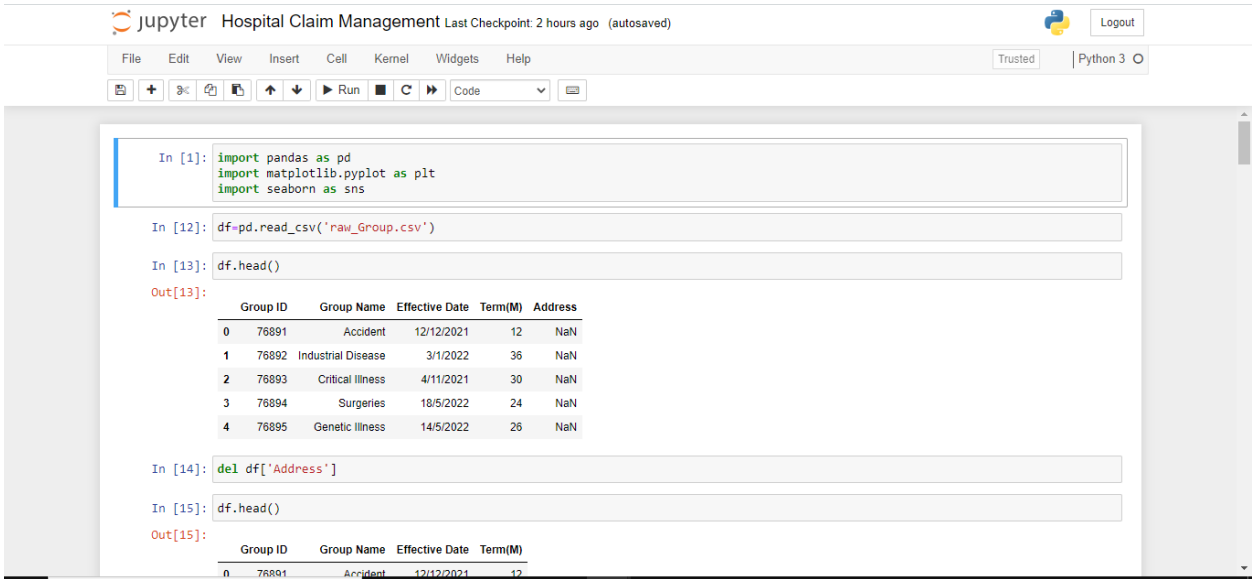
Cleansing of Group csv file



jupyter Hospital Claim Management Last Checkpoint: 2 hours ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run Code

```
In [1]: import pandas as pd
        import matplotlib.pyplot as plt
        import seaborn as sns
```

```
In [12]: df=pd.read_csv('raw_Group.csv')
```

```
In [13]: df.head()
```

Out[13]:

	Group ID	Group Name	Effective Date	Term(M)	Address
0	76891	Accident	12/12/2021	12	NaN
1	76892	Industrial Disease	3/1/2022	36	NaN
2	76893	Critical Illness	4/11/2021	30	NaN
3	76894	Surgeries	18/5/2022	24	NaN
4	76895	Genetic Illness	14/5/2022	26	NaN

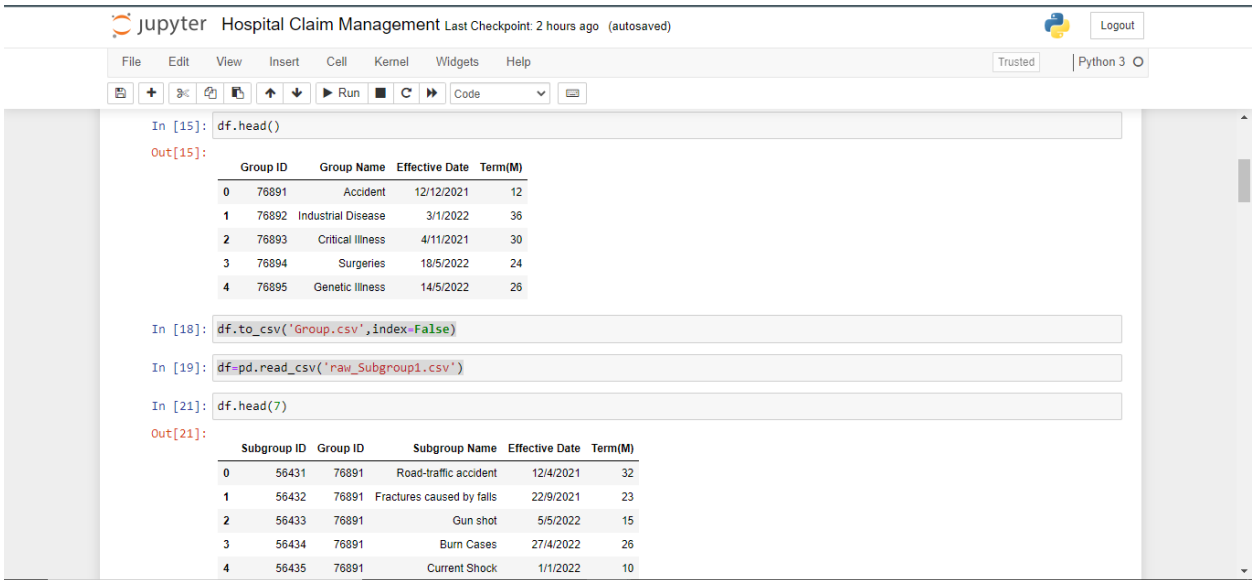
```
In [14]: del df['Address']
```

```
In [15]: df.head()
```

Out[15]:

	Group ID	Group Name	Effective Date	Term(M)
0	76891	Accident	12/12/2021	12

Cleansing of Subgroup1 csv file



jupyter Hospital Claim Management Last Checkpoint: 2 hours ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run Code

```
In [15]: df.head()
```

Out[15]:

	Group ID	Group Name	Effective Date	Term(M)
0	76891	Accident	12/12/2021	12
1	76892	Industrial Disease	3/1/2022	36
2	76893	Critical Illness	4/11/2021	30
3	76894	Surgeries	18/5/2022	24
4	76895	Genetic Illness	14/5/2022	26

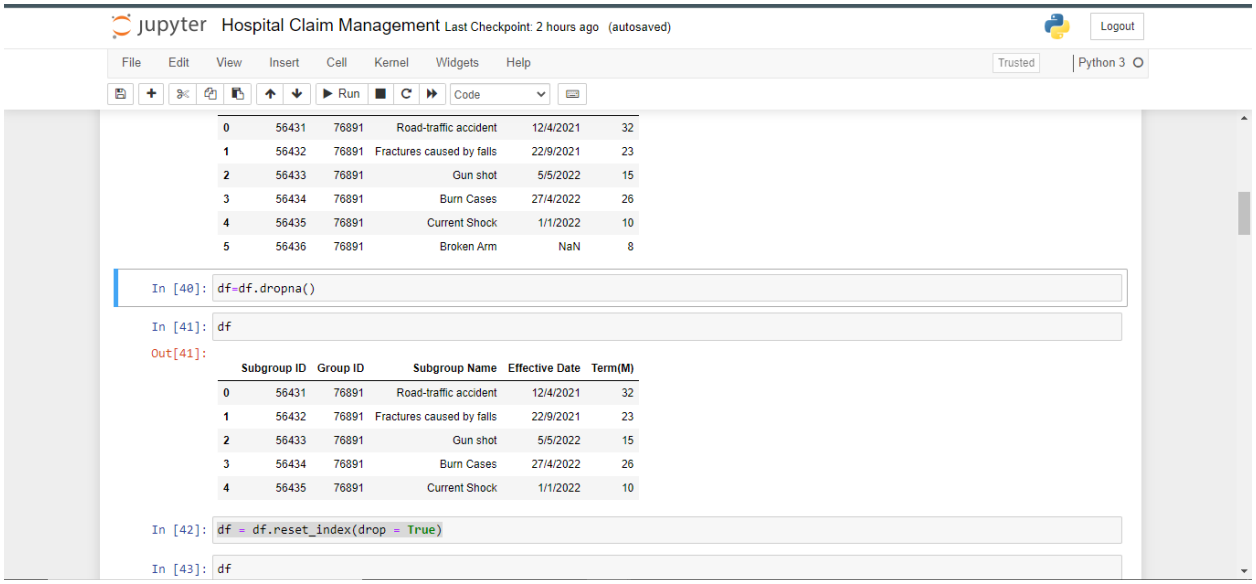
```
In [18]: df.to_csv('Group.csv',index=False)
```

```
In [19]: df=pd.read_csv('raw_Subgroup1.csv')
```

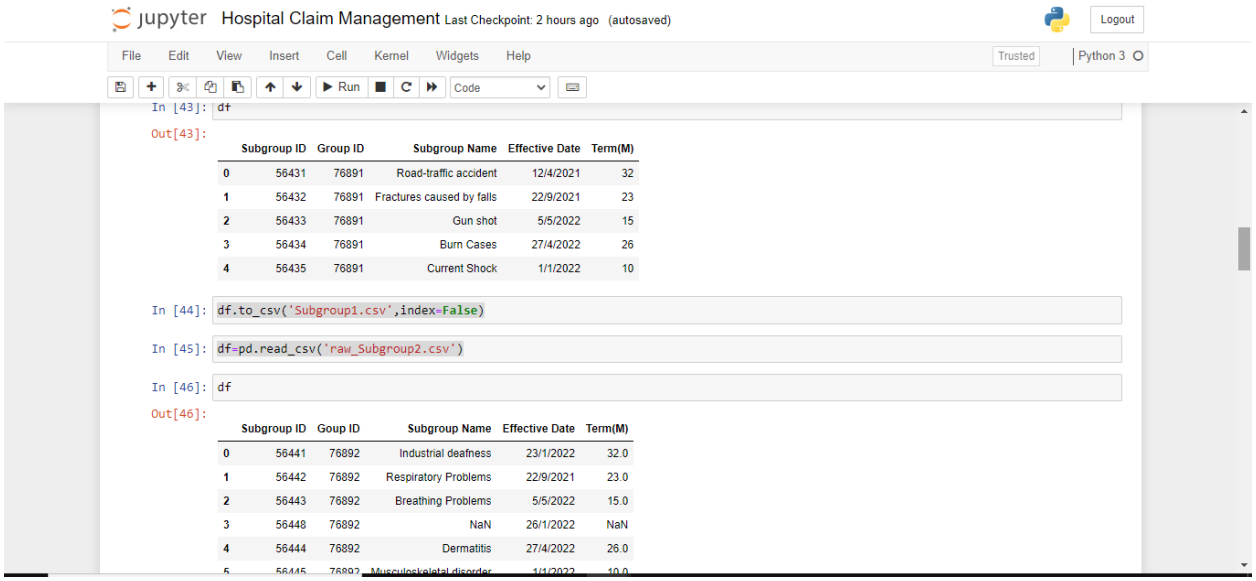
```
In [21]: df.head(7)
```

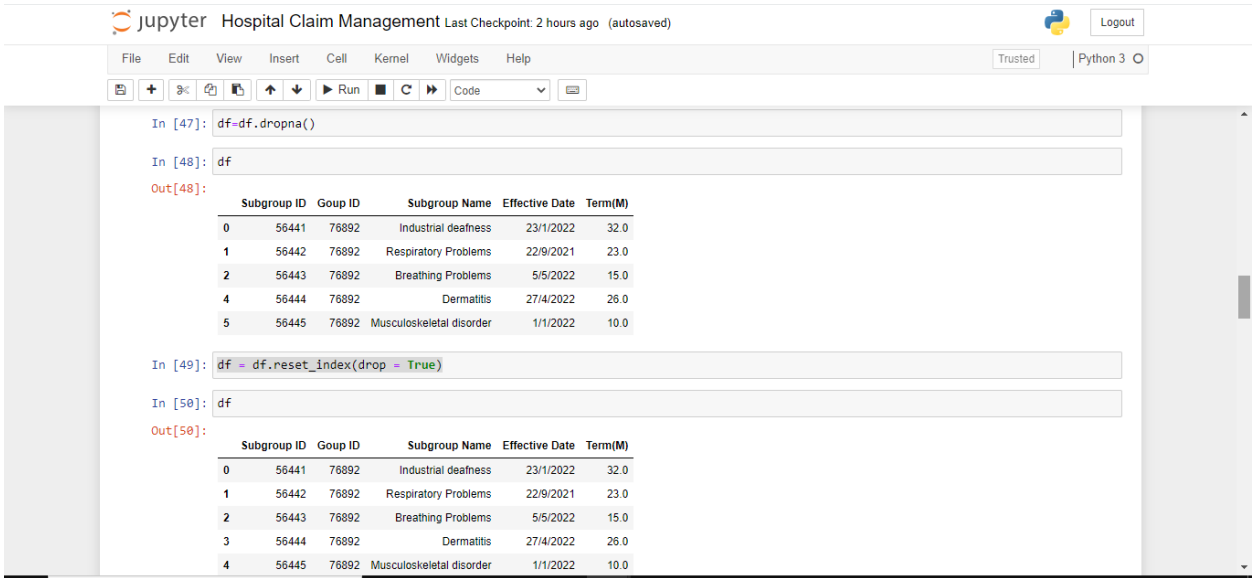
Out[21]:

	Subgroup ID	Group ID	Subgroup Name	Effective Date	Term(M)
0	56431	76891	Road-traffic accident	12/4/2021	32
1	56432	76891	Fractures caused by falls	22/9/2021	23
2	56433	76891	Gun shot	5/5/2022	15
3	56434	76891	Burn Cases	27/4/2022	26
4	56435	76891	Current Shock	1/1/2022	10

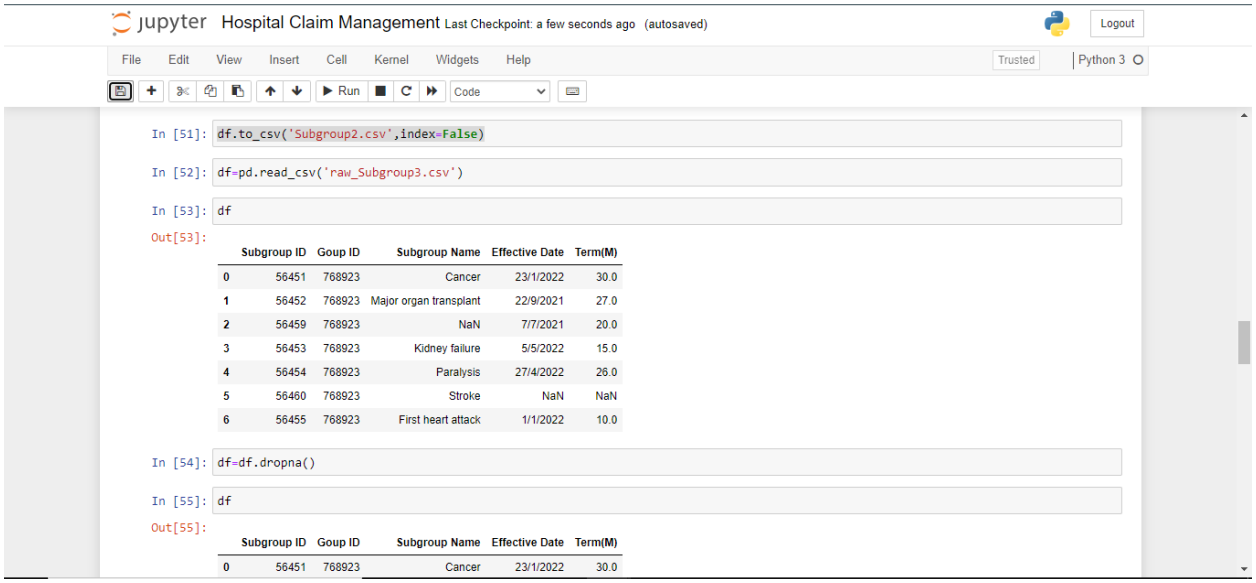


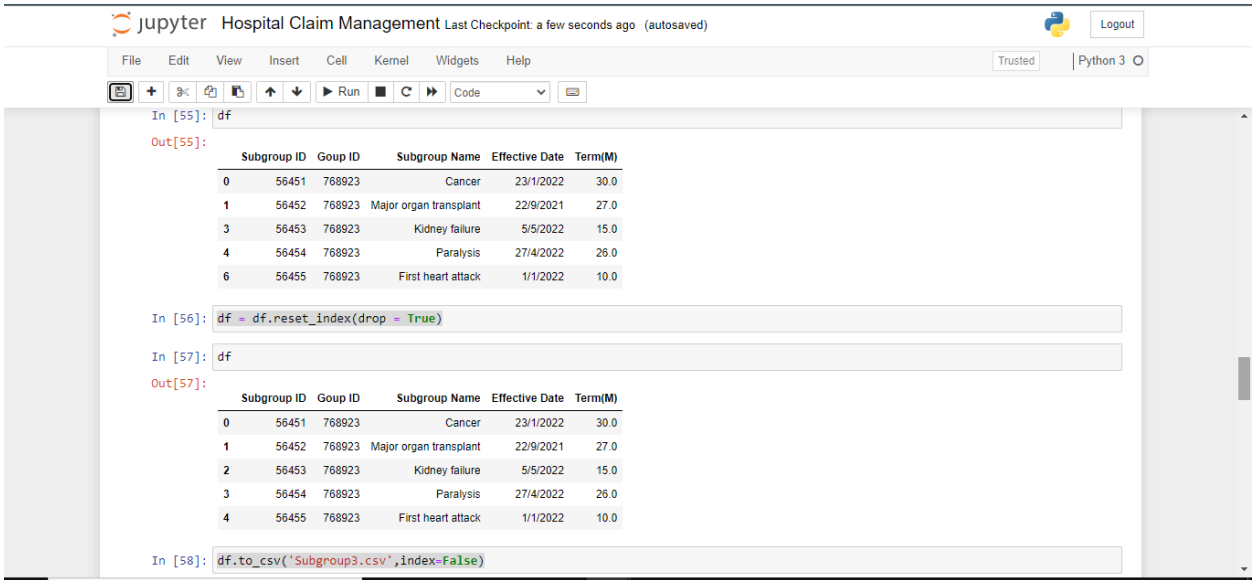
Cleansing of Subgroup2 csv file





Cleansing of Subgroup3 csv file

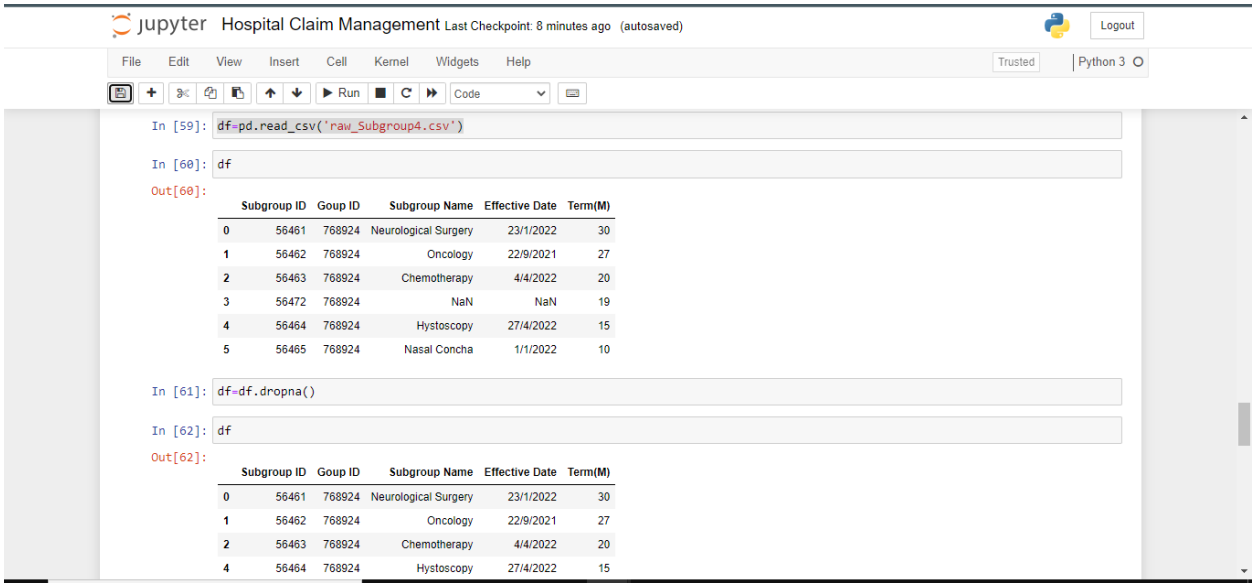




	Subgroup ID	Goup ID	Subgroup Name	Effective Date	Term(M)
0	56451	768923	Cancer	23/1/2022	30.0
1	56452	768923	Major organ transplant	22/9/2021	27.0
3	56453	768923	Kidney failure	5/5/2022	15.0
4	56454	768923	Paralysis	27/4/2022	26.0
6	56455	768923	First heart attack	1/1/2022	10.0

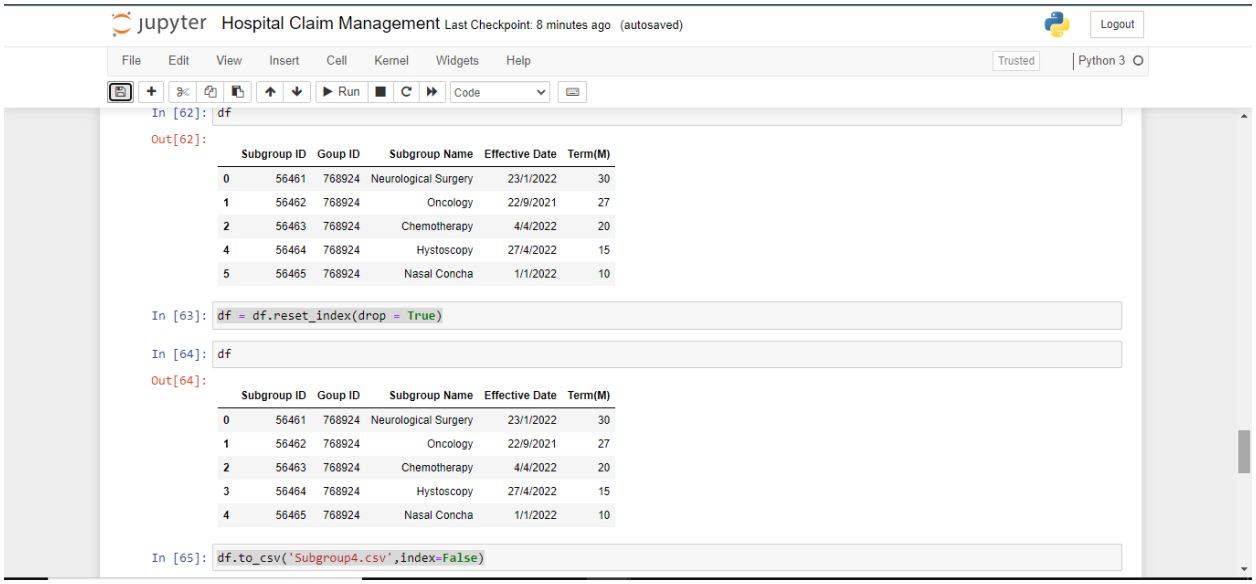
	Subgroup ID	Goup ID	Subgroup Name	Effective Date	Term(M)
0	56451	768923	Cancer	23/1/2022	30.0
1	56452	768923	Major organ transplant	22/9/2021	27.0
2	56453	768923	Kidney failure	5/5/2022	15.0
3	56454	768923	Paralysis	27/4/2022	26.0
4	56455	768923	First heart attack	1/1/2022	10.0

Cleansing of Subgroup4 csv file

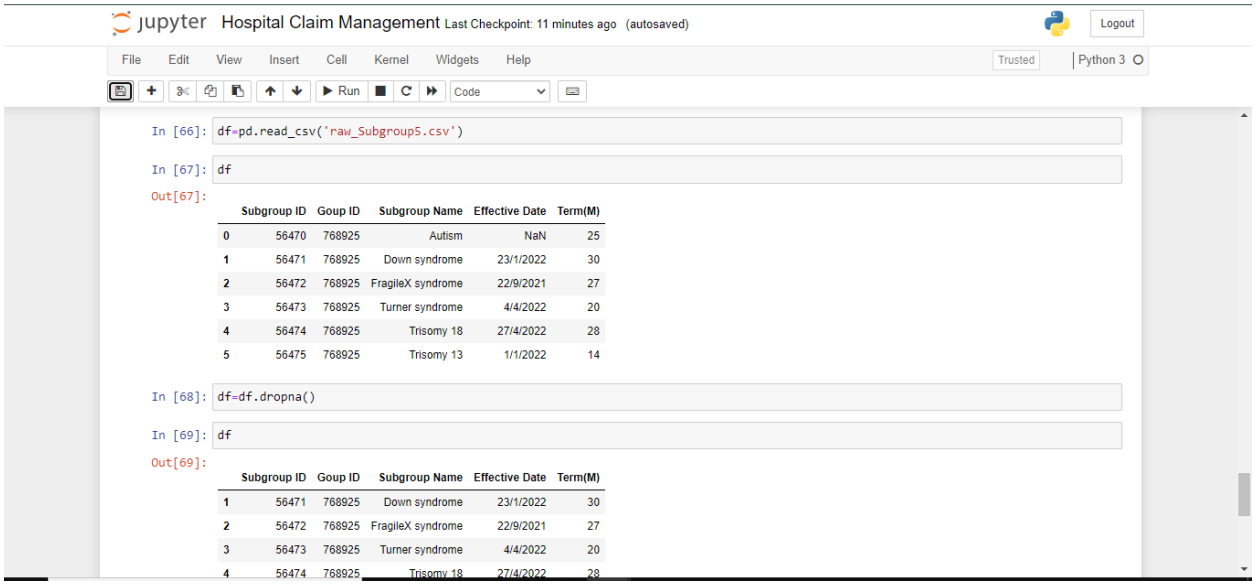


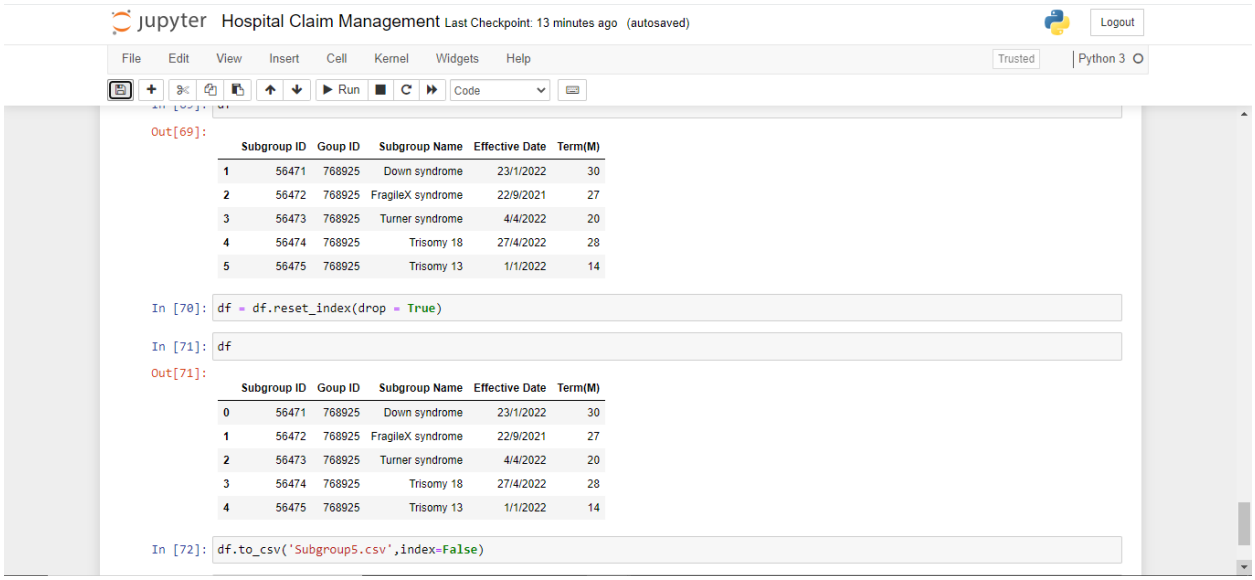
	Subgroup ID	Goup ID	Subgroup Name	Effective Date	Term(M)
0	56461	768924	Neurological Surgery	23/1/2022	30
1	56462	768924	Oncology	22/9/2021	27
2	56463	768924	Chemotherapy	4/4/2022	20
3	56472	768924	NaN	NaN	19
4	56464	768924	Hystoscopy	27/4/2022	15
5	56465	768924	Nasal Concha	1/1/2022	10

	Subgroup ID	Goup ID	Subgroup Name	Effective Date	Term(M)
0	56461	768924	Neurological Surgery	23/1/2022	30
1	56462	768924	Oncology	22/9/2021	27
2	56463	768924	Chemotherapy	4/4/2022	20
4	56464	768924	Hystoscopy	27/4/2022	15



Cleansing of Subgroup5 csv file




jupyter Hospital Claim Management Last Checkpoint: 13 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3
Run
Code
Out[69]:
Subgroup ID Goup ID Subgroup Name Effective Date Term(M)
1 56471 768925 Down syndrome 23/1/2022 30
2 56472 768925 FragileX syndrome 22/9/2021 27
3 56473 768925 Turner syndrome 4/4/2022 20
4 56474 768925 Trisomy 18 27/4/2022 28
5 56475 768925 Trisomy 13 1/1/2022 14
In [70]: df = df.reset_index(drop = True)
In [71]: df
Out[71]:
Subgroup ID Goup ID Subgroup Name Effective Date Term(M)
0 56471 768925 Down syndrome 23/1/2022 30
1 56472 768925 FragileX syndrome 22/9/2021 27
2 56473 768925 Turner syndrome 4/4/2022 20
3 56474 768925 Trisomy 18 27/4/2022 28
4 56475 768925 Trisomy 13 1/1/2022 14
In [72]: df.to_csv('Subgroup5.csv',index=False)