**Homework 2. Due by Wednesday, Sept. 28 at 11:59 PM.**

**Please copy all your queries (sql code) from MySQL Workbench and paste them into this Word document for each question. Upload this Word file with your answers to the HW2 assignment folder on Blackboard. Check the file after uploading it.**

The create\_safety.sql script creates a database which contains the 6 tables described below. The data contains safety-related information for 179 company branch locations. To test your queries, create those tables using the create\_safety.sql script.

|  |  |
| --- | --- |
| **Locations** | **Field Description** |
| Location\_ID (pk) | 4 digit location id number |
| Headcount | Number of employees at locations |
| Division | 2 character division code |
| State | Location state |
| City | Location city |
| Employee\_Safety\_Committee | Yes/No indicator of safety committee at location |

|  |  |
| --- | --- |
| **Audits** | **Field Description** |
| Location\_ID (pk and fk) | 4 digit location id number |
| Audit\_Date (pk) | Date of audit |
| Auditor | Name of head auditor |
| Audit\_Findings | Pass/Fail indicator of audit results |
| Corrective\_Action (Unique) | 5 digit CA number, if any (default = NULL) |

|  |  |
| --- | --- |
| **Correctiveactions** | **Field Description** |
| Corrective\_Action (pk and fk) | 5 digit CA number, if any (default = NULL) |
| Audit\_Date | Date of audit |
| Complete\_Date | Date CA was completed |

|  |  |
| --- | --- |
| **Trainings** | **Field Description** |
| Location\_ID (pk and fk) | 4 digit location id number |
| Training\_Date (pk) | Date of training |
| Training\_Location | Location of training (Onsite/Offsite) |

|  |  |  |
| --- | --- | --- |
| **lti (Lost Time Incidents)** | **Field Description** | |
| Location\_ID (pk and fk) | 4 digit location id number | |
| Incident\_ID (pk) | 4 digit incident id number | |
| Department | 2 digit department number | |
| EmployeeNumber | 5 digit employee number | |
| InjuryCode | Description of injury type | |
| DaysLost | # of days lost due to injury | |
|  |  | |
| **nmi (Near Miss Incidents)** | | **Field Description** | |
| Location\_ID (pk and fk) | | 4 digit location id number | |
| Incident\_ID (pk) | | 4 digit incident id number | |
| Dept | | 2 digits department number | |
| IncidentCode | | Description of incident type (root cause, etc.) | |
| EmployeeNumber | | 5 digit employee number | |

Diagram

Description automatically generated

**1. (10 points) What is the total headcount for division "KJ"? (hint: use locations table). The query needs to produce the total headcount**

USE safety;

SELECT Division, SUM(Headcount) AS Total\_Headcount

FROM locations

WHERE Division = "KJ";

**2. (10 points) Show the most recent audit date for location 2415. USE STR\_TO\_DATE function (as I show in Panopto video 3) to work with dates**

SELECT Location\_ID, MAX(STR\_TO\_DATE(Audit\_Date, "%m/%d/%Y")) AS recent\_date

FROM audits

WHERE Location\_ID = 2415;

**3. (10 points) How many audits did the auditor make whose name includes “ain” as part of the name? The query needs to produce the total count of audits**

SELECT COUNT(Auditor) AS auditor\_ain

FROM audits

WHERE Auditor LIKE ("%ain%");

**4. (10 points) What is the total number of audits for each auditor?**

SELECT COUNT(Audit\_Date) AS total\_audits, Auditor

FROM audits

GROUP BY Auditor;

**5. (10 points) How many audits did NOT have any corrective actions?**

SELECT COUNT(Auditor) AS no\_correctiveaction

FROM audits

WHERE Corrective\_Action IS NULL;

**6. (10 points) Use only NMI table. Count NMI incidents by location id and modify query to show only results for locations where the number of incidents is more than 110.**

SELECT Location\_ID, COUNT(Incident\_ID) AS incident\_number

FROM nmi

GROUP BY Location\_ID

HAVING COUNT(Incident\_ID) > 110;

**7. (10 points) Use only LTI table. Count LTI incidents by location id and modify query to show only results for locations where location\_id starts with the number “2” or starts with the number “3”. You may use WHERE or HAVING for that.**

SELECT Location\_ID, COUNT(Incident\_ID) AS incident\_number\_lti

FROM lti

GROUP BY Location\_ID

HAVING Location\_ID REGEXP "^[2,3]";

**8. (10 points) USE Subquery. Use only ‘locations’ table. Show headcount by location id and modify query to show only results for locations where headcount is less than the average headcount for all locations. You may use WHERE or HAVING to show that report. Think about if you need to use GROUP BY or not.**

SELECT Location\_ID, Headcount

FROM locations

WHERE Headcount < (SELECT AVG(Headcount) FROM locations);

**9. (10 points) Create a report showing the number of LTI (Lost Time Incidents) by location as well as LTI Percent (LTI / Location Headcount). Only show results where LTI Percent is more than 20% (i.e., more than 0.2). To be able to confirm the answer by looking at the query results, you need to group by Location\_ID, Division, Headcount and use the same attributes in SELECT part.**

SELECT lo.Location\_ID, Headcount, Division, COUNT(Incident\_ID) AS LTI, COUNT(Incident\_ID)/Headcount AS LTI\_Percent

FROM locations AS lo JOIN lti

ON lo.Location\_ID = lti.Location\_ID

GROUP BY lo.Location\_ID, Headcount, Division

HAVING COUNT(Incident\_ID)/Headcount > 0.2;

**10. (10 points) Create the same report for NMI and NMI% by location. But for that report, only show results for the division “PK” (you may use WHERE or HAVING to filter the results).**

SELECT lo.Location\_ID, Headcount, Division, COUNT(Incident\_ID) AS NMI, COUNT(Incident\_ID)/Headcount AS NMI\_Percent

FROM locations AS lo JOIN nmi

ON lo.Location\_ID = nmi.Location\_ID

GROUP BY lo.Location\_ID, Headcount, Division

HAVING Division = "PK";