

내가 사용한 데이터는 excel 데이터입니다.

이 데이터는 1년 동안 사람들의 문화예술관람 유형에 대한
 _____내용의 데이터로 구성되었습니다.

기본적인 내용을 탐색해보았습니다.

```
> df_pro
```

	통계분류(1)	통계분류(2)	2020...3	2020...4	2020...5	2020...6
1	통계분류(1)	통계분류(2)	표본수	영화관람	전시회관람	박물관 관람
2	전체	소계	10088.0	60.0	9.6	9.2
3	성별	남성	5003.0	60.2	8.0	8.5
4	<NA>	여성	5085.0	59.7	11.2	9.9
5	연령별	15~19세	619.0	80.2	13.0	11.1
6	<NA>	20대	1536.0	86.5	14.5	8.3
7	<NA>	30대	1644.0	77.5	13.2	14.7
8	<NA>	40대	1888.0	72.5	12.2	14.8
9	<NA>	50대	1858.0	58.1	7.9	7.2
10	<NA>	60대	1377.0	31.6	5.0	4.9
11	<NA>	70세이상	1166.0	11.3	1.3	2.2
12	학력별	초졸 이하	1101.0	18.6	2.2	3.6
13	<NA>	중졸	1067.0	38.7	5.5	5.2
14	<NA>	고졸	4039.0	59.2	7.5	7.5
15	<NA>	대졸이상	3881.0	77.3	14.8	13.6
16	동거가구원수별	1인	1902.0	42.9	5.0	4.4
17	<NA>	2인	2581.0	39.0	5.0	5.3
18	<NA>	3인이상	5605.0	68.8	11.6	11.1
19	혼인상태별	미혼	3041.0	82.4	13.2	8.8
20	<NA>	기혼	5713.0	55.6	9.2	10.3
21	<NA>	사별/이혼/기타	1334.0	27.6	2.5	3.3

19	혼인상태별	미혼	3041.0	82.4	13.2	8.8
20	<NA>	기혼	5713.0	55.6	9.2	10.3
21	<NA>	사별/이혼/기타	1334.0	27.6	2.5	3.3
22	가구주여부별	가구주	5073.0	50.0	6.9	8.0
23	<NA>	가구주 아님	5015.0	67.9	11.7	10.2
24	종사자지위별	상용근로자	2980.0	74.7	12.1	11.9
25	<NA>	임시근로자	518.0	57.0	11.0	5.8
26	<NA>	일용근로자	164.0	37.7	5.7	8.5
27	<NA>	고용원이 있는 자영업자	395.0	63.4	10.6	7.8
28	<NA>	고용원이 없는 자영업자	1755.0	48.8	6.0	6.6
29	<NA>	무급가족종사자	173.0	33.3	4.9	6.3
30	<NA>	기타종사자	59.0	57.0	15.9	11.9
31	<NA>	해당없음(무직)	4044.0	56.5	9.4	9.2
32	가구소득별	100만원 미만	980.0	18.7	2.3	1.7
33	<NA>	100~200만원	923.0	24.1	3.7	4.2
34	<NA>	200~300만원	1405.0	47.5	5.1	5.9
35	<NA>	300~400만원	1824.0	60.4	8.6	9.1
36	<NA>	400~500만원	1782.0	67.9	10.5	10.4
37	<NA>	500~600만원	1407.0	70.0	10.6	10.9
38	<NA>	600만원 이상	1767.0	77.0	15.6	13.0
39	지역규모별	대도시	4373.0	66.0	9.4	9.0
40	<NA>	중소도시	3776.0	61.2	11.0	10.3
41	<NA>	읍면지역	1939.0	44.0	7.3	7.4
42	권역별	수도권	4895.0	61.1	10.8	9.6
43	<NA>	강원/제주권	440.0	52.0	8.4	5.7

46	<NA>	대경관	1047.0	61.9	5.4	4.7
47	<NA>	동남권	1562.0	55.3	8.9	7.9
48	17개 시도별	서울	1931.0	63.6	9.1	8.2
49	<NA>	부산	681.0	62.2	10.0	8.5
50	<NA>	대구	490.0	71.3	4.7	2.3
51	<NA>	인천	556.0	67.7	9.7	10.8
52	<NA>	광주	294.0	70.9	9.6	11.2
53	<NA>	대전	303.0	72.0	11.1	12.2
54	<NA>	울산	221.0	59.2	9.2	9.6
55	<NA>	세종	61.0	71.0	40.6	57.6
56	<NA>	경기	2408.0	57.8	12.3	10.4
57	<NA>	강원	320.0	42.2	5.7	3.9
58	<NA>	충북	321.0	64.7	11.3	17.1
59	<NA>	충남	430.0	58.9	7.9	8.3
60	<NA>	전북	375.0	54.2	8.0	13.0
61	<NA>	전남	360.0	49.6	4.4	3.5
62	<NA>	경북	557.0	53.4	6.0	6.9
63	<NA>	경남	660.0	46.8	7.6	6.8
64	<NA>	제주	120.0	74.8	15.0	10.0
65	장애여부별	장애등록	225.0	27.0	5.7	8.7
66	<NA>	미등록	233.0	31.5	7.9	7.1
67	<NA>	해당사항없음	9630.0	61.3	9.7	9.3
	2020...7	2020...8	2020...9	2020...10	2020...11	
	연극공연 관람	연예공연 관람	음악연주회 관람	전통예술공연 관람	무용공연 관람	
2	8.0	5.4	3.4	2.9	0.9	

5	9.2	7.9	4.3	1.9	0.6
6	14.9	9.3	4.7	2.0	1.5
7	12.1	7.0	4.2	2.4	1.1
8	8.4	5.0	4.5	3.3	1.1
9	5.5	4.5	3.4	3.3	0.9
10	3.3	3.5	1.5	2.8	0.6
11	1.7	1.0	1.1	4.5	0.3
12	2.6	1.7	1.4	4.5	0.5
13	3.8	4.7	2.2	3.1	0.1
14	6.5	4.8	2.7	2.6	0.9
15	11.9	7.0	5.0	2.8	1.3
16	4.5	3.4	1.5	1.8	0.7
17	4.7	4.1	2.2	3.6	0.4
18	9.4	6.0	4.0	2.9	1.1
19	12.3	9.0	4.3	1.8	1.2
20	6.8	4.4	3.3	3.4	0.9
21	3.4	1.9	1.3	3.2	0.6
22	5.4	3.7	2.6	3.3	0.9
23	10.0	6.7	4.0	2.6	1.0
24	11.0	7.0	4.4	2.7	1.3
25	6.9	5.7	4.1	3.1	1.4
26	3.8	5.8	4.1	3.3	1.4
27	5.9	6.6	3.0	3.5	1.7
28	4.6	3.4	1.5	2.7	0.8
29	2.5	3.0	1.3	5.4	0.5
30	11.1	5.9	9.2	7.6	0.0

32	1.4	1.7	0.2	1.5	0.3
33	2.4	2.5	1.9	3.7	0.3
34	5.0	4.4	2.0	2.8	0.8
35	6.2	3.3	3.5	3.5	0.7
36	7.8	5.4	3.1	2.1	1.0
37	10.0	5.2	3.3	2.7	1.2
38	13.6	9.7	6.0	3.6	1.4
39	8.6	5.6	3.2	2.0	1.0
40	9.0	5.8	4.2	3.6	0.8
41	4.5	3.8	2.3	3.7	1.0
42	10.7	6.8	4.4	3.1	1.0
43	4.2	3.3	3.6	4.0	1.7
44	6.8	4.1	2.0	2.3	0.7
45	3.2	5.2	2.5	4.6	0.7
46	8.6	2.4	1.9	2.6	1.2
47	3.2	3.9	2.7	1.6	0.6
48	11.3	6.4	4.1	2.2	1.2
49	3.4	3.7	2.3	1.3	0.3
50	10.9	1.7	1.4	0.7	0.7
51	5.6	8.6	2.8	1.7	0.8
52	7.1	8.2	5.0	5.1	0.8
53	8.8	3.8	2.9	1.4	2.4
54	2.0	3.5	3.1	1.3	0.6
55	10.5	5.8	0.0	2.2	1.4
56	11.4	6.8	4.9	4.1	1.0
57	4.7	2.1	4.1	2.5	2.5

58	8.3	4.7	2.3	2.2	0.0
59	3.8	3.6	1.6	2.9	0.0
60	2.2	5.2	1.5	7.3	0.6
61	1.1	2.7	1.4	1.5	0.7
62	6.5	2.9	2.4	4.2	1.6
63	3.3	4.3	2.9	2.1	0.8
64	3.1	6.0	2.5	7.5	0.0
65	7.4	3.2	3.6	8.7	0.7
66	3.5	1.6	0.3	0.4	0.7
67	8.1	5.5	3.5	2.9	0.9

67 개의 행과 11 개의 변수들로 구성되었으며,

모든 변수 -- 통계분류(1) 통계분류(2) 2020...3 2020...4 2020...5 2020...6 2020...7
2020...8 2020...9 2020...10 2020...11 변수는 문자형으로 구성되어 있습니다.

summary함수로 기본적인 통계값을 확인해보면 결과는 아래와 같습니다.

-summary(df_pro)

통계분류(1)	통계분류(2)	2020...3	2020...4
Length:67	Length:67	Length:67	Length:67

Class :character	Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character	Mode :character
2020...5	2020...6	2020...7	2020...8
Length:67	Length:67	Length:67	Length:67
Class :character	Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character	Mode :character
2020...9	2020...10	2020...11	
Length:67	Length:67	Length:67	
Class :character	Class :character	Class :character	
Mode :character	Mode :character	Mode :character	

names함수로 변수명을 살펴보면 "통계분류(1)" "통계분류(2)" "2020...3" "2020...4"
"2020...5" "2020...6" "2020...7" "2020...8" "2020...9" "2020...10"
"2020...11" 이런 변수명을 가지고 있습니다.

dim함수로 67 11 이다.