**Memo: # 0 (Bears In Canada Project)**

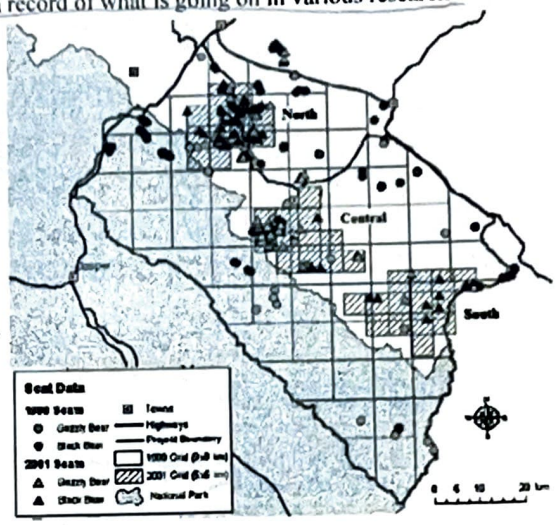
# DBDS DataBase Design Specialists, Inc.

"We are always 5NF "

**To:** Pozos Garcia, Adolfo NN:70, DA
**From:** Design Team Leader
**Subject:** New project, communications, and deliverables
**Date:** October 16, 2024

***New Project:***

We have just won the contract to design and implement a prototype database for the Bears In Canada Project – BIC project. You have been selected as the DA (Data Architect) for this project. The motivation for this project is to keep a record of what is going on in various research studies involved with bears where samples are located by telemetry, hair snags, and scat located by specially trained canines (dogs).




The motivation for the project is declining populations of bears. The project is about tracking the location and health of bears in or near a national park in Canada. Tracking is done by three different methods: telemetry tracking (after a bear has been captured and a tracking device has been attached), hair snags (found on bushes and trees in the area), and scats (animal feces) located by specially trained dogs. DNA data from hair and scat samples can be used to determine the sex of the bear and some health data. The scat samples provide individualized data on the physiological health of that animal.

You will be receiving a series of memos to detail the project as information becomes available. Since we are off-site, much of the information will be gathered through electronic mail communication with the principals in the project.

There are many regions in which studies are being conducted. Three regions are to be part of the test data: North Region (NR), Central Region (CR) and South Region (SR) (see diagram). Regions are divided into cells (which are numbered) and a x, y position location within a cell.

***Communications:***

The BIC contact person is Bob Bureaucrat. He has just joined the BIC project as project director. He has selected Sam Supervisor, the North Region manager, as an additional contact person. Remember, these folks are our customers and know the details of the project; however, they are not database experts so do not ask them technical (DBMS) questions. We have covered those topics in our design meetings.

You may communicate with Bob Bureaucrat and Sam Supervisor using electronic mail (**actually Canvas messages**). Send your messages destined for Bob or Sam to me (with their name added as a prefix to the subject line) and I will forward the mail to them and their responses back to you. They have asked that you limit each piece of mail to **one question** of 10 lines or less. They have also asked that you limit your questions to 1 or 2 per week, as they are both working on other projects. Also, Bob Bureaucrat is out of the office much of the time. If he is not in the office, questions will be returned. He will try to let us know in advance when he will be available.

I would suggest that you number your questions or use descriptive subject headings on the mail messages so you can correlate the answer with the question. I am sure they will use the REPLY function to answer your questions.

As an aid to the customer, I have supplied some predefined answer files and have installed them on their computers. Some of the responses are:

The question is not relevant to the BIC effort.
The question is too complex, please revise it.
I do not understand the question. I am confused by ...
Bob Bureaucrat is out of the office. Please resubmit later.

***Deliverables:***

We are going to use the methodology described in our design meetings. You are to supply the usual deliverables at the end of the project, i.e.,

- List of (charts of) functional dependencies
- A relational database design (5NF) - a Logical Data Model (LDM), and a corresponding Physical Data Model (PDM)
- A prototype of the relational database design (RDD) generated by ERWin to an SQL Server 2022 database
- Results, from the prototype database, for typical queries (questions)

Your technical report should coordinate the problem statement, and the items listed above. Please prepare your reports in machine-readable form (as a Word docx file). I will use portions of your work in my final report to the customer. In addition, please keep an organized record of your development efforts. Periodically I will have to ask you for the current status so that I can present interim reports to the customer.

***Completion Date:***

The project completion date will be announced later. The first deliverable will be requested in a few days.

**Memo: # 1 (Bears In Canada Project)**

# DBDS DataBase Design Specialists, Inc.

"We are always 5NF."

**To:** Pozos Garcia, Adolfo, NN:70, DA
**From:** Design Team Leader
**Subject:** Enterprise overview
**Date:** October 16, 2024

The basis of the Bears In Canada (BIC) project is the location and analysis of samples that are part of studies to determine the location and health of the animals (Bears) that are the target population of the studies. Studies are carried out in various regions within a National Park in Canada. The basic unit of investigation within a particular study is a sample. That sample is associated with a particular animal using (one or more of) telemetry, hair snags, or scats to determine the location, and DNA analysis (if applicable) to gauge the health of the animal at the time of the sample.

Interviews were conducted with BIC personnel, including Bob Bureaucrat, and Sam Supervisor (North Region manager) and other region managers. Excerpts from those interviews are listed below:

**Question** - How do you identify a specific animal?
**Sam Supervisor** - Each animal is numbered to help us keep track of each individual animal.

**Question** - How many regions are used in the research studies?
**Bob Bureaucrat** - There are 3. But we currently have a proposal under consideration to expand the research.

**Question** - What species of animals are in the North Region?
**Sam Supervisor** - Right now there are black bears and grizzly bears.

**Question** - What kinds of animals are studied in the Central Region?
**Bob Bureaucrat** - Black bears
**Mary Manager** - There are black bears and grizzly bears, but we have other kinds of animals too, for which we sometimes get samples (the dogs sometimes cannot tell the difference).

**Question** - What kinds of animals are in the South Region?
**Fred Foreman** - We are following the same animals as the other regions.

**Question** - What does the sample data look like?
**Sam Supervisor** - I do know that the PHT (Physiological Health Threshold) and location values are part of the data for my region. I do know that data are collected by study and each sample is classified based on the type of data collected.
**Bob Bureaucrat** - I don't know all the details, but I can have the staff look it up for you if you'd like.

... time passes ...

**Bob Bureaucrat** - Here is an example of data that has been collected. (The # column is **not part of the data**, only for reference)

| # | Species Code | Region ID | Study Name | Study ID | Size | Class ID | Average PHT | PHT value | Animal Number | Location | Sample Date | Sample Number | Sex | Status |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | B | NR | North 2022 | N22 | 9x9 | S | 113 | 109 | 42 | 05:8:3 | Jul 2022 | 17 | M | 1 |
| 2 | B | SR | South 2022 | S22 | 5x5 | T | 113 | | 89 | 93:2:4 | Nov 2022 | 22 | F | 1 |
| 3 | B | CR | Central 2022 | C22 | 9x9 | T | 113 | | 59 | 32:1:9 | Sep 2022 | 44 | M | 0 |
| 4 | G | CR | Central 2022 | C22 | 9x9 | H | 142 | | 113 | 40:1:1 | Oct 2022 | 45 | F | 0 |
| 5 | B | CR | Central 2022 | C22 | 9x9 | T | 113 | | 59 | 41:2:3 | Sep 2022 | 47 | M | 0 |
| 6 | B | CR | Central 2023 | C23 | 9x9 | S | 113 | 100 | 59 | 34:4:4 | Sep 2023 | 48 | F | 1 |
| 7 | B | CR | Central 2023 | C23 | 9x9 | S | 113 | 103.5 | 50 | 40:1:1 | Jul 2023 | 56 | ? | 1 |
| 8 | B | NR | North 2022 | N22 | 9x9 | S | 113 | 120 | 118 | 07:1:2 | Jun 2022 | 59 | F | 1 |
| 9 | G | CR | Central 2023 | C23 | 9x9 | S | 142 | 135 | 112 | 32:5:5 | Jul 2023 | 79 | M | 1 |
| 10 | G | CR | Central 2022 | C22 | 9x9 | T | 142 | | 66 | 31:5:8 | Nov 2022 | 82 | F | 0 |
| 11 | U | NR | North 2022 | N22 | 9x9 | S | | | 66 | 01:1:9 | Jul 2022 | 100 | ? | 0 |
| 12 | B | SR | South 2022 | S22 | 5x5 | H | 113 | | 66 | 80:3:2 | Jul 2022 | 68 | M | 0 |
| 13 | B | NR | North 2022 | N22 | 9x9 | S | 113 | 115 | 42 | 15:2:6 | Aug 2022 | 27 | M | 1 |
| 14 | G | NR | North 2022 | N22 | 9x9 | S | 142 | 135 | 113 | 19:4:7 | Jul 2022 | 11 | F | 0 |

**Question** - What do these abbreviations mean?
**Bob Bureaucrat** - Here are the codes we use

**Species (animal type):**
B = Black bear
G = Grizzly bear
U = Undetermined

**Region (name):**
NR = North Region
CR = Central Region
SR = South Region

**Classification (name):**
T = Telemetry
H = Hair snag
S = Scat

**Status**
1 = Sample exists
0 = Sample used up

**Question** - What are Status and Location?
**Bob Bureaucrat** - Status has to do with whether there is any sample left (1) and available for further analysis or not, 0 . (for telemetry 1 means unit is still working, 0 it is not). Location is a grid number followed by the x, y position within the grid.

**Memo: # 2 (Bears In Canada Project)**

# DBDS DataBase Design Specialists, Inc.

"We are always 5NF."

**To:** Pozos Garcia, Adolfo NN: 70, DA
**From:** Design Team Leader
**Subject:** LDM
**Date:** October 28, 2024

I assume you have completed your functional dependency analysis – if you have, include a screen snapshot of your analysis. For now, just to check that we are on the way with this project, I will need to see your ERwin LDM (Logical Data Model). It should be based on your FD analysis using the information in Memo #0 and Memo #1. Name your ERwin file **BICnn.erwin** where **nn** is your number.

**Note: When you set up the ERwin model, pick the option for both Logical and Physical data models.**
**Do not generate a database yet.**

For the LDM all that is needed is a readable snapshot (use the Snipping Tool to capture the LDM - inside of the ERWin LDM frame – images of the entire screen display are hard to read). Place the ERWin model snapshot in a Word document (along with the snapshot of your FD analysis if you have one), name your file **MemoTwoNN.docx** where NN is your number

Then look for a submission message on Canvas.

## Due date: Saturday, November 2, 2024

## Memo: # 3 (Bears In Canada Project)

# DBDS DataBase Design Specialists, Inc.

"We are always 5NF."

**To:** Pozos Garcia, Adolfo NN: 70, DA
**From:** Design Team Leader
**Subject:** Queries and ERWin Model
**Date:** November 11, 2024

**Personnel:**

It has been decided that the database must contain information about the participants in the BIC project. The participants include management personnel, the laboratory technicians who analyze the samples, and dogs that retrieve scat samples. The laboratory technicians are assigned duties based on sample classification – they only process one kind of sample. Each participant has been given a participant ID, a beginning date and ending date of their contact with the project has been included if known. The Participant ID for Bob Bureaucrat is P0000 and the NR, CR and SR manager's IDs are P0101, P0102, and P0103. The managers have been with the project from 5/23/2023 and Bob started 2 months ago.

It has also been requested that the date of Bob Bureaucrat's next project review be recorded.

**Data Access:**

There are several users of the information that is stored in the database. In order to increase the value of this information, it should be possible to maintain the quality of the data entered into the database. Region managers will need to see the data concerning the samples in their region. Laboratory technicians will need to be able to enter and/or change data for the samples assigned to them. **Provide as much of this as possible using the normal integrity constraints automatically enforced by the database engine,** do not add an authentication system – SQL Server security can deal with that part later.

The project director wishes to know what is happening with regard to the project, therefore he wants the system to record information on the last time each manager accessed the database. And he wants to know how many times each dog locates a sample.

## The new information about the data received from Bob is on the next page

After the design is complete and the data from the various memos is loaded into the database, the following list of questions is to be answered, using the data that has been loaded. This is only preliminary. Other information may be needed as the project progresses. These questions may be helpful in your design for the Logical Data Model.

**Questions:**

1) What is the largest physiological health value observed for a black bear?

2) For each animal, list all its sample classifications in chronological (date) order.

3) In what region is the Central 2022 study and what size grid pattern is used on that study?

4) List the animals that are within 5 units of their average physiological health value.

5) List the sample information for the Central 2022 studies made in September 2022 and in November 2022.

6) List the names of managers and their last access to the database.

7) List the study name, animal number, animal species code, sample number and date found for samples that were found in the Central Region in cell number 40?

8) Who (Name and ID) manages the Central Region?

9) Who (Name and ID) has access data on South Region sample data?

10) List the name of each dog and the number of scat samples found by that dog.

11) List the samples analyzed by P2045.

12) List the names of the dogs that have worked in each study.

13) List the IDs and names of participants who analyzed telemetry samples

## General data values:

| # | Species Code | RegionID | StudyName | StudyID | Size | ClassID | Average PHT | PTH Value | Animal Number | Location | Sample Date | Sample Number | Sex | Status | Participant ID |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | B | NR | North 2022 | N22 | 9x9 | S | 113.0 | 109 | 42 | 05:8:3 | Jul 2022 | 17 | M | 1 | D0004 |
| 2 | B | SR | South 2022 | S22 | 5x5 | T | 113.0 | | 89 | 93:2:4 | Nov 2022 | 22 | F | 1 | P2045 |
| 3 | B | CR | Central 2022 | C22 | 9x9 | T | 113.0 | | 59 | 32:1:9 | Sep 2022 | 44 | M | 0 | P2001 |
| 4 | G | CR | Central 2022 | C22 | 9x9 | H | 142.0 | | 113 | 40:1:1 | Oct 2022 | 45 | F | 0 | P2046 |
| 5 | B | CR | Central 2022 | C22 | 9x9 | T | 113.0 | | 59 | 41:2:3 | Sep 2022 | 47 | M | 0 | P2045 |
| 6 | B | CR | Central 2023 | C23 | 9x9 | S | 113.0 | 100 | 59 | 34:4:4 | Sep 2023 | 48 | F | 1 | D0013 |
| 7 | B | CR | Central 2023 | C23 | 9x9 | S | 113.0 | 103.5 | 50 | 40:1:1 | Jul 2023 | 56 | ? | 1 | D0004 |
| 8 | B | NR | North 2022 | N22 | 9x9 | S | 113.0 | 120 | 118 | 07:1:2 | Jun 2022 | 59 | F | 1 | D0022 |
| 9 | G | CR | Central 2023 | C23 | 9x9 | S | 142.0 | 135 | 112 | 32:5:5 | Jul 2023 | 79 | M | 1 | D0004 |
| 10 | G | CR | Central 2022 | C22 | 9x9 | T | 142.0 | | 66 | 31:5:8 | Nov 2022 | 82 | F | 0 | P2045 |
| 11 | U | NR | North 2022 | N22 | 9x9 | S | | | 66 | 01:1:9 | Jul 2022 | 100 | ? | 0 | D0022 |
| 12 | B | SR | South 2022 | S22 | 5x5 | H | 113.0 | | 66 | 80:3:2 | Jul 2022 | 68 | M | 0 | P2004 |
| 13 | B | NR | North 2022 | N22 | 9x9 | S | 113.0 | 115 | 42 | 15:2:6 | Aug 2022 | 27 | M | 1 | D0008 |
| 14 | G | NR | North 2022 | N22 | 9x9 | S | 142.0 | 135 | 113 | 19:4:7 | Jul 2022 | 11 | F | 0 | D0008 |
| 15† | B | SR | South 2022 | S22 | 5x5 | S | 113.0 | 117 | 63 | 90:3:4 | Jul 2022 | 45 | M | 0 | D0013 |
| 16† | G | CR | Central 2022 | C22 | 9x9 | T | 142.0 | | 114* | 40:4:1 | Oct 2022 | 17 | ? | 1 | P30nn |
| 17† | G | CR | Central 2022 | C22 | 9x9 | S | 142.0 | 150 | 114* | 40:4:1 | Oct 2022 | 18 | F | 1 | D0004 |

† new data * Same animal

## Data for Laboratory Technicians and Dogs:

**Technicians**

| # | ParticipantID | Name | StartDate | EndDate | ClassID |
|---|---|---|---|---|---|
| 1 | P2001 | Bill Brown | 2022-02-14 | | T |
| 2 | P2004 | Jane Smith | 2022-02-14 | | H |
| 3 | P2036 | Frank Martin | 2020-08-15 | 2022-01-01 | T |
| 4 | P2045 | Anne Dough | 2021-06-12 | | T |
| 5 | P2046 | Mike Green | 2020-10-28 | | H |
| 6 | P30nn | <your name> | <current date> | | T |

**Dogs**

| # | ParticipantID | Name | StartDate | EndDate | Number Samples |
|---|---|---|---|---|---|
| 1 | D0004 | Max | 2022-06-01 | | 3 |
| 2 | D0008 | Sampson | 2022-02-05 | | 3 |
| 3 | D0013 | Cindy | 2021-12-10 | 2022-12-20 | 2 |
| 4 | D0022 | Rover | 2022-05-20 | | 2 |

**Notes:** The # columns to the left of the data rows are **not part of the data.** They are for reference only.
New sample data has been added, marked by †. Except for the last column, the previous data has not changed.
Add "your" row to the Participant table where **nn** is your number and change the ParticipantID in row 16 above (may involve adding rows to subtypes too).

**Deliverables:**

I have to give a presentation to Bob Bureaucrat in an upcoming meeting. I need a Logical Data Model and Physical Data Model for the presentation. Give me your current thinking on this on or before **Noon, Tuesday, November 26, 2024.** The design is to be based on information supplied in this Memo and including the example data provided by Bob Bureaucrat in Memo #1, the associated functional dependencies, and the queries included in this memo. Do not create the database yet, but you can begin to prepare the INSERTs for the data to be loaded.

Your model should have been constructed using ERWin. The deliverable is your ERWin file **BICnn.erwin** where **nn** is your number. **Look for an assignment on Canvas.**

**Memo: # 4 (Bears In Canada Project)**

# DBDS DataBase Design Specialists, Inc.

"We are always 5NF."

**To:** Pozos Garcia, Adolfo, NN: 70 DA
**From:** Design Team Leader
**Subject:** Database and Final Report
**Date:** November 25, 2024

***Database:***

Make any necessary adjustments to your LDM and/or PDM based on feedback from Memo 3. Use ERwin to generate your database (RDD) on the SQL Server 2022 server. Name the database BICnn where nn is your number.

Be sure you added yourself as a Technician with Participant ID of P30nn where nn is your NN number. In row 16 of the Memo 3 data, change the P30nn to match your nn number, and row 6 in the Technician data.

Load the data from the memos into your SQL Server 2022 database.

Carry out your analysis.

***Final Report:***

Please complete and submit to me your **final report** that will go to the **customer**.

- Be sure to include deliverables for this project were **detailed in Memo #0**
- In the LDM, try to draw it so the items that come "first" in the enterprise are at the top (i.e.., independent entities at the top)
- Include as an appendix to your report the two draft (old) LDMs and the PDM (along with other supporting material relevant to the project but not of high interest to the customer).
- Be sure all the data items from the memos are in your test data.
  - Include a listing of the DML (INSERT statements) for the data load
  - and include SELECT * for ALL the tables in your test set.
- **Name your report file BICFinalNN.docx**

Your prototype database in SQL Server 2022 on the server will be examined in detail in conjunction with the final report.

**Deliverable: Final Report Noon on Saturday, December 7, 2024. Look for a Canvas Assignment.**