



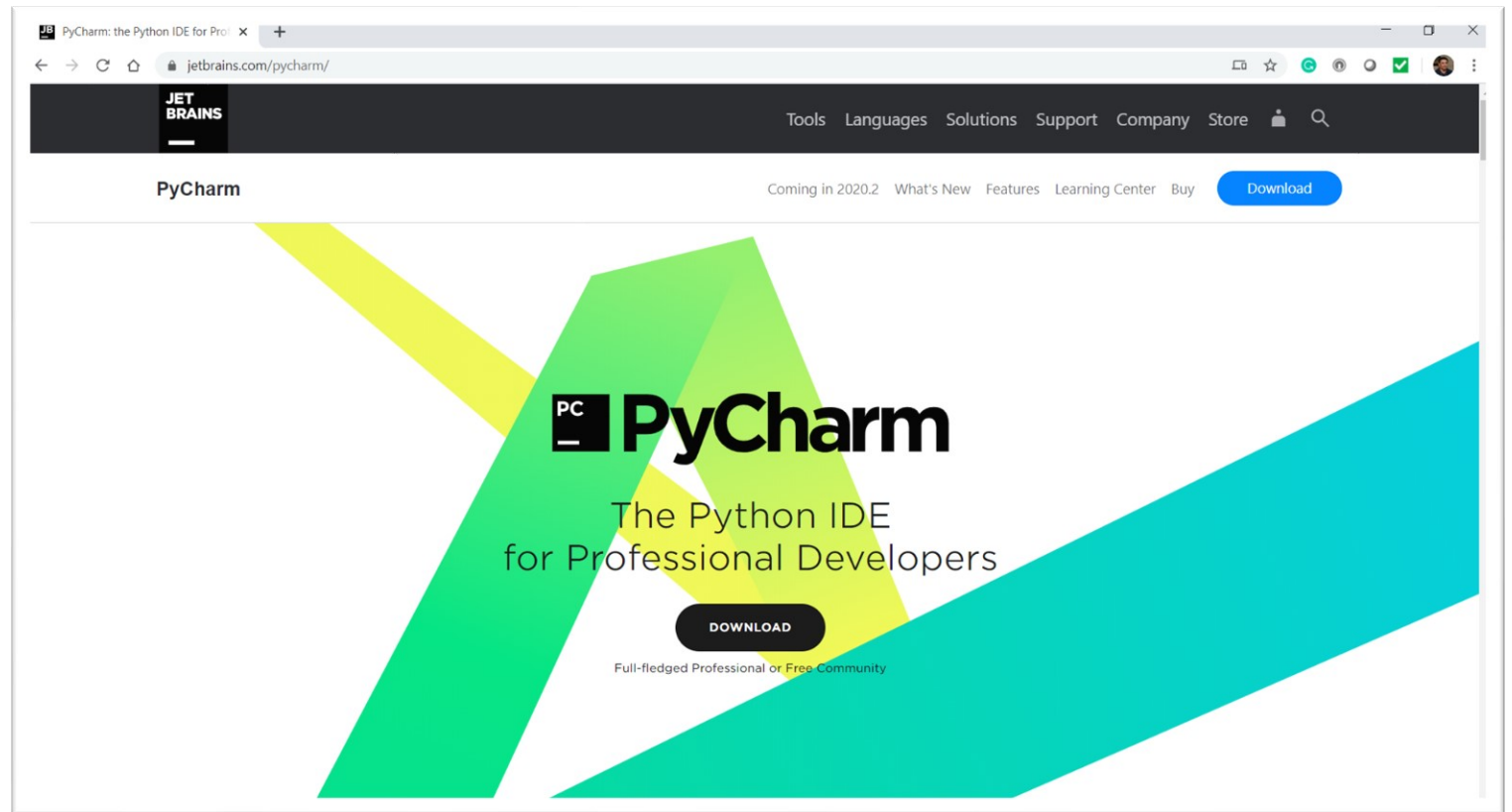
# Apache Spark

Setup your IDE

# Setup your IDE



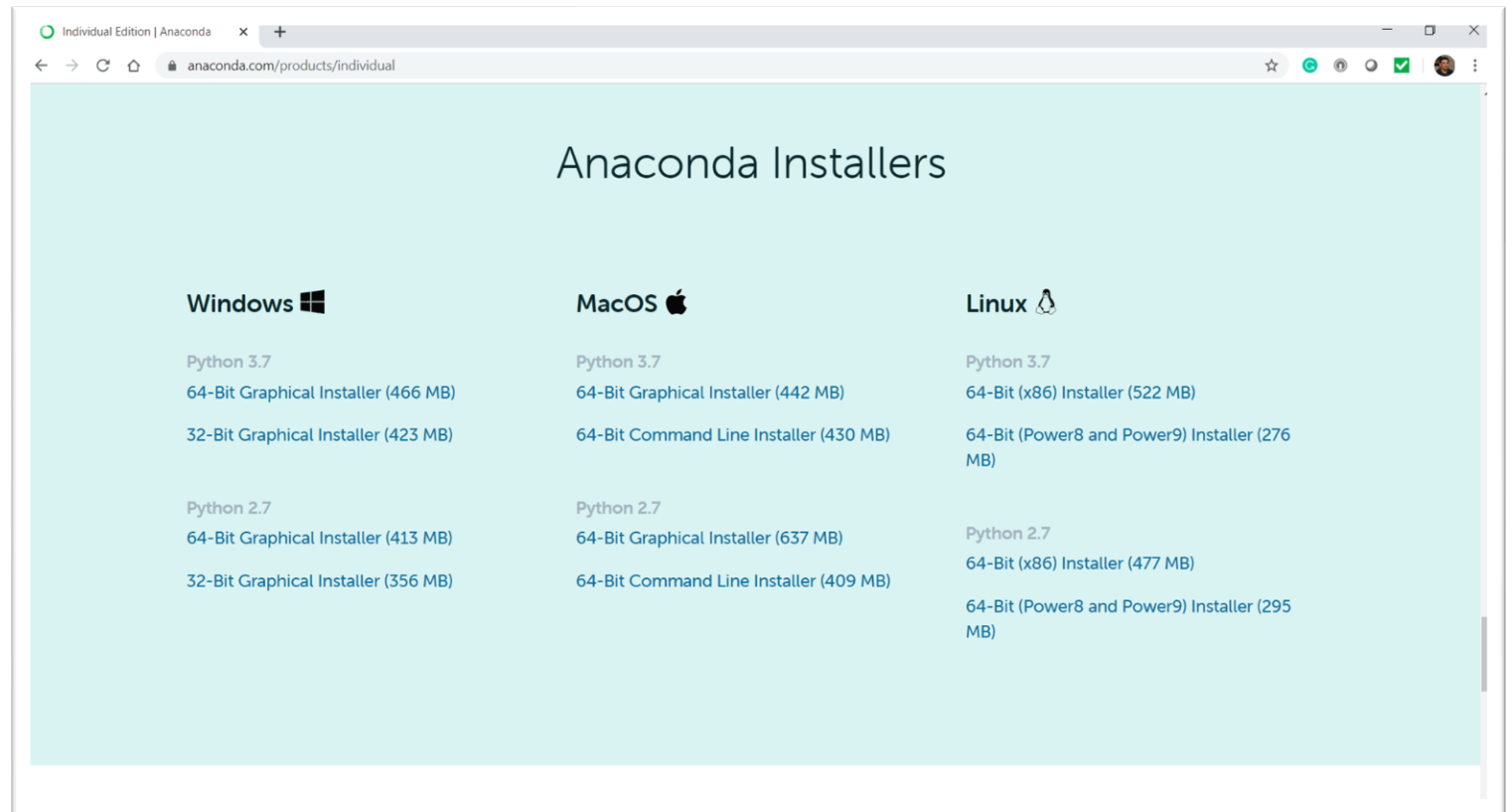
<https://www.jetbrains.com/pycharm/>



# Setup your IDE



<https://www.anaconda.com/products/individual>



# Setup your IDE



<https://spark.apache.org/downloads.html>

A screenshot of the Apache Spark Downloads page in a web browser. The browser's address bar shows the URL "spark.apache.org/downloads.html". The page features the Apache Spark logo and the tagline "Lightning-fast unified analytics engine". A navigation bar includes links for "Download", "Libraries", "Documentation", "Examples", "Community", "Developers", and "Apache Software Foundation". The main content area is titled "Download Apache Spark™" and provides instructions for downloading a release. It includes two dropdown menus for selecting a Spark release (2.4.5) and a package type (Pre-built for Apache Hadoop 2.7). A download link for "spark-2.4.5-bin-hadoop2.7.tgz" is provided, along with a note about Scala versions. A section for "Latest Preview Release" explains that preview releases are for testing upcoming features and may contain bugs. A "Link with Spark" section shows the Maven coordinates for Spark. A "Download Spark" button is prominently displayed. On the right side, there is a "Latest News" section with links to recent releases and an "Archive" link. Below the "Download Spark" button, there is a "Built-in Libraries" section with links to SQL and DataFrames, Spark Streaming, MLlib (machine learning), GraphX (graph), and Third-Party Projects.

Downloads | Apache Spark

spark.apache.org/downloads.html

**APACHE Spark™** Lightning-fast unified analytics engine

Download Libraries Documentation Examples Community Developers Apache Software Foundation

### Download Apache Spark™

1. Choose a Spark release:
2. Choose a package type:
3. Download Spark: [spark-2.4.5-bin-hadoop2.7.tgz](#)
4. Verify this release using the 2.4.5 [signatures](#), [checksums](#) and [project release KEYS](#).

Note that, Spark is pre-built with Scala 2.11 except version 2.4.2, which is pre-built with Scala 2.12.

### Latest Preview Release

Preview releases, as the name suggests, are releases for previewing upcoming features. Unlike nightly packages, preview releases have been audited by the project's management committee to satisfy the legal requirements of Apache Software Foundation's release policy. Preview releases are not meant to be functional, i.e. they can and highly likely will contain critical bugs or documentation errors. The latest preview release is Spark 3.0.0-preview2, published on Dec 23, 2019. You can select and download it above.

### Link with Spark

Spark artifacts are [hosted in Maven Central](#). You can add a Maven dependency with the following coordinates:

```
groupId: org.apache.spark  
artifactId: spark-core_2.11  
version: 2.4.5
```

### Installing with PyPi

#### Latest News

Spark 2.4.5 released (Feb 08, 2020)  
Preview release of Spark 3.0 (Dec 23, 2019)  
Preview release of Spark 3.0 (Nov 06, 2019)  
Spark 2.3.4 released (Sep 09, 2019) [Archive](#)

**APACHE EVENTS** [LEARN MORE](#)

**Download Spark**

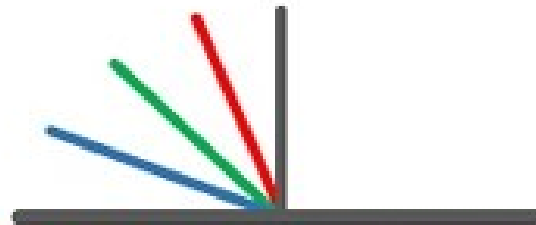
#### Built-in Libraries:

[SQL and DataFrames](#)  
[Spark Streaming](#)  
[MLlib \(machine learning\)](#)  
[GraphX \(graph\)](#)  
[Third-Party Projects](#)

# Setup your IDE



1. Anaconda Platform
2. PyCharm
3. Spark Binaries
4. SPARK\_HOME
5. HADOOP\_HOME



LEARNING JOURNAL

[www.learningjournal.guru](http://www.learningjournal.guru)