

Getting started with Databricks Notebooks

Understanding Spark, July 29, 2017

Renat Bekbolatov, [Easy Eight Consulting](#)

Databricks

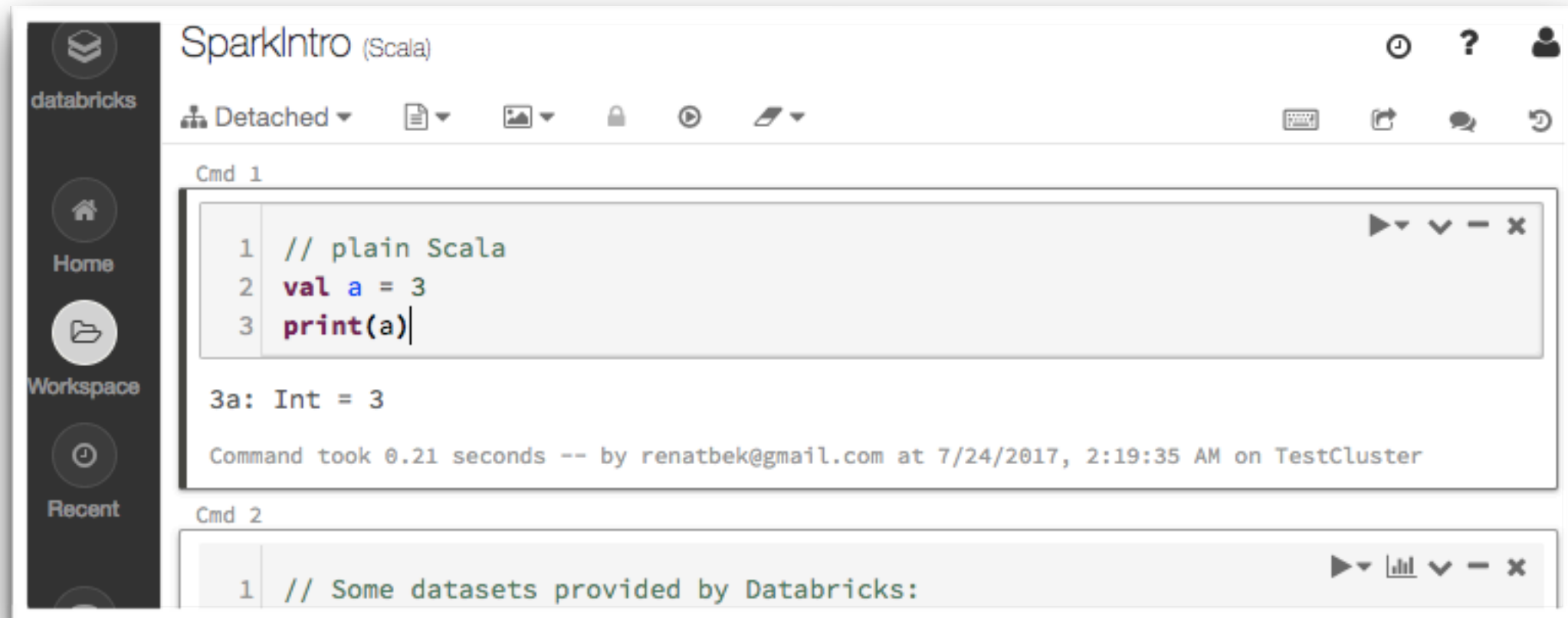
Unified Analytics Platform



From the team that created Apache Spark

Renat Bekbolatov, [Easy Eight Consulting](#)

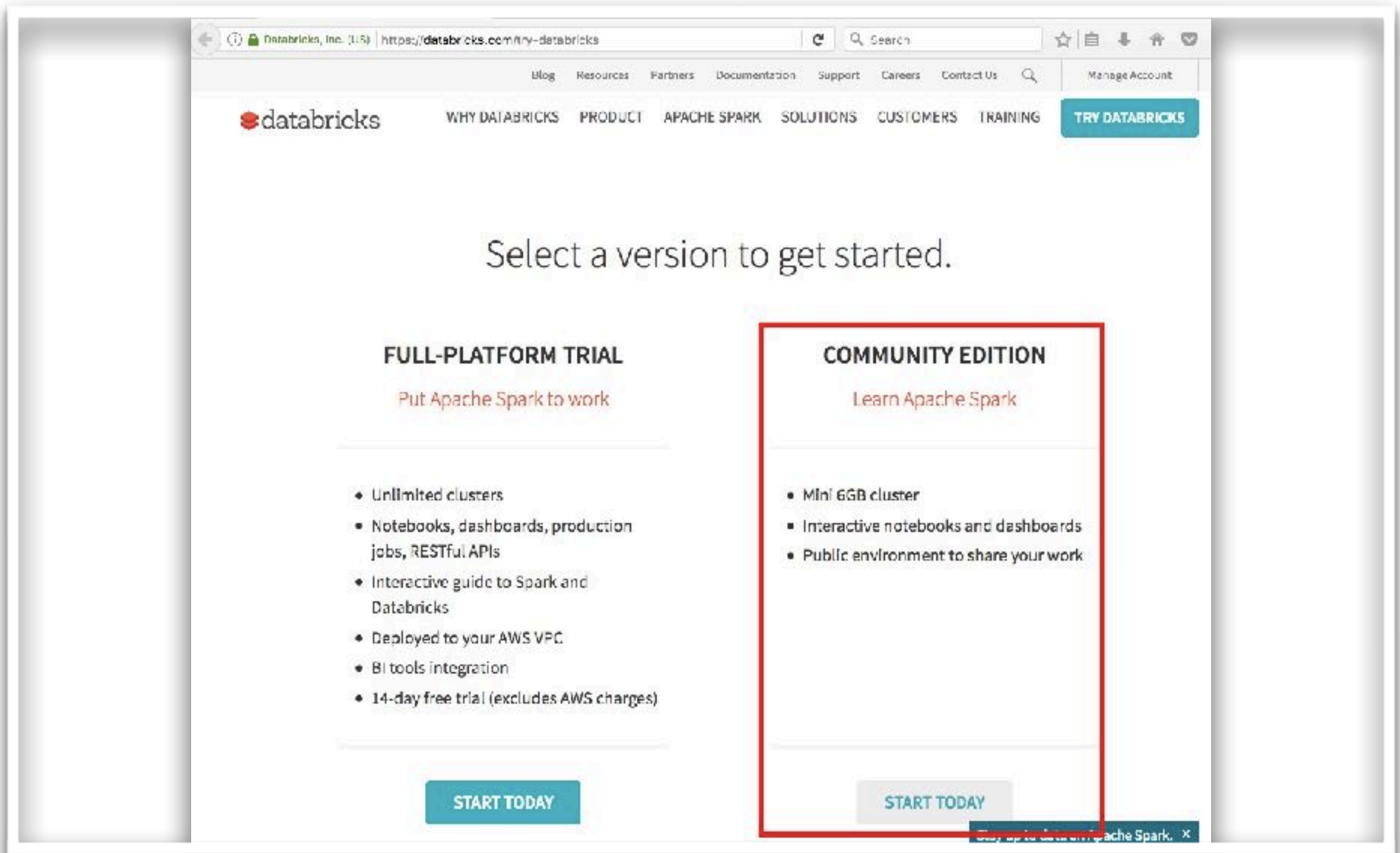
Databricks Notebooks



- interactive coding environment
- combine code and execution
- Scala, Python, R
- share/import notebooks
- an much more ...

Now, let's sign up for a Databricks **account**, create a **cluster** and import a **notebook**.

1) Go to: databricks.com/try-databricks



2) Sign up for "Community Edition"

3) Create a cluster

Create Cluster

New Cluster Cancel Create Cluster 0 Workers: 0.0 GB Memory, 0 Cores, 0 DBU
1 Driver: 6.0 GB Memory, 0.88 Cores, 1 DEU; ?

Cluster Name

Databricks Runtime Version ?
Spark 2.1.1-db5 (Scala 2.11)

Instance

Free 6GB Memory: As a Community Edition user, your cluster will automatically terminate after an idle period of two hours. For [more configuration options](#), please [upgrade your Databricks subscription](#).

Instances Spark

Availability Zone ?

Currently supported browsers: Chrome or Firefox

4) Import a notebook:

The screenshot shows the Databricks web interface. On the left sidebar, the 'Workspace' icon is circled in red. In the main area, the 'Clusters' tab is selected, and a dropdown menu is open with the 'Import' option highlighted by a red rectangle. Below this, an 'Import Notebooks' dialog box is shown, which is outlined in orange. The dialog has 'Import from:' with radio buttons for 'File' and 'URL' (selected). A text input field contains the URL: `n/Bekbolatov/presentation/blob/master/understanding_spark/SparkIntro.ipynb`. Below the input field, it lists accepted formats: `.dbc`, `.scala`, `.py`, `.sql`, `.r`, `.ipynb`, `.html`. At the bottom right of the dialog are 'Cancel' and 'Import' buttons. An orange arrow points from the URL in the dialog to a box at the bottom of the slide.

Workspace

Shared

Clusters

Import

Import Notebooks

Import from: ☐ File ☒ URL

`n/Bekbolatov/presentation/blob/master/understanding_spark/SparkIntro.ipynb`

Accepted formats: `.dbc`, `.scala`, `.py`, `.sql`, `.r`, `.ipynb`, `.html`
(To import a library such as a jar or egg, [click here](#))

Cancel Import

https://github.com/Bekbolatov/presentation/blob/master/understanding_spark/SparkIntro.ipynb