# Project 3 - Part One

## Team 6

## Team 6 - Project 3

Create a short document, with the names of group members. You should briefly describe your collaboration tool(s) you'll use as a group, including for communication, code sharing, and project documentation. You should have identified your data sources, where the data can be found, and how to load it. And you should have created at least a logical model for your normalized database, and produced an Entity-Relationship (ER) diagram documenting your database design.

### Team Members

- Donald Butler
- Jeff Parks
- Mark Schmalfeld
- Nick Oliver

### Collaboration Tools

- Slack for team communication
- Shared GitHub repository for editing code and uploading documents

### Data Sources

- Data will be obtained from publicly available job postings for Data Scientist positions on job aggregation websites. The commonly used section in job postings where required skills are listed will be used as the source of the data for what Data Science skills are in demand.
- Data will be pulled from APIs or scrapped from the postings
- Data will be loaded into the a database using CSV files or direct insert statements

### Database Design

**Tables**

`Job Listing`

| column_name | data_type | attrs |
|---|---|---|
| id | int | pk |
| source | text | not null |
| date_posted | date | |
| company | text | |

`Job Listing Requirements`

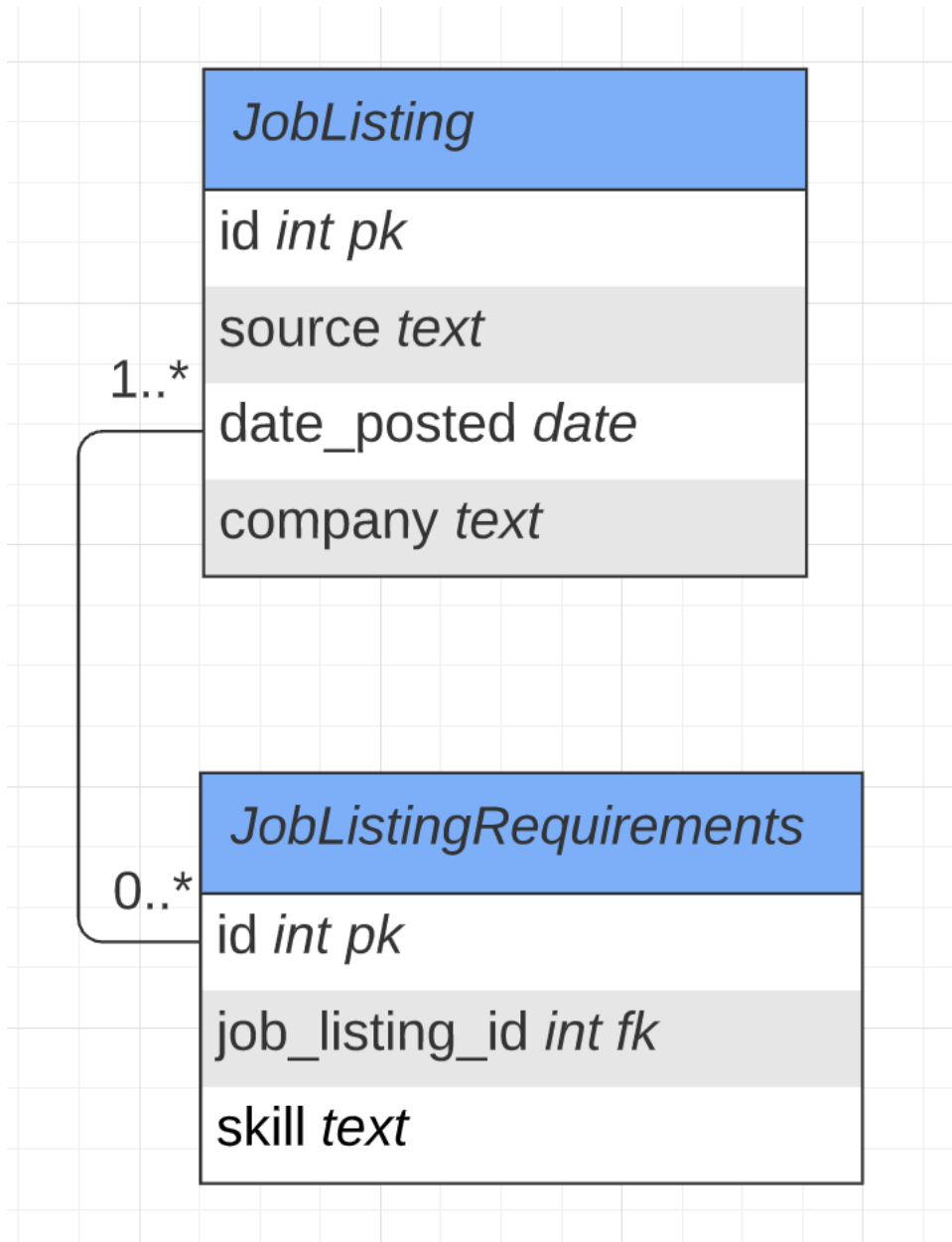| column_name    | data_type | attrs    |
|----------------|-----------|----------|
| id             | int       | pk       |
| job_listing_id | int       | fk       |
| skill          | text      | not null |



Figure 1: Entity Relationship Diagram