**TECHNICAL REPORT**

DATABASE DESIGN AND IMPLEMENTATION

|  |  |
| --- | --- |
| PREPARED BY | Varadharajan, Aashika |
| NN | 14 |
|  |  |

**Contents**

[**Project Objective 3**](#_Toc38648209)

[**List of Functional Dependencies 3**](#_Toc38648210)

[**Logical Data Model 6**](#_Toc38648211)

[**Physical Data Model 7**](#_Toc38648212)

[**Schema For The Generation Of Relational Database Design From Erwin 8**](#_Toc38648213)

[**DML Insert Commands Used For Loading The Data 11**](#_Toc38648214)

[**Select Statements On All The Tables 15**](#_Toc38648215)

[**Results For The Questions From The Prototype Db 19**](#_Toc38648216)

[**APPENDIX – Previous versions of the designs 24**](#_Toc38648217)

# **PROJECT OBJECTIVE:**

The main aim of the BIC Project was to create a 5 NF database that captured the animal tracking information from the study groups focused on collecting the sample data about the location and health of bears in a Canadian National Park.

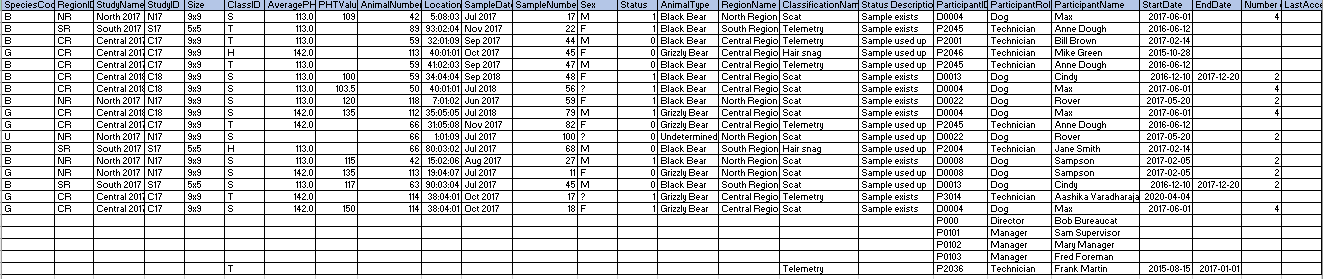
The enterprise problem statement, the example data presented during the interview and the additional data and questions provided during the project were the main inputs for the design

# **List of Functional Dependencies:**

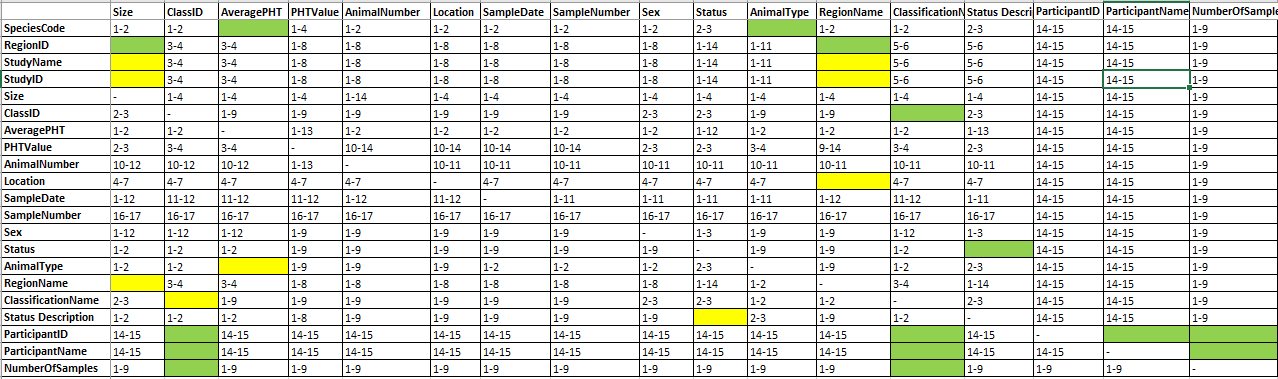
The statements and the sample data provided in the memos aided in deriving a Universal Relation. Using the Universal relation, Functional dependencies were found and marked in a FD sheet. Attached here is the Universal Relation and the FD sheet for reference.

**Universal Relation:**

\*LastAccessDate only applicable for managers



**FD Sheet:**



The following is the list of valid FDs derived from the Universal Relation:

**Singleton:**

{SpeciesCode} -> {AveragePHT}

{SpeciesCode} -> {AnimalType}

{StudyID} -> {RegionID}

{StudyID} -> {StudyName}

{RegionID} -> {RegionName}

{RegionID} -> {Size}

{ClassID} -> {ClassificationName}

{StatusCode} -> {StatusDescription}

{ParticipantID} -> {ParticipantName}

{ParticipantID} -> {PartcipantRole}

{ParticipantID} -> {StartDate}

{ParticipantID} -> {EndDate}

{ParticipantID} -> {NumberOfSamples}

{ParticipantID} -> {LastAccessDate}

{ParticipantID} -> {NextVisitDate}

**Composite:**

{AnimalNumber,StudyID} -> {SpeciesCode}

{SampleNumber,AnimalNumber,StudyID} -> {Sex}

{SampleNumber,AnimalNumber,StudyID} -> {PHTValue}

{SampleNumber,AnimalNumber,StudyID} -> {Location}

{SampleNumber,AnimalNumber,StudyID} -> {StatusCode}

{SampleNumber,AnimalNumber,StudyID} -> {ClassID}

Ideas behind choosing the FDs and entities to derive an LDM:

1. From details given in Memo 1 and from the FDs, we have the following:

{SpeciesCode} -> {AveragePHT}

{SpeciesCode} -> {AnimalType}

The AnimalType and AveragePHT are determined by the SpeciesCode. Hence, these attributes become part of the **Species** entity.

1. From the following FDs:

{RegionID} -> {RegionName}

{RegionID} -> {Size}

We know that the Size and the regionName are determined by the RegionID. Hence, **Region** entity will have these three attributes.

1. From the FDs, we know that StudyID and StudyName are equivalent.

{StudyID} -> {RegionID}

{StudyID} -> {StudyName}

Hence, we choose StudyID to determine the StudyName. Also, Studies are specific to regions, studies determine regions. Hence, the Study entity will have these 3 attributes. Since RegionID which is already part of the Region enity is determined by StudyID in **Study** entity, there will be a **non-identifying relationship** between the Region and Study entities where RegionID becomes a non-key attribute of the study table as Foreign Key.

1. ClassID and ClassificationName are equivalent FDs. Hence, the **Classification**

entity will have the ClassID and ClassificationName which satisfies the FD:

{ClassID} -> {ClassificationName}

1. **Status** entity consists of FD: {StatusCode} -> {StatusDescription}
2. **Participant** entity is derived from the following FDs:

{ParticipantID} -> {ParticipantName}

{ParticipantID} -> {PartcipantRole}

{ParticipantID} -> {StartDate}

{ParticipantID} -> {EndDate}

{ParticipantID} -> {NumberOfSamples}

{ParticipantID} -> {LastAccessDate}

{ParticipantID} -> {NextVisitDate}

Here the participantID determines all other attributes.

1. AnimalNumber and StudyID together defines the Species code. So, the **Animal** entity will have these three attributes satisfying the FD:

{AnimalNumber,StudyID} -> {SpeciesCode}

1. **Sample** entity, until the new information was provided in memo 3, the AnimalNumber and Study ID used to determine Sex, SampleNumber used to be unique to uniquely identify all other attributes. After inclusion of the value 114 in the Animal Number, the AnimalNumber and StudyID no more determines Sex, but together with the SampleNumber, the following FDs are determined:

{SampleNumber,AnimalNumber,StudyID} -> {Sex}

{SampleNumber,AnimalNumber,StudyID} -> {PHTValue}

{SampleNumber,AnimalNumber,StudyID} -> {Location}

{SampleNumber,AnimalNumber,StudyID } -> StatusCode

{SampleNumber,AnimalNumber,StudyID } -> ClassID

This means, there is an identifying relationship from Animal to Sample. In addition, there are 2 non identifying relationships from Status and Classification entities to the Sample.

### Logical Data Model:

\*Zoom in to have a clear picture of the model

1. This logical model defines the structure of the data elements and the relationships between them. This model was formed using the FDs.
2. There were supporting information from the enterprise statements which led to the formation of 6 more entities, which were subcategories of 2 parent entities.
3. From the statement, we understand that, the Samples are classified based on the ClassID. The example data proved that, PHTValues were there only for ScatSample, and the HairAndTelemetrySample didn’t have PHT values set.
4. This resulted in 2 sub-categories of the Sample entity:
5. ScatSample – which had the PHTValues
6. HairAndTelemetrySample which didn’t have PHTValues.

This is to eliminate NULLs in the design, in place of PHTValues.

1. Again, from the enterprise statement it was understood that, there are 4 kinds of participants who are involved in the study of samples.
2. Number of Samples is an attribute that has values only for the Dog Participants. Hence, the Dog entity will have NumberOfSamples. Also, the statement Dog retrieves ScatSample establishes a non-identifying relation from the dog to ScatSample entity, thereby making the Dog participantID as a non-key attribute.
3. The example information proved that; the technicians retrieve the H&T samples. This gives a non-identifying relationship from the Technician to HairAndTelemetrySample entity. Therefore, 4 subcategories of Participants: Dog, Technician, Manager, Director.
4. Also, as part of data access, the Managers must have access to who works on which samples in their region. This establishes a non-identifying relationship between the Region entity and the Manager Entity. In addition, the Director must have access to the last time a manager accessed the database, which is also part of the manager entity.
5. As part of the statement, there will soon be a date attribute that records the director’s next visit.

### Physical Data Model:

\*Zoom in to have a clear picture of the model



In the physical data model the logical model is implemented as a database system. The Domains and datatypes are defined for the columns in the tables.

There are 2 User defined domains created: 1. LocationDom14 – for location and 2. Datedom14 – for Sample date.

Using the forward engineering schema generation, the RDD schema was generated, which helped in creating the tables in the Database – BIC14.

### SCHEMA FOR THE GENERATION OF RELATIONAL DATABASE DESIGN FROM ERWin

USE master

CREATE DATABASE BIC14

CREATE TYPE LocationDom14

FROM CHAR(8) NULL

go

CREATE TYPE DateDom14

FROM CHAR(8) NULL

go

CREATE TABLE Region

(

RegionID char(2) NOT NULL ,

RegionName char(14) NULL ,

Size char(3) NULL ,

PRIMARY KEY CLUSTERED (RegionID ASC)

)

go

CREATE TABLE Study

(

StudyID char(3) NOT NULL ,

StudyName char(15) NULL ,

RegionID char(2) NULL ,

PRIMARY KEY CLUSTERED (StudyID ASC),

FOREIGN KEY (RegionID) REFERENCES Region(RegionID)

)

go

CREATE TABLE Species

(

SpeciesCode char(1) NOT NULL ,

AnimalType char(14) NULL ,

AveragePHT float(2) NULL ,

PRIMARY KEY CLUSTERED (SpeciesCode ASC)

)

go

CREATE TABLE Animal

(

AnimalNumber int NOT NULL ,

StudyID char(3) NOT NULL ,

SpeciesCode char(1) NULL ,

PRIMARY KEY CLUSTERED (AnimalNumber ASC,StudyID ASC),

FOREIGN KEY (StudyID) REFERENCES Study(StudyID),

FOREIGN KEY (SpeciesCode) REFERENCES Species(SpeciesCode)

)

go

CREATE TABLE Classification

(

ClassID char(1) NOT NULL ,

ClassificationName char(10) NULL ,

PRIMARY KEY CLUSTERED (ClassID ASC)

)

go

CREATE TABLE Status

(

StatusCode int NOT NULL ,

StatusDescription char(15) NULL ,

PRIMARY KEY CLUSTERED (StatusCode ASC)

)

go

CREATE TABLE Sample

(

SampleNumber int NOT NULL ,

AnimalNumber int NOT NULL ,

StudyID char(3) NOT NULL ,

ClassID char(1) NULL ,

StatusCode int NULL ,

SampleDate char(8) NULL ,

Location char(8) NULL ,

Sex char(1) NULL ,

PRIMARY KEY CLUSTERED (SampleNumber ASC,AnimalNumber ASC,StudyID ASC),

FOREIGN KEY (AnimalNumber,StudyID) REFERENCES Animal(AnimalNumber,StudyID),

FOREIGN KEY (ClassID) REFERENCES Classification(ClassID),

FOREIGN KEY (StatusCode) REFERENCES Status(StatusCode)

)

go

CREATE TABLE Participant

(

ParticipantID char(5) NOT NULL ,

ParticipantName char(20) NULL ,

ParticipantRole char(12) NULL ,

StartDate date NULL ,

EndDate date NULL ,

PRIMARY KEY CLUSTERED (ParticipantID ASC)

)

go

CREATE TABLE Dog

(

ParticipantID char(5) NOT NULL ,

NumberOfSamples int NULL ,

PRIMARY KEY CLUSTERED (ParticipantID ASC),

FOREIGN KEY (ParticipantID) REFERENCES Participant(ParticipantID)

)

go

CREATE TABLE ScatSample

(

SampleNumber int NOT NULL ,

AnimalNumber int NOT NULL ,

StudyID char(3) NOT NULL ,

PHTValue float(1) NULL ,

ParticipantID char(5) NULL ,

PRIMARY KEY CLUSTERED (SampleNumber ASC,AnimalNumber ASC,StudyID ASC),

FOREIGN KEY (SampleNumber,AnimalNumber,StudyID) REFERENCES Sample(SampleNumber,AnimalNumber,StudyID),

FOREIGN KEY (ParticipantID) REFERENCES Dog(ParticipantID)

)

go

CREATE TABLE Manager

(

ParticipantID char(5) NOT NULL ,

LastAccessDate datetime NULL ,

RegionID char(2) NULL ,

PRIMARY KEY CLUSTERED (ParticipantID ASC),

FOREIGN KEY (ParticipantID) REFERENCES Participant(ParticipantID),

FOREIGN KEY (RegionID) REFERENCES Region(RegionID)

)

go

CREATE TABLE Technician

(

ParticipantID char(5) NOT NULL ,

PRIMARY KEY CLUSTERED (ParticipantID ASC),

FOREIGN KEY (ParticipantID) REFERENCES Participant(ParticipantID)

)

go

CREATE TABLE HairAndTelemetrySample

(

SampleNumber int NOT NULL ,

AnimalNumber int NOT NULL ,

StudyID char(3) NOT NULL ,

ParticipantID char(5) NULL ,

PRIMARY KEY CLUSTERED (SampleNumber ASC,AnimalNumber ASC,StudyID ASC),

FOREIGN KEY (SampleNumber,AnimalNumber,StudyID) REFERENCES Sample(SampleNumber,AnimalNumber,StudyID),

FOREIGN KEY (ParticipantID) REFERENCES Technician(ParticipantID)

)

go

CREATE TABLE Director

(

ParticipantID char(5) NOT NULL ,

NextVisitDate date NULL ,

PRIMARY KEY CLUSTERED (ParticipantID ASC),

FOREIGN KEY (ParticipantID) REFERENCES Participant(ParticipantID)

)

go

### DML INSERT COMMANDS USED FOR LOADING THE DATA:

There were a total of 14 tables in the database. The data of independent tables were loaded first so that, there is no error related to the Foreign Key constraints observed while loading the data.

The following is the order in which the data was loaded:

1. **Species:**

INSERT INTO Species([SpeciesCode],[AnimalType],[AveragePHT])

VALUES ('B','Black Bear',113.0),

('G','Grizzly Bear',142.0),

('U','Undetermined',NULL)

Result:

(3 rows affected)

Completion time: 2020-04-20T14:29:15.2259440-04:00

1. **Region:**

INSERT INTO Region(RegionID,RegionName,Size)

VALUES ('NR','North Region','9x9'),

('SR','South Region','5x5'),

('CR','Central Region','9x9')

Result:

(3 rows affected)

Completion time: 2020-04-20T15:30:20.7747872-04:00

1. **Study:**

INSERT INTO Study(StudyID,StudyName,RegionID)

VALUES ('N17','North 2017','NR'),

('S17','South 2017','SR'),

('C17','Central 2017','CR'),

('C18','Central 2018','CR')

Result:

(4 rows affected)

Completion time: 2020-04-20T15:35:34.5711466-04:00

1. **Classification:**

INSERT INTO Classification (ClassID,ClassificationName)

VALUES ('S','Scat'),

('T','Telemetry'),

('H','Hair snag')

Result:

(3 rows affected)

Completion time: 2020-04-20T15:43:39.3262316-04:00

1. **Status:**

INSERT INTO Status(StatusCode,StatusDescription)

VALUES (0,'Sample exists'),

(1,'Sample used up')

Result:

(2 rows affected)

Completion time: 2020-04-20T15:54:28.8982544-04:00

1. **Participant:**

INSERT INTO Participant (ParticipantID,ParticipantName,ParticipantRole,StartDate,EndDate)

VALUES

('D0004','Max','Dog','2017-06-01',NULL),

('P2045','Anne Dough','Technician','2016-06-12',NULL),

('P2001','Bill Brown','Technician','2017-02-14',NULL),

('P2046','Mike Green','Technician','2015-10-28',NULL),

('D0013','Cindy','Dog','2016-12-10','2017-12-20'),

('D0022','Rover','Dog','2017-05-20',NULL),

('P2004','Jane Smith','Technician','2017-02-14',NULL),

('D0008','Sampson','Dog','2017-02-05',NULL),

('P3014','Aashika Varadharajan','Technician','2020-04-04',NULL),

('P2036','Frank Martin','Technician','2015-08-15','2017-01-01'),

('P0000','Bob Bureaucat','Director', NULL,NULL),

('P0101','Sam Supervisor','Manager', NULL,NULL),

('P0102','Mary Manager','Manager', NULL,NULL),

('P0103','Fred Foreman','Manager', NULL,NULL)

Result:

(14 rows affected)

Completion time: 2020-04-20T16:19:09.0671184-04:00

1. **Animal:**

INSERT INTO Animal(AnimalNumber,StudyID,SpeciesCode)

VALUES

(42,'N17','B'),

(89,'S17','B'),

(59,'C17','B'),

(113,'C17','G'),

(59,'C18','B'),

(50,'C18','B'),

(118,'N17','B'),

(112,'C18','G'),

(66,'C17','G'),

(66,'N17','U'),

(66,'S17','B'),

(113,'N17','G'),

(63,'S17','B'),

(114,'C17','G')

Result:

(14 rows affected)

Completion time: 2020-04-20T16:35:27.6606487-04:00

1. **Sample:**

INSERT INTO Sample (SampleNumber,AnimalNumber,StudyID,ClassID,StatusCode,SampleDate,Location,Sex)

VALUES (17,42,'N17','S',1,'Jul 2017','05:08:03','M'),

(22,89,'S17','T',1,'Nov 2017','93:02:04','F'),

(44,59,'C17','T',0,'Sep 2017','32:01:09','M'),

(45,113,'C17','H',0,'Oct 2017','40:01:01','F'),

(47,59,'C17','T',0,'Sep 2017','41:02:03','M'),

(48,59,'C18','S',1,'Sep 2018','34:04:04','F'),

(56,50,'C18','S',1,'Jul 2018','40:01:01','?'),

(59,118,'N17','S',1,'Jun 2017','07:01:02','F'),

(79,112,'C18','S',1,'Jul 2018','35:05:05','M'),

(82,66,'C17','T',0,'Nov 2017','31:05:08','F'),

(100,66,'N17','S',0,'Jul 2017','01:01:09','?'),

(68,66,'S17','H',0,'Jul 2017','80:03:02','M'),

(27,42,'N17','S',1,'Aug 2017','15:02:06','M'),

(11,113,'N17','S',0,'Jul 2017','19:04:07','F'),

(45,63,'S17','S',0,'Jul 2017','90:03:04','M'),

(17,114,'C17','T',1,'Oct 2017','38:04:01','?'),

(18,114,'C17','S',1,'Oct 2017','38:04:01','F')

Result:

(17 rows affected)

Completion time: 2020-04-20T20:39:31.0067825-04:00

1. **Manager:**

INSERT INTO Manager(ParticipantID,LastAccessDate,RegionID)

VALUES

('P0101','4/14/20 13:00','NR'),

('P0102','4/12/20 17:00','CR'),

('P0103','4/17/20 11:30', 'SR')

Result:

(3 rows affected)

Completion time: 2020-04-20T21:40:07.5846260-04:00

1. **Director:**

INSERT INTO Director(ParticipantID,NextVisitDate)

VALUES('P0000',NULL)

Result:

(1 row affected)

Completion time: 2020-04-20T21:37:34.5985904-04:00

1. **Technician:**

INSERT INTO Technician(ParticipantID)

VALUES

('P2045'),

('P2001'),

('P2046'),

('P2004'),

('P3014'),

('P2036')

Result:

(6 rows affected)

Completion time: 2020-04-20T22:07:58.1502257-04:00

1. **Dog:**

INSERT INTO DOG(ParticipantID,NumberOfSamples)

VALUES('D0004',4),

('D0013',2),

('D0022',2),

('D0008',2)

RESULT:

(4 rows affected)

Completion time: 2020-04-20T22:07:58.1502257-04:00

1. **ScatSample:**

INSERT INTO ScatSample(SampleNumber,AnimalNumber,StudyID,PHTValue,ParticipantID)

VALUES

(17,42,'N17',109,'D0004'),

(48,59,'C18',100,'D0013'),

(56,50,'C18',103.5,'D0004'),

(59,118,'N17',120,'D0022'),

(79,112,'C18',135,'D0004'),

(100,66,'N17',NULL,'D0022'),

(27,42,'N17',115,'D0008'),

(11,113,'N17',135,'D0008'),

(45,63,'S17',117,'D0013'),

(18,114,'C17',150,'D0004')

Result:

(10 rows affected)

Completion time: 2020-04-20T22:07:58.1502257-04:00

1. **HairAndTelemetrySample:**

INSERT INTO HairAndTelemetrySample(SampleNumber,AnimalNumber,StudyID,ParticipantID)

VALUES (22,89,'S17','P2045'),

(44,59,'C17','P2001'),

(45,113,'C17','P2046'),

(47,59,'C17','P2045'),

(82,66,'C17','P2045'),

(68,66,'S17','P2004'),

(17,114,'C17','P3014')

Result:

(7 rows affected)

Completion time: 2020-04-20T22:07:58.1502257-04:00

### SELECT STATEMENTS ON ALL THE TABLES:

1. **SPECIES:**

SELECT \* FROM SPECIES;

SpeciesCode AnimalType AveragePHT

----------- -------------- -------------

B Black Bear 113

G Grizzly Bear 142

U Undetermined NULL

1. rows affected)

Completion time: 2020-04-21T00:49:15.6277802-04:00

1. **REGION:**

select \* from REGION;

RegionID RegionName Size

-------- -------------- ----

CR Central Region 9x9

NR North Region 9x9

SR South Region 5x5

(3 rows affected)

Completion time: 2020-04-21T00:52:53.0529045-04:00

1. **STUDY**

SELECT \* FROM STUDY;

StudyID StudyName RegionID

------- --------------- --------

C17 Central 2017 CR

C18 Central 2018 CR

N17 North 2017 NR

S17 South 2017 SR

1. rows affected)

Completion time: 2020-04-21T00:54:38.2496978-04:00

1. **CLASSIFICATION**

SELECT \* FROM CLASSIFICATION;

ClassID ClassificationName

------- ------------------

H Hair snag

S Scat

T Telemetry

(3 rows affected)

Completion time: 2020-04-21T00:56:15.3897729-04:00

1. **STATUS**

SELECT \* FROM STATUS;

StatusCode StatusDescription

----------- -----------------

0 Sample exists

1 Sample used up

(2 rows affected)

Completion time: 2020-04-21T00:59:56.4010900-04:00

1. **PARTICIPANT:**

SELECT \* FROM PARTICIPANT;

ParticipantID ParticipantName ParticipantRole StartDate EndDate

------------- -------------------- --------------- ---------- ----------

D0004 Max Dog 2017-06-01 NULL

D0008 Sampson Dog 2017-02-05 NULL

D0013 Cindy Dog 2016-12-10 2017-12-20

D0022 Rover Dog 2017-05-20 NULL

P0000 Bob Bureaucat Director NULL NULL

P0101 Sam Supervisor Manager NULL NULL

P0102 Mary Manager Manager NULL NULL

P0103 Fred Foreman Manager NULL NULL

P2001 Bill Brown Technician 2017-02-14 NULL

P2004 Jane Smith Technician 2017-02-14 NULL

P2036 Frank Martin Technician 2015-08-15 2017-01-01

P2045 Anne Dough Technician 2016-06-12 NULL

P2046 Mike Green Technician 2015-10-28 NULL

P3014 Aashika Varadharajan Technician 2020-04-04 NULL

(14 rows affected)

Completion time: 2020-04-22T19:41:30.6157743-04:00

1. **ANIMAL**

SELECT \* FROM ANIMAL;

AnimalNumber StudyID SpeciesCode

------------ ------- -----------

42 N17 B

50 C18 B

59 C17 B

59 C18 B

63 S17 B

66 C17 G

66 N17 U

66 S17 B

89 S17 B

112 C18 G

113 C17 G

113 N17 G

114 C17 G

118 N17 B

(14 rows affected)

Completion time: 2020-04-21T01:04:39.7065455-04:00

1. **SAMPLE:**

SELECT \* FROM SAMPLE;

SampleNumber AnimalNumber StudyID ClassID StatusCode SampleDate Location Sex

------------ ------------ ------- ------- ----------- ---------- -------- ----

11 113 N17 S 0 Jul 2017 19:04:07 F

17 42 N17 S 1 Jul 2017 05:08:03 M

17 114 C17 T 1 Oct 2017 38:04:01 ?

18 114 C17 S 1 Oct 2017 38:04:01 F

22 89 S17 T 1 Nov 2017 93:02:04 F

27 42 N17 S 1 Aug 2017 15:02:06 M

44 59 C17 T 0 Sep 2017 32:01:09 M

45 63 S17 S 0 Jul 2017 90:03:04 M

45 113 C17 H 0 Oct 2017 40:01:01 F

47 59 C17 T 0 Sep 2017 41:02:03 M

48 59 C18 S 1 Sep 2018 34:04:04 F

56 50 C18 S 1 Jul 2018 40:01:01 ?

59 118 N17 S 1 Jun 2017 07:01:02 F

68 66 S17 H 0 Jul 2017 80:03:02 M

79 112 C18 S 1 Jul 2018 35:05:05 M

82 66 C17 T 0 Nov 2017 31:05:08 F

100 66 N17 S 0 Jul 2017 01:01:09 ?

(17 rows affected)

Completion time: 2020-04-21T01:08:19.9393499-04:00

1. **SCATSAMPLE**

SELECT \* FROM SCATSAMPLE;

SampleNumber AnimalNumber StudyID PHTValue ParticipantID

------------ ------------ ------- ------------- -------------

11 113 N17 135 D0008

17 42 N17 109 D0004

18 114 C17 150 D0004

27 42 N17 115 D0008

45 63 S17 117 D0013

48 59 C18 100 D0013

56 50 C18 103.5 D0004

59 118 N17 120 D0022

79 112 C18 135 D0004

100 66 N17 NULL D0022

(10 rows affected)

Completion time: 2020-04-21T01:09:55.8920200-04:00

1. **HAIRANDTELEMETRYSAMPLE:**

SELECT \* FROM HAIRANDTELEMETRYSAMPLE;

SampleNumber AnimalNumber StudyID ParticipantID

------------ ------------ ------- -------------

17 114 C17 P3014

22 89 S17 P2045

44 59 C17 P2001

45 113 C17 P2046

47 59 C17 P2045

68 66 S17 P2004

82 66 C17 P2045

(7 rows affected)

Completion time: 2020-04-21T01:18:08.8636126-04:00

1. **DOG**

SELECT \* FROM DOG;

ParticipantID NumberOfSamples

------------- ---------------

D0004 4

D0008 2

D0013 2

D0022 2

(4 rows affected)

Completion time: 2020-04-21T01:23:02.3221796-04:00

1. **TECHNICIAN**

SELECT \* FROM TECHNICIAN;

ParticipantID

-------------

P2001

P2004

P2036

P2045

P2046

P3014

(6 rows affected)

Completion time: 2020-04-21T01:43:40.7048478-04:00

1. **MANAGER:**

SELECT \* FROM MANAGER;

ParticipantID LastAccessDate RegionID

------------- ----------------------- --------

P0101 2020-04-14 13:00:00.000 NR

P0102 2020-04-12 17:00:00.000 CR

P0103 2020-04-17 11:30:00.000 SR

(3 rows affected)

Completion time: 2020-04-21T01:44:31.2705266-04:00

1. **DIRECTOR:**

SELECT \* FROM DIRECTOR;

ParticipantID NextVisitDate

------------- -------------

P0000 NULL

(1 row affected)

Completion time: 2020-04-21T01:45:19.0470442-04:00

### RESULTS FOR THE QUESTIONS FROM THE PROTOTYPE DB

**Question 1:** What is the largest physiological health value observed for a black bear?

**Query:**

SELECT max(ss.PHTValue) AS MaxPHTValueofBlackBear

FROM ScatSample ss,Species sp,Animal a

WHERE ss.AnimalNumber = a.AnimalNumber

AND a.SpeciesCode = sp.SpeciesCode

AND sp.AnimalType = 'Black Bear'

**Result:**

MaxPHTValueofBlackBear

----------------------

120

Warning: Null value is eliminated by an aggregate or other SET operation.

(1 row affected)

Completion time: 2020-04-21T01:48:53.6806753-04:00

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

**Query:**

SELECT s.AnimalNumber, s.SampleDate, c.ClassificationName

FROM Sample s, Classification c

WHERE s.ClassID = c.ClassID

ORDER BY s.AnimalNumber, CAST(s.SampleDate AS date)

**Result:**

AnimalNumber SampleDate ClassificationName

------------ ---------- ------------------

42 Jul 2017 Scat

42 Aug 2017 Scat

50 Jul 2018 Scat

59 Sep 2017 Telemetry

59 Sep 2017 Telemetry

59 Sep 2018 Scat

63 Jul 2017 Scat

66 Jul 2017 Hair snag

66 Jul 2017 Scat

66 Nov 2017 Telemetry

89 Nov 2017 Telemetry

112 Jul 2018 Scat

113 Jul 2017 Scat

113 Oct 2017 Hair snag

114 Oct 2017 Telemetry

114 Oct 2017 Scat

118 Jun 2017 Scat

(17 rows affected)

Completion time: 2020-04-21T01:50:16.7090367-04:00

**Question 3:** In what region is the Central 2017 study and what size grid pattern is used on that study?

**Query:**

SELECT S.StudyName, R.RegionID, R.Size

FROM Study S, Region R

WHERE S.StudyName = 'Central 2017'

AND S.RegionID = R.RegionID

Result:

StudyName RegionID Size

--------------- -------- ----

Central 2017 CR 9x9

(1 row affected)

Completion time: 2020-04-21T01:51:16.5645994-04:00

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

**Query:**

SELECT DISTINCT ss.AnimalNumber

FROM ScatSample ss, Animal a, Species s, Sample sa

WHERE ss.SampleNumber = sa.SampleNumber

AND ss.AnimalNumber = a.AnimalNumber

AND a.SpeciesCode = s.SpeciesCode

AND ABS(ss.PHTValue - s.AveragePHT)<5

**Result:**

AnimalNumber

------------

42

63

(2 rows affected)

Completion time: 2020-04-21T01:52:35.5091551-04:00

**QUESTION 5:** List the sample information for the Central 2017 studies made in September 2017 and in November 2017.

**QUERY:**

SELECT Sa.SampleNumber, sa.AnimalNumber, sa.StudyID,sa.ClassID, sa.StatusCode,sa.SampleDate, sa.Location, sa.Sex

FROM Sample sa, Study st

WHERE Sa.StudyID = St.StudyID

AND St.StudyName = 'Central 2017'

AND Sa.SampleDate in ('Sep 2017', 'Nov 2017')

**Result:**

SampleNumber AnimalNumber StudyID ClassID StatusCode SampleDate Location Sex

------------ ------------ ------- ------- ----------- ---------- -------- ----

44 59 C17 T 0 Sep 2017 32:01:09 M

47 59 C17 T 0 Sep 2017 41:02:03 M

82 66 C17 T 0 Nov 2017 31:05:08 F

(3 rows affected)

Completion time: 2020-04-21T01:54:34.1392087-04:00

**QUESTION 6:** What studies have animals for which samples were gathered in July 2017?

Query:

SELECT Study.StudyName, Sample.AnimalNumber, Sample.SampleNumber

FROM Sample,Study

WHERE SampleDate = 'Jul 2017'

AND Sample.StudyID = Study.StudyID

**Result:**

StudyName AnimalNumber SampleNumber

--------------- ------------ ------------

North 2017 113 11

North 2017 42 17

South 2017 63 45

South 2017 66 68

North 2017 66 100

(5 rows affected)

Completion time: 2020-04-21T01:58:36.4708252-04:00

**QUESTION 7:** What type of samples were collected in the South 2017 study?

Query:

SELECT Cl.ClassificationName, Sa.SampleNumber,Sa.AnimalNumber

FROM SAMPLE Sa, Classification Cl, Study st

WHERE sa.ClassID = Cl.ClassID

AND sa.StudyID = st.StudyID

AND st.StudyName = 'South 2017'

**Result:**

ClassificationName SampleNumber AnimalNumber

------------------ ------------ ------------

Telemetry 22 89

Scat 45 63

Hair snag 68 66

(3 rows affected)

Completion time: 2020-04-21T02:01:56.3575015-04:00

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

Query:

SELECT M.ParticipantID, P.ParticipantName

FROM Manager M, Participant P, Region R

WHERE M.ParticipantID = P.ParticipantID

AND M.RegionID = R.RegionID

AND R.RegionName = 'Central Region'

**Result:**

ParticipantID ParticipantName

------------- --------------------

P0102 Mary Manager

(1 row affected)

Completion time: 2020-04-21T02:03:25.1842849-04:00

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

Query:

SELECT M.ParticipantID, P.ParticipantName

FROM Manager M, Participant P, Region R

WHERE M.ParticipantID = P.ParticipantID

AND M.RegionID = R.RegionID

AND R.RegionName = 'South Region'

Result:

ParticipantID ParticipantName

------------- --------------------

P0103 Fred Foreman

(1 row affected)

Completion time: 2020-04-21T02:05:01.3602317-04:00

**QUESTION 10:** How many times has each animal has been sampled?

Query:

SELECT AnimalNumber, COUNT(AnimalNumber) AS NumberOfTimesSampled

FROM Sample

GROUP BY AnimalNumber

Result:

AnimalNumber NumberOfTimesSampled

------------ --------------------

42 2

50 1

59 3

63 1

66 3

89 1

112 1

113 2

114 2

118 1

(10 rows affected)

Completion time: 2020-04-21T02:06:18.7696944-04:00

**QUESTION 11:** List the samples analyzed by P2045.

Query:

SELECT ht.SampleNumber

FROM HairAndTelemetrySample ht, Participant P, Sample S

WHERE ht.ParticipantID = 'P2045'

AND ht.ParticipantID = p.ParticipantID

AND ht.SampleNumber = s.SampleNumber

Result:

SampleNumber

------------

22

47

82

(3 rows affected)

Completion time: 2020-04-21T02:10:51.4811001-04:00

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

**Query:**

SELECT st.StudyName, P.ParticipantName

FROM Dog d, Participant P, ScatSample S, Study st

WHERE d.ParticipantID = P.ParticipantID

AND d.ParticipantID = s.ParticipantID

AND S.StudyID = st.StudyID

**Result:**

StudyName ParticipantName

--------------- --------------------

North 2017 Sampson

North 2017 Max

Central 2017 Max

North 2017 Sampson

South 2017 Cindy

Central 2018 Cindy

Central 2018 Max

North 2017 Rover

Central 2018 Max

North 2017 Rover

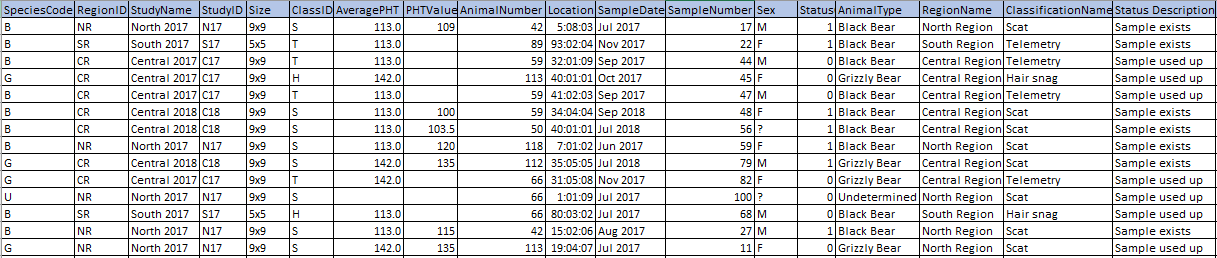
(10 rows affected)

Completion time: 2020-04-21T02:11:56.7007547-04:00

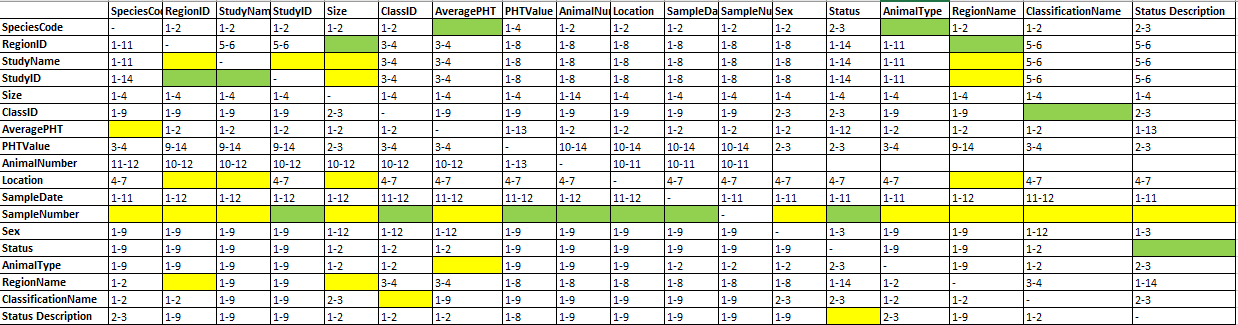
## APPENDIX – Previous versions of the designs

VERSION 1:

**UNIVERSAL RELATION:**



**FD SHEET:**



**List of valid FDs that makes the design 3NF:**

{SpeciesCode} -> {AveragePHT}

{SpeciesCode} -> {AnimalType}

{StudyID} -> {RegionID}

{StudyID} -> {StudyName}

{RegionID} -> {RegionName}

{RegionID} -> {Size}

{ClassID} -> {ClassificationName}

{StatusCode} -> {StatusDescription}

{AnimalNumber,StudyID} -> {SpeciesCode}

{AnimalNumber,StudyID} -> {Sex}

{SampleNumber, AnimalNumber,StudyID } -> {ClassID}

{SampleNumber, AnimalNumber,StudyID } -> {Status}

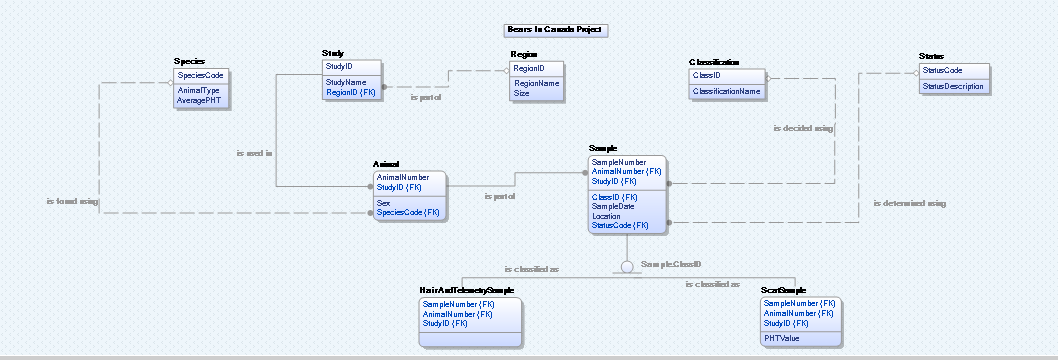
{SampleNumber, AnimalNumber,StudyID } -> {Location}

{SampleNumber, AnimalNumber,StudyID } -> {SampleDate}

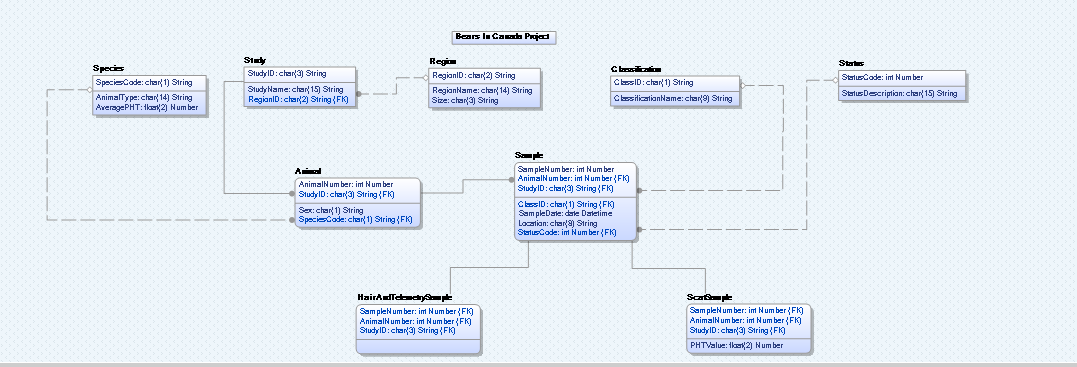
{SampleNumber, AnimalNumber,StudyID } -> {PHTValue}

{SampleNumber, AnimalNumber,StudyID } -> {Status}

Logical Data Model:



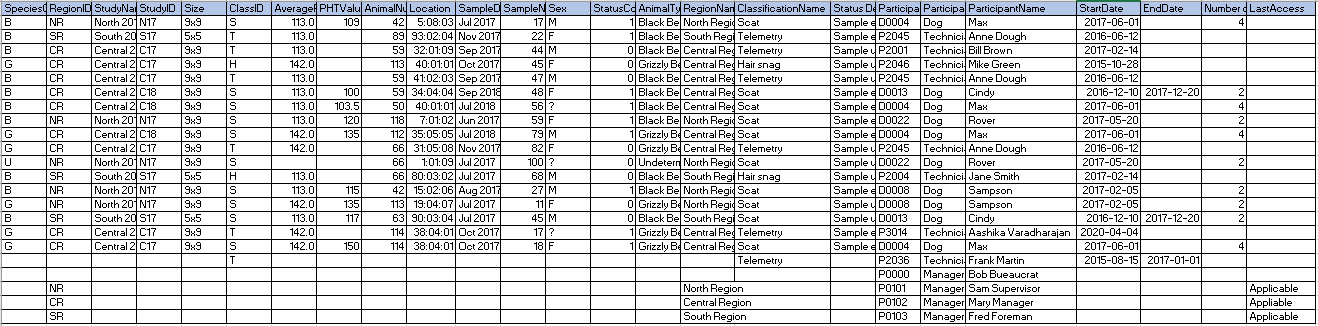
Physical Data Model:



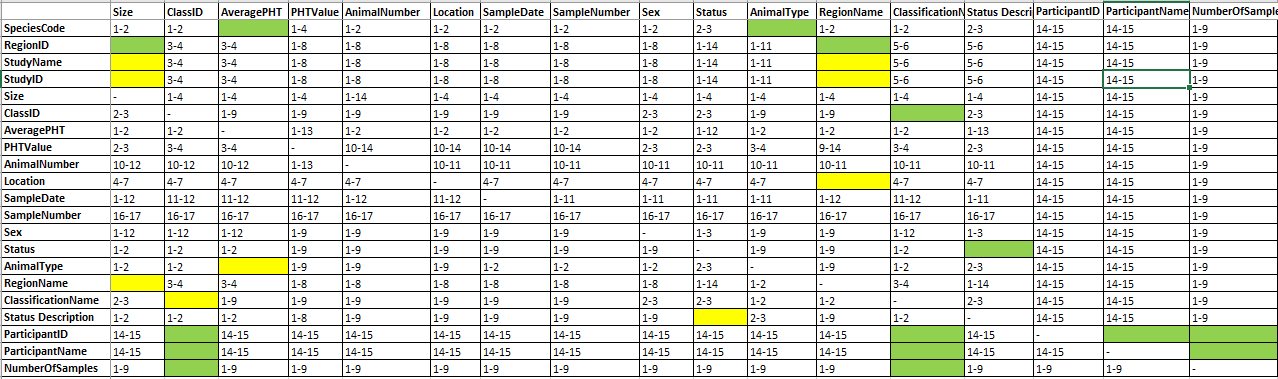
VERSION 2:

**UNIVERSAL RELATION:**

\*Zoom in to look into the data



**FD Sheet:**



**List of valid FDs:**

{SpeciesCode} -> {AveragePHT}

{SpeciesCode} -> {AnimalType}

{StudyID} -> {RegionID}

{StudyID} -> {StudyName}

{RegionID} -> {RegionName}

{RegionID} -> {Size}

{ClassID} -> {ClassificationName}

{StatusCode} -> {StatusDescription}

{ParticipantID} -> {ParticipantName}

{ParticipantID} -> {PartcipantRole}

{ParticipantID} -> {StartDate}

{ParticipantID} -> {EndDate}

{ParticipantID} -> {NumberOfSamples}

{ParticipantID} -> {LastAccessDate}

{AnimalNumber,StudyID} -> {SpeciesCode}

{SampleNumber,AnimalNumber,StudyID} -> {Sex}

{SampleNumber,AnimalNumber,StudyID} -> {PHTValue}

{SampleNumber,AnimalNumber,StudyID} -> {Location}

{SampleNumber,AnimalNumber,StudyID} -> {StatusCode}

{SampleNumber,AnimalNumber,StudyID} -> {ClassID}