## Upcoming HPC Classes

<table>
<thead>
<tr>
<th>Quick Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPC Home</td>
</tr>
<tr>
<td>Getting an account</td>
</tr>
<tr>
<td>Getting started on Prince</td>
</tr>
<tr>
<td>Prince How-to Articles</td>
</tr>
<tr>
<td>Logging in</td>
</tr>
<tr>
<td>Windows</td>
</tr>
<tr>
<td>Mac / Linux</td>
</tr>
<tr>
<td>Clusters and Storage</td>
</tr>
<tr>
<td>Prince (HPC)</td>
</tr>
<tr>
<td>Dumbo (Hadoop)</td>
</tr>
<tr>
<td>Brooklyn (OpenStack)</td>
</tr>
<tr>
<td>Dalmatia (NYU Abu Dhabi)</td>
</tr>
</tbody>
</table>
Transfering data to/from the clusters
Transfering data to/from Prince cluster using Globus
Submiting jobs with sbatch
Available software
Licensed Software Available on the HPC Cluster
Building Software
Slurm Tutorial
Tutorials
FAQs
Scratch Area Cleanup
Programming for Biologists
Acknowledgement Statement
Research Gallery
HPC People
Big Data Tutorial 1: MapReduce
Date: Tuesday, June 12, 2018
Time: 2:00pm - 4:00pm
Presenter: Tatiana Polunina
Location: Bobst Library, Rm. 617, 6th Floor
Libraries: Bobst Library
This class will provide a brief overview of what Hadoop is and the various components that are involved in the Hadoop ecosystem. There will be a hands on showcase for the users on how to use the dumbo(Hadoop) cluster to run basic map-reduce jobs. Various hands on exercises have been incorporated for the users to get a better understanding.

Class Materials

The pre-requisites of this class:

1. HPC user account is mandatory.
2. The user needs to have a basic knowledge of Unix and Java/python.

Big Data Tutorial 2: Using Hive
Date: Wednesday, June 20, 2018
Time: 2:00pm - 4:00pm
Presenter: Tatiana Polunina
Location: Bobst Library, Rm. 617, 6th Floor
Libraries: Bobst Library
This tutorial provides a basic understanding of Apache Hive and its usage in the Hadoop eco-system. There will be hands-on examples on how to use Apache Hive and a step by step instructional on how to run Hive jobs using NYU's Dumbo(Hadoop) Cluster.

Class Materials

The pre-requisites of this class:

1. HPC user account is mandatory.
2. Experience using SQL

Big Data Tutorial 3: Intro to Spark
Date: Wednesday, June 27, 2018
Time: 2:00pm - 4:00pm
Presenter: Wensheng Deng
Location: Bobst Library, Rm. 617, 6th Floor
Libraries: Bobst Library
This tutorial provides a basic understanding of Apache Spark and its usage in the Hadoop eco-system. There will be hands-on examples on how to use Apache Spark and a step by step instructional on how to run Spark jobs using NYU's Dumbo (Hadoop) Cluster.

Class Materials
**The pre-requisites of this class:**

1. NYU HPC user account is mandatory.
2. Experience using Scala/Java.

---

**Introduction to Unix/Linux and the shell**

*b*Date: Tuesday, July 3, 2018  
*Time:* 2:00pm - 4:00pm  
*Presenter:* Eugene Dedats  
*Location:* Bobst Library, Rm. 617, 6th Floor  
*Libraries:* Bobst Library

A hands-on introduction to using the Unix command line interface. Aimed at HPC users with little to no Linux experience. We'll cover sessions 1 and 2 of the "Introduction to Unix/Linux and the HPC environment at NYU" tutorial on the NYU HPC Wiki. The tutorial is available online at [https://wikis.nyu.edu/pages/viewpage.action?pageId=53859101](https://wikis.nyu.edu/pages/viewpage.action?pageId=53859101)

Most of the activities in this tutorial can be completed without an NYU HPC account, but you will get more out of this class if you are already an NYU HPC user. You can apply for an account at [https://wikis.nyu.edu/display/NYUHPC/High+Performance+Computing+at+NYU (see "Getting and Renewing Access")](https://wikis.nyu.edu/display/NYUHPC/High+Performance+Computing+at+NYU (see "Getting and Renewing Access"))

---

**Using Slurm on Prince cluster**

*b*Date: Wednesday, July 11, 2018  
*Time:* 2:00pm - 4:00pm  
*Presenter:* Wensheng Deng  
*Location:* Bobst Library, Rm. 617, 6th Floor  
*Libraries:* Bobst Library

A hands-on introduction to using the Slurm. Aimed at HPC users with knowledge of Linux.

---

**Big Data Tutorial 1: MapReduce**

*b*Date: Wednesday, July 18, 2018  
*Time:* 2:00pm - 4:00pm  
*Presenter:* Tatiana Polunina  
*Location:* Bobst Library, Rm. 617, 6th Floor  
*Libraries:* Bobst Library

This class will provide a brief overview of what Hadoop is and the various components that are involved in the Hadoop ecosystem. There will be a hands on showcase for the users on how to use the dumbo(Hadoop) cluster to run basic map-reduce jobs. Various hands on exercises have been incorporated for the users to get a better understanding.

**Class Materials**

---

**Big Data Tutorial 2: Using Hive**

*b*Date: Tuesday, July 24, 2018  
*Time:* 2:00pm - 4:00pm  
*Presenter:* Tatiana Polunina  
*Location:* Bobst Library, Rm. 617, 6th Floor  
*Libraries:* Bobst Library

This tutorial provides a basic understanding of Apache Hive and its usage in the Hadoop eco-system. There will be hands-on examples on how to use Apache Hive and a step by step instructional on how to run Hive jobs using NYU's Dumbo(Hadoop) Cluster.

**Class Materials**

---
This tutorial provides a basic understanding of Apache Spark and its usage in the Hadoop eco-system. There will be hands-on examples on how to use Apache Spark and a step by step instructional on how to run Spark jobs using NYU's Dumbo (Hadoop) Cluster.

**Class Materials**

The pre-requisites of this class:

1. NYU HPC user account is mandatory.
2. Experience using Scala/Java.