Job Market Analysis Platform

Baochen Hu, Ning Zhang

Schedule

- Crawl Job Info from Website (Craiglist, Glassdoor)

- Ingest the data using any pipeline(kafka)

- distribute to backend

- Hbase for report

- Spring Boot Web app to demo

- Spark(Real Time Analysis)

Technology Stack

- Hbase
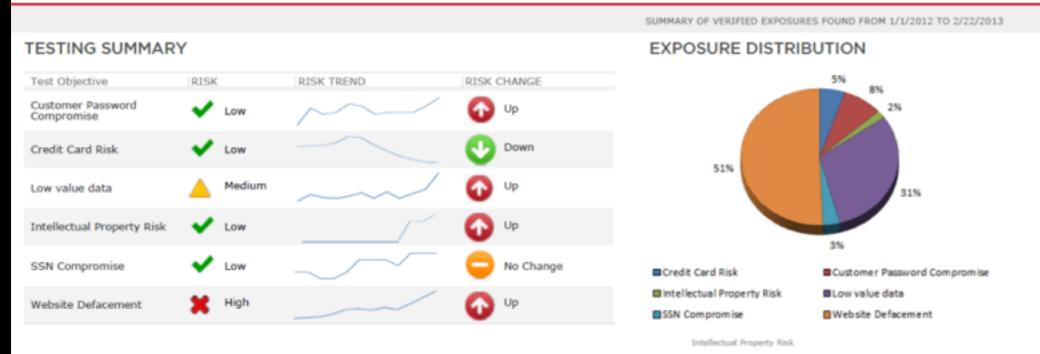
- Kafka

- Spark-Streaming

- Spark-ML

- Scrapy Crawler

Questions

- What is highest salary

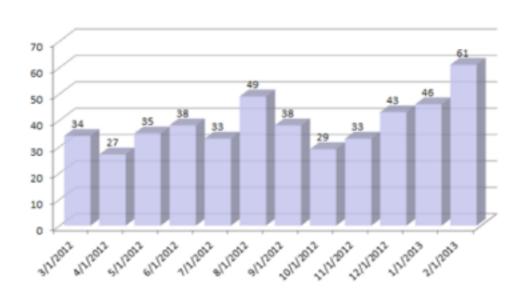- Which company has the best review

- Which company is the best match

Expectation: Report

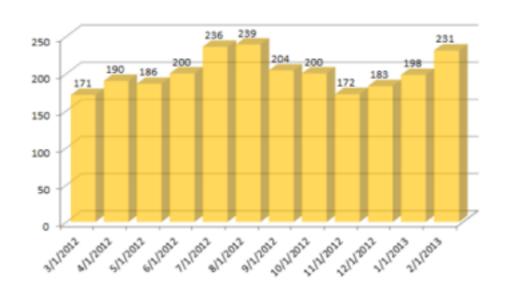# Recommendation