



# Formação Engenheiro de Dados

Colunas Ordenadas

# Colunas Ordenadas

- ◊ Conhecido também como Wide Column Store, Column Oriented Database etc.
- ◊ Dados armazenados em colunas, não em linhas
- ◊ Keyspace: equivalente a schema do relacional: agrupa
- ◊ Column families: equivalente a uma tabela do relacional: dados
- ◊ Column Family: várias linhas
- ◊ Cada linha pode conter diferentes números de colunas, não precisa haver uma equivalência
- ◊ Cada coluna contém uma chave/valor mais data e hora
- ◊ Cada linha tem uma chave unica

# Colunas Ordenadas

Família				
Colunas				
Chave		Nome	Cidade	Profissão
→	223454	Pedro	Novo Hamburgo	Engenheiro
Linhas		Nome	Cidade	Idade
→	456786	Maria	Boa Vista	26
	899775	Ana		Cientista

# Por quê?

- ◇ Imagine uma tabela com centenas de colunas e milhões de registros
- ◇ Normalmente, você não vai precisar de todas as colunas, ou seja:
  - ◇ ~~Select \* from tabela~~
  - ◇ Select col1,col2,col3 from tabela
- ◇ Porém mesmo havendo índices, o gerenciados de banco de dados vai precisar carregar todas as colunas com todos os dados!



# Modelo

- ◇ No modelo de colunas, cada coluna é armazenada em um arquivo fisicamente separado
- ◇ O gerenciador de banco de dados, precisa carregar apenas as colunas da consulta
- ◇ É um ótimo modelo para compressão (mais dados repetidos do que linhas), reduzindo espaço de armazenamento

Col1	Col2	Col3	Col4	Col5
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...

Col1	...	...	...	...	...
Col2	...	...	...	...	...
Col3	...	...	...	...	...
Col4	...	...	...	...	...
Col5	...	...	...	...	...