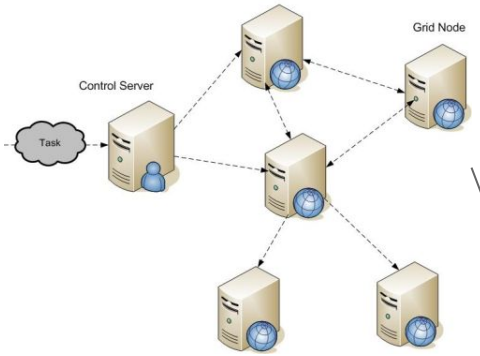
# Spark support

1. Remote execution at Zeppelin
2. Sending jobs to Livy Server
3. User-friendly data representation

*JetBrains spring practice 2018*
*Dmitriy Naidanov, Roman Shein, Anton Yalyshev*





Scala

akka

Spark

kafka

lagom

https://insights.stackoverflow.com/survey/2016#technology-top-paying-tech

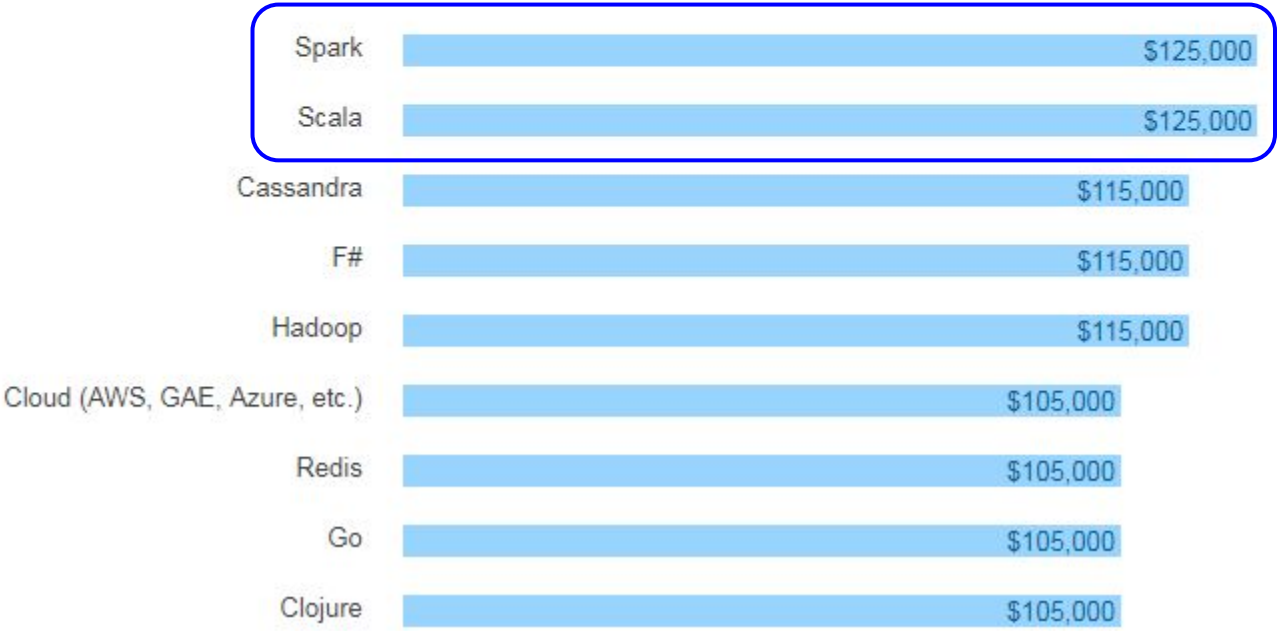
stackoverflow



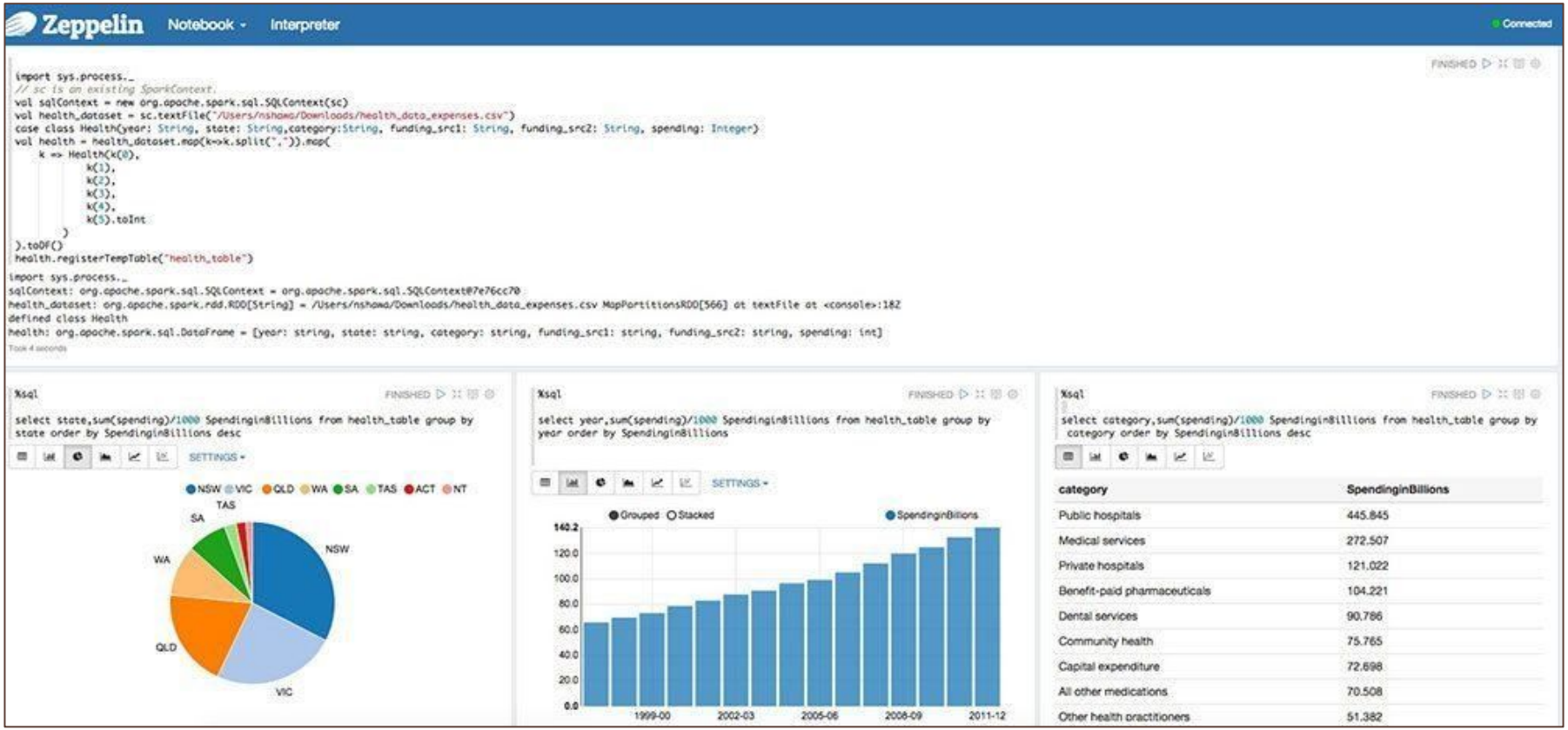
Back to top

# V. Top Paying Tech

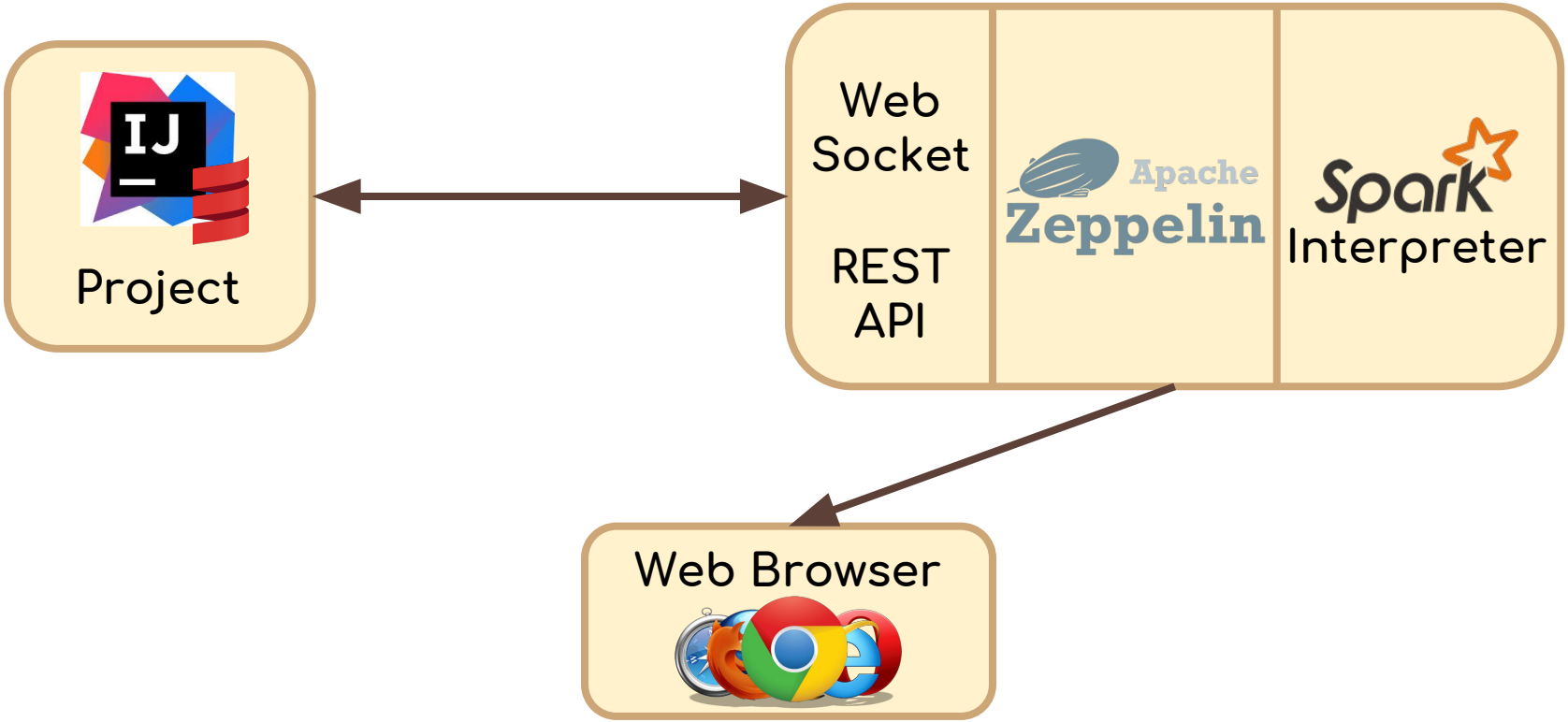
Top Paying Tech in US | Top Paying Tech Worldwide



# 1. Удалённое исполнение кода на Zeppelin



# 1. Удалённое исполнение кода на Zeppelin



# 1. Удалённое исполнение кода на Zeppelin

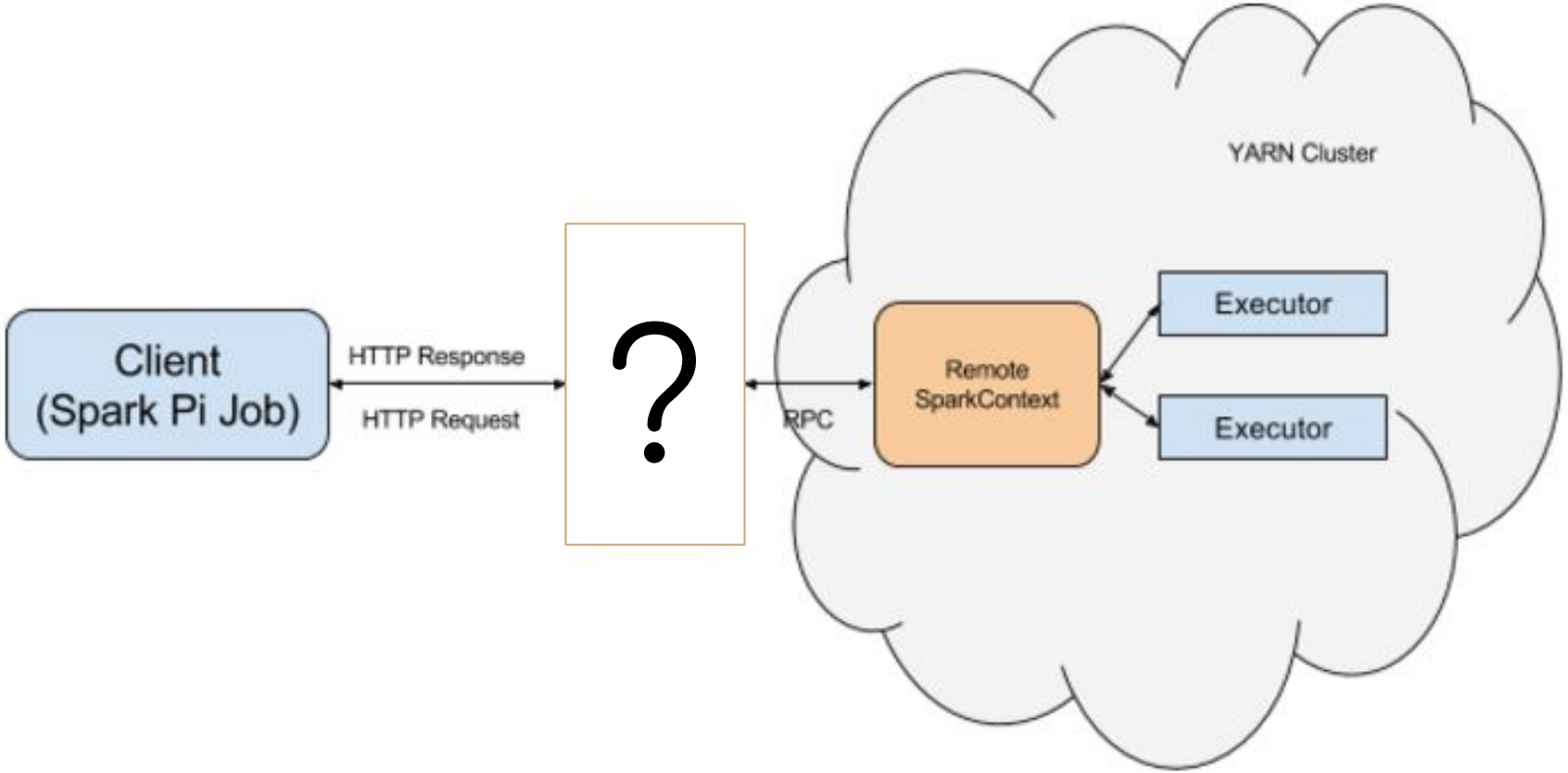
**Что делать:**

1. Механизм работы ран. конфигураций в IntelliJ IDEA
2. Установка, настройка Apache Zeppelin server
3. Взаимодействие IDEA и Zeppelin через REST API

**Что потребуется:**

1. Java, Scala
2. Функциональное программирование
3. REST взаимодействие

# 2. Отправка задач на Livy server



# 2. Отправка задач на Livy server

**Что делать:**

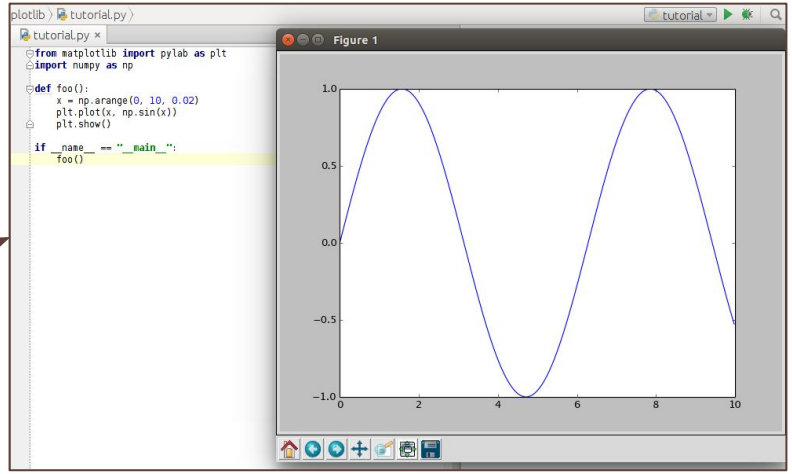
1. Вычислительный кластер
2. Установка, настройка Apache Livy server
3. Механизм работы IDEA с Livy через REST API

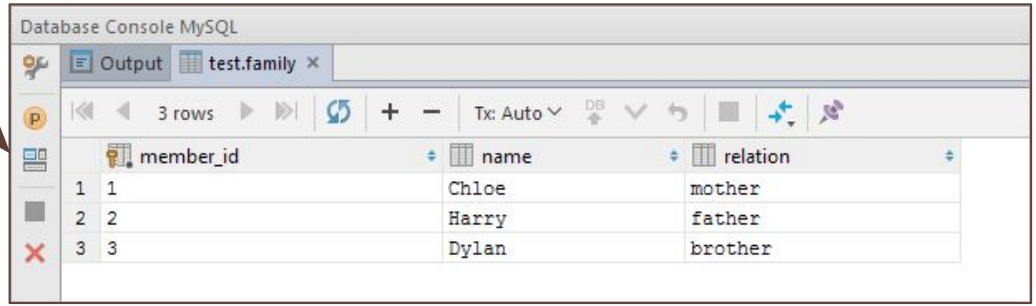
**Что потребуется:**

1. Java, Scala
2. Функциональное программирование
3. REST взаимодействие

# 3. Визуализация расчётов

- RDD [ T ]
- DataFrame
- Dataset



# 3. Визуализация расчётов

**Что делать:**

1. Разобраться с хранением результатов
2. Наиболее удобный вариант визуализации

**Что потребуется:**

1. Java, Scala
2. Функциональное программирование
3. Опыт с Plotting libs

# References

- Тестовое задание:
  - Удалённое исполнение Zeppelin: https://stepik.org/lesson/68757
  - Отправка Job'ов на Livy server: https://stepik.org/lesson/68758
  - Графическое отображение: https://stepik.org/lesson/68759
- Описание проектов:
  - https://jetbrains.ru/students/practice/themes/zeppelin-spark-code/
  - https://jetbrains.ru/students/practice/themes/spark-jobs-to-apache-server/
  - https://jetbrains.ru/students/practice/themes/spark-data-visualization/
- E-mail: anton.yalyshev@jetbrains.com