

E20-007

データ・サイエンスとビッグデータ分析試験

試験の概要



所要時間

90分
60問

受験可能な言語

英語
日本語

模擬試験

アソシエイト - E20-007試験

[E20-007\(英語\)](#)

[E20-007\(日本語\)](#)

EMCトレーニングセンタージャパン

〒212-0013

神奈川県川崎市幸区堀川町580-1

6 川崎テックセンター20階

EMCトレーニングセンタージャパン

<http://japan.emc.com/training/cus>

[tomers-training-offerings/index.htm](http://japan.emc.com/training/cus-tomer-training-offerings/index.htm)

概要

この試験は、データ分析の実践、データ・サイエンティストの役割、データ分析ライフサイクルの主な段階、Rを使用したデータの分析と調査、モデル構築と評価のための統計、理論と方法を使用した高度な分析、分析プロジェクトの運用可能化、データのビジュアル化技術に重点をおいています。試験に合格した受験者はEMC Proven Professionalデータ・サイエンス・アソシエイト技術者認定が与えられます。

EMCでは、試験に備えて知識を評価するための模擬試験を無償で提供しています。模擬試験により、実際の認定試験問題の内容や質問のタイプに慣れることができます。模擬試験の結果は認定試験のこれまでの準備の目安の1つとなり、学習が必要な内容が明らかになります。模擬試験での合格点は、認定試験での合格点を保証するものではありません。

試験問題の内容

この試験の出題範囲となる内容は次のとおりです。

- **ビッグデータ分析と、データ・サイエンティストの役割**
 - ビッグデータの特徴
 - 分析での実践
 - データ・サイエンティストの役割および必要なスキル
- **データ分析のライフサイクル**
 - 検出
 - データの準備
 - モデル・プランニングと構築
 - 結果の報告
 - データ分析プロジェクトの運用可能化
- **データの初期分析**
 - 基本Rコマンドを使用したデータ分析
 - 統計指標とデータを理解するためのビジュアル化
 - 理論、方法そして、分析結果のモデル評価
- **ビッグデータのための高度な分析 - 理論と方法**
 - K平均法クラスタリング
 - 相関ルール
 - 線形回帰
 - 単純ベイズ分類器
 - 決定木
 - 時系列分析
 - テキスト分析
- **データサイエンスのための高度な分析 - テクノロジーとツール**

- MapReduce
 - Hadoopエコシステム
 - SQL OLAP拡張、Windows機能、ユーザー定義機能そして、集計
- **分析プロジェクトの運用可能化とデータのビジュアル化技術**
 - 分析プロジェクトの運用可能化
 - プランニングと有効データ・ビジュアル化の作成

推奨トレーニング

以下は、この試験の受験者に推奨されるカリキュラムです。

| コース | モード | 状況 |
|--------------------|------|---------|
| データ・サイエンスとビッグデータ分析 | VILT | 2012年3月 |

注:これらの詳細は、2012年3月30日時点の試験の内容を反映しています。

認定プロフェッショナル・プログラムでは、技術的な傾向の主流と重要性が反映されるよう定期的に試験を更新しています。最新の情報については、Proven ProfessionalのWebサイトを定期的に確認してください。

EMC²、EMC、およびwhere information livesは、EMC Corporationの登録商標です。他のすべての名称ならびに製品についての商標は、それぞれの所有者の商標または登録商標です。© Copyright 2010 EMC Corporation. All Rights Reserved. 1/11