

The American Statistical Association DataFest

- 최우수상 수상 (2023)

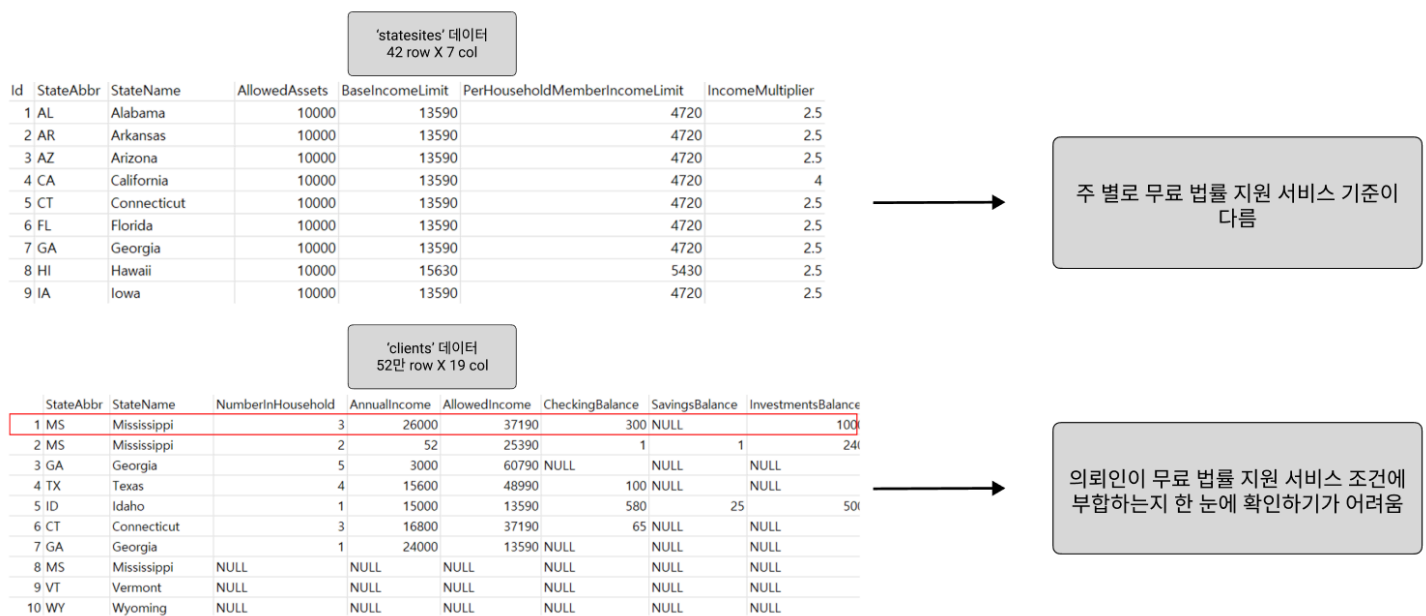
- Team Megabyte (Younseo Kim, Yong Jun Choi, Woosung Lim, Hyeongkwon Kim & Taehwan Lee)
- 2023.03.25 ~ 2023.03.26
- 프로그래밍 언어 및 도구: R (tidyverse, dplyr, ggplot2), Tableau, Excel, Figma
- 활용한 기술: EDA, 데이터 전처리, 알고리즘 개발, 프로토타입 UI/UX 디자인

[프로젝트 소개]

- 미국 변호사 협회에서 제공한 50만 건 이상의 의뢰인 데이터, 2만 명의 변호사 데이터, 30만 건 이상의 상담 데이터를 활용하여 협회의 무료 법률 지원 상담 서비스의 효율성을 개선한 프로젝트

Step 1: 의뢰인의 무료 법률 서비스 자격 판별

[문제점]



- 주 마다 무료 법률 서비스 기준이 상이합니다.
- 무료 법률 서비스를 받기 위해서는 자산과 수입 두 가지 기준을 모두 충족해야 하며, 관련 변수들이 나뉘어져 있어 변호사들이 의뢰인이 무료 법률 서비스 기준에 부합하는지 한눈에 판단하기 어렵습니다.

[데이터 전처리]

Raw 'clients' 데이터
52만 row X 19 col

AnnualIncome	AllowedIncome	CheckingBalance	SavingsBalance	InvestmentsBalance
NULL	NULL	NULL	NULL	NULL
46000	60790	300	NULL	NULL
15000	60790	500	NULL	NULL
NULL	NULL	NULL	NULL	NULL
5000	25390	1	NULL	NULL
NULL	NULL	NULL	NULL	NULL
19000	13590	NULL	NULL	NULL
36000	13590	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL
57000	51350	NULL	NULL	NULL
13000	37190	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL
22000	48990	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL
9000	48990	1000	NULL	NULL
31680	37190	-45	0	NULL
1800	51350	0	NULL	NULL

AnnualIncome 값이 NULL인 의뢰인 제거

```
# convert it to numeric value
clients$AnnualIncome = as.numeric(clients$AnnualIncome)
clients$AllowedIncome = as.numeric(clients$AllowedIncome)
clients$CheckingBalance = as.numeric(clients$CheckingBalance)
clients$SavingsBalance = as.numeric(clients$SavingsBalance)
clients$InvestmentsBalance = as.numeric(clients$InvestmentsBalance)
```

Categorical 형식의 자산과 수입에 관련된 변수들을 Numeric 형식으로 변환

'clients' 데이터

Id	StateAbbr	StateName	AllowedAssets
1	AL	Alabama	10000
2	AR	Arkansas	10000
3	AZ	Arizona	10000
4	CA	California	10000
5	CT	Connecticut	10000
6	FL	Florida	10000
7	GA	Georgia	10000
8	HI	Hawaii	10000
9	IA	Iowa	10000
10	ID	Idaho	10000
11	IL	Illinois	500000
12	IN	Indiana	10000
13	KS	Kansas	10000
14	LA	Louisiana	500000
15	MA	Massachusetts	10000

+

'statesites' 데이터

Id	StateAbbr	AnnualIncome	AllowedIncome
26642	FL	13000.0	13590
9357	NH	0.0	51350
15793	VA	361920.0	13590
19979	UT	600.0	37190
26422	MI	38000.0	41910
5943	MA	0.0	13590
14185	CA	1500.0	32470
11423	KS	5000.0	84390
1951	IA	56643.0	25390
24735	MD	33000.0	37190
20079	AR	12000.0	13590
23634	AL	48.0	37190
2153	LA	10400.0	32470
629	PA	14000.0	48990
20130	US	12000.0	13590

=

Id	StateAbbr	AnnualIncome	AllowedIncome	CheckingBalance	SavingsBalance	InvestmentsBalance	AllowedAssets
26642	FL	13000.0	13590	400.00	2000.00	N/A	10000
9357	NH	0.0	51350	0.00	500.00	N/A	10000
15793	VA	361920.0	13590	72.00	N/A	N/A	10000
19979	UT	600.0	37190	N/A	N/A	N/A	10000
26422	MI	38000.0	41910	N/A	N/A	89.06	10000
5943	MA	0.0	13590	N/A	N/A	N/A	10000
14185	CA	1500.0	32470	N/A	N/A	N/A	10000
11423	KS	5000.0	84390	N/A	N/A	N/A	10000
1951	IA	56643.0	25390	N/A	N/A	N/A	10000
24735	MD	33000.0	37190	N/A	N/A	N/A	10000
20079	AR	12000.0	13590	700.00	N/A	N/A	10000
23634	AL	48.0	37190	N/A	N/A	N/A	10000
2153	LA	10400.0	32470	5.00	N/A	N/A	500000
629	PA	14000.0	48990	450.00	15.00	N/A	10000
20130	US	12000.0	13590	-114.00	N/A	N/A	10000

의뢰인 데이터에 LEFT JOIN을 통하여 의뢰인의 주별 무료 법률 서비스 기준 자산 변수를 생성

Cleaned 'clients' 데이터
43만 row X 21 col

AnnualIncome	AllowedIncome	CheckingBalance	SavingsBalance	InvestmentsBalance	AllowedAssets	sum_asset
13000.0	13590	400.00	2000.00	N/A	10000	2400.00
0.0	51350	0.00	500.00	N/A	10000	500.00
361920.0	13590	72.00	N/A	N/A	10000	72.00
600.0	37190	N/A	N/A	N/A	10000	0.00
38000.0	41910	N/A	N/A	89.06	10000	89.06
0.0	13590	N/A	N/A	N/A	10000	3.00
1500.0	32470	N/A	N/A	N/A	10000	1900.00
5000.0	84390	N/A	N/A	N/A	10000	0.00
56643.0	25390	N/A	N/A	N/A	10000	1142.00
33000.0	37190	N/A	N/A	N/A	10000	0.00
12000.0	13590	700.00	N/A	N/A	10000	700.00
48.0	37190	N/A	N/A	N/A	10000	0.00
10400.0	32470	5.00	N/A	N/A	500000	5.00
14000.0	48990	450.00	15.00	N/A	10000	465.00
12000.0	13590	-114.00	N/A	N/A	10000	-114.00

자산과 관련된 변수들을 하나의 변수로 통합한 sum_assets 파생변수 생성

1. clients 데이터셋에서 'AnnualIncome' (수입 변수)이 NULL인 경우, 자산 관련 변수들도 모두 NULL이므로 'AnnualIncome'이 NULL인 의뢰인들을 제거합니다.
2. clients 데이터셋의 자산과 수입에 관련된 변수들이 Categorical Variable (범주형 변수)로 저장되어 있어 Numeric Variable (숫자형 변수)로 변환합니다.
3. clients 데이터셋의 'StateAbbr' 변수를 통해 statesites (주별 무료 법률 서비스 기준 데이터셋)과 LEFT JOIN하여 통합된 데이터셋을 생성합니다.
4. clients 데이터셋의 자산과 관련된 변수들을 하나의 변수로 통합한 'sum_assets' 파생변수를 생성합니다.

(대회 진행요원이 자산과 관련된 변수들에서 의뢰인들이 0을 적는 대신 기입을 하지 않은 경우가 많다고 하여, 자산과 관련된 변수들 중 1개 이상 기입한 의뢰인들에 한해서는 NULL 값을 0으로 간주합니다)

Pro Bono Classified

AnnualIncor	AllowedInco
26000	37190
52	25390
3000	60790
15600	48990
15000	13590
16800	37190
24000	13590
NA	NA
NA	NA
NA	NA

CheckingBalance	SavingsBalar	Investments	sum_assest	AllowedAssets
300	NA	NA	300	10000
1	1	NA	2	10000
NA	NA	NA	0	10000
100	NA	NA	100	10000
580	25	NA	605	10000
65	NA	NA	65	10000
NA	NA	NA	0	10000
NA	NA	NA	0	10000
NA	NA	NA	0	10000
NA	NA	NA	0	10000

AnnualIncor	AllowedInco	ProBono	CheckingBalance	SavingsBalar	Investments	sum_assest	AllowedAssets	ProBono2	ProBono_Final
26000	37190	Y	300	NA	NA	300	10000	Y	Y
52	25390	Y	1	1	NA	2	10000	Y	Y
3000	60790	Y	NA	NA	NA	0	10000	NC	Y
15600	48990	Y	100	NA	NA	100	10000	Y	Y
15000	13590	N	580	25	NA	605	10000	Y	N
16800	37190	Y	65	NA	NA	65	10000	Y	Y
24000	13590	N	NA	NA	NA	0	10000	NC	N
NA	NA	NC	NA	NA	NA	0	10000	NC	NC
NA	NA	NC	NA	NA	NA	0	10000	NC	NC
NA	NA	NC	NA	NA	NA	0	10000	NC	NC

- 수입 기준:
 - 'AnnualIncome'이 'AllowedIncome'보다 낮으면 'ProBono_income'을 Y, 높으면 N으로 저장합니다.
- 자산 기준:
 - statesites (주별 무료 법률 기준 데이터셋)에 없는 주 출신의 경우 'ProBono_assets'를 NAA (No Allowed Assets)로 저장합니다.
 - 자산 관련 변수들이 모두 NULL이면 'ProBono_assets'를 NC (Not Classified)로 저장합니다.
 - 'sum_assests'가 'AllowedAssets'보다 낮으면 'ProBono_assets'를 Y, 높으면 N으로 저장합니다.
- 최종 판정:
 - 'ProBono_income'과 'ProBono_assets'가 모두 Y이면 'ProBono_final'을 Y로 저장합니다.
 - 'ProBono_income' 또는 'ProBono_assets' 중 하나라도 N이면 'ProBono_final'을 N으로 저장합니다.
 - 'ProBono_income'이 Y이고 'ProBono_assets'가 NAA이면 'ProBono_final'을 Y로 저장합니다. (무료 법률 서비스 기준 자산이 없는 주들이 존재)
 - 'ProBono_income'과 상관없이 'ProBono_assets'가 NC이면 'ProBono_final'을 NC로 저장합니다.

Step 2: 효율적인 변호사와 의뢰인 매칭

[문제점]

'attorneys' 데이터
2만 row X 8 col

Id	StateAbbr	AttorneyUno	City	County	StateName	PostalCode	CreatedUtc
1	NH	7A2C12BE-9F37-4C9D-A6ED-41B916820CE0	Concord	Cheshire	New Hampshire	3301	2019-02-14 16:21:52
2	NH	29A49AA6-047E-4008-ABBD-3DE57156B526	Concord	Merrimack	New Hampshire	3229	2019-02-14 18:14:30
3	MS	9DF11865-7F29-452C-A69E-54CD72E4A116	Jackson	Rankin	Mississippi	39203	2016-08-19 17:07:23
4	VA	693FD55F-5F58-405A-B389-5E78B3406497	Richmond	Richmond (city)	Virginia	23219	2016-08-19 18:22:53
5	CT	0AC2A866-19CD-4D1A-94B7-8F41A1430553	Wethersfield	Hartford	Connecticut	6410	2016-08-21 01:21:13
6	LA	231AB5D6-6431-4BAE-8AC3-0E4538FD7062	Covington	St. Tammany Parish	Louisiana	70433	2016-08-16 19:45:15
7	MS	0FAD5B41-848C-40E9-A57F-9AFF731D4993	Jackson	Hinds	Mississippi	39207	2016-08-22 21:51:32
8	VA	0D0CF105-0EAB-4A0B-966E-B291AB00186E	Richmond	Richmond city	Virginia	23220	2016-08-22 14:48:23
9	AR	3AFFD9E0-2246-4125-AB80-76A5B8F888A	Little Rock	Pulaski	Arkansas	72212	2016-08-25 14:04:46
10	CT	2C68DD67-A823-42DD-A645-D5549652CDC7	New Haven	New Haven	Connecticut	6511	2016-08-23 15:28:23
11	LA	D3C78D34-1A3A-4E1A-B2A0-7F10D2F6031E	Ruston	Lincoln	Louisiana	71270	2016-08-16 20:14:09
12	MI	2180B56F-642E-461F-92FE-7EEE0FAD9E41	Lansing	Ingham	Michigan	48933	2016-11-16 19:43:59
13	VA	74B15243-2169-4062-BDF3-C7A24748B21B	Falls Church	Fairfax	Virginia	22042	2016-08-22 15:24:33
14	WY	323CC3D0-917D-4F86-80FD-5F856419A42E	Cheyenne	Laramie	Wyoming	82002	2016-08-16 17:17:25
15	LA	8EA03754-959A-4796-98E0-786F8A9D1634	Metairie	Orleans	Louisiana	70001	2016-08-17 20:49:31

변호사의 전문 분야 또는
실력을 판단할 수 있는 요소가
없음

'questions' 데이터
30만 row X 8 col

Category	SubcategoryUno	Subcategory	AskedByClientUno	AskedOnUtc	TakenByAttorneyUno
Family and Children	E9F87919-0C05-4D9D-AA6C-3BC49132F1D	Family/Divorce/Custody/Visitation/Child Support	9C35423C-914A-42BC-8B8C-A370E76E2178	2019-06-19 22:10:06	NULL
Individual Rights	73A59926-BA08-44EC-ABF3-34C9181EEAB3	Civil/Constitutional Rights	DF636AEF-7008-4F90-9A5C-0A2A726C6859	2019-06-19 23:34:36	NULL
Individual Rights	73A59926-BA08-44EC-ABF3-34C9181EEAB3	Civil/Constitutional Rights	DF636AEF-7008-4F90-9A5C-0A2A726C6859	2019-06-19 23:35:18	FA341428-1575-4720-981C-2D28C3560137
Family and Children	E9F87919-0C05-4D9D-AA6C-3BC49132F1D	Family/Divorce/Custody/Visitation/Child Support	C9C3FA40-0060-4989-97F1-63AFEF4FA484	2019-06-20 02:33:14	FA341428-1575-4720-981C-2D28C3560137
Consumer Financial Questions	73FE406E-2B51-456A-A6D9-78A7D05802AA	Bankruptcy/ Debts & Purchases	ED7DEA10-CD87-4A05-891D-889D1EA710BE	2019-06-20 17:04:47	NULL
Family and Children	E9F87919-0C05-4D9D-AA6C-3BC49132F1D	Family/Divorce/Custody/Visitation/Child Support	19A1157F-8AF1-4378-84FA-41C5FD6DE97	2019-06-20 18:11:56	NULL
Housing and Homelessness	F4B89ABA-D699-48FD-A6E7-58F895CC4021	Housing/Eviction/Foreclosure/Tenants	32D8D09B-C6CF-481B-82A9-985F199C8C6	2019-06-20 19:20:02	NULL
Other	E3257971-6837-4B4D-800D-1E404D0C8B8A	Civil Procedure	714BF984-8485-48F5-93D9-4742C862643B	2019-06-20 18:46:40	FA341428-1575-4720-981C-2D28C3560137
Other	A911AEC6-4E0E-4583-8007-024EC32CADE	Expungement	3A361F5F-C49D-41AF-88CA-8C211D14D6E1	2019-06-20 21:26:13	9514BCF0-8AD4-4948-812E-70688F451492
Family and Children	8C7CE01E-1EDE-42CA-8C99-0233E568585F	Estate/Wills/Inheritance/Power of Attorney	11262CC7-711D-4387-86D7-975C18D8DB96	2019-06-21 17:32:08	3431BD06-C607-4541-9348-3EBA71BD08BA
Housing and Homelessness	F4B89ABA-D699-48FD-A6E7-58F895CC4021	Housing/Eviction/Foreclosure/Tenants	BE3FC73C-50CF-4175-A721-2BC448688CA	2019-06-21 20:19:52	NULL
Other	81D88EAF-586A-AC9B-8E9A-CAEAED081D70	Personal Injury	10C82829-7592-4D22-86AE-D483DABF4074	2019-06-21 21:27:20	9514BCF0-8AD4-4948-812E-70688F451492
Family and Children	E9F87919-0C05-4D9D-AA6C-3BC49132F1D	Family/Divorce/Custody/Visitation/Child Support	C78C3854-9155-4985-AF51-0241610E4AE1	2019-06-22 07:46:44	6A82E739-8D68-4FA7-8651-DC29564324B
Income Maintenance	E7C4D9A8-8153-4F8D-9A52-FAC7378E28CA	Government Benefits (SNAP, TANF, SSI, SSDI)	3A3EAD02-8233-4262-9700-39D062D18314	2019-06-22 21:05:47	NULL
Housing and Homelessness	F4B89ABA-D699-48FD-A6E7-58F895CC4021	Housing/Eviction/Foreclosure/Tenants	EB012D48-46CA-408F-A625-38DA21E8D088	2019-06-23 12:15:22	3431BD06-C607-4541-9348-3EBA71BD08BA

변호사의 상담 시간, 가장 많이
상담한 분야, 처리한 의뢰 수를
한눈에 파악 할 수 없음

'attorneytimeentries' 데이터
20만 row X 6 col

Id	StateAbbr	TimeEntryUno	AttorneyUno	Hours	EnteredOnUtc
84703	TX	888B9751-87DF-467D-967E-322A9789A48B	0023FF7C-934D-4974-87A9-B90B52D4D5B	0.6	2020-03-09 18:25:06
84793	TX	4389E156-9508-4558-898E-1F6C28CA070A	0023FF7C-934D-4974-87A9-B90B52D4D5B	1.0	2020-03-11 19:56:02
85193	TX	36C8D519-5633-4530-98C0-FF732E36235	0023FF7C-934D-4974-87A9-B90B52D4D5B	1.0	2020-03-16 18:13:46
52869	IL	EF822A21-D79F-45F1-ACC9-E2826977F2A7	003CA876-C682-4E9A-8161-48E485AD0F8D	0.5	2020-02-13 19:16:23
9101	TX	598F4A40-CB46-42C5-A3EE-95F4B34817BE	00401285-58EE-4E39-9807-87A570C4B9ED	0.5	2017-09-12 19:59:35
9441	TX	4C7AE2A5-5C52-4FA7-8730-1067906D8272	00401285-58EE-4E39-9807-87A570C4B9ED	0.5	2017-09-15 14:19:42
13892	TX	29839026-92DA-4AC5-93C6-0A01067FC433	00401285-58EE-4E39-9807-87A570C4B9ED	0.6	2017-10-24 14:05:44
12670	GA	72D8B823-E168-4B76-9D1E-D31AD74567D8	00609E0E-87BE-496A-9847-A711F4569A25	0.2	2020-11-16 18:18:28
12703	GA	E18579C1-9641-439D-81AE-D235D03C56F5	00609E0E-87BE-496A-9847-A711F4569A25	0.2	2020-11-16 18:25:45
12736	GA	69093379-ACCF-4F30-8943-F85E002CSA9E	00609E0E-87BE-496A-9847-A711F4569A25	0.2	2020-11-16 18:36:08
12802	GA	51C21AA2-E6E7-4FCB-9093-1E6B5E2FA08D	00609E0E-87BE-496A-9847-A711F4569A25	0.3	2020-11-16 18:41:56
12835	GA	D6E681AC-CE1A-4C84-A97F-67FED980BEEC	00609E0E-87BE-496A-9847-A711F4569A25	0.1	2020-11-16 19:25:14
52456	MA	C32A7ASA-ABC3-4198-AD26-0EEAF5679635	00671667-CEAB-498B-B206-1035682D833B	0.5	2020-07-17 20:48:37
62074	MA	7EB99280-0689-4327-A006-878EBCDE063	00671667-CEAB-498B-B206-1035682D833B	0.4	2021-01-24 06:05:04
62087	MA	825F2849-8978-48B2-A537-8E217B24703E	00671667-CEAB-498B-B206-1035682D833B	0.2	2021-01-24 06:32:09

- 의뢰인이 변호사의 전문 분야와 실력을 판단할 수 있는 요소가 없습니다.
- 변호사가 지금까지 의뢰인들과 얼마나 상담했는지, 어떤 전문 분야에서 가장 많이 상담했는지, 얼마나 많은 의뢰를 처리했는지 한눈에 파악할 수 없습니다.

[데이터 전처리]

Raw 'questions' 데이터
30만 row X 14 col

Category	SubcategoryUno	Subcategory	AskedByClientUno	AskedOnUtc	TakenByAttorneyUno	TakenOnUtc
Family and Children	E9F87919-0CD5-4D9D-AAC6-3CB49132F1D	Family/Divorce/Custody/Visitation/Child Support	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-19 22:10:06	NULL	NULL
Individual Rights	73A59926-BA08-44EC-ABF3-F4C9181EEA83	Civil/Constitutional Rights	DF636AEE-7D08-4F90-9ASC-0A2A726C6B59	2019-06-19 23:34:36	NULL	NULL
Individual Rights	73A59926-BA08-44EC-ABF3-F4C9181EEA83	Civil/Constitutional Rights	DF636AEE-7D08-4F90-9ASC-0A2A726C6B59	2019-06-19 23:57:18	FA34142B-1575-4720-981C-2D28C3560137	2019-06-22 02:37:16
Family and Children	E9F87919-0CD5-4D9D-AAC6-3CB49132F1D	Family/Divorce/Custody/Visitation/Child Support	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-20 02:33:14	FA34142B-1575-4720-981C-2D28C3560137	2019-07-27 18:04:08
Consumer Financial Questions	73F4866E-2B51-456A-A6D9-7BA7D05802AA	Banking/Consumer Financial Questions	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-20 17:04:47	NULL	NULL
Family and Children	E9F87919-0CD5-4D9D-AAC6-3CB49132F1D	Family/Divorce/Custody/Visitation/Child Support	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-20 18:11:56	NULL	NULL
Housing and Homelessness	F48B9A8A-D699-48FD-A6E7-58F895CC4021	Housing/Divorce/Custody/Visitation/Child Support	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-20 19:20:02	NULL	NULL
Other	E3257971-6837-4BAD-800D-E404F0C888A	Other	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-20 18:46:40	FA34142B-1575-4720-981C-2D28C3560137	2019-06-24 19:04:14
Other	A911AEC6-4E0E-45B3-8807-024EC32EAEDE	Expulsion/Other	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-20 21:26:13	9514BCF0-BAD4-4948-B12E-7068BF451492	2019-07-18 17:39:46
Family and Children	8C7CE01E-1EDE-42C4-8C99-0233E565855F	Establishment/Other	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-21 17:32:08	3431BD06-C607-4541-934B-3E8AA718DB8A	2019-07-08 20:00:18
Housing and Homelessness	F48B9A8A-D699-48FD-A6E7-58F895CC4021	Housing/Divorce/Custody/Visitation/Child Support	9C35423C-914A-42BC-8EBC-A370E76E2178	2019-06-21 20:19:52	NULL	NULL
Other	81D88E4F-586A-4C9B-8E9A-C4AEAD081D7D	Personal Injury	10C82829-7592-4D22-86AE-D482DABF4074	2019-06-21 21:27:20	9514BCF0-BAD4-4948-B12E-7068BF451492	2019-07-18 17:31:51
Family and Children	E9F87919-0CD5-4D9D-AAC6-3CB49132F1D	Family/Divorce/Custody/Visitation/Child Support	C7BC3B54-9155-49B5-AF51-D241610E4AE1	2019-06-22 07:46:44	6A82E739-BD68-4FA7-8651-DC2956423248	2019-07-18 17:32:48
Income Maintenance	E7CAD9A8-B153-4FBD-9A52-FAC737BE28C4	Government Benefits (SNAP, TANF, SOS, SSI/D)	3A3EADD2-8233-4262-9700-39D062D1B314	2019-06-22 21:05:47	NULL	NULL
Housing and Homelessness	F48B9A8A-D699-48FD-A6E7-58F895CC4021	Housing/Divorce/Custody/Visitation/Child Support	EB012D4B-45CA-408F-A625-3BDA21E80D88	2019-06-23 12:15:22	3431BD06-C607-4541-934B-3E8AA718DB8A	2019-06-24 15:34:20

TakenByAttorneyUno가 NULL인
기록들을 제거
->
Cleaned 'questions' 데이터
22만 row X 14 col



Category	Subcategory1	Subcategory2	Subcategory3	TotalHours	NumberofCases
NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA
Work, Employment and Unemployment	Employment/Labor Law	Other	Landlord/Tenant	9.70	34
NA	NA	NA	NA	NA	NA
Family and Children	Family/Divorce/Custody/Adoption			1.80	9
NA	NA			NA	NA
NA	NA			NA	NA
Housing and Homelessness	Housing or Property Owned			41.20	126
Consumer Financial Questions	Debts & Purchases			0.40	3
Family and Children	Family Law or Divorce or Custody	Other	NA	0.90	3
Family and Children	Estates and Wills	Other	NA	49.10	72
Other	Wildfire or COVID-19: Other related issue	Wildfire or COVID-19: Personal Injury	NA	5.70	2
NA	NA	NA	NA	NA	NA
Family and Children	Family/Safety/Divorce/Custody	Civil/Constitutional Rights & Discrimination	Housing/Renting/Homelessness	9.00	14

'attorneys' 데이터셋에 변호사의 전문분야,
세부 전문분야들, 상담에 할애한 총 시간과
상담 기록의 횟수를 알려주는 변수들을 추가

- questions (상담 기록 데이터셋)에서 변호사가 답을 하지 않은 기록들을 제거합니다. ('TakenByAttorneyUno'가 NULL인 경우)
 - questions에서 변호사가 가장 많이 답변한 'Category' (전문분야)와 'Subcategory' (세부 전문분야)를 분석하여 변호사의 전문분야와 세부 전문분야를 알려주는 변수들을 기존 attorneys (변호사 데이터셋)에 생성합니다.
 - questions에서 변호사가 맡은 상담 기록의 횟수를 'NumberofCases'라는 변수를 attorneys에 생성합니다.
 - attorneytimeentries (상담 당 변호사가 할애한 시간 데이터셋)에서 변호사가 할애한 시간의 총합을 'TotalHours'라는 변수를 attorneys에 생성합니다.
- 모든 변수들이 NULL인 변호사는 등록만 되어 있을 뿐, 상담을 진행하지 않은 변호사입니다.

[알고리즘 개발]

New Attorney database

	AttorneyUnd	Category	Subcat1	Subcat2	Subcat3	TotalHours	NumberofCases	StateAbbr	TotalHoursLevel	NumberOfCasesLevel	Level
121	0554530A-7	Family and (Family/Safety, NA			NA	2.5	6	IL	2	3	2.5
122	055D91CE-0	Work, Empli Employment C Housing Discri			Unemployment	0.9	4	IN	1	2	1.5
123	05829D2F-D	Family and (Family/Divorce NA			NA	1.9	16	NC	2	3	2.5
124	0591A8ED-0	Family and (Family/Divorce NA			NA	0.3	2	HI	1	2	1.5
125	05A87916-3	Family and (Family/Divorce Landlord/Tena			Constitutional	2	21	TX	2	3	2.5
126	05D3D641-8	Family and (Family/Divorce NA			NA	0.1	1	TN	1	1	1
127	05DD8A38-5	Family and (Family/Divorce Rental Propert			Debts & Purch	625.9	3163	MO	3	3	3
128	05E7DFD3-6	Family and (Family/Divorce Landlord/Tena			Expunction/Ex	5.5	12	TX	3	3	3
129	0609175C-D	Housing anc Housing/Evicti Work/Employr			NA	NA	2	NC	1	2	1.5
130	06260DD3-9	Family and (Family/Divorce Domestic Viole			Employment	14.5	20	TX	3	3	3
131	063CD5F6-D	Work, Empli Work/Employ NA			NA	NA	1	NC	1	1	1
132	064EFF3C-6	Housing anc Housing or Res Debts, Bankrup			Other	97.3	124	ME	3	3	3
133	0661E696-8	Housing anc Home Owners NA			NA	NA	1	NE	1	1	1
134	06677FFB-8	Family and (Wills/Inheritar Civil/Constituti			Family/Divorce	3.9	7	AL	2	3	2.5

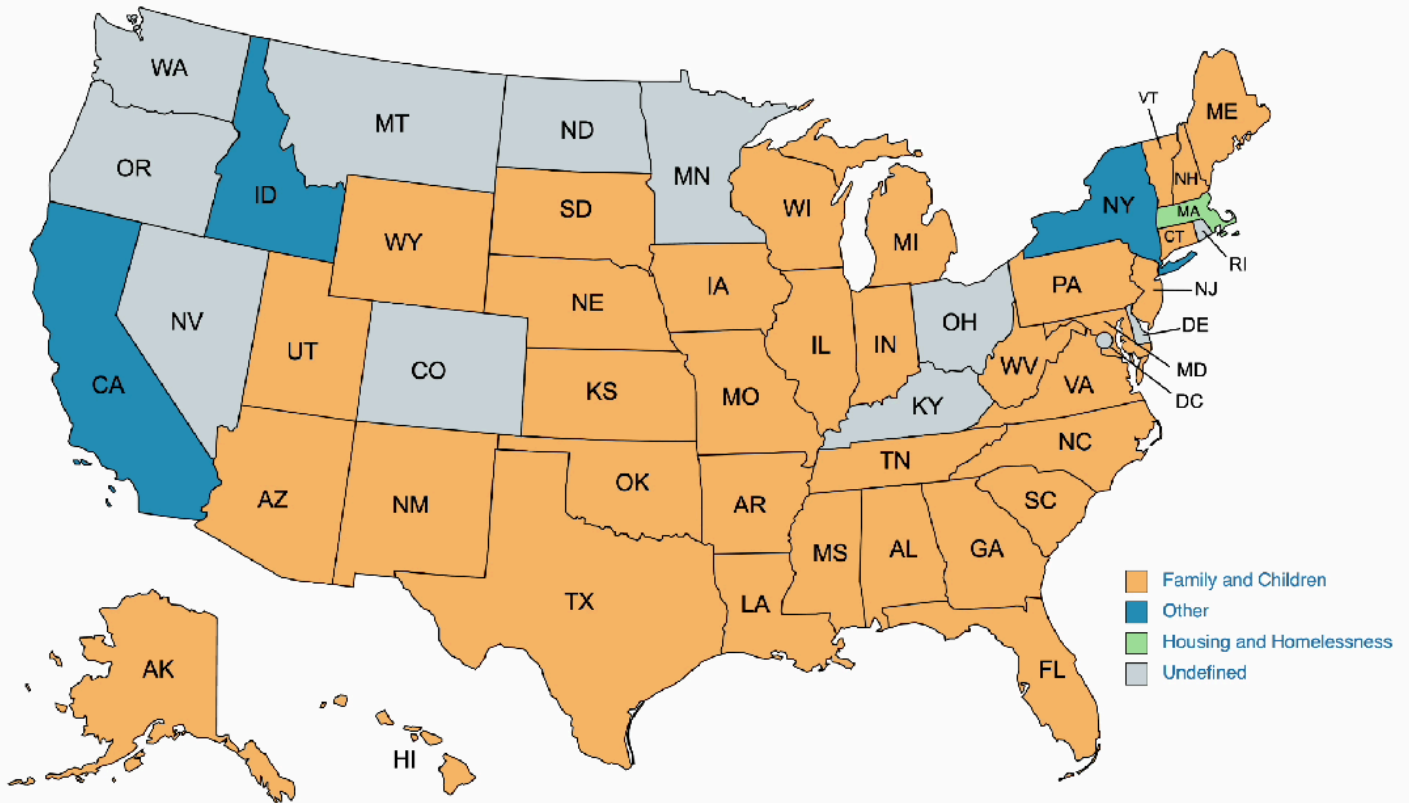


1. 'TotalHours'와 'NumberofCases'의 하위 40%, 60%, 100%의 기준값을 구하여 'TotalHoursLevel'과 'NumberofCasesLevel'의 기준점을 설정합니다.
2. 'TotalHours'에 대해 하위 40%는 Level 1, 40%에서 80% 사이는 Level 2, 80%에서 100% 사이는 Level 3으로 설정하며, 'TotalHours'가 0인 경우는 Level 1로 간주합니다.
3. 'NumberofCases'에 대해서도 동일하게 적용하여 'NumberofCasesLevel' 변수를 생성하되, 'NumberofCases'가 0인 경우는 Level 1로 간주합니다.
4. 'TotalHoursLevel'과 'NumberofCasesLevel'을 종합하여 변호사의 전체적인 실력을 판단할 수 있는 Level 변수를 생성합니다.

Step 3: 미국 내 법률 상담 트렌드 분석

[가장 많이 상담된 법률 카테고리]

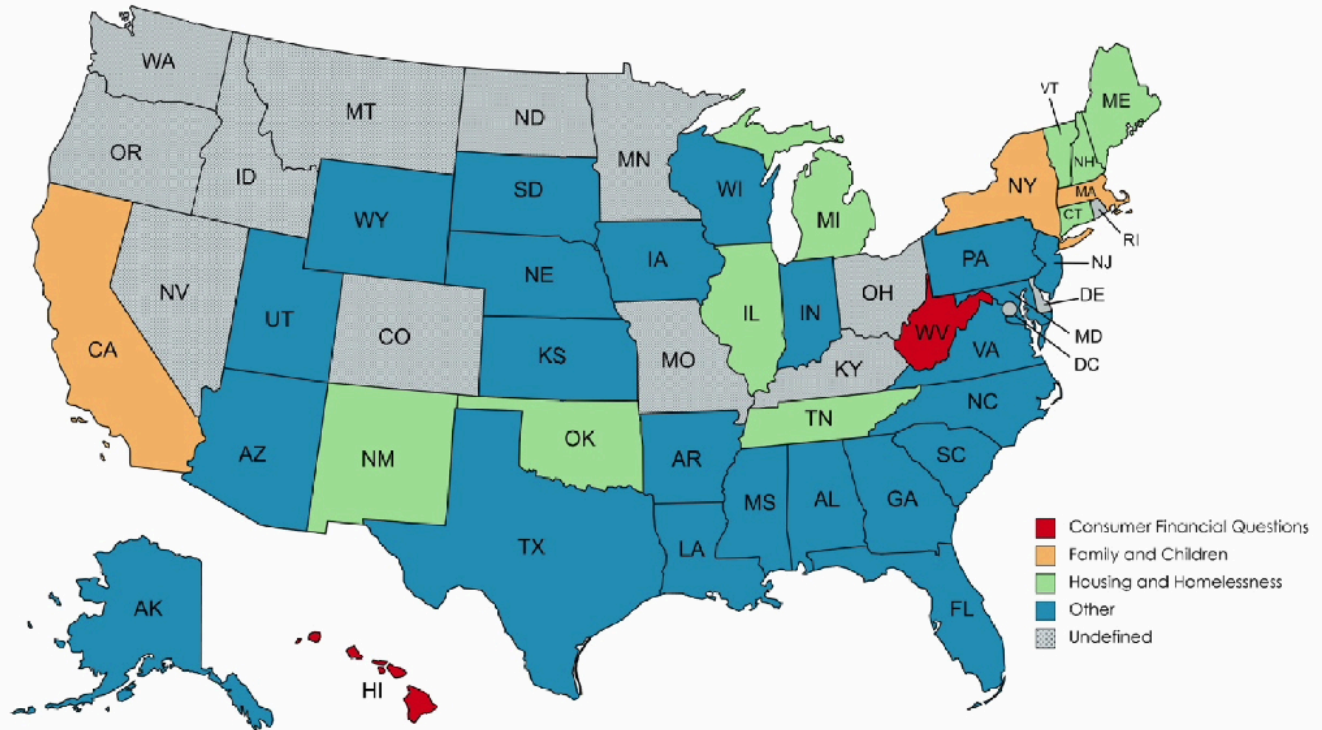
The Most Frequent Category by states



- 주 별로 미국 변호사 협회의 무료 법률 상담 서비스에서 가장 많이 상담된 카테고리는 가족 및 아동(**Family and Children**) 으로 나타납니다.
- 이는 많은 주에서 이혼, 양육권 분쟁, 가정 폭력 등 가족 갈등과 관련된 법률 상담이 많이 이루어지고 있음을 나타냅니다.

【두번째로 많이 상담된 법률 카테고리】

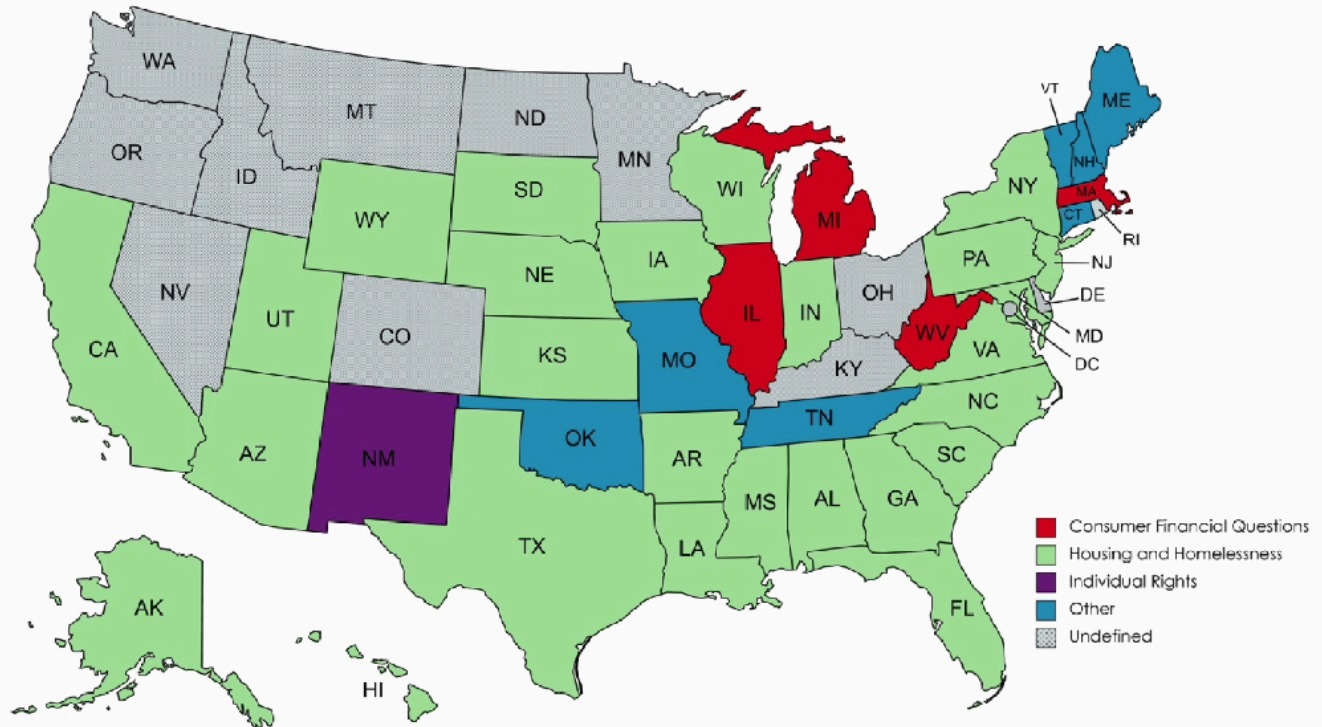
2nd most frequent Category by states



- 주별로 미국 변호사 협회의 무료 법률 상담 서비스에서 두 번째로 많이 상담된 카테고리는 기타(**Others**)로 나타났습니다.
- 기타 카테고리에는 상해, 자연재해, 이민, 세금과 관련된 상담들이 포함되며, 이는 많은 주에서 일상 생활과 관련된 법률 상담이 많이 이루어지고 있음을 나타냅니다.

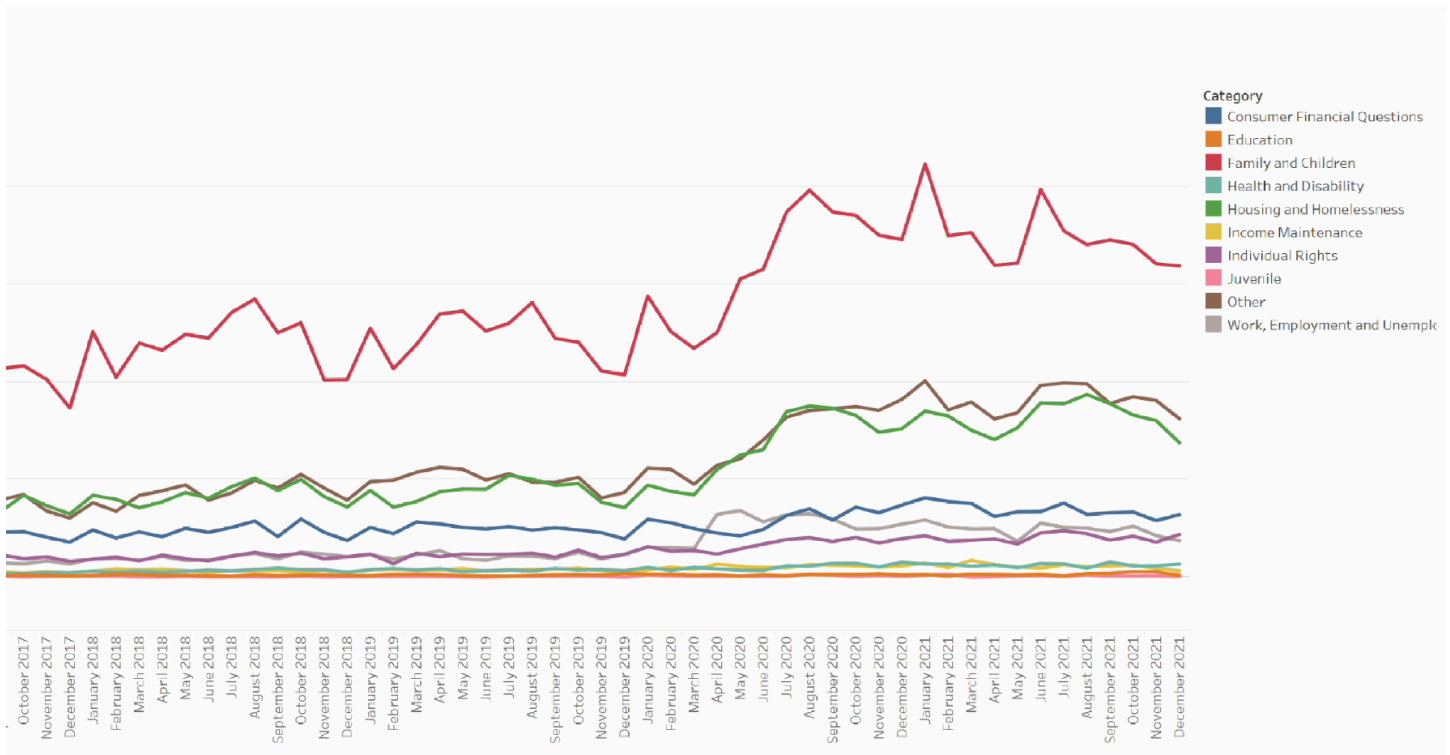
【세번째로 많이 상담된 법률 카테고리】

3rd most frequent Category by states



- 주별로 미국 변호사 협회의 무료 법률 상담 서비스에서 세 번째로 많이 상담된 카테고리는 주거 및 노숙자(**Housing and Homelessness**)로 나타났습니다.
- 이는 여러 주에서 주거 안정성, 세입자 권리, 노숙자 문제와 관련된 법률 상담이 많이 이루어지고 있음을 나타냅니다.

[시간에 따른 법률 상담 카테고리별 추이]





- 2020년도 초반부터 가족 및 아동, 기타 (상해, 자연재해, 이민, 세금과 관련된 상담), 주거 및 노숙자, 노동, 고용 및 실업에 관련 법률 상담의 급격한 증가가 관찰되었습니다.
- 가족 및 아동 상담의 증가는 코로나로 인해 재택근무가 증가하면서 가족 간의 갈등이 많아졌기 때문으로 추측됩니다.
- 기타 카테고리 상담의 증가는 코로나 관련 법률 상담이 자연재해로 분류되었기 때문입니다.
- 주거 및 노숙자 상담의 증가는 팬데믹 동안 경제적 어려움으로 주거 불안정성이 심화되었기 때문으로 추정됩니다.
- 노동, 고용 및 실업 관련 상담의 증가는 팬데믹 동안 실업을 증가와 재택근무로 인한 변화 때문으로 추정됩니다.

Step 4: 실제 적용 프로토타입

[의뢰인의 무료 법률 서비스 자격 판별, 효율적인 변호사와 의뢰인 매칭]

Data Search

 Nick Colson

Profile Status

Income

Assets

Your Cases

Case1

Case2

Case3

Best Attorney for you

Lv.3

James

Best



Expert at: Family and Children
Hours: 200
Total Cases: 3432

Lv.3

Caithy

Expert at: Family and Children
Hours: 20
Total Cases: 1800

Data Search

 Nick Colson

< Back

Log Out

Attorney information

Name	James	Expert*	Family and Children
State	TN	hour	1.8
County	Montgomery	Cases done	12332
Mail	qwer1234@asdf.com	Total Level	3

Details

No	Category	Sub Category	Description	Date
1	Family and Children	Divorce	N/A	YYYY.MM.DD
2	Family and Children	Abuse	Blah Blah	YYYY.MM.DD
3	Family and Children	Divorce	Blah Blah	YYYY.MM.DD
4	Individual Rights	Abuse	Blah Blah	YYYY.MM.DD

- 프로토타입을 사용함으로써 의뢰인이 무료 법률 지원 상담 서비스의 수입과 자산 기준에 부합하는지 확인할 수 있습니다.
- 또한, 의뢰인이 신청한 상담의 법률 카테고리를 전문 분야로 가지고 있는 변호사를 추천해주며, 변호사가 지금까지 얼마나 많은 의뢰인들과 상담했는지, 어떤 전문 분야에서 가장 많이 상담했는지, 얼마나 많은 의뢰를 처리했는지 한눈에 파악할 수 있습니다.

[미국 내 법률 상담 트렌드 분석]

