New York Times Article Database

Timeline

- Due Date: April 15, 2019
- Cutoff Date: April 17, 2019
- Late penalty: Five points will be deducted each calendar day submitted after the due date.
- No projects will be accepted after the cutoff date.

Objective

- Create a database to manage articles, topics and reporters for *The New York Times* newspaper
- Create reports

New York Times database requirements

- Reporters write articles. Track the reporter name, salary and office assigned. Reporters can be assigned to many offices.
- Offices are located throughout the world. Track the office street, city, state, zip code and country.
- Articles are written by reporters. Track the article title, publish date, body (which includes all text of the article), author and status (draft, waiting approval, published, etc.) One article can be written by many reporters.
- Articles are assigned at least one topic. For instance, some topics are *baseball, space shuttle, Mars, artificial intelligence, Middle East, Brexit, election, etc*. One article can be assigned many topics.
- Articles can optionally include a photo. Articles can include one or several photos.
- Articles are assigned to a section of the newspaper. For instance, an article about the Yankees will be in the Sports section. Some other newspaper sections are international, politics, business, New York metropolitan, opinion, obituary, etc.

Data

You must enter at least the following data

- At least 15 articles
- At least 10 reporters
- At least 5 offices
- At least 20 topics
- At least 10 sections
- At least 50% of the articles will have photos
- At least 50% of the articles will be assigned to more than one reporter

Questions

- Create SQL to answer the following queries.
- Replace underlined items with values of your own choosing. For instance, replace the underlined terms Brexit, the last six months and New York reporters with your own values.
- Create data to insure all questions generate output.
- Format all output. For instance, all numbers will display with commas, dollar values will display with a $ prefix and create descriptive labels for all columns.

1. Identify published articles in the last six months about Brexit and immigration. Display the article title, publish date and authors.
2. Identify articles in the last year about the topic education and the article body includes references to CUNY and financial terms. Display the article title and publish date. Use a nested select to answer this question. Order the articles chronologically.
3. Identify topics not assigned to published articles in the last six months. Display the topics. Use a nested select to answer this question.
4. Identify the top five topics in the last year. Display two columns: topic and number of articles. Display the topic with the most articles first. Display one row for each distinct topic. Use a function to answer this question.
5. Identify the number of articles by reporter in the last five years. Display three columns: Reporter, year published and number of articles. Display one row for each distinct reporter and year. Order the output by reporter name. Use a function to answer this question.
6. Identify photo only articles about the Yankees in the last six months. Display the article title and publish date. Use a nested select to answer this question. Note, a photo only article has no text in the body and a photo.
7. Identify New York reporters with no completed or draft articles in the last six months. Display the reporter. Order the output by reporter name. Use a nested select to answer this question.
8. Identify all articles published today. Display the section, article title, authors, topics and page numbers. Order the output by section, page number (where the article appears in the newspaper) and article title.
9. Identify articles published today without a topic. Display the article and authors. Use a nested select to answer this question.
10. Identify number of articles and the total salary by office in the last year. Display three columns: office location, number of articles and total salary. Display one row for each distinct office. The office with the highest salary will be displayed first. Use a function to answer this question.
11. Close the Arlington Virginia office and re-assign all reporters to the Washington DC office. Identify the SQL code to implement.
12. Display the structure of ALL tables using SQL Describe.
13. Display the version of Oracle. Enter:

```
SELECT *
FROM v$version;
```

Additional Design Requirements

- Include all SQL commands to create your database and answer the questions including create tables, select, update, insert data, alter column names and alter column types.
- Create your database using Oracle version 18c. Projects created with other databases will be rejected and not graded. Utilizing other databases requires prior instructor approval.
- Normalize your database to third normal form.
- All multi-value columns must be saved to their own table.
- Output for all questions must include at least one row displayed.
- Identify and create primary keys for all tables.
- Create foreign keys to enforce referential integrity. For instance, you must have foreign keys with references to ***at least*** the following:
    a. Articles and reporters
    b. Reporter and offices
    c. Article and topics
    d. Article and photos
- Include the question, SQL command to answer the question and output from the SQL command.
- Create descriptive column labels for all output. For instance, don't display a column label named *count(*)*

Formatting

- The column output should be displayed in a `non-proportional font such as courier`. This will display the columns vertically straight.
- All columns in your search must display on one line. Don't wrap columns to two lines.
- Your project must be typed.
- All pages of your output must include the following in the header: name, class, date and project number.
- The first page of your project must include your name, the last four digits of your student id, class, submission date and the project number.

Submission

- Review the grading rubric on Blackboard to identify how the project will be evaluated and graded.
- Projects are due on the due date. No projects will be accepted after the cutoff date. Five points will be deducted for each calendar day, including weekends a project is submitted after the due date.
- An electronic copy of your project will be submitted to Blackboard on or before the due date. The file name uploaded to Blackboard will be in the format: [last name] [first name] Project2.docx or [last name] [first name] Project2.pdf. For example, *Smith Sally Project2.pdf*.
- Submit one MS Word or one Adobe PDF file. For instance, don't submit separate files for create tables, insert and output.
- No projects will be accepted if sent to my email, left in my office mailbox or delivered to any other member of the department.
- If you submit multiple versions of the project, the last submitted project will be graded. Unless you receive prior approval, a project submitted before the due date and re-submitted after the due date is late.
- Unless you receive prior approval, projects submitted after the due date is late.
- Projects not in compliance with the submission requirements will be rejected and not graded.

Academic Integrity

Projects and examinations must represent your own work. Group projects and exams are not permitted. Although you are encouraged to ask other students for information, you should neither copy another student's project nor permit another student to see your work. You can be asked to perform specific procedures and operations in the presence of the instructor. A student who submits a project that is too similar to another student's work will receive a ZERO for the project. Additional penalties may be imposed. Students found guilty of any form of academic dishonesty such as plagiarism or cheating on an exam or computer project are subject to discipline, including, but not limited to, failure in the course and suspension or dismissal from the College. You are required to comply with the CUNY Policy on Academic Integrity available at
http://www.cuny.edu/about/administration/offices/la/Academic_Integrity_Policy.pdf