



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

سرفصل های
دوره آموزشی

دوره جامع علم داده





جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

SQL Server

پیش زمینه در مورد T-SQL Querying

▪ آشنایی با مفاهیم اولیه T-SQL:

- SQL
- Set theory
- Predicate logic
- Relational model

▪ معماری SQL Server

▪ اجرای دستور Select:

- عبارت From
- عبارت Where
- عبارت Group by
- عبارت Having
- عبارت Select
- عبارت Order by
- فیلترهای Top و Offset-Fetch
- آشنایی اولیه با Window Function ها



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

▪ شرط ها و عملگرها

▪ عبارت **Case**

▪ **Three-Valued Logic**

▪ کار با داده های کاراکتری:

- **Data Type**

- **Collation**

- عملگرها و توابع

- عملگر شرطی **Like**

▪ کار با داده های تاریخ و زمان:

- **Data Type** ها

- ثابت ها

- کار به صورت جداگانه با تاریخ و زمان

- محدود کردن بازه های تاریخی

- توابع تاریخ و زمان

▪ آشنایی با **Metadata**:

- **Catalog views**

- آشنایی با **Information Schema view**

- آشنایی با توابع و روال های سیستمی



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

Join ها

- **Cross join**
- **Inner joins**
- مثال های مختلفی از **join** ها
- **Outer Join**

Subquery ها

- **Subquery** های غیروابسته:
 - **Scalar Subqueries**
 - **Multivalued Subqueries**
- **Subquery** های وابسته:
 - عبارت شرطی **Exists**
- ساختار **Subquery** ها:
 - برگرداندن مقدار بعدی و قبلی
 - محاسبه مان



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

عبارت Table

▪ **Derived Table** :

- جداول تودرتو

- ارجاعات چندگانه

▪ **CTE** یا **Common Table Expression** :

- **Recursive CTE**

- تعریف چند **CTE**

▪ **Apply Operator**

کار با مجموعه ها

▪ عملگر **Union** :

- ترکیب با **Union All**

▪ عملگر **Intersect**

▪ عملگر **Except**



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

فراتر از اصول اولیه Query نویسی

▪ **Window Function** ها:

- **Ranking** توابع

- **Offset** توابع

- **Aggregate** توابع

▪ **Pivot Data**:

- **Pivoit** با **SQL Standard**

- **Pivoit** با دستور **Pivoit**

▪ **Unpivoting Data**

▪ **Grouping Set**:

- **Grouping set**

- **Cube**

- **Rollup**

- **Grouping**

تغییرات داده ها:

▪ **Insert Data**:

- **Insert Values**

- **Insert Select**

- **Insert Exec**

- **Select into**

- **Bulk insert**

- **Identity property**



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

▪ **Delete Data**

- دستور *Delete*

- دستور *Truncate*

▪ **Update Data**

- دستور *Updates*

▪ **Merge Data**

Tableau

▪ آشنایی با نرم افزار *Getting Started*

▪ آماده سازی داده ها *Connecting to Data*

▪ بصری سازی و ویژوالیزشن *Visual Analytics*

▪ ساخت داشبوردهای مدیریتی *Dashboards and Stories*



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

Microsoft Power BI

- معرفی سرویس ها و ابزارهای *Power BI* و کاربرد آنها
- تشریح مراحل چهارگانه سیکل گزارش های تحلیلی و داشبوردها با ابزار *Power BI*
- مروری بر دریافت و اصلاح داده ها از منابع مختلف داده
- مروری بر مدل سازی داده ها و معرفی توابع *DAX* برای ساخت محاسبه های کاربردی
- مروری بر مصورسازی (تحلیل) داده ها و معرفی ویژوال های کاربردی
- معرفی *Power BI Service* برای به اشتراک گذاری داشبوردها
- اشتراک گذاری گزارش ها به کمک *Power BI Report Server*
- اشتراک گذاری داشبوردها و گزارش ها روی موبایل و تبلت



جهد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

R For Data Engineering

Installation:

- **Data Types**
- **Control Flow in R**
- **Vectorized Thinking**
- **introduction to packages**
- **Introduction to Date & Time manipulation**
- **Introduction to String Manipulation**
- **Importing Data**
 - **Importing Data from Flat Files**
 - **Importing Excel Files**
- **Data Manipulation**
 - **Data Transformation**



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

Python For Data Engineering

Installation:

- *Python & Anaconda installation*
- *Basic Data in python*
- *Variables in python*
- *package Management*
- *introduction to numpy*
- *introduction to pandas*
- *introduction to control flows*
- *data importing in python*
 - *importing data from text files*
 - *importing data from csv*
 - *importing data from excel*
- *data manipulation*
 - *datetime manipulation fundamentals*
 - *string manipulation fundamentals*
 - *table manipulation fundamentals*
- *introduction to matplotlib*



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

[*BIG DATA*]

▪ مقدمه ای بر کلان داده ها و سیستم های ذخیره سازی توزیع شده

- بررسی انواع داده ها و منابع تولید کننده آنها در سیستم ها و محیط های عملیاتی
- بررسی انواع داده ها و منابع تولید کننده آنها در سیستم ها و محیط های عملیاتی
- سیستم های مدیریت داده توزیع شده و پایگاه داده های توزیع شده
- ویژگی های پایگاه داده های توزیع شده
- تعریف، توصیف و بررسی و ویژگی های کلان داده ها (*Big Data*)
- جایگاه کلان داده ها و پردازش کلان داده ها
- کاربرد کلان داده ها و تحلیل آنها
- منابع تولید کننده داده های کلان

▪ معرفی و شروع کار با هادوپ و اکوسیستم هادوپ

- معرفی *Hadoop* به عنوان سکوی پردازش و ذخیره سازی داده در ابعاد کلان
- اجزاء هادوپ
- قابلیت ها و توانمندی های *Hadoop*
- نقاط قوت و ضعف هادوپ در مقایسه با سیستم های پردازشی موجود
- معرفی مدل پردازشی نگاشت کاهش
- مراحل انجام کار در اجرای وظایف نگاشت کاهش



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

- مسائل قابل حل در مدل پردازشی نگاشت کاهش و مسائل سازگار با این مدل
- اجرای وظایف نگاشت کاهش در هادوپ و بررسی مراحل اجرا
- معرفی سیستم فایل توزیع شده در **Hadoop**
- ویژگی های **HDFS** و نحوه عملکرد **HDFS**
- معرفی **YARN** و وظیفه **YARN**
- جزاء **YARN**
- منابع قابل مدیریت و نحوه مدیریت منابع و وظایف توسط **YARN**

- معرفی آپاچی پیگ **Apache Pig**
 - معرفی پیگ
 - مزایا و معایب **Pig**
 - توسعه نمونه برنامه ها و استفاده از **Pig** جهت تعامل با داده های کلن و انجام پردازش دسته ای

- مقدمه ای بر آپاچی اسپارک **Spark Apache**
 - معرفی اسپارک
 - مدل پردازش توزیع شده در اسپارک
 - اجزاء اسپارک
 - ویژگی ها و کاربردهای اسپارک
 - منابع پردازشی، مدیریت منابع و انواع روش های مدیریت منابع در اسپارک



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

- مدل اجرایی وظایف در اسپارک
- *(python)* و پایتون *(scala)* ارتباط با اسپارک توسط اسکالا
- مدل داده ای *RDD*
- انواع منابع داده ای و تبدیل آنها به مدل داده ای اسپارک
- نحوه حل مسئله و توسعه نمونه برنامه های اسپارک برای انجام فرآیند های پردازش دسته
- حل مسئله و توسعه نمونه برنامه های اسپارک برای تحلیل برخط کلان داده

Installation:

- **Data Types**
- **Control Flow in R**
- **Vectorized Thinking**
- **introduction to packages**
- **Introduction to Date & Time manipulation**
- **Introduction to String Manipulation**
- **Importing Data**
 - Importing Data from Flat Files
 - Importing Excel Files
- **Data Manipulation**
 - Data Transformation



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

تحلیل اکتشافی

- تحلیل داده اکتشافی چیست؟
- اهمیت و نقش تحلیل داده اکتشافی
- مراحل کار
- توصیف داده
- شیوه های تحلیل تک متغیره و چند متغیره
- شیوه های تحلیل بصری و غیربصری
- شاخص های پر کاربرد در تحلیل *Numerical*
- اهمیت تحلیل بصری
- قواعد تحلیل بصری
- مشکلات تحلیل داده های چند بعدی



جهاد دانشگاهی صنعتی شریف
مرکز آموزش های تخصصی کاربردی

Python For Data Scientist

Installation:

- *Python & Anaconda installation*
- *Basic Data in python*
- *Variables in python*
- *package Management*
- *introduction to numpy*
- *introduction to pandas*
- *introduction to control flows*
- *data importing in python*
 - *importing data from text files*
 - *importing data from csv*
 - *importing data from excel*
- *data manipulation*
 - *datetime manipulation fundamentals*
 - *string manipulation fundamentals*
 - *table manipulation fundamentals*
- *introduction to matplotlib*

R For Data Scientist

www.sctae.net

۰۲۱-۶۷۶۴۱۹۹۹