

Lecture 2: Build your working environment Part II

COSC 526: Introduction to Data Mining



THE UNIVERSITY OF
TENNESSEE
KNOXVILLE

Instructors:



Michela Taufer

Assistants to the Instructor:



Ian Lumsden



Nigel Tan



Kae Suarez



Paula Olaya



Leo Valera

Course Repository



Clone our Class Repository

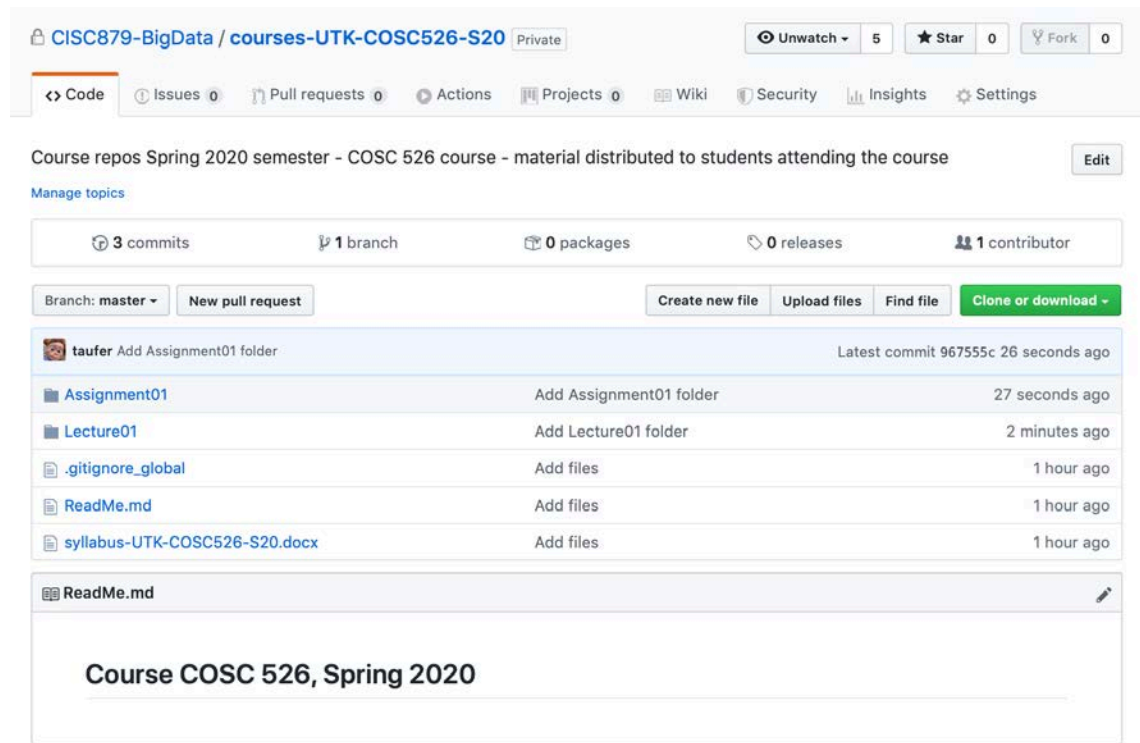
- Using either the GitHub Desktop app or the git CLI, clone the class repo:

<https://github.com/CISC879-BigData/2021S-COSC526>



Clone the course repository

<https://github.com/CISC879-BigData/courses-UTK-COSC526-S21>



CISC879-BigData / **courses-UTK-COSC526-S20** Private

Unwatch 5 Star 0 Fork 0

Code Issues 0 Pull requests 0 Actions Projects 0 Wiki Security Insights Settings

Course repos Spring 2020 semester - COSC 526 course - material distributed to students attending the course [Edit](#)

[Manage topics](#)

3 commits 1 branch 0 packages 0 releases 1 contributor

Branch: master New pull request Create new file Upload files Find file Clone or download

tauer Add Assignment01 folder Latest commit 967555c 26 seconds ago

File	Commit	Time
Assignment01	Add Assignment01 folder	27 seconds ago
Lecture01	Add Lecture01 folder	2 minutes ago
.gitignore_global	Add files	1 hour ago
ReadMe.md	Add files	1 hour ago
syllabus-UTK-COSC526-S20.docx	Add files	1 hour ago

ReadMe.md

Course COSC 526, Spring 2020

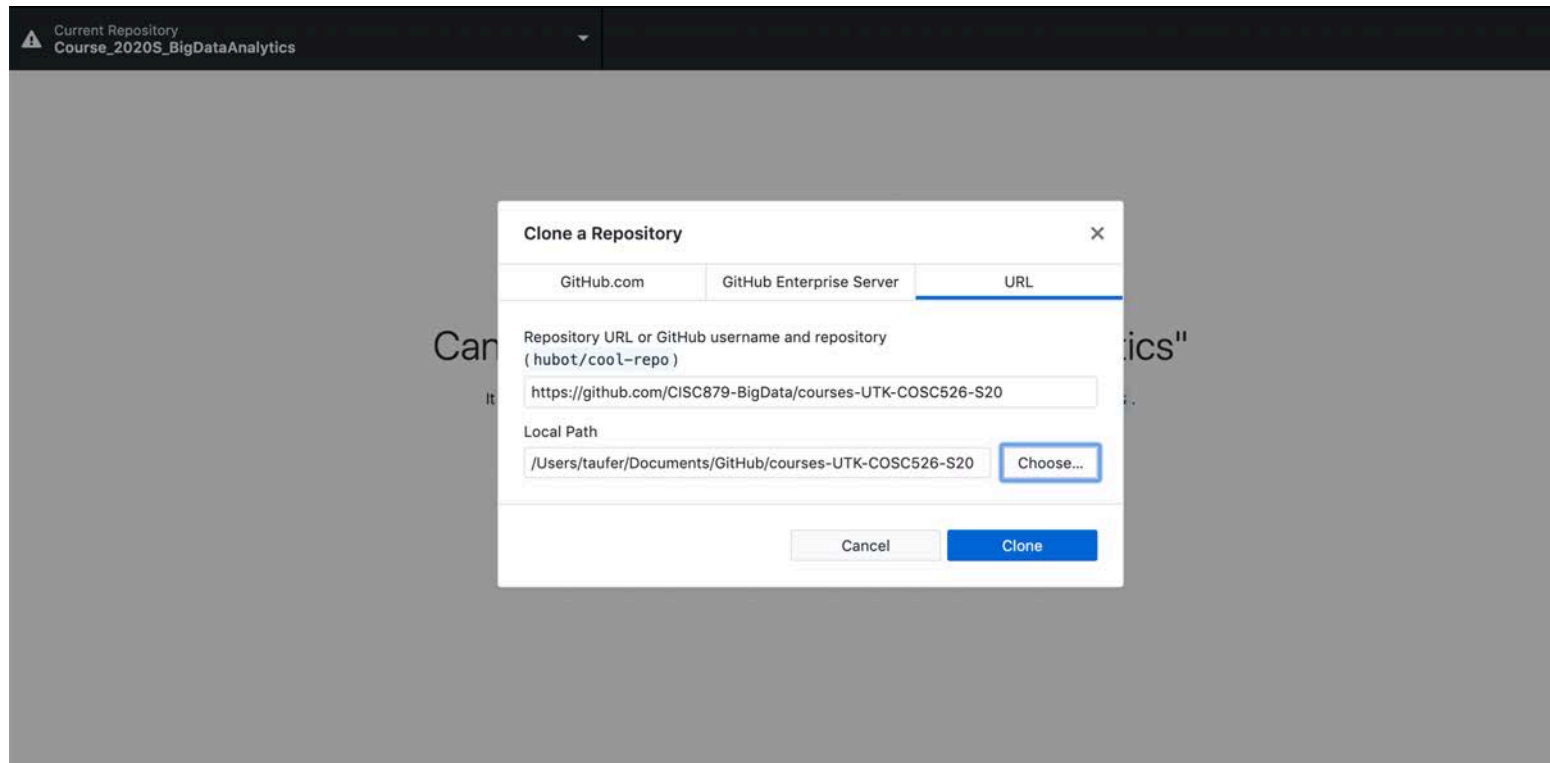
Clone the course repository

The screenshot shows a GitHub repository page for 'CISC879-BigData / courses-UTK-COSC526-S20'. The repository is private and has 5 watches, 0 stars, and 0 forks. The 'Code' tab is selected, showing the repository description: 'Course repos Spring 2020 semester - COSC 526 course - material distributed to students attending the course'. Below the description, there are statistics: 3 commits, 1 branch, 0 packages, 0 releases, and 1 contributor. A 'Clone or download' button is visible. A dropdown menu is open, showing options to 'Clone with HTTPS' (using the URL 'https://github.com/CISC879-BigData/courses-UTK-COSC526-S20'), 'Open in Desktop', and 'Download ZIP'. The repository contains several files and folders: 'Assignment01' (Add Assignment01 folder), 'Lecture01' (Add Lecture01 folder), '.gitignore_global' (Add files), 'ReadMe.md' (Add files), and 'syllabus-UTK-COSC526-S20.docx' (Add files). The 'ReadMe.md' file is expanded, showing the text 'Course COSC 526, Spring 2020'.

Open from CLI

```
$ git clone https://github.com/CISC879-BigData/courses-UTK-COSC526-S21.git
```


Open from GitHub Desktop



Get Ready for Lecture 2


- Get your GitHub repository ready for Assignment

🔗 master ▾ 🔗 1 branch 🏷 0 tags Go to file Add file ▾ 📄 Code ▾

 **taufer** Rename files

5532335 4 days ago ⌚ 3 commits

📁 Assignment01	Add lecture and assignment 1	4 days ago
📁 Lecture01	Rename files	4 days ago
📄 README.md	Initial commit	last month

README.md 

courses-UTK-COSC526-S21

Python, Anaconda, Jupiter Notebook



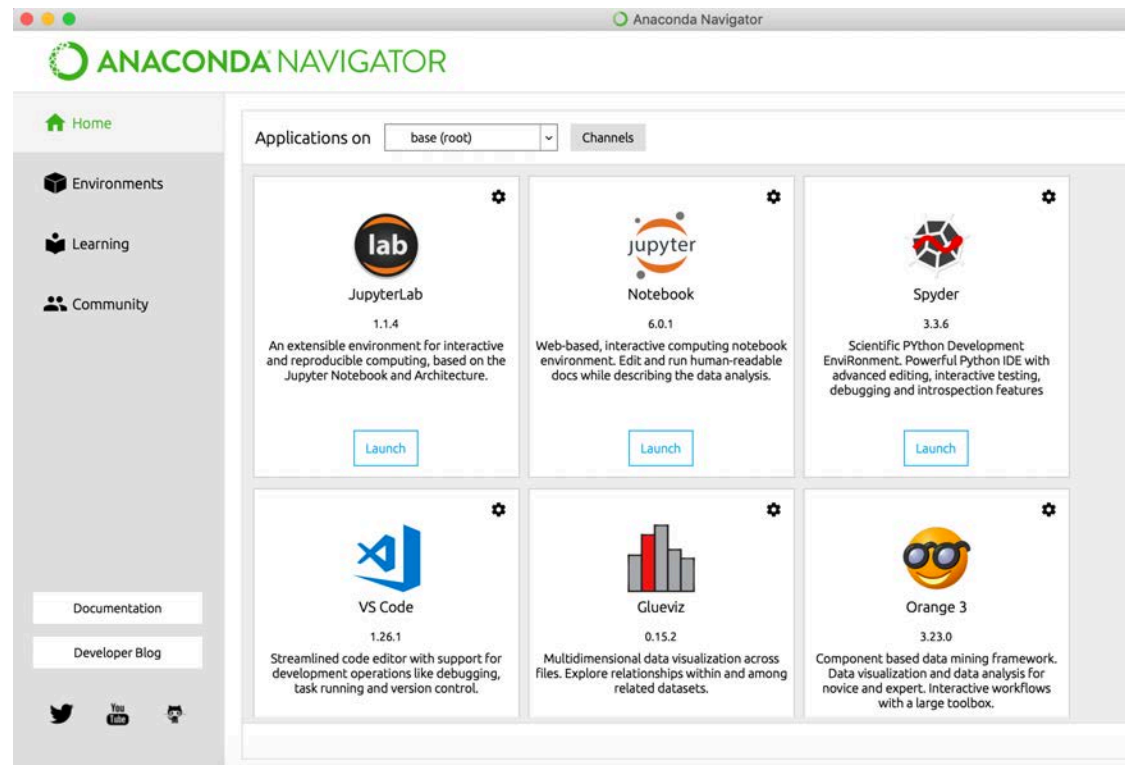
Python and Anaconda

- **Python:** It is Python 3.7!
- **Anaconda:** Python distribution
 - Include many popular packages by default, including Jupyter notebook
 - Make installing additional packages easy
- Your task:
 - Install Anaconda and launch the Jupyter notebook

Install Anaconda

- To install Anaconda (Distribution, not Enterprise or Cloud), follow the installation instructions for your operating system:
- Windows
<https://docs.anaconda.com/anaconda/install/windows/>
- Mac
<https://docs.anaconda.com/anaconda/install/mac-os/>
- Linux
<https://docs.anaconda.com/anaconda/install/linux/>

Anaconda Navigator



Jupyter Notebook

- **Jupyter notebook:** Our notebook for data analytics
- Programming in a browser
 - Create code in a cell – code in edit mode
 - Run code in a cell – code in command mode
 - Write text before and after code cells – markdown
- Your task:
 - Launch Jupyter from Anaconda GUI (or command line)



Open your assignment directory

Search or jump to... Pull requests Issues Marketplace Explore

CISC879-BigData / 2018F-CS594-CS690 Private

Unwatch 5 Star 0 Fork 0

<> Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

Branch: master 2018F-CS594-CS690 / Assignment01 / Create new file Upload files Find file History

imnasnainaec StartHere edits. Latest commit 7873682 3 days ago

..		
images	StartHere edits.	3 days ago
Assignment01.ipynb	Initial commit	4 days ago
StartHere.ipynb	StartHere edits.	3 days ago
StartHere.pdf	StartHere edits.	3 days ago
data.csv	Initial commit	4 days ago
data.tsv	Initial commit	4 days ago

© 2018 GitHub, Inc. Terms Privacy Security Status Help Contact GitHub Pricing API Training Blog About

Open your assignment directory


[Quit](#)[Logout](#)[Files](#)[Running](#)[Clusters](#)

Select items to perform actions on them.

[Upload](#)[New ▾](#)

<input type="checkbox"/> 0 ▾ / 00_git_repos / 2018F-CS594-CS690 / Assignment01			Name ▾	Last Modified	File size
<input type="checkbox"/>	..			seconds ago	
<input type="checkbox"/>	images			3 hours ago	
<input type="checkbox"/>	Assignment01.ipynb	Running		3 hours ago	11.1 kB
<input type="checkbox"/>	StartHere.ipynb	Running		an hour ago	6.87 kB
<input type="checkbox"/>	data.csv			3 hours ago	1.06 kB
<input type="checkbox"/>	data.tsv			3 hours ago	1.05 kB
<input type="checkbox"/>	StartHere.pdf			3 hours ago	289 kB

Create your own notebook

 jupyter

QuitLogout

FilesRunningClusters

Select items to perform actions on them.

UploadNew↺

☐ 0 ▾ / 00_git_repos / 2018F-CS594-CS690 / Assignment01

Name ▾

<input type="checkbox"/>	..		
<input type="checkbox"/>	images		
<input type="checkbox"/>	Assignment01.ipynb	Running	
<input type="checkbox"/>	StartHere.ipynb	Running	
<input type="checkbox"/>	data.csv		3 hours ago 1.06 kB
<input type="checkbox"/>	data.tsv		3 hours ago 1.05 kB
<input type="checkbox"/>	StartHere.pdf		3 hours ago 289 kB

Notebook:

Python 3

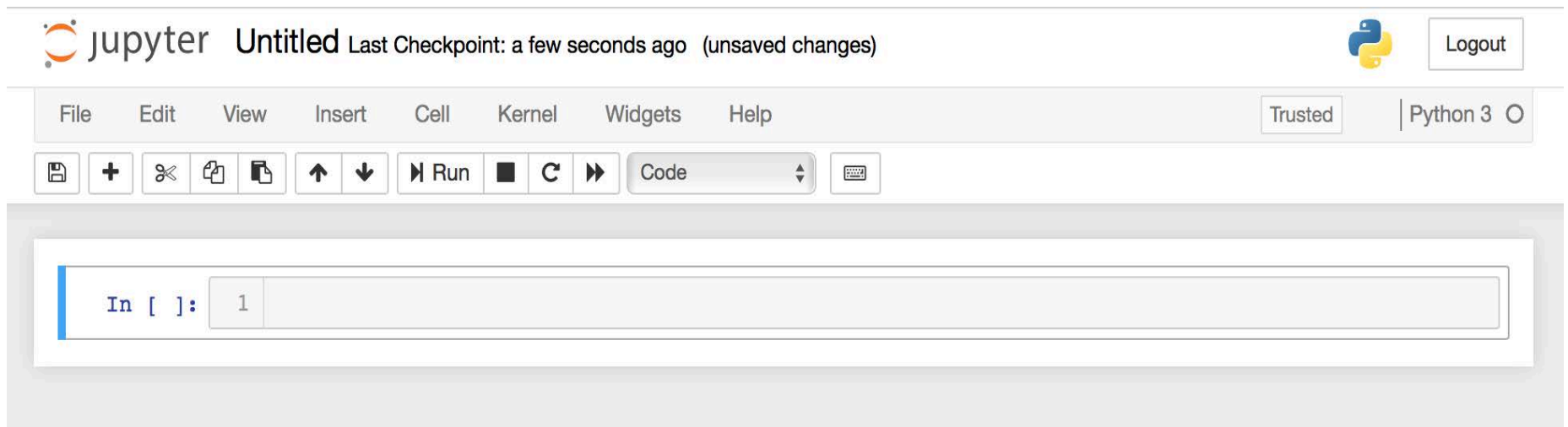
Other:

Text File

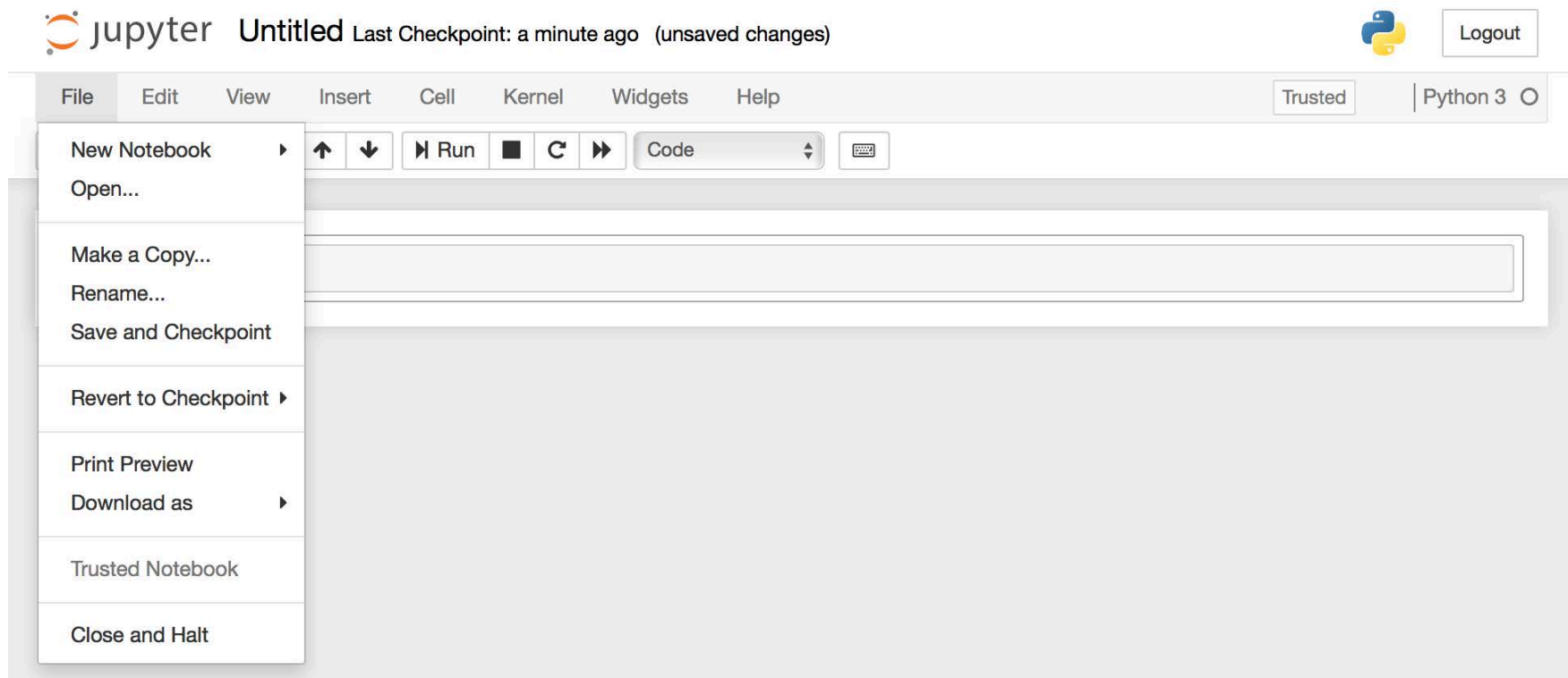
Folder

Terminal

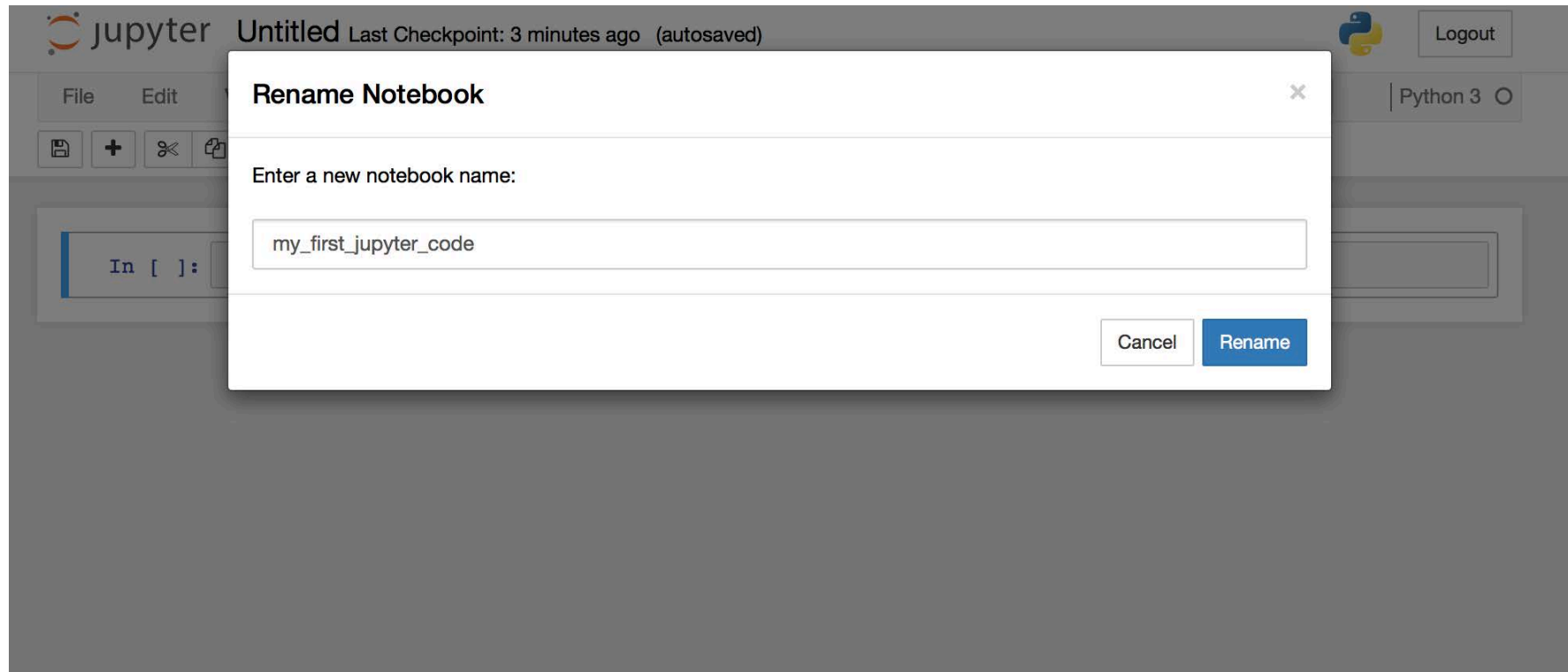
Create your own notebook



Create your own notebook



Rename the file



Create your own notebook

 jupyter

QuitLogout

FilesRunningClusters

Select items to perform actions on them.

UploadNew ↕

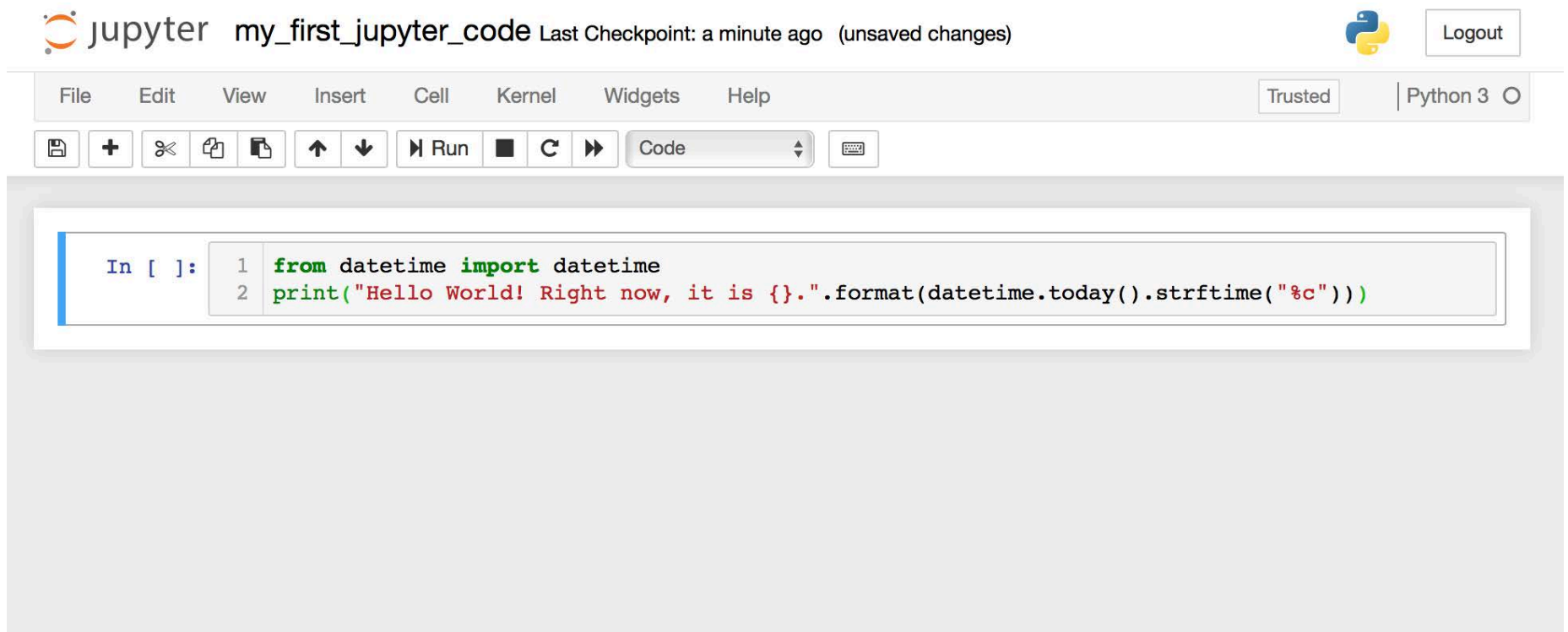
☐ 0 ▾

/ 00_git_repos / 2018F-CS594-CS690 / Assignment01

Name ▾Last ModifiedFile size

<input type="checkbox"/> ..		seconds ago	
<input type="checkbox"/> images		3 hours ago	
<input type="checkbox"/> Assignment01.ipynb	Running	3 hours ago	11.1 kB
<input type="checkbox"/> my_first_jupyter_code.ipynb	Running	seconds ago	555 B
<input type="checkbox"/> StartHere.ipynb	Running	an hour ago	6.87 kB
<input type="checkbox"/> data.csv		3 hours ago	1.06 kB
<input type="checkbox"/> data.tsv		3 hours ago	1.05 kB
<input type="checkbox"/> StartHere.pdf		3 hours ago	289 kB


Create a code cell



The screenshot displays the JupyterLab web interface. At the top, the header shows the Jupyter logo, the text "jupyter my_first_jupyter_code", and a status message "Last Checkpoint: a minute ago (unsaved changes)". To the right of the header is a Python logo and a "Logout" button. Below the header is a menu bar with options: File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. To the right of the menu bar are buttons for "Trusted" and "Python 3". Below the menu bar is a toolbar with icons for saving, creating a new file, cutting, copying, pasting, undo, redo, running the cell, and a dropdown menu currently set to "Code". The main area contains a single code cell with the following Python code:

```
In [ ]: 1 from datetime import datetime
        2 print("Hello World! Right now, it is {}".format(datetime.today().strftime("%c")))
```

Run your code code

jupyter my_first_jupyter_code Last Checkpoint: 2 minutes ago (unsaved changes)  Logout

File Edit View Insert Cell Kernel Widgets Help Trusted | Python 3

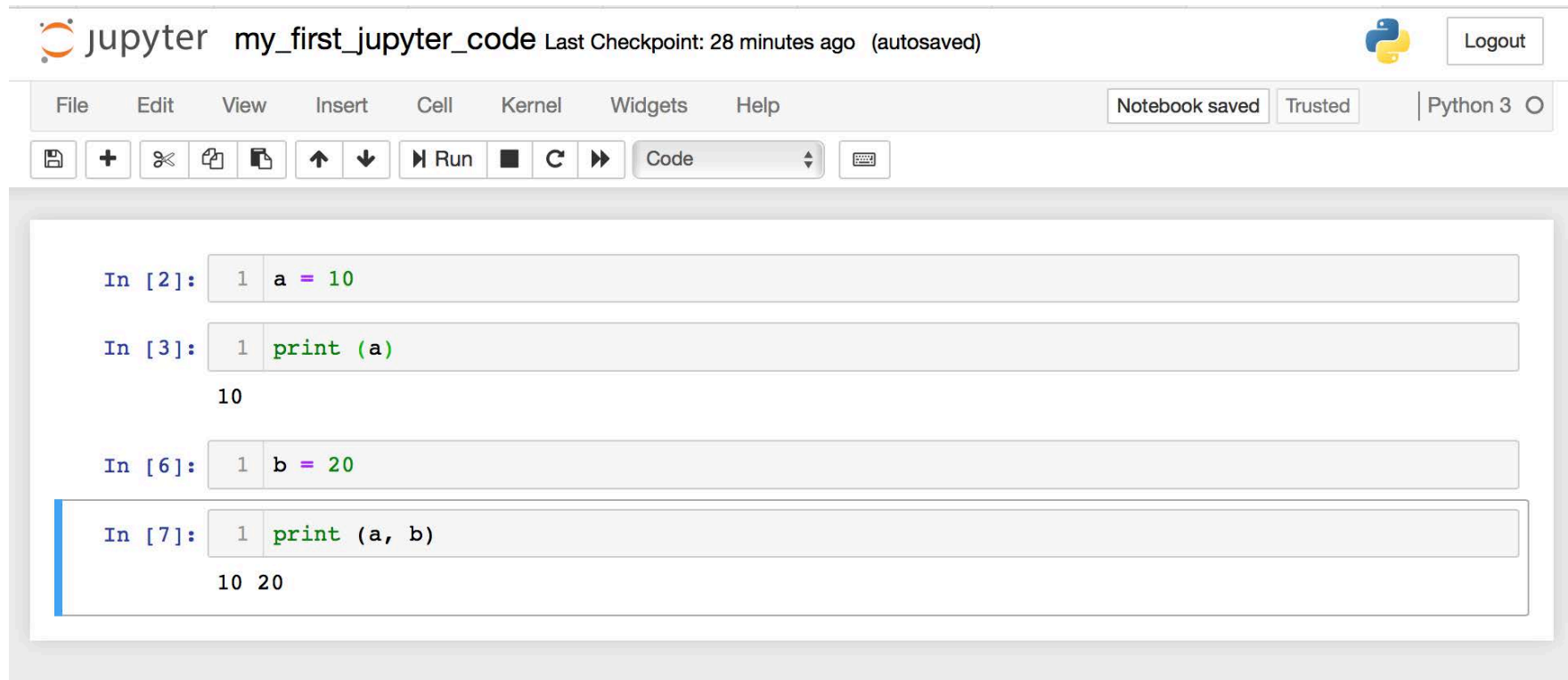
Save Add Split Cell Up Down Run Stop Restart Code Keyboard


```
In [1]: 1 from datetime import datetime
        2 print("Hello World! Right now, it is {}".format(datetime.today().strftime("%c")))

Hello World! Right now, it is Sat Aug 25 13:47:49 2018.
```

```
In [ ]: 1
```

Propagations: from cell to cell



jupyter my_first_jupyter_code Last Checkpoint: 28 minutes ago (autosaved)  Logout

File Edit View Insert Cell Kernel Widgets Help Notebook saved Trusted Python 3

⏏ + 🔍 📄 ⬆ ⬆ ⏏ Run ⏏ ⏏ Code

```
In [2]: 1 a = 10
```

```
In [3]: 1 print (a)
```

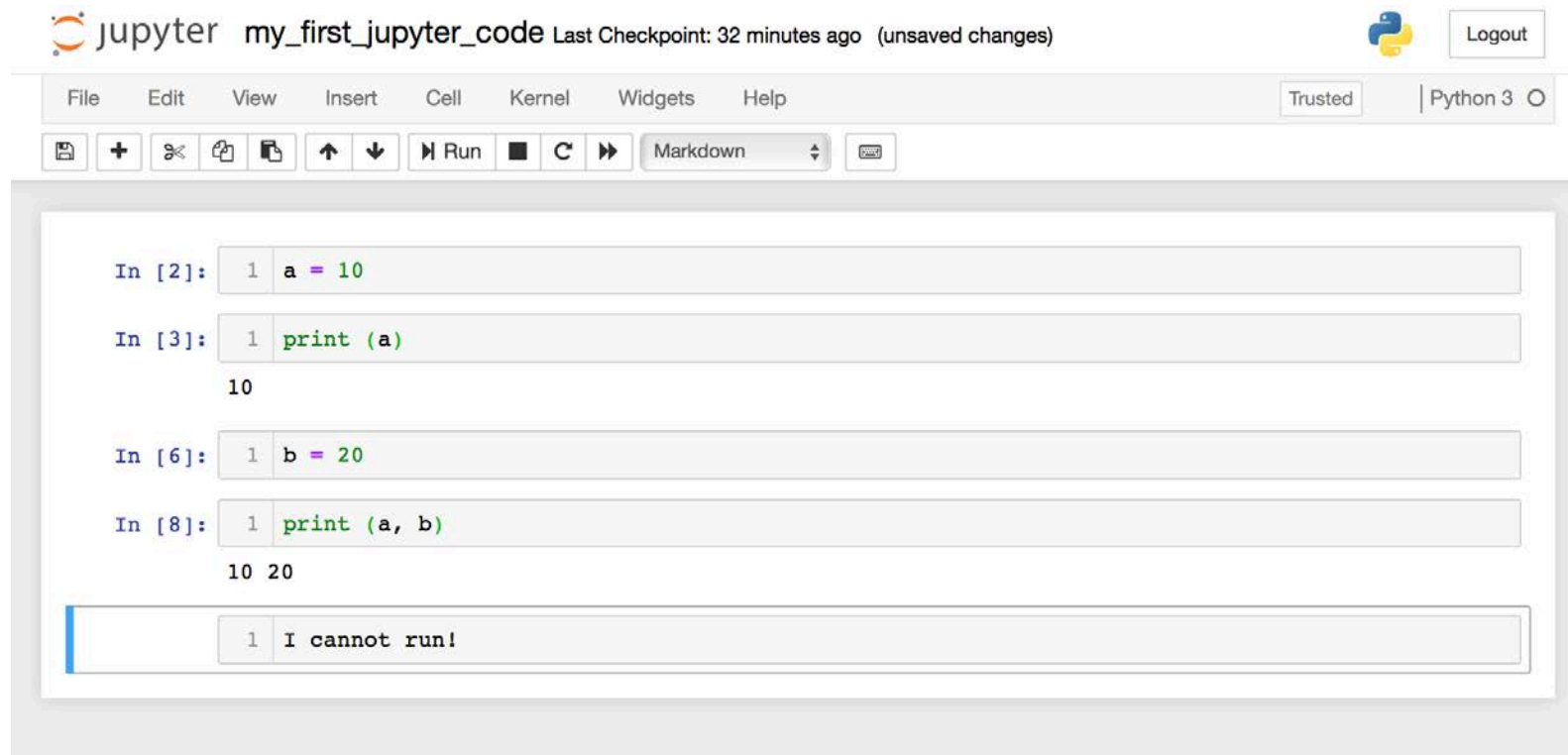
10

```
In [6]: 1 b = 20
```

```
In [7]: 1 print (a, b)
```

10 20

Add text to your notebook



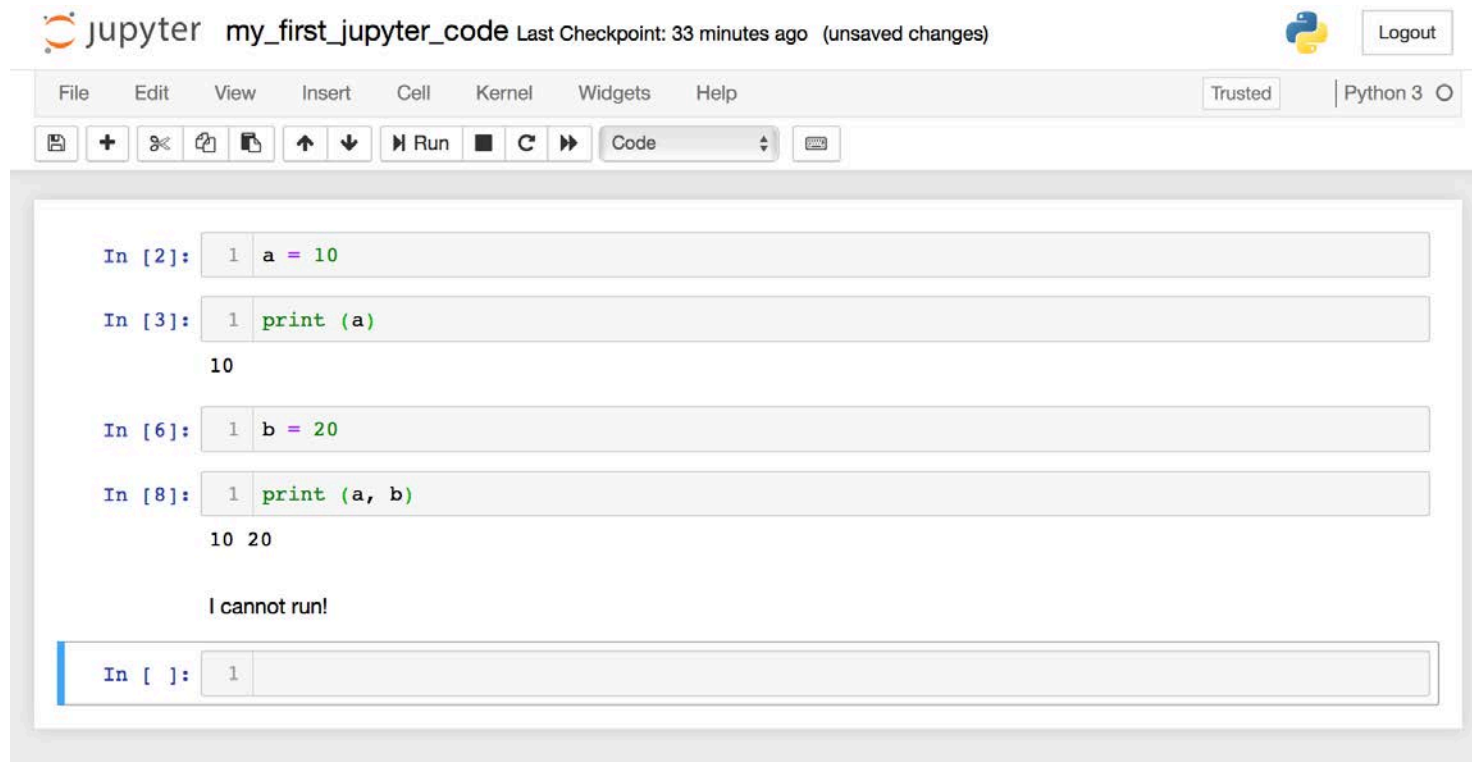
The screenshot displays a Jupyter Notebook titled "my_first_jupyter_code". The top bar shows the Jupyter logo, the notebook name, and the last checkpoint time: "Last Checkpoint: 32 minutes ago (unsaved changes)". On the right, there is a Python logo and a "Logout" button. Below the top bar is a menu bar with options: File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. To the right of the menu bar are "Trusted" and "Python 3" indicators. Below the menu bar is a toolbar with icons for saving, adding a new cell, deleting a cell, duplicating a cell, moving a cell up/down, running a cell, and a dropdown menu currently set to "Markdown".

The notebook contains four code cells:

- In [2]:** `1 a = 10`
- In [3]:** `1 print (a)`
Output: `10`
- In [6]:** `1 b = 20`
- In [8]:** `1 print (a, b)`
Output: `10 20`

Below these cells is an empty code cell with the text `1 I cannot run!` entered.

Add text to your notebook



The screenshot displays a Jupyter Notebook titled "my_first_jupyter_code". The top bar includes the Jupyter logo, the notebook name, and a status message: "Last Checkpoint: 33 minutes ago (unsaved changes)". On the right, there is a Python logo and a "Logout" button. Below the title bar is a menu bar with options: File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. To the right of the menu bar are "Trusted" and "Python 3" indicators. A toolbar with various icons for file operations and execution is located below the menu bar. The main area contains four code cells. The first cell, labeled "In [2]:", contains the code `a = 10`. The second cell, labeled "In [3]:", contains `print (a)` and has an output of `10`. The third cell, labeled "In [6]:", contains `b = 20`. The fourth cell, labeled "In [8]:", contains `print (a, b)` and has an output of `10 20`. Below these is a text prompt "I cannot run!". At the bottom, there is an empty code cell labeled "In []:" with a cursor.

```
Jupyter my_first_jupyter_code Last Checkpoint: 33 minutes ago (unsaved changes) Logout
```

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

In [2]: 1 `a = 10`

In [3]: 1 `print (a)`
10

In [6]: 1 `b = 20`

In [8]: 1 `print (a, b)`
10 20

I cannot run!

In []: 1

We are ready to start

