

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sb
```

```
In [6]: df = pd.read_csv(r'C:\Users\VARBRU\OneDrive\Desktop\FFINAL_FILES\FFINAL_FILES_CSV\FFINAL_FILES_CSV\Insurer _dummy.csv')#c
df=pd.DataFrame(df)
(df)
```

Out[6]:

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
0	1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
1	1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
2	1001	NaN	Individual Health Insurance	NaN	1/1/2036	NaN	200000.0	Active	15	200000	
3	1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
4	1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
5	1002	NaN	Family Health Insurance	3/1/2020	NaN	15/3/2020	NaN	NaN	18	300000	

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
6	1003	Birla Sun Life Insurance Co.	Critical illness Insurance	5/2/2020	5/2/2041	17/05/2020	400000.0	Active	20	150000	
7	1003	Birla Sun Life Insurance Co.	Critical illness Insurance	5/2/2020	5/2/2041	17/05/2020	400000.0	Active	20	150000	
8	1003	NaN	Critical illness Insurance	NaN	NaN	17/05/2020	NaN	Active	20	150000	



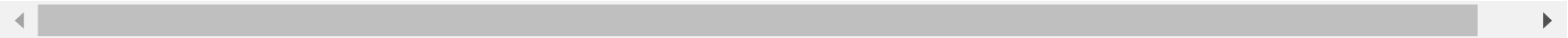
In [7]:

```
df = pd.read_csv(r'C:\Users\VARBRU\OneDrive\Desktop\FFINAL_FILES\FFINAL_FILES_CSV\FFINAL_FILES_CSV\Insurer _dummy.csv')#c
df=pd.DataFrame(df)
(df)
```

Out[7]:

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
0	1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
1	1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
2	1001	NaN	Individual Health Insurance	NaN	1/1/2036	NaN	200000.0	Active	15	200000	

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
3	1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
4	1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
5	1002	NaN	Family Health Insurance	3/1/2020	NaN	15/3/2020	NaN	NaN	18	300000	
6	1003	Birla Sun Life Insurance Co.	Critical illness Insurance	5/2/2020	5/2/2041	17/05/2020	400000.0	Active	20	150000	
7	1003	Birla Sun Life Insurance Co.	Critical illness Insurance	5/2/2020	5/2/2041	17/05/2020	400000.0	Active	20	150000	
8	1003	NaN	Critical illness Insurance	NaN	NaN	17/05/2020	NaN	Active	20	150000	



```
In [8]: df.isnull() #checking the null values
```

Out[8]:

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
0	False	False	False	False	False	False	False	False	False	False	
1	False	False	False	False	False	False	False	False	False	False	
2	False	True	False	True	False	True	False	False	False	False	

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
3	False	False	False	False	False	False	False	False	False	False	
4	False	False	False	False	False	False	False	False	False	False	
5	False	True	False	False	True	False	True	True	False	False	
6	False	False	False	False	False	False	False	False	False	False	
7	False	False	False	False	False	False	False	False	False	False	
8	False	True	False	True	True	False	True	False	False	False	

In [9]:

```
duplicate=df[df.duplicated()]
duplicate
```

Out[9]:

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
1	1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
4	1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
7	1003	Birla Sun Life Insurance Co.	Critical illness Insurance	5/2/2020	5/2/2041	17/05/2020	400000.0	Active	20	150000	

In [10]:

```
remove_duplicates=df.drop_duplicates()
remove_duplicates #drop duplicates
```

Out[10]:

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
0	1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
2	1001	NaN	Individual Health Insurance	NaN	1/1/2036	NaN	200000.0	Active	15	200000	
3	1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
5	1002	NaN	Family Health Insurance	3/1/2020	NaN	15/3/2020	NaN	NaN	18	300000	
6	1003	Birla Sun Life Insurance Co.	Critical illness Insurance	5/2/2020	5/2/2041	17/05/2020	400000.0	Active	20	150000	
8	1003	NaN	Critical illness Insurance	NaN	NaN	17/05/2020	NaN	Active	20	150000	



In [13]:

```
drop_null=remove_duplicates.dropna()  
drop_null #Drop Null Values
```

Out[13]:

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
0	1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	

	Insurance ID	Company Name	Insurance Name	Start Date(dd/mm/yyyy)	End Date(dd/mm/yyyy)	Effective date(dd/mm/yyyy)	Insurance Base Amount	Insurance Status	Term(Yrs)	TopUp Amount	Insur C
3	1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
6	1003	Birla Sun Life Insurance Co.	Critical illness Insurance	5/2/2020	5/2/2041	17/05/2020	400000.0	Active	20	150000	



```
In [14]: drop_null.to_csv('Insurer_Final')
```

```
In [ ]:
```