



Addition Financial Project

Data Import

See Data Import



Member Data

(1661365, 24)

	Unique_Member_Identifier	EOM_TRANS_DATE	age
0	360faa7e8e345d2725f92460ca39696ad885b8768053284a51e60a620b6552c9	2022-01-31 00:00:00	
1	4ef271022b929dd2ea56f3c0e9a7efae346a07e40f6ea1515c9e115beb91b099	2022-01-31 00:00:00	
2	f868c8fb32d526c00024e408fdd712e881de936372bef85a00a59aa38e8eff45	2022-01-31 00:00:00	
3	4fe6ec028a8a5234bb02fbe97488fb9431109a292ee8bc5c9a47f0fe1d69c273	2022-01-31 00:00:00	
4	f46df5083939ce116be51669af3315c474655c936204c2b25566d05405263ace	2022-01-31 00:00:00	

	age	n_accts	n_checking_accts	n_savings_accts	n_open_loans	n_open_cds	n_open_
count	1,661,365	1,661,365	1,661,365	1,661,365	1,661,365	1,661,365	
mean	49.4313	2.5145	0.9374	1.3166	0.175	0.2153	
std	20.3868	1.3098	0.6152	0.7064	0.4443	0.833	
min	0	1	0	0	0	0	
25%	32	2	1	1	0	0	
50%	52	2	1	1	0	0	
75%	65	3	1	1	0	0	
max	111	34	7	17	4	32	

Post Import Size

Member Data = (1661365, 24)

Branch Data = (731, 12)

Grouped Branch Data = (39, 12)

Pre-Processing

Member Data Processed

Member Size = (1661365, 24)

	Unique_Member_Identifier	EOM_TRANS_DATE	age
0	360faa7e8e345d2725f92460ca39696ad885b8768053284a51e60a620b6552c9	2022-01-31 00:00:00	
1	4ef271022b929dd2ea56f3c0e9a7efae346a07e40f6ea1515c9e115beb91b099	2022-01-31 00:00:00	
2	f868c8fb32d526c00024e408fdd712e881de936372bef85a00a59aa38e8eff45	2022-01-31 00:00:00	
3	4fe6ec028a8a5234bb02fbe97488fb9431109a292ee8bc5c9a47f0fe1d69c273	2022-01-31 00:00:00	
4	f46df5083939ce116be51669af3315c474655c936204c2b25566d05405263ace	2022-01-31 00:00:00	

Grouped Member Data (Branch & Date)

Grouped Size = (703, 19)

	BranchCategory	EOM_TRANS_DATE	ATMCount	BillPaymentCount	CashCount	DraftCount
14	Altamonte Springs	2023-03-31 00:00:00	155	301	87	256
15	Altamonte Springs	2023-04-30 00:00:00	102	136	0	184
16	Altamonte Springs	2023-05-31 00:00:00	111	243	62	475
17	Altamonte Springs	2023-06-30 00:00:00	188	329	95	286
18	Altamonte Springs	2023-07-31 00:00:00	87	287	49	406
19	Apopka	2022-01-31 00:00:00	158	494	105	485
20	Apopka	2022-02-28 00:00:00	183	728	114	447
21	Apopka	2022-03-31 00:00:00	189	413	81	311
22	Apopka	2022-04-30 00:00:00	253	386	41	114
23	Apopka	2022-05-31 00:00:00	170	402	147	510
24	Apopka	2022-06-30 00:00:00	220	346	96	339

Grouped Member Data

Grouped Size = (37, 18)

	BranchCategory	ATMCount	BillPaymentCount	CashCount	DraftCount	ACHCount	Fee
0	Altamonte Springs	2,158	5,120	937	5,292	18,828	
1	Apopka	3,677	7,080	1,736	7,619	27,922	
2	Boone High School	46	150	4	34	332	
3	Clermont	2,167	5,835	1,161	6,035	22,535	
4	Colonial High School	27	49	5	78	150	
5	Downtown Campus	331	449	112	507	1,531	
6	East Orlando	3,183	8,934	1,694	9,535	39,947	
7	Edgewater High School	5	2	1	0	5	
8	Eustis	1,105	2,239	535	2,959	9,504	
9	Fern Park	1,812	4,835	889	5,595	20,093	

Post Processing Size

Member Data

Member Data = (1661365, 24)

Grouped Member Data (Branch & Date) = (703, 19)

Grouped Member Data = (37, 18)

Branch Data

Branch Data = (731, 12)

Grouped Branch Data = (39, 12)

Exploratory Analysis

See Exploratory Analysis

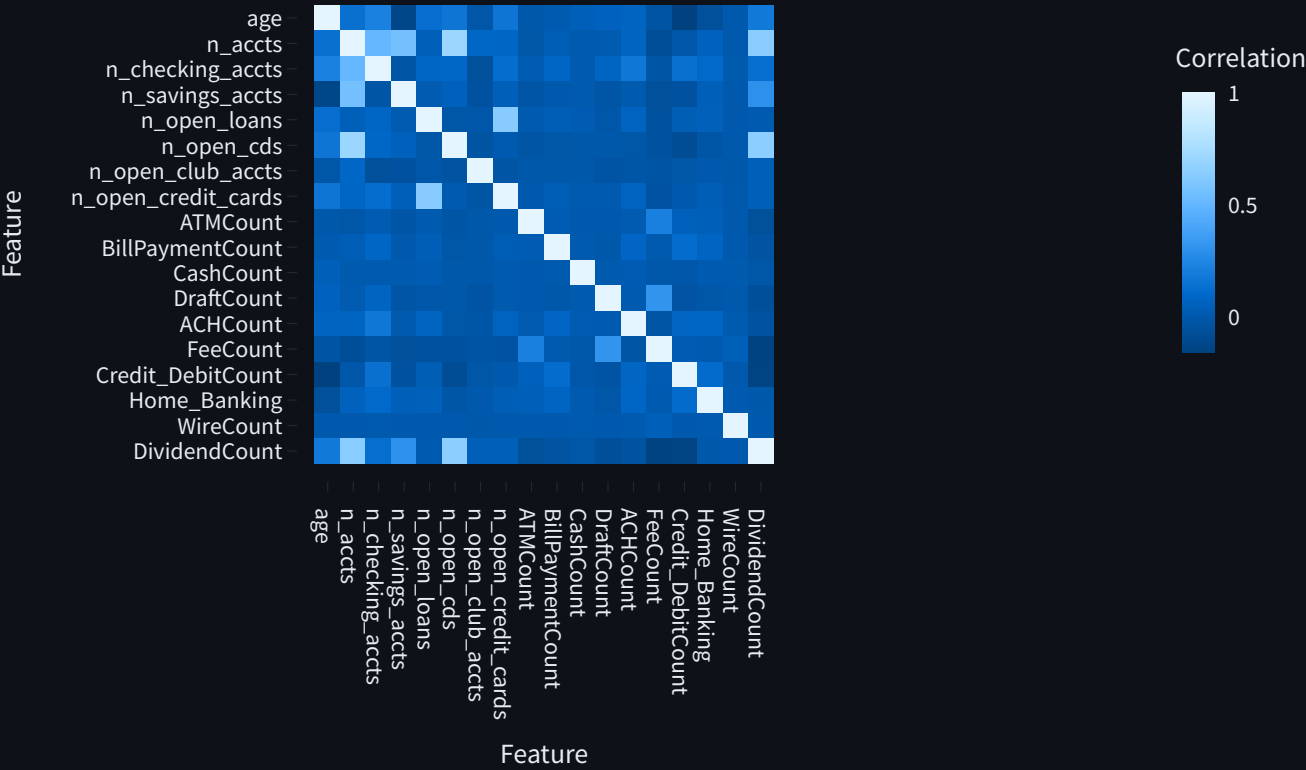
Branch Data

Member Data

Member Data

	age	n_accts	n_checking_accts	n_savings_accts	n_open_loans	n_open_cds	n_open_
count	1,661,365	1,661,365	1,661,365	1,661,365	1,661,365	1,661,365	
mean	49.4313	2.5145	0.9374	1.3166	0.175	0.2153	
std	20.3868	1.3098	0.6152	0.7064	0.4443	0.833	
min	0	1	0	0	0	0	
25%	32	2	1	1	0	0	
50%	52	2	1	1	0	0	
75%	65	3	1	1	0	0	
max	111	34	7	17	4	32	

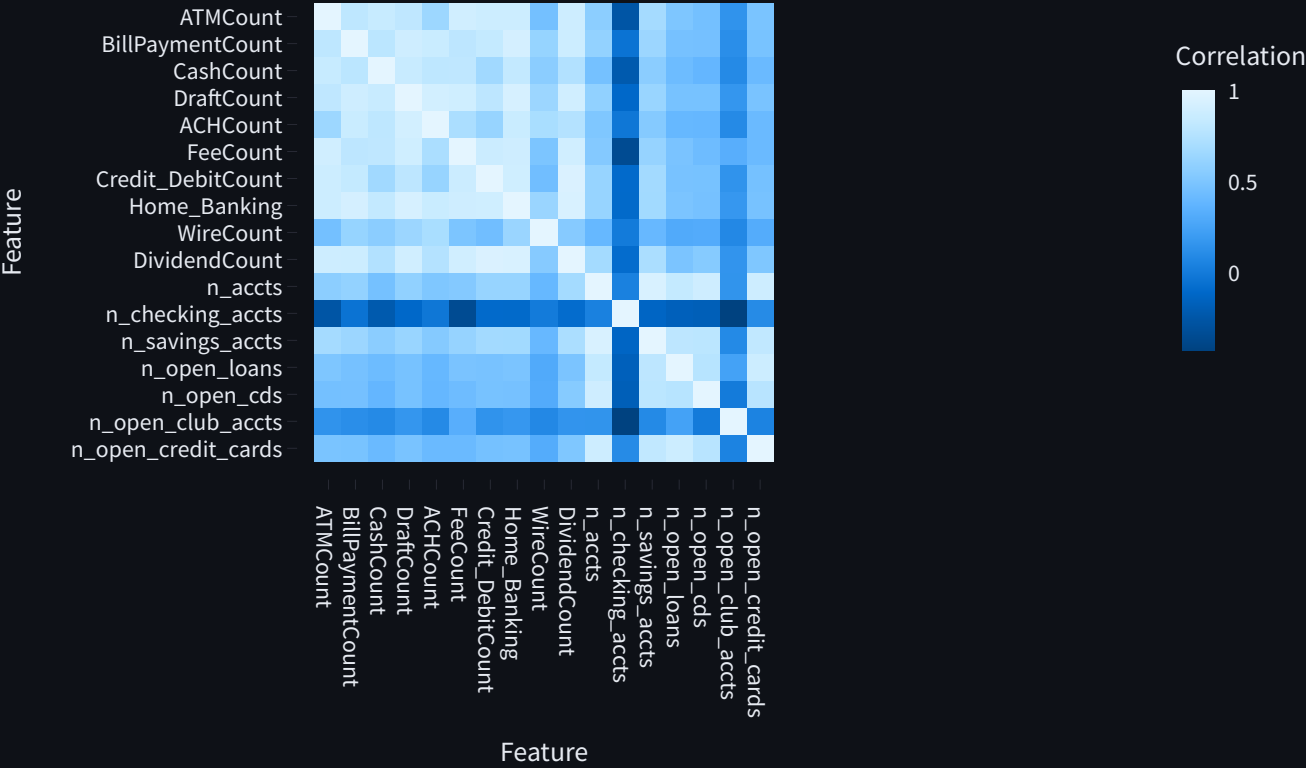
Correlation Matrix



Grouped Member Data (Branch & Date)

	ATMCount	BillPaymentCount	CashCount	DraftCount	ACHCount	FeeCount	Credit_DebitCou
count	703	703	703	703	703	703	703
mean	77.4666	168.0754	38.9417	192.973	697.1337	119.3442	1,150.514
std	74.3005	176.6726	46.1049	197.2047	872.9319	109.6924	1,108.033
min	0	0	0	0	0	0	0
25%	4	8	0	6	7	6	111
50%	61	130	20	148	305	99	93
75%	140	260	65	307	1,167.5	210.5	1,76
max	341	1,123	208	831	4,422	512	5,28

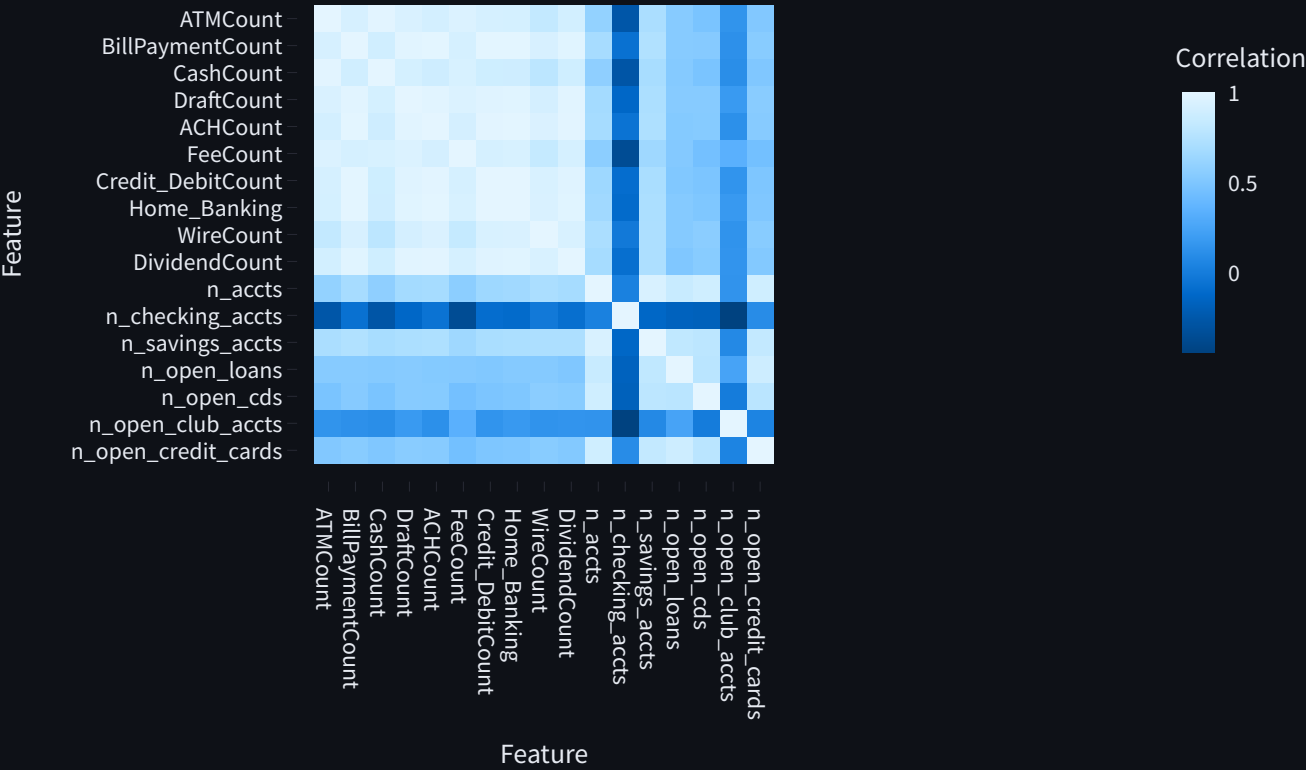
Correlation Matrix



Grouped Member Data

	ATMCount	BillPaymentCount	CashCount	DraftCount	ACHCount	FeeCount	Credit_DebitCount
count	37	37	37	37	37	37	
mean	1,471.8649	3,193.4324	739.8919	3,666.4865	13,245.5405	2,267.5405	21,859.4324
std	1,381.1688	3,019.8987	734.9749	3,402.4423	12,808.6507	2,023.5412	20,584.4324
min	5	2	1	0	5	5	
25%	60	146	12	86	393	116	1,471.8649
50%	1,169	2,495	551	3,053	10,524	1,797	18,645.5405
75%	2,915	4,963	1,571	5,792	20,050	3,960	30,584.4324
max	4,361	10,942	2,139	10,981	44,858	6,506	80,584.4324

Correlation Matrix



Branch & Member Data Time Series Alignment

See Time Series Alignment ⤴

Branch Data Columns Member Data Columns

	0
0	EOM_TRANS_DATE
1	BranchCategory
2	ATM
3	Bill Payment
4	Cash
5	Draft
6	ACH
7	Fee
8	Credit/Debit Card
9	Home Banking

	0
0	BranchCategory
1	EOM_TRANS_DATE
2	ATMCount
3	BillPaymentCount
4	CashCount
5	DraftCount
6	ACHCount
7	FeeCount
8	Credit_DebitCount
9	Home_Banking

Branch Data Types

	0
EOM_TRANS_DATE	datetime64[ns]
BranchCategory	object
ATM	int64
Bill Payment	int64
Cash	int64
Draft	int64
ACH	int64
Fee	int64
Credit/Debit Card	int64
Home Banking	int64

Member Data Types

	0
BranchCategory	object
EOM_TRANS_DATE	datetime64[ns]
ATMCount	int64
BillPaymentCount	int64
CashCount	int64
DraftCount	int64
ACHCount	int64
FeeCount	int64
Credit_DebitCount	int64
Home_Banking	int64

Modeling

Notes



correlation (matrix or pairplot)
stats(mean,median, etc)

residuals vs fitted [depends on model]

feature importance (ranking, score, coef (confusion matrix), etc) [depends on model]

Branch Data (Date Column)

Member Data (More Date Column)

Member Data - Groupby Date then Groupby Branch Category (sum all attributes) (keep)
Merge Datasets based on date

Grouping Member Data, you'll find multiple zip codes, getting a unique count for each

Find all zip codes that don't equal branch zip, count each, greatest number means that