



## Fundamentals of US Healthcare Data Analytics

Nature of Course: Theory + Practical

Total Hours per Day: 2 Hours

Course Duration: 2 Months (80 Hours)

### Overview of US Health Care

#### Part 1: Key Data Components

- Eligibility  
Objective: Provide overview of Eligibility file, processing and its analysis steps
  - Introduction to Eligibility
  - Analysis of Eligibility Data
  - Processing Raw File
- Medical Claims  
Objective: Provide overview of Claims file, processing and its analysis steps
  - Introduction to Claims
  - Analysis of Claims Data
  - Processing Raw File
- Pharmacy Claims  
Objective: Provide overview of Pharmacy file, processing and its analysis steps
  - Introduction to Pharmacy
  - Analysis of Pharmacy Data
  - Processing Raw File
- Provider  
Objective: Provide overview of Provider file, processing and its analysis steps
  - Introduction to Provider
  - Analysis of Provider Data
  - Processing Raw File

#### Part 2: MRF, Benefit and Cost Calculation Logic

- Introduction to MRF(Machine Readable Files)  
Objective: Provide overview of the role and significance of MRF in the US healthcare system
  - Introduction to MRF in HealthCare
  - Key Aspects of MRF
  - Challenges and Considerations

- Components of a MRF  
Objective: Provide overview of various components of MRF components and its significance
  - Table Of Contents
  - In - Network Components
  - Out of Network Components
- Plan Benefit  
Objective: Provide overview of plan's various components and its significance
  - Premium
  - Deductible and Out of Pocket
  - Coinsurance and Copayment
  - Provider Tier or Network Provider
  - Services
- Cost Calculation Logic  
Objective: Provide insight of how MRF and Plan Benefit work around for cost calculation

### Part 3: SQL

- Introduction to SQL  
Objective: Introduce Students to the basics of SQL
  - Introduction to SQL
  - Install and Setup Database
- Basic SQL Queries
  - Data Types, Primary-Foreign Keys & Constraints
  - DDL and DML Statements
- Create Table In SQL & Create Database
  - INSERT UPDATE, DELETE & ALTER Table
  - SELECT Statement & WHERE Clause
  - Aggregate Functions
  - Group By and Having clause
- JOINS in SQL
  - Different Types of JOINS
  - SELF JOIN, UNION & UNION ALL
  - Use cases of JOIN operation
- Managing Database Objects
  - Working with Indexes, Views, Synonyms, and Sequences
  - Partitioning and Materialized Views
  - Introduction of PL SQL, Stored Procedure, Functions, Trigger

## Part 4: SQL

- Working with Eligibility Data
  - Learn about eligibility data
  - Get meaningful insights about eligibility
- Working with Medical Data
  - Learn about medical data
  - Get meaningful insights about medical data
- Working with Pharmacy Data
  - Learn about pharmacy data
  - Get meaningful insights about pharmacy data
- MM, PMPM and Units per K concepts
  - Calculate member month
  - Calculate per member per month
  - Calculate units per k

## Part 5: PySpark

- Introduction to Apache Spark and PySpark  
Objective: Introduce students to the basics of Apache Spark and PySpark
  - Introduction to Big Data and Apache Spark
  - Introduction to PySpark
  - Setting Up a PySpark Environment
- Pyspark Spark Session and Spark Context  
Objective: Learn about Spark Session and its configurations
  - Introduction to Spark Session and Spark Context
  - Creating Spark Session
  - Load and Save Files using PySpark
- PySpark DataFrames  
Objective: Learn about PySpark DataFrames and their advantages over RDDs.
  - Introduction to DataFrames
  - DataFrame Actions and Usages
- Working with RDDs (Resilient Distributed Datasets)  
Objective: Understand the fundamentals building blocks of PySpark - RDD
  - Introduction to RDDs
  - Creating and Transforming RDDs
  - Actions on RDDs
- SQL with Pyspark  
Objective: Learn sql with PySpark
  - Creating Tables
  - SQL query in PySpark

## Part 6: Working with PySpark

- Working with Eligibility Data
- Working with Medical Data
- Working with Pharmacy Data
- Working with MRF Data

## Part 7: Data Quality

- Introduction to Data Quality
  - Importance of Data Quality in Healthcare Analytics
  - Understanding Data Quality Metrics and Parameters
  - Data Correctness
- Definition and Significance
  - Methods for Assessing Data Correctness
  - Accuracy of Data Entry
  - Data Validation Techniques
  - Error Identification and Rectification Processes
- Data Completeness
  - Definition and Importance in Healthcare Data
  - Measures to Assess Data Completeness
  - Missing Data Handling Techniques
  - Estimation and Imputation Methods
  - Evaluating the Impact of Incomplete Data on Analysis
- Data Quality Control
  - Overview of Quality Control in Data Management
  - Establishing Data Quality Standards and Procedures
  - Implementation of Quality Control Processes
  - Quality Control Reporting Process

## Part 8: AWS

- Introduction to AWS
  - Objective: Provide an overview of AWS services and tools relevant to data analysis
    - Overview of AWS
    - Introduction to Cloud computing
- Data Storage in AWS
  - Objective: Explore AWS services for Data Storage
    - Amazon S3 (Simple Storage Service) (With Practical)
    - Amazon RDS (Relational Database Service)
- Data Processing with AWS
  - Objective: Introduce Amazon services for data processing and analysis

- AWS Athena for Querying S3 Data (With practical)
- AWS Glue Job for ETL (With Practical)
- Additional AWS Services  
Objective: Overview of different aws provided services used in data analysis and processing

## **Labs**

Lab assignments will focus on the practice and mastery of contents covered in the lectures and introduce critical and fundamental problem-solving techniques to the students.