[](https://scitexas.edu/)**SYLLABUS**

DSO107-AWS: Big Data Analytics - AWS

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **COURSE DESCRIPTION** | The Big Data Analytics - AWS course introduces students to Big Data at a hands-on level and gives students exposure and practice with several skills and tools currently in use on Amazon Web Services (AWS). These skills start with learning cloud foundations, and then broadening methods will be used to help students analyze substantial amounts of data utilizing AWS. The course helps learners develop skills with AWS services that are critical for conducting analysis of big data problems. The course consists of a series of labs that supports data engineering, data analysis, and data science. | | | | | |
| **COURSE LENGTH /**  **COURSE HOURS** | 3 Weeks  75 contact hours (30 hours lecture, 30 hours laboratory) | | | | | |
| **CLASS SCHEDULE** | Start Date: | Meeting Days: | | | Meeting Times: | |
| End Date: |
| **INSTRUCTOR INFORMATION** | Name: | | | | | |
| Phone: | | | | | |
| Email: | | | | | |
| Office Hours (Days/Times): | | | | | |
| **TOTAL OUT-OF-CLASS HOURS** | 15 hours  Out-of-class assignments are those activities identified to be completed outside of the scheduled class meeting time. These assignments will be graded as part of the overall course grade. For specific  assignments and required/estimated time on task, see Weekly Content and Activities grid below. | | | | | |
| **COURSE QUARTER CREDITS** | 4.5 | | | | | |
| **PREREQUISITES** | DSO102, DSO104, DSO108, & DSO109 | | | | | |
| **COURSE LEARNING OBJECTIVES** | 1. Big Data Overview and Introduction to Amazon Web Services (AWS) 2. AWS Global Infrastructure and Cloud Security 3. AWS Networking Content Delivery and Compute 4. AWS Storage and Databases 5. AWS Cloud Architectures Autoscaling and Monitoring 6. AWS Big Data Analytics S3 Athena[¶](http://localhost:8889/lab/tree/Documents/GitHub/DS-Student-Resources/DS107-Big-Data/New-Course-Update/DS107L6-AWS-Big-Data-Analytics-S3-Athena.ipynb#DS107-Big-Data-:-AWS-Big-Data-Analytics-S3-Athena) 7. AWS Query Data S3 Athena Glue Analyze Data Redshift 8. AWS Analyze Data SageMaker Jupyter Bokeh Automate Data Loading Data Pipeline 9. AWS Analyze Streaming Data Kinesis Firehose OpenSearch Kibana Analyze IoT Data & Analytics 10. Data Analytics Final Exam | | | | | |
| **REQUIRED TEXTBOOKS AND**  **MATERIALS** | Students are required to have a laptop or desktop computer with internet access.  Minimum: PC (Windows) or MacBook laptop. 8GB ram, 256GB HD, Core i5/AMD Ryzen 5+/M1  Recommended: PC (Windows) or MacBook laptop. 8GB ram, 256GB SSD, Core i7/AMD Ryzen 7/M1+ It is requirement that you download programming resources to your laptop/desktop for this class. | | | | | |
| **REFERENCE MATERIALS** | In addition to textbooks, workbooks, lab manuals or other required materials, classes may utilize sources from the learning resource center, Internet sites, periodicals, newspapers, professional (or business) publications, state-specific laws or codes, magazines, personal interview, guest speakers, publisher provided information (via CD, DVD, or website), instructor work experience, video, audio or other visual files/documents to convey and aid in obtaining course objectives. | | | | | |
| **INSTRUCTIONAL METHODS** | Instructors will utilize a variety of delivery methods, teaching strategies, and instructional process to deliver course content. These may include, but are not limited to, lecture, demonstration, presentation, case studies, internet (on-line) activities, video, audio, group projects, guest speakers, field trips, interactive exercises, and/or simulated work experiences. Due to the intensive nature of the course,  student participation is necessary for the successful completion of this module. | | | | | |
| **GRADING** | Professionalism & Class Participation 5% Assignments, Hands-On & Homework 72% Projects, Competencies & Research 23%  **TOTAL 100%** | | | Students must achieve a grade of “C” or better in each course to progress through the program.  Students who receive a grade other than “A,” “B,” or “C” must repeat the course, consistent with SCI policies on course repetition | | |
| **GRADES** | **Grade**  A  B | | **Percentage**  90-100%  80-89% | | | **Grade Points**  4  3 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | C  F | 70-79%  Below 70% | 2  0 |
| **ATTENDANCE / MAKE-UP WORK** | See the SCI School College Catalog to review the Attendance and Make-up Work policies. | | |
| **PROFESSIONALISM & CLASS PARTICIPATION** | Students are expected to act professionally by preparing for class, meeting deadlines, cooperating with the instructor and classmates, and contributing in a positive way to the learning experience. Class preparation will require time spent on assignments outside of scheduled class time. Working in teams, communicating respectfully, and adhering to the Code of Conduct and Dress Code in the SCI School catalog are essential to success in your education and your field. Professionalism is a part of the grading scale for each class.  Observables:   1. Timeliness (is on time to class, returns to class on time after breaks, informs instructor about absences ahead of time, submits work prior to or on the due date) 2. Communicates with the instructor and peers in a respectful manner 3. Prepares for active and meaningful class participation 4. Adheres to the Student Code of Conduct (in the SCI Catalog) 5. Adheres to the Student Dress Code (in the SCI Catalog) | | |
| **ACADEMIC SUPPORT** | Tutoring is available upon request | | |
| **TECHNICAL SUPPORT** | [exeter-helpdesk-woz@woz-u.com](mailto:exeter-helpdesk-woz@woz-u.com) | | |

**WEEKLY CONTENT & ACTIVITIES**

|  |  |
| --- | --- |
| **Week One** | |
| Objectives: 1, 2, 3 | |
| **In-Class** | **Lesson 1 Big Data Overview and Introduction to AWS**  - Define different types of cloud computing  - Describe six advantages of cloud computing  - Recognize the main AWS service categories and core services  - Review the AWS Cloud Adoption Framework (AWS CAF)  - Explain the AWS pricing philosophy  - Recognize fundamental pricing characteristics  - Indicate the elements of total cost of ownership  - Discuss the results of the AWS Pricing Calculator  - Identify how to set up an organizational structure that simplifies billing and account visibility to review cost data.  - Identify the functionality in the AWS Billing Dashboard  -Describe how to use AWS Bills, AWS Cost Explorer, AWS Budgets, and AWS Cost and Usage Reports  - Identify the various AWS technical support plans and features  **Lesson 2 AWS Global Infrastructure and Cloud Security**  - Recognize the shared security model  - Identify the responsibility of the customer and AWS  - Recognize IAM users, groups and roles  - Describe different types of security credentials in IAM  - Identify the steps to securing a new AWS Account  - Explore IAM users and groups  - Recognize how to secure AWS data  - Recognize AWS compliance programs  - Identify the difference between AWS Regions, Availability Zones, and Edge locations  - Identify the various AWS Services and Service Categories  - Recognize the shared security model  - Identify the responsibility of the customer and AWS  - Recognize IAM users, groups, and roles  - Describe the different types of security credentials in IAM  - Identify the steps in securing a new AWS account  - Explore IAM users and groups  - Recognize how to secure AWS data  - Recognize AWS compliance programs  **Lesson 3 AWS Networking Content Delivery and Compute**  - Recognize the basics of networking  - Describe virtual networking in the cloud with Amazon VPC  - Label a network diagram  - Design a basic VPC architecture  - Indicate the steps to build a VPC  - Identify security groups  - Create your own VPC and add additional components to it to produce a customized network  - Identify the fundamentals of Amazon Route 53  - Recognize the benefits of Amazon CloudFront  - Provide an overview of different AWS compute services in the cloud  - Demonstrate why to use Amazon Elastic Compute Cloud (Amazon EC2)  - Identify the functionality in the EC2 console  - Perform basic functions in EC2 to build a virtual computing environment  - Identify EC2 cost optimization elements  - Demonstrate when to use AWS Elastic Beanstalk  - Demonstrate when to use AWS Lambda  - Identify how to run containerized applications in a cluster of managed servers |
| **Homework** | **Total Out-of-Class Hours –5 hours** |
| **Laboratories** | - CF Lab #1 Introduction to AWS IAM  - CF Lab #2 Build a VPC and Webserver  - CF Lab #3 Introduction to Amazon EC2 |

|  |  |
| --- | --- |
| **Week Two** | |
| Objectives: 4, 5, 6 | |
| **In-Class** | **Lesson 4 AWS Storage and Databases**  - Identify the different types of storage  - Explain Amazon S3  - Identify the functionality in Amazon S3  - Explain Amazon EBS  - Identify the functionality in Amazon EBS  - Perform functions in Amazon EBS to build an Amazon EC2 storage solution  - Explain Amazon EFS  - Identify the functionality in Amazon EFS  - Explain Amazon S3 Glacier  - Identify the functionality in Amazon S3 Glacier  - Differentiate between Amazon EBS, Amazon S3, Amazon EFS, and Amazon S3 Glacier  - Amazon Relational Database Service (Amazon RDS)  - Amazon DynamoDB  - Amazon Redshift  - Amazon Aurora  **Lesson 5 AWS Cloud Architectures Autoscaling and Monitoring**  - Describe the AWS Well-Architected Framework, including the six pillars  - Identify the design principles of the AWS Well-Architected - Framework  - Explain the importance of reliability and high availability  - Identify how AWS Trusted Advisor helps customers  - Interpret AWS Trusted Advisor recommendations  - Indicate how to distribute traffic across Amazon Elastic Compute Cloud (Amazon EC2) instances by using Elastic Load Balancing  - Identify how Amazon CloudWatch enables you to monitor AWS resources and applications in real time  - Explain how Amazon EC2 Auto Scaling launches and releases servers in response to workload changes  - Perform scaling and load balancing tasks to improve an architecture  **Lesson 6 AWS Big Data Analytics S3 Athena**  - Describe the role of the Amazon S3 service in big data solutions.  - Explain how AWS Identity and Access Management (IAM) controls access to Amazon S3.  - Create an S3 bucket.  - Load data into an S3 bucket.  - Query resources in Amazon S3.  - Access Amazon Athena in the AWS Management Console.  - Create an Amazon Athena database.  - Create a table in Amazon Athena.  - Optimize an Amazon Athena database.  - Run moderately complex queries with Amazon Athena. |

|  |  |
| --- | --- |
|  |  |
| **Homework** | **Total Out-of-Class Hours –5 hours** |
| **Laboratories** | - CF Lab #3 Working with EBS  - CF Lab #4 Build a Database Server  - CF Lab #5 Scale & Load Balance your Architecture  - DA Lab #1 Store data in Amazon S3  - DA Lab #2 Query Data in Amazon Athena |

|  |  |
| --- | --- |
| **Week Three** | |
| Objectives: 7,8,9,10 | |
| **In-Class** | **Lesson 7 AWS Query Data S3 Athena Glue Analyze Data Redshift**  - Access AWS Glue in the AWS Management Console.  - Create a crawler with AWS Glue.  - Create a database and table with AWS Glue.  - Query data in Amazon S3 with Amazon Athena and AWS Glue.  - Access Amazon Redshift in the AWS Management Console.  - Create an Amazon Redshift cluster.  - Load data from Amazon S3 into Amazon Redshift.  - Query data in Amazon Redshift.  **Lesson 8 AWS Analyze Data SageMaker Jupyter Bokeh Automate Data Loading Data Pipeline**  - Describe Jupyter notebooks and the Bokeh visualization package.  - Create a Jupyter notebook with Amazon SageMaker.  - Import data into a Jupyter notebook.  - Create a presentation with a Jupyter notebook.  - Visualize data with the open-source Bokeh Python package.  - Access AWS Data Pipeline in the AWS Management Console.  - Create a data pipeline.  - Load data from Amazon S3 into Amazon Redshift with a data pipeline.  - Troubleshoot a data pipeline.  - Export data from Amazon Redshift to a Jupyter notebook.  **Lesson 9 AWS Analyze Streaming Data Kinesis Firehose OpenSearch Kibana Analyze IoT Data & Analytics**  - Access Amazon Kinesis Data Firehose and Amazon Elasticsearch Service (Amazon ES) in the AWS Management Console.  - Create a Kinesis Data Firehose delivery stream.  - Integrate a Kinesis Data Firehose delivery stream with Amazon ES.  - Build visualizations with Kibana.  - Access the AWS IoT Analytics Service in the AWS Management Console  - Create an AWS IoT Analytics channel  - Create an AWS IoT Analytics data store  - Create an AWS IoT Analytics pipeline  - Create an AWS IoT Core rule  - Query an AWS IoT Analytics data store  **Lesson 10 Final Exam**  - Data Analytics Final Exam |
| **Homework** | **Total Out-of-Class Hours –5 hours** |
| **Laboratories** | - DA Lab #3 Query data in Amazon S3 with Amazon Athena and AWS Glue  - DA Lab #4 Analyze Data with Amazon Redshift  - DA Lab #5 Analyze Data with Amazon SageMaker, Jupyter Notebooks and Bokeh  - DA Lab #6 Automate Loading Data with the AWS Data Pipeline  - DA Lab #7 Analyze Streaming Data with Amazon Kinesis Firehose, Amazon Elasticsearch and Kibana  - DA Lab #8 Analyze IoT Data with AWS IoT Analytics |