

문자 타입

▣ 문자 타입

- 하나의 문자를 작은따옴표(')로 감싼 것을 문자 리터럴이라고 한다.
- 문자 리터럴은 유니코드로 변환되어 저장되는데,
유니코드는 세계 각국의 문자를 0~65535 숫자로 매핑한 국제 표준 규약이다.
- 자바는 이러한 유니코드를 저장할 수 있도록 char 타입을 제공한다.

```
char var1 = 'A';    //'A' 문자와 매핑되는 숫자: 65로 대입  
char var3 = '가';   //'가' 문자와 매핑되는 숫자: 44032로 대입
```

● char 변수에 값의 저장은 문자표를 참고해 정수로 저장되어 지며, 출력은 정수와 매핑된 문자로 출력됩니다.

```
char c = 65;        //10진수 65와 매핑되는 문자: 'A'  
char c = 0x0041;    //16진수 0x0041과 매핑되는 문자: 'A'
```

```
char c = '';        //컴파일 에러  
char c = ' ';       //공백 하나를 포함해서 초기화
```

```
char c4 = 44032;    //유니코드 직접 저장
```

[아스키코드]

DEC	HEX	OCT	Char	DEC	HEX	OCT	Char	DEC	HEX	OCT	Char
0	00	000	Ctrl-@ NUL	43	2B	053	+	86	56	126	V
1	01	001	Ctrl-A SOH	44	2C	054	,	87	57	127	W
2	02	002	Ctrl-B STX	45	2D	055	-	88	58	130	X
3	03	003	Ctrl-C ETX	46	2E	056	.	89	59	131	Y
4	04	004	Ctrl-D EOT	47	2F	057	/	90	5A	132	Z
5	05	005	Ctrl-E ENQ	48	30	060	0	91	5B	133	[
6	06	006	Ctrl-F ACK	49	31	061	1	92	5C	134	\
7	07	007	Ctrl-G BEL	50	32	062	2	93	5D	135]
8	08	010	Ctrl-H BS	51	33	063	3	94	5E	136	^
9	09	011	Ctrl-I HT	52	34	064	4	95	5F	137	_
10	0A	012	Ctrl-J LF	53	35	065	5	96	60	140	`
11	0B	013	Ctrl-K VT	54	36	066	6	97	61	141	a
12	0C	014	Ctrl-L FF	55	37	067	7	98	62	142	b
13	0D	015	Ctrl-M CR	56	38	070	8	99	63	143	c
14	0E	016	Ctrl-N SO	57	39	071	9	100	64	144	d
15	0F	017	Ctrl-O SI	58	3A	072	:	101	65	145	e
16	10	020	Ctrl-P DLE	59	3B	073	;	102	66	146	f
17	11	021	Ctrl-Q DC1	60	3C	074	<	103	67	147	g
18	12	022	Ctrl-R DC2	61	3D	075	=	104	68	150	h
19	13	023	Ctrl-S DC3	62	3E	076	>	105	69	151	i
20	14	024	Ctrl-T DC4	63	3F	077	?	106	6A	152	j
21	15	025	Ctrl-U NAK	64	40	100	@	107	6B	153	k
22	16	026	Ctrl-V SYN	65	41	101	A	108	6C	154	l
23	17	027	Ctrl-W ETB	66	42	102	B	109	6D	155	m
24	18	030	Ctrl-X CAN	67	43	103	C	110	6E	156	n
25	19	031	Ctrl-Y EM	68	44	104	D	111	6F	157	o
26	1A	032	Ctrl-Z SUB	69	45	105	E	112	70	160	p
27	1B	033	Ctrl-[ESC	70	46	106	F	113	71	161	q
28	1C	034	Ctrl-\ FS	71	47	107	G	114	72	162	r
29	1D	035	Ctrl-] GS	72	48	110	H	115	73	163	s
30	1E	036	Ctrl-^ RS	73	49	111	I	116	74	164	t
31	1F	037	Ctrl_ US	74	4A	112	J	117	75	165	u
32	20	040	Space	75	4B	113	K	118	76	166	v
33	21	041	!	76	4C	114	L	119	77	167	w
34	22	042	"	77	4D	115	M	120	78	170	x
35	23	043	#	78	4E	116	N	121	79	171	y
36	24	044	\$	79	4F	117	O	122	7A	172	z
37	25	045	%	80	50	120	P	123	7B	173	{
38	26	046	&	81	51	121	Q	124	7C	174	
39	27	047	'	82	52	122	R	125	7D	175	}
40	28	050	(83	53	123	S	126	7E	176	~

[유니코드]

U+AC00 to U+AD00

		0	1	2	3
UTF8: 234, 176, 128;	UNICODE: AC0	가	각	갈	갸
UTF8: 234, 176, 144;	UNICODE: AC1	감	갑	값	갇
UTF8: 234, 176, 160;	UNICODE: AC2	갸	갯	갼	갾
UTF8: 234, 176, 176;	UNICODE: AC3	갸	갻	갼	갽
UTF8: 234, 177, 128;	UNICODE: AC4	갈	갇	갸	갹
UTF8: 234, 177, 144;	UNICODE: AC5	갸	갻	갼	갽
UTF8: 234, 177, 160;	UNICODE: AC6	갸	갻	갼	갽
UTF8: 234, 177, 176;	UNICODE: AC7	거	걱	귀	견
UTF8: 234, 178, 128;	UNICODE: AC8	검	갑	견	건
UTF8: 234, 178, 144;	UNICODE: AC9	갸	갻	갼	갽
UTF8: 234, 178, 160;	UNICODE: ACA	갸	갻	갼	갽
UTF8: 234, 178, 176;	UNICODE: ACB	결	겉	겉	겉
UTF8: 234, 179, 128;	UNICODE: ACC	결	겉	겉	겉
UTF8: 234, 179, 144;	UNICODE: ACD	겉	겉	겉	겉
UTF8: 234, 179, 160;	UNICODE: ACE	고	곡	곡	곡
UTF8: 234, 179, 176;	UNICODE: ACF	곰	곰	곰	곰
		0	1	2	3
기타 등 등					

한글 '가' 자는
유니코드에서 16진수로 AC00(10진수의 44032)

<예>

```
char var1 = 0xAC00; //16진수
char var3 = 44032; //10진수
char var4 = '가'; //문자
char var2 = '\uAC00'; //유니코드
```