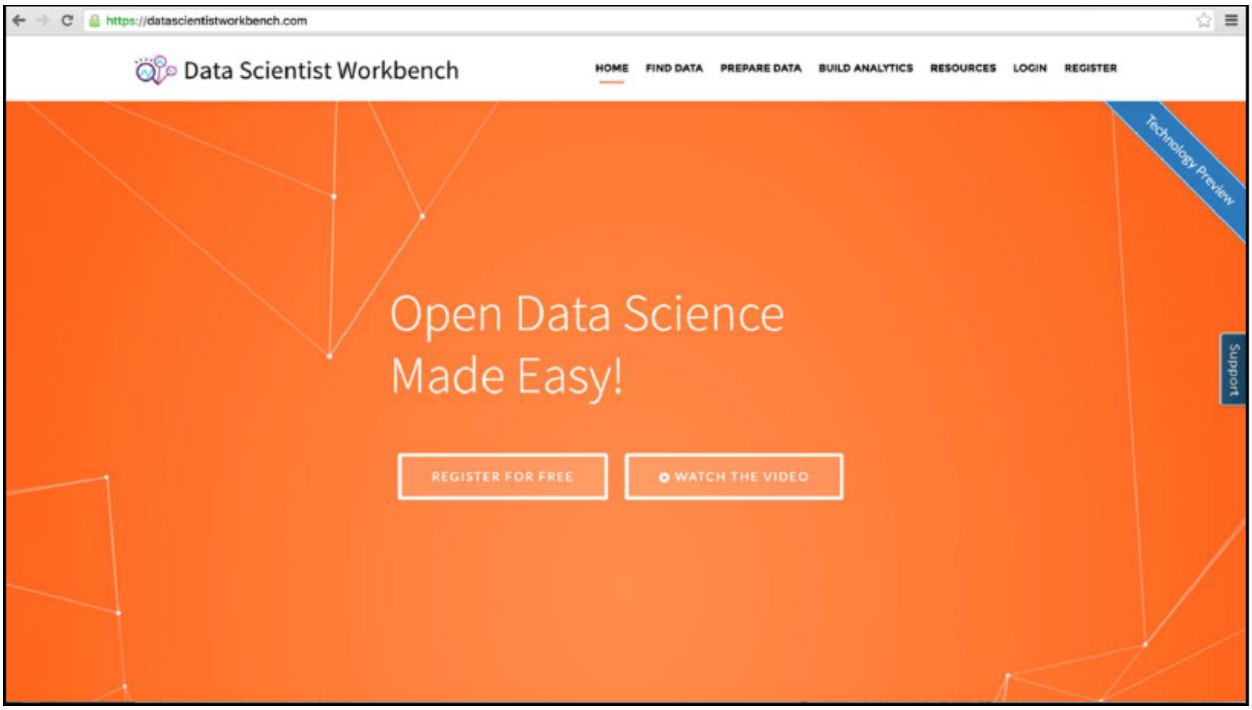
Lab: Introducing Data Scientist Workbench

This Lab must be done on the Data Scientist Workbench. If you have not signed up yet, go to datascientistworkbench.com and click REGISTER FOR FREE. Then login (2nd link from the right):



Contents

Lab 1 Introducing Data Scientist Workbench

After completing this hands-on lab, you will know:

- How to register an account on Data Scientist Workbench (DSWB).
- How to upload and work with your data on DSWB.
- How to use OpenRefine to prepare your data on DSWB.
- How to open and do simple tasks with the build analytics tools on DSWB, Jupyter & Zeppelin Notebbooks, RStudio IDE, and Seahorse.
- How to change your profile settings on DSWB.
- How to open the Feedback Forum and vote on your favourite ideas.
- How to see the courses available for free on Big Data University.

Allow 30 minutes to complete this section of the lab.

NOTE: This is a guided Lab. Solutions are found at the end of this document

1.1 Getting Started with Data Scientist Workbench

__1. Register an account on DSWB

__2. Upload data by going to My Data, and Upload Data

__3. Explore each of the tools: Click OpenRefine

__4. Click on Jupyter Notebook and explore the interface, like clicking on one of the tutorial notebooks

__5. Click on Zeppelin Notebook and import one of the sample notebooks

__6. Click on RStudio IDE and run 1+1 in the console

__7. Open Seahorse and open one of the tutorials

__8. Change your profile settings by toggling the left-hand sidebar and clicking on your username (https://my.datascientistworkbench.com/profile). Select a preferred data center to optimize your account's performance (upload/download speeds for you)

__9. Click on the Feedback Forum and read through some of the top ideas. Vote for your favorite idea.

__10. Click on Online Learning to check out courses on Big Data University

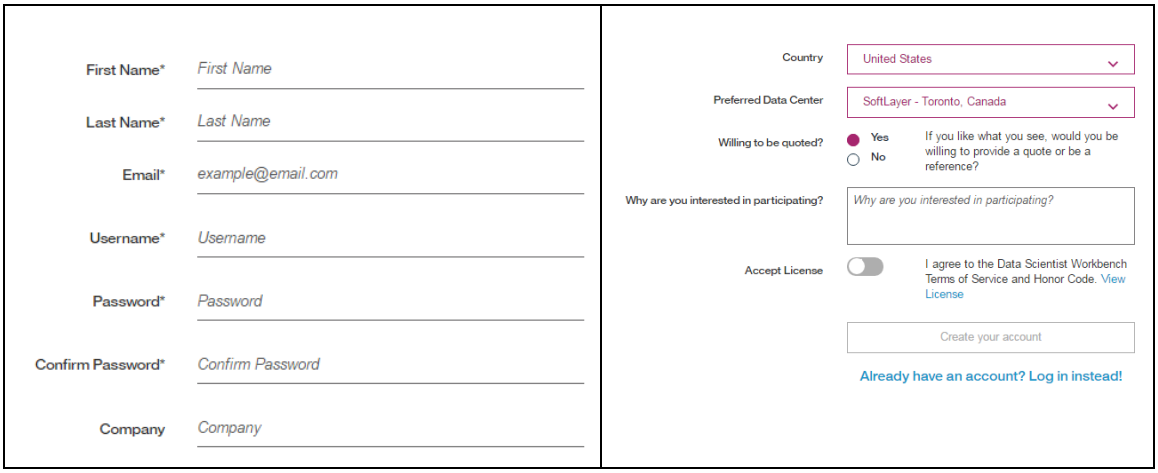
1.2 Summary

Congratulations! ...

1.3 Getting Started with Data Scientist Workbench - Solutions

__1. **Register an account on DSWB:** *Go to datascientistworkbench.com and click **SIGN UP FOR FREE**. If you have already registered, simply login. You do not need to do this step. To register, create an account using one of your social media accounts (follow the instructions on the screens that open up for each of the social media, as they differ somewhat):*

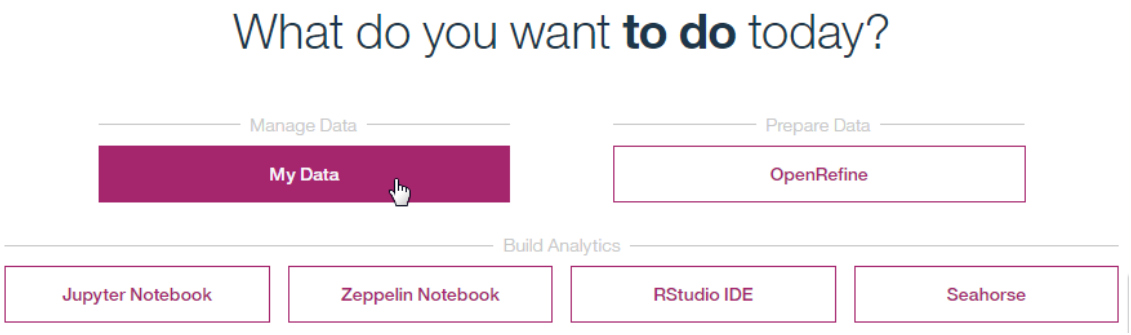
Or create your account specifying the required information:



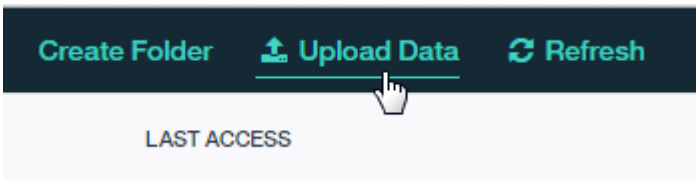
Based on your location, use the default Data Center, or select one from the dropdown.

Then view the License terms, and click Accept License to create your account.

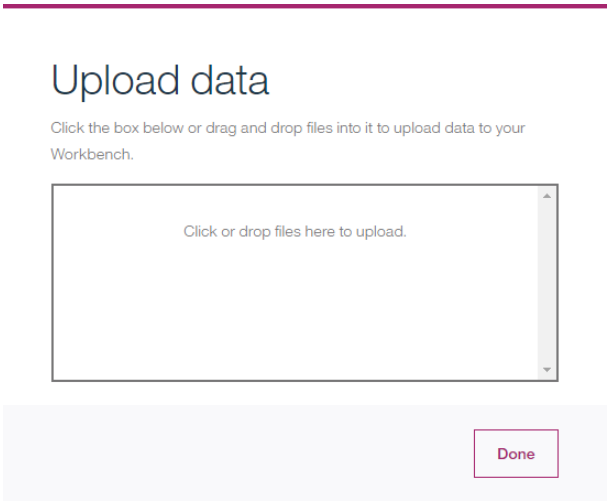
__2. **Upload data by going to My Data, and Upload Data:** From the DSWB home page, click My Data:



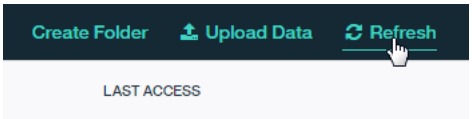
*To upload a file to the default /resources/data directory, click **Upload Data** (top right):*



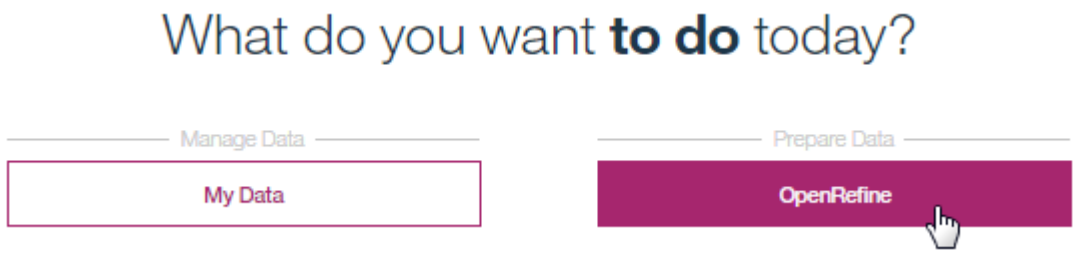
*Drag and drop the files you want to upload and click **Done**:*



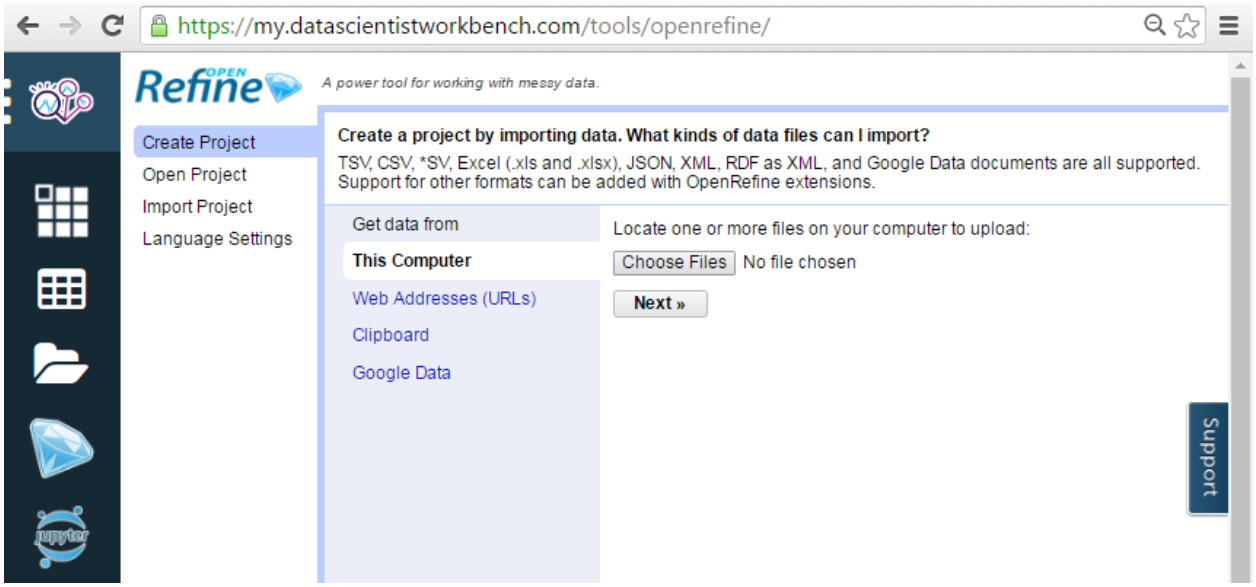
*To check if your files have been uploaded, click **Refresh**:*



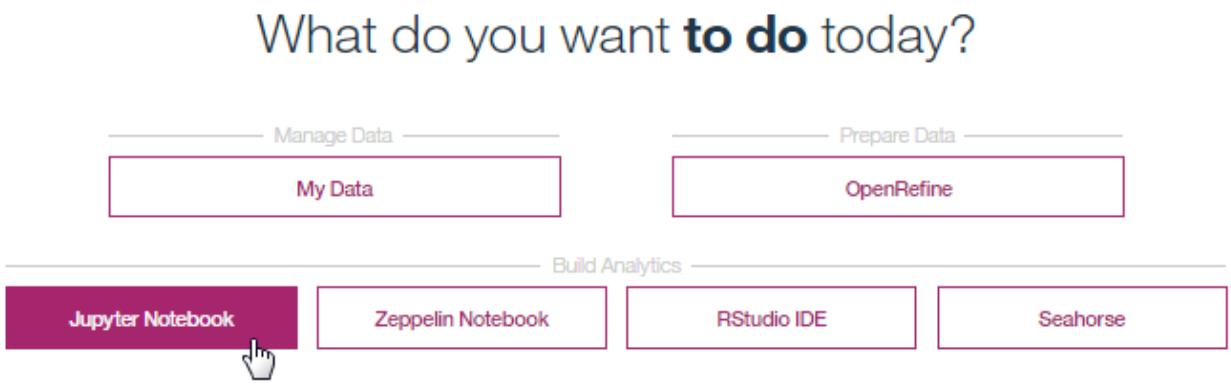
__3. **Explore each of the tools: Click OpenRefine:**



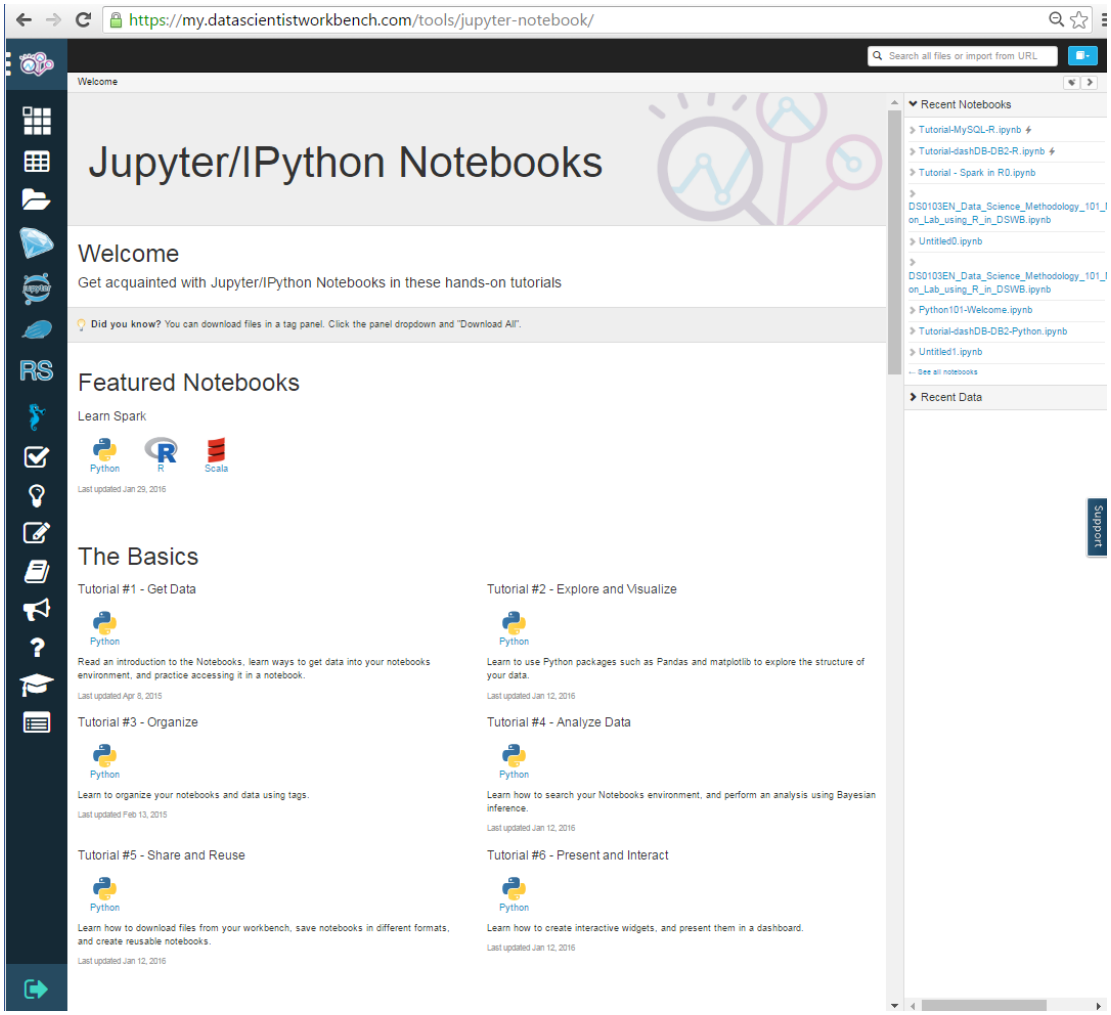
OpenRefine opens – explore some of the options:



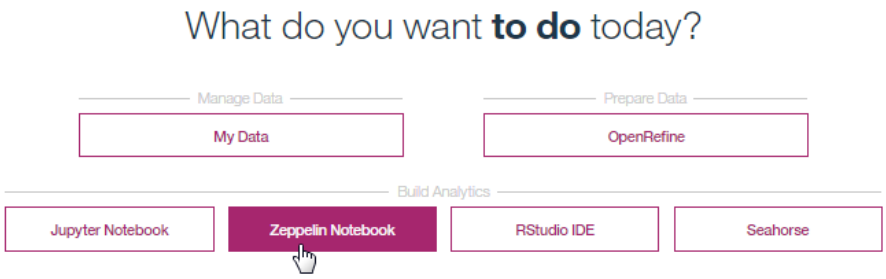
__4. **Click on Jupyter Notebook and explore the interface, like clicking on one of the tutorial notebooks:**



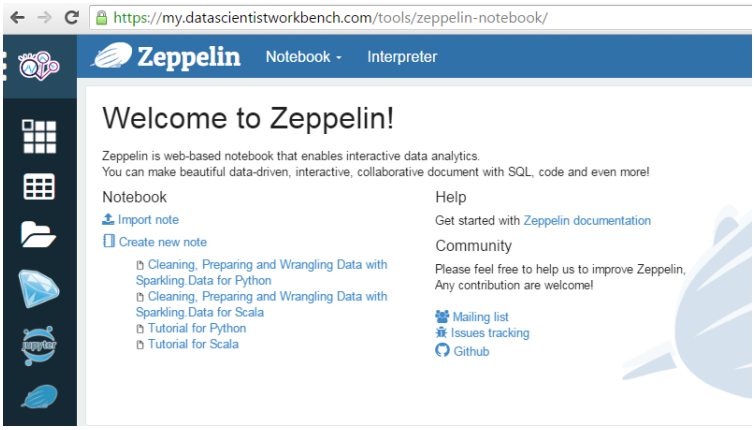
The Jupyter Notebook opens – scroll down and explore some of the tutorials:



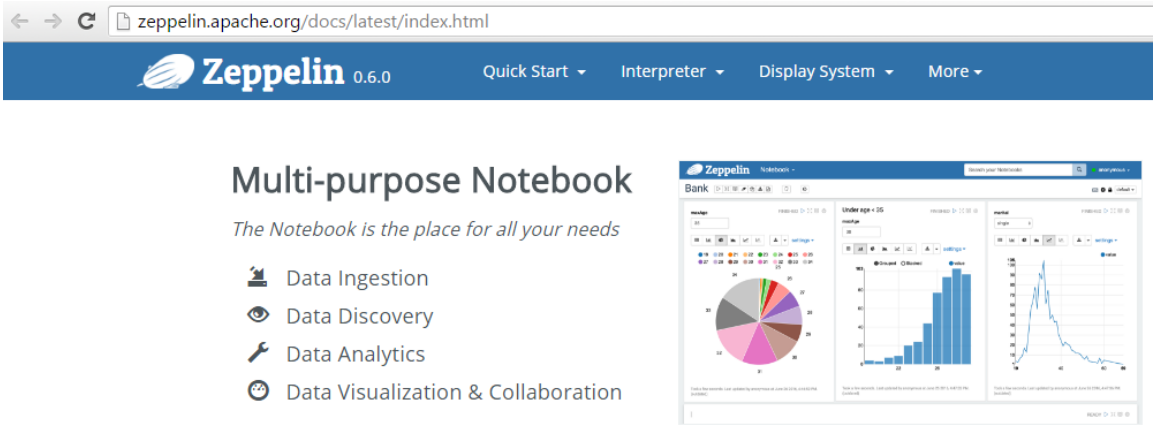
__5. **Click on Zeppelin Notebook and import one of the sample notebooks:**



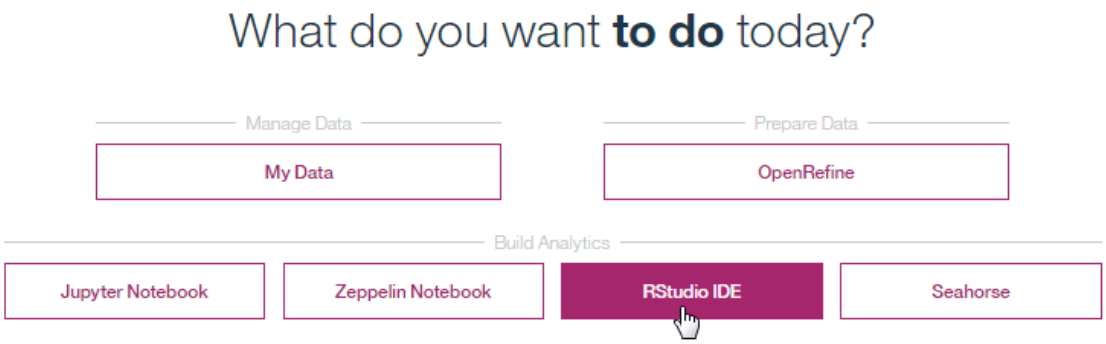
The Zeppelin Notebook opens:



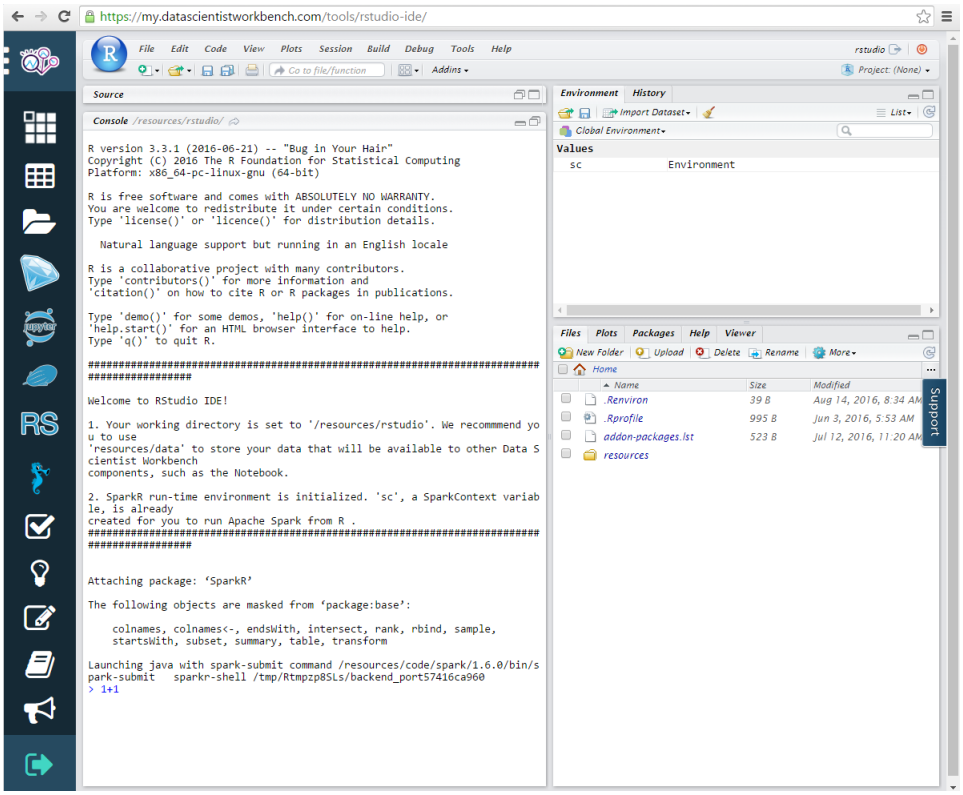
Click Help to explore what this Notebook does:



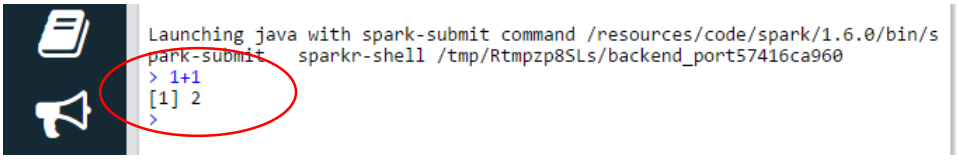
__6. **Click on RStudio IDE and run 1+1 in the console:**



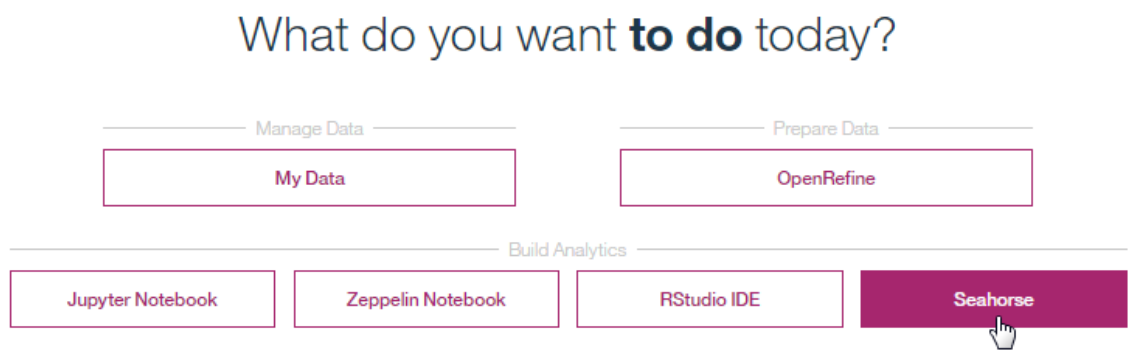
RStudio IDE opens; type 1+1 on the Console (left pane) – as shown below; right pane shows environment variables, history, Files, Plots, Packages, Help, and Viewer:



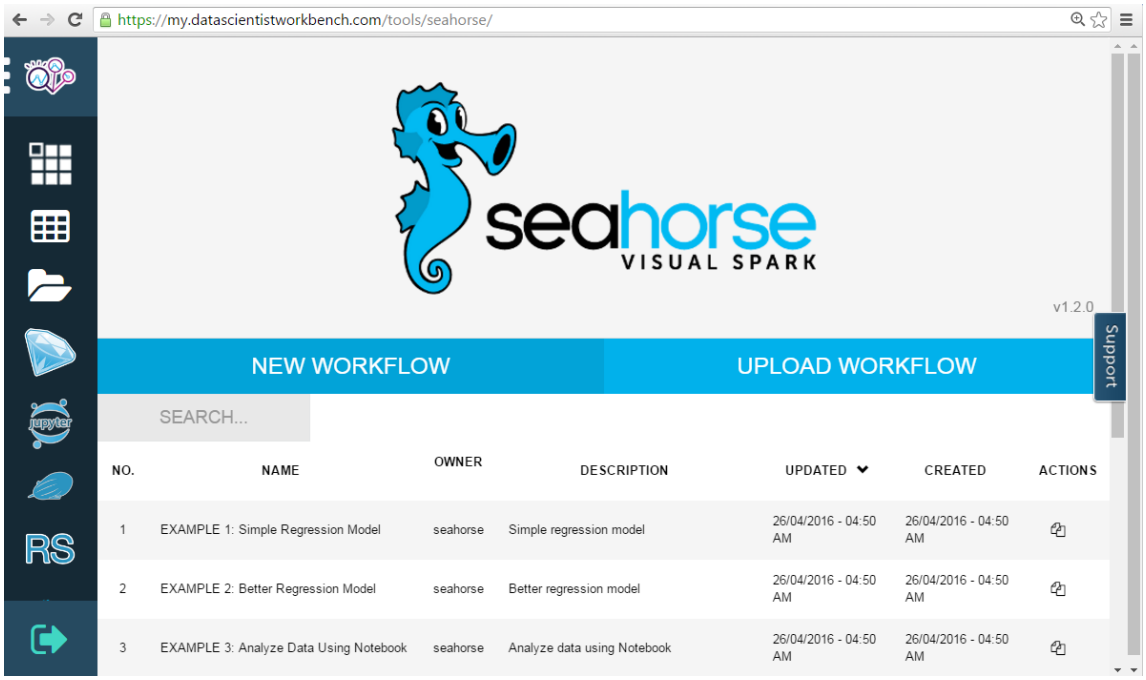
To run the code, press Enter; the output displays below the code:



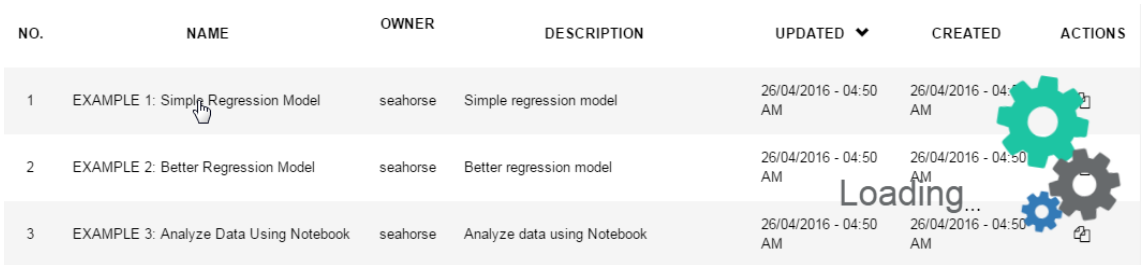
__7. **Open Seahorse and open one of the tutorials:**



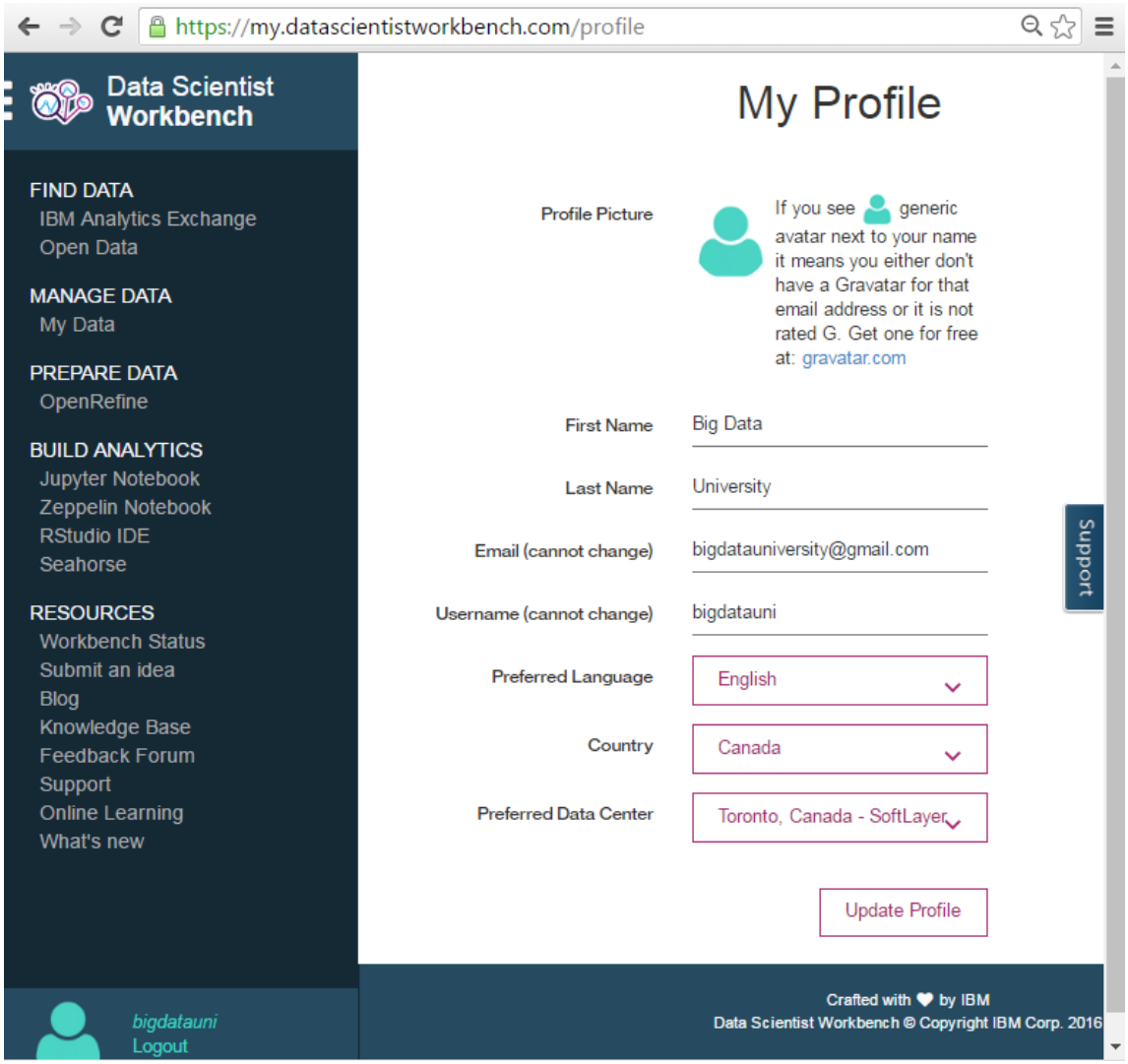
Seahorse opens:



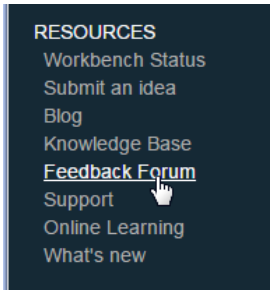
Explore one of the tutorials (by clicking on it):

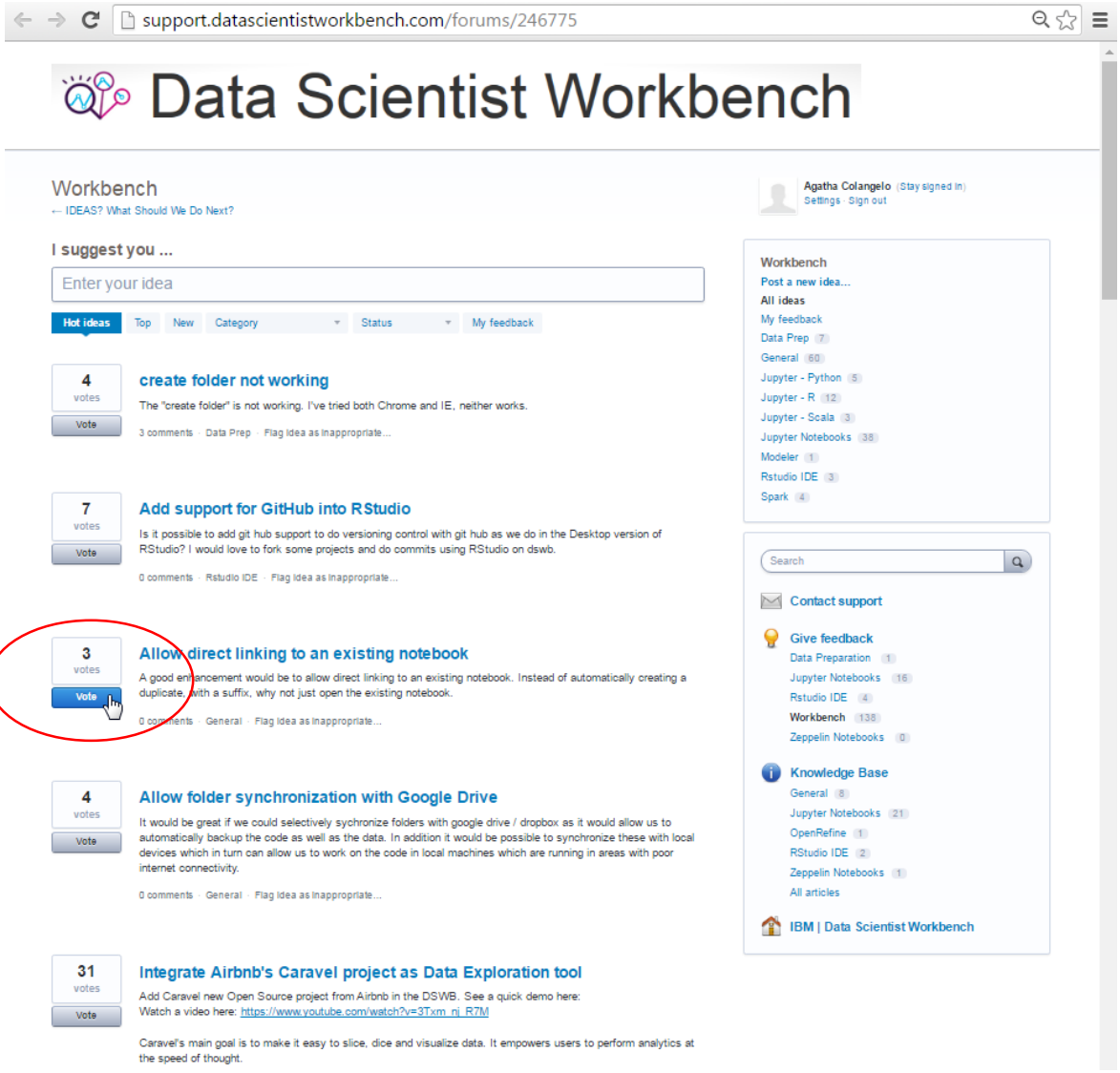


__8. **Change your profile settings** *by toggling the left-hand sidebar (bottom left) and clicking on your username:*

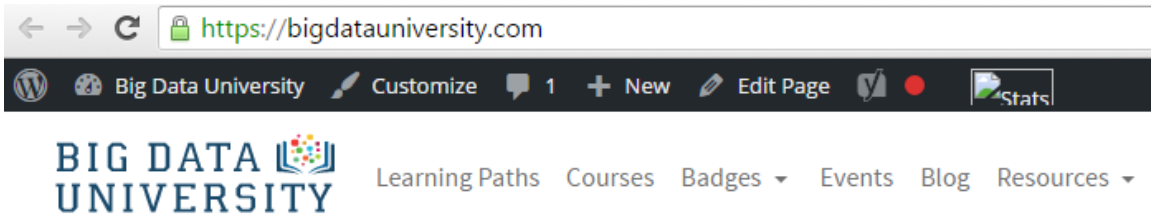


__9. **Click on the Feedback Forum and read through some of the top ideas. Vote for your favorite idea:**

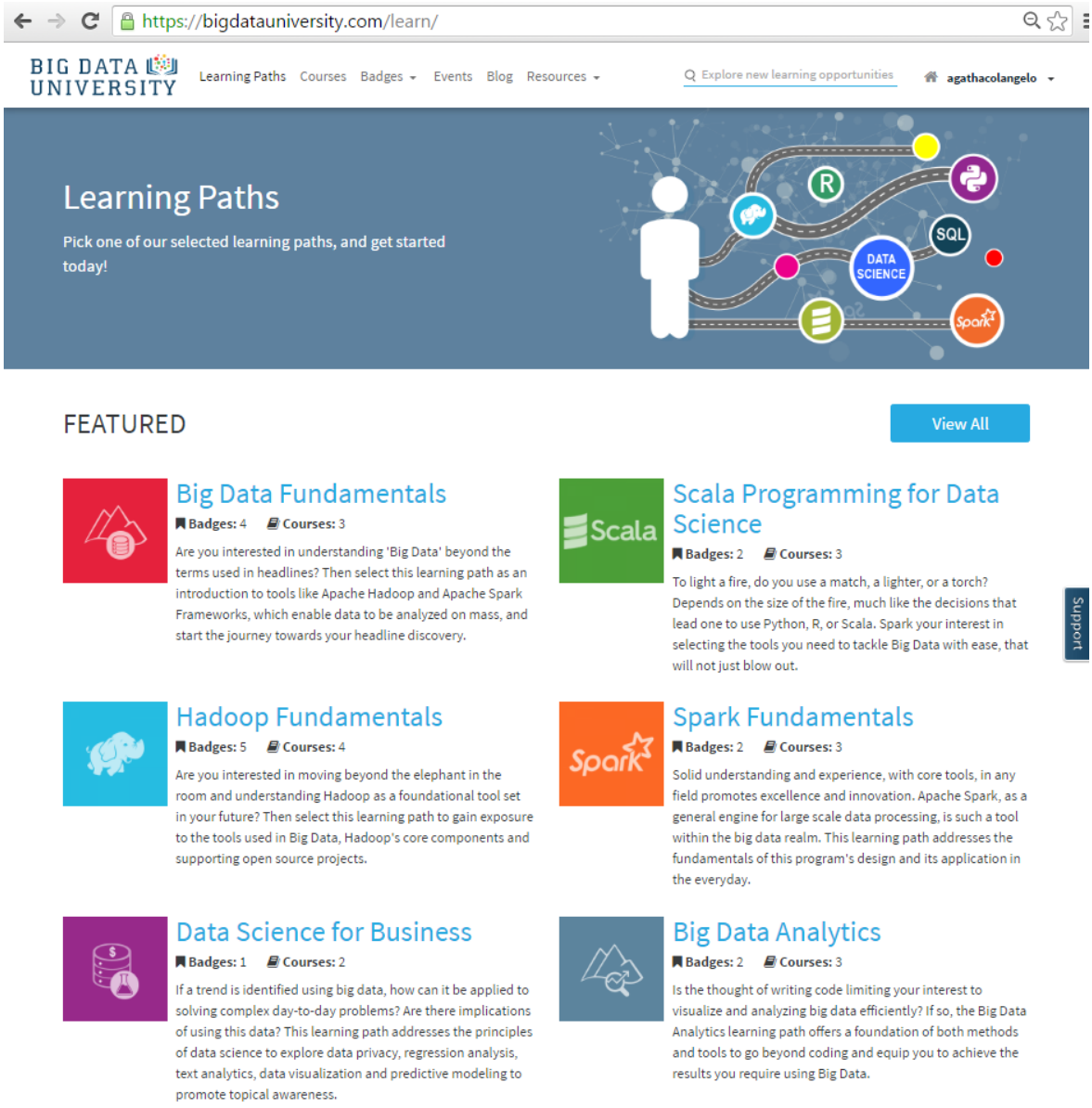




__10. **Click on Online Learning to check out courses on Big Data University:**



Check out the Learning paths:



And some of the free courses: