

Getting Started with Watson Studio

Watson Studio is a cloud environment for working with data using Jupyter Notebooks. These instructions will help you learn the basic operations within Watson Studio. The official documentation for Watson Studio is found [here](#).

Note: Your own display and the screenshots in this guide may have different resolutions, so you may have to scroll down to see options that are already visible in the screenshots. I have tried to highlight buttons in reds when necessary.

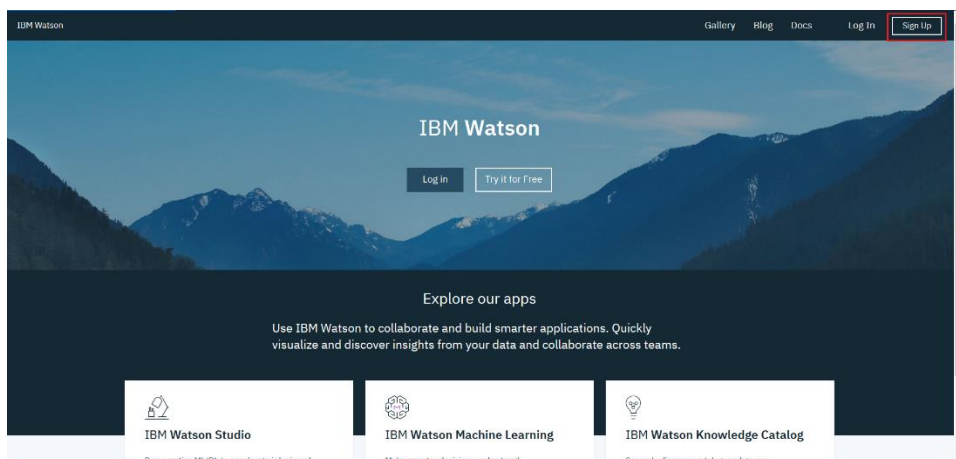
Note: Some of the screenshots may have names which vary from the directions, since this was adapted from a different course. Overall names do not have to strictly correspond to what these instructions say.

Table of Contents

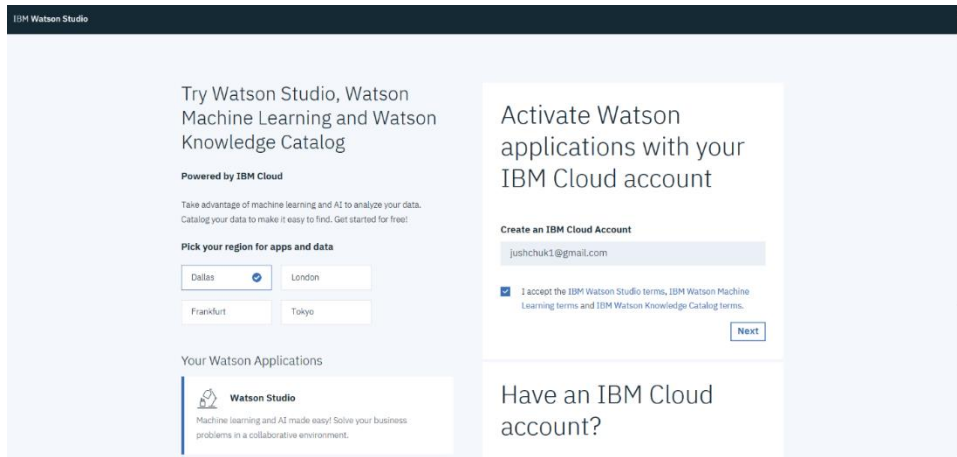
1. Creating a Watson Studio account
2. Creating a project
3. Creating a notebook
4. Importing data into a project
5. Using data in a notebook
6. Adding from gallery to project

1. Creating a Watson Studio Account

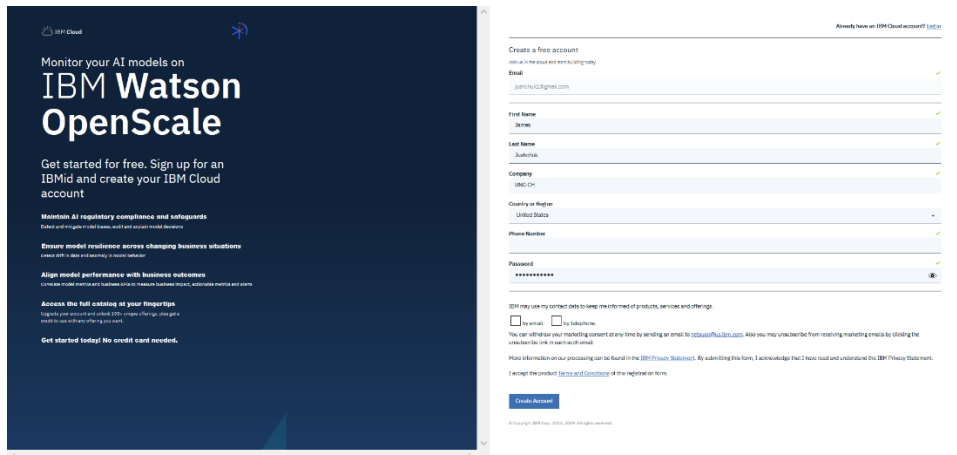
Visit <https://dataplatform.cloud.ibm.com/>, and click **Sign Up** in the top right.



Select **Dallas** region, provide your email, accept the terms and services and click **Next**.

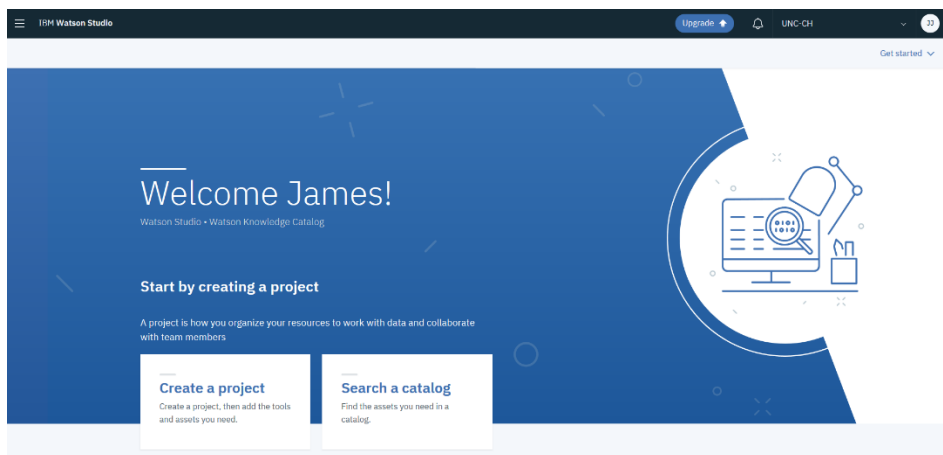


Next, fill in your information (for Company you can put ECU). Phone Number is optional. Uncheck the by email and by telephone marketing boxes. Click **Create account** and wait for a confirmation email.



In the confirmation email, click **Confirm account** and then log into Watson Studio again.

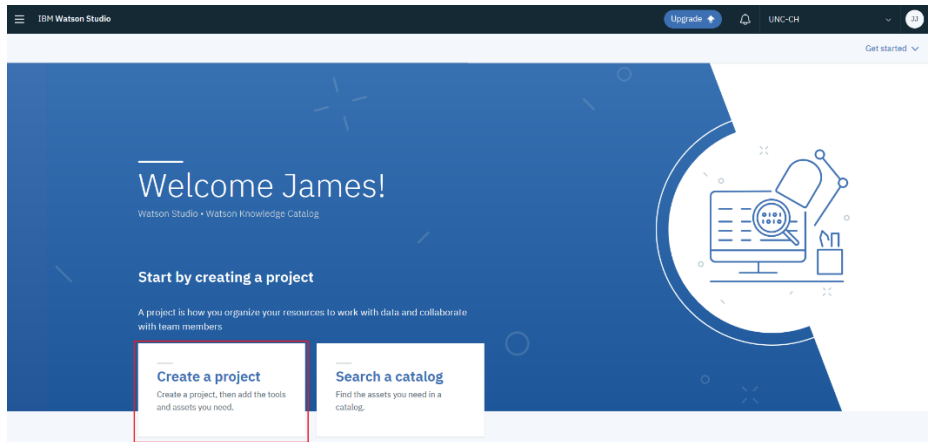
Once the Watson Apps are setup, click **Done** and you should be brought to the Watson Studio homepage. The **Start tour** popup is optional (if it appears). You now have your Watson Studio account.



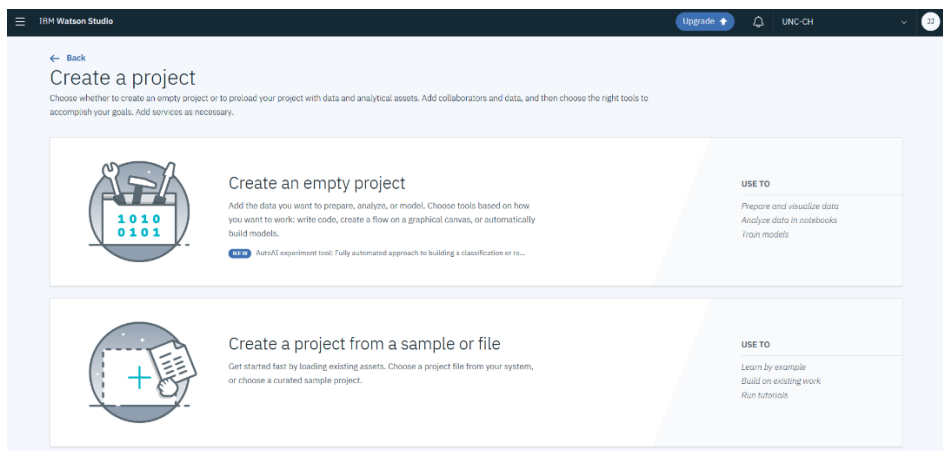
2. Creating a project

Projects are structured to organize your work and data. You can start by creating a single project.

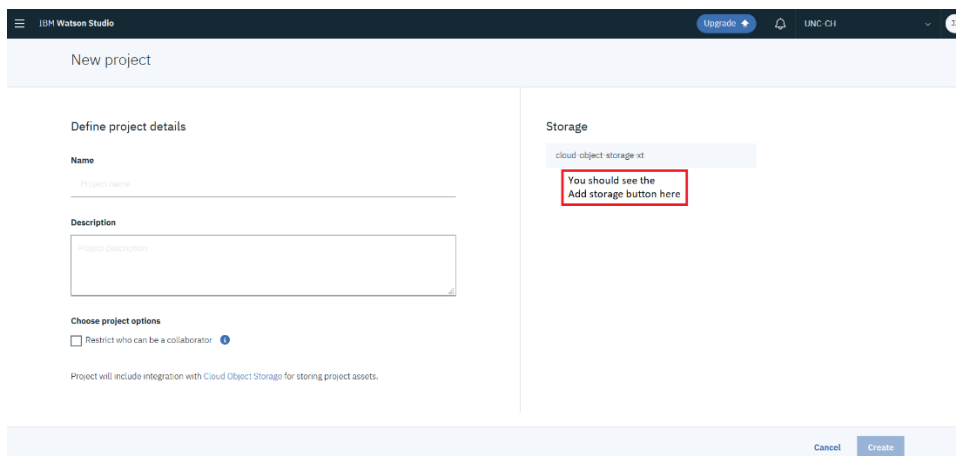
Visit <https://dataplatform.cloud.ibm.com/>, and login if you have not already.



Click **Create a project** and then select **Create an empty project**.



You need storage for this project which will be created by clicking **Add** below Storage.



This will take you to a new page which will create a Lite tier of Cloud Object Storage for you. This will be completely free and associated with your existing Watson Studio account. Keep all default options and click **Create** at the bottom.

The screenshot shows the IBM Watson Studio interface for creating Cloud Object Storage. The top navigation bar includes the IBM logo, 'IBM Watson Studio', an 'Upgrade' button, a notification bell, 'UNC-CH', and a user profile icon. Below the navigation bar, a message states: 'Pricing Plan: Monthly Process shown above reflect the: United States'. The main content area is a table with three columns: PLAN, FEATURES, and PRICING. The 'PLAN' column has two options: 'Lite' (selected with a radio button) and 'Standard'. The 'FEATURES' column for the 'Lite' plan lists: '1 COS Service Instance', 'Storage up to 25 GB/mo.', 'Up to 20,000 GET requests/mo.', 'Up to 2,000 PUT requests/mo.', 'Up to Data Retrieval 10 GB/mo.', 'Up to 5GB Public Outbound', and 'Applies to aggregate total across all storage bucket classes'. The 'PRICING' column for the 'Lite' plan shows 'Free'. A note at the bottom of the 'Lite' plan states: 'The Lite service plan for Cloud Object Storage includes Regional and Cross Regional resiliency, flexible data classes, and built in security.' Below the table, there is a note: 'There is no minimum fee, so you pay only for what you use.' and a link 'Expand each section to view details'. At the bottom right, there are 'Cancel' and 'Create' buttons.

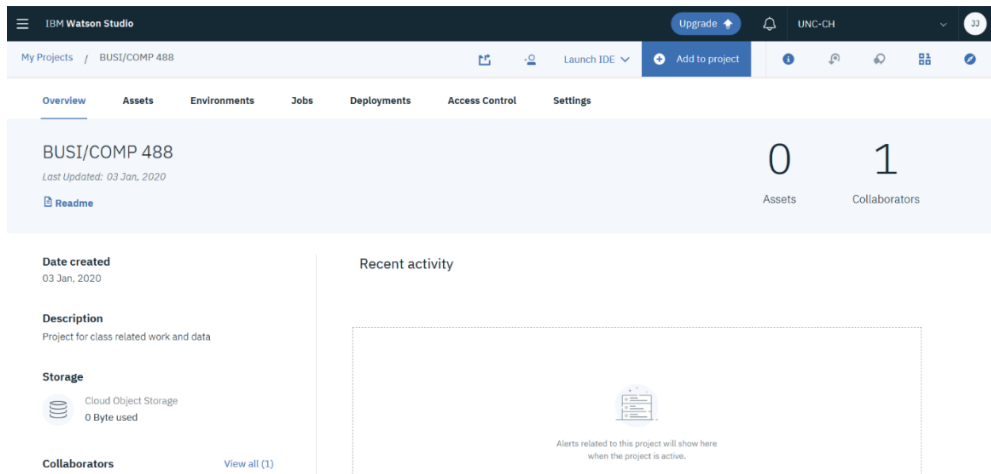
Again, keep all default options and click **Confirm** (your Service name may be different but that is ok).

The screenshot shows the 'Confirm Creation' dialog box overlaid on the pricing page. The dialog box has a title bar with a close button (X). It contains three sections: 'Plan' with a dropdown menu showing 'Lite', 'Resource group' with a dropdown menu showing 'Default', and 'Service name' with a text input field containing 'cloud-object-storage-xt'. At the bottom of the dialog box, there are 'Cancel' and 'Confirm' buttons.

Go back to the New Project page and refresh if your new storage does not appear automatically. Then fill in a name (such as NLP First Project) and an optional description. Make the Restrict who can collaborate box unchecked and click **Create**.

The screenshot shows the 'New project' page in IBM Watson Studio. The page has a header with the IBM logo, 'IBM Watson Studio', an 'Upgrade' button, a notification bell, 'UNC-CH', and a user profile icon. The main content area is divided into two sections: 'Define project details' and 'Storage'. The 'Define project details' section has a 'Name' field with the text 'BUSI/COMP 488 First Project', a 'Description' field with the text 'Project description', and a 'Choose project options' section with a checkbox labeled 'Restrict who can be a collaborator' which is unchecked. Below this, a note states: 'Project will include integration with Cloud Object Storage for storing project assets.' The 'Storage' section shows a dropdown menu with the text 'cloud-object-storage-xt'. At the bottom right, there are 'Cancel' and 'Create' buttons.

You now have your first project in Watson Studio. This is the main interface from which you will add data, create notebooks, and run them. You are free to create more projects (perhaps one for each assignment, etc.), but you do not have to.

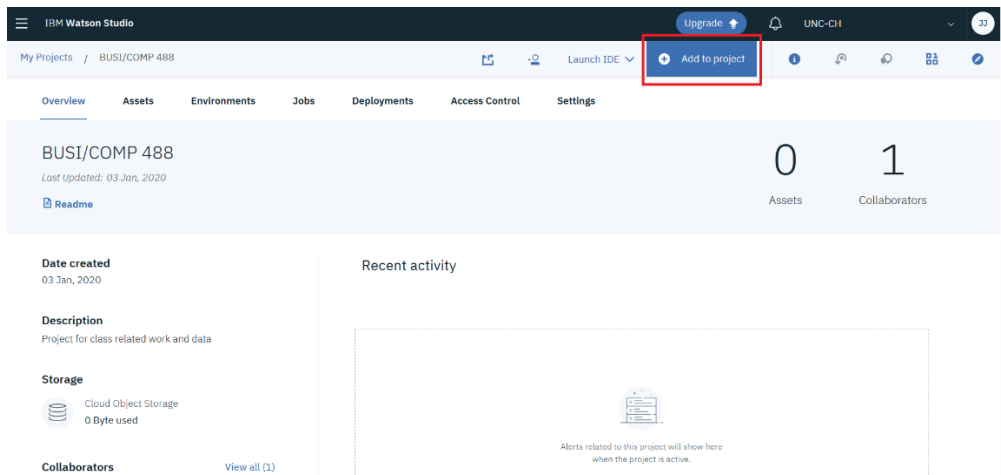


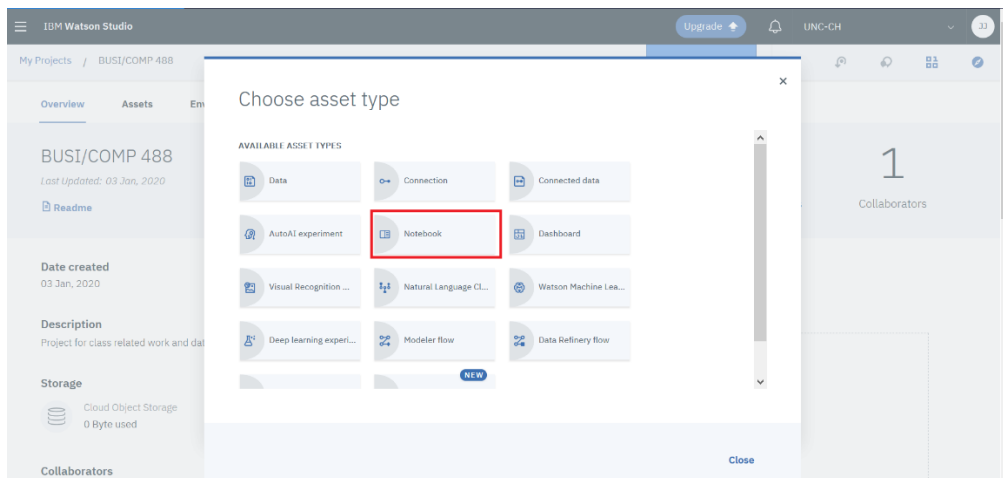
The official documentation on projects can be found [here](#).

3. Creating a notebook

Jupyter Notebooks are a programming environment that contains blocks of code, text, diagrams, or images. The code blocks can be run independent of one another, but variables and functions can be referenced between them.

To create your first notebook, go to the project and select **+ Add to project** and then **Notebook**. When adding anything to your project (data or notebooks) this is a consistent way to add things.

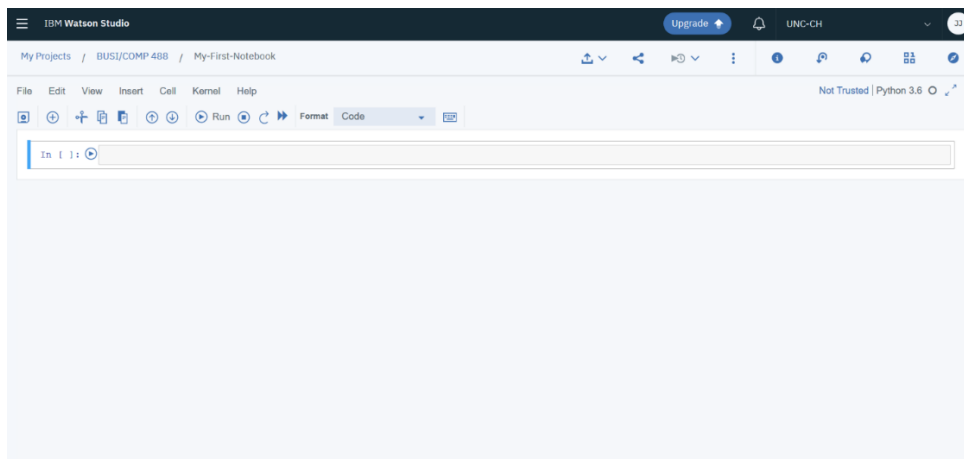




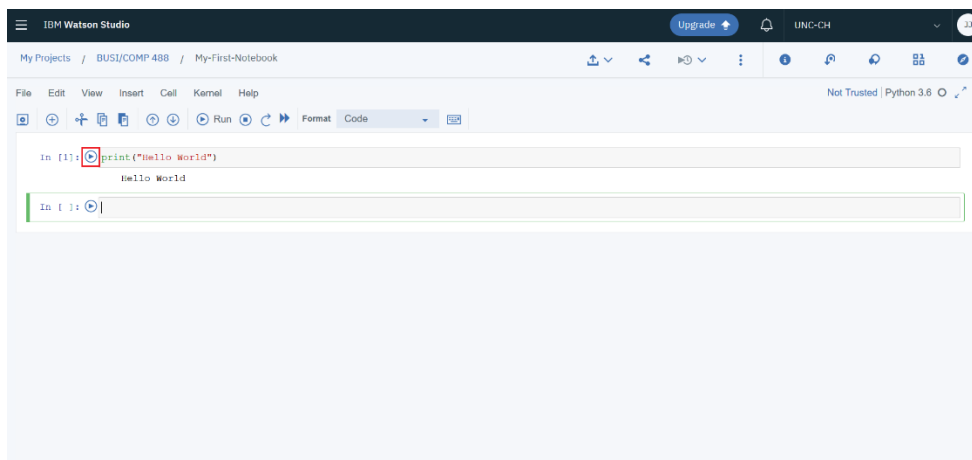
When creating a notebook, you can either create a new blank notebook, upload an existing notebook from your computer, or import a notebook from a URL. These options are available below New Notebook title. For now, create a blank notebook (later on you will import an existing notebook). Create a name, optional description, select a runtime of **Default Python 3.6 Free (1 vCPU and 4 GB RAM)**, and then click **Create Notebook**.

Note: Every notebook has a runtime environment associated with it. Unless otherwise specified, the runtime you select should be **Default Python 3.6 Free (1 vCPU and 4 GB RAM)**. Other runtimes will have more resources, but they consume usage credits called capacity unit hours. In the free Watson Studio account, you have a finite amount of these per month to use. If you were to accidentally use all your allotted credits, you will not be charged anything, but will simply have to use the free runtime. The free runtime is sufficient for most exercises.

Once you create the notebook, it should look similar to this:



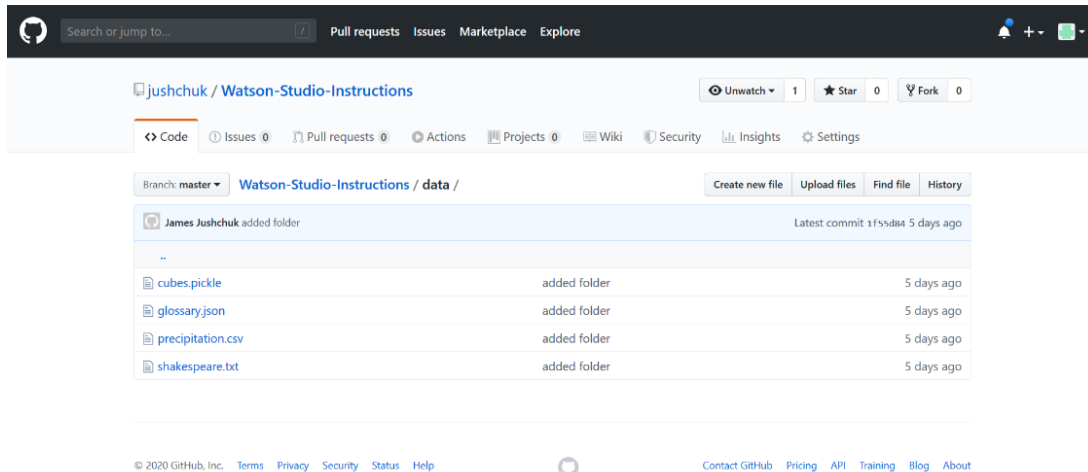
Click on the textbox next to `In []:` (this is called a code cell) and type `print("Hello World")`. Then click the play button (next to `In []:`) or press Shift+Enter. You should now see the output below the cell. Further on these instructions you will work more with Jupyter Notebooks.



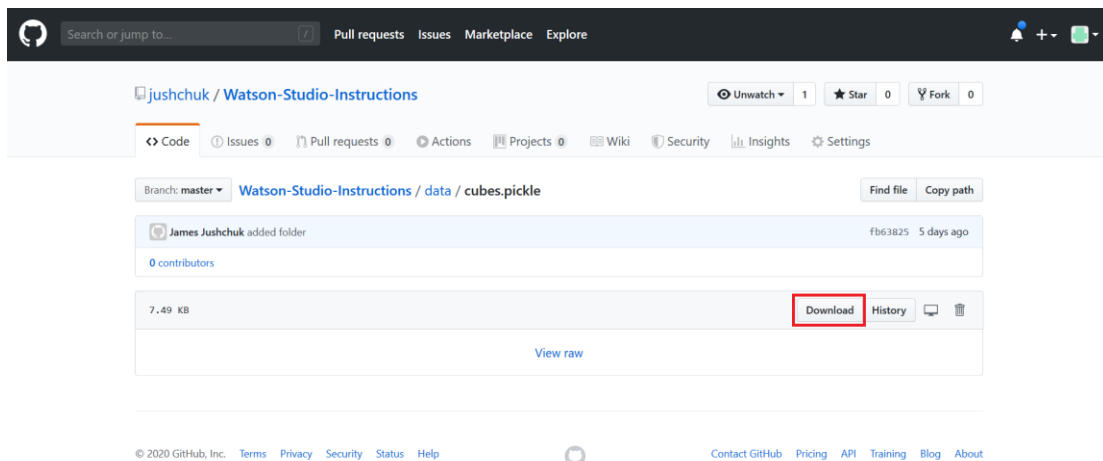
More information about working with notebooks can be found in the official documentation [here](#).

4. Importing data into a project

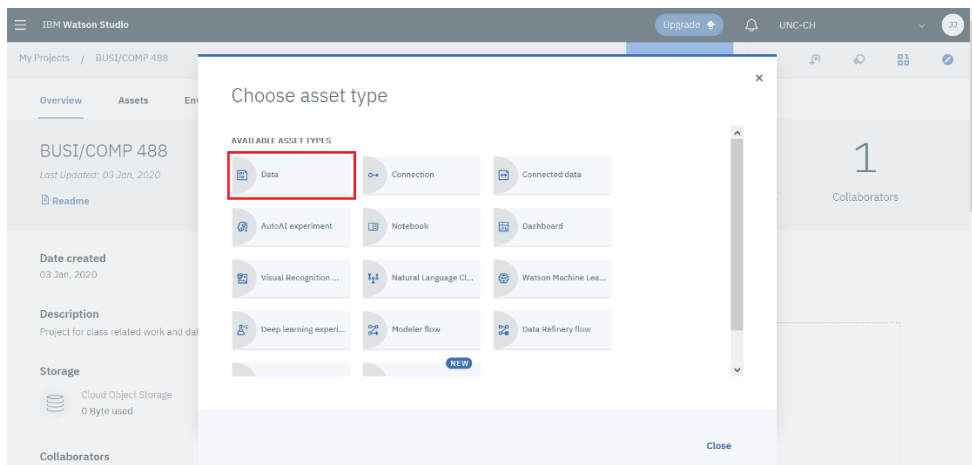
As a prerequisite for this section, visit <https://github.com/jushchuk/Watson-Studio-Instructions/tree/master/data>.



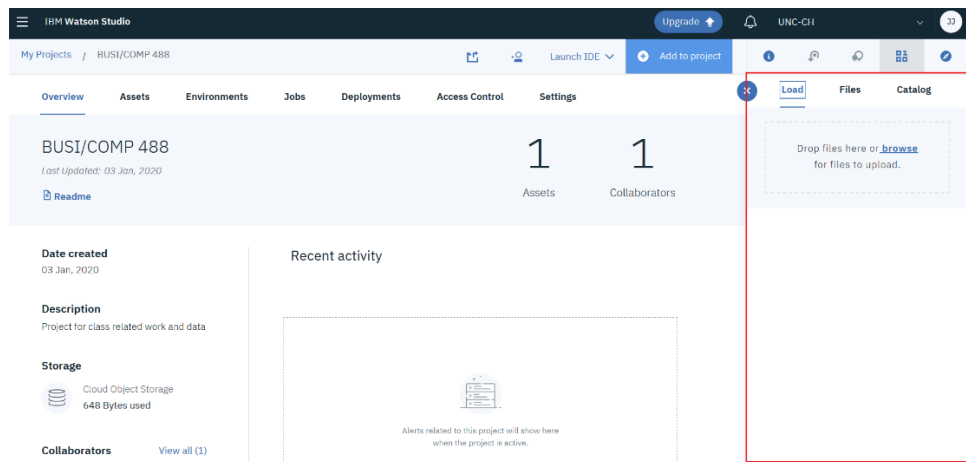
Click on and download the four separate files called cubes.pickle, glossary.json, precipitation.csv, and shakespeare.txt. The screenshot is for cubes.pickle.



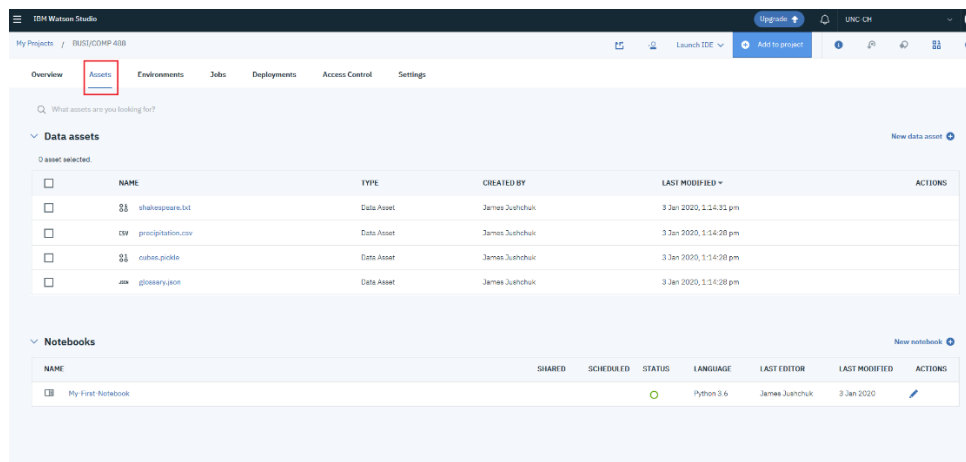
To add data to a project, go to the project and select **+ Add to project** and **Data**.



A pane on the right will appear that allows you to add data from your local machine. Use the **Load** option and **browse** your local machine for any number of files. Select these four files you previously downloaded to load them as data assets.

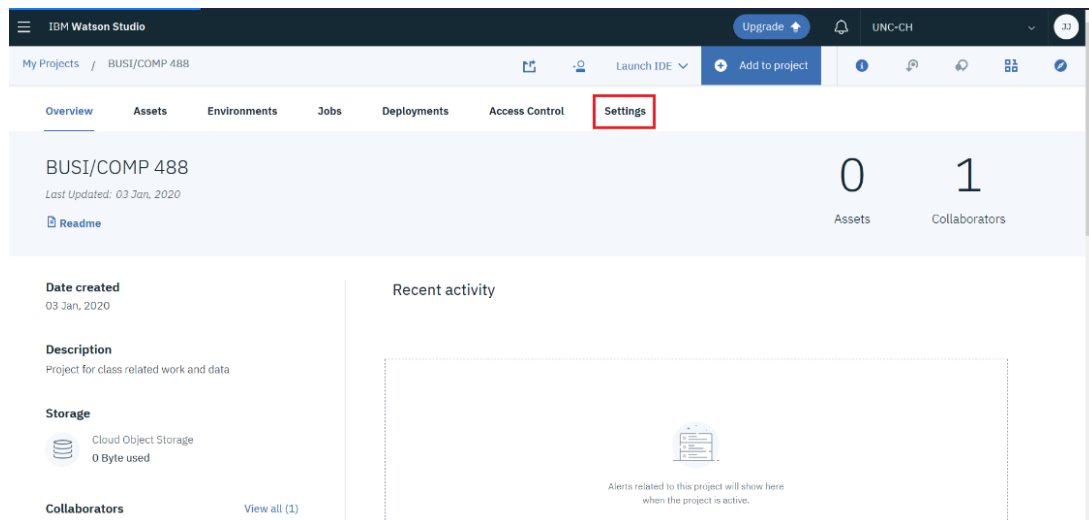


If you click on the **Assets** tab on the project page, you should see your new data assets along with any notebooks you have created.

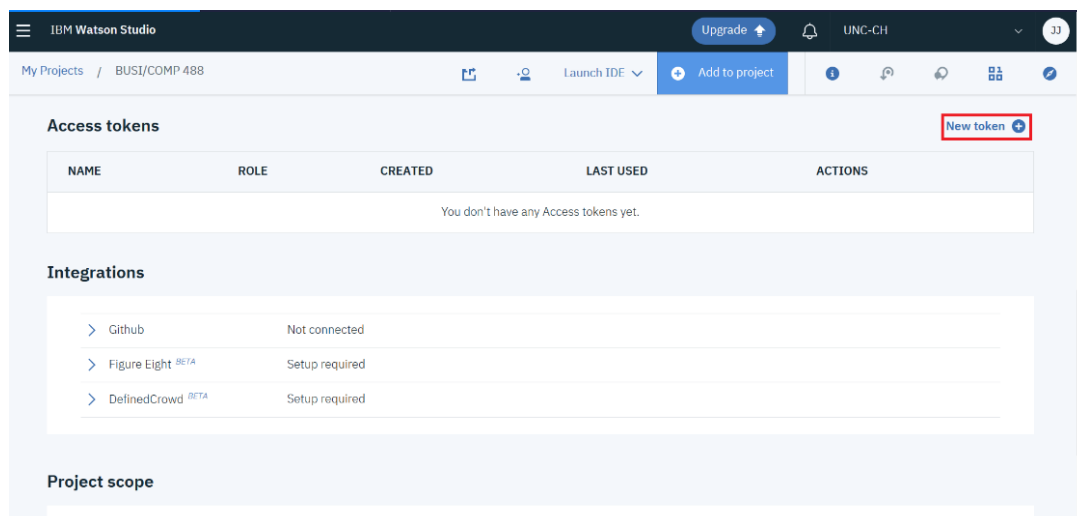


5. Using data in a notebook

Data that is added to a project can then be used in notebooks in the same project. This requires an access token associated with the project, which you must create. To do this, go to **Settings** tab in the project and scroll down to **Access Token**.

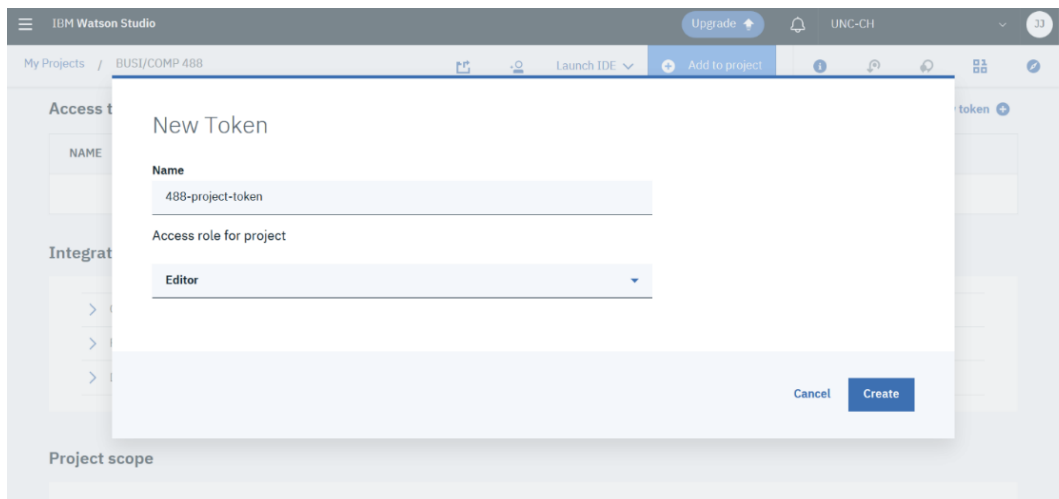


Click **New token +**, which will open up a popup with more options.



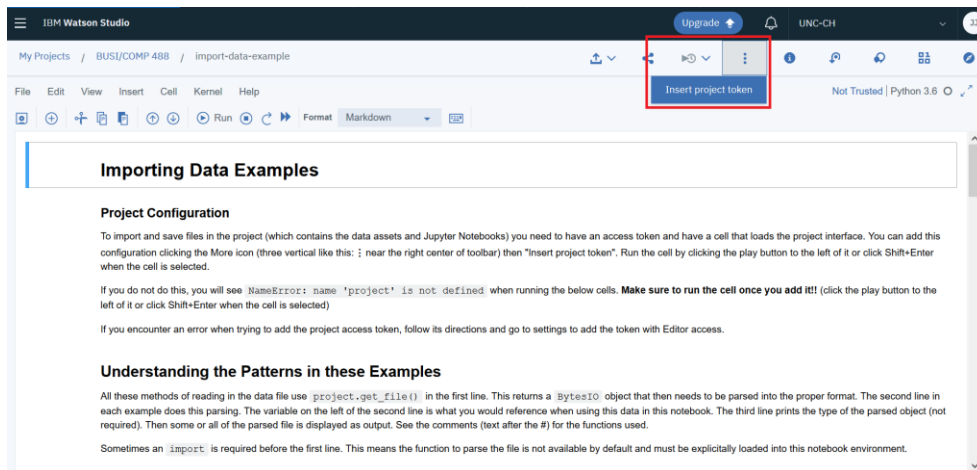
Give it a name like `nlp-project-token` and **Editor** access role and then click **Create**. This will allow you to not only read in data from your project in notebooks, but also create or edit data assets in your project.

Note: You will have to perform this step of creating an access token for every new project you create.



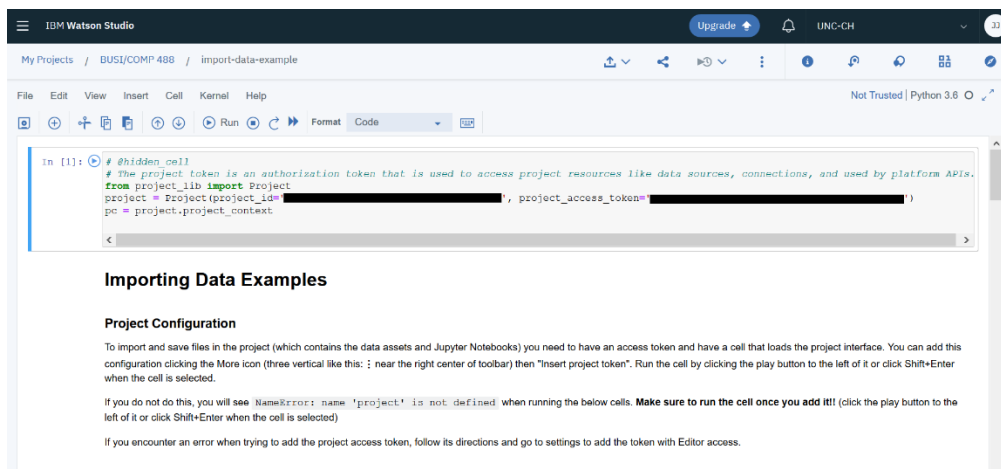
Now, you will import an existing notebook that will give examples on how to read in data and use in a notebook. The notebook is located at <https://raw.githubusercontent.com/jushchuk/Watson-Studio-Instructions/master/import-data-examples.ipynb>. In the project, click **+ Add to project** and then **From URL**. Fill in a name, optional description, runtime of **Default Python 3.6 Free (1 vCPU and 4 GB RAM)**, and URL.

Once the notebook is loaded, the project access token must be added to the notebook, so it can access data assets in your project. Click on the **three vertical dots** button and then **Insert Project Token**.



If you have followed the steps prior to this step, a new code cell will be created at the top of the notebook. Make sure to run this newly create cell by clicking the play button or pressing Shift+Enter.

Note: This project token cell contains credentials that allow access to your project data assets. Keep this cell information private as anyone with these credentials could access your project data assets.



Go through the examples to make sure they run without error. Don't worry about the technical details of how they each work, it is just important that you can load the data. Feel free to use this notebook to mess around with the data you import. Note that what you do to the data won't be saved beyond this notebook session unless you save it as a data asset as the last examples shows. The content of the example data you imported and anything you create is not relevant to the course, but they serve as simple examples for how to do imports / exports that you can use as a guide throughout this course.

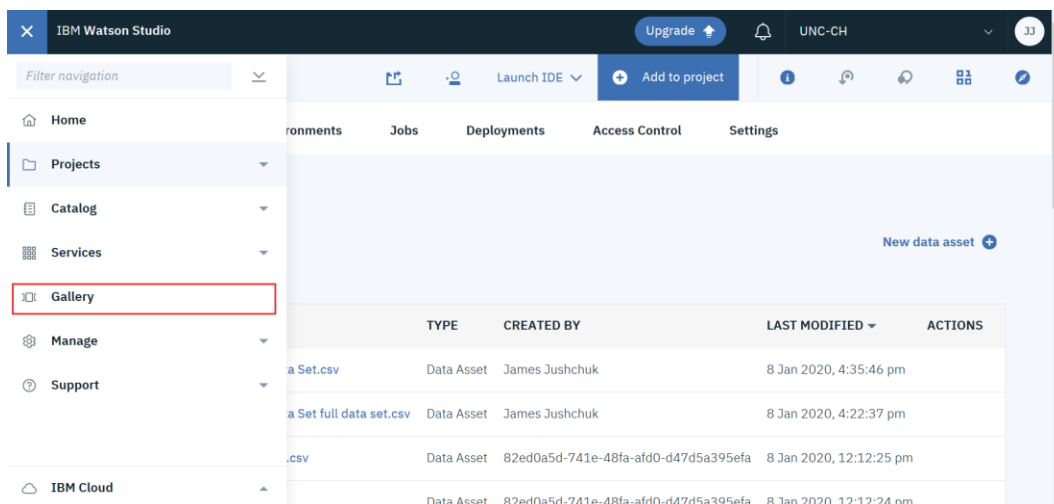
The documentation for loading and accessing data can be found [here](#).

6. Adding from Gallery to Project

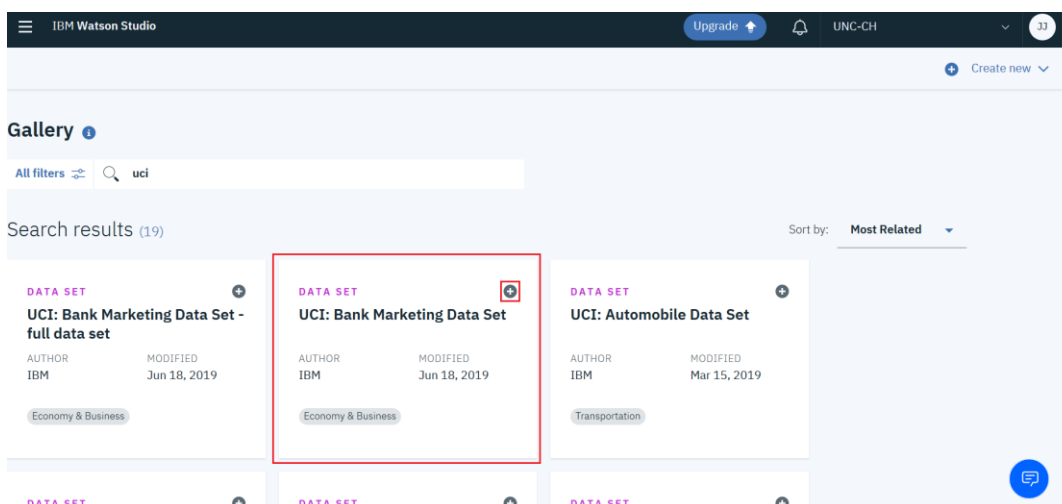
This short instruction will show you how to add a dataset (or notebook) to your project from the Gallery in IBM Watson Studio. OPTIONAL: If you want to try the notebook titled **Python-Introduction.ipynb** (found at <https://raw.githubusercontent.com/jushchuk/Watson-Studio-Instructions/master/Python-Introduction.ipynb>), then you should follow these instructions to add the UCI Banking data set to your project.

Gallery is collection of datasets and notebooks create by IBM to jumpstart or supplement your projects. You can search through the gallery and easily add data or notebooks to your projects.

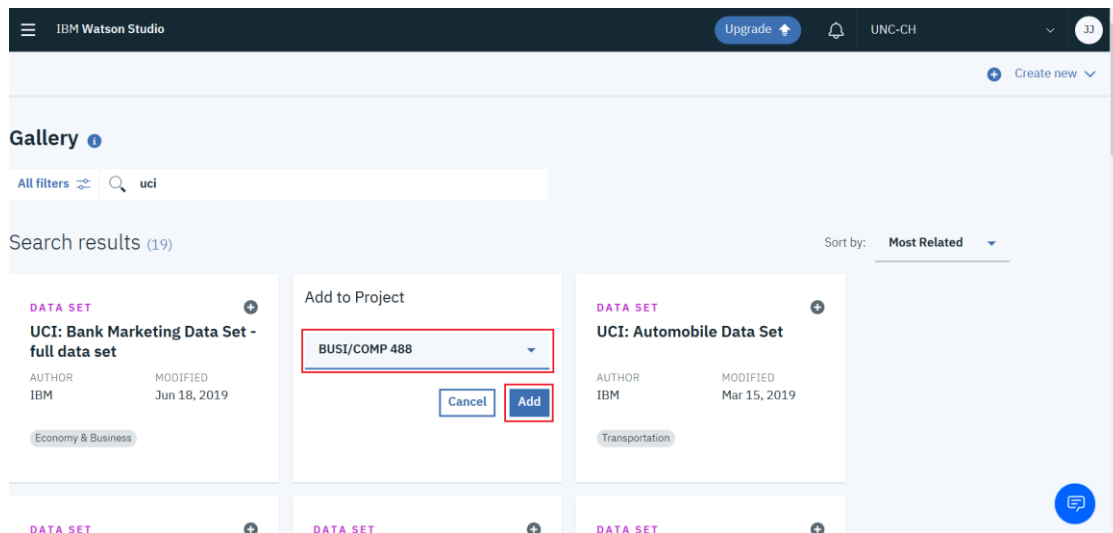
First step is to access the gallery inside Watson Studio. Go to the expanded menu in the top left of your Watson Studio project, then click **Gallery**.



From here, use the search to find a data set called "UCI: Bank Marketing Data Set". This should be the second result when searching for "UCI". Click on the + to choose which project to add the data set to.



Select the project of your choice and then click **Add**.



When you return to your project and go to the Assets tab, you should see the data set **UCI Bank Marketing Data Set.csv** under data assets. Now this data asset can be used in this project's notebooks.

Feel free to explore the gallery and add other notebooks and datasets to your projects.

Adding Gallery Data Set Documentation can be found [here](#).

Note: Notebooks autosave every minute or so, however, if a notebook is left idle for a period of time with no activity, the environment is automatically stopped. The code you have written will be saved, but environment variables will not persist between sessions. This means when you resume work after the environment is stopped, you will have to re-execute any code cells to repopulate the environment with any packages, functions, or variables. If you have a variable that would be slow or difficult to reproduce by simply rerunning the notebook cells, save it as a data asset in the form of a pickle file before you let the notebook go idle. Then reload it at the start of your next session and continue working with it.