```
In [1]:   pip install pymongo[srv]
```

Requirement already satisfied: pymongo[srv] in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (3.11.2)
Requirement already satisfied: dnspython<2.0.0,>=1.16.0; extra == "srv" in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (from pymongo[srv]) (1.16.0)
WARNING: You are using pip version 20.2.3; however, version 20.3.3 is available.
You should consider upgrading via the '/Library/Frameworks/Python.framework/Versions/3.9/bin/python3.9 -m pip install --upgrade pip' command.
Note: you may need to restart the kernel to use updated packages.

```
In [2]:   pip install tabula-py
```

Requirement already satisfied: tabula-py in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (2.2.0)
Requirement already satisfied: numpy in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (from tabula-py) (1.19.4)
Requirement already satisfied: pandas>=0.25.3 in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (from tabula-py) (1.1.4)
Requirement already satisfied: distro in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (from tabula-py) (1.5.0)
Requirement already satisfied: pytz>=2017.2 in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (from pandas>=0.25.3->tabula-py) (2020.4)
Requirement already satisfied: python-dateutil>=2.7.3 in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (from pandas>=0.25.3->tabula-py) (2.8.1)
Requirement already satisfied: six>=1.5 in /Library/Frameworks/Python.framework/Versions/3.9/lib/python3.9/site-packages (from python-dateutil>=2.7.3->pandas>=0.25.3->tabula-py) (1.15.0)
WARNING: You are using pip version 20.2.3; however, version 20.3.3 is available.
You should consider upgrading via the '/Library/Frameworks/Python.framework/Versions/3.9/bin/python3.9 -m pip install --upgrade pip' command.
Note: you may need to restart the kernel to use updated packages.

```
In [8]:   import tabula
          import pymongo
          from pymongo import MongoClient
          import pandas as pd
          import sys
          import csv
```

Creating Mongodb client

```
In [39]:  client=pymongo.MongoClient("localhost",27017)
          db = client.test
```

```
In [40]:  def csv_to_json(filename, header=None):
              data = pd.read_csv(filename, header=None)
              return data.to_dict('records')

          path = str(input("Enter file path: "))
          name = str(input("Enter  Collection name for database: "))

          records = db[name]
          tabula.convert_into(path , f"./{name}.csv" , output_format="csv", pages='all')
          file_data=(csv_to_json(f"./{name}.csv"))
          print(file_data)
```

[{0: 'Date', 1: 'Particulars', 2: 'Type of Interactlon'}, {0: 'April 06, 2018', 1: 'SBICAP Securities Ltd', 2: 'One on One'}]

```
In [41]:  l=[]
          for i in range(len(file_data)):
              FILE={}
              for key in file_data[i]:
                  FILE[str(key)]=file_data[i][key]
              l.append(FILE)
```

```
In [42]:  l
```

Out[42]: [{'0': 'Date', '1': 'Particulars', '2': 'Type of Interactlon'},
          {'0': 'April 06, 2018', '1': 'SBICAP Securities Ltd', '2': 'One on One'}]

```
In [43]:  if isinstance(l, list):
              records.insert_many(l)
          else:
              records.insert_one(l)
```

```
In [44]:  abc=records.find()
```

```
In [45]:  list(abc)
```

Out[45]: [{'_id': ObjectId('5ff0558378564f9b0a33e0b4'),
           '0': 'Date',
           '1': 'Particulars',
           '2': 'Type of Interactlon'},
          {'_id': ObjectId('5ff0558378564f9b0a33e0b5'),
           '0': 'April 06, 2018',
           '1': 'SBICAP Securities Ltd',
           '2': 'One on One'}]

```
In [46]:  import tabula
          import pandas as pd
          import numpy as np
          from IPython.display import display
```

```
In [48]:  doc='1c1edeee-a13e-4b2e-90be-eb1dd03c3384.pdf'
          df = tabula.read_pdf(doc, pages="all")
```

```
In [49]:  for x in df:
              dfobj=pd.DataFrame(x)
              display(dfobj)
```

	Date	Particulars	Type of Interactlon
0	April 06, 2018	SBICAP Securities Ltd	One on One

```
In [ ]:
```