**14th January 2023**

**Big Bang** DATA SCIENCE

# SQL for *Data science*!

**16-Weeks | 3 times a week | 2.5 hours per session | $699**

BBDS's SQL for Data Science and Machine Learning program is designed for graduates ; and working professionals looking to make their mark in this high-demand industry.

We are looking to enroll candidates with the right blend of academic spirit and soft skills and guide them through our cohort-based learning path for 15 weeks with virtual internships, and ultimately get them hired.

## What you will learn?

## Scan to Register?

# Ready to advance your career ?

**Module 1 : SQL for Data Science**
RDS | Database Engine | SQL components | Selecting | Joining | Queries & sub-queries | Rename|Filtering | Grouping | Sorting | Functions | Data Analysis with SQL

**Module 2 : Data Visualization with Power BI**
Data Viz | Data Profiling | Reports & Dashboards | Data Modeling

**Module 3 : Information Security Management**
Security Documentations | Security Frameworks | Decision-making process |Risk Assessment | Risk Mitigation | Evaluation and Assessment | Risk Equation

**Module 4: Python - Data Science & ML**
Pandas | NumPy |Matplotlib | Seaborn| Data Preprocessing | EDA | Classification | Regression | Clustering

**Module 5: Capstone Project**

You can retake the program unlimited times with no additional cost

## Meet your mentors
Our network of mentors will guide you with through your journey to **learn, upskill and land your dream role!**

**Mo Medwani**
PhD Candidate
4 - Master Degrees
23 yrs of Experience

**Ed. Bujak**
25 yrs. of Teaching Exp
2 Master Degrees

**Mohammed Mjeeb**
MBA - Microsoft Certified
24 yrs. Of Experience

We are **Big Bang Data Science Solutions**

Internet 2.0 Outstanding Leadership Award - Dubai 2022

# Program Agenda

| Week 1 | Program Orientation |
|---|---|

**Learning Objectives:** The program begins with a lecture on "Why Learning Data Science is an Absolute Must!" where you will be introduced to the Data Science process, Data Science portfolio, analytics types, day-to-day activities, quantitative, and statistical techniques required, and the career path to start the journey.

| Session 1 | Program Orientation |
|---|---|
| | • Program Orientation (Agenda – Curriculum)<br>• BBDS Website \| Canvas \| Slack walk-through |
| Session 2 | Why should you become a Data Scientist? |
| | • Data Explosion<br>• Why Data Science? \| What is Data Science? \| Type of Analytics<br>• Data Science Portfolio \| Data Science Process \| Career in Data Science |
| Session 3 | Projects, Teams & Team Leads |
| | • Introduction to Data Science Project<br>• Projects Discussion (Milestone Projects Assignment)<br>• Projects List \| Project Templates \| Project Team |

| Week 2 | SQL for Data Science: Relational Database Systems |
|---|---|

**Learning Objectives:** In this module, we will learn about Introduction to Databases, Brief History of Database Application, Advantages of Using the DBMS, Characteristics of the Database Approach, what are Relational Databases and What is SQL; installation of SQL server application and client tools.as well as exploring reporting, analysis, and integration services.

| Session 1 | Relational Database Systems: An Introduction |
|---|---|
| | • Introduction to Databases \| Relational Databases and Non-Relational Databases<br>• Advantages of Databases<br>• Characteristics of Database |
| Session 2 | Planning the Installation and Installing SQL Server |
| | • Install the SQL Server Engine on the desktop/laptop<br>• Install SQL Server Management Studio (SSMS)<br>• Exploring Reporting, Analysis, and Integration Services |
| Session 3 | Connecting to Server and Create Database |
| | • Connecting to SQL Server<br>• Creating Altering and Dropping a Database<br>• Intro to SQL |

| Week 3 | SQL for Data Science: Tables and Constraints |
|---|---|

**Learning Objectives:** In this module, we will learn about creating and working with tables, Databases Data Definition Language (DDL) and Data Manipulation Language (DML), Data Types SELECT, INSERT, UPDATE, AND DELETE Statements in SQL.

| Session 1 | Tables and Constraints |
|---|---|
| | • Creating, and Working with Tables<br>• Default Constraints<br>• Cascading referential integrity |

| Session 2 | Tables and Constraints |
|---|---|
| | <ul><li>Check Constraint | Identity column</li><li>Get last generated id in SQL server</li><li>Unique key constraint</li></ul> |
| Session 3 | SQL Statements Group by Clause |
| | <ul><li>SQL Statements (Select, Insert, Delete, Update)</li><li>Group by Clause</li><li>Difference between where and having in SQL server</li></ul> |
| Week 4 | SQL for Data Science: Functions |

**Learning Objectives:** In this module, we will learn about Joins, System Functions, User Define functions, Stored Procedures, Advantages of the stored procedure, Date, Mathematical functions, and Temp tables.

| Session 1 | Joins and Functions |
|---|---|
| | <ul><li>Basic Joins | Advance Joins | Self Joins</li><li>Different ways to replace NULL | Coalesce function</li><li>Union and Union All</li></ul> |
| Session 2 | Stored Procedures, Built in Functions |
| | <ul><li>Advantages of stored procedures | Stored Procedures with output parameters |</li><li>Build in string functions | Left, Right, Charindex, and Substring functions</li><li>Replicate, Space, Patindex, Replace and Stuff functions</li></ul> |
| Session 3 | Date and Mathematical Functions, Temporary Tables |
| | <ul><li>DateTime functions | IsDate, Day, Month, Year and DateName functions</li><li>DatePart, DateAdd and DateDiff functions | Convert and Cast functions</li><li>Mathematical functions | Scalar User Defined functions | Temporary tables</li></ul> |
| Week 5 | SQL for Data Science: Indexes | Vies |Triggers |

**Learning Objectives:** In this module we will learn about different types of Indexes (Cluster, NonCluster indexes), Views, Triggers, Derived Tables, Sub and Correlated queries and fundamentals of accessing SQL databases with Python.

| Session 1 | Indexes |
|---|---|
| | <ul><li>Clustered and NonClustered Indexes</li><li>Unique and Non-Unique Indexes</li><li>Advantages and disadvantages of Indexes</li></ul> |
| Session 2 | View and Triggers |
| | <ul><li>Views | Updateable Views</li><li>Indexed Views | View Limitations</li><li>DML Triggers (Instead of Insert trigger, Instead of Update trigger, instead of delete trigger)</li></ul> |
| Session 3 | Derived Table, Subqueries, Accessing data in Python |
| | <ul><li>Using SQL within Python to access database</li><li>Derived table | Subqueries in SQL</li><li>Correlated subqueries</li></ul> |
| Week 6 | SQL for Data Science: Data Warehousing 1 |

**Learning Objectives:** In this module, we will learn what data Warehousing and the difference between Data Warehouse, database, Data Lake, and data warehouse architecture is. we will learn the difference between OLTP and OLAP. Also, how to import and export data from any data source to and from a database. What is Extract, Transform and Load (ETL) process.

| Session 1 | Introduction to Data Warehousing |
|---|---|

|  | |
|---|---|
| | • What is a Data Warehouse \| Benefits of data Warehouse |
| | • Difference between data warehouse, database, and data lake |
| | • Data warehouse architecture and design |
| **Session 2** | **Data Warehousing** |
| | • Why do we need a Data Warehouse? |
| | • What does a Data Warehouse look like? |
| | • The Ideal Data Warehouse |
| **Session 3** | **Data Warehousing** |
| | • The Difference Between OLTP and OLAP |
| | • Import and Export data from data sources using SQL Server Integration service and python scripts. |
| | • What is Extract, Transform and Load (ETL) process |
| **Week 7** | **SQL for Data Science: Data Warehousing 2** |

**Learning Objectives:** In this module, we will introduce SQL Server analysis services, Create/backup/restore Data warehouse; What are Star and snowflake Schemas, and differences between them; the Fact tables, Dimension tables and difference between them.

| **Session 1** | **Creating, back up, and Restore Data warehouse** |
|---|---|
| | • Introduction to SQL Server analysis services |
| | • Create data warehouse |
| | • Backup and Restore data warehouse |
| **Session 2** | **Star and Snowflake Schema** |
| | • What is Star Schema |
| | • What is Snowflake Schema |
| | • What is the difference between Star and snowflake Schema |
| **Session 3** | **Fact and Dimension tables** |
| | • What are Facts tables |
| | • What are Dimension tables |
| | • Create Facts and Dimension tables |
| **Week 8** | **Data Visualization - Introduction Power BI** |

**Learning Objectives:** In this module, we will learn Intro to Power BI, what is Power BI, Power BI Features. loading data from various sources into Power BI.

| **Session 1** | **Introduction to Power BI** |
|---|---|
| | • Intro to Power BI |
| | • What is Power BI |
| | • Installing Power BI |
| **Session 2** | **Power BI Features, Components, Services** |
| | • Power BI Feature |
| | • Power BI Components |
| | • Power BI Services |
| **Session 3** | **Loading Data from Various source into power BI** |
| | • Loading data from the database into Power BI. |
| | • Loading data from Excel into Power BI |
| | • Loading data from an API into Power BI |
| **Week 9** | **Data Visualization - Power BI Transform Data** |

**Learning Objectives:**  In this module, we will learn Power BI ETL process, how to Transform data, how to create data models in Powe BI; and Dax Functions, Data Modeling and create a financial report.

| Session 1 | **Power BI ETL Process / Transform Data** |
|---|---|
| | • ETL Process<br>• Transform Data<br>• Using Power Query Editor where we massage and scope the data |
| **Session 2** | **Dax Functions** |
| | • Naming Column, Adding Columns headers, Editing data types<br>• Working with Applied Steps area. (Records every action taken) Adding, Merging, reordering Columns<br>• Cleaning the data (replacing, upper/lower, rtrim, ltrim, Prefix, Suffix), |
| **Session 3** | **Dax Functions and Data Modeling** |
| | • Data parsing<br>• Conditional columns<br>• Data modeling (joining data tables) |
| **Week 10** | **Data Visualization - Power BI ETL Process** |
| **Learning Objectives:** In this module, we will create charts and graphs, profession Power BI Reports | |
| **Session 1** | **Build Data warehouse and ETL Process 1** |
| | • Building Financial Report using Excel as Data source<br>• Adding Cards, Data tables, Cluster Bar, Pie, donut charts<br>• Adding Maps and tree maps. |
| **Session 2** | **Build Data warehouse and ETL Process 2** |
| | • Build Financial Report using SQL server database as a data source<br>• Adding Sliders (date and dropdown)<br>• Adding Dax code to display last refreshed report |
| **Session 3** | **Building Reports with Data source as Data warehouse** |
| | • Create Assessment Report<br>• Work with Aggregates; Measures (Maxx, Avrageex, Countax)<br>• Create Small Multiples |
| **Week 11** | **Information Security Management** |
| **Learning Objectives:** In this model, we will learn Information security management. We will learn how to define and manage controls that an organization needs to implement to ensure that it is sensibly protecting the confidentiality, availability, and integrity of assets from threats and vulnerabilities | |
| **Session 1** | **Core Tenants, Security Documentations, Frameworks** |
| | • Core Tenants<br>• Security Documentations<br>• Security Frameworks |
| **Session 2** | **Risk Assessment, Risk Mitigation** |
| | • Decision-making process<br>• Risk Assessment | Risk Mitigation | Evaluation and Assessment<br>• Risk Equation | Elements of Risk |
| **Session 3** | **Types of Risk, Risk Handling and Security Controls** |
| | • Types of Risk<br>• Risk Handling<br>• Security Controls |
| **Week 12** | **Python for Data Science (NumPy)** |
| **Learning Objectives:** In this module we will explore the fundamentals of the Python programming language. Additionally, we will introduce the NumPy Python package which is a core dependency for many other Python packages by examining some of NumPy's capabilities and strengths: broadcasting, indexing, slicing, math, statistics, matrices, axis, stacking, joining, searching, sort, count, Boolean arrays, and much more. | |

| Session 1 | Introduction to Python programming (Fundamentals) |
|---|---|
| | <ul><li>Data Types - is implicitly and dynamically typed</li><li>Operators</li><li>ds Loops</li><li>Flow Control</li><li>Functions</li><li>Importing packages and modules</li></ul> |
| Session 2 | NumPy Array \| The Shape and Reshaping of NumPy Array |
| | <ul><li>NumPy ndarray data type, features, and array Dtype</li><li>matrices, dimensions, axis, shape, reshape</li><li>Sequences (index a slice), Iterable</li></ul> |
| Session 3 | Expanding, Squeezing, Indexing and Slicing of NumPy Array |
| | <ul><li>Filtering</li><li>Statistics</li><li>Correlation</li><li>Regression</li></ul> |
| Week 13 | Python for Data Science (Pandas) |

**Learning Objectives:** In this module we will introduce the pandas Python package which is the go-to library for EDA (Exploratory Data Analysis) with pandas DataFrame, Series, and index data types We will examine some of the capabilities of pandas including manipulating/wrangling, statistics, data visualization.

| Session 1 | Pandas Fundamentals |
|---|---|
| | <ul><li>Pandas' data types: Index, Series, DataFrame</li><li>Operations and methods on Pandas' data types</li><li>Axis</li></ul> |
| Session 2 | Using Pandas for Data Analysis & Data Wrangling |
| | <ul><li>EDA, data cleaning, filtering, row/record, and column/feature manipulations</li><li>Statistical foundations</li><li>Problematic data - outliers, missing data</li></ul> |
| Session 3 | Data Wrangling & Visualizing with Pandas |
| | <ul><li>Data transformations</li><li>Slicing, changing index, data conversion, joining and merging, concatenation, columns, pivoting, melting, handling duplicates</li></ul> |
| Week 14 | Exploratory Data Analysis & Statistical Analysis |

**Learning Objectives:** In this module, students will be introduced to the second phase of the Cross- Industry Standard Process for Data Mining (CRISP-DM) process model. Students obtain data and verify that it is appropriate for the needs. Students learn to identify issues that cause the analyst to return to the 'Business Understanding' phase of the project and revise the plan. Students may even discover flaws in the 'Business Understanding', another reason to rethink goals and plans.

The Data Understanding phase includes four tasks. These are Gathering data – Describing data – Exploring data – Verifying data quality.

| Session 1 | Data Loading & Manipulation in Python |
|---|---|
| | <ul><li>Read data from different sources</li><li>Get to Know the Data<ul><li>Identifying Categorical Data: Nominal, Ordinal and Continuous</li></ul></li></ul> |

|  |  |
|---|---|
|  | • Univariate \| Bivariate \| Multivariate Analysis |
|  | • Types of Data \| Type of Central Tendency \| Data distribution \| Data Density |
| **Session 2** | **Data Types- Measure of Shape - Position - Dispersion** |
|  | • Measures of the Spread: Range – IQR – Variance – Standard Deviation<br>• Measures of Dispersion<br>• Measures of Position: Statistical Analysis (John T. 5 Numbers Summary)<br>• Measures of Relationships: Correlation \| Variance & Covariance<br>• Measures of Shapes: Skewness & Kurtosis<br>• Python hands-on training with real-world business problem |
| **Session 3** | **Data Visualization (Numerical Data & Graphical Descriptive Statistics)** |
|  | • Data Visualization (Histogram, Box Plot, Bar Plot, Scatter Plot)<br>• Data Quality Report \| Develop the Code Book<br>    o Summary of Data \| Type of variable \| Ranges of variables \| Missing fields \| Identify the primary list of variables to solve the business problem<br>• Detecting missing values & outliers \| Duplicates & redundant records<br>• Python hands-on training with real-world business problems |
| **Week 15** | **Data Preparation & Preprocessing** |

**Learning Objectives:** In the Data Preparation phase, students learn the process of cleaning and transforming raw data prior to processing and analysis. This is an important step prior to processing and often involves reformatting data, making corrections to data and the combining of data sets to enrich data.

Students also learn how to fix data quality issues discovered through EDA & Data visualization from the Data Understanding phase.

| | |
|---|---|
| **Session 1** | **Data Preparation & Processing (Categorical Attributes)** |
|  | • Encoding Categorical Data<br>    o Replacing values \| Creating Dummy Variables \| Encoding labels \| One-Hot encoding \|Binary encoding \| Backward difference encoding \| Miscellaneous features<br>• Data Quality Report<br>• Develop the Code Book<br>    o Summary of Data \| Type of variable \| Ranges of variables \| Missing fields \| Identify the primary list of variables to solve the business problem |
| **Session 2** | **Data Preparation & Processing (Numeric Attributes)** |
|  | • Statistical Analysis (John T. 5 Numbers Summary)<br>• Variance & Covariance \| Correlation Analysis<br>• Python hands-on training with real-world business problems |
| **Session 3** | **Data Preparation & Processing (Missing Values \| Outliers \| Duplicates)** |
|  | • Dealing with Skewness & Kurtosis \| Missing Values & Outliers \| Duplicates & redundant<br>• Variable Conversion<br>    o Discretization<br>    o Binning<br>• Variable Transformation<br>    o Normalization \| Standardization<br>    o Factorizations \| Binarization<br>• Box-Cox Transformation<br>• Data Partitioning |
| **Week 16** | **Introduction to Machine Learning** |

**Learning Objectives:** This module broadens concepts learned in Data Understanding, Data Preparation & Exploratory Data Analysis (EDA) by extending to Machine Learning where you will learn about the models &

methods used in machine learning & apply them to the real-world.

The aim of supervised machine learning is to build a model that makes predictions based on evidence in the presence of uncertainty. In this session, we will learn about Machine Learning, Introduction to Classification, Introduction to Regression and introduction to Clustering.

| Session 1 | Introduction to Classification Analysis |
|---|---|
| | <ul><li>Algorithm for Decision Tree Induction</li><li>Naïve Bayes Assumption</li><li>Logistic Regression as a Generalized Linear Model</li><li>KNN Algorithm</li><li>Linear SVM</li><li>Soft-margin SVM</li><li>Nonlinear SVM</li></ul> |
| Session 2 | Introduction to Regression Analysis |
| | <ul><li>Making Predictions with Linear Regression</li><li>Polynomial Algorithm</li><li>Multiple Linear Regression</li><li>The F-Statistic</li><li>Interpreting results of Categorical variables</li><li>Heteroscedasticity</li><li>Backward Elimination \| Backward Elimination \| Automatic Backward Elimination</li><li>OLS Regression- Theory \| Implementation</li><li>Confidence Interval and OLS Regressions</li><li>Ridge Regression \| LASSO Regression</li></ul> |
| Session 3 | Introduction to Clustering Analysis |
| | <ul><li>Bisecting K-means</li><li>K-means & Different Types of Clusters</li><li>K-means as an Optimization Problem</li><li>Basic Agglomerative Hierarchical Clustering Algorithm</li><li>BIRCH \| DBSCAN</li><li>Mini-Batch K-Means</li><li>Mean Shift \| OPTICS</li><li>Spectral Clustering</li><li>Gaussian Mixture Model</li></ul> |
| Week 17 | Individual Capstone Project (1 Month) |

**Learning Objectives:** During the final week, you will be transitioned into full-time focus on your final, passion individual capstone assignment that will require you to apply the tools you have learned in the program.

**Duration:** The capstone will run for ONE month starting from the 29th of April 2023 till the 29th of May 2023.

**Grading:** The capstone project is a program requirement. To successfully complete the capstone project and receive the certificate, students must achieve "Meets Expectations" OR "Exceeds Expectations".

## Data Science Interview Questions & Answers

**Learning Objectives:** In this module, you will be introduced to 120+ Data Science Interview questions and answers covering the topics like:

- The Big Picture | Optimization | Data Pre-processing | Sampling & Splitting
- Supervised Learning | Unsupervised Learning | Model Evaluation | Ensemble Learning
- Business Applications

## Data Science Interview Real Challenges)

**Learning Objectives:** In this module, you will be introduced to some 24 hours real take-home challenges used as hiring process with some big companies:

- Capgemini Challenge | SDSC Challenge
- Foot Locker Challenge | Fatality Crashes Challenge
- PayPal Challenge

## Resume Preparation

**Learning Objectives:** In this module, you will learn how to build an affective Data Science resume:

- Structure of your Data Science Resume
- Adding Content and Information to your Data Science Resume
- Get Feedback from Industry Experts
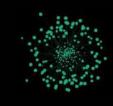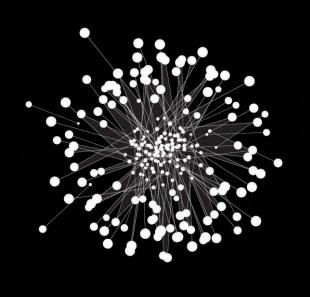- Build your Digital Presence
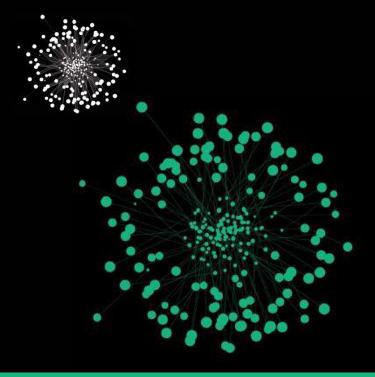
# Career Services

## Resume Preparation | Interview Preparation | Employment Assistance Certificate Analytics Professional Training

# Certificate of Completion

Upon successful completion of the program, Big Bang Data Science Institute grants a verified/certified digital certificate of completion to participants. This program is graded as pass or fail; participants must receive 80% to pass and obtain the certificate of completion.



After successful completion of the program, your verified digital certificate will be emailed to you in the name you used when registering for the program. All certificate images are for illustrative purposes only and may be subject to change at the discretion of Big Bang Data Science Solutions

# About BBDS...

We are Big Bang Data Science Solutions (BBDS), a holistic up-skilling platform driven by a unique, cohesive "Learn-Apply-Solve" framework.

This innovative solution provides application-oriented immersive and interactive learning experience with extensive real-industry courses, cases, datasets, and projects. It also ensures a blended pathway between industry and academia through simulation and contextualization.

BBDS regularly presents at numerous conference workshops and until recently held regular monthly Meetups with industry experts as speakers.

We currently offer a few multi-week, multi session courses that are live (then recorded) programs that participants have thoroughly enjoyed since we support our participants with almost endless one-on-one or group live support sessions.

## *CONNECT WITH A PROGRAM ADVISOR*

Have questions about the program or how it fits inwith your career goals?

### +1 (315) 975-1661

✉ **register@bbds.ma**    🌐 **www.bbds.ma**

📅 **DURATION**
16 Weeks

🖥 **FORMAT**
Online

💲 **PROGRAM FEE**
US $699

**ENROLL**

**SCHEDULE A CALL**