**Harvard University Extension School**

**"Principles of Big Data Processing"**

**CSCI E-88, Fall 2019**

**Final Project**

**by FirstName LastName**

## Project Goal and Problem Statement

Explain the problem you are going to solve or the use case you will be implementing in your project

## YouTube Video URL

Include your YouTube video URL here

## Big Data Source

Specify the data source for you project - it should be either a big public data source set or self-generated datasets if needed.

Describe the data, size, schema and any other details that are need to understand the data.

## Expected results

Describe the expected results/end state of the project

## Processing Pipeline

Explain design of your pipeline and reasons for your choices of technologies for each tier.

Describe all tiers that will be used and their configuration if applicable. Provide more details for the new technology[s] that you used.



## Implementation

This is the main section. It should contain all implementation details for each tier. Copy all relevant configuration and code. Code should be well documented.

## Results

Demonstrate your results. Provide all relevant screenshots of the log files, data stored in DB and GUI tier if applicable.

## Conclusions and Lesson Learned

Describe what you have learned during this project:

* What issues did you have?
* What limitations, if any, did you run into with the technologies used?
* What would you do differently next time?
* What would you like to improve if you had more time?
* What alternatives to technologies you used you might consider ?
* Where would you take your project next ?