Formation Reproductibilité en science de données

Programme de la formation

1. Introduction
2. Environnement de travail : Installation des logiciels & tests
3. Le versioning des codes et documents avec Git
   1. Git: commandes de base
   2. Git: La gestion des conflits
4. Les containers Docker
   1. La technologie des containers
   2. Docker : commandes de base
5. R studio
6. Jupyter notebook & lab
   1. Jupyter pour R
   2. Jupyter pour python
   3. le framework spark en python (pyspark)
7. Le deploiement dans le cloud avec Openstack
   1. Pourquoi le cloud
   2. Horizon, l’interface d’openstack
   3. Deploiement
      1. configuration et lancement de VM
      2. Acces en SSH
      3. Installation de Docker sur VM
      4. Deploiement Jupyter et interface

Introduction

diaporama : doc/f1-1\_introduction.pdf

Installation

document : doc/f2\_installation.pdf

Le versioning des codes et documents avec Git

Les containers Docker

R studio

Jupyter notebook & lab

Jupyter pour R

Jupyter pour python

Le framework spark en Python (pyspark)

Le deploiement dans le cloud avec Openstack