Exam #2 Review
Baseball Statistics Problem

Background Information

The sport of baseball is known as America's favorite pasttime. Each spring, teams begin a season-long drive to win the championship pennant for their league. One goes on to win the World Series.

As part of the quest, the winningest team in each division along with the team with the most wins that was not a division champion participates in post-season playoffs.

Problem Statement

In this assignment, students will explore team stats from Major League Baseball teams in 2007 through 2009. By analyzing this data, trends and relations between different statistics can be investigated.

Instructions

IMPORTANT: This is not the actual Exam for your section. You will not receive any credit for completing this project.

IMPORTANT: This assignment requires the Windows version of Microsoft Office.

IMPORTANT: Complete the steps below in the order they are given. Completing the steps out of order may complicate the assignment or result in an incorrect result.

1. Download and extract the provided Data Files ZIP file. It contains the following files for use in this assignment:

    a. **seasons.xml** – Team statistics for the 2007 through 2009 Major League Baseball seasons [1]–[3].

Table: *Seasons*		
Field Name	**Type**	**Description**
Season	Number	Part of a composite key. Year of the baseball season.
TeamAbbrv	Short Text	Part of a composite key. Team name abbreviation.
Wins	Number	Games won.
Losses	Number	Games lost.
RunsScored	Number	Number of runs (points) this team scored against opponents.
RunsAllowed	Number	Number of runs (points) opponents scored against this team.
DivisionChamp	Yes/No	Indicates if the team was a division champion.
Wildcard	Yes/No	Indicates if this team was a wildcard selection for the playoffs.

    b. **teams.xml** – List of teams for each league and division of Major League Baseball [4], [5].

Table: *Teams*		
Field Name	**Type**	**Description**
TeamAbbrv	Short Text	Primary key. Team name abbreviation.
TeamName	Short Text	Full name of the baseball team.
LeagueAbbrv	Short Text	Abbreviation of league to which this team belongs.
Division	Short Text	Division to which this team belongs.
Established	Number	Year the team entered major league play.
Valuation	Currency	2016 financial valuation of the team.

2. Begin by creating a new Microsoft Access database named ***lastname_firstname*_er2_bsp.accdb**.

3. Import the following items into the database:

    a. **seasons.xml** file – Import structure and data into a new table.

    b. **teams.xml** file – Import structure and data into a new table.

4. We need to create a new table to store information about each Major League Baseball league.

a.   Create a table named *Leagues* with the fields below. Use appropriate field types and designate a good primary key.

Table: *Leagues*	
Field Name	**Description**
LeagueAbbrv	Abbreviation of the league.
LeagueName	Full name of the league.

b.   Enter records for all leagues below.

HINT: The *Leagues* table will contain 2 records.

LeagueAbbrv	LeagueName
AL	American League
NL	National League

5.   Create the appropriate relationships for the following tables. Enforce referential integrity, but do not enable cascade updates or cascade deletes.

a.   *Seasons* and *Teams*

b.   *Teams* and *Leagues*

6.   Create separate queries to provide the information requested below. Name each query after the step in which it appears (e.g., name the query in Step 6a as *Query6A*).

HINT: Run your queries to test them. Make sure that they display all and only the records that you would expect to appear.

a.   Create a query to compare team offenses and defenses. List the season, team name, runs scored, and runs allowed. Also, include a field to calculate the ratio of runs scored to runs allowed.

You can calculate the ratio of runs scored to runs allowed using the formula:

$$\frac{[Seasons.RunsScored]}{[Seasons.RunsAllowed]}$$

Format the calculated field as a standard-type number with 2 decimal places. Sort by the ratio in descending order.

HINT: This query will show 90 records and 5 fields.

b. We wish to calculate the win percentage for each team. Create a query listing the team name, wins, and losses. Also, include a field to calculate the win percentage.

You can calculate the win percentage using the formula:
$$\frac{[Seasons.Wins]}{([Seasons.Wins] + [Seasons.Losses])}$$

Only display records for the 2009 season. Do not show the season in the results.

Format the calculated field as a fixed-type number with 3 decimal places. Sort by win percentage in descending order.

HINT: This query will show 30 records and 4 fields.

c. Create a query to summarize team records. List, for each team name, the sum of wins and sum of losses.

Format the sums as standard-type numbers with no decimal places. Sort by team name in ascending order.

HINT: This query will show 30 records and 3 fields.

d. We wish to calculate statistics on division champions. Create a query listing, for each season, if the teams are division champions, the average of wins, and average of losses.

Only display records for division champions.

Format the averages as standard-type numbers with 1 decimal place.

HINT: This query will show 3 records and 4 fields.

e. We want to compare records for each league division. Copy-and-paste this SQL code into a new query:

```
SELECT Leagues.LeagueName, Teams.Division,
Avg(Teams.Valuation) AS AvgOfValuation, Round(Avg([Wins]),1)
AS AvgOfWins, Round(Avg([Losses]),1) AS AvgOfLosses,
Round(Avg([RunsScored]),1) AS AvgOfRunsScored,
Round(Avg([RunsAllowed]),1) AS AvgOfRunsAllowed
FROM (Leagues INNER JOIN Teams ON Leagues.LeagueAbbrv =
Teams.LeagueAbbrv) INNER JOIN Seasons ON Teams.TeamAbbrv =
Seasons.TeamAbbrv
GROUP BY Leagues.LeagueName, Teams.Division
ORDER BY Leagues.LeagueName, Teams.Division;
```

IMPORTANT: Do not make any modifications to the query other than entering the above SQL code.

HINT: This query will show 6 records and 7 fields.

7. Using the Form Wizard, create a form with a subform. The main form should display the league name. The subform must display a datasheet with all *Teams* table fields.

   Name the main form *LeagueData* and the subform *LeagueDataSubform*.

8. Using the Report Wizard, create a report to show the results of the *Query6E* query. Display all fields from the query.

   Group by league name and sort by division in ascending order. Use a stepped layout and landscape page orientation. Name the report *LeagueSummary*.

   Ensure the full widths of all columns are visible on the report.

9. We need to create a new table to store analysis questions responses.

   a. Create a table named *AnalysisQuestions* with the fields below. Use appropriate field types and designate a good primary key.

Table: *AnalysisQuestions*	
Field Name	**Description**
QuestionNumber	Question being answered.
Response	Response to the analysis question prompt.

10. In the *AnalysisQuestions* table, answer four of the five analysis questions below. Respond to one question per record.

    a. While most teams play 162 games throughout a single season, why do some play 163?

    b. While most teams play 162 games, why do some teams not play all 161 games?

    c. What does the Runs Scored-to-Runs Allowed ratio tell us? Do think this statistic closely correlates with that of win percentage?

    d. Based on previous season data, make a claim as to what win percentage a team needs to make the playoffs. Explore previous season data on win percentages for each teams, and see what the worst win percentages were for teams that made the playoffs.

    e. What is a "good" win percent? Explore previous season data on win percentages for each team, and see what win percentages the best teams had.

11. Run the Compact and Repair Database utility on your database. Ignore any errors you receive when running the utility.

Grading Rubric

This is a practice assignment and is worth no points. A comparable Exam would be worth 100 points and graded using this rubric, with partial credit awarded as appropriate:

Step 3	7.5 points	Step 7	7.5 points
Steps 4a-b	10 points total	Step 8	7.5 points
Step 5	7.5 points	Step 9a	2.5 points
Steps 6a-e	7.5 points each	Steps 10a-e (pick 4 of 5)	5 points each

The analysis questions in Steps 10a-e can be evaluated using this rubric:

Standard	Meets Requirements (2.5 points)	Does Not Meet Requirements (0 points)
Answer is reasonable.	Answer addresses the question prompt and is factually correct or a reasonable interpretation of available data.	Answer does not address the question prompt, is factually incorrect, or is an unreasonable interpretation of available data.
Answer is supported.	Logical rationale is provided to support the given answer.	Logical rationale is not provided to support the given answer.

References

[1] "2007 MLB Season History," *ESPN.com*. Available: *http://espn.go.com/mlb/history/season/_/year/2007*.

[2] "2008 MLB Season History," *ESPN.com*. Available: *http://espn.go.com/mlb/history/season/_/year/2008*.

[3] "2009 MLB Season History," *ESPN.com*. Available: *http://espn.go.com/mlb/history/season/_/year/2009*.

[4] "Current Major League Baseball franchises," *Wikipedia*. Available: *https://en.wikipedia.org/w/index.php?title=Current_Major_League_Baseball_franchises&oldid=675940837*.

[5] M. Ozanian, "Baseball's Most Valuable Teams," *Forbes*, Mar. 23, 2016. Available: *http://www.forbes.com/sites/mikeozanian/2016/03/23/baseballs-most-valuable-teams/*.