 20 10:22:31 cselab708-OptiPlex-3050 systemd[1]: Started MongoDB Da>
Aug 20 10:22:46 cselab708-OptiPlex-3050 mongod[1434]: {"t":{$date:"2>
lines 1-12/12 (END)...skipping...
● mongod.service - MongoDB Database Server
   Loaded: loaded (/lib/systemd/system/mongod.service; enabled; vendor preset: enabled)
   Active: active (running) since Tue 2024-08-20 10:22:31 IST; 35min ago
     Docs: https://docs.mongodb.org/manual

```

Activities Terminal Aug 20 10:53

```

Activities Terminal Aug 20 10:53
cselab7-08@cselab708-OptiPlex-3050:~$ wget https://downloads.mongodb.com/compass/mongodb-compass_1.43.5_amd64.deb
--2024-08-20 10:51:26-- https://downloads.mongodb.com/compass/mongodb-compass_1.43.5_amd64.deb
Resolving downloads.mongodb.com (downloads.mongodb.com)... 108.157.238.17, 108.157.238.41, 108.157.238.92, ...
Connecting to downloads.mongodb.com (downloads.mongodb.com)|108.157.238.17|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 109811260 (105M) [application/octet-stream]
Saving to: 'mongodb-compass_1.43.5_amd64.deb'

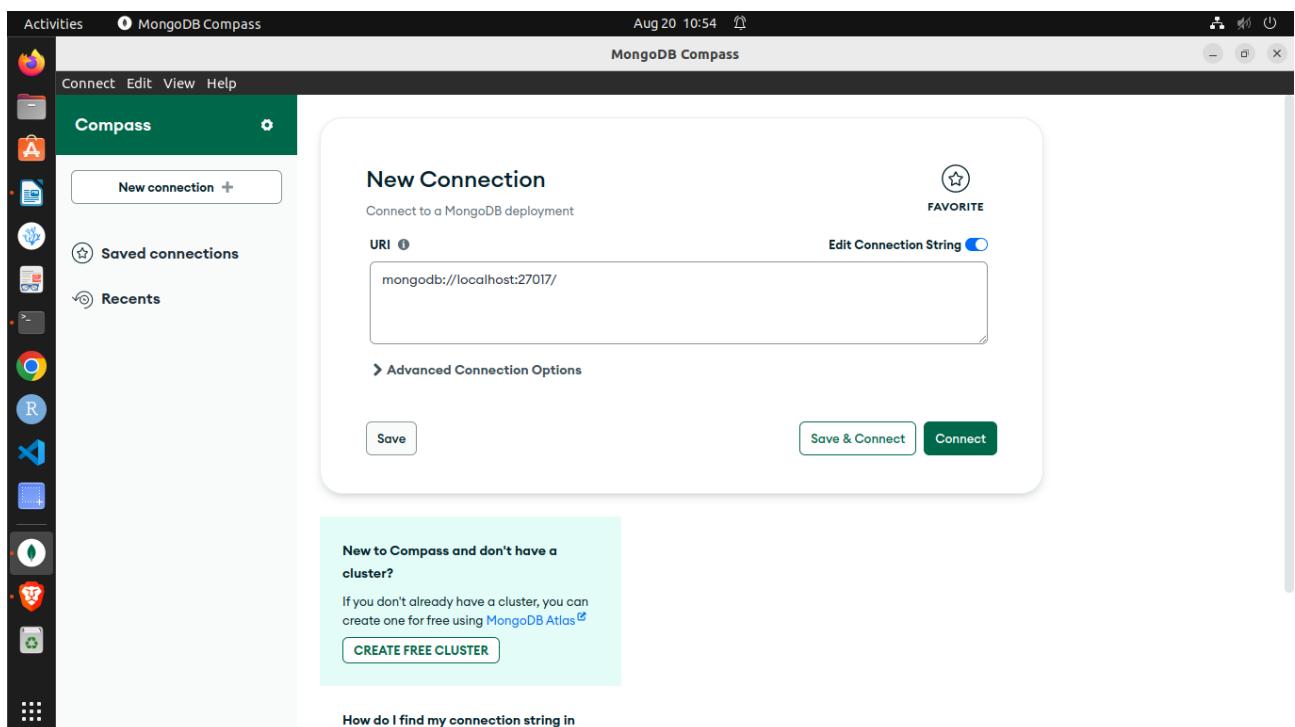
mongodb-compass_1 100%[=====] 104.72M  10.7MB/s  in 9.7s

2024-08-20 10:51:37 (10.8 MB/s) - 'mongodb-compass_1.43.5_amd64.deb' saved [109811260/109811260] ution does not support using apt for installing
cselab7-08@cselab708-OptiPlex-3050:~$ sudo apt install ./mongodb-compass_1.43.5_amd64.deb
[sudo] password for cselab7-08:
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
Note, selecting 'mongodb-compass' instead of './mongodb-compass_1.43.5_amd64.deb'
Suggested packages:
  gir1.2-gnomekeyring-1.0 libgnome-keyring0
The following NEW packages will be installed:
  mongodb-compass
0 upgraded, 1 newly installed, 0 to remove and 576 not upgraded.
Need to get 0 B/110 MB of archives.
After this operation, 438 MB of additional disk space will be used.
Get:1 /home/cselab7-08/mongodb-compass_1.43.5_amd64.deb mongodb-compass amd64 1.43.5 [110 MB] compass
Selecting previously unselected package mongodb-compass.
(Reading database ... 307861 files and directories currently installed.)
Preparing to unpack .../mongodb-compass_1.43.5_amd64.deb ...
Unpacking mongodb-compass (1.43.5) ...
Setting up mongodb-compass (1.43.5) ...
Processing triggers for desktop-file-utils (0.26-1ubuntu3) ...
Processing triggers for gnome-menus (3.36.0-1ubuntu3) ...
Processing triggers for mailcap (3.70+nmuiubuntu1) ...
N: Download is performed unsandboxed as root as file '/home/cselab7-08/mongodb-compass_1.43.5_amd64.deb' couldn't be accessed by user '_apt'. - pkgAcquire::Run (13: Permission denied)
cselab7-08@cselab708-OptiPlex-3050:~$ sudo dpkg -i mongodb-compass_1.43.5_amd64.deb
sudo apt-get install -f # This installs required compass dependencies
Install Compass
1. Double-click the installer file.
2. Follow the prompts to install Compass. You can safely ignore the warning about the file being owned by root.
Once installed, Compass launches and prompts you to configure it.

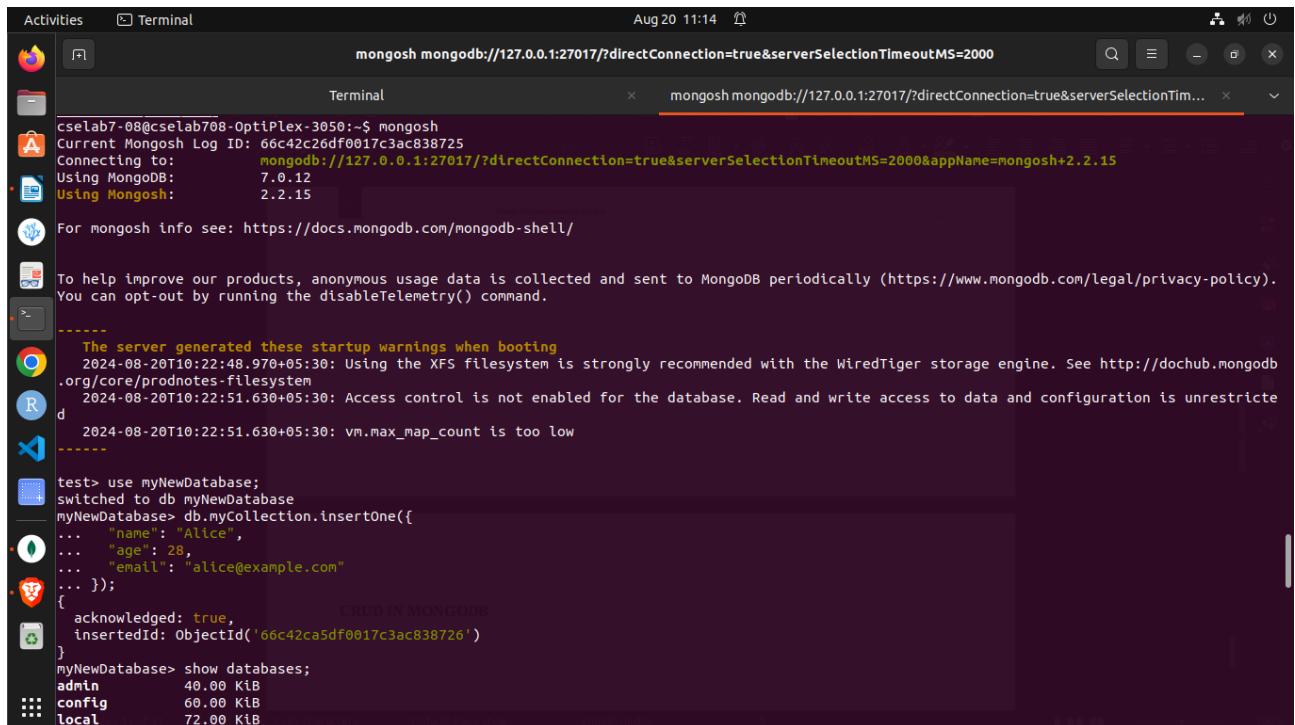
```

Activities Terminal Aug 20 10:54

```
[sudo] password for cselab7-08:
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
Note, selecting 'mongodb-compass' instead of './mongodb-compass_1.43.5_amd64.deb'
Suggested packages:
  gir1.2-gnomekeyring-1.0 libgnome-keyring0
The following NEW packages will be installed:
  mongodb-compass
0 upgraded, 1 newly installed, 0 to remove and 576 not upgraded.
Need to get 0 B/110 MB of archives.
After this operation, 438 MB of additional disk space will be used.
Get:1 /home/cselab7-08/mongodb-compass_1.43.5_amd64.deb mongodb-compass amd64 1.43.5 [110 MB]
Selecting previously unselected package mongodb-compass.
(Reading database ... 307861 files and directories currently installed.)
Preparing to unpack .../mongodb-compass_1.43.5_amd64.deb ...
Unpacking mongodb-compass (1.43.5) ...
Setting up mongodb-compass (1.43.5) ...
Processing triggers for desktop-file-utils (0.26-1ubuntu3) ...
Processing triggers for gnome-menus (3.36.0-1ubuntu3) ...
Processing triggers for mailcap (3.70+mu1ubuntu1) ...
N: Download is performed unsandboxed as root as file '/home/cselab7-08/mongodb-compass_1.43.5_amd64.deb' couldn't be accessed by user '_apt'.
pkgAcquire::Run (13: Permission denied)
cselab7-08@cselab708-OptiPlex-3050:~$ sudo dpkg -i mongodb-compass_1.43.5_amd64.deb
sudo apt-get install -f # This installs required compass dependencies
(Reading database ... 308291 files and directories currently installed.)
Preparing to unpack mongodb-compass_1.43.5_amd64.deb ...
Unpacking mongodb-compass (1.43.5) over (1.43.5) ...
Setting up mongodb-compass (1.43.5) ...
Processing triggers for mailcap (3.70+mu1ubuntu1) ...
Processing triggers for gnome-menus (3.36.0-1ubuntu3) ...
Processing triggers for desktop-file-utils (0.26-1ubuntu3) ...
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
0 upgraded, 0 newly installed, 0 to remove and 576 not upgraded.
cselab7-08@cselab708-OptiPlex-3050:~$ mongodb-compass
```

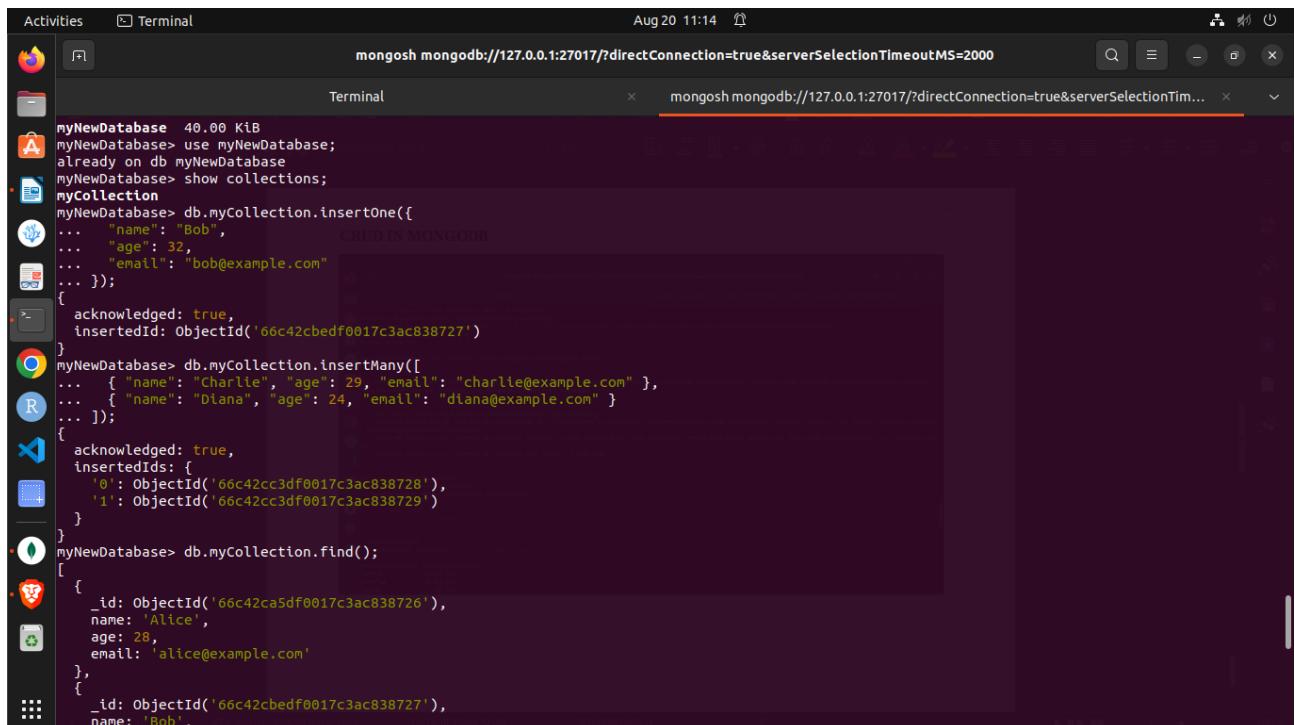


CRUD IN MONGODB



```
Activities Terminal Aug 20 11:14 mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTim...
cselab7-0@cselab708-OptiPlex-3050:~$ mongosh
Current Mongosh Log ID: 66c42c26df0017c3ac838725
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.2.15
Using MongoDB:      7.0.12
Using Mongosh:      2.2.15
For mongosh info see: https://docs.mongodb.com/mongodb-shell/
To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy). You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting
2024-08-20T10:22:48.970+05:30: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2024-08-20T10:22:51.630+05:30: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2024-08-20T10:22:51.630+05:30: vm.max_map_count is too low
-----
test> use myNewDatabase;
switched to db myNewDatabase
myNewDatabase> db.myCollection.insertOne({
...   "name": "Alice",
...   "age": 28,
...   "email": "alice@example.com"
... });
{
  acknowledged: true,          CRUD IN MONGODB
  insertedId: ObjectId('66c42ca5df0017c3ac838726')
}
myNewDatabase> show databases;
admin      40.00 KiB
config    60.00 KiB
local     72.00 KiB
```



```
Activities Terminal Aug 20 11:14 mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTim...
myNewDatabase 40.00 KiB
myNewDatabase> use myNewDatabase;
already on db myNewDatabase
myNewDatabase> show collections;
myCollection
myNewDatabase> db.myCollection.insertOne({
...   "name": "Bob",
...   "age": 32,
...   "email": "bob@example.com"
... });
{
  acknowledged: true,
  insertedId: ObjectId('66c42cbef0017c3ac838727')
}
myNewDatabase> db.myCollection.insertMany([
...   { "name": "Charlie", "age": 29, "email": "charlie@example.com" },
...   { "name": "Diana", "age": 24, "email": "diana@example.com" }
... ]);
{
  acknowledged: true,
  insertedIds: {
    '0': ObjectId('66c42cc3df0017c3ac838728'),
    '1': ObjectId('66c42cc3df0017c3ac838729')
  }
}
myNewDatabase> db.myCollection.find();
[
  {
    _id: ObjectId('66c42ca5df0017c3ac838726'),
    name: 'Alice',
    age: 28,
    email: 'alice@example.com'
  },
  {
    _id: ObjectId('66c42cbef0017c3ac838727'),
    name: 'Bob',
  }
]
```

Activities Terminal Aug 20 11:14

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
```

myNewDatabase> use myNewDatabase;
already on db myNewDatabase
myNewDatabase> show collections;
myCollection
myNewDatabase> db.myCollection.insertOne({
... "name": "Bob",
... "age": 32,
... "email": "bob@example.com"
... });
{
 acknowledged: true,
 insertedId: ObjectId('66c42cbef0017c3ac838727')
}
myNewDatabase> db.myCollection.insertMany([
... { "name": "Charlie", "age": 29, "email": "charlie@example.com" },
... { "name": "Diana", "age": 24, "email": "diana@example.com" }
...]);
{
 acknowledged: true,
 insertedIds: {
 '0': ObjectId('66c42cc3df0017c3ac838728'),
 '1': ObjectId('66c42cc3df0017c3ac838729')
 }
}
myNewDatabase> db.myCollection.find();
[
 {
 _id: ObjectId('66c42ca5df0017c3ac838726'),
 name: 'Alice',
 age: 28,
 email: 'alice@example.com'
 },
 {
 _id: ObjectId('66c42cbef0017c3ac838727'),
 name: 'Bob',
 age: 32,
 email: 'bob@example.com'
 }
]

Activities Terminal Aug 20 11:15

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
```

myNewDatabase> db.myCollection.find({ "age": { \$gt: 30 } });
[
 {
 _id: ObjectId('66c42cbef0017c3ac838727'),
 name: 'Bob',
 age: 32,
 email: 'bob@example.com'
 }
]
myNewDatabase> db.myCollection.find({ "age": { \$gt: 30 } }).pretty();
[
 {
 _id: ObjectId('66c42cbef0017c3ac838727'),
 name: 'Bob',
 age: 32,
 email: 'bob@example.com'
 }
]
myNewDatabase> db.myCollection.updateOne(
... { "name": "Alice" }, // filter
... { \$set: { "age": 29 } } // update operation
...);
{
 acknowledged: true,
 insertedId: null,
 matchedCount: 1,
 modifiedCount: 1,
 upsertedCount: 0
}
myNewDatabase> db.myCollection.deleteOne({ "name": "Bob" });
{ acknowledged: true, deletedCount: 1 }
myNewDatabase>

EXPT-5

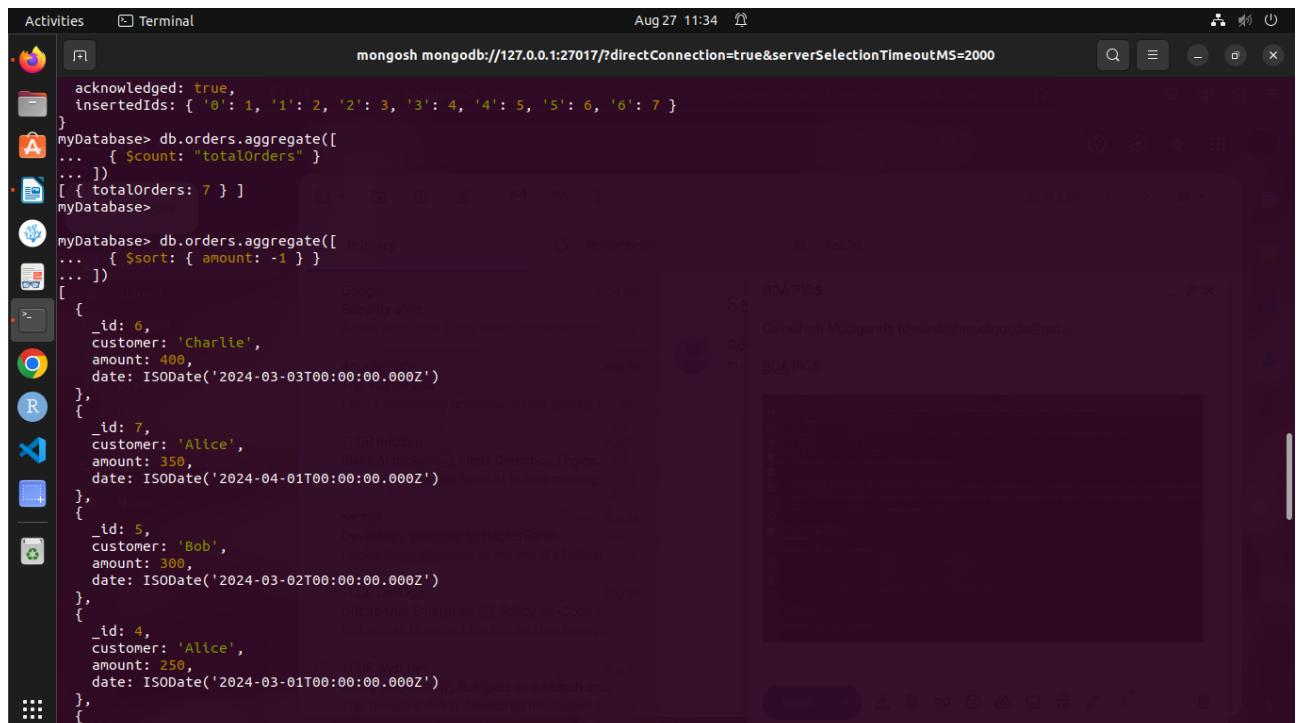
AIM: Implement Functions: Count – Sort – Limit – Skip – Aggregate using MongoDB.

Description:

In MongoDB, you can implement various functions for data manipulation:

1. Count: Use `count_documents()` to count the number of documents matching a query.
2. Sort: Use `sort()` to arrange documents in ascending or descending order based on specified fields.
3. Limit: Use `limit()` to restrict the number of documents returned by a query.
4. Skip: Use `skip()` to bypass a specified number of documents before returning results.
5. Aggregate: Use the `aggregate()` method to perform complex data processing operations like filtering, grouping, and sorting through a pipeline of stages.

CODE AND OUTPUTS:



```
Activities Terminal Aug 27 11:34 mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
{
  acknowledged: true,
  insertedIds: { '0': 1, '1': 2, '2': 3, '3': 4, '4': 5, '5': 6, '6': 7 }
}
myDatabase> db.orders.aggregate([
...   { $count: "totalOrders" }
... ])
[ { totalOrders: 7 } ]
myDatabase>
myDatabase> db.orders.aggregate([
...   { $sort: { amount: -1 } }
... ])
[ {
  _id: 6,
  customer: 'Charlie',
  amount: 400,
  date: ISODate('2024-03-03T00:00:00.000Z')
},
{
  _id: 7,
  customer: 'Alice',
  amount: 350,
  date: ISODate('2024-04-01T00:00:00.000Z')
},
{
  _id: 5,
  customer: 'Bob',
  amount: 300,
  date: ISODate('2024-03-02T00:00:00.000Z')
},
{
  _id: 4,
  customer: 'Alice',
  amount: 250,
  date: ISODate('2024-03-01T00:00:00.000Z')
},
{
  _id: 3,
  customer: 'Charlie',
  amount: 200,
  date: ISODate('2024-02-28T00:00:00.000Z')
},
{
  _id: 2,
  customer: 'Bob',
  amount: 150,
  date: ISODate('2024-02-27T00:00:00.000Z')
},
{
  _id: 1,
  customer: 'Alice',
  amount: 100,
  date: ISODate('2024-02-26T00:00:00.000Z')
} ]
```

Activities Terminal Aug 27 11:33

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
Cselab7-08@cselab708-OptiPlex-3050:~$ mongosh
Current Mongosh Log ID: 66cd6bf448287aea4c838725
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.2.15
Using MongoDB:     7.0.12
Using Mongosh:    2.2.15
mongosh 2.3.0 is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://docs.mongodb.com/mongodb-shell/
-----
The server generated these startup warnings when booting
2024-08-27T10:14:18.030+05:30: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2024-08-27T10:14:20.941+05:30: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2024-08-27T10:14:20.941+05:30: vm.max_map_count is too low
-----
test> use myDatabase
switched to db myDatabase
myDatabase>
myDatabase> db.orders.insertMany([
...   { _id: 1, customer: "Alice", amount: 100, date: new ISODate("2024-01-01") },
...   { _id: 2, customer: "Bob", amount: 200, date: new ISODate("2024-01-02") },
...   { _id: 3, customer: "Alice", amount: 150, date: new ISODate("2024-02-01") },
...   { _id: 4, customer: "Alice", amount: 250, date: new ISODate("2024-03-01") },
...   { _id: 5, customer: "Bob", amount: 300, date: new ISODate("2024-03-02") },
...   { _id: 6, customer: "Charlie", amount: 400, date: new ISODate("2024-03-03") },
...   { _id: 7, customer: "Alice", amount: 350, date: new ISODate("2024-04-01") }
... ])
{
  acknowledged: true,
  insertedIds: { '0': 1, '1': 2, '2': 3, '3': 4, '4': 5, '5': 6, '6': 7 }
}
myDatabase> db.orders.aggregate([
...   { $count: "totalOrders" }
... ])
{
  totalOrders: 7
}
```

Activities Terminal Aug 27 11:34

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
myDatabase> db.orders.aggregate([
...   { $limit: 3 }
... ])
[ {
  _id: 1,
  customer: 'Alice',
  amount: 100,
  date: ISODate('2024-01-01T00:00:00.000Z')
},
{
  _id: 2,
  customer: 'Bob',
  amount: 200,
  date: ISODate('2024-01-02T00:00:00.000Z')
},
{
  _id: 3,
  customer: 'Alice',
  amount: 150,
  date: ISODate('2024-02-01T00:00:00.000Z')
}]
myDatabase> db.orders.aggregate([
...   { $skip: 2 }
... ])
[ {
  _id: 3,
  customer: 'Alice',
  amount: 150,
  date: ISODate('2024-02-01T00:00:00.000Z')
},
{
  _id: 4,
  customer: 'Alice',
  amount: 250,
  date: ISODate('2024-03-01T00:00:00.000Z')
}]
```

Activities Terminal Aug 27 11:34 mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000

```
customer: 'Alice', amount: 150, date: ISODate('2024-02-01T00:00:00.000Z')
},
{
  _id: 4,
  customer: 'Alice',
  amount: 250,
  date: ISODate('2024-03-01T00:00:00.000Z')
},
{
  _id: 5,
  customer: 'Bob',
  amount: 300,
  date: ISODate('2024-03-02T00:00:00.000Z')
},
{
  _id: 6,
  customer: 'Charlie',
  amount: 400,
  date: ISODate('2024-03-03T00:00:00.000Z')
},
{
  _id: 7,
  customer: 'Alice',
  amount: 350,
  date: ISODate('2024-04-01T00:00:00.000Z')
]
myDatabase> db.orders.aggregate([
...   { $group: { _id: "$customer", totalAmount: { $sum: "$amount" } } },
...   { $sort: { totalAmount: -1 } },
...   { $skip: 1 },
...   { $limit: 1 },
...   { $count: "finalCount" }
... ],
... [ { finalCount: 1 } ]
myDatabase>
```