Column Stores: Motivation

- Consider a table:
 - Foo (a INTEGER, b INTEGER, c VARCHAR(255), ...)
- And the query`:
 - SELECT a FROM Foo WHERE a > 10
- What happens with the previous record format in terms of the bytes that have to be read from the IO subsystem?

Column Stores: Motivation

- Store data "vertically"
- Contrast that with a "row-store" that stores all the attributes of a tuple/record contiguously
 - The previous record formats are "row stores"

111	212	lt wa	as a cold morning		111	212	It was a cold morning
222	222	War	m and sunny here		222	222	Warm and sunny here
333	232	Artic	winter conditions		333	232	Artic winter conditions
444	242	Tr	opical weather		444	242	Tropical weather
		•			• • File 1	• • File 2	File 3
	Each file is a set of pages. Columns can be stored in compressed form						

Column Stores: Motivation

- Are there any disadvantages associated with column stores?
 - 1. Updates are slower
 - 2. Retrieving back more than one attribute can be slower, e.g. Queries like SELECT * are slower