# Big Data Analytics and Management Final Project Proposal

By Farnoosh Koleini, Seyedhadi Seyed, Allie Rodriguez

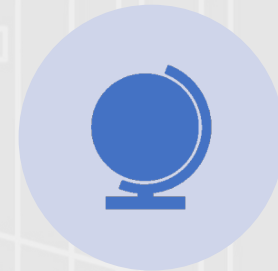# Salaries Of ML/AI Tech Jobs Around The World
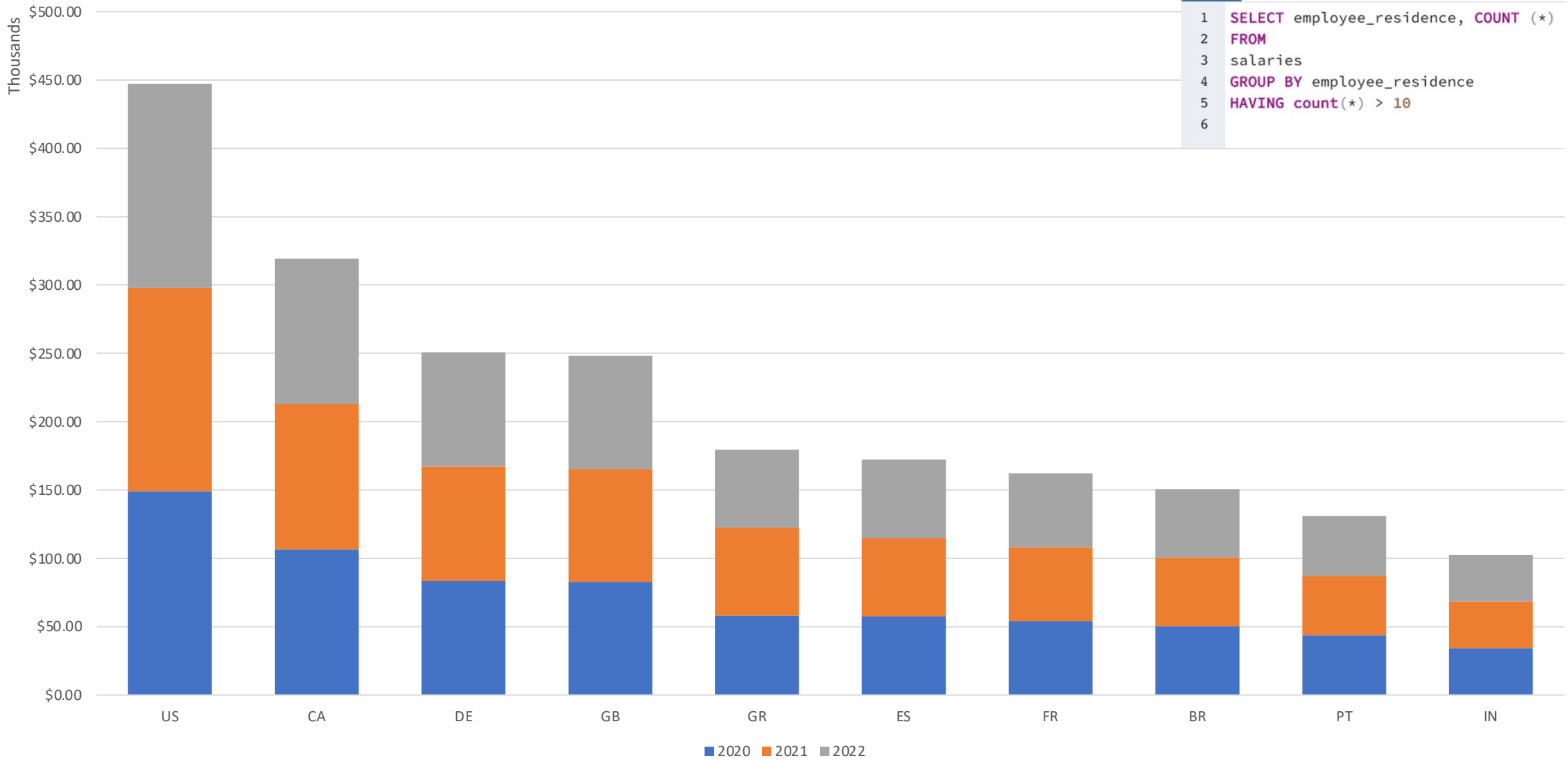


TECH FIELD

JOBS

SALARIES

GLOBAL

# Research Question

What are the factors that influence salaries for ML/AI professionals in the US?

Which countries offer competitive salaries that can be leveraged to lower the cost of talent acquisition?
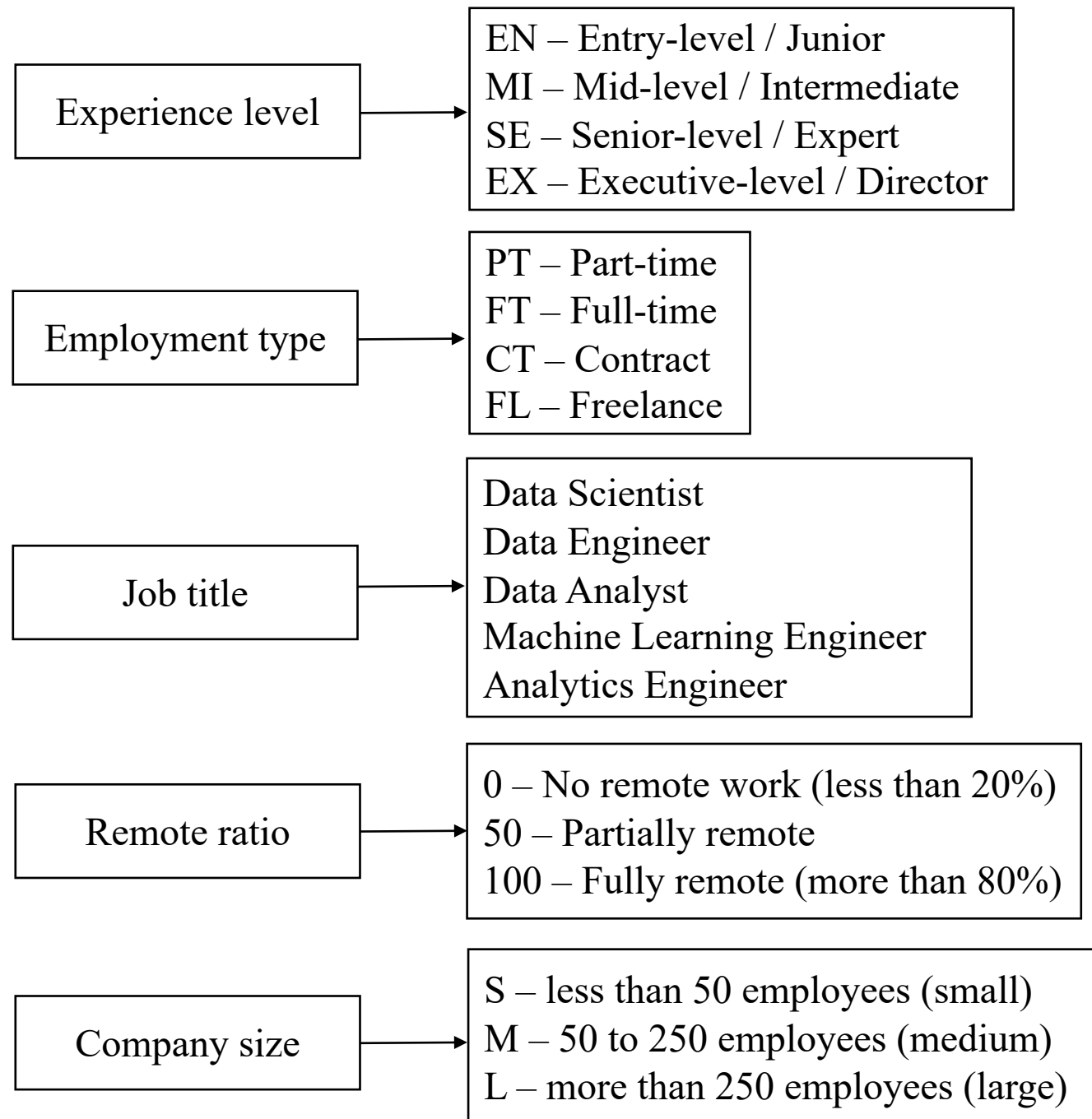
# Average Salaries of Top 10 earning countries



```
Query    Query History
1    SELECT employee_residence, COUNT (*)
2    FROM
3    salaries
4    GROUP BY employee_residence
5    HAVING count(*) > 10
6
```

Thousands

$500.00
$450.00
$400.00
$350.00
$300.00
$250.00
$200.00
$150.00
$100.00
$50.00
$0.00

US    CA    DE    GB    GR    ES    FR    BR    PT    IN

■2020   ■2021   ■2022

# Dataset

- [ML/AI Salaries](#)

- Collected and made available at [ai-jobs.net.](#)

- 11 columns and 2532 rows
  - Work year
  - Salary
  - Salary currency
  - Salary in usd
  - Employee residence
  - Company location

| Experience level | → | EN – Entry-level / Junior<br>MI – Mid-level / Intermediate<br>SE – Senior-level / Expert<br>EX – Executive-level / Director |

| Employment type | → | PT – Part-time<br>FT – Full-time<br>CT – Contract<br>FL – Freelance |

| Job title | → | Data Scientist<br>Data Engineer<br>Data Analyst<br>Machine Learning Engineer<br>Analytics Engineer |

| Remote ratio | → | 0 – No remote work (less than 20%)<br>50 – Partially remote<br>100 – Fully remote (more than 80%) |

| Company size | → | S – less than 50 employees (small)<br>M – 50 to 250 employees (medium)<br>L – more than 250 employees (large) |

# Method



➤ Analyze and compare the salaries (results) of ML/AI professionals in other countries with those in the United States to identify the countries that offer competitive salaries.

➤ We will use the SQL for analyzing this dataset. (Update the results based on the updated dataset which includes 2023 salary information.)

## Query    Query History

```
1   SELECT work_year
2   job_title,
3   salary_in_usd,
4   employee_residence
5   FROM salaries
6   WHERE employee_residence = "US"
7   ORDER BY work_year, salary_in_usd DESC
8
```

| | work_year<br>integer | job_title<br>character varying (50) | salary_in_usd<br>integer | employee_residence<br>character varying (50) |
|---|---|---|---|---|
| 1 | 2020 | Research Scientist | 450000 | US |
| 2 | 2020 | Data Scientist | 412000 | US |
| 3 | 2020 | Director of Data Science | 325000 | US |
| 4 | 2020 | Machine Learning Engineer | 250000 | US |
| 5 | 2020 | Data Science Manager | 190200 | US |
| 6 | 2020 | Lead Data Scientist | 190000 | US |
| 7 | 2020 | Data Engineer | 188000 | US |
| 8 | 2020 | Machine Learning Engineer | 150000 | US |
| 9 | 2020 | Data Scientist | 138350 | US |
| 10 | 2020 | Machine Learning Engineer | 138000 | US |

# Visualizations

- [Tableau](Tableau) is a data visualization and business intelligence software

- Offers an intuitive interface for creating interactive dashboards, reports, and charts

- Connects to various data sources, including spreadsheets, databases, and cloud platforms

- Supports collaboration and knowledge sharing through teamwork features

# Timeline

- Feb 6th - Feb 10th: Found research paper and created slides for proposal
  - Feb 6th: Collect and explore a new dataset
  - Feb 7th - Feb 10th: Worked on slides for final project proposal and defining a new analysis task

- Feb 13th - Feb 24th: Proposal presentations

- Feb 28th - March 23rd: Implementing and developing algorithms and features
  - Farnoosh, Seyedhadi and Allie: Test dataset to get main results
  - Farnoosh: Find other approaches have been proposed and test them, finally develop a new strategy
  - Seyedhadi and Allie: Investigate, analyze and come to research findings

- March 27th - April 11th: Work on final project presentation