

CIS 2334 Semester Project

Part 1

Abalone is a common name for any of a group of small to very large marine gastropod mollusks in the family Haliotidae. The flesh of abalones is widely considered to be a desirable food and is consumed raw or cooked by a variety of cultures. The haliotid family has a worldwide distribution. The majority of abalone species are found in cold waters, such as off the coasts of New Zealand, South Africa, Australia, Western North America, and Japan.



In February 2021, a group of marine biologists (collectors) want to study the diversity of the abalone in the US and monitor the species' biological statistics. Funded by different universities, 15 marine biologists have teamed up and formed an investigation squad. They have traveled across the country and collected thousands of abalone samples from different coastal area in the US.

After an abalone sample is collected the marine biologist would measure a group of biological statistics and document the collector's information, as well as the water region information where the abalone has been collected.

The biological statistics the marine biologists measured include:

Gender (text) – abalone's gender: M (male), F(female), I(infant)

Length (continuous) – Longest shell measurement

Diameter (continuous) – Perpendicular to length

Height (continuous) – abalone's height with meat in shell

Whole weight (continuous) – abalone's total weight
Shucked weight (continuous) – weight of meat
Viscera weight (continuous) – abalone's gut weight
Shell weight (continuous) – abalone's dry shell weight
Rings (integer) – the number of rings on abalone

The collector information includes collector first and last names and the collector organization. Water region information includes the name of the region where the abalone has been collected and the water temperature in degree centigrade (degree Celsius) when the abalone has been collected.

As data collection work going to an end, the research team realized that they have collected a huge amount of data, however, none of them are experts in data management. The research team is using a merged csv file to store the information on the cloud. Soon they noticed that they are facing with many problems, such as how to store the data safely; how to handle the data input simultaneously from different team members; how to avoid human errors while data entering; how to search efficiently in the data; etc...

For the above reasons, the research team has contacted “you”, an expert in database, to get some professional help. They would like you to create a database for them to handle their valuable (expensive and useful) data. They would like to see a prototype, and they already have made a few simple requests with the prototype database.

To satisfy the research team's requirements, you must use your DB knowledge and you must demonstrate that your database is fulfilling the requirements. For that purpose, perform the following tasks:

1. Create the database:
 - a. Import the csv file into a MySQL database.
 - b. Validation: you need to use a SQL command to check the import results and verify that the data has been successfully imported into the database. The SQL command to use will show the tables in the database.
 - c. File submission: Export the database (dump the database out of the server). You need to submit the dumped file (.sql). File name: Firstname_ Lastname_FullDB.sql (with FirstName and LastName being your own first name and last name).
 - d. Report submission:
 - i. Create an MS Word document
 - ii. Title: CIS 2334 Semester Project - Part 1.
 - iii. Include your Last Name, your First Name and your cougar ID.
 - iv. Use the subtitle “Question 1”.
 - v. Paste all the SQL commands you have used for this step.
 - vi. Paste Screen shots of all the results.
 - vii. Name your word file: Firstname_ Lastname_ProjectPart1.doc(x).
2. Conduct a query and find all the abalones collected in Florida.
 - a. File submission: Export the results in a csv file.
Name your csv file Firstname_ Lastname_Florida.csv
 - b. You need to use a SQL command to find the all the desired results.
 - c. Report submission:
 - i. Use the same MS Word document.

- ii. Use the subtitle "Question 2".
 - iii. Paste all the SQL commands you have used for this step.
 - iv. Paste Screen shots of all the results.
- 3. Conduct a query and find all the abalones that are infants.
 - a. File submission: Export the results in a csv file.
Name your csv file Firstname_ Lastname_Infant.csv.
 - b. You need to use a SQL command to find the all the desired results.
 - c. Report submission:
 - i. Use the same MS Word document.
 - ii. Use the subtitle "Question 3".
 - iii. Paste all the SQL commands you have used for this step.
 - iv. Paste Screen shots of all the results.
- 4. Conduct a query and find all the abalones that are collected in the water with a temperature warmer than x degrees Celsius, x being the last 2 digits of your student ID divided by 5. For example, if your student id is 12345678, x is $78/5 = 15.6$. You will need to find all the abalones that are collected from the water that is warmer than 15.6 degrees.
 - a. File submission: Export the results in a csv file.
Name your csv file Firstname_ Lastname_Warmer.csv
 - b. You need to use a SQL command to find the all the desired results.
 - c. Report submission:
 - i. Use the same MS Word document.
 - ii. Use the subtitle "Question 4".
 - iii. Paste all the SQL commands you have used for this step.
 - iv. Paste Screen shots of all the results.

5. Insert the following new data into the database

Gender	Length	Diameter	Height	Whole_weight	Shucked_weight	Viscera_weight	Shell_weight	Rings
M	0.71	0.555	0.195	1.9485	0.9455	0.3765	0.495	12

Collector: Jan Odom, University of Houston

Water info: Florida Atlantic Coast; 71.8 degrees Fahrenheit.

- a. File submission: Export the results in a csv file.
Name your csv file Firstname_ Lastname_New.csv.
 - b. You need to use a SQL command to find the all the desired results.
 - c. Report submission:
 - i. Use the same MS Word document.
 - ii. Use the subtitle "Question 5".
 - iii. Paste all the SQL commands you have used for this step.
 - iv. Paste Screen shots of all the results.
- 6. Delete the data for which the "Whole_weight" is larger than y, y being the last digit of your student id divided by 4. For example, if your student id is 12345678, y is $8/4 = 2$. You will need to delete all the data for which the "Whole_weight" is larger than 2.

- a. File submission: Export the results after the requested deletion in a csv file.
Name your csv file Firstname_Lastname_Weight.csv.
- b. You need to use a SQL command to find the all the desired results.
- c. Report submission:
 - i. Use the same MS Word document.
 - ii. Use the subtitle "Question 6".
 - iii. Paste all the SQL commands you have used for this step.
 - iv. Paste Screen shots of all the results.

7. Bonus (1% of total grade).

In this DB, we first consider that the only **Entity** is abalone. Define the **Primary Key** for that **Entity**.

- a. Find the shorthand representation of the corresponding table. Find all the functional dependences.
- b. Is the table in the first NF? If not, explain why not and then perform the first normalization.
- c. Is the table in the second NF? If not, explain why not and then perform the second normalization.
- d. Is the table in the third NF? If not, explain why not and then perform the third normalization.
- e. Create a CSV files for each table that has been created by the normalizations.
- f. File submission: Name your csv files Firstname_Lastname_Abalone.csv, Firstname_Lastname_Collectors.csv and Firstname_Lastname_Water.csv.
- g. Draw by hand an Entity-Relationship Diagram using the crow foot representation. Take a picture of it.
- h. Report submission:
 - i. Use the same MS Word document.
 - ii. Use the subtitle "Question 7".
 - iii. Insert all the tables and Entity-Relationship Diagram.

8. Export the word document into a PDF file.

Name your PDF file: Firstname_Lastname_ProjectPart1.pdf.

Submit all the 7 files (.sql, .pdf and .csv files) to Blackboard.

(10 files if you choose to do the Bonus question).