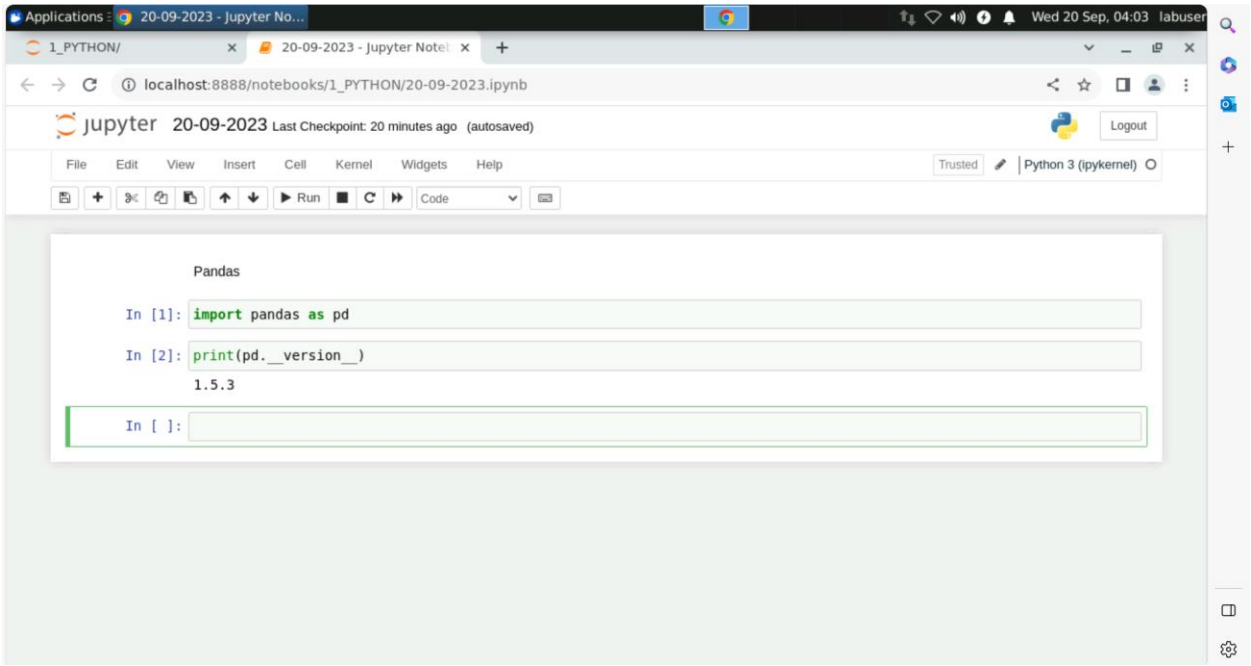
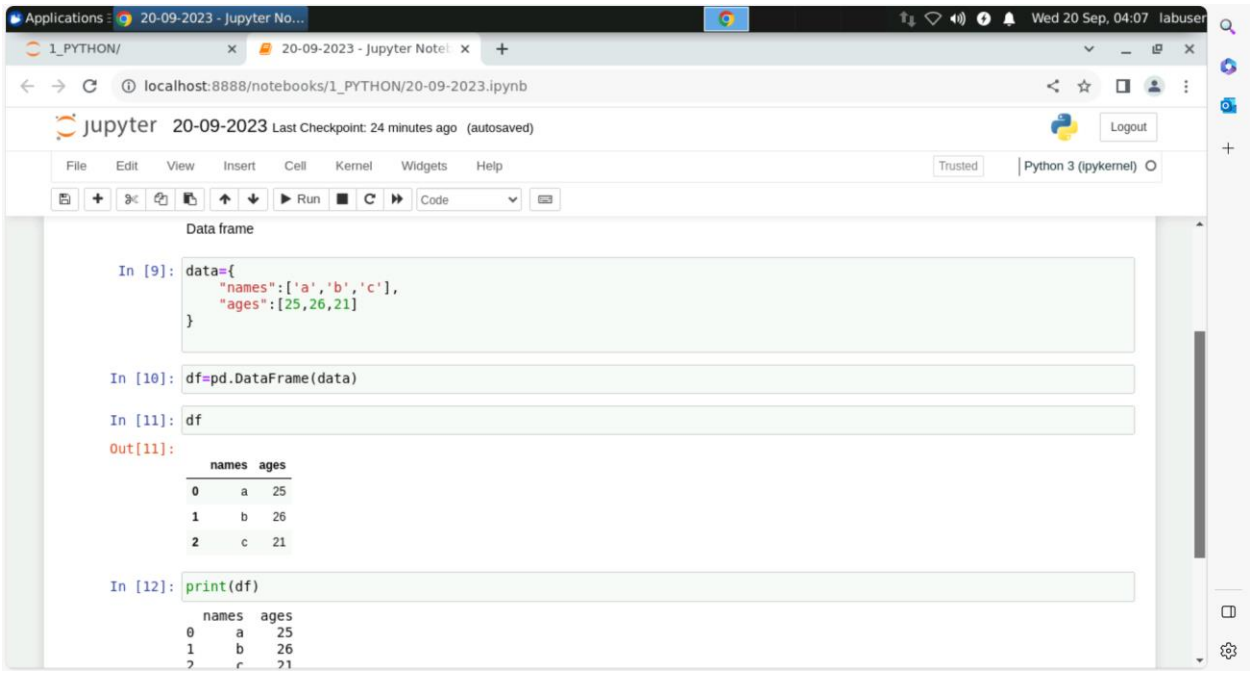
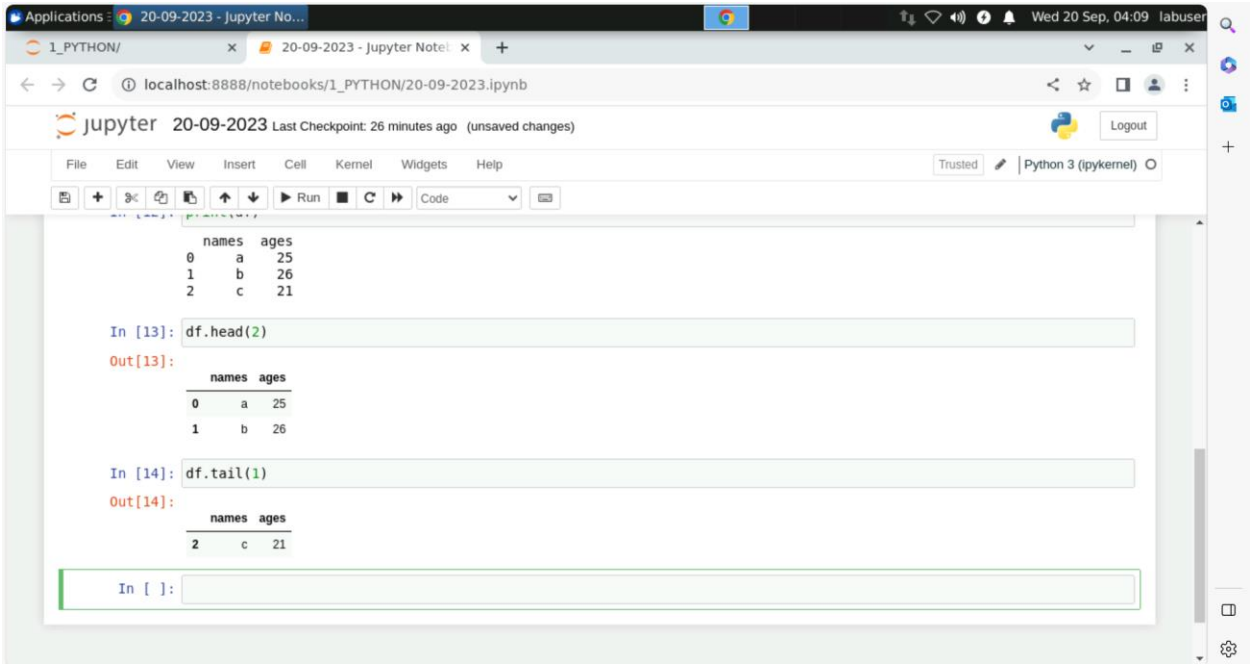
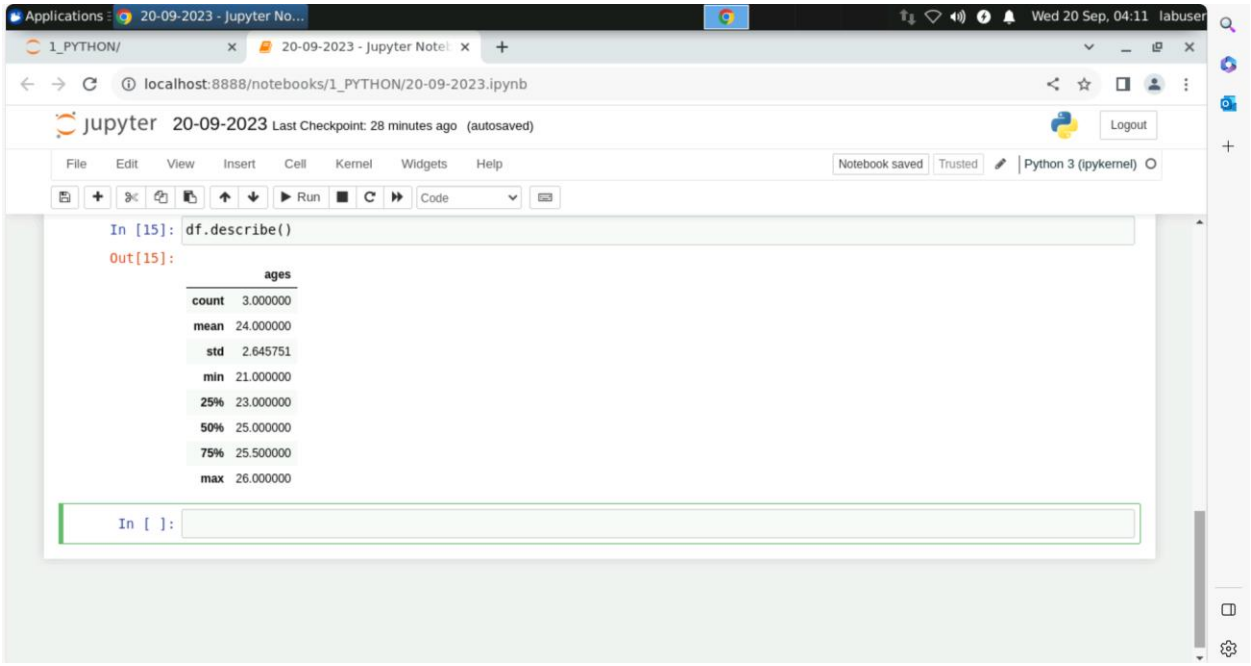
Pandas

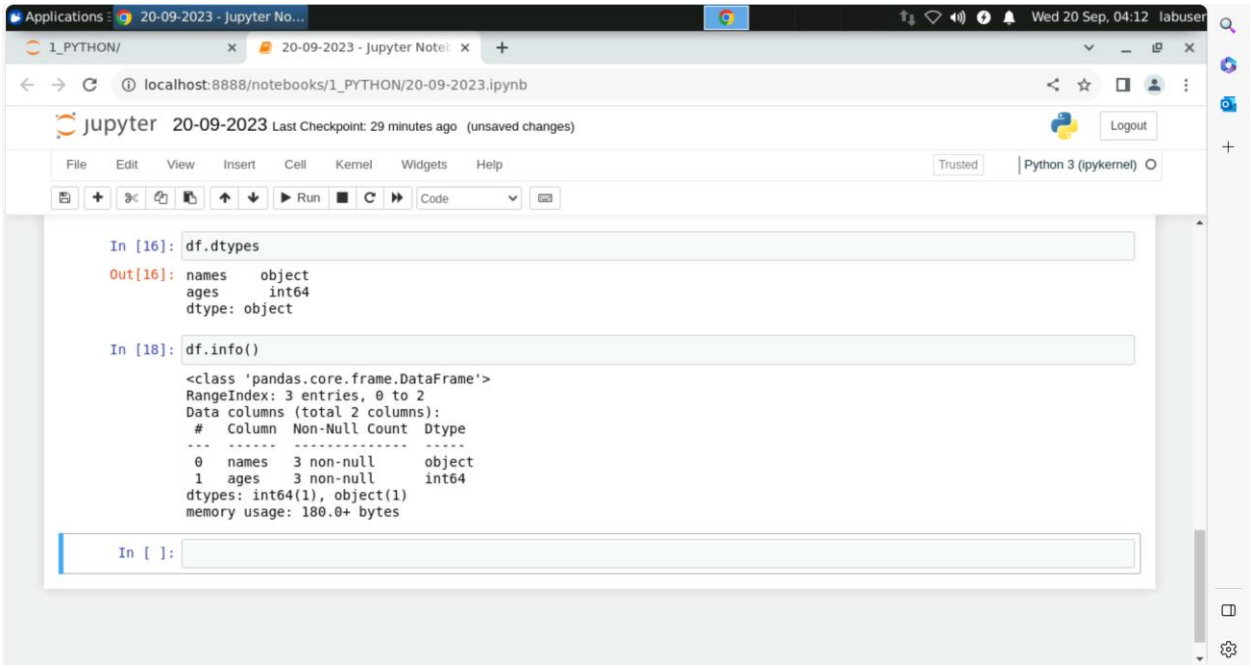
https://pandas.pydata.org/docs/

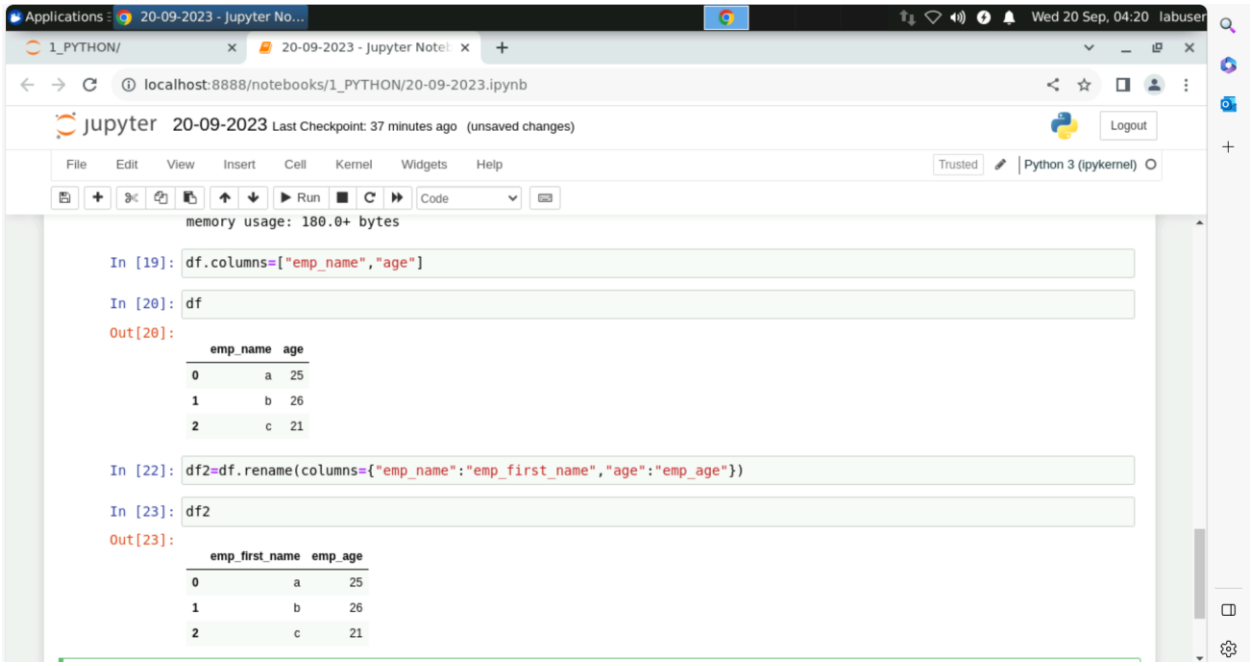


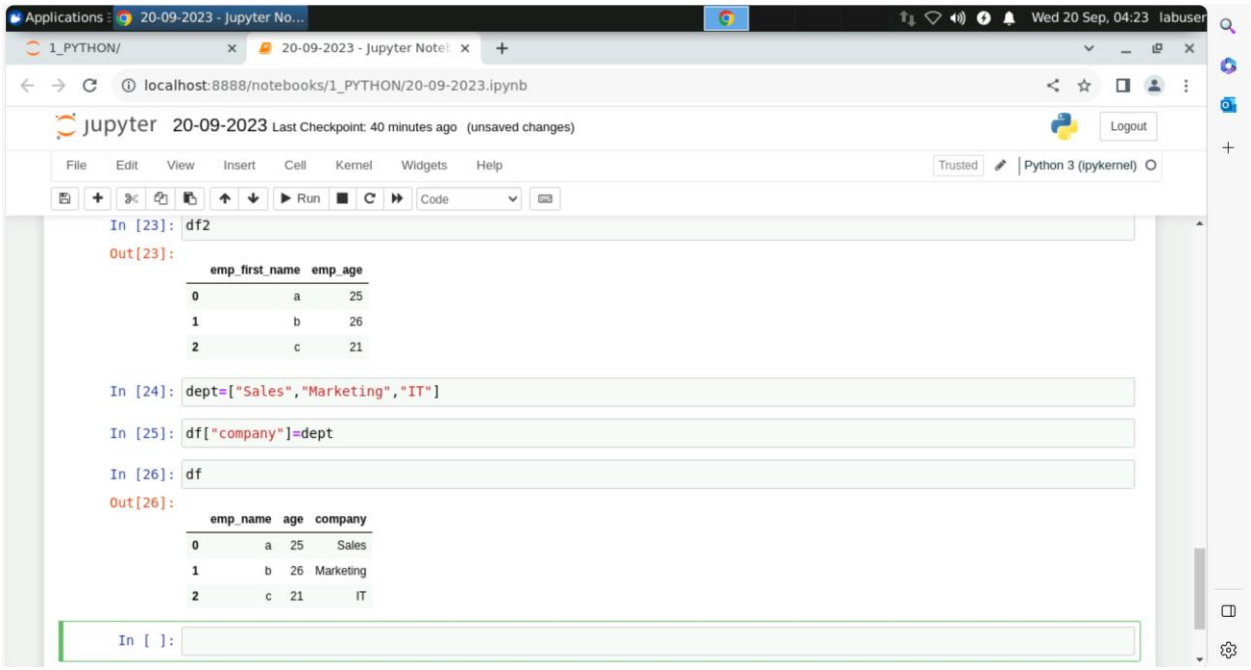


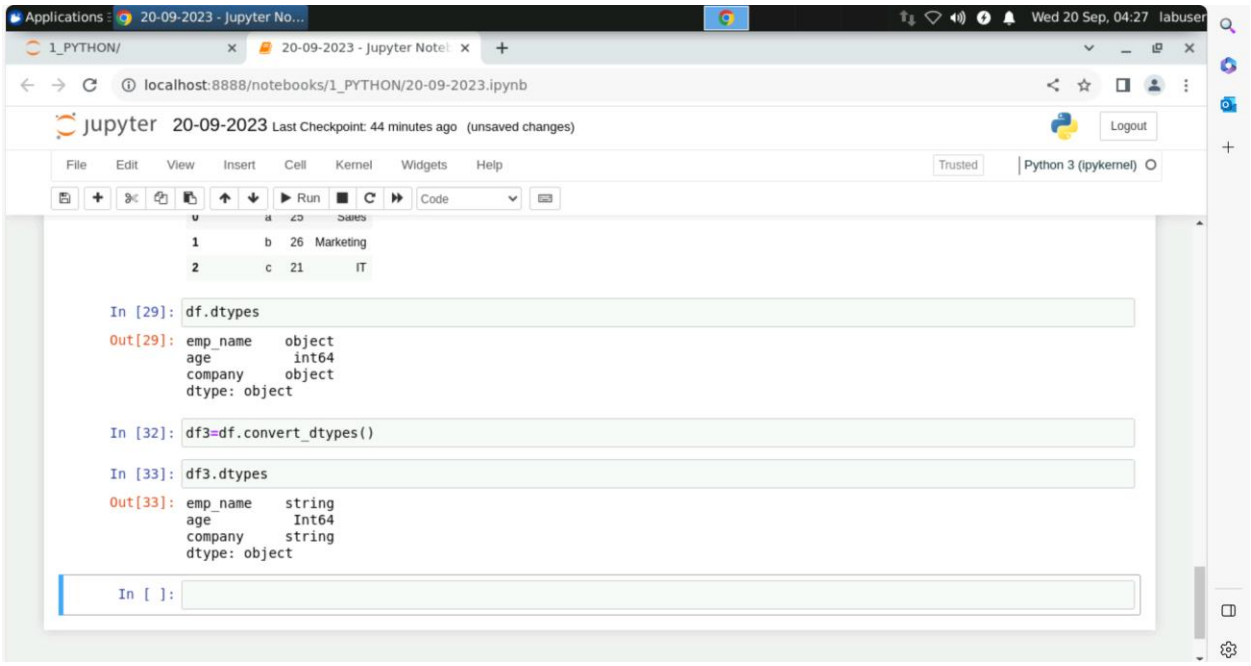

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:09
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 26 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
names ages
0 a 25
1 b 26
2 c 21
In [13]: df.head(2)
Out[13]:
names ages
0 a 25
1 b 26
In [14]: df.tail(1)
Out[14]:
names ages
2 c 21
In [ ]:

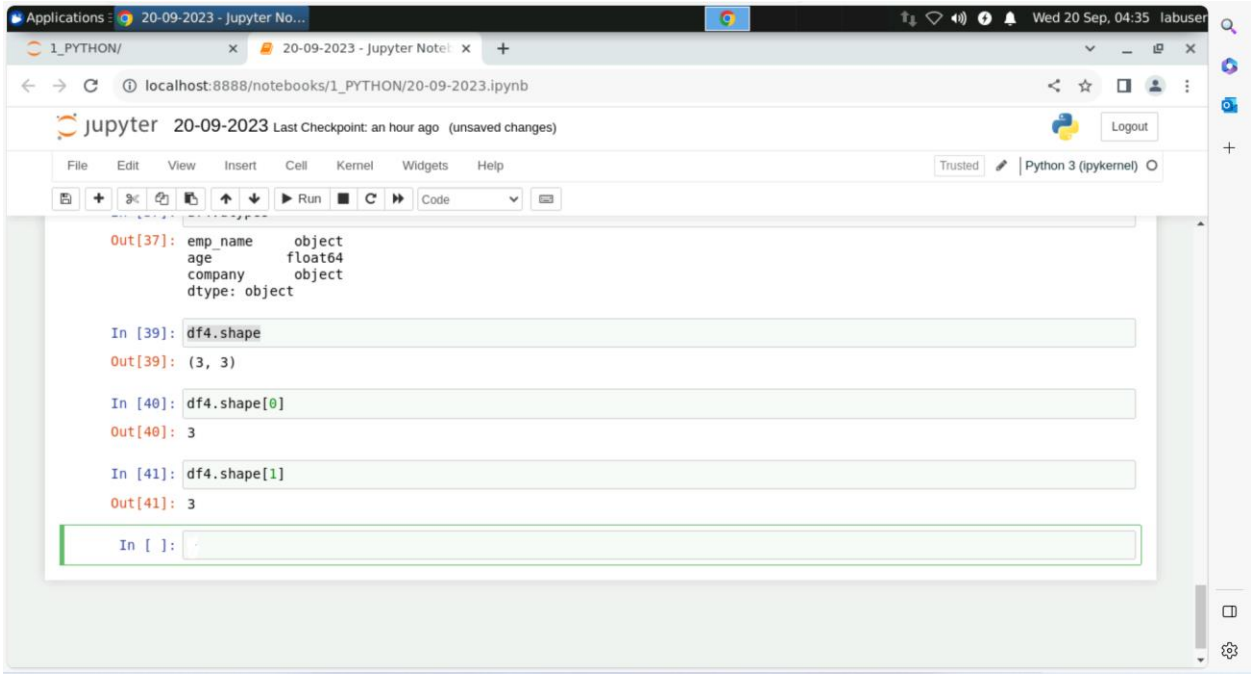

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:11
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 28 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Notebook saved
Trusted
Python 3 (ipykernel)
Run
Code
In [15]: df.describe()
Out[15]:
ages
count 3.000000
mean 24.000000
std 2.645751
min 21.000000
25% 23.000000
50% 25.000000
75% 25.500000
max 26.000000
In [ ]:

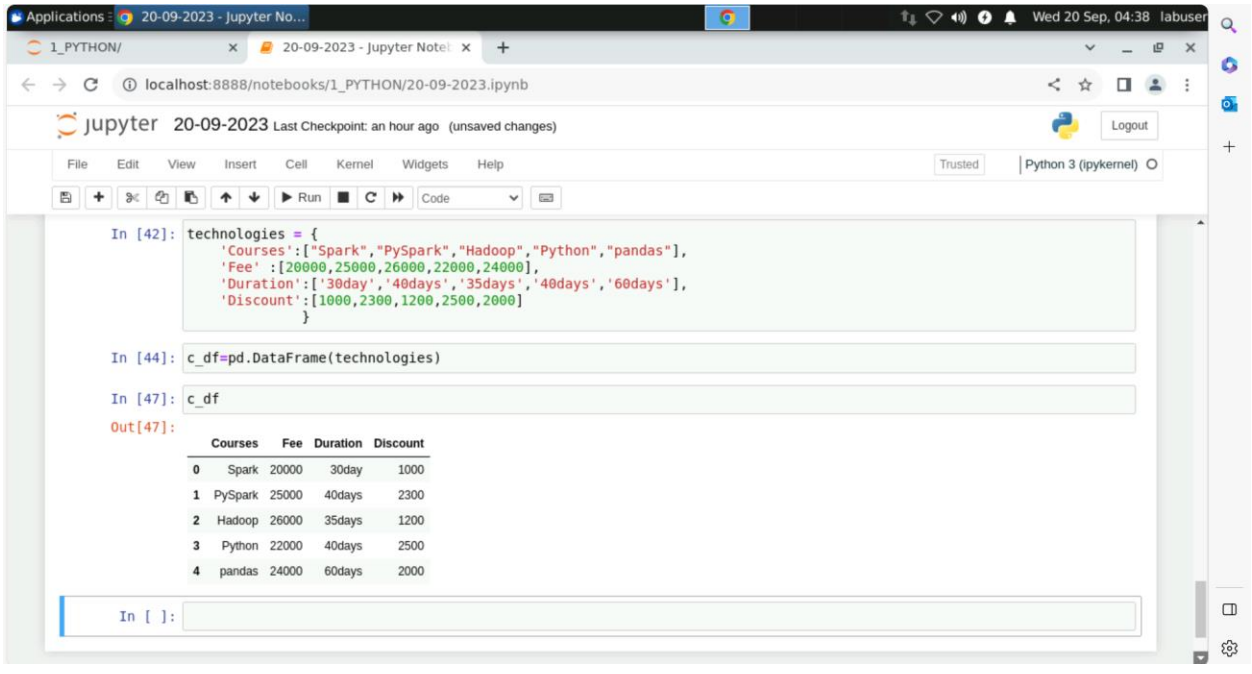

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:12
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 29 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
In [16]: df.dtypes
Out[16]: names object
ages int64
dtype: object
In [18]: df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3 entries, 0 to 2
Data columns (total 2 columns):
Column Non-Null Count Dtype
--- ------ -------------- -----
0 names 3 non-null object
1 ages 3 non-null int64
dtypes: int64(1), object(1)
memory usage: 180.0+ bytes
In [ ]:

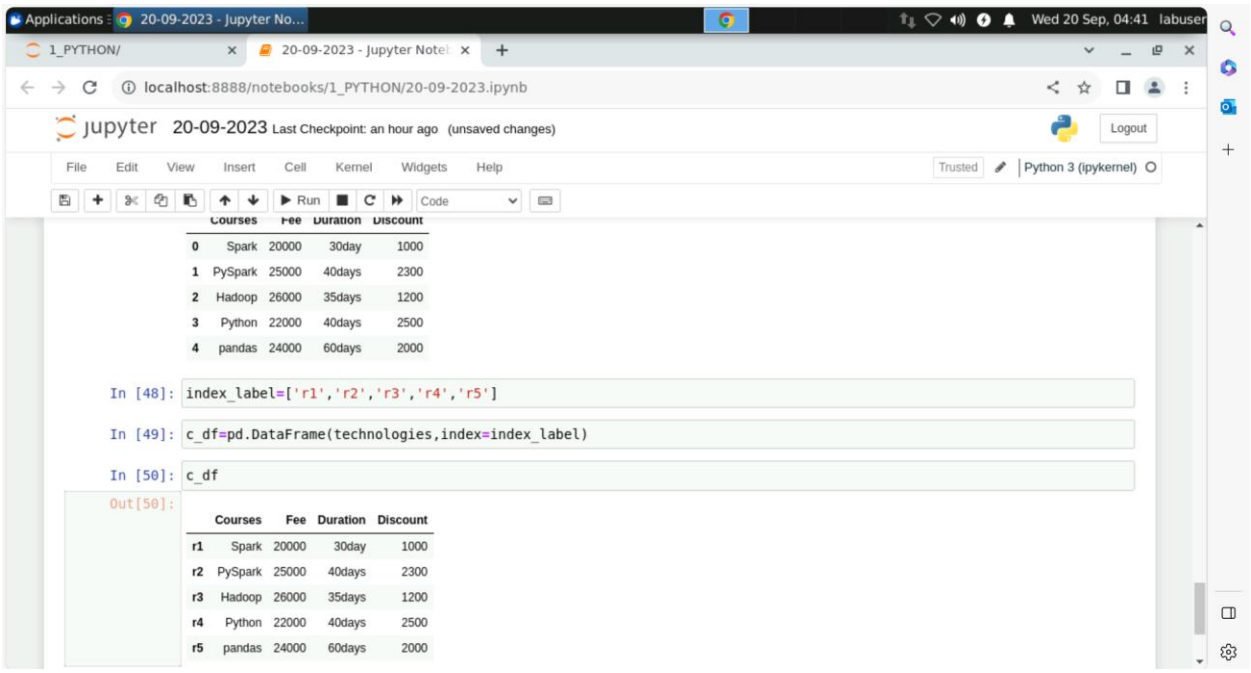

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:20
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 37 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
memory usage: 180.0+ bytes
In [19]: df.columns=["emp_name","age"]
In [20]: df
Out[20]:
emp_name age
0 a 25
1 b 26
2 c 21
In [22]: df2=df.rename(columns={"emp_name":"emp_first_name","age":"emp_age"})
In [23]: df2
Out[23]:
emp_first_name emp_age
0 a 25
1 b 26
2 c 21

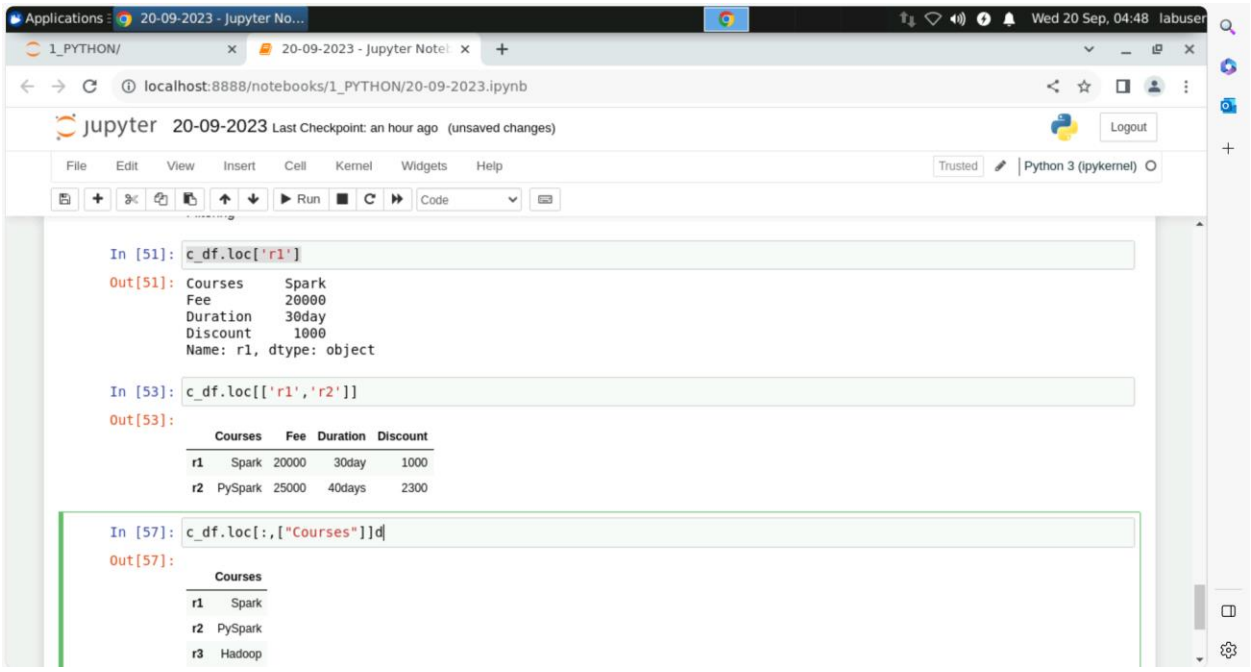

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:23
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 40 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
In [23]: df2
Out[23]:
emp_first_name emp_age
0 a 25
1 b 26
2 c 21
In [24]: dept=["Sales","Marketing","IT"]
In [25]: df["company"]=dept
In [26]: df
Out[26]:
emp_name age company
0 a 25 Sales
1 b 26 Marketing
2 c 21 IT
In [ ]:

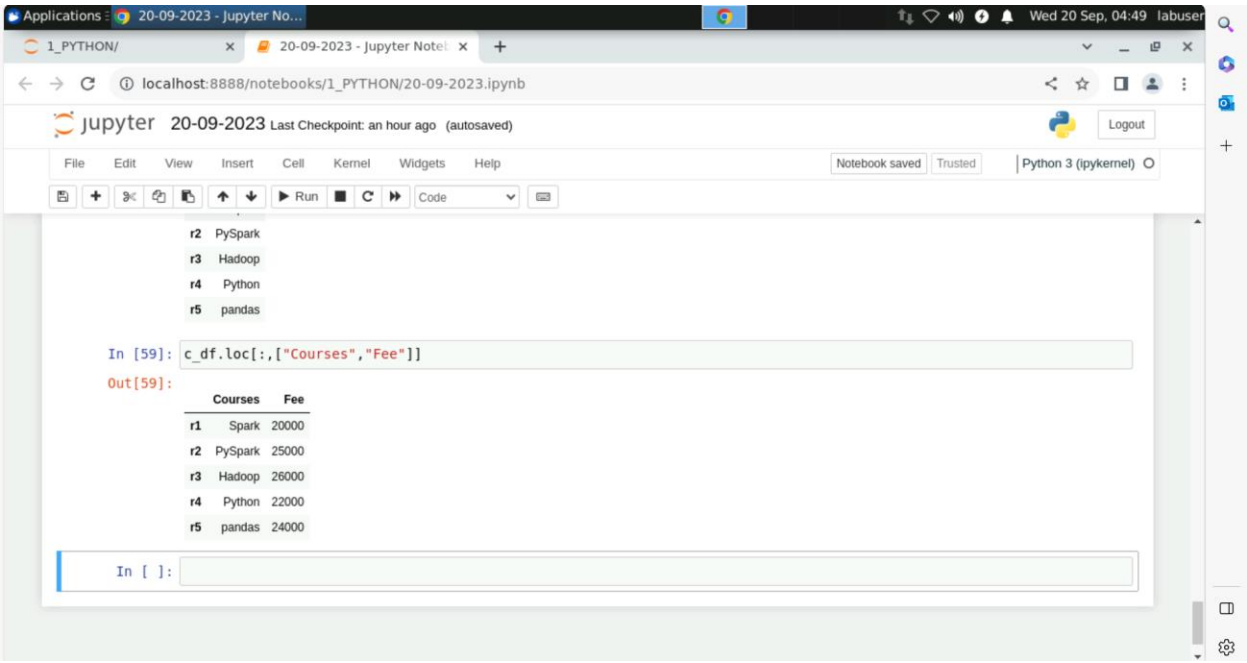

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:27
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 44 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
1 b 26 Marketing
2 c 21 IT
In [29]: df.dtypes
Out[29]: emp_name object
age int64
company object
dtype: object
In [32]: df3=df.convert_dtypes()
In [33]: df3.dtypes
Out[33]: emp_name string
age Int64
company string
dtype: object
In [ ]:

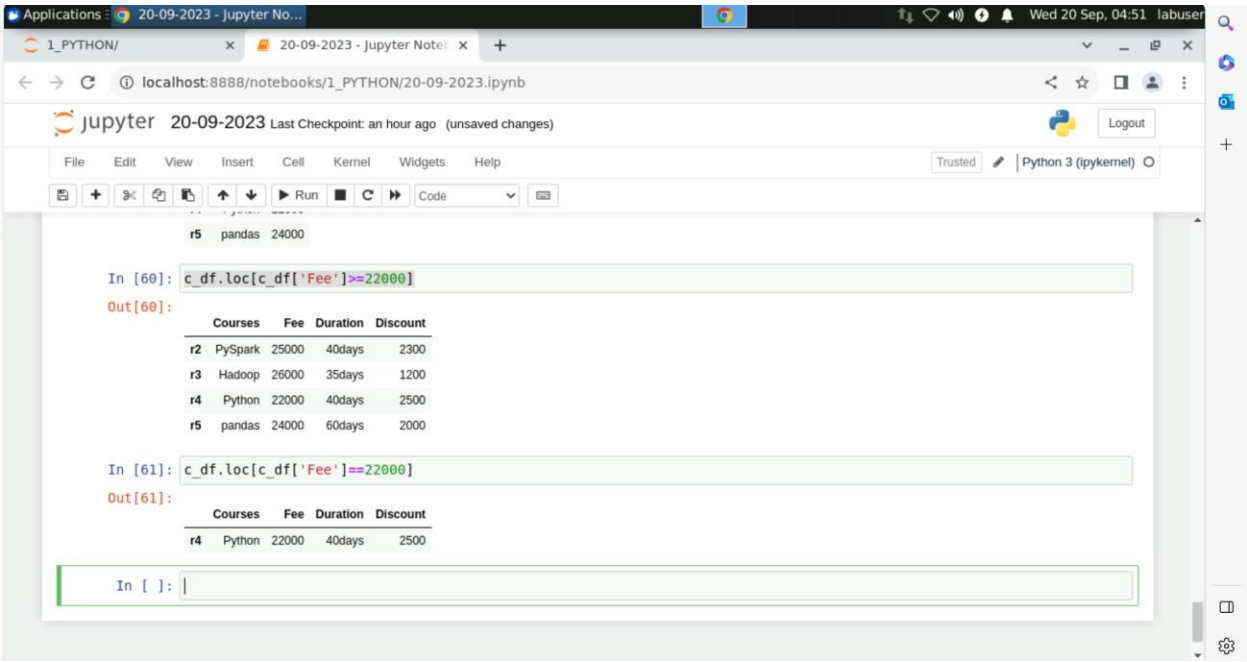

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:35 labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: an hour ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
Out[37]:
emp_name object
age float64
company object
dtype: object
In [39]: df4.shape
Out[39]: (3, 3)
In [40]: df4.shape[0]
Out[40]: 3
In [41]: df4.shape[1]
Out[41]: 3
In [ ]:

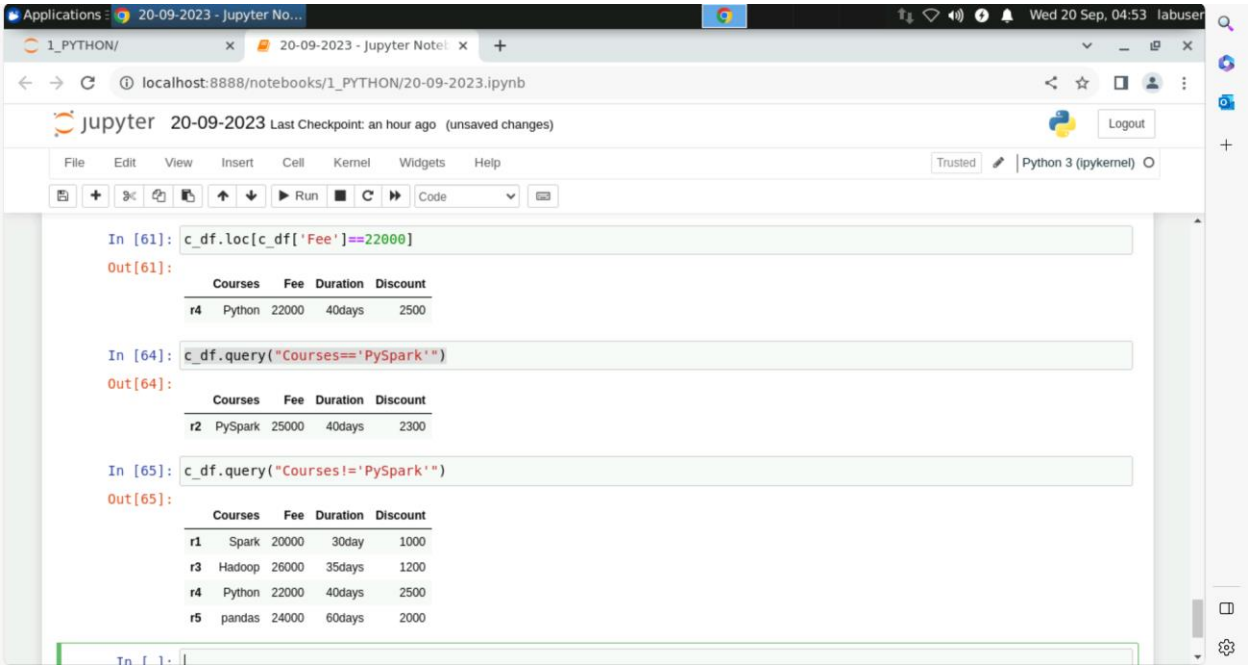

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:38 labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: an hour ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
In [42]:
technologies = {
'Courses':["Spark","PySpark","Hadoop","Python","pandas"],
'Fee' :[20000,25000,26000,22000,24000],
'Duration':['30day','40days','35days','40days','60days'],
'Discount':[1000,2300,1200,2500,2000]
}
In [44]: c_df=pd.DataFrame(technologies)
In [47]: c_df
Out[47]:
Courses Fee Duration Discount
0 Spark 20000 30day 1000
1 PySpark 25000 40days 2300
2 Hadoop 26000 35days 1200
3 Python 22000 40days 2500
4 pandas 24000 60days 2000
In [ ]:


Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:41
labuser
1_PYTHON/
20-09-2023 - Jupyter Noteb
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: an hour ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
In [48]: index_label=['r1','r2','r3','r4','r5']
In [49]: c_df=pd.DataFrame(technologies,index=index_label)
In [50]: c_df
Out[50]:
Courses Fee Duration Discount
r1 Spark 20000 30day 1000
r2 PySpark 25000 40days 2300
r3 Hadoop 26000 35days 1200
r4 Python 22000 40days 2500
r5 pandas 24000 60days 2000

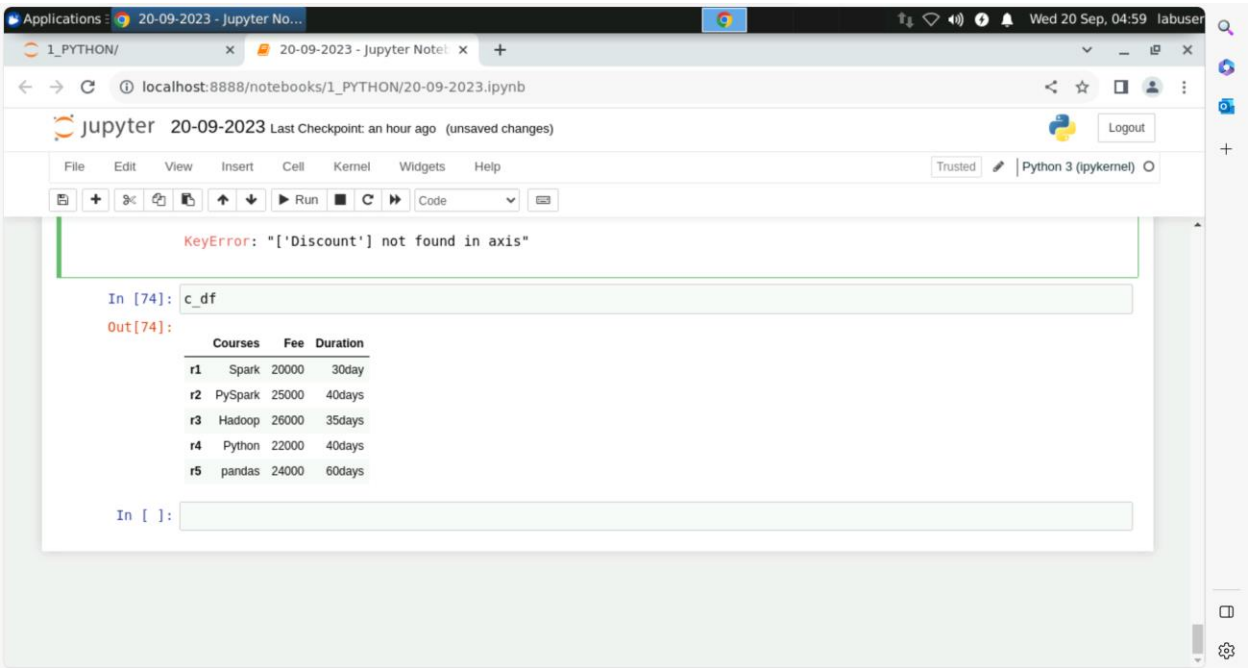

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:48
labuser
1_PYTHON/
20-09-2023 - Jupyter Noteb
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: an hour ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
In [51]: c_df.loc['r1']
Out[51]:
Courses Spark
Fee 20000
Duration 30day
Discount 1000
Name: r1, dtype: object
In [53]: c_df.loc[['r1','r2']]
Out[53]:
Courses Fee Duration Discount
r1 Spark 20000 30day 1000
r2 PySpark 25000 40days 2300
In [57]: c_df.loc[:,["Courses"]]d
Out[57]:
Courses
r1 Spark
r2 PySpark
r3 Hadoop

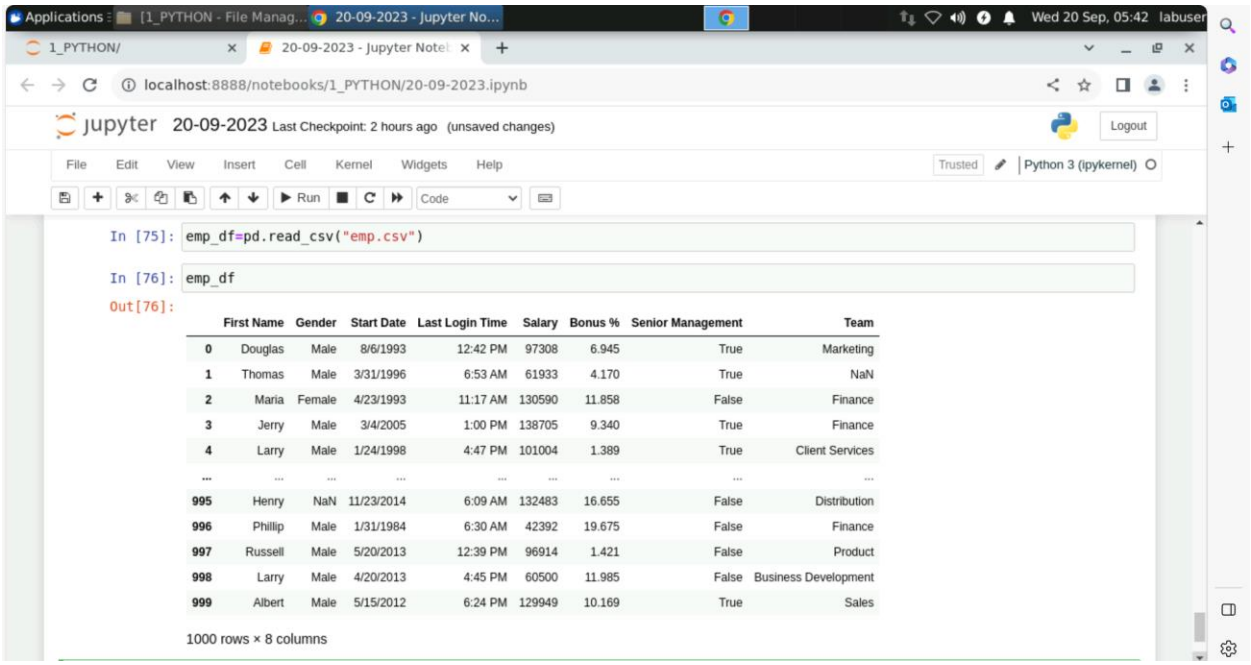

Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:49 labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: an hour ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Notebook saved
Trusted
Python 3 (ipykernel)
Run
Code
r2 PySpark
r3 Hadoop
r4 Python
r5 pandas
In [59]: c_df.loc[:,["Courses","Fee"]]
Out[59]:
Courses Fee
r1 Spark 20000
r2 PySpark 25000
r3 Hadoop 26000
r4 Python 22000
r5 pandas 24000
In [ ]:


Applications
20-09-2023 - Jupyter No...
Wed 20 Sep, 04:51 labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: an hour ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
r5 pandas 24000
In [60]: c_df.loc[c_df['Fee']>=22000]
Out[60]:
Courses Fee Duration Discount
r2 PySpark 25000 40days 2300
r3 Hadoop 26000 35days 1200
r4 Python 22000 40days 2500
r5 pandas 24000 60days 2000
In [61]: c_df.loc[c_df['Fee']==22000]
Out[61]:
Courses Fee Duration Discount
r4 Python 22000 40days 2500
In [ ]:

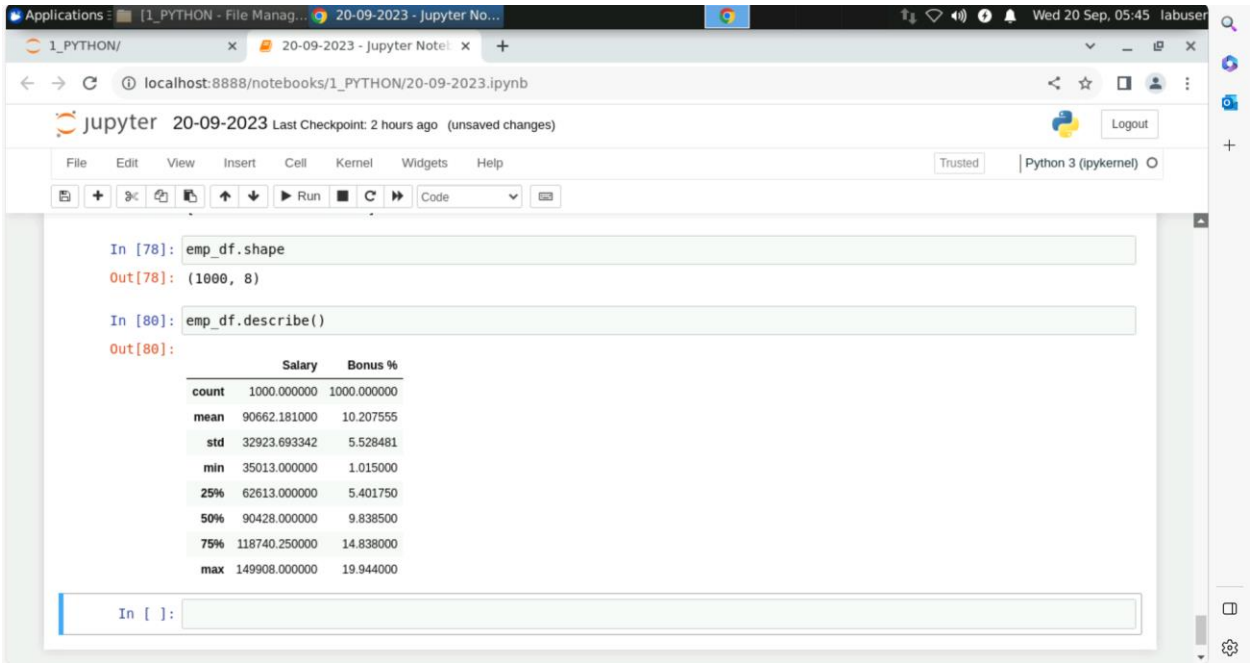


c_df=c_df.drop(['Discount'], axis=1) ----- TO DROP A COLUMN

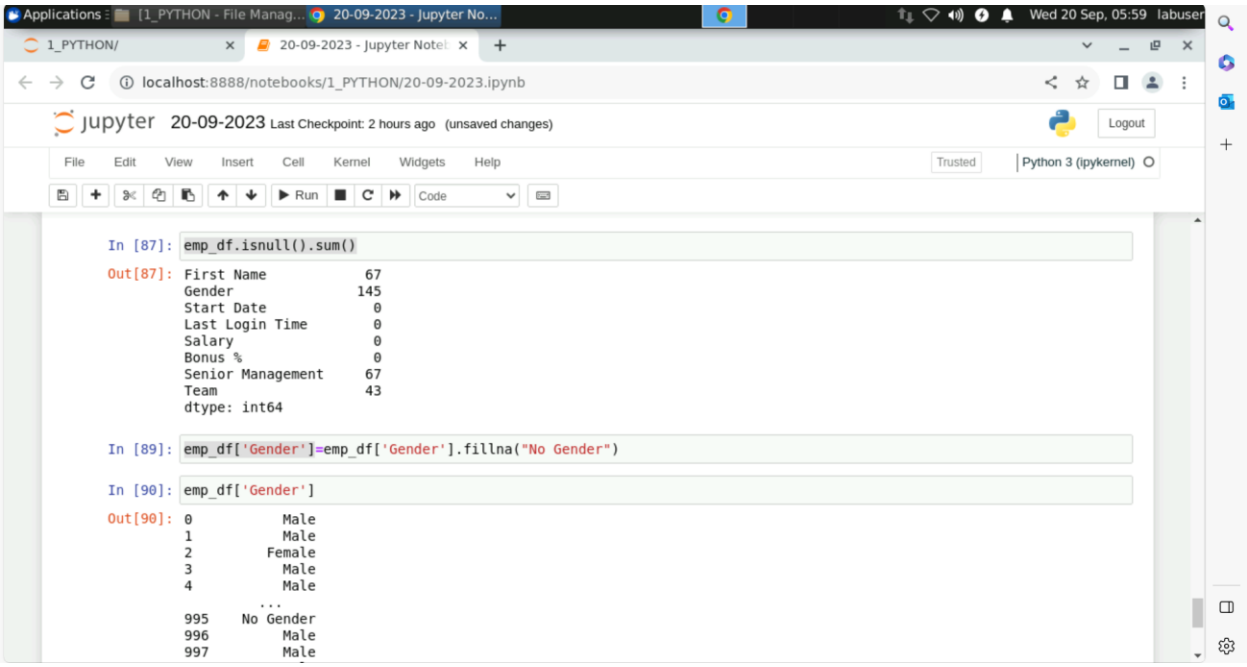


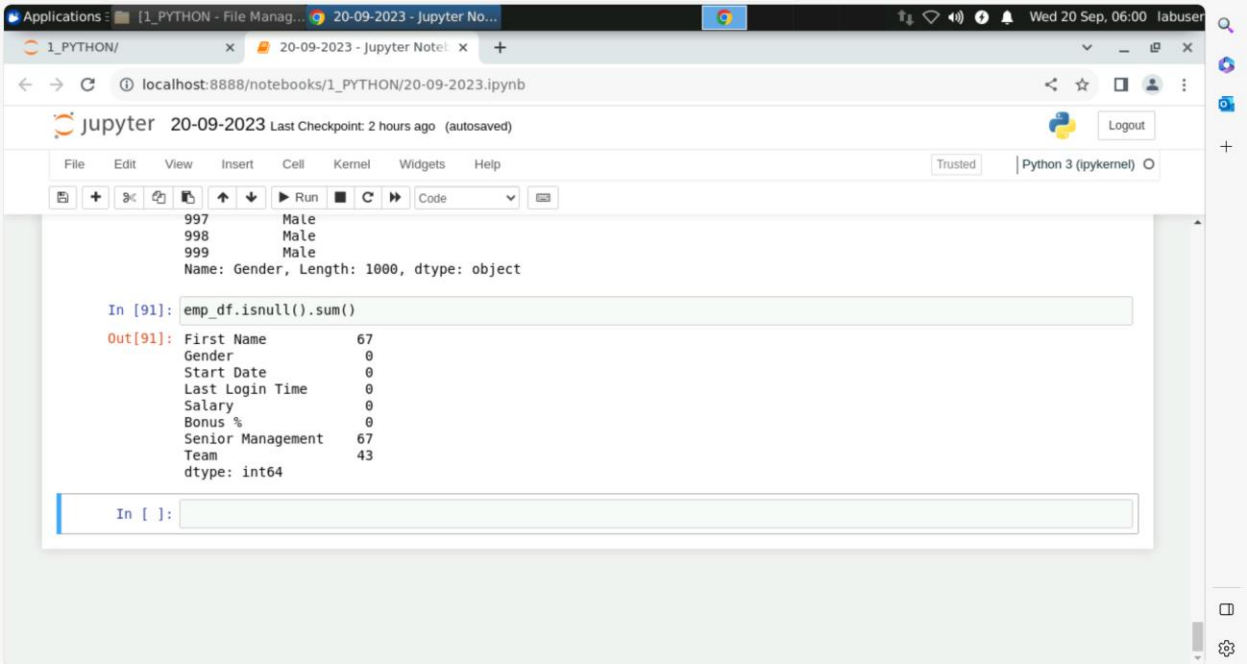


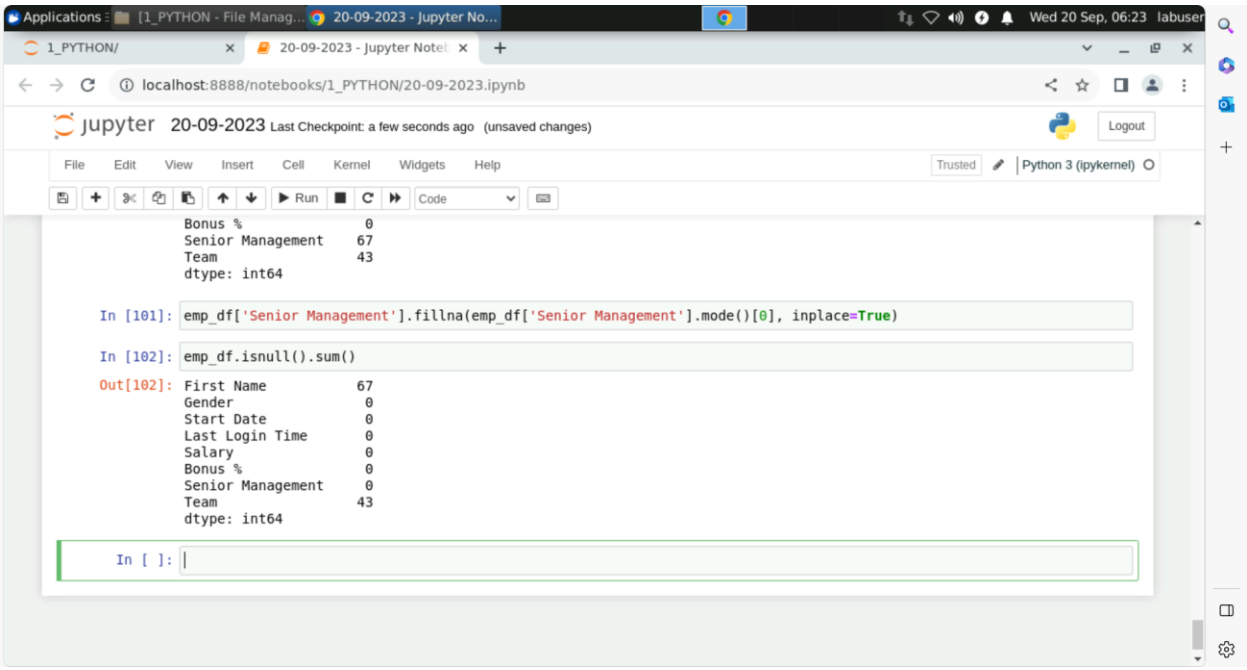
	First Name	Gender	Start Date	Last Login Time	Salary	Bonus %	Senior Management	Team
0	Douglas	Male	8/6/1993	12:42 PM	97308	6.945	True	Marketing
1	Thomas	Male	3/31/1996	6:53 AM	61933	4.170	True	NaN
2	Maria	Female	4/23/1993	11:17 AM	130590	11.858	False	Finance
3	Jerry	Male	3/4/2005	1:00 PM	138705	9.340	True	Finance
4	Larry	Male	1/24/1998	4:47 PM	101004	1.389	True	Client Services
...	...	...	...	...	...	...	...	...
995	Henry	NaN	11/23/2014	6:09 AM	132483	16.655	False	Distribution
996	Phillip	Male	1/31/1984	6:30 AM	42392	19.675	False	Finance
997	Russell	Male	5/20/2013	12:39 PM	96914	1.421	False	Product
998	Larry	Male	4/20/2013	4:45 PM	60500	11.985	False	Business Development
999	Albert	Male	5/15/2012	6:24 PM	129949	10.169	True	Sales

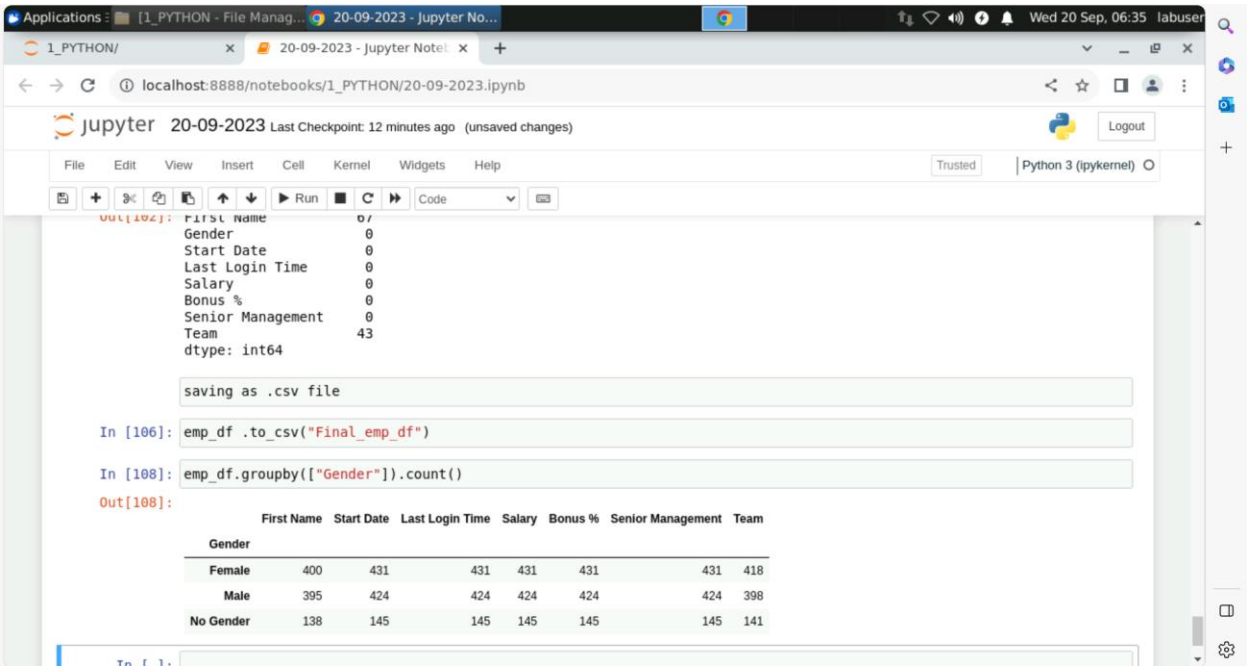


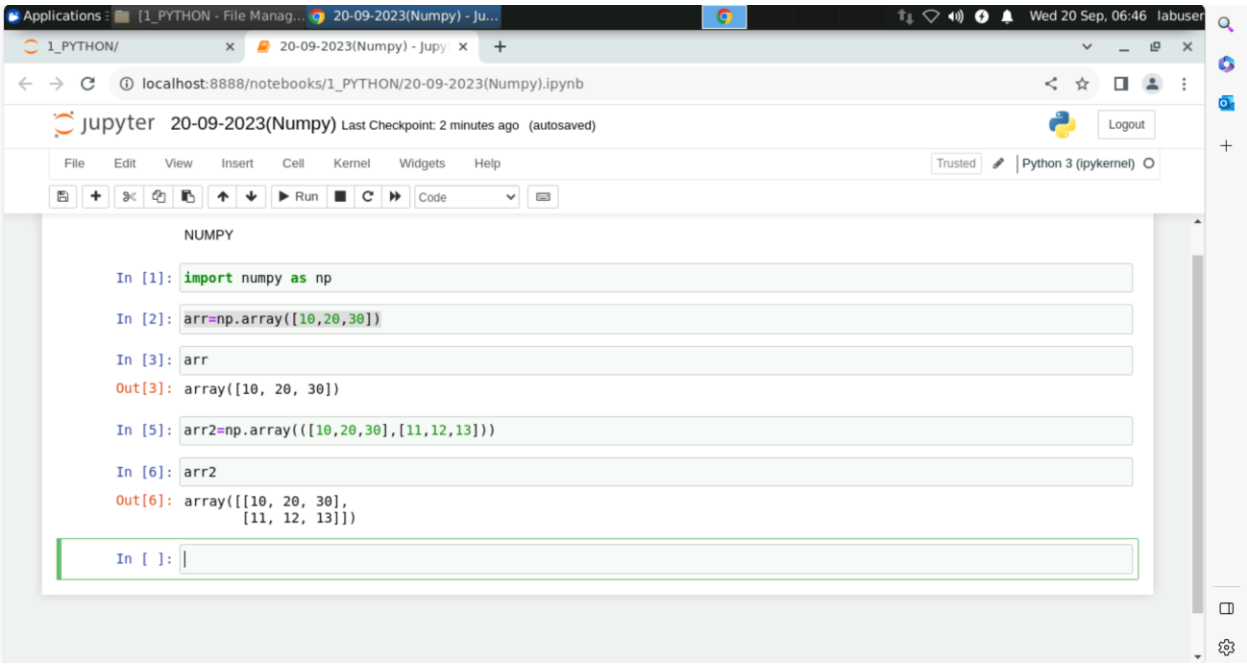
	Salary	Bonus %
count	1000.000000	1000.000000
mean	90662.181000	10.207555
std	32923.693342	5.528481
min	35013.000000	1.015000
25%	62613.000000	5.401750
50%	90428.000000	9.838500
75%	118740.250000	14.838000
max	149908.000000	19.944000

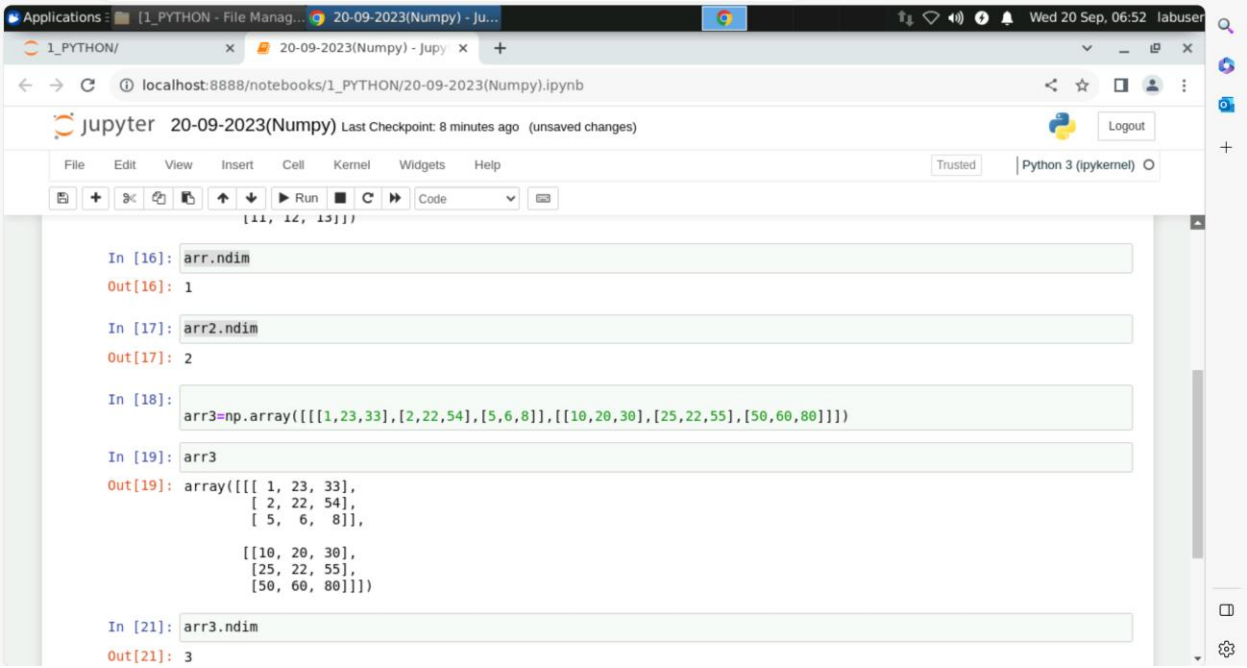

Applications
[1_PYTHON - File Manag...
20-09-2023 - Jupyter No...
Wed 20 Sep, 05:59
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 2 hours ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
In [87]: emp_df.isnull().sum()
Out[87]: First Name 67
Gender 145
Start Date 0
Last Login Time 0
Salary 0
Bonus % 0
Senior Management 67
Team 43
dtype: int64
In [89]: emp_df['Gender']=emp_df['Gender'].fillna("No Gender")
In [90]: emp_df['Gender']
Out[90]: 0 Male
1 Male
2 Female
3 Male
4 Male
...
995 No Gender
996 Male
997 Male

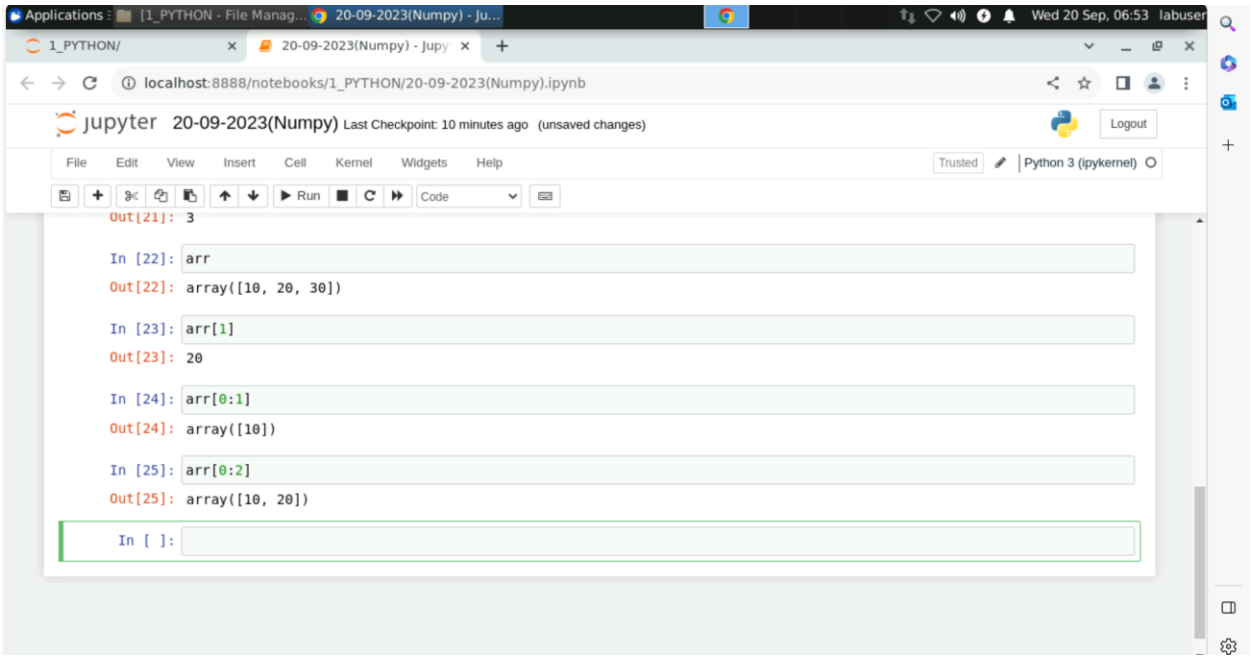

Applications
[1_PYTHON - File Manag...
20-09-2023 - Jupyter No...
Wed 20 Sep, 06:00
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 2 hours ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
997 Male
998 Male
999 Male
Name: Gender, Length: 1000, dtype: object
In [91]: emp_df.isnull().sum()
Out[91]: First Name 67
Gender 0
Start Date 0
Last Login Time 0
Salary 0
Bonus % 0
Senior Management 67
Team 43
dtype: int64
In [ ]:

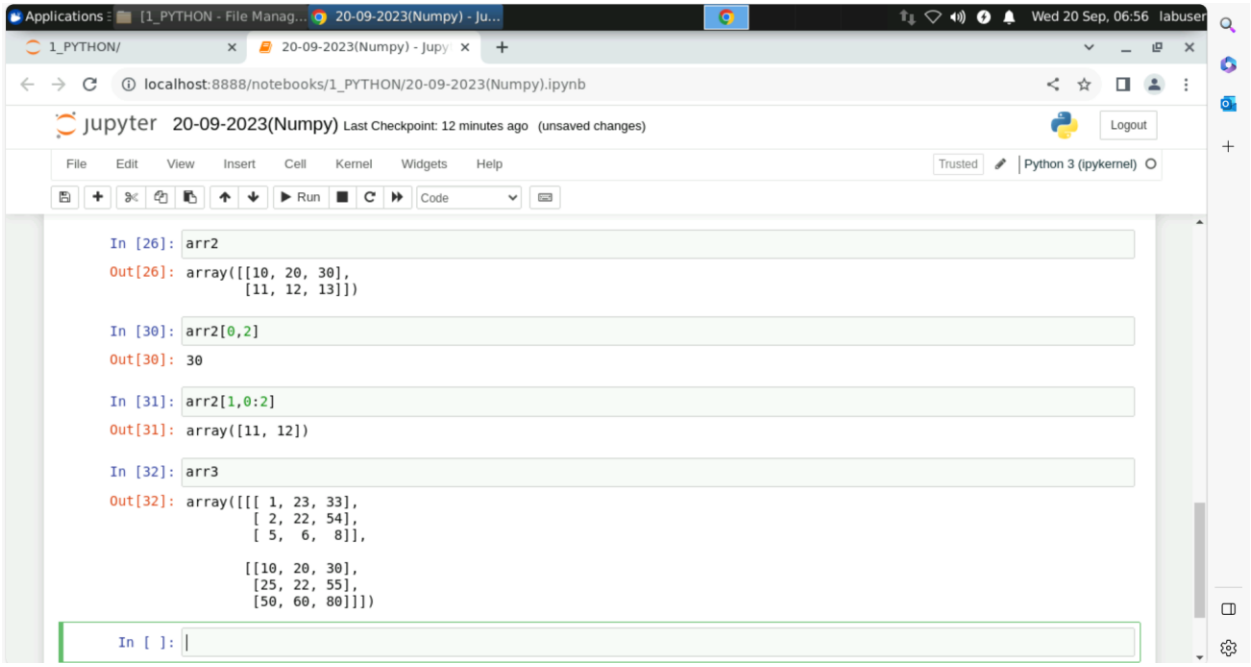

Applications
[1_PYTHON - File Manag...
20-09-2023 - Jupyter No...
Wed 20 Sep, 06:23
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: a few seconds ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
Bonus % 0
Senior Management 67
Team 43
dtype: int64
In [101]: emp_df['Senior Management'].fillna(emp_df['Senior Management'].mode()[0], inplace=True)
In [102]: emp_df.isnull().sum()
Out[102]: First Name 67
Gender 0
Start Date 0
Last Login Time 0
Salary 0
Bonus % 0
Senior Management 0
Team 43
dtype: int64
In [ ]:

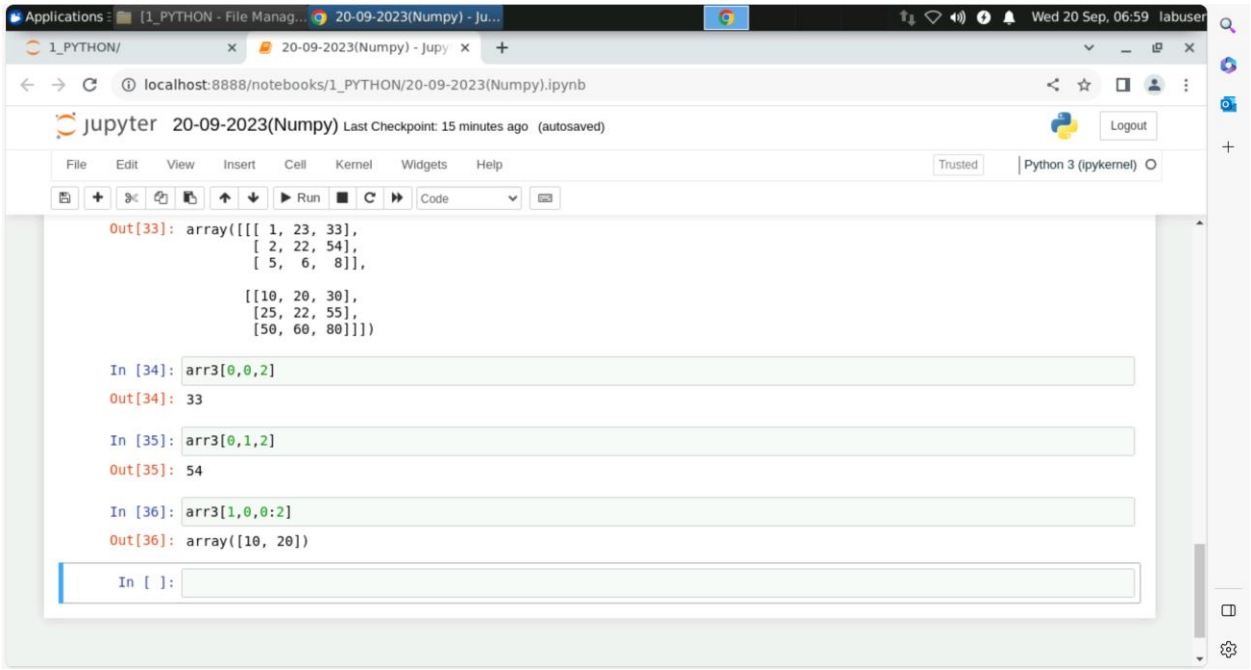

Applications
[1_PYTHON - File Manag...
20-09-2023 - Jupyter No...
Wed 20 Sep, 06:35
labuser
1_PYTHON/
20-09-2023 - Jupyter Note
localhost:8888/notebooks/1_PYTHON/20-09-2023.ipynb
jupyter 20-09-2023 Last Checkpoint: 12 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
Gender 0
Start Date 0
Last Login Time 0
Salary 0
Bonus % 0
Senior Management 0
Team 43
dtype: int64
saving as .csv file
In [106]: emp_df .to_csv("Final_emp_df")
In [108]: emp_df.groupby(["Gender"]).count()
Out[108]:
First Name Start Date Last Login Time Salary Bonus % Senior Management Team
Gender
Female 400 431 431 431 431 431 418
Male 395 424 424 424 424 424 398
No Gender 138 145 145 145 145 145 141

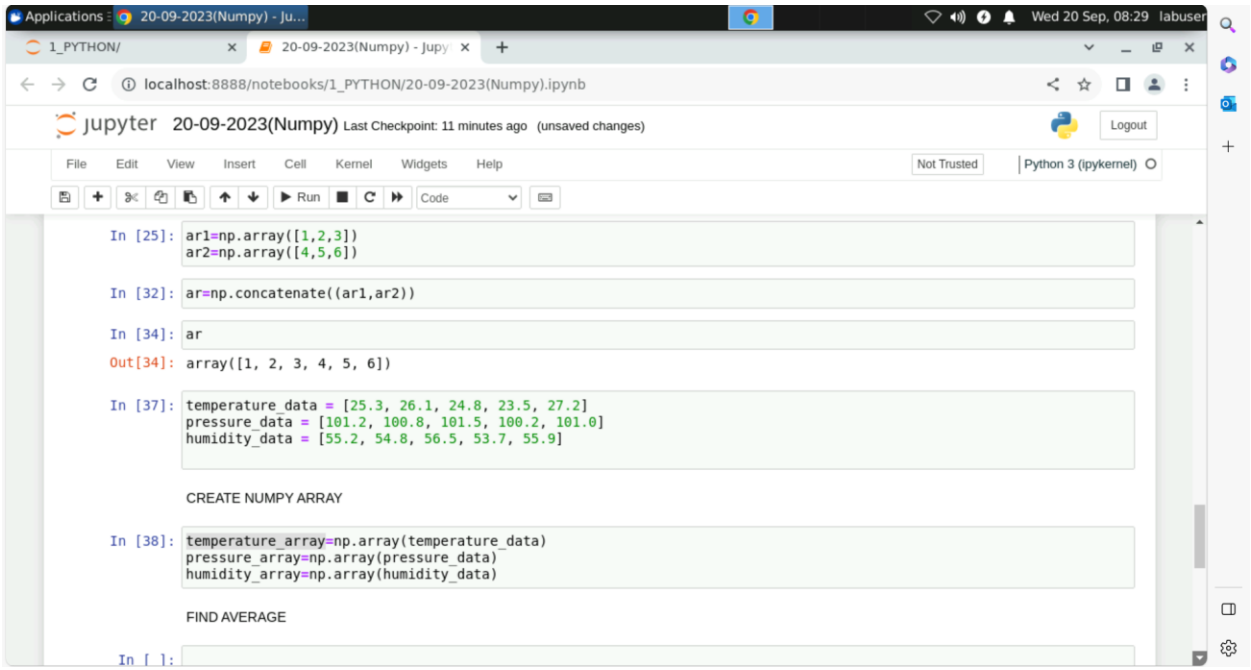

Applications
[1_PYTHON - File Manag...
20-09-2023(Numpy) - Ju...
Wed 20 Sep, 06:46
labuser
1_PYTHON/
20-09-2023(Numpy) - Jupy
localhost:8888/notebooks/1_PYTHON/20-09-2023(Numpy).ipynb
jupyter 20-09-2023(Numpy) Last Checkpoint: 2 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
NUMPY
In [1]: import numpy as np
In [2]: arr=np.array([10,20,30])
In [3]: arr
Out[3]: array([10, 20, 30])
In [5]: arr2=np.array(([10,20,30],[11,12,13]))
In [6]: arr2
Out[6]: array([[10, 20, 30],
[11, 12, 13]])
In [ ]:

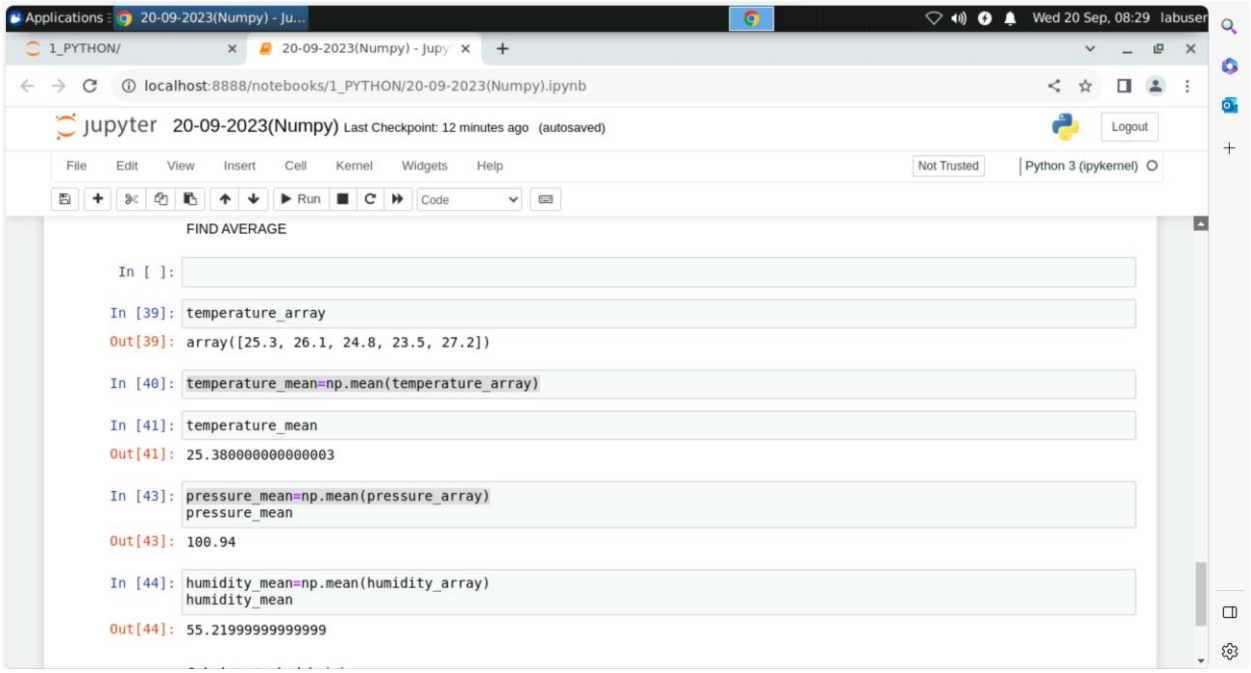

Applications
[1_PYTHON - File Manag...
20-09-2023(Numpy) - Ju...
Wed 20 Sep, 06:52
labuser
1_PYTHON/
20-09-2023(Numpy) - Jupy
localhost:8888/notebooks/1_PYTHON/20-09-2023(Numpy).ipynb
jupyter 20-09-2023(Numpy) Last Checkpoint: 8 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
In [16]: arr.ndim
Out[16]: 1
In [17]: arr2.ndim
Out[17]: 2
In [18]: arr3=np.array([[[1,23,33],[2,22,54],[5,6,8]],[[10,20,30],[25,22,55],[50,60,80]]])
In [19]: arr3
Out[19]: array([[[1, 23, 33],
[2, 22, 54],
[5, 6, 8]],
[[10, 20, 30],
[25, 22, 55],
[50, 60, 80]]])
In [21]: arr3.ndim
Out[21]: 3

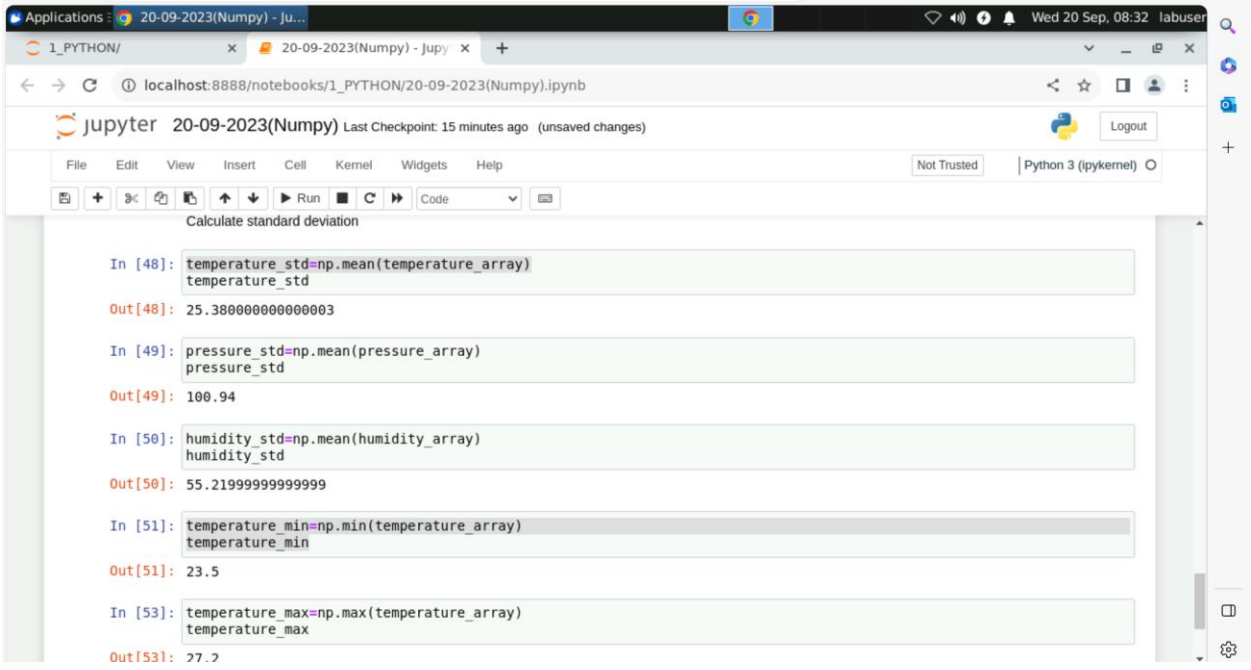

Applications
[1_PYTHON - File Manag...
20-09-2023(Numpy) - Ju...
Wed 20 Sep, 06:53
labuser
1_PYTHON/
20-09-2023(Numpy) - Jupy
localhost:8888/notebooks/1_PYTHON/20-09-2023(Numpy).ipynb
jupyter 20-09-2023(Numpy) Last Checkpoint: 10 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
Out[21]: 3
In [22]: arr
Out[22]: array([10, 20, 30])
In [23]: arr[1]
Out[23]: 20
In [24]: arr[0:1]
Out[24]: array([10])
In [25]: arr[0:2]
Out[25]: array([10, 20])
In [ ]:


Applications
[1_PYTHON - File Manag...
20-09-2023(Numpy) - Ju...
Wed 20 Sep, 06:56
labuser
1_PYTHON/
20-09-2023(Numpy) - Jupy
localhost:8888/notebooks/1_PYTHON/20-09-2023(Numpy).ipynb
jupyter 20-09-2023(Numpy) Last Checkpoint: 12 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
In [26]: arr2
Out[26]: array([[10, 20, 30],
[11, 12, 13]])
In [30]: arr2[0,2]
Out[30]: 30
In [31]: arr2[1,0:2]
Out[31]: array([11, 12])
In [32]: arr3
Out[32]: array([[[1, 23, 33],
[2, 22, 54],
[5, 6, 8]],
[[10, 20, 30],
[25, 22, 55],
[50, 60, 80]]])
In [ ]:

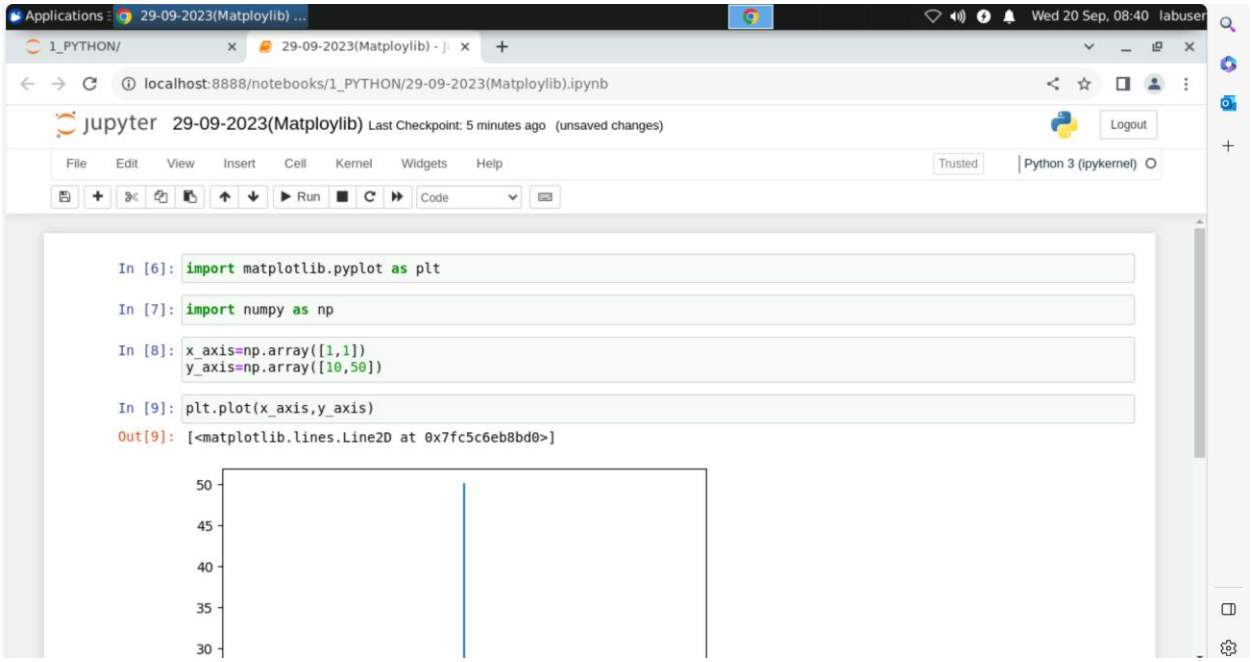

Applications
[1_PYTHON - File Manag...
20-09-2023(Numpy) - Ju...
Wed 20 Sep, 06:59
labuser
1_PYTHON/
20-09-2023(Numpy) - Jupy
localhost:8888/notebooks/1_PYTHON/20-09-2023(Numpy).ipynb
jupyter 20-09-2023(Numpy) Last Checkpoint: 15 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Run
Code
Out[33]: array([[[1, 23, 33],
[2, 22, 54],
[5, 6, 8]],
[[10, 20, 30],
[25, 22, 55],
[50, 60, 80]]])
In [34]: arr3[0,0,2]
Out[34]: 33
In [35]: arr3[0,1,2]
Out[35]: 54
In [36]: arr3[1,0,0:2]
Out[36]: array([10, 20])
In [ ]:

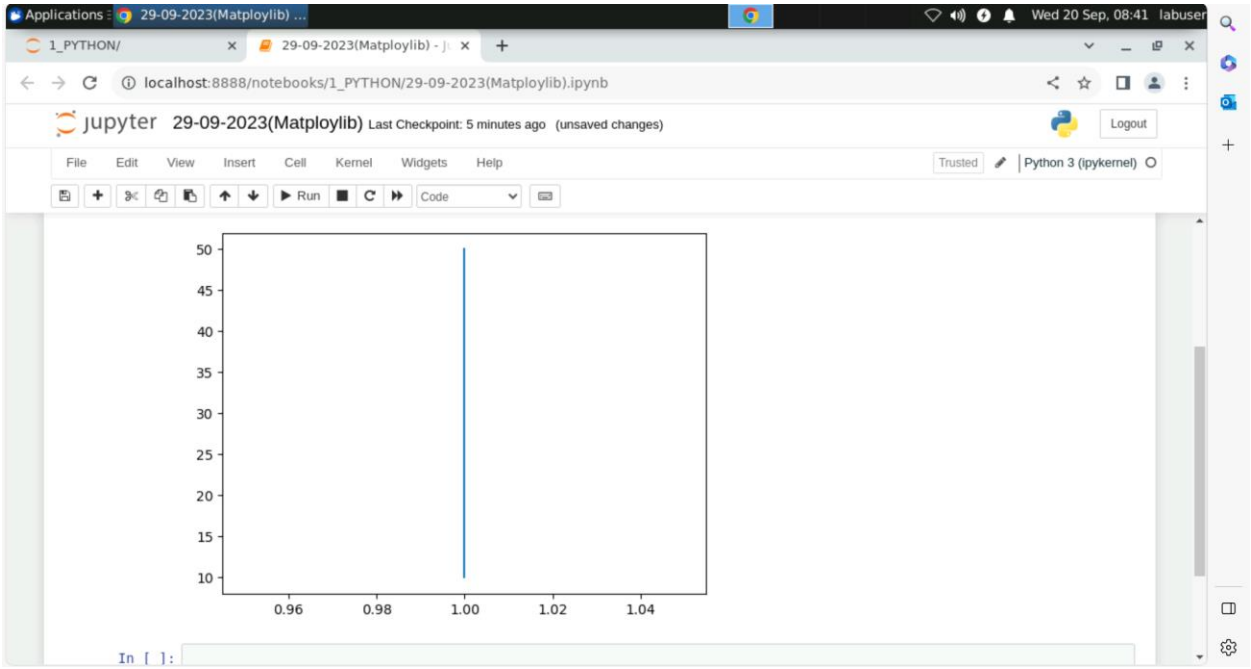

Applications
20-09-2023(Numpy) - Ju...
Wed 20 Sep, 08:29
labuser
1_PYTHON/
20-09-2023(Numpy) - Jupy
localhost:8888/notebooks/1_PYTHON/20-09-2023(Numpy).ipynb
jupyter 20-09-2023(Numpy) Last Checkpoint: 11 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
In [25]: ar1=np.array([1,2,3])
ar2=np.array([4,5,6])
In [32]: ar=np.concatenate((ar1,ar2))
In [34]: ar
Out[34]: array([1, 2, 3, 4, 5, 6])
In [37]: temperature_data = [25.3, 26.1, 24.8, 23.5, 27.2]
pressure_data = [101.2, 100.8, 101.5, 100.2, 101.0]
humidity_data = [55.2, 54.8, 56.5, 53.7, 55.9]
CREATE NUMPY ARRAY
In [38]: temperature_array=np.array(temperature_data)
pressure_array=np.array(pressure_data)
humidity_array=np.array(humidity_data)
FIND AVERAGE
In [ ]:


Applications
20-09-2023(Numpy) - Ju...
Wed 20 Sep, 08:29
labuser
1_PYTHON/
20-09-2023(Numpy) - Jupy
localhost:8888/notebooks/1_PYTHON/20-09-2023(Numpy).ipynb
jupyter 20-09-2023(Numpy) Last Checkpoint: 12 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
FIND AVERAGE
In [ ]:
In [39]: temperature_array
Out[39]: array([25.3, 26.1, 24.8, 23.5, 27.2])
In [40]: temperature_mean=np.mean(temperature_array)
In [41]: temperature_mean
Out[41]: 25.380000000000003
In [43]: pressure_mean=np.mean(pressure_array)
pressure_mean
Out[43]: 100.94
In [44]: humidity_mean=np.mean(humidity_array)
humidity_mean
Out[44]: 55.21999999999999


Applications
20-09-2023(Numpy) - Ju...
Wed 20 Sep, 08:32
labuser
1_PYTHON/
20-09-2023(Numpy) - Jupy
localhost:8888/notebooks/1_PYTHON/20-09-2023(Numpy).ipynb
jupyter 20-09-2023(Numpy) Last Checkpoint: 15 minutes ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Calculate standard deviation
In [48]: temperature_std=np.mean(temperature_array)
temperature_std
Out[48]: 25.380000000000003
In [49]: pressure_std=np.mean(pressure_array)
pressure_std
Out[49]: 100.94
In [50]: humidity_std=np.mean(humidity_array)
humidity_std
Out[50]: 55.21999999999999
In [51]: temperature_min=np.min(temperature_array)
temperature_min
Out[51]: 23.5
In [53]: temperature_max=np.max(temperature_array)
temperature_max
Out[53]: 27.2

https://numpy.org/devdocs/reference/index.html





Applications 29-09-2023(Matplotlib) ... Wed 20 Sep, 08:45 labuser

1_PYTHON/ | 29-09-2023(Matplotlib) - J

localhost:8888/notebooks/1_PYTHON/29-09-2023(Matplotlib).ipynb

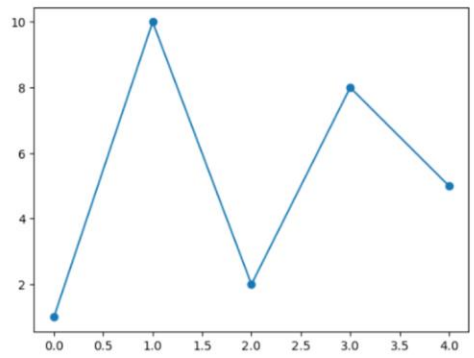
jupyter 29-09-2023(Matplotlib) Last Checkpoint: 10 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

Run Code

In [13]:
```
y_point=np.array([1,10,2,8,5])
plt.plot(y_point,marker="o")
```

Out[13]: [<matplotlib.lines.Line2D at 0x7fc58d1c2ed0>]



Applications 29-09-2023(Matplotlib) ... Wed 20 Sep, 08:47 labuser

1_PYTHON/ | 29-09-2023(Matplotlib) - J

localhost:8888/notebooks/1_PYTHON/29-09-2023(Matplotlib).ipynb

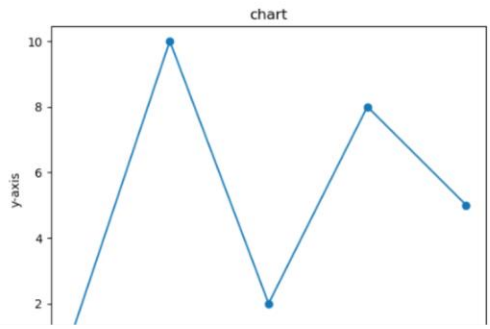
jupyter 29-09-2023(Matplotlib) Last Checkpoint: 12 minutes ago (unsaved changes) Logout

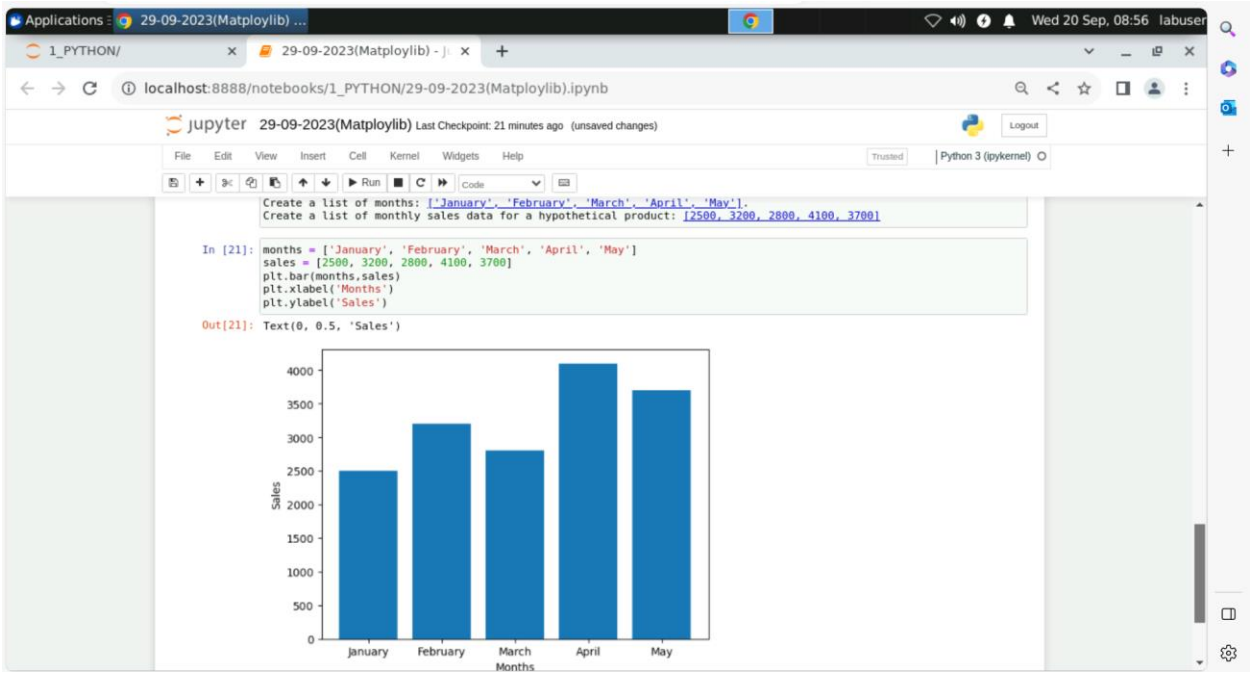
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

Run Code

In [17]:
```
y_point=np.array([1,10,2,8,5])
plt.plot(y_point,marker="o")
plt.xlabel("x-axis")
plt.ylabel("y-axis")
plt.title("chart")
```

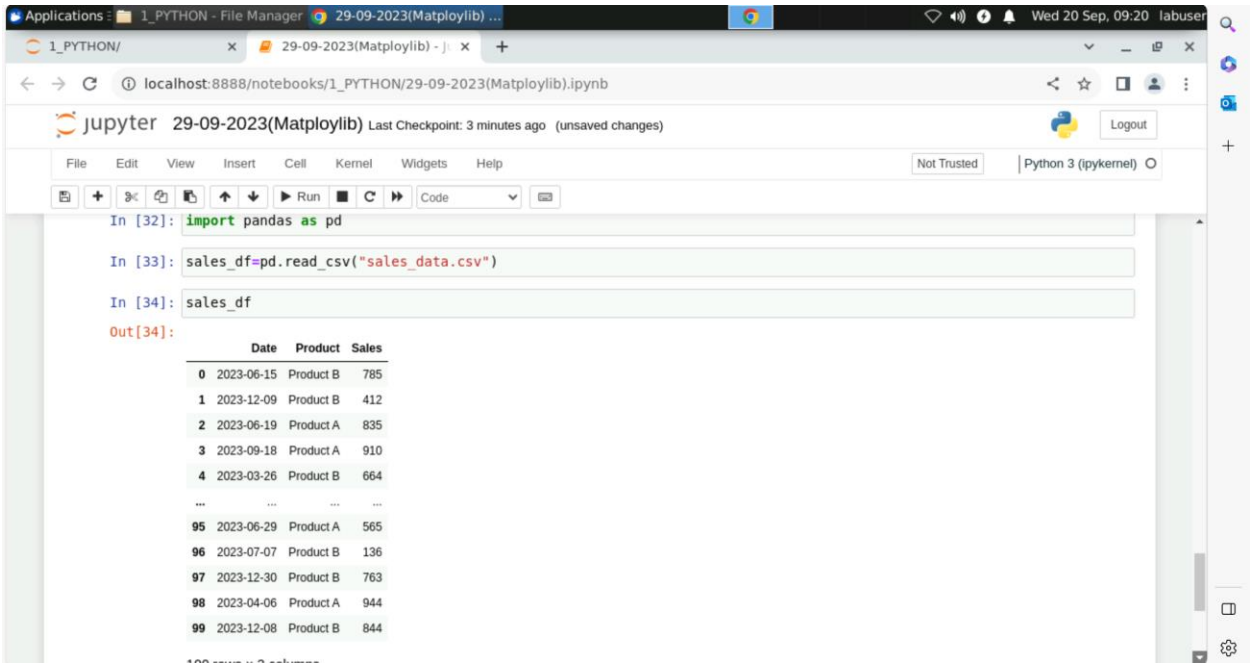
Out[17]: Text(0.5, 1.0, 'chart')

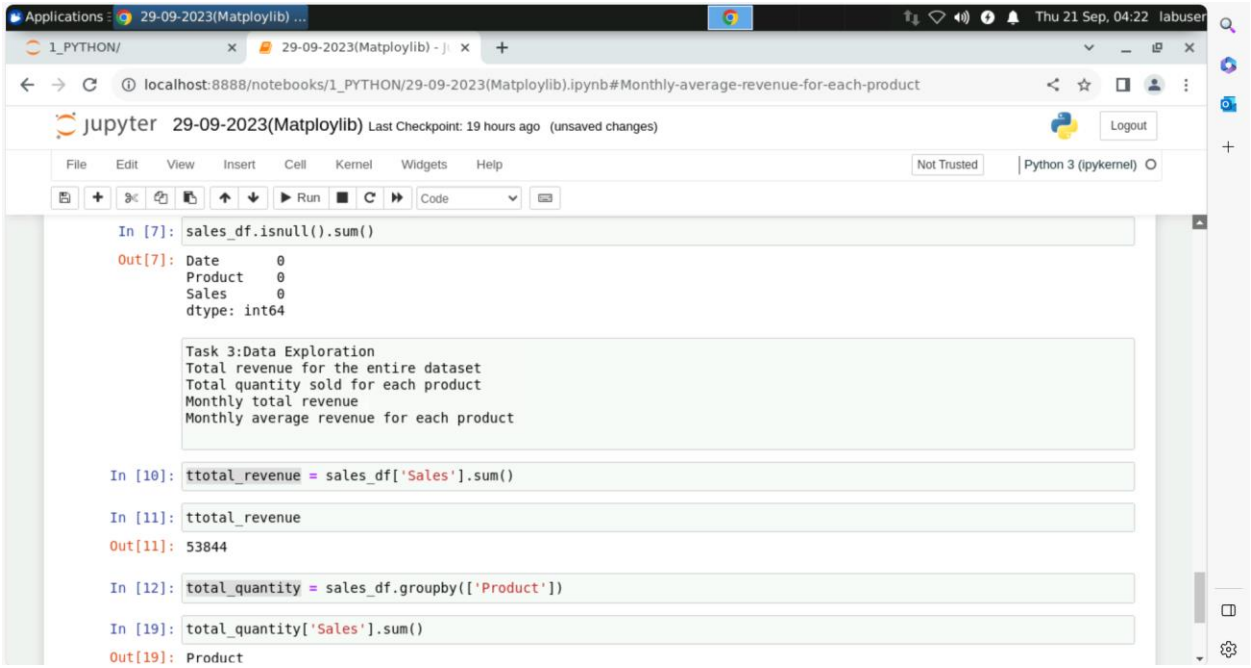


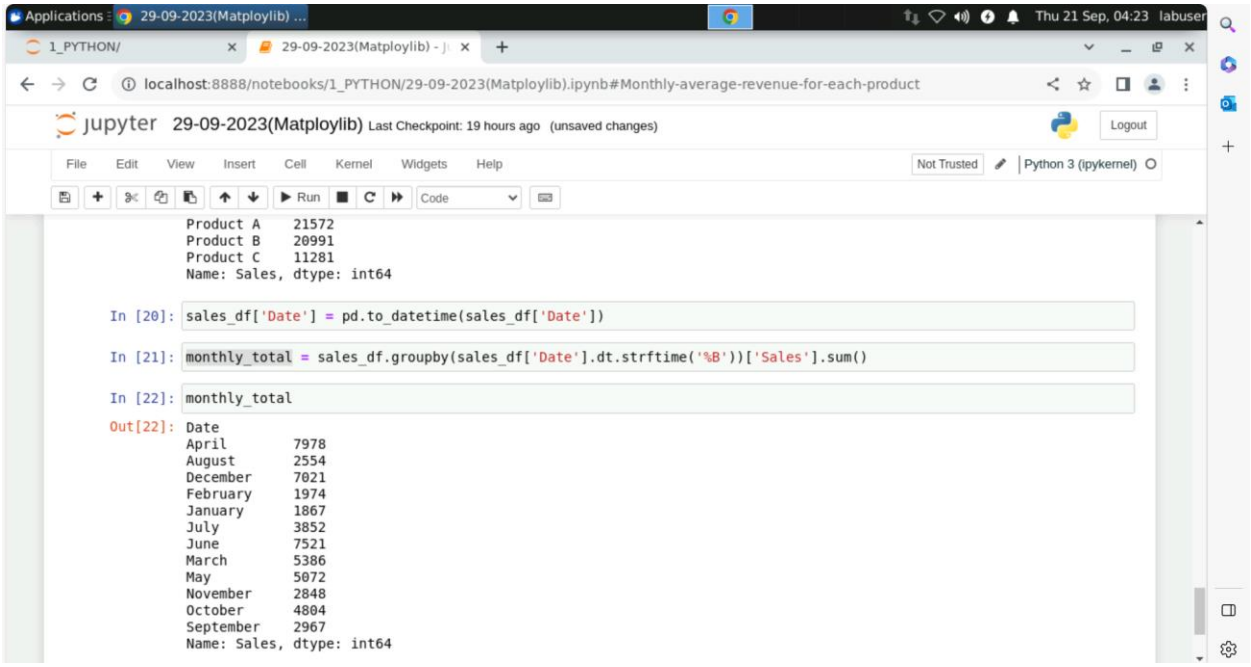


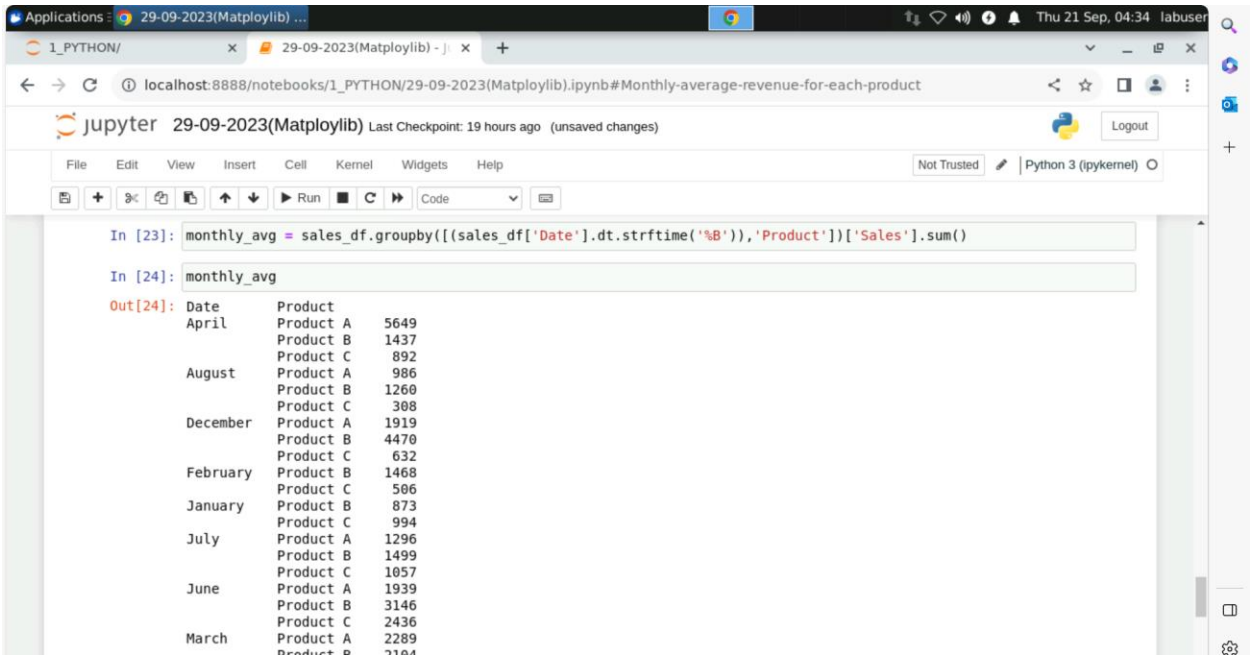
plt.savefig("Bargraph.png")

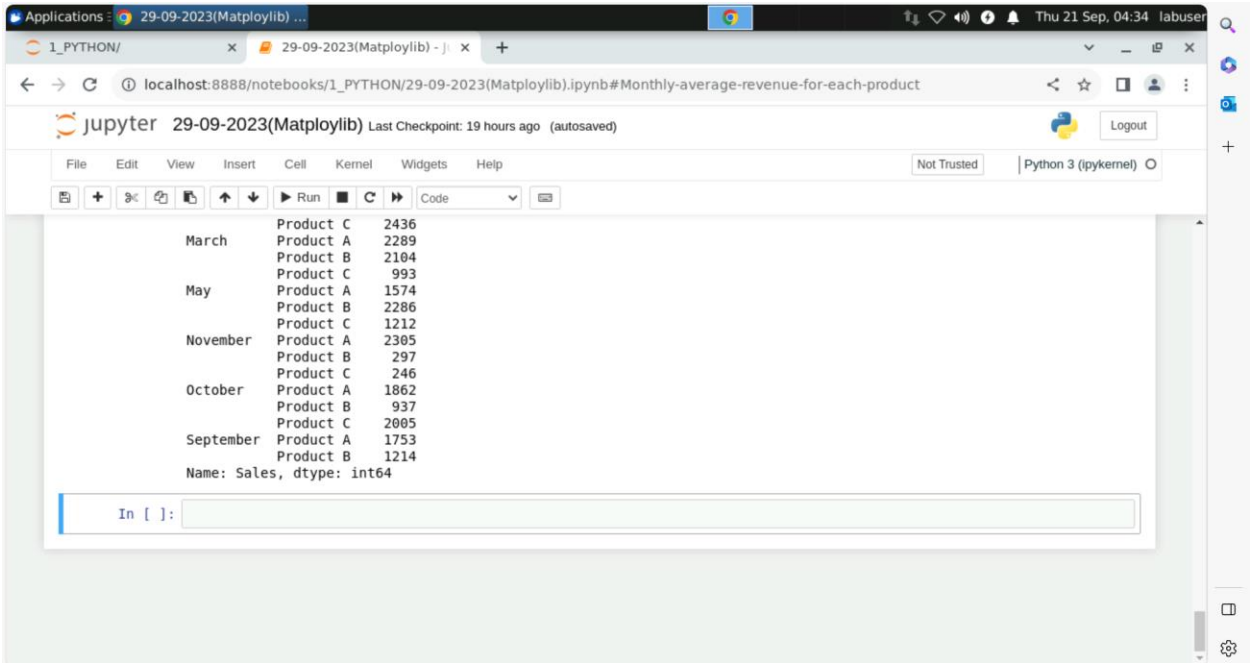
ACTIVITY:




Applications
29-09-2023(Matplotlib) ...
Thu 21 Sep, 04:22 labuser
1_PYTHON/
29-09-2023(Matplotlib) - J
localhost:8888/notebooks/1_PYTHON/29-09-2023(Matplotlib).ipynb#Monthly-average-revenue-for-each-product
jupyter 29-09-2023(Matplotlib) Last Checkpoint: 19 hours ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
In [7]: sales_df.isnull().sum()
Out[7]: Date 0
Product 0
Sales 0
dtype: int64
Task 3:Data Exploration
Total revenue for the entire dataset
Total quantity sold for each product
Monthly total revenue
Monthly average revenue for each product
In [10]: ttotal_revenue = sales_df['Sales'].sum()
In [11]: ttotal_revenue
Out[11]: 53844
In [12]: total_quantity = sales_df.groupby(['Product'])
In [19]: total_quantity['Sales'].sum()
Out[19]: Product


Applications
29-09-2023(Matplotlib) ...
Thu 21 Sep, 04:23 labuser
1_PYTHON/
29-09-2023(Matplotlib) - J
localhost:8888/notebooks/1_PYTHON/29-09-2023(Matplotlib).ipynb#Monthly-average-revenue-for-each-product
jupyter 29-09-2023(Matplotlib) Last Checkpoint: 19 hours ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Product A 21572
Product B 20991
Product C 11281
Name: Sales, dtype: int64
In [20]: sales_df['Date'] = pd.to_datetime(sales_df['Date'])
In [21]: monthly_total = sales_df.groupby(sales_df['Date'].dt.strftime('%B'))['Sales'].sum()
In [22]: monthly_total
Out[22]: Date
April 7978
August 2554
December 7021
February 1974
January 1867
July 3852
June 7521
March 5386
May 5072
November 2848
October 4804
September 2967
Name: Sales, dtype: int64


Applications
29-09-2023(Matplotlib) ...
Thu 21 Sep, 04:34 labuser
1_PYTHON/
29-09-2023(Matplotlib) - J
localhost:8888/notebooks/1_PYTHON/29-09-2023(Matplotlib).ipynb#Monthly-average-revenue-for-each-product
jupyter 29-09-2023(Matplotlib) Last Checkpoint: 19 hours ago (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
In [23]: monthly_avg = sales_df.groupby([(sales_df['Date'].dt.strftime('%B')),'Product'])['Sales'].sum()
In [24]: monthly_avg
Out[24]:
Date Product
April Product A 5649
Product B 1437
Product C 892
August Product A 986
Product B 1260
Product C 308
December Product A 1919
Product B 4470
Product C 632
February Product B 1468
Product C 506
January Product B 873
Product C 994
July Product A 1296
Product B 1499
Product C 1057
June Product A 1939
Product B 3146
Product C 2436
March Product A 2289


Applications
29-09-2023(Matplotlib) ...
Thu 21 Sep, 04:34 labuser
1_PYTHON/
29-09-2023(Matplotlib) - J
localhost:8888/notebooks/1_PYTHON/29-09-2023(Matplotlib).ipynb#Monthly-average-revenue-for-each-product
jupyter 29-09-2023(Matplotlib) Last Checkpoint: 19 hours ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Product C 2436
March Product A 2289
Product B 2104
Product C 993
May Product A 1574
Product B 2286
Product C 1212
November Product A 2305
Product B 297
Product C 246
October Product A 1862
Product B 937
Product C 2005
September Product A 1753
Product B 1214
Name: Sales, dtype: int64
In [ ]: