

Figure 1

Survey Instrument

What is the current overall level of big data analytics usage at your company?

Very Frequently

Frequently

Occasionally

Rarely

Never

How frequently are these big data skills used at your organization? Please answer all questions.

	Very Frequently	Frequently	Occasionally	Rarely	Never
Artificial Intelligence	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
C/C++ programming	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cryptography	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data Mining	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data Visualization	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Java	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Machine learning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Natural Language Processing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Parallel processing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Structured Query Language	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Python	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statistical Analysis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Indicate which relational databases are in use at your organization. Select all that apply.

- IBM DB2
- Microsoft SQL Server
- MySQL
- Oracle database
- PostgreSQL
- SAP HANA
- Teradata
- Other please specify

Indicate which NOSQL non-relational databases are in use at your organization. Select all that apply.

- Apache Cassandra
- Couchbase
- ArangoDB
- Apache Hadoop MapReduce
- Apache CouchDB – document db
- Apache HBase
- MongoDB – document db
- Redis
- Other please specify

Indicate which big data tools are in use at your organization. Select all that apply.

- Apache Pig programming
- Apache Hadoop HDFS distributed file system
- Apache Hive Query Language
- Apache HBase column-oriented database
- Apache Avro data serialization
- Dryad
- Hortonworks Data Platform
- JAQL query language
- Jaspersoft BI Suite
- IBM InfoSphere
- Karmasphere studio and analyst (Hadoop)
- Apache Mahout machine learning algorithms
- Pentaho business analytics
- Skytree Server
- Tableau Desktop and Server
- Talend Open Studio
- IBM Apache Zookeeper
- Other please specify

Indicate which stream processing tools are in use at your organization. Select all that apply.

- Apache Kafka
- Apache Spark Streaming
- IBM InfoSphere Streams
- SAP HANA
- SQLstream s-Server
- StreamCloud
- Apache Storm

Other please specify

Indicate which data analysis tools are in use at your organization. Select all that apply.

Microsoft Dryad – parallel computing

IBM SPSS Modeler

IBM Watson Analytics

Konstanz Information Miner (KNMINE)

Apache MapReduce – parallel computing

R

RapidMiner

SAS Enterprise Miner

Waikato Environment for Knowledge Analysis (Weka)/Pentaho

Other please specify

Indicate which statistical analysis tools are in use at your organization. Select all that apply.

R

SAS JMP

Minitab

Matlab

SAS

SPSS

Stata

Statsoft Statistica

Other please specify

Indicate which data visualization tools are in use at your organization. Select all that apply.

FusionCharts

Google Analytics

Highcharts

IBM Watson Analytics

Microsoft Power BI

Oracle Visual Analyzer

Qlikview

SAP Analytics Cloud

Tableau

Other please specify