# Guidelines when solving the below problems -

* Problem #1 - SQL
  1. Please send us a word document with the SQL and the tested output *or* send us a GitHub URL where the SQL and tested output results have been saved.
* Problem #2 - Concordance
  1. Should be purely your work in Java and tested. Cite if the solution was taken from a website as a reference.
  2. Zip the solution and input/output results and send it to us or send us a GitHub URL where the source code and input/output results have been saved.

# Problem #1 - SQL

Given the following tables…

PERSON

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

PK person\_id NUMBER(10) NOT NULL

first\_name VARCHAR2(100)

preferred\_first\_name VARCHAR2(100)

last\_name VARCHAR2(100) NOT NULL

date\_of\_birth DATE

hire\_date DATE

occupation VARCHAR2(1)

ADDRESS

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

PK address\_id NUMBER(10) NOT NULL

FK person\_id NUMBER(10) NOT NULL

address\_type VARCHAR2(4) NOT NULL

street\_line\_1 VARCHAR2(100)

city VARCHAR2(100)

state VARCHAR2(100)

zip\_code VARCHAR2(30)

1) Write a query to select all rows from *person*. If the person row has a value in *preferred\_first\_name*, select the preferred name instead of the value in *first name*. Alias the column as *REPORTING\_NAME*.

2) Write a query to select all rows from *person* that have a NULL *occupation*.

3) Write a query to select all rows from *person* that have a *date\_of\_birth* before August 7th, 1990.

4) Write a query to select all rows from *person* that have a *hire\_date* in the past 100 days.

5) Write a query to select rows from *person* that also have a row in *address* with *address\_type* = 'HOME'.

6) Write a query to select all rows from *person* and only those rows from address that have a matching billing address (*address\_type* = 'BILL'). If a matching billing address does not exist, display 'NONE' in the *address\_type* column.

7) Write a query to count the number of addresses per address type.

Output:

**address\_type count**

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

HOME 99

BILL 150

8) Write a query to select data in the following format:

**last\_name home\_address billing\_address**

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

Smith 89 Lyon Circle, Clifton, VA 12345 25 Science Park, New Haven, CT 06511

Jones 212 Maple Ave, Manassas, VA 22033 275 Winchester Ave, New Haven, CT 06511

9) Write a query to update the *person.occupation* column to ‘X’ for all rows that have a ‘BILL’ address in the *address* table.

# Problem #2 - Concordance

*Please provide a code sample to answer the following question. The code can be in the Java.*

DHM Concordance

Given an arbitrary text document written in English, write a program that will generate a concordance, i.e. an alphabetical list of all word occurrences, labeled with word frequencies. Bonus: label each word with the sentence numbers in which each occurrence appeared.

* 1. a {2:1,1}
  2. all {1:1}
  3. alphabetical {1:1}
  4. an {2:1,1}
  5. appeared {1:2}
  6. arbitrary {1:1}
  7. bonus {1:2}
  8. concordance {1:1}
  9. document {1:1}
  10. each {2:2,2}
  11. english {1:1}
  12. frequencies {1:1}
  13. generate {1:1}
  14. given {1:1}
  15. i.e. {1:1}
  16. in {2:1,2}
  17. label {1:2}
  18. labeled {1:1}
  19. list {1:1}
  20. numbers {1:2}
  21. occurrence {1:2}
  22. occurrences {1:1}
  23. of {1:1}
  24. program {1:1}
  25. sentence {1:2}
  26. text {1:1}
  27. that {1:1}
  28. the {1:2}
  29. which {1:2}
  30. will {1:1}
  31. with {2:1,2}
  32. word {3:1,1,2}
  33. write {1:1}
  34. written {1:1}

# Problem #3 – DataWarehouse

PetUniversity is an institution for pets and their owner to enroll in training courses, we are building a pet data warehouse to move the operational reporting analysis from transactional system to the data warehouse for all their reporting needs.

* PetUniversity enrolls pets of different types - dogs, cats, birds, horses. Pets and their owners have to together enroll in the courses.
* Each course starts 1st day of the month, is a 4 weeks course, and has a PASS or FAIL grade
* The courses and prices are -
  + For all pets (each course is $100)
    - First Aid and Safety
    - CPR
    - Pet Grooming
  + Dogs (each course is $150)
    - Pet Sitting and Dog Walking
    - Dog Grooming
    - Dog Training 101,
    - Advanced Dog Training
* Depending on the type of the pet, there are different training locations
  + New Haven [dog],
  + Hamden [cat],
  + Woodbridge [dog, cat],
  + Orange [birds, horses]
* There are three holds that are held at the course level -
  + Financial Hold (full tuition is not paid)
  + Registration Hold (registration information is missing)
  + Vaccination Hold (if missing vaccination information)

Your assignment is to build a data warehouse, a star schema data model, add test data, and write SQL queries with results that will answer the below questions -

1. How many new pets are enrollment in each course per month?
   * Pet\_first\_name, Pet\_last\_name, Pet\_type, Owner\_fist\_name, Owner\_last\_name, Enrollment\_date, Month, location.
2. Missing grades report?
   * Pet\_first\_name, Pet\_last\_name, Owner\_fist\_name, Owner\_last\_name, course\_name, course\_instructor
3. Are there holds (financial hold, vaccinations) on the pets that would prevent them from continuing with their training?
   * Pet\_first\_name, Pet\_last\_name, Owner\_fist\_name, Owner\_last\_name, hold\_type, hold\_description, month, hold\_date, course\_price