

🌸 16. uniq 명령어

uniq: 중복된 라인을 제거하는 명령어

예제 1. 직업을 출력하는데 중복을 제거해서 출력하시오.

```
awk -F ',' '{print $3}' emp2.csv | sort -k 1 | uniq
```

=> 정렬을 먼저 해놓고 중복을 제거해야 중복이 제대로 제거된다.

예제 2. 부서번호를 출력하는데 중복을 제거해서 출력하시오.

```
awk -F ',' '{print $8}' emp2.csv | sort -nk 1 | uniq
```

예제 3. 직업, 직업별 인원수를 출력하시오.

```
awk -F ',' '{print $3}' emp2.csv | sort | uniq -c
```

예제 4. 위의 결과를 다시 출력하는데 직업별 인원수가 높은 것부터 출력하시오.

```
awk -F ',' '{print $3}' emp2.csv | sort | uniq -c | sort -nr
```

문제 1. 아래의 SQL을 리눅스 명령어로 수행하시오.

```
select job, count(*)
from emp
where job != 'SALESMAN'
group by job
order by count(*) desc;
```

```
awk -F ',' '$3!="SALESMAN" {print $3}' emp2.csv | sort | uniq -c | sort -nr
```

문제 2. 아래의 SQL을 리눅스 명령어로 수행하시오.

```
select deptno, count(*)
from emp
group by deptno
order by deptno asc;
```

```
awk -F ',' '{print $8}' emp2.csv | sort | uniq -c
```

문제:

다음 텍스트 파일에 uniq 명령어를 실행했을 때, 어떤 결과가 출력될까요?

```
apple
apple
banana
banana
banana
orange
orange
apple
```

a)

```
apple
banana
orange
apple
```

b)

```
apple
banana
orange
```

c)

```
apple
apple
banana
orange
apple
```

d)

```
apple
banana
banana
orange
orange
apple
```

B

문제: emp2.csv 파일에 있는 직책(job)별로 레코드의 수를 세는 리눅스 명령어 시퀀스는 무엇일까요?

```
7839,KING,PRESIDENT,,1981-11-17,5000,,10
7698,BLAKE,MANAGER,7839,1981-05-01,2850,,30
7782,CLARK,MANAGER,7839,1981-05-09,2450,,10
7566,JONES,MANAGER,7839,1981-04-01,2975,,20
...
```

- a) awk -F, '{print \$3}' emp2.csv | sort | uniq
- b) cat emp2.csv | cut -d, -f3 | sort | uniq -c
- c) awk -F, '{print \$2}' emp2.csv | sort | uniq -c
- d) awk -F, '{print \$3}' emp2.csv | sort | uniq -c

D