Exam Review
SQL $\rightarrow$ RA $\rightarrow$ Iter
$\text{union [all]}$

$\text{limit (limit)}$

$\text{select}$

$\text{distinct (select distinct)}$

$\text{order by (order by)}$

$\text{select items (\pi)}$

$\text{having (\sigma)}$

$\text{agg (for agg Qs)}$

$\text{where (\sigma)}$

$\text{source (X, \n)}$

$\text{output}$
Iterators

void open() {
    // call open() on child iterators
    // prepare the iterator
}

Tuple getNext() {
    // read, process, and return a tuple
}

void close() {
    // clean-up the iterator
    // call close() on child iterators
}
CREATE TABLE PLAYERS(
    ID string, FIRSTNAME string, LASTNAME string,
    FIRSTSEASON int, LASTSEASON int, WEIGHT int,
    BIRTHDATE date
);

INSERT INTO
    PLAYERS(ID, FIRSTNAME, LASTNAME, FIRSTSEASON, LASTSEASON,
    WEIGHT, BIRTHDATE)
VALUES
    ('MARTIPHO1', 'Phil', 'Martin', 1954, 1954, 190, 1928-04-02)
    ('GREENLA01', 'Lamar', 'Green', 1969, 1974, 210, 1947-03-22)
    ('BIANCAL01', 'Al', 'Bianchi', 1956, 1965, 185, 1932-03-26)
    ('SMITHTO01', 'Tony', 'Smith', 1990, 2000, 185, 1968-06-14);

SELECT FIRSTNAME, LASTNAME, FIRSTSEASON, LASTSEASON
FROM PLAYERS
WHERE LASTSEASON-FIRSTSEASON>5;

SELECT P1_FIRSTNAME, P2_FIRSTNAME
FROM (
    SELECT P1.FIRSTNAME AS P1_FIRSTNAME, P2.FIRSTNAME AS P2_FIRSTNAME
    P1.FIRSTSEASON AS P1_FIRSTSEASON, P1.LASTSEASON AS P1_LASTSEASON,
    P2.FIRSTSEASON AS P2_FIRSTSEASON, P2.LASTSEASON AS P2_LASTSEASON
    FROM PLAYERS P1, PLAYERS P2
    WHERE P1.ID<>P2.ID
) SUB_Q
WHERE P1_FIRSTSEASON<P2_FIRSTSEASON
    AND P1_LASTSEASON>P2_LASTSEASON;

SELECT EXPERIENCE, COUNT(*)
FROM (
    SELECT ID, FIRSTNAME, LASTNAME,
    (LASTSEASON-FIRSTSEASON) AS EXPERIENCE
    FROM PLAYERS
)
GROUP BY EXPERIENCE;
SELECT FIRSTNAME, LASTNAME, FIRSTSEASON, LASTSEASON
FROM PLAYERS
WHERE LASTSEASON-FIRSTSEASON > 5

Players

<table>
<thead>
<tr>
<th>Lamar</th>
<th>AC</th>
<th>Michael</th>
</tr>
</thead>
</table>

Players

| Farmer | DE | Marti | Phi | 01 |
SELECT P1_FIRSTNAME, P2_FIRSTNAME
FROM (SELECT P1_FIRSTNAME AS P1_FIRSTNAME,
       P2_FIRSTNAME AS P2_FIRSTNAME,
       P1_FIRSTSEASON AS P1_FIRSTSEASON,
       P1_LASTSEASON AS P1_LASTSEASON,
       P2_FIRSTSEASON AS P2_FIRSTSEASON,
       P2_LASTSEASON AS P2_LASTSEASON
    FROM PLAYERS P1, PLAYERS P2
    WHERE P1.ID<>P2.ID
    ) SubQ
WHERE P1_FIRSTSEASON<P2_FIRSTSEASON
AND P1_LASTSEASON>P2_LASTSEASON;
\((P_1 \times P_2) \times P_3\)
Aggregate

```sql
SELECT EXPERIENCE, COUNT(*)
FROM (SELECT ID, FIRSTNAME, LASTNAME, (LASTSEASON-FIRSTSEASON) AS EXPERIENCE
      FROM PLAYERS
    ) Exp
GROUP BY EXPERIENCE;
```
<table>
<thead>
<tr>
<th>R</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>42</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>28</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>32</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>26</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>9</td>
<td>3</td>
</tr>
</tbody>
</table>

\[ A = ?, C = ? \]

1. \((4, 7)\)
2. \((28)\)
3. \((13, 4, 5, 9, 10)\)

Unclustered

Clustered

Pages
Find $\theta$

$\theta = 2$
<table>
<thead>
<tr>
<th></th>
<th>Heap File</th>
<th>Sorted File</th>
<th>B+Tree</th>
<th>Hash</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lookup A=?</td>
<td>N</td>
<td>Log₂ N</td>
<td>Log₂ N</td>
<td>1</td>
</tr>
<tr>
<td>Scan A&lt;40A23</td>
<td>N</td>
<td>Log₂ N</td>
<td>Log₂ N</td>
<td>N</td>
</tr>
<tr>
<td>Insert</td>
<td>A=0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delete</td>
<td>A=0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>A</td>
<td>B</td>
<td>S</td>
<td>B</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\text{GHS}
\]
\[
\text{Hash } S, \text{ Stream } R
\]
\[
\text{S} \quad \text{B} \quad (B, C)
\]
\[
\begin{align*}
12 & \rightarrow (12, 12, 12, 12) \\
2 & \rightarrow (2, 2, 2, 2) \\
4 & \rightarrow (4, 4) \\
6 & \rightarrow (6, 6) \\
8 & \rightarrow (8, 8)
\end{align*}
\]
\[
\text{RM} \quad S
\]
\[
\begin{align*}
\text{NLJ}_{VR} & \rightarrow VS \\
\langle 1, 2 \rangle & \rightarrow \langle 12, 8 \rangle \\
& \rightarrow \langle 12, 9 \rangle \\
& \rightarrow \langle 2, 3, 7 \rangle \\
& \rightarrow \langle 4, 7 \rangle \\
& \rightarrow \langle 5, 6 \rangle \\
\end{align*}
\]
\[
\begin{align*}
\text{BNLJ} & \rightarrow \begin{bmatrix} 1, 2 \\ 3, 7 \\ 5, 6 \end{bmatrix} \\
& \rightarrow \begin{bmatrix} 12, 8 \\ 12, 9 \\ 2, 2 \\ 2, 7 \end{bmatrix} \\
& \rightarrow \langle 12, 8 \rangle \rightarrow \langle 12, 9 \rangle \rightarrow 2 \rightarrow \emptyset \\
& \rightarrow \langle 2, 3, 7 \rangle \rightarrow \langle 4, 7 \rangle \rightarrow \emptyset \\
& \rightarrow \langle 5, 6 \rangle \rightarrow \langle 2, 7 \rangle \\
& \rightarrow \langle 4, 4 \rangle \rightarrow \langle 6, 2 \rangle \rightarrow \emptyset
\end{align*}
\]
SELECT  
  lineitem.orderkey,  
  sum(lineitem.extendedprice*(1-lineitem.discount)) as revenue,  
  orders.orderdate,  
  orders.shippriority  
FROM  
  customer,  
  orders,  
  lineitem  
WHERE  
  customer.mktsegment = 'BUILDING' and customer.custkey = orders.custkey  
  and lineitem.orderkey = orders.orderkey  
  and orders.orderdate < DATE('1995-03-15')  
  and lineitem.shipdate > DATE('1995-03-15')  
GROUP BY  
  lineitem.orderkey, orders.orderdate,  
  orders.shippriority  
ORDER BY  
  revenue desc, orders.orderdate;