# Apache Zeppelin

## Team 19

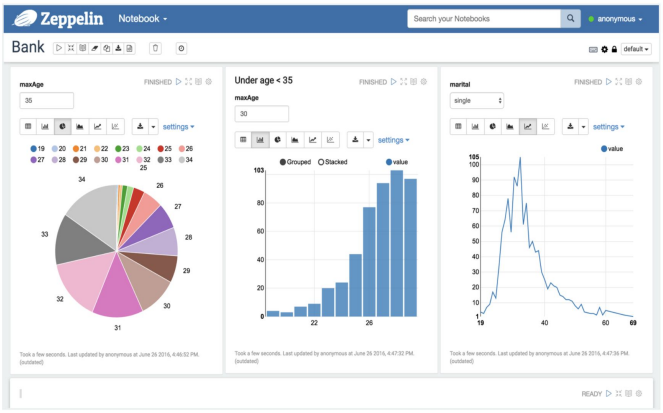
Carlson School of Management

As the world continues to become increasingly digital, data science pipelines are steadily moving online. Apache Zeppelin is a relatively new web-based notebook that adds data exploration, visualization, sharing, and collaboration capabilities to Spark. Python is supported, as well as an increasing number of other programming languages including Scala, Hive, SparkSQL, shell, and markdown. Apache Zeppelin provides a “Modern Data Science Studio” for Data Ingestion, Data Discovery, Data Analytics, Data Visualization and collaboration.

## Language Support

cassandra, R, HIVE, TAJO, LENS, ALLUXIO, Scalding, Flink, PostgreSQL, elasticsearch, APACHE GEODE, APACHE HBASE, hadoop HDFS, apache Ignite, python

## Dashboarding




## Use Cases

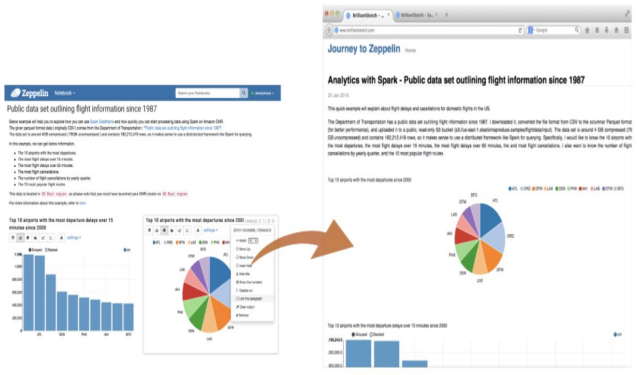
Delta Lakes

Social Network Analysis

amazon web services

Integration with AWS

## Collaboration



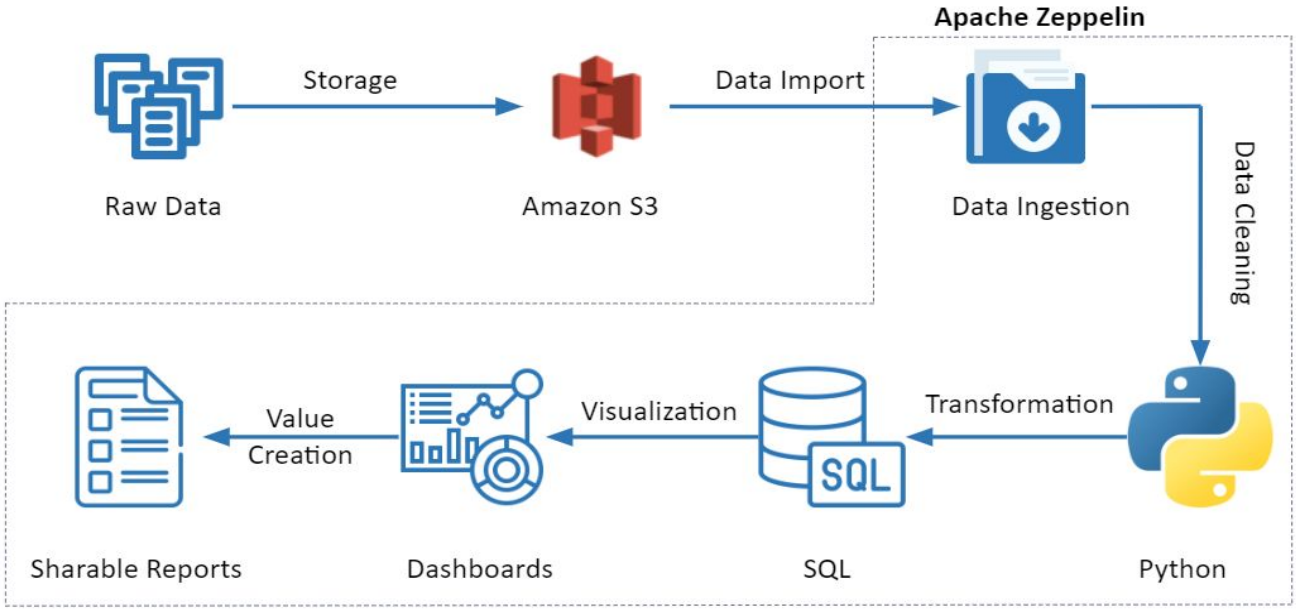

- Now collaborate with ease with security settings and report sharing

## Advantages

VS

| | | |
|---|---|---|
| Restrictions and Grant Access | **Security** | No Security Functions |
| A wide range of interpreters | **Interpreters** | No interpreters added in |
| Multi-user config on one server | **Multi-user Capability** | No support multi-user config |
| Interactive visuals at faster rate | **Plotting of graph** | Output charts using plotly lib |

## Our Workflow




## Adopters

Uber

Anurag Mishra mishr212@umn.edu
Arkan Chatterjee chatt140@umn.edu
Havish Gutta gutta006@umn.edu
Jiatong Li li002180@umn.edu
Songhan Lei lei00049@umn.edu
Github