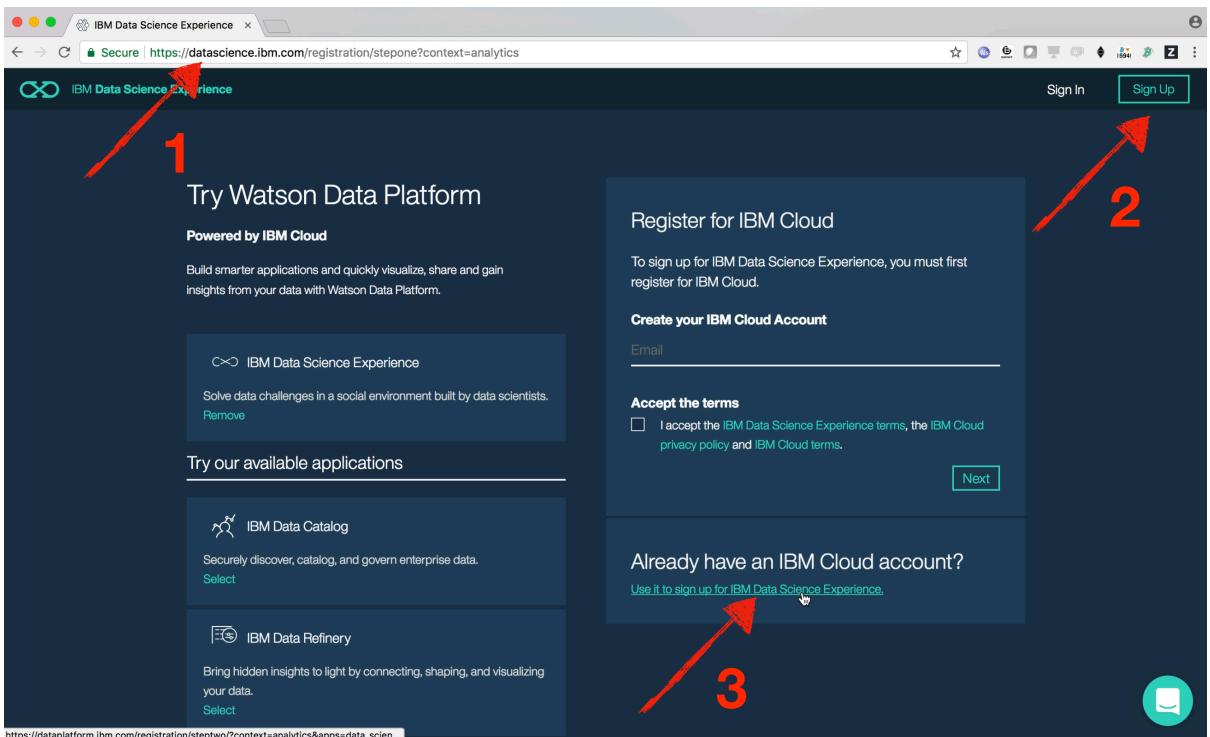


Create and setup a Data Science Experience (DSX) project for jupyter and Apache Spark in the IBM Cloud

Step 1: Signup for DSX

1. Start the process



- a. Open the following URL (1)
<http://datascience.ibm.com>
*HINT: In case you've accidentally created your resourced in United Kingdom instead of US South you can just use the following link instead:
<https://eu-gb.datascience.ibm.com/>.*
- b. Click on “Sign Up” (2)
- c. Click on “Use it to sign up for IBM Data Science Experience (3)

2. Make use all dropdowns are filled

Select Organization and Space

Confirm your IBM Cloud organization and space information below.
[Or create new organization and space](#)

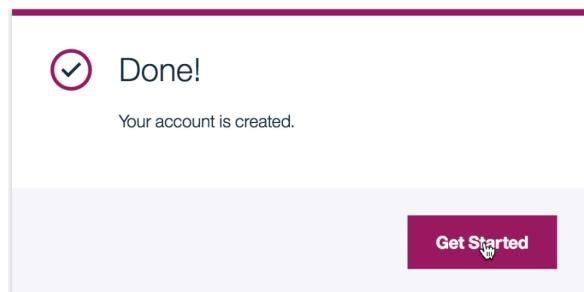
Select IBM Cloud account
Dummyuser23.ormium@spamgourmet.com ▾

IBM Cloud Organization
Dummyuser23.ormium@spamgourmet.com ▾

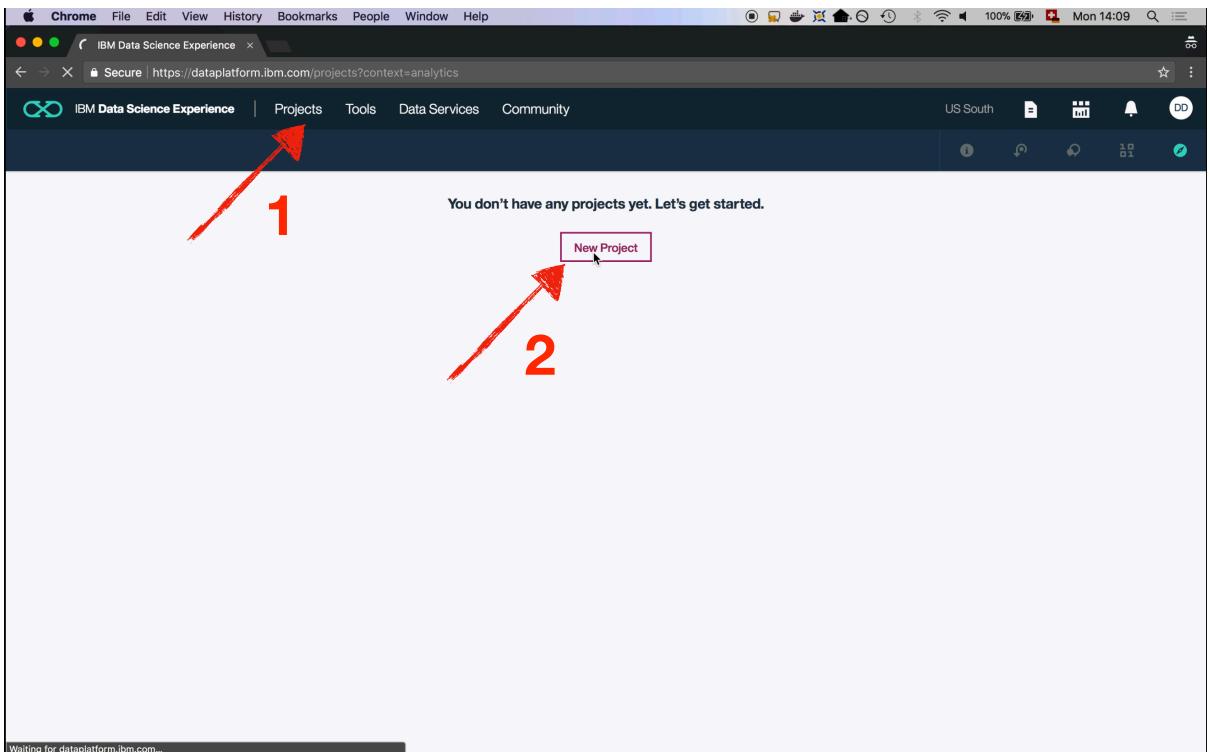
IBM Cloud Space
dev ▾

 Continue

3. Once the process is finished (it may take up some time) click on “Get Started”



4. In the DSX UI, create a new Project



- a. Click on “Projects”
- b. Click on “New Project”

5. Configure the project

New project

Define project details

Name
default

Description
Project description

Choose project options

Restrict who can be a collaborator

Define storage

① Select storage type

Object Storage (Swift API) IBM Cloud Object Storage

② Add

Add an object storage instance and then return to this page and click Refresh.

③ Refresh

Define compute engine

① Select Spark service

Spark service Add

Add IBM Analytics for Apache Spark, then re Refresh.

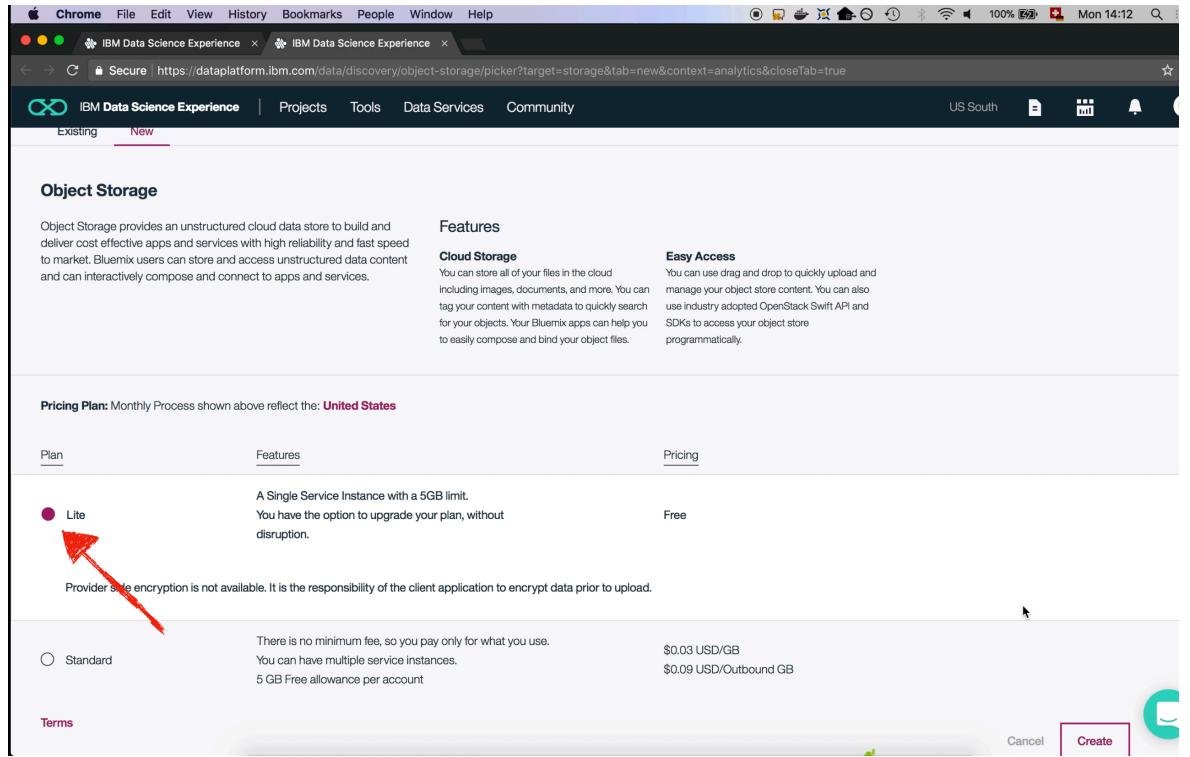
② Refresh

Armand from IBM
Hi
Dummyuser231.ormium@spamgourme...

Cancel Create

- a. Enter a name (1)
- b. Select “Object Storage (Swift API) (2)

c. Click on “Add” (3)



The screenshot shows the 'Object Storage' creation page in the IBM Data Science Experience. At the top, there are tabs for 'Existing' and 'New'. Below the tabs, the title 'Object Storage' is displayed. A brief description of Object Storage follows. To the right, there are two sections: 'Features' and 'Cloud Storage' (under 'Easy Access'). Under 'Cloud Storage', it says: 'You can store all of your files in the cloud including images, documents, and more. You can tag your content with metadata to quickly search for your objects. Your Bluemix apps can help you to easily compose and bind your object files.' In the 'Features' section, it says: 'Provider side encryption is not available. It is the responsibility of the client application to encrypt data prior to upload.' Below this, there is a table comparing 'Plan', 'Features', and 'Pricing' for 'Lite' and 'Standard' plans.

Plan	Features	Pricing
<input checked="" type="radio"/> Lite	A Single Service Instance with a 5GB limit. You have the option to upgrade your plan, without disruption. Provider side encryption is not available. It is the responsibility of the client application to encrypt data prior to upload.	Free
<input type="radio"/> Standard	There is no minimum fee, so you pay only for what you use. You can have multiple service instances. 5 GB Free allowance per account	\$0.03 USD/GB \$0.09 USD/Outbound GB

Terms

Cancel Create

- i. Select “Lite”
- ii. Click on “Create”

iii. Make sure all fields are filled up and click on “Confirm”

Confirm Creation

Organization: Dummyuser231.ormium@spamgourmet.com

Plan
Lite

Space
dev

Service name
object-storage-kp

[Cancel](#) [Confirm](#)

d. Click on “Add” again

The screenshot shows the 'New project' page in the IBM Data Science Experience. The left panel contains 'Define project details' (Name: default, Description: Project description) and 'Choose project options' (checkbox for Restrict who can be a collaborator). The right panel contains 'Define storage' (Select storage type: Object Storage (Swift API) selected, IBM Cloud Object Storage), '② Add' (button highlighted with a red arrow), '③ Refresh', and 'Define compute engine' (Select Spark service: Spark service, Add button highlighted with a red arrow). A tooltip for the 'Add' button in the compute engine section indicates: 'Add IBM Analytics for Apache Spark, then re Refresh.' A message from Armand from IBM is visible at the bottom right.

New project

Define project details

Name
default

Description
Project description

Choose project options

① Select storage type

Object Storage (Swift API) IBM Cloud Object Storage

② Add

Add an object storage instance and then return to this page and click Refresh.

③ Refresh

Define compute engine

① Select Spark service

Spark service

Add

Add IBM Analytics for Apache Spark, then re Refresh.

② Refresh

Armand from IBM
Hi
Dummyuser231.ormium@spamgourme...

Cancel Create

- e. Select “Lite” and click on “Create”

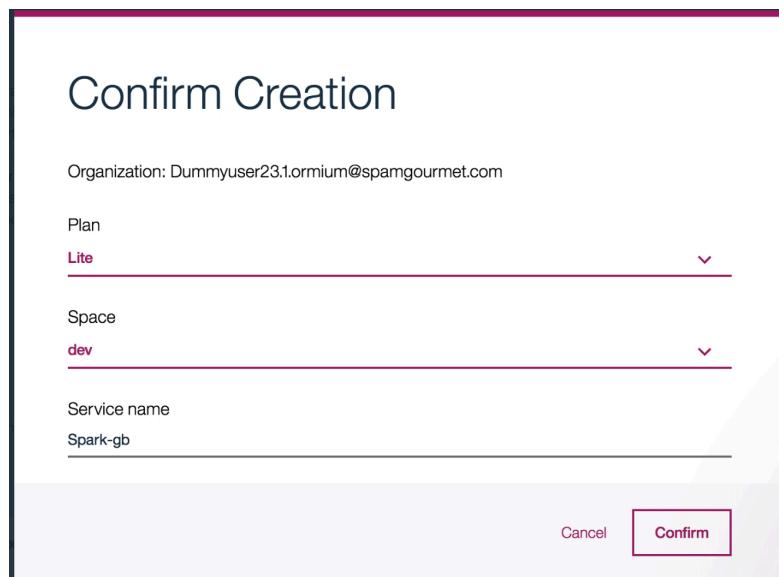
The screenshot shows the IBM Data Science Experience interface. At the top, there's a navigation bar with links for Chrome, File, Edit, View, History, Bookmarks, People, Window, Help, and a search bar. Below that is a secondary navigation bar with links for Projects, Tools, Data Services, and Community. On the right side, it shows 'US South' and some icons.

The main content area has tabs for 'Existing' and 'New'. The 'New' tab is selected, and under it, there's a section for 'Apache Spark'. This section contains a brief description of Apache Spark, followed by three columns of features:

- Incredibly Fast**: Apache Spark delivers 100x the performance of Apache Hadoop for certain workloads because of its advanced in-memory computing engine.
- Easy to Use and Powerful**: Apache Spark's Streaming and SQL programming models backed by MLlib and GraphX make it incredibly easy for developers and data scientists to build apps that exploit machine learning and graph analytics. Because the service is 100% compatible with Apache Spark, developers can build their apps and run them against the IBM managed service to benefit from operational, maintenance, and hardware excellence.
- Convenient Data Storage**: Object Storage enables a convenient way to upload your data from a file for immediate use by your Spark instance. You can set up Object Storage directly from the Spark service interface.

Below this, there's a 'Pricing Plan' section. It shows a table with three columns: 'Plan', 'Features', and 'Pricing'. The 'Plan' column has a radio button next to 'Lite'. The 'Features' column lists '2 Spark Executors'. The 'Pricing' column shows 'Free'. A red arrow points to the 'Lite' plan. At the bottom of the table, it says 'An entry level plan to run programs using up to 2 Spark executors'. There are also 'Terms' and 'Cancel' buttons at the bottom right.

- f. Again, make sure all fields are filled and click on “Confirm”



6. Please make sure that an object store and Apache Spark service have been deployed.
You might have to click on “refresh”, then click on “Confirm”

The screenshot shows the 'Create Project' wizard in Watson Studio. It consists of three main sections:

- Define project details:** Includes fields for 'Name' (set to 'default') and 'Description' (with placeholder 'Project description').
- Define storage:** Shows 'Select storage type' (Object Storage (Swift API) is selected), 'Target object storage instance' (set to 'object-storage-kp'), and 'Target container' (set to 'default').
- Define compute engine:** Shows 'Select Spark service' (Spark service is set to 'Spark-gb'), a warning about associating the same service with multiple projects, and a message 'Failed, please try again later'.

At the bottom right, there are 'Cancel' and 'Create' buttons, with 'Create' being highlighted by a red box. Two red arrows point from the 'Define storage' and 'Define compute engine' sections towards the 'Create' button.

7. Congratulations, your DSX setup is complete!

Select Organization and Space

Confirm your IBM Cloud organization and space information below.
[Or create new organization and space](#)

Select IBM Cloud account
Dummyuser231.ormium@spamgourmet.com ▾

IBM Cloud Organization
Dummyuser231.ormium@spamgourmet.com ▾

IBM Cloud Space
dev ▾

Continue

Fill out the following form. Please make sure you select “United States” for Country or Region.

The screenshot shows the 'Sign up for IBM Cloud' registration page. The left side features a blue sidebar with various promotional offers like 'Guaranteed free development with Lite plans' and 'Start on your projects right away'. The main form on the right has fields for Email*, First Name*, Last Name*, Company, and Country or Region*. A red arrow points to the 'Country or Region*' dropdown menu, which is set to 'United States'. There are also fields for Password* and checkboxes for 'Keep me informed of products, services, and offerings from IBM companies worldwide' and 'By email'. At the bottom right is a 'Privacy - Terms' link.

Sign up for an IBMid and create your IBM Cloud account
Build on IBM Cloud for free with no time restrictions

Guaranteed free development with Lite plans
Develop worry-free and at no cost with cap based Lite plan services for as long as you like.

Start on your projects right away
Skip entering your credit card info and get working in just a few short steps.

Get \$200 on us to try paid services
Ease into cloud pricing or try something new with \$200 in credit available for 1 month upon upgrade.

Ready to get started? Sign up today!

Email*
thisis@test.com

First Name*
myname

Last Name*
mylastname

Company

Country or Region*
United States

Password*
1qQ_dsdfsdfsdfsdfsf

Keep me informed of products, services, and offerings from IBM companies worldwide.
 By email

Privacy - Terms

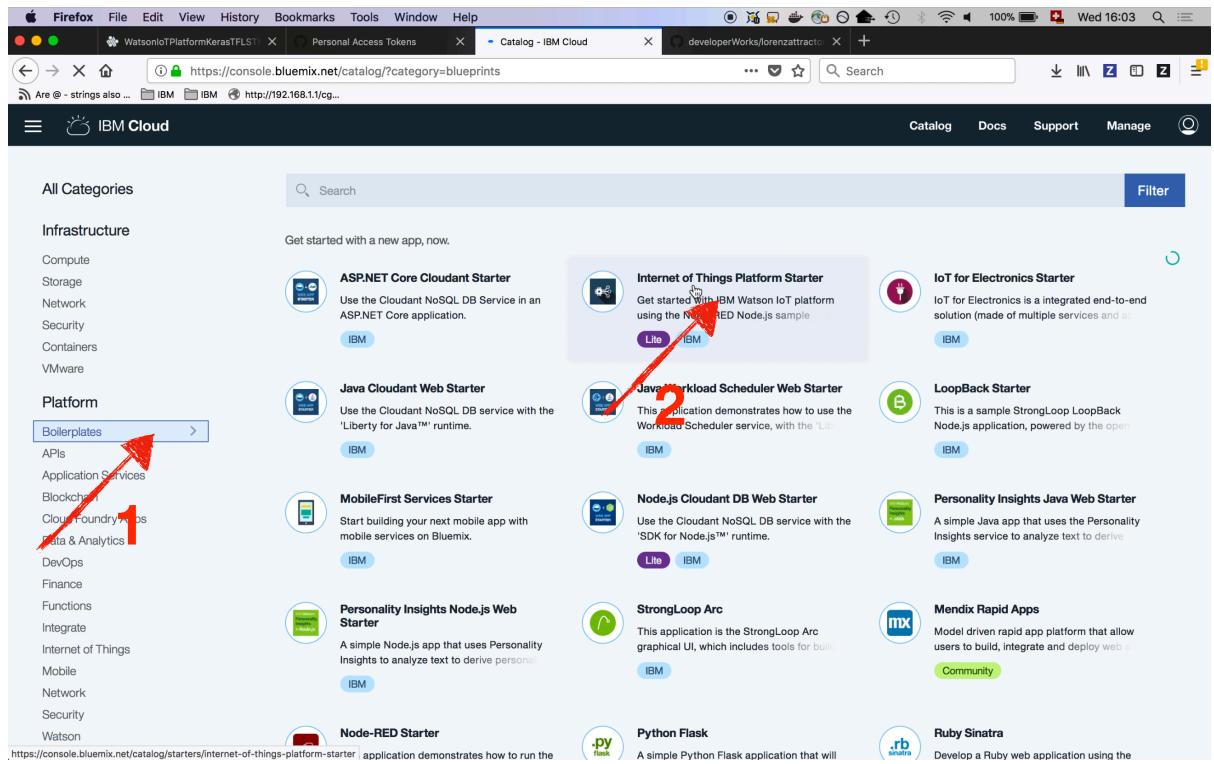
You'll receive an email for confirmation, please click on that link in that email and login.

2. Create the NodeRED Boilerplate

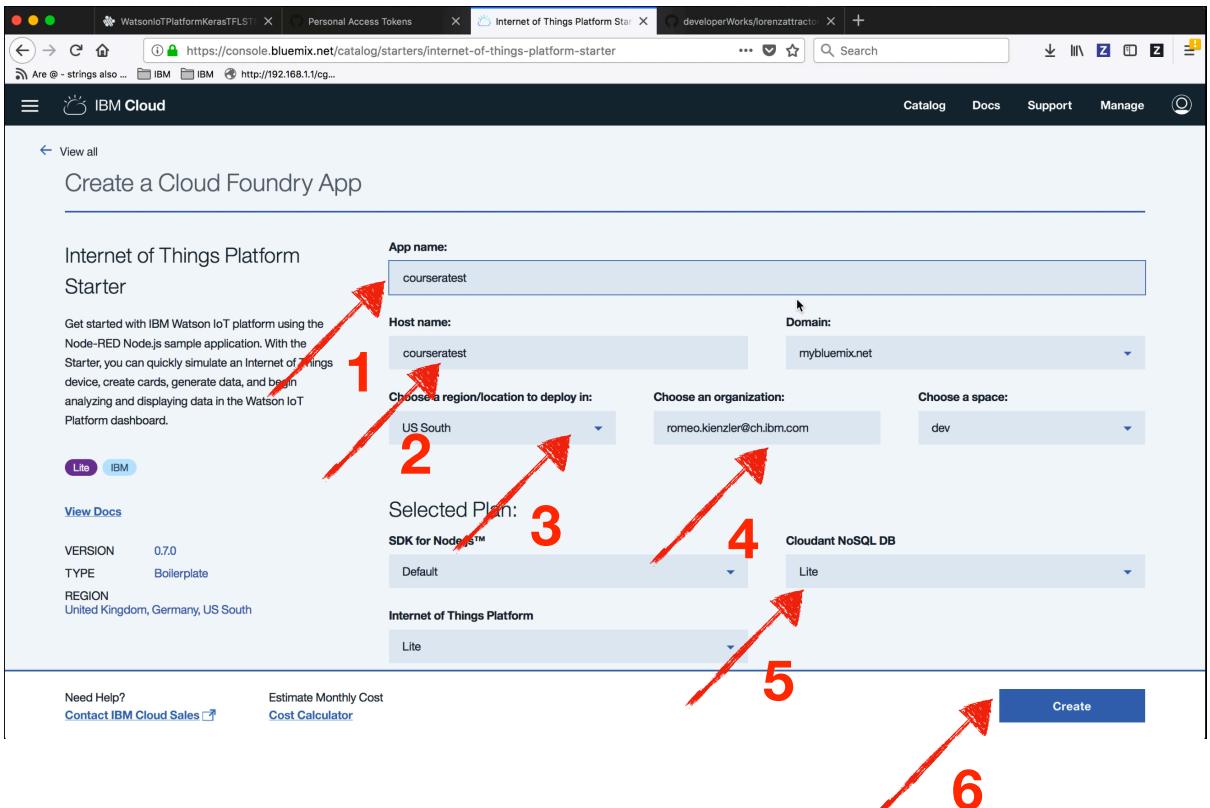
1. Click on “Create Resource”

The screenshot shows the IBM Cloud Dashboard. At the top, there are navigation links: Catalog, Docs, Support, Manage, and a user icon. Below the header, the dashboard is divided into sections: "Cloud Foundry Apps" and "Cloud Foundry Services". In the "Cloud Foundry Apps" section, there are two entries: "ccana" and "lorenz", both listed as "Running (1/1)". In the "Cloud Foundry Services" section, there are five entries: "Apache Spark-p4", "ccana-cloudantNoSQLDB", "lorenz-cloudantNoSQLDB", "lorenz-iotf-service", and "Object Storage-bt", all listed as "Lite". A prominent blue button labeled "Create resource" is located in the top right corner of the dashboard area. A large red arrow points from the bottom left towards this "Create resource" button.

2. Under “Boilerplates” click on “Internet of Things Platform Starter”



3. Create the Boilerplate

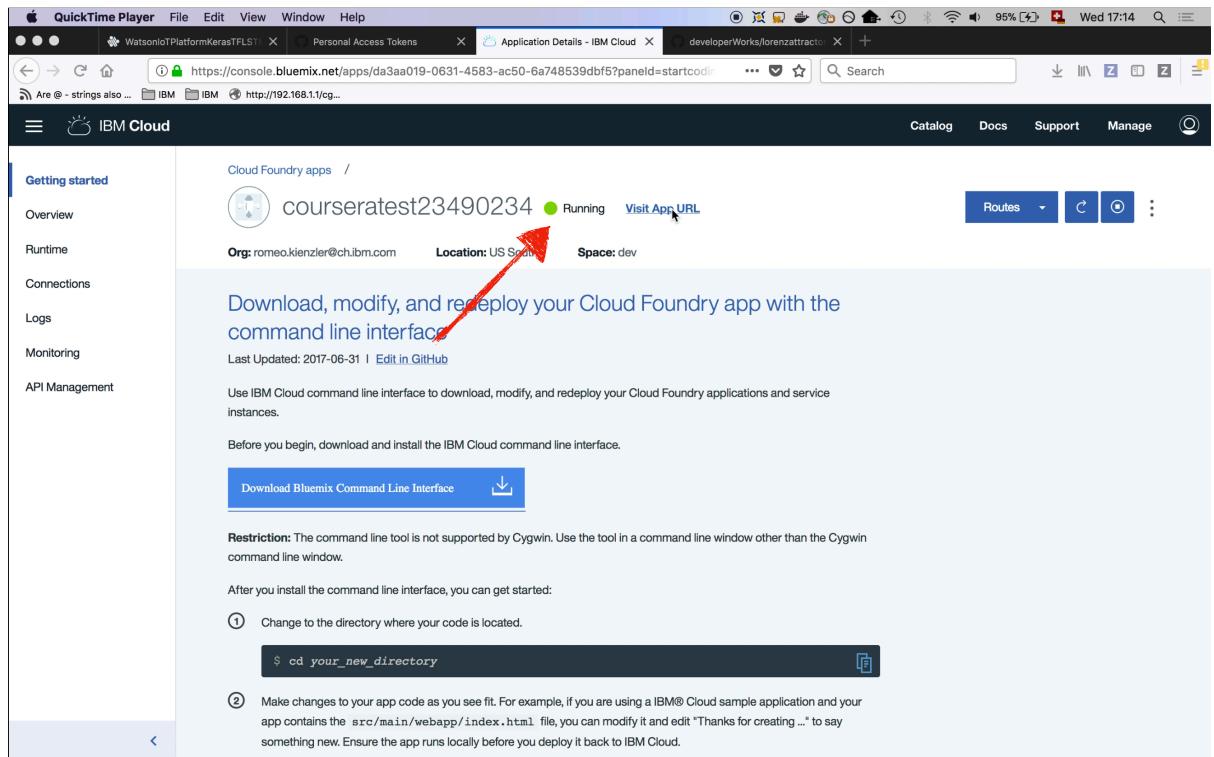


- Choose a unique name for your application (1)
- This name must be unique since it will be part of the URL (2)
In case it is not unique you'll receive an error message and can start over
- Please make sure you select "US South" as region (3)
- Please make sure Organization and Space is non-empty (4)
- Please make sure all drop downs are showing "Lite" as plan (5)
- Click on Create (6)

4. Wait until the app is started

The screenshot shows a Firefox browser window with several tabs open. The active tab is 'Application Details - IBM Cloud' at the URL <https://console.bluemix.net/apps/b1c34954-0f42-451c-8e29-8c51ed6bf778?panelId=startcodin...>. The main content area displays an application named 'courseratest23490234'. The status of the app is 'Starting'. A red arrow points from the text 'Wait until the app is started' in the previous step to this 'Starting' status indicator. The sidebar on the left is titled 'Getting started' and includes links for Overview, Runtime, Connections, Logs, Monitoring, and API Management.

5. Once the app has started click on “Visit App URL”



6. Setup your NodeRED instance with a user name and password

Secure your Node-RED editor

Secure your editor so only authorised users can access it

Username

Password good

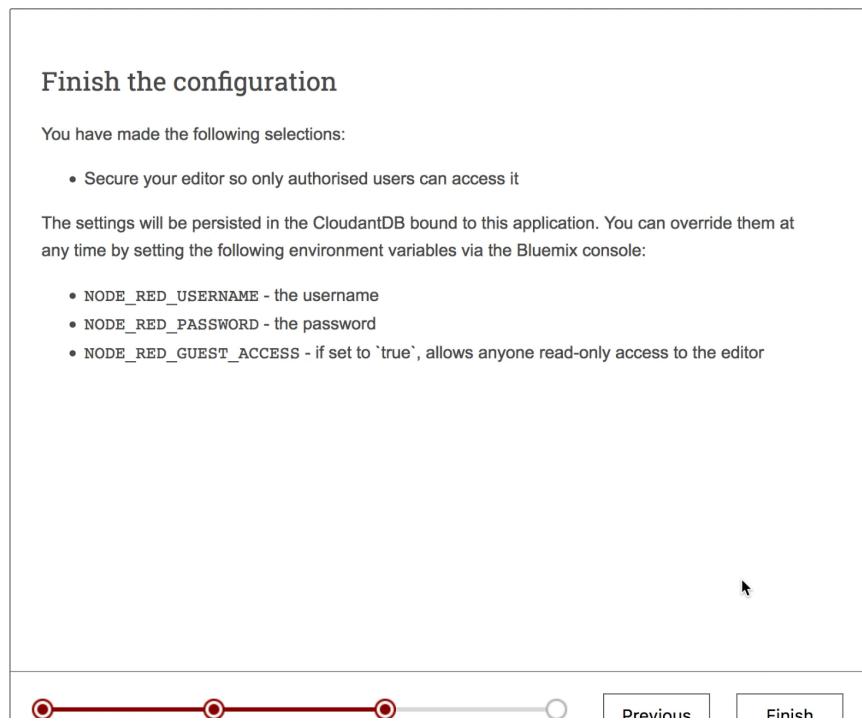
Allow anyone to view the editor, but not make any changes

Not recommended: Allow anyone to access the editor and make changes

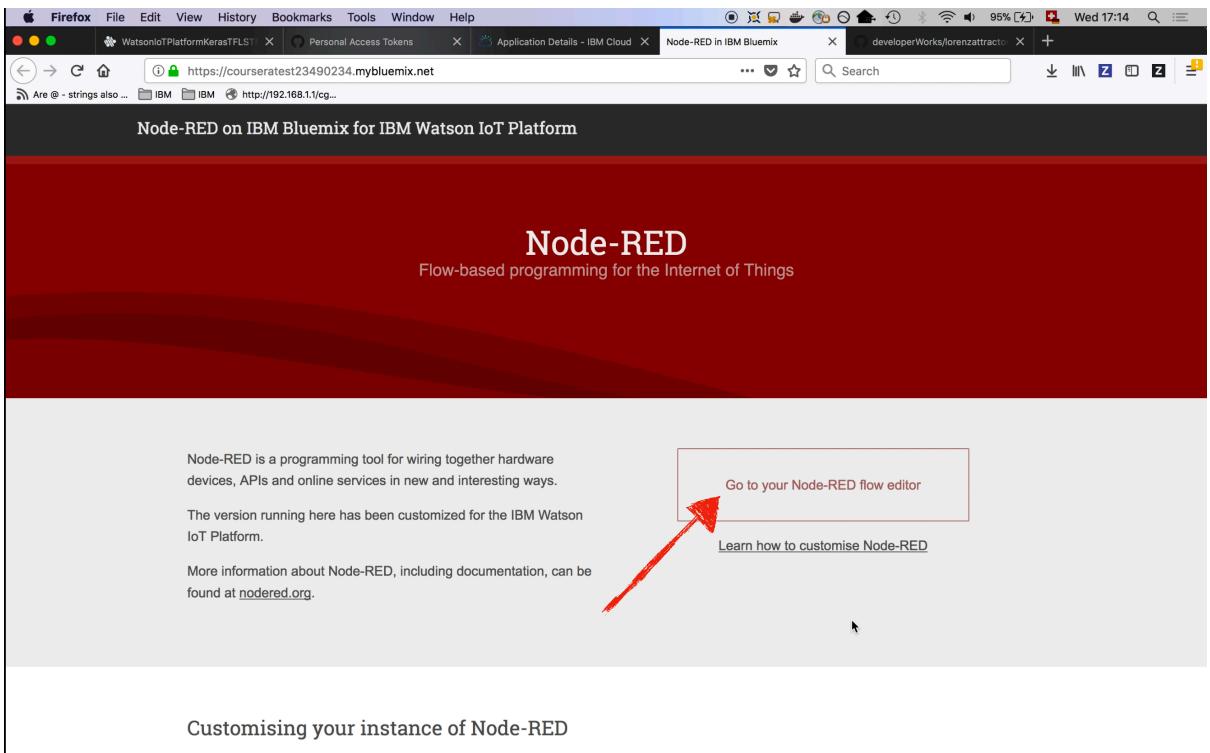
»

Previous Next

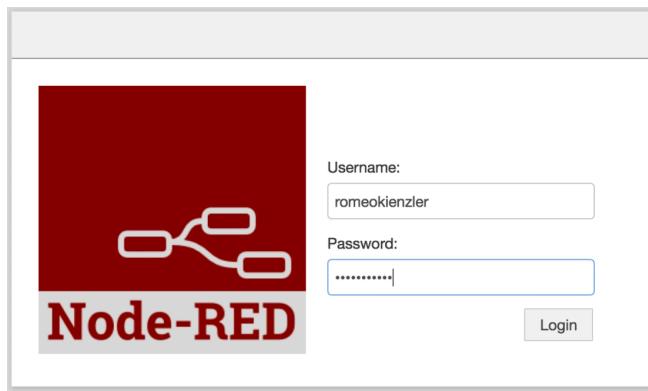
7. Click on “finish”



8. Click on “Go to your NodeRED flow editor”



9. Login with the credentials you've setup before



10. Congratulations, if you see the following screen, you are done!

