Install Apache Spark on Windows 10

# Overview

# We cover installation, testing, and initial exploration of Apache Spark in Windows 10.

# Prerequisites

# Dffsdf

# Dsffdsdf

# Winutils

The “winutils.exe” utility provides POSIX command support for Hadoop code and is used my Spark as well. This utility, along with the supporting libraries and dll files, must be provided for the Windows 10 operating system. The adventurous may wish to build their own winutils artifacts on Windows with C++; and they may reference the *2021-04-14 – Build your own WINUTILS* below.

# Dependencies

# Runtime Dependencies

* **Spark Versioning Policy**: <https://spark.apache.org/versioning-policy.html>
* **Spark 3.4 Dependencies**: <https://spark.apache.org/docs/latest/>
* **Historic Spark Version Differences:** <https://www.educba.com/spark-versions/>

|  |  |  |
| --- | --- | --- |
| **Language** | **Supported** | **Deprecated** |
| Spark Runtime - JVM | 8/11/17 | 8 |
| R | 3.5+ |  |
| Python | 3.7+ | 3.7 |
| Scala | 2.12/2.13 | Scala App compiled version matches Spark compiled version |

# Spark Downloads and Utilities

* Spark Distribution Downloads: <https://spark.apache.org/downloads.html>
* *Original* WINDUTILS site: <https://github.com/steveloughran/winutils>
* *New* WINUTILS site: <https://github.com/cdarlint/winutils>
* Spark 3.4 Hadoop 3.3.1 WINUTILS site: <https://kontext.tech/article/1221/hadoop-3-4-0-winutils-for-windows-10-x64>
* 2021-04-14 – Build your own WINUTILS: <https://pivotalbi.com/build-your-own-winutils-for-spark/>

# References

Spark Installation Instructions:

* 2023-06-14: <https://www.knowledgehut.com/blog/big-data/how-to-install-apache-spark-on-windows> (v3.4.0)
* 2023-01-27: <https://sparkbyexamples.com/spark/apache-spark-installation-on-windows/> (v3.0.0)
* 2021-Spark 3.1.2: <https://denisecase.github.io/starting-spark/>

# Sss

# Ssss

# Sss

# s