

# DATA SCIENCE

11 WEEK PART TIME COURSE

**Week 1 – Data Science & Python**  
**Wednesday 9th December**

1. What are the steps in a Data Science project?
2. What is Python and why is it Popular with Data Scientists?
3. Bit more Git
4. Python Lab
5. Discussion
6. Course Projects
7. Homework

**DATA SCIENCE PART TIME COURSE**

---

# **WHAT ARE THE STEPS IN A DATA SCIENCE PROJECT?**

# What is Data Science?



“Data analysis projects today rely on databases, computer and network hardware, and computer and network software. A collection of models and methods for data analysis will be used only if the collection is implemented in a computing environment that makes the models and methods sufficiently efficient to use”



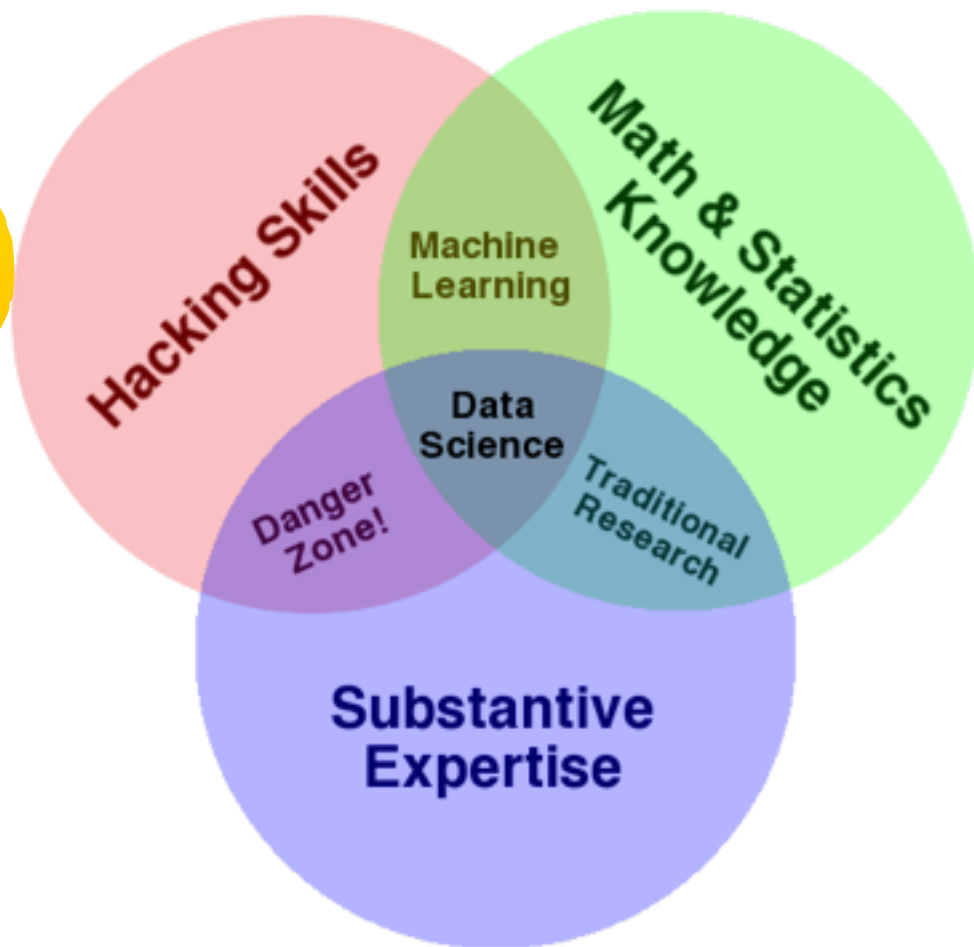
- Created by Guido Van Rossem in 1991 and emphasizes productivity and code readability
- Version 3 (but 2.7 is still very popular)
- “Python is an interpreted, object-oriented, high-level programming language with dynamic semantics”



- Batteries Included: Large collection of built in libraries e.g. SciKit, Pandas, Theano, etc
- Simple and clean syntax
- General purpose language: lots of people outside of data science will be able to work with it



# PYTHON & R





---

**DATA SCIENCE PART TIME COURSE**

---

# **USING DATA SCIENCE PACKAGES**

- Packages are libraries of code written to solve a particular set of problems
- In Python there are many related to data science including Pandas, SciKit Learn, Numpy
- These are installed and managed with PIP (Pip Installs Packages)

```
pip install some-package-name
```

- pandas: manipulate data
- SciPy / NumPy: scientific computing and numerical calculations
- scikit-learn: use machine learning methods
- matplotlib: visualise data
- statsmodels: perform statistical tests
- BeautifulSoup: read in XML and HTML data
- iPython: interactive programming

---

**DATA SCIENCE PART TIME COURSE**

---

# **A BIT MORE GIT!**

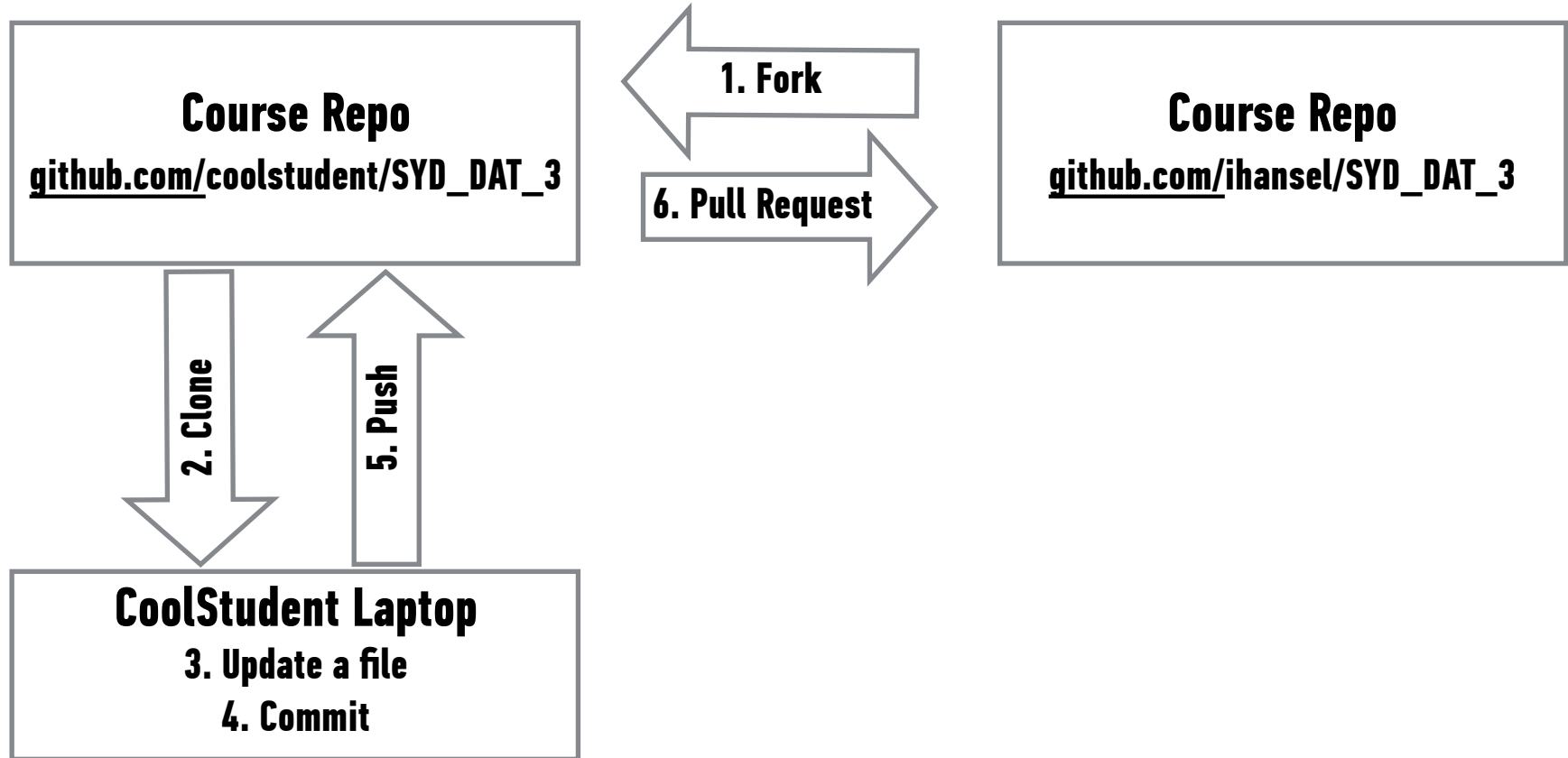


- Forking, making a copy of someone else's repository so you can work on it



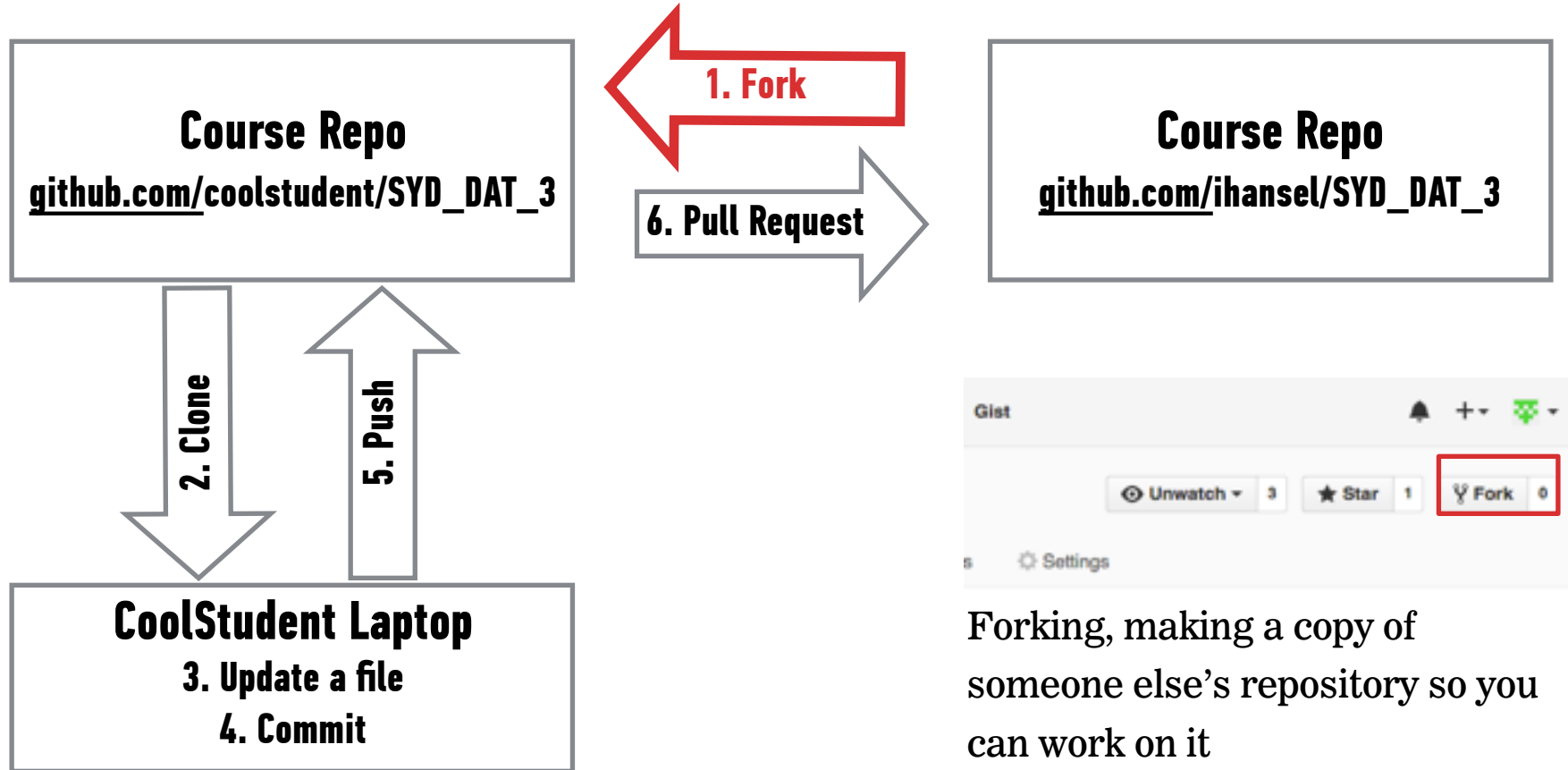
# HOW DO WE ACTUALLY COLLABORATE WITH GIT?

14



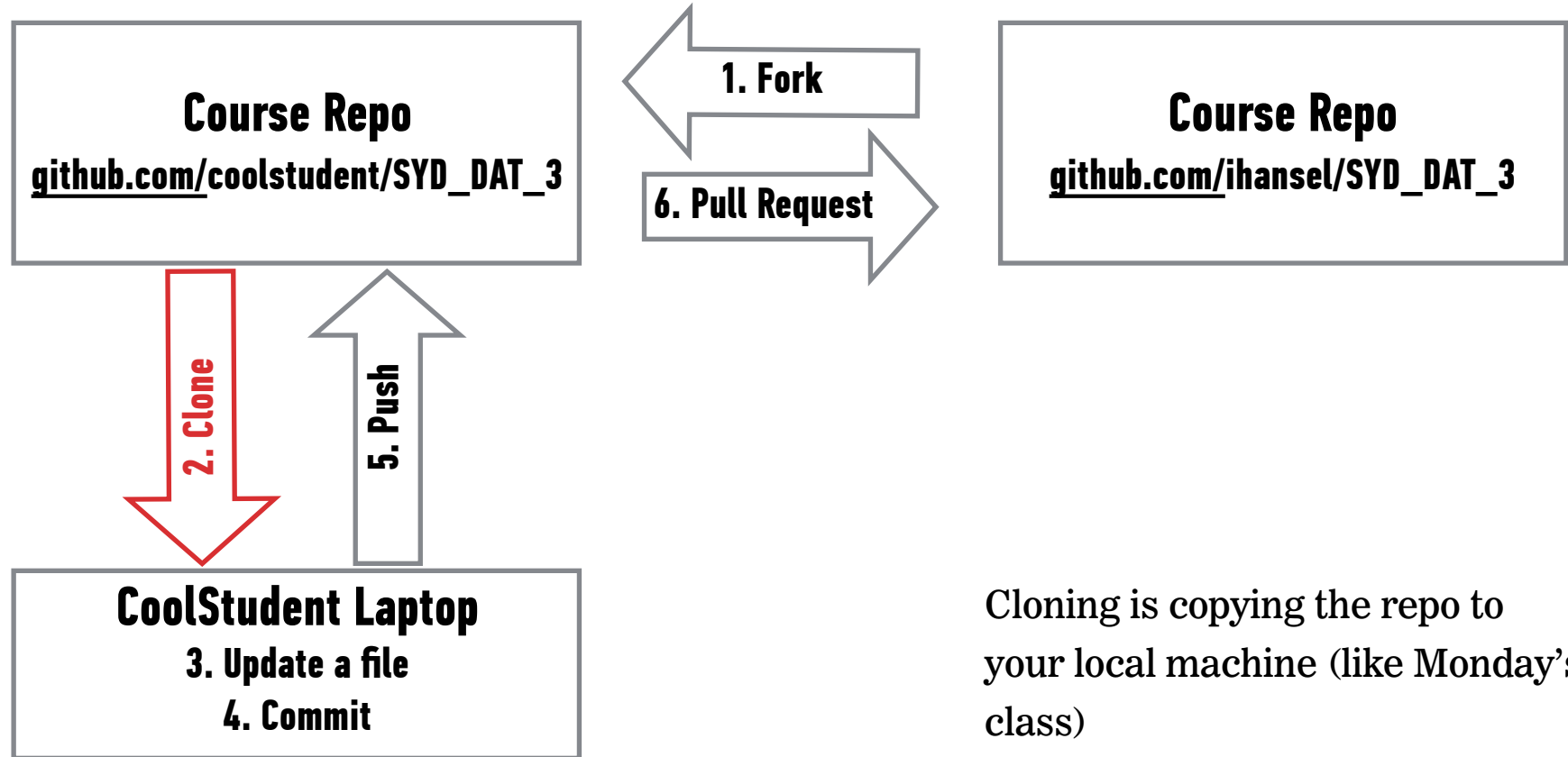
# HOW DO WE ACTUALLY COLLABORATE WITH GIT?

15

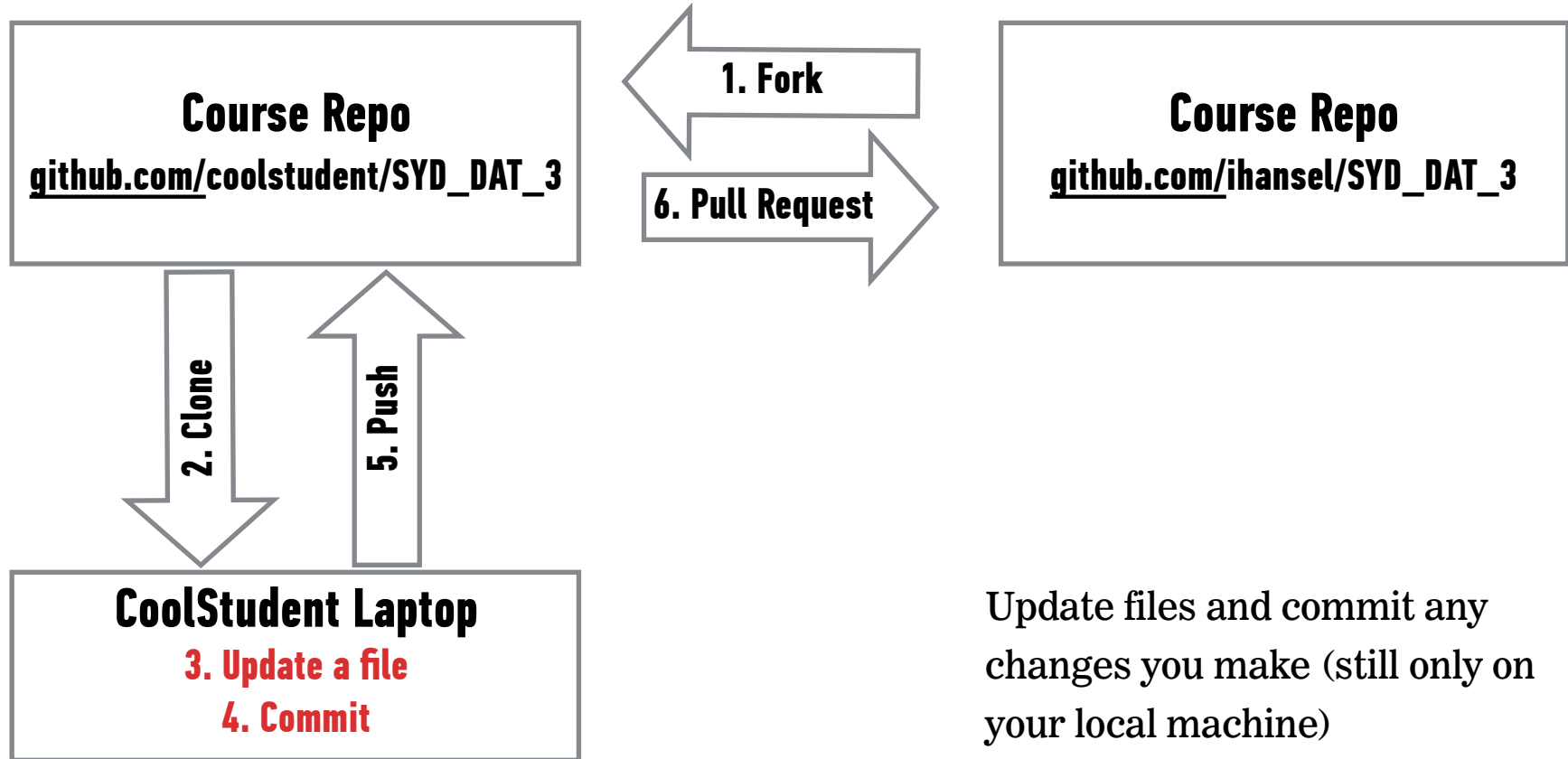


# HOW DO WE ACTUALLY COLLABORATE WITH GIT?

16

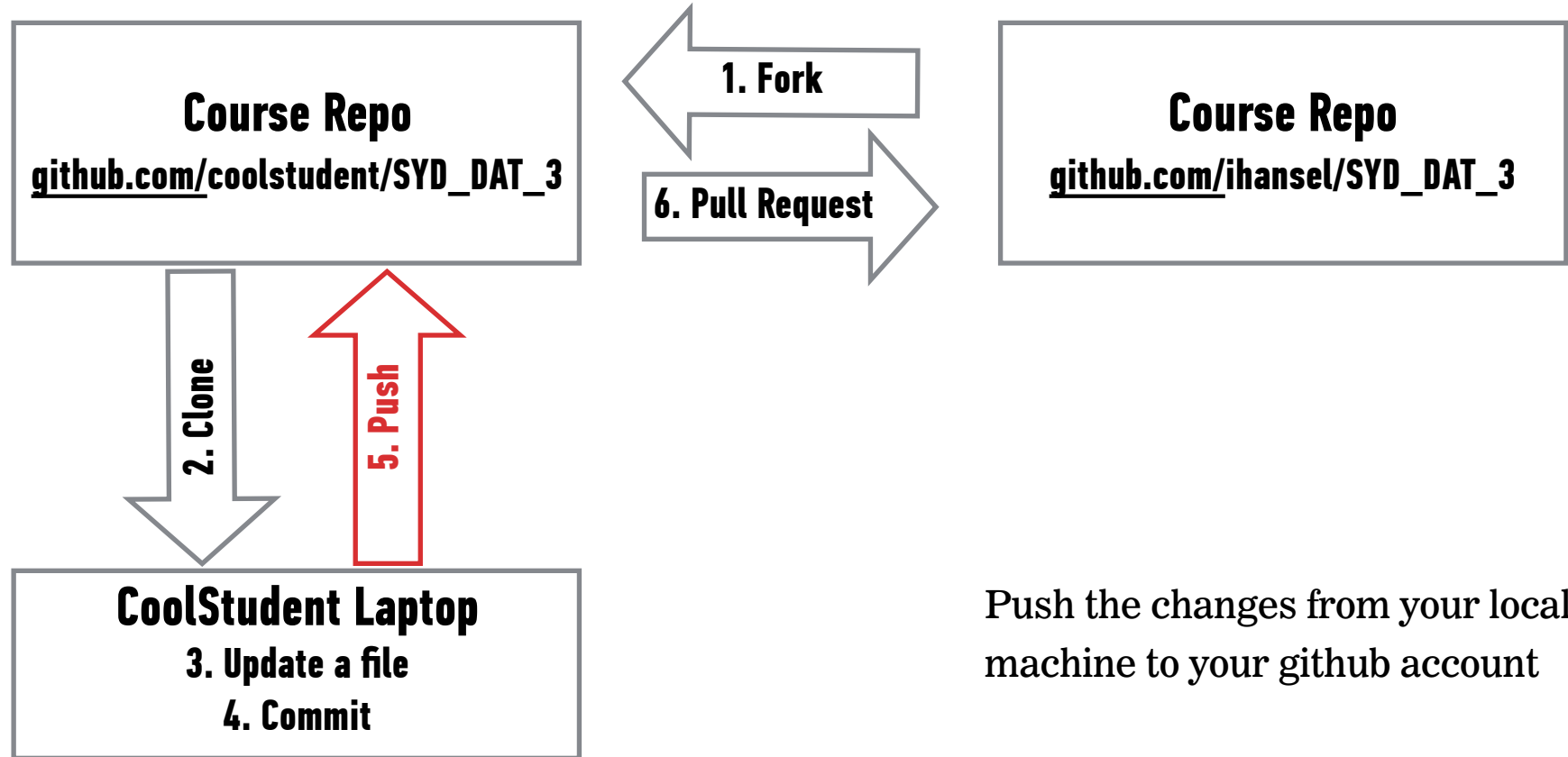






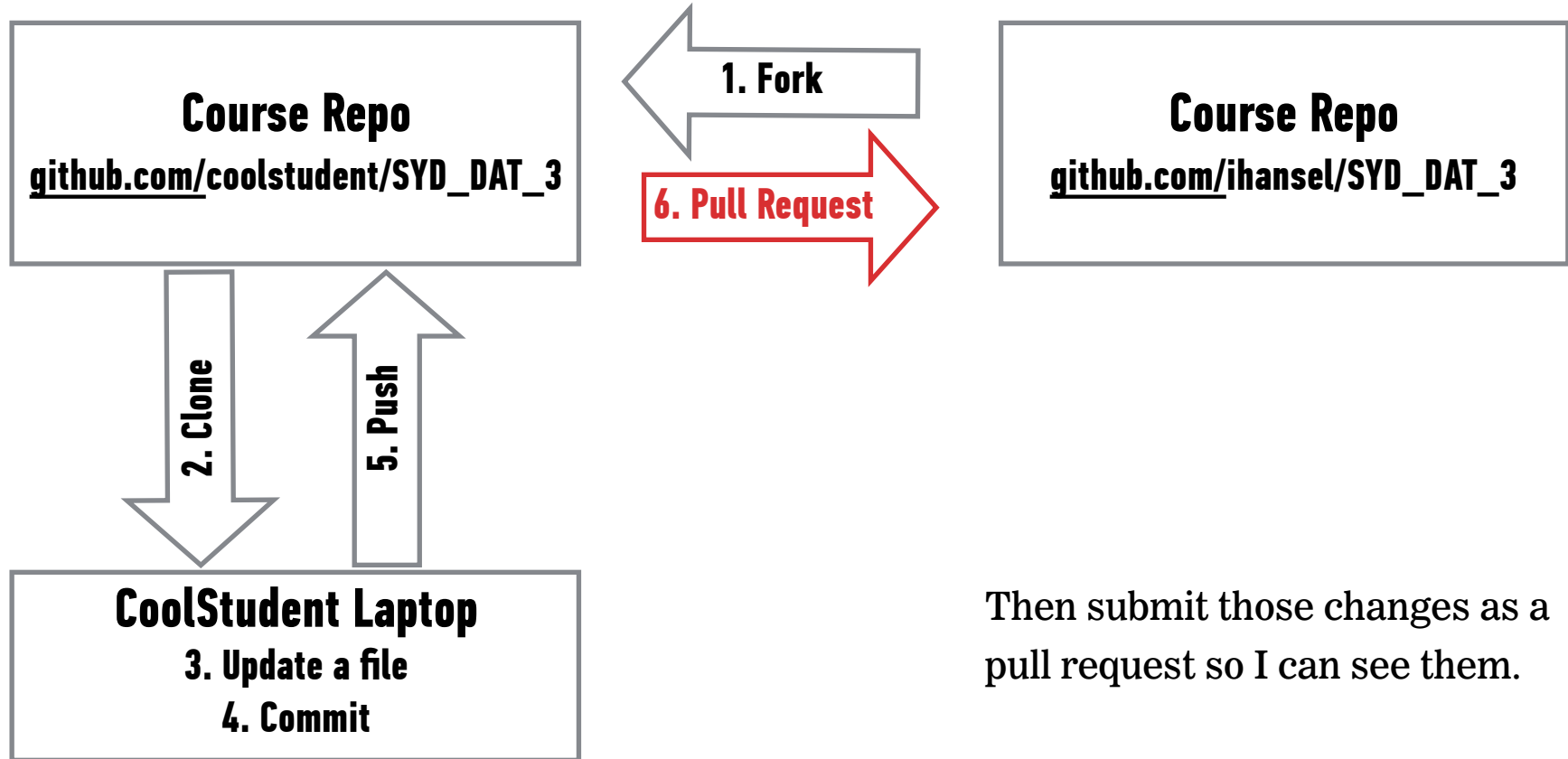
## HOW DO WE ACTUALLY COLLABORATE WITH GIT?

18



## HOW DO WE ACTUALLY COLLABORATE WITH GIT?

19

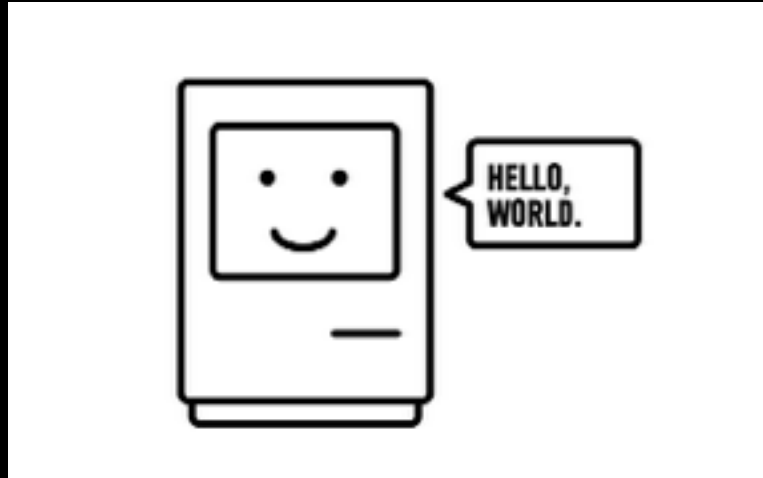


---

**DATA SCIENCE PART TIME COURSE**

---

# **PYTHON LAB**



**DATA SCIENCE - Week 1 Day 2**

---

# **DISCUSSION TIME**

## Data Science Week 1 Monday

- ☒ Course Expectations
- ☒ What the Course Involves
- ☒ Course Project Information (Basics)
- ☐ Lab-Setup
  - ☐ Anaconda Python
  - ☒ Git/Github
  - ☐ Slack

---

**DATA SCIENCE - Week 1 Day 2**

---

# **DISCUSSION TIME**

## **Pework**

### **▸ Readings**

- Metacademy Learning Plan**
- Data Science Handbook**
- An Introduction to Statistical Learning**

---

**DATA SCIENCE - Week 1 Day 2**

---

# **DISCUSSION TIME**

## **Homework**

- **Due next Friday**
- **I will review within 3-4 days**
- **Counts to letter of completion**
- **(Will be easy)**



- # books-resources
- # general
- # homework
- # instructional-team
- # meetups-events
- # news
- # presentations

- # random
- # software

- slackbot
- eladair
- jared
- meggan
- mikaela
- sofia\_alfonso
- thelocker

## #prework

This is the very beginning of the [#prework](#) channel, which you created yesterday. The purpose of this channel is: *Ask any questions about the pre-work for the course.* ([edit](#)).

+ Add a service integration    + Invite others to this channel

Yesterday



ihansel 1:53 PM  
joined #prework



ihansel 1:53 PM  
set the channel purpose: Ask any questions about the pre-work for the course.



mikaela 1:53 PM  
joined #prework by invitation from @ihansel, along with @eladair, @jared, @meggan, @sofia\_alfonso, @thelocker



ihansel 1:54 PM  
Hi @channel, hope you're all excited for the course and have made some progress on the pre-work! I've made a slack-channel just for the pre-work for the course. If you've got any questions or are having any difficulty getting through it feel free to message in here or pm me. ([edited](#))



ihansel 1:54 PM  
uploaded a file +



Welcome to Data Science - DAT3.pdf  
1MB PDF



ihansel 1:55 PM  
And the good news - all of you have completed part of the pre-work just by being on slack 😊