

Tuần 3: Biểu đồ và đồ thị (tiếp theo)

Giảng viên: ThS. Lê Minh Tân

Nội dung

- I. Notebook và trình thông dịch
- II. Môi trường Spark trong Zeppelin

I. Notebook và trình thông dịch

- Notebook gồm: paragraph chứa code.
- Một notebook gắn với một hoặc nhiều trình thông dịch.
- Zeppelin hỗ trợ:
 - Spark
 - Python
 - R
 - SQL
 - Scala
 - Markdown
 - JDBC

```
%<tên trình thông dịch>  
%pyspark  
%md
```

- Trình biên dịch chính.
 - Hoạt động độc lập.
- Trình biên dịch trừu tượng.
 - Không thể chạy độc lập.
 - Có thể kế thừa.
- Trình biên dịch đầu ra.
 - Không thể chạy độc lập.
 - Chạy sau khi trình biên dịch đã trả kết quả.

II. Môi trường Spark trong Zeppelin

- Chia sẻ dữ liệu giữa các trình thông dịch.
- Cung cấp các API từ Spark thông qua:
 - SparkContext
 - SQLContext
 - **SparkSession**
 - **ZeppelinContext**