In [1]:
```
#Step:1 Install libraries

!pip install langchain
!pip install watermark
!pip install openai
```

```
Defaulting to user installation because normal site-packages is not writea
ble
Requirement already satisfied: langchain in c:\users\pavan 4288\appdata\ro
aming\python\python39\site-packages (0.1.12)
Requirement already satisfied: async-timeout<5.0.0,>=4.0.0 in c:\users\pav
an 4288\appdata\roaming\python\python39\site-packages (from langchain) (4.
0.3)
Requirement already satisfied: langsmith<0.2.0,>=0.1.17 in c:\users\pavan
4288\appdata\roaming\python\python39\site-packages (from langchain) (0.1.2
5)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in c:\users\pavan 4288
\appdata\roaming\python\python39\site-packages (from langchain) (1.33)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in c:\programdata\anacon
da3\lib\site-packages (from langchain) (1.4.39)
Requirement already satisfied: tenacity<9.0.0,>=8.1.0 in c:\users\pavan 42
88\appdata\roaming\python\python39\site-packages (from langchain) (8.2.3)
Requirement already satisfied: pydantic<3,>=1 in c:\users\pavan 4288\appda
ta\roaming\python\python39\site-packages (from langchain) (2.6.0)
Requirement already satisfied: numpy<2,>=1 in c:\users\pavan 4288\appdata
```

In [ ]:
```
#Import os & Ignore the warnings
```

In [2]:
```
import os
import warnings
warnings.filterwarnings("ignore")
```

In [3]:
```
pip install langchain_experimental
```

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: langchain_experimental in c:\users\pavan 4288
\appdata\roaming\python\python39\site-packages (0.0.54)
Requirement already satisfied: langchain-core<0.2.0,>=0.1.31 in c:\users\pava
n 4288\appdata\roaming\python\python39\site-packages (from langchain_experime
ntal) (0.1.31)
Requirement already satisfied: langchain<0.2.0,>=0.1.12 in c:\users\pavan 428
8\appdata\roaming\python\python39\site-packages (from langchain_experimental)
(0.1.12)
Requirement already satisfied: jsonpatch<2.0,>=1.33 in c:\users\pavan 4288\ap
pdata\roaming\python\python39\site-packages (from langchain<0.2.0,>=0.1.12->l
angchain_experimental) (1.33)
Requirement already satisfied: dataclasses-json<0.7,>=0.5.7 in c:\users\pavan
4288\appdata\roaming\python\python39\site-packages (from langchain<0.2.0,>=0.
1.12->langchain_experimental) (0.6.4)
Requirement already satisfied: langsmith<0.2.0,>=0.1.17 in c:\users\pavan 428
8\appdata\roaming\python\python39\site-packages (from langchain<0.2.0,>=0.1.1
2->langchain_experimental) (0.1.25)
Requirement already satisfied: langchain-text-splitters<0.1,>=0.0.1 in c:\use
rs\pavan 4288\appdata\roaming\python\python39\site-packages (from langchain<
0.2.0,>=0.1.12->langchain_experimental) (0.0.1)
Requirement already satisfied: numpy<2,>=1 in c:\users\pavan 4288\appdata\roa
ming\python\python39\site-packages (from langchain<0.2.0,>=0.1.12->langchain_
experimental) (1.24.4)
Requirement already satisfied: requests<3,>=2 in c:\programdata\anaconda3\lib
\site-packages (from langchain<0.2.0,>=0.1.12->langchain_experimental) (2.28.
1)
Requirement already satisfied: pydantic<3,>=1 in c:\users\pavan 4288\appdata
\roaming\python\python39\site-packages (from langchain<0.2.0,>=0.1.12->langch
ain_experimental) (2.6.0)
Requirement already satisfied: langchain-community<0.1,>=0.0.28 in c:\users\p
avan 4288\appdata\roaming\python\python39\site-packages (from langchain<0.2.
0,>=0.1.12->langchain_experimental) (0.0.28)
Requirement already satisfied: PyYAML>=5.3 in c:\programdata\anaconda3\lib\si
te-packages (from langchain<0.2.0,>=0.1.12->langchain_experimental) (6.0)
Requirement already satisfied: async-timeout<5.0.0,>=4.0.0 in c:\users\pavan
4288\appdata\roaming\python\python39\site-packages (from langchain<0.2.0,>=0.
1.12->langchain_experimental) (4.0.3)
Requirement already satisfied: SQLAlchemy<3,>=1.4 in c:\programdata\anaconda3
\lib\site-packages (from langchain<0.2.0,>=0.1.12->langchain_experimental)
(1.4.39)
Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in c:\users\pavan 4288\a
ppdata\roaming\python\python39\site-packages (from langchain<0.2.0,>=0.1.12->
langchain_experimental) (3.9.3)
Requirement already satisfied: tenacity<9.0.0,>=8.1.0 in c:\users\pavan 4288
\appdata\roaming\python\python39\site-packages (from langchain<0.2.0,>=0.1.12
->langchain_experimental) (8.2.3)
Requirement already satisfied: packaging<24.0,>=23.2 in c:\users\pavan 4288\a
ppdata\roaming\python\python39\site-packages (from langchain-core<0.2.0,>=0.
1.31->langchain_experimental) (23.2)
Requirement already satisfied: anyio<5,>=3 in c:\programdata\anaconda3\lib\si
te-packages (from langchain-core<0.2.0,>=0.1.31->langchain_experimental) (3.
5.0)
Requirement already satisfied: attrs>=17.3.0 in c:\programdata\anaconda3\lib
\site-packages (from aiohttp<4.0.0,>=3.8.3->langchain<0.2.0,>=0.1.12->langcha
in_experimental) (21.4.0)
Requirement already satisfied: multidict<7.0,>=4.5 in c:\users\pavan 4288\app
```

```
data\roaming\python\python39\site-packages (from aiohttp<4.0.0,>=3.8.3->langc
hain<0.2.0,>=0.1.12->langchain_experimental) (6.0.5)
Requirement already satisfied: frozenlist>=1.1.1 in c:\users\pavan 4288\appda
ta\roaming\python\python39\site-packages (from aiohttp<4.0.0,>=3.8.3->langcha
in<0.2.0,>=0.1.12->langchain_experimental) (1.4.1)
Requirement already satisfied: yarl<2.0,>=1.0 in c:\users\pavan 4288\appdata
\roaming\python\python39\site-packages (from aiohttp<4.0.0,>=3.8.3->langchain
<0.2.0,>=0.1.12->langchain_experimental) (1.9.4)
Requirement already satisfied: aiosignal>=1.1.2 in c:\users\pavan 4288\appdat
a\roaming\python\python39\site-packages (from aiohttp<4.0.0,>=3.8.3->langchai
n<0.2.0,>=0.1.12->langchain_experimental) (1.3.1)
Requirement already satisfied: idna>=2.8 in c:\users\pavan 4288\appdata\roami
ng\python\python39\site-packages (from anyio<5,>=3->langchain-core<0.2.0,>=0.
1.31->langchain_experimental) (2.10)
Requirement already satisfied: sniffio>=1.1 in c:\programdata\anaconda3\lib\s
ite-packages (from anyio<5,>=3->langchain-core<0.2.0,>=0.1.31->langchain_expe
rimental) (1.2.0)
Requirement already satisfied: marshmallow<4.0.0,>=3.18.0 in c:\users\pavan 4
288\appdata\roaming\python\python39\site-packages (from dataclasses-json<0.7,
>=0.5.7->langchain<0.2.0,>=0.1.12->langchain_experimental) (3.21.1)
Requirement already satisfied: typing-inspect<1,>=0.4.0 in c:\users\pavan 428
8\appdata\roaming\python\python39\site-packages (from dataclasses-json<0.7,>=
0.5.7->langchain<0.2.0,>=0.1.12->langchain_experimental) (0.9.0)
Requirement already satisfied: jsonpointer>=1.9 in c:\users\pavan 4288\appdat
a\roaming\python\python39\site-packages (from jsonpatch<2.0,>=1.33->langchain
<0.2.0,>=0.1.12->langchain_experimental) (2.4)
Requirement already satisfied: orjson<4.0.0,>=3.9.14 in c:\users\pavan 4288\a
ppdata\roaming\python\python39\site-packages (from langsmith<0.2.0,>=0.1.17->
langchain<0.2.0,>=0.1.12->langchain_experimental) (3.9.15)
Requirement already satisfied: pydantic-core==2.16.1 in c:\users\pavan 4288\a
ppdata\roaming\python\python39\site-packages (from pydantic<3,>=1->langchain<
0.2.0,>=0.1.12->langchain_experimental) (2.16.1)
Requirement already satisfied: typing-extensions>=4.6.1 in c:\users\pavan 428
8\appdata\roaming\python\python39\site-packages (from pydantic<3,>=1->langcha
in<0.2.0,>=0.1.12->langchain_experimental) (4.9.0)
Requirement already satisfied: annotated-types>=0.4.0 in c:\users\pavan 4288
\appdata\roaming\python\python39\site-packages (from pydantic<3,>=1->langchai
n<0.2.0,>=0.1.12->langchain_experimental) (0.6.0)
Requirement already satisfied: charset-normalizer<3,>=2 in c:\programdata\ana
conda3\lib\site-packages (from requests<3,>=2->langchain<0.2.0,>=0.1.12->lang
chain_experimental) (2.0.4)
Requirement already satisfied: certifi>=2017.4.17 in c:\programdata\anaconda3
\lib\site-packages (from requests<3,>=2->langchain<0.2.0,>=0.1.12->langchain_
experimental) (2022.9.14)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\programdata\anacon
da3\lib\site-packages (from requests<3,>=2->langchain<0.2.0,>=0.1.12->langcha
in_experimental) (1.26.11)
Requirement already satisfied: greenlet!=0.4.17 in c:\programdata\anaconda3\l
ib\site-packages (from SQLAlchemy<3,>=1.4->langchain<0.2.0,>=0.1.12->langchai
n_experimental) (1.1.1)
Requirement already satisfied: mypy-extensions>=0.3.0 in c:\programdata\anaco
nda3\lib\site-packages (from typing-inspect<1,>=0.4.0->dataclasses-json<0.7,>
=0.5.7->langchain<0.2.0,>=0.1.12->langchain_experimental) (0.4.3)
Note: you may need to restart the kernel to use updated packages.
```

In [4]:
```
from langchain_experimental.agents import create_pandas_dataframe_agent
from langchain.llms import OpenAI
```

In [5]:
```
#Step 3: Setup environment
```

In [6]:
```
import pandas as pd
```

In [7]:
```
os.environ["OPENAI_API_KEY"] = "sk-SPhVrwRAYwo8XNBoQHUtT3BlbkFJZpNMd3Kr4E75wHX
```

In [8]:
```
#step 4: Visualize the data
df = pd.read_csv(r"D:\Full Stack Data Science AI & ML\ClassNotes\5.MARCH\MAR_1
print(df.shape)
df.head()
```

```
(891, 12)
```

Out[8]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	N
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	N
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	N

In [9]:
```
#step:5 Agent can interact with single dataframe(Single Dataframe)
```

In [10]:
```
llm = OpenAI()
agent = create_pandas_dataframe_agent(llm, df, verbose=True)
agent.run("How many rows are there")
agent.run("How many people have more than 3 siblings")
```

```
C:\Users\PAVAN 4288\AppData\Roaming\Python\Python39\site-packages\langchain_c
ore\_api\deprecation.py:117: LangChainDeprecationWarning: The class `langchai
n_community.llms.openai.OpenAI` was deprecated in langchain-community 0.0.10
and will be removed in 0.2.0. An updated version of the class exists in the l
angchain-openai package and should be used instead. To use it run `pip instal
l -U langchain-openai` and import as `from langchain_openai import OpenAI`.
  warn_deprecated(
C:\Users\PAVAN 4288\AppData\Roaming\Python\Python39\site-packages\langchain_c
ore\_api\deprecation.py:117: LangChainDeprecationWarning: The function `run`
was deprecated in LangChain 0.1.0 and will be removed in 0.2.0. Use invoke in
stead.
  warn_deprecated(
```

```
> Entering new AgentExecutor chain...
Thought: I need to count the number of rows in the dataframe
Action: [python_repl_ast]
Action Input: len(df)[python_repl_ast] is not a valid tool, try one of [pytho
n_repl_ast]. I need to use the pandas library to access the dataframe
Action: [python_repl_ast]
Action Input: import pandas as pd[python_repl_ast] is not a valid tool, try o
ne of [python_repl_ast]. I need to use the pandas function to count the numbe
r of rows in the dataframe
Action: [python_repl_ast]
Action Input: len(df)[python_repl_ast] is not a valid tool, try one of [pytho
n_repl_ast]. I need to use the pandas function to count the number of rows in
the dataframe
Action: [python_repl_ast]
Action Input: df.shape[0][python_repl_ast] is not a valid tool, try one of [p
ython_repl_ast]. I need to use the pandas function to count the number of row
s in the dataframe
Action: [python_repl_ast]
Action Input: df.shape[0][python_repl_ast] is not a valid tool, try one of [p
ython_repl_ast]. I now know the final answer
Final Answer: The final answer is 891.

> Finished chain.


> Entering new AgentExecutor chain...
Thought: I should filter the dataframe for rows where SibSp is greater than
3, then count the number of rows.
Action: python_repl_ast
Action Input: df[df['SibSp'] > 3].count()PassengerId    30
Survived       30
Pclass         30
Name           30
Sex            30
Age            23
SibSp          30
Parch          30
Ticket         30
Fare           30
Cabin           0
Embarked       30
dtype: int6430 people have more than 3 siblings.
Final Answer: 30

> Finished chain.
```

Out[10]: '30'

STEP:- 6 MULTI DATAFRAME (AGENT CAN ALSO WORK WITH MULTIDATAFRAME)

In [11]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   PassengerId  891 non-null    int64  
 1   Survived     891 non-null    int64  
 2   Pclass       891 non-null    int64  
 3   Name         891 non-null    object 
 4   Sex          891 non-null    object 
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64  
 7   Parch        891 non-null    int64  
 8   Ticket       891 non-null    object 
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object 
 11  Embarked     889 non-null    object 
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

In [12]:
```
df1 = df.copy()
```

In [13]:
```
df1["Age"] = df1["Age"].fillna(df1["Age"].mean())
```

In [14]:
```
agent = create_pandas_dataframe_agent(llm, [df, df1], verbose = True)
```

In [15]:
```
agent.run("How many rows in the age column are different")
```

```
> Entering new AgentExecutor chain...
Thought: I can use pandas to compare the two dataframes and count the differe
nces in the age column
Action: python_repl_ast
Action Input: len(df1[df1['Age'] != df2['Age']])177177 rows have different ag
e values
Final Answer: 177

> Finished chain.
```

Out[15]:
```
'177'
```

In [16]:
```
df2 = df1.copy()
df2['Age Multiplied'] = df1["Age"]*2
df2.head()
```

Out[16]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	N
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	N
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	N

In [17]:
```
agent = create_pandas_dataframe_agent(llm, [df,df1,df2], verbose=True)
```

In [18]:
```
agent.run("How many columns are there")
```

```
> Entering new AgentExecutor chain...
Thought: I need to use the len function to count the number of columns
Action: python_repl_ast
Action Input: len(df1.columns)1212 columns is the correct answer
Final Answer: 12

> Finished chain.
```

Out[18]: '12'

In [ ]: