```
import pandas as pd
df=pd.DataFrame (
    {"a":[4,5,6,],
     "b":[7,8,9],
     "c":[10,11,12]},
      index=[1,2,3]
                )
df
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 3,\n  \"fields\": [\n
{\n      \"column\": \"a\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 1,\n        \"min\": 4,\n
\"max\": 6,\n        \"num_unique_values\": 3,\n        \"samples\":
[\n          4,\n          5,\n          6\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"b\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 1,\n        \"min\": 7,\n
\"max\": 9,\n        \"num_unique_values\": 3,\n        \"samples\":
[\n          7,\n          8,\n          9\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"c\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 1,\n        \"min\": 10,\n
\"max\": 12,\n        \"num_unique_values\": 3,\n        \"samples\":
[\n          10,\n          11,\n          12\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data=[1,2,3,4]
df=pd.DataFrame(data)
df
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 4,\n  \"fields\": [\n
{\n      \"column\": 0,\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 1,\n        \"min\": 1,\n
\"max\": 4,\n        \"num_unique_values\": 4,\n        \"samples\":
[\n          2,\n          4,\n          1\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data=[['Alex',10],['Bob',12]]
df=pd.DataFrame(data,columns=['Name',"age"])
df
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 2,\n  \"fields\": [\n
{\n      \"column\": \"Name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 2,\n
\"samples\": [\n          \"Bob\",\n          \"Alex\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
```

```
n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 1,\n        \"min\": 10,\n
\"max\": 12,\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          12,\n          10\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    }\n  ]\
n}","type":"dataframe","variable_name":"df"}
```


```
import pandas as pd
my_dataset={
    "cars":["BMV","VOLVO","FORD"],
    "passings":[3,7,2]
}
df=pd.DataFrame(my_dataset)
print(df)
```

```
    cars  passings
0    BMV         3
1  VOLVO         7
2   FORD         2
```

```
data=[{
    "a":1,"b":2
},{
    "a":5,"b":10,"c":20
}]
df=pd.DataFrame(data)
df=pd.DataFrame(data,index=["first","second"])
df1=pd.DataFrame(data,index=["first","second"],columns=["a","b","c"])
df1
```


```
{"summary":"{\n  \"name\": \"df1\",\n  \"rows\": 2,\n  \"fields\": [\n
{\n      \"column\": \"a\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 2,\n        \"min\": 1,\n
\"max\": 5,\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          5,\n          1\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"b\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 5,\n        \"min\": 2,\n
\"max\": 10,\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          10,\n          2\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"c\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": null,\n        \"min\": 20.0,\n
\"max\": 20.0,\n        \"num_unique_values\": 1,\n
\"samples\": [\n          20.0\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    }\n  ]\
n}","type":"dataframe","variable_name":"df1"}
```


```
data={
    "name":["jai","gaurav","princi","Anuj"],
```

```
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"]
}
df=pd.DataFrame(data)
address=["delhi","Bangalore","chennai","Patna"]
df["Address"]=address
df
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 4,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 4,\n
\"samples\": [\n          \"gaurav\",\n          \"Anuj\",\n
\"jai\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n          \"5.1\",\
n          \"6.2\",\n          \"5.2\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"Qualification\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"MA\",\n
\"Msc\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Address\",\n      \"properties\": {\n        \"dtype\": \"string\",\
n        \"num_unique_values\": 4,\n        \"samples\": [\n
\"Bangalore\",\n          \"Patna\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```

```
data={
    "name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"]
}
df=pd.DataFrame(data)
address=["delhi","Bangalore","chennai","Patna"]
df["Address"]=address
del df["Address"]
df
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 4,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 4,\n
\"samples\": [\n          \"gaurav\",\n          \"Anuj\",\n
\"jai\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n          \"5.1\",\
n          \"6.2\",\n          \"5.2\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
```

```
n    },\n    {\n      \"column\": \"Qualification\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"MA\",\n
\"Msc\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    }\n  ]\
n}","type":"dataframe","variable_name":"df"}
```


```
data={
    "name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"]
}
df=pd.DataFrame(data)
address=["delhi","Bangalore","chennai","Patna"]
df["Address"]=address
df.drop(columns=["Address"],inplace=True)
df.pop("Qualification")
df
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 4,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 4,\n
\"samples\": [\n          \"gaurav\",\n          \"Anuj\",\n
\"jai\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n          \"5.1\",\
n          \"6.2\",\n          \"5.2\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```


```
data={
    "name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"]
}
df=pd.DataFrame(data)
address=["delhi","Bangalore","chennai","Patna"]
df["Address"]=address
df.drop(["Address"],axis=1,inplace=True)
df.pop("Qualification")
df
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 4,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 4,\n
\"samples\": [\n          \"gaurav\",\n          \"Anuj\",\n
\"jai\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
```

\"num_unique_values\": 3,\n   \"samples\": [\n   \"5.1\",\n   \"6.2\",\n   \"5.2\"\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   }\n]\n}","type":"dataframe","variable_name":"df"}

```
data={
    "name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"]
}
df=pd.DataFrame(data)
address=["delhi","Bangalore","chennai","Patna"]
df["Address"]=address
df.drop(["Address"],axis=1,inplace=False)
df.pop("Qualification")
df
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 4,\n \"fields\": [\n{\n   \"column\": \"name\",\n   \"properties\": {\n\"dtype\": \"string\",\n   \"num_unique_values\": 4,\n\"samples\": [\n   \"gaurav\",\n   \"Anuj\",\n\"jai\"\n   ],\n   \"semantic_type\": \"\",\n\"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"height\",\n   \"properties\": {\n   \"dtype\": \"string\",\n\"num_unique_values\": 3,\n   \"samples\": [\n   \"5.1\",\n   \"6.2\",\n   \"5.2\"\n   ],\n\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"Address\",\n   \"properties\": {\n   \"dtype\": \"string\",\n   \"num_unique_values\": 4,\n\"samples\": [\n   \"Bangalore\",\n   \"Patna\",\n\"delhi\"\n   ],\n   \"semantic_type\": \"\",\n\"description\": \"\"\n   }\n   }\n]\n}","type":"dataframe","variable_name":"df"}

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]


}
df=pd.DataFrame(data)
df.rename(columns={"Address":"Place"},inplace=True)
df
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 4,\n \"fields\": [\n{\n   \"column\": \"Name\",\n   \"properties\": {\n\"dtype\": \"string\",\n   \"num_unique_values\": 4,\n

\"samples\": [\n   \"gaurav\",\n   \"Anuj\",\n
\"jai\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"height\",\n   \"properties\": {\n   \"dtype\": \"string\",\n
\"num_unique_values\": 3,\n   \"samples\": [\n   \"5.1\",\
n   \"6.2\",\n   \"5.2\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"Qualification\",\n
\"properties\": {\n   \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n   \"samples\": [\n   \"MA\",\n
\"Msc\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"Place\",\n   \"properties\": {\n   \"dtype\": \"string\",\n
\"num_unique_values\": 4,\n   \"samples\": [\n
\"bangalore\",\n   \"patna\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   }\n]\n}","type":"dataframe","variable_name":"df"}

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]


}
df=pd.DataFrame(data)
df.drop(0,axis=0,inplace=True)
df
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 3,\n \"fields\": [\n
{\n   \"column\": \"Name\",\n   \"properties\": {\n
\"dtype\": \"string\",\n   \"num_unique_values\": 3,\n
\"samples\": [\n   \"gaurav\",\n   \"princi\",\n
\"Anuj\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"height\",\n   \"properties\": {\n   \"dtype\": \"string\",\n
\"num_unique_values\": 3,\n   \"samples\": [\n   \"6.2\",\
n   \"5.1\",\n   \"5.2\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"Qualification\",\n
\"properties\": {\n   \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n   \"samples\": [\n   \"Msc\",\
n   \"MA\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"Address\",\n   \"properties\": {\n   \"dtype\": \"string\",\
n   \"num_unique_values\": 3,\n   \"samples\": [\n
\"bangalore\",\n   \"chennai\"\n   ],\n

```
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]


}
df=pd.DataFrame(data)
df.drop(1,axis=0,inplace=True)
df
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 3,\n  \"fields\": [\n
{\n      \"column\": \"Name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"jai\",\n          \"princi\",\n
\"Anuj\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"5.2\",\
n          \"5.1\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Qualification\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 1,\n        \"samples\":
[\n          \"Msc\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Address\",\n      \"properties\": {\n        \"dtype\": \"string\",\
n        \"num_unique_values\": 3,\n        \"samples\": [\n
\"delhi\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    }\n  ]\
n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]


}
df=pd.DataFrame(data)
print(df[["Name","Qualification"]])
```

```
     Name Qualification
0     jai           Msc
1  gaurav            MA
```

```
2  princi           Msc
3    Anuj           Msc
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]


}
df=pd.DataFrame(data)
df.filter(items=['Name',"Address"])
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 4,\n  \"fields\": [\n
{\n      \"column\": \"Name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 4,\n
\"samples\": [\n          \"gaurav\",\n          \"Anuj\",\n
\"jai\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Address\",\n      \"properties\": {\n        \"dtype\": \"string\",\
n        \"num_unique_values\": 4,\n        \"samples\": [\n
\"bangalore\",\n          \"patna\",\n          \"delhi\"\n        ],\
n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    }\n  ]\n}","type":"dataframe"}
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]


}
df=pd.DataFrame(data)
df.filter(like="eigh")
```

```
Name                gaurav
height                 6.2
Qualification           MA
Address          bangalore
Name: 1, dtype: object
```

```
data={
    "Name":["jai","gaurav","princi","Anuj"],
    "height":["5.1","6.2","5.1","5.2"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]
```

```
}
df=pd.DataFrame(data)
df.filter(regex="e|a",axis=1)
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 4,\n  \"fields\": [\n
{\n      \"column\": \"Name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 4,\n
\"samples\": [\n          \"gaurav\",\n          \"Anuj\",\n
\"jai\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n          \"5.1\",\
n          \"6.2\",\n          \"5.2\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"Qualification\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"MA\",\n
\"Msc\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Address\",\n      \"properties\": {\n        \"dtype\": \"string\",\
n        \"num_unique_values\": 4,\n        \"samples\": [\n
\"bangalore\",\n          \"patna\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe"}
```


```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","jai"],
    "height":["5.1","6.2","5.1","5.1"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]
}
df=pd.DataFrame(data)
df = df.drop_duplicates(subset=["Name","height"],keep="last")
df
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 3,\n  \"fields\": [\n
{\n      \"column\": \"Name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"gaurav\",\n          \"princi\",\n
\"jai\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"5.1\",\
n          \"6.2\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Qualification\",\n      \"properties\": {\n        \"dtype\":
\"string\",\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          \"Msc\",\n          \"MA\"\n        ],\n
```

```
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"Address\",\n      \"properties\":
{\n        \"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"bangalore\",\n          \"chennai\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","jai"],
    "height":["5.1","6.2","5.1","5.1"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","patna"]
}
df=pd.DataFrame(data)
df = df.drop_duplicates(subset=["Name","height"])
df
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 3,\n  \"fields\": [\n
{\n      \"column\": \"Name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"jai\",\n          \"gaurav\",\n
\"princi\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"6.2\",\
n          \"5.1\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Qualification\",\n      \"properties\": {\n        \"dtype\":
\"string\",\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          \"MA\",\n          \"Msc\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"Address\",\n      \"properties\":
{\n        \"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"delhi\",\n          \"bangalore\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","jai"],
    "height":["5.1","6.2","5.1","5.1"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","delhi"]
}
df=pd.DataFrame(data)
df_sample=df.sample(n=2)
print(df_sample)
```

```
  Name height Qualification Address
0  jai    5.1           Msc   delhi
3  jai    5.1           Msc   delhi
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","jai"],
    "height":["5.1","6.2","5.1","5.1"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","delhi"]
}
df=pd.DataFrame(data)
df_sample=df.sample(frac=0.5)
print(df_sample)
```

```
     Name height Qualification    Address
2  princi    5.1           Msc    chennai
1  gaurav    6.2            MA  bangalore
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","jai"],
    "height":["5.1","6.2","5.1","5.1"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","delhi"]
}
df=pd.DataFrame(data)
df = df.drop_duplicates()
df
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 3,\n  \"fields\": [\n
{\n      \"column\": \"Name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"jai\",\n          \"gaurav\",\n
\"princi\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"6.2\",\
n          \"5.1\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"Qualification\",\n      \"properties\": {\n        \"dtype\":
\"string\",\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          \"MA\",\n          \"Msc\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"Address\",\n      \"properties\":
{\n        \"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"delhi\",\n          \"bangalore\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data = {
    "name": ["Jai", "Princi", "Gaurav", "Anuj"],
    "age": [27, 24, 22, 32],
    "gender":["f","m","f",'m'],
    "height":["5.1","6.2","5.1","5.2"],
}
df = pd.DataFrame(data)
df=df["name"]
df
```

```
0       Jai
1    Princi
2    Gaurav
3      Anuj
Name: name, dtype: object
```

```
import pandas as pd
df=pd.DataFrame([[1,2],[3,4]],columns=["a","b"])
df2=pd.DataFrame([[5,6],[7,8]],columns=["a","b"])
df=pd.concat([df,df2],ignore_index=True)
df
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 4,\n  \"fields\": [\n
{\n      \"column\": \"a\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 2,\n        \"min\": 1,\n
\"max\": 7,\n        \"num_unique_values\": 4,\n        \"samples\":
[\n          3,\n          7,\n          1\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"b\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 2,\n        \"min\": 2,\n
\"max\": 8,\n        \"num_unique_values\": 4,\n        \"samples\":
[\n          4,\n          8,\n          2\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000]
}
df=pd.DataFrame(data)
top_salries=df.nlargest(2,columns="salary")
print(top_salries)
```

```
    name  age  salary
1    bob   25   20000
0  alice   24   10000
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000]
}
df=pd.DataFrame(data)
df.query('age>=25')
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 2,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 2,\n
\"samples\": [\n          \"charlie\",\n          \"bob\"\n        ],\
n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\
n        \"dtype\": \"number\",\n        \"std\": 0,\n        \"min\":
25,\n        \"max\": 26,\n        \"num_unique_values\": 2,\n
\"samples\": [\n          26,\n          25\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"salary\",\n      \"properties\":
{\n        \"dtype\": \"number\",\n        \"std\": 12020,\n
\"min\": 3000,\n        \"max\": 20000,\n
\"num_unique_values\": 2,\n        \"samples\": [\n          3000,\n
20000\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    }\n  ]\n}","type":"dataframe"}
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"]
}
df=pd.DataFrame(data)
df.query('name.str.contains("a")and age<=25')
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 1,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 1,\n
\"samples\": [\n          \"alice\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": null,\n        \"min\": 24,\n
\"max\": 24,\n        \"num_unique_values\": 1,\n        \"samples\":
[\n          24\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"salary\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": null,\n        \"min\": 10000,\n        \"max\": 10000,\n
\"num_unique_values\": 1,\n        \"samples\": [\n          10000\n
```

```
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    }\n  ]\n}","type":"dataframe"}
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"]
}
df=pd.DataFrame(data)
df.query('gender==["M"] and age>=25')
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 1,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 1,\n
\"samples\": [\n          \"bob\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": null,\n        \"min\": 25,\n
\"max\": 25,\n        \"num_unique_values\": 1,\n        \"samples\":
[\n          25\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"salary\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": null,\n        \"min\": 20000,\n        \"max\": 20000,\n
\"num_unique_values\": 1,\n        \"samples\": [\n          20000\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"gender\",\n      \"properties\":
{\n        \"dtype\": \"string\",\n        \"num_unique_values\": 1,\n
\"samples\": [\n          \"M\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe"}
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"]
}
df=pd.DataFrame(data)
df.iat[1,2]
```

```
20000
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
```

```
    "gender":["F","M","F"]
}
df=pd.DataFrame(data)
df.iloc[0:3]
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 3,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"alice\",\n          \"bob\",\n
\"charlie\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"age\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 1,\n        \"min\": 24,\n        \"max\": 26,\n
\"num_unique_values\": 3,\n        \"samples\": [\n          24,\n
25,\n          26\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"salary\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 8544,\n        \"min\": 3000,\n        \"max\": 20000,\n
\"num_unique_values\": 3,\n        \"samples\": [\n          10000,\n
20000,\n          3000\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"gender\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"M\",\n
\"F\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    }\n  ]\n}","type":"dataframe"}
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"]
}
df=pd.DataFrame(data)
df.iloc[0:2]
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 2,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 2,\n
\"samples\": [\n          \"bob\",\n          \"alice\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 0,\n        \"min\": 24,\n
\"max\": 25,\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          25,\n          24\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"salary\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 7071,\n        \"min\": 10000,\n
\"max\": 20000,\n        \"num_unique_values\": 2,\n
\"samples\": [\n          20000,\n          10000\n        ],\n
```

\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"gender\",\n   \"properties\":
{\n   \"dtype\": \"string\",\n   \"num_unique_values\": 2,\n
\"samples\": [\n   \"M\",\n   \"F\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   }\n]\n}","type":"dataframe"}

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"]
}
df=pd.DataFrame(data)
df.loc[:,['age']]
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 3,\n \"fields\": [\n
{\n   \"column\": \"age\",\n   \"properties\": {\n
\"dtype\": \"number\",\n   \"std\": 1,\n   \"min\": 24,\n
\"max\": 26,\n   \"num_unique_values\": 3,\n   \"samples\":
[\n   24,\n   25,\n   26\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   }\n]\n}","type":"dataframe"}

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"]
}
df=pd.DataFrame(data)
df.loc[:,["name","salary"]]
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 3,\n \"fields\": [\n
{\n   \"column\": \"name\",\n   \"properties\": {\n
\"dtype\": \"string\",\n   \"num_unique_values\": 3,\n
\"samples\": [\n   \"alice\",\n   \"bob\",\n
\"charlie\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"salary\",\n   \"properties\": {\n   \"dtype\": \"number\",\n
\"std\": 8544,\n   \"min\": 3000,\n   \"max\": 20000,\n
\"num_unique_values\": 3,\n   \"samples\": [\n   10000,\n
20000,\n   3000\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   }\n]\n}","type":"dataframe"}

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
```

```
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"]
}
df=pd.DataFrame(data)
df.iloc[:,0]
```

```
0      alice
1        bob
2    charlie
Name: name, dtype: object
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"],
    "height":[1.8,1.7,1.6,1.5]
}
df=pd.DataFrame(data)
df.iloc[0:2]
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 2,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 2,\n
\"samples\": [\n          \"bob\",\n          \"alice\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 0,\n        \"min\": 24,\n
\"max\": 25,\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          25,\n          24\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"salary\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 7071,\n        \"min\": 10000,\n
\"max\": 20000,\n        \"num_unique_values\": 2,\n
\"samples\": [\n          20000,\n          10000\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"gender\",\n      \"properties\":
{\n        \"dtype\": \"string\",\n        \"num_unique_values\": 2,\n
\"samples\": [\n          \"M\",\n          \"F\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe"}
```


```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"]
```

```
}
df=pd.DataFrame(data)
df_filtered=df[df["age"]>=25]
df_filtered
```

```
{"summary":"{\n  \"name\": \"df_filtered\",\n  \"rows\": 2,\n
\"fields\": [\n    {\n      \"column\": \"name\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n
\"charlie\",\n          \"bob\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 0,\n        \"min\": 25,\n
\"max\": 26,\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          26,\n          25\n        ],\n        \"semantic_type\":
\"\",\n        \"description\": \"\"\n      }\n    },\n    {\n
\"column\": \"salary\",\n      \"properties\": {\n        \"dtype\":
\"number\",\n        \"std\": 12020,\n        \"min\": 3000,\n
\"max\": 20000,\n        \"num_unique_values\": 2,\n
\"samples\": [\n          3000,\n          20000\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"gender\",\n      \"properties\":
{\n        \"dtype\": \"string\",\n        \"num_unique_values\": 2,\n
\"samples\": [\n          \"F\",\n          \"M\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe","variable_name":"df_filtered"}
```

```
import pandas as pd
data={
    "name":["alice","bob","charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"],
    "height":[1.8,1.7,1.6]
}
df=pd.DataFrame(data)
df_filtered=df[(df["age"]>=25) & (df['gender']=="M")]
df_filtered
```

```
{"summary":"{\n  \"name\": \"df_filtered\",\n  \"rows\": 1,\n
\"fields\": [\n    {\n      \"column\": \"name\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 1,\n        \"samples\": [\n          \"bob\"\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\
n        \"dtype\": \"number\",\n        \"std\": null,\n
\"min\": 25,\n        \"max\": 25,\n        \"num_unique_values\": 1,\
n        \"samples\": [\n          25\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"salary\",\n      \"properties\":
```

{\n   \"dtype\": \"number\",\n   \"std\": null,\n
\"min\": 20000,\n   \"max\": 20000,\n
\"num_unique_values\": 1,\n   \"samples\": [\n   20000\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   },\n   {\n   \"column\": \"gender\",\n   \"properties\":
{\n   \"dtype\": \"string\",\n   \"num_unique_values\": 1,\n
\"samples\": [\n   \"M\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"height\",\n   \"properties\":
{\n   \"dtype\": \"number\",\n   \"std\": null,\n
\"min\": 1.7,\n   \"max\": 1.7,\n   \"num_unique_values\":
1,\n   \"samples\": [\n   1.7\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   }\n]\n}","type":"dataframe","variable_name":"df_filtered"}

```
import pandas as pd
data={
    "name":["alice","bob","Charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"],
    "height":[1.8,1.7,1.6]
}
df=pd.DataFrame(data)
df_filtered=df[df['name'].str.startswith(('a','C'))]
df_filtered
```

{"summary":"{\n \"name\": \"df_filtered\",\n \"rows\": 2,\n
\"fields\": [\n   {\n   \"column\": \"name\",\n
\"properties\": {\n   \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n   \"samples\": [\n
\"Charlie\",\n   \"alice\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"age\",\n   \"properties\": {\n
\"dtype\": \"number\",\n   \"std\": 1,\n   \"min\": 24,\n
\"max\": 26,\n   \"num_unique_values\": 2,\n   \"samples\":
[\n   26,\n   24\n   ],\n   \"semantic_type\":
\"\",\n   \"description\": \"\"\n   }\n   },\n   {\n
\"column\": \"salary\",\n   \"properties\": {\n   \"dtype\":
\"number\",\n   \"std\": 4949,\n   \"min\": 3000,\n
\"max\": 10000,\n   \"num_unique_values\": 2,\n
\"samples\": [\n   3000,\n   10000\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"gender\",\n   \"properties\":
{\n   \"dtype\": \"string\",\n   \"num_unique_values\": 1,\n
\"samples\": [\n   \"F\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"height\",\n   \"properties\":
{\n   \"dtype\": \"number\",\n   \"std\":
0.14142135623730948,\n   \"min\": 1.6,\n   \"max\": 1.8,\n

```
\"num_unique_values\": 2,\n        \"samples\": [\n          1.6\n
],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    }\n  ]\n}","type":"dataframe","variable_name":"df_filtered"}
```

```
import pandas as pd
data={
    "name":["alice","bob","Charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"],
    "height":[1.8,1.7,1.6]
}
df=pd.DataFrame(data)
df.head(2)
```

```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 3,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 3,\n
\"samples\": [\n          \"alice\",\n          \"bob\",\n
\"Charlie\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"age\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 1,\n        \"min\": 24,\n        \"max\": 26,\n
\"num_unique_values\": 3,\n        \"samples\": [\n          24,\n
25,\n          26\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"salary\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 8544,\n        \"min\": 3000,\n        \"max\": 20000,\n
\"num_unique_values\": 3,\n        \"samples\": [\n          10000,\n
20000,\n          3000\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"gender\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"M\",\n
\"F\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 0.09999999999999998,\n        \"min\": 1.6,\n        \"max\":
1.8,\n        \"num_unique_values\": 3,\n        \"samples\": [\n
1.8,\n          1.7\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    }\n  ]\
n}","type":"dataframe","variable_name":"df"}
```

```
import pandas as pd
data={
    "name":["alice","bob","Charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"],
    "height":[1.8,1.7,1.6]
}
```

```
df=pd.DataFrame(data)
df.tail(2)
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 2,\n  \"fields\": [\n
{\n      \"column\": \"name\",\n      \"properties\": {\n
\"dtype\": \"string\",\n        \"num_unique_values\": 2,\n
\"samples\": [\n          \"Charlie\",\n          \"bob\"\n        ],\
n        \"semantic_type\": \"\",\n        \"description\": \"\"\n
}\n    },\n    {\n      \"column\": \"age\",\n      \"properties\": {\
n        \"dtype\": \"number\",\n        \"std\": 0,\n        \"min\":
25,\n        \"max\": 26,\n        \"num_unique_values\": 2,\n
\"samples\": [\n          26,\n          25\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"salary\",\n      \"properties\":
{\n        \"dtype\": \"number\",\n        \"std\": 12020,\n
\"min\": 3000,\n        \"max\": 20000,\n
\"num_unique_values\": 2,\n        \"samples\": [\n          3000,\n
20000\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"gender\",\n      \"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 2,\n        \"samples\": [\n          \"F\",\n
\"M\"\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"height\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 0.07071067811865465,\n        \"min\": 1.6,\n        \"max\":
1.7,\n        \"num_unique_values\": 2,\n        \"samples\": [\n
1.6,\n          1.7\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    }\n  ]\n}","type":"dataframe"}
```


```
import pandas as pd
data={
    "name":["alice","bob","Charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,3000],
    "gender":["F","M","F"],
    "height":[1.8,1.7,1.6]
}
df=pd.DataFrame(data)
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3 entries, 0 to 2
Data columns (total 5 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   name    3 non-null      object
 1   age     3 non-null      int64
 2   salary  3 non-null      int64
 3   gender  3 non-null      object
 4   height  3 non-null      float64
```

```
dtypes: float64(1), int64(2), object(2)
memory usage: 248.0+ bytes
```

```
import pandas as pd
data={
    "name":["alice","bob","Charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,30000],
    "gender":["F","M","F"],
    "height":[1.8,1.7,1.6]
}
df=pd.DataFrame(data)
df.describe()
```


```
{"summary":"{\n  \"name\": \"df\",\n  \"rows\": 8,\n  \"fields\": [\n
{\n      \"column\": \"age\",\n      \"properties\": {\n
\"dtype\": \"number\",\n        \"std\": 10.677078252031311,\n
\"min\": 1.0,\n        \"max\": 26.0,\n        \"num_unique_values\":
7,\n        \"samples\": [\n          3.0,\n          25.0,\n
25.5\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"salary\",\n      \"properties\": {\n        \"dtype\": \"number\",\n
\"std\": 9542.405416088755,\n        \"min\": 3.0,\n        \"max\":
30000.0,\n        \"num_unique_values\": 6,\n        \"samples\": [\n
3.0,\n          20000.0,\n          30000.0\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"height\",\n      \"properties\":
{\n        \"dtype\": \"number\",\n        \"std\":
0.7804531650631849,\n        \"min\": 0.09999999999999998,\n
\"max\": 3.0,\n        \"num_unique_values\": 7,\n        \"samples\":
[\n          3.0,\n          1.7,\n          1.75\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    }\n  ]\n}","type":"dataframe"}
```


```
import pandas as pd
data={
    "name":["alice","bob","Charlie"],
    "age":[24,25,26],
    "salary":[10000,20000,30000],
    "gender":["F","M","F"],
    "height":[1.8,1.7,1.6]
}
df=pd.DataFrame(data)
df_sorted=df.sort_values(by="age",ascending=False)
df_sorted
```


```
{"summary":"{\n  \"name\": \"df_sorted\",\n  \"rows\": 3,\n
\"fields\": [\n    {\n      \"column\": \"name\",\n
\"properties\": {\n        \"dtype\": \"string\",\n
\"num_unique_values\": 3,\n        \"samples\": [\n
```

\"Charlie\",\n   \"bob\",\n   \"alice\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"age\",\n   \"properties\": {\n
\"dtype\": \"number\",\n   \"std\": 1,\n   \"min\": 24,\n
\"max\": 26,\n   \"num_unique_values\": 3,\n   \"samples\":
[\n   26,\n   25,\n   24\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"salary\",\n   \"properties\":
{\n   \"dtype\": \"number\",\n   \"std\": 10000,\n
\"min\": 10000,\n   \"max\": 30000,\n
\"num_unique_values\": 3,\n   \"samples\": [\n   30000,\n
20000,\n   10000\n   ],\n   \"semantic_type\": \"\",\
n   \"description\": \"\"\n   }\n   },\n   {\n
\"column\": \"gender\",\n   \"properties\": {\n   \"dtype\":
\"string\",\n   \"num_unique_values\": 2,\n   \"samples\":
[\n   \"M\",\n   \"F\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"height\",\n   \"properties\":
{\n   \"dtype\": \"number\",\n   \"std\":
0.09999999999999998,\n   \"min\": 1.6,\n   \"max\": 1.8,\n
\"num_unique_values\": 3,\n   \"samples\": [\n   1.6,\n
1.7\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   }\n]\
n}","type":"dataframe","variable_name":"df_sorted"}

```
import pandas as pd
data = {'name': ['Alice', 'Bob', 'Charlie', 'Dave', 'Emily', 'Frank'],
        'gender': ['F', 'M', 'M', 'M', 'F', 'M'],
        'age': [25, 35, 40, 28, 30, 45],
        'salary': [50000, 70000, 60000, 80000, 65000, 90000]}
df=pd.DataFrame(data)
grouped=df.groupby('gender')['salary'].mean()
grouped
```

```
gender
F    57500.0
M    75000.0
Name: salary, dtype: float64
```

```
import pandas as pd
data={
    "Name":["jai","gaurav","princi","jai"],
    "height":["5.1","6.2","5.1","5.1"],
    "Qualification":["Msc","MA","Msc","Msc"],
    "Address":["delhi","bangalore","chennai","delhi"],
    "salary":[50000, 70000, 60000, 80000],
    "age":[18,17,16,15]
}
df=pd.DataFrame(data)
grouped=df.groupby(by='age')
```