

SRPCE

2022-2025



NAME - A.Kannimuthu

DEPARTMENT-CSE

ROLL-NO - 39

SUBJECT - cloud Application Development

REG-NO - 422021104304

SEMESTER- 5



IBM Cloud Database Analytics



Provide a introduction and module and functions and program for big data analytics IBM cloud database phase 4



Certainly! Big Data Analytics on IBM Cloud Database involves several phases and typically uses various tools and services provided by IBM. In Phase 4, you might be focused on data analysis and reporting. Here's an introduction to Phase 4, along with a sample program outline for data analytics on the IBM Cloud Database:

Introduction to Phase 4: Data Analysis and Reporting

In Phase 4 of Big Data Analytics on IBM Cloud Database, the primary focus is on extracting valuable insights from the vast amount of data stored in IBM Cloud databases. This phase involves the use of data analysis tools and techniques to uncover trends, patterns, and correlations within the data. Additionally, it includes the creation of reports and visualizations to communicate these findings effectively.

Module and Functions:

1. **Data Extraction:** Create functions to extract data from the IBM Cloud Database using SQL queries or database connectors.
2. **Data Transformation:** Develop functions to clean, preprocess, and transform the data for analysis.
3. **Data Analysis:** Implement various analytical functions and algorithms to derive insights from the data. You can use tools like Apache Spark or IBM Watson Studio for this.
4. **Visualization:** Use libraries like Matplotlib, Seaborn, or IBM Cognos Analytics to create visualizations that convey the results of your analysis.
5. **Reporting:** Develop functions to generate reports and dashboards summarizing the key findings.

Sample Program Outline:

Below is a simplified program outline for Phase 4 of Big Data Analytics on IBM Cloud Database:

```
# Import necessary libraries and modules
import ibm_db          # for database connection
import pandas as pd    # for data manipulation
import matplotlib.pyplot as plt  # for data visualization

# Function to connect to the IBM Cloud Database
def connect_to_db():
    conn = ibm_db.connect("your_connection_string", "", "")
    return conn

# Function to extract data from the database
def extract_data(conn, sql_query):
    result = ibm_db.exec_immediate(conn, sql_query)
    data = ibm_db.fetch_assoc(result)
    return data

# Function to transform and clean data
def transform_data(data):
    # Implement data transformation logic here
    cleaned_data = ...

# Function for data analysis
def analyze_data(data):
    # Implement data analysis algorithms
    insights = ...

# Function for data visualization
def visualize_data(insights):
    # Create visualizations based on the analysis results
    # For example, create bar charts, scatter plots, etc.

# Function to generate reports
def generate_report(data, insights):
    # Create a report or dashboard summarizing the analysis findings
```



```
# You can use Jupyter Notebooks or reporting tools like IBM C

# Main program
if __name__ == "__main__":
    # Connect to the database
    conn = connect_to_db()

    # Extract data using SQL query
    sql_query = "SELECT * FROM your_table"
    data = extract_data(conn, sql_query)

    # Transform and clean data
    cleaned_data = transform_data(data)

    # Analyze data
    analysis_results = analyze_data(cleaned_data)

    # Visualize data
    visualize_data(analysis_results)

    # Generate reports
    generate_report(cleaned_data, analysis_results)

    # Close the database connection
    ibm_db.close(conn)
```

Please note that this is a simplified outline, and in a real-world scenario, you may need to adapt and extend the functions and modules to suit your specific requirements and use IBM Cloud services such as IBM Db2, IBM Watson Studio, and other relevant tools for a comprehensive Big Data Analytics project.