In [2]:

```
import datetime
import matplotlib.pyplot as plt
from matplotlib import gridspec
import matplotlib.cm as cm
import numpy as np
import pandas as pd
import seaborn as sns
plt.rcParams["figure.figsize"] = (10,5)

import warnings
sns.set_style("whitegrid")
warnings.filterwarnings('ignore')
```

In [3]:

```
#Loading the data from csv files.
df=pd.read_csv('Rilders_List2020-2.csv')
df.head()
```

Out[3]:

	LeadID	State	City	RiderName	LL EMP Code	RiderNumber	DOJ	Client	
0	31	Karnataka	Bengaluru	SRIKANTH P	LLBBC001	6363376901	04-04-2019	Bigbasket Bangalore	
1	32	Karnataka	Bengaluru	SUPREEM .	LLBBC003	7483505921	4/16/2019	Bigbasket Bangalore	
2	33	Tamil Nadu	Chennai	Silambarasan V	LLBBC004	8695084040	03-01-2019	Bigbasket Chennai	C
3	35	Karnataka	Bengaluru	VEERESH U	LLBBC011	9110444430	6/23/2019	Bigbasket Bangalore	
4	37	Karnataka	Bengaluru	Harish M	LLBBC027	9008816586	6/17/2019	Bigbasket Bangalore	

In [4]:

```
df.shape
```

Out[4]:

```
(7349, 15)
```

In [5]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7349 entries, 0 to 7348
Data columns (total 15 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   LeadID                7349 non-null   int64  
 1   State                 7299 non-null   object 
 2   City                  7299 non-null   object 
 3   RiderName             7349 non-null   object 
 4   LL EMP Code           7344 non-null   object 
 5   RiderNumber           7349 non-null   object 
 6   DOJ                   7257 non-null   object 
 7   Client                7299 non-null   object 
 8   Hub                   7299 non-null   object 
 9   Previous_Status       7295 non-null   object 
 10  Current_Status        7294 non-null   object 
 11  Status_Changed        7294 non-null   object 
 12  Modified_by           7294 non-null   object 
 13  ReasonForLeaving      5737 non-null   float64
 14  RelevingDate_Entered  5769 non-null   object 
dtypes: float64(1), int64(1), object(13)
memory usage: 861.3+ KB
```

In [6]:

```
df.isnull().sum()
```

Out[6]:

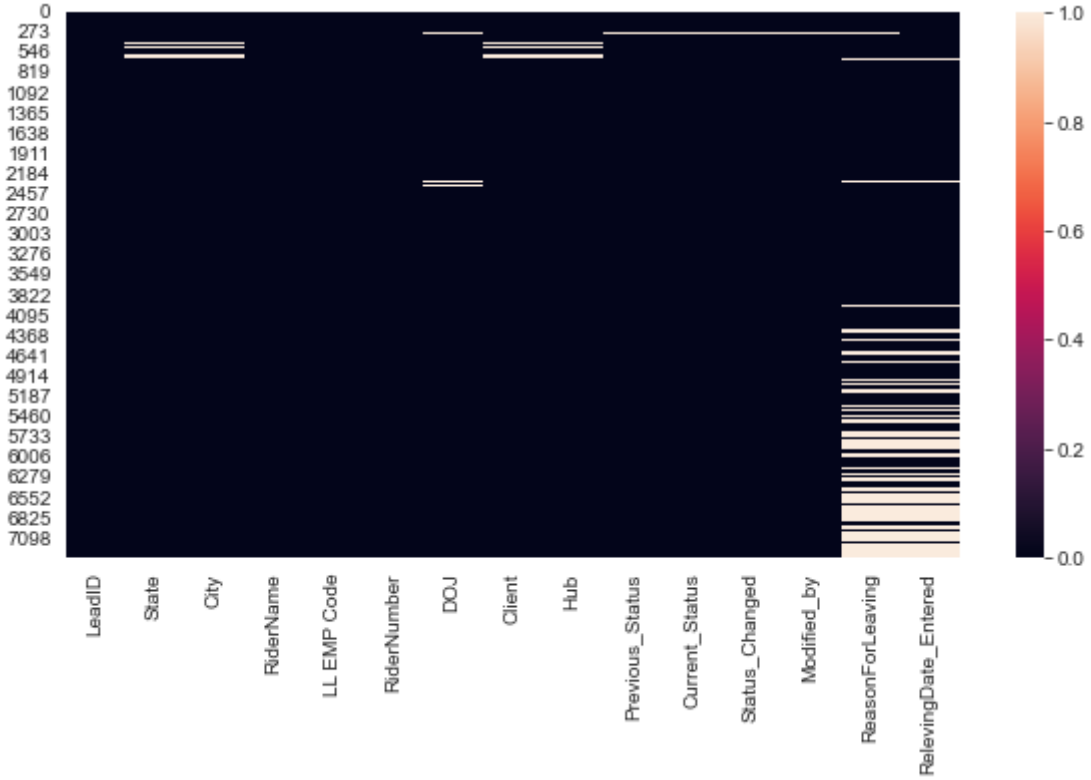
```
LeadID                   0
State                   50
City                    50
RiderName                0
LL EMP Code              5
RiderNumber              0
DOJ                     92
Client                  50
Hub                     50
Previous_Status         54
Current_Status          55
Status_Changed          55
Modified_by             55
ReasonForLeaving      1612
RelevingDate_Entered  1580
dtype: int64
```

In [7]:

```
sns.heatmap(df.isnull())
```

Out[7]:

```
<AxesSubplot:>
```



In [8]:

```
df1=df.copy()
```

In [9]:

```
df1.shape
```

Out[9]:

```
(7349, 15)
```

In [10]:

```
# def format_color_groups(df1):
#     colors = ['violet', 'pink']
#     x = df1.copy()
#     factors = list(x['RiderNumber'].unique())
#     i = 0
#     for factor in factors:
#         style = f'background-color: {colors[i]}'
#         x.loc[x['RiderNumber'] == factor, :] = style
#         i = not i
#     return x
```

In [11]:

```
# df1.style.apply(format_color_groups, axis=None)
```

In [12]:

```
df1=df1.drop_duplicates(subset='RiderNumber').copy()
```

In [13]:

```
df1.head()
```

Out[13]:

	LeadID	State	City	RiderName	LL EMP Code	RiderNumber	DOJ	Client	
0	31	Karnataka	Bengaluru	SRIKANTH P	LLBBC001	6363376901	04-04-2019	Bigbasket Bangalore	
1	32	Karnataka	Bengaluru	SUPREEM .	LLBBC003	7483505921	4/16/2019	Bigbasket Bangalore	
2	33	Tamil Nadu	Chennai	Silambarasan V	LLBBC004	8695084040	03-01-2019	Bigbasket Chennai	C
3	35	Karnataka	Bengaluru	VEERESH U	LLBBC011	9110444430	6/23/2019	Bigbasket Bangalore	
4	37	Karnataka	Bengaluru	Harish M	LLBBC027	9008816586	6/17/2019	Bigbasket Bangalore	

In [15]:

```
df1.shape
```

Out[15]:

```
(6685, 15)
```

In [13]:

```
df1[df1['DOJ'] == '0000-00-00'].shape
```

Out[13]:

```
(1093, 15)
```

In [14]:

```
df1=df1.drop([2,1349,1428,4400,5748],axis=0).copy()
df1.head()
```

Out[14]:

	LeadID	State	City	RiderName	LL EMP Code	RiderNumber	DOJ	Client	
0	31	Karnataka	Bengaluru	SRIKANTH P	LLBBC001	6363376901	04-04-2019	Bigbasket Bangalore	B H
1	32	Karnataka	Bengaluru	SUPREEM .	LLBBC003	7483505921	4/16/2019	Bigbasket Bangalore	B H
3	35	Karnataka	Bengaluru	VEERESH U	LLBBC011	9110444430	6/23/2019	Bigbasket Bangalore	B H
4	37	Karnataka	Bengaluru	Harish M	LLBBC027	9008816586	6/17/2019	Bigbasket Bangalore	B H
5	38	Karnataka	Bengaluru	ANANDH .	LLBBC028	9095272783	2/26/2019	TOW Bangalore	HS

In [15]:

```
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 6680 entries, 0 to 7348
Data columns (total 15 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   LeadID                6680 non-null   int64  
 1   State                 6631 non-null   object 
 2   City                  6631 non-null   object 
 3   RiderName             6680 non-null   object 
 4   LL EMP Code           6675 non-null   object 
 5   RiderNumber           6680 non-null   object 
 6   DOJ                   6588 non-null   object 
 7   Client                6631 non-null   object 
 8   Hub                   6631 non-null   object 
 9   Previous_Status       6626 non-null   object 
 10  Current_Status        6625 non-null   object 
 11  Status_Changed        6625 non-null   object 
 12  Modified_by           6625 non-null   object 
 13  ReasonForLeaving      5111 non-null   float64
 14  RelevingDate_Entered  5143 non-null   object 
dtypes: float64(1), int64(1), object(13)
memory usage: 835.0+ KB
```

In [16]:

```
df1['DOJ'] = pd.to_datetime(df1['DOJ'],errors='coerce')
```

In [17]:

```
df1['RelevingDate_Entered'] = pd.to_datetime(df1['RelevingDate_Entered'], errors='coerce')
```

In [18]:

```
df1['Status_Changed'] = pd.to_datetime(df1['Status_Changed'], errors='ignore')
```

In [19]:

```
df1.sample(5)
```

Out[19]:

	LeadID	State	City	RiderName	LL EMP Code	RiderNumber	DOJ	Client	
4539	21767	Karnataka	Bengaluru	Anita Shekhar Deb	10056	8340231034	2022-01-21	Bigbasket Bangalore	Bi
5043	22950	Tamil Nadu	Chennai	Raghu Nathan	10809	7200173245	2022-02-16	Shadowfax E-Commerce (Chennai)	
2196	15309	Karnataka	Bengaluru	Gopinath M B	6771	6301703944	2022-03-15	1kart ROK	Chik
719	8108	Karnataka	Bengaluru	B M VENUKUMAR	2650	8861400676	NaT	TOW Bangalore	
2595	16753	Telangana	Hyderabad	Md Javeed	7291	9951783205	2021-10-10	Bigbasket Hyderabad	

In [20]:

```
df1.isna().sum()
```

Out[20]:

```
LeadID                0
State                49
City                 49
RiderName             0
LL EMP Code           5
RiderNumber           0
DOJ                1187
Client               49
Hub                  49
Previous_Status      54
Current_Status       55
Status_Changed       55
Modified_by          55
ReasonForLeaving   1569
RelevingDate_Entered 3612
dtype: int64
```

In [21]:

```
df1.isna().any(axis=1).value_counts()
df1.isna().sum()
```

Out[21]:

```
LeadID                   0
State                   49
City                    49
RiderName                0
LL EMP Code              5
RiderNumber              0
DOJ                   1187
Client                  49
Hub                     49
Previous_Status         54
Current_Status          55
Status_Changed          55
Modified_by             55
ReasonForLeaving      1569
RelevingDate_Entered  3612
dtype: int64
```

In [22]:

```
df2=df1.dropna().copy()
df2.head()
```

Out[22]:

	LeadID	State	City	RiderName	LL EMP Code	RiderNumber	DOJ	Client	
0	31	Karnataka	Bengaluru	SRIKANTH P	LLBBC001	6363376901	2019-04-04	Bigbasket Bangalore	Bigba Hulim
1	32	Karnataka	Bengaluru	SUPREEM .	LLBBC003	7483505921	2019-04-16	Bigbasket Bangalore	Bigba Hulim
6	39	Karnataka	Bengaluru	Deepak Badiya	LLBBC029	9382654318	2022-05-10	BB_Now Bangalore	BBI E Bellai
8	46	Karnataka	Bengaluru	Naveen Kumar V	LLBBC088	7349716619	2019-08-07	Bigbasket Bangalore	Bigba Siddaj
10	48	Karnataka	Bengaluru	BANAJ KUMAR SAHOO .	LLBBA011	9337077633	2019-07-04	Bigbasket Bangalore	Bigba Siddaj

In [23]:

```
df2.shape
```

Out[23]:

```
(2673, 15)
```

In [24]:

```
df2['no_of_working_days']=(df2['RelevingDate_Entered']-df2['DOJ']).dt.days
```

In [25]:

```
df2.head()
```

Out[25]:

	LeadID	State	City	RiderName	LL EMP Code	RiderNumber	DOJ	Client	
0	31	Karnataka	Bengaluru	SRIKANTH P	LLBBC001	6363376901	2019-04-04	Bigbasket Bangalore	Bigba Hulir
1	32	Karnataka	Bengaluru	SUPREEM .	LLBBC003	7483505921	2019-04-16	Bigbasket Bangalore	Bigba Hulir
6	39	Karnataka	Bengaluru	Deepak Badiya	LLBBC029	9382654318	2022-05-10	BB_Now Bangalore	BBI E Bellar
8	46	Karnataka	Bengaluru	Naveen Kumar V	LLBBC088	7349716619	2019-08-07	Bigbasket Bangalore	Bigba Siddaj
10	48	Karnataka	Bengaluru	BANAJ KUMAR SAHOO .	LLBBA011	9337077633	2019-07-04	Bigbasket Bangalore	Bigba Siddaj

In [26]:

```
df2[df2['no_of_working_days'] <= 0]
```

Out[26]:

	LeadID	State	City	RiderName	LL EMP Code	RiderNumber	DOJ	Client
16	58	Karnataka	Bengaluru	Rudresh R	LLBBA066	8660622218	2021-10-20	Grofers Bangalore
53	115	Tamil Nadu	Chennai	Maria Victas .	LLC0115	9940274285	2021-07-22	Bigbasket Chennai
60	132	Telangana	Hyderabad	Nitin s	LLB0040	7483388888	2022-03-18	Bigbasket Hyderabad
69	147	Tamil Nadu	Chennai	VIGNESH D	LLC0221	9087902184	2021-01-18	Bigbasket Chennai
72	157	Tamil Nadu	Chennai	Dineshkumar R	LLC0294	7339697155	2022-04-03	Bigbasket Chennai
...	...	...	...	...	...	...	...	...
7045	27084	Karnataka	Bengaluru	Devana Dharaneedhar Reddy	13480	7348867507	2022-05-31	1kart Bangalore
7162	27303	Tamil Nadu	Chennai	Karthikeyan K	13723	9360139369	2022-06-07	Bigbasket Chennai
7163	27304	Tamil Nadu	Chennai	S.Gopi Krishnan	13724	8939247195	2022-06-07	Bigbasket Chennai
7167	27309	Karnataka	Bengaluru	Srihari R	13634	9108097481	2022-07-04	BB_Now Bangalore
7180	27337	Tamil Nadu	Chennai	B.Aravinda Pandian	13662	9659075005	2022-06-07	Bigbasket Chennai

250 rows × 16 columns

In [27]:

```
df2= df2[df2['no_of_working_days'] >= 0]
```

In [28]:

```
df3= df2.drop(['LeadID','State','City','LL EMP Code','RiderNumber',
    'Client','Hub','Modified_by','ReasonForLeaving'], axis=1).copy()
df3.head()
```

Out[28]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entered
0	SRIKANTH P	2019-04-04	Active	Terminated	2020-05-19	2020-04-02
1	SUPREEM .	2019-04-16	Active	Resigned	2020-05-19	2020-02-20
6	Deepak Badiya	2022-05-10	Active	Terminated	2022-05-14	2022-05-14
8	Naveen Kumar V	2019-08-07	Active	Drop-outs	2022-02-10	2022-02-10
10	BANAJ KUMAR SAHOO .	2019-07-04	Active	Terminated	2020-07-20	2020-04-29

In [29]:

```
df3.shape
```

Out[29]:

```
(2479, 7)
```

In [30]:

```
df3.head()
```

Out[30]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entered
0	SRIKANTH P	2019-04-04	Active	Terminated	2020-05-19	2020-04-02
1	SUPREEM .	2019-04-16	Active	Resigned	2020-05-19	2020-02-20
6	Deepak Badiya	2022-05-10	Active	Terminated	2022-05-14	2022-05-14
8	Naveen Kumar V	2019-08-07	Active	Drop-outs	2022-02-10	2022-02-10
10	BANAJ KUMAR SAHOO .	2019-07-04	Active	Terminated	2020-07-20	2020-04-29

In [31]:

```
max(df3['no_of_working_days'])
```

Out[31]:

```
976
```

In [32]:

```
sum(df3['no_of_working_days'])
```

Out[32]:

```
170710
```

In [33]:

```
df3['Riders_Joined_by_year'] = pd.DatetimeIndex(df3['DOJ']).year
df3['Riders_Quitted_by_year'] = pd.DatetimeIndex(df3['RelevingDate_Entered']).year
```

In [34]:

```
df3.head()
```

Out[34]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entered
0	SRIKANTH P	2019-04-04	Active	Terminated	2020-05-19	2020-04-02
1	SUPREEM .	2019-04-16	Active	Resigned	2020-05-19	2020-02-20
6	Deepak Badiya	2022-05-10	Active	Terminated	2022-05-14	2022-05-14
8	Naveen Kumar V	2019-08-07	Active	Drop-outs	2022-02-10	2022-02-10
10	BANAJ KUMAR SAHOO .	2019-07-04	Active	Terminated	2020-07-20	2020-04-29

In [35]:

```
df3[df3['Current_Status']=='Drop-outs'].sample(5)
```

Out[35]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entere
641	Saravana Muthu Nagarajan	2020-09-21	Active	Drop-outs	2022-02-10	2021-12-1
1978	Shahrukh Khan	2021-08-18	Active	Drop-outs	2022-01-28	2022-01-0
2753	MAIBU SABU	2021-10-14	Active	Drop-outs	2022-03-16	2022-03-1
6113	Jagadish Jaga	2022-04-27	Active	Drop-outs	2022-05-30	2022-05-1
5381	Balaji E	2022-05-13	Active	Drop-outs	2022-06-08	2022-05-3

In [36]:

```
df3[df3['Current_Status'] == 'Active'].sample(5)
```

Out[36]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Ent
5631	Kushal B.M	2022-03-22	Undertraining	Active	2022-03-31	2022-(
4899	Ashok Kumar R	2022-02-10	Undertraining	Active	2022-02-10	2022-(
4416	Nagaraju V	2022-01-21	Undertraining	Active	2022-01-31	2022-(
2976	Sharanabasapp	2021-11-01	Undertraining	Active	2021-11-01	2022-(
2014	Abhinandan C R	2021-08-16	Undertraining	Active	2021-08-18	2022-(

In [37]:

```
df3['Current_Status'].value_counts()
```

Out[37]:

```
Drop-outs     2158
Resigned       184
Active          88
Terminated      49
Name: Current_Status, dtype: int64
```

In [38]:

```
df3['Current_Status'].unique()
```

Out[38]:

```
array(['Terminated', 'Resigned', 'Drop-outs', 'Active'], dtype=object)
```

In [39]:

```
np.array(pd.Categorical(df3['Current_Status']).categories)
```

Out[39]:

```
array(['Active', 'Drop-outs', 'Resigned', 'Terminated'], dtype=object)
```

In [40]:

```
def stats(a):
    mean=a['no_of_working_days'].mean()
    median=a['no_of_working_days'].median()
    mode=a['no_of_working_days'].mode()[0]
    return [mean,mode,median]
```

In [41]:

```
df2['Riders_Joined_by_month'] = pd.DatetimeIndex(df3['DOJ']).month
df2['Riders_Quitted_by_month'] = pd.DatetimeIndex(df3['RelevingDate_Entered']).month
df2['Riders_Joined_by_year'] = pd.DatetimeIndex(df3['DOJ']).year
df2['Riders_Quitted_by_year'] = pd.DatetimeIndex(df3['RelevingDate_Entered']).year
```

In [42]:

```
df2.head()
```

Out[42]:

	LeadID	State	City	RiderName	LL EMP Code	RiderNumber	DOJ	Client	
0	31	Karnataka	Bengaluru	SRIKANTH P	LLBBC001	6363376901	2019-04-04	Bigbasket Bangalore	Bigba Hulir
1	32	Karnataka	Bengaluru	SUPREEM .	LLBBC003	7483505921	2019-04-16	Bigbasket Bangalore	Bigba Hulir
6	39	Karnataka	Bengaluru	Deepak Badiya	LLBBC029	9382654318	2022-05-10	BB_Now Bangalore	BBI E Bellaı
8	46	Karnataka	Bengaluru	Naveen Kumar V	LLBBC088	7349716619	2019-08-07	Bigbasket Bangalore	Bigba Siddaı
10	48	Karnataka	Bengaluru	BANAJ KUMAR SAHOO .	LLBBA011	9337077633	2019-07-04	Bigbasket Bangalore	Bigba Siddaı

In [43]:

```
stats(df2)
```

Out[43]:

```
[68.86244453408632, 7, 32.0]
```

In [44]:

```
stats(df2[df2['Riders_Joined_by_year']==2019])
```

Out[44]:

```
[273.8, 81, 187.0]
```

In [45]:

```
stats(df2[df2['Riders_Joined_by_year']==2020])
```

Out[45]:

```
[378.675, 0, 473.5]
```

In [46]:

```
stats(df2[df2['Riders_Joined_by_year']==2021])
```

Out[46]:

```
[124.5506419400856, 137, 112.0]
```

In [47]:

```
stats(df2[df2['Riders_Joined_by_year']==2022])
```

Out[47]:

```
[27.651524208009565, 7, 20.0]
```

In [48]:

```
stats(df2[df2['Riders_Quitted_by_year']==2020])
```

Out[48]:

```
[103.2, 0, 69.0]
```

In [49]:

```
stats(df2[df2['Riders_Quitted_by_year']==2021])
```

Out[49]:

```
[147.6764705882353, 11, 64.0]
```

In [50]:

```
stats(df2[df2['Riders_Quitted_by_year']==2022])
```

Out[50]:

```
[64.6699613235926, 7, 31.0]
```

For 2019-2022

No of joiner and quitter by Year wise

In [51]:

```
df_joinbyyear=pd.DataFrame(df3.groupby(pd.Grouper(key='DOJ', freq='Y')).size().sort_index(a
df_joinbyyear.reset_index(inplace=True)
```

In [52]:

```
df_joinbyyear
```

Out[52]:

	DOJ	0
0	2019-12-31	25
1	2020-12-31	80
2	2021-12-31	701
3	2022-12-31	1673

In [53]:

```
df_quitbyyear=pd.DataFrame(df3.groupby(pd.Grouper(key='RelevingDate_Entered', freq='Y')).si
```

In [54]:

```
df_quitbyyear.reset_index(inplace=True)
```

In [55]:

```
df_quitbyyear
```

Out[55]:

	RelevingDate_Entered	0
0	2020-12-31	50
1	2021-12-31	102
2	2022-12-31	2327

In [56]:

```
df_cumulative_year19=df_joinbyyear.merge(df_quitbyyear, left_on='DOJ', right_on='RelevingDa
df_cumulative_year19
```

Out[56]:

	DOJ	0_x	RelevingDate_Entered	0_y
0	2019-12-31	25	NaT	NaN
1	2020-12-31	80	2020-12-31	50.0
2	2021-12-31	701	2021-12-31	102.0
3	2022-12-31	1673	2022-12-31	2327.0

In [57]:

```
df_cumulative_year19=df_cumulative_year19.drop('RelevingDate_Entered',axis=1)
df_cumulative_year19= df_cumulative_year19.rename(columns={'0_x':'joiners', '0_y':'quitters
```

In [58]:

```
df_cumulative_year19=df_cumulative_year19.rename(columns={'DOJ':'Month'})
```

In [59]:

```
df_cumulative_year19
```

Out[59]:

	Month	joiners	quitters
0	2019-12-31	25	NaN
1	2020-12-31	80	50.0
2	2021-12-31	701	102.0
3	2022-12-31	1673	2327.0

In [60]:

```
df_cumulative_year19.fillna(0, inplace=True)
```

In [61]:

```
df_cumulative_year19.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 4 entries, 0 to 3
Data columns (total 3 columns):
 #   Column    Non-Null Count  Dtype         
---  ------    --------------  -----         
 0   Month     4 non-null      datetime64[ns]
 1   joiners   4 non-null      int64         
 2   quitters  4 non-null      float64       
dtypes: datetime64[ns](1), float64(1), int64(1)
memory usage: 128.0 bytes
```

In [62]:

```
df_cumulative_year19['joiners'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[62]:

```
count         4.00
mean        619.75
median      390.50
<lambda>     25.00
Name: joiners, dtype: float64
```

In [63]:

```
df_cumulative_year19['quitters'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[63]:

```
count         4.00
mean        619.75
median       76.00
<lambda>      0.00
Name: quitters, dtype: float64
```

In [64]:

```
df_cumulative_year19['quitters']=df_cumulative_year19['quitters']*-1
```

In [65]:

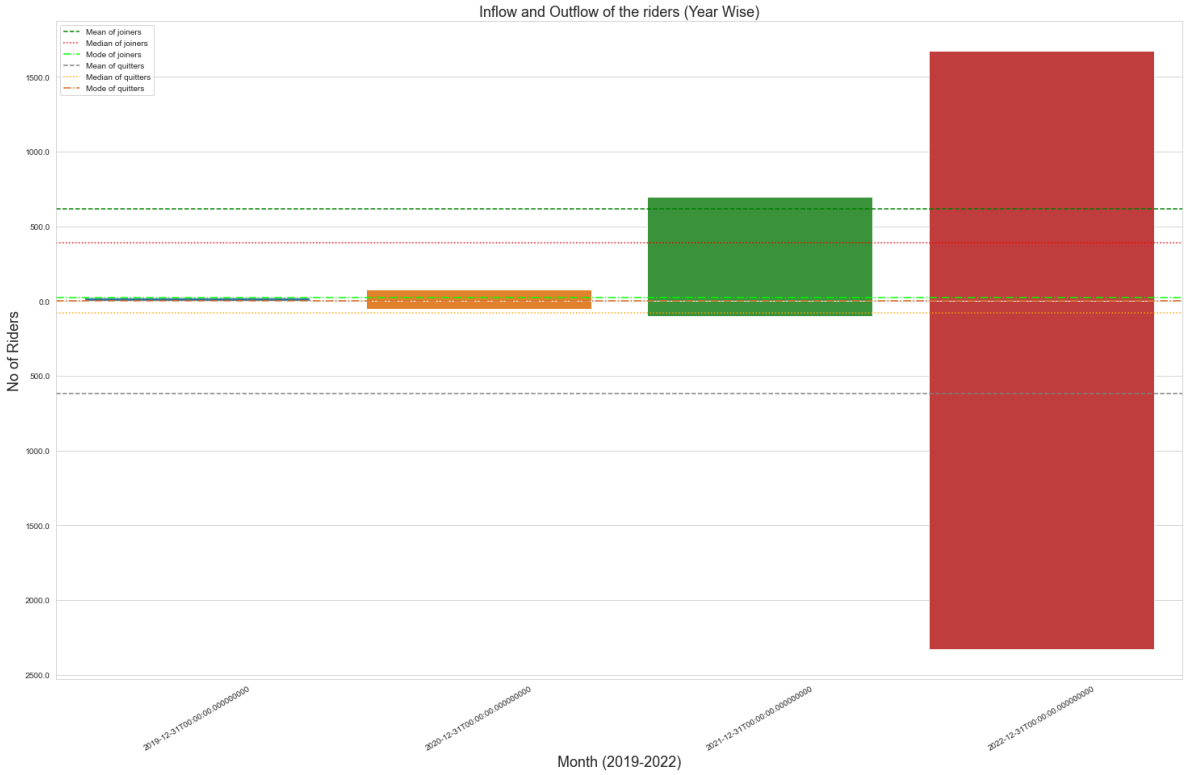
```
df_cumulative_year19
```

Out[65]:

	Month	joiners	quitters
0	2019-12-31	25	-0.0
1	2020-12-31	80	-50.0
2	2021-12-31	701	-102.0
3	2022-12-31	1673	-2327.0

In [66]:

```
plt.figure(figsize=(25,15))
bar_plot=sns.barplot(data=df_cumulative_year19, x=df_cumulative_year19['Month'], y=df_cumul
            dodge=False)
bar_plot=sns.barplot(data=df_cumulative_year19, x=df_cumulative_year19['Month'], y=df_cumul
            dodge=False)
plt.xticks(rotation=30)
plt.ylabel('No of Riders',fontsize=18)
plt.xlabel('Month (2019-2022)',fontsize=18)
plt.title(label="Inflow and Outflow of the riders (Year Wise)",fontsize=18)
bar_plot.axhline(y =619.75,color ="green", linestyle ="--",label='Mean of joiners')
bar_plot.axhline(y = 390.50,color ="red", linestyle =":",label='Median of joiners')
bar_plot.axhline(y = 25,color ="lime", linestyle ="-.",label='Mode of joiners')
bar_plot.axhline(y =-619.75,color ="grey", linestyle ="--",label='Mean of quitters')
bar_plot.axhline(y = -76.00,color ="orange", linestyle =":",label='Median of quitters')
bar_plot.axhline(y = 0,color ="chocolate", linestyle ="-.",label='Mode of quitters')
bar_plot.set_yticklabels([str(abs(bar_plot)) for bar_plot in bar_plot.get_yticks()])
plt.legend()
plt.show()
```



No of joiner and quitter by Month wise

In [67]:

```
df_joinbymonth=pd.DataFrame(df3.groupby(pd.Grouper(key='DOJ', freq='M')).size().sort_index(
df_joinbymonth.reset_index(inplace=True)
```

In [68]:

```
df_quitbymonth=pd.DataFrame(df3.groupby(pd.Grouper(key='RelevingDate_Entered', freq='M')).s
```

In [69]:

```
df_quitbymonth.reset_index(inplace=True)
```

In [70]:

```
df_cumulative_month=df_joinbymonth.merge(df_quitbymonth, left_on='DOJ', right_on='RelevingD
```

In [71]:

```
df_cumulative_month=df_cumulative_month.drop('RelevingDate_Entered',axis=1)
df_cumulative_month= df_cumulative_month.rename(columns={'0_x':'joiners', '0_y':'quitters'}
```

In [72]:

```
df_cumulative_month=df_cumulative_month.rename(columns={'DOJ':'Month'})
```

In [73]:

```
df_cumulative_month.head()
```

Out[73]:

	Month	joiners	quitters
0	2019-04-30	2	NaN
1	2019-05-31	0	NaN
2	2019-06-30	2	NaN
3	2019-07-31	1	NaN
4	2019-08-31	4	NaN

In [74]:

```
df_cumulative_month.fillna(0, inplace=True)
```

In [75]:

```
df_cumulative_month.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 39 entries, 0 to 38
Data columns (total 3 columns):
 #   Column    Non-Null Count  Dtype         
---  ------    --------------  -----         
 0   Month     39 non-null     datetime64[ns]
 1   joiners   39 non-null     int64         
 2   quitters  39 non-null     float64       
dtypes: datetime64[ns](1), float64(1), int64(1)
memory usage: 1.2 KB
```

In [76]:

```
df_cumulative_month['joiners'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[76]:

```
count       39.000000
mean        63.564103
median       8.000000
<lambda>     4.000000
Name: joiners, dtype: float64
```

In [77]:

```
df_cumulative_month['quitters'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[77]:

```
count       39.000000
mean        63.564103
median       1.000000
<lambda>     0.000000
Name: quitters, dtype: float64
```

In [78]:

```
df_cumulative_month['quitters']=df_cumulative_month['quitters']*-1
```

In [79]:

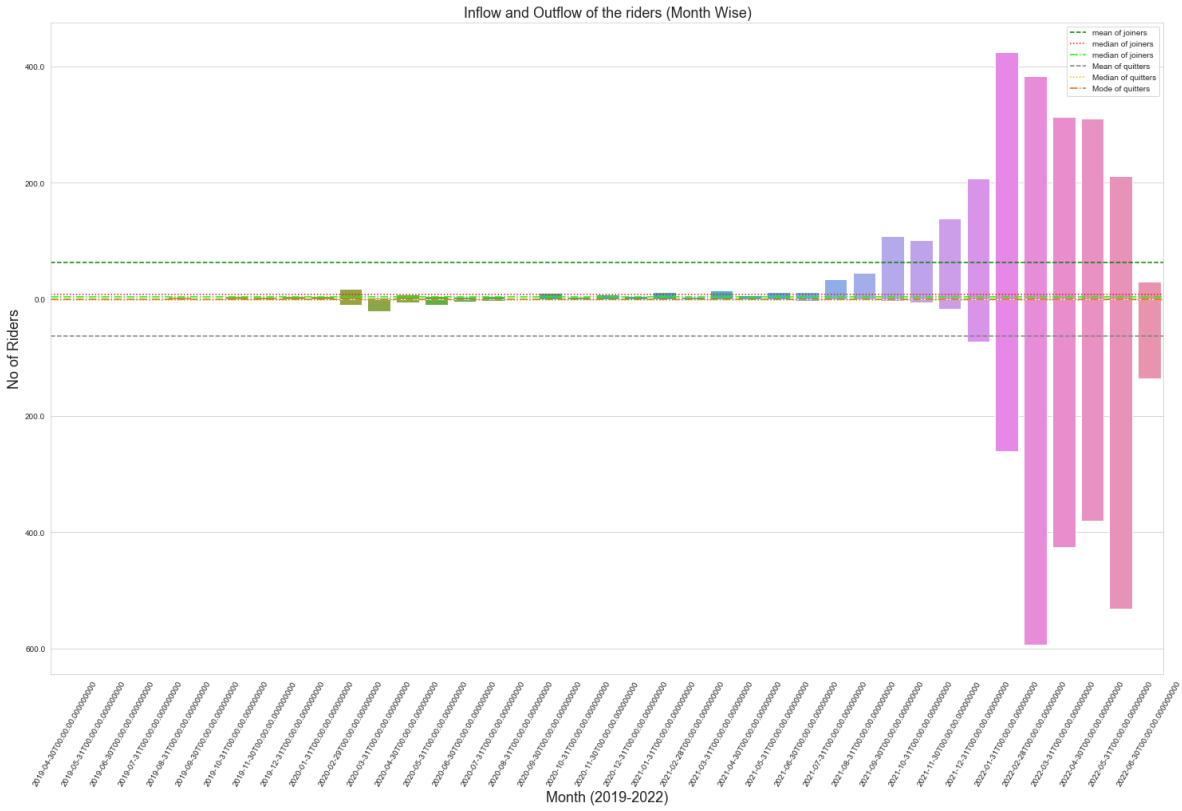
```
df_cumulative_month.head()
```

Out[79]:

	Month	joiners	quitters
0	2019-04-30	2	-0.0
1	2019-05-31	0	-0.0
2	2019-06-30	2	-0.0
3	2019-07-31	1	-0.0
4	2019-08-31	4	-0.0

In [80]:

```
plt.figure(figsize=(25,15))
bar_plot=sns.barplot(data=df_cumulative_month, x=df_cumulative_month['Month'], y=df_cumulat
            dodge=False)
bar_plot=sns.barplot(data=df_cumulative_month, x=df_cumulative_month['Month'], y=df_cumulat
            dodge=False)
plt.ylabel('No of Riders',fontsize=18)
plt.xlabel('Month (2019-2022)',fontsize=18)
plt.title(label="Inflow and Outflow of the riders (Month Wise)",fontsize=18)
plt.xticks(rotation=60)
bar_plot.axhline(y =63.6,color ="green", linestyle ="--",label='mean of joiners')
bar_plot.axhline(y = 8,color ="red", linestyle =":",label='median of joiners')
bar_plot.axhline(y = 4,color ="lime", linestyle ="-.",label='median of joiners')
bar_plot.axhline(y =-63.6,color ="grey", linestyle ="--",label='Mean of quitters')
bar_plot.axhline(y = -1,color ="orange", linestyle =":",label='Median of quitters')
bar_plot.axhline(y = 0,color ="chocolate", linestyle ="-.",label='Mode of quitters')
bar_plot.set_yticklabels([str(abs(bar_plot)) for bar_plot in bar_plot.get_yticks()])
plt.legend()
plt.show()
```



For 2020-2022

In [81]:

```
df3.head()
```

Out[81]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entered
0	SRIKANTH P	2019-04-04	Active	Terminated	2020-05-19	2020-04-02
1	SUPREEM .	2019-04-16	Active	Resigned	2020-05-19	2020-02-20
6	Deepak Badiya	2022-05-10	Active	Terminated	2022-05-14	2022-05-14
8	Naveen Kumar V	2019-08-07	Active	Drop-outs	2022-02-10	2022-02-10
10	BANAJ KUMAR SAHOO .	2019-07-04	Active	Terminated	2020-07-20	2020-04-29

In [82]:

```
df_20_22=df3[df3['Riders_Joined_by_year']>2019]
```

In [83]:

```
df_20_22.head()
```

Out[83]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entered
6	Deepak Badiya	2022-05-10	Active	Terminated	2022-05-14	2022-05-14
45	ESAKKI RAJAN I	2020-01-03	Active	Resigned	2022-02-01	2021-11-09
65	Bala Krishnan .	2020-01-02	Active	Terminated	2020-05-19	2020-02-21
66	NAVEEN KUMAR .	2020-01-02	Active	Terminated	2020-05-19	2020-03-04
75	Kishore K	2020-01-23	Active	Terminated	2020-07-20	2020-03-19

No of joiner and quitter by Year wise

In [84]:

```
df_joinbyyear=pd.DataFrame(df_20_22.groupby(pd.Grouper(key='DOJ', freq='Y')).size().sort_in
df_joinbyyear.reset_index(inplace=True)
```

In [85]:

```
df_quitbyyear=pd.DataFrame(df_20_22.groupby(pd.Grouper(key='RelevingDate_Entered', freq='Y'
```

In [86]:

```
df_quitbyyear.reset_index(inplace=True)
```

In [87]:

```
df_cumulative_year=df_joinbyyear.merge(df_quitbyyear, left_on='DOJ', right_on='RelevingDate
```

In [88]:

```
df_cumulative_year=df_cumulative_year.drop('RelevingDate_Entered',axis=1)
df_cumulative_year= df_cumulative_year.rename(columns={'0_x':'joiners', '0_y':'quitters'})
```

In [89]:

```
df_cumulative_year=df_cumulative_year.rename(columns={'DOJ':'Month'})
```

In [90]:

```
df_cumulative_year
```

Out[90]:

	Month	joiners	quitters
0	2020-12-31	80	28
1	2021-12-31	701	101
2	2022-12-31	1673	2325

In [91]:

```
df_cumulative_year.fillna(0, inplace=True)
```

In [92]:

```
df_cumulative_year.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 3 entries, 0 to 2
Data columns (total 3 columns):
 #   Column    Non-Null Count  Dtype         
---  ------    --------------  -----         
 0   Month     3 non-null      datetime64[ns]
 1   joiners   3 non-null      int64         
 2   quitters  3 non-null      int64         
dtypes: datetime64[ns](1), int64(2)
memory usage: 96.0 bytes
```

In [93]:

```
df_cumulative_year['joiners'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[93]:

```
count         3.0
mean        818.0
median      701.0
<lambda>     80.0
Name: joiners, dtype: float64
```

In [94]:

```
df_cumulative_year['quitters'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[94]:

```
count         3.0
mean        818.0
median      101.0
<lambda>     28.0
Name: quitters, dtype: float64
```

In [95]:

```
df_cumulative_year['quitters']=df_cumulative_year['quitters']*-1
```

In [96]:

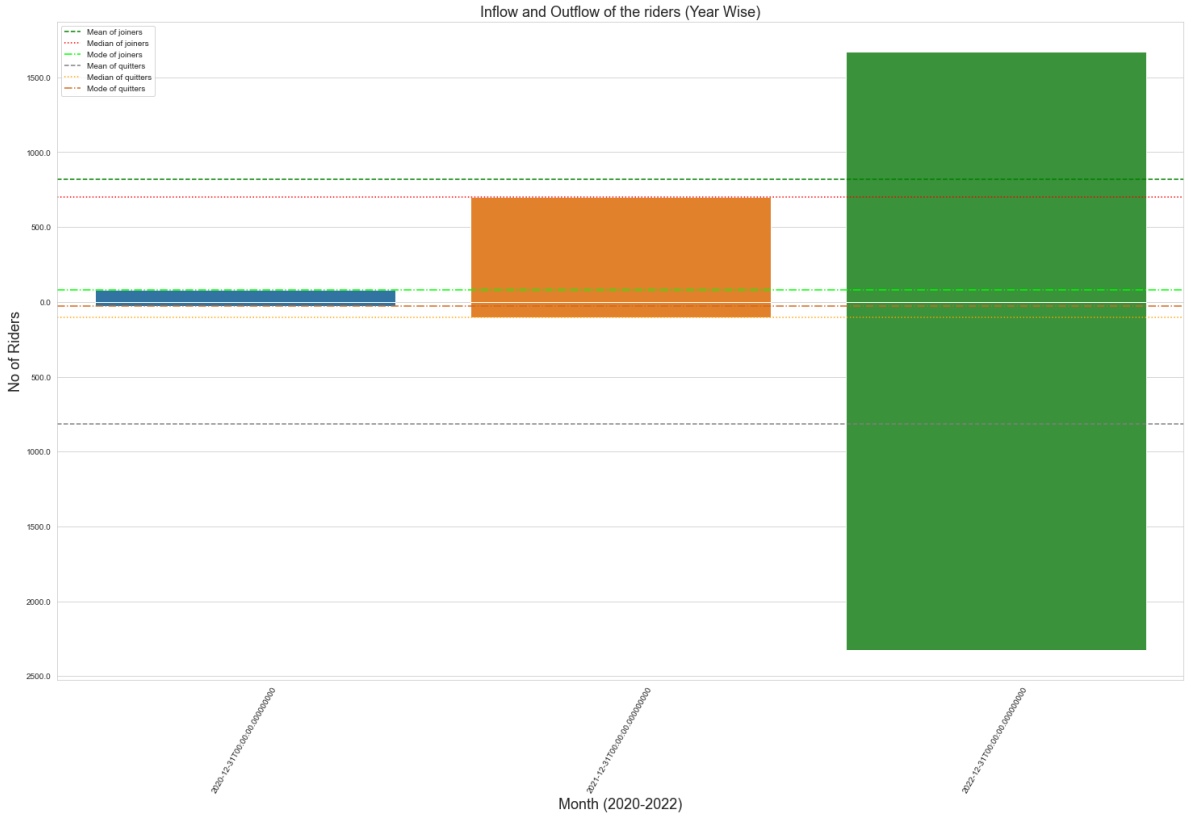
```
df_cumulative_year
```

Out[96]:

	Month	joiners	quitters
0	2020-12-31	80	-28
1	2021-12-31	701	-101
2	2022-12-31	1673	-2325

In [97]:

```
plt.figure(figsize=(25,15))
bar_plot=sns.barplot(data=df_cumulative_year, x=df_cumulative_year['Month'], y=df_cumulativ
            dodge=False)
bar_plot=sns.barplot(data=df_cumulative_year, x=df_cumulative_year['Month'], y=df_cumulativ
            dodge=False)
plt.ylabel('No of Riders',fontsize=18)
plt.xlabel('Month (2020-2022)',fontsize=18)
plt.title(label="Inflow and Outflow of the riders (Year Wise)",fontsize=18)
plt.xticks(rotation=60)
bar_plot.axhline(y =818,color ="green", linestyle ="--",label='Mean of joiners')
bar_plot.axhline(y = 701,color ="red", linestyle =":",label='Median of joiners')
bar_plot.axhline(y = 80,color ="lime", linestyle ="-.",label='Mode of joiners')
bar_plot.axhline(y =-818,color ="grey", linestyle ="--",label='Mean of quitters')
bar_plot.axhline(y = -101,color ="orange", linestyle =":",label='Median of quitters')
bar_plot.axhline(y = -28,color ="chocolate", linestyle ="-.",label='Mode of quitters')
bar_plot.set_yticklabels([str(abs(bar_plot)) for bar_plot in bar_plot.get_yticks()])
plt.legend()
plt.show()
```



No of joiner and quitter by Month wise

In [98]:

```
df_joinbymonth=pd.DataFrame(df_20_22.groupby(pd.Grouper(key='DOJ', freq='M')).size().sort_i
df_joinbymonth.reset_index(inplace=True)
```

In [99]:

```
df_quitbymonth=pd.DataFrame(df_20_22.groupby(pd.Grouper(key='RelevingDate_Entered',
                                                        freq='M')).size().sort_index(ascend
```

In [100]:

```
df_quitbymonth.reset_index(inplace=True)
```

In [101]:

```
df_cumulative_month=df_joinbymonth.merge(df_quitbymonth, left_on='DOJ', right_on='RelevingD
```

In [102]:

```
df_cumulative_month=df_cumulative_month.drop('RelevingDate_Entered',axis=1)
df_cumulative_month= df_cumulative_month.rename(columns={'0_x':'joiners', '0_y':'quitters'}
```

In [103]:

```
df_cumulative_month=df_cumulative_month.rename(columns={'DOJ':'Month'})
```

In [104]:

```
df_cumulative_month.head()
```

Out[104]:

	Month	joiners	quitters
0	2020-01-31	6	NaN
1	2020-02-29	18	6.0
2	2020-03-31	3	9.0
3	2020-04-30	9	1.0
4	2020-05-31	5	7.0

In [105]:

```
df_cumulative_month.fillna(0, inplace=True)
```

In [106]:

```
df_cumulative_month.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 30 entries, 0 to 29
Data columns (total 3 columns):
 #   Column    Non-Null Count  Dtype         
---  ------    --------------  -----         
 0   Month     30 non-null     datetime64[ns]
 1   joiners   30 non-null     int64         
 2   quitters  30 non-null     float64       
dtypes: datetime64[ns](1), float64(1), int64(1)
memory usage: 960.0 bytes
```

In [107]:

```
df_cumulative_month['joiners'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[107]:

```
count       30.0
mean        81.8
median      13.0
<lambda>     4.0
Name: joiners, dtype: float64
```

In [108]:

```
df_cumulative_month['quitters'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[108]:

```
count       30.0
mean        81.8
median       2.5
<lambda>     0.0
Name: quitters, dtype: float64
```

In [109]:

```
df_cumulative_month['quitters']=df_cumulative_month['quitters']*-1
```

In [110]:

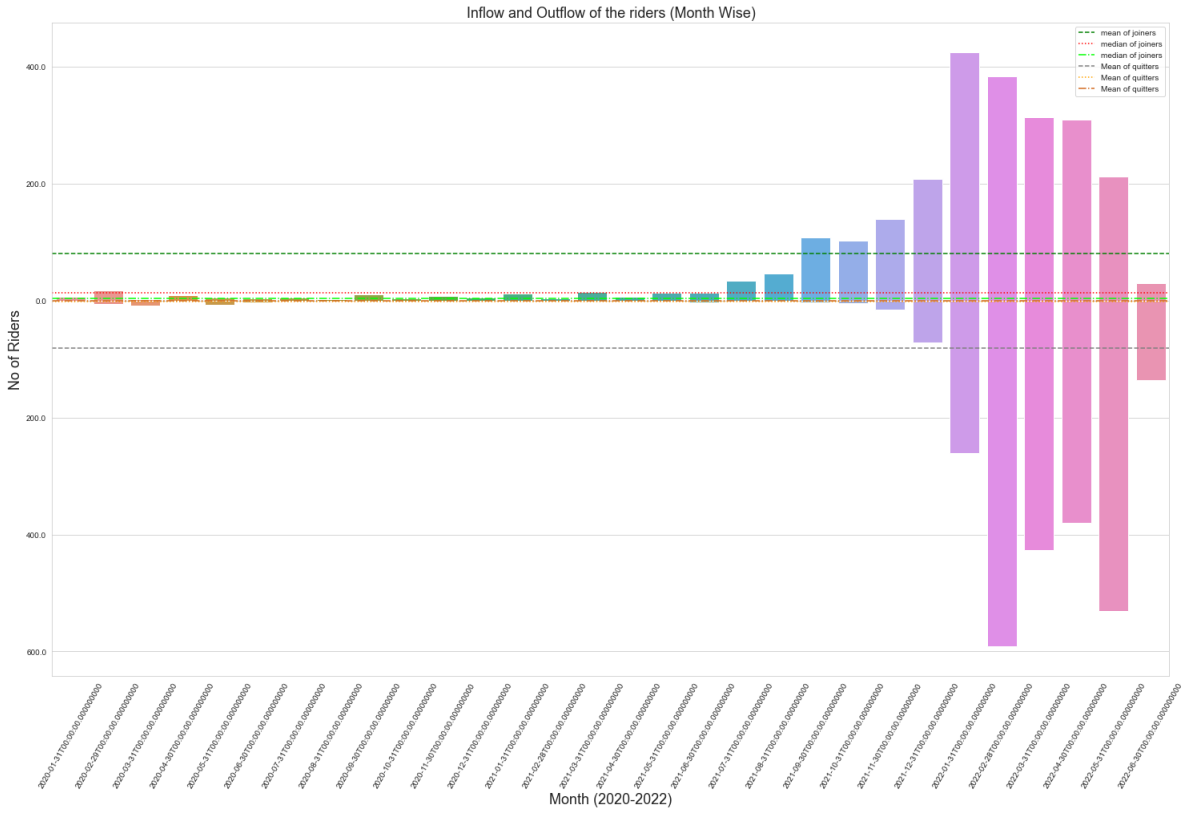
```
df_cumulative_month.head()
```

Out[110]:

	Month	joiners	quitters
0	2020-01-31	6	-0.0
1	2020-02-29	18	-6.0
2	2020-03-31	3	-9.0
3	2020-04-30	9	-1.0
4	2020-05-31	5	-7.0

In [111]:

```
plt.figure(figsize=(25,15))
bar_plot=sns.barplot(data=df_cumulative_month, x=df_cumulative_month['Month'], y=df_cumulat
            dodge=False)
bar_plot=sns.barplot(data=df_cumulative_month, x=df_cumulative_month['Month'], y=df_cumulat
            dodge=False)
plt.ylabel('No of Riders',fontsize=18)
plt.xlabel('Month (2020-2022)',fontsize=18)
plt.title(label="Inflow and Outflow of the riders (Month Wise)",fontsize=18)
plt.xticks(rotation=60)
bar_plot.axhline(y =81,color ="green", linestyle ="--",label='mean of joiners')
bar_plot.axhline(y = 13,color ="red", linestyle =":",label='median of joiners')
bar_plot.axhline(y = 4,color ="lime", linestyle ="-.",label='median of joiners')
bar_plot.axhline(y =-81,color ="grey", linestyle ="--",label='Mean of quitters')
bar_plot.axhline(y = -2,color ="orange", linestyle =":",label='Mean of quitters')
bar_plot.axhline(y = 0,color ="chocolate", linestyle ="-.",label='Mean of quitters')
bar_plot.set_yticklabels([str(abs(bar_plot)) for bar_plot in bar_plot.get_yticks()])
plt.legend()
plt.show()
```



Clients and Hub

In [112]:

```
df2['Client']
```

Out[112]:

```
0       Bigbasket Bangalore
1       Bigbasket Bangalore
6          BB_Now Bangalore
8       Bigbasket Bangalore
10      Bigbasket Bangalore
               ...
7287    Bigbasket Bangalore
7289        1kart Bangalore
7290              1kart ROK
7292        1kart Bangalore
7325        1kart Bangalore
Name: Client, Length: 2479, dtype: object
```

In [113]:

```
df2['Hub'].unique()
```

Out[113]:

```
array(['Bigbasket Hulimavu', 'BBNow BLR-Bellandur', 'Bigbasket Siddapura',
       'Bigbasket Sarjapura', 'Bigbasket HRBR', 'Bigbasket JP Nagar',
       'Delhivery Wilson Garden', 'Bigbasket Mylapore',
       'Bigbasket Ittmadugu', 'Bigbasket Chitlapakkam',
       'Bigbasket Jayanagar', 'TOW Whitefield', 'Bigbasket Vanagaram',
       'Amazon Aramghar HYDK', '1kart Peenya', '1kart BEL',
       '1Kart RajajiNagar', 'ShadowFax Food LL Villa',
       '1Kart Mysore Road', '1Kart Ganga Nagar', 'Bigbasket Mahadevpura',
       'SFX E-Com CHN_Velachery', '1kart kanakapura',
       'Bigbasket Thoraipakkam', 'SFX E-Com CHN_Pallikaranai',
       '1Kart Srinagar', 'Grofers Jayanagar', '1Kart Vijayanagar',
       'Flipkart Minjur', '1kart_KRPuramHub_BLR', 'Grofers Velachery',
       'Myntra HSR Layout', '1kart Bilekahalli', 'BigBasket KRpuram',
       'BigBasket Kazhipattur', 'Shadowfax Food Chennai',
       '1kart Devanahalli', 'sowkea - valasaravakkam',
       'Grofers Nungambakkam', 'Bigbasket Manikonda',
       '1kart_DomlurHub_BLR', '1Kart Mahadevpura', 'Porter Guindy Office',
       'Flipkart Vandalur', 'Amazon Balanagar HYDI', 'sowkea- Royapettah',
```

In [114]:

```
df2['Hub'].nunique()
```

Out[114]:

```
353
```

In [115]:

```
df2['Client'].unique()
```

Out[115]:

```
array(['Bigbasket Bangalore', 'BB_Now Bangalore', 'Delhivery Bangalore',
       'Bigbasket Chennai', 'TOW Bangalore', 'Amazon Hyderabad',
       '1kart Bangalore', 'Shadowfax Food Bangalore',
       'Shadowfax E-Commerce ( Chennai )', 'Grofers Bangalore',
       'Flipkart Chennai', 'Grofers Chennai', 'Myntra Bangalore',
       'Shadowfax Food - Chennai', '1kart ROK', 'Sowkea Chennai',
       'Bigbasket Hyderabad', 'Porter 2 W Chennai', 'NAAMDHARI Bangalore',
       'PORTER BLR- SMARTSHIFT LOGISTICS SOLUTIONS P LTD',
       'Porter 2W Bangalore', 'OTP Express B2B',
       'shadowfax Bangalore FLKRT QUICK',
       'Shadowfax E-Commerce Bangalore',
       'shadowfax Hyderabad FLKRT QUICK', 'Flipkart Delhi',
       'Blue Dart-Chennai', 'Flipkart Hyderabad', 'Porter_Dwarka Sec-14',
       'Porter DELHI', 'Flipkart Kerala', 'Grofers Delhi',
       'Flipkart Noida', 'Village Milk', 'Porter 3 W Chennai',
       'Bigbasket Coimbatore', 'ShadowFax E-Comm Kerala',
       'Grofers Noida UP', 'Swiggy Bangalore', 'Flipkart Mumbai',
       'Shadowfax E-Commerce (Delhi)', 'Gourmet Garden',
```

In [116]:

```
df2['Client'].nunique()
```

Out[116]:

```
55
```

In [117]:

```
df_clients=df2.groupby("Client")['no_of_working_days'].agg(['count','mean','median', lambda
df_clients.columns=['no_of_rider','average', 'median', 'mode']
```

In [118]:

```
df_clients.head(5)
```

Out[118]:

Client	no_of_rider	average	median	mode
1kart Bangalore	533	66.275797	23.0	7
1kart ROK	50	58.660000	30.0	15
Amazon Fresh	2	13.000000	13.0	5
Amazon Hyderabad	12	295.583333	308.5	17
BB_Now Bangalore	56	15.750000	11.0	2

In [119]:

```
df_hub=df2.groupby("Hub")['no_of_working_days'].agg(['count','mean','median',lambda x :x.mo
df_hub.columns=['no_of_rider','average', 'median', 'mode']
```

In [120]:

```
df_hub.head()
```

Out[120]:

Hub	no_of_rider	average	median	mode
1Kart Bangalore Anekal	1	13.000000	13.0	13
1Kart Bangalore Jakkasandra	1	4.000000	4.0	4
1Kart Bangalore Sarjapur	3	35.000000	36.0	31
1Kart Banshankari	9	24.444444	24.0	24
1Kart E_City	16	50.125000	39.5	44

Testing

In [152]:

```
df3.head()
```

Out[152]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entered
0	SRIKANTH P	2019-04-04	Active	Terminated	2020-05-19	2020-04-02
1	SUPREEM .	2019-04-16	Active	Resigned	2020-05-19	2020-02-20
6	Deepak Badiya	2022-05-10	Active	Terminated	2022-05-14	2022-05-14
8	Naveen Kumar V	2019-08-07	Active	Drop-outs	2022-02-10	2022-02-10
10	BANAJ KUMAR SAHOO .	2019-07-04	Active	Terminated	2020-07-20	2020-04-29

In [153]:

```
df_19=df3[df3['Riders_Joined_by_year']==2019]
```

In [154]:

```
df_19.head()
```

Out[154]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_En
0	SRIKANTH P	2019-04-04	Active	Terminated	2020-05-19	2020-
1	SUPREEM .	2019-04-16	Active	Resigned	2020-05-19	2020-
8	Naveen Kumar V	2019-08-07	Active	Drop-outs	2022-02-10	2022-
10	BANAJ KUMAR SAHOO .	2019-07-04	Active	Terminated	2020-07-20	2020-
15	SHANTHAKUMAR .	2019-08-13	Active	Terminated	2020-07-20	2020-

In [155]:

```
df_joinbymonth=pd.DataFrame(df_19.groupby(pd.Grouper(key='DOJ', freq='M')).size().sort_inde
df_joinbymonth.reset_index(inplace=True)
```

In [156]:

```
df_quitbymonth=pd.DataFrame(df_19.groupby(pd.Grouper(key='RelevingDate_Entered',
                                                     freq='M')).size().sort_index(ascend
```

In [157]:

```
df_quitbymonth.reset_index(inplace=True)
```

In [158]:

```
df_cumulative_month=df_joinbymonth.merge(df_quitbymonth, left_on='DOJ', right_on='RelevingD
```

In [159]:

```
df_cumulative_month=df_cumulative_month.drop('RelevingDate_Entered',axis=1)
df_cumulative_month= df_cumulative_month.rename(columns={'0_x':'joiners', '0_y':'quitters'}
```

In [160]:

```
df_cumulative_month=df_cumulative_month.rename(columns={'DOJ':'Month'})
```

In [161]:

```
df_cumulative_month.head()
```

Out[161]:

	Month	joiners	quitters
0	2019-04-30	2.0	NaN
1	2019-05-31	0.0	NaN
2	2019-06-30	2.0	NaN
3	2019-07-31	1.0	NaN
4	2019-08-31	4.0	NaN

In [162]:

```
df_cumulative_month.fillna(0, inplace=True)
```

In [163]:

```
df_cumulative_month.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 34 entries, 0 to 33
Data columns (total 3 columns):
 #   Column    Non-Null Count  Dtype  
---  ------    --------------  -----  
 0   Month     34 non-null     object 
 1   joiners   34 non-null     float64
 2   quitters  34 non-null     float64
dtypes: float64(2), object(1)
memory usage: 1.1+ KB
```

In [164]:

```
df_cumulative_month['joiners'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[164]:

```
count       34.000000
mean         0.735294
median       0.000000
<lambda>     0.000000
Name: joiners, dtype: float64
```

In [165]:

```
df_cumulative_month['quitters'].agg(['count','mean','median',lambda x :x.mode()[0]])
```

Out[165]:

```
count       34.000000
mean         0.735294
median       0.000000
<lambda>     0.000000
Name: quitters, dtype: float64
```

In [166]:

```
df_cumulative_month['quitters']=df_cumulative_month['quitters']*-1
```

In [167]:

```
df_cumulative_month.head()
```

Out[167]:

	Month	joiners	quitters
0	2019-04-30 00:00:00	2.0	-0.0
1	2019-05-31 00:00:00	0.0	-0.0
2	2019-06-30 00:00:00	2.0	-0.0
3	2019-07-31 00:00:00	1.0	-0.0
4	2019-08-31 00:00:00	4.0	-0.0

Testing 1

In [121]:

```
df3.head()
```

Out[121]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entered
0	SRIKANTH P	2019-04-04	Active	Terminated	2020-05-19	2020-04-02
1	SUPREEM .	2019-04-16	Active	Resigned	2020-05-19	2020-02-20
6	Deepak Badiya	2022-05-10	Active	Terminated	2022-05-14	2022-05-14
8	Naveen Kumar V	2019-08-07	Active	Drop-outs	2022-02-10	2022-02-10
10	BANAJ KUMAR SAHOO .	2019-07-04	Active	Terminated	2020-07-20	2020-04-29

In [122]:

```
df_2020_joiner=df3[df3['Riders_Joined_by_year']==2020]
```

In [123]:

```
df_2020_joiner.head()
```

Out[123]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_Entered
45	ESAKKI RAJAN I	2020-01-03	Active	Resigned	2022-02-01	2021-11-09
65	Bala Krishnan .	2020-01-02	Active	Terminated	2020-05-19	2020-02-21
66	NAVEEN KUMAR .	2020-01-02	Active	Terminated	2020-05-19	2020-03-04
75	Kishore K	2020-01-23	Active	Terminated	2020-07-20	2020-03-19
76	Rohith Kumar	2020-01-20	Active	Drop-outs	2022-05-06	2022-05-01

In [124]:

```
df_2020_quitter=df3[df3['Riders_Quitted_by_year']==2020]
```

In [125]:

```
df_2020_quitter.head()
```

Out[125]:

	RiderName	DOJ	Previous_Status	Current_Status	Status_Changed	RelevingDate_En
0	SRIKANTH P	2019-04-04	Active	Terminated	2020-05-19	2020-
1	SUPREEM .	2019-04-16	Active	Resigned	2020-05-19	2020-
10	BANAJ KUMAR SAHOO .	2019-07-04	Active	Terminated	2020-07-20	2020-
15	SHANTHAKUMAR .	2019-08-13	Active	Terminated	2020-07-20	2020-
18	Somashekar Navalur Charlie	2019-08-22	Active	Terminated	2020-05-19	2020-

In [126]:

```
stats(df_2020_joiner)
```

Out[126]:

```
[378.675, 0, 473.5]
```

In [127]:

```
stats(df_2020_quitter)
```

Out[127]:

```
[103.2, 0, 69.0]
```

In [128]:

```
df_2021_joiner=df3[df3['Riders_Joined_by_year']==2021]
```

In [129]:

```
df_2021_quitter=df3[df3['Riders_Quitted_by_year']==2021]
```

In [130]:

```
stats(df_2021_joiner)
```

Out[130]:

```
[124.5506419400856, 137, 112.0]
```

In [131]:

```
stats(df_2021_quitter)
```

Out[131]:

```
[147.6764705882353, 11, 64.0]
```

In [132]:

```
df_2022_joiner=df3[df3['Riders_Joined_by_year']==2022]
```

In [133]:

```
df_2022_quitter=df3[df3['Riders_Quitted_by_year']==2022]
```

In [134]:

```
stats(df_2022_joiner)
```

Out[134]:

```
[27.651524208009565, 7, 20.0]
```

In [135]:

```
stats(df_2022_quitter)
```

Out[135]:

```
[64.6699613235926, 7, 31.0]
```

In [ ]:

```

```