In [1]:
```
import os
import numpy as np
import pandas as pd
## import matplotlib.pyplot as plt
## %matplotlib inline
## import seaborn as sns
import sqlalchemy
from sqlalchemy import text
from datetime import datetime
pd.set_option('display.max_columns', None)
```

In [8]:
```
share_path = '//10.10.32.223/Risk Management/Analytics and Forecasting/Global PP Metrics/M
local_path = 'C:/Users/dustinoneil/Desktop/Global Payment Processing Metrics/'
```

Irvine

In [3]:
```
print(datetime.now().strftime("%H:%M:%S"), "  --  Start reading Irvine data")
Irvine = pd.read_csv(local_path + 'Irvine_Merchant_Portfolio_rev_3.2 - add zeus CRM.csv'
            , dtype={'Month' : 'str'})
```

```
09:14:01   --  Start reading Irvine data
C:\ProgramData\Anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3071: DtypeWar
ning: Columns (5) have mixed types.Specify dtype option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

In [ ]:
```
```

In [4]:
```
Irvine_1 = Irvine.copy()
#Irvine_1 = Irvine_1.loc[~Irvine_1['Month'].isna()]
Irvine_1 = Irvine_1.rename(columns={"Month":"YM"})
Irvine_1['Month'] = pd.to_datetime(Irvine_1['Date'])
Irvine_1['Business Country'] = 'US'
Irvine_1 = Irvine_1.drop(columns=['Date'])

status_dict_Irvine_1 = {'MS - Cancellation':'Cancelled',
                        'MS - Active':'Active'}
Irvine_1['current Status'] = Irvine_1['current Status'].replace(status_dict_Irvine_1)

Irvine_1['Business Group'] = 'Meritus'
#Irvine_1 = Irvine_1.drop(columns=['Activation Flag.1']) #can remove the duplicate column
Irvine_1
```

Out[4]:

	MID	Business DBA Name	Business Legal Name	current Status	MCC	CRM_VendorUID	Partner DBA	Bank	BI
0	3286000000008060	Active Ride Shop 1002	Active Sports Lifestyle USA LLC	Cancelled	5699.0	NaN	Pro Music Group	Woodforest	32860
1	3286000000008094	Two Trees Inn	Mashantucket Pequot Gaming Enterprise	Cancelled	7011.0	NaN	Primary Merchant Solutions	Woodforest	32860

	MID	Business DBA Name	Business Legal Name	current Status	MCC	CRM_VendorUID	Partner DBA	Bank	BII
2	3286000000008094	Two Trees Inn	Mashantucket Pequot Gaming Enterprise	Cancelled	7011.0	NaN	Primary Merchant Solutions	Woodforest	32860
3	3286000000008094	Two Trees Inn	Mashantucket Pequot Gaming Enterprise	Cancelled	7011.0	NaN	Primary Merchant Solutions	Woodforest	32860
4	3286000000008094	Two Trees Inn	Mashantucket Pequot Gaming Enterprise	Cancelled	7011.0	NaN	Primary Merchant Solutions	Woodforest	32860
...	...	...	...	...	...	...	...	...	
564377	9109000089863889	Merrick Mirror Hosting	Merrick Mirror Hosting Inc	Cancelled	4816.0	NaN	DL Ventures, Inc	BMO - Harris	91090
564378	9109000089863889	Merrick Mirror Hosting	Merrick Mirror Hosting Inc	Cancelled	4816.0	NaN	DL Ventures, Inc	BMO - Harris	91090
564379	9109000089863889	Merrick Mirror Hosting	Merrick Mirror Hosting Inc	Cancelled	4816.0	NaN	DL Ventures, Inc	BMO - Harris	91090
564380	9109000089863889	Merrick Mirror Hosting	Merrick Mirror Hosting Inc	Cancelled	4816.0	NaN	DL Ventures, Inc	BMO - Harris	91090
564381	9109000089863889	Merrick Mirror Hosting	Merrick Mirror Hosting Inc	Cancelled	4816.0	NaN	DL Ventures, Inc	BMO - Harris	91090

564382 rows × 38 columns

In [ ]:

In [5]:
```
Irvine_1['Tenure_Months'].unique()
```

Out[5]:
```
array([  0,   5,   1,  14,  11,   2,   3,  24,  51,  69,  22,  57,  35,
         7,  28,  37,  41,  63,  61,  18,  62,  31,  48,  32,  55,   6,
        60,  42,  50,  58,  20,  10,  46,  15,   4,  44,  19,  59,  43,
         9,  25,  36,  52,  26,  33,  30,  54,  29,  13,  21,  40,  49,
        39,  17,   8,  27,  23,  45,  47,  12,  16,  34,  70,  53,  56,
        68,  65,  38,  66,  67,  64, 134,  84], dtype=int64)
```

In [10]:
```
Irvine_1['current Status'].unique()
```

Out[10]:
```
array(['Cancelled', 'Active', 'MS - Inactive',
       'MS - Conducted Welcome Call'], dtype=object)
```

UK and Montreal

Talk to Denise O about BINs to get to processor

In [11]:
```
#UK Montreal
print(datetime.now().strftime("%H:%M:%S"), "  --  Start reading UK and Montreal data")
UK_Montreal_2019 = pd.read_excel(share_path + 'Portfolio Metrics - Dustin UK MTL 2019 v2.x
UK_Montreal_2020 = pd.read_excel(share_path + 'Portfolio Metrics - Dustin UK MTL 2020 v2.x
UK_Montreal_2021 = pd.read_excel(share_path + 'Portfolio Metrics - Dustin UK MTL 2021 v2.x
```

```
09:33:31   --  Start reading UK and Montreal data
```

In [12]:
```
UK_Montreal = pd.concat([UK_Montreal_2019, UK_Montreal_2020, UK_Montreal_2021], axis=0)
UK_Montreal
```

Out[12]:

	ACCOUNT_ID	BUSINESSDBANAME	BUSINESSLEGALNAME	MCC	CURRENTSTATUS	MCCDESCRIPTION	PAR
0	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Closed	SPORTING GOODS STORES	
1	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Closed	SPORTING GOODS STORES	
2	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Closed	SPORTING GOODS STORES	
3	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Closed	SPORTING GOODS STORES	
4	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Closed	SPORTING GOODS STORES	
...	...	...	...	...	...	...	
321770	1003073224	cc - Fraser Arms	Second Street Fund Inc.	6513.0	Open	REAL ESTATE AGENTS AND MANAGERS-RENTALS	
321771	1003073244	cc - McNiven Place	Second Street Fund Inc.	6513.0	Open	REAL ESTATE AGENTS AND MANAGERS-RENTALS	
321772	1003073254	cc - Ciara Manor	Second Street Fund Inc.	6513.0	Open	REAL ESTATE AGENTS AND MANAGERS-RENTALS	
321773	1003073304	cc - Fraser Tower	Second Street Fund Inc.	6513.0	Open	REAL ESTATE AGENTS AND MANAGERS-RENTALS	
321774	1003073344	Cardinal Sports Academy Ltd.	Cardinal Sports Academy	8299.0	Open	SCHOOLS + EDUCATIONAL SVC-NOT ELSEWHERE CLASSI...	Sc

879183 rows × 37 columns

In [ ]:

In [13]:
```
UK_Montreal_1 = UK_Montreal.copy()
UK_Montreal_1 = UK_Montreal_1.rename(columns={
                                    "ACCOUNT_ID":"MID",
                                    "BUSINESSDBANAME": "Business DBA Name",
                                    "BUSINESSLEGALNAME": "Business Legal Name",
                                    "CURRENTSTATUS": "current Status",
                                    "PARTNERDBA": "Partner DBA",
                                    "OFFICENAME": "Office Name",
                                    "BANK": "Bank",
                                    "BUSINESSCOUNTRY": "Business Country",
                                    "BUSINESSGROUP" : "Business Group",
                                    "MNTH":"Month",
                                    "ACTIVE_MONTH_MAX":"Active Month Max",
                                    "INACTIVE_MONTH_MAX":"Inactive Month Max",
                                    "CC_AVG_TICKET_VALUE":"CC Avg Ticket Value",
                                    "CU_APPROVED_VOLUME":"CU Approved Volume",
                                    "PAYMENT_CATEGORY":"Payment Category",
                                    "CANCEL_DATE": "Cancel Month Max"
                                   }) #inplace = True

UK_Montreal_1 = UK_Montreal_1.drop(columns=['MCCDESCRIPTION', 'Tenure_Months', 'CANCEL_MON
UK_Montreal_1["MID"] = UK_Montreal_1["MID"].astype(str)
UK_Montreal_1["YM"] = UK_Montreal_1["Month"].dt.strftime('%Y') + UK_Montreal_1["Month"].dt
UK_Montreal_1["Active Month Max"] = UK_Montreal_1["Active Month Max"].dt.strftime('%Y') +
#UK_Montreal_1["Inactive Month Max"] =
UK_Montreal_1["Cancel Month Max"] = UK_Montreal_1["Cancel Month Max"].dt.strftime('%Y') +


#Flags
UK_Montreal_1["Activation Flag"] = (UK_Montreal_1["Active Month Max"] == UK_Montreal_1["YM
#UK_Montreal_1["Inactive Flag -1"] =
UK_Montreal_1["Cancelation Flag"] = (UK_Montreal_1["Cancel Month Max"] == UK_Montreal_1["Y
UK_Montreal_1["Cancelation Flag -1"] = (UK_Montreal_1["Cancel Month Max"] == UK_Montreal_1

UK_Montreal_1["Actively Processing"] = (UK_Montreal_1['Sales Volume'] >= 50).astype('int')
UK_Montreal_1['Tenure_Months'] = UK_Montreal_1[["MID", "Actively Processing"]].groupby('MI

status_dict = {'Closed': 'Cancelled', 'Open': 'Active'}
UK_Montreal_1['current Status'] = UK_Montreal_1['current Status'].replace(status_dict)
UK_Montreal_1
```

Out[13]:

	MID	Business DBA Name	Business Legal Name	MCC	current Status	Partner DBA	ISV_PARTNER_NAME	Bank	BIN	Office Name
0	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Cancelled	77778075 - AAA-No Agent(,/)	NaN	MOC	NaN	Montreal
1	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Cancelled	77778075 - AAA-No Agent(,/)	NaN	MOC	NaN	Montreal
2	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Cancelled	77778075 - AAA-No Agent(,/)	NaN	MOC	NaN	Montreal

	MID	Business DBA Name	Business Legal Name	MCC	current Status	Partner DBA	ISV_PARTNER_NAME	Bank	BIN	Office Name
3	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Cancelled	77778075 - AAA-No Agent(,/)	NaN	MOC	NaN	Montreal
4	500046	Ad Mats Corp. (Usd)	Ad Mats Corp. (Usd)	5941.0	Cancelled	77778075 - AAA-No Agent(,/)	NaN	MOC	NaN	Montreal
...	...	...	...	...	...	...	...	...	...	...
321770	1003073224	cc - Fraser Arms	Second Street Fund Inc.	6513.0	Active	77778075 - AAA-No Agent(,/)	Rentmoola - Propertii	PET	NaN	Montreal
321771	1003073244	cc - McNiven Place	Second Street Fund Inc.	6513.0	Active	77778075 - AAA-No Agent(,/)	Rentmoola - Propertii	PET	NaN	Montreal
321772	1003073254	cc - Ciara Manor	Second Street Fund Inc.	6513.0	Active	77778075 - AAA-No Agent(,/)	Rentmoola - Propertii	PET	NaN	Montreal
321773	1003073304	cc - Fraser Tower	Second Street Fund Inc.	6513.0	Active	77778075 - AAA-No Agent(,/)	Rentmoola - Propertii	PET	NaN	Montreal
321774	1003073344	Cardinal Sports Academy Ltd.	Cardinal Sports Academy	8299.0	Active	77778063 - Softerware (Doug Schoenberg) (Peopl...	Softerware	PET	NaN	Montreal

879183 rows × 36 columns

In [14]:
```
UK_Montreal_1['Bank'].unique()
```

Out[14]:
```
array(['MOC', 'PET', 'MER', 'VAN', 'AIB', 'BOV', 'OPL', 'NAB', 'BBVA'],
      dtype=object)
```

In [15]:
```
#str replace works but I'd like to use map instead
#UK_Montreal_1['current Status'] = UK_Montreal_1['current Status'].str.replace('Closed',
#UK_Montreal_1['current Status'] = UK_Montreal_1['current Status'].str.replace('Open', 'M
UK_Montreal_1['current Status'].unique()
```

Out[15]:
```
array(['Cancelled', 'Active'], dtype=object)
```

In [16]:
```
#Alternate methods

#UK_Montreal_1["YM"].dt.year.astype(str) + UK_Montreal_1["MNTH"].dt.month.astype(str)
#UK_Montreal_1["Actively Processing"] = UK_Montreal_1['Sales Volume'].ge(50).astype('int')

#Tenure_Months
#I found 3 methods to perform this. The better methods generate a series on which to work

#1 Transform method: Generate a sub-df which is a series, MID is the index, groupby and t
    #used above

#2 Map method
```

```
    #UK_Montreal_1['Tenure_Months_2'] = UK_Montreal_1['MID'].map(UK_Montreal_1.groupby('MI

#3 Change index method
    #More complicated, can be done easier
    #changes the index to MID and uses Group by.sum()...sum() sums that column on the inde
    #UK_Montreal_1 = UK_Montreal_1.set_index('MID')
    #UK_Montreal_1 = UK_Montreal_1.assign(Tenure_Months_3 = UK_Montreal_1["Actively Proces
    #UK_Montreal_1 = UK_Montreal_1.reset_index()
```

WLV

In [17]:

```
#Houston and WLV are US acq Business GROUP
#I don't know how to get Business Unit

#WLV
print(datetime.now().strftime("%H:%M:%S"), "  --  Start reading WLV data")
WLV_2019 = pd.read_excel(share_path + 'WLV Processing 2019.xlsx',
                    #nrows = 100,
                    sheet_name='Sheet1', #will have to iterate through them, or explicitly
                    converters={'MID':str, 'FirstProcessDate':pd.to_datetime,
                                'ClosureDate':pd.to_datetime, 'month':pd.to_datetime}) #'m

WLV_2020 = pd.read_excel(share_path + 'WLV Processing 2020.xlsx',
                    #nrows = 100,
                    sheet_name='Sheet1', #will have to iterate through them, or explicitly
                    converters={'MID':str, 'FirstProcessDate':pd.to_datetime,
                                'ClosureDate':pd.to_datetime, 'month':pd.to_datetime}) #'m

WLV_2021 = pd.read_excel(share_path + 'WLV Processing 2021.xlsx',
                    #nrows = 100,
                    sheet_name='Sheet1', #will have to iterate through them, or explicitly
                    converters={'MID':str, 'FirstProcessDate':pd.to_datetime,
                                'ClosureDate':pd.to_datetime, 'month':pd.to_datetime}) #'m
```

```
09:43:39   --  Start reading WLV data
```

In [ ]:

In [18]:

```
WLV = pd.concat([WLV_2019, WLV_2020, WLV_2021], axis=0)  #add WLV_2019,
```

In [ ]:

In [19]:

```
WLV_1 = WLV.copy()
WLV_1 = WLV_1.rename(columns={
                            #"ACCOUNT_ID":"MID",
                            "MerchantDBA": "Business DBA Name",
                            "MerchantLegal": "Business Legal Name",
                            "AccountStatus": "current Status",
                            "Partner": "Partner DBA",
                            "Office": "Office Name",
                            "AcquiringBank": "Bank",
                            #"BUSINESSCOUNTRY": "Business Country",
                            "BusinessGroup" : "Business Group",
                            "month":"Month",
                            "FirstProcessDate":"Active Month Max",
                            #"INACTIVE_MONTH_MAX":"Inactive Month Max",
                            "ClosureDate":"Cancel Month Max",
                            "CCApprovedAverageTicket":"CC Avg Ticket Value",
```

```
                                "CCApprovedVolume":"CU Approved Volume",
                                "CP/CNPMerchant":"Payment Category",
                                "VisaSaleCount":"Sales Count",
                                "VisasSaleAmount":"Sales Volume",
                                "VisaReturnCount":"Refund Count",
                                "VisaReturnAmount":"Refund Volume",
                                "Chargebackcount":"CB Count",
                                "ChargebackAmount":"CB Volume",
                                "ClosureReason": "Closure Reason Name"
                               })

#WLV_1 = WLV_1.drop(columns=['ISV_PARTNER_NAME'])

#WLV_1["Active Month Max"]= pd.to_datetime(WLV_1["Active Month Max"])
#WLV_1["Cancel Month Max"]= pd.to_datetime(WLV_1["Cancel Month Max"])
WLV_1["Active Month Max"] = WLV_1["Active Month Max"].dt.strftime('%Y') + WLV_1["Active Mo
WLV_1["Cancel Month Max"] = WLV_1["Cancel Month Max"].dt.strftime('%Y') + WLV_1["Cancel Mo

#Flags
WLV_1["YM"] = WLV_1["Month"].dt.strftime('%Y') + WLV_1["Month"].dt.strftime('%m')
WLV_1["Activation Flag"] = (WLV_1["Active Month Max"] == WLV_1["YM"]).astype('int')
#WLV_1["Inactive Flag -1"] =
WLV_1["Cancelation Flag"] = (WLV_1["Cancel Month Max"] == WLV_1["YM"]).astype('int')
WLV_1["Cancelation Flag -1"] = (WLV_1["Cancel Month Max"] == WLV_1["YM"]).astype('int')*-1

WLV_1["Actively Processing"] = (WLV_1['Sales Volume'] >= 50).astype('int')
WLV_1['Tenure_Months'] = WLV_1[["MID", "Actively Processing"]].groupby('MID').transform('s

WLV_1['Business Country'] = 'US'
WLV_1['Business Group'] = 'iPayment'
#WLV_1['Office Name'] = 'West Lake Village'
#WLV_1['Month'] = pd.to_datetime(WLV_1['YM'].str[0:4] + '-' + WLV_1['YM'].str[4:6]+  '-01
WLV_1['Refund Volume'] = WLV_1['Refund Volume']*-1
WLV_1['CB Volume'] = WLV_1['CB Volume']*-1

status_dict_WLV = {0: 'Cancelled', 1: 'Active'}
WLV_1['current Status'] = WLV_1['current Status'].replace(status_dict_WLV)

WLV_1
```

Out[19]:

	MID	Business Legal Name	Business DBA Name	Business Group	Office Name	BIN	Bank	
0	4200091110000880	FIRST DIRECT CORP.	FIRST DIRECT CORP.	iPayment	WLV.iPayment	420009	BBVA	IPI
1	4200091110037296	CONNECTION TECH COMMUNICATIONS	CONNECTION TECH COMMUNICATIONS	iPayment	WLV.iPayment	420009	BBVA	DSHO04 COMM
2	4200091110044037	THE CRUISE DEPOT, INC.	THE CRUISE DEPOT, INC.	iPayment	WLV.iPayment	420009	BBVA	EODE0
3	4200097990003589	ZONK DESIGNS INC	SNOOZER BODY PILLOWS	iPayment	WLV.iPayment	420009	BBVA	IPI
4	4200098980002805	BODYVED, INC.	BODYVED	iPayment	WLV.iPayment	420009	BBVA	IPI
...	...	...	...	...	...	...	...	...

	MID	Business Legal Name	Business DBA Name	Business Group	Office Name	BIN	Bank	
519273	4228997800000221	A & W Oil & Tire Company Inc	A & W One Stop	iPayment	WLV.PCS	422899	BBVA	PCS-264 Enterp
519274	4431637061000001	Joelle Team INC	D 'N A Smoke Shop	iPayment	WLV.PCS	443163	Merrick Bank	PCS-30 Proc
519275	4228997800000205	Jeffery Hatcher	Jeffery Hatcher Sole Proprietorship	iPayment	WLV.PCS	422899	BBVA	PCS-Hous
519276	4431632110007719	Hernando Sportsmans Club Inc	Hernando Sportsmans Club	iPayment	WLV.PCS	443163	Merrick Bank	PCS-34
519277	4431637060000036	D&PUSA2 INC	Smokin Joes	iPayment	WLV.PCS	443163	Merrick Bank	PCS-Hous

1933952 rows × 32 columns

In [ ]:

In [20]:
```
WLV_1['current Status'].unique()
```

Out[20]:
```
array(['Active', 'Cancelled'], dtype=object)
```

In [21]:
```
WLV_1['Tenure_Months'].unique()
```

Out[21]:
```
array([ 2, 29, 17, 33, 34, 15, 13, 32, 24,  3, 12, 31,  6, 14,  4,  1,  0,
       26,  5,  9, 23, 19, 21, 11, 16, 25,  7, 20, 28, 10,  8, 30, 18, 27,
       22])
```

Houston_BBVA

This is all TYSY - no BIN given

In [22]:
```
print(datetime.now().strftime("%H:%M:%S"), "  --  Start reading Houston BBVA data")
Houston_BBVA_2019 = pd.read_excel(share_path + 'Houston BBVA Processing 2019.xlsx', sheet_
                    converters={'Month':str, 'MID':str})
Houston_BBVA_2020 = pd.read_excel(share_path + 'Houston BBVA Processing 2020.xlsx', sheet_
                    converters={'Month':str, 'MID':str})
Houston_BBVA_2021 = pd.read_excel(share_path + 'Houston BBVA Processing 2021.xlsx', sheet_
                    converters={'Month':str, 'MID':str})

#Houston BBVA Processing 2021.xlsx and 2020
```

```
10:09:00   --  Start reading Houston BBVA data
```

In [23]:

```
Houston_BBVA = pd.concat([Houston_BBVA_2019, Houston_BBVA_2020, Houston_BBVA_2021], axis=0
```

In [24]:

```
Houston_BBVA_1 = Houston_BBVA.copy()
Houston_BBVA_1 = Houston_BBVA_1.rename(columns={
                            #"ACCOUNT_ID":"MID",
                            #"MerchantDBA": "Business DBA Name",
                            #"MerchantLegal": "Business Legal Name",
                            "Current Status": "current Status",
                            #"Partner": "Partner DBA",
                            #"Office": "Office Name",
                            #"AcquiringBank": "Bank",
                            #"BUSINESSCOUNTRY": "Business Country",
                            "BusinessGroup" : "Business Group",
                            "Month":"YM",
                            "Credit Decision Date":"Active Month Max",  #Would prefer firs
                            #"INACTIVE_MONTH_MAX":"Inactive Month Max",
                            "Cancel Month":"Cancel Month Max",
                            "CC Avg Ticket":"CC Avg Ticket Value",
                            #"CCApprovedVolume":"CU Approved Volume",
                            #"CP/CNPMerchant":"Payment Category",
                            #"VisaSaleCount":"Sales Count",
                            #"VisasSaleAmount":"Sales Volume",
                            "Return Count":"Refund Count",
                            "Return Volume":"Refund Volume",
                            #"Chargebackcount":"CB Count",
                            #"ChargebackAmount":"CB Volume",
                            #"ClosureReason": "Closure Reason Name" #not available
                           })



#Houston_BBVA_1 = Houston_BBVA_1.drop(columns=['ISV_PARTNER_NAME'])

Houston_BBVA_1["Active Month Max"] = Houston_BBVA_1["Active Month Max"].dt.strftime('%Y')
Houston_BBVA_1["Cancel Month Max"] = Houston_BBVA_1["Cancel Month Max"].dt.strftime('%Y')


#Flags
Houston_BBVA_1["Activation Flag"] = (Houston_BBVA_1["Active Month Max"] == Houston_BBVA_1
#Houston_BBVA_1["Inactive Flag -1"] =
Houston_BBVA_1["Cancelation Flag"] = (Houston_BBVA_1["Cancel Month Max"] == Houston_BBVA_1
Houston_BBVA_1["Cancelation Flag -1"] = (Houston_BBVA_1["Cancel Month Max"] == Houston_BBV

Houston_BBVA_1["Actively Processing"] = (Houston_BBVA_1['Sales Volume'] >= 50).astype('int
Houston_BBVA_1['Tenure_Months'] = Houston_BBVA_1[["MID", "Actively Processing"]].groupby('

Houston_BBVA_1['Business Country'] = 'US'
Houston_BBVA_1['Business Group'] = 'MCPS'
Houston_BBVA_1['Bank'] = 'BBVA'
Houston_BBVA_1['Month'] = pd.to_datetime(Houston_BBVA_1['YM'].str[0:4] + '-' + Houston_BBV
#Houston_BBVA_1['Refund Volume'] = Houston_BBVA_1['Refund Volume']*-1
#Houston_BBVA_1['CB Volume'] = Houston_BBVA_1['CB Volume']*-1


status_dict_Houston_BBVA = {#nan:'MS - Active', gets handled by replace(np.nan...
                            'V':'Diverted',
                            'C':'Cancelled',
                            'C,V':'Cancelled',
                            'V,C':'Cancelled',
                            'S':'Active'}

Houston_BBVA_1['current Status'] = Houston_BBVA_1['current Status'].replace(status_dict_Ho
```

```
Houston_BBVA_1['current Status'] = Houston_BBVA_1['current Status'].replace(np.nan, 'Activ
Houston_BBVA_1
```

Out[24]:

	MID	Business DBA Name	Business Legal Name	current Status	MCC	Partner DBA	Bank	Office Name	Active Month Max	Cancel Month Max	CC Avg Ticket Value	Ap
0	220000019	COMPASS BANK	COMPASS BANK	Active	6010	BBVA	BBVA	Houston	197610	NaN	NaN	
1	220000019	COMPASS BANK	COMPASS BANK	Active	6010	BBVA	BBVA	Houston	197610	NaN	NaN	
2	220000043	BBVA COMPASS BELTLINE RD	BBVA COMPASS BELTLINE RD	Active	6010	BBVA	BBVA	Houston	197610	NaN	NaN	
3	220000043	BBVA COMPASS BELTLINE RD	BBVA COMPASS BELTLINE RD	Active	6010	BBVA	BBVA	Houston	197610	NaN	NaN	
4	220001694	COMPASS GOVERNORS DR	COMPASS GOVERNORS DR	Active	6010	BBVA	BBVA	Houston	197610	NaN	NaN	
...	...	...	...	...	...	...	...	...	...	...	...	...
17290	220379581	BBVA COMPASS	BBVA COMPASS	Active	6010	BBVA	BBVA	Houston	200101	NaN	NaN	
17291	221172115	DANCE ET CETERA	DANCE ET CETERA LLC	Active	7911	BBVA Portfolio Purchase 2018	BBVA	Houston	201408	NaN	44.42	
17292	220510971	LAWRENCE A WRIGHT MD	LAWRENCE A WRIGHT MD	Active	8043	BBVA Portfolio Purchase 2018	BBVA	Houston	200409	NaN	10.00	2
17293	221072471	CORRAL TRAN SINGH	CORRAL TRAN SINGH LLP	Active	8111	BBVA Portfolio Purchase 2018	BBVA	Houston	201205	NaN	1500.00	4
17294	220553549	HOTEL LIQUIDATORS	GARY S WILLIAMS	Active	5712	BBVA Portfolio Purchase 2018	BBVA	Houston	200505	NaN	2611.54	

42718 rows × 25 columns

In [25]:
```
Houston_BBVA_1['current Status'].unique()
```

Out[25]:
```
array(['Active', 'Cancelled', 'Diverted'], dtype=object)
```

Houston_WF

In [26]:
```
#Houston_WF is in primus SQL
print(datetime.now().strftime("%H:%M:%S"), "  --  Start reading Houston WF data")
Houston_WF = pd.read_csv(local_path + 'Houston WF_Portfolio_Processing rev 1.0.csv'
            , dtype={'Month' : 'str'})
```

```
10:09:25   --  Start reading Houston WF data
```

In [ ]:

In [27]:
```
Houston_WF_1 = Houston_WF.copy()
#Houston_WF_1 = Houston_WF_1.loc[~Houston_WF_1['Month'].isna()]
Houston_WF_1 = Houston_WF_1.rename(columns={"Month":"YM"})
Houston_WF_1['Month'] = pd.to_datetime(Houston_WF_1['Date'])
Houston_WF_1['Business Country'] = 'US'
Houston_WF_1['Business Group'] = 'MCPS'
Houston_WF_1 = Houston_WF_1.drop(columns=['Date'])

status_dict_Houston_WF_1 = {'MS - Cancellation':'Cancelled', 'MS - Active':'Active'}
Houston_WF_1['current Status'] = Houston_WF_1['current Status'].replace(status_dict_Housto

#Houston_WF_1 = Houston_WF_1.drop(columns=['Activation Flag.1']) #can remove the duplicate
Houston_WF_1
```

Out[27]:

	MID	Business DBA Name	Business Legal Name	current Status	MCC	Partner DBA	Bank	BIN	Office Name
0	296201636887	ABSOLUTE PROCESS INSTRUMENTS INC	ABSOLUTE PROCESS INSTRUMENTS I	Active	5999	CBS06	Woodforest	296201	Houston
1	296201636887	ABSOLUTE PROCESS INSTRUMENTS INC	ABSOLUTE PROCESS INSTRUMENTS I	Active	5999	CBS06	Woodforest	296201	Houston
2	296201636887	ABSOLUTE PROCESS INSTRUMENTS INC	ABSOLUTE PROCESS INSTRUMENTS I	Active	5999	CBS06	Woodforest	296201	Houston
3	296201636887	ABSOLUTE PROCESS INSTRUMENTS INC	ABSOLUTE PROCESS INSTRUMENTS I	Active	5999	CBS06	Woodforest	296201	Houston
4	296201636887	ABSOLUTE PROCESS INSTRUMENTS INC	ABSOLUTE PROCESS INSTRUMENTS I	Active	5999	CBS06	Woodforest	296201	Houston
...	...	...	...	...	...	...	...	...	...
1493050	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975	Merchants Bancard Network, Inc	Woodforest	548298	Houston
1493051	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975	Merchants Bancard Network, Inc	Woodforest	548298	Houston
1493052	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975	Merchants Bancard Network, Inc	Woodforest	548298	Houston

	MID	Business DBA Name	Business Legal Name	current Status	MCC	Partner DBA	Bank	BIN	Office Name
1493053	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975	Merchants Bancard Network, Inc	Woodforest	548298	Houston
1493054	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975	Merchants Bancard Network, Inc	Woodforest	548298	Houston

1493055 rows × 34 columns

In [28]:
```
Houston_WF_1['current Status'].unique()
```

Out[28]:
```
array(['Active', 'Cancelled'], dtype=object)
```

MCC Groups

In [29]:
```
MCC_group = pd.WLV = pd.read_excel(share_path + 'MCC Group Map.xlsx', sheet_name = 'Sheet1
```

In [30]:
```
MCC_group
```

Out[30]:

	MCC	MCC Group	MCC Description
0	742	Other Business Services	VETERINARY SERVICES
1	763	Other Business Services	AGRICULTURAL COOPERATIVES
2	780	Maintenance & Repair	LANDSCAPING AND HORTICULTURAL SERVICES
3	1520	Maintenance & Repair	GENERAL CONTRACTORS-RESIDENTIAL BUILDINGS
4	1711	Maintenance & Repair	HEATING PLUMBING AIR CONDITIONING CONTRACTORS
...	...	...	...
898	3816	Travel, Airlines, Rental and Ticketing	HOME2SUITES
899	7231	Other Business Services	BEAUTY SHOPS AND BARBER SHOPS
900	7539	Maintenance & Repair	AUTO REPAIR (NON-DEALER)
901	8221	Health, Legal, Government & Education	COLLEGES UNIV PRO SCHOOLS JUNIOR COLLEGES
902	5423	Retail / Tangible Goods	FREEZER LOCKER MEAT PROVISIONERS

903 rows × 3 columns

Concat Offices

In [31]:
```
merge_1 = pd.concat([Irvine_1, UK_Montreal_1, WLV_1, Houston_BBVA_1, Houston_WF_1], axis=0
merge_1.reset_index()
```

```
bank_dict_merge_1 = {'Merrick Bank': 'MER'}
merge_1['Bank'] = merge_1['Bank'].replace(bank_dict_merge_1)
merge_1 = merge_1.merge(MCC_group, on='MCC', how='left')
merge_1
```

Out[31]:

	MID	Business DBA Name	Business Legal Name	current Status	MCC	CRM_VendorUID	Partner DBA	Bank	
0	3286000000008060	Active Ride Shop 1002	Active Sports Lifestyle USA LLC	Cancelled	5699.0	NaN	Pro Music Group	Woodforest	32
1	3286000000008094	Two Trees Inn	Mashantucket Pequot Gaming Enterprise	Cancelled	7011.0	NaN	Primary Merchant Solutions	Woodforest	32
2	3286000000008094	Two Trees Inn	Mashantucket Pequot Gaming Enterprise	Cancelled	7011.0	NaN	Primary Merchant Solutions	Woodforest	32
3	3286000000008094	Two Trees Inn	Mashantucket Pequot Gaming Enterprise	Cancelled	7011.0	NaN	Primary Merchant Solutions	Woodforest	32
4	3286000000008094	Two Trees Inn	Mashantucket Pequot Gaming Enterprise	Cancelled	7011.0	NaN	Primary Merchant Solutions	Woodforest	32
...	...	...	...	...	...	...	...	...	
4913285	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975.0	NaN	Merchants Bancard Network, Inc	Woodforest	54
4913286	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975.0	NaN	Merchants Bancard Network, Inc	Woodforest	54
4913287	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975.0	NaN	Merchants Bancard Network, Inc	Woodforest	54
4913288	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975.0	NaN	Merchants Bancard Network, Inc	Woodforest	54
4913289	5482986000100146	BELTONE MODESTO	DENNIS THOMAS	Active	5975.0	NaN	Merchants Bancard Network, Inc	Woodforest	54

4913290 rows × 43 columns

In [32]:
```
merge_1['current Status'].unique()
```

Out[32]:
```
array(['Cancelled', 'Active', 'MS - Inactive',
       'MS - Conducted Welcome Call', 'Diverted'], dtype=object)
```

In [33]:
```
merge_1['Bank'].unique()
```

Out[33]:
```
array(['Woodforest', 'BBVA', 'Headquarter', 'NMC', 'LL Group Wells Fargo',
       'Wells Meritus', 'ACH/DD Only', 'NCAL', 'TransNational',
       'BMO - Harris', 'MOC', 'PET', 'MER', 'VAN', 'AIB', 'BOV', 'OPL',
       'NAB', 'Wells Fargo'], dtype=object)
```

In [34]:
```
#local folder

print(datetime.now().strftime("%H:%M:%S"), "  --  Start local save")
merge_1.to_csv(local_path + "Global Merchant Portfolio_rev 10.0 add zeus CRM.csv") #, shee
print(datetime.now().strftime("%H:%M:%S"), "  --  End local save")
```

```
10:10:33   --  Start local save
10:14:14   --  End local save
```

In [35]:
```
#share drive (R drive)

#print(datetime.now().strftime("%H:%M:%S"), "  --  Start shared drive save")
#merge_1.to_csv(share_path + "python_merge_9_added Houston.csv") #, sheet_name='Data')
#print(datetime.now().strftime("%H:%M:%S"), "  --  End shared drive save")
```

In [36]:
```
#import tkinter as tk

#root= tk.Tk()

#canvas1 = tk.Canvas(root, width = 300, height = 300)
#canvas1.pack()

#label1 = tk.Label(root, text="Done with python_merge_9_added Houston.csv")
#canvas1.create_window(150, 150, window=label1)

#root.mainloop()
```