**ETL Project**

**Title: Data Science Skills in Job Descriptions**

The goal of this project is to collect data about popular skillsets required in data science-related job descriptions.

**Team:**

* Lidio Meireles
* Christopher Reutz
* Saurabh Bhatt
* Pavitra Bagirthi

**Extract**

* Extract list of tags/skills (e.g. python, mongodb etc) by scraping stackoverflow <https://stackoverflow.com/tags>
* Search for “data scientist” related jobs in Glassdoor and scrape the job descriptions.

**Transformation**

* Find the tags/skills present on each job description

**Load**

* Insert into mondodb the following information:
* Job title
* Company
* City
* Salary
* Job description
* Tags (list of tags)