



دورة تدريبية في هندسة البيانات



دورة تدريبية في هندسة البيانات

مقدمة:

تعد هندسة البيانات أمراً بالغ الأهمية في تحويل البيانات الخام إلى رؤى مفيدة تدعم قرارات الأعمال، والابتكارات، والتطورات التكنولوجية. توفر هذه الدورة التدريبية في هندسة البيانات، أساساً لتصميم وبناء وإدارة هياكل البيانات التي تسهل معالجة البيانات وتخزينها واسترجاعها بشكل فعال. سيستكشف المشاركون المفاهيم الرئيسية والنماذج والعمليات، بما في ذلك تخزين البيانات، خطوط أنابيب ETL للاستخراج، التحويل، التحميل، إدارة بحيرات البيانات، والحلول المبتكرة على السحابة.

تركز دورة هندسة البيانات، على التطبيق العملي، وتعد المتعلمين للتعامل مع تحديات هندسة البيانات المعقدة، وتبسيط سير العمل في البيانات، ودعم بيئات البيانات القابلة للتوسع وعالية الأداء. كما توفر الدورة فهماً شاملاً لهندسة البيانات، بما في ذلك المبادئ الأساسية، والتعريفات، والمفاهيم الأساسية. سيكتسب المشاركون الخبرة في مهارات هندسة البيانات، مثل إتقان SQL لهندسة البيانات، واستكشاف الأدوات الحديثة في هندسة البيانات، وتنفيذ حلول فعالة في هندسة البيانات.

تتناول هذه الدورة التدريبية في هندسة البيانات، على مهارات هندسة البيانات، وأفضل الممارسات، والتقنيات المتقدمة في هندسة البيانات السحابية لتلبية احتياجات الأعمال الحديثة. الدورة مثالية للمحترفين الطموحين، وتتضمن إرشادات حول كيفية الحصول على الشهادة، مما يجهز الحضور بالمعرفة اللازمة للتفوق كمدربين لهندسة البيانات. هذه الدورة تعد الأفراد لمواجهة التحديات الواقعية وتقديم حلول مبتكرة.

الفئات المستهدفة:

- المهندسون الطموحون في مجال البيانات.
- محللو البيانات الذين يسعون لتطوير مهارات متقدمة.
- مسؤولو تكنولوجيا المعلومات وقواعد البيانات.
- مطورو البرمجيات الذين ينوون الانتقال إلى أدوار بيانات.
- محترفو ذكاء الأعمال BI.
- مهندسو البيانات والمعماريون والحلول.
- مدراء المشاريع التقنية في المجالات المبتكرة على البيانات.
- الخريجون الجدد في علوم الكمبيوتر أو علوم البيانات.

أهداف الدورة التدريبية:

بنهاية هذه الدورة التدريبية في هندسة البيانات، سيكون المشاركون قادرين على:

- فهم المفاهيم الأساسية في هندسة البيانات.
- تطوير تدفقات ETL فعالة لمعالجة البيانات.
- تصميم وتنفيذ هياكل بيانات وخطوط أنابيب مثينة.
- اكتساب الكفاءة في حلول تخزين البيانات، بما في ذلك بحيرات البيانات ومستودعات البيانات.
- أتمتة سير العمل في البيانات لتحسين كفاءة المعالجة.
- إتقان SQL لمعالجة البيانات وإدارة قواعد البيانات.
- الاستفادة من منصات السحابة لهندسة البيانات القابلة للتوسع.
- ضمان جودة البيانات، والحكم عليها، وتأمينها عبر الأنظمة.
- تطبيق أدوات البيانات الضخمة مثل Hadoop و Spark في سيناريوهات واقعية.

الكفاءات المستهدفة:

بنهاية برنامج هندسة البيانات، سيكون المشاركون قادرين على:

- تصميم هندسة البيانات.
- تطوير عملية ETL الاستخراج، التحويل، التحميل.
- تحسين التخزين و مستودعات البيانات.
- أتمتة وتنظيم خطوط أنابيب البيانات.
- إتقان SQL وإدارة قواعد البيانات.
- هندسة البيانات السحابية AWS، Azure، GCP.
- ممارسات جودة البيانات والحكم عليها.
- معالجة البيانات الضخمة Hadoop، Spark.
- تأمين البيانات والامتثال للقوانين.

محتوى الدورة:

الوحدة 1: مقدمة في مبادئ هندسة البيانات:

- استكشاف دور وأهمية هندسة البيانات في الأعمال الحديثة.
- فهم الفرق بين هندسة البيانات وعلوم البيانات وتحليل البيانات.
- تعلم المفاهيم الأساسية مثل النمذجة، وخطوط الأنابيب، والهندسة.
- فحص دورة حياة البيانات من الحصول على البيانات الخام إلى توليد الرؤى.
- مناقشة التطبيقات الصناعية ودراسات الحالة في هندسة البيانات.

الوحدة 2: بناء خطوط أنابيب ETL:

- تعلم أساسيات عمليات الاستخراج والتحويل والتحميل ETL.
- تصميم وتنفيذ خطوط أنابيب ETL لتنظيف البيانات وتحويلها.
- أتمتة خطوط أنابيب البيانات لتبسيط سير العمل.
- معالجة تحديات تكامل البيانات من مصادر متعددة.
- استكشاف الأدوات والإطارات المستخدمة في عمليات ETL، مثل NiFi و Talend Apache.

الوحدة 3: تصميم مستودعات البيانات وبحيرات البيانات:

- فهم المفاهيم والفروقات بين مستودعات البيانات وبحيرات البيانات.
- تصميم أنظمة التخزين التي تحسن استرجاع البيانات ومعالجتها.
- استكشاف تصميم المخططات، والفهرسة، والتقسيم للاستعلام فعال.
- تنفيذ تخزين البيانات الموزعة على السحابة باستخدام Amazon Redshift و Google BigQuery و Snowflake.
- مناقشة أفضل الممارسات في تخزين البيانات من حيث القابلية للتوسع والمرونة.

الوحدة 4: معالجة البيانات الضخمة والأنظمة الموزعة:

- مقدمة في مفاهيم البيانات الضخمة وأطر المعالجة.
- العمل مع أدوات الحوسبة الموزعة مثل Hadoop و Spark Apache.
- تنفيذ خطوط أنابيب بيانات قابلة للتوسع لمعالجة مجموعات البيانات الكبيرة.
- تحسين معالجة البيانات باستخدام التوازي والحوسبة داخل الذاكرة.
- فهم تقسيم البيانات، والتجزئة، وإدارة الكتل.

الوحدة 5: حوكمة البيانات والجودة والأمان:

- تعلم مبادئ حوكمة البيانات وإدارة البيانات.
- تنفيذ إطار العمل لضمان جودة البيانات للحصول على مخرجات موثوقة.
- تطبيق أفضل ممارسات الأمان على البيانات، بما في ذلك التشفير وضوابط الوصول.
- استكشاف متطلبات الامتثال GDPR، HIPAA في التعامل مع البيانات.
- إنشاء خط بيانات وممارسات تحقيق لتتبع تدفق البيانات والتغييرات.