import pandas as pd
import numpy as np
import matplotlib.pyplot as plp
import seaborn as sb

In [6]:
 df = pd.read\_csv(r'C:\Users\VARBRU\OneDrive\Desktop\FFINAL\_FILES\_CSV\FFINAL\_FILES\_CSV\Insurer \_dummy.csv')#d
 df=pd.DataFrame(df)
 (df)

$\cap$	rt I	6	
$\cup$ $\cup$		U	
			4

	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
(	<b>)</b> 1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
1	<b>I</b> 1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
2	2 1001	NaN	Individual Health Insurance	NaN	1/1/2036	NaN	200000.0	Active	15	200000	
3	<b>3</b> 1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
4	<b>1</b> 1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
į	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	
4											•

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	
4											•

In [8]:

df.isnull() #checking the null values

Out[8]:

Insura	ince 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
<b>0</b> F	alse	False	False	False	False	False	False	False	False	False	
<b>1</b> F	alse	False	False	False	False	False	False	False	False	False	
<b>2</b> F	alse	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	
4											•

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
1001	Aegon Life Insurance Co. Ltd.	Individual Health Insurance	1/1/2020	1/1/2036	15/1/2020	200000.0	Active	15	200000	
1002	Edelweiss Life Insurance Co.	Family Health Insurance	3/1/2020	3/1/2039	15/3/2020	300000.0	Active	18	300000	
1003	Birla Sun Life Insurance Co.	Critical illness Insurance	5/2/2020	5/2/2041	17/05/2020	400000.0	Active	20	150000	
	1001 1002	Aegon Life Insurance Co. Ltd. Edelweiss Life Insurance Co. Birla Sun Life Insurance	Aegon Life Insurance Co. Ltd.  Edelweiss Life Insurance Co. Birla Sun Life Insurance Co. Birla Sun Life Insurance Co. Critical illness Insurance	Aegon Life Insurance Co. Ltd.  Edelweiss Life Insurance Co.  Birla Sun Life Insurance Co. Life Insurance Co. Life Insurance Co. Life Insurance Co.  Birla Sun Life Insurance Life Insurance Co.  Critical illness Insurance	IDNameNameDate(dd/mm/yyyy)Date(dd/mm/yyyy)1001Aegon Life Insurance Co. Ltd.Individual Health Insurance1/1/20201/1/20361002Edelweiss Life Insurance Co.Family Health Insurance Insurance3/1/20203/1/20391003Birla Sun Life Insurance InsuranceCritical illness Insurance5/2/20205/2/2041	IDNameNameDate(dd/mm/yyyy)Date(dd/mm/yyyy)date(dd/mm/yyyy)1001Aegon Life Insurance Co. Ltd.Individual Health Insurance1/1/20201/1/203615/1/20201002Edelweiss Life Insurance Co.Family Health Insurance Co.3/1/203915/3/20201003Birla Sun Life Insurance Insurance InsuranceCritical illness Insurance Insurance5/2/20205/2/204117/05/2020	Insurance ID Name Name Name Date(dd/mm/yyyy) Date(dd/mm/yyyy) date(dd/mm/yyyy) Base Amount  Aegon Life Insurance Co. Ltd.  Edelweiss Life Insurance Co. Ltd.  Edelweiss Co. Life Insurance Co. Size Co. Life Insurance Co. Life Insurance Co. Size Co. Size Co. Life Insurance Insurance Co. Size	Insurance ID Name Name Date(dd/mm/yyyy)	Insurance ID Name Name Name Date(dd/mm/yyyy) Date(dd/mm/yyyy) date(dd/mm/yyyy) Amount  Aegon Life Insurance Co. Ltd.  Edelweiss Life Insurance Co. Life Insurance Insu	Insurance Company Name Name Name Date(dd/mm/yyyy) Date(dd/mm/yyyy) date(dd/mm/yyyy) Ramount  Aegon Life Insurance Co. Ltd.  Edelweiss Life Insurance Co.  Local Life Insurance Co.  Life Insurance Insurance C

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	
4											•

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
1	<b>0</b> 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		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	
	4											•
In [14]:	dr	rop_null.	to_csv('Ir	surer_Fin	al')							
In [ ]:												