Plan...

- Class preparation… 2/7 questions

- Group work (7 groups, 5-10 minutes)
  - Each group reviews responses to one question
  - Determines “best” response to share with class
  - Identifies observations about different solutions

- Presentations (~30 minutes)
  - 3-5 minute presentation
  - Present solution / observations

- Mini-Lecture (~10 minutes)
  - Queries with multiple tables…
Single Table Queries…

Set 1: Using technical language..

• 1A. What unique IP addresses visited from Germany?
• 1B. How many hits did each unique IP address have?
• 1C. What are the IP addresses and number of hits for IP addresses that visited the Arthritis Source between 2:00 AM and 3:00 AM?
• 1D. What did the following IP addresses do when they accessed the site: (1) cust9.max12.washington.dc.ms.uu.net, (2) festpc02.ib.be, (3) gatekeeper.volvo.se, (4) goldentrout.execpc.com

Set 2: Requiring more inference...

• 2A. Who accessed the Arthritis Source from European Countries?
• 2B. Which movies have been accessed the most?
• 2C. Have users been visiting the Arthritis Source in the early morning hours? If so, who and how much?
Queries with Multiple Tables

- Until now, we have focused on queries with single tables

- What if question requires attributes from multiple tables
  - INNER JOIN
  - LEFT JOIN
  - RIGHT JOIN
JOIN TYPES

• Syntax
  
  SELECT field1, field2, …
  FROM TABLE1, TABLE2
  TABLE1 ________ JOIN TABLE2
  ON [TABLE1].field = [TABLE2].field

• Join Types
  – INNER:
    • Includes only rows where the joined fields from both tables are equal.
  – LEFT:
    • TABLE1 drives the query results
    • Includes all records from TABLE1 and only records from TABLE2 when the joined fields are equal.
  – RIGHT:
    • TABLE2 drives the query results
    • Includes all records from TABLE2 and only records from TABLE1 when the joined fields are equal.
Example

• Imagine a new table – regUsers
  – Design
    • regUserIP
    • regUserName
    • ...
  – Explanation
    • Customers do **not** need to be registered to use
      – hitData will contain IP addresses **not** on regUser

• Registered customers **do not** need to use
  – regUser may have IPs which have no hits in hitData
Question: What resources did each registered user access?

- “Where clause” approach

- JOIN approach
Question: What resources did each registered user access?

- “Where clause” approach

```sql
SELECT [CUSTOMER].customerName, [HITDATA].resourceName
FROM CUSTOMER, HITDATA
WHERE [CUSTOMER].customerIP = [HITDATA].ip
```

- JOIN approach

```sql
SELECT [CUSTOMER].customerName, [HITDATA].resourceName
FROM CUSTOMER, HITDATA
CUSTOMER INNER JOIN HITDATA
ON [CUSTOMER].customerIP = [HITDATA].ip
```
Question: What resources did each customer access?

Variations

• Previous:
  
  SELECT [CUSTOMER].customerName, [HITDATA].resourceName
  FROM CUSTOMER, HITDATA
  CUSTOMER INNER JOIN HITDATA
  ON [CUSTOMER].customerIP = [HITDATA].ip

• Create a list of resources accessed by EACH and EVERY customer – if customers accessed no resources, include that on the list.

  SELECT [CUSTOMER].customerName, [HITDATA].resourceName
  FROM CUSTOMER, HITDATA
  CUSTOMER ________ JOIN HITDATA
  ON [CUSTOMER].customerIP = [HITDATA].ip
Question: What resources did each customer access?

Variations

• Previous:
  SELECT [CUSTOMER].customerName, [HITDATA].resourceName
  FROM CUSTOMER, HITDATA
  CUSTOMER INNER JOIN HITDATA
  ON [CUSTOMER].customerIP = [HITDATA].ip

• Create a list of all resources accessed, and customer Name if it is known.

  SELECT [CUSTOMER].customerName, [HITDATA].resourceName
  FROM CUSTOMER, HITDATA
  CUSTOMER JOIN HITDATA
  ON [CUSTOMER].customerIP = [HITDATA].ip
Question: What resources did each customer access?

Variations

• Create a list of resources accessed by EACH and EVERY customer – if customers accessed no resources, include that on the list.

• Create a list of all resources accessed, and customer Name if it is known.