

# ÍNDICE

<b>Introducción .....</b>	<b>XIII</b>
<b>Capítulo 1. Conceptos de Big Data .....</b>	<b>1</b>
Definición, necesidad y características de Big Data .....	1
Aplicaciones típicas de Big Data .....	4
Patrones de detección del fraude.....	4
Patrones de Social Media .....	6
Patrones de modelado y gestión de riesgo .....	7
Big Data y el sector de la energía.....	8
Big Data en el Call Center .....	9
<b>Capítulo 2. Componentes de una plataforma de Big Data .....</b>	<b>11</b>
Plataforma de código abierto Hadoop .....	11
Hadoop Distributed File System (HDFS) .....	12
Hadoop MapReduce .....	13
Hadoop Common .....	14
Aplicaciones de desarrollo en Hadoop.....	14
Avro .....	15
Cassandra .....	15
Chukwa .....	15
Flume.....	15
Hbase (NoSQL).....	16
Hive.....	16
Jaql .....	17
Lucene .....	17
Oozie.....	17
Pig.....	18
ZooKeeper .....	18
Hadoop streaming.....	19
Situar datos en Hadoop.....	19

Copia de datos básica.....	19
Big Data y el campo de la investigación .....	20
<b>Capítulo 3. Big Data con herramientas de IBM .....</b>	<b>23</b>
IBM Power Systems.....	23
Hardware .....	24
Sistemas operativos.....	25
Software System.....	26
Herramientas.....	26
Analytics con Power System.....	27
IBM Solution for Hadoop Power Systems Edition.....	27
IBM Solution for Analytics Power Systems Edition .....	27
IBM BLU Acceleration Solution Power Systems Edition.....	29
IBM AIX Solution Editions para Cognos y SPSS .....	29
IBM PureData System for Operational Analytics (US) .....	31
Big Data Solution with InfoSphere BigInsights and Streams.....	32
IBM i para Business Intelligence (US).....	33
IBM DB2 Web Query for i .....	33
Cloud Computing en Power Systems .....	34
Virtualization Foundation Solutions .....	36
IBM SmartCloud Entry for Power Systems .....	37
IBM Power Systems Solution Edition for Cloud .....	37
IBM Power Systems Solution Edition for Scale Out Cloud .....	37
Herramientas Avanzadas en la nube .....	38
IBM SPSS Modeler.....	38
Usando el ratón .....	42
Ayuda en Modeler .....	42
El menú Herramientas de Modeler .....	43
Ejemplo de trabajo con IBM SPSS Modeler.....	48
Insertar un nodo fuente (origen) de datos en el área de trabajo .....	48
Enlazar un nodo con una fuente de datos .....	50
Controlar la carga de datos con un nodo Tabla .....	51
Definir variables predictoras con el nodo Tipo .....	53
Utilizar un nodo de modelado .....	56
Ejecutar de una ruta .....	57
Predecir con un modelo .....	59
Guardar un modelo .....	59
Nodos de orígenes de datos.....	60
Nodos de operaciones con registros .....	61

Nodos de operaciones con campos .....	62
Nodos para gráficos .....	63
Nodos para modelado.....	64
Nodos de resultado.....	66
Nodos de exportación .....	67
IBM SPSS Modeler e IBM SPSS Statistics .....	67
<b>Capítulo 4. Big Data con herramientas de Oracle.....</b>	<b>69</b>
Oracle y el Big Data .....	69
Oracle Big Data Appliance.....	71
Oracle Big Data Connectors.....	73
Oracle NoSQL Database.....	76
Oracle Exadata Database .....	78
Oracle Exalytics In-Memory Machine.....	79
Oracle Business Analytics.....	80
Oracle Business Intelligence Foundation Suite.....	80
Enterprise Performance Management .....	84
Aplicaciones analíticas.....	84
Information Discovery .....	87
Advanced Analytics.....	87
Nube .....	87
Soluciones de datos rápidos de Oracle .....	88
Oracle Social Cloud.....	89
<b>Capítulo 5. Big Data con herramientas de Microsoft.....</b>	<b>91</b>
Microsoft y el Big Data .....	91
Solución Big Data de Microsoft .....	92
Acceso a Hadoop.....	93
Adaptación de Hadoop para la empresa.....	94
Aprovechamiento de información .....	95
El papel de SQL Server .....	96
Los orígenes de Hadoop. La nube .....	97
HDInsight.....	98
Escalamiento con total flexibilidad a petición .....	99
Análisis de datos semiestructurados, estructurados y no estructurados .....	99
Desarrollo en el lenguaje favorito. Hardware.....	100
Excel para visualizar datos de Hadoop .....	100
Los clusters locales de Hadoop y la nube .....	100

HDInsight y HBase .....	101
Conceptos esenciales en Azure HDInsight .....	102
Datos de gran tamaño .....	102
Apache Hadoop .....	102
MapReduce .....	102
HDInsight .....	103
El ecosistema Hadoop en Azure .....	104
Pig .....	105
Hive .....	105
Sqoop .....	106
Herramientas de Business Intelligence y conectores .....	106
Escenarios de datos de gran tamaño en HDInsight .....	106
Introducción al uso de HDInsight de Azure .....	107
Requisitos previos .....	107
Configuración de un entorno local para ejecutar PowerShell .....	108
Aprovisionamiento de un cluster de HDInsight .....	108
Ejecución de un trabajo WordCount de MapReduce .....	111
Conexión a las herramientas de inteligencia empresarial de Microsoft .....	118
Uso de MapReduce con HDInsight .....	121
Escenario .....	121
Ejecución de la muestra con Azure PowerShell .....	122
El código Java para el programa de recuento de palabras de MapReduce ..	126
Carga de datos en HDInsight .....	129
Carga de datos en el almacenamiento de blobs usando AzCopy .....	130
Carga de datos en el almacenamiento de blobs usando Azure PowerShell .	131
Carga de datos en el almacenamiento de blobs usando el explorador de almacenamiento de Azure .....	132
Carga de datos en el almacenamiento de blobs usando la línea de comandos de Hadoop .....	134
Importación de datos a HDFS desde base de datos SQL o SQL Server usando Sqoop .....	136
Administración de HDInsight con PowerShell .....	138
Aprovisionamiento de un cluster de HDInsight .....	138
Enumeración y visualización de clusters .....	141
Eliminación de un cluster .....	141
Concesión/Revocación del acceso a los servicios de HTTP .....	142
Envío de trabajos de MapReduce .....	142
Envío de trabajos de Hive .....	144
Introducción al emulador de HDInsight .....	145
Instalación del emulador de HDInsight .....	145

HDInsight y HBase .....	101
Conceptos esenciales en Azure HDInsight .....	102
Datos de gran tamaño .....	102
Apache Hadoop .....	102
MapReduce .....	102
HDInsight .....	103
El ecosistema Hadoop en Azure .....	104
Pig.....	105
Hive.....	105
Sqoop.....	106
Herramientas de Business Intelligence y conectores .....	106
Escenarios de datos de gran tamaño en HDInsight.....	106
Introducción al uso de HDInsight de Azure .....	107
Requisitos previos.....	107
Configuración de un entorno local para ejecutar PowerShell .....	108
Aprovisionamiento de un cluster de HDInsight .....	108
Ejecución de un trabajo WordCount de MapReduce .....	111
Conexión a las herramientas de inteligencia empresarial de Microsoft.....	118
Uso de MapReduce con HDInsight.....	121
Escenario .....	121
Ejecución de la muestra con Azure PowerShell.....	122
El código Java para el programa de recuento de palabras de MapReduce ..	126
Carga de datos en HDInsight.....	129
Carga de datos en el almacenamiento de blobs usando AzCopy.....	130
Carga de datos en el almacenamiento de blobs usando Azure PowerShell .	131
Carga de datos en el almacenamiento de blobs usando el explorador de almacenamiento de Azure.....	132
Carga de datos en el almacenamiento de blobs usando la línea de comandos de Hadoop .....	134
Importación de datos a HDFS desde base de datos SQL o SQL Server usando Sqoop .....	136
Administración de HDInsight con PowerShell .....	138
Aprovisionamiento de un cluster de HDInsight .....	138
Enumeración y visualización de clusters.....	141
Eliminación de un cluster.....	141
Concesión/Revocación del acceso a los servicios de HTTP .....	142
Envío de trabajos de MapReduce .....	142
Envío de trabajos de Hive .....	144
Introducción al emulador de HDInsight .....	145
Instalación del emulador de HDInsight.....	145

Múltiplos: un conjunto de gráficos con los mismos ejes .....	266
Matrices.....	266
Tarjetas.....	266
Mosaicos.....	267
Analysis Services, Integration Services y Reporting Services.....	268
Informes de Reporting Services (SSRS) .....	269
SQL Server Integration Services.....	271
Analysis services .....	274
<b>Capítulo 8. Herramientas de Big Data en SAS .....</b>	<b>279</b>
Hadoop y Big Data en SAS .....	279
SAS, Hadoop y el proceso analítico .....	281
Big Data y soluciones Hadoop de SAS .....	282
Acceso y administración de datos de Hadoop.....	282
Explorar, visualizar y tratar datos científicos.....	283
Analizar y modelizar .....	283
Implementar e integrar .....	284
SAS/ACCESS interface to Hadoop.....	284
Características .....	286
Requisitos del sistema .....	287
Software de SAS requerido.....	288
SAS Data Management .....	288
Características .....	292
Requisitos del sistema .....	300
SAS servidor de federación .....	302
Características .....	305
Requisitos del sistema .....	306
Software SAS Base.....	307
Características .....	309
Requisitos del sistema .....	313
Herramientas para explorar y visualizar datos científicos.....	313
SAS Visual Analytics.....	314
Características .....	318
Requisitos del sistema. Entorno de servidor .....	322
Requisitos del sistema. Entorno del cliente.....	323
Software necesario .....	323
Soporte al cliente de BI móvil SAS .....	324

**BIG DATA. TÉCNICAS, HERRAMIENTAS Y APLICACIONES**

<b>SAS In-Memory Statistics for Hadoop .....</b>	<b>325</b>
Características .....	327
Requisitos del sistema .....	332
<b>Índice analítico .....</b>	<b>335</b>