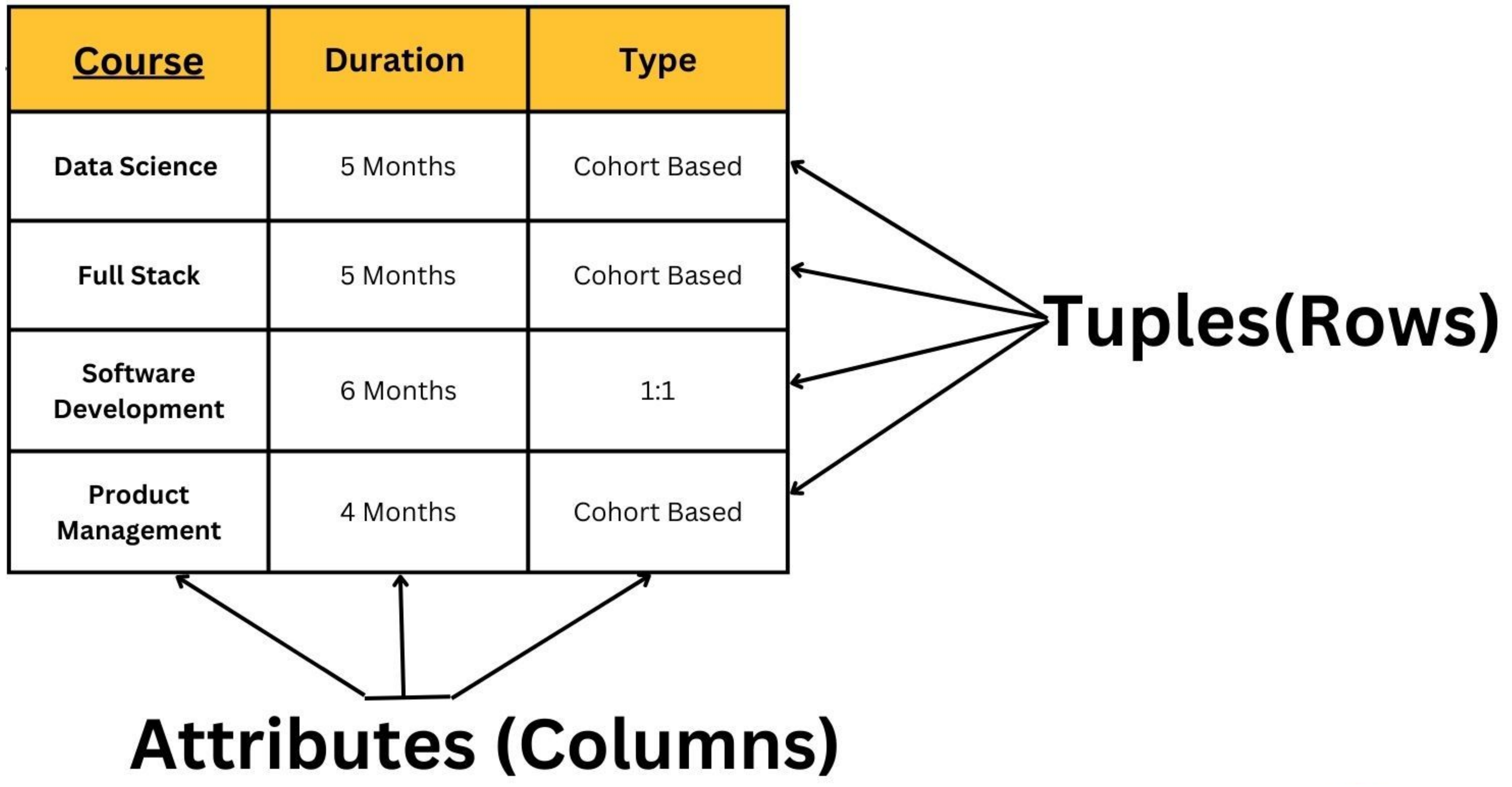
TECH VAULT

COLUMNAR DATABASES

PART (1)

AMR ELHELW

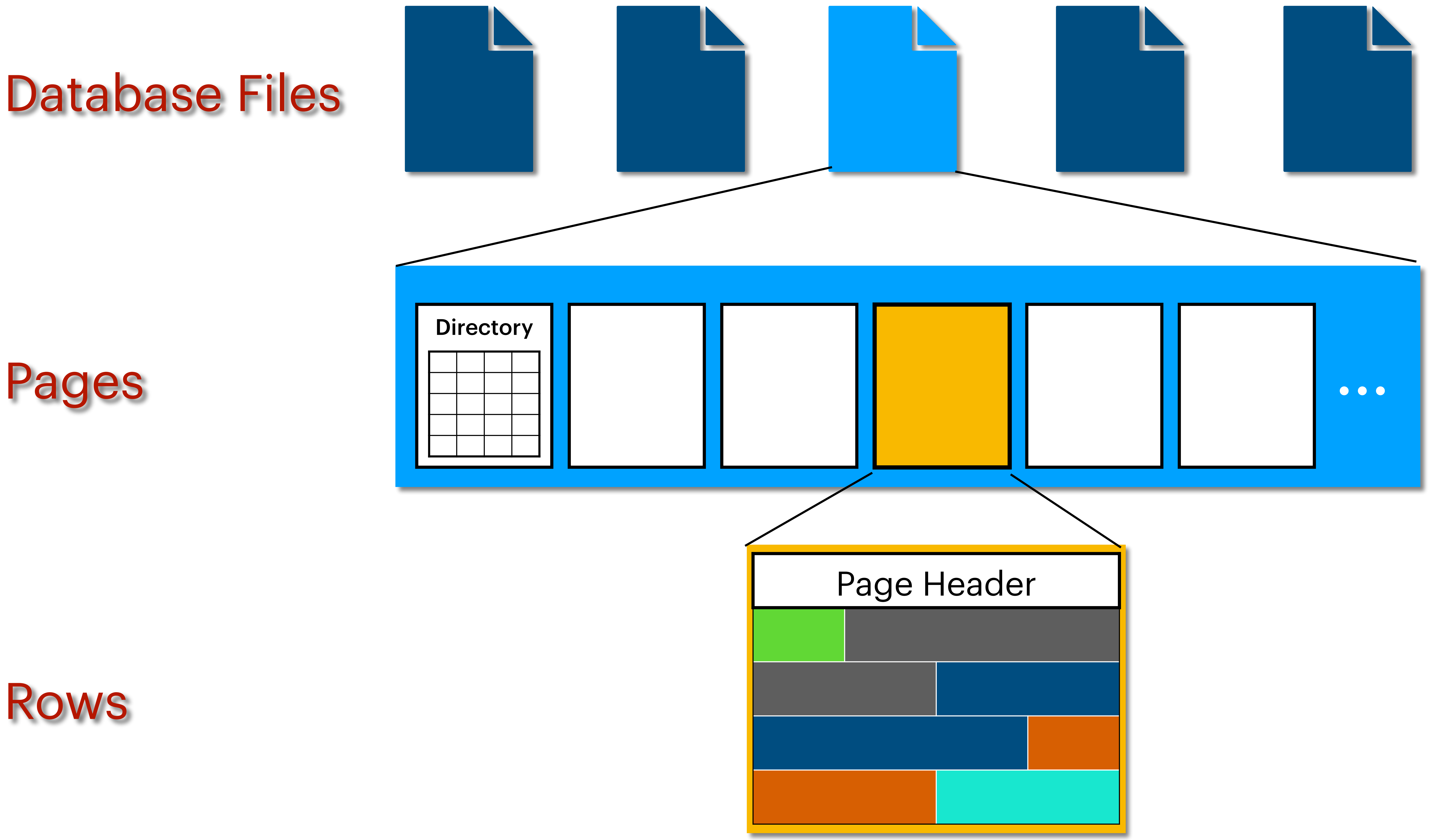
Relational Model



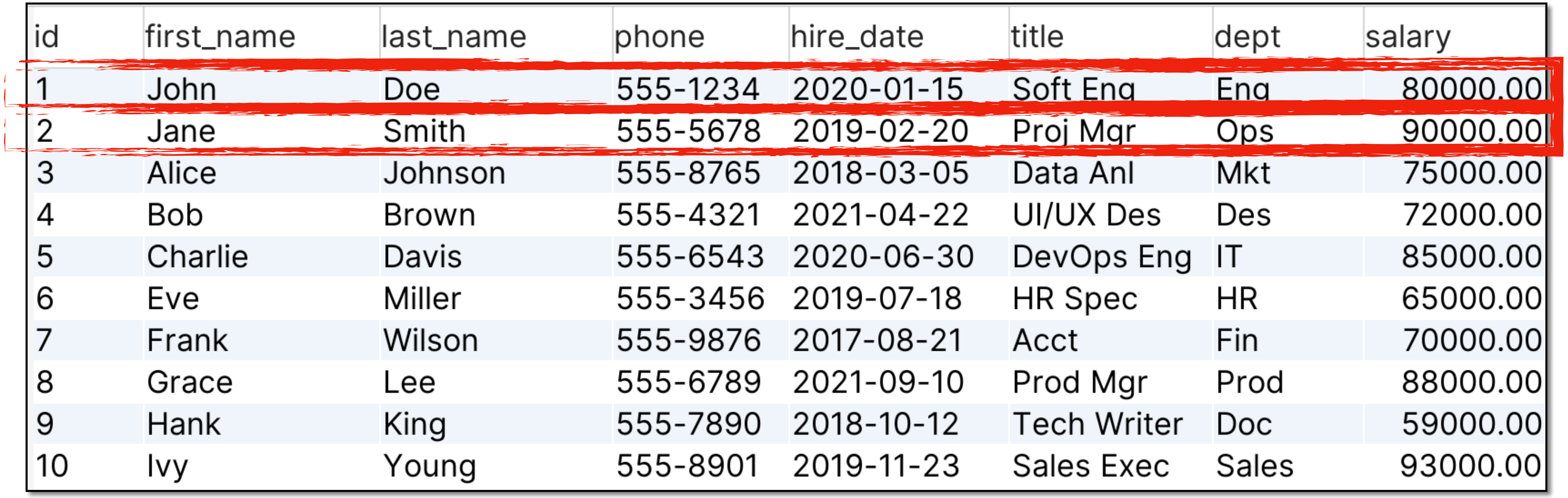
Course	Duration	Type
Data Science	5 Months	Cohort Based
Full Stack	5 Months	Cohort Based
Software Development	6 Months	1:1
Product Management	4 Months	Cohort Based

Amr Elhelw's TECH VAULT

Data Storage



Amr Elhelw's TECH VAULT



id	first_name	last_name	phone	hire_date	title	dept	salary
1	John	Doe	555-1234	2020-01-15	Soft Eng	Eng	80000.00
2	Jane	Smith	555-5678	2019-02-20	Proj Mgr	Ops	90000.00
3	Alice	Johnson	555-8765	2018-03-05	Data Anl	Mkt	75000.00
4	Bob	Brown	555-4321	2021-04-22	UI/UX Des	Des	72000.00
5	Charlie	Davis	555-6543	2020-06-30	DevOps Eng	IT	85000.00
6	Eve	Miller	555-3456	2019-07-18	HR Spec	HR	65000.00
7	Frank	Wilson	555-9876	2017-08-21	Acct	Fin	70000.00
8	Grace	Lee	555-6789	2021-09-10	Prod Mgr	Prod	88000.00
9	Hank	King	555-7890	2018-10-12	Tech Writer	Doc	59000.00
10	Ivy	Young	555-8901	2019-11-23	Sales Exec	Sales	93000.00

1	John	Doe	555-1234	2020-01-15	Soft Eng	Eng	80000.00
2	Jane	Smith	555-5678	2019-02-20	Proj Mgr	Ops	90000.00

3	Alice	Johnson	555-8765	2018-03-05	Data Anl	Mkt	75000.00
4	Bob	Brown	555-4321	2021-04-22	UI/UX Des	Des	72000.00

5	Charlie	Davis	555-6543	2020-06-30	DevOps Eng	IT	85000.00
6	Eve	Miller	555-3456	2019-07-18	HR Spec	HR	65000.00

7	Frank	Wilson	555-9876	2017-08-21	Acct	Fin	70000.00
8	Grace	Lee	555-6789	2021-09-10	Prod Mgr	Prod	88000.00

9	Hank	King	555-7890	2018-10-12	Tech Writer	Doc	59000.00
10	Ivy	Young	555-8901	2019-11-23	Sales Exec	Sales	93000.00

Read all data of product ‘P801’

Create a new order

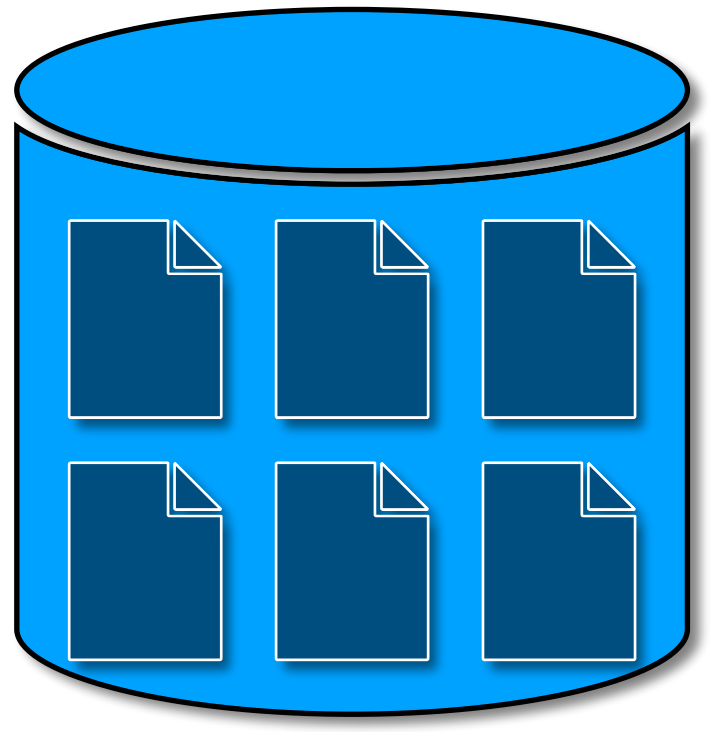
Return all employees who work in ‘Sales’ and hired after ‘2025-01-01’

Update salary of employee ‘156’ to 6,000

OLTP

Amr Elhelw’s TECH VAULT

```sql
SELECT first_name, last_name, hire_date
FROM employees
WHERE dept = 'HR'
```



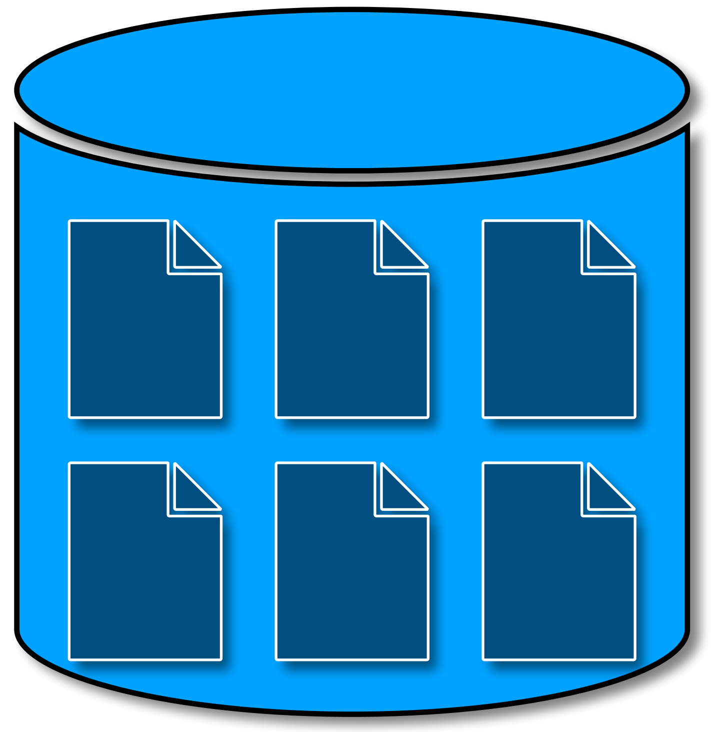
1	John	Doe	555-1234	2020-01-15	Soft Eng	Eng	80000.00
2	Jane	Smith	555-5678	2019-02-20	Proj Mgr	Ops	90000.00

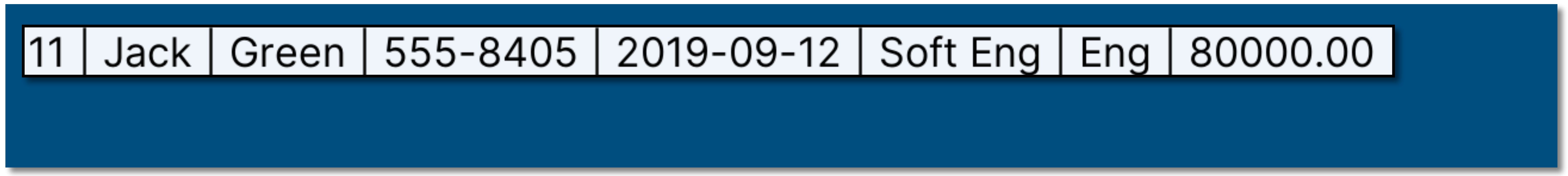
3	Alice	Johnson	555-8765	2018-03-05	Data Anl	Mkt	75000.00
4	Bob	Brown	555-4321	2021-04-22	UI/UX Des	Des	72000.00

5	Charlie	Davis	555-6543	2020-06-30	DevOps Eng	IT	85000.00
6	Eve	Miller	555-3456	2019-07-18	HR Spec	HR	65000.00

Amr Elhelw's
TECH
VAULT

```sql
INSERT INTO employees
VALUES (...)
```





11	Jack	Green	555-8405	2019-09-12	Soft Eng	Eng	80000.00

Amr Elhelw's TECH VAULT

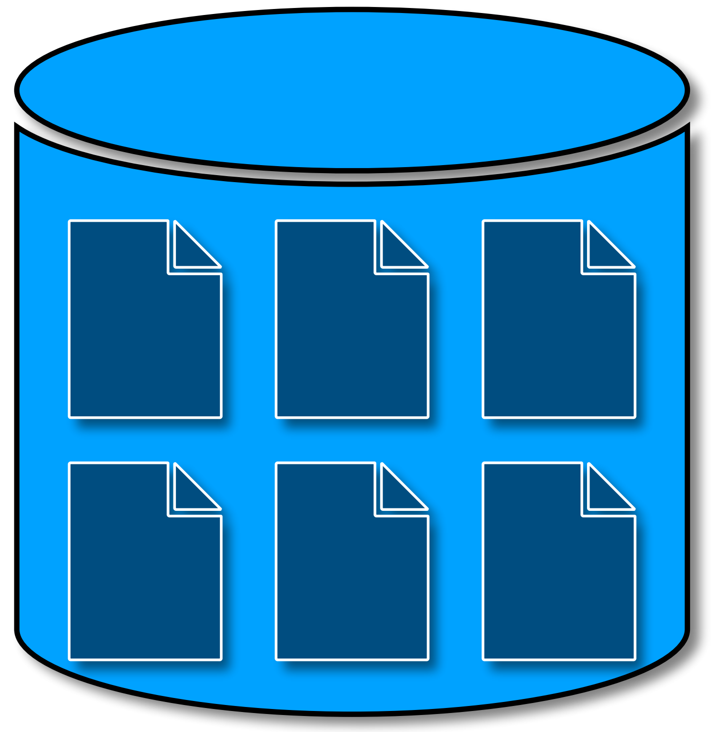
Return sum of all employees' salaries

Return highest salary in each department

"OLAP"

Amr Elhelw's TECH VAULT

```sql
SELECT dept, max(salary)
FROM employees
GROUP BY dept
```



1	John	Doe	555-1234	2020-01-15	Soft Eng	Eng	80000.00
2	Jane	Smith	555-5678	2019-02-20	Proj Mgr	Ops	90000.00

3	Alice	Johnson	555-8765	2018-03-05	Data Anl	Mkt	75000.00
4	Bob	Brown	555-4321	2021-04-22	UI/UX Des	Des	72000.00

5	Charlie	Davis	555-6543	2020-06-30	DevOps Eng	IT	85000.00
6	Eve	Miller	555-3456	2019-07-18	HR Spec	HR	65000.00

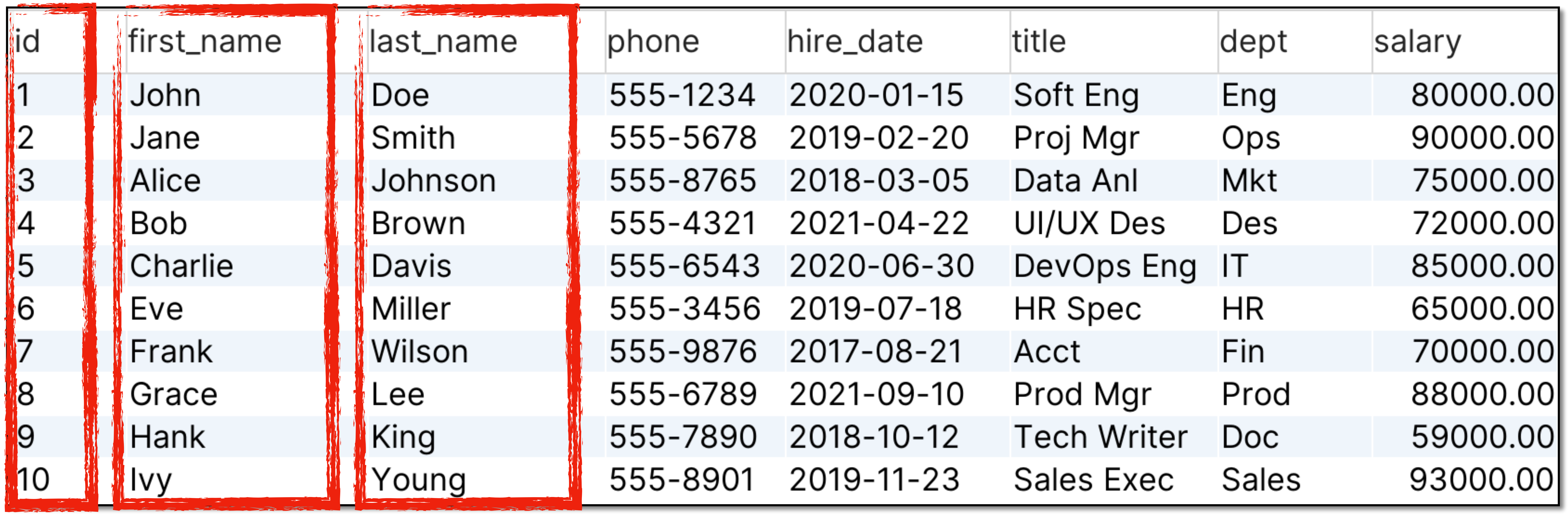
7	Frank	Wilson	555-9876	2017-08-21	Acct	Fin	70000.00
8	Grace	Lee	555-6789	2021-09-10	Prod Mgr	Prod	88000.00

9	Hank	King	555-7890	2018-10-12	Tech Writer	Doc	59000.00
10	Ivy	Young	555-8901	2019-11-23	Sales Exec	Sales	93000.00

Amr Elhelw's TECH VAULT

Row Stores

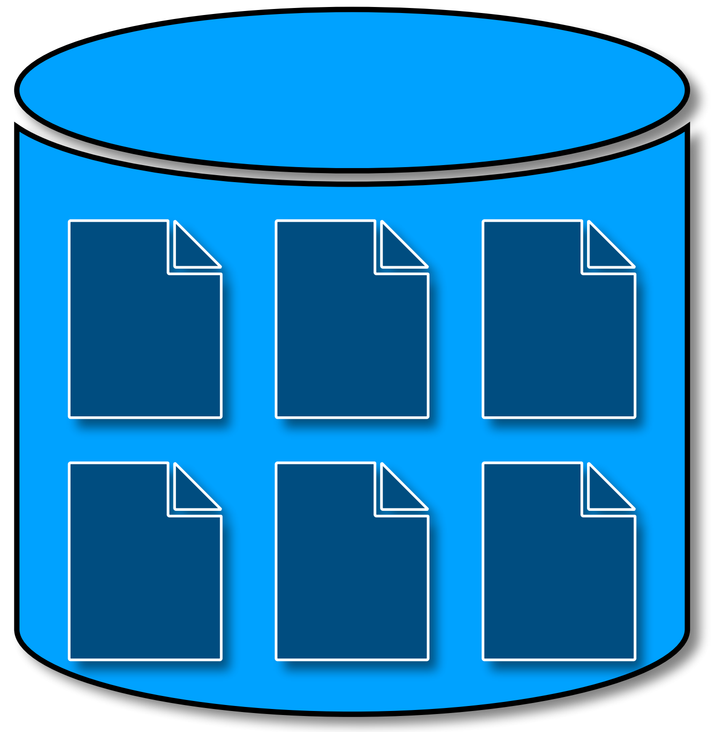
- Pros
  - *Good for queries that read a limited number of rows, with many attributes*
  - *Efficient writes (inserts, updates, and deletes)*
- Cons
  - *Not suitable for queries that need only a few columns from a large number of rows (e.g. analytical queries)*



id	first_name	last_name	phone	hire_date	title	dept	salary
1	John	Doe	555-1234	2020-01-15	Soft Eng	Eng	80000.00
2	Jane	Smith	555-5678	2019-02-20	Proj Mgr	Ops	90000.00
3	Alice	Johnson	555-8765	2018-03-05	Data Anl	Mkt	75000.00
4	Bob	Brown	555-4321	2021-04-22	UI/UX Des	Des	72000.00
5	Charlie	Davis	555-6543	2020-06-30	DevOps Eng	IT	85000.00
6	Eve	Miller	555-3456	2019-07-18	HR Spec	HR	65000.00
7	Frank	Wilson	555-9876	2017-08-21	Acct	Fin	70000.00
8	Grace	Lee	555-6789	2021-09-10	Prod Mgr	Prod	88000.00
9	Hank	King	555-7890	2018-10-12	Tech Writer	Doc	59000.00
10	Ivy	Young	555-8901	2019-11-23	Sales Exec	Sales	93000.00

1	2	3	4	5	6	7	8	9	10
John	Jane	Alice	Bob	Charlie	Eve	Frank	Grace	Hank	Ivy
Doe	Smith	Johnson	Brown	Davis	Miller	Wilson	Lee	King	Young
555-1234	555-5678	555-8765	555-4321	555-6543	555-3456	555-9876	555-6789	555-7890	555-8901
2020-01-15	2019-02-20	2018-03-05	2021-04-22	2020-06-30	2019-07-18	2017-08-21	2021-09-10	2018-10-12	2019-11-23
Soft Eng	Proj Mgr	Data Anl	UI/UX Des	DevOps Eng	HR Spec	Acct	Prod Mgr	Tech Writer	Sales Exec
Eng	Ops	Mkt	Des	IT	HR	Fin	Prod	Doc	Sales
80000.00	90000.00	75000.00	72000.00	85000.00	65000.00	70000.00	88000.00	59000.00	93000.00

```sql
SELECT dept, max(salary)
FROM employees
GROUP BY dept
```



Eng	Ops	Mkt	Des	IT	HR	Fin	Prod	Doc	Sales

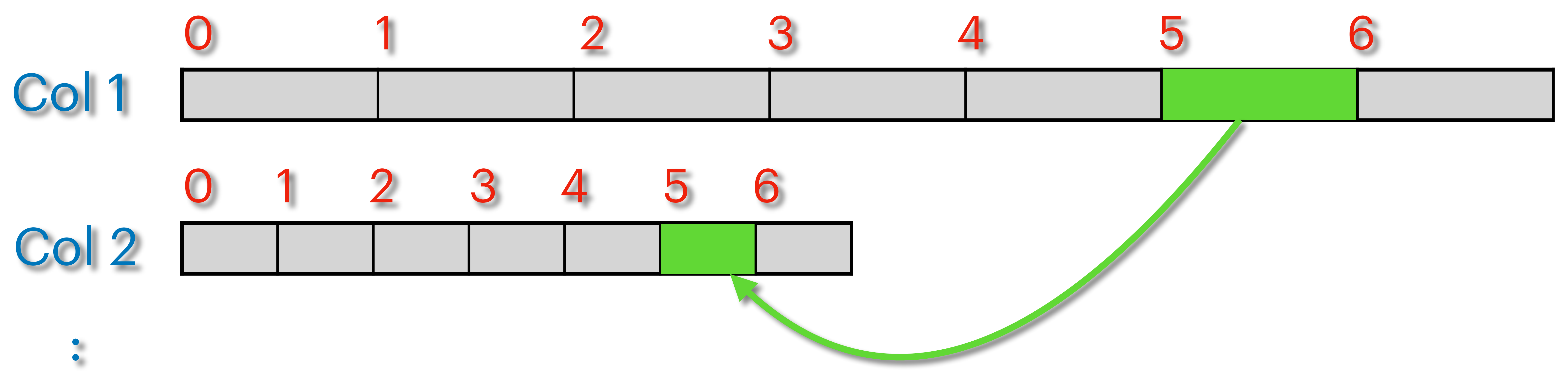
80000.00	90000.00	75000.00	72000.00	85000.00	65000.00	70000.00
88000.00	59000.00	93000.00				

How to identify values from the same row?

Amr Elhelw's TECH VAULT

Connecting values

(1) Offsets - for fixed length values

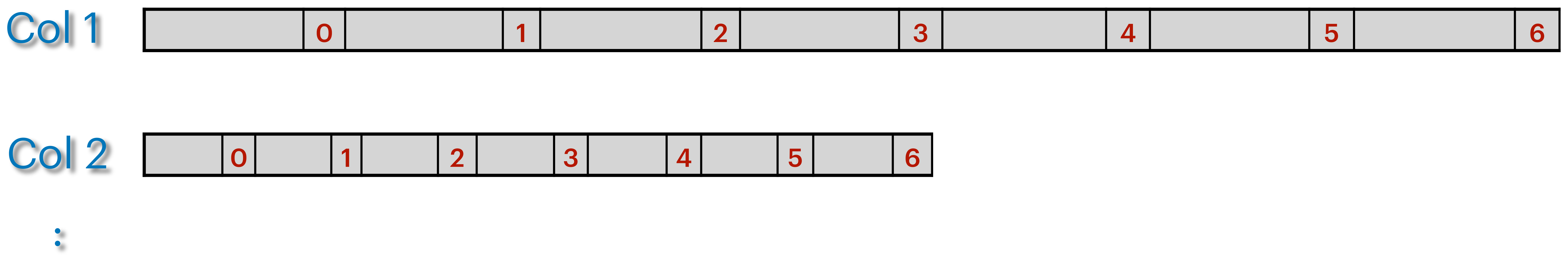


- **Advantages**
  - *Simple implementation*
  - *No additional storage needed*
- **Limitations**
  - *Only works with fixed-length data types*
  - *All columns must be stored in the same order*

Amr Elhelw's TECH VAULT

Connecting values

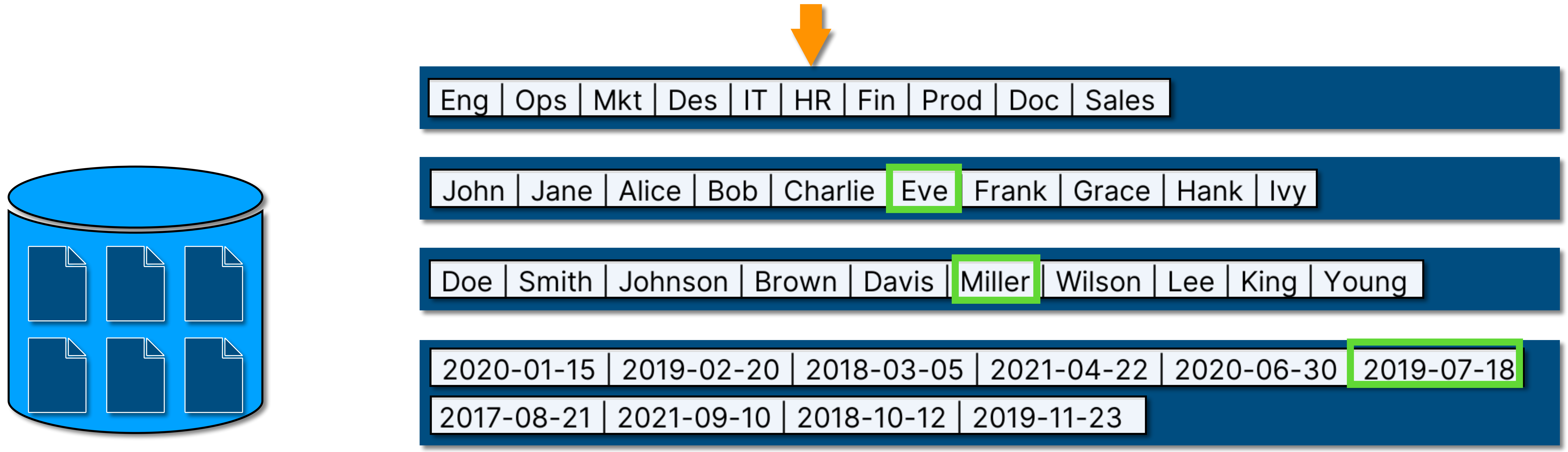
(2) Explicit row ids



- **Advantages**
  - *Any data type*
- **Limitations**
  - *Row ids are stored with every column - Additional storage (and I/O) needed*
  - *Limits data compression*

Amr Elhelw's TECH VAULT

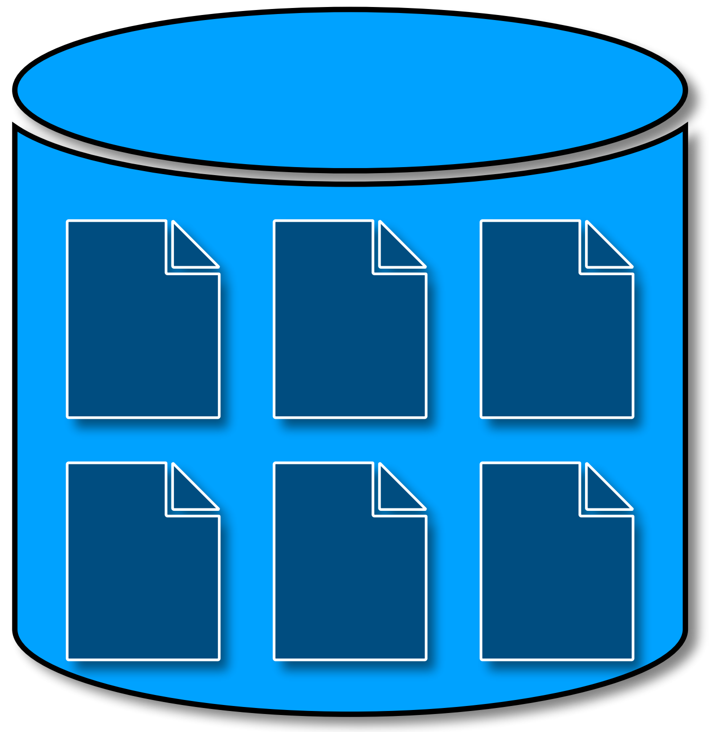
```
SELECT first_name, last_name, hire_date
FROM employees
WHERE dept = 'HR'
```

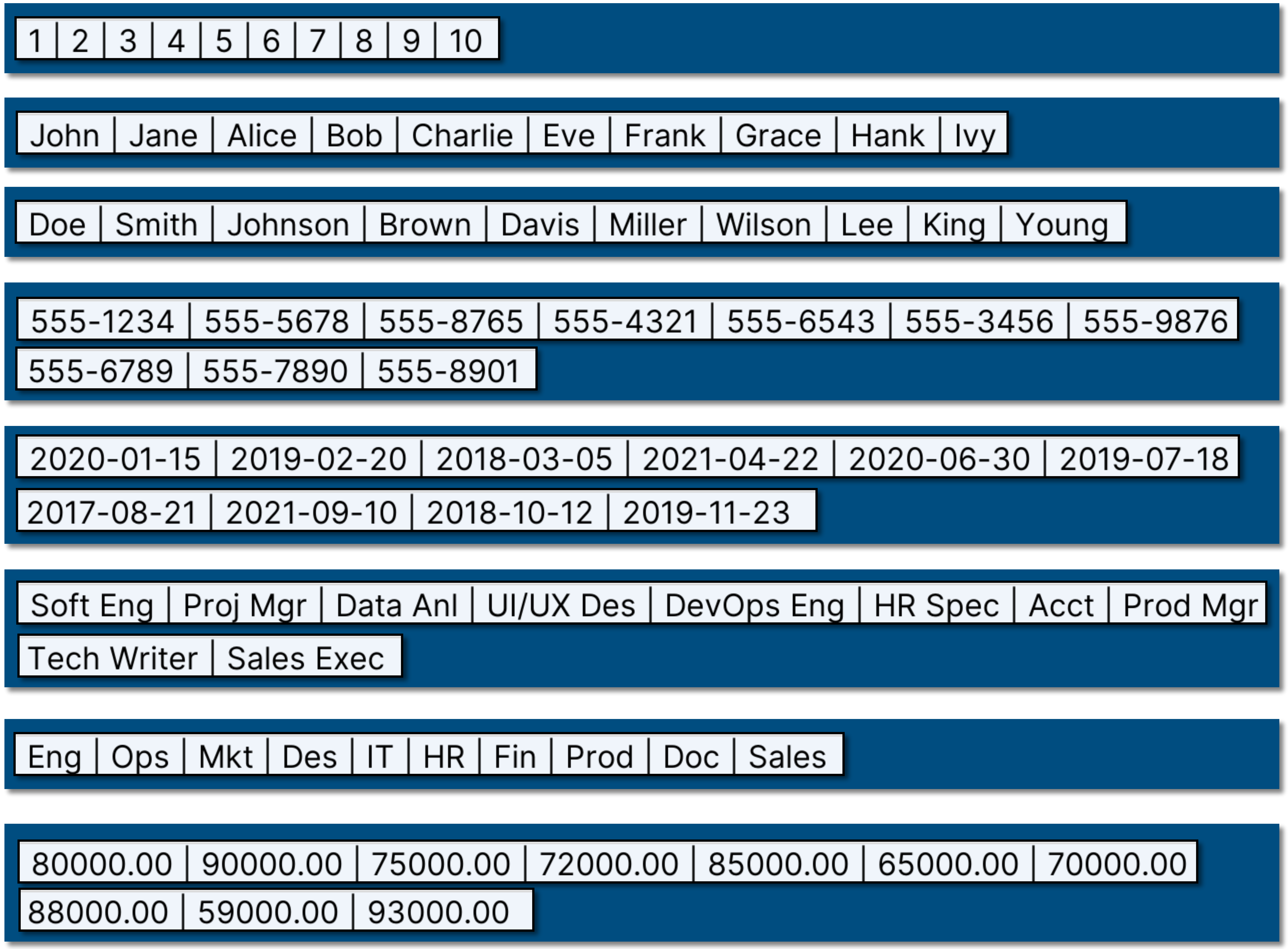


Eng | Ops | Mkt | Des | IT | HR | Fin | Prod | Doc | Sales
John | Jane | Alice | Bob | Charlie | Eve | Frank | Grace | Hank | Ivy
Doe | Smith | Johnson | Brown | Davis | Miller | Wilson | Lee | King | Young
2020-01-15 | 2019-02-20 | 2018-03-05 | 2021-04-22 | 2020-06-30 | 2019-07-18
2017-08-21 | 2021-09-10 | 2018-10-12 | 2019-11-23

Amr Elhelw's TECH VAULT

```sql
INSERT INTO employees
VALUES (...)
```





1	2	3	4	5	6	7	8	9	10
John	Jane	Alice	Bob	Charlie	Eve	Frank	Grace	Hank	Ivy
Doe	Smith	Johnson	Brown	Davis	Miller	Wilson	Lee	King	Young
555-1234	555-5678	555-8765	555-4321	555-6543	555-3456	555-9876	555-6789	555-7890	555-8901
2020-01-15	2019-02-20	2018-03-05	2021-04-22	2020-06-30	2019-07-18	2017-08-21	2021-09-10	2018-10-12	2019-11-23
Soft Eng	Proj Mgr	Data Anl	UI/UX Des	DevOps Eng	HR Spec	Acct	Prod Mgr	Tech Writer	Sales Exec
Eng	Ops	Mkt	Des	IT	HR	Fin	Prod	Doc	Sales
80000.00	90000.00	75000.00	72000.00	85000.00	65000.00	70000.00	88000.00	59000.00	93000.00

Amr Elhelw's TECH VAULT

Column Stores

- Pros
    - *Suitable for queries that need only a few columns from a large number of rows*
    - *Enables data compression (even better I/O) **
- Cons
    - *Not suitable for point queries and writes due to having to touch many columns*

Amr Elhelw's TECH VAULT

	Row Stores	Column Stores
Storage	Data of the same row is stored together	Data of the same column is stored together
Point Queries	👍	👎
Inserts/Updates/Deletes	👍	👎
Aggregation/Analytical Queries	👎	👍
Compression Support	👎	👍

Amr Elhelw's TECH VAULT