PYTHON TRAINING MODULE

(CLASSROOM)

PYTHON TRAINING MODULE

The Python Training module will make the reader accustomed to python language. This material will help the reader in understanding the basics of the python language, Python libraries and the use of python for the analytics.

DELIVERY METHOD

25 % Self-paced Learning

75 % Instructor led training

VERSION

2019

LEARNING OBJECTIVES

- Explain what Python is
- Advantages and disadvantages of Python
- Getting started with Python and its different versions
- Explain variables, strings and functions
- Use of mathematical operators and functions
- Explain different statements like if, for etc.
- Explain the python libraries
- Explain Details of the Pandas library
  - o Series and Data Frames
  - o Grouping and aggregating
  - o Merging and joining
- Define error handling in Python
- Define RE objects
- Define pattern matching and Parsing of data
- Define regression with Use case study
- Define exploratory data analysis
- Define correlation matrix
- Define visualization using matplotlib
- Define churn analysis with Use case
- Define advance Machine learning Algorithms
- Define Support vector machine
- Define Random forest

PREREQUISITES SKILLS

- Computer Science fundamentals
- Basic knowledge of applied math, algorithms, and data modelling
- Basic knowledge of statistics

DURATION

40 Hours

SKILL LEVEL

Basic – Intermediate

HARDWARE REQUIREMENTS

Processor	2 GHz or Higher
GB RAM	8 GB
GB Disk Free	80 GB
Network Requirements	Yes

Notes

The following unit and exercise durations are estimates and might not reflect every class experience. The estimates do not include the duration of optional exercises or sections. Students in this course use an IBM Cloud Lite account to perform the exercises. This account will never expire; therefore, students can continue working on the optional exercises after the class

COURSE AGENDA

UNIT I. Introduction to PYTHON

Duration: 6 Hrs.

Overview	This unit explains what is Python, its advantages and disadvantages, how to run python scripts, how to use variables, string operator and functions.
Learning Objectives	After completing this unit, you should be able to: - Explain what Python is - How to install and get start with python - How to use basic variables and stings in python - Work with Mathematical operators in python

UNIT II. Deep dive into PYTHON

Duration: 8 Hrs.

Overview	This unit consist more in depth working of Python like inputting the data, working with Boolean and other statements.
Learning Objectives	After completing this unit, you should be able to: - How to input data in Python - Use Boolean with python - Use If and elif statement in python - Use while loop in python - Work with lists - Use For statement

UNIT III. Python Libraries

Duration: 8 Hrs.

Overview	This unit explains the use of pandas library for data analysis
Learning Objectives	After completing this unit, you should be able to: • Install Pandas • Work with series and data frames • Work on grouping, aggregating and applying different functions on data • Merge and Join the data

UNIT IV. Error Handling

Duration: 4 hrs.

Overview	This unit explains how to deal with different type of errors that one can encounter while working with Python.
Learning Objectives	After completing this unit, you should be able to: • Deal with Syntax errors • Deal with the exceptions

UNIT V. Other Topics

Duration: 4 hrs.

Overview	This unit explains how to deal with miscellaneous things in python
Learning Objectives	After completing this unit, you should be able to: • Work with regular expression • Work with Pattern matching • Parse data

UNIT VI. Regression (Use case study)

Duration: 3 hrs.

Overview	This unit explains regression analysis with the help of a use case.
Learning Objectives	After completing this unit, you should be able to: • Define regression analysis • Work with regression analysis

UNIT VII. Other Regression related topics

Duration: 4 hrs.

Overview	This unit explains different topics which are important from the point of view of data analytics.
Learning Objectives	After completing this unit, you should be able to: • Define exploratory analysis • Define correlation matrix • Perform visualization using matplotlib • Implement linear regression

UNIT VIII. Advance

Duration: 3 hrs.

Overview	This unit explains some advance data analytics techniques.
Learning Objectives	After completing this unit, you should be able to: • Apply advanced Machine learning algorithms • Work on Support vector machines • Define Random forest

RDBMS TRAINING MODULE

(CLASSROOM)

RDBMS TRAINING MODULE

The RDBMS Training module will get the reader accustomed with RDBMS concepts. This material will help the reader in understanding the basics of RDBMS, what are Entities and Relationships, Overview on Normalization, Database Design and Performance Tuning, JDBC and Advanced concept in RDBMS like Database Security and Database backup and Restore.

DELIVERY METHOD

100% Instructor led training

VERSION

2020

LEARNING OBJECTIVES

- Understanding Database Concepts
    - Introduction
    - Tables
    - Primary Keys
    - Foreign Keys
    - Installation of SQLite
    - Installation of Docker based MySQL and DB2 database

- Understanding Database Storage
    - Introduction
    - Database normalization
    - Indexes and how they are used in databases
    - Configure non-clustered indexes
    - Configure clustered indexes

- Entities and Relationships
    - Introduction
    - Entities and Their Attributes
    - Domains
    - Basic Data Relationships
    - Documenting Relationships
    - Dealing with Many-to-Many Relationships
    - Relationships and Business Rules
    - Data Modeling Versus Data Flow
    - Schemas

- The Relational Data Model
    - Introduction
    - Understanding Relations
    - Primary Keys
    - Representing Data Relationships
    - Views
    - The Data Dictionary

- Normalization
    - Introduction
    - Translating an ER Diagram into Relations
    - Normal Forms
    - First Normal Form
    - Second Normal Form
    - Third Normal Form
    - Boyce–Codd Normal Form
    - Fourth Normal Form
    - Fifth Normal Form
    - Sixth Normal Form

- Database Design and Performance Tuning
    - Introduction
    - Indexing
    - Clustering
    - Partitioning

- Creating Database Objects
    - Understand data definition language (DDL)
    - Choose appropriate data types

- Manipulating Data
    - Introduction
    - Understand data manipulation language (DML)

- JDBC As the Fundamental Java API
    - Introduction
    - JDBC basics

- JPA as the JAVA ORM API
    - Introduction
    - From JDBC to JPA

- Database Security
    - Introduction
    - Sources of External Security Threats
    - Sources of Internal Threats
    - External Remedies
    - Internal Solutions

- Understanding Database Backup and Restore
    - Introduction
    - Understand different types of backups
    - Define a backup and recovery strategy

- Introduction of MySQL
    - Create Tables
    - Drop Tables
    - Insert Query
    - Select Query
    - Where Query
    - Update Query
    - Delete Query

- o Like Clause
- o MySQL Joins

PREREQUISITES SKILLS

- Computer Science fundamentals
- Basic hands-on experience in Unix Operating system
- Basic understanding of Database concepts
- Basic knowledge on Containerization & Virtualization concepts

DURATION

36 Hours

SKILL LEVEL

Basic – Intermediate

HARDWARE REQUIREMENTS

Processor	2 GHz or Higher
GB RAM	8 GB
GB Disk Free	20 GB
Network Requirements	Yes

NOTES

The following unit and exercise durations are estimates and might not reflect every class experience. The estimates do not include the duration of optional exercises or sections.

COURSE AGENDA

CHAPTER I. Understanding Database Concepts

Duration: 1 Hr.

Overview	This chapter provides an overview on database concepts, provides details on database tables, primary keys and foreign keys. Also, it gives detailed steps on installation of SQLite and Dockerized MySQL and DB2 databases.
Learning Objectives	After completing this unit, you should be able to: • Explain the database concepts like tables and different keys. • Install SQLite, MySQL and DB2.

CHAPTER II. Understanding Database Storage

Duration: 2 Hrs.

Overview	In this chapter, you will get an overview on Database normalization, Indexes and how they are used along with configuring clustered as well as non-clustered indexes in databases.
Learning Objectives	After completing this unit, you should be able to: • Explain database normalization, clustered as well as non-clustered indexes. • Create Indexes in database.

CHAPTER III. Entities and Relationships

Duration: 2 Hrs.

Overview	This chapter provide an overview on entities and relationships in RDBMS and explains concepts like Domains, Relationship and Business rules, Data Modelling and Schemas etc.
Learning Objectives	After completing this unit, you should be able to: • Understand concepts like entities, attributes, data modelling and relationship in RDBMS.

CHAPTER IV. The Relational Data Model

Duration: 2 hrs.

Overview	This chapter provides an overview on Database relations, primary keys and how to represent the data relationships in RDBMS. This chapter also explains about views and Data dictionary in RDBMS.
Learning Objectives	After completing this unit, you should be able to: • Understand the Relational Data Model in RDBMS • Understand concepts like database views and data dictionary.

CHAPTER V. Normalization

Duration: 3 hrs.

Overview	This chapter provides an overview on the Database Normalization and all the database normal forms (from first till sixth) and Boyce-Codd Normal form.
Learning Objectives	After completing this unit, you should be able to: • Understand the Database Normalization and all the database normal forms.

CHAPTER VI. Database Design and Performance Tuning

Duration: 2 hrs.

Overview	This chapter provides an overview on the performance and tuning of a database.
Learning Objectives	After completing this unit, you should be able to: • Understand the tuning of database

CHAPTER VII. Creating Database Objects

Duration: 4 hrs.

Overview	This chapter provides an overview of Data definition language and its operation.
Learning Objectives	After completing this unit, you should be able to: • Understand how DDLs used to create or modify the Schema, tables index etc.

CHAPTER VIII. Manipulating Data

Duration: 4 hrs.

Overview	This chapter provides an overview of Data manipulation language and its operation.
Learning Objectives	After completing this unit, you should be able to: • Understand how DMLs used to store, manipulate, retrieve data from tables.

CHAPTER IX. JDBC As the Fundamental Java API

Duration: 3 hrs.

Overview	This chapter provides an overview on JDBC Basics.
Learning Objectives	After completing this unit, you should be able to: • Understand on how to create Database and to connect through Java API, and CRUD operations using Java API

CHAPTER X. JPA as the JAVA ORM API

Duration: 4 hrs.

Overview	This chapter provides an overview of JPA
Learning Objectives	After completing this unit, you should be able to: • Understand on how to adopt JPA from JDBC and CRUD operation using JPA

CHAPTER XI. Database Security

Duration: 3 hrs.

Overview	This chapter provides an overview of Database Security on External Security threats, Internal threats.
Learning Objectives	After completing this unit, you should be able to: • Understand on external security threats, internal threats and the social remedies.

CHAPTER XII. Understanding Database Backup and Restore

Duration: 3 hrs.

Overview	This chapter provides an overview of Data Backup and Restore
Learning Objectives	After completing this unit, you should be able to: • Understand on different backups, how to take a Database backup, restore point.

CHAPTER XIII. Introduction to MySQL

Duration: 3 hrs.

Overview	This chapter provides an overview of basics of MySQL, basic queries
Learning Objectives	After completing this unit, you should be able to: • Understand how to insert, delete, select, update, where, drop, create queries in MySQL Database.

Cloud Application Developer

(Classroom)

Career path description

The Cloud Application Developer career path prepares students to develop, build, deploy, and test applications using a cloud platform to build Software as a Service (SaaS) solutions. This will require cloud application development skills, such as Node.js, REST architecture, JSON, Kubernetes, AI, Cloud Foundry and DevOps services. The courses in the Cloud Application Developer career path use tools to build, deploy, run, and manage applications on a cloud platform.

ibm.com/training

General information

Delivery method

15% self-paced and 85% instructor-led

Version

2019

Product

IBM Cloud

Audience

Undergraduate senior students from IT related academic programs such as computer science, software engineering, information systems and similar others

Learning objectives

After completing this course, you should be able to:
- Define cloud computing
- Describe the choices that are available to developers when creating cloud applications
- Describe infrastructure as a service, platform as a service, and software as a service
- Describe IBM Cloud
- Distinguish between the various IBM Cloud Platform compute options
  - IBM Cloud Foundry
  - IBM Cloud Kubernetes Service
  - IBM Cloud Function
  - Virtual Machines
  - VMware
- Understand the Twelve-Factor App methodology
- Describe how to create a Cloud Foundry application in IBM Cloud
- Describe how to create services in IBM Cloud
- Explain how to manage your IBM Cloud account with IBM Cloud CLI
- Deploy an application using IBM Cloud App Service (Web Apps)
- Explain containers and the difference between containers and Virtual Machines
- Describe container orchestration (Kubernetes)
- List Kubernetes key capabilities
- Realize the importance of using Kubernetes to prevent vendor lock-in
- Describe the Kubernetes building blocks:
  - Pod
  - Deployment
  - Service
- Describe DevOps
- Describe the capabilities of IBM Continuous Delivery
  - Identify the Web IDE features in IBM Continuous Delivery
  - Describe how to use Git Repos and Issue Tracking
  - Explain the pipeline build and deploy processes
- Create a Git repository to manage your source code
- View and edit code in the Eclipse Orion Web Integrated Development Environment (IDE)
- Build and Deploy code to IBM Cloud using DevOps
- Describe the characteristics of REST APIs
- Describe IBM Watson
- Provide examples of REST APIs using IBM Watson
- List the different database options available in IBM Cloud
- Explain the benefits of IBM Cloudant
- Describe how to design a cloud solution and enrich application using cloud services
- Explain the origin and purpose of the Node.js JavaScript framework
- Write a simple web server with Node.js
- Deploy an IBM SDK for Node.js application on an IBM Cloud account
- Create a Node.js module and use it in your code
- Understand asynchronous callbacks and be able to write the code in a Node.js application
- Create a Hello World Express application
- Use third-party modules in Node.js
- Use React to create interactive web pages
- Explain the core concepts of ECMAScript and React

Prerequisites Skills

- Basic knowledge of JavaScript.
- Basic understanding of HTML
- Basic understanding of the IBM Skills Academy learning environment

Duration

40 hours

Skill level

Basic – Intermediate

Hardware requirements

Classroom (ILT) setup requirements	
Processor	2.5 GHz or faster Core 2 Duo
GB RAM	4 GB
GB free disk space	30 GB
Network requirements	Yes
Other requirements	Valid Email address

Notes

The following unit and exercise durations are estimates, and might not reflect every class experience. If the course is customized or abbreviated, the duration of unchanged units will probably increase.

Course Agenda

MODULE I – CLOUD APPLICATION FOUNDATIONS

Course I – Introduction to HTML5 and JavaScript Programming

Duration: 5 hours

Course introduction
Duration: 5 minutes

Unit 1. HTML5 overview
Duration: 45 minutes

Overview	This unit describes some of the overall goals and features of HTML5.
Learning objectives	After completing this unit, you should be able to: • Describe what HTML does • List the objectives of HTML5 • Describe the document types that are supported in HTML5 • Describe the document object model (DOM) tree • Describe some of the differences between HTML4 and HTML5 • List some HTML document API properties and methods • Describe how scripting is enabled in browsers • Describe browser support for HTML5 features

Unit 2. JavaScript programming for web applications
Duration: 1 hour and 30 minutes

Overview	This unit covers some of the essential features of JavaScript. JavaScript is an important programming language for building the next generation of interactive web applications, including mobile web applications. In this unit, you learn how to use JavaScript to manipulate the document object model elements of an HTML page.
Learning objectives	After completing this unit, you should be able to: • Describe JavaScript primitives and objects • Explain how variables are declared and used in JavaScript • Describe JavaScript control structures • Describe functions in JavaScript • Describe the document object model (DOM) hierarchy • Describe the window and document objects • Identify the DOM objects that are commonly used in JavaScript applications for working with HTML documents

Exercise 1. Working with JavaScript in HTML documents
Duration: 45 minutes

Overview	In this exercise, you work with the Web Application Server Developer Tools for Eclipse environment to develop HTML documents and JavaScript functions.
Learning objectives	After completing this exercise, you should be able to: • Create HTML web pages • Use style statements in HTML documents • Connect scripts to documents • Write JavaScript functions • Create interactive alert and confirm window objects • Use JavaScript to modify the document object model (DOM)

Unit 3. HTML5 features
Duration: 45 minutes

Overview	This unit covers the new structural and form features of HTML5.
Learning objectives	After completing this unit, you should be able to: • List new elements in HTML5 • Describe HTML5 structural elements: section, article, header, footer, figure, figcaption • Describe the attributes of the HTML5 input element: tel, email, datetime, number, range, color

Exercise 2. Working with HTML5 features
Duration: 45 minutes

Overview	In this exercise, you explore some of the features available in HTML5.
Learning objectives	After completing this exercise, you should be able to: • Create a web page and insert a simple HTML5 form layout • Add new markup elements • Use input types that include attributes such as email to perform client-side validation • Test the application

Unit 4. Course summary
Duration: 5 minutes

Overview	This unit provides a summary of the course, a description of the class evaluation process, and information for future study.
Learning objectives	After completing this unit, you should be able to: • Explain how the course met its learning objectives • Submit an evaluation of the class • Identify other Web Application Server Education courses that are related to this course • Access the Web Application Server Education website • Locate appropriate resources for further study

MODULE II – CLOUD APPLICATION DEVELOPER

Course I – Essentials of Cloud Application Development

Duration: 19 hours

Course introduction
Duration: 15 minutes

Unit 1. Introduction to cloud computing
Duration: 60 minutes

Overview	This unit provides an overview about cloud computing. It lists characteristics and benefits of cloud computing and describes cloud computing service and deployment models.
Learning objectives	After completing this unit, you should be able to: • Define cloud computing. • Describe the characteristics of cloud. • Describe the benefits of cloud and the factors that contribute to its growth. • Describe cloud services models (IaaS, PaaS and SaaS). • Describe the cloud deployment options (Private, Public, Hybrid). • Describe cloud native applications and development methods. • Explain the Twelve-Factor App methodology • Describe the choices that developers have when building cloud applications.

Unit 2. Introduction to IBM Cloud
Duration: 120 minutes

Overview	This unit gets you started with IBM Cloud. It provides an overview of IBM Cloud services and the type of applications you can build on IBM Cloud. It explains how to create and manage cloud apps on IBM Cloud and how to manage users and resources.

Learning objectives	After completing this unit, you should be able to:
	• Describe IBM Cloud.
	• Identify the runtimes and services that IBM Cloud offers.
	• Distinguish among the various compute options on IBM Cloud.
	    o IBM Cloud Foundry
	    o IBM Cloud Kubernetes Service
	    o IBM Cloud Functions
	    o Virtual Machines
	    o VMware
	• Describe IBM Cloud regions, zones and multi-availability zones.
	• Describe the IBM Cloud dashboard, catalog, and documentation features.
	• Work with IBM Cloud resources.
	• Explain starter kits and Cloud Foundry boilerplates.
	• Describe how to manage your IBM Cloud users and resources (Optional).
	• Explain Identity and Access Management (IAM) and Resource Groups (Optional).
	• Describe how the application route is used to test an application in the browser.
	• Bind services to an application in IBM Cloud.
	• Describe the environmental variables that are used with IBM Cloud services.
	• Explain function as a service.

Unit 3. Deploying applications to Cloud Foundry on IBM Cloud
Duration: 60 minutes

Overview	This unit introduces Cloud Foundry and describes how to deploy applications to Cloud Foundry on IBM Cloud by using the IBM Cloud CLI. This unit describes the basic structure of Node.js apps.

Learning objectives	After completing this unit, you should be able to:
	• Explain Cloud Foundry basic concepts, including:
	    o Organizations and spaces
	    o Buildpacks
	    o Resiliency
	    o Logging and debugging
	    o Domains and routes
	    o Services binding
	• Explain the capabilities of the IBM Cloud command-line interface (CLI).
	• Describe the structure of the sample Node.js application.
	• Deploy the sample Node.js application by using the IBM Cloud CLI

Exercise 1. Getting started with Cloud Foundry apps on IBM Cloud
Duration: 75 minutes

Overview	This exercise describes how you can deploy a web app without downloading or configuring a runtime environment, or framework or setting up a server. This exercise also covers how to test and run the app when it is deployed.
Learning objectives	After completing this exercise, you should be able to: • Create an IBM Cloud application by using one of the available run times. • Install the IBM Cloud command-line interface (CLI). • Sign on to IBM Cloud from the CLI. • Deploy an application from a local workstation by using the IBM Cloud CLI. • Test the application with its endpoint after the application is deployed and started.

Unit 4. Adopting a DevOps approach by using IBM Continuous Delivery
Duration: 90 minutes

Overview	This unit introduces the features and functions of the DevOps services on the cloud development platform, IBM Cloud.
Learning objectives	After completing this unit, you should be able to: • Describe DevOps. • Describe the capabilities of IBM Cloud Continuous Delivery. • Identify the web-based integrated development environment (Web IDE) features in IBM Cloud Continuous Delivery. • Describe how to use source code management (such as Git) and Issue tracking. • Explain how to build and deploy applications using DevOps tools on IBM Cloud.

Exercise 2. Developing IBM Cloud applications with IBM Cloud Continuous Delivery
Duration: 90 minutes

Overview	In this exercise, you will explore DevOps capabilities in IBM Cloud. You work with the IBM Cloud Continuous Delivery services to explore, develop, build, and deploy IBM Cloud applications.
Learning objectives	After completing this exercise, you should be able to: • Enable an application to use IBM Cloud Continuous Delivery. • Create a Git repository to manage source code • View and edit code in the Eclipse Orion Web Integrated Development Environment (IDE). • Build and deploy code to IBM Cloud. • Test the application in IBM Cloud.

Unit 5. REST architecture and Watson APIs
Duration: 90 minutes

Overview	This unit introduces Representational State Transfer (REST) and Resources Representation and JavaScript Object Notation (JSON). It describes how to apply REST architecture concepts to server-side applications. This unit introduces Watson services and provides examples that show how to call Watson services by using REST APIs.
Learning objectives	After completing this unit, you should be able to: • Describe the main characteristics of REST APIs. • Explain the REST architecture style for designing networked applications. • List best practices to follow when using REST in your application. • Describe the representation format of data in REST. • Explain the advantages of the JSON data format. • Describe the security options for REST in your application • Describe IBM Watson services in IBM Cloud • Provide examples of REST APIs using IBM Watson.

Unit 6. Introduction to data services on IBM Cloud
Duration: 90 minutes

Overview	This unit provides an overview of the types of data stores that are used in cloud computing. You will also learn about the data services offerings that are available through the cloud development platform.
Learning objectives	After completing this unit, you should be able to: • Describe different databases types and capabilities • Describe the main types of data services in IBM Cloud. • Explain the benefits of IBM Cloudant. • Access Cloudant databases and documents on IBM Cloud. • Use HTTP APIs to interact with Cloudant database.

Exercise 3. IBM Cloud with Cloudant
Duration: 45 minutes

Overview	This exercise demonstrates how you can create a Cloudant database service on IBM Cloud without installing or configuring the database instance on your workstation. You use an HTTP API client such as Postman to create, read, update, and delete Cloudant documents. You create indexes and query data by using Cloudant API end points.
Learning objectives	After completing this exercise, you should be able to: • Create an instance of the Cloudant service on IBM Cloud. • Create service credentials by using IBM Cloud Identity and Access Management (IAM).. • Access the Cloudant documentation. • Explore the features of the Cloudant dashboard. • Create, read, update, and delete Cloudant documents by using HTTP APIs. • Verify the data that is stored in the database by using the Cloudant dashboard. • Create indexes and query Cloudant documents by using HTTP APIs.

Unit 7. Enriching your application with IBM Cloud services
Duration: 60 minutes

Overview	This unit presents an example of a cloud application to solve a business problem. It introduces functional and non-functional requirements, application architecture, and services on IBM Cloud that you can integrate with your application logic to implement a solution.

Learning objectives	After completing this unit, you should be able to:
	• Explain functional and non-functional requirements.
	• Design a simple architecture for cloud applications.
	• Identify services listed in the IBM Cloud catalog that you can integrate in your applications, such as:
	   o Database
	   o App ID
	   o Watson Natural Language Understanding
	   o Watson Tone Analyzer
	   o Log Analysis
	   o Monitoring
	• Integrate cloud services in a Node.js app by using REST APIs.

Exercise 4. Securing a web application with single sign-on (optional)
Duration: 60 minutes

Overview	In this exercise, you secure an application by using the App ID service for single sign-on by authenticating your application through trusted server providers.
Learning objectives	After completing this exercise, you should be able to: • Create an App ID service. • Bind the App ID service to an application to add single sign-on capability. • List and implement different configuration options for the App ID service.

Unit 8. Developing containerized applications on Kubernetes
Duration: 120 minutes

Overview	This unit introduces containers and containers orchestration. It provides an overview of the Kubernetes platform and describes basic concepts such as Kubernetes architecture, Kubernetes objects, and management of Kubernetes objects.

Learning objectives	After completing this unit, you should be able to:
	• Explain containers and the difference between containers and virtual machines.
	• Describe container orchestration.
	• List the key capabilities of Kubernetes.
	• Articulate the importance of using Kubernetes to prevent vendor lock-in.
	• Describe the Kubernetes building blocks:
	    o Pod
	    o Deployment
	    o Service
	• Scale and auto-scale your deployment for high availability.

Unit 9. IBM Cloud Kubernetes Service overview
Duration: 60 minutes

Overview	This unit introduces IBM Cloud Kubernetes Service and describes how a Kubernetes cluster lets you securely manage the resources that you need to deploy, update, and scale applications.
Learning objectives	After completing this unit, you should be able to: • Explore IBM Cloud Kubernetes Service on Kubernetes platforms. • Create a Kubernetes cluster by using the IBM Cloud Kubernetes Service. • Create containers and build on the IBM Cloud Container Registry. • Use Vulnerability Advisor to scan for vulnerabilities. • Manage Kubernetes clusters by using the kubectl command-line interface. • Deploy an application from a local workstation by using kubectl.

Exercise 5. Managing IBM Kubernetes Service clusters
Duration: 30 minutes

Overview	This exercise demonstrates how to create an IBM Kubernetes Service cluster and manage it by using the kubectl CLI.
Learning objectives	After completing this exercise, you should be able to: • Create an IBM Kubebernetes Service cluster. • Connect to a cluster on IBM Cloud Kubernetes Service. • List the worker nodes in a cluster.

Exercise 6. Deploying an application on Kubernetes
Duration: 90 minutes

Overview	In this exercise, you build a containerized application and deploy it to IBM Cloud Kubernetes Service.
Learning objectives	After completing this exercise, you should be able to: • Create a containerized Node.js application and build it on IBM Cloud Container Registry. • Explain how the container security analysis capability of Vulnerability Advisor can identify security vulnerabilities by scanning an image. • Create a deployment and scale it. • Expose your application on the internet.

Course II – Developing Node.js Applications on IBM Cloud

Duration: 11 hours

Course introduction
Duration: 15 minutes

Unit 1. Introduction to server-side JavaScript
Duration: 60 minutes

Overview	This unit introduces server-side JavaScript and Node.js. It describes how to create a Node.js server and implement Node.js modules.
Learning objectives	After completing this unit, you should be able to: • Explain the origin and purpose of the Node.js JavaScript framework • Write a simple web server with Node.js • Import Node.js modules into your script

---

Exercise 1. Developing a Hello World Node.js app on IBM Cloud
Duration: 90 minutes

Overview	In this exercise, you create a Node.js Cloud Foundry application on IBM Cloud. You will develop a Node.js-based server application (by using the Eclipse Orion Web IDE) that responds to web browser requests.
Learning objectives	After completing this exercise, you should be able to: • Create an IBM SDK for Node.js application. • Write your first Node.js application. • Deploy an IBM SDK for Node.js application on IBM Cloud. • Create a Node.js module and use it in your code.

Unit 2. Asynchronous I/O with callback programming
Duration: 90 minutes

Overview	The Node.js SDK relies on callback functions to handle network calls in an asynchronous manner. In this unit, you will learn how to write anonymous callback functions to act upon network events and listen, and intercept network traffic.
Learning objectives	After completing this unit, you should be able to: • Explain synchronous and asynchronous calls. • Write asynchronous calls code in Node.js applications. • Explain request flows that are sent to Node.js applications that use the http module.

Exercise 2. Understanding asynchronous callback
Duration: 90 minutes

Overview	This exercise shows how to use callback functions to call an external service. This exercise uses the IBM Watson Language Translator service in IBM Cloud. You create a Node.js module that contains the logic for these calls.
Learning objectives	After completing this exercise, you should be able to: • Write asynchronous callbacks code in Node.js applications.

Unit 3. Express web application framework
Duration: 90 minutes

Overview	This unit describes the Express web application framework, which provides a structured way to handle HTTP actions on server resources. You will learn how to write a REST service with Express and parse JSON data from an HTTP message.
Learning objectives	After completing this unit, you should be able to: • Explain the difference between code that is written in "pure" JavaScript and code that is written with the Express framework. • Explain what Express is and its benefits. • Use Express as a third-party npm package. • Explain the use of middleware functions. • Handle routes and requests..

Exercise 3. Creating your first Express application
Duration: 90 minutes

Overview	In this exercise, you create an application that uses the Express framework and the IBM Watson Natural Language Understanding service to extract the author name from articles that are published on the web. You provide the web address (URL) of the article to the application, and it outputs the name of the author (or multiple names if the article has multiple authors).
Learning objectives	After completing this exercise, you should be able to: • Create a Hello World Express application. • Create a simple HTML view for your application. • Explain Express routing. • Use third-party modules in Node.js. • Use the Watson Natural Language Understanding service in your applications. • Use a Git repository in DevOps on IBM Cloud.

Unit 4. Async patterns with ECMAScript
Duration: 60 minutes

Overview	This unit describes async patterns in ECMAScript including callbacks, promises, and async/await.
Learning objectives	After completing this unit, you should be able to: • Explain async patterns in ECMAScript such as callbacks, promises, and async/await.

Exercise 4. Building a rich UI application by using React and ES8
Duration: 90 minutes

Overview	This exercise guides you through building an interactive and rich client-side application by using React. You also explore the async/await feature of ECMAScript 2017, which is commonly known as ES8, and some features of ES8 through a server-side application by using Node.js.
Learning objectives	After completing this exercise, you should be able to: • Deploy a React application on IBM Cloud. • Deploy a Node.js application on IBM cloud. • Explain the structure of a React application. • Use ES8 features in Node.js applications.

Unit 5. Building rich UI applications with React (optional)
Duration: 60 minutes

Overview	This unit describes async patterns in JavaScript. It introduces React and basic React concepts such as: • Components • Props • State • Events • Component lifecycle • Nested components • Lists and keys.
Learning objectives	After completing this unit, you should be able to: • Explain the React component lifecycle. • Explain React states, props, and events. • Create lists of nested components. • Explain React deployment options on IBM Cloud.

IBM Official Badges and Associated Job Roles

IBM Official Badges	• Cloud Application Developer: Explorer • Cloud Application Developer: Mastery Award
Associated Job Roles	• Cloud Solution Administrator • Mobile Application Developer • Mobile Application Administrator • Business Process Developer

For more information

To learn more about this career path and others, see www.ibm.biz/ibmskillsacademy

To learn more about validating your technical skills with IBM Open Badges, see www.youracclaim.com

To stay informed about the IBM Skills Academy, see the following sites:

Facebook: www.facebook.com/ibmskillsacademy

WATSON STUDIO TRAINING MODULE

(CLASSROOM)

WATSON STUDIO TRAINING MODULE

The training module will make the reader accustomed to Watson Studio . This material will help the reader in understanding the basics of the Watson Studio and the use of Watson Studio for analytics.

DELIVERY METHOD

25 % Self-paced Learning

75 % Instructor led

training

VERSION

2019

LEARNING OBJECTIVES

- Watson Studio – Introduction
- Watson Knowledge Catalog
- Creating Projects
- Administering Projects
- Environment Definition
- Ingesting Streaming data
- Data refinery – Preparing the data
- Streams flow
- Using Notebooks in Watson Studio
- Decision Optimization
- Machine learning & AI using Watson Studio
- Natural Language Processing using Watson Studio
- Visual recognition Models using Watson Studio
- Deploying Models
- Deploying apps using Python Flask
- Catalogs – Administration
- Auto AI

PREREQUISITES SKILLS

- Computer Science fundamentals
- Basic knowledge of applied math, algorithms, and data modelling
- Basic knowledge of statistics

DURATION

32 Hours

SKILL LEVEL

Basic – Intermediate

HARDWARE REQUIREMENTS (For local version)

Processor	2 GHz or Higher
GB RAM	8 GB
GB Disk Free	80 GB
Network Requirements	Yes

Notes

The following unit and exercise durations are estimates and might not reflect every class experience. The estimates do not include the duration of optional exercises or sections. Students in this course use an IBM Cloud Lite account to perform the exercises. This account will never expire; therefore, students can continue working on the optional exercises after the class

COURSE AGENDA

UNIT I. Introduction to Watson Studio – Getting Started

Duration: 6 Hrs.

Overview	This unit explains what is Watson Studio. How this product provides an end to end framework for analytics – Development to deployment of models.
Learning Objectives	After completing this unit, you should be able to: • Understand the product • How to install and get started with Watson Studio • Overview of the product

UNIT II. Projects

Duration: 8 Hrs.

Overview	A project is how you organize your resources to work with data. How do we create projects?
Learning Objectives	After completing this unit, you should be able to: • Understand Assets • Understand Environments to run Watson Studio models • Access Control • Collaboration in Watson Studio

UNIT III. Preparing data

Duration: 8 Hrs.

Overview	After you create a project, or join one, the next step is to add data to the project and prepare the data for analysis.
Learning Objectives	After completing this unit, you should be able to: • Adding data to a project • Refining data with data refinery • Cleaning data within Watson Studio

UNIT IV. Data Science

Duration: 4 hrs.

Overview	This unit explains how to apply data science using Watson Studio.
Learning Objectives	After completing this unit, you should be able to: • Analyze data by writing code in notebooks. • Develop applications in RStudio. • Visualize your data without coding with analytic dashboards.

UNIT V. Machine learning and AI in Watson Studio

Duration: 4 hrs.

Overview	This units give an in depth details about Machine Leaning & AI built in within Watson Studio.
Learning Objectives	After completing this unit, you should be able to: • Visual recognition • Natural language classification • Watson Machine Learning

UNIT VI. Catalog

Overview	This units give an in depth details about Machine Leaning & AI built in within Watson Studio.
Learning Objectives	After completing this unit, you should be able to: • Creating a catalog • Finding and viewing an asset • Adding an asset to a catalog • Adding catalog data assets to a project.

UNIT VII. Governance & Administration

Overview	This units give an in depth details about Machine Leaning & AI built in within Watson Studio.
Learning Objectives	After completing this unit, you should be able to: • Policies • Data protection rules • Business terms and data classification • Managing Watson apps • Determining roles

Predictive Analytics Modeler

(Classroom)

Career path description

The Predictive Analytics Modeler career path prepares students to learn the essential analytics models to collect and analyze data efficiently. This will require skills in predictive analytics models, such as data mining, data collection and integration, nodes, and statistical analysis. The Predictive Analytics Modeler will use tools for market research and data mining in order to predict problems and improve outcomes.

ibm.com/training

General information

Delivery method

20% web-based and 80% instructor led

Version

2020

Product

IBM SPSS Modeler
IBM Watson Studio, IBM Watson Machine Learning

Audience

Undergraduate senior students from IT related academic programs i.e. computer science, software engineering, information systems and similar others

Learning objectives

After completing this course, you should be able to:
- The importance of analytics and how its transforming the world today
- Understand how analytics provided a solution to industries using real case studies
- Explain what is analytics, the various types of analytics, and how to apply it
- Improve efficiency, sample records, and work with sequence data
- Explain data transformations, and functions
- Understand modeling, relationships, derive and reclassify fields
- Integrate and collect data
- Understand the principles of data mining
- Use the user interface of modeler to create basic program streams
- Read a statistics data file into modeler and define data characteristics
- Review and explore data to look at data distributions and to identify data problems, including missing values
- Use the automated data prep node to further prepare data for modeling
- User a partition node to create training and testing data subsets

Prerequisites Skills
- English Proficiency
- Basic Internet and web browser usage experience
- Basic analytics experience
- Exposure to the IBM Skills Academy Portal learning environment
- Exposure to the IBM Skills Academy Cloud hands-on labs platform

Duration

57.75 hours

Skill level

Basic – Intermediate

Hardware requirements

Classroom (ILT) setup requirements	
Processor	Intel Core i7 CPU @ 2.7 GHz
GB RAM	8 GB
GB free disk space	60 GB
Network requirements	No
Other requirements	IBM ID

Notes

The following unit and exercise durations are estimates, and might not reflect every class experience. If the course is customized or abbreviated, the duration of unchanged units will probably increase.

Course Agenda

MODULE I – ANALYTICS OVERVIEW

Course I – Business Analytics Overview

Duration: 1 hour

Course introduction
Duration: 5 minutes

Unit 1. Analytics overview
Duration: 10 minutes

Overview	This unit provides an understanding of the importance of business analytics in our world, society, and life.
Learning objectives	After completing this unit, you should be able to: • Understand how analytics is transforming the world • Understand the profound impact of analytics in business decisions • Understand what is analytics and how it works • Understand why business analytics has become important in various industries

---

Unit 2. Analytics trends: Past, present & future
Duration: 15 minutes

Overview	This unit explains how analytics has evolved over time.
Learning objectives	After completing this unit, you should be able to: • Understand the history of analytics and how it has changed today • Understand how to analyze unstructured data • Understand how analytics is making the world smarter • Understand where the future of analytics lies

---

Unit 3. Towards a predictive enterprise
Duration: 5 minutes

Overview	This unit explains the effects of business analytics in the corporate world that has led to its global adoption across geographies and industries.
Learning objectives	After completing this unit, you should be able to: • Explain why successful enterprises need business analytics • Understand how business analytics can help turn data into insight

Unit 4. Analytics: Industry domains
Duration: 5 minutes

Overview	This unit highlights the application of analytics across major industries.
Learning objectives	After completing this unit, you should be able to: • Understand how predictive analytics is transforming all types of organizations • Explain how analytics supports retail companies • Understand how analytics can reduce crime rates and accidents • Explain the use of analytics in law enforcement and insurance companies • Understand how analytics can affect the future of education

Unit 5. Case studies and solutions
Duration: 15 minutes

Overview	This unit covers real case studies and solutions of the adoption of business analytics across the world.
Learning objectives	After completing this unit, you should be able to: • Understand the importance of business analytics • Comprehend how big data and analytics can help in understanding consumer/customer behavior • Explain how analytics can help manage assets • Understand how analytics can help combat fraud • Explain how analytics can help us to understand social sentiments

MODULE II – Business Analytics Foundations

Course I – Business Intelligence and Analytics 101

Duration: 1 hour

Course introduction **Duration: 5 minutes**	

Business Intelligence and Analytics 101 **Duration: 1 hour**	
Overview	This course provides a collection of resources designed for participants to become familiar with business intelligence (BI) and analytics concepts. Participants will review materials to introduce themselves to terminology and practical business use cases for a high level understanding of BI and analytics. The course includes a pre-assessment for participants to measure their understanding of the content before taking the course, and a post-assessment for participants to gauge their learning after reviewing the materials.
Learning objectives	After completing this course, you should be able to: • Explain what is analytics • Define various types of analytics • Demonstrate how to apply analytics • Describe business intelligence • Demonstrate how to apply business intelligence

MODULE III – PREDICTIVE ANALYTICS MODELER

Course I – Introduction to A Predictive Analytics Platform & Data Mining

Duration: 27.2 hours

Course introduction **Duration: 10 minutes**	

Unit 1. Introduction to data mining **Duration: 1 hour**	
Overview	In this unit, you will learn about data mining and its applications.
Learning objectives	After completing this unit, you should be able to: • List two applications of data mining • Explain the stages of the CRISP-DM process model • Describe successful data-mining projects and the reasons why projects fail • Describe the skills needed for data mining

Exercise 1. The ACME business case: Modeling response for a campaign
Duration: 1 hour

Overview	In this exercise, you will learn how to apply data mining.
Learning objectives	After completing this exercise, you should be able to: • Understand data mining • Describe how to apply data mining in different scenarios

Unit 2. Working with IBM SPSS Modeler
Duration: 1 hour

Overview	In this unit, you will learn about objects such as streams and nodes and you will acquire experience with the software.
Learning objectives	After completing this exercise, you should be able to: • Describe the MODELER user-interface • Work with nodes • Run a stream or a part of a stream • Open and save a stream • Use the online Help

Exercise 1. Work with IBM SPSS Modeler
Duration: 45 minutes

Overview	In this exercise, you will learn about MODELER's user-interface to create streams.
Learning objectives	After completing this exercise, you should be able to: • Create streams • Change streams • Generate a select node from the Table output

Unit 3. Creating a data-mining project
Duration: 1 hour

Overview	In this unit you will learn about building a model and then applying that model to future cases of a data-mining project.
Learning objectives	After completing this unit, you should be able to: • Explain the basic framework of a data-mining project • Build a model • Deploy a model

Exercise 1. Create a data mining project to predict response in an ACME campaign
Duration: 45 minutes

Overview	In this exercise, you will build a model using data of the test mailing. This model (hopefully) identifies groups with high response rates. You will then use this model to select the groups with high response rates in the rest of the customer database (only these groups will be included in the actual mailing for the XL Original Orange Baseball Cap).
Learning objectives	After completing this exercise, you should be able to: • Build a model using historical data • Deploy the model

Unit 4. Collecting initial data
Duration: 1 hour

Overview	In this unit, you will learn how to collect initial data. You will also learn how to describe data.
Learning objectives	After completing this exercise, you should be able to: • Explain the concepts of data structure, unit of analysis, field storage and field measurement level • Import Microsoft Excel files • Import text files • Import from databases • Export data to various formats

Exercise 1. Collect initial data for ACME
Duration: 45 minutes

Overview	In this exercise you will learn how to collect initial data for ACME
Learning objectives	After completing this exercise, you should be able to: • Collect initial data for ACME

Unit 5. Understanding your data
Duration: 1 hour

Overview	In this unit, you will learn how to explore data and assess it's quality.
Learning objectives	After completing this exercise, you should be able to: • Audit the data • Explain how to check for invalid values • Take action for invalid values • Explain how to define blanks

Exercise 1. Understand the ACME data
Duration: 1 hour

Overview	In this exercise, you will learn how to understand the ACME data
Learning objectives	After completing this exercise, you should be able to: • Import and examine data

Unit 6. Setting the unit of analysis
Duration: 1 hour

Overview	In this exercise, you will learn how to set unit of analysis in three different methods.
Learning objectives	After completing this exercise, you should be able to: • Set the unit of analysis by removing duplicate records • Set the unit of analysis by aggregating records • Set the unit of analysis by expanding a categorical field into a series of flag fields

Exercise 1. Set the unit of analysis for the ACME data
Duration: 45 minutes

Overview	In this exercise, you will learn how to set the unit of analysis for the ACME data.
Learning objectives	After completing this exercise, you should be able to: • Cleanse data by removing duplicate records • Expand a categorical field into a series of flag fields

Unit 7. Integrating data
Duration: 1 hour

Overview	In this exercise you will learn how to combine different datasets into a single dataset for analysis.
Learning objectives	After completing this exercise, you should be able to: • Integrate data by appending records from multiple datasets • Integrate data by merging fields from multiple datasets • Sample records

Exercise 1. Integrate ACME data
Duration: 45 minutes

Overview	In this exercise, you will learn how to combine a number of datasets into a single dataset as a preparation for analysis and modeling.
Learning objectives	After completing this exercise, you should be able to: • Append records from two datasets • Merge fields from different datasets • Enrich a dataset with aggregated data • Sample records

Unit 8. Deriving and reclassifying fields
Duration: 1 hour

Overview	In this unit, you will learn how to construct the final dataset for modeling by cleansing and enriching your data.
Learning objectives	After completing this exercise, you should be able to: • Use the Control Language for Expression Manipulation (CLEM) • Derive new fields • Reclassify field values

Exercise 1. Derive and reclassify fields for the ACME data
Duration: 45 minutes

Overview	In this exercise, you will learn how to cleanse and enrich a dataset to build models.
Learning objectives	After completing this exercise, you should be able to: • Cleanse data and derive fields for modeling • Cleanse data and reclassify fields for modeling

Unit 9. Identifying relationships
Duration: 1 hour

Overview	In this unit, you will learn methods used to examine the relationship between two fields.
Learning objectives	After completing this exercise, you should be able to: • Examine the relationship between two categorical fields • Examine the relationship between a categorical field and a continuous field • Examine the relationship between two continuous fields

Exercise 1. Identify relationships in the ACME data
Duration: 45 minutes

Overview	In this exercise you will learn how to assess relationships and determine its strength by doing a demo.
Learning objectives	After completing this exercise, you should be able to: • Assess the relationship between churn and handset • Assess the relationship between churn and number of dropped calls • Assess the relationship between number of products and revenues

Unit 10. Introduction to modeling
Duration: 1 hour

Overview	In this unit, you will learn about the modeling stage of the CRISP-DM process model.
Learning objectives	After completing this exercise, you should be able to: • List three modeling objectives • Use a classification model • Use a segmentation model

Exercise 1. Predict response in ACME campaigns
Duration: 45 minutes

Overview	In this exercise you will learn about classification and segmentation using a synthetic dataset from a telecommunications firm.
Learning objectives	After completing this exercise, you should be able to: • Predict churn by running a CHAID model • Predict churn by running a Neural Net model • Compare the accuracy of these models • Find groups of similar customers, based on usage.

Course II – Advanced data preparation using IBM SPSS Modeler

Duration: 13.75 hours

Course introduction
Duration: 10 minutes

Unit 1. Using functions to cleanse and enrich data
Duration: 1 hour

Overview	In this unit, you will learn how to use various different kinds of functions.
Learning objectives	After completing this unit, you should be able to: • Use date functions • Use conversion functions • Use string functions • Use statistical functions • Use missing value functions

---

Exercise 1. Using functions to clean and enrich travel agency data
Duration: 30 minutes

Overview	In this exercise you will work with data about customers and their holiday destinations. You will derive new fields to answer questions such as "What is the mean age of the customers?", "What was the most popular month to travel?", "What was the most popular destination?", and "What was the mean amount of money spent?"
Learning objectives	After completing this exercise, you should be able to: • Use functions to clean and enrich travel data

---

Unit 2. Using additional field transformations
Duration: 1 hour

Overview	In this unit, you will learn about using additional field transformations.
Learning objectives	After completing this exercise, you should be able to: • Replace values with the Filler node • Recode continuous fields with the Binning node • Change a field's distribution with the Transform node

Exercise 1. Use additional field transformations to prepare travel agency data for modeling
Duration: 1 hour

Overview	In this exercise, you will use additional field transformations to prepare travel agency data for modeling
Learning objectives	After completing this exercise, you should be able to: • Use additional field transformations to prepare travel agency data

Unit 3. Working with sequence data
Duration: 1 hour

Overview	In this unit, you will learn how to work with sequenced data.
Learning objectives	After completing this exercise, you should be able to: • Use cross-record functions • Count an event across records • Expand a continuous field into a series of continuous fields with the Restructure node • Use geospatial and time data with the Space-Time-Boxes node

Exercise 1. Prepare sequence data of a travel agency for analysis
Duration: 45 minutes

Overview	In this exercise you will work with a dataset storing customers and their holidays. You will derive new fields to answer questions such as "What is the mean age of the customers?", "What was the most popular country?" and so forth.
Learning objectives	After completing this exercise, you should be able to: • Prepare sequence data of a travel agency for analysis

Exercise 2. Determine the availability or taxis
Duration: 1 hour

Overview	In this exercise, you will learn how to determine the availability of taxis.
Learning objectives	After completing this exercise, you should be able to: • Determine the availability of taxis

Unit 4. Sampling, partitioning, and balancing data
Duration: 1 hour

Overview	In this unit, you will learn how to use the Sample node and various reasons for sampling records.
Learning objectives	After completing this exercise, you should be able to: • Use the Sample node to draw simple and complex samples • Partition the data into a training and a testing set • Reduce or boost the number of records

Exercise 1. Sample, partition, and balance house property data
Duration: 45 minutes

Overview	In this exercise, you will sample, partition, and balance charity data.
Learning objectives	After completing this exercise, you should be able to: • Sample, partition, and balance charity data

Unit 5. Improving efficiency
Duration: 1 hour

Overview	In this exercise, you will learn how to work with SQL pushback, Set Globals node and parameters to optimize efficiency.
Learning objectives	After completing this exercise, you should be able to: • Use database scalability by SQL pushback • Use the Data Audit node to process outliers and missing values • Use the Set Globals node • Use parameters • Use looping and conditional execution

Exercise 1.Improve efficiency with travel agency data
Duration: 45 minutes

Overview	In this exercise you will process outliers, extremes and missing values, using the Data Audit node. You will use the Set Globals node to replace missing values, and you will be introduced to automation by using parameters and looping.
Learning objectives	After completing this exercise, you should be able to: • Use the Data Audit node to process outliers, extremes and missing values • Compute standardized scores using globals • Use parameters • Create a loop through values

Course III – Predictive Analytics with IBM Watson Studio

Duration: 12 hours

Course introduction
Duration: 15 minutes

Unit 1. Introduction to IBM Watson Studio
Duration: 30 minutes

Overview	This unit provides a high level overview of IBM Watson Studio, its components, key applications and the value added by the IBM offering.
Learning objectives	After completing this unit, you should be able to: • Describe Watson Studio. • Identify industry use cases. • List Watson Studio offerings. • Create Watson Studio projects. • Describe Watson Studio and Spark environment. • Describe Watson Studio and Object Storage. • Explain Watson Studio high availability considerations. • Prepare and analyze data. • Use Jupyter notebooks.

Exercise 1. Getting started with Watson Studio
Duration: 1 hour 30 minutes

Overview	This exercise introduces you to the basic tasks that you perform when you use Watson Studio.
Learning objectives	After completing this exercise, you should be able to: • Create a Watson Studio project. • Manage the project. • Assign collaborators. • Load a data set into the project's object store. • Manage Cloud Object Storage. • Analyze data by using Watson Studio. • Use scikit-learn for linear regression.

Unit 2. Introduction to IBM Watson Machine Learning
Duration: 1 hour

Overview	This unit provides an overview of the IBM Watson Machine Learning service available on IBM Cloud. It explains the process of preparing the data before it is provided to machine learning algorithms. This unit describes the use of the Data Refinary tool to cleanse and shape tabular data with a graphical flow editor.
Learning objectives	After completing this unit, you should be able to: • Describe data preparation before feeding into machine learning algorithms. • Describe Watson Machine Learning features and capabilities. • Describe the Data Refinery tool. • List the data formats and sizes that Data Refinery operates on. • List the Data Refinery features and explain each feature.

Exercise 2. Getting started with Data Refinery and AutoAI
Duration: 1 hour

Overview	In the first part, this exercise introduces you to the basic tasks of refining data with Data Refinery on Watson Studio. In the second part, it introduces you to AutoAI on Watson Studio for machine learning models creation.
Learning objectives	After completing this exercise, you should be able to: • Import data into Data Refinery in an existing project. • Review the data with the Profile and Visualizations features • Refine the data by that use various shaping operations. • Run a job for the Data Refinery flow. • Create a model by using the AutoAI graphical tool in Watson Studio.

Unit 3. Introduction to neural networks and deep learning
Duration: 1 hour 90 minutes

Overview	This unit introduces machine learning models that are inspired by the structure of the human brain, which is known as neural networks. Then, this unit provides an overview to deep learning, which is a machine learning technique that uses neural networks to learn.
Learning objectives	After completing this unit, you should be able to: • Describe neural networks. • Explain the concepts of Perceptron and back propagation. • Explain what an activation function is and identify the most common examples. • Articulate the difference between deep and shallow neural networks. • List the reasons for the current deep learning emergence. • Identify the basic architectures of deep neural networks and their applications. • Describe the functions of IBM Watson Studio Neural Network Modeler.

Exercise 3. Exploring deep learning and neural network modeling with Watson Studio
Duration: 1 hour 30 minutes

Overview	This exercise guides you step-by-step through the design of a deep learning neural network architecture based on a sample flow that is provided to you. You create your own convolutional neural network with Watson Studio.
Learning objectives	After completing this exercise, you should be able to: • Build a neural network to recognize handwritten digits. • Create a neural network design flow by using the neural network modeler. • Train models with experiment builder.

Unit 4. IBM Watson Studio Jobs
Duration: 30 minutes

Overview	This unit describes how to use IBM Watson Studio Jobs to run data operations assets such as notebooks and Data Refinery flows. This unit explains how to create and run jobs from a Data Refinery flow and a notebook.
Learning objectives	After completing this unit, you should be able to: • Explain the purpose and function of jobs. • Explain how to create jobs by using a project, Data Refinery flow, or a Notebook. • Describe how to view, edit, and run jobs. • Use Watson Studio Jobs on a practical use case to automate model training and deployment.

Exercise 4. Automating data preparation, model training, and deployment with Watson Studio jobs
Duration: 1 hour 30 minutes

Overview	This exercise demonstrates how to use Watson Studio jobs to run an end to end scenario in a data science project. You use previously developed artifacts to create Watson Studio jobs to prepare your data, train the model, deploy the trained model, and score your data.
Learning objectives	After completing this exercise, you should be able to: • Create data preparation jobs from Data Refinery flows. • Create jobs from notebooks to train, evaluate, and deploy models. • Create jobs from notebooks to score your data.

IBM Official Badges and Associated Job Roles

IBM Official Badges	Predictive Analytics Modeler: Explorer I Mastery Award
Associated Job Roles	• Business Intelligence Analyst • Big Data Developer • Data Warehouse Developer

For more information

To learn more about this career path and others, see **ibm.biz/ibmskillsacademy**

To learn more about validating your technical skills with IBM Open Badges, see www.youracclaim.com

To stay informed about the IBM Skills Academy, see the following sites:

Facebook: www.facebook.com/ibmskillsacademy

Big Data Engineer

(Classroom)

Career path description

The Big Data Engineer career path prepares students to use the Big Data platform and methodologies in order to collect and analyze large amounts of data from different sources. This will require skills in Big Data architecture, such as Apache Hadoop, Ambari, Spark, Big SQL, HDFS, YARN, MapReduce, ZooKeeper, Knox, Sqoop, and HBase.

ibm.com/training

General information

Delivery method

95% instructor led and 5% web-based

Version

2018

Product

HDP Open Source and IBM Watson Studio

Audience

Undergraduate senior students from IT related academic programs i.e. computer science, software engineering, information systems and similar others

Learning objectives

After completing this course, you should be able to understand the following topics:
- Big Data and Data Analytics
- Hortonworks Data Platform (HDP)
- Apache Ambari
- Hadoop and the Hadoop Distributed File System
- MapReduce and YARN
- Apache Spark
- Storing and Querying data
- ZooKeeper, Slider, and Knox
- Loading data with Sqoop
- DataPlane Service
- Stream Computing
- Data Science essentials
- Drew Conway's Venn Diagram - and that of others
- The Scientific Process applied to Data Science
- The steps in running a Data Science project
- Languages used for Data Science (Python, R, Scala, Julia, …)
- Survey of Data Science Notebooks
- Markdown language with notebooks
- Resources for Data Science, including GitHub
- Jupyter Notebook
- Essential packages: NumPy, SciPy, Pandas, Scikit-learn, NLTK, BeautifulSoup…
- Data visualizations: matplotlib, …, PixieDust
- Using Jupyter "Magic" commands
- Using Big SQL to access HDFS data
- Creating Big SQL schemas and tables
- Querying Big SQL tables
- Configuring Big SQL security
- Data federation with Big SQL
- IBM Watson Studio
- Analyzing data with Watson Studio

Prerequisites Skills

- Basic knowledge of Linux
- Basic SQL knowledge
- Working knowledge with big data and Hadoop technologies
- Have a basic understanding of notebook technologies for data science
- Students can attend free courses at www.bigdatauniversity.com to acquire the necessary requirements
- Exposure to the IBM Skills Academy Portal learning environment
- Exposure to the IBM Skills Academy Cloud hands-on labs platform

Duration

32.7 hours

Skill level

Basic – Intermediate

Classroom (ILT) setup requirements	
Processor	3 GHz or higher
GB RAM	20 GB
GB free disk space	80 GB
Network requirements	Yes
Other requirements	IBM ID

Notes

The following unit and exercise durations are estimates, and might not reflect every class experience. If the course is customized or abbreviated, the duration of unchanged units will probably increase.

Course Agenda

MODULE I – BIG DATA OVERVIEW

Course I – Introduction to the Big Data Ecosystem

Duration: 1.6 hours

Course introduction Duration: 5 minutes	

Unit 1. Introduction to Big Data Duration: 90 minutes	

Overview	In this unit you will learn about Big Data and understand why it's important.
Learning objectives	After completing this unit, you should be able to: • Understand what Big Data is • Develop an understanding of the complete open-source Hadoop ecosystem and its near-term future directions • Understand the major challenges of data • Understand how the growth of interconnected devices helps big data • List some real life examples of Big Data • Learn the types of Big Data • Student some Big Data use cases

MODULE II – Prerequisites
This course does not have any prerequisites

MODULE III – Big Data Engineer

Course I – Introduction to the Big Data Ecosystem

Duration: 19.5 hours

Course introduction Duration: 5 minutes	

Unit 1. Introduction to Big Data
Duration: 30 minutes

Overview	In this unit you will learn about Big Data and understand why it's important.
Learning objectives	After completing this unit, you should be able to: • Develop an understanding of the complete open-source Hadoop ecosystem and its near-term future directions • Be able to compare and evaluate the major Hadoop distributions and their ecosystem components, both their strengths and their limitations • Gain hands-on experience with key components of various big data ecosystem components and their roles in building a complete big data • solution to common business problems • Learning the tools that will enable you to continue your big data education after the course

Unit 2. Introduction to Hortonworks Data Platform (HDP)
Duration: 30 minutes

Overview	In this unit you will learn about the Hortonworks Data Platform (HDP).
Learning objectives	After completing this unit, you should be able to: • Describe the functions and features of HDP • List the IBM value-add components • Explain what IBM Watson Studio is • Give a brief description of the purpose of each of the value-add components

Lab 1. Exploration of the lab environment
Duration: 1 hour

Overview	In this lab, you will explore the lab environment. You will access your lab environment and launch Apache Ambari. You will startup a variety of services by using the Ambari GUI. You will also explore some of the directory structure on the Linux system that you will be using.
Learning objectives	After completing this lab, you should be able to: • Explore the lab environment • Launch Apache Ambari • Start a variety of services using Apache GUI • Explore some of the directory structure on the Linux system

Unit 3. Apache Ambari
Duration: 30 minutes

Overview	In this section you will learn about Ambari, which is one of the operations tools that come with HDP.
Learning objectives	After completing this unit, you should be able to: • Understand the purpose of Apache Ambari in the HDP stack • Understand the overall architecture of Ambari, and Ambari's relation to other services and components of a Hadoop cluster • List the functions of the main components of Ambari • Explain how to start and stop services from Ambari Web Console

Lab 1. Managing Hadoop clusters with Apache Ambari
Duration: 1 hour

Overview	In this lab you will explore the Apache Ambari web console and perform basic starting and stopping of services, giving you experience in using Apache Ambari to manage your Hadoop cluster.

Learning objectives	After completing this lab, you should be able to:
	• Manage Hadoop clusters with Apache Ambari
	   o Start the Apache Ambari web console and perform basic start/stop services
	   o Explore other aspects of the Ambari web server

Unit 4. Hadoop and HDFS
Duration: 1 hour

Overview	This unit will explain the underlying technologies that are important to solving the bigdata challenge.
Learning objectives	After completing this unit, you should be able to: • Understand the basic need for a big data strategy in terms of parallel reading of large data files and internode network speed in a cluster • Describe the nature of the Hadoop Distributed File System (HDFS) • Explain the function of the NameNode and DataNodes in an Hadoop cluster • Explain how files are stored and blocks ("splits") are replicated

Lab 1. File access and basic commands with HDFS
Duration: 1 hour

Overview	This lab is intended to provide you with experience in using the Hadoop Distributed File System (HDFS). The basic HDFS file system commands learned here will be used throughout the remainder of the course. You will also be moving some data into HDFS that will be used in later units of this course. The files that you will need are stored in the Linux directory /home/labfiles.
Learning objectives	After completing this lab, you should be able to: • File access and basic commands with HDFS

Unit 5. MapReduce and YARN
Duration: 2 hours

Overview	In this unit you will learn about MapReduce and YARN.
Learning objectives	After completing this unit, you should be able to: • Describe the MapReduce model v1 • List the limitations of Hadoop 1 and MapReduce 1 • Review the Java code required to handle the Mapper class, the • Reducer class, and the program driver needed to access MapReduce • Describe the YARN model • Compare Hadoop 2/YARN with Hadoop 1

Lab 1. Running MapReduce and YARN jobs
Duration: 1 hour

Overview	In this lab, you will run Java programs using Hadoop v2, YARN, and related technologies.
Learning objectives	After completing this lab, you should be able to: • Run MapReduce and YARN jobs

Lab 2. Creating and coding a simple MapReduce job
Duration: 1 hour

Overview	In this lab, you will compile and run a more complete version of WordCount that has been written specifically for MapReduce2.
Learning objectives	After completing this lab, you should be able to: • Create and code a simple MapReduce job

Unit 6. Apache Spark
Duration: 2 hours

Overview	In this unit you will learn about Apache Spark.
Learning objectives	After completing this unit, you should be able to: • Understand the nature and purpose of Apache Spark in the Hadoop ecosystem • List and describe the architecture and components of the Spark unified stack • Describe the role of a Resilient Distributed Dataset (RDD) • Understand the principles of Spark programming • List and describe the Spark libraries • Launch and use Spark's Scala and Python shells

Lab 1. Working with a Spark RDD with Scala
Duration: 1 hour

Overview	In this lab, you will learn to use some of the fundamental aspects of running Spark in the HDP environment.
Learning objectives	After completing this lab, you should be able to: • Work with Spark RDD with Scala

Unit 7. Storing and querying data
Duration: 2 hours

Overview	In this unit you will learn about storing and querying data.
Learning objectives	After completing this unit, you should be able to: • List the characteristics of representative data file formats, including flat/text files, CSV, XML, JSON, and YAML • List the characteristics of the four types of NoSQL datastores • Describe the storage used by HBase in some detail • Describe and compare the open source programming languages, Pig and Hive • List the characteristics of programming languages typically used by • Data Scientists: R and Python

Lab 1. Using Hive to access Hadoop/HBase data
Duration: 30 minutes

Overview	In this lab, you will use Hive to access Hadoop/HBase data.
Learning objectives	After completing this lab, you should be able to: • Use Hive to access Hadoop/HBase data

Unit 8. ZooKeeper, Slider, and Knox
Duration: 1 hour

Overview	In this unit you will learn about ZooKeeper, Slider and Knox.
Learning objectives	After completing this unit, you should be able to: • Understand the challenges posed by distributed applications and how ZooKeeper is designed to handle them • Explain the role of ZooKeeper within the Apache Hadoop infrastructure and the realm of Big Data management • Explore generic use cases and some real-world scenarios for ZooKeeper • Define the ZooKeeper services that are used to manage distributed systems • Explore and use the ZooKeeper CLI to interact with ZooKeeper services • Understand how Apache Slider works in conjunction with YARN to deploy distributed applications and to monitor them • Explain how Apache Knox provides peripheral security services to an Hadoop cluster

Lab 1. Explore ZooKeeper
Duration: 30 minutes

Overview	In this lab, you will connect to ZooKeeper and explore the ZooKeeper files.
Learning objectives	After completing this exercise, you should be able to: • Connect to ZooKeeper and explore the ZooKeeper files

Unit 9. Loading data with Sqoop
Duration: 30 minutes

Overview	In this unit you will learn how to load data with Sqoop.
Learning objectives	After completing this unit, you should be able to: • List some of the load scenarios that are applicable to Hadoop • Understand how to load data at rest • Understand how to load data in motion • Understand how to load data from common sources such as a data warehouse, relational database, web server, or database logs • Explain what Sqoop is and how it works • Describe how Sqoop can be used to import data from relational systems into Hadoop and export data from Hadoop into relational systems • Brief introduction to what Flume is and how it works

Lab 1. Moving data into HDFS with Sqoop
Duration: 30 minutes

Overview	In this lab, you will learn how to move data into an HDFS cluster from a relational database.
Learning objectives	After completing this lab, you should be able to: • Move data into HDFS with Sqoop

Unit 10. Security and Governance
Duration: 1 hour

Overview	In this unit you will learn about the need of data governance and the role of data security in it.
Learning objectives	After completing this unit, you should be able to: • Explain the need for data governance and the role of data security in this governance • List the Five Pillars of security and how they are implemented with HDP • Discuss the history of security with Hadoop • Identify the need for and the methods used to secure Personal & Sensitive Information • Describe the function of the Hortonworks DataPlane Service (DPS)

Unit 11. Stream Computing
Duration: 1 hour

Overview	In this unit you will learn about stream computing.
Learning objectives	After completing this unit, you should be able to: • Define streaming data • Describe IBM as a pioneer in streaming data - with System S IBM Streams • Explain streaming data - concepts & terminology • Compare and contrast batch data vs streaming data • List and explain streaming components & Streaming Data Engines (SDEs)

Course II – Introduction to Data Science

Duration: 1.75 hours

Course introduction
Duration: 5 minutes

Unit 1. Data Science and Data Science Notebooks
Duration: 45 minutes

Overview	In this unit, you will learn about data science and data science notebooks.
Learning objectives	After completing this unit, you should be able to: • Have a better understanding of methodology "scientific approach" methods used & skills practiced by Data Scientists • Recognize the iterative nature of a data science project • Outline the benefits of using Data Science Notebooks • Describe the mechanisms and tools used with Data Science Notebooks • Compare and contrast the major Notebooks used by Data Scientists

Unit 2. Data Science with Open Source Tools
Duration: 30 minutes

Overview	In this unit, we will concentrate on the Jupyter Notebook and Python
Learning objectives	After completing this unit, you should be able to: • Getting started with Jupyter Notebook • Data and notebooks in Jupyter • How notebooks help data scientists • Essential packages: NumPy, SciPy, Pandas, Scikit-learn, NLTK, BeautifulSoup, … • Data visualizations: matplotlib, …, PixieDust • Using Jupyter "Magic" commands

Lab 1. Introduction to Jupyter Notebooks
Duration: 30 minutes

Overview	In this lab you will be introduced to Jupyter Notebooks.

Learning objectives	After completing this exercise, you should be able to:
	• Start Jupyter - it will open in a web browser
	• Import the lab file (all Jupyter files have a .ipynb suffix) into your default workspace
	    o This is now a copy of the provided lab file and you can do anything with it
	    o If you mess it up, you can re-import again later
	• Explore the component panels - some are markdown, some are code, some are results of running the code (output data, visualizations, …)
	• Learn how to run single panels - and then the whole script
	    o You may need to adjust the provided script to locate the data files thataccompany the Jupyter.ipynb file
	    o Add some additional panels, as described in the lab script

Course III – Big SQL

Duration: 7.25 hours

Course introduction
Duration: 5 minutes

Unit 1. Using Big SQL to access data residing in the HDFS
Duration: 40 minutes

Overview	In this unit, you will learn about Big SQL, and how to use it to access data residing in the HDFS
Learning objectives	After completing this unit, you should be able to: • Overview of Big SQL • Understand how Big SQL fits in the Hadoop architecture • Start and stop Big SQL using Ambari and command line • Connect to Big SQL using command line • Connect to Big SQL using IBM Data Server Manager

Lab 1. Connecting to the IBM Big SQL Server
Duration: 30 minutes

Overview	In this lab you will connect to the Big SQL Server using multiple techniques. You will first explore the lab environment. You will then learn how to set up JSqsh and use it to connect to the Big SQL server. You will also explore the Big SQL service using the Data Server Manager (DSM) graphical web interface.
Learning objectives	After completing this exercise, you should be able to: • Configure images • Start Hadoop components • Start up the Big SQL and DSM services • Connect to Big SQL using JSqsh • Execute basic Big SQL statements • Explore Big SQL through Ambari using DSM

Unit 2. Creating Big SQL schemas and tables
Duration: 55 minutes

Overview	In this unit, you will learn how to create Big SQL schemas and tables
Learning objectives	After completing this unit, you should be able to: • Describe and create Big SQL schemas and tables • Describe and list the Big SQL data types • Work with various Big SQL DDLs • Load data into Big SQL tables using best practices

Lab 1. Creating and managing Big SQL schemas and tables
Duration: 35 minutes

Overview	In this lab you will start off by creating and dropping a simple Big SQL table. You then will create multiple Big SQL tables using a variety of data types and load the tables with data. You will also work with views, external tables, and other methods of creating Big SQL tables.
Learning objectives	After completing this exercise, you should be able to: • Create and drop simple Big SQL table • Create sample tables • Move data into HDFS • Load data into Big SQL tables • Create and work with views • Create external tables

Unit 3. File formats and querying Big SQL tables
Duration: 1 hour

Overview	In this unit, you will learn about file formats and querying Big SQL tables.
Learning objectives	After completing this unit, you should be able to: • Describe Big SQL supported file formats • Query Big SQL tables using various DMLs

Lab 1. Querying Big SQL tables
Duration: 30 minutes

Overview	In this lab you will experiment with some more advanced SQL queries. You will then explore Big SQL's ARRAY type. You will also create a user-defined function (UDF) and write queries that call the UDF. Finally, you will store data in an alternate file format (Parquet).
Learning objectives	After completing this exercise, you should be able to: • Connect to Big SQL • Query data with Big SQL • Work with the ARRAY type • Work with Big SQL functions • Store data in an alternate file format (Parquet)

Unit 4. Configuring Big SQL security
Duration: 1 hour

Overview	In this unit, you will learn about how to configure Big SQL security
Learning objectives	After completing this unit, you should be able to: • Configure authentication for Big SQL • Manage security with Apache Ranger • Enable SSL encryption • Configure authorization of Big SQL objects • Configure impersonation in Big SQL

Lab 1. Configuring Big SQL security
Duration: 30 minutes

Overview	In this lab you will work with Big SQL authorization techniques.
Learning objectives	After completing this lab, you should be able to: • Use column masking and row based access control to restrict access to your data

Lab 2. Configuring impersonation in Big SQL
Duration: 30 minutes

Overview	In this lab you will enable and configure impersonation with Big SQL
Learning objectives	After completing this lab, you should be able to: • Configure impersonation in Big SQL

Unit 5. Data federation with Big SQL
Duration: 45 minutes

Overview	In this unit, you will learn data federation with Big SQL
Learning objectives	After completing this unit, you should be able to: • Understand the concept of Big SQL federation • List the supported data sources • Set up and configure a federation server to use different data sources

Lab 1. Using Fluid Query with Big SQL
Duration: 15 minutes

Overview	In this lab you will configure Fluid Query with Big SQL
Learning objectives	After completing this lab, you should be able to: • Configure Fluid Query with Big SQL

Course introduction
Duration: 5 minutes

Unit 1. Introduction to IBM Watson Studio
Duration: 30 minutes

Overview	In this unit, you will learn about Watson Studio.
Learning objectives	After completing this unit, you should be able to: • What is Watson Studio? • Setting up a project • Working with collaborators • Managing data assets

---

Lab 1. Getting started with Watson Studio
Duration: 1 hour

Overview	In this lab, you will create and manage a project, add collaborators, and load a data set to the object store.
Learning objectives	After completing this lab, you should be able to: • Sign up for a Watson Studio account • Create a new project • Manage a project • Add collaborators • Load data • Manage the object storage

---

Unit 2. Analyzing data with Watson Studio
Duration: 30 minutes

Overview	In this unit, you will learn how to analyze data with Watson Studio.
Learning objectives	After completing this unit, you should be able to: • Overview of Jupyter notebooks • Creating notebooks • Coding and running notebooks • Sharing and publishing notebooks

Lab 1. Analyzing data with Watson Studio
Duration: 30 minutes

Overview	In this lab, you will run through a sample notebook in Watson Studio and use PixieDust for data visualization
Learning objectives	After completing this lab, you should be able to: • Create a notebook • Use notebooks • Work with external data

IBM Official Badges and Associated Job Roles

IBM Official Badges	Big Data Engineer 2018: Explorer | Mastery Award
Associated Job Roles	• Business Intelligence Analyst • Artificial Intelligence Analyst

For more information

To learn more about this career path and others, see ibm.biz/ibmskillsacademy

To learn more about validating your technical skills with IBM Open Badges, see www.youracclaim.com

To stay informed about the IBM Skills Academy, see the following sites:

Facebook: www.facebook.com/ibmskillsacademy

AI Analyst

(Classroom)

Career path description

The Artificial Intelligence Analyst career path prepares students to apply AI concepts to build real-life solutions. This career path introduces students to basic concepts of AI, machine learning algorithms, natural language processing, chatbots, and computer vision. Students apply the concepts they learn to practical examples by using IBM Watson services and tools on IBM Cloud.

ibm.com/training

General information

Delivery method

25% self-placed training

75% Instructor led training

Version

2020

Products

IBM Watson Discovery, IBM Watson Assistant, IBM Watson Visual Recognition, IBM Watson Tone Analyzer, IBM Watson Natural Language Understanding, IBM Watson Studio, IBM Watson Knowledge Studio, IBM Cloud

Audience

Undergraduate senior students from IT related academic programs such as computer science, software engineering, information systems and similar others.

Learning objectives

After completing this course, you should be able to:
- Explain what artificial intelligence (AI) is.
- Describe the field of AI and its subfields: Machine learning, natural language processing (NLP), and computer vision.

- List applications of AI in the industry and government.
- Describe machine learning.
- Describe different type of machine learning algorithms.
- Apply machine learning algorithms to specific problems.
- Explain deep learning.
- Explain convolutional neural networks and neural networks.
- Describe examples of unsupervised and supervised learning.
- Describe IBM Watson.
- Explain how IBM Watson technology is applied to solve real world problems.
- Explain the capabilities of each IBM Watson service.
- Describe IBM Watson Studio, its components, and key applications.
- Describe the CRISP-DM process model and explain where machine learning fits in the CRISP-DM process.
- Create machine learning models for different machine learning algorithms by using IBM Watson Studio.
- Explain domain adaptation.
- Describe the purpose of training the various IBM Watson services.
- Describe IBM Watson Knowledge Studio capabilities and use.
- Explain what NLP is.
- List tools and services for NLP.
- Identify NLP use cases.
- Explain main NLP concepts.
- Explain how to evaluate the quality of an NLP algorithm.
- Identify the IBM Watson services based on NLP technology.
- Use IBM Watson Discovery to build a cognitive query application.
- Describe chatbot applications and chatbots design guidelines.
- Explain core concepts and artifacts needed to build a chatbot application.
- Build chatbot applications with IBM Watson Assistant and Node-RED.
- Explain what computer vision is.
- Identify computer vision use cases.
- Explain how computer vision analyzes and processes images and describe commonly used computer vision techniques.
- Use the IBM Watson Visual Recognition service to classify an image, detect faces, and recognize text in an image.
- Create custom models with IBM Watson Visual Recognition.
- Train the IBM Watson Visual Recognition service with IBM Watson Studio.
- Integrate multiple IBM Watson services to build a comprehensive intelligent solution.

Prerequisites Skills

- Computer science fundamentals
- Basic knowledge of applied math, algorithms, and data modeling
- Basic knowledge of probability and statistics
- Basic knowledge of Node.js and cloud computing
- Access to IBM Cloud
- Exposure to the IBM Skills Academy Portal learning environment

Duration

36 hours

Skill level

Basic – Intermediate

Classroom (ILT) setup requirements	
Processor	2 GHz or higher
GB RAM	8 GB
GB free disk space	80 GB
Network requirements	Yes
Other requirements	IBM ID

Notes

The following unit and exercise durations are estimates, and might not reflect every class experience. If the course is customized or abbreviated, the duration of unchanged units will probably increase.

Course Agenda

MODULE I – AI OVERVIEW

Course I – AI Overview

Duration: 3 hours and 30 minutes

Course Overview
Duration: 5 minutes

Unit 1. Introduction to artificial intelligence

Unit 1. Introduction to artificial intelligence Duration: 3 hours	
Overview	This unit explains what artificial intelligence (AI) is, its history and evolution, AI types, integral components of AI systems, factors that influenced the evolution of AI, and applications of AI in the industry, government, and science.
Learning objectives	After completing this unit, you should be able to: • Explain what AI is. • Describe the types of AI. • List the factors that influenced the advancement of AI in recent years. • List the applications of AI in the industry, science, and government. • List the subfields that are the focus of AI research.

Unit 2. Business analytics

Unit 2. Business analytics Duration: 30 minutes	
Overview	This unit introduces business analytics and describes different approaches and types of business analytics.
Learning objectives	After completing this unit, you should be able to: • Explain what business analytics is. • Describe different approaches and types of business analytics. • Describe analytical solutions. • Explain the challenges of analytical solutions.

MODULE II – Prerequisites

Course I – IBM Watson overview

Duration: 6 hours and 30 minutes

Unit 1. Introduction to IBM Watson

Unit 1. Introduction to IBM Watson Duration: 1 hour	
Overview	This unit introduces IBM Watson and its history.
Learning objectives	After completing this unit, you should be able to: • Explain what IBM Watson is and how it works. • Explain how Watson technology is made available to developers and organizations.

Unit 2. IBM Watson applied to industry, business and science

Unit 2. IBM Watson applied to industry, business and science Duration: 1 hour and 30 minutes	
Overview	This unit provides several examples that demonstrate how IBM Watson is transforming industry, business and science.
Learning objectives	After completing this unit, you should be able to: • Provide examples of Watson AI technologies applied to several industries.

Unit 3. IBM Watson use cases

Unit 3. IBM Watson use cases Duration: 30 minutes	
Overview	This unit presents two use cases showing organizations that successfully implemented AI solutions, based on IBM Watson technology.
Learning objectives	After completing this unit, you should be able to: • Describe how IBM Watson technology is being applied to solve real world problems.

Unit 4. Evolution from DeepQA to IBM Watson services

Unit 4. Evolution from DeepQA to IBM Watson services Duration: 1 hour	
Overview	This unit describes the evolution of Watson technology from the original DeepQA architecture to the present.
Learning objectives	After completing this unit, you should be able to: • Explain what the DeepQA architecture was. • Explain why IBM decided to commercialize Watson. • Describe the evolution of Watson services from the original DeepQA architecture to the present. • Recognize the Watson services available today on the IBM Cloud.

Unit 5. IBM Watson services overview

Unit 5. IBM Watson services overview Duration: 2 hours	
Overview	This unit provides an overview of the Watson services available in IBM Cloud.
Learning objectives	After completing this unit, you should be able to: • List the Watson services. • Explain the capabilities of each Watson service.

Exercise 0. Setting up your hands-on environment

Exercise 1. Setting up your hands-on environment Duration: 30 min	
Overview	This exercise guides you through the setup of your workstation before you perform the exercises in this course.
Learning objectives	After completing this exercise, you should have: • An IBM Cloud Lite account. • cURL installed on your workstation. • Node.js installed on your workstation. • Git installed on your workstation. • A code/text editor installed on your workstation.

MODULE III – AI Analyst (Classroom)

Duration: 26 hours
Course introduction

Course introduction
Duration: 30 minutes

Unit 1. Introduction to machine learning

Unit 1. Introduction to machine learning
Duration: 1 hour and 30 min

Overview	This unit recaps the main topics in Module I, AI overview and provides a deeper view into complex subjects such as: • Machine learning • Machine learning algorithms • Neural networks • Deep learning
Learning objectives	After completing this unit, you should be able to: • Explain what machine learning is. • Describe machine learning types and approaches. • List different machine learning algorithms. • Explain what neural networks and deep learning are, and why they are important in today's AI field. • Explain how to evaluate your machine learning model.

Exercise 1. Applying machine learning algorithms

Exercise 1. Applying machine learning algorithms
Duration: 1 hour 30 min

Overview	In this exercise, you apply machine learning algorithms to solve real problems.
Learning objectives	After completing this exercise, you should be able to: • Determine the centroids of a data set with the K-means clustering algorithm. • Predict the class of an object with the Naïve Bayes classifier. • Apply the linear regression algorithm to solve supervised learning problems. • Construct a decision tree to predict outcomes.

Unit 2. Introduction to IBM Watson

Unit 2. Introduction to IBM Watson Duration: 1 hour	
Overview	This unit provides an overview of key IBM Watson services, their purpose, how they work, and helps you get started with Watson services on IBM Cloud.
Learning objectives	After completing this unit, you should be able to: • Explain what IBM Watson is. • List IBM Watson services offerings. • List IBM Cloud Watson services. • Explain the capabilities of each Watson service. • List the Watson services that can be trained. • List the Watson services that cannot be trained. • Create a Watson service instance on IBM Cloud.

Exercise 2. Exploring IBM Watson services

Exercise 2. Exploring IBM Watson services Duration: 1 hour 15 min	
Overview	This exercise introduces you to Watson REST APIs. You will use cURL commands to submit requests to and receive responses from several Watson services.
Learning objectives	After completing this exercise, you should be able to: • Create Watson service instances. • Copy credentials from a service instance. • Submit API calls with the appropriate parameters. • Analyze the response returned from the Watson service. • Use Watson API Reference documentation.

Unit 3. Introduction to IBM Watson Studio

Unit 3. Introduction to IBM Watson Studio
Duration: 30 minutes

Overview	This unit provides a high level overview of Watson Studio, its components, key applications and the value added by the IBM offering.
Learning objectives	After completing this unit, you should be able to: • Describe Watson Studio. • Identify industry use cases. • List Watson Studio offerings. • Create Watson Studio projects. • Describe Watson Studio and Spark. • Describe Watson Studio and Object Storage. • Explain Watson Studio high availability considerations. • Prepare and analyze data. • Use Jupyter notebooks.

Exercise 3. Getting started with IBM Watson Studio

Exercise 3. Getting started with IBM Watson Studio
Duration: 1 hour and 30 min

Overview	This exercise introduces you to the basic tasks that you have to perform when using Watson Studio.
Learning objectives	After completing this exercise, you should be able to: • Create a Watson Studio project. • Manage the project. • Assign collaborators. • Load a data set into the project's object store. • Manage Object Storage. • Analyze data by using Watson Studio. • Use PixieDust for data visualization.

Unit 4. Introduction to IBM Watson Machine Learning

Unit 4. Introduction to IBM Watson Machine Learning Duration: 30 minutes	

Overview	This unit describes the CRoss Industry Standard Process for Data Mining known as CRISP-DM and explains the process of preparing data for a machine learning algorithm. This unit provides an overview of the IBM Watson Machine Learning service available on IBM Cloud.
Learning objectives	After completing this unit, you should be able to: • Describe the CRISP-DM process model. • Explain where machine learning fits in the CRISP-DM process. • Describe data preparation before feeding into machine learning algorithms. • Describe Watson Machine Learning features and capabilities.

Exercise 4. Getting started with IBM Watson Machine Learning

Exercise 4. Getting started with IBM Watson Machine Learning Duration: 1 hour and 30 minutes	

Overview	This exercise introduces you to the basic tasks that you have to perform while building machine learning models for different algorithms using Watson Machine Learning and Watson Studio.
Learning objectives	After completing this exercise, you should be able to: • Create a machine learning model by using Watson Studio and Watson Machine Learning. • Use data sets to train the model. • Use different estimators to train the machine learning model representing different machine learning algorithms. • Deploy machine learning models. • Evaluate the deployed models. • Call the deployed models from your applications. • Test the model with your data.

Exercise 5. Exploring deep learning and neural network modeling with IBM Watson Studio

Exercise 5. Exploring Deep Learning and Neural Network Modeler with IBM Watson Studio Duration: 1 hour	

Overview	This exercise guides you through designing, building, and training a deep learning model to recognize handwritten digits. The optional exercise guides you through using the MNIST computer vision data set to train a TensorFlow model to recognize handwritten digits.
Learning objectives	After completing this exercise, you should be able to: • Build a neural network to recognize handwritten digits. • Create a neural network design flow by using the neural network modeler. • Train models with experiment builder. • Work with Watson Machine Learning experiments to train deep learning models (TensorFlow).

Unit 5. Introduction to natural language processing (NLP)

Unit 5. Introduction to natural language processing (NLP) Duration: 30 minutes	

Overview	This unit introduces NLP. It covers key applications of NLP, basics concepts and terminology, tools and services and NLP challenges.
Learning objectives	After completing this unit, you should be able to: • Explain what NLP is. • Identify NLP use cases. • Explain basic NLP concepts and terminology. • List the tools and services for NLP.

Unit 6. NLP concepts and components

Unit 6. NLP concepts and components
Duration: 30 minutes

Overview	This unit covers NLP components, the NLP pipeline, natural language understanding, natural language generation, information retrieval, and information extraction.
Learning objectives	After completing this unit, you should be able to: • Define the NLP categories. • Describe the NLP pipeline. • Explain the challenges in natural language understanding. • Explain the concepts of information retrieval and extraction. • Describe sentiment analysis.

Unit 7. NLP evaluation metrics

Unit 7. NLP evaluation metrics
Duration: 30 minutes

Overview	This unit explains how to evaluate the quality of your NLP algorithm.
Learning objectives	After completing this unit, you should be able to: • Define various metrics to measure the quality of your NLP algorithm. • Understand the difference between these metrics.

Unit 8. NLP and IBM Watson

Unit 8. NLP and IBM Watson
Duration: 30 minutes

Overview	This unit lists the Watson services and software that are based on NLP and explains the main capabilities of Watson Natural Language Classifier, Watson Natural Language Understanding, Watson Discovery.
Learning objectives	After completing this unit, you should be able to: • List the NLP Watson services • List the Watson services that perform information extraction. • Describe the capabilities of IBM Watson Natural Language Classifier. • Describe the capabilities of the IBM Watson Natural Language Understanding. • Describe the capabilities of IBM Watson Discovery.

Exercise 6. Ingest, Convert, Enrich and Query with IBM Watson Discovery Service

Exercise 6. Ingest, Convert, Enrich and Query with IBM Watson Discovery Service
Duration: 1 hour 30 minutes

Overview	This exercise takes you through the process of preparing a collection of documents and running queries to extract insights from the documents. In the optional exercise you will work with the Discovery API.
Learning objectives	After completing this exercise, you should be able to: • Create a Watson Discovery service instance. • Create a collection. • Add content to a collection. • Create a custom configuration. • Build queries. • Use the Discovery API.

Unit 9. Introduction to IBM Watson Knowledge Studio

Unit 9. Introduction to IBM Watson Knowledge Studio
Duration: 45 minutes

Overview	This unit introduces Watson Knowledge Studio, its capabilities, and features. This unit explains the end-to-end domain adaptation process.
Learning objectives	After completing this unit, you should be able to: • Describe IBM Watson Knowledge Studio. • List the Watson services that are trained by Knowledge Studio. • List the Knowledge Studio workspace resources. • Explain the process to build Knowledge Studio models that can be deployed and used with other Watson services.

Exercise 7. Creating a machine learning model with Watson Knowledge Studio

Exercise 7. Creating a machine learning model with Watson Knowledge Studio.
Duration: 1 hour and 15 minutes

Overview	This exercise takes you through the process of building a machine learning model with Knowledge Studio that you can deploy and use with Watson services. In the optional exercise, you will create a rule-based model that you can use to find text patterns in documents.
Learning objectives	After completing this exercise, you should be able to: • Create a workspace for Watson Knowledge Studio. • Configure the workspace resources. • Create document sets. • Pre-annotate documents. • Create tasks for human annotators. • Analyze inter-annotator agreement and adjudicate conflicts in annotated documents. • Create machine learning models.

Unit 10.Introduction to chatbots

Unit 10.Introduction to chatbots
Duration: 30 minutes

Overview	This unit provides a high level introduction to chatbots, chatbot applications and guidelines to consider when designing a chatbot.
Learning objectives	After completing this unit, you should be able to: • Explain what a chatbot is. • Describe common applications of chatbots. • Identify factors that drive the growing popularity of chatbots. • Recognize the guidelines to consider when designing a chatbot. • List examples of tools and services that you can use to create chatbots.

Unit 11. Introduction to IBM Watson Assistant

Unit 11. Introduction to IBM Watson Assistant
Duration: 1 hour

Overview	This unit covers the core concepts that you need to understand to build a chatbot with Watson Assistant.
Learning objectives	After completing this unit, you should be able to: • Explain assistants and skills. • Explain intents. • Explain entities. • Explain context variables. • Describe how the nodes in a dialog are triggered. • Describe how the dialog flow is processed. • Describe the features that can be used to enrich the chatbot.

Exercise 8. Getting started with IBM Watson Assistant

Exercise 8. Getting started with IBM Watson Assistant
Duration: 45 minutes

Overview	This exercise introduces IBM Watson Assistant and walks you through the process of creating a very simple chatbot with Watson Assistant.
Learning objectives	After completing this exercise, you should be able to: • Create a Watson Assistant service instance. • Create a Watson Assistant skill. • Add intents. • Build a dialog.

Exercise 9. Help Desk chatbot

Exercise 9. Help Desk chatbot
Duration: 1 hour 30 minutes

Overview	In this exercise you will create a chatbot application with Node-RED without coding and integrate it with the Watson Assistant service.
Learning objectives	After completing this exercise, you should be able to: • Create a Watson Assistant skill. • Add intents and entities. • Build a dialog. • Create a Node-RED application that integrates with the Watson Assistant service. • Set up Slack as a front-end chat service for the Help Desk chatbot.

Unit 12. Introduction to computer vision

Unit 12. Introduction to computer vision
Duration: 30 minutes

Overview	This unit provides a high level introduction to computer vision.
Learning objectives	After completing this unit, you should be able to: • Define computer vision. • Explain the history of computer vision and its advancement with AI. • Identify computer vision use cases. • List tools and services for computer vision.

Unit 13. Computer vision fundamentals

Unit 13. Computer vision fundamentals
Duration: 30 minutes

Overview	This unit explains the basic steps of a typical computer vision pipeline, how CV analyzes and processes images, and explores commonly used techniques in CV.
Learning objectives	After completing this unit, you should be able to: • Describe image representation for computers. • Describe the computer vision pipeline. • Describe different preprocessing techniques. • Explain image segmentation. • Explain feature extraction and selection. • Describe when object recognition takes place.

Unit 14. Introduction to IBM Watson Visual Recognition

Unit 14. Introduction to IBM Watson Visual Recognition Duration: 45 min	

Overview	This unit introduces the Watson Visual Recognition service, describes its capabilities and how to train the service.
Learning objectives	After completing this unit, you should be able to: • Describe the IBM Watson Visual Recognition service. • List the features available with Watson Visual Recognition. • Describe the output provided by the Watson Visual Recognition service. • Explain the capabilities of the default classifier. • Explain the difference between a default and a custom classifier. • Describe how to train a custom classifier.

Exercise 10. IBM Watson Visual Recognition model builder in IBM Watson Studio

Exercise 10. IBM Watson Visual Recognition model builder in IBM Watson Studio Duration: 1hour	

Overview	In this exercise, you learn how to create, train, and test a custom model. With a custom model, you can train the Watson Visual Recognition service to classify images to suit your business needs.
Learning objectives	After completing this exercise, you should be able to: • Create a Watson Visual Recognition service. • Create and train a custom Watson Visual Recognition model by using Watson Studio. • Edit and retrain the trained model. • Describe the effect of adding a negative class when training a Watson Visual Recognition model.

Unit 15. Designing and building an intelligent solution

	Unit 15. Designing and building an intelligent solution Duration: 45 minutes
Overview	This unit explains the benefits of integrating multiple Watson services to build a comprehensive intelligent solution. This unit presents two intelligent solutions use cases: Cognitive banking FAQ chatbot and Intelligent procurement system.
Learning objectives	After completing this unit, you should be able to: • Explain the need to integrate multiple IBM Watson services to build an intelligent solution. • Describe the general outline for the integration of IBM Watson Assistant with other services and applications. • Explain the key concepts that enable Watson Assistant integration. • Describe the integration flow between Watson Assistant, Watson Discovery, Watson Natural Language Understanding, and Watson Tone Analyzer to build the cognitive banking chatbot. • Describe the integration flow between Watson Knowledge Studio and Watson Discovery.

Exercise 11. Creating a cognitive banking FAQ chatbot

	Exercise 11. Creating a cognitive banking FAQ chatbot Duration: 1 hour and 30 minutes
Overview	This exercise introduces you to IBM Watson Node.js SDK to include conversation interactions, anger detection, natural language understanding, and answer discovery in your FAQ chatbot application.
Learning objectives	After completing this exercise, you should be able to: • Create a chatbot using Watson Assistant and Node.js. • Use Watson Discovery with passage retrieval to find answers in FAQ documents. • Use Watson Tone Analyzer to detect emotion in a conversation. • Identify entities in the user's input with Watson Natural Language Understanding.

Duration: 1 hour

Overview	In this exercise, you will create a Discovery collection with procurement documents initially enriched by the Discovery Default Configuration. Then, you will create a Knowledge Studio machine learning model trained for the procurement domain and deploy the model to Discovery. Finally, you will evaluate the results.
Learning objectives	After completing this exercise, you should be able to: • Create a machine learning model in Watson Knowledge Studio and deploy it to Watson Discovery. • Create a Watson Discovery custom configuration and leverage a Watson Knowledge Studio model to enrich entities and relations. • Integrate a custom model from Watson Knowledge Studio with the Discovery service to provide custom entity and relations enrichments customized for a specific procurement domain.