Global Data Platform On-Boarding

# Pre-Requisites

1. Python training
2. Big data training

# Environment Setup

1. Install Pycharm.
2. Download code from github (<https://github.com/adriandolha/bigdata>). To do that use Pycharm -> Checkout from version control -> Git
3. Build the code using the following command:
   1. **pip install .**
4. Install Kafka
5. Install Spark