

Data Analysis CA 4 - Database Design & SQL Querying

Part A: Deliverable 2 - Query

James Bunt (D00262403)

January 2023

```
# Clear the console  
cat("\014")
```

```
# Load necessary packages
library(pacman)
pacman::p_load(knitr, dbplyr, RSQLite, rmarkdown, tidyverse)

# Connect to the database
dbfile <- "ca4-db.sqlite"
ca4_db <- dbConnect(RSQLite::SQLite(), dbfile)
```

Database Queries

1. SELECT with WHERE, LIKE, and OR

```
# Select data from table using where, like and or clauses
result <- dbGetQuery(ca4_db, "SELECT
    LibraryName AS [Library Name],
    Description AS [Asset],
    Name AS [Type]
FROM
    AssetsInLibrary
WHERE
    Description LIKE '%Red%' OR
    Description LIKE '%Blue%'" )

# Display results in table using kable (in knitr package)
kable(result,
    col.names = c("Library Name", "Asset", "Type"),
    caption = "List of assets that are red or blue")
```

Table 1: List of assets that are red or blue

Library Name	Asset	Type
3D Props	Blue Sword	3D Prop
3D Characters	Red Villian	3D Character
3D Concepts	Red Armour	3D Prop
3D Concepts	Blue Armour	3D Prop

2. SELECT with DISTINCT and ORDER BY

```
# Select data from table using distinct and order by
result <- dbGetQuery(ca4_db, "SELECT DISTINCT
    Name AS [Asset Type]
FROM
    AssetType")
```

```

ORDER BY
  Name")

# Display results in table using kable (in knitr package)
kable(result,
  col.names = "Asset Types",
  caption = "List of distinct asset types")

```

Table 2: List of distinct asset types

Asset Types
3D Character
3D Prop
Concept Art

3. Inner Join

```

# Select data from table using distinct and order by
result <- dbGetQuery(ca4_db, "SELECT
    Project.Name AS [Project Name],
    Asset.Description AS [Asset Description],
    AssetType.Name AS [Asset Type]
FROM
    Asset
INNER JOIN
    AssetType ON Asset.AssetType_Id = AssetType.Id
INNER JOIN
    AssetsInProject ON Asset.Id = AssetsInProject.AssetId
INNER JOIN
    Project ON AssetsInProject.AssetId = Project.Id
ORDER BY
    Project.Name,
    Asset.Description")

# Display results in table using kable (in knitr package)
kable(result,
  col.names = c("Project Name", "Asset Description", "Asset Type"),
  caption = "List of unique asset types")

```

Table 3: List of unique asset types

Project Name	Asset Description	Asset Type
Easter Hunt	Golden Hero	3D Character
Summer Mayhem	Red Villian	3D Character

4. Subquery with SELECT

```
# Select data from tables using a select subquery
result <- dbGetQuery(ca4_db, "SELECT S.Name as SkillName, T.Name, TM.FirstName || ' ' || TM.LastName as EmployeeName
FROM Team T
INNER JOIN TeamMember TM ON T.Id = TM.Team_Id
INNER JOIN (SELECT * FROM SkillsPerMember WHERE Skill_Id = (SELECT Id FROM Skill S)) SPM ON TM.Team_Id = SPM.Team_Id
INNER JOIN Skill S ON SPM.Skill_Id = S.Id")

# Display results in table using kable (in knitr package)
kable(result, col.names = c("Skill", "Team", "Employee Name"), caption = "List of people who can code")
```

Table 4: List of people who can code

Skill	Team	Employee Name
Coding	Design	Andrew Apple
Coding	Design	Beth Barker
Coding	Development	Candy Curtis

5. SELECT across a Date Range

```
# select data from table between a range of dates
result <- dbGetQuery(ca4_db, "SELECT P.Name, P.DueDate, COUNT(DISTINCT AP.Asset_Id) as AssetCount
FROM Project P
INNER JOIN AssetsinProject AP ON P.Id = AP.Project_Id
WHERE P.DueDate <= '2023-06-30'
GROUP BY P.Name, P.DueDate")

kable(result, col.names = c("Project Name", "Due Date", "Asset Count"), caption = "List of projects and their asset count, due by end of June 2023")
```

Table 5: List of projects and their asset count, due by end of June 2023

Project Name	Due Date	Asset Count
Easter Hunt	2023-03-01	2
Summer Mayhem	2023-06-01	4

Close Database Connection

```
# close the connection to the database to avoid hitting a connection limit
dbDisconnect(ca4_db)
```