ZEPPELIN

A Notebook Interface to your Big Data
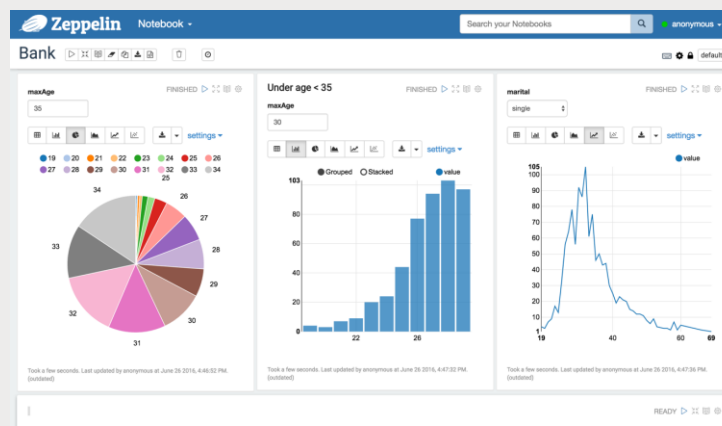
What is Zeppelin?

- If you're familiar with iPython notebooks – it's like that
    - *Lets you interactively run scripts / code against your data*
    - *Can interleave with nicely formatted notes*
    - *Can share notebooks with others on your cluster*
- If you're not familiar with iPython notebooks – well, you kind of just have to see it.
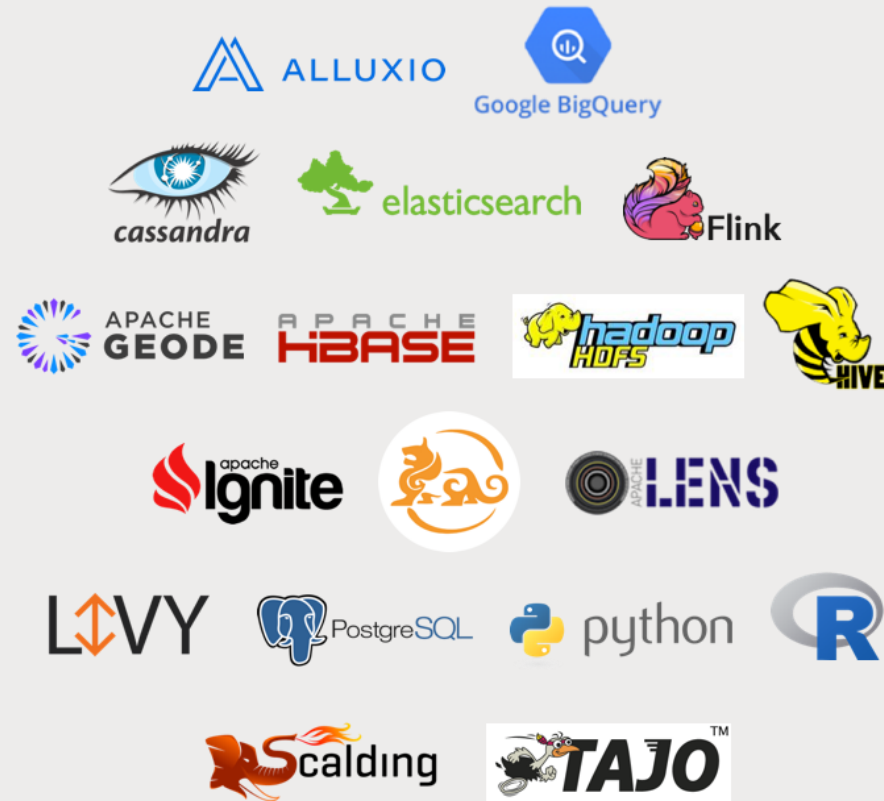
Apache Spark integration

- Can run Spark code interactively (like you can in the Spark shell)
  - *This speeds up your development cycle*
  - *And allows easy experimentation and exploration of your big data*
- Can execute SQL queries directly against SparkSQL
- Query results may be visualized in charts and graphs
- Makes Spark feel more like a data science tool!

Zeppelin can do much more than Spark

It'll make more sense if we just play with it.

- Zeppelin comes pre-installed on Hortonworks Data Platform
- So let's jump in