



# GALLUP 데이터베이스

Gallup 데이터베이스는 한국갤럽조사연구소가 제공하는 국내 최대의 본격적인 조사정보은행(Research Information Data Bank)이다. 정치, 경제, 사회조사로부터 기업경영 및 마케팅의 핵심정보인 국민의 생활지표(Life Style Index)에 이르기까지 광범위하고 다양한 주제가 교차분석형태로 자세하게 제공된다.

그 주제가 다양한 만큼 메뉴검색으로는 원하는 자료를 골라 내기가 쉽지 않을 것이다. 때문에 Gallup 데이터베이스에서는 명령어검색을 제공한다. 키워드 기본검색으로부터 부울리안 검색, 절단 검색(전방일치검색), 필드 제한검색에 이르기까지 각각의 검색방법은 세밀한 단위의 정보검색을 가능하게 한다. 따라서 이용자는 자신이 원하는 정보를 빠르고 정확하게 얻을 수 있을 것이다.

Gallup 데이터베이스에서 사용하는 주요 명령어는 다음의 4가지이다.

명령어	의 미	형 식	예
Select	검색 수행	S 키워드?/필드명	s 직장인/re
Display	결과화면 출력	D 집합번호(set번호)	and 여가?/de
TRace	검색식 열람	TR	d1tr
Expand	색인어 열람	E 키워드	e 직업

각각에 대해 자세히 살펴보도록 하자.

## 1. 검색 (Select)

① 기본검색 : 모든 검색 필드를 대상으로 하여 키워드가 포함된 모든 레코드를 검색한다.

(예) Gallup> S 강

위 예는 '강'이라는 단어가 존재하는 자료를 검색한 경우이다. 해당 자료의 검색이 완료되면 아래와 같은 검색 결과가 나타난다.

번호	건수	검색식
S001	7	강/ALL

위의 내용은 데이터베이스내에 '강'이라는 단어가 존재하는 자료가 모두 7건이며 이는 S1(Set 1 즉, 1번 집합)에 모여 있음을 보여준다. '강'과 '강/ALL'은 같은 의미로 모든 검색 필드를 대상으로 검색한다.

② 부울리안검색 : and, or, ( )의 연산자를 이용하여 두 개 이상의 단어를 조



합검색하는 것으로 복합주제 검색시 유용하다. 'A and B'는 A와 B를 모두 포함하는 자료의 집합(교집합)을 말하고, 'A or B'는 A나 B중 어느 하나를 포함하는 자료의 집합(합집합)을 말한다. 부울연산자의 실행순서는 동일하나 괄호와 함께 사용했을 때는 괄호안의 것이 먼저 실행된다.

(예) Gallup> s (강 or 산) and 한국  
번호    건수    검색식

S001	7	강/ALL
S002	12	산/ALL
S003	19	(강/ALL OR 산/ALL)
S004	244	한국/ALL
S005	5	(강/ALL OR 산/ALL) AND 한국/ALL

위의 예는 '강'과 '산', '한국'이라는 세 단어를 모두 포함하는 자료를 검색한 경우이다. 시스템은 이 검색식을 분석하여 우선 괄호안의 각 단어에 대해 검색한다. 각 단어의 자료집합인 S1과 S2를 합집합한 결과로서 S3을 출력하고, 이어서 '한국'이라는 단어의 자료집합인 S4를 만든다. 그리고 이 S3과 S4의 교집합을 구해 최종적으로 집합 S5를 보여준다.

③ 절단검색(전방일치검색) : 단어의 우측을 절단하는 것으로 절단기호(전방일치기호:?) 앞에 오는 문자열로 시작하는 단어가 포함된 자료를 검색한다

(예) Gallup> s 한국?  
번호    건수    검색식

S001	1202	한국?/ALL
------	------	---------

위의 예는 한국으로 시작되는 모든 단어를 검색한 경우이다. '한국'과 앞부분이 일치하는 단어로 '한국인', '한국사람', '한국전쟁' 등이 있을 때 '한국' 뒤에 절단기호(?)를 사용하면 '한국'을 포함하여 '한국'으로 시작되는 모든 단어를 포함한 자료를 검색하게 된다.

④ 제한검색 : 특정 필드를 접미어로 제한하여 검색한다. 즉, 모든 필드를 검색 대상으로 하는 것이 아니라, 특정 필드만을 대상으로 제한하여 검색한다.

(예) Gallup> s (강 or 산)/ti

번호	건수	검색식
S001	3	강/TI
S002	10	산/TI
S003	13	(강 OR 산)/TI

위의 예는 제목을 나타내는 접미어인 '/ti'를 사용하여 제한검색한 경우이다. 이때 검색 대상은 제목 필드로 제한되며, 따라서 제목에서 '강' 또는 '산'을 포함하는 자료를 검색하게 된다. 제한할 수 있는 접미어는 아래와 같다. 색인어라 함은 해당 레코드의 주제를 나타내는 어휘를 말한다.

필드	의미
/TI	제목(Title)
/SR	응답항목(Survey Result)
/AR	조사지역(Area)
/IM	조사방법(Interview Method)
/QT	질문내용(QuesTion)
/DE	색인어(DEscriptor)
/RE	응답자(REspondent)
/ST	조사시기(Survey Term)

## 2. 결과화면 출력 (Display)

검색한 자료에 대한 내용을 보고자 할 때 사용한다.

(예) Gallup> d 1

위의 예는 Select 명령어로 검색된 1번집합(S1)을 출력하는 경우이다. 이 때는 해당자료의 조사시기와 제목이 출력된다. 여기에서 각각의 조사자료에 대해 자세한 내용을 보고자 할 경우에는 해당자료의 번호를 선택하면 된다. 제목, 질문, 전체응답률, 조사개요, 비교분석자료, 인구특성별 응답률이 출력된다. 다음자료를 보고자 할 경우에는 'f' 명령어를 사용하면 된다.

## 3. 검색식 열람 (TRace)

Select 명령어를 사용해 형성된 집합번호들을 열람하는 것을 말한다. 검색식 열람을 끝내고자 할 경우 '\* 계속 <Y/n> :'에 'n'을 입력하면 'Gallup>' 프롬프트 상태로 돌아간다.

(예) Gallup> tr

번호	건수	검색식
S001	7	강/ALL
S002	4	바다/TI
S003	37	동양/DE
...		
S008	36	S7 OR 한국

위의 예에서 'tr' 명령어를 사용하면 Gallup DB에 접속한 후 형성된 검색식의 내용을 처음부터 차례로 보여준다.

## 4. 색인어 열람 (Expand)

특정 색인어와 인접한 단어들을 열람하는 것을 말한다. 단어의 정확한 표기나 그 단어의 인접 색인어를 알고자 할 때 사용할 수 있다.

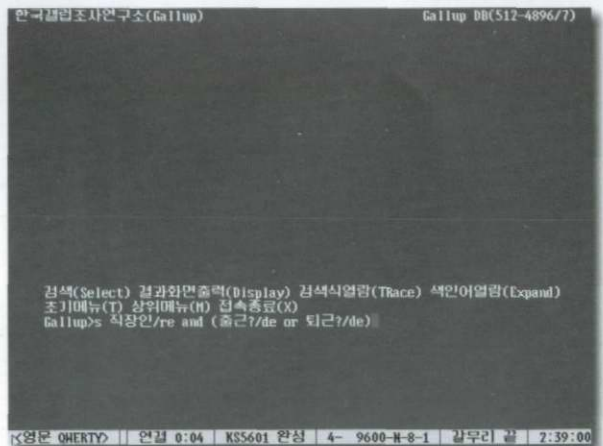
한번에 99개의 인접색인어를 보여 준다. 색인어 열람을 끝내고자 할 경우 '\* 계속 <Y/n> :'에 'n'을 입력하면 'Gallup>' 프롬프트 상태로 돌아간다.

(예) Gallup> e 강

번호	건수	검색식
E01	7	강
E02	3	강물
E03	5	강원도
...		
E15	1	강화

이제, 검색을 해 보도록 하자  
직장인의 라이프스타일에 대해 알고 싶다. 출퇴근시간 및 식사시간, 여가활동, 직장생활 등. 다음과 같이 검색할 수 있을 것이다.

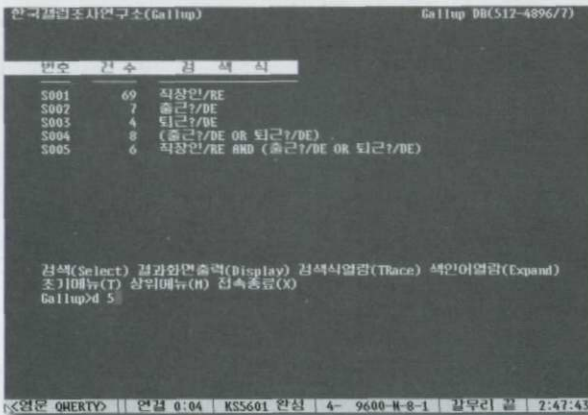
<그림 1>





그러면 ‘\* 작업 중입니다. 잠시 기다려 주십시오’ 라는 메시지가 나타난 후 다음과 같은 검색 결과 화면이 보여질 것이다

〈그림 2〉



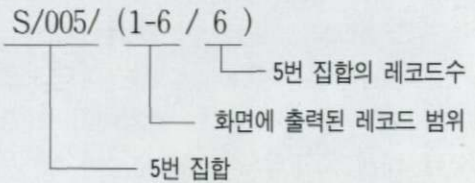
직장인의 출퇴근 관련 조사자료에 대한 검색 결과 화면이다. 응답대상자를 직장인으로 한정하기 위해서 /re 필드 제한 검색을 하였다. 또한 출근 및 퇴근 관련 자료를 얻고자 하는 것이므로 ‘출근?’ 과 ‘퇴근?’ 을 or로 조합하였다. 이 때 주의할 점은 ‘출근?/DE’ 과 ‘퇴근?/DE’ 의 or 조합이 ‘직장인/RE’ 과의 and 조합에 앞서 연산되어야 하므로 괄호를 사용하여야 한다는 것이다. 출근, 퇴근에 ‘?’ 를 이용, 절단검색을 함으로써

〈그림 3〉



출근, 출근시간 등 출근과 관련된 여러 색인어를 포함하도록 하였다. ‘S001, S002...’ 는 집합번호를 의미한다. Display 명령어를 이용하여 검색된 집합에 수록된 내용을 본다. ‘d 5’ 의 5는 5번 집합을 의미한다.

제목 출력 화면이다. ‘S/005/ (1-6 / 6)’ 의 의미는 다음과 같다.



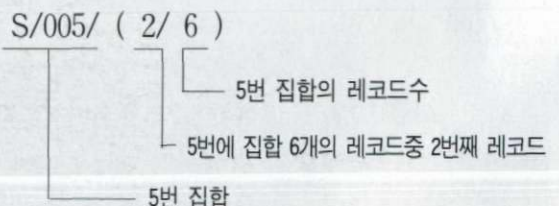
즉, 5번 집합 총 6개의 레코드중 1-6번이 화면에 출력되고 있다는 의미이다.

‘2’ 를 선택하면 2번 레코드의 내용을 볼 수 있다. 전체응답률, 조사개요, 비교분석자료, 인구 특성별응답률이 포함되어 있다.

〈그림 4〉



전체응답률 화면이다. ‘S/005/ ( 2/6 )’ 의 의미는 다음과 같다.



즉, 5번 집합 총 6개의 레코드중 2번째 레코드가 화면에 출력되고 있다는 의미이다.

2번 레코드의 다음 화면을 보고 싶으면 enter를, 3번째 레코드를 보고 싶으면 'f'를 입력하면 된다.

〈그림 5〉



조사개요 화면이다. 조사자료에서는 조사개요가 매우 중요하다.

같은 조사라도 조사지역, 조사대상, 표본크기, 표본추출법, 조사방법, 조사시기 등에 따라 결과가 달라질 수 있기 때문이다. 따라서 조사자료를 볼 때에는 항상 조사개요를 같이 읽어 주어야 한다.

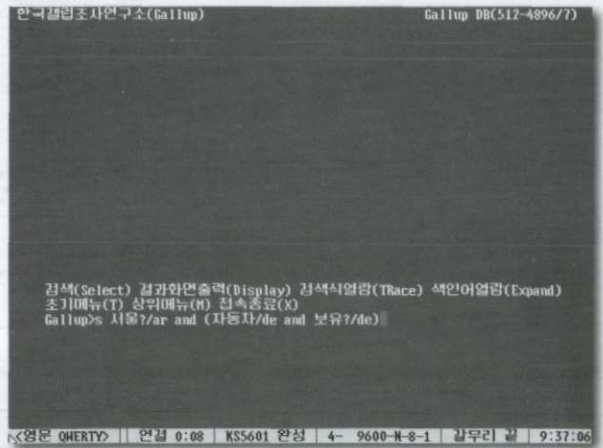
특히 조사결과를 해석할 때 명심하여야 할 것이 표본오차이다. 일반적으로 표본오차는 표본크기에 따라 달라진다.

위의 경우와 같이 비확률 표본추출의 경우에는 표본오차를 계산할 수 없다.

두번째 검색을 해보도록 하자.

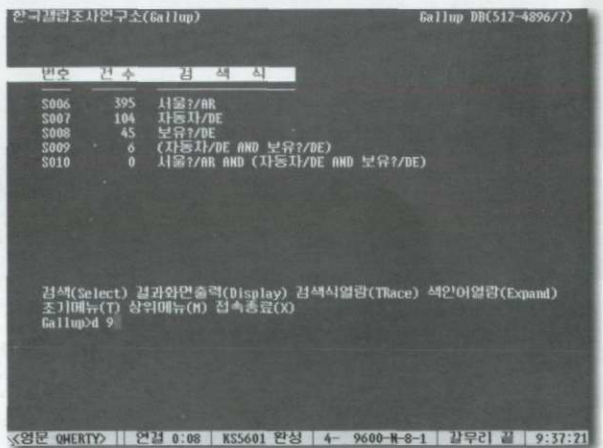
서울시 거주자의 자동차 보유상태에 대해 알고 싶다. 자동차 보유율, 자동차 보유대수 등. 다음과 같이 검색할 수 있을 것이다.

〈그림 6〉



'서울'로 지역을 제한하였다. 이때 서울, 서울시, 서울특별시 등, 서울을 나타내는 모든 단어를 포괄하기 위해서 '?'를 사용하였다. 또한 자동차 보유상태에 대한 것이므로 '자동차'와 '보유?'를 색인어 필드에서 검색하여 and 조합하였다. 역시 보유상태, 보유율, 보유대수, 보유수 등을 포괄하기 위해서 '?'를 사용하였다. \*작업 중입니다. 잠시 기다려 주십시오'라는 메시지가 나타난 후 다음과 같은 검색결과 화면이 보여진다.

〈그림 7〉



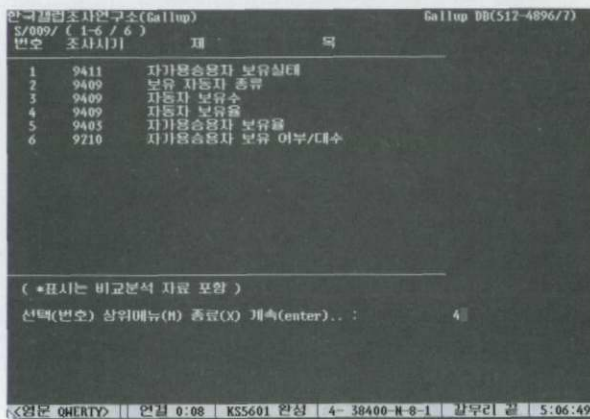
10번 집합이 0건이라는 것은 서울지역을 대상으로 한 자동차 보유상태 조사자료가 수록되어



있지 않다는 것이다. Gallup 데이터베이스는 전체응답률 외에 성별, 연령별, 교육수준별, 직업별, 지역별, 지역크기별, 생활수준별 등의 인구특성별 응답률을 제공한다. 따라서 서울지역 대상 조사 자료가 없다면 지역을 제한하지 않고 자동차 보유실태에 대한 검색을 하도록 한다. 지역별 응답률을 이용해 서울지역 조사자료를 얻을 수 있을 것이다. 이 때 주의할 점은 전체응답률보다 사례수가 적어짐으로써 오차한계가 더 커진다는 것이다.

9번 집합이 이에 해당되므로 다시 검색식을 입력하지 않고 'd 9'를 한다.

<그림 8>



9번 집합 6개의 레코드가 화면에 출력되었다. 여기에서 4번 레코드를 보고자 한다.

<그림 9>



9번 집합 총 6개의 레코드중 현재 4번 레코드를 보고 있다.

<그림 10>



조사개요를 읽도록 한다. 전국의 만 18세이상 남녀 1582명을 대상으로 '94년 9월 27일부터 94년 10월 10일까지 가구방문 개별면접에 의해 조사된 자료이다.

표본은 확률 표본추출법인 다단계 무작위 방법에 의해 선정되었으며, 표본오차는 95% 신뢰수준일 때  $\pm 2.5\%$ 이다. '95% 신뢰수준'이라는 의미는 다음과 같다. 즉, 똑같은 표본추출과정과 동일한 면접원, 동일한 질문지를 사용한다고 가정할 때, 같은 시간에 반복하여 표본추출해서 100번 조사할 경우 95번은 그 범위내에서 같은 결과(응답률)를 얻을 수 있다는 것이다.

따라서 전체응답률 '소유 49.0%'라 하는 것은, 똑같은 조사를 100번 하였을 경우 그중 95번의 조사에서 응답률이 46.5 - 51.5%내에 나타난다는 것을 말한다.



<그림 11>

한국갤럽조사연구소(Gallup) GALLUP DB(512-4896/7)

[ S/009/ ( 4/6 ) 자동차 보유율]

	# 성 별(%)	
	남자	여자
사례수(명)	(802)	(780)
소유	49.4	48.6
비소유	50.6	51.4
	100.0	100.0

\* 상위메뉴(H) 초기메뉴(T) 다음자료(F) 종료(X) 계속(enter) : █

<엔터 QHERTY> | 연결 0:09 KSS601 완성 4- 38400-N-8-1 갈무리 끝 5:07:01

<그림 14>

한국갤럽조사연구소(Gallup) GALLUP DB(512-4896/7)

[ S/009/ ( 4/6 ) 자동차 보유율]

	# 직 업 별(%)						
	농/임/어업	자영업	블루칼라	화이트칼라	가정주부	학생	무직/기타
사례수(명)	(200)	(215)	(233)	(214)	(483)	(110)	(127)
소유	33.2	69.5	34.7	57.7	53.3	55.3	79.2
비소유	66.8	30.5	65.3	42.3	46.7	44.7	70.8
	100.0	100.0	100.0	100.0	100.0	100.0	100.0

\* 상위메뉴(H) 초기메뉴(T) 다음자료(F) 종료(X) 계속(enter) : █

<엔터 QHERTY> | 연결 0:09 KSS601 완성 4- 38400-N-8-1 갈무리 끝 5:07:07

<그림 12>

한국갤럽조사연구소(Gallup) GALLUP DB(512-4896/7)

[ S/009/ ( 4/6 ) 자동차 보유율]

	# 연 령 별(%)				
	18-25세	26-30세	31-40세	41-50세	51-59세
사례수(명)	(424)	(266)	(423)	(277)	(192)
소유	47.7	56.0	54.4	49.8	29.4
비소유	52.3	44.0	45.6	50.2	70.6
	100.0	100.0	100.0	100.0	100.0

\* 상위메뉴(H) 초기메뉴(T) 다음자료(F) 종료(X) 계속(enter) : █

<엔터 QHERTY> | 연결 0:09 KSS601 완성 4- 38400-N-8-1 갈무리 끝 5:07:03

<그림 15>

한국갤럽조사연구소(Gallup) GALLUP DB(512-4896/7)

[ S/009/ ( 4/6 ) 자동차 보유율]

	# 지 역 별(%)						
	서울	경기	강원	충청	전라	경북	경남
사례수(명)	(421)	(332)	(56)	(158)	(191)	(169)	(255)
소유	51.5	56.1	45.4	40.5	33.1	54.5	50.0
비소유	48.5	43.9	54.6	59.5	66.9	45.5	50.0
	100.0	100.0	100.0	100.0	100.0	100.0	100.0

\* 상위메뉴(H) 초기메뉴(T) 다음자료(F) 종료(X) 계속(enter) : █

<엔터 QHERTY> | 연결 0:09 KSS601 완성 4- 38400-N-8-1 갈무리 끝 5:07:10

<그림 13>

한국갤럽조사연구소(Gallup) GALLUP DB(512-4896/7)

[ S/009/ ( 4/6 ) 자동차 보유율]

	# 교 육 수 준 별(%)			
	국졸이하	중졸	고졸	대재이상
사례수(명)	(223)	(213)	(739)	(406)
소유	24.1	44.3	51.3	61.0
비소유	75.9	55.7	48.7	39.0
	100.0	100.0	100.0	100.0

\* 상위메뉴(H) 초기메뉴(T) 다음자료(F) 종료(X) 계속(enter) : █

<엔터 QHERTY> | 연결 0:09 KSS601 완성 4- 38400-N-8-1 갈무리 끝 5:07:05

지역별 응답률이다. 보이는 바와 같이 서울지역의 응답률이 나온다. 따라서 서울시 거주자의 자동차 보유상태에 대해 알고 싶다면 이 응답률을 이용하면 된다. 단, 주의할 점은 앞서도 언급했듯이 각 계층별 응답률의 경우 표본수가 줄어들어 표본오차가 더 커진다는 것이다. 표본오차의 한계에 대한 자세한 내용은 Gallup 데이터베이스 접속전의 '11. 한국갤럽 소개' 란에 수록되어 있으니 참조하기 바란다.

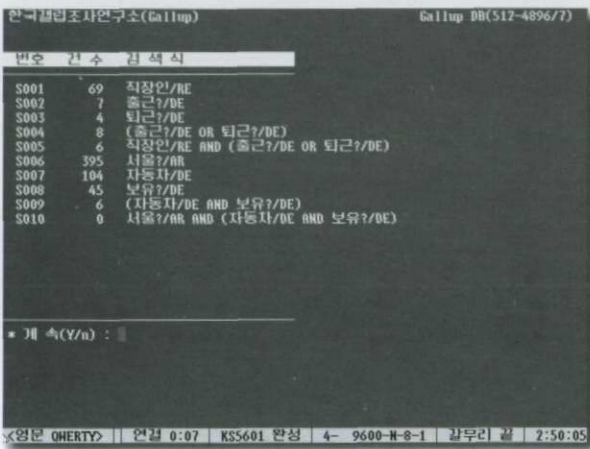
이렇듯 검색에 이용한 검색식들은 TRace 명령어를 이용하여 열람할 수 있다.



〈그림 16〉



〈그림 17〉

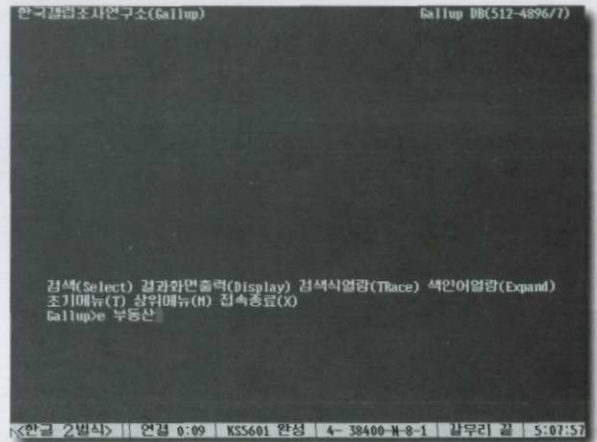


이때 보여지는 집합번호는 Display 명령어를 이용하여 출력할 수 있다. 즉, '출근?/DE' 라는 검색식을 다시 입력하지 않고도 'd 2' 명령을 사용함으로써 출근에 관한 자료를 볼 수 있다.

'e 부동산' 은 '부동산' 이라는 특정 색인어와 그에 인접한 단어들을 열람하는 명령이다. 색인어는 한번에 99개까지 볼 수 있으며, 그 이상 보고자 할 경우에는 Expand 명령어를 다시 사용하여야 한다. 이용자가 입력한 특정 색인어가 없을 때에는 그 다음 순서(자모순)의 색인어부터 화면에 출력된다.

다음은 Expand 명령어에 대한 것이다.

〈그림 18〉



〈그림 19〉



지금까지 Gallup 데이터베이스의 명령어검색 방법 및 검색 예제에 대해 살펴 보았다.

명령어검색은 검색의 효율과 빠른 검색결과를 위해서 익숙해져야 할 부분이다. 메뉴검색에 비해 복잡한 주제일지라도 검색식을 잘 작성하면 원하는 정보를 쉽게 얻을 수 있다.

주요 명령어인 Select, Display, TRace, Expand와 친숙해지면 Gallup 데이터베이스를 보다 다양하게 활용할 수 있을 것이다. **DrC**