

Identifying customer segments based on transactional behavior and demographic information

In branding and customer analytics, it is standard practise to identify customer categories based on transactional behaviour and demographic data. Businesses may modify marketing tactics, enhance customer experiences, and improve product offerings by analysing these data points to get insights into consumer preferences, requirements, and behaviours. Following is a step-by-step procedure for segmenting customers based on transactional behaviour and statistical data:

1. Importing Libraries

```
In [1]: import numpy as np
import pandas as pd
import glob
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

2. Processing the data

Loading the transactional data

```
In [2]: for num in range(132):  
        filename= f'/content/sample_data/dataset/transactions_{num}.csv'  
  
        data = pd.read_csv(filename, delimiter='|')  
        print(f'{filename}')  
        print(data)  
  
        new_file = f'/content/sample_data/dataset/new_transactions_{num}.csv'  
        data.to_csv(new_file, index=False)
```

/content/sample_data/dataset/transactions_0.csv

	cc_num	acct_num	trans_num	\
0	4218196001337	888022315787	91ab12e73ef38206e1121e9648d2408d	
1	4218196001337	888022315787	071553d533a6822a4431c354c434ddcb	
2	4218196001337	888022315787	0cfad38ef15e4749eff68dc83f62c151	
3	4218196001337	888022315787	5782693d7c70f062f258cb30bfa8900f	
4	4218196001337	888022315787	35fd7db657d7e30dd608c37f7798186e	
...	
108123	3592927524372804	632265815336	b4f6bc50e8c25083554f7a6365fbd020	
108124	3592927524372804	632265815336	a2eb246a2082a232256fdb91a115f574	
108125	3592927524372804	632265815336	58f1916fef29a345f45d3f99728b8bea	
108126	3592927524372804	632265815336	974d131d9a8894661dfd6bd034aa6511	
108127	3592927524372804	632265815336	90f6b22704e30cc8cfb7068104cc8b22	

	unix_time	category	amt	is_fraud	\
0	1558719550	gas_transport	69.12	0	
1	1569425519	grocery_pos	68.11	0	
2	1577205601	misc_net	40.35	0	
3	1571428238	grocery_pos	96.22	0	
4	1549840400	gas_transport	71.89	0	
...	
108123	1632055405	health_fitness	7.92	0	
108124	1656648528	kids_pets	7.13	0	
108125	1653812600	kids_pets	8.88	0	
108126	1668993474	personal_care	9.03	0	
108127	1640397559	home	5.29	0	

	merchant	merch_lat	merch_long
0	Phillips Group	39.491416	-75.588522
1	Tucker Ltd	40.890319	-75.573359
2	Dixon PLC	39.244958	-74.475327
3	Lambert-Cooper	39.656925	-75.802342
4	Griffith LLC	40.313342	-74.220434
...
108123	Moran, James and Jennings	41.214453	-87.737857
108124	Monroe, Martin and Johnson	42.584501	-88.876296
108125	Chang-Moore	41.315911	-87.371073
108126	Logan, Fuentes and Hoffman	42.624020	-88.170830
108127	Thompson-Wu	41.246885	-88.128293

[108128 rows x 10 columns]

/content/sample_data/dataset/transactions_1.csv

	cc_num	acct_num	trans_num	\
0	4065133387262473	950105810049	8bc87b4a938a94de66c73312a70b955e	
1	4065133387262473	950105810049	a091d0441f62b8195a005b698a9b25d7	
2	4065133387262473	950105810049	e458e6934892d3871fc49483f6744eab	
3	4065133387262473	950105810049	ec9286554b98b0f6272e7ad3e79ac46a	
4	4065133387262473	950105810049	ff9109f3fcfc870213143c7cd55dee1c	
...	
45852	3519712847818919	500229982535	55c9149ce75c1c8ce0cb2dfc5c8dcc7c	
45853	3519712847818919	500229982535	20132ba48cbdc99f4d4549e4312b502e	
45854	3519712847818919	500229982535	f3df236e33b081e59897c81dd08f4d30	
45855	3519712847818919	500229982535	b1e80670cfd826deec25b069631de3a6	
45856	3519712847818919	500229982535	286772de74e8766544de7b488f092fe3	

	unix_time	category	amt	is_fraud	\
0	1648587377	grocery_pos	43.25	0	

1	1651692302	misc_pos	98.94	0
2	1642543681	gas_transport	53.33	0
3	1650742723	misc_net	85.15	0
4	1650746122	misc_net	61.35	0
...
45852	1670203235	kids_pets	3.02	0
45853	1654748840	shopping_net	8.24	0
45854	1668827415	kids_pets	1.00	0
45855	1657111141	shopping_pos	1.38	0
45856	1663161443	health_fitness	7.97	0
		merchant	merch_lat	merch_long
0	Olson, Montgomery and Vaughn		42.070264	-83.150048
1	Mills Ltd		43.586871	-83.310010
2	Reynolds-Gamble		42.564447	-83.768322
3	Rivera-Evans		43.403984	-82.304605
4	Santiago Group		41.733693	-82.953968
...
45852	Espinoza, Allen and Ramos		36.709414	-77.110391
45853	Hendrix PLC		36.451140	-78.815838
45854	Fuller-Hodge		36.424812	-77.777825
45855	Thomas, Moore and Porter		35.158257	-77.847093
45856	Schwartz Ltd		36.440027	-78.247122

[45857 rows x 10 columns]

/content/sample_data/dataset/transactions_2.csv

	cc_num	acct_num	trans_num \
0	180032174357488	790229857846	0a4c98ebb051e261cc9bfb6e613f1836
1	180032174357488	790229857846	a28189df860527740af3659d5ecd388c
2	180032174357488	790229857846	c6c3ab5903a35055218c2d230c46b9ce
3	180032174357488	790229857846	05961e3f5b0554340229d8782111626c
4	180032174357488	790229857846	e01ce2b309474c4a7f97cbdedab26e7e
...
23355	4646726368154185	44386431879	66fe860ddf8077e7bcc7d60606893c92
23356	4646726368154185	44386431879	319088b8fc37909de223cd7beb6177c2
23357	4646726368154185	44386431879	c70810355ef6878b86c34dbb6f8efbae
23358	4646726368154185	44386431879	7cfcf2b4d3453c594612e825a823ad72
23359	4646726368154185	44386431879	c2b1d6abaf4e734e7c6484311d936d82

	unix_time	category	amt	is_fraud	merchant
0	1561474903	gas_transport	83.57	0	Howard and Sons
1	1571490659	grocery_pos	49.35	0	Thomas-Harmon
2	1550168294	gas_transport	71.08	0	Stone-Mcdonald
3	1552668951	gas_transport	42.14	0	Ward, Ferrell and Ward
4	1554572400	gas_transport	63.92	0	Clark, Boyd and Rosales
...
23355	1612006901	shopping_pos	4.30	0	Green and Sons
23356	1632054304	food_dining	9.45	0	Roberts-Guerra
23357	1639017649	health_fitness	6.53	0	Sanchez Group
23358	1649415714	home	9.40	0	Bell Ltd
23359	1637113534	health_fitness	4.41	0	Koch-Escobar

	merch_lat	merch_long
0	40.266129	-73.331157
1	41.241958	-74.315249
2	41.661700	-74.406380

3	39.828078	-74.410418
4	40.651339	-73.964047
...
23355	35.486360	-84.845374
23356	34.293966	-85.474455
23357	34.625914	-86.131112
23358	34.566233	-85.004455
23359	34.995944	-84.248997

```
[23360 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_3.csv
```

Empty DataFrame

```
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
```

Index: []

```
/content/sample data/dataset/transactions 4.csv
```

	cc_num	acct_num	trans_num \
0	4490797796124	802714479522	550713c1e27c321e318dd29cad98562d
1	4490797796124	802714479522	1641826f2b0a514acef428ba699f8378
2	4490797796124	802714479522	953a5cda97f4c4f5182fb2a1bc1b0a35
3	4490797796124	802714479522	bbff06ecd49a41be36d50dcb26111492
4	4490797796124	802714479522	50a37f49b6d5499eed447820c1eaac4
...
8757	30306396384526	278085326817	cf3e6bc11e99994ded018b446e118a61
8758	30306396384526	278085326817	ad47580a8c0e3dd94b88a2c8c138e85b
8759	30306396384526	278085326817	8117444aa0260c5d6bf6e938a41e89b9
8760	30306396384526	278085326817	53b67a02e9aa415e8d836ad957ff37f8
8761	30306396384526	278085326817	0bed1913a1409ff9cdd8e71aba82a887

	unix_time	category	amt	is_fraud	merchant
\	1558209673	gas_transport	47.91	0	Kennedy Group
1	1582636924	misc_net	66.40	0	Fischer and Sons
2	1608842400	shopping_pos	59.49	0	Nelson PLC
3	1600193977	shopping_net	51.75	0	Brown, Wilson and Ritter
4	1577408337	misc_net	40.70	0	Wiggins-Ward
...
8757	1615166857	entertainment	8.69	0	Miller PLC
8758	1665490007	personal_care	1.04	0	Hughes-Chavez
8759	1651819143	kids_pets	2.23	0	Lopez, Wilson and Rodriguez
8760	1639817992	home	1.38	0	Sheppard and Sons
8761	1616416527	entertainment	1.89	0	Garrett, Grant and Parker

```

      merch_lat merch_long
0      35.493488 -78.905707
1      35.711074 -80.162494
2      35.854523 -79.568793
3      35.131479 -79.708353
4      34.498605 -80.309223
...      ...      ...
8757  43.690830 -68.433787
8758  44.958377 -68.435919
8759  44.024758 -68.530387
8760  43.889071 -68.273949
8761  45.246473 -69.243795

```

```
[8762 rows x 10 columns]
```

/content/sample_data/dataset/transactions_5.csv

	cc_num	acct_num	trans_num	\
0	4192832764832	718172762479	8a4d89c0ae4b4a71c152228044c3b961	
1	4192832764832	718172762479	c87f888a1fd591da993897d6cf4e9952	
2	4192832764832	718172762479	8cd70b485a39d65dcfa40e44de20e17a	
3	4192832764832	718172762479	792aeaf8e6ba696fcd3eca5d34240999	
4	4192832764832	718172762479	c0bc143521b61d39d65d2af206fd7388	
...	
43677	4201027621267969279	401907373048	ca9e2800b5fe2c45a95cf02d278b9220	
43678	4201027621267969279	401907373048	323b672c49d492dda1b369d3926b83a7	
43679	4201027621267969279	401907373048	31ee88e75b9abfd9d417761adabc6059	
43680	4201027621267969279	401907373048	86b18493f7fe301b938f03e6eb7b0631	
43681	4201027621267969279	401907373048	95838cfd39619673162fec65f918be90	

	unix_time	category	amt	is_fraud	\
0	1641665906	gas_transport	49.97	0	
1	1647649712	grocery_net	55.14	0	
2	1650124751	grocery_pos	78.27	0	
3	1650751913	grocery_pos	77.22	0	
4	1645829935	gas_transport	72.08	0	
...	
43677	1669294296	home	4.36	0	
43678	1671433473	personal_care	4.51	0	
43679	1668771977	home	3.03	0	
43680	1664885395	health_fitness	1.09	0	
43681	1664335616	health_fitness	8.48	0	

	merchant	merch_lat	merch_long
0	Schneider-Gibson	33.177548	-80.657003
1	Maynard, Parrish and Harrison	33.186199	-81.561961
2	Stephens, Shelton and Graham	33.536041	-80.204725
3	Johnson-Benton	33.992088	-80.101204
4	Bell, Campos and Orr	32.710811	-80.119715
...
43677	Hall, Williams and Smith	34.808224	-85.016350
43678	Ortega-Porter	34.609911	-84.306357
43679	Roberts PLC	35.814669	-84.937601
43680	Reyes, Houston and Ingram	35.835944	-84.322738
43681	Smith-Maynard	34.708033	-85.938286

[43682 rows x 10 columns]

/content/sample_data/dataset/transactions_6.csv

	cc_num	acct_num	trans_num	\
0	676326113989	361825349415	9af15d890839df0ec68921ab2dbd4532	
1	676326113989	361825349415	b057598e15ff0306177042210b55a2e6	
2	676326113989	361825349415	3d6b4bf271df2b2af148ca65688b0906	
3	676326113989	361825349415	3e2c50d7e437bcf94d7c26677afc128b	
4	676326113989	361825349415	015f68a3736d0f4fc85e602353284036	
...	
4375	676326113989	361825349415	715face6363ebde21438efa6dc92d8b5	
4376	676326113989	361825349415	4a9ef60ce4f108ec7615a1d0e81106fe	
4377	676326113989	361825349415	5044c813273dd4f72cbe1be47ee7a3cd	
4378	676326113989	361825349415	85c83c34d6fa8678c5b1452e61c097f1	
4379	676326113989	361825349415	c17b509da040dad8e5be16fb57bd7170	

	unix_time	category	amt	is_fraud	\
0	1606224811	shopping_pos	51.74	0	

1	1558117716	gas_transport	53.74	0
2	1555271952	gas_transport	58.42	0
3	1584657322	misc_net	55.22	0
4	1561817570	gas_transport	80.43	0
...
4375	1648973306	kids_pets	4.22	0
4376	1640856169	home	2.34	0
4377	1667561835	personal_care	9.83	0
4378	1635154316	health_fitness	3.49	0
4379	1629173176	health_fitness	9.43	0
		merchant	merch_lat	merch_long
0	Flores, Bennett and Morrow		42.944651	-76.511538
1	Bell, Nielsen and Lawrence		42.075807	-77.495079
2	Hale-Jones		41.887659	-76.040542
3	Goodman-Griffith		42.497446	-76.868471
4	Harris, Curtis and Nash		43.702032	-76.658720
...
4375	Harris, Hunter and Jimenez		42.064270	-76.779892
4376	Woodard, Henson and Beltran		42.397125	-76.912150
4377	Reyes and Sons		41.881057	-76.006675
4378	Thomas-Martinez		42.113324	-75.804057
4379	Mcgee, Wheeler and Patterson		42.980819	-75.690579

[4380 rows x 10 columns]

/content/sample_data/dataset/transactions_7.csv

	cc_num	acct_num	trans_num	\
0	4514627048281480	888335239225	bc0e07fc536edd091733a497ccc93632	
1	4514627048281480	888335239225	fd02b1c2a5cccebd7440f79e47130596	
2	4514627048281480	888335239225	45da402c86bc81c2822d211647bbaee7	
3	4514627048281480	888335239225	a0bcfd87dc3865f4775c792b083dbb	
4	4514627048281480	888335239225	75e087d7ad585e72c93c73807162b4f4	
...
3946	5227674783667429	293218817973	29ded59b0983f8f2baa6143e0f3697c6	
3947	5227674783667429	293218817973	52289f9387730d365b0b2c731ec108db	
3948	5227674783667429	293218817973	4437ce90cdfa47c29cbbbe2e8ba8e91a	
3949	5227674783667429	293218817973	2a7cc2bc13ae6144759d6b4aea789128	
3950	5227674783667429	293218817973	3d05f3d6c482c66359394dfe3e3c9e9a	

	unix_time	category	amt	is_fraud	merchant	\
0	1654698170	shopping_net	68.96	0	Lee, Moore and Ortega	
1	1652223082	misc_net	79.96	0	Lopez Group	
2	1642611256	gas_transport	99.82	0	Smith-James	
3	1648851832	grocery_pos	54.59	0	Kim Inc	
4	1646404650	gas_transport	46.76	0	Gray, Eaton and Brown	
...
3946	1671278689	personal_care	9.20	0	White-Martin	
3947	1671953405	travel	9.78	0	Valdez Inc	
3948	1665231829	home	1.85	0	Cordova-Walker	
3949	1670667495	personal_care	3.33	0	Hayes, Fleming and Perez	
3950	1665748035	home	6.28	0	Greene-Harris	

	merch_lat	merch_long
0	30.943815	-82.346113
1	30.822666	-83.294197
2	30.721466	-81.953963
3	31.470061	-82.246027

```

4      31.459783 -82.259630
...      ...
3946  35.145819 -84.148263
3947  34.083742 -85.776226
3948  34.743825 -84.270978
3949  34.576165 -86.026268
3950  35.281210 -85.074635

```

```
[3951 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_8.csv
```

```

      cc_num      acct_num      trans_num \
0      180075851268619  406674938298  65e70cda0ed5404aa5a0a23a29c4fc4a
1      180075851268619  406674938298  014fc38f9497a5161138184803ea0198
2      180075851268619  406674938298  2dc312d56e88080af894cb282ccf2bf5
3      180075851268619  406674938298  d879251f9157585d3d25b84e937281a4
4      180075851268619  406674938298  6109470955b860339f76f42dd31a7dfb
...      ...
36534  4361337605230458  349734538563  d28dd9e760375a7138cf79f775e48f22
36535  4361337605230458  349734538563  0d5daf8a8e20f7612bc838c779490a4a
36536  4361337605230458  349734538563  b21c4c0145f0ab08ef65f3925f46bb6f
36537  4361337605230458  349734538563  5a8e1ed6c8141abc8cc5e518b2b82776
36538  4361337605230458  349734538563  daf1eb54b806b553c5be2b80881e7dae

```

```

      unix_time      category      amt      is_fraud \
0      1645535543      gas_transport  105.99      0
1      1643041238      gas_transport  60.40      0
2      1649709546      grocery_pos   70.69      0
3      1646177416      grocery_net   61.95      0
4      1650122092      grocery_pos   62.78      0
...      ...
36534  1669977096      kids_pets     2.41      0
36535  1668155467      home         3.57      0
36536  1665132559      health_fitness  5.04      0
36537  1663560039      food_dining   5.59      0
36538  1661085581      entertainment  1.31      0

```

```

      merchant      merch_lat      merch_long
0      Gomez-Salinas  46.639550  -122.725185
1      Perez Group   46.825744  -121.333638
2      Gonzalez, Yang and Lee  47.882733  -122.170536
3      Peterson-Gonzales  46.203355  -122.181655
4      Adkins-Strickland  47.262153  -121.543893
...      ...
36534      Brooks Inc  41.778975  -79.203435
36535      Anthony-Hernandez  41.518858  -81.006001
36536      Golden, Morales and Farmer  43.026416  -81.019862
36537      Ross, Davis and Stanley  42.854417  -80.031098
36538      Gomez-Black  41.916547  -79.197116

```

```
[36539 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_9.csv
```

```
Empty DataFrame
```

```
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
```

```
Index: []
```

```
/content/sample_data/dataset/transactions_10.csv
```

```
Empty DataFrame
```


Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_11.csv

	cc_num	acct_num	trans_num	\
0	6584732365337749	843643900137	1c3d17397d7149e2a955b4b2423b0748	
1	6584732365337749	843643900137	c62ca420fef5ea197b496f861af29cd9	
2	6584732365337749	843643900137	54bcc5a58d8bf272c64be31f2420e87e	
3	6584732365337749	843643900137	d660da6115de2a5b4621f5e64af9087f	
4	6584732365337749	843643900137	e5901413138fc03a46fb79ddad1c7122	
...	
1434	6584732365337749	843643900137	c76318fbc6da3df3fd05bd9792569247	
1435	6584732365337749	843643900137	f44df6522d0e47fa85040302aaa638b6	
1436	6584732365337749	843643900137	323f46e15bb1abd5b5cf166d3a4cc5ee	
1437	6584732365337749	843643900137	822b09773728785f57da86ea55c1d810	
1438	6584732365337749	843643900137	0cd080ad84d246760bfdb68e9971cf2f	

	unix_time	category	amt	is_fraud	\
0	1645882376	shopping_pos	9.38	1	
1	1645888457	gas_transport	1119.10	1	
2	1646050072	home	573.21	1	
3	1646052298	personal_care	553.55	1	
4	1646047130	travel	23.62	1	
...	
1434	1665309982	health_fitness	6.82	0	
1435	1670036724	kids_pets	3.32	0	
1436	1657633210	shopping_pos	3.00	0	
1437	1671427385	personal_care	3.18	0	
1438	1657177587	shopping_net	5.05	0	

	merchant	merch_lat	merch_long
0	Huffman-Washington	44.251847	-91.168351
1	Morrison-Brown	44.433155	-92.259518
2	Pierce, Gardner and Rich	42.990796	-91.185804
3	Farley, Herrera and Phillips	43.232627	-91.020145
4	Smith-Aguilar	44.340716	-90.968698
...
1434	Willis-Matthews	42.593388	-90.880847
1435	Green-Montgomery	42.594710	-90.929426
1436	Johnson, Henry and Williams	44.297690	-91.058760
1437	Ray-Torres	42.835389	-91.921160
1438	Lambert Inc	42.947908	-92.186299

[1439 rows x 10 columns]

/content/sample_data/dataset/transactions_12.csv

	cc_num	acct_num	trans_num	\
0	4896331812335761701	149852234418	f3ec0819590302134f03ffdc2f44697f	
1	4896331812335761701	149852234418	c1607c993e41f2c3b42d72d1506bef7b	
2	4896331812335761701	149852234418	6f530db25d20fe351249a54491fd3fde	
3	4896331812335761701	149852234418	6d11805f2acd938fec99376001afafe8	
4	4896331812335761701	149852234418	605342f297c575cb1ccf2c08cad082ee	
...	
1069	4896331812335761701	149852234418	a364d5f15455c18ce1e67be537cdcb10	
1070	4896331812335761701	149852234418	92485a9bffb538ec786d2ba7fabbe1f9	
1071	4896331812335761701	149852234418	edf9841d2d3a4811bd1095daf10d2b1e	
1072	4896331812335761701	149852234418	4e948b5f09059ecf1b055e7dec7de70c	
1073	4896331812335761701	149852234418	6ebc93610fcb26c458718ed6135b3f72	

	unix_time	category	amt	is_fraud	\
0	1646060228	gas_transport	65.17	0	
1	1644848624	gas_transport	47.58	0	
2	1645632153	gas_transport	64.43	0	
3	1645311286	gas_transport	82.47	0	
4	1641571926	gas_transport	50.28	0	
...
1069	1662799728	food_dining	8.45	0	
1070	1652606637	misc_pos	6.98	0	
1071	1659163922	shopping_pos	2.37	0	
1072	1662696738	food_dining	2.42	0	
1073	1662293196	food_dining	7.25	0	

	merchant	merch_lat	merch_long
0	Larson, Ryan and Huang	38.143430	-90.327335
1	Myers-Reed	39.119498	-90.760379
2	Baker-Bullock	39.384368	-90.361517
3	Spencer-Hall	39.443567	-89.752400
4	King, Rodriguez and Hancock	38.857278	-89.609525
...
1069	Hughes-Glenn	39.162530	-91.208575
1070	Bryant LLC	39.564906	-89.882200
1071	Mccarthy, Pierce and Valentine	38.771408	-89.703457
1072	Rodriguez, Taylor and Nguyen	38.899203	-91.100170
1073	Moore and Sons	37.938364	-90.454253

[1074 rows x 10 columns]

/content/sample_data/dataset/transactions_13.csv

	cc_num	acct_num	trans_num	\
0	347138067130732	199437219108	b6cb7abe1bc9ad9fa86c461c174fa4e8	
1	347138067130732	199437219108	c4e32dc6d1c0f68501636e2730d613be	
2	347138067130732	199437219108	1a9effbcbf75d1181a564c05feb571614	
3	347138067130732	199437219108	c5aa27a125d333c63b494ca61b2d3021	
4	347138067130732	199437219108	794c84f0b003b61e00b478f6153d9c48	
...
2501	6011969874869415	976596062877	97b863f99dd3cadede6ca601b72e2cb6	
2502	6011969874869415	976596062877	4b0ba0d002e572c128ed85ccc83fcfb1	
2503	6011969874869415	976596062877	e79337d9fd3cc284182b04e66a1db89f	
2504	6011969874869415	976596062877	a6f79fc5bab4c1e954d033e1427e54c5	
2505	6011969874869415	976596062877	dea4b39761d91c4e848ae5c547e61d7d	

	unix_time	category	amt	is_fraud	merchant
\					
0	1654464208	shopping_net	65.29	0	Lam-Kent
1	1646770233	grocery_pos	65.53	0	Hoffman-Martinez
2	1652553951	misc_pos	71.96	0	Rush Group
3	1649344489	grocery_pos	67.72	0	Davis LLC
4	1654194345	shopping_net	71.31	0	Castro-Wyatt
...
2501	1661670130	food_dining	451.02	0	Morgan, Blake and Hardy
2502	1666937558	home	567.34	0	Jimenez PLC
2503	1665887033	home	485.78	0	Walker, Davis and Monroe
2504	1658575219	entertainment	502.96	0	Holt-Carter
2505	1666417125	home	478.51	0	Arnold-Kane

merch_lat merch_long

```

0      35.127424 -85.628540
1      34.538931 -86.731857
2      34.948486 -85.065414
3      33.498267 -84.985591
4      33.374870 -85.680996
...
2501   33.040991 -93.282939
2502   33.801778 -93.248803
2503   33.015969 -93.261945
2504   34.146619 -93.335327
2505   33.021204 -93.925613

```

[2506 rows x 10 columns]

/content/sample_data/dataset/transactions_14.csv

```

      cc_num      acct_num      trans_num \
0      3531261685071052  897566796191  92b8567ed8083c8f7a961b714b0c0e95
1      3531261685071052  897566796191  66307bd29d2eb80a602d166a5386d7f4
2      3531261685071052  897566796191  433edec3965b52d6133ab4500eb92824
3      3531261685071052  897566796191  370c812827fe5c53437b2a7354dd010f
4      3531261685071052  897566796191  e6f46188515e2464b19396ee156a3448
...
99324  5280926166201653  108886290366  8a8a73546a69a94a4f378607ab15862a
99325  5280926166201653  108886290366  544f8deeebaab28e1f5c594abf67554d
99326  5280926166201653  108886290366  1b885b7c62cdbfb46599595cddd76a6a
99327  5280926166201653  108886290366  d7f143f433dd08dc7b3856985d9fde0f
99328  5280926166201653  108886290366  bb82ce0bf6feedcef562f08bdf308135

```

```

      unix_time      category      amt  is_fraud \
0      1671807596      grocery_pos    11.01      1
1      1671801586      gas_transport  290.72      1
2      1671813269      grocery_pos    295.04      1
3      1671805643      grocery_pos    323.37      1
4      1671808961      grocery_pos    303.03      1
...
99324  1662619661      personal_care    6.79      0
99325  1630234167      health_fitness    7.52      0
99326  1634865177      home            10.00      0
99327  1662731332      personal_care    6.30      0
99328  1671595070      travel          3.40      0

```

```

      merchant      merch_lat  merch_long
0      Snyder, Smith and Mcdonald  42.124712  -91.669751
1      Perez-Brown      42.574693  -92.050436
2      Johns Inc        41.533761  -90.825622
3      Webster-Austin   42.727905  -91.198276
4      Thompson-Stafford  42.903856  -91.146959
...
99324      Santiago Inc  36.951204  -115.184555
99325      Roberts Group  37.016928  -115.343758
99326      Short Ltd     35.427305  -115.633600
99327      Ortiz, Jackson and Welch  37.080096  -114.279567
99328      Butler Inc    35.553051  -115.802843

```

[99329 rows x 10 columns]

/content/sample_data/dataset/transactions_15.csv

```

      cc_num      acct_num      trans_num \
0      675986887866  987989291479  8dacd18c72c38612270ad9f8ef82ab1b

```

1	675986887866	987989291479	902c2b932505c1662144d32c8202c833
2	675986887866	987989291479	0587accc799e08e3a0a906f6dbdb1b0e
3	675986887866	987989291479	47c7d6220fd72f804db1804f65772837
4	675986887866	987989291479	ebc1386e9dccbe7bcaf314888cf40ac6
...
1455	675986887866	987989291479	c5329f75af098d80621e27abe7de1abe
1456	675986887866	987989291479	62eb81510f1ddb176ab06b7dc956914f
1457	675986887866	987989291479	8589d7898b9912f7f8441057f05ad704
1458	675986887866	987989291479	5c75106a01c68fecc50c7b8ba9e2aa54
1459	675986887866	987989291479	ce4648ec89717e00090df2578fed6980

	unix_time	category	amt	is_fraud	\
0	1559687300	gas_transport	68.55	0	
1	1561732187	gas_transport	59.12	0	
2	1562509875	gas_transport	67.97	0	
3	1567465604	grocery_net	86.73	0	
4	1577806875	grocery_pos	65.08	0	
...
1455	1653545472	health_fitness	8.25	0	
1456	1614763746	shopping_pos	175.61	0	
1457	1638583783	food_dining	8.96	0	
1458	1656754621	home	7.36	0	
1459	1634381810	food_dining	6.22	0	

	merchant	merch_lat	merch_long
0	Whitney-Moore	38.143625	-89.276767
1	Mason and Sons	37.747138	-89.206174
2	English, Trevino and Bryant	37.148030	-88.884998
3	Watson, Fisher and Williams	37.579065	-88.879816
4	Perez, Fleming and Smith	37.016670	-88.802980
...
1455	Little Inc	37.501431	-88.308732
1456	Clark and Sons	37.862228	-88.629577
1457	Davis, Jones and Paul	37.904766	-87.803223
1458	Stephenson-Lopez	37.530567	-88.538473
1459	Higgins and Sons	36.583893	-88.466304

[1460 rows x 10 columns]

/content/sample_data/dataset/transactions_16.csv

	cc_num	acct_num	trans_num	\
0	4891077119918081211	402211201664	232dd835841f8f3acedbc2066597875c	
1	4891077119918081211	402211201664	039f2e2bf57a6cdbccb124e8d55ae9d9	
2	4891077119918081211	402211201664	8c90717dfa02a7fb5bd79ccfb9c80670	
3	4891077119918081211	402211201664	f1a77d09a334c50f7b5a523e59697e69	
4	4891077119918081211	402211201664	f002286bee82ab48ec68b0f723697ff9	
...
10226	3534679121416143	103258999645	23540ddb4eb750360bff8e9d3eb667f3	
10227	3534679121416143	103258999645	09c674f1427cfa95b8e6ddc464224694	
10228	3534679121416143	103258999645	d8209986ac5b4d0cf97f0eab2b56aba9	
10229	3534679121416143	103258999645	16737fba0d1bc2894c287724c0fdb570	
10230	3534679121416143	103258999645	588ed2336bba6aa43964121e2ae57ddd	

	unix_time	category	amt	is_fraud	\
0	1576966714	grocery_pos	62.15	0	
1	1557010309	gas_transport	49.49	0	
2	1585754712	misc_net	67.48	0	
3	1551396617	gas_transport	62.01	0	

4	1585425018	misc_net	34.66	0
...
10226	1620902323	entertainment	3.80	0
10227	1657869554	kids_pets	7.59	0
10228	1608809405	shopping_pos	2.57	0
10229	1625630522	food_dining	4.32	0
10230	1663496308	personal_care	8.78	0

		merchant	merch_lat	merch_long
0		Miller and Sons	38.572578	-81.774686
1		Reynolds-Bryant	40.074063	-81.643343
2		Porter PLC	40.383163	-81.745601
3		Blackburn and Sons	39.488174	-81.957905
4		Valdez Inc	39.149621	-81.795189
...	
10226		Jackson LLC	61.209822	-158.430297
10227		Rivera-Ray	62.288254	-158.324646
10228	Caldwell, Valentine and Sellers		61.120091	-157.476436
10229	Turner-Poole		62.196653	-157.272488
10230	Caldwell-Burnett		61.606561	-157.408837

[10231 rows x 10 columns]

/content/sample_data/dataset/transactions_17.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_18.csv

	cc_num	acct_num	trans_num	\
0	630447468723	940251608167	252687585221df5dc6e4954f1eb9f410	
1	630447468723	940251608167	4b77b098972d0ba7243283ea226ef896	
2	630447468723	940251608167	f79a211254333abdd107e222d30a2a32	
3	630447468723	940251608167	cc23d5dc069a6181047fcafd118f8db8	
4	630447468723	940251608167	5b2724c12c03e725a98d0fc13c1b098f	
...
7888	379902355770272	167373214626	4cbbeec4dc900d1573abafa5639dfd5c	
7889	379902355770272	167373214626	36e92e2242565189e40629f1ba58d895	
7890	379902355770272	167373214626	856c41011b8ccf773752aca1cd888e21	
7891	379902355770272	167373214626	85a2091585ca3368ce1679fe080c7e37	
7892	379902355770272	167373214626	e1339031a70569295694204be9113cdd	

	unix_time	category	amt	is_fraud	merchant
\					
0	1643289027	grocery_pos	287.75	1	Lam Inc
1	1643383555	misc_net	288.23	1	Martin Ltd
2	1643297375	grocery_pos	300.00	1	Ferguson-Gonzalez
3	1643289214	grocery_pos	288.53	1	Robinson LLC
4	1643290844	misc_net	338.48	1	Gonzalez Inc
...
7888	1670742520	personal_care	4.57	0	Nunez, Sandoval and Wood
7889	1655549997	shopping_net	6.42	0	Thomas-Perkins
7890	1661601758	entertainment	2.53	0	Miller Inc
7891	1664018740	food_dining	9.69	0	Gomez, Stewart and Hicks
7892	1671865187	personal_care	5.62	0	Taylor Inc

	merch_lat	merch_long
0	34.139096	-117.159437

```

1      33.258724 -117.967951
2      33.251874 -118.409882
3      34.567360 -117.388140
4      34.555893 -117.820108
...
7888   35.554274 -85.334827
7889   34.667397 -84.911506
7890   35.998485 -84.564365
7891   34.474062 -85.108057
7892   35.352308 -85.330240

```

[7893 rows x 10 columns]

/content/sample_data/dataset/transactions_19.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_20.csv

	cc_num	acct_num	trans_num \
0	4909619797978	326474132970	29629a76b171a60f6e6ff4fe64ae8988
1	4909619797978	326474132970	5f6ba65125911200eaed78c719e9c85e
2	4909619797978	326474132970	b425a7741b21c7c99d3f34d86a512bf6
3	4909619797978	326474132970	b6327192de7f3704e72d25474859faca
4	4909619797978	326474132970	e403dd5d734c35c06dae623347288b65
...
154753	180047909863618	921460407733	31ef21c712f1ee444a06751a8dfcbb56
154754	180047909863618	921460407733	0e4f4ea87366f0a6ff6a77d6d69ae5da
154755	180047909863618	921460407733	e189e882db43eea50713602b153ee223
154756	180047909863618	921460407733	d685c55c32e970233576f2f3aaaa77e3
154757	180047909863618	921460407733	0907994749e9078ef4a9e08ac2fcdd4b

	unix_time	category	amt	is_fraud \
0	1553878401	grocery_net	85.88	0
1	1579711131	misc_pos	86.60	0
2	1559159358	grocery_net	72.39	0
3	1549556432	gas_transport	50.67	0
4	1555014318	grocery_net	99.98	0
...
154753	1608361075	shopping_pos	4.88	0
154754	1590582548	shopping_net	9.51	0
154755	1669863951	personal_care	2.32	0
154756	1608254329	shopping_pos	1.91	0
154757	1607477122	shopping_pos	9.16	0

	merchant	merch_lat	merch_long
0	Zamora-Hernandez	30.066512	-90.723646
1	Simmons LLC	29.858211	-89.535477
2	Reeves-Walker	30.078241	-90.908793
3	Luna Ltd	29.410496	-89.870197
4	Patterson, Cooper and Brown	30.005284	-89.700016
...
154753	Smith Group	34.685320	-119.348250
154754	Miller-Moore	35.988114	-119.057271
154755	Joseph Inc	35.458715	-119.632409
154756	Gardner, Chan and Russell	36.060534	-120.283720
154757	Cole, Fisher and Wilson	34.910037	-119.213412

```
[154758 rows x 10 columns]
/content/sample_data/dataset/transactions_21.csv
Empty DataFrame
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
Index: []
/content/sample_data/dataset/transactions_22.csv
```

	cc_num	acct_num	trans_num	
0	3536881013827217	589419928804	c31c6e90cf191756db56d617fcceebeb	
1	3536881013827217	589419928804	126e88be5b03ddf484d89b78756c2cc7	
2	3536881013827217	589419928804	21712361deb77da6fbd08486ce1f1722	
3	3536881013827217	589419928804	0c4403834584258a8741080e3038f9af	
4	3536881013827217	589419928804	2856a9a7e1cfbed037d50a238c766dba	
...	
353	3536881013827217	589419928804	f037ed72e9f65c28d346da04d1c2f53d	
354	3536881013827217	589419928804	cd52e2a01fe907f6e1b6adddc6fe7400	
355	3536881013827217	589419928804	2e48bbccc0e658be32fa71ae64367dfc	
356	3536881013827217	589419928804	17fa12ccce8a4f9084a8607c23c8546c	
357	3536881013827217	589419928804	7d4db99aefa15e99c7650a172583e728	

	unix_time	category	amt	is_fraud	merchant
0	1648217767	grocery_pos	55.66	0	Hart, Avila and Johns
1	1641689364	gas_transport	54.25	0	Hopkins-Bradley
2	1657217626	shopping_pos	82.38	0	Butler, Parker and Gilbert
3	1650644329	grocery_pos	42.50	0	Ortiz, Malone and Barnes
4	1648651442	grocery_pos	51.89	0	Kim Group
...
353	1664426232	health_fitness	6.80	0	Davis, Morgan and Valdez
354	1667632321	home	797.46	0	Pierce Group
355	1665278189	home	8.93	0	Nolan-Mullins
356	1661770492	food_dining	5.51	0	Robertson-Rojas
357	1670414555	kids_pets	7.26	0	Scott, Mckenzie and Dalton

	merch_lat	merch_long
0	35.037296	-92.833021
1	35.363728	-92.534079
2	34.698186	-92.493184
3	36.272352	-93.227659
4	35.033026	-92.044381
...
353	36.259677	-91.808943
354	35.740895	-91.801359
355	34.749097	-91.943809
356	36.121232	-92.025980
357	35.287492	-93.034909

```
[358 rows x 10 columns]
/content/sample_data/dataset/transactions_23.csv
```

	cc_num	acct_num	trans_num	
0	347138067130732	199437219108	dd49036b665b24e2674e11d926b85e02	
1	347138067130732	199437219108	60902edf24604e129a0ce69c5ecd9965	
2	347138067130732	199437219108	b62d43dc73cd7f7863dbde722d45e420	
3	347138067130732	199437219108	05ac4d015616bdbecaed8435fadab202	
4	347138067130732	199437219108	ca34e8d10ba09ea598407cea655a1bd4	
...	
7295	6011969874869415	976596062877	55d5bce8a4fb44a738e4161e8a84354b	

```

7296 6011969874869415 976596062877 e0c275b51e16a50ce16c621c580d75d7
7297 6011969874869415 976596062877 316db17a0b0888dec9e9d29c2e1aff0e
7298 6011969874869415 976596062877 ca81fd9d698069a12aebad3474c472a1
7299 6011969874869415 976596062877 bc79c5f69d61bae52f2296c098070c45

```

```

      unix_time      category      amt  is_fraud  \
0      1576617639      misc_pos    69.68         0
1      1584650002      misc_pos    62.65         0
2      1590944755      shopping_net  70.73         0
3      1597193308      shopping_pos  77.97         0
4      1601931764      shopping_pos  72.56         0
...      ...      ...      ...      ...
7295  1661488043      kids_pets    524.40         0
7296  1661228576      kids_pets    475.71         0
7297  1631875114      health_fitness  478.33         0
7298  1644737532         home    560.93         0
7299  1671868982      travel    536.41         0

```

```

      merchant      merch_lat      merch_long
0      Vargas-Obrien    33.377761    -85.964407
1      Nichols, Bell and Sanders    33.544674    -85.323726
2      Johnson Group    33.756144    -86.099000
3      Henderson, Watkins and Rose    33.710669    -86.504719
4      Zamora-Jones    35.330195    -85.307917
...      ...      ...      ...
7295      Yu, George and Ryan    33.205210    -92.552918
7296      Norman-Hanson    34.060344    -92.933577
7297      Trevino Group    33.631931    -93.740141
7298      Brown, Watkins and Moody    32.809922    -94.270471
7299      Hanson LLC    33.197554    -93.290357

```

[7300 rows x 10 columns]

/content/sample_data/dataset/transactions_24.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_25.csv

```

      cc_num      acct_num      trans_num  \
0      4006875868240516653  935451073243  6c0df2592d98de08267ccb7a6a68c1a6
1      4006875868240516653  935451073243  5f7263ba4f3f774f7e13af6b78a94173
2      4006875868240516653  935451073243  3f82bd3a96e28eea2e81b6ad255da192
3      4006875868240516653  935451073243  a6e16acb24d845fd61a5ecb57b419200
4      4006875868240516653  935451073243  692b4310fe9e09f9c632cce5c681c2a1
...      ...      ...      ...
24771      343251790447085  298026601683  2a1e7680d91e51a60716064068577268
24772      343251790447085  298026601683  09267380cfb24c33c7048fc9e9f87a6f
24773      343251790447085  298026601683  1f70991b54083954b077b5711ebbf0a
24774      343251790447085  298026601683  096f2c10cf3589b727d39bf1a3e36916
24775      343251790447085  298026601683  5bc5506499295498f7e658b111450129

```

```

      unix_time      category      amt  is_fraud  \
0      1653062644      misc_pos    78.43         0
1      1659648627      entertainment    58.97         0
2      1654739563      shopping_net    90.59         0
3      1653070601      misc_pos    79.88         0
4      1658353827      entertainment    59.26         0

```



```

...      ...      ...      ...      ...
24771  1670908006  personal_care  2.94      0
24772  1669943329   kids_pets     6.43      0
24773  1665710553    home         6.41      0
24774  1670748794  personal_care 1252.96    0
24775  1671940413   travel        76.76     0

      merchant  merch_lat  merch_long
0      Rice Group  41.503185  -73.668424
1      Perez-Leach 40.234917  -74.028145
2      Smith PLC   40.122477  -73.712372
3      Johnson, White and Kemp 40.908808  -74.628287
4      Simmons-Hood 41.057622  -74.489124

...      ...      ...      ...      ...
24771      Monroe Ltd 28.019383  -81.794612
24772      Jimenez Group 27.065622  -82.343797
24773      Coleman-Fry 28.505515  -82.288554
24774  Snow, Johnston and Barron 27.451606  -81.956534
24775      Rose-Jones 27.860629  -82.725985

```

[24776 rows x 10 columns]

/content/sample_data/dataset/transactions_26.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_27.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_28.csv

```

      cc_num      acct_num      trans_num \
0      6011034324451072  347300118003  9d69254235162c2622a18f40428be484
1      6011034324451072  347300118003  f7a363dcd4061c5509b898062145eb4
2      6011034324451072  347300118003  53f1ece7775bd7922604359cf60ad390
3      6011034324451072  347300118003  a93fe93ef304d209954a6c439776c3e6
4      6011034324451072  347300118003  8ab8aebc64059ac090589ba2fede2939

...      ...      ...      ...
10215  6011416389448467  18557023491  088595e4705b1e8104b63b0769a61522
10216  6011416389448467  18557023491  7080ccaf3726725fbbd65125d04c872b
10217  6011416389448467  18557023491  465454d97aa3bf6b3e8798adef040ab5
10218  6011416389448467  18557023491  52cd84669bf4b041dd496d879f307981
10219  6011416389448467  18557023491  e53cddbfff8d7505310fd2a43016c29f6

      unix_time      category      amt  is_fraud \
0      1604613096  shopping_pos  67.75      0
1      1595623862  shopping_net  60.24      0
2      1556572938  grocery_net  65.82      0
3      1558486442  grocery_net  69.30      0
4      1579184692  misc_pos     76.81      0

...      ...      ...      ...      ...
10215  1626232418  food_dining  569.36      0
10216  1648353549  home        527.15      0
10217  1632556536  health_fitness  566.27      0
10218  1627638220  food_dining  603.77      0
10219  1641625053  home        429.79      0

```

	merchant	merch_lat	merch_long
0	Huff, Smith and Underwood	32.088436	-108.388786
1	White-Brown	32.763459	-108.349340
2	Cruz, Cook and Avery	32.135833	-108.979088
3	Perez, Jones and Smith	32.990473	-109.442091
4	Schroeder-Schultz	31.968292	-108.185316
...
10215	Martinez Group	38.082922	-85.562103
10216	Williams-Harrington	39.795780	-84.700282
10217	Green-Alexander	39.611931	-85.384566
10218	Hernandez Ltd	38.281048	-85.700447
10219	Ali, Newman and Porter	39.406757	-85.259344

[10220 rows x 10 columns]

/content/sample_data/dataset/transactions_29.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_30.csv

	cc_num	acct_num	trans_num \
0	180032174357488	790229857846	1cafb20a1a84ba4ef98888bb1bd9e104
1	180032174357488	790229857846	910f9d9b6a76145ef7225a54b2f4800a
2	180032174357488	790229857846	e82e4d4d0a9d6f116b15e306247efcb4
3	180032174357488	790229857846	db27226cde3b25c94742a61316011b05
4	180032174357488	790229857846	fdf36b30e3e123fcb0d49016685fbb73
...
3575	4646726368154185	44386431879	c8b3b1151ca6774cfc471a966a6ae7
3576	4646726368154185	44386431879	cd484149b737daf479cce7f02319ba4d
3577	4646726368154185	44386431879	af306a24bea2ed97109db7f92f5d011d
3578	4646726368154185	44386431879	e1050109cbac5b5102e007f7cb1d268c
3579	4646726368154185	44386431879	039897ed83867640d42167756565c845

	unix_time	category	amt	is_fraud \
0	1643297157	gas_transport	46.11	0
1	1644155219	gas_transport	58.44	0
2	1649279492	grocery_pos	53.30	0
3	1644593323	gas_transport	75.60	0
4	1646506847	gas_transport	83.86	0
...
3575	1665295134	health_fitness	3.47	0
3576	1653623794	misc_pos	3.83	0
3577	1656670373	shopping_net	3.85	0
3578	1660139117	entertainment	9.16	0
3579	1657937110	shopping_pos	2.76	0

	merchant	merch_lat	merch_long
0	Mora, Holland and James	40.666351	-74.271049
1	Smith-Hudson	40.265087	-73.968016
2	Campbell, Taylor and Frank	41.364445	-74.144440
3	Bass, Sanders and Young	40.324152	-73.190458
4	Costa-Woodard	41.769712	-73.335726
...
3575	Cox-Rodriguez	35.789782	-86.013622
3576	Graves, Young and Schneider	34.101284	-85.980451
3577	Simmons Group	34.728192	-85.290111

```

3578           Mendoza PLC  35.655890  -84.330756
3579           Cook-Arnold  34.646292  -84.866490

```

[3580 rows x 10 columns]

/content/sample_data/dataset/transactions_31.csv

```

      cc_num      acct_num      trans_num      unix_time
\
0      676326113989      361825349415      363a18bd03b717e553ec9413ffff4ec57      1658853267
1      676326113989      361825349415      90440b82da299e708b4fc52532a139ef      1658850097
2      676326113989      361825349415      82322ae6999755ded3a2c4b85b6ba57c      1658854048
3      676326113989      361825349415      b5d4e8970733a47283f033194f300994      1658845921
4      676326113989      361825349415      03049744730b57552449d4f7cb7e46a2      1659011771
..      ...      ...      ...      ...
364      676326113989      361825349415      89ed706943724343c94b16cf77a07cbf      1662299751
365      676326113989      361825349415      794dd469be04479c37a6b6140f026dc2      1659345112
366      676326113989      361825349415      122c567315bfa3a84bbe96e4a8da0c20      1660279232
367      676326113989      361825349415      905a908e5b77666a2d72c4ed6b0de01c      1660555163
368      676326113989      361825349415      7d4f881036358f0d5e3fa757deda8144      1670824901

```

```

      category      amt      is_fraud      merchant      merch_lat
\
0      gas_transport      22.51      1      Contreras PLC      42.863460
1      gas_transport      20.56      1      Walker-Casey      43.237965
2      grocery_pos      339.13      1      Nash and Sons      42.984440
3      grocery_pos      328.75      1      Carr Inc      42.702707
4      travel      324.40      1      Roth, Vasquez and Whitaker      43.399213
..      ...      ...      ...      ...
364      health_fitness      9.15      0      Decker LLC      43.544410
365      entertainment      210.83      0      Brown and Sons      42.805614
366      entertainment      4.27      0      Wheeler, Gomez and Avery      43.069129
367      entertainment      7.95      0      Lane-Craig      43.148275
368      personal_care      6.40      0      Welch, Phillips and Diaz      42.909643

```

merch_long

```

0      -76.681985
1      -76.036416
2      -76.425987
3      -77.406702
4      -76.272895
..      ...
364      -77.079779
365      -76.940384
366      -75.710918
367      -77.040719
368      -76.856706

```

[369 rows x 10 columns]

/content/sample_data/dataset/transactions_32.csv

```

      cc_num      acct_num      trans_num      \
0      30415152638403      373790097990      a4340496ad3ddcb9021da9019f320f23
1      30415152638403      373790097990      8666af3e5b17f0f00e3dc8809f282711
2      30415152638403      373790097990      6c69d7c41c2ef9876dfa45fffb42e92e2
3      30415152638403      373790097990      00ed85f84252b09a7a6dce2a1debc02e
4      30415152638403      373790097990      8144f0c46268916a4a3c73c7e8443845
..      ...      ...      ...
8755      3572989376297459      59703614299      80a0a715756da73a2f59e18226df6d06
8756      3572989376297459      59703614299      f4916b8279f260df57d5ea3cebf8884d

```

```

8757 3572989376297459 59703614299 720b8ee232fe166d50671cd6e8dcee32
8758 3572989376297459 59703614299 119a03bbc83e3428b6fdeb17a7d841db
8759 3572989376297459 59703614299 f3551e44f6b743caf82d76317ed40051

```

```

      unix_time      category      amt  is_fraud  \
0      1552669266  gas_transport  59.19         0
1      1584207641    misc_net    71.30         0
2      1571068578  grocery_pos   54.69         0
3      1588379277    misc_pos   52.36         0
4      1567537065  grocery_pos   63.96         0
...          ...          ...      ...      ...
8755  1669807343  personal_care    6.30         0
8756  1621830688   food_dining    6.55         0
8757  1671272867     travel     5.95         0
8758  1610004097  entertainment    9.66         0
8759  1614848269  entertainment    5.19         0

```

```

              merchant  merch_lat  merch_long
0      Mcguire, Nolan and Hernandez  43.404637  -84.632098
1      Copeland, Turner and Patterson  42.826796  -83.336394
2      Flynn, Gonzalez and Smith    43.416714  -83.967022
3      Miles-Dominguez              43.084354  -84.755591
4      Lopez-Jenkins                42.926032  -84.468170
...          ...          ...      ...
8755              Gallagher-Nelson  39.994475  -76.064208
8756              Thomas LLC        39.545295  -76.920495
8757              Turner-Swanson    38.907395  -75.238238
8758              Jones-Sanchez     39.973384  -75.329904
8759              Thompson-Jackson  39.082620  -75.400270

```

[8760 rows x 10 columns]

/content/sample_data/dataset/transactions_33.csv

```

      cc_num      acct_num      trans_num  \
0      2287201859626305  125069449315  84b9c7a314a0bf940c43884abb8cd35d
1      2287201859626305  125069449315  9c0a861873f5f6258a904dc955bb6892
2      2287201859626305  125069449315  5354a960f9310916f16164e17d4a64a1
3      2287201859626305  125069449315  6efa3c7b27fc87de70349bfd4586a882
4      2287201859626305  125069449315  d503d94901f1083c5391c027221f627a
...          ...          ...      ...
2501   347208496498560   11546128003  5a45673ae9ac120c8e3d3b7d25ebad61
2502   347208496498560   11546128003  3f100be4b07655cef7de797536ab7aec
2503   347208496498560   11546128003  b60834eaa7c5d1c4a06a4b6f1651fa23
2504   347208496498560   11546128003  6f52ab9daff01772bddf23b7eb76ab76
2505   347208496498560   11546128003  c7f7e0b030c061cb9e34e63f9d9e8cc3

```

```

      unix_time      category      amt  is_fraud  \
0      1644602241  gas_transport  57.70         0
1      1649112904  grocery_pos   83.59         0
2      1648648881  grocery_net   50.89         0
3      1648506490  grocery_net   75.65         0
4      1648942951  grocery_pos   56.93         0
...          ...          ...      ...
2501  1664176669   food_dining    3.66         0
2502  1664184943   food_dining    5.56         0
2503  1661573937  entertainment    8.64         0
2504  1653907733    misc_pos     3.01         0
2505  1667614459  health_fitness    9.86         0

```

	merchant	merch_lat	merch_long
0	Lamb-Hart	37.673781	-87.599695
1	Hunt, Bishop and Mcdowell	36.528417	-86.916169
2	Blackburn Inc	37.433935	-87.574003
3	Oneill-Hernandez	37.184247	-86.197379
4	Walsh and Sons	36.771541	-87.185186
...
2501	Mcintyre, Kemp and Rhodes	40.223824	-96.714782
2502	Martinez Ltd	39.730509	-97.917259
2503	Nichols PLC	39.630807	-97.340039
2504	Jacobs, Williams and Patel	40.734086	-96.178885
2505	Ferrell, Thompson and Torres	40.139678	-96.743013

[2506 rows x 10 columns]

/content/sample_data/dataset/transactions_34.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_35.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_36.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_37.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_38.csv

	cc_num	acct_num	trans_num \
0	4841962332605860	3005591724	0a9c8afdc2ca6b4bce8ecc2be7d2e935
1	4841962332605860	3005591724	e336c0796eb5df533b742430e8232faf
2	4841962332605860	3005591724	e0f09f622f8e154b1e93ae42ed8b0a46
3	4841962332605860	3005591724	ed2091c4c6c27c2d3d6a1bf1b998f9d5
4	4841962332605860	3005591724	10817e38ab04f577f9af4c4aa52762a6
...
94972	6011473650032991	272682708416	f98739a67968bbcd0db5dfef319be0
94973	6011473650032991	272682708416	176c218c4eb475a314d8f96c24a3a134
94974	6011473650032991	272682708416	bad376f22bedf7eeb9337107fd17cbb4
94975	6011473650032991	272682708416	02547d11a54578ee07f21b5cbf3e3575
94976	6011473650032991	272682708416	4d4f434185457f48849c693cfed1c711

	unix_time	category	amt	is_fraud	merchant
0	1595972601	shopping_pos	75.51	0	Warren-Nelson
1	1609268401	entertainment	76.15	0	Boone-Cruz
2	1589568716	shopping_net	78.83	0	Nguyen-Moore
3	1556754723	gas_transport	75.85	0	Hood Inc
4	1614554682	entertainment	66.32	0	Hurst, Ramirez and Cohen
...

94972	1661867636	personal_care	7.71	0	Rodriguez-Wade
94973	1671540752	travel	1.33	0	Alexander PLC
94974	1661862482	personal_care	4.70	0	Roman and Sons
94975	1661507536	personal_care	2.14	0	Mccarthy, Lopez and Allen
94976	1627977969	health_fitness	9.70	0	Porter and Sons

	merch_lat	merch_long
0	40.571481	-77.148282
1	40.978736	-76.871023
2	39.861767	-76.008964
3	39.836262	-76.962103
4	40.780294	-75.876983
...
94972	34.393317	-78.130479
94973	35.594774	-78.301459
94974	35.675828	-78.365079
94975	35.657101	-79.234460
94976	35.005893	-78.670954

[94977 rows x 10 columns]

/content/sample_data/dataset/transactions_39.csv

	cc_num	acct_num	trans_num \
0	30160398863650	652941429071	9c6f65115717064fd0349d1dbd02d9be
1	30160398863650	652941429071	f1a327c9e7a6638f9cb8199de0a7a731
2	30160398863650	652941429071	dac2627d627e8f346b364d8da8f555a3
3	30160398863650	652941429071	d14a889f8447f3cd418d45d2092fdb27
4	30160398863650	652941429071	7b904cdea24bf406502fa215ef023101
...
40882	4594853400127623	997039479727	b70a19e841df4001d42fc2ae6d8787ae
40883	4594853400127623	997039479727	21023ec33977cd3acbe6bb67d99683d6
40884	4594853400127623	997039479727	31b44d814a2f14d9c55b9676d4d53775
40885	4594853400127623	997039479727	7df193ce532b682c914e1c70e16a2a54
40886	4594853400127623	997039479727	695b2833b4f2e3bfc45b76d9e441f7c3

	unix_time	category	amt	is_fraud \
0	1650897750	grocery_pos	9.89	1
1	1650899044	gas_transport	9.89	1
2	1650903412	gas_transport	14.00	1
3	1650907403	grocery_pos	311.63	1
4	1650936529	gas_transport	291.33	1
...
40882	1665837516	health_fitness	1.99	0
40883	1669622506	home	7.13	0
40884	1654772674	misc_pos	8.95	0
40885	1660371690	entertainment	4134.95	0
40886	1655359503	shopping_net	1.52	0

	merchant	merch_lat	merch_long
0	Vazquez LLC	40.273363	-75.822797
1	Hall PLC	39.345601	-75.875301
2	Johnson LLC	40.812858	-75.677036
3	Turner LLC	39.817482	-74.938757
4	Rivera, Chavez and Davis	39.207293	-74.229907
...
40882	Hall Group	30.699113	-98.531273
40883	Wade-Rodriguez	29.464685	-97.663554
40884	Rojas, Barnes and Erickson	30.393799	-97.872988

40885 Sherman-Daniels 30.516804 -97.664019
 40886 Murphy, Fischer and Woods 30.567516 -97.793668

[40887 rows x 10 columns]

/content/sample_data/dataset/transactions_40.csv

	cc_num	acct_num	trans_num \
0	377480957394058	256845189467	5c010d5684d75b59d6aa02d81a3613c1
1	377480957394058	256845189467	beeac075680cfecbb3ecfb6894a3a9a90
2	377480957394058	256845189467	960dbd5058152c1c08994fd12a992d00
3	377480957394058	256845189467	d7678c49fefe6fc8b4120e8c2ec99540
4	377480957394058	256845189467	32d071025482e41ddc1963a6445ac61d
...
224896	3519925692476886	617852071065	0bde701a54af99fb7e5aa5a8e4411859
224897	3519925692476886	617852071065	30ea09b2542bcc81e80104ca441db16a
224898	3519925692476886	617852071065	ed38c86092c9b84af71c300f5e285a13
224899	3519925692476886	617852071065	7c48e12366ced59e5013e5cd4594a400
224900	3519925692476886	617852071065	e94d589b549fa03c32775859729bc958

	unix_time	category	amt	is_fraud \
0	1583184159	misc_net	80.46	0
1	1555701663	gas_transport	58.58	0
2	1575072903	grocery_pos	60.11	0
3	1575297118	grocery_pos	36.41	0
4	1563209257	grocery_net	60.17	0
...
224896	1594734506	misc_pos	8.36	0
224897	1607933927	shopping_pos	22.76	0
224898	1646037880	home	7.66	0
224899	1630226477	food_dining	8.32	0
224900	1668860469	personal_care	7.40	0

	merchant	merch_lat	merch_long
0	Marshall LLC	41.919802	-87.592037
1	Phelps-Campbell	41.032331	-88.592018
2	Hopkins, Jordan and Blair	41.865136	-88.818713
3	Williams, Bray and Aguilar	41.984380	-87.540684
4	Fitzgerald Group	42.206876	-87.130330
...
224896	Callahan, Anthony and Watson	42.989883	-71.580746
224897	Allison, Escobar and Camacho	42.328317	-71.614956
224898	Wright, Williams and Powers	43.260820	-70.486761
224899	Perry and Sons	41.794446	-71.901264
224900	Johnson and Sons	42.673450	-70.828146

[224901 rows x 10 columns]

/content/sample_data/dataset/transactions_41.csv

	cc_num	acct_num	trans_num \
0	180008357017032	31024004064	673f1c201c366bf8d15efaf0d40b60ed
1	180008357017032	31024004064	77fa9f0909067bddd6a6ec8c113185b5
2	180008357017032	31024004064	c76de38615a44eab1463e4255cc2223f
3	180008357017032	31024004064	3b63230a2f60d3e29f3b36721adccb92
4	180008357017032	31024004064	0644f49ea5a50bf5d0c2e87b47f9ca05
...
9317	4371450311809	499348764671	b8aba4793b7b506b4e4b9039cfff83b8
9318	4371450311809	499348764671	df294c8b5f0902e94e316c638d51bfba
9319	4371450311809	499348764671	e169b334bc866ef1c13493a434838596
9320	4371450311809	499348764671	11f9c9e77ea153b3805df3ba452beeb6

9321 4371450311809 499348764671 d99651c07db1b1e6b80d6cc24c5b5f60

	unix_time	category	amt	is_fraud	\
0	1650587403	grocery_pos	8.64	1	
1	1650560529	gas_transport	10.18	1	
2	1650562830	grocery_pos	10.02	1	
3	1650555511	grocery_pos	320.41	1	
4	1650559022	gas_transport	342.77	1	
...	
9317	1665649601	health_fitness	2.24	0	
9318	1656998186	shopping_pos	1.71	0	
9319	1656142845	shopping_net	7.48	0	
9320	1653643832	misc_pos	1.49	0	
9321	1665624222	health_fitness	5.40	0	

	merchant	merch_lat	merch_long
0	Ibarra-Benson	40.736928	-74.723498
1	Meyers, Diaz and Ashley	40.005778	-73.324198
2	Le, Kane and Myers	41.246197	-74.785883
3	Chapman and Sons	40.942670	-73.477831
4	Howell-Smith	41.329498	-74.897860
...
9317	Reid and Sons	34.067930	-116.303940
9318	Hernandez Ltd	32.359374	-116.286063
9319	Conley-Jones	33.007067	-116.871852
9320	Simmons, Fields and Brown	32.246416	-117.408449
9321	Harper, Richards and Nolan	32.442428	-116.882172

[9322 rows x 10 columns]

/content/sample_data/dataset/transactions_42.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_43.csv

	cc_num	acct_num	trans_num	\
0	4461571732030931	956395061779	bf9ade501f4778661e7368b52177f6ee	
1	4461571732030931	956395061779	d5aadd8495cd99dd1fe5c60371bb31aa	
2	4461571732030931	956395061779	76613b1d3443bb3f4e7c8cc02033e4fb	
3	4461571732030931	956395061779	d56b3e77f05471aefc85448ac1656ce2	
4	4461571732030931	956395061779	de1b00ded79a1c57816d42976efeeda0	
...	
21179	180041314675360	293307311995	a2a40ef46f4d5b8bfb3be53872f1c7dd	
21180	180041314675360	293307311995	8211df8713e8455caf0a5ef067191b7c	
21181	180041314675360	293307311995	7c8735ccd7387c5bbcc92a3f85b02564	
21182	180041314675360	293307311995	eca39927455a6e5964f540f846215be2	
21183	180041314675360	293307311995	adf3d7a644baf46f9a6e504c3339d7a3	

	unix_time	category	amt	is_fraud	\
0	1657734210	shopping_pos	91.24	0	
1	1655503605	shopping_pos	59.37	0	
2	1650734789	misc_net	75.49	0	
3	1646927059	grocery_pos	71.73	0	
4	1652105391	misc_pos	68.24	0	
...	
21179	1669980493	kids_pets	6.17	0	
21180	1669781586	kids_pets	8.94	0	

21181	1670579742	personal_care	567.45	0
21182	1661597101	food_dining	26.21	0
21183	1671625509	personal_care	7.86	0

		merchant	merch_lat	merch_long
0		Brown, Tran and Zhang	38.419841	-85.184004
1		Reed Group	37.629021	-85.562092
2		Carr Group	37.943825	-86.336896
3		Perkins, Riley and Robles	38.954940	-85.472901
4		Kline, Nguyen and Williams	39.181538	-85.538001
...	
21179		Holmes Ltd	37.735402	-83.032099
21180		Marshall and Sons	37.476081	-82.994869
21181		Burton and Sons	39.400767	-82.255915
21182		Jenkins, Williams and Jenkins	39.100287	-83.450935
21183		Green, Rasmussen and Macdonald	39.040886	-81.758927

[21184 rows x 10 columns]

/content/sample_data/dataset/transactions_44.csv

	cc_num	acct_num	trans_num	\
0	4514627048281480	888335239225	35d7371e85a6127a37914f355e822f37	
1	4514627048281480	888335239225	ed0c6c50e77af6eaea30046da23dbc9f	
2	4514627048281480	888335239225	479826e20b113849ad21a8767e35c168	
3	4514627048281480	888335239225	fec14935fd2cd63343d028f31f51f578	
4	4514627048281480	888335239225	260564c0cf03b7df91d739e69861fde3	
...
23368	5227674783667429	293218817973	2e4a763702f09da95d125bdd9a0d196f	
23369	5227674783667429	293218817973	16b148721ef2d5f6a856437e138f701a	
23370	5227674783667429	293218817973	bff18e361784a838b83f3b1262dbc298	
23371	5227674783667429	293218817973	f0b4b81b51e460293282d0b63e4cbdfb	
23372	5227674783667429	293218817973	f9d977abc9e03d1cb7762a3bbcb77bca	

	unix_time	category	amt	is_fraud	\
0	1582937132	misc_net	29.07	0	
1	1599268936	shopping_net	69.30	0	
2	1563297868	gas_transport	69.05	0	
3	1566427872	grocery_net	67.08	0	
4	1558308578	gas_transport	55.57	0	
...
23368	1610165283	entertainment	4.73	0	
23369	1633339149	health_fitness	6.13	0	
23370	1657776546	kids_pets	4.18	0	
23371	1631695258	health_fitness	8.24	0	
23372	1665918923	personal_care	8.86	0	

		merchant	merch_lat	merch_long
0		Adams, Herrera and Mcpherson	30.850678	-82.531775
1		Gonzalez PLC	31.460715	-81.748748
2		Fisher, Harris and Reese	29.792599	-83.395039
3		Rodriguez-Conrad	30.868097	-83.489898
4		Stephens Inc	30.215806	-82.030570
...	
23368		Wong LLC	35.028432	-84.990501
23369		Butler-Cook	34.816789	-84.846744
23370		Johns-Jackson	33.882807	-85.167481
23371		Daniel Group	35.140199	-85.554704
23372		Barrera-Ramos	35.026677	-84.678734

[23373 rows x 10 columns]

/content/sample_data/dataset/transactions_45.csv

	cc_num	acct_num	trans_num	\
0	3524295633432445	625080730349	ed19e604e86163678851bc0e4623a4cb	
1	3524295633432445	625080730349	31e5a664f221ebf650853ba33ba6c196	
2	3524295633432445	625080730349	7cb9b1e961ec57ae662109077d84f8f5	
3	3524295633432445	625080730349	e45af9031b92acc50ce14c9a9782a5f2	
4	3524295633432445	625080730349	5c77128edc2c4c8b3091990020cdf0c9	
...	
19374	4096496683496043	835714725252	0089045cb0b555d2c77a9690fc24b34c	
19375	4096496683496043	835714725252	ddfb86b21f289317c0a28baadcd9c41c	
19376	4096496683496043	835714725252	6a005cc875e9ca03484e1c6acab85056	
19377	4096496683496043	835714725252	885fe7dfedcf27e4bee88738e1492d30	
19378	4096496683496043	835714725252	150534828e98c70891800bb9c5936459	

	unix_time	category	amt	is_fraud	merchant
0	1654894071	shopping_net	56.66	0	Lam-Kent
1	1655650540	shopping_pos	75.72	0	Underwood, Pena and Garza
2	1652970337	misc_pos	85.82	0	Thompson Group
3	1652045256	misc_pos	42.95	0	Salazar-Atkins
4	1655085412	shopping_pos	99.33	0	Stark, Jones and Davis
...
19374	1670375210	kids_pets	8.67	0	Sherman-Williams
19375	1668257994	kids_pets	3.40	0	Potter-Brown
19376	1668416879	kids_pets	7.65	0	Moreno Group
19377	1665710998	home	3.87	0	Shaw LLC
19378	1669892933	kids_pets	7.49	0	Stephenson, Grimes and Beck

	merch_lat	merch_long
0	39.522730	-88.944528
1	40.656170	-88.450713
2	40.146614	-89.003980
3	38.965339	-89.196551
4	38.824846	-89.215172
...
19374	34.772597	-79.779781
19375	34.267421	-80.144453
19376	34.085209	-80.410626
19377	34.866067	-79.432673
19378	33.585393	-79.881444

[19379 rows x 10 columns]

/content/sample_data/dataset/transactions_46.csv

	cc_num	acct_num	trans_num	\
0	4461571732030931	956395061779	0c2c9125193656cdb6962322e53793c3	
1	4461571732030931	956395061779	daee0d9e88371d08c6eb92b8eea5b4ad	
2	4461571732030931	956395061779	1b259ea9330f611e7713a514037a93c8	
3	4461571732030931	956395061779	76e076c0069fa8e4ebc65be99596b212	
4	4461571732030931	956395061779	10ec88e46ef61702a374b4cc334279c6	
...	
111011	180041314675360	293307311995	3aee41f075c3b91f1c79c3b506706d78	
111012	180041314675360	293307311995	395eeef95f39ea08cc70fadacf237aac	
111013	180041314675360	293307311995	227e5490a7e1ecc5a68e8cfa57dc421f	
111014	180041314675360	293307311995	4b2afc428145b342dbf9c2a3e1541c6b	
111015	180041314675360	293307311995	f40db2cf31f5d215fafc8853c6cff7d8	

	unix_time	category	amt	is_fraud	\
0	1551561675	gas_transport	56.06	0	
1	1556565567	gas_transport	61.20	0	
2	1547671879	gas_transport	86.57	0	
3	1587600766	misc_pos	86.17	0	
4	1576781795	misc_net	79.15	0	
...	
111011	1639121467	home	7.54	0	
111012	1650336861	kids_pets	1.56	0	
111013	1626503315	food_dining	1.02	0	
111014	1630241835	health_fitness	4.56	0	
111015	1661082099	personal_care	4.70	0	

	merchant	merch_lat	merch_long
0	Johnston-Acosta	38.839306	-85.972665
1	Knox Group	39.164956	-85.099271
2	Wilson, Jacobs and Lawson	38.385797	-85.364645
3	Thompson, Wolfe and Gould	37.653371	-86.482074
4	Edwards-Richards	38.382941	-85.866553
...
111011	Moreno Ltd	38.881050	-81.946841
111012	Harrison, Soto and Rodriguez	38.887786	-82.129429
111013	Adams-Smith	39.435715	-83.387099
111014	Luna, Bowen and Carter	39.365652	-82.315247
111015	Fields, Gentry and Soto	39.431753	-82.361134

[111016 rows x 10 columns]

/content/sample_data/dataset/transactions_47.csv

	cc_num	acct_num	trans_num	\
0	4795133053194979696	585020232980	f10c7863504e1389f482196daa3d4203	
1	4795133053194979696	585020232980	d2bfd26c68bd826121d297eab812c73	
2	4795133053194979696	585020232980	9efc7f4a8db9b08d12ad82bef03710b9	
3	4795133053194979696	585020232980	c5fbd98f50629af20fbaef8d80b0d1ab	
4	4795133053194979696	585020232980	90980cb3a2dca599cc63a025de81aa07	
...	
23368	4191260221323	679124338217	0f0c18fcc1981be836346a270d2f7708	
23369	4191260221323	679124338217	853bf1cdf71c30a52fcc3d91740345d9	
23370	4191260221323	679124338217	6cffe1cd3e050e89a808eec2e2e0a1f8	
23371	4191260221323	679124338217	3e8f7bb8698bcd9a45883206a863774c	
23372	4191260221323	679124338217	f6756583c4240d8d70e4a8f158c27577	

	unix_time	category	amt	is_fraud	\
0	1583685552	gas_transport	13.61	1	
1	1583683240	misc_net	301.98	1	
2	1583686780	grocery_pos	327.75	1	
3	1583686008	grocery_pos	278.63	1	
4	1583761619	misc_net	293.73	1	
...	
23368	1586510379	misc_pos	3.48	0	
23369	1577415265	misc_net	8.37	0	
23370	1605413663	shopping_pos	5.11	0	
23371	1635409224	health_fitness	8.95	0	
23372	1620036259	entertainment	7.97	0	

	merchant	merch_lat	merch_long
0	Reynolds-Smith	40.164988	-79.456743

1	Livingston and Sons	41.290426	-80.622981
2	Davis, Martin and Burton	41.075397	-79.718986
3	Williams, Hudson and Buckley	40.586551	-79.489719
4	Morales, Jones and Cardenas	39.851332	-80.117580
...
23368	Oneal, Allen and Hawkins	39.011339	-84.715278
23369	Rivera, Morris and Glenn	38.961994	-84.676587
23370	Banks-Mooney	39.235679	-84.998581
23371	Perez-Burke	40.489000	-85.486662
23372	Morgan-Lee	39.821872	-85.818531

[23373 rows x 10 columns]

/content/sample_data/dataset/transactions_48.csv

	cc_num	acct_num	trans_num \
0	30292842222314	491227929210	f67c4b47cc6a418f9f1edb683b5af74d
1	30292842222314	491227929210	d1b5ccc3f29477e14df905eda95908c4
2	30292842222314	491227929210	a9aac8888d033bd07e90851638dd84fe
3	30292842222314	491227929210	f2518539ccd7c1f2f41384345c9a768b
4	30292842222314	491227929210	0fcc82ac9036388700264ad603fa516c
...
49778	6520575639836526	665775741850	dbbe98eb3dfd69af176aa19890a11691
49779	6520575639836526	665775741850	1ff2e161aaf26da3d8a989c6965a4e94
49780	6520575639836526	665775741850	b102539b935a46f601b6969fefe41f23
49781	6520575639836526	665775741850	11452ef8cbf04c63b7a39ef889ff12c1
49782	6520575639836526	665775741850	d939ec086d2bfe5615e213c10d6041eb

	unix_time	category	amt	is_fraud \
0	1644271062	gas_transport	72.70	0
1	1644085966	gas_transport	73.14	0
2	1647014898	grocery_net	98.90	0
3	1646754763	grocery_net	85.81	0
4	1643408303	gas_transport	72.68	0
...
49778	1668488944	home	4.10	0
49779	1671861500	personal_care	2.56	0
49780	1665317575	health_fitness	1.64	0
49781	1668322046	home	6.64	0
49782	1668828394	home	2.05	0

	merchant	merch_lat	merch_long
0	Miller, Solis and Vega	33.756152	-118.180690
1	Dunn and Sons	34.684152	-118.072794
2	Massey-Morgan	32.980310	-118.407788
3	Rodriguez, Thomas and Sexton	33.196887	-118.124583
4	George, Hicks and Krueger	33.009362	-118.740220
...
49778	Mitchell-Brown	35.582542	-118.261995
49779	Le-Morris	34.167867	-119.154091
49780	Avery, Gonzalez and George	35.375374	-119.251738
49781	Sweeney, White and Esparza	34.265296	-117.764426
49782	Morales-Wilson	34.110858	-118.985261

[49783 rows x 10 columns]

/content/sample_data/dataset/transactions_49.csv

	cc_num	acct_num	trans_num \
0	4244717465421409	194959880423	a523504f0f51872fc637fa671fafe22a
1	4244717465421409	194959880423	86681c12479c6f089b05617857cbfa24

2	4244717465421409	194959880423	4730bfb9c7de0d69698c6c0481e2fd75
3	4244717465421409	194959880423	1ca10c651b2b95a83a3d16db94c556d7
4	4244717465421409	194959880423	e3fba2df323456e272d3ae0a111498ab
...
204448	4604409079772781	485572028745	deba8e5a5f4662ef0623134f98f154a2
204449	4604409079772781	485572028745	8dcd03f7d9fbbae845793ffb0a2ded36
204450	4604409079772781	485572028745	90cd01cb27e1130f54260ddc4e06cf14
204451	4604409079772781	485572028745	fab57b57a52c77c463c0b8133068c292
204452	4604409079772781	485572028745	bc11ff95e69fca3b8df6662611f310a4

	unix_time	category	amt	is_fraud	\
0	1573915518	gas_transport	8.55	1	
1	1573916219	food_dining	355.21	1	
2	1573914320	shopping_pos	887.15	1	
3	1573914786	shopping_pos	793.91	1	
4	1573917423	shopping_net	713.53	1	
...	
204448	1605935870	shopping_pos	4.11	0	
204449	1616306903	entertainment	9.24	0	
204450	1614479113	entertainment	7.86	0	
204451	1604886291	shopping_pos	4.59	0	
204452	1611985153	shopping_pos	4.25	0	

	merchant	merch_lat	merch_long
0	Perez, English and Cox	49.068333	-123.051367
1	Cruz, Marsh and Burke	48.967123	-121.265437
2	Phillips Group	48.244516	-122.237457
3	Brown-Morton	48.133023	-122.916003
4	Kelly and Sons	49.054959	-121.837657
...
204448	Ball-Gomez	42.681869	-83.617938
204449	Heath, Bass and Reyes	42.776807	-82.059483
204450	Wiggins LLC	43.346101	-83.777671
204451	Nguyen Ltd	43.235394	-83.302738
204452	Holloway, Rodgers and Rose	43.319392	-82.432722

[204453 rows x 10 columns]

/content/sample_data/dataset/transactions_50.csv

	cc_num	acct_num	trans_num
\			
0	4358841834374016	521500975371	311b225b2a1927eb72511fc852ef7ceb
1	4358841834374016	521500975371	3c428030f737c2c35c48b72128a5e41b
2	4358841834374016	521500975371	8623a8d0f5c7886879bba04cc58a9f6c
3	4358841834374016	521500975371	ecd25d4ef83350fcf87094648de37acb
4	4358841834374016	521500975371	ece42ca5e3b4633b5cb578df72d7e860
...
183972	4601080181030683173	117887236764	f8d5f94ababa2fefaf2d278ad9687bcf8
183973	4601080181030683173	117887236764	1a1d67cb98ad6c22ea6c88bd3cec3f9e
183974	4601080181030683173	117887236764	90a940e1796532cd0df18ce7c51f9a80
183975	4601080181030683173	117887236764	63d4946335cd750b2f241c838d7289ac
183976	4601080181030683173	117887236764	e6ae5a6210e717209029833e63ea5f5e

	unix_time	category	amt	is_fraud	\
0	1566321512	grocery_pos	100.39	0	
1	1563656523	grocery_pos	77.78	0	
2	1558045617	grocery_net	83.35	0	
3	1554825671	grocery_net	61.17	0	

4	1551360953	gas_transport	63.27	0
...
183972	1632578098	health_fitness	5.08	0
183973	1662185613	personal_care	6.67	0
183974	1632897034	health_fitness	4.59	0
183975	1628254695	food_dining	6.12	0
183976	1617000819	entertainment	2.45	0

	merchant	merch_lat	merch_long
0	Allen PLC	37.708531	-84.840628
1	Hughes-Nelson	39.002725	-86.028650
2	Rosales, Quinn and Johnson	37.458960	-85.401004
3	Watson, Fisher and Williams	37.630174	-86.404157
4	Johnson-Lawson	38.104203	-85.293647
...
183972	Long, Gibson and McIntosh	41.096302	-74.461268
183973	Allen, Rhodes and Olson	40.930486	-75.114876
183974	Schultz-Jackson	39.933608	-75.423625
183975	Brown, Ellis and Santana	39.884383	-75.564335
183976	Terry, Miller and Perry	39.551217	-74.806508

[183977 rows x 10 columns]

/content/sample_data/dataset/transactions_51.csv

	cc_num	acct_num	trans_num \
0	6584732365337749	843643900137	087215e7b1284599ca53ae8e76c320d2
1	6584732365337749	843643900137	cc48a9e7916ff67719624b6fe07dcb3a
2	6584732365337749	843643900137	1752827103b6d8e56f305deeff142086
3	6584732365337749	843643900137	c609108203b47a710fb93e7093036585
4	6584732365337749	843643900137	4406bcc991b17e425bfe2185da6c7f2d
...
1455	6584732365337749	843643900137	7b2b63508d92d5dd61446fb5cb03e301
1456	6584732365337749	843643900137	1df5a10b715d536a666143ebb429a818
1457	6584732365337749	843643900137	67f8305d28b564b34ccf264b9f4ea764
1458	6584732365337749	843643900137	4d3333cd26ba3ac993151b30b474e94e
1459	6584732365337749	843643900137	e9126d852f94ab72fa6b40ee7e44c6f1

	unix_time	category	amt	is_fraud	merchant
\					
0	1555686450	gas_transport	66.58	0	Griffith LLC
1	1548958402	gas_transport	74.28	0	Blackburn, Owen and Morris
2	1552419151	gas_transport	63.64	0	Brown PLC
3	1574291651	grocery_pos	54.93	0	Watson, Parsons and Schmidt
4	1546735016	gas_transport	41.67	0	Black LLC
...
1455	1667282786	personal_care	5.74	0	Dennis Inc
1456	1591184669	misc_pos	8.28	0	Beck, Mcneil and Benjamin
1457	1625045557	food_dining	9.98	0	Lee-Schneider
1458	1667465864	personal_care	9.42	0	Benson Inc
1459	1620636587	entertainment	6.57	0	Arnold Group

	merch_lat	merch_long
0	44.020551	-91.532724
1	43.731354	-90.399642
2	43.814645	-90.661504
3	43.045640	-90.615255
4	43.925290	-90.478813
...

```

1455 44.172820 -90.422277
1456 43.413139 -91.058986
1457 43.853512 -90.834174
1458 43.025016 -91.550511
1459 43.820191 -92.239134

```

```
[1460 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_52.csv
```

	cc_num	acct_num	trans_num	\
0	180075851268619	406674938298	e9fcbb1d4dbff0f4aae226b6482ebd68	
1	180075851268619	406674938298	f833c319359b932240f438e262788218	
2	180075851268619	406674938298	60c52efb89781938ab5804811748e1ae	
3	180075851268619	406674938298	dbe851acd5ed6f9d12795cbb74c718e4	
4	180075851268619	406674938298	c8f3eca4d022e7d97ad7b8e265258865	
...
159141	4361337605230458	349734538563	cf819676e6deac266040781a7d6a40dd	
159142	4361337605230458	349734538563	51676b80557a0ec6a58122224f48ef4c	
159143	4361337605230458	349734538563	16e7ec150e49ac5f88d51cfb0e1656a8	
159144	4361337605230458	349734538563	a3a910e67f17ee06ac36f25b6297c914	
159145	4361337605230458	349734538563	dfd4b5e5f00f984cc81a9a1370d262a8	

	unix_time	category	amt	is_fraud	\
0	1575552380	grocery_pos	56.91	0	
1	1570147581	grocery_pos	69.71	0	
2	1569097181	grocery_pos	74.64	0	
3	1572018037	grocery_pos	66.69	0	
4	1554943830	gas_transport	89.14	0	
...
159141	1602319564	shopping_net	8.63	0	
159142	1624776626	food_dining	3.99	0	
159143	1644139064	home	3.11	0	
159144	1641523890	home	1.09	0	
159145	1622290438	entertainment	301.93	0	

	merchant	merch_lat	merch_long
0	King, Olson and Smith	47.753778	-123.190368
1	Kim PLC	47.967845	-122.247512
2	Johnson-Maddox	46.621623	-122.891241
3	Morrow, Goodman and Holland	46.836755	-121.562162
4	Cline, Jensen and Carey	47.066125	-122.904399
...
159141	James Inc	43.053330	-80.170939
159142	Reid, Sanchez and Fisher	42.381655	-79.302993
159143	Key Group	41.769652	-79.765810
159144	Johnson LLC	42.980108	-80.064378
159145	Leblanc, Vargas and Barnes	42.003362	-79.090451

```
[159146 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_53.csv
```

```
Empty DataFrame
```

```
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
```

```
Index: []
```

```
/content/sample_data/dataset/transactions_54.csv
```

```
Empty DataFrame
```

```
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
```

Index: []

/content/sample_data/dataset/transactions_55.csv

	cc_num	acct_num	trans_num	\
0	377480957394058	256845189467	e7230464158f9c7b3a95923b5c823b51	
1	377480957394058	256845189467	0b133e54ec978323484101526f755f42	
2	377480957394058	256845189467	13c29592d84ed12d984bd2e576036b65	
3	377480957394058	256845189467	b5f0f7be02536085856020229a958336	
4	377480957394058	256845189467	be6643f660ef4407b2067cd9d6870562	
...	
56244	3519925692476886	617852071065	dcea0125d8b8a4ceea162331e05464af	
56245	3519925692476886	617852071065	764240285d808c202cb8e69374d64bde	
56246	3519925692476886	617852071065	fbdd75988830c0a6487aa0a373a6cb9a	
56247	3519925692476886	617852071065	558632fb264b35fd843cb3c82adf7bca	
56248	3519925692476886	617852071065	cc3b6d5230a39f7831d9316434e9d0cb	

	unix_time	category	amt	is_fraud	\
0	1661612144	grocery_pos	287.93	1	
1	1661614535	misc_net	741.26	1	
2	1661694841	shopping_net	638.63	1	
3	1661688865	shopping_net	1121.18	1	
4	1661694681	shopping_net	898.54	1	
...	
56244	1659516575	entertainment	1.14	0	
56245	1665545494	health_fitness	3.93	0	
56246	1657944451	shopping_pos	5.52	0	
56247	1653803805	misc_pos	1.02	0	
56248	1670650801	kids_pets	1.78	0	

	merchant	merch_lat	merch_long
0	Stewart Ltd	41.005191	-88.628379
1	Carr Group	42.698003	-88.813593
2	Garcia-Hunt	42.501350	-87.422601
3	Boyer, Price and Chan	42.333992	-87.232013
4	Kent-Ortiz	41.010402	-88.028739
...
56244	Phillips-Maynard	43.158028	-70.113956
56245	White, Fritz and Morgan	41.583169	-70.674287
56246	Baker Ltd	42.826378	-70.673580
56247	White, Gonzalez and Tanner	43.231622	-71.105659
56248	Young, Fowler and Velez	41.896761	-70.647887

[56249 rows x 10 columns]

/content/sample_data/dataset/transactions_56.csv

	cc_num	acct_num	trans_num	\
0	3544928787570307	185865727840	e5f6b59ebc648b371e37a3999369a2a5	
1	3544928787570307	185865727840	13dbd1168a202d54b4828f94a3a8d953	
2	3544928787570307	185865727840	7ff130853ef15fc8b1784c76da83171c	
3	3544928787570307	185865727840	f099306ef13c9783af1c262052a6ff57	
4	3544928787570307	185865727840	cb15fb65e2c82f9c1f3d87c0249c1560	
...	
1069	4178955990984125	243041877024	037bf0483937b7902914950a60a55077	
1070	4178955990984125	243041877024	8c2442058520d73c34bcf0f39dc0e289	
1071	4178955990984125	243041877024	ae70e9e6082f9a96638709132cc415f0	
1072	4178955990984125	243041877024	d3536ac22e851e912b4086861b756d04	
1073	4178955990984125	243041877024	fd0274f72fa2a2914e76a3f911d2b326	

unix_time	category	amt	is_fraud	merchant
-----------	----------	-----	----------	----------

\						
0	1651785219	misc_pos	73.58	0		Hodge-Duncan
1	1652574109	misc_pos	78.42	0		Salazar-Baker
2	1642175961	gas_transport	75.61	0		Clark Group
3	1649812410	misc_net	72.75	0		Morgan-Aguirre
4	1655848061	shopping_pos	74.08	0		Paul-Kelley
...
1069	1667373899	home	555.56	0		Wu-Riley
1070	1671959256	travel	534.43	0		Myers-Clarke
1071	1672227726	travel	475.14	0	Murray, Smith and Carroll	
1072	1671879052	personal_care	562.93	0		Williamson PLC
1073	1668215909	home	485.36	0		Morgan and Sons

	merch_lat	merch_long
0	48.984414	-119.306075
1	48.030228	-120.077876
2	49.126907	-120.041089
3	48.479925	-119.211692
4	48.270640	-120.077041
...
1069	40.900847	-88.802848
1070	40.603441	-90.004287
1071	41.228287	-89.828474
1072	40.677603	-90.200682
1073	40.096792	-90.367690

[1074 rows x 10 columns]

/content/sample_data/dataset/transactions_57.csv

	cc_num	acct_num	trans_num	\
0	30292842222314	491227929210	f4ae4473d12990d77fe2b56ee9a0c13b	
1	30292842222314	491227929210	c6aa18ad4b545adf1fa146beba77e225	
2	30292842222314	491227929210	5b9e4200fca68957d5c71379f9acf1a1	
3	30292842222314	491227929210	4a45d4b140caa15701b4755c16e40c44	
4	30292842222314	491227929210	b77922b6aad203ec05d5f386b1fd09bf	
...	
173790	6520575639836526	665775741850	0a1da5c0324c90f57c9378b106878461	
173791	6520575639836526	665775741850	32b6f6f07e6139a47e5255d0e22bc508	
173792	6520575639836526	665775741850	25d9c7d3bee03804f02bc31c2e517d8f	
173793	6520575639836526	665775741850	718b34b992d8123b3689a9b864d61747	
173794	6520575639836526	665775741850	39bb05b49878b5fadd75f8d7cdfd07d7	

	unix_time	category	amt	is_fraud	\
0	1564178003	grocery_net	76.66	0	
1	1559003655	gas_transport	68.71	0	
2	1555266330	gas_transport	81.90	0	
3	1579475917	misc_net	52.63	0	
4	1574529955	grocery_pos	65.43	0	
...	
173790	1611994967	shopping_pos	8.01	0	
173791	1589165504	misc_pos	4.13	0	
173792	1668844953	personal_care	5.63	0	
173793	1612952464	shopping_pos	5.81	0	
173794	1643948631	home	4.93	0	

	merchant	merch_lat	merch_long
0	Howard PLC	34.473275	-119.121134
1	Coleman, Lopez and Mitchell	34.799318	-117.654393

2	Gentry, Mclean and Mcconnell	33.986136	-118.110989
3	Powers-Norris	34.601638	-118.930636
4	Romero PLC	33.490563	-119.198029
...
173790	Watkins-Griffin	35.995568	-119.388749
173791	Melton, Flores and Wilson	34.277599	-118.751288
173792	Green, Mason and Kirby	35.766405	-119.340650
173793	Smith Ltd	36.019133	-118.078766
173794	Freeman, Gutierrez and Torres	35.299565	-118.941630

[173795 rows x 10 columns]

/content/sample_data/dataset/transactions_58.csv

	cc_num	acct_num	trans_num	\
0	4795133053194979696	585020232980	fa7bbdd3e4c9846f5f7f60bff95212a7	
1	4795133053194979696	585020232980	9ea8416638d897e3d8cf6919aac2b000	
2	4795133053194979696	585020232980	35d8d7728dd16719e9f0a1c7160cae68	
3	4795133053194979696	585020232980	e06abd1ef51726dec0753071c39e6aae	
4	4795133053194979696	585020232980	fb198a51057341da6cc7173340be4cb3	
...	
6081	4191260221323	679124338217	7c7e0a2246a45b748ff8d0f420524b8e	
6082	4191260221323	679124338217	0596be5c40693047a49d1a0a6d7309ac	
6083	4191260221323	679124338217	56aa11b6d612eb29dab0cddb5d6cb7fa0	
6084	4191260221323	679124338217	0445d755c120d8a79882a7842fb540b2	
6085	4191260221323	679124338217	8106bce75aae8e5ad035f6bf0d28cb87	

	unix_time	category	amt	is_fraud	\
0	1647633509	grocery_net	59.42	0	
1	1641585223	gas_transport	46.50	0	
2	1647024899	gas_transport	53.53	0	
3	1648233139	grocery_net	69.10	0	
4	1649469538	grocery_pos	38.78	0	
...	
6081	1669432580	kids_pets	5.37	0	
6082	1663315939	food_dining	9.34	0	
6083	1663938677	food_dining	9.75	0	
6084	1658207205	shopping_pos	3139.17	0	
6085	1655438796	shopping_net	2.37	0	

	merchant	merch_lat	merch_long
0	Lopez, Bradley and Bates	40.173928	-79.871141
1	Nicholson-Perry	40.791194	-80.059102
2	Galloway-Morales	39.742987	-79.874543
3	Young Ltd	41.191978	-80.379249
4	Greene, Wright and Smith	41.441403	-79.074076
...
6081	Maynard-Ochoa	38.869306	-84.425202
6082	Martinez Group	40.711375	-85.508575
6083	Williams, Perez and Morgan	40.036933	-84.589223
6084	Callahan-Ferguson	39.457236	-84.859666
6085	Fletcher-Weaver	39.259544	-85.998638

[6086 rows x 10 columns]

/content/sample_data/dataset/transactions_59.csv

	cc_num	acct_num	trans_num	\
0	675942799072	231045526966	cafd8c447e3983c8f5d73b5de1821ee2	
1	675942799072	231045526966	18d6b71917c3d041bb94ae53bae37ac0	
2	675942799072	231045526966	4243c81d3c9bcc37bb70714d538464ef	

3	675942799072	231045526966	50b4601eac43404306e034068a00c846
4	675942799072	231045526966	6ad675adf0defff2ba5f5407f1717b8b
...
4325	4328348315203771070	417495431869	c8bcdaa7c2e4bcd0ccc4a070ce5221aa
4326	4328348315203771070	417495431869	252a144c392237cddb1934087dd48574d
4327	4328348315203771070	417495431869	657a1ec51d22e821c2d2cae828a112d2
4328	4328348315203771070	417495431869	3e1b0fca55ccf4aa5553f450a168c523
4329	4328348315203771070	417495431869	7961449032a3ddb2d1f570c9b29d8a9d

	unix_time	category	amt	is_fraud	\
0	1666623583	misc_net	10.86	1	
1	1666620037	grocery_pos	257.65	1	
2	1666618505	grocery_net	310.55	1	
3	1666629433	misc_net	314.22	1	
4	1666623088	grocery_pos	317.22	1	
...	
4325	1660386331	entertainment	2.66	0	
4326	1651551616	misc_net	6.71	0	
4327	1661518826	entertainment	970.54	0	
4328	1654950663	shopping_net	4.86	0	
4329	1658045997	shopping_pos	5.33	0	

	merchant	merch_lat	merch_long
0	Nguyen LLC	35.409342	-78.313878
1	Johnson Inc	34.773208	-79.835370
2	Shaffer, Reynolds and Marshall	34.866932	-79.355794
3	Hodge Ltd	35.030842	-78.586686
4	Hayes, Sanchez and Schmidt	35.895279	-78.515180
...
4325	Smith-Bradley	33.748939	-119.230074
4326	Ramirez LLC	34.352800	-118.745790
4327	Arnold-Evans	34.296766	-117.688748
4328	Hester-Allen	34.215635	-117.831760
4329	Flores-Jones	32.918629	-119.026339

[4330 rows x 10 columns]

/content/sample_data/dataset/transactions_60.csv

	cc_num	acct_num	trans_num	\
0	6011113885034236	330792868853	3c70ae43b79fec58feb50409985cbc4	
1	6011113885034236	330792868853	37bbef728177c90132d4292cdf316bcd	
2	6011113885034236	330792868853	9572e06be613f2cba6ab191f2d8ada5c	
3	6011113885034236	330792868853	57155d6ec9dae224ddcd3e5a8b1494e0	
4	6011113885034236	330792868853	eee07a5a5a3858848739ec90dd743f2e	
...	
1786	4886407880267	328429374890	63f7148b6d9e587b6877143954736a6b	
1787	4886407880267	328429374890	fc3fd0c63ce6481876814a527d14a75b	
1788	4886407880267	328429374890	4612036ba54a0c2d2e992d526ea88df6	
1789	4886407880267	328429374890	1362136504659a4b13dc9be775a8c6c2	
1790	4886407880267	328429374890	bd432ecd9b17512f7ecd143c423bd8a2	

	unix_time	category	amt	is_fraud	merchant	\
0	1653671135	misc_pos	69.79	0	Wood, Kim and Robbins	
1	1657294312	shopping_pos	56.04	0	Houston LLC	
2	1658178912	shopping_pos	72.87	0	Miller Ltd	
3	1649122316	grocery_pos	51.50	0	Frazier-Jackson	
4	1658349965	shopping_pos	56.06	0	Mcdonald-Fernandez	
...	

1786	1666923788	home	4.93	0	Roberts-Conrad
1787	1661948723	food_dining	7.34	0	Miller-Parks
1788	1669513592	kids_pets	6.88	0	Harris-Roach
1789	1668421079	kids_pets	9.81	0	Sanders-Williams
1790	1670907652	personal_care	158.48	0	Keller-Taylor

	merch_lat	merch_long
0	32.722064	-92.987196
1	32.249173	-93.664175
2	31.669115	-93.269693
3	32.112034	-93.313454
4	31.839536	-92.646897
...
1786	39.767663	-96.905805
1787	40.092526	-97.450274
1788	40.994119	-98.841362
1789	41.425032	-98.636565
1790	39.931386	-97.167508

[1791 rows x 10 columns]

/content/sample_data/dataset/transactions_61.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_62.csv

	cc_num	acct_num	trans_num
0	4896331812335761701	149852234418	cfaa36a23c2ff02f15b0eebaa6ca8bcc
1	4896331812335761701	149852234418	d9e01f88713d2fdc85210e3bd1edf5ca
2	4896331812335761701	149852234418	28bde53d664ea62d22e98564bbf9b167
3	4896331812335761701	149852234418	c65a7e15a3946c074d54a9ba5ec8c1e3
4	4896331812335761701	149852234418	85edebb320838c76acd461c53265bfd6
...
4375	4896331812335761701	149852234418	1d805cb3e45935d143b35b4b12389b24
4376	4896331812335761701	149852234418	c88416239fdb8a8b6bdc8619819fd46a
4377	4896331812335761701	149852234418	1e9807c4d9404aaa7befc0604059559e
4378	4896331812335761701	149852234418	59ac185742551c4733c9698ad463f0a1
4379	4896331812335761701	149852234418	e41ebaa52b597f91e717f11027576d1f

	unix_time	category	amt	is_fraud	merchant
0	1564607292	gas_transport	62.63	0	Bradford, Ramos and Conway
1	1550273508	gas_transport	62.99	0	Ross Inc
2	1552755692	gas_transport	77.60	0	Valenzuela Inc
3	1565551047	gas_transport	51.56	0	Brown PLC
4	1548103546	gas_transport	63.70	0	Olsen-Ruiz
...
4375	1662088542	kids_pets	5.18	0	Hartman-Ramirez
4376	1587873533	misc_net	3.36	0	Wilson, Thomas and Mills
4377	1602656713	misc_pos	9.70	0	Ayala Group
4378	1633664280	food_dining	35.69	0	Meyer, Ellison and Roth
4379	1621165528	shopping_pos	1.84	0	Wolfe, Sandoval and Oconnor

	merch_lat	merch_long
0	38.182233	-89.965908
1	38.942561	-90.298534
2	38.249016	-90.114256

```

3      38.041679  -90.885558
4      39.399866  -89.780760
...      ...      ...
4375   39.507472  -90.998448
4376   39.806989  -91.289944
4377   39.593064  -91.535484
4378   39.044702  -89.772646
4379   38.939376  -90.320582

```

[4380 rows x 10 columns]

/content/sample_data/dataset/transactions_63.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_64.csv

```

      cc_num      acct_num      trans_num  \
0      4490797796124  802714479522  a3ef4d47d2b13a27e6dc2d7bbf92acba
1      4490797796124  802714479522  d9e9892fdb9da03c1f5a67954e0f706
2      4490797796124  802714479522  9d5c79b9c0f56f2bf3eabaf55c3e0f26
3      4490797796124  802714479522  25d7001e082762c03d7cdc5fcbe9828a
4      4490797796124  802714479522  7a606f10d5125d12114883a465b766d9
...      ...      ...      ...
2869   30306396384526  278085326817  573c4e6cc3cb4106da410649e6c53a5e
2870   30306396384526  278085326817  88238029e25c7c4408fde01158aac43d
2871   30306396384526  278085326817  1ce11f828e8cd58d0b120d342ee0d3d0
2872   30306396384526  278085326817  509a171b7a64fe0e2e95a78bc59548a8
2873   30306396384526  278085326817  22b79700cddc9ef7c3460d0b552a2fb7

```

```

      unix_time      category      amt  is_fraud      merchant
\
0      1660488349  gas_transport  19.30         1      Reed-Dixon
1      1660496359  grocery_pos   22.19         1      Rosario LLC
2      1660497745  gas_transport  328.33        1      Pena Inc
3      1660496465  grocery_pos   308.33        1      Brooks LLC
4      1660651700  misc_pos     50.74         1      Wise-Johnson
...      ...      ...      ...      ...
2869   1672454516  travel        4.01         0  Sanchez, Nunez and Hoover
2870   1667385637  home          2.96         0      Lowe Ltd
2871   1670126530  kids_pets     3.93         0      Carter Group
2872   1671153032  personal_care  1.48         0      Smith-Gonzales
2873   1663233003  health_fitness  5.01         0      Willis Inc

```

```

      merch_lat  merch_long
0      35.213592  -79.597141
1      35.507334  -78.394947
2      35.000853  -78.868041
3      35.116081  -79.666710
4      36.081634  -79.115669
...      ...      ...
2869   45.004826  -69.510524
2870   44.162219  -68.361270
2871   44.028680  -68.971669
2872   45.444450  -68.402506
2873   45.502235  -69.823759

```

[2874 rows x 10 columns]

/content/sample_data/dataset/transactions_65.csv

	cc_num	acct_num	trans_num	\
0	379389701550295	720358760888	51c2c33f56af3ca7b207cbc5e6933668	
1	379389701550295	720358760888	f69c07ac21c1ae536911f0aaa6ca074b	
2	379389701550295	720358760888	3523c738b557595e2477a6beb5b99a64	
3	379389701550295	720358760888	63067a920b6670184721eaf582a503b6	
4	379389701550295	720358760888	f525f173c6621f617f5ff34a163da714	
...	
36504	503870615922	312931364007	a616c1375af13b73bec59603ee7b90d0	
36505	503870615922	312931364007	484ca40b92166d0fd4008b0481de1bc8	
36506	503870615922	312931364007	35d053b3fd44daff576394479da2dcdd	
36507	503870615922	312931364007	07d0ef3fd7ed75a5f06203bc04769306	
36508	503870615922	312931364007	4a0edfc3f984c8206dd08a97a9bb2517	

	unix_time	category	amt	is_fraud	\
0	1569099128	grocery_net	11.11	1	
1	1569084042	gas_transport	7.56	1	
2	1569076081	misc_net	7.05	1	
3	1569088314	gas_transport	884.55	1	
4	1569215939	shopping_pos	975.47	1	
...	
36504	1647257242	home	5.50	0	
36505	1576394828	misc_net	4.08	0	
36506	1586594639	misc_pos	8.05	0	
36507	1657698275	kids_pets	6.19	0	
36508	1639971099	health_fitness	1.30	0	

	merchant	merch_lat	merch_long
0	Bray-Hart	25.911466	-97.042386
1	Miller, Small and Young	26.785446	-98.135088
2	Fleming PLC	26.838448	-97.278552
3	Shaw, Clark and Murphy	25.953636	-97.609938
4	Marshall Group	25.119819	-97.308718
...
36504	Sullivan-Jenkins	42.069913	-73.413071
36505	Lowe, Andrade and Griffin	40.976257	-72.822337
36506	Espinoza-Diaz	41.092525	-73.864965
36507	Martinez, Hernandez and Davis	41.022724	-74.143981
36508	Grimes LLC	41.833573	-73.961377

[36509 rows x 10 columns]

/content/sample_data/dataset/transactions_66.csv

	cc_num	acct_num	trans_num	\
0	4891077119918081211	402211201664	3e83e0a77771558c6daf6bdd272f6c4c	
1	4891077119918081211	402211201664	20819ecd9109e9cf302f355517626bc7	
2	4891077119918081211	402211201664	92a533083ef1d32858926edf549935a1	
3	4891077119918081211	402211201664	b516532ac635ab0c2359796d25ba33e1	
4	4891077119918081211	402211201664	49d86d3aca8789a0cbec73d3190044a5	
...	
3575	3534679121416143	103258999645	519fc7722f82ff9fc29c676afa67f865	
3576	3534679121416143	103258999645	56d127ec0b4deb9176dc101b1755a763	
3577	3534679121416143	103258999645	d7b1569b28d7c9934ee7af084ec2a463	
3578	3534679121416143	103258999645	e9437059d6d6adce422fd97e19a4f740	
3579	3534679121416143	103258999645	3071d2954fd52e7c5feb825f40f32721	

	unix_time	category	amt	is_fraud	merchant
\					

0	1650832775	grocery_pos	57.81	0	Johnson and Sons
1	1650129428	grocery_pos	69.54	0	Williams, Vega and White
2	1645891540	gas_transport	56.06	0	Gardner-Mcdonald
3	1645202314	gas_transport	74.61	0	Scott, Anderson and Allison
4	1644528165	gas_transport	55.61	0	Johnston Ltd
...
3575	1671348577	personal_care	9.65	0	Allen LLC
3576	1653892071	misc_pos	4.62	0	Burton Inc
3577	1666326940	home	2.03	0	Russo, Banks and Proctor
3578	1654691124	misc_pos	1.39	0	Terrell-Armstrong
3579	1659517047	shopping_pos	4.38	0	Hernandez-Wiggins

	merch_lat	merch_long
0	38.537479	-82.777755
1	40.320563	-82.720365
2	39.115711	-81.762917
3	38.785439	-81.963553
4	40.022723	-83.285232
...
3575	62.469827	-158.189798
3576	62.322803	-158.338653
3577	61.834744	-157.759693
3578	61.283282	-158.060932
3579	60.955730	-157.012305

[3580 rows x 10 columns]

/content/sample_data/dataset/transactions_67.csv

	cc_num	acct_num	trans_num \
0	4244717465421409	194959880423	1e0409c48dcc893417ea565d36b1b90a
1	4244717465421409	194959880423	507fd974ce97d761ad79dec19a6c38b4
2	4244717465421409	194959880423	b415120ff27b61e0d51e66946369f561
3	4244717465421409	194959880423	271419a884510f6b8e4c4dd4eeb1e8d8
4	4244717465421409	194959880423	b829f3c437e73913bcb65c4260ee087c
...
56217	4604409079772781	485572028745	ed5db362b2c2eb67428b1cb5ec669a33
56218	4604409079772781	485572028745	fc430ede5f46d50a2378a9d8f281b338
56219	4604409079772781	485572028745	1f23fc0ca9f25a9e12b29f6fdc23faec
56220	4604409079772781	485572028745	f7fb83871e65d2019437dcb85deb53f9
56221	4604409079772781	485572028745	e41e3811e6f64cc689038cd81b2ea084

	unix_time	category	amt	is_fraud \
0	1644446321	grocery_net	67.75	0
1	1651244514	misc_pos	70.62	0
2	1652193320	misc_pos	61.11	0
3	1650147611	misc_net	44.73	0
4	1643151808	grocery_net	84.11	0
...
56217	1665895301	home	1.35	0
56218	1655523368	shopping_net	8.92	0
56219	1661841286	food_dining	7.65	0
56220	1671884538	personal_care	5.36	0
56221	1671588352	personal_care	3.20	0

	merchant	merch_lat	merch_long
0	Curtis, Miller and Wade	48.447986	-122.210094
1	Murray-Cohen	49.415702	-122.538354
2	Atkins, Cuevas and Wilson	48.755264	-122.252770

```

3      Rodriguez, Garcia and Patterson 48.287324 -121.639599
4      Gallegos and Sons 49.446720 -121.861955
...
56217      Adkins Ltd 43.183331 -82.862532
56218      Henderson-Allen 42.544910 -83.454984
56219      Ford-Farmer 43.151680 -83.135118
56220      Morrow Inc 42.513979 -83.330267
56221      Brown, Pineda and Carson 42.923344 -83.671483

```

[56222 rows x 10 columns]

/content/sample_data/dataset/transactions_68.csv

```

      cc_num      acct_num      trans_num \
0      345331586923222 903253916954 e27acf7ff2e796804d189d09b6fc4318
1      345331586923222 903253916954 8738514f11d4bd8efeedbe72830bbb13
2      345331586923222 903253916954 21c4b7105a1665323d3d2905a08f0c6d
3      345331586923222 903253916954 99b32dab08deb551530fb4122bae8593
4      345331586923222 903253916954 b0e5828a5ec6e503e93b269a5834b149
...
10763 6011405800851409 925069791808 39cfc9db546430fe7c21353c4a2d9b96
10764 6011405800851409 925069791808 91dbb05fdb0048d2d58cb374bc78568d
10765 6011405800851409 925069791808 39bd617ba45c361230ed8683132ac358
10766 6011405800851409 925069791808 fce6c9ae46fe12a25b8e4dab62ee245b
10767 6011405800851409 925069791808 883333ebc70665c35c8dffbcd4676b13

```

```

      unix_time      category      amt      is_fraud \
0      1648136912      grocery_net 54.57      0
1      1645461229      gas_transport 55.63      0
2      1643817753      gas_transport 35.27      0
3      1642861828      gas_transport 81.31      0
4      1648942439      grocery_pos 61.92      0
...
10763 1667358414      home 6.30      0
10764 1667979257      home 9.69      0
10765 1664690620      health_fitness 4.33      0
10766 1671933758      travel 7.03      0
10767 1672371079      travel 6.72      0

```

```

      merchant      merch_lat      merch_long
0      Solis-Carlson 39.261204 -74.917412
1      Collins, Kramer and Shaw 40.320181 -74.315647
2      Mcgee, Smith and Schroeder 39.192935 -74.892697
3      Jacobs, Miller and Hayes 39.275767 -75.813709
4      Sullivan PLC 40.254785 -76.188640
...
10763      Gonzalez-Young 35.776956 -108.764270
10764      Nelson-Keller 37.383857 -108.277691
10765      Hill LLC 37.185631 -108.199015
10766      Dickerson PLC 36.087391 -108.210665
10767      Thomas, Gonzalez and Garza 36.471286 -108.315140

```

[10768 rows x 10 columns]

/content/sample_data/dataset/transactions_69.csv

```

      cc_num      acct_num      trans_num \
0      345331586923222 903253916954 bc4413f1f92b513b6392aeb9006dbf96
1      345331586923222 903253916954 e1588aa4d518768032f8e8900d65dd11
2      345331586923222 903253916954 26c78862ff75c6ab7a4fb63b64c7b66e
3      345331586923222 903253916954 3fb2c522c3c25d54ebf46623959ed79e

```



```

4      345331586923222  903253916954  fe9b7fe85674bbb66d5928bfb4172dd3
...
40904  6011405800851409  925069791808  e6df96a4304d2c13bfc3be5ad419720a
40905  6011405800851409  925069791808  22ce1c3c153b407e6b047427c6e2d805
40906  6011405800851409  925069791808  2d32c47e5cee7078471438ab2778395f
40907  6011405800851409  925069791808  76107e5d6bb482f133cf032ed0280965
40908  6011405800851409  925069791808  4e34799c93a706b1d228c9af1e6fa325

```

	unix_time	category	amt	is_fraud	merchant
\					
0	1568128791	grocery_pos	9.14	1	Barnes-Weaver
1	1568130339	grocery_pos	300.92	1	Fritz-Sharp
2	1568214698	misc_net	276.87	1	Boyer-Rollins
3	1568133633	gas_transport	864.38	1	Martinez-Hamilton
4	1568134389	misc_net	781.08	1	Fuentes PLC
...
40904	1648181862	home	1.97	0	Davidson, Myers and Adams
40905	1662273483	personal_care	2.69	0	Smith and Sons
40906	1648966196	home	1.06	0	Sheppard and Sons
40907	1595815616	shopping_net	6.40	0	Moore, Walker and Villa
40908	1608986594	shopping_pos	9.39	0	Thomas-Williamson

	merch_lat	merch_long
0	39.721820	-75.227242
1	39.649788	-76.102056
2	40.438444	-75.813218
3	40.338957	-74.697064
4	39.422171	-74.817241
...
40904	37.170950	-109.133669
40905	37.513214	-108.209409
40906	36.261900	-108.462216
40907	35.769100	-108.917904
40908	36.008118	-107.227243

[40909 rows x 10 columns]

/content/sample_data/dataset/transactions_70.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_71.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_72.csv

	cc_num	acct_num	trans_num
\			
0	4006875868240516653	935451073243	fcd2abf4456c72a66c020a105bb3e952
1	4006875868240516653	935451073243	3e424caec4f7789e4762f2ad44e6e22d
2	4006875868240516653	935451073243	566f0253aa820665fbdf7ed4466ca028
3	4006875868240516653	935451073243	9ec7df691704e6489e6f76598df831a8
4	4006875868240516653	935451073243	0ac2f2cad60b52d77358fa16fd25bc87
...
113931	343251790447085	298026601683	4eafc5b16da625c3436f2484d8dcabcb
113932	343251790447085	298026601683	041b18f68b893b68e7aa29c16a08a08d

```

113933      343251790447085  298026601683  a26df5348dfee4839f0fa00155e2ed1c
113934      343251790447085  298026601683  e388fc8055b1426c62640baaae10108f
113935      343251790447085  298026601683  ed07c01568cceb91b94d7572d3dd5e0e

```

```

      unix_time      category      amt  is_fraud  \
0      1548682878  gas_transport    8.15         1
1      1548761403    misc_net   878.70         1
2      1548829254  shopping_net   813.30         1
3      1548761472    misc_net   808.83         1
4      1548847967  entertainment   907.10         1
...      ...      ...      ...      ...
113931  1627636943  health_fitness    7.47         0
113932  1624880326    food_dining    2.30         0
113933  1629797869  health_fitness    5.82         0
113934  1629976595  health_fitness    9.16         0
113935  1669255468  personal_care    8.29         0

```

```

      merchant  merch_lat  merch_long
0      Hudson PLC  39.874427  -74.584194
1      Moyer-Robinson  41.649017  -73.722520
2      Jenkins LLC  39.818060  -74.310105
3      Brown and Sons  39.825515  -74.093356
4      Sanchez, Morgan and Bell  40.636104  -74.117059
...      ...      ...      ...
113931  Horton, Ware and Reynolds  27.707908  -82.320308
113932      Hanson and Sons  27.896372  -83.387356
113933      White LLC  27.110602  -82.426821
113934  Snyder-Johnson  28.110510  -81.891797
113935  Ewing-Russell  27.211428  -83.514176

```

[113936 rows x 10 columns]

/content/sample_data/dataset/transactions_73.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_74.csv

```

      cc_num      acct_num      trans_num  \
0      4841962332605860  3005591724  bb103e57d99cf4fcbec951ba3b846881
1      4841962332605860  3005591724  5eedd746912a0364b29fcd8d55609c83
2      4841962332605860  3005591724  974a31926aa77a7d5cf7827ff1328bfb
3      4841962332605860  3005591724  9648ceb5076615b7ec6defb03ce4927d
4      4841962332605860  3005591724  c3dbac8c46e1f3b554446ce363f644a8
...      ...      ...      ...
20467  6011548720124025  930152599735  d4955be1597f9b201a531fcae62a3b29
20468  6011548720124025  930152599735  6749b73039dfc537b122aaf4fa923439
20469  6011548720124025  930152599735  da6e0fa068612f325fd80b32baff493d
20470  6011548720124025  930152599735  e7aaefd6e8f2c61fef4cdb5b06bdf4d3
20471  6011548720124025  930152599735  eb0b53a361674cabccfcbbc8df6a34e2b

```

```

      unix_time      category      amt  is_fraud  \
0      1660435821    food_dining   69.20         0
1      1654110816  shopping_net   76.31         0
2      1647019155  grocery_pos   44.24         0
3      1659749892  entertainment  101.06         0
4      1656458797  shopping_pos   53.30         0
...      ...      ...      ...      ...

```

20467	1661657234	food_dining	7.23	0
20468	1670743041	personal_care	9.34	0
20469	1669083655	kids_pets	9.77	0
20470	1664931641	home	8.48	0
20471	1662017533	food_dining	4.03	0

	merchant	merch_lat	merch_long
0	Ingram Ltd	40.282097	-75.363992
1	Campbell-Stone	40.165998	-75.945560
2	Duarte Group	40.739818	-77.124587
3	Brown-Washington	39.540628	-76.076041
4	Blair-Warner	40.697782	-75.323660
...
20467	Ward, Castillo and Parrish	38.252115	-76.959677
20468	Alvarez-Bender	38.351350	-75.828709
20469	Baker and Sons	38.556988	-76.348362
20470	Miller-Rivers	38.859558	-76.368137
20471	Wood LLC	38.799449	-76.962863

[20472 rows x 10 columns]

/content/sample_data/dataset/transactions_75.csv

	cc_num	acct_num	trans_num \
0	4909619797978	326474132970	2f300d84888d398856f18462ca81048d
1	4909619797978	326474132970	0d252e7c0b6274ce60c697780e10e93c
2	4909619797978	326474132970	69de7889eac55aa065a5f69eb5eb910d
3	4909619797978	326474132970	54bf046b1b57c1bb821625be5873d9f8
4	4909619797978	326474132970	80ca47f8a6cf1c8e0efeeba1a0d18752
...
33290	180047909863618	921460407733	907df46c5286cdfdfd279a5b5189f386
33291	180047909863618	921460407733	030a03d6b02a1041c2420d55d03018c9
33292	180047909863618	921460407733	72a181181885cd827900bad00f0267cf
33293	180047909863618	921460407733	79da6714843b609dc97999ac1537b456
33294	180047909863618	921460407733	98648b331b2edd19cefa6094d651888d

	unix_time	category	amt	is_fraud	merchant
0	1651329207	misc_pos	41.03	0	Torres-Davis
1	1652121210	misc_pos	66.17	0	Sanchez-Fuentes
2	1650120435	misc_net	65.99	0	Cook, Gordon and Cameron
3	1650759278	misc_net	66.69	0	Hall Group
4	1643827462	grocery_net	80.73	0	Dickerson-Macdonald
...
33290	1671964556	personal_care	8.28	0	Brown and Sons
33291	1667882725	home	4.97	0	Cole Inc
33292	1670129552	kids_pets	5.23	0	Martin and Sons
33293	1666510778	home	4.86	0	Buchanan PLC
33294	1670657615	kids_pets	1.41	0	Henson, Gonzalez and Hall

	merch_lat	merch_long
0	30.473613	-89.006731
1	29.034377	-89.356216
2	29.585810	-90.207869
3	29.813094	-89.504049
4	29.732909	-90.873800
...
33290	35.921557	-119.437841
33291	36.377996	-120.250180

```
33292 35.289740 -119.937551
33293 35.271043 -119.901133
33294 35.578425 -119.564850
```

```
[33295 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_76.csv
```

	cc_num	acct_num	trans_num \
0	30415152638403	373790097990	fcf84ea7b47b8cee956f10734c7a4181
1	30415152638403	373790097990	26d920284fc26391471bbd928c036926
2	30415152638403	373790097990	1f817119d16f0d7ca104eca25045964d
3	30415152638403	373790097990	ad7d18abadbddd5c0b234079eb480fe
4	30415152638403	373790097990	f0bd078b70ef754671d78200c016639f
...
1791	3572989376297459	59703614299	da0d4e86b708db49efe072fd202e06a0
1792	3572989376297459	59703614299	9a476bbebe2dfa760f2429f57f984ff1
1793	3572989376297459	59703614299	eb829d1859acc697d0ee9f112f7157e0
1794	3572989376297459	59703614299	dfedcc469c16b74a66ca33b0dc5177e0
1795	3572989376297459	59703614299	b7bce79f0726606fb28f8b26472ec97c

	unix_time	category	amt	is_fraud \
0	1661731079	grocery_pos	18.74	1
1	1661697591	gas_transport	16.03	1
2	1661702577	grocery_net	351.43	1
3	1661838296	misc_net	334.18	1
4	1661862716	shopping_net	291.39	1
...
1791	1667708604	home	2.58	0
1792	1661071699	food_dining	1.12	0
1793	1670573842	personal_care	9.97	0
1794	1664278959	health_fitness	9.08	0
1795	1658987869	entertainment	4.29	0

	merchant	merch_lat	merch_long
0	Johnson, Shelton and Harris	42.897667	-83.576580
1	Blanchard, Perez and Rodriguez	43.068607	-83.239612
2	Nichols-Williams	43.391165	-84.257335
3	Carr, Douglas and Owens	42.510365	-84.564436
4	Jenkins Group	43.842388	-84.202128
...
1791	Moore-Johnson	39.249580	-76.073873
1792	Gallagher, Johnson and Williams	38.629854	-75.793765
1793	Green, Rasmussen and Macdonald	38.727454	-76.476168
1794	Gonzales Inc	39.679489	-75.992101
1795	Farley-Burns	38.629429	-76.511813

```
[1796 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_77.csv
```

```
Empty DataFrame
```

```
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
```

```
Index: []
```

```
/content/sample_data/dataset/transactions_78.csv
```

```
Empty DataFrame
```

```
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
```

```
Index: []
```

```
/content/sample_data/dataset/transactions_79.csv
```

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_80.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_81.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_82.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_83.csv

	cc_num	acct_num	trans_num	unix_time	category	amt	is_fraud	merchant	merch_lat	merch_long
0	3536881013827217	589419928804	40aaa3e84b730c666fb3721ddccbc777	1557797559	gas_transport	48.96	0	Riddle Ltd	35.665809	-93.282838
1	3536881013827217	589419928804	3dd2e9f8ea6ceaaeda49be55e1e9f62e	1574004878	grocery_pos	59.64	0	Allen-Hall	35.570645	-92.522491
2	3536881013827217	589419928804	267cabe65ab40eddc5049d80af6e9f	1581877582	misc_net	50.17	0	Keith PLC	34.986871	-91.946619
3	3536881013827217	589419928804	55aca6de88e6d73d24fbc5131ed1cebc	1574892875	grocery_pos	54.37	0	Sanchez Ltd	35.916252	-92.669138
4	3536881013827217	589419928804	88a0f255f0a2716dea587182458e2e66	1597448354	shopping_net	52.43	0	Garcia PLC	34.650732	-93.100173
...
4375	3536881013827217	589419928804	22b2ad5ffba1a8562500f948718f6b98	1637898135	home	4.45	0	Thomas, Tran and Mcgee	35.743057	-92.554091
4376	3536881013827217	589419928804	01d7456fa5f3c82147943527230156f3	1667799646	personal_care	2.42	0	Brown-Fernandez	35.016217	-92.004737
4377	3536881013827217	589419928804	1d7533c7a2c6743c50b1f606f00ce3d1	1632888739	health_fitness	5.56	0	Baker, Peters and Oneal	34.551987	-91.889966
4378	3536881013827217	589419928804	5941989c4ab5e8811efa85c47b5bd6ae	1671082031	travel	1.07	0	Fleming, Reyes and Sampson		
4379	3536881013827217	589419928804	c4dfd17e6d95a9efe85785aca3ebde5b	1661426364	personal_care	2.81	0	Lopez-Lowe		

4378 34.404222 -93.455622
 4379 35.359074 -92.585130

[4380 rows x 10 columns]

/content/sample_data/dataset/transactions_84.csv

	cc_num	acct_num	trans_num	\
0	379389701550295	720358760888	ec7682ca6d4f6d98444d33197c2f19e3	
1	379389701550295	720358760888	49b87f27b6b877361bdc97c466aad678	
2	379389701550295	720358760888	d7506c10a64d68586b7171a04b9f29e2	
3	379389701550295	720358760888	f4da3d4e64088a930f38b731b999eb47	
4	379389701550295	720358760888	592d538cc5ecef6f93b3305bdcf87f9	
...	
7539	503870615922	312931364007	1827a3b6983c1bc08e477c3827a9f95c	
7540	503870615922	312931364007	89b743c7e0cda37fbda6e75425432fff	
7541	503870615922	312931364007	4ec87ce6a8c30f10f3c623d5a445bc1d	
7542	503870615922	312931364007	2e883e2a3caf967abcad00f26d9a4857	
7543	503870615922	312931364007	d9b6130112b12ea08bb88418cb746dc8	

	unix_time	category	amt	is_fraud	merchant
0	1646758862	gas_transport	52.15	0	Vasquez, Salazar and Tapia
1	1644348855	gas_transport	89.67	0	Castillo-Hall
2	1646436504	gas_transport	60.22	0	Garcia LLC
3	1641913450	gas_transport	53.98	0	Walker-Price
4	1642194543	gas_transport	34.23	0	Jones PLC
...
7539	1671868944	personal_care	1.21	0	Gonzales-Cook
7540	1661949949	food_dining	7.92	0	Daugherty-Oliver
7541	1671607752	personal_care	9.70	0	Potts, Clark and Davis
7542	1664605605	health_fitness	4.12	0	Warren LLC
7543	1657266200	shopping_pos	9.64	0	Armstrong LLC

	merch_lat	merch_long
0	26.514536	-98.392710
1	25.904754	-97.983460
2	25.141672	-98.243726
3	26.726523	-96.546297
4	26.384450	-97.296645
...
7539	40.971637	-73.022455
7540	41.863649	-72.760656
7541	41.790641	-74.375547
7542	41.757276	-73.316177
7543	40.480041	-72.955262

[7544 rows x 10 columns]

/content/sample_data/dataset/transactions_85.csv

	cc_num	acct_num	trans_num	\
0	4065133387262473	950105810049	ff8ab95f260a50933698120eae2d7823	
1	4065133387262473	950105810049	6e8ac989e247eda22263339099089df0	
2	4065133387262473	950105810049	7c0b387965f87cde4dc735feba081096	
3	4065133387262473	950105810049	f70a510f021294076064a8801d359d20	
4	4065133387262473	950105810049	05e38f18c520d85e4b739b4495b74433	
...	
200068	3519712847818919	500229982535	b4fd4442cb01554618e76c9ba983f14d	
200069	3519712847818919	500229982535	bd9863f6e864130c806346506b959f2c	
200070	3519712847818919	500229982535	3a738476c33c4eb9cb80d3a85b534634	

```

200071 3519712847818919 500229982535 6cd299b37e890c4b4ae8e9caa12237d7
200072 3519712847818919 500229982535 ba5f6f7ee5f84aaf6c1b58b4f01b3a2a

```

```

      unix_time      category      amt  is_fraud  \
0      1549578636  gas_transport  74.77         0
1      1581086984    misc_pos    66.67         0
2      1575647588    misc_net    53.58         0
3      1550100312  gas_transport  90.56         0
4      1558202807   grocery_net  103.19         0
...      ...      ...      ...      ...
200068  1659879624    kids_pets    2.62         0
200069  1659257094    kids_pets    2.32         0
200070  1610699031  shopping_pos  568.13         0
200071  1656581841    kids_pets    4.95         0
200072  1590154009  shopping_net  819.56         0

      merchant  merch_lat  merch_long
0      Hampton PLC  42.451489  -83.546091
1      Rangel-Watkins  43.602999  -83.103037
2      Gilmore-Campbell  43.074693  -82.385364
3      Richardson-Shields  42.457372  -82.997935
4      Murray Ltd  42.745360  -82.521541
...      ...      ...      ...
200068      Fox, Watts and Bowman  34.919506  -77.680559
200069      Weber-Hanson  35.405033  -78.110351
200070      Marshall Group  36.549067  -77.580453
200071  Maldonado, Smith and Young  35.356964  -78.659657
200072      Dalton-Greer  35.145457  -77.568429

```

[200073 rows x 10 columns]

/content/sample_data/dataset/transactions_86.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_87.csv

```

      cc_num      acct_num      trans_num
\
0      675925178633  585807126672  3064e75f27ea9bf568f4b855ce15a101
1      675925178633  585807126672  de4d6ce993e9c101c79d705114e15b7e
2      675925178633  585807126672  bc201f507085da049961906dd7311155
3      675925178633  585807126672  b922b342e491398b3206cd5d9c389a9f
4      675925178633  585807126672  574338de6faed91d4e513dd380cd40fa
...      ...      ...      ...
131537  4966177182498980888  433444305880  48433a3687ce7b29251726553af29724
131538  4966177182498980888  433444305880  560cdcfeef8962a68d937f0ad1bf53e16
131539  4966177182498980888  433444305880  63e2b6f0f41aff9fd8b9c5d78c2caaa5
131540  4966177182498980888  433444305880  9576a45ad0bd53ada1578c318ee62588
131541  4966177182498980888  433444305880  974c647b2f9cbf7cbfbbf35236e2c408

```

```

      unix_time      category      amt  is_fraud  \
0      1556895594   grocery_pos  296.71         1
1      1556900929   grocery_pos  292.38         1
2      1556972138    misc_net    810.89         1
3      1557055338     travel    806.48         1
4      1557062129     travel    991.09         1
...      ...      ...      ...      ...

```

131537	1634453744	home	2.63	0
131538	1672463560	travel	4.89	0
131539	1668150825	personal_care	7.70	0
131540	1645774467	home	6.29	0
131541	1633915691	home	2.94	0

	merchant	merch_lat	merch_long
0	Todd, Lopez and Jordan	43.341577	-90.693991
1	Cole Ltd	42.820622	-90.128354
2	Williams, Rose and Peterson	43.586800	-90.619265
3	Nelson Inc	42.503685	-89.560121
4	Henderson, Ramirez and Powers	43.749382	-89.464017
...
131537	Franco LLC	33.211005	-97.360643
131538	Peterson-Robinson	32.540356	-97.572537
131539	Miller-Wood	32.195652	-96.934660
131540	Nicholson-Rodriguez	32.481082	-97.320664
131541	Baker-West	32.489247	-97.208223

[131542 rows x 10 columns]

/content/sample_data/dataset/transactions_88.csv

	cc_num	acct_num	trans_num \
0	4351161559407816183	917558277935	5860638de86ecfd7fe8b0971b0fa0753
1	4351161559407816183	917558277935	850d93533b340e3e8f20108a33c791e9
2	4351161559407816183	917558277935	5d6dc4e19693c5069e1f549350b391f8
3	4351161559407816183	917558277935	342dc8a176bf6b9e0a8092f4c3835339
4	4351161559407816183	917558277935	2fb538547276deafeed07a356061bf8c
...
24789	378452182307683	724845224567	f90406ba58889ea4aee857ab131aa020
24790	378452182307683	724845224567	011e9cf117a08472777382e8407c1194
24791	378452182307683	724845224567	667f61c73d082d40a53f3a6e52bc6a94
24792	378452182307683	724845224567	bdd897abaa2c73628b050b8e67a1e245
24793	378452182307683	724845224567	d24b4879a559157e29f5260b804d1922

	unix_time	category	amt	is_fraud \
0	1650040378	gas_transport	6.87	1
1	1650041572	grocery_pos	305.89	1
2	1650037327	grocery_pos	295.07	1
3	1650034661	grocery_pos	317.69	1
4	1650115762	misc_net	786.93	1
...
24789	1662875249	health_fitness	5.84	0
24790	1668130318	kids_pets	6.83	0
24791	1667710246	kids_pets	7.91	0
24792	1666153849	home	1.55	0
24793	1665304511	home	1.79	0

	merchant	merch_lat	merch_long
0	Ward, Ferrell and Ward	34.177218	-78.603664
1	Perkins Inc	34.511869	-78.068596
2	Duke-Reid	34.908458	-77.204782
3	Walker-Jimenez	34.162100	-77.779570
4	Cox-Ward	35.010678	-77.970874
...
24789	Walker PLC	41.688481	-95.925128
24790	Evans, Watts and Hartman	40.670948	-95.162493
24791	Allen-Harper	41.762152	-95.106997

24792 Dyer, Harrison and Walters 40.536461 -96.443960
 24793 Harris, Brooks and Parrish 41.521267 -95.808758

[24794 rows x 10 columns]

/content/sample_data/dataset/transactions_89.csv

	cc_num	acct_num	trans_num
\			
0	4351161559407816183	917558277935	2933e877e767c1303be04733a09783b9
1	4351161559407816183	917558277935	79faec57489343a6fe37865c27915790
2	4351161559407816183	917558277935	ef9fce06ada096457180db111d8f73b9
3	4351161559407816183	917558277935	3181199b4c57e5d067844f9977cd93c9
4	4351161559407816183	917558277935	1affc9046546aeac9092e9f39e00e2e
...
105249	6539385258001066	95004578288	424f79d008337da08f352d834dade4eb
105250	6539385258001066	95004578288	f31d15c8b717957ec1aa10cda2ad89c2
105251	6539385258001066	95004578288	04c394d2668aa51a8f2382bf52f5917e
105252	6539385258001066	95004578288	c820e69e4e1cc0bfea4618669dd04428
105253	6539385258001066	95004578288	6e099282fd887e45ce06178fb39a6587

	unix_time	category	amt	is_fraud	\
0	1671133371	grocery_pos	8.77	1	
1	1671119713	gas_transport	9.99	1	
2	1671117183	grocery_pos	260.56	1	
3	1671111799	gas_transport	365.12	1	
4	1671123202	grocery_pos	297.61	1	
...
105249	1652774509	kids_pets	7.74	0	
105250	1658725869	kids_pets	5.82	0	
105251	1668153251	personal_care	5.45	0	
105252	1628774079	health_fitness	3.65	0	
105253	1658822910	kids_pets	9.36	0	

	merchant	merch_lat	merch_long
0	Figueroa-Oneill	33.825049	-77.517009
1	Ray and Sons	34.315535	-78.561810
2	Evans, Robertson and Johnson	34.584630	-79.046950
3	Phillips Group	34.630518	-78.141085
4	Wells, Larsen and Gutierrez	34.324341	-78.752507
...
105249	Humphrey-Tate	26.257294	-81.004965
105250	Hudson and Sons	26.465893	-81.906150
105251	Ortega Group	26.057277	-82.459215
105252	Lee and Sons	26.969636	-82.240450
105253	Evans-Lyons	26.722952	-81.915298

[105254 rows x 10 columns]

/content/sample_data/dataset/transactions_90.csv

	cc_num	acct_num	trans_num	\
0	675925178633	585807126672	63431ec72fe61163125634f521cb88ca	
1	675925178633	585807126672	7e8813f38be8c77be3e6e2d09d53538b	
2	675925178633	585807126672	841f8c81ce9d32263ca3a897bc42a8a0	
3	675925178633	585807126672	8ef81d22ea82b2cbb40b3e712f84bb42	
4	675925178633	585807126672	c6b4413112caf11acef957b8faaa6a3b	
...
33015	4966177182498980888	433444305880	5ed63a423729c5031ff5fd408ca86433	
33016	4966177182498980888	433444305880	dfac2840e0919846c64ee63750335e17	
33017	4966177182498980888	433444305880	6878ac8d885f8851e2c029267407bb5c	

```

33018 4966177182498980888 433444305880 6a3c86eb818fdab401f2733d31a1f64c
33019 4966177182498980888 433444305880 ae325a1c84d52ddcb993c22ac1a65355

```

```

      unix_time      category      amt      is_fraud  \
0      1644944468      gas_transport  52.67          0
1      1647103618      grocery_pos    68.10          0
2      1648731713      grocery_pos    53.13          0
3      1656018931      shopping_pos   50.09          0
4      1642099591      gas_transport   76.34          0
...      ...      ...      ...      ...
33015 1663920310      health_fitness   6.67          0
33016 1671878374          travel        5.41          0
33017 1667641716      kids_pets       2.56          0
33018 1672058951          travel        7.90          0
33019 1664678938          home         3.87          0

```

```

      merchant      merch_lat      merch_long
0      Page-Peterson  43.151961  -90.606901
1      Schmitt, Ramos and Gomez  44.325106  -90.791779
2      Colon, Bryan and Simon  44.351147  -89.546342
3      Monroe, Jones and Blanchard  43.255307  -90.186900
4      Richardson, Evans and Jensen  43.964112  -90.348451
...      ...      ...      ...
33015      Mccarthy and Sons  32.980991  -96.328492
33016      Freeman-White  32.142827  -96.746618
33017      Stone Inc  33.438912  -96.750165
33018      Hobbs, Bartlett and Gibson  32.811325  -96.614513
33019      Scott-Potter  32.861076  -97.657350

```

[33020 rows x 10 columns]

/content/sample_data/dataset/transactions_91.csv

```

      cc_num      acct_num      trans_num  \
0      3524295633432445  625080730349  23a5e62b53e76521e584265728865d77
1      3524295633432445  625080730349  8d596e978f026e175daa359e815a41b6
2      3524295633432445  625080730349  c40a39e0e3e7b0b970cce1dd3912dd2a
3      3524295633432445  625080730349  483ac0a659669c52432b2a1ac2141297
4      3524295633432445  625080730349  d9e48a1e7718cf1fa9e5e9548d02e0e2
...      ...      ...      ...
87647 4096496683496043  835714725252  bf09200ab7f786504c782a33c08c5fec
87648 4096496683496043  835714725252  9b89d3033899ef724358504d53763162
87649 4096496683496043  835714725252  1505155e11be2f228ce532d7a639ed8a
87650 4096496683496043  835714725252  f43c8fa98c19bc07c9b08508498344fb
87651 4096496683496043  835714725252  076112c603252e1c9fcd4ae61c5ecdda

```

```

      unix_time      category      amt      is_fraud  \
0      1622592239      food_dining  72.79          0
1      1623696023      food_dining  61.43          0
2      1613417841      entertainment  56.29          0
3      1622907041      food_dining  74.45          0
4      1595602049      shopping_pos   59.80          0
...      ...      ...      ...      ...
87647 1633686460      health_fitness   6.64          0
87648 1668401788      personal_care   4.00          0
87649 1655784557      kids_pets       1.71          0
87650 1642310208          home         3.09          0
87651 1672396430          travel        1.16          0

```

	merchant	merch_lat	merch_long
0	West LLC	39.164615	-88.423145
1	Powell-Tran	40.045994	-88.036423
2	Delgado-Jones	39.320959	-89.004575
3	Brady Ltd	39.486753	-89.238607
4	Kaiser, Olson and Adams	40.669113	-87.675833
...
87647	Friedman, Wilson and Hernandez	33.617729	-80.103060
87648	Johnson LLC	35.007299	-79.408784
87649	Garner, Coffey and Yates	33.215123	-79.091935
87650	Mendez, Wolfe and Brown	33.677937	-79.821182
87651	Dixon-Sanders	33.661943	-79.399971

[87652 rows x 10 columns]

/content/sample_data/dataset/transactions_92.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_93.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_94.csv

	cc_num	acct_num	trans_num
\			
0	4192832764832	718172762479	fc48cba496d5b589f817d518998885db
1	4192832764832	718172762479	afb1fb76bcb8aab50e4f8949722d7b59
2	4192832764832	718172762479	3e30e7b02637c194aaae37e39c25238e
3	4192832764832	718172762479	efdb9081506530fa363cd2a5a93e6582
4	4192832764832	718172762479	961be24900e708be83c66d1f52d87eda
...
173764	4201027621267969279	401907373048	f6d392550293c08745f033781e588388
173765	4201027621267969279	401907373048	44743f29f922c15f308fdf488d75451c
173766	4201027621267969279	401907373048	de1f64c24baebaeff9c61020b4250828
173767	4201027621267969279	401907373048	b3981e07f8b72a76a615cfcb8d16cb26
173768	4201027621267969279	401907373048	17f2571589ac049493d02d0915379282

	unix_time	category	amt	is_fraud	\
0	1577538432	grocery_pos	50.80	0	
1	1575489918	grocery_pos	94.82	0	
2	1585862952	misc_net	76.60	0	
3	1570292598	grocery_pos	82.11	0	
4	1569780888	grocery_pos	73.76	0	
...	
173764	1668428964	personal_care	9.24	0	
173765	1670677174	travel	6.30	0	
173766	1671159194	travel	9.97	0	
173767	1655527110	kids_pets	1.26	0	
173768	1587447550	misc_pos	4.42	0	

	merchant	merch_lat	merch_long
0	Fisher, Anderson and Meyers	33.841844	-80.943043
1	Spencer-Johnson	34.367517	-79.968788
2	Lopez, Finley and Short	34.352641	-80.446669
3	Harris, Phillips and Reed	32.552003	-80.443405

```

4      Rodriguez, Greene and Kline  34.053885  -81.564384
...
173764      Velasquez-Johnson  34.914276  -84.969137
173765      Parker LLC  35.747809  -86.149930
173766      Williams-Norris  35.163733  -84.818804
173767  Williams, Castro and Henderson  35.042970  -85.192788
173768      Campbell-Knox  34.749912  -85.841008

```

[173769 rows x 10 columns]

/content/sample_data/dataset/transactions_95.csv

```

      cc_num      acct_num      trans_num \
0      630447468723  940251608167  9c43708b2a5a61dacf52522ff88dd724
1      630447468723  940251608167  6a9865ac7a906769a820b750bd11e199
2      630447468723  940251608167  b432123cf05f8e696c5a15aca18c1f42
3      630447468723  940251608167  18e3888c4b9ca4fd2fc4bdcdcb0762716
4      630447468723  940251608167  5b7be22d747ae841a025ad17e13cf130
...
23380  2612982665702796  413003624262  5d65e0d5651d7846a6d5ec35c4042ed8
23381  2612982665702796  413003624262  e191d0ccc54a58e4c11165e81c78f109
23382  2612982665702796  413003624262  431d592d72b07392917492bd9b9cbe7f
23383  2612982665702796  413003624262  2280496f30d84b8df5532a0c19788d8f
23384  2612982665702796  413003624262  0321517491b05876f6449d739a87c91f

```

```

      unix_time      category      amt      is_fraud \
0      1559958242  gas_transport  60.68          0
1      1562371305  gas_transport  46.64          0
2      1571405119  grocery_pos   50.86          0
3      1554476313  gas_transport  42.87          0
4      1552495770  gas_transport  62.06          0
...
23380  1621069708  entertainment  3.35          0
23381  1654938519  kids_pets     26.60          0
23382  1591773970  misc_pos      6.24          0
23383  1608515920  shopping_pos   7.72          0
23384  1607142455  shopping_pos   3.40          0

```

```

      merchant      merch_lat      merch_long
0      Casey, Escobar and Rojas  33.666560  -118.987850
1      Barnes Inc  34.188372  -118.306563
2      Nelson LLC  34.805603  -118.588272
3      Haynes LLC  34.297527  -118.404049
4      Skinner-Bryant  33.661120  -117.525006
...
23380      Bates, Davis and Carter  41.449025  -92.002697
23381      Dennis PLC  42.627426  -91.237140
23382      Schroeder-Schultz  41.833960  -91.484653
23383      Long, Wilson and Richards  42.628400  -90.884021
23384  Robinson, Maldonado and Anderson  41.776129  -92.504073

```

[23385 rows x 10 columns]

/content/sample_data/dataset/transactions_96.csv

```

      cc_num      acct_num      trans_num \
0      180087859964867  328032182625  0b705f8873d7cfa2a980638e6f4cc3f3
1      180087859964867  328032182625  d259e0270b6ff362785ba7b23d97ae19
2      180087859964867  328032182625  70ee83664d3d0d41fbef04e8965e6a15
3      180087859964867  328032182625  0143f8d6987746e5c37a6e17a62c51de
4      180087859964867  328032182625  87b6f5a0a14ca0ecd8b4cdee7ad56ed1

```

```

...
2863 4244626325144889 843963062529 7051acefaff52a7d9252cccaecd3f4ba
2864 4244626325144889 843963062529 ce6d4360977952c72ff942c9b322edb4
2865 4244626325144889 843963062529 c6b7ca42ada4a3f1383b6bc903725e3e
2866 4244626325144889 843963062529 9cbb3c986618fbe82207f5e47c6cffd5
2867 4244626325144889 843963062529 ac815ec84c0074a31cb1bb82472bf019

```

```

      unix_time      category      amt  is_fraud  \
0      1645797842  gas_transport  50.23         0
1      1648743519   grocery_pos  58.27         0
2      1644271505  gas_transport  69.41         0
3      1647043714   grocery_net  48.17         0
4      1648132671   grocery_net  37.62         0

```

```

...
2863 1661938645  entertainment  4.53         0
2864 1669430500         home      2.46         0
2865 1653023556   misc_pos      7.28         0
2866 1657353400   shopping_pos  3.49         0
2867 1662874451   food_dining  8.36         0

```

```

      merchant      merch_lat  merch_long
0      Kerr Inc      42.644852 -112.156222
1      Hernandez-Carter 44.094398 -110.881596
2      Martin, Jackson and Vasquez 43.403739 -111.494580
3      Tucker-Mora    44.206890 -110.437842
4      Patton, Russell and Sheppard 42.624122 -111.702467
...
2863 Robinson, Torres and Phelps 36.925977 -88.902496
2864      Ramos Group 38.738437 -89.187771
2865 Downs, Osborne and Dickson 37.461170 -89.219890
2866 Johnson, Henry and Williams 38.055376 -89.189429
2867      Evans-Roberts 38.711213 -88.786730

```

[2868 rows x 10 columns]

/content/sample_data/dataset/transactions_97.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_98.csv

```

      cc_num      acct_num      trans_num  \
0      6011113885034236 330792868853 3c03a4f7de801ad4ba8da91a176b20ea
1      6011113885034236 330792868853 f300d6d4708e36d0faa1398878fd2538
2      6011113885034236 330792868853 81aae47377e403449fd58ebc3c0a864b
3      6011113885034236 330792868853 f9f55938cc1cd7a202c237af18c67d95
4      6011113885034236 330792868853 7451829cbbc4210eb219d3aeb0286112
...
7296      4886407880267 328429374890 2da3787f308e241bd8fbab190ba6f0ac
7297      4886407880267 328429374890 ded7f4e0357769a3dbbe92c6177b233f
7298      4886407880267 328429374890 bece254c1c4af6d1abab7d4cf303769b
7299      4886407880267 328429374890 a1e915fe0913e38530ed6663f5700381
7300      4886407880267 328429374890 6c32eb5591b3a584d333e7841b1a584e

```

```

      unix_time      category      amt  is_fraud  \
0      1647207959  gas_transport  18.21         1
1      1647329870   shopping_net 345.96         1
2      1647253005   grocery_pos 1001.54         1

```

3	1647344398	shopping_pos	1139.34	1
4	1647344946	shopping_net	1102.35	1
...
7296	1659272960	kids_pets	5.67	0
7297	1661311202	personal_care	5.49	0
7298	1661422108	personal_care	5.28	0
7299	1664595501	personal_care	4.59	0
7300	1628476624	health_fitness	3.65	0

		merchant	merch_lat	merch_long
0		Parsons Inc	32.805508	-91.841875
1		Armstrong-Davis	32.328367	-92.949235
2		Burnett-Moore	31.843032	-93.624804
3		Young-Parker	32.879412	-93.361085
4		Powell, Zuniga and Lopez	32.118928	-92.230792
...	
7296	Schmitt, Kaiser and Hines		39.960422	-97.771690
7297	Martin Inc		40.943384	-97.239755
7298	Jones-Green		39.657437	-97.731889
7299	Reyes-Sparks		40.767574	-98.530162
7300	Finley and Sons		41.475084	-98.596396

[7301 rows x 10 columns]

/content/sample_data/dataset/transactions_99.csv

	cc_num	acct_num	trans_num \
0	4379006548708470	386831851262	56a5bcce8624516e74e11de68943f691
1	4379006548708470	386831851262	5a9a3f8eb577775a9fe9c6594d9f2f24
2	4379006548708470	386831851262	5255b69138b7040b8126b7375c0938c9
3	4379006548708470	386831851262	951ec04dfdd7f44af81b390a9e0f58b1
4	4379006548708470	386831851262	bbdadbac0a614aad393a249a222d2936
...
27769	4246163034387	21414345345	7432f399b000172fcc1e8790f6fc965b
27770	4246163034387	21414345345	7b6dda3778381981a53b9aa5043fff15
27771	4246163034387	21414345345	a42a8a1fa3919010412c0302912acf67
27772	4246163034387	21414345345	45f5c90dd51fe578857b2c6fa8e55c64
27773	4246163034387	21414345345	6ba78d99dcb30dbb5f3611eee87af76a

	unix_time	category	amt	is_fraud \
0	1563727568	gas_transport	70.82	0
1	1554075005	gas_transport	60.84	0
2	1551377103	gas_transport	47.58	0
3	1548960789	gas_transport	46.75	0
4	1565995626	grocery_net	76.61	0
...
27769	1590472417	misc_pos	1.71	0
27770	1624189162	entertainment	1.85	0
27771	1607057141	shopping_pos	5.22	0
27772	1621153529	entertainment	7.79	0
27773	1668858175	personal_care	1.45	0

		merchant	merch_lat	merch_long
0		Watts-Harris	36.852282	-79.605825
1		Ferguson, Turner and Wilson	37.886491	-78.965399
2		Lucero, Hopkins and Hamilton	37.933907	-78.055451
3		Johnson Inc	37.692286	-78.691388
4		Ramirez, Burns and Leblanc	36.738468	-78.607819
...	

```

27769           Moore and Sons  39.947898  -73.705782
27770           Orr, Miller and Lester  40.986267  -73.403986
27771           Petersen-Davis  40.973092  -74.058395
27772  Castillo, Valenzuela and Rodriguez  41.133557  -73.279871
27773           Zhang, Webb and Richard  40.930376  -73.042878

```

[27774 rows x 10 columns]

/content/sample_data/dataset/transactions_100.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_101.csv

```

           cc_num      acct_num      trans_num \
0      3544928787570307  185865727840  5fac62fbd3791bc809a273975220ee0c
1      3544928787570307  185865727840  396cd81280601f061675deb6af85a30f
2      3544928787570307  185865727840  fd7039c9fed1e459c11677f60f492b5b
3      3544928787570307  185865727840  b0a5d253fa0910a80ef706b3e0d55a45
4      3544928787570307  185865727840  697b361d674cc6e91058988a0a3e182e
...
297      4178955990984125  243041877024  1a70e472db3c03eccc3ab638522d98f3
298      4178955990984125  243041877024  04557b186433cf13adf1a6a80f262497
299      4178955990984125  243041877024  3cd8bbbeb24c129c3c42e84f2dd313a8c
300      4178955990984125  243041877024  cff9f45cdaa3e2684b5f6b2e525791a7
301      4178955990984125  243041877024  cf729f31ba1638f4aa49f325677028b3

           unix_time      category      amt  is_fraud \
0      1572201650      grocery_pos    64.75         0
1      1567819529      grocery_pos    79.77         0
2      1554938910      grocery_net    66.64         0
3      1602513055      shopping_pos    67.15         0
4      1560724723      grocery_pos    69.76         0
...
297      1609209137      entertainment    515.46         0
298      1622862456      food_dining    524.15         0
299      1672228741      travel        502.08         0
300      1632123644      health_fitness    439.25         0
301      1637916293      health_fitness    522.45         0

```

```

           merchant  merch_lat  merch_long
0           Perkins Inc  48.493631 -120.667740
1           Jones-Jones  49.400357 -119.465310
2      Hatfield-Schmidt  48.945681 -119.734216
3      Harrison, Buchanan and Bell  48.687851 -120.992058
4      Martinez-Lindsey  48.418225 -119.423310
...
297      Rogers Inc  39.855081 -89.494252
298      Perry, Perkins and Joseph  41.486614 -89.722622
299      Mckinney Inc  40.887878 -89.756732
300      Khan, Collins and Tyler  40.162219 -88.961615
301      Walker, Brown and Evans  40.229859 -89.963685

```

[7302 rows x 10 columns]

/content/sample_data/dataset/transactions_102.csv

```

           cc_num      acct_num      trans_num \
0      2287201859626305  125069449315  d652fefe28c5ea3be49d11aec5618577
1      2287201859626305  125069449315  0ac2e73a9d136ac8aa952771f9c5a09d

```

2	2287201859626305	125069449315	9aec8374317c1a9764e63ec3104dbf5d
3	2287201859626305	125069449315	c1181252f773aa1df971d76976f6bd78
4	2287201859626305	125069449315	a7fb687cda76057fe583e53f68fc6bae
...
11674	347208496498560	11546128003	73b5a42df5c5a3d8e42875b454e04895
11675	347208496498560	11546128003	137123a29fdd9bb1ea234efaa3a89fcb
11676	347208496498560	11546128003	376740d25f9166260ef1ca5a5a48ae65
11677	347208496498560	11546128003	1fd91c31642c82b1ecaf28b534bc3c31
11678	347208496498560	11546128003	3fcf3a8496fd07270c085a222673d0f6

	unix_time	category	amt	is_fraud	\
0	1639759546	grocery_pos	311.48	1	
1	1639756204	grocery_pos	271.92	1	
2	1639801937	shopping_net	969.32	1	
3	1639914588	shopping_net	1065.66	1	
4	1639829443	shopping_net	919.27	1	
...	
11674	1655882162	home	6.50	0	
11675	1668859462	personal_care	7.86	0	
11676	1617958283	shopping_pos	4.29	0	
11677	1672465404	travel	15.25	0	
11678	1649385537	health_fitness	9.37	0	

	merchant	merch_lat	merch_long
0	Rosales PLC	36.318181	-86.818869
1	Fowler, Hill and Gomez	36.825902	-87.710929
2	Garrett LLC	35.954698	-87.010617
3	Kelly-Estrada	37.092787	-86.351365
4	Jordan, Pacheco and Powell	35.842687	-87.327846
...
11674	Sullivan, Garza and Walsh	40.617972	-97.498096
11675	Weaver-Martinez	39.338293	-96.687467
11676	Mitchell-Guerrero	38.882256	-96.321067
11677	Murray Inc	40.022834	-96.246063
11678	Adams, Duke and Davis	39.821922	-96.841142

[11679 rows x 10 columns]

/content/sample_data/dataset/transactions_103.csv

	cc_num	acct_num	trans_num	\
0	4894204664896676	860928862867	cdb7c08664a78598dfc849b21a5d7649	
1	4894204664896676	860928862867	405c518b09f65841239cc609c2f3298f	
2	4894204664896676	860928862867	f154bab56eb53c02ab8ca31d3c34824d	
3	4894204664896676	860928862867	cd5e20e7a2b2d75926ba10487a1922d8	
4	4894204664896676	860928862867	10a903e1243236ef5e0dfe063bd26e9f	
...	
49443	4058416806134954	682282682147	820b140ca394d6e0d8d3cc585231e68a	
49444	4058416806134954	682282682147	990376a269f9856e406b0ea0f7f2fe70	
49445	4058416806134954	682282682147	b2b31f9c3f1842e71ba7bfee71820b42	
49446	4058416806134954	682282682147	81a3ea31f31ee1b8c6d5f50ebec2cf0a	
49447	4058416806134954	682282682147	7354cbb70101ed7117d974f7bcc8959c	

	unix_time	category	amt	is_fraud	\
0	1661362747	misc_net	325.66	1	
1	1661358273	shopping_net	867.95	1	
2	1661437332	shopping_pos	6.86	1	
3	1661355891	grocery_pos	1044.24	1	
4	1661354876	misc_pos	989.23	1	


```

...
49443 1666869881 home 8.03 0
49444 1653560378 shopping_net 9.72 0
49445 1662776866 food_dining 166.67 0
49446 1670909554 personal_care 2525.43 0
49447 1664276695 health_fitness 3.44 0

```

```

merchant merch_lat merch_long
0 Nicholson Inc 43.244987 -86.948124
1 Gibson LLC 43.079094 -85.973611
2 Hayden-Cervantes 42.697161 -86.683374
3 French Ltd 43.479563 -87.002490
4 Parker, Baker and Hernandez 43.298682 -86.264522

```

```

...
49443 Yu-Coleman 29.832867 -94.746088
49444 Ware PLC 29.534841 -94.625824
49445 Stone-Cantu 29.445606 -95.393905
49446 Valenzuela-Atkins 28.965707 -94.869416
49447 Martin-James 30.095859 -96.487658

```

[49448 rows x 10 columns]

/content/sample_data/dataset/transactions_104.csv

```

cc_num acct_num trans_num \
0 180087859964867 328032182625 ec7a4a26ea49e310ac7d5bca4736ab77
1 180087859964867 328032182625 10c62c004cc374d274d765f0f38a01f5
2 180087859964867 328032182625 a2eed92745bb5472ca999ad4e386d079
3 180087859964867 328032182625 175a8da686eaf84fb6d38d9d591879ae
4 180087859964867 328032182625 0949128d17bc9c50c4a890035379e58b
...
13135 4244626325144889 843963062529 8c85184b18bc3c49664933f05843b8e6
13136 4244626325144889 843963062529 91e73436c45f1566807adf8ee028746c
13137 4244626325144889 843963062529 b4c79a70240d2588f1d8f2bd1ce5d3b3
13138 4244626325144889 843963062529 201fb3ba7b5670286a1c03dabff71b4f
13139 4244626325144889 843963062529 902bb08ab24d92d8db96182d6e1c728a

```

```

unix_time category amt is_fraud \
0 1560785234 gas_transport 56.89 0
1 1572824070 grocery_net 52.02 0
2 1573259849 grocery_net 66.51 0
3 1554314156 gas_transport 73.58 0
4 1563202970 gas_transport 59.38 0
...
13135 1664542969 personal_care 5.37 0
13136 1592447779 misc_net 5.18 0
13137 1588319921 misc_net 9.82 0
13138 1671328486 travel 5.69 0
13139 1610852373 shopping_net 4.03 0

```

```

merchant merch_lat merch_long
0 Ray, Schultz and Miller 44.299301 -110.364047
1 Carter, Merritt and Hernandez 43.311914 -111.600285
2 Baker Inc 43.987965 -110.695481
3 Garcia-Ruiz 44.039282 -111.710095
4 Smith, Jones and Weber 42.448859 -110.850013
...
13135 Jackson-Salazar 38.640153 -90.201743
13136 Daniels-Brown 37.526660 -90.227453

```

```

13137      Baird, Reynolds and Dunn  38.828703  -89.374977
13138      Robertson-Rollins  38.870555  -89.294867
13139      Medina, Moore and Knight  38.233597  -90.125744

```

```
[13140 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_105.csv
```

```
Empty DataFrame
```

```
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
```

```
Index: []
```

```
/content/sample_data/dataset/transactions_106.csv
```

```

      cc_num      acct_num      trans_num \
0      4623546622650  163959980994  1834ca43383a8c34692156d3a7e8754a
1      4623546622650  163959980994  fcf4bd62606bec2c79f9553ca291511d
2      4623546622650  163959980994  346baf5f16c2e360b639d86586b2145b
3      4623546622650  163959980994  9b81413107db95bac834513f7b6f2f30
4      4623546622650  163959980994  b9bec7575192206b5ad7455c69ca902b
...      ...      ...      ...
30689  4209140169332843  295677321717  29bbdfb27b5e962772797684a0697e11
30690  4209140169332843  295677321717  24dd8418c00b3606c994e35bf63402f6
30691  4209140169332843  295677321717  9c654ac247fe6e16e1205c4cbacb780d
30692  4209140169332843  295677321717  698bc6bd07fb6537d08de77dabca4a7a
30693  4209140169332843  295677321717  14e6da7743217d47bb5409b25ceda334

```

```

      unix_time      category      amt      is_fraud      merchant \
0      1549836559  gas_transport  35.45      0      Owen-Smith
1      1572700469  grocery_pos   46.77      0      Williams-Clements
2      1570298672  grocery_pos   49.56      0      Griffin-Wright
3      1561838139  gas_transport  66.28      0      Booth and Sons
4      1555454488  gas_transport  60.27      0      Diaz, Terrell and Cooke
...      ...      ...      ...      ...
30689  1655011253      kids_pets    6.05      0      Hernandez Inc
30690  1612521535  shopping_pos   8.10      0      Nelson-Martinez
30691  1582880061      misc_net    6.38      0      Robinson Ltd
30692  1591323998      misc_pos    5.05      0      Lopez Ltd
30693  1632832498  food_dining    8.67      0      Parsons-Moss

```

```

      merch_lat  merch_long
0      38.687477 -120.463476
1      39.250744 -120.946621
2      38.335985 -121.124639
3      38.303503 -120.352800
4      37.756225 -121.156036
...      ...      ...
30689  31.339078 -97.164701
30690  31.022857 -98.289674
30691  31.313630 -97.683709
30692  31.222730 -97.606216
30693  30.604240 -97.798320

```

```
[30694 rows x 10 columns]
```

```
/content/sample_data/dataset/transactions_107.csv
```

```
Empty DataFrame
```

```
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
```

```
Index: []
```

```
/content/sample_data/dataset/transactions_108.csv
```

	cc_num	acct_num	trans_num	\
0	4358841834374016	521500975371	5c5c65ca7fc8502f04c80ffb939c0ce8	
1	4358841834374016	521500975371	284b205d38c37eb202dea700301f8f3f	
2	4358841834374016	521500975371	cefc1abc961b1283b9acd4ca0cf879cf	
3	4358841834374016	521500975371	9d024808ed79734d891a79290c6348ec	
4	4358841834374016	521500975371	153cf8da6f4c4bffe8c17300a21996bd	
...	
42630	4601080181030683173	117887236764	2bd920f0aef85600b64452295a862099	
42631	4601080181030683173	117887236764	eea7e9929d03906916434299e58a9fef	
42632	4601080181030683173	117887236764	99e0bbb8d16b8fedae9b829609abaa8c	
42633	4601080181030683173	117887236764	f3108c8846dc0e9b73e5d1983738ecf3	
42634	4601080181030683173	117887236764	7aa314a083dedf06f1d800f5143ba7a3	

	unix_time	category	amt	is_fraud	\
0	1654187487	misc_net	672.09	1	
1	1654278310	food_dining	109.59	1	
2	1654348459	kids_pets	17.97	1	
3	1651333213	misc_pos	87.32	0	
4	1649101604	grocery_pos	65.67	0	
...	
42630	1660974836	food_dining	6.52	0	
42631	1671849242	personal_care	8.05	0	
42632	1660360763	entertainment	8.59	0	
42633	1671784186	personal_care	6.15	0	
42634	1662725818	food_dining	8.15	0	

	merchant	merch_lat	merch_long
0	Mckee, Mathis and Vasquez	37.842202	-86.674674
1	Garcia-Brooks	37.373642	-85.691648
2	Mcfarland, Singh and Chang	38.796781	-85.000260
3	Rodriguez, Henry and Mercer	37.676216	-85.006481
4	Austin PLC	39.072061	-86.662571
...
42630	Little-Franklin	40.433492	-75.304137
42631	Campbell, Rogers and Delacruz	40.952608	-74.571212
42632	Christensen-Lewis	40.859246	-74.362744
42633	Brown, Jackson and Aguirre	40.358288	-74.649146
42634	Mora LLC	40.320032	-73.772590

[42635 rows x 10 columns]

/content/sample_data/dataset/transactions_109.csv

	cc_num	acct_num	trans_num	\
0	6574625916011247	438009262556	8046a8ef1ece4983ce2040c92721fddc	
1	6574625916011247	438009262556	3db171b01dc386b87a501ee17809a63e	
2	6574625916011247	438009262556	20c4deccc8f76b2a66b54dfc61e79efe	
3	6574625916011247	438009262556	9911c55f4cd2472ead7451d83dc86b31	
4	6574625916011247	438009262556	afd0a6b31190f26188c6be987d5e1ced	
...	
143140	371833597487680	397728875125	287eb521afd28d09a5d30e461911ca70	
143141	371833597487680	397728875125	835f5f809e4c5c305f39a7c952379dc1	
143142	371833597487680	397728875125	0f033f4417a989ae84161bdc35aef9c9	
143143	371833597487680	397728875125	40e97299c8f8e20ad8aae419fb978ad1	
143144	371833597487680	397728875125	e01f04094495dabb735e8f75e7aeca435	

	unix_time	category	amt	is_fraud	\
0	1597329674	shopping_pos	64.48	0	
1	1579962218	misc_pos	68.75	0	

2	1582937393	misc_pos	64.55	0
3	1582748537	misc_pos	42.92	0
4	1603719294	shopping_pos	90.82	0
...
143140	1642847156	home	5.14	0
143141	1640782502	home	9.98	0
143142	1627796362	health_fitness	8.50	0
143143	1635989021	home	2.89	0
143144	1671363547	travel	8.10	0

	merchant	merch_lat	merch_long
0	York, Collins and McIntosh	35.757206	-80.558868
1	Crawford, Moody and Fisher	36.649869	-79.910345
2	Spencer Ltd	36.043182	-80.078021
3	Cole, Marsh and Washington	36.173966	-80.140223
4	Goodwin-Clarke	35.347004	-80.348457
...
143140	Graham Group	39.249913	-75.095242
143141	Diaz Ltd	39.266721	-74.402007
143142	Stanton-Aguilar	39.900388	-73.996303
143143	Reese-Cook	40.912703	-74.393432
143144	Rogers Inc	40.554934	-73.511701

[143145 rows x 10 columns]

/content/sample_data/dataset/transactions_110.csv

	cc_num	acct_num	trans_num \
0	4894204664896676	860928862867	dbf4212423288ae3672641583ff0626a
1	4894204664896676	860928862867	9ecccc128ee0da46b727de526ea4b73a0
2	4894204664896676	860928862867	aee2f26d63aaa03e1d75773b851c4940
3	4894204664896676	860928862867	6129e99f45d514485f21f069056472b0
4	4894204664896676	860928862867	a731f778bd6585bc8f2f1f076739fbdb
...
201516	4058416806134954	682282682147	37ca7d2e1081bd5fe4dac9cd6a723156
201517	4058416806134954	682282682147	df82b77fa18ccdf9ba72711f9342a6ee
201518	4058416806134954	682282682147	fcd382b5de488198aeaecafd5a2db09e
201519	4058416806134954	682282682147	cc9368920d51359383189100ec2a84dd
201520	4058416806134954	682282682147	8f68ed23b717715c520054c408fd766b

	unix_time	category	amt	is_fraud \
0	1549308866	gas_transport	59.86	0
1	1566051632	grocery_pos	57.98	0
2	1582810972	misc_pos	99.42	0
3	1565472062	grocery_pos	64.41	0
4	1564710915	grocery_pos	49.81	0
...
201516	1602839780	shopping_pos	7.52	0
201517	1613887278	entertainment	6.63	0
201518	1594988768	shopping_net	1.40	0
201519	1640431673	home	7.73	0
201520	1599824350	shopping_pos	7.17	0

	merchant	merch_lat	merch_long
0	Manning and Sons	43.261216	-85.306298
1	Wilcox-Costa	43.766524	-86.149271
2	Wood-Johnson	42.438595	-85.508920
3	Morrison, Newman and Miller	43.594697	-85.829157
4	Hall, Roth and Smith	43.249256	-85.421506

```

...
201516 Thompson, Davis and Trevino 28.734463 -95.918218
201517 Ramirez LLC 28.958458 -96.296083
201518 Williams Inc 29.689095 -96.520683
201519 Stewart, Rivera and Fitzgerald 30.389369 -95.739505
201520 Thomas, Moore and Porter 30.067951 -96.394274

```

[201521 rows x 10 columns]

/content/sample_data/dataset/transactions_111.csv

```

      cc_num      acct_num      trans_num \
0      6571589042902575  212660437303  7f415fecb09dec96c66587ab515a4a8
1      6571589042902575  212660437303  bbd52050ad2ddb74526b330f0518d38b
2      6571589042902575  212660437303  d1b1253920b45d5dd3c9e41db4f8c031
3      6571589042902575  212660437303  ee869c387419b380ff096ae1793c7033
4      6571589042902575  212660437303  7c903196b1a8d8d3370a1471cce97e3d
...
40896  180026732044133  481110497017  fd700a2516362a21f2ade78153584583
40897  180026732044133  481110497017  37c59387d70dc43399aebf2d5eab350f
40898  180026732044133  481110497017  b4cc1693e56065186d94e7d28d84c322
40899  180026732044133  481110497017  a034ea5eb14316e4c80d50fc3e40b638
40900  180026732044133  481110497017  0d4c7e32697c589939a978bb6eaa5534

```

```

      unix_time      category      amt      is_fraud \
0      1558123341  gas_transport  81.10          0
1      1550241939  gas_transport  51.12          0
2      1565721301  grocery_net   51.02          0
3      1567278363  grocery_net   48.34          0
4      1557100130  gas_transport  48.95          0
...
40896  1601798021  shopping_net   6.25          0
40897  1577092700  misc_net       6.97          0
40898  1603540921  shopping_pos   4.46          0
40899  1671967105  travel        150.91         0
40900  1654863504  kids_pets      7.18          0

```

```

      merchant      merch_lat      merch_long
0      Smith and Sons  41.712924  -86.879877
1      Olson-Phillips  41.684306  -87.430849
2      Allen Ltd      41.106043  -88.074070
3      Miller, Armstrong and Hansen  42.021945  -88.348569
4      Sanford, Smith and Miller  42.047215  -87.670626
...
40896  Simmons-Adams  28.951126  -97.049569
40897  Escobar Inc    28.626805  -97.899080
40898  Hale-Velasquez  27.744856  -96.965882
40899  Anderson, Gutierrez and Hernandez  28.581825  -97.625857
40900  Patel-Hernandez  29.280659  -97.369667

```

[40901 rows x 10 columns]

/content/sample_data/dataset/transactions_112.csv

```

      cc_num      acct_num      trans_num \
0      4218196001337  888022315787  1702559e2eb52864f4c88693a2fffe4d
1      4218196001337  888022315787  3b743f5774d47a9d16ca217107f12f3d
2      4218196001337  888022315787  71cacdd6c4d4f5f6b390849238c9f58b
3      4218196001337  888022315787  9f51e27f41796ccf1916a88c3cdd741c
4      4218196001337  888022315787  880e04b08afa88461a99d4b6c2455083
...

```

23700	3592927524372804	632265815336	fac6a11c5bf1e3788bf8032b5fa7100f
23701	3592927524372804	632265815336	c5b4c24681f467d2431dd050747d5e9c
23702	3592927524372804	632265815336	8ded2d088fbc3899faaa1a43de7d24fa
23703	3592927524372804	632265815336	56bad7a26b51c74337096f380335e4fc
23704	3592927524372804	632265815336	fd160a87fdff6dd95d71a0ebbe62b343

	unix_time	category	amt	is_fraud	\
0	1656633080	shopping_pos	81.91	0	
1	1658342808	entertainment	57.62	0	
2	1652194483	misc_pos	66.46	0	
3	1655690394	shopping_pos	76.14	0	
4	1653235243	misc_pos	62.09	0	
...	
23700	1669020076	kids_pets	3.02	0	
23701	1671082396	personal_care	5.29	0	
23702	1670128691	kids_pets	1210.16	0	
23703	1665825923	home	9.97	0	
23704	1664100575	home	2.99	0	

	merchant	merch_lat	merch_long
0	Hansen, Heath and Wood	39.294324	-75.339080
1	Myers, Dorsey and Wallace	39.918987	-75.227456
2	Tyler, Kemp and Woodard	39.600319	-74.319276
3	Brown, Dixon and Reyes	39.230300	-74.581277
4	Johnson-Smith	40.520865	-74.884797
...
23700	Holmes Group	41.813111	-88.601853
23701	Johnson, Walker and Harvey	42.778239	-88.205736
23702	Davenport-Suarez	41.444327	-89.041171
23703	Sandoval, Clark and Suarez	42.549121	-88.123709
23704	Beard-Sims	41.602967	-87.702483

[23705 rows x 10 columns]

/content/sample_data/dataset/transactions_113.csv

	cc_num	acct_num	trans_num	\
0	6571589042902575	212660437303	505f59be167d528e37c5f110e88b5fc4	
1	6571589042902575	212660437303	0bdae9dc9094d47562a00cb345ad5714	
2	6571589042902575	212660437303	bc961acfb1203833779450e0c4a9b191	
3	6571589042902575	212660437303	3cdc6783d3db8479bf92abd99af931f3	
4	6571589042902575	212660437303	2bf163f54d24ea416877b0340c18b0b6	
...	
9317	180026732044133	481110497017	d0f2301604032b3ee028ee098fa25691	
9318	180026732044133	481110497017	453bd655e755a3f1dc6c5f67afeae11e	
9319	180026732044133	481110497017	41087f61d2249d656c64c40f62f4cca7	
9320	180026732044133	481110497017	66795cac29a9951a3e19fb437fd45dc0	
9321	180026732044133	481110497017	548db1d711d63638d7886c76d14080fd	

	unix_time	category	amt	is_fraud	\
0	1647189095	grocery_net	71.05	0	
1	1645897902	gas_transport	39.14	0	
2	1648673429	grocery_net	100.41	0	
3	1643483703	gas_transport	63.72	0	
4	1648334245	grocery_net	59.71	0	
...	
9317	1651491628	misc_net	3.46	0	
9318	1663486148	food_dining	7.77	0	
9319	1659581663	entertainment	7.82	0	

9320	1665308290	health_fitness	3.24	0
9321	1658652065	shopping_pos	6.74	0
		merchant	merch_lat	merch_long
0		Hall LLC	41.098790	-87.974668
1		Spencer-Strong	42.086824	-87.291339
2		Berg-Allen	41.250491	-87.861449
3		Bailey LLC	41.877105	-87.131671
4	Allen, Fernandez and Berry		41.683339	-86.911524
...	
9317		Jones Ltd	28.196350	-96.916082
9318		Rodriguez-Mendez	28.262135	-97.067160
9319		Martin-Moore	28.403518	-97.671567
9320		Aguilar PLC	28.254176	-97.913409
9321	Cox, Jacobs and Gallagher		28.056610	-97.224771

[9322 rows x 10 columns]

/content/sample_data/dataset/transactions_114.csv

	cc_num	acct_num	trans_num	\
0	180008357017032	31024004064	a73e97ef3caee7db2c3732e8f235e48a	
1	180008357017032	31024004064	7c18977208e040f1a75e3d728a13275b	
2	180008357017032	31024004064	c080180694b58c682e99808c859dcc2a	
3	180008357017032	31024004064	2927381652960bb54d9a25b75ceb20aa	
4	180008357017032	31024004064	9e1642d285ba28546dac493fe2adec9c	
...
52590	4371450311809	499348764671	93411717667e0343552cf469b26684c3	
52591	4371450311809	499348764671	3dc863510aff3a4766e5e5cd83367a28	
52592	4371450311809	499348764671	cc0372cd9dfbc86f009c43165bc7d4fa	
52593	4371450311809	499348764671	3eace47d14d148b44dcc0d31c013e56	
52594	4371450311809	499348764671	71d056b9bc685dd15b34fd62f69d5fdc	

	unix_time	category	amt	is_fraud	\
0	1586012017	gas_transport	8.75	1	
1	1586040635	misc_net	307.72	1	
2	1586013352	grocery_pos	921.96	1	
3	1586092067	shopping_net	1117.20	1	
4	1586067506	shopping_net	933.17	1	
...
52590	1667129972	personal_care	3.27	0	
52591	1629462810	food_dining	9.04	0	
52592	1636621013	health_fitness	8.96	0	
52593	1657257659	kids_pets	6.19	0	
52594	1634111706	food_dining	7.66	0	

	merchant	merch_lat	merch_long
0	Booker-Edwards	41.467300	-73.390131
1	Taylor-Lewis	41.527065	-74.507437
2	Schneider-Cook	41.113090	-73.458387
3	Mcclain, Richardson and Hill	39.715148	-74.568753
4	Jenkins Group	41.491582	-74.221104
...
52590	Hall, Jackson and Burton	33.968948	-117.594945
52591	Martin-Lee	32.844903	-117.015118
52592	Walker-Johnson	32.398619	-116.624057
52593	Rodriguez Ltd	32.522398	-116.594915
52594	Li, Santiago and Hall	33.834480	-116.953995

```
[52595 rows x 10 columns]
/content/sample_data/dataset/transactions_115.csv
Empty DataFrame
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
Index: []
/content/sample_data/dataset/transactions_116.csv
```

	cc_num	acct_num	trans_num
0	30160398863650	652941429071	2daeb1bc14a5fa9defe39da127c5f05c
1	30160398863650	652941429071	7cc91990b8e39a92126c543f554e3deb
2	30160398863650	652941429071	10f0a4060d14be628a0cd78283876f6a
3	30160398863650	652941429071	f0121658e727fa016c9eac0ba93f3ed5
4	30160398863650	652941429071	35d5ee9e8cab076c30440511cd58a583
...
191279	4594853400127623	997039479727	35c382c3e0d91cd0d86e9bf23024b65c
191280	4594853400127623	997039479727	0c042fe6684e38010b64a4a0931be3b3
191281	4594853400127623	997039479727	a9965ca82f1d6747b46cc9c3283c0fde
191282	4594853400127623	997039479727	b9c97543e1db4b0e534a7251d9943a47
191283	4594853400127623	997039479727	3643c517c18aa62b6e8fe6e1a09f6ff3

	unix_time	category	amt	is_fraud
0	1586108647	grocery_pos	297.97	1
1	1586189618	misc_net	848.72	1
2	1586263382	kids_pets	1111.78	1
3	1586262804	shopping_net	1008.57	1
4	1586261439	shopping_net	1001.60	1
...
191279	1656768217	kids_pets	5.90	0
191280	1620800637	entertainment	5.89	0
191281	1658289072	kids_pets	5.05	0
191282	1632298233	food_dining	3.67	0
191283	1657850873	kids_pets	1.05	0

	merchant	merch_lat	merch_long
0	Marshall PLC	39.582247	-74.921040
1	Murray, Anderson and Nelson	39.761380	-73.973780
2	Deleon-Casey	40.631310	-73.947432
3	Jones-Lambert	40.258404	-75.319419
4	Rice-Thomas	41.014704	-75.289434
...
191279	Moody, Sanford and Henderson	30.880549	-98.049659
191280	Potter-Irwin	30.003433	-98.301710
191281	Stevens Inc	29.849981	-98.745169
191282	Hanna-Rice	30.768154	-98.685435
191283	Herrera-Gomez	30.260646	-98.144484

```
[191284 rows x 10 columns]
/content/sample_data/dataset/transactions_117.csv
```

	cc_num	acct_num	trans_num
0	4623546622650	163959980994	25cad0b9ac03035bca3504222c349ebb
1	4623546622650	163959980994	6251db113d897772f0f11a45cd4a7d80
2	4623546622650	163959980994	7d4d9ff7faf6c2381f2a789259525d97
3	4623546622650	163959980994	a05b247ed54d415b24cb1e29d633e88a
4	4623546622650	163959980994	294e0966d5af69ea8aaf277b81df8fd3
...
8250	4209140169332843	295677321717	8cf1bee57a7a99b0cb6502b0e215c1d7
8251	4209140169332843	295677321717	bf2182cbd28941794598fa3474c60d3d


```

8252 4209140169332843 295677321717 b572da053250d22bc5ae68abc999eb9d
8253 4209140169332843 295677321717 aebfdff305f435168b62dff41ebf27da
8254 4209140169332843 295677321717 27e480adb5fa2d1985cb5358d76a17ae

```

	unix_time	category	amt	is_fraud	merchant
\					
0	1647381659	grocery_net	62.33	0	Spencer PLC
1	1648141143	grocery_net	54.68	0	Lewis PLC
2	1648510775	grocery_net	33.51	0	Nelson, Salazar and Harris
3	1644091223	gas_transport	76.76	0	Rogers, Smith and Murphy
4	1645921272	gas_transport	46.81	0	Rios, Lutz and Arnold
...
8250	1666180439	health_fitness	5.56	0	Moore, Fleming and Collins
8251	1661603898	entertainment	9.29	0	May, Hampton and Allen
8252	1662104567	food_dining	7.17	0	Mason-Becker
8253	1653548875	misc_pos	4.92	0	Weber Ltd
8254	1651134505	misc_net	5.33	0	Fleming and Sons

	merch_lat	merch_long
0	38.073267	-120.783607
1	38.853961	-121.034676
2	39.608384	-121.480749
3	38.320699	-121.620389
4	38.682314	-122.015948
...
8250	30.065224	-97.720181
8251	29.994879	-97.120455
8252	29.967950	-97.668996
8253	30.861101	-98.115220
8254	30.754427	-96.783791

[8255 rows x 10 columns]

/content/sample_data/dataset/transactions_118.csv

	cc_num	acct_num	trans_num	\
0	3531261685071052	897566796191	1c4b6f9417d0208c1e2714da020b9795	
1	3531261685071052	897566796191	f17176165273d2762cf8a6b0e0f71b70	
2	3531261685071052	897566796191	1d4f75322bb04868a2b816a335657c49	
3	3531261685071052	897566796191	96a49da5d1b141669bb9a2751bfbd1ad	
4	3531261685071052	897566796191	97485fc5fadd9720084a55886d92dc81	
...
23716	6011235158501532	483297892147	fc185ef2e4a5da500c478e74d0c0007e	
23717	6011235158501532	483297892147	dabaf12caf2e40c9caa53e2b76a0e9af	
23718	6011235158501532	483297892147	7eae3a49eacc65aadcc230dbd414db771	
23719	6011235158501532	483297892147	c2d3addcc9d0eda05bbcb81f6a240fe4	
23720	6011235158501532	483297892147	36550b0c8f34d1325c78618f112c09b1	

	unix_time	category	amt	is_fraud	\
0	1644239931	gas_transport	87.10	0	
1	1646491824	grocery_net	54.29	0	
2	1654893818	shopping_net	60.84	0	
3	1655306146	shopping_pos	66.65	0	
4	1647536554	grocery_pos	91.89	0	
...
23716	1659490822	grocery_pos	318.04	1	
23717	1659451753	grocery_pos	289.34	1	
23718	1659620446	shopping_pos	806.60	1	
23719	1659534864	misc_net	753.01	1	

23720	1659528259	misc_net	843.64	1
-------	------------	----------	--------	---

	merchant	merch_lat	merch_long
0	Ferguson-Powers	42.281935	-91.394942
1	Castro PLC	43.280428	-91.640818
2	Johnson LLC	43.359863	-92.286408
3	Marsh, Wang and Farley	42.539713	-91.278176
4	Moses-Watkins	41.663150	-91.271195
...
23716	Robinson LLC	40.390987	-86.305156
23717	Campbell, Morales and Lewis	41.491843	-84.516130
23718	Anderson-Stout	40.187452	-84.651411
23719	Wright, Brooks and Ellis	40.702016	-84.384925
23720	Cox-Morrow	42.037215	-86.134604

[23721 rows x 10 columns]

/content/sample_data/dataset/transactions_119.csv

	cc_num	acct_num	trans_num	\
0	6011034324451072	347300118003	5fd8114f0a9e6173c4b6b2d4a1357629	
1	6011034324451072	347300118003	cc6a2a82b2e153d17a61a89b3998f010	
2	6011034324451072	347300118003	036a0c03732a15f6e14a78ebc3954786	
3	6011034324451072	347300118003	ea3861ae99d1115d28df7e9c3a93e819	
4	6011034324451072	347300118003	1106c5d5f386d895e1f159a1f0f25699	
...
2152	6011416389448467	18557023491	e121499b46bed48c715b3654a7249efd	
2153	6011416389448467	18557023491	dcbc02583f1837bc3abaf29ea519b951	
2154	6011416389448467	18557023491	72ab6569c38eeefcf5dbd610309f2484	
2155	6011416389448467	18557023491	65b789cc8346bf10ac28f4448ae2f2c0	
2156	6011416389448467	18557023491	a0f101b191dcb7a9290e4365229f6297	

	unix_time	category	amt	is_fraud	merchant
\					
0	1642086410	gas_transport	67.37	0	Kent PLC
1	1655946207	shopping_pos	71.92	0	Smith, Taylor and Palmer
2	1647697745	grocery_pos	76.54	0	Johnson Group
3	1647995663	grocery_pos	63.04	0	Adams Ltd
4	1649186549	grocery_pos	83.62	0	Pratt, Joyce and Barajas
...
2152	1664795003	health_fitness	526.09	0	Roberts Group
2153	1663043203	health_fitness	463.47	0	Dunn LLC
2154	1667614022	home	445.14	0	Jones, Copeland and Flynn
2155	1657616465	shopping_pos	487.63	0	Smith, Howell and Singh
2156	1662365473	food_dining	446.57	0	Rojas-Decker

	merch_lat	merch_long
0	31.722756	-109.043992
1	32.666096	-108.996380
2	31.681885	-110.087164
3	31.643283	-108.291944
4	31.581803	-110.153007
...
2152	38.131983	-84.541069
2153	40.006959	-85.546131
2154	38.110988	-86.240769
2155	39.878313	-85.992831
2156	39.330204	-85.030692

```
[2157 rows x 10 columns]
/content/sample_data/dataset/transactions_120.csv
Empty DataFrame
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
Index: []
/content/sample_data/dataset/transactions_121.csv
Empty DataFrame
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
Index: []
/content/sample_data/dataset/transactions_122.csv
Empty DataFrame
Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]
Index: []
/content/sample_data/dataset/transactions_123.csv
```

	cc_num	acct_num	trans_num
\			
0	180039174101196	2468061102	2e4c63d64c0053d94b980fd521b598f3
1	180039174101196	2468061102	f783f65aec14f996cea63b29549fdaf2
2	180039174101196	2468061102	ca224cd04c5bd6f9e4f0dd108bd1c0d3
3	180039174101196	2468061102	aa59638ccf1d76b5c273d2043d073ed2
4	180039174101196	2468061102	da5cf20ea3c48795490ba7d9861e2fa3
...
119783	4851579197111212590	832138782905	93d76617de21d5feefed08f326ff4898
119784	4851579197111212590	832138782905	34c22aae7d1fa7e212d8263bbc3e8dd5
119785	4851579197111212590	832138782905	d7fe8e5746ea89330fcd82735ea8a663
119786	4851579197111212590	832138782905	302acb8d54ae95cc58d25d15c89f79fc
119787	4851579197111212590	832138782905	47d3592b3097730463a7bef3d172a3a0

	unix_time	category	amt	is_fraud	\
0	1555186188	gas_transport	56.22	0	
1	1591494957	shopping_net	76.53	0	
2	1589646172	shopping_net	47.62	0	
3	1609188911	entertainment	58.82	0	
4	1599846641	shopping_pos	50.83	0	
...
119783	1667119394	personal_care	4.28	0	
119784	1629786349	health_fitness	3.38	0	
119785	1632211461	health_fitness	2.91	0	
119786	1653706304	kids_pets	6.71	0	
119787	1671877236	travel	4.33	0	

	merchant	merch_lat	merch_long
0	White and Sons	33.384458	-117.418197
1	Jordan Group	32.679111	-116.997700
2	Garcia, Miller and Schmidt	33.260507	-118.211902
3	Martinez Group	34.025481	-118.152173
4	Warren LLC	33.197729	-116.843387
...
119783	Benitez-Johnson	26.442585	-79.274969
119784	Mcgee-Torres	26.115670	-79.304866
119785	White-Gardner	25.600866	-80.576894
119786	Young, Bray and Williams	25.128220	-79.890427
119787	Dudley-Foley	25.638094	-80.878658

[119788 rows x 10 columns]

/content/sample_data/dataset/transactions_124.csv

	cc_num	acct_num	trans_num	\
0	675942799072	231045526966	4fad6ab602d7252581daec24aea0cd37	
1	675942799072	231045526966	1b2c2695612a4ebc496ed500289752b7	
2	675942799072	231045526966	59fb5b9acb46de6c494735638a1d1b27	
3	675942799072	231045526966	0649b455737390fdff71ad2d48929c83	
4	675942799072	231045526966	f6bcd1d64a73276060f5565cebf845ea	
...	
37968	4328348315203771070	417495431869	0a1c5325549a054d10869961daf4c6ff	
37969	4328348315203771070	417495431869	804271d1498bd9c1bd0b7d2cd1cc5fea	
37970	4328348315203771070	417495431869	b6743e55b3fe64648949b77b2dfd96f1	
37971	4328348315203771070	417495431869	f1f3bf1be673c553439d faafcaa88b36	
37972	4328348315203771070	417495431869	6db9afa9eaf26834dafdbfd9bf517058	

	unix_time	category	amt	is_fraud	\
0	1592489610	grocery_pos	6.65	1	
1	1592606893	misc_net	300.38	1	
2	1592575413	misc_net	279.30	1	
3	1592496484	gas_transport	1005.81	1	
4	1592500835	grocery_pos	733.25	1	
...	
37968	1620529747	entertainment	7.61	0	
37969	1657972693	kids_pets	5.82	0	
37970	1584274905	misc_pos	30.19	0	
37971	1650188760	home	8.55	0	
37972	1581246487	misc_net	7.71	0	

	merchant	merch_lat	merch_long
0	Caldwell-Becker	35.493504	-79.592587
1	Mitchell PLC	34.871254	-79.981769
2	Jenkins-Becker	35.443005	-79.688249
3	Briggs PLC	34.590573	-79.339442
4	Roman-Edwards	34.192906	-78.957848
...
37968	Meyer Inc	33.082472	-117.421648
37969	Houston, Sutton and Porter	33.113990	-118.544472
37970	Miller, Campbell and Collins	33.903606	-117.747451
37971	Benitez-Chandler	33.995792	-119.072749
37972	Wiggins-Ward	33.590210	-117.408844

[37973 rows x 10 columns]

/content/sample_data/dataset/transactions_125.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_126.csv

	cc_num	acct_num	trans_num	\
0	4379006548708470	386831851262	43028f629cc929b691de83bf57e20cd4	
1	4379006548708470	386831851262	00efa3471e368ad849831c94da04d69b	
2	4379006548708470	386831851262	02fa6cf3372ff1cc1919a41c6c7bb69f	
3	4379006548708470	386831851262	ec0fbda723956c6dce37bbfabd287cb3	
4	4379006548708470	386831851262	111634de854d656d4caf10458f5cf6c3	
...	
17204	4246163034387	21414345345	7f3bb2e1c41f8368f051f3b1ddbc00e6	
17205	4246163034387	21414345345	772dc226e2ba06ef89a048a556002744	

17206	4246163034387	21414345345	4b2bf4a26c72c5e819b302dcaf5914a7
17207	4246163034387	21414345345	05e184263981bc11634b160fcea07def
17208	4246163034387	21414345345	ce9c813feae7995654972143908253a0

	unix_time	category	amt	is_fraud	\
0	1666621971	grocery_pos	13.70	1	
1	1666622753	gas_transport	267.20	1	
2	1666747314	shopping_net	784.18	1	
3	1666755425	misc_pos	7.36	1	
4	1666787362	shopping_pos	839.96	1	
...	
17204	1656769455	shopping_pos	6.74	0	
17205	1671621799	personal_care	9.26	0	
17206	1665816098	health_fitness	6.97	0	
17207	1660221147	entertainment	6.72	0	
17208	1669034611	kids_pets	9.74	0	

	merchant	merch_lat	merch_long
0	Jimenez-Cole	37.466919	-78.238190
1	Hansen, Schultz and Jones	36.864743	-79.457201
2	Sanchez, Lopez and Long	37.685351	-78.204569
3	Bauer, Garcia and Frazier	38.394966	-79.976426
4	Johnson-Mendez	38.107685	-79.827665
...
17204	Thompson-Johnston	40.496869	-73.281103
17205	Mendez, Gardner and Gray	41.680521	-74.438121
17206	Thompson Group	39.783465	-74.230034
17207	Schmitt-Baker	40.626174	-72.750100
17208	Rios PLC	41.139579	-72.608642

[17209 rows x 10 columns]

/content/sample_data/dataset/transactions_127.csv

	cc_num	acct_num	trans_num	\
0	4129957047264261	424039957450	603ed7b0864bb0ba9d4aa7ca371951a5	
1	4129957047264261	424039957450	d44eaf448dbb0cbbbec8089080571dc4	
2	4129957047264261	424039957450	6550b19529b307a7aa69ebdcd852c140	
3	4129957047264261	424039957450	646f2aae6c759e618f651eb73aa78ec1	
4	4129957047264261	424039957450	fc2c3e5ab34acfd1923149e3979aba1	
...	
25873	4413395479406453	929002904576	660c294fe8d540431cc54d25077dc4d1	
25874	4413395479406453	929002904576	4be02c8434ff59abf1d72f2b8372be0f	
25875	4413395479406453	929002904576	0b9f714b93b5b1369cdd4fbbc111297d	
25876	4413395479406453	929002904576	db1d2149e9c939413b0a33b672b496f2	
25877	4413395479406453	929002904576	fa5d1154d8add5ac13931ab0008df642	

	unix_time	category	amt	is_fraud	merchant	\
0	1652477805	misc_pos	66.80	0	Lee-Stewart	
1	1656786392	shopping_pos	66.85	0	Jones, Goodman and Mason	
2	1646320811	grocery_net	61.94	0	Cook, Marsh and Gomez	
3	1655396260	shopping_pos	61.21	0	George-Morgan	
4	1655428844	shopping_pos	72.38	0	Porter-Ferguson	
...	
25873	1670377413	kids_pets	2.89	0	Morris-Tran	
25874	1664342570	home	9.63	0	Lopez PLC	
25875	1672056151	travel	1.78	0	Sanchez Ltd	
25876	1668741851	kids_pets	8.18	0	Garcia PLC	
25877	1672208268	travel	7.44	0	Lee-Carter	

```

      merch_lat merch_long
0      42.724568 -83.525088
1      42.575795 -83.731731
2      42.932297 -84.086652
3      42.974256 -84.128618
4      42.646975 -82.265299
...      ...
25873  34.687688 -118.891984
25874  34.381254 -117.017665
25875  33.151970 -118.836058
25876  33.925622 -118.417269
25877  33.684247 -118.620519

```

[25878 rows x 10 columns]

/content/sample_data/dataset/transactions_128.csv

```

      cc_num      acct_num      trans_num      unix_time
\
0  675986887866  987989291479  631a1b576d2800cb3d35f9834811ec81  1662647087
1  675986887866  987989291479  258762d9c6eee3785db82a2c14c2ae8a  1662649679
2  675986887866  987989291479  0a54b1479ceb7a29d4103ed633c1ae94  1662656302
3  675986887866  987989291479  26f6569188051ca461c410a9589b6c20  1662657503
4  675986887866  987989291479  e86b36e3dc4b594d9abedc0066e93f87  1662815833
5  675986887866  987989291479  8e2759bc257d45fce0a64d92cfddb33e  1662816108
6  675986887866  987989291479  4793f473f9de850f16d3f3f62141f124  1662813693
7  675986887866  987989291479  f641ddaf172e3a0f643df4a9528b5866  1662813294
8  675986887866  987989291479  ba90e5f6ebcb351a966a1c924cb6f99b  1662776020

```

```

      category      amt  is_fraud      merchant  merch_lat \
0  gas_transport    7.08         1  Henry, Cummings and Warner  37.168597
1  grocery_pos    288.19         1    Johnson, Cobb and Page  36.499252
2  grocery_pos    272.18         1      Henderson-Powell  37.492935
3  misc_net      797.21         1        Watts Group  38.004762
4  shopping_net   918.03         1      Simpson-Douglas  37.943482
5  shopping_pos  1197.08         1  Smith, White and Goodwin  36.791794
6  shopping_pos   894.79         1    Peterson, King and Khan  36.473072
7  shopping_net  1172.11         1      Kennedy-Davis  37.619328
8  shopping_net  1109.35         1      Bell-Odonnell  38.316452

```

```

      merch_long
0  -88.613281
1  -89.705726
2  -89.451845
3  -88.950043
4  -88.210426
5  -89.237758
6  -88.103133
7  -89.404065
8  -89.336295

```

/content/sample_data/dataset/transactions_129.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

/content/sample_data/dataset/transactions_130.csv

```

      cc_num      acct_num      trans_num \
0      180039174101196      2468061102  d2460713987ae1fef2b26e52f35da717

```

1	180039174101196	2468061102	38726478064eaf255f6b37bb78c763bd
2	180039174101196	2468061102	87775267b929bfa6568c3c8ad89ffa2e
3	180039174101196	2468061102	10d9746eb3277135feadd8b03231abc
4	180039174101196	2468061102	83167ba4ffbf3bf1335dc3fe4ba90f33
...
32667	2704390552958068	756767582811	1f091e64bd3f2382b3a4f90889329c32
32668	2704390552958068	756767582811	971aa63530f8f8fb1de28472e6aa53bc
32669	2704390552958068	756767582811	555941e364d1f591a6ae4088557f8ed3
32670	2704390552958068	756767582811	b756330bb5da439e28412f40e7a5f9e3
32671	2704390552958068	756767582811	d8b2caf59c06c36fea6980964f0a1fd6

	unix_time	category	amt	is_fraud	\
0	1652797224	misc_pos	68.66	0	
1	1645401070	gas_transport	88.33	0	
2	1653411429	misc_pos	79.26	0	
3	1649373563	grocery_pos	56.46	0	
4	1658430154	entertainment	59.83	0	
...
32667	1662431076	health_fitness	8.21	0	
32668	1665826484	home	4.79	0	
32669	1662292853	health_fitness	1.99	0	
32670	1670561624	personal_care	6.64	0	
32671	1671856381	travel	12.74	0	

	merchant	merch_lat	merch_long
0	Mathews-Ortiz	34.090562	-116.528455
1	Bowen Inc	34.152683	-116.728265
2	Foster, Palmer and Murray	33.444484	-116.883501
3	Hunt-Miller	32.409793	-117.475535
4	Sullivan, Erickson and Frazier	32.423665	-116.833080
...
32667	Anderson-King	33.831954	-116.587500
32668	Vang, Higgins and Decker	33.591525	-117.962834
32669	Ibarra LLC	34.370334	-117.628700
32670	Casey, Gonzales and Mueller	33.690374	-117.301976
32671	Cruz-Johnson	34.012253	-116.333944

[32672 rows x 10 columns]

/content/sample_data/dataset/transactions_131.csv

Empty DataFrame

Columns: [cc_num, acct_num, trans_num, unix_time, category, amt, is_fraud, merchant, merch_lat, merch_long]

Index: []

Merging data to one file

```
In [3]: import os
import csv

directory = "/content/sample_data/dataset"
merged_filename = "transactions.csv"

# Creating a CSV writer object with '|' as the delimiter for the merged file
with open(os.path.join(directory, merged_filename), 'w', newline='') as outfile:
    writer = csv.writer(outfile, delimiter='|')

    # Loop through all the files that match the pattern
    for filename in glob.glob(directory + "/transactions_[0-9]*.csv"):

        # Open the file for reading
        with open(filename, 'r') as infile:

            # Creating a CSV reader object with '|' as the delimiter
            reader = csv.reader(infile, delimiter='|')

            # Loop through each row in the file and write it to the merged file
            for row in reader:
                writer.writerow(row)

trans= pd.read_csv('/content/sample_data/dataset/transactions.csv', delimiter
='|')
trans.to_csv('transactions.csv', index=False)
trans
```


Out[3]:

	cc_num	acct_num	trans_num	unix_time	
0	345331586923222	903253916954	e27acf7ff2e796804d189d09b6fc4318	1648136912	gr
1	345331586923222	903253916954	8738514f11d4bd8efeedbe72830bbb13	1645461229	gas_
2	345331586923222	903253916954	21c4b7105a1665323d3d2905a08f0c6d	1643817753	gas_
3	345331586923222	903253916954	99b32dab08deb551530fb4122bae8593	1642861828	gas_
4	345331586923222	903253916954	b0e5828a5ec6e503e93b269a5834b149	1648942439	grc
...
4261030	6584732365337749	843643900137	7b2b63508d92d5dd61446fb5cb03e301	1667282786	pers
4261031	6584732365337749	843643900137	1df5a10b715d536a666143ebb429a818	1591184669	
4261032	6584732365337749	843643900137	67f8305d28b564b34ccf264b9f4ea764	1625045557	fo
4261033	6584732365337749	843643900137	4d3333cd26ba3ac993151b30b474e94e	1667465864	pers
4261034	6584732365337749	843643900137	e9126d852f94ab72fa6b40ee7e44c6f1	1620636587	ente
4261035 rows × 10 columns					

Loading customer data

```
In [4]: data = pd.read_csv('/content/sample_data/dataset/customers.csv', delimiter='|')  
        print(data)  
        data.to_csv('customers.csv', index=False)
```

	ssn	cc_num	first	last	gender	\
0	115-04-4507	4218196001337	Jonathan	Johnson	M	
1	715-55-5575	4351161559407816183	Elaine	Fuller	F	
2	167-48-5821	4192832764832	Michael	Cameron	M	
3	406-83-7518	4238849696532874	Brandon	Williams	M	
4	697-93-1877	4514627048281480	Lisa	Hernandez	F	
..	
995	392-96-7670	30125158904184	Sarah	Martin	F	
996	594-17-7993	180047909863618	Erin	Wells	F	
997	196-93-1156	4371450311809	Michael	James	M	
998	895-65-9304	3519925692476886	Michael	Lewis	M	
999	107-40-0160	343251790447085	Isaac	Smith	M	

	street	city	state	zip	lat
\					
0	863 Lawrence Valleys	Ambler	PA	19002	40.1809
1	310 Kendra Common Apt. 164	Leland	NC	28451	34.2680
2	05641 Robin Port	Cordova	SC	29039	33.4275
3	26916 Carlson Mountain	Birmingham	AL	35242	33.3813
4	809 Burns Creek	Fargo	GA	31631	30.7166
..
995	1666 Jenna Unions	Denver	CO	80236	39.6535
996	444 Alexis Estate Suite 824	Wasco	CA	93280	35.6480
997	914 Cassandra Gateway Suite 061	Escondido	CA	92026	33.1605
998	34141 Katelyn Path	Mattapan	MA	2126	42.2739
999	67148 Rose Cliff Apt. 314	Saint Petersburg	FL	33710	27.7898

	long	city_pop	job	dob	\
0	-75.2156	32412	Accounting technician	1959-10-03	
1	-78.0578	27112	Professor Emeritus	1963-06-07	
2	-80.8857	4215	International aid/development worker	1973-05-30	
3	-86.7046	493806	Seismic interpreter	1942-12-26	
4	-82.5801	559	Medical laboratory scientific officer	1939-05-22	
..	
995	-105.0376	990452	Colour technologist	1993-08-31	
996	-119.4487	27152	Software engineer	1982-05-27	
997	-117.0978	171802	Agricultural consultant	1999-05-14	
998	-71.0939	25562	Civil Service administrator	1994-02-23	
999	-82.7243	341043	Manufacturing systems engineer	1942-07-21	

	acct_num
0	888022315787
1	917558277935
2	718172762479
3	947268892251
4	888335239225
..	...
995	201218361926
996	921460407733
997	499348764671
998	617852071065
999	298026601683

[1000 rows x 15 columns]

```
In [5]: cus = pd.read_csv('customers.csv')
cus.head(10)
```

Out[5]:

	ssn	cc_num	first	last	gender	street	city	state	zip
0	115-04-4507	4218196001337	Jonathan	Johnson	M	863 Lawrence Valleys	Ambler	PA	19002
1	715-55-5575	4351161559407816183	Elaine	Fuller	F	310 Kendra Common Apt. 164	Leland	NC	28451
2	167-48-5821	4192832764832	Michael	Cameron	M	05641 Robin Port	Cordova	SC	29039
3	406-83-7518	4238849696532874	Brandon	Williams	M	26916 Carlson Mountain	Birmingham	AL	35242
4	697-93-1877	4514627048281480	Lisa	Hernandez	F	809 Burns Creek	Fargo	GA	31631
5	246-08-3947	3517182278248964	Vincent	Hickman	M	465 Lam Mission	Forsyth	GA	31029
6	347-03-9639	213193010310510	Kayla	Ramirez	F	82997 Kristen Valley Suite 311	Patterson	LA	70392
7	479-94-0888	4065133387262473	Danielle	Leblanc	F	108 Tanya Terrace Apt. 267	Rochester	MI	48307
8	061-22-6210	630447468723	Cheryl	Robinson	F	805 Brendan Neck	Lakewood	CA	90715
9	162-08-8321	2222913619399092	Bradley	Sandoval	M	3534 Anderson Rue Suite 751	Glendale	CA	91202

Exploring the data

Finding null values

```
In [6]: trans.isnull().sum()
```

```
Out[6]: cc_num      0
        acct_num    0
        trans_num    0
        unix_time    0
        category     0
        amt          0
        is_fraud     0
        merchant     0
        merch_lat    0
        merch_long    0
        dtype: int64
```

```
In [7]: cus.isnull().sum()
```

```
Out[7]: ssn        0
        cc_num      0
        first       0
        last        0
        gender      0
        street      0
        city        0
        state       0
        zip         0
        lat         0
        long        0
        city_pop    0
        job         0
        dob         0
        acct_num    0
        dtype: int64
```

```
In [8]: trans.dtypes
```

```
Out[8]: cc_num      object
        acct_num    object
        trans_num    object
        unix_time    object
        category     object
        amt          object
        is_fraud     object
        merchant     object
        merch_lat    object
        merch_long    object
        dtype: object
```

```
In [9]: cus.dtypes
```

```
Out[9]: ssn          object
cc_num      int64
first       object
last        object
gender      object
street      object
city        object
state       object
zip         int64
lat         float64
long        float64
city_pop    int64
job         object
dob         object
acct_num    int64
dtype: object
```

Merging customers and transactional data

```
In [10]: merged_data = pd.merge(trans,cus,on='acct_num')
print(merged_data.columns)
```

```
Index(['cc_num_x', 'acct_num', 'trans_num', 'unix_time', 'category', 'amt',
      'is_fraud', 'merchant', 'merch_lat', 'merch_long', 'ssn', 'cc_num_y',
      'first', 'last', 'gender', 'street', 'city', 'state', 'zip', 'lat',
      'long', 'city_pop', 'job', 'dob'],
      dtype='object')
```

Saving the dataframe in a new csv file

```
In [11]: merged_data.to_csv('merged_data.csv', index=False)
m_data = pd.read_csv('merged_data.csv')
m_data.head(7)
```

Out[11]:

	cc_num_x	acct_num	trans_num	unix_time	category	
0	676305278100	930558644123	34821a75f2a6ef420bfe6b775ef7ebb6	1558633432	grocery_net	1!
1	676305278100	930558644123	47504e5f0f060d0ba136241a7381d78c	1572390467	grocery_pos	1!
2	676305278100	930558644123	330ac8785c386bcfedccda9450c6fcf1	1547506176	gas_transport	2:
3	676305278100	930558644123	bf1b90eb153d61172bdaec4698e71b90	1553981956	grocery_net	2
4	676305278100	930558644123	663a9e929cebc05428ec8a4293b0c6	1547666941	gas_transport	2:
5	676305278100	930558644123	b9055adc45accd9f0abf0a3bbbac137	1576777975	misc_net	1!
6	676305278100	930558644123	218cfc86fddc0af02f557cee48df8465	1559252707	grocery_net	2:

7 rows × 24 columns

Exploring merged data

Finding missing values in merged data

```
In [12]: m_data.isnull().sum()
```

```
Out[12]: cc_num_x      0
         acct_num     0
         trans_num    0
         unix_time    0
         category     0
         amt          0
         is_fraud     0
         merchant     0
         merch_lat    0
         merch_long   0
         ssn         0
         cc_num_y     0
         first        0
         last         0
         gender       0
         street       0
         city         0
         state        0
         zip          0
         lat          0
         long         0
         city_pop     0
         job          0
         dob          0
         dtype: int64
```


In [13]: `m_data.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1574059 entries, 0 to 1574058
Data columns (total 24 columns):
#   Column          Non-Null Count  Dtype
---  -
0   cc_num_x        1574059 non-null  int64
1   acct_num        1574059 non-null  int64
2   trans_num       1574059 non-null  object
3   unix_time       1574059 non-null  int64
4   category        1574059 non-null  object
5   amt             1574059 non-null  float64
6   is_fraud        1574059 non-null  int64
7   merchant        1574059 non-null  object
8   merch_lat       1574059 non-null  float64
9   merch_long      1574059 non-null  float64
10  ssn             1574059 non-null  object
11  cc_num_y        1574059 non-null  int64
12  first           1574059 non-null  object
13  last            1574059 non-null  object
14  gender          1574059 non-null  object
15  street          1574059 non-null  object
16  city            1574059 non-null  object
17  state           1574059 non-null  object
18  zip             1574059 non-null  int64
19  lat             1574059 non-null  float64
20  long            1574059 non-null  float64
21  city_pop        1574059 non-null  int64
22  job             1574059 non-null  object
23  dob            1574059 non-null  object
dtypes: float64(5), int64(7), object(12)
memory usage: 288.2+ MB
```

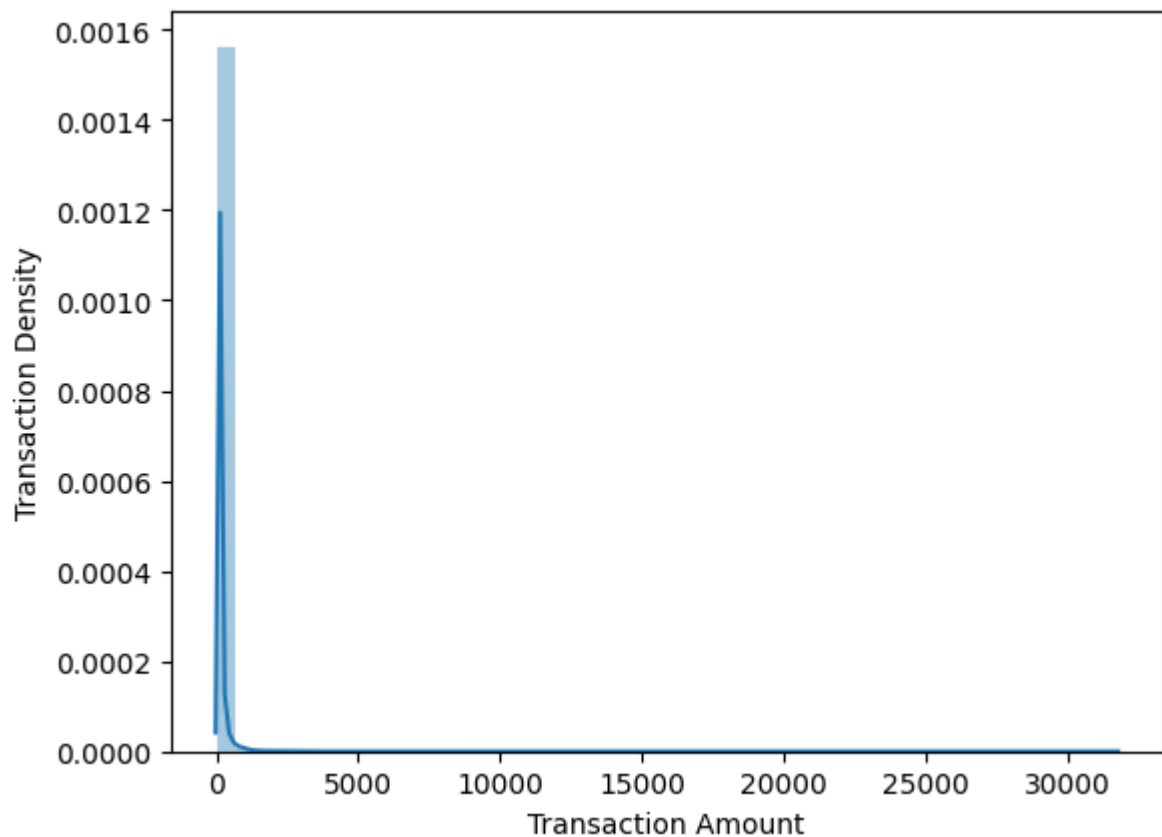
In [14]: `m_data.describe()`

Out[14]:

	cc_num_x	acct_num	unix_time	amt	is_fraud	merch_lat	n
count	1.574059e+06	1.574059e+06	1.574059e+06	1.574059e+06	1.574059e+06	1.574059e+06	1.5
mean	3.411746e+17	4.765663e+11	1.610985e+09	7.165052e+01	6.689711e-04	3.741340e+01	-9.1
std	1.192472e+18	2.843779e+11	3.653670e+07	1.554710e+02	2.585583e-02	5.332681e+00	1.6
min	6.040027e+10	2.348758e+09	1.546261e+09	1.000000e+00	0.000000e+00	1.870250e+01	-1.6
25%	2.131042e+14	2.307161e+11	1.577574e+09	9.710000e+00	0.000000e+00	3.368105e+01	-9.8
50%	3.533223e+15	4.559282e+11	1.609393e+09	4.617000e+01	0.000000e+00	3.843453e+01	-8.6
75%	4.609380e+15	7.173788e+11	1.641035e+09	8.358000e+01	0.000000e+00	4.132451e+01	-7.9
max	4.983666e+18	9.993899e+11	1.672492e+09	3.177259e+04	1.000000e+00	6.263027e+01	-6.8

Visualizing Transaction Density

```
In [15]: sns.distplot(m_data['amt'])
plt.xlabel("Transaction Amount")
plt.ylabel("Transaction Density");
```



Deriving customer's age from date of birth

```
In [16]: m_data['dob'] = pd.to_datetime(m_data['dob'])
m_data['age'] = (pd.to_datetime('today') - m_data['dob']).astype('<m8[Y]')
m_data['age']
```

```
Out[16]: 0      26.0
1      26.0
2      26.0
3      26.0
4      26.0
...
1574054 28.0
1574055 28.0
1574056 28.0
1574057 28.0
1574058 28.0
Name: age, Length: 1574059, dtype: float64
```

Creating bar plot to visualize the number of customers in different age group

```

In [17]: ages = np.array(m_data['age'])
transaction_amounts = np.array(m_data['amt'])

age_groups = [(10,20),(20,30),(30,40),(40,50),(50,60),(60,70),(70,80),(80,90),
(90,100)]
avg_transaction_amounts = list()
midpoints = list()

for age_group in age_groups:
    lower_bound, upper_bound = age_group
    mask = (ages >= lower_bound) & (ages < upper_bound)
    group_avg = np.mean(transaction_amounts[mask])
    avg_transaction_amounts.append(group_avg)
    midpoints.append((lower_bound + upper_bound) / 2)

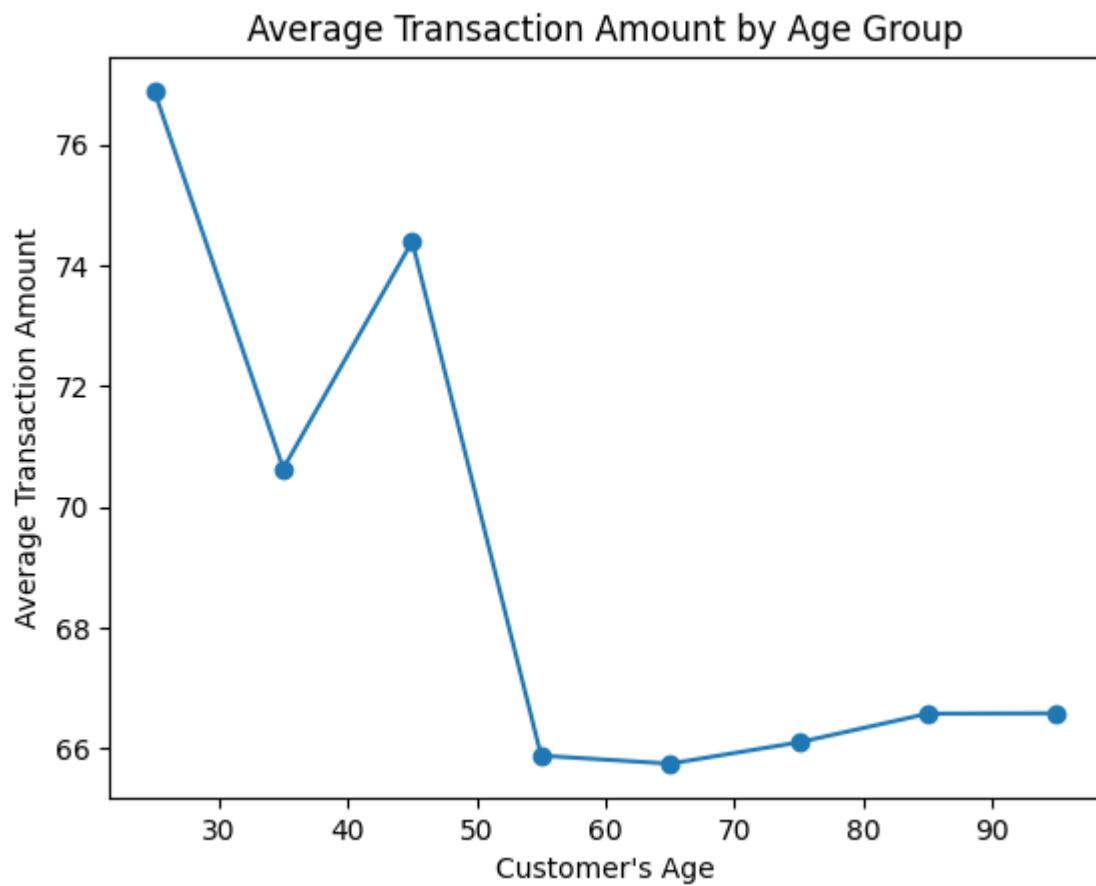
```

Creating line plot to identify any shifts or patterns in the amount of transactions as age increases.

```

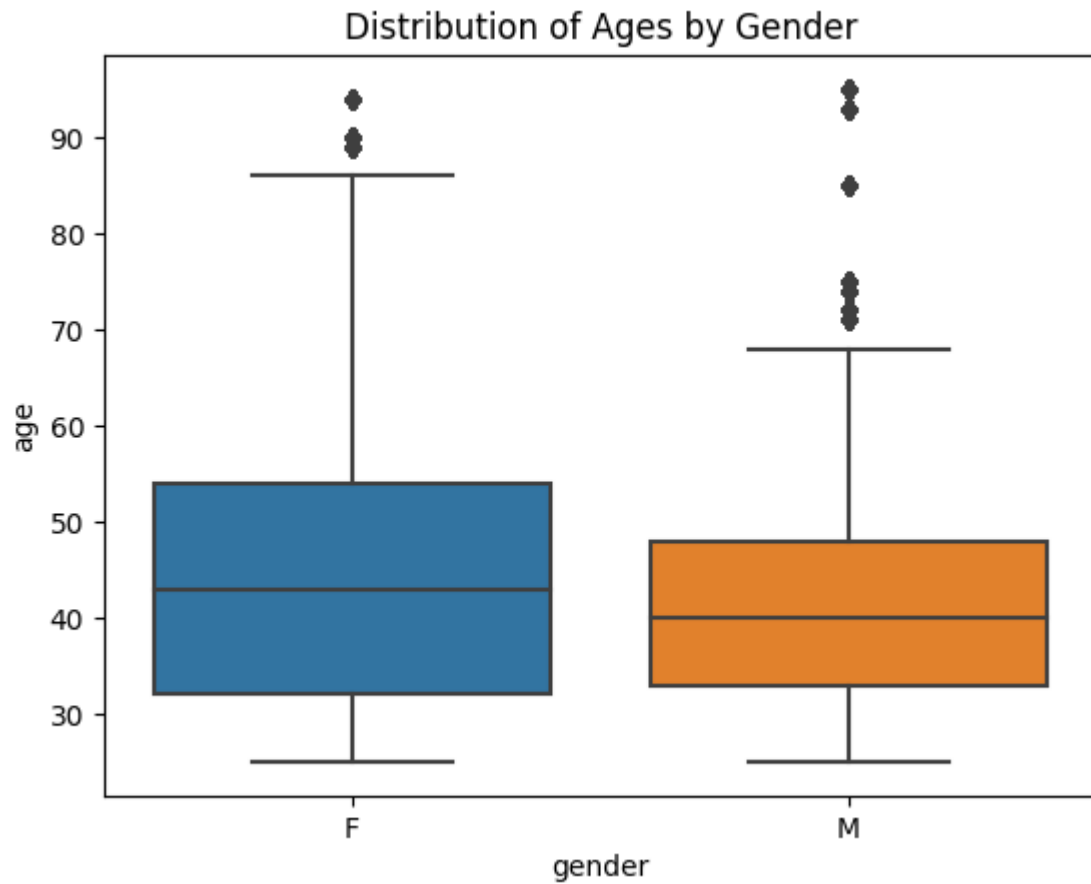
In [18]: plt.plot(midpoints, avg_transaction_amounts, marker='o')
plt.xlabel("Customer's Age")
plt.ylabel('Average Transaction Amount')
plt.title('Average Transaction Amount by Age Group')
plt.show()

```



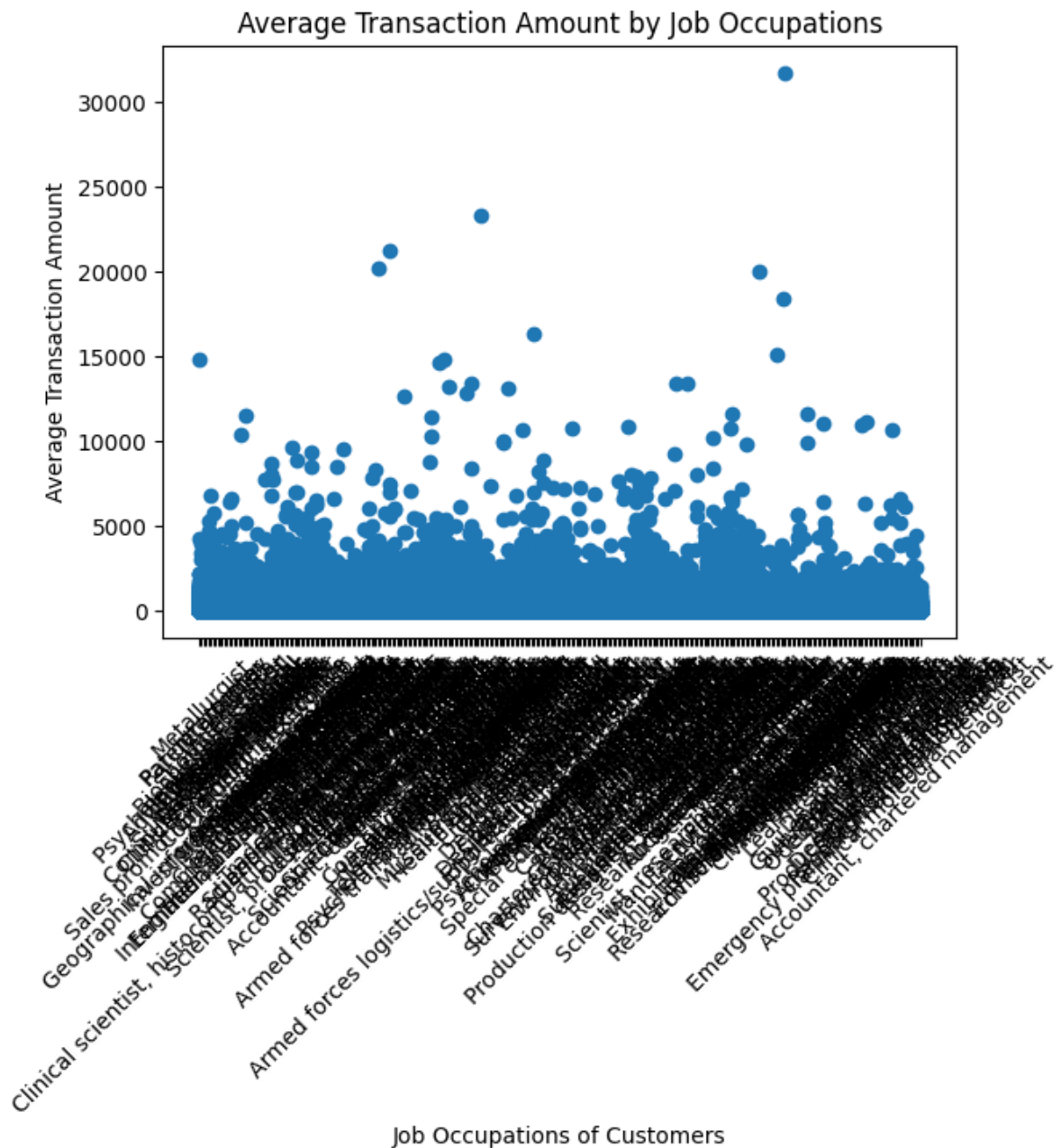
Creating bar chart to illustrate the distribution of ages by gender

```
In [19]: sns.boxplot(data=m_data,x='gender',y='age')  
plt.title("Distribution of Ages by Gender");
```



Creating scatter plot to illustrate the average transaction amounts between various employment occupations

```
In [20]: plt.scatter(data=m_data, x='job', y='amt')
plt.xlabel('Job Occupations of Customers')
plt.ylabel('Average Transaction Amount')
plt.title('Average Transaction Amount by Job Occupations')
plt.xticks(rotation=45)
plt.show()
```



Creating new features for segmenting customers based on their transactions

```
In [21]: frequency_of_purchase = m_data.groupby('acct_num')['trans_num'].count().reset_index()
frequency_of_purchase.rename(columns={'trans_num': 'frequency_of_purchase'}, in
place=True)
m_data = pd.merge(m_data, frequency_of_purchase, on='acct_num', how='left')
```

```
In [22]: m_data['frequency_of_purchase']
```

```
Out[22]: 0          2550
         1          2550
         2          2550
         3          2550
         4          2550
         ...
        1574054      1195
        1574055      1195
        1574056      1195
        1574057      1195
        1574058      1195
        Name: frequency_of_purchase, Length: 1574059, dtype: int64
```

```
In [23]: average_purchase_value = m_data.groupby('acct_num')['amt'].mean().reset_index()
         average_purchase_value.rename(columns={'amt': 'average_purchase_value'}, inplace=True)
         m_data = pd.merge(m_data, average_purchase_value, on='acct_num', how='left')
```

```
In [24]: m_data['average_purchase_value']
```

```
Out[24]: 0          83.110467
         1          83.110467
         2          83.110467
         3          83.110467
         4          83.110467
         ...
        1574054      51.540929
        1574055      51.540929
        1574056      51.540929
        1574057      51.540929
        1574058      51.540929
        Name: average_purchase_value, Length: 1574059, dtype: float64
```

```
In [25]: total_spending = m_data.groupby('acct_num')['amt'].sum().reset_index()
         total_spending.rename(columns={'amt': 'total_spending'}, inplace=True)
         m_data = pd.merge(m_data, total_spending, on='acct_num', how='left')
```

```
In [26]: m_data['total_spending']
```

```
Out[26]: 0          211931.69
         1          211931.69
         2          211931.69
         3          211931.69
         4          211931.69
         ...
        1574054      61591.41
        1574055      61591.41
        1574056      61591.41
        1574057      61591.41
        1574058      61591.41
        Name: total_spending, Length: 1574059, dtype: float64
```

```
In [27]: m_data['unix_time'] = pd.to_datetime(m_data['unix_time'], unit='s')
first_transaction_time = m_data.groupby('acct_num')['unix_time'].min().reset_index()
first_transaction_time['recency_of_purchase'] = (pd.to_datetime('today') - first_transaction_time['unix_time']).dt.days
m_data = pd.merge(m_data, first_transaction_time[['acct_num', 'recency_of_purchase']], on='acct_num', how='left')
```

```
In [28]: m_data['recency_of_purchase']
```

```
Out[28]: 0          1605
1          1605
2          1605
3          1605
4          1605
...
1574054    1601
1574055    1601
1574056    1601
1574057    1601
1574058    1601
Name: recency_of_purchase, Length: 1574059, dtype: int64
```

Using K-means clustering for customer segmentation based on transactional behavior and demographic information

Assigning numerical labels to each unique gender

```
In [29]: from sklearn.preprocessing import LabelEncoder

label_encoder = LabelEncoder()
m_data['gender'] = label_encoder.fit_transform(m_data['gender'])
```

```
In [30]: selected_features = m_data[['frequency_of_purchase', 'average_purchase_value',
'total_spending', 'recency_of_purchase', 'age', 'gender']]
```

```
In [31]: from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()
scaled = scaler.fit_transform(selected_features)
```

Applying Elbow Method to determine the optimum number of clusters to which the data maybe clustered

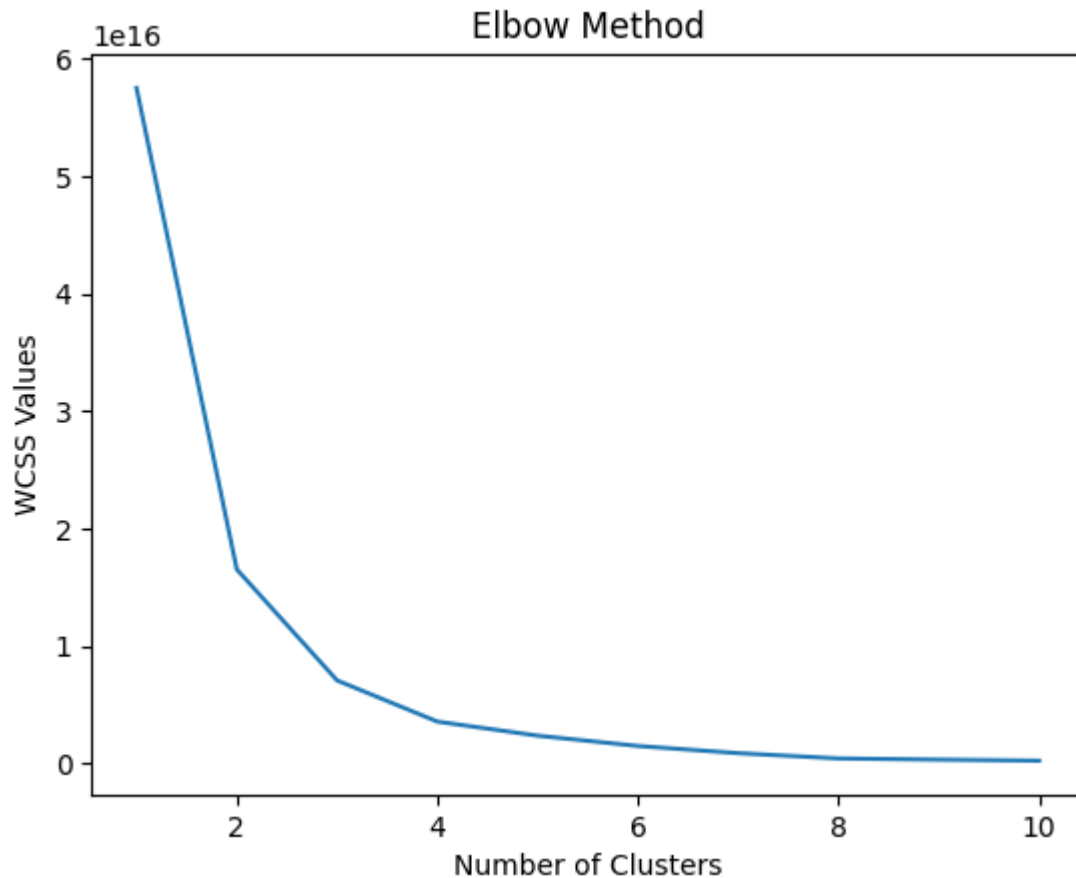
```
In [32]: wcss = [] #within clusters sum of squares

for iter in range(1,11):
    kmeans = KMeans(n_clusters=iter, init='k-means++', random_state = 0)
    kmeans.fit(selected_features)
    wcss.append(kmeans.inertia_)

print("Within Cluster Sum of Squares (WCSS):", wcss)
```

Within Cluster Sum of Squares (WCSS): [5.74906976623925e+16, 1.6541108181070272e+16, 7092002487194571.0, 3590850486201102.0, 2394022680913940.5, 1517826131720211.8, 905518767908902.2, 456518081870780.5, 355596196169811.4, 267636239908384.88]

```
In [33]: plt.plot(range(1,11),wcss)
plt.title("Elbow Method")
plt.xlabel("Number of Clusters")
plt.ylabel("WCSS Values")
plt.show()
```



We can see the plot reduces drastically from cluster number 1-3 and flattens from 4-10. Hence, we are getting an elbow at K=3. Thus, we can assign number of clusters as K=3.

```
In [34]: #Initializing model
kmeansmodel = KMeans(n_clusters = 3, init="k-means++", random_state=0)
```



```
In [35]: y_km = kmeansmodel.fit_predict(selected_features)
```

```
In [36]: X = np.array(selected_features)
```

```
In [69]: c_labels = kmeans.fit_predict(X)

# Calculating the number of customers in each cluster
cluster_counts = {}
for label in c_labels:
    if label in cluster_counts:
        cluster_counts[label] += 1
    else:
        cluster_counts[label] = 1

# Sorting the clusters based on the number of customers
sorted_clusters = sorted(cluster_counts.items(), key=lambda x: x[1], reverse=True)

# Selecting the top 100 customers from the clusters
selected_customers = []
count = 0
for cluster, _ in sorted_clusters:
    cluster_indices = [i for i, label in enumerate(c_labels) if label == cluster]
    cluster_customers = X[cluster_indices]
    selected_customers.extend(cluster_customers)
    count += len(cluster_customers)
    if count >= 100:
        break

# selected_customers now contains the top 100 customers from the clusters
```

```
In [62]: top_customers = np.array(selected_customers)
```

```
In [63]: from sklearn.cluster import KMeans
from sklearn.metrics import silhouette_score

# Performing K-means clustering on the top 100 customers

kmeans = KMeans(n_clusters=3)
cluster_labels_top_c = kmeans.fit_predict(top_customers)

# Calculating the silhouette score
silhouette_avg = silhouette_score(top_customers, cluster_labels_top_c)

print("The silhouette score for the top 100 customers is:", silhouette_avg)
```

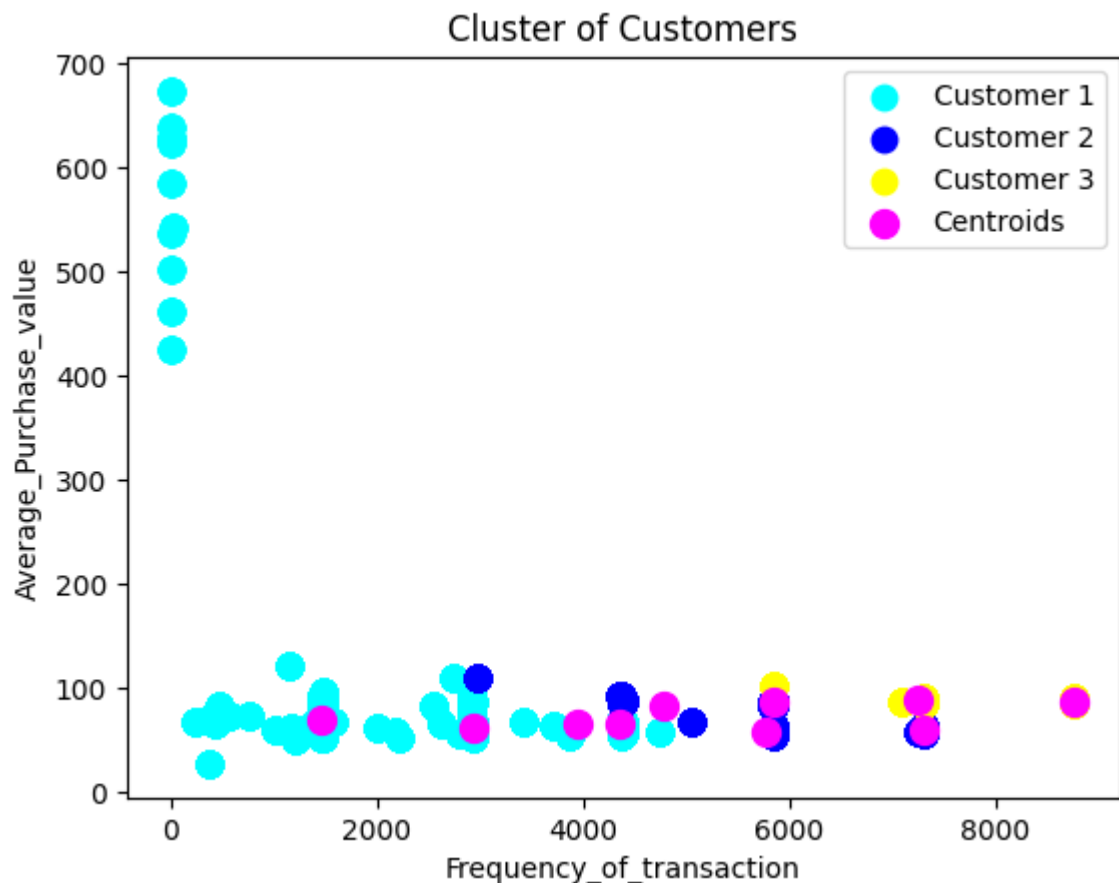
The silhouette score for the top 100 customers is: 0.7447968981996144

```
In [37]: clusters = kmeansmodel.cluster_centers_
clusters
```

```
Out[37]: array([[3.25951528e+03, 6.60043545e+01, 2.12062975e+05, 1.59594192e+03,
 5.12508051e+01, 5.32469141e-01],
 [5.98339039e+03, 7.07676126e+01, 4.13514416e+05, 1.60146643e+03,
 3.72660538e+01, 5.93732546e-01],
 [7.90042941e+03, 8.85418635e+01, 6.98918173e+05, 1.60577185e+03,
 3.70967813e+01, 1.15246701e-12]])
```

```
In [38]: plt.scatter(X[y_km == 0,0], X[y_km == 0,1], s=80, c='cyan', label="Customer
1")
plt.scatter(X[y_km == 1,0], X[y_km == 1,1], s=80, c='blue', label="Customer
2")
plt.scatter(X[y_km == 2,0], X[y_km == 2,1], s=80, c='yellow', label="Customer
3")

plt.scatter(kmeans.cluster_centers_[0, 0], kmeans.cluster_centers_[0, 1], s=100,
c='magenta', label='Centroids')
plt.title("Cluster of Customers")
plt.xlabel("Frequency_of_transaction")
plt.ylabel("Average_Purchase_value")
plt.legend()
plt.show()
```



Exploring the clusters

```
In [39]: from pandas.io.parsers.readers import MANDATORY_DIALECT_ATTRS
cluster_km = kmeansmodel.labels_

m_data['cluster_km'] = cluster_km

#Analyzing the characteristics of each cluster to understand customer segments
for cluster in range(3):
    cluster_data = m_data[m_data['cluster_km'] == cluster]
    print(f"Cluster {cluster}:\n")
    print(cluster_data.describe()) # You can print statistics or perform further analysis for each cluster
```

Cluster 0:

	cc_num_x	acct_num	amt	is_fraud \
count	7.589360e+05	7.589360e+05	758936.000000	758936.000000
mean	3.025036e+17	4.881637e+11	66.004355	0.000895
std	1.124582e+18	2.760079e+11	155.421528	0.029898
min	6.040027e+10	2.348758e+09	1.000000	0.000000
25%	3.866271e+13	2.687429e+11	8.870000	0.000000
50%	2.717544e+15	4.759571e+11	45.030000	0.000000
75%	4.467046e+15	7.173788e+11	80.700000	0.000000
max	4.983666e+18	9.993899e+11	31772.590000	1.000000

	merch_lat	merch_long	cc_num_y	gender \
count	758936.000000	758936.000000	7.589360e+05	758936.000000
mean	37.565845	-92.825955	3.025036e+17	0.532469
std	5.664289	17.343759	1.124582e+18	0.498945
min	18.702948	-160.367675	6.040027e+10	0.000000
25%	33.477743	-102.168127	3.866271e+13	0.000000
50%	38.333749	-87.635717	2.717544e+15	1.000000
75%	41.434933	-80.630192	4.467046e+15	1.000000
max	62.630265	-68.979098	4.983666e+18	1.000000

	zip	lat	long	city_pop \
count	758936.000000	758936.000000	758936.000000	7.589360e+05
mean	52801.402222	37.566728	-92.826518	3.079856e+05
std	29583.262663	5.634324	17.333752	5.999702e+05
min	1571.000000	19.702500	-159.368300	8.670000e+02
25%	28206.000000	33.608200	-102.352300	1.943300e+04
50%	49837.000000	38.547600	-87.747400	6.554000e+04
75%	79761.000000	41.186700	-80.819500	2.197520e+05
max	99645.000000	61.630300	-69.977900	2.906700e+06

	age	frequency_of_purchase	average_purchase_value \
count	758936.000000	758936.000000	758936.000000
mean	51.250805	3259.515279	66.004355
std	16.582167	1097.099351	11.212576
min	25.000000	7.000000	28.229653
25%	38.000000	2920.000000	60.063630
50%	50.000000	2922.000000	64.285356
75%	63.000000	4380.000000	67.167202
max	95.000000	4742.000000	673.189000

	total_spending	recency_of_purchase	cluster_km
count	758936.000000	758936.000000	758936.0
mean	212062.974937	1595.941919	0.0
std	69147.727204	79.009025	0.0
min	3821.360000	180.000000	0.0
25%	167773.190000	1605.000000	0.0
50%	213536.420000	1606.000000	0.0
75%	274215.670000	1606.000000	0.0
max	312080.050000	1606.000000	0.0

Cluster 1:

	cc_num_x	acct_num	amt	is_fraud \
count	5.335500e+05	5.335500e+05	533550.000000	533550.000000
mean	5.163890e+17	4.700902e+11	70.767613	0.000472
std	1.446709e+18	2.915475e+11	160.552236	0.021728

min	6.040616e+10	6.672441e+09	1.000000	0.000000
25%	3.424027e+14	2.112621e+11	9.510000	0.000000
50%	4.063578e+15	4.719310e+11	43.650000	0.000000
75%	4.900533e+15	7.234382e+11	79.660000	0.000000
max	4.956253e+18	9.944193e+11	23347.070000	1.000000

	merch_lat	merch_long	cc_num_y	gender \
count	533550.000000	533550.000000	5.335500e+05	533550.000000
mean	36.675119	-89.151974	5.163890e+17	0.593733
std	5.147380	15.596696	1.446709e+18	0.491136
min	18.702502	-156.093612	6.040616e+10	0.000000
25%	33.276127	-97.218231	3.424027e+14	0.000000
50%	37.577634	-83.953243	4.063578e+15	1.000000
75%	40.804991	-77.481957	4.900533e+15	1.000000
max	46.152564	-69.937212	4.956253e+18	1.000000

	zip	lat	long	city_pop \
count	533550.000000	533550.000000	533550.000000	5.335500e+05
mean	45828.831242	36.674721	-89.14975	3.086762e+05
std	28915.162574	5.114385	15.58646	5.959204e+05
min	1581.000000	19.702500	-155.09390	2.693000e+03
25%	22015.000000	33.065000	-97.21270	1.855700e+04
50%	40031.000000	37.287600	-84.14330	5.565900e+04
75%	77055.000000	40.685000	-77.29500	2.156190e+05
max	96720.000000	45.152800	-70.93720	2.906700e+06

	age	frequency_of_purchase	average_purchase_value \
count	533550.000000	533550.000000	533550.000000
mean	37.266054	5983.390393	70.767613
std	7.480467	1004.514227	14.641348
min	25.000000	2968.000000	54.907505
25%	31.000000	5840.000000	58.170890
50%	38.000000	5840.000000	61.039904
75%	44.000000	7299.000000	87.645780
max	50.000000	7304.000000	110.957210

	total_spending	recency_of_purchase	cluster_km
count	533550.000000	533550.000000	533550.0
mean	413514.416009	1601.466429	1.0
std	64376.765882	44.653468	0.0
min	320659.830000	1145.000000	1.0
25%	356534.080000	1606.000000	1.0
50%	407523.200000	1606.000000	1.0
75%	444865.400000	1606.000000	1.0
max	525216.840000	1606.000000	1.0

Cluster 2:

	cc_num_x	acct_num	amt	is_fraud \
count	2.815730e+05	2.815730e+05	281573.000000	281573.000000
mean	1.133942e+17	4.575788e+11	88.541863	0.000433
std	6.785760e+17	2.913534e+11	144.212713	0.020811
min	6.043496e+10	3.614025e+10	1.000000	0.000000
25%	2.131463e+14	2.291333e+11	18.700000	0.000000
50%	3.512590e+15	4.012396e+11	53.900000	0.000000
75%	4.498392e+15	7.017294e+11	111.920000	0.000000
max	4.272742e+18	9.763247e+11	14805.230000	1.000000

	merch_lat	merch_long	cc_num_y	gender	zip \
count	281573.000000	281573.000000	2.815730e+05	281573.0	281573.000000
mean	38.401446	-90.768206	1.133942e+17	0.0	50300.554435
std	4.496300	15.532411	6.785760e+17	0.0	28929.602618
min	25.231542	-123.312667	6.043496e+10	0.0	6443.000000
25%	34.817487	-97.243388	2.131463e+14	0.0	26726.000000
50%	39.685288	-86.922757	3.512590e+15	0.0	48328.000000
75%	41.646220	-79.389040	4.498392e+15	0.0	75021.000000
max	48.392670	-71.616005	4.272742e+18	0.0	98198.000000

	lat	long	city_pop	age \
count	281573.000000	281573.000000	2.815730e+05	281573.000000
mean	38.400821	-90.767259	3.838910e+05	37.096781
std	4.457535	15.520864	7.213736e+05	7.302174
min	26.231500	-122.312900	2.682000e+03	25.000000
25%	34.750900	-97.641900	1.826900e+04	29.000000
50%	39.766500	-87.167800	7.152800e+04	39.000000
75%	41.721200	-79.328300	1.939330e+05	44.000000
max	47.392900	-72.615300	2.680484e+06	48.000000

	frequency_of_purchase	average_purchase_value	total_spending \
count	281573.000000	281573.000000	281573.000000
mean	7900.429413	88.541863	698918.172869
std	783.267280	2.505475	66638.402400
min	5844.000000	85.254465	599912.360000
25%	7300.000000	87.191679	639903.840000
50%	7303.000000	88.377894	653459.920000
75%	8760.000000	89.142116	770719.250000
max	8765.000000	102.654408	805301.640000

	recency_of_purchase	cluster_km
count	281573.000000	281573.0
mean	1605.771850	2.0
std	0.477441	0.0
min	1604.000000	2.0
25%	1606.000000	2.0
50%	1606.000000	2.0
75%	1606.000000	2.0
max	1606.000000	2.0

```
In [40]: m_data['Cluster'] = y_km

# Analyzing the differences between clusters
cluster_analysis = m_data.groupby('Cluster').mean()

# Displaying the cluster analysis results
print(cluster_analysis)
```

	cc_num_x	acct_num	amt	is_fraud	merch_lat	\
Cluster						
0	3.025036e+17	4.881637e+11	66.004355	0.000895	37.565845	
1	5.163890e+17	4.700902e+11	70.767613	0.000472	36.675119	
2	1.133942e+17	4.575788e+11	88.541863	0.000433	38.401446	

	merch_long	cc_num_y	gender	zip	lat	\
Cluster						
0	-92.825955	3.025036e+17	0.532469	52801.402222	37.566728	
1	-89.151974	5.163890e+17	0.593733	45828.831242	36.674721	
2	-90.768206	1.133942e+17	0.000000	50300.554435	38.400821	

	long	city_pop	age	frequency_of_purchase	\
Cluster					
0	-92.826518	307985.594590	51.250805	3259.515279	
1	-89.149750	308676.164082	37.266054	5983.390393	
2	-90.767259	383891.015424	37.096781	7900.429413	

	average_purchase_value	total_spending	recency_of_purchase	\
Cluster				
0	66.004355	212062.974937	1595.941919	
1	70.767613	413514.416009	1601.466429	
2	88.541863	698918.172869	1605.771850	

	cluster_km
Cluster	
0	0.0
1	1.0
2	2.0

Applying BIRCH Algorithm (Balanced Iterative Reducing and Clustering using Hierarchies)

```
In [41]: from sklearn.cluster import Birch

selected_features = m_data[['frequency_of_purchase', 'average_purchase_value',
'total_spending', 'recency_of_purchase', 'age', 'gender']]
scaler = StandardScaler()
scaled = scaler.fit_transform(selected_features)

birch = Birch(n_clusters=3)
birch.fit(scaled)

y_pred_b = birch.predict(scaled)
```



```
In [42]: m_data['cluster_birch'] = y_pred_b

# Analyzing the clusters
cluster_centers = birch.subcluster_centers_
for i, center in enumerate(cluster_centers):
    print(f"Cluster {i+1} Center:")
    print(center)
    print()

# Printing the count of customers in each cluster
print(m_data['cluster_birch'].value_counts())
```

Cluster 1 Center:

[-0.99681616 -0.85379866 -1.0072781 0.0939972 0.13523105 1.08787676]

Cluster 2 Center:

[0.41629828 -0.95427045 -0.13515728 0.10432395 -0.94693431 1.08787676]

Cluster 3 Center:

[-1.72689899 -0.76955742 -1.47357415 0.06994274 0.15283026 1.08787676]

Cluster 4 Center:

[-1.10815027 -0.97358007 -1.10224301 0.09152295 -0.61462293 1.08787676]

Cluster 5 Center:

[1.09804249 -0.89047505 0.33087614 0.10351446 -0.69858832 1.08787676]

Cluster 6 Center:

[0.58633928 -0.86596141 0.01582774 0.10406708 0.01703867 1.08787676]

Cluster 7 Center:

[-0.28841503 -0.89978162 -0.56264372 0.09738153 -0.05890766 1.08787676]

Cluster 8 Center:

[-0.30400666 -0.88422938 -0.56767322 0.1046307 -0.95954564 1.08787676]

Cluster 9 Center:

[-1.66562382 -0.6088624 -1.41419641 0.06702248 1.12231189 1.08787676]

Cluster 10 Center:

[-0.90913563 -0.48614252 -0.86373815 0.09398024 0.90833 1.08787676]

Cluster 11 Center:

[-0.30390887 -0.42099969 -0.41701498 0.09939203 0.54322698 1.08787676]

Cluster 12 Center:

[-1.70769154 -0.28532722 -1.40571232 0.08921843 2.85102858 0.96012432]

Cluster 13 Center:

[-0.4725218 -0.29516797 -0.49717995 0.09591307 1.82907314 1.08787676]

Cluster 14 Center:

[-1.0075295 -0.39976979 -0.91342094 0.09314599 1.11200808 -0.91922177]

Cluster 15 Center:

[-1.7072199 -0.37209806 -1.41476031 0.07059816 2.49645638 -0.91922177]

Cluster 16 Center:

[-0.32680126 -0.4516921 -0.44408557 0.09886392 1.82163098 -0.91922177]

Cluster 17 Center:

[-1.70725425 -0.4290683 -1.42097382 0.07768916 1.46998534 -0.91922177]

Cluster 18 Center:

[-1.73538244 0.38793903 -1.35326894 0.0810182 1.14359543 -0.91922177]

Cluster 19 Center:

[-0.30400188 -0.34742046 -0.39317228 0.09315788 3.08217303 -0.91922177]

Cluster 20 Center:

[-0.30453746 -0.44854733 -0.4264544 0.09978588 0.70577447 -0.91922177]

Cluster 21 Center:

[-0.94816671 -0.37625266 -0.86887742 0.09663669 2.49273778 -0.91922177]

Cluster 22 Center:

[-1.68300073 -0.75683541 -1.44024383 0.08352051 -1.06626431 1.08787676]

Cluster 23 Center:

[-1.71764147 -0.56552568 -1.44238969 0.08291556 0.59112577 -0.91922177]

Cluster 24 Center:

[-1.70841353 -1.26534836 -1.51242122 0.07770737 -0.30859761 1.08787676]

Cluster 25 Center:

[-2.24262417 -0.1373848 -1.79545031 0.10565799 0.13938506 1.08787676]

Cluster 26 Center:

[-1.70841353 -0.51770092 -1.43142439 0.10565799 -0.41223962 1.08787676]

Cluster 27 Center:

[-1.70841353 -0.23231502 -1.40050694 0.07277888 0.27729123 1.08787676]

Cluster 28 Center:

[-1.00639055 -0.34722434 -0.90316637 0.10565799 1.72530601 -0.91922177]

Cluster 29 Center:

[-1.00639055 -0.3512866 -0.90404654 0.09743962 3.44912133 1.08787676]

Cluster 30 Center:

[-1.15304604 -0.45158236 -1.02991303 0.10565799 1.86321218 1.08787676]

Cluster 31 Center:

[-1.83583551 -1.41815042 -1.60044706 0.0234602 -1.10177047 -0.91922177]

Cluster 32 Center:

[-1.34330935 0.96477798 -0.93431886 0.08375977 -1.17817244 -0.91922177]

Cluster 33 Center:

[-0.90853901 1.29628181 -0.45332252 0.10130498 -0.41583655 -0.91922177]

Cluster 34 Center:

[1.80177562 1.16554918 2.11937473 0.10059813 -0.16299117 -0.91922177]

Cluster 35 Center:

[-0.29962274 1.26617195 0.13556591 0.10062407 -0.93060415 -0.91922177]

Cluster 36 Center:

[-1.00391464 1.332006 -0.53659403 0.09797306 0.02422073 -0.91922177]

Cluster 37 Center:

[0.71382894 1.15651067 1.06830381 0.10565799 -0.153788 -0.91922177]

Cluster 38 Center:

[0.39673885 1.19040847 0.7819017 0.1005385 -0.96459676 -0.91922177]

Cluster 39 Center:

[-1.00572722 1.1335025 -0.58170454 0.085095 -0.96383185 -0.91922177]

Cluster 40 Center:

[1.1669652 1.20630207 1.53305346 0.10173627 -1.06601744 -0.91922177]

Cluster 41 Center:

[-1.70680553 1.32332186 -1.23033525 0.07412243 -0.42616452 -0.91922177]

Cluster 42 Center:

[1.08823497 1.13207903 1.41670689 0.10382446 -0.15877841 -0.91922177]

Cluster 43 Center:

[-1.70656481 1.08797167 -1.25589801 0.08008831 0.03211131 -0.91771814]

Cluster 44 Center:

[0.39957876 2.18642899 1.21650033 0.10565799 -0.89491122 -0.91922177]

Cluster 45 Center:

[0.39820537 1.15958022 0.76990179 0.10565799 0.04087969 -0.91922177]

Cluster 46 Center:

[-0.99574739 2.77095198 -0.21468172 0.10565799 0.17168918 -0.91922177]

Cluster 47 Center:

[-1.00639055 0.91690395 -0.6292662 0.09739719 0.31159571 -0.91922177]

Cluster 48 Center:

[-1.73050831 0.86631579 -1.29731529 0.08134587 -0.97934112 -0.91922177]

Cluster 49 Center:

[-1.21591531 2.88154204 -0.47191919 0.10050446 -1.23967664 -0.91922177]

Cluster 50 Center:

[-0.30557692 1.14027712 0.08882641 0.09967479 0.08935713 -0.91922177]

Cluster 51 Center:

[-1.85939655 3.52000037 -1.19374194 0.07277888 -1.23967664 -0.91922177]

Cluster 52 Center:

[-1.83583551 -1.41815042 -1.60044706 0.0234602 -1.10177047 -0.91922177]

Cluster 53 Center:

[-1.3368831 -1.34493729 -1.30817362 -7.5022142 0.35195984 1.07986972]

Cluster 54 Center:

[-1.9392025 -0.73567267 -1.6105045 -14.72282403 1.57562584
-0.91922177]

Cluster 55 Center:

[-2.1049628 0.45604477 -1.66589072 -14.72972894 2.81752535
-0.91922177]

Cluster 56 Center:

[-0.55584567 -1.14858941 -0.80534313 -7.48941816 0.4151974 1.08787676]

Cluster 57 Center:

```
[-1.6980084 -0.62372771 -1.4311146 -7.48870288 0.24128962 -0.91922177]
```

Cluster 58 Center:

```
[-2.23012239 -3.0620875 -1.86734958 -7.52229728 -1.17072356 1.08787676]
```

Cluster 59 Center:

```
[-1.65744474 -0.22139699 -1.36132234 -14.72282403 0.48415048  
1.08787676]
```

Cluster 60 Center:

```
[ 0.02115952 -0.23557421 -0.11509261 -7.47297861 -0.55014579 -0.91922177]
```

Cluster 61 Center:

```
[-2.40707065 40.01334065 -1.89933688 -11.50067051 0.75996282  
-0.91922177]
```

Cluster 62 Center:

```
[-2.40610897 32.72371877 -1.89751701 -1.47253965 2.82855538 -0.91922177]
```

Cluster 63 Center:

```
[-2.40658981 39.32112556 -1.89640377 -8.32783566 1.3115875 -0.91922177]
```

Cluster 64 Center:

```
[-2.40466646 33.18802194 -1.88869423 -3.14937465 1.58739984 -0.91922177]
```

Cluster 65 Center:

```
[-2.40610897 36.12267968 -1.8952471 -23.27139459 0.89786899  
-0.91922177]
```

Cluster 66 Center:

```
[-2.40658981 30.3605367 -1.90172295 -0.23957274 2.34588378 -0.91922177]
```

Cluster 67 Center:

```
[-2.40562813 27.46955895 -1.89861261 -2.86990215 2.20797761 -0.91922177]
```

Cluster 68 Center:

```
[-2.40610897 38.83970143 -1.89343262 -21.34796622 1.44949367  
-0.91922177]
```

Cluster 69 Center:

```
[-2.40562813 42.42116314 -1.88751815 -23.33715283 0.69100974  
-0.91922177]
```

Cluster 70 Center:

```
[-2.40610897 24.89007688 -1.90274849 -17.09012049 0.4151974  
1.08787676]
```

2 1557359

1 16609

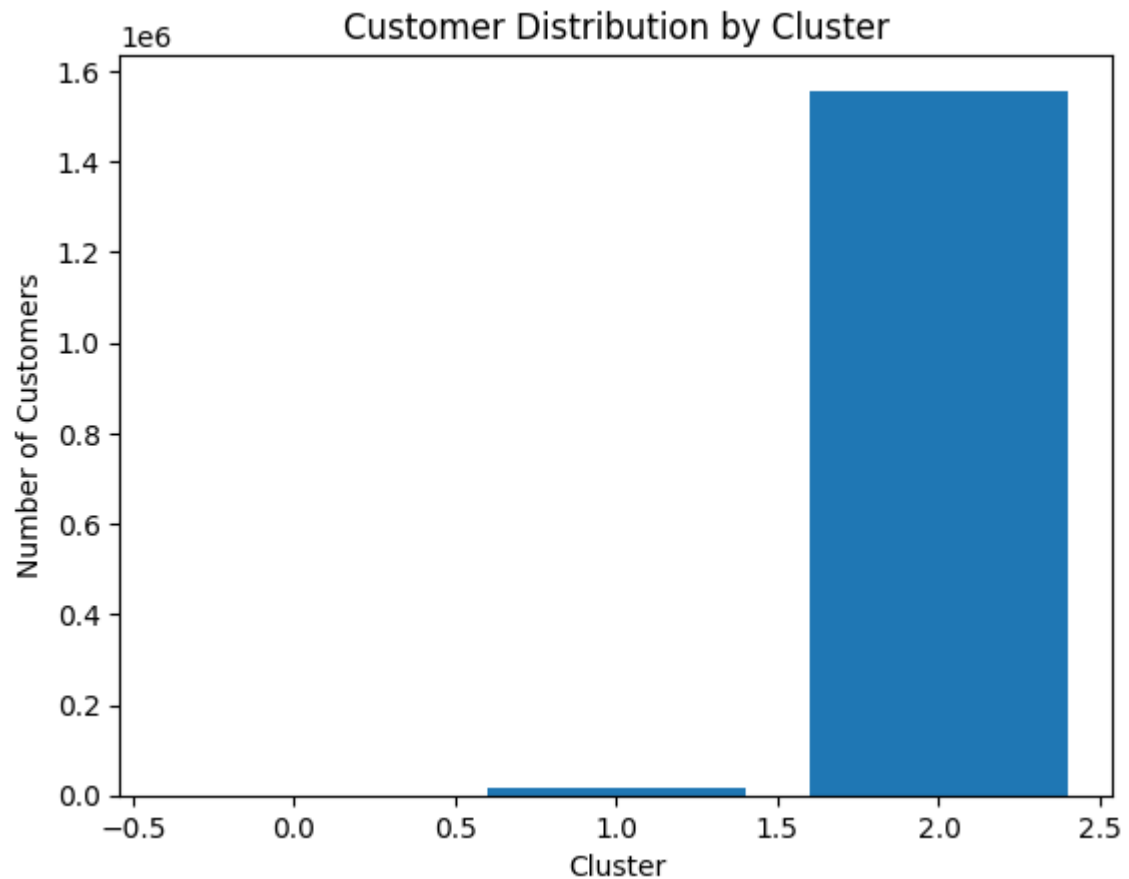
0 91

Name: cluster_birch, dtype: int64

```
In [43]: cluster_counts = m_data['cluster_birch'].value_counts()
# Creating bar plot
plt.bar(cluster_counts.index, cluster_counts.values)

# Setting plot labels and title
plt.xlabel('Cluster')
plt.ylabel('Number of Customers')
plt.title('Customer Distribution by Cluster')

# Showing the plot
plt.show()
```

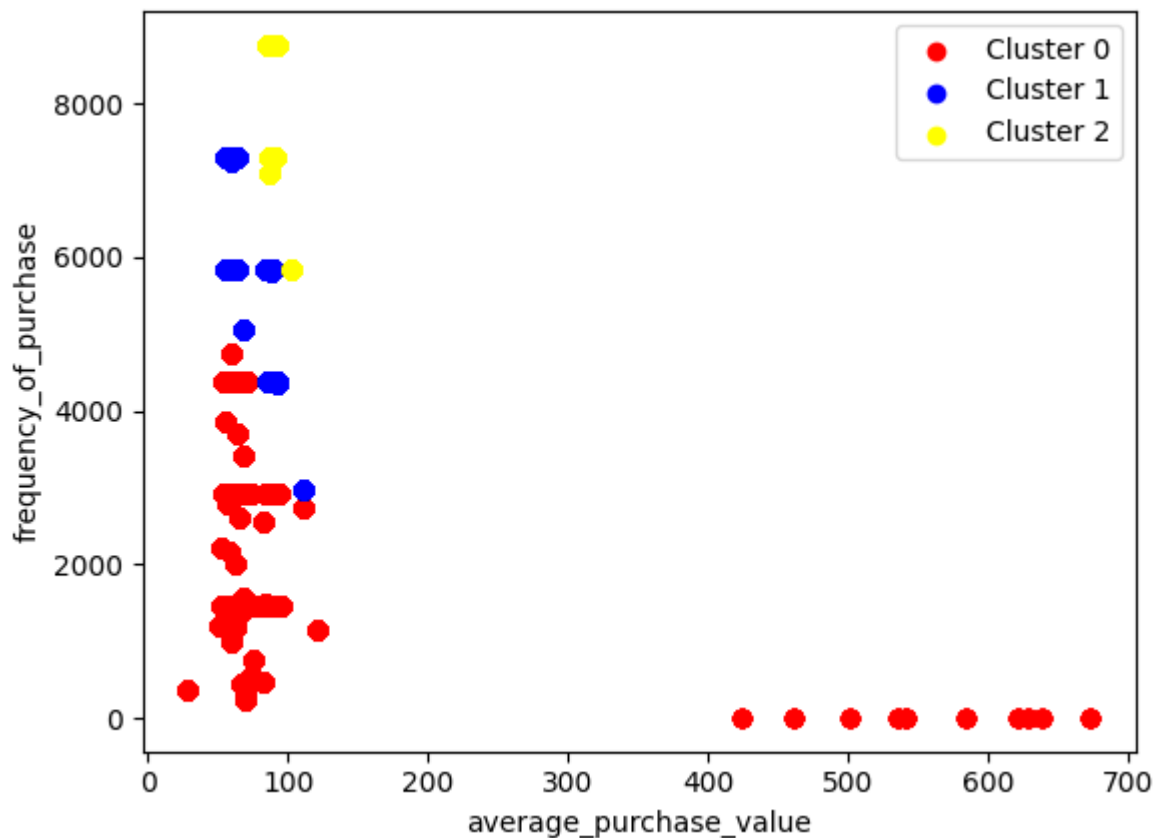


```
In [44]: # Create scatter plot
feature1 = 'average_purchase_value'
feature2 = 'frequency_of_purchase'
feature3 = 'total_spending'
plt.scatter(m_data[m_data['Cluster'] == 0][feature1], m_data[m_data['Cluster']
== 0][feature2], color='red', label='Cluster 0')
plt.scatter(m_data[m_data['Cluster'] == 1][feature1], m_data[m_data['Cluster']
== 1][feature2], color='blue', label='Cluster 1')
plt.scatter(m_data[m_data['Cluster'] == 2][feature1], m_data[m_data['Cluster']
== 2][feature2], color='yellow', label='Cluster 2')

# Adding scatter plots for other clusters

# Setting plot labels and legend
plt.xlabel(feature1)
plt.ylabel(feature2)
plt.legend()

# Showing the plot
plt.show()
```



In [45]: m_data

Out[45]:

	cc_num_x	acct_num	trans_num	unix_time	ca
0	676305278100	930558644123	34821a75f2a6ef420bfe6b775ef7ebb6	2019-05-23 17:43:52	groce
1	676305278100	930558644123	47504e5f0f060d0ba136241a7381d78c	2019-10-29 23:07:47	groce
2	676305278100	930558644123	330ac8785c386bcfedccda9450c6fcf1	2019-01-14 22:49:36	gas_tr
3	676305278100	930558644123	bf1b90eb153d61172bdaec4698e71b90	2019-03-30 21:39:16	groce
4	676305278100	930558644123	663a9e929cebc05428ec8a4293b0c6	2019-01-16 19:29:01	gas_tr
...
1574054	6584732365337749	843643900137	7b2b63508d92d5dd61446fb5cb03e301	2022-11-01 06:06:26	person
1574055	6584732365337749	843643900137	1df5a10b715d536a666143ebb429a818	2020-06-03 11:44:29	mi
1574056	6584732365337749	843643900137	67f8305d28b564b34ccf264b9f4ea764	2021-06-30 09:32:37	food
1574057	6584732365337749	843643900137	4d3333cd26ba3ac993151b30b474e94e	2022-11-03 08:57:44	person
1574058	6584732365337749	843643900137	e9126d852f94ab72fa6b40ee7e44c6f1	2021-05-10 08:49:47	enterte

1574059 rows × 32 columns



Performing clustering using Gaussian Mixture Model (GMM)


```
In [46]: from sklearn.mixture import GaussianMixture

# Performing Gaussian Mixture Model clustering
gmm = GaussianMixture(n_components=3) # Specifying the number of desired clusters
gmm.fit(scaled)

# Getting the predicted cluster labels
cluster_gm = gmm.predict(scaled)

# Adding the cluster labels to your original dataset
m_data['Cluster_gm'] = cluster_gm

# Analyzing the results

cluster_ct = m_data['Cluster_gm'].value_counts()
print(cluster_ct)
```

```
1    840431
0    711977
2     21651
Name: Cluster_gm, dtype: int64
```

```
In [ ]: new_data = m_data[['frequency_of_purchase', 'average_purchase_value', 'total_s
pending', 'recency_of_purchase', 'age', 'gender', 'Cluster_gm']]
```

```
In [48]: # Grouping the data by cluster labels
grouped = new_data.groupby('Cluster_gm')

# Iterating over each feature and create a box plot for each cluster
for feature in new_data.columns[:-1]: # Exclude the last column (cluster labels)
    plt.figure()
    plt.title(f"Box Plot for {feature}")

    # Creating a list of data for each cluster
    df = [grouped.get_group(Cluster_gm)[feature] for Cluster_gm in grouped.groups]

    # Creating the box plot
    plt.boxplot(df, labels=grouped.groups.keys())
    plt.xlabel('Cluster Label')
    plt.ylabel('Feature Value')

    # Showing the plot
    plt.show()
```

