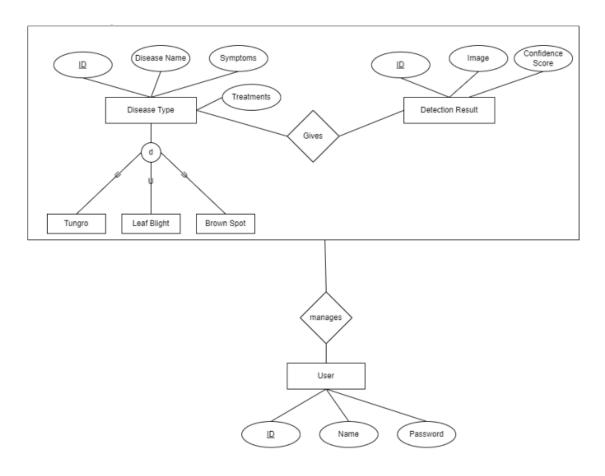
1. Develop an EER model for following.	2
2. Apply PL/SQL for processing databases. Give examples of	
Procedure, Functions and Triggers.	3
3. Illustrate different data types in PostgreSQL.	6
4. Apply CRUD operations and retrieve data in NoSQL	
environment (Use MongoDB or any NoSQL database).	7
5. Illustrate Planning and Optimizing Performance (Use	
PostgreSQL).	9
6. Understand the basic storage architecture of distributed file	
systems. Setup Apache Hadoop in your local machine.	11
7. Distributed Horizontal Fragmentation.	14
8. Distributed Vertical Fragmentation.	15
9. Using prolog, perform the following tasks for the Family	
Relations:	17

## 1. Develop an EER model for following.

A university maintains records of its students and programs in which they have enrolled. It stores student id, name, address, and phone number of student and programs code, program name and duration of a program. A student is either a full time or a part time student (only one of the types). A student can register for many programs and programs can have many students.



# 2. Apply PL/SQL for processing databases. Give examples of Procedure, Functions and Triggers.

## PL/SQL(procedural constructs to the SQL statements)

= The PL/SQL language is capable of performing various different operations at any given time.

## **SQL(Structured Query Language)**

= The SQL language is capable of executing only a single operation at any given time.

#### PL/SQL PROCEDURES & Functions

= Procedures and functions! Procedures act like mini-programs, automating tasks and keeping your code clean. Functions are value-driven workhorses, performing calculations and returning results.

### PL/SQL TRIGGERS

=PL/SQL triggers are special procedures that automatically execute (or "trigger") in response to certain events, such as data changes.

## **Examples:**

#### **Procedures**:

```
postgres=# CREATE TABLE orders (
   order_id INT,
   customer_id INT,
   order_date DATE,
   order_total DECIMAL(10,2)
CREATE TABLE
postgres=# SELECT add_order(123, 456, '2024-06-08', 100.50);
add order
(1 row)
postgres=# SELECT * FROM orders;
order_id | customer_id | order_date | order_total
                   456 | 2024-06-08 |
     123 |
                                           100.50
(1 row)
postgres=#
```

#### **Functions**:

```
sangharsha@sangharsha-Inspiron-5570:/home$ sudo -u postgres psql
[sudo] password for sangharsha:
psql (14.12 (Ubuntu 14.12-Oubuntu0.22.04.1))
Type "help" for help.
postgres=# CREATE OR REPLACE FUNCTION calculate_order_total(quantity INT, unit_price DECIMAL(10,2))
RETURNS DECIMAL(10,2)
LANGUAGE SQL
AS $$
    SELECT quantity * unit_price;
$$;
CREATE FUNCTION
postgres=# SELECT calculate_order_total(5, 10.99);
calculate_order_total
                 54.95
(1 row)
postgres=# SELECT calculate_order_total(45, 10.99);
calculate_order_total
                 494.55
(1 row)
postgres=#
```

### **Trigger:**

```
sangharsha@sangharsha-Inspiron-5570:/home$ sudo -u postgres psql
[sudo] password for sangharsha:
psql (14.12 (Ubuntu 14.12-Oubuntu0.22.04.1))
Type "help" for help.
postgres=# -- Create a table
CREATE TABLE example_table (
   id SERIAL PRIMARY KEY,
     data TEXT,
creation_timestamp TIMESTAMP
);
-- Create the trigger functionCREATE OR REPLACE FUNCTION update_creation_timestamp()
RETURNS TRIGGER
LANGUAGE plpgsql
AS $$
BEGIN
     NEW.creation_timestamp := CURRENT_TIMESTAMP;
     RETURN NEW;
END;
$$;
-- Create the trigger
CREATE TRIGGER update_timestamp_trigger
BEFORE INSERT ON example_table
FOR EACH ROW
EXECUTE FUNCTION update_creation_timestamp();
CREATE TABLE
CREATE FUNCTION
CREATE TRIGGER
postgres=# -- Insert a row
INSERT INTO example_table (data) VALUES ('Example Data')
RETURNING *;
             data
                                   creation_timestamp
 id |
   1 | Example Data | 2024-06-05 12:01:42.113717
(1 row)
INSERT 0 1
postgres=#
```

## 3. Illustrate different data types in PostgreSQL.

```
sangharsha@sangharsha-Inspiron-5570:/home$ sudo -u postgres psql
[sudo] password for sangharsha:
psql (14.12 (Ubuntu 14.12-0ubuntu0.22.04.1))
Type "help" for help.

postgres=# SELECT typname, typlen
postgres-# FROM pg_catalog.pg_type
postgres-# WHERE typtype = 'b'
postgres-# ORDER BY typname;
```

```
sangharsha@sangharsha-Inspiron-5570: /home
                                                                      | typlen
                           typname
__pg_foreign_data_wrappers
__pg_.oretgn_data_wrappers
__pg_foreign_servers
__pg_foreign_table_columns
__pg_foreign_tables
__pg_user_mappings
_aclitem
 _administrable_role_authorizations
_applicable_roles
_attributes
_bit
                                                                                 -1
_bool
_box
_bpchar
                                                                                 -1
-1
_bytea
_cardinal_number
_char
_
_character_data
__character_sets
_check_constraint_routine_usage
_check_constraints
_
_cidr
_cidr
_
_circle
_crcle
_collation_character_set_applicability
_collations
_column_column_usage
_column_domain_usage
_column_options
_column_privileges
_column_udt_usage
_columns
                                                                                 -1
_constraint_column_usage
_constraint_table_usage
_cstring
_data_type_privileges
_datemultirange
_daterange
_domain_constraints
_domain_udt_usage
_domains
_element_types
_enabled_roles
_float4
_float8
__floats
_foreign_data_wrapper_options
_foreign_data_wrappers
_foreign_server_options
_foreign_servers
_foreign_table_options
```

# 4. Apply CRUD operations and retrieve data in NoSQL environment (Use MongoDB or any NoSQL database).

=> Here I use COUCHDB.

Couchdb=Apache CouchDB (link resides outside ibm.com) is an open source NoSQL document database that collects and stores data in JSON-based document formats.



Fig: DashBoard Of CouchDB.

In this lab4. I inject the localhost by using Python Script and I am doing CRUD using python script on COUCHDB.

#### OUTPUT:

```
gharsha@sangharsha-Inspiron-5370:-/sanghu$ python3 crud1.py
ose an operation:
croad a document
croad a document
blotate a document
croad a document
sea an operation:
freed document with ID: 527fe9b7a032805c928758dca8008eb4,Name: Sangharsha Pyakuret
sea an operation:
freate a document
croad a do
                     tt vocament.

the number of the operation: 2
document ID to read: 527fe907a922865c929758dca8008eb4
ent defails:
527fe907a922065c929758dca804674
529fe907a922065c929758dca804670b34
Sampharsha Pyakurel
tyy; 95c.CSIT
                                              : Bs.Csit
6 : ('street': 'Baniyatar', 'city': 'Kathmandu', 'state': 'Bagmati', 'zip': 44700)
an operation:
te a document
a document
te a document
te a document
                                    ete document.

The number of the operation: 3

document ID to update: $27fe8b7a692665c920758dca8868eb4

document ID to update: $27fe8b7a692665c920758dca8868eb4

document (D to update: $27fe8b7a692665c920758dca8868eb4'@'2-060634128161ad36c5c4albd99c6cf5c' {'name': 'Sangharsha Pyakurel', 'faculty': 'BCA', 'hours': 6, 'address': {'str
'Baniyatar', 'city': 'Kathmandu', 'state': 'Bagmati', 'zip': 44700}}>

an operation:
ate a document
                         lete a document the operation: 4 document D to delete: $27fe9b7a822065c920758dca8008eb4 ded document D to delete: $27fe9b7a822065c920758dca8008eb4 ean operation: earle adocument data document data document data document lete a document le
                                    t
the number of the operation:
```

# 5. Illustrate Planning and Optimizing Performance (Use PostgreSQL).

- => Managing the performance of a PostgreSQL database with millions of transactions every week requires careful planning, monitoring, and optimization. Here are some strategies that can help improve the performance of your database:
  - **Indexing:** Proper indexing can significantly speed up the performance of queries that are frequently used in your app. Indexes allow the database to quickly find the relevant data without scanning the entire table. You can create indexes on columns that are frequently used in WHERE clauses or JOIN statements.

```
CREATE INDEX idx_user_id ON transactions (user_id);
```

• Query optimization: Review the queries used in your app to identify any inefficient or slow-performing queries. Use EXPLAIN and ANALYZE to understand the query plan and performance.

```
SELECT user_id, SUM(amount)
FROM transactions
WHERE transaction_date >= '2022-01-01'
GROUP BY user_id;
```

• **Partitioning:** If your app generates millions of transactions every week, consider partitioning the data into smaller and more manageable chunks.

```
CREATE TABLE transactions (
    transaction_id SERIAL PRIMARY KEY,
    transaction_date DATE,
    user_id INTEGER,
    amount DECIMAL(10, 2),
    ...
)
PARTITION BY RANGE (transaction_date);
```

- Scaling: As the volume of data grows, scaling your database can help maintain performance.
- **Monitoring:** Regularly monitor your database to identify any performance bottlenecks or issues.
- Archiving: Consider archiving older data to reduce the size of the database and improve performance.
- **Reporting:** For reports that require data from the last year, consider pre-aggregating the data into a separate table or using a reporting tool like Tableau or Looker.

# 6. Understand the basic storage architecture of distributed file systems. Setup Apache Hadoop in your local machine.

Distributed file systems are designed to store and manage large amounts of data across multiple nodes in a network. They distribute data across different servers to enhance performance, reliability, and scalability. Key components typically include:

- NameNode: Manages the file system metadata and controls access to files by clients
- DataNodes: Store and retrieve blocks of data.
- Client: Requests service from NameNode and DataNodes.

Hadoop's Distributed File System (HDFS) is a prime example, offering high throughput access to application data and is suitable for applications that require high aggregate bandwidth across many nodes.

## **Installing Hadoop on Ubuntu 22.04**

### **Step 1: Update System Repositories**

```
sangharsha@sangharsha-Inspiron-5570:~$ sudo apt update
[sudo] password for sangharsha:
Hit:1 https://download.docker.com/linux/ubuntu jammy InRelease
Hit:2 https://dl.google.com/linux/chrome/deb stable InRelease
Hit:3 http://np.archive.ubuntu.com/ubuntu jammy InRelease
Get:4 http://np.archive.ubuntu.com/ubuntu jammy-updates InRelease [128 kB]
Ign:5 https://apache.bintray.com/couchdb-deb focal InRelease
Get:6 http://security.ubuntu.com/ubuntu jammy-security InRelease [129 kB]
Hit:7 https://apache.jfrog.io/artifactory/couchdb-deb jammy InRelease
Hit:8 http://np.archive.ubuntu.com/ubuntu jammy-backports InRelease
Get:9 http://np.archive.ubuntu.com/ubuntu jammy-backports InRelease
```

## Step 2:Installing Java Development Kit (JDK)

```
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$ sudo apt install default-jdk
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
    ca-certificates-java default-jdk-headless default-jre default-jre-headless fonts-dej
    java-common libatk-wrapper-java libatk-wrapper-java-jni libice-dev libpthread-stubs0
    libsm-dev libx11-dev libxau-dev libxcb1-dev libxdmcp-dev libxt-dev openjdk-11-jdk
    openjdk-11-jdk-headless openjdk-11-jre openjdk-11-jre-headless x11proto-dev xorg-sgm
    xtrans-dev
Suggested packages:
```

**Step 3: Check Java Version** 

```
(venv) <a href="mailto:sangharsha@sangharsha-Inspiron-5570:/opt">sangharsha@sangharsha-Inspiron-5570:/opt</a> java -version openjdk version "11.0.22" 2024-01-16
OpenJDK Runtime Environment (build 11.0.22+7-post-Ubuntu-0ubuntu222.04.1)
OpenJDK 64-Bit Server VM (build 11.0.22+7-post-Ubuntu-0ubuntu222.04.1, mixed mode, sharing)
(venv) <a href="mailto:sangharsha@sangharsha-Inspiron-5570:/opt">sangharsha@sangharsha-Inspiron-5570:/opt</a>
```

#### **Step 3: Downloading Hadoop**

```
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$ sudo wget https://dlcdn.apache.
org/hadoop/common/hadoop-3.4.0/hadoop-3.4.0.tar.gz
--2024-06-05 16:31:39-- https://dlcdn.apache.org/hadoop/common/hadoop-3.4.0/had
oop-3.4.0.tar.gz
Resolving dlcdn.apache.org (dlcdn.apache.org)... 2a04:4e42::644, 151.101.2.132
Connecting to dlcdn.apache.org (dlcdn.apache.org)|2a04:4e42::644|:443... connect
ed.
HTTP request sent, awaiting response... 200 OK
Length: 965537117 (921M) [application/x-gzip]
Saving to: 'hadoop-3.4.0.tar.gz'
hadoop-3.4.0.tar.gz 100%[==============] 920.81M 5.52MB/s in 2m 34s

2024-06-05 16:34:13 (5.97 MB/s) - 'hadoop-3.4.0.tar.gz' saved [965537117/9655371
17]
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$
```

## **Step 4: Extracting the Content**

```
hadoop-3.4.0/sbin/start-dfs.cmd
hadoop-3.4.0/sbin/kms.sh
hadoop-3.4.0/sbin/yarn-daemon.sh
hadoop-3.4.0/sbin/workers.sh
hadoop-3.4.0/sbin/stop-all.cmd
hadoop-3.4.0/sbin/stop-all.sh
hadoop-3.4.0/sbin/stop-dfs.cmd
hadoop-3.4.0/sbin/hadoop-daemon.sh
hadoop-3.4.0/sbin/stop-secure-dns.sh
hadoop-3.4.0/sbin/httpfs.sh
hadoop-3.4.0/sbin/start-dfs.sh
hadoop-3.4.0/sbin/start-all.cmd
hadoop-3.4.0/sbin/hadoop-daemons.sh
hadoop-3.4.0/sbin/refresh-namenodes.sh
hadoop-3.4.0/sbin/start-balancer.sh
hadoop-3.4.0/sbin/start-all.sh
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$
```

Step 5: Moving the Extracted Hadoop Directory to the Installation Location

```
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$ ls
containerd couchdb google hadoop-3.4.0.tar.gz venv
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$
```

#### **Step 6: Finding the Installed Java Path**

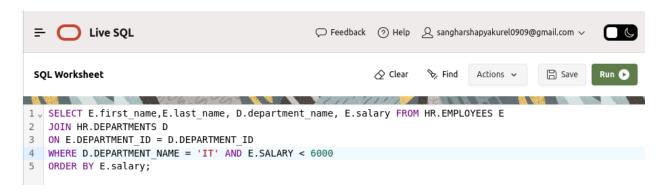
```
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$ readlink -f /usr/bin/java | sed
"s:bin/java::"
/usr/lib/jvm/java-11-openjdk-amd64/
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$
```

### Step 7: Editing Hadoop Environment Configuration File

## **Step 8: Verifying Hadoop Installation**

```
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$ /usr/local/hadoop/hadoop-3.4.0/bin/hadoop version
Hadoop 3.4.0
Source code repository git@github.com:apache/hadoop.git -r bd8b77f398f626bb7791783192ee7a5dfaeec760
Compiled by root on 2024-03-04T06:35Z
Compiled on platform linux-x86_64
Compiled with protoc 3.21.12
From source with checksum f7fe694a3613358b38812ae9c31114e
This command was run using /usr/local/hadoop/hadoop-3.4.0/share/hadoop/common/hadoop-common-3.4.0.jar
(venv) sangharsha@sangharsha-Inspiron-5570:/opt$
```

## 7. Distributed Horizontal Fragmentation.



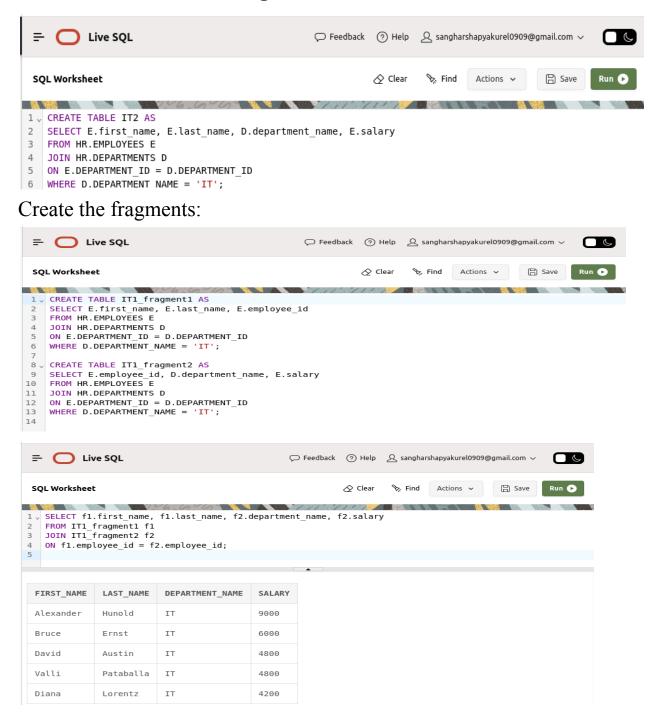
## **Output:**

FIRST_NAME	LAST_NAME	DEPARTMENT_NAME	SALARY
Diana	Lorentz	IT	4200
David	Austin	IT	4800
Valli	Pataballa	IT	4800

Download CSV

3 ------

## 8. Distributed Vertical Fragmentation.



- 9. Using prolog, perform the following tasks for the Family Relations:
- i. Define facts to represent parent-child relationships in the family tree.

```
sang... ×
                                             sang... ×
                 sang... ×
                                                           sang... ×
   sang... ×
  GNU nano 6.2
                                                        family.pl
Define facts for parent-child relationships
parent_child(krishna, manisha).
parent_child(krishna, rohit).
parent child(hari, manisha).
parent_child(hari, rohit).
parent_child(manisha, anita).
parent_child(manisha, dipesh).
parent_child(rohit, anjali).
parent_child(rohit, ravi).
parent_child(gopal, anita).
parent_child(gopal, dipesh).
parent_child(gopal, anjali).
parent_child(gopal, ravi).
```

ii. Define rules to represent the ancestor and descendant relationships.

```
% Define rules for ancestor relationship
ancestor(X, Y) :-
    parent_child(X, Y).
ancestor(X, Y) :-
    parent_child(X, Z),
    ancestor(Z, Y).

% Define rules for descendant relationship
descendant(X, Y) :-
    parent_child(Y, X).
descendant(X, Y) :-
    parent_child(Z, X),
    descendant(Z, Y).
```

- iii. Write queries to find all ancestors and descendants of a given person.
  - -> Ancestor

```
?- ancestor(X, anita).
X = manisha .
```

#### -> Descendant

```
?- descendant(X, krishna).
X = manisha .
```

iv. Define rules to represent sibling relationships

```
% Define rules for sibling relationship
sibling(X, Y) :-
    parent_child(Z, X),
    parent_child(Z, Y),
    X \= Y.
```

- v. Write queries to find all siblings.
- -> Query to find all siblings of "anita":

```
?- sibling(anita, X).
X = dipesh .
```

vi. Define rules to represent the aunt and uncle relationships.

```
% Define rules for aunt and uncle relationships
aunt(X, Y) :-
    parent_child(Z, Y),
    sibling(X, Z).
uncle(X, Y) :-
    parent_child(Z, Y),
    sibling(X, Z).
```

- vii. Write queries to find all aunts and uncles of a given person.
- -> Query to find all aunts of "anjali":

```
?- aunt(X, anjali).
X = manisha .
```

-> Query to find all uncles of "dipesh":

```
?- uncle(X, dipesh).
X = rohit .
```

viii. Define a rule to represent the cousin relationship.

ix. Write a query to find all cousins of a given person.

-> Query to find all cousins of "anita"

```
?- cousin(anita, X).
X = anjali .
```

### Complete Code in 'family.pl'

```
sang...
                sang...
                                            sang...
                                                          sang...
                                                                         sang
  GNU nano 6.2
                                                        family.pl
Define facts for parent-child relationships
parent_child(krishna, manisha).
parent_child(krishna, rohit).
parent_child(hari, manisha).
parent_child(hari, rohit).
parent_child(manisha, anita).
parent_child(manisha, dipesh).
parent_child(rohit, anjali).
parent_child(rohit, ravi).
parent_child(gopal, anita).
parent_child(gopal, dipesh).
parent_child(gopal, anjali).
parent_child(gopal, ravi).
% Define rules for ancestor relationship
ancestor(X, Y) :-
    parent_child(X, Y).
ancestor(X, Y) :-
    parent_child(X, Z),
    ancestor(Z, Y).
% Define rules for descendant relationship
descendant(X, Y) :-
   parent_child(Y, X).
descendant(X, Y) :-
    parent_child(Z, X),
    descendant(Z, Y).
% Define rules for sibling relationship
sibling(X, Y) :-
    parent_child(Z, X),
parent_child(Z, Y),
    X \= Y.
% Define rules for aunt and uncle relationships
aunt(X, Y) :-
    parent_child(Z, Y),
     sibling(X, Z).
uncle(X, Y) :-
    parent_child(Z, Y),
    sibling(X, Z).
% Define a rule for cousin relationship
cousin(X, Y) :-
    parent_child(Z, X),
    parent_child(W, Y),
    sibling(Z, W).
```