

# Apache Spark Installation Guide (Windows)

## Installing Apache Spark

1. Go to the Apache Spark official website. Here is the link: <https://spark.apache.org/>
2. Click on the "**Download**" tab in the top menu.
3. Scroll down to the "**Latest stable release**" section and click on the "**Download Spark**" link.
4. On the Spark download page, click on the "**Download Spark**" button for the latest version of Apache Spark for Windows.
5. Save the Spark binary distribution ZIP file to your computer.
6. Extract the contents of the ZIP file to a directory of your choice. This will be your Spark installation directory.
7. Open a command prompt with administrative privileges.
8. Navigate to the Spark installation directory using the "**cd**" command. For example, if you extracted the ZIP file to C:\spark, you can navigate to that directory with the following command: `cd C:\spark`.
9. Inside the Spark installation directory, locate the "**conf**" folder and go to its subdirectory "**templates**". Make a copy of the "**spark-env.cmd.template**" file and rename it to "**spark-env.cmd**".
10. Open the "**spark-env.cmd**" file with a text editor.
11. Uncomment the line that sets the "**SPARK\_HOME**" environment variable by removing the "#" character at the beginning of the line. Set the value of "SPARK\_HOME" to the path of your Spark installation directory. For example: `set SPARK_HOME=C:\spark`.
12. Save and close the "spark-env.cmd" file.
13. Return to the command prompt with administrative privileges.
14. Navigate to the "bin" directory inside the Spark installation directory. For example: `cd C:\spark\bin`.
15. Run the "**spark-shell.cmd**" script to start the Spark shell.
16. Apache Spark is now installed on your Windows computer. You can start using it by running Spark applications or executing Spark commands in the Spark shell. Note that you may need to configure additional settings, such as setting the JAVA\_HOME environment variable, depending on your specific setup. Refer to the Apache Spark documentation for more information.

### Note:

If all steps have been followed, MySQL should run fine on your machine. If, however, troubleshooting and support is required, please contact that Data Science Helpdesk to resolve<sup>1</sup>.

Click [HERE](#) to schedule a meeting with the Helpdesk!

---

<sup>1</sup> This document was last updated on April 24, 2023

# Apache Spark Installation Guide (MacOS)

## Installing Apache Spark

1. Go to the Apache Spark official website. Here is the link: <https://spark.apache.org/>
2. Click on the "**Download**" tab in the top menu.
3. Scroll down to the "**Latest stable release**" section and click on the "**Download Spark**" link.
4. On the Spark download page, click on the "**Download Spark**" button for the latest version of Apache Spark for macOS.
5. Save the Spark binary distribution TAR file to your computer.
6. Open a terminal window on your macOS.
7. Navigate to the directory where you saved the Spark TAR file using the "**cd**" command. For example, if you saved the file in your Downloads folder, you can navigate to that directory with the following command: `cd ~/Downloads`.
8. Extract the contents of the TAR file to a directory of your choice. This will be your Spark installation directory. You can use the "tar" command to extract the TAR file. For example: `tar -xzf spark-<version>.tgz`, where <version> is the version number of Spark that you downloaded.
9. Move the extracted Spark directory to a desired location, if needed.
10. Inside the Spark installation directory, locate the "**conf**" folder and go to its subdirectory "**templates**". Make a copy of the "**spark-env.sh.template**" file and rename it to "**spark-env.sh**".
11. Open the "spark-env.sh" file with a text editor.
12. Uncomment the line that sets the "SPARK\_HOME" environment variable by removing the "#" character at the beginning of the line. Set the value of "SPARK\_HOME" to the path of your Spark installation directory. For example: `export SPARK_HOME=/Users/<username>/spark`, where <username> is your macOS username.
13. Save and close the "spark-env.sh" file.
14. In the terminal, navigate to the "bin" directory inside the Spark installation directory. For example: `cd ~/spark/bin`.
15. Run the "**./spark-shell**" script to start the Spark shell.
16. Apache Spark is now installed on your macOS. You can start using it by running Spark applications or executing Spark commands in the Spark shell. Note that you may need to configure additional settings, such as setting the JAVA\_HOME environment variable, depending on your specific setup. Refer to the Apache Spark documentation for more information

### Note:

If all steps have been followed, MySQL should run fine on your machine. If, however, troubleshooting and support is required, please contact that Data Science Helpdesk to resolve<sup>2</sup>.

Click [HERE](#) to schedule a meeting with the Helpdesk!

---

<sup>2</sup> This document was last updated on April 24, 2023