

Loan_data

June 8, 2024

```
[49]: from pymongo import MongoClient
import pandas as pd
```

```
[50]: client = MongoClient('mongodb://localhost:27017/')
```

```
[51]: db = client.loan_db
collection = db.loan_data
```

We perform some simple task, beginwith finding people that has exeptional high fico scores

```
[52]: query = {"fico": {"$gt": 800}}
high_fico_docs = collection.find(query)
df_high_fico = pd.DataFrame(list(high_fico_docs))
print(df_high_fico)
```

	_id	int	installment	\
0	6664f0cda92ab3700544585a	{'rate': 0.0743}	155.38	
1	6664f0cda92ab3700544585d	{'rate': 0.0743}	155.38	
2	6664f0cda92ab37005445886	{'rate': 0.0743}	155.38	
3	6664f0cda92ab370054458b1	{'rate': 0.0712}	64.96	
4	6664f0cda92ab370054458da	{'rate': 0.0901}	795.11	
..	
140	6664f0cda92ab370054476bb	{'rate': 0.0743}	108.77	
141	6664f0cda92ab37005447cbb	{'rate': 0.0859}	31.62	
142	6664f0cda92ab37005447ceb	{'rate': 0.0859}	379.33	
143	6664f0cda92ab37005447cf1	{'rate': 0.1913}	918.02	
144	6664f0cda92ab37005447d07	{'rate': 0.1148}	725.31	

	log	dti	fico	\
0	{'annual': {'inc': 11.08214255}}	0.28	802	
1	{'annual': {'inc': 12.20607265}}	3.72	812	
2	{'annual': {'inc': 11.15625052}}	4.56	812	
3	{'annual': {'inc': 11.60823564}}	0.22	812	
4	{'annual': {'inc': 10.95080655}}	8.93	822	
..	
140	{'annual': {'inc': 12.61153775}}	1.00	802	
141	{'annual': {'inc': 11.69524702}}	12.05	807	
142	{'annual': {'inc': 12.01370075}}	10.21	807	
143	{'annual': {'inc': 11.14908203}}	17.56	802	

```

144 {'annual': {'inc': 11.47210347}} 19.98 807

                                days                                revol \
0   {'with': {'cr': {'line': 4649.958333}}} {'bal': 1576, 'util': 5.7}
1   {'with': {'cr': {'line': 6778.958333}}} {'bal': 85607, 'util': 0.7}
2       {'with': {'cr': {'line': 9809}}} {'bal': 813, 'util': 1}
3       {'with': {'cr': {'line': 16213}}} {'bal': 50, 'util': 0.3}
4   {'with': {'cr': {'line': 9696.958333}}} {'bal': 185, 'util': 0.5}
..
140 {'with': {'cr': {'line': 8593.958333}}} {'bal': 150786, 'util': 17}
141 {'with': {'cr': {'line': 6149.958333}}} {'bal': 269726, 'util': 4.3}
142 {'with': {'cr': {'line': 8191.041667}}} {'bal': 198058, 'util': 5.1}
143       {'with': {'cr': {'line': 12344}}} {'bal': 190050, 'util': 8}
144 {'with': {'cr': {'line': 8580.041667}}} {'bal': 158000, 'util': 0}

```

```

                                not
0   {'fully': {'paid': 0}}
1   {'fully': {'paid': 0}}
2   {'fully': {'paid': 0}}
3   {'fully': {'paid': 0}}
4   {'fully': {'paid': 0}}
..
140 {'fully': {'paid': 0}}
141 {'fully': {'paid': 0}}
142 {'fully': {'paid': 0}}
143 {'fully': {'paid': 1}}
144 {'fully': {'paid': 0}}

```

[145 rows x 9 columns]

We found 145 records

We query with additional condition, will lower the bar to fico score with 750 and interest rate that below 0.06

```

[53]: query = {"fico": {"$gt": 750 }, "int.rate": {"$lt": 0.07}}
      high_fico_low_rate_docs = collection.find(query)

      df_high_fico_low_rate = pd.DataFrame(list(high_fico_low_rate_docs))
      print(df_high_fico_low_rate)

```

```

                                _id                                int  installment \
0   6664f0cda92ab37005447305 {'rate': 0.0639}                76.51
1   6664f0cda92ab370054473dd {'rate': 0.0676}                49.24
2   6664f0cda92ab370054473e2 {'rate': 0.0676}                76.93
3   6664f0cda92ab37005447507 {'rate': 0.0676}               147.70
4   6664f0cda92ab3700544752f {'rate': 0.0676}               103.08
5   6664f0cda92ab3700544755a {'rate': 0.0676}               193.85
6   6664f0cda92ab370054475a8 {'rate': 0.0676}               129.24

```

7	6664f0cda92ab370054475ba	{'rate': 0.0676}	107.70
8	6664f0cda92ab370054475bb	{'rate': 0.0639}	107.11
9	6664f0cda92ab370054475c6	{'rate': 0.0676}	30.77
10	6664f0cda92ab370054475cd	{'rate': 0.0639}	45.91
11	6664f0cda92ab37005447618	{'rate': 0.0639}	100.99
12	6664f0cda92ab3700544763d	{'rate': 0.0676}	147.70

		log	dti	fico \
0	{'annual': {'inc': 10.42228135}}	3.39	782	
1	{'annual': {'inc': 11.33260191}}	3.65	782	
2	{'annual': {'inc': 10.28192399}}	2.38	752	
3	{'annual': {'inc': 11.48246626}}	13.40	787	
4	{'annual': {'inc': 10.87804719}}	23.55	757	
5	{'annual': {'inc': 11.65268741}}	6.25	777	
6	{'annual': {'inc': 10.59663473}}	19.83	777	
7	{'annual': {'inc': 10.858999}}	14.54	777	
8	{'annual': {'inc': 11.22390917}}	14.18	782	
9	{'annual': {'inc': 11.00209984}}	10.18	752	
10	{'annual': {'inc': 11.2515607}}	2.38	772	
11	{'annual': {'inc': 9.655026193}}	21.46	777	
12	{'annual': {'inc': 9.552226498}}	6.05	772	

		days	revol \
0	{'with': {'cr': {'line': 6279}}}	{'bal': 3375, 'util': 7.8}	
1	{'with': {'cr': {'line': 3333}}}	{'bal': 6955, 'util': 14.6}	
2	{'with': {'cr': {'line': 5610.041667}}}	{'bal': 1988, 'util': 13.3}	
3	{'with': {'cr': {'line': 5490}}}	{'bal': 80723, 'util': 27.1}	
4	{'with': {'cr': {'line': 6330.041667}}}	{'bal': 18350, 'util': 29.9}	
5	{'with': {'cr': {'line': 5160}}}	{'bal': 12345, 'util': 17.4}	
6	{'with': {'cr': {'line': 11370}}}	{'bal': 12308, 'util': 22.8}	
7	{'with': {'cr': {'line': 3390}}}	{'bal': 1498, 'util': 2.7}	
8	{'with': {'cr': {'line': 7171.041667}}}	{'bal': 5256, 'util': 15.6}	
9	{'with': {'cr': {'line': 10020.04167}}}	{'bal': 17028, 'util': 11.7}	
10	{'with': {'cr': {'line': 3510.041667}}}	{'bal': 3947, 'util': 7.6}	
11	{'with': {'cr': {'line': 7712}}}	{'bal': 3071, 'util': 16.5}	
12	{'with': {'cr': {'line': 3480.041667}}}	{'bal': 2621, 'util': 17}	

	not
0	{'fully': {'paid': 0}}
1	{'fully': {'paid': 0}}
2	{'fully': {'paid': 0}}
3	{'fully': {'paid': 0}}
4	{'fully': {'paid': 0}}
5	{'fully': {'paid': 0}}
6	{'fully': {'paid': 0}}
7	{'fully': {'paid': 0}}
8	{'fully': {'paid': 0}}
9	{'fully': {'paid': 0}}

```

10 {'fully': {'paid': 0}}
11 {'fully': {'paid': 0}}
12 {'fully': {'paid': 0}}

```

As the result we only found 13 records that meeting the condition

Next I can perform range query on same field. We can do range between fic score 500 to 700. The low to moderate fico profiles

```

[54]: query = {"fico": {"$gt": 500, "$lt": 700}}
      fico_range_docs = collection.find(query)

      df_fico_range = pd.DataFrame(list(fico_range_docs))
      print(df_fico_range)

```

	_id	int	installment	\
0	6664f0cda92ab37005445842	{'rate': 0.1357}	366.86	
1	6664f0cda92ab37005445844	{'rate': 0.1426}	102.92	
2	6664f0cda92ab37005445846	{'rate': 0.1496}	194.02	
3	6664f0cda92ab37005445848	{'rate': 0.1134}	87.19	
4	6664f0cda92ab3700544584a	{'rate': 0.1347}	360.43	
...	
4216	6664f0cda92ab37005447da3	{'rate': 0.1568}	161.01	
4217	6664f0cda92ab37005447da4	{'rate': 0.1565}	69.98	
4218	6664f0cda92ab37005447da5	{'rate': 0.1461}	344.76	
4219	6664f0cda92ab37005447da7	{'rate': 0.1071}	97.81	
4220	6664f0cda92ab37005447da8	{'rate': 0.16}	351.58	

	log	dti	fico	\
0	{'annual': {'inc': 10.37349118}}	11.63	682	
1	{'annual': {'inc': 11.29973224}}	14.97	667	
2	{'annual': {'inc': 10.71441777}}	4.00	667	
3	{'annual': {'inc': 11.40756495}}	17.25	682	
4	{'annual': {'inc': 10.4341158}}	22.09	677	
...	
4216	{'annual': {'inc': 11.22524339}}	8.00	677	
4217	{'annual': {'inc': 10.11047245}}	7.02	662	
4218	{'annual': {'inc': 12.18075484}}	10.39	672	
4219	{'annual': {'inc': 10.59663473}}	13.09	687	
4220	{'annual': {'inc': 10.81977828}}	19.18	692	

	days	revol	\
0	{'with': {'cr': {'line': 4710}}}	{'bal': 3511, 'util': 25.6}	
1	{'with': {'cr': {'line': 4066}}}	{'bal': 4740, 'util': 39.5}	
2	{'with': {'cr': {'line': 3180.041667}}}	{'bal': 3839, 'util': 76.8}	
3	{'with': {'cr': {'line': 3989}}}	{'bal': 69909, 'util': 51.1}	
4	{'with': {'cr': {'line': 6713.041667}}}	{'bal': 13846, 'util': 71}	
...	
4216	{'with': {'cr': {'line': 7230}}}	{'bal': 6909, 'util': 29.2}	

```

4217 {'with': {'cr': {'line': 8190.041667}}} {'bal': 2999, 'util': 39.5}
4218 {'with': {'cr': {'line': 10474}}} {'bal': 215372, 'util': 82.1}
4219 {'with': {'cr': {'line': 3450.041667}}} {'bal': 10036, 'util': 82.9}
4220 {'with': {'cr': {'line': 1800}}} {'bal': 0, 'util': 3.2}

```

```

not
0 {'fully': {'paid': 0}}
1 {'fully': {'paid': 0}}
2 {'fully': {'paid': 1}}
3 {'fully': {'paid': 0}}
4 {'fully': {'paid': 0}}
...
4216 {'fully': {'paid': 1}}
4217 {'fully': {'paid': 1}}
4218 {'fully': {'paid': 1}}
4219 {'fully': {'paid': 1}}
4220 {'fully': {'paid': 1}}

```

[4221 rows x 9 columns]

There is 4222 record fall in this category. Next will add another condition interest rate that is less than 10%

```

[55]: query = {"fico": {"$gt": 500, "$lt": 700 }, "int.rate": {"$lt": 0.10}}
mid_fico_low_rate_docs = collection.find(query)

df_mid_fico_low_rate_docs = pd.DataFrame(list(mid_fico_low_rate_docs))
print(df_mid_fico_low_rate_docs)

```

	_id	int	installment \
0	6664f0cda92ab370054458cb	{'rate': 0.0996}	38.70
1	6664f0cda92ab370054458cf	{'rate': 0.0996}	32.25
2	6664f0cda92ab370054458ed	{'rate': 0.0996}	96.75
3	6664f0cda92ab37005445919	{'rate': 0.0996}	290.24
4	6664f0cda92ab37005445924	{'rate': 0.0964}	154.08
5	6664f0cda92ab3700544592a	{'rate': 0.0933}	95.87
6	6664f0cda92ab3700544592b	{'rate': 0.0996}	95.94
7	6664f0cda92ab37005445948	{'rate': 0.0964}	192.60
8	6664f0cda92ab3700544596f	{'rate': 0.0996}	253.16
9	6664f0cda92ab37005445a05	{'rate': 0.06}	467.45
10	6664f0cda92ab37005445c47	{'rate': 0.06}	231.03
11	6664f0cda92ab37005445db8	{'rate': 0.06}	216.38
12	6664f0cda92ab37005446f04	{'rate': 0.06}	380.27
13	6664f0cda92ab37005447214	{'rate': 0.074}	248.48
14	6664f0cda92ab37005447674	{'rate': 0.0964}	32.11
15	6664f0cda92ab37005447675	{'rate': 0.0996}	103.20
16	6664f0cda92ab37005447679	{'rate': 0.0996}	64.50
17	6664f0cda92ab3700544768a	{'rate': 0.0964}	141.25
18	6664f0cda92ab370054476bc	{'rate': 0.0996}	80.63

19	6664f0cda92ab3700544776d	{'rate': 0.0996}	257.99
20	6664f0cda92ab37005447777	{'rate': 0.0996}	51.60
21	6664f0cda92ab370054477b1	{'rate': 0.0996}	25.80
22	6664f0cda92ab370054477c3	{'rate': 0.0933}	31.96
23	6664f0cda92ab370054477d6	{'rate': 0.0996}	54.02
24	6664f0cda92ab37005447a2b	{'rate': 0.06}	184.28

		log	dti	fico \
0	{'annual': {'inc': 10.37349118}}	1.13	677	
1	{'annual': {'inc': 10.46310334}}	5.83	687	
2	{'annual': {'inc': 11.03488966}}	11.98	682	
3	{'annual': {'inc': 10.54534144}}	3.38	697	
4	{'annual': {'inc': 11.00209984}}	12.94	697	
5	{'annual': {'inc': 10.50780352}}	2.69	697	
6	{'annual': {'inc': 10.1266311}}	14.16	692	
7	{'annual': {'inc': 9.798127037}}	2.87	697	
8	{'annual': {'inc': 10.16585182}}	6.14	697	
9	{'annual': {'inc': 11.04292184}}	15.28	687	
10	{'annual': {'inc': 11.1844214}}	14.07	672	
11	{'annual': {'inc': 10.23995979}}	12.09	667	
12	{'annual': {'inc': 11.35040654}}	9.98	677	
13	{'annual': {'inc': 11.19821472}}	0.44	697	
14	{'annual': {'inc': 9.392661929}}	10.00	697	
15	{'annual': {'inc': 11.91839057}}	10.00	692	
16	{'annual': {'inc': 8.699514748}}	10.00	687	
17	{'annual': {'inc': 10.30895266}}	3.72	697	
18	{'annual': {'inc': 9.798127037}}	2.33	692	
19	{'annual': {'inc': 9.903487553}}	1.50	697	
20	{'annual': {'inc': 9.798127037}}	10.33	682	
21	{'annual': {'inc': 8.699514748}}	0.00	682	
22	{'annual': {'inc': 9.878169745}}	0.92	692	
23	{'annual': {'inc': 8.881836305}}	17.33	687	
24	{'annual': {'inc': 10.35964577}}	7.49	647	

		days	revol \
0	{'with': {'cr': {'line': 1920}}}	{'bal': 613, 'util': 76.6}	
1	{'with': {'cr': {'line': 2158.958333}}}	{'bal': 1188, 'util': 3}	
2	{'with': {'cr': {'line': 1799}}}	{'bal': 12640, 'util': 83.2}	
3	{'with': {'cr': {'line': 1782}}}	{'bal': 3305, 'util': 55.1}	
4	{'with': {'cr': {'line': 2969.041667}}}	{'bal': 2822, 'util': 9.9}	
5	{'with': {'cr': {'line': 4680}}}	{'bal': 1492, 'util': 33.2}	
6	{'with': {'cr': {'line': 1918.041667}}}	{'bal': 7121, 'util': 33.7}	
7	{'with': {'cr': {'line': 1199.041667}}}	{'bal': 1554, 'util': 29.3}	
8	{'with': {'cr': {'line': 2040.041667}}}	{'bal': 991, 'util': 34.2}	
9	{'with': {'cr': {'line': 4652.041667}}}	{'bal': 26044, 'util': 34.5}	
10	{'with': {'cr': {'line': 2910}}}	{'bal': 15885, 'util': 46.9}	
11	{'with': {'cr': {'line': 2550}}}	{'bal': 14320, 'util': 67.5}	
12	{'with': {'cr': {'line': 5705.041667}}}	{'bal': 10844, 'util': 68.6}	

```

13 {'with': {'cr': {'line': 5190.041667}}} {'bal': 666, 'util': 8.5}
14     {'with': {'cr': {'line': 1050}}} {'bal': 0, 'util': 48.7}
15 {'with': {'cr': {'line': 6030.041667}}} {'bal': 0, 'util': 19.8}
16     {'with': {'cr': {'line': 810}}} {'bal': 0, 'util': 70}
17     {'with': {'cr': {'line': 1050}}} {'bal': 3493, 'util': 63.5}
18 {'with': {'cr': {'line': 928.9583333}}} {'bal': 2937, 'util': 43.3}
19 {'with': {'cr': {'line': 479.0416667}}} {'bal': 358, 'util': 6}
20     {'with': {'cr': {'line': 2910}}} {'bal': 6646, 'util': 100.7}
21 {'with': {'cr': {'line': 480.0416667}}} {'bal': 0, 'util': 0}
22     {'with': {'cr': {'line': 360}}} {'bal': 362, 'util': 25.9}
23 {'with': {'cr': {'line': 540.0416667}}} {'bal': 2469, 'util': 29.7}
24 {'with': {'cr': {'line': 2369.958333}}} {'bal': 7884, 'util': 63.1}

```

```

not
0 {'fully': {'paid': 0}}
1 {'fully': {'paid': 1}}
2 {'fully': {'paid': 0}}
3 {'fully': {'paid': 0}}
4 {'fully': {'paid': 0}}
5 {'fully': {'paid': 0}}
6 {'fully': {'paid': 0}}
7 {'fully': {'paid': 0}}
8 {'fully': {'paid': 0}}
9 {'fully': {'paid': 0}}
10 {'fully': {'paid': 0}}
11 {'fully': {'paid': 0}}
12 {'fully': {'paid': 0}}
13 {'fully': {'paid': 0}}
14 {'fully': {'paid': 0}}
15 {'fully': {'paid': 0}}
16 {'fully': {'paid': 0}}
17 {'fully': {'paid': 0}}
18 {'fully': {'paid': 1}}
19 {'fully': {'paid': 0}}
20 {'fully': {'paid': 1}}
21 {'fully': {'paid': 0}}
22 {'fully': {'paid': 0}}
23 {'fully': {'paid': 1}}
24 {'fully': {'paid': 0}}

```

We found only 25 records that has interest rate lower than 10%, and they are all close to 700 fico score

[]: