

LanguageManual VirtualColumns

- [Virtual Columns](#)
 - [Simple Examples](#)

Virtual Columns

Hive 0.8.0 provides support for two virtual columns:

One is `INPUT__FILE__NAME`, which is the input file's name for a mapper task.

the other is `BLOCK__OFFSET__INSIDE__FILE`, which is the current global file position.

For block compressed file, it is the current block's file offset, which is the current block's first byte's file offset.

Since Hive 0.8.0 the following virtual columns have been added:

- `ROW__OFFSET__INSIDE__BLOCK`
- `RAW__DATA__SIZE`
- `ROW__ID`
- `GROUPING__ID`

It is important to note, that all of the **virtual columns listed here cannot be used for any other purpose** (i.e. table creation with columns having a virtual column will fail with "SemanticException Error 10328: Invalid column name..")

Simple Examples

```
select INPUT__FILE__NAME, key, BLOCK__OFFSET__INSIDE__FILE from src;
```

```
select key, count(INPUT__FILE__NAME) from src group by key order by key;
```

```
select * from src where BLOCK__OFFSET__INSIDE__FILE > 12000 order by key;
```