## Zepellin Notebook

Table of Contents :

1. What is Zeppelin
2. Installation steps
3. E

* Currently Zeppelin supports many interpreters such as Scala(with Apache Spark), Python(with Apache Spark), SparkSQL, Hive, Markdown and Shell.
* Some basic charts are already included in Zeppelin. Visualizations are not limited to SparkSQL's query, any output from any language backend can be recognized and visualized.
* Zeppelin can dynamically create some input forms into your notebook.
* Notebook URL can be shared among collaborators. Zeppelin can then broadcast any changes in real-time, just like the collaboration in Google docs.
* Zeppelin provides an URL to display the result only, that page does not include Zeppelin's menu and buttons. This way, you can easily embed it as an iframe inside of your website.

Note: Zeppelin provides built-in Apache Spark integration. You don't need to build a separate module, plugin or library for it.

**Zeppelin's Spark integration provides**

* + Automatic SparkContext and SQLContext injection
  + Runtime jar dependency loading from local filesystem or maven repository. Learn more about [dependency loader](https://zeppelin.apache.org/docs/0.5.5-incubating/interpreter/spark.html).
  + Canceling job and displaying its progress

### Text

Zeppelin prints output of language backend in text, by default.

Html

With '%html' directive, Zeppelin treats your output as html

* Practical :

1. New Note ( default interpreters )
   * Choose correct interpreter for your notebook

* See what are available
  + How to create folders

**Pyspark\_advaced/Zeppelin\_tutorial/lession1**

**Import existing notes :**

**Hortonworks gallery for zeppelin notebooks : show how to import**

**First note :**

**Show all the interpreters and what are the directives**

**Show example of %md**

**“”**

**# Zeppelin tutorial :**

**## table of contents**

1. **First note ( hello note )**
2. **Interprters**
3. **Add dependency**

**“”**

**ADD PARAGRAPH ABOVE + BELOW**

**CTRL + SHIFT + d TO REMOVE**

**USE SHIFT + ENTER TO RUN THE CURRENT PARAGRAPH**

**Start with if not interpreter specified ; then default is spark ( or the one chosen while you created the note**

**EXAMPLE OF SPARK PARAGRAPH:**

**select \* from ${table1}**

**EXAMPLE OF SHELL + hdfs commands**

echo "my name is ${newvar}"

echo "my name is ${newvar}"

hadoop fs -ls /user/root

**EXAMPLE OF SQL**

**Show tables;**

**select \* from sample\_07 limit 5;**

**show charts : data |**

**change measures:**

**download data “**

**First line shold be interpreter ( show cheatsheet for interpreters )**

1. %md
2. Markdown interpreters – show some example ( 20 sec )
3. Example of pyspark interpreter

%md

Zeppelin notebooks support multiple language as backend tool

1) default : scala with spark

2) %sh : shell

3) %md : markdown

4) %sql : spark SQL

5) %pyspark : pyspark

* print(sc)
* print(sc.version)

%sql

select to\_date(payment\_date), sum(amount) amt from payment group by to\_date(payment\_date) order by to\_date(payment\_date) limit 30

%pyspark

%pyspark

df = sqlContext.sql("select to\_date(payment\_date), sum(amount) amt from payment group by to\_date(payment\_date) order by to\_date(payment\_date) limit 30")

df.show()

df.registerTempTable("inventory")

%sql

select \* from inventory