**Python & Pyspark Data processing**

* Get data from multiple sources:

1. SQL Server
2. MySQL
3. PostgreSQL
4. MongoDB
5. Noe4j
6. Oracle
7. CSV, Excel and textFile
8. Hadoop
9. HTML Website (Beautiful sup, selenium)
10. Facebook API
11. Google Analytics
12. Sensor Bluetooth
13. Socket IP

* Exploratory Data Analysis (rfm)
* Data cleaning
* Data visualizations
* Write data to multiple sources:

Machine learning, liner regression, Classification and Clustering