

# LondonLAB The Great Resignation

Group A11

# OUR TEAM

---



**Kazmer Nagy-Betegh**

knagybetegh.mam2022@london.edu



**Yuan Gao**

ygao.mam2022@london.edu



**Nikos Panayotou**

npanayotou.mam2022@london.edu



**Linli Ding**

lding.mam2022@london.edu



**Hanlu Lin**

hlin.mam2022@london.edu

# Agenda

1. Background and Approach
2. Data Analysis
3. Recommendations, Limitations and Conclusions

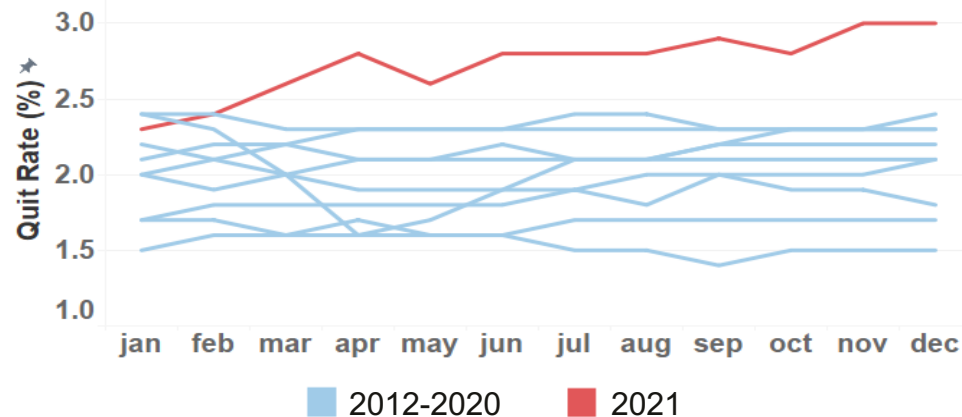
# 1. Background

Problem Statement, Project Steps, Implications

### “The Great Resignation” trend in the news



**The last 10 months of 2021 saw the highest quits rate in record**  
Monthly quits rate over 2012 – 2021 in US



Source of graph data: US Bureau of Labor Statistics – Quits Rate: <https://www.bls.gov/jlt/>

### The Great Resignation

We believe that the risk of someone changing jobs is related to both **their area of expertise** and **geographic area**. We want to use data to understand where, both geographically and in markets, we're most at risk of losing staff so that we can proactively look to engage with them and keep them at Jacobs.

Jacobs and similar companies hold a lot of their value in form of their employees' experience and expertise.

# Background Description

## Problem Statement and Project Steps

Problem Statement	Implications	Project Steps	
<ul style="list-style-type: none"> <li>From mid 2021, businesses across the world have seen record levels of voluntary resignations.</li> <li>Job openings increased because of recovery of COVID-19</li> <li>Big challenge for employers to engage and keep their employees</li> </ul>	<p><b>Keep the Employees</b></p> <p>By using data to understand where, both geographically and in markets, companies are most at risk of losing staff so that we can proactively look to engage with them and keep them at Jacobs</p>	<ul style="list-style-type: none"> <li><b>Quantify the size</b> of the problem</li> <li><b>Define</b> project scope</li> <li><b>Deep dive</b> into data and analyze</li> </ul>	
		<b>Data Scraping</b>	Create data scrapers for 6 selected companies to acquire job openings data from official website
		<b>Data Analysis</b>	Analyze scraped datasets to understand trends in companies, industrial, and geographical areas
		<b>Future Implications</b>	Define metrics to monitor in the long run for future implications in business

## Quantify The Problem

# Jacobs

Acquire  
Resignation Data  
from Jacobs

By Departments

By Location

By Job Types

By Job Functions

Understand

Compare

Clients

Competitors

Subsidiary

nationalgrid



Palantir

SWECO



CONSULTING  
TECHNOLOGY  
INNOVATION

## Project Deep Dive - Why do employees move?

### The Resignation Problem

#### Reasons Behind Movement

- Remuneration
- Promotion
- Preferred Locations
- More Exciting Projects
- Better Relationship with Clients

#### Negative Outcomes

- Decreased work quality
- Decreased efficiency
- Loss of knowledge & capability
- Loss of clients & projects
- Loss of market share

#### Solutions

Define metrics from data to monitor – **signal employee movements**

#### Internal Metrics

- Employee Profile
- 3 years without promotion
  - Lower benefits
  - Age

#### External Metrics

- Market Changes
- Market growing
  - Competitors' expanding
  - Seasonal movement

Try **keeping employees** in each scenarios

## 2. Data Analysis

Distribution, trend and category analysis



## 2. Data Analysis

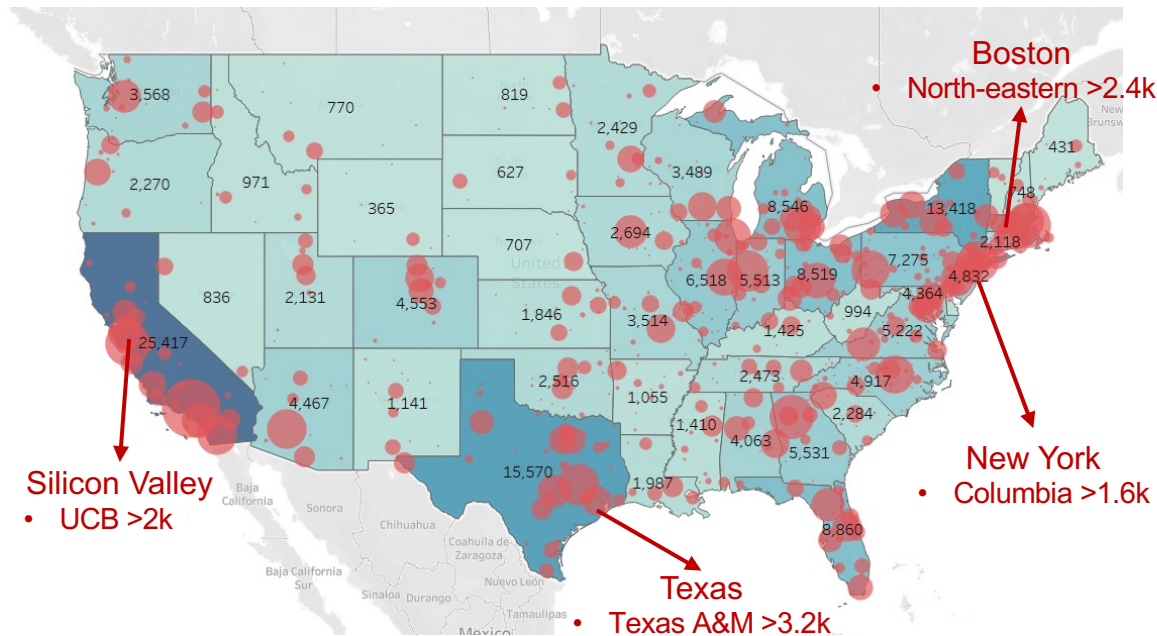
- I. Geographical analysis**
- II. Industrial analysis
- III. Company analysis
- IV. Job Description Analysis

## Geographical distribution of engineer supply and demand in the US

## How does the recent job market look for engineers in US?

## Eastern US and west coast have the most engineer graduates

## Engineer degrees awarded in 2020

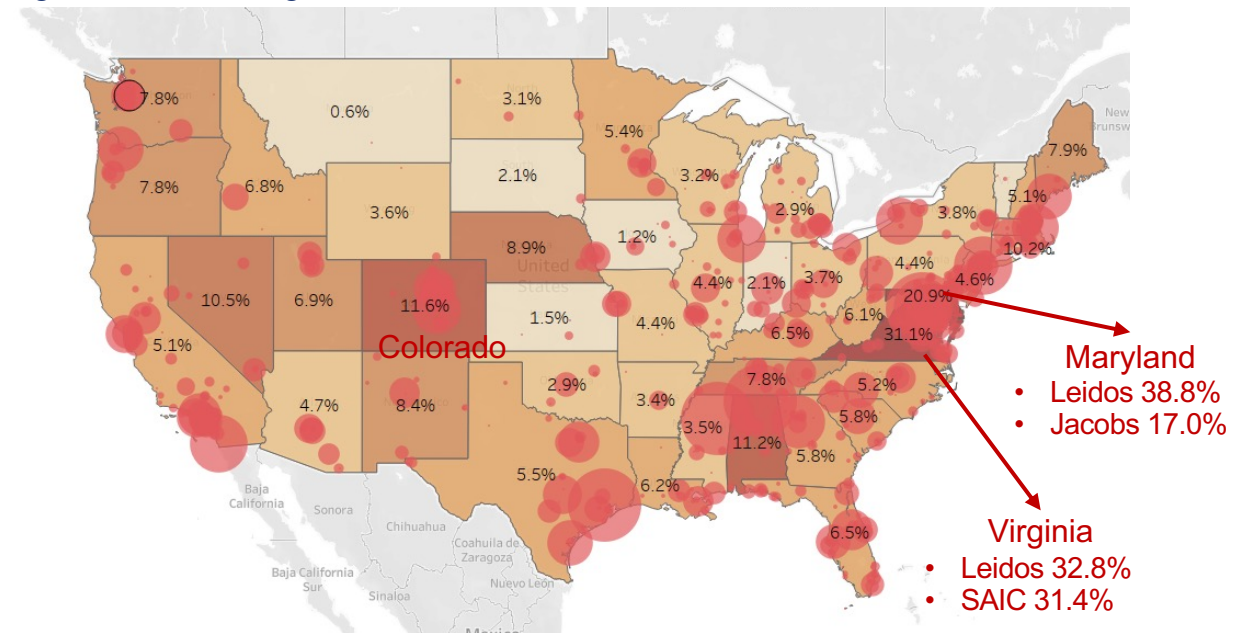


- California has the most engineer graduates among all states, followed by Texas and New York
- Top engineering hubs include Silicon Valley (4.5% engineer per employee) and Houston (2.2% engineer per employee)
- New York and Boston have one of the most engineer supply per city, mainly because of the university hubs

Source: US engineer degrees awarded data <https://datausa.io/profile/cip/engineering/>

## East coast has the most engineer job opportunities

Number of engineer posting (Dec/2021 to Apr/2022) per engineer graduate among states



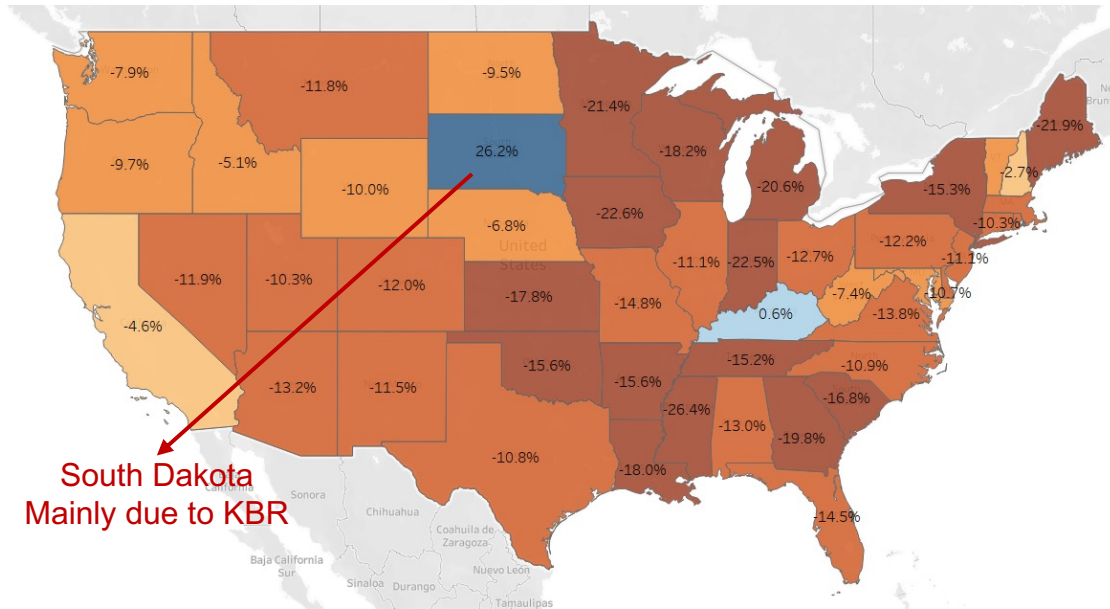
- Virginia (31.1% post per grad) has the highest job opportunities for engineers based on normalised state-wise comparison, followed by Maryland (20.9% post per grad) and Colorado (11.6% post per grad)
- Most hiring in Virginia comes from Leidos (32.8% of total posts) and SAIC (31.4% of total posts), which are both head-quartered in Virginia

Source: Jacobs internal scraped data from 12 companies

# Job opening number change from Mar/2022 to Apr/2022

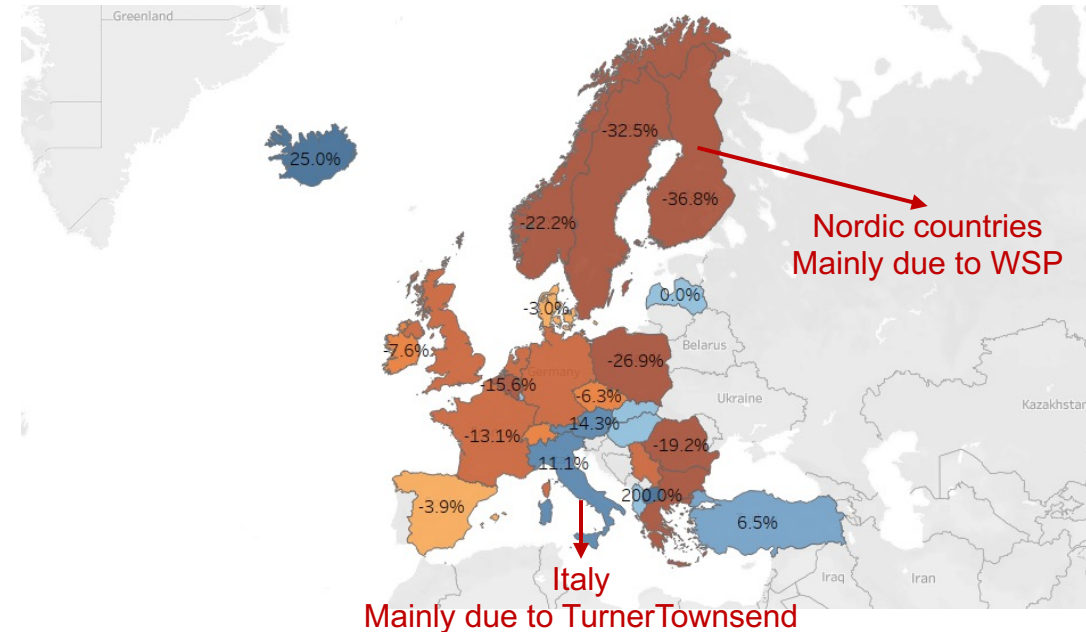
How many more/less jobs are there in main regions?

## Less job openings in most of the US states except South Dakota Job opening change over Mar-Apr 2022 in US



- South Dakota has 26.2% more openings, mainly thanks to 74% more postings added by KBR
- KBR is the main hiring company in South Dakota, accounting for >50% of total postings

## Much less job openings in Nordic countries Job opening change over Mar-Apr 2022 in EU



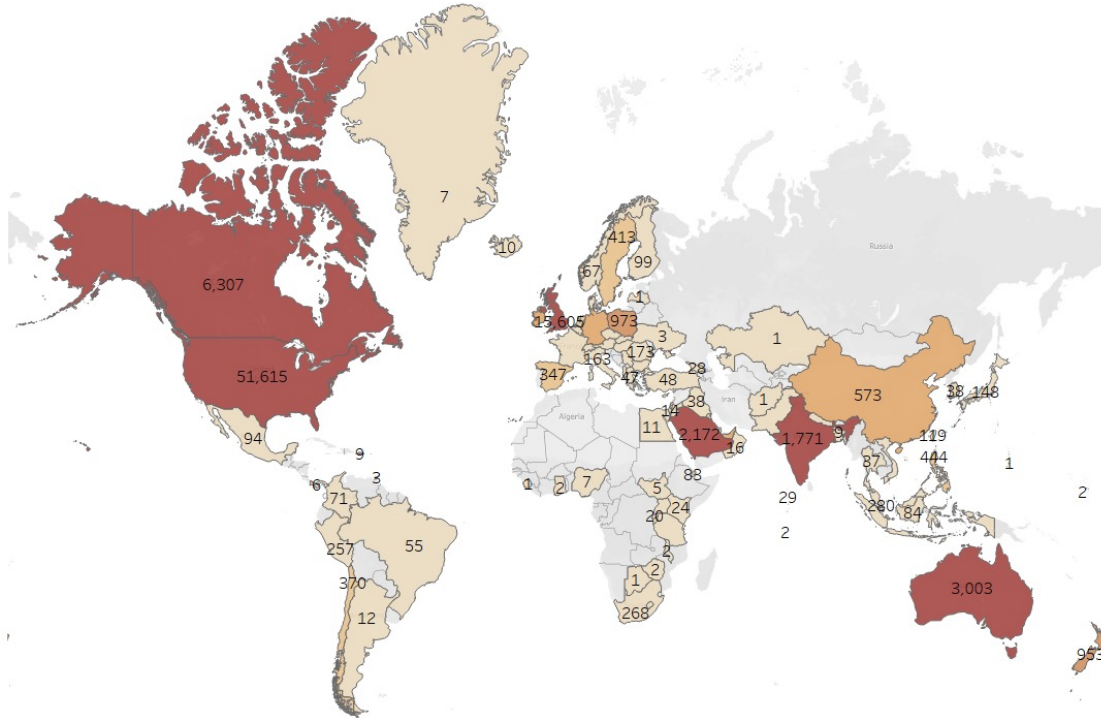
- Most increase in job openings occurs in Italy and Turkey, most decrease in job openings occurs in Nordic countries
- TurnerTownsend increased 375% openings to 15 in Italy, WSP decreased 70% openings to 195 in Nordic countries
- Outliers include Iceland (5 postings in April), North Macedonia (3 postings in April), which are excluded from the analysis

# Different geo-distribution of job postings of existing and new dataset

How can our additional scraped companies enrich Jacob's existing dataset?

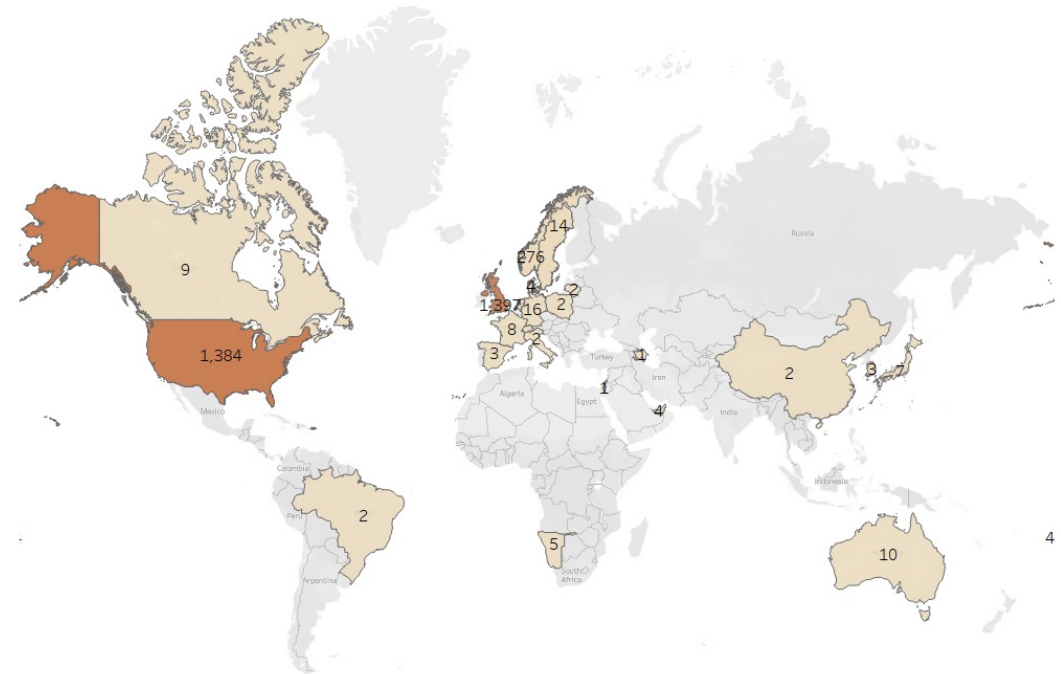
**Existing dataset shows that most hirings cluster in north America, UK and Australia**

Geo-distribution of job postings of the existing dataset



**Our new dataset supplements mainly US and Europe**

Geo-distribution of job postings of the new dataset



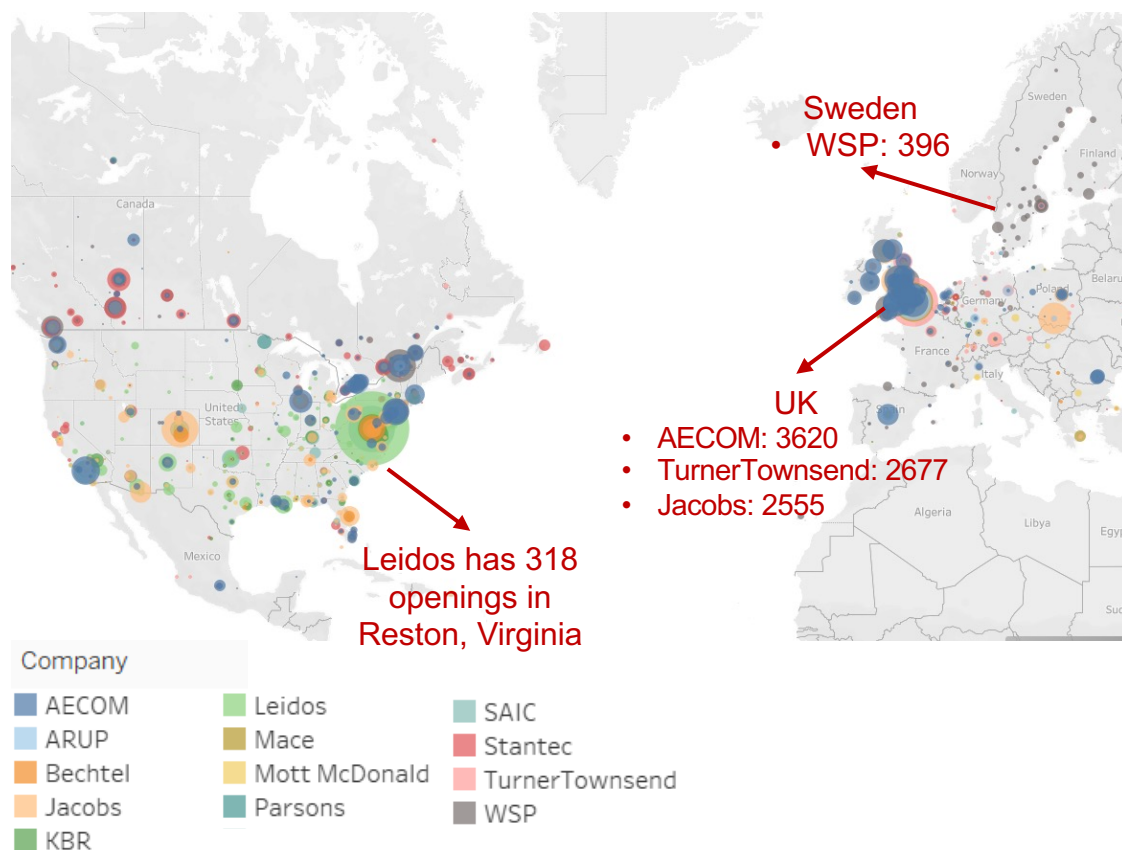


# Geographical distribution of hiring activities per company

How hiring is distributed across regions differentiated by companies?

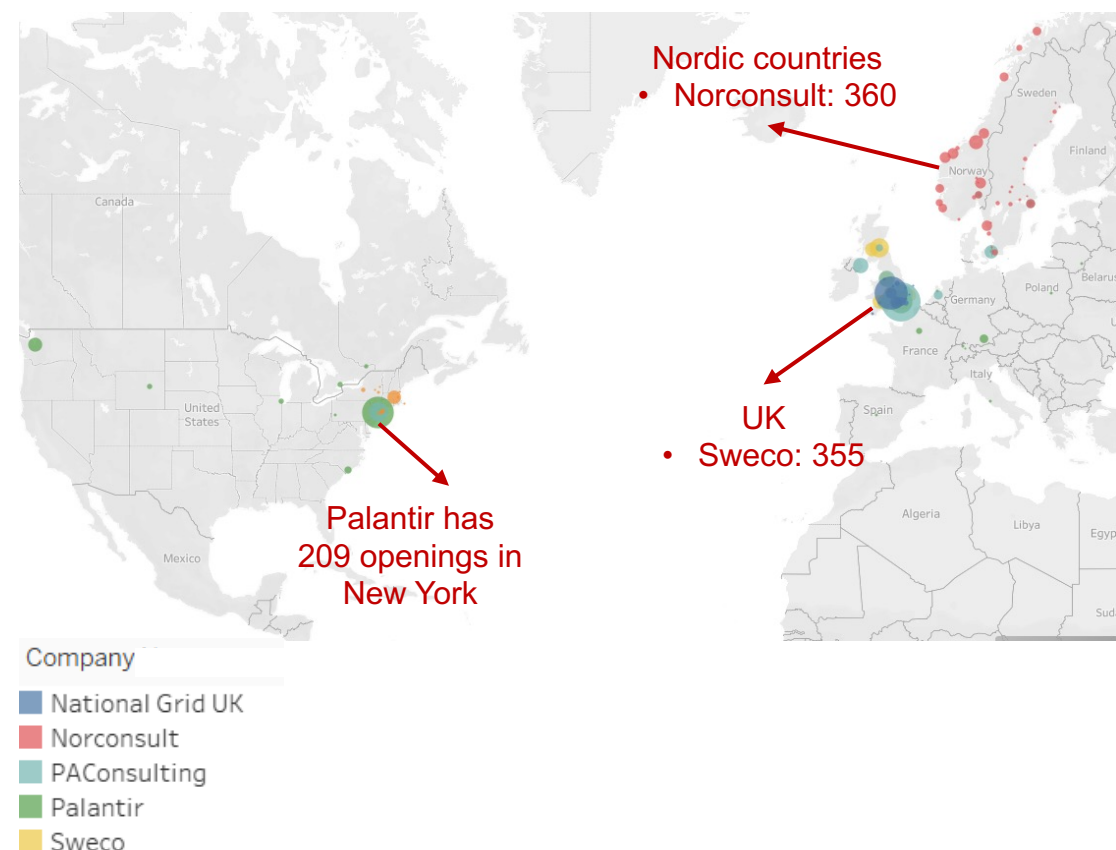
**Hirings in UK and US are diverse while some regions are dominated by only a few firms**

Job openings in May per company in existing dataset



**New dataset mostly supplements Nordic countries, UK and US east coast**

Job openings in May per company in new dataset



## 2. Data Analysis and Recommendation

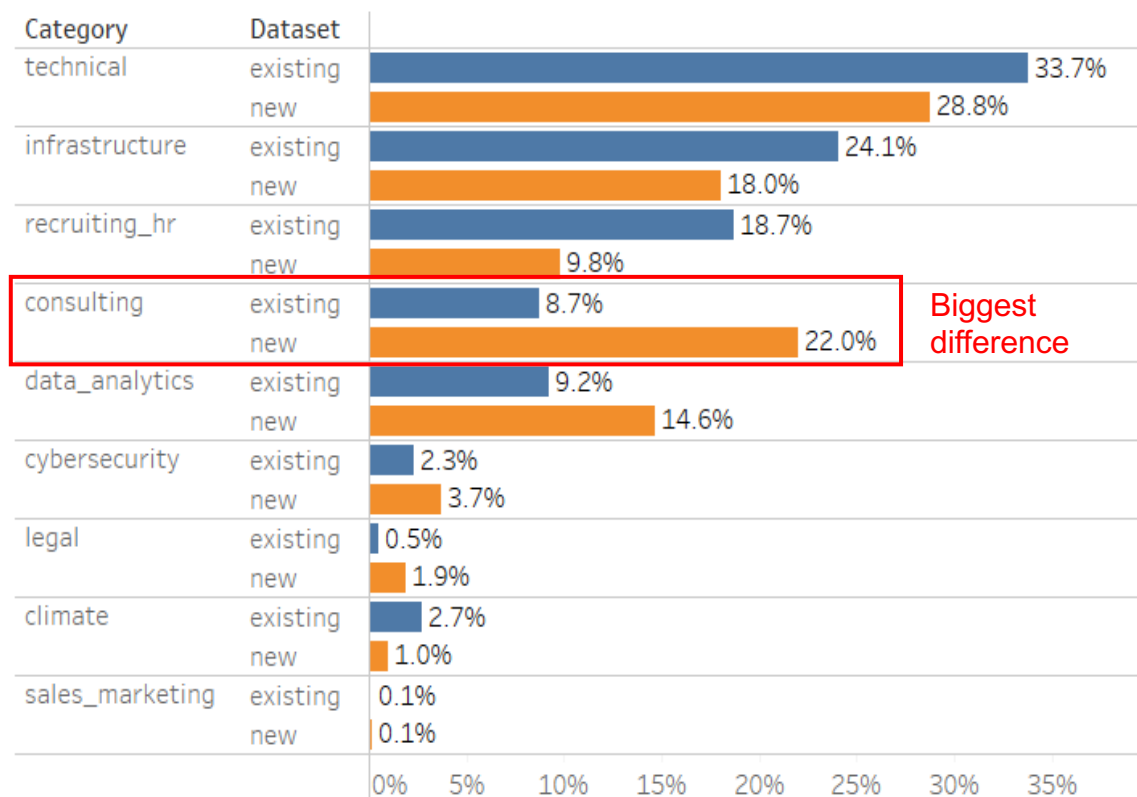
- I. Geographical analysis
- II. Industrial analysis**
- III. Company analysis
- IV. Job Description Analysis

# Comparison between existing and new dataset

How does the composition of new dataset look different from existing dataset?

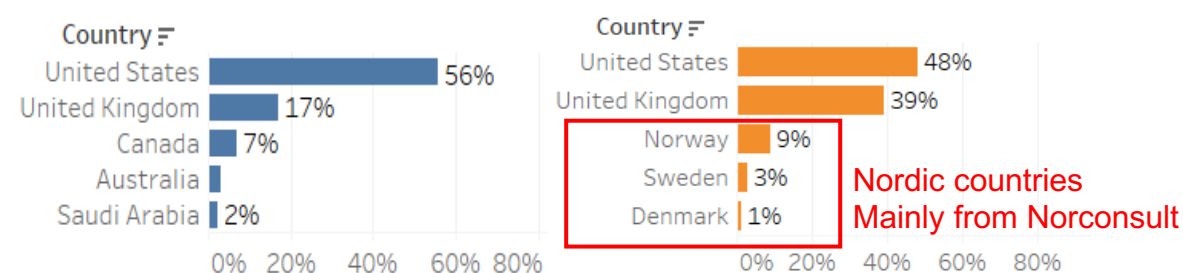
## Much more pct openings of consulting in new dataset

Comparison of job categories between existing and new dataset



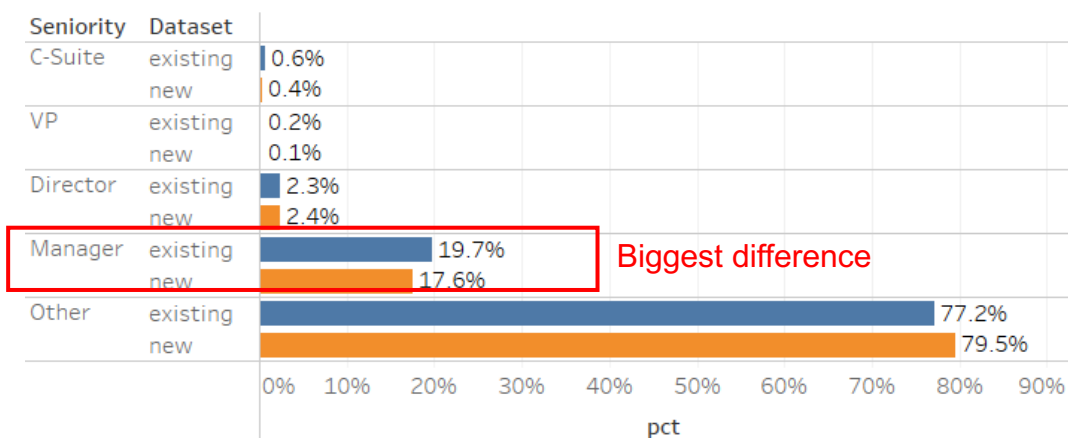
## Much more pct openings from Nordic countries in new dataset

Comparison of top 5 hiring countries between existing and new dataset



## Biggest difference exist in manager level job openings

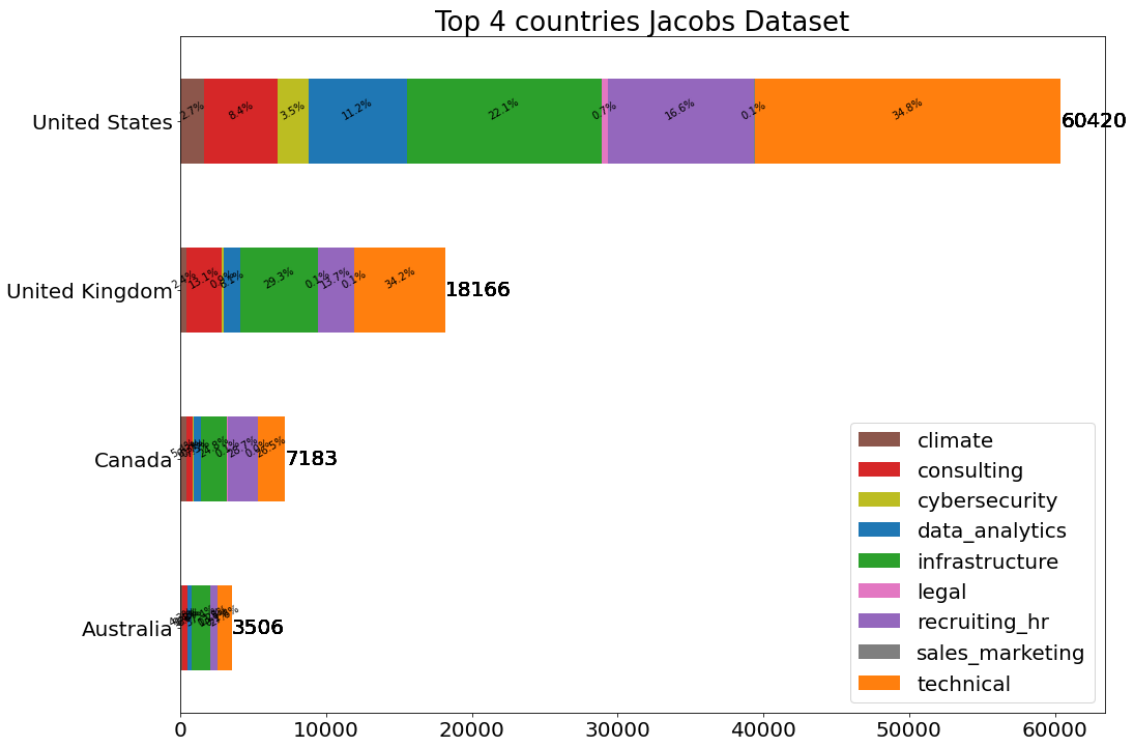
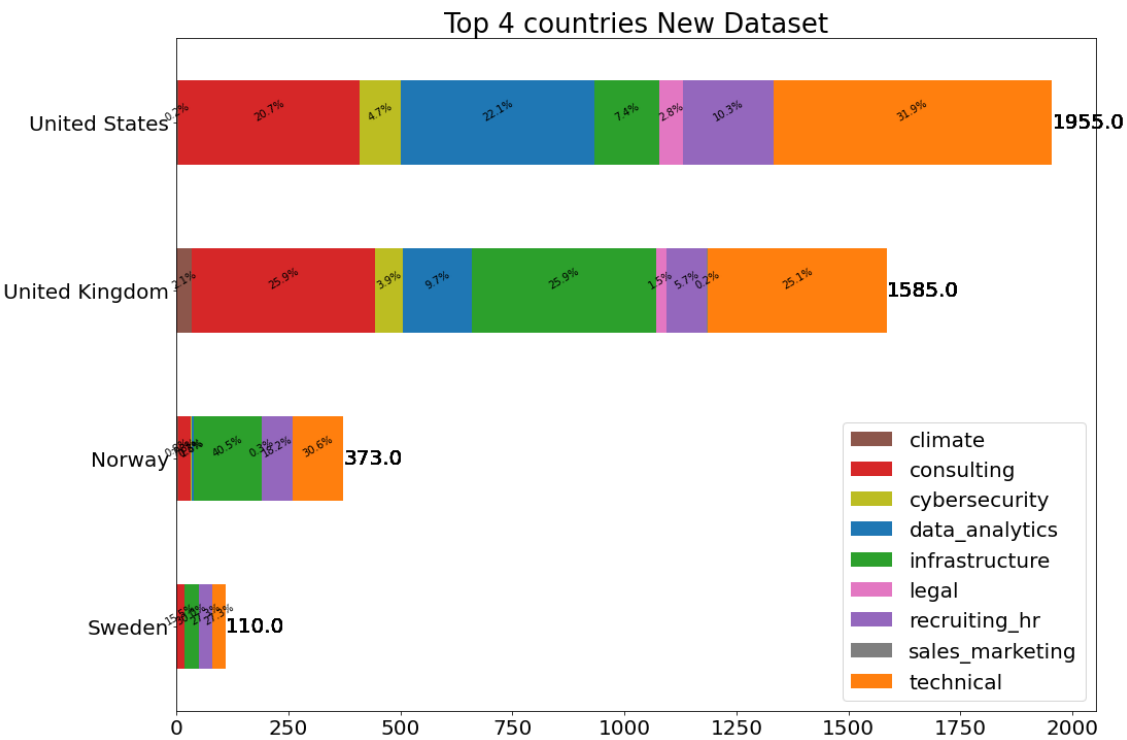
Comparison of seniority percentage between existing and new dataset



# Categories Recruited for within Countries, New vs Old Dataset

Technical, infrastructure and data analytics roles are leading the market

- **US and UK** have most job openings, followed by **Norway and Sweden**
- **Technical** roles is leading among 4 markets, especially in US and UK, while US have significant percentage of **Data Analytics** roles opening
- **US and UK** have most job openings, followed by **Canada and Australia**
- Technical roles is leading among 4 markets, especially in US and UK, while US have significant percentage of **technical** and **infrastructure** roles opening





## 2. Data Analysis and Recommendation

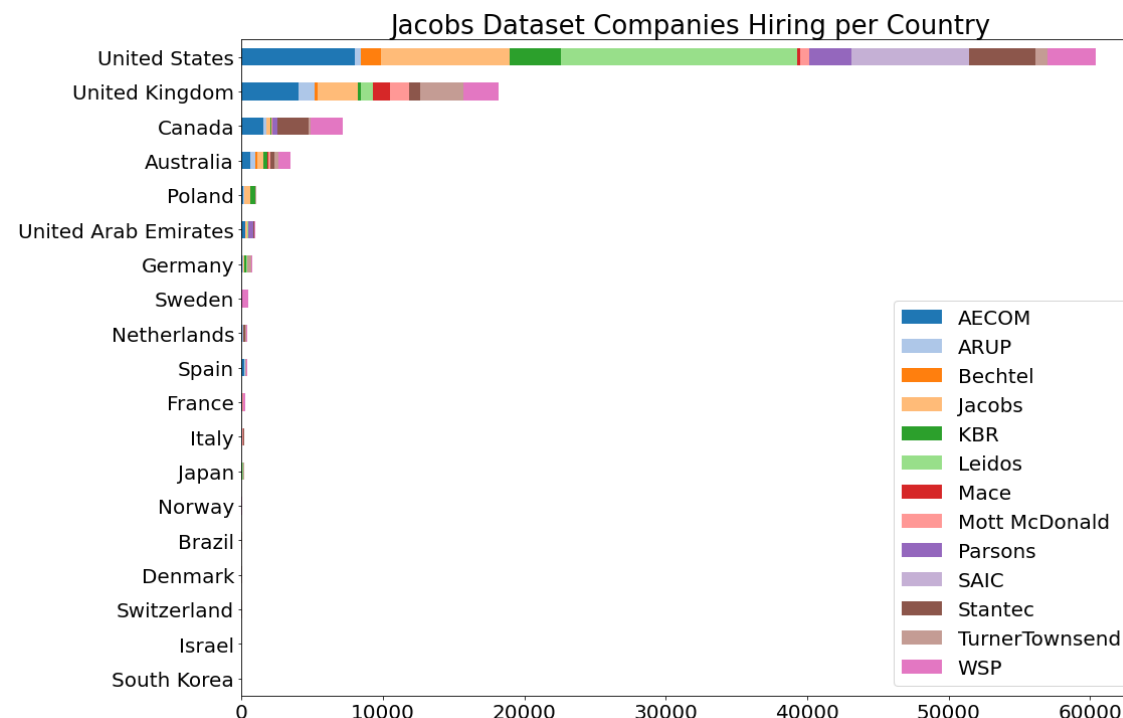
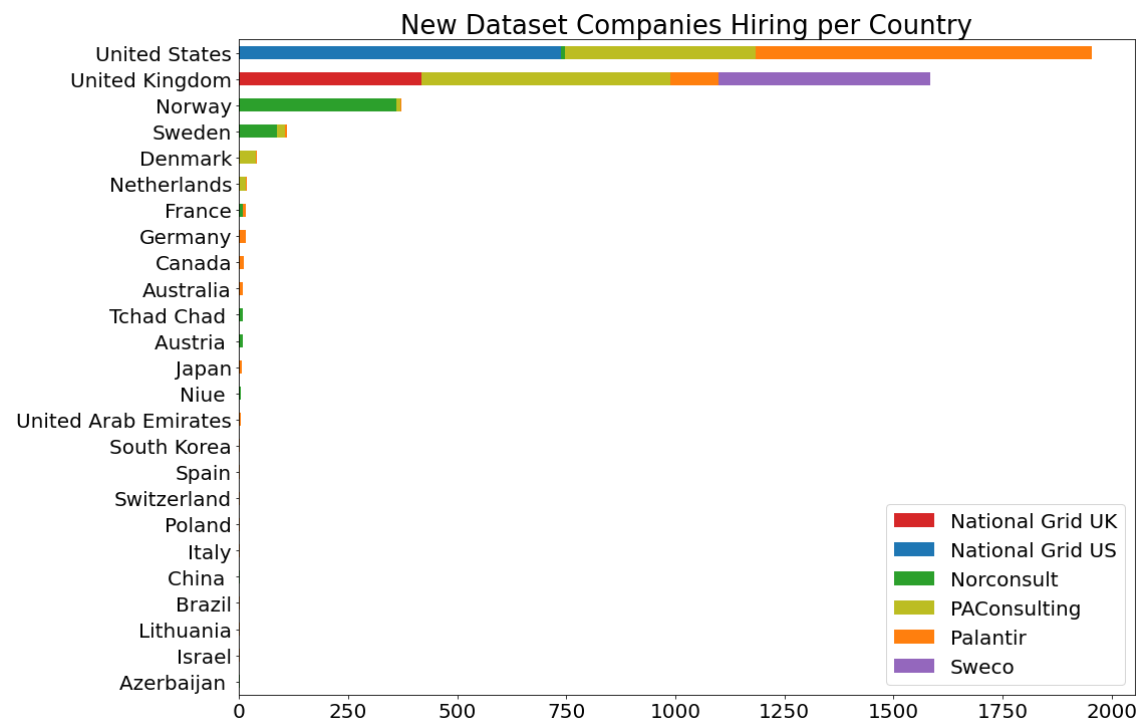
- I. Geographical analysis
- II. Industrial analysis
- III. Company analysis**
- IV. Job Description Analysis

# Companies Recruiting for within Countries, New vs Old Dataset

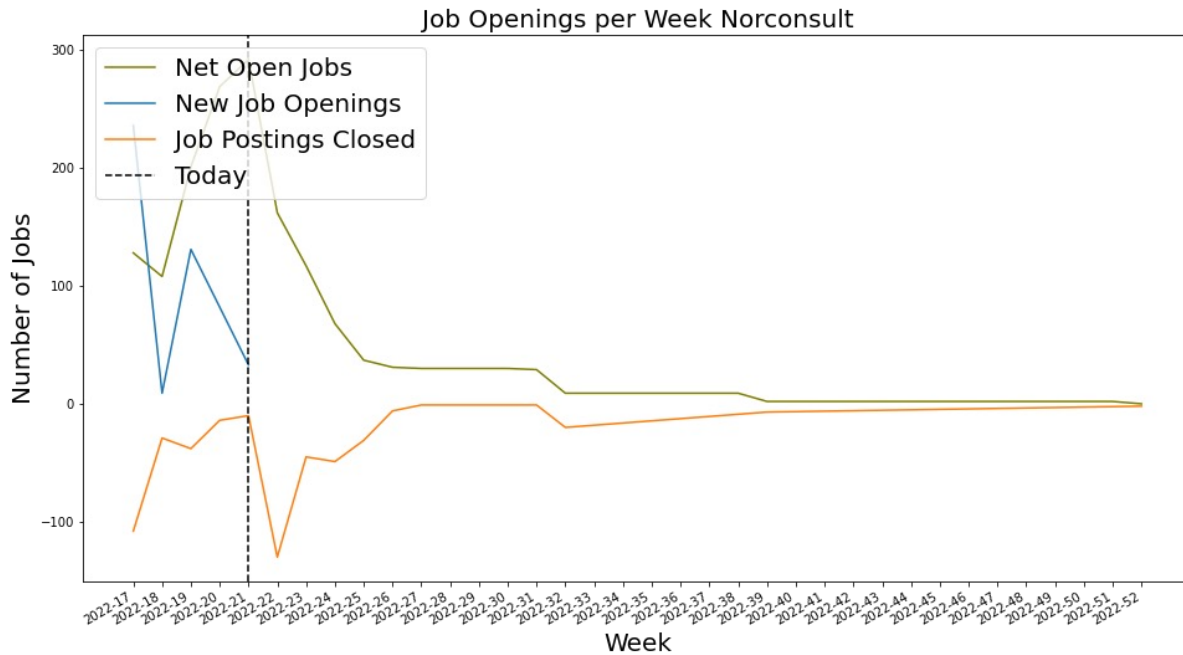
Comparison between newly scraped and original Jacobs datasets

- **Palantir** has most job openings in US
- **PA Consulting** has most job openings in UK and among the first two leading markets
- **National Grid** is hiring **more in US** than UK

- **US has more job openings** than other regions
- **Leidos** has most job openings in US
- **ARUP/Mace/Mott McDonald/TurnerTownsend** are hiring more in UK than US



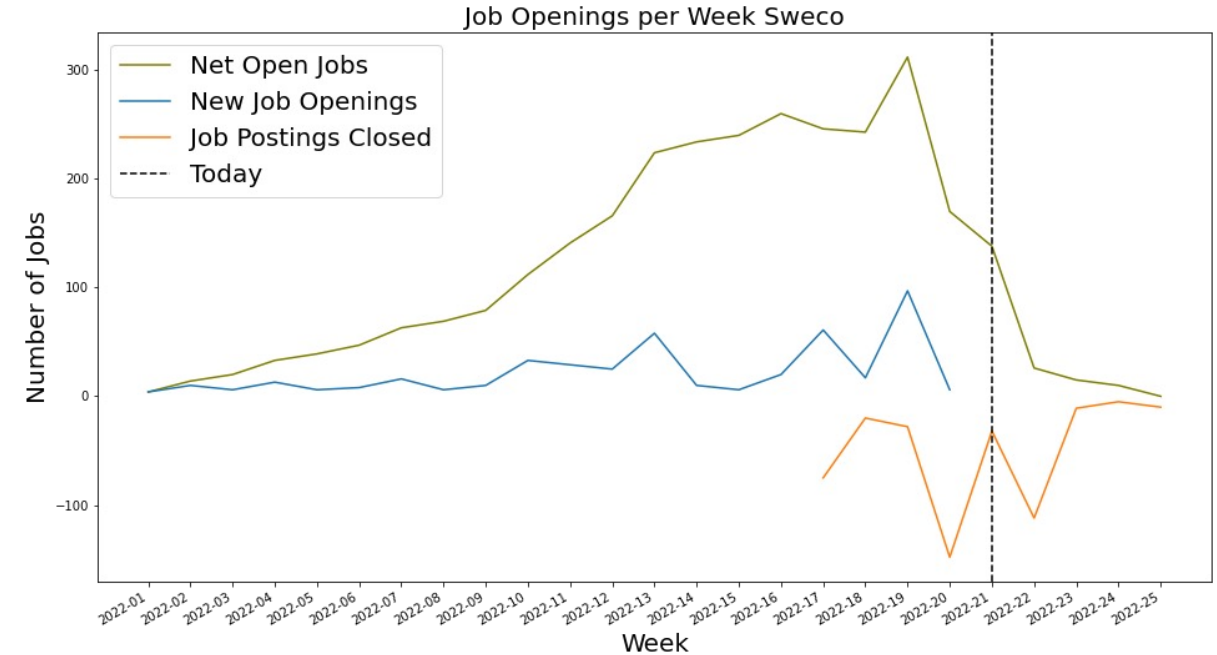
- Net Open Jobs increased from **early May until May 23<sup>rd</sup>** in 2022 and will keep decreasing afterwards
- Closed Job Postings decreased in first few weeks
- New Job Openings posted decreased in first few weeks



## Job Opening Trend Analysis – By Company

Track job openings per week for Norconsult and Sweco

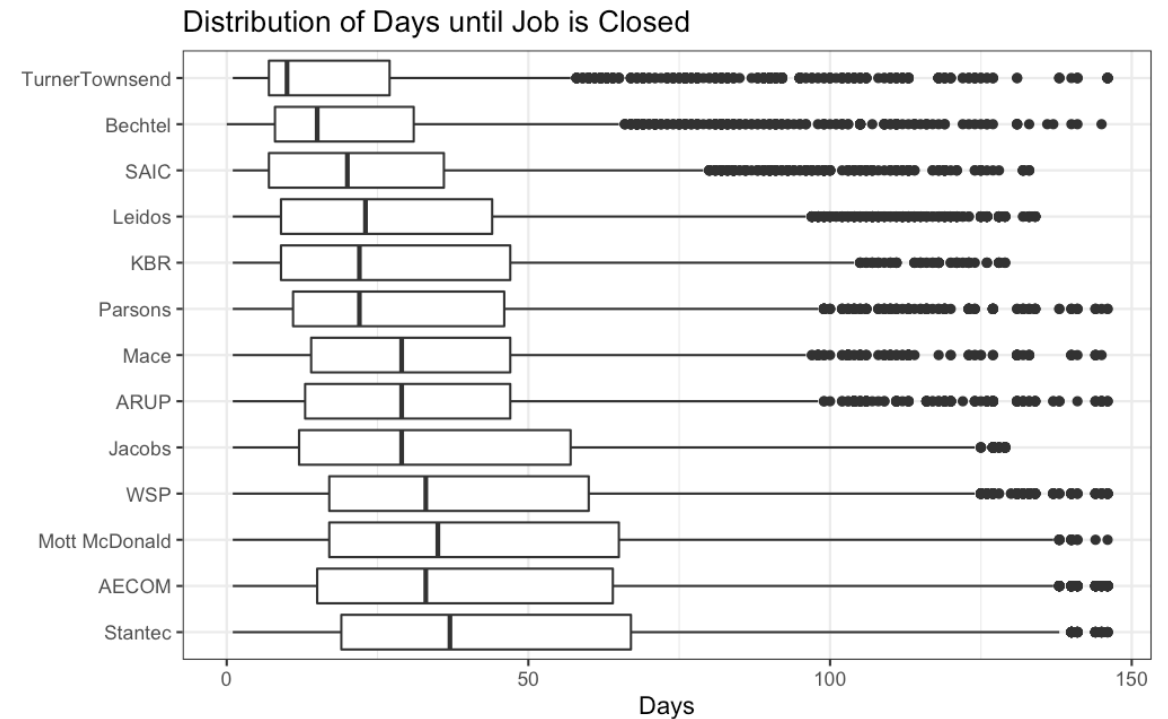
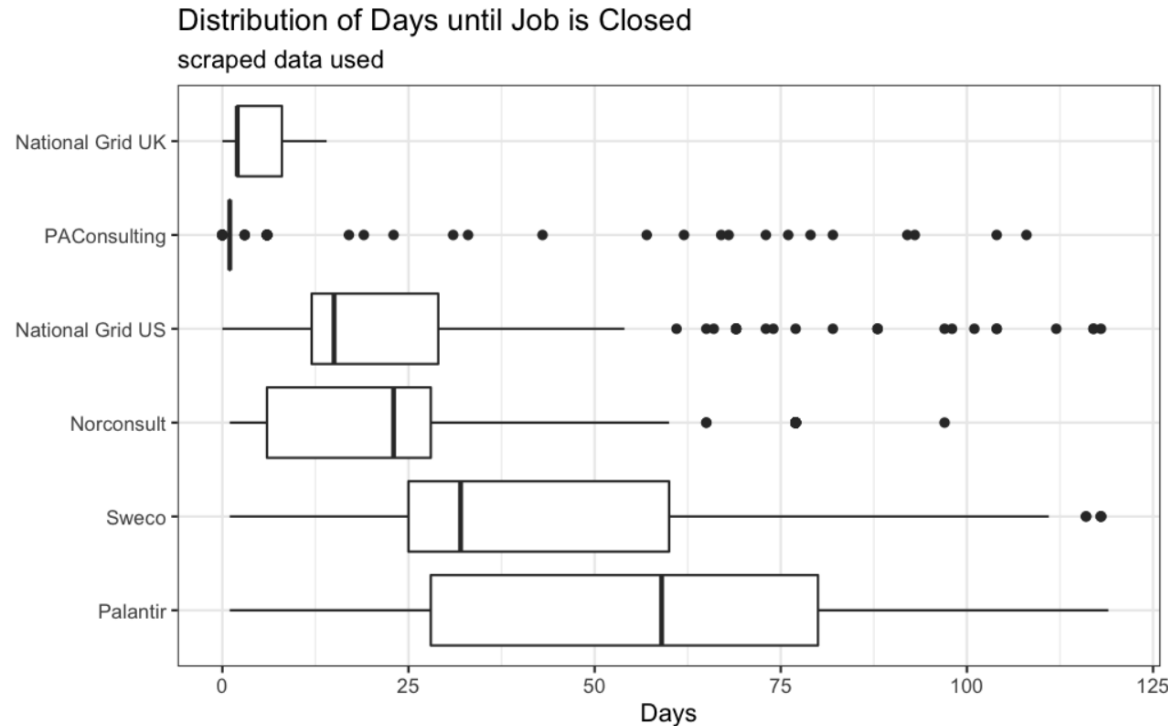
- The Net Open Jobs increased from **January 1st until the last week of May in 2022**, and will keep decreasing afterwards
- Number of New Job Openings remains stable till March 21st, dropped on May 9<sup>th</sup> in 2022



# How long jobs take to be filled? – By Company

Distribution of Duration of Job Openings for each company

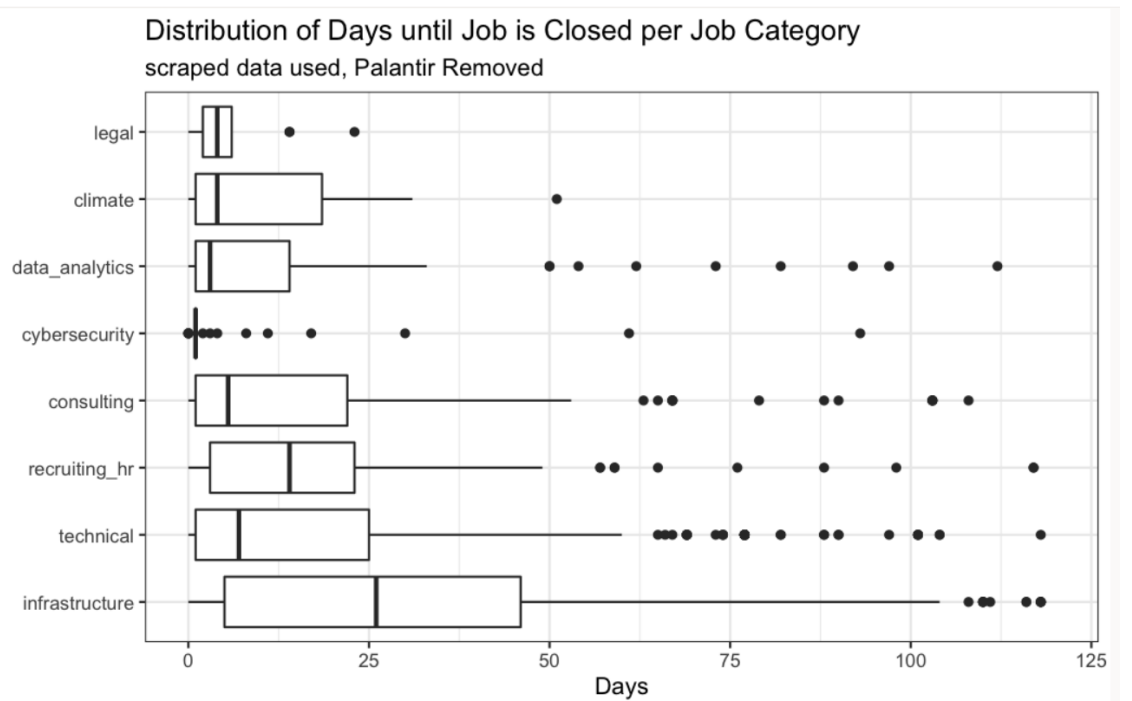
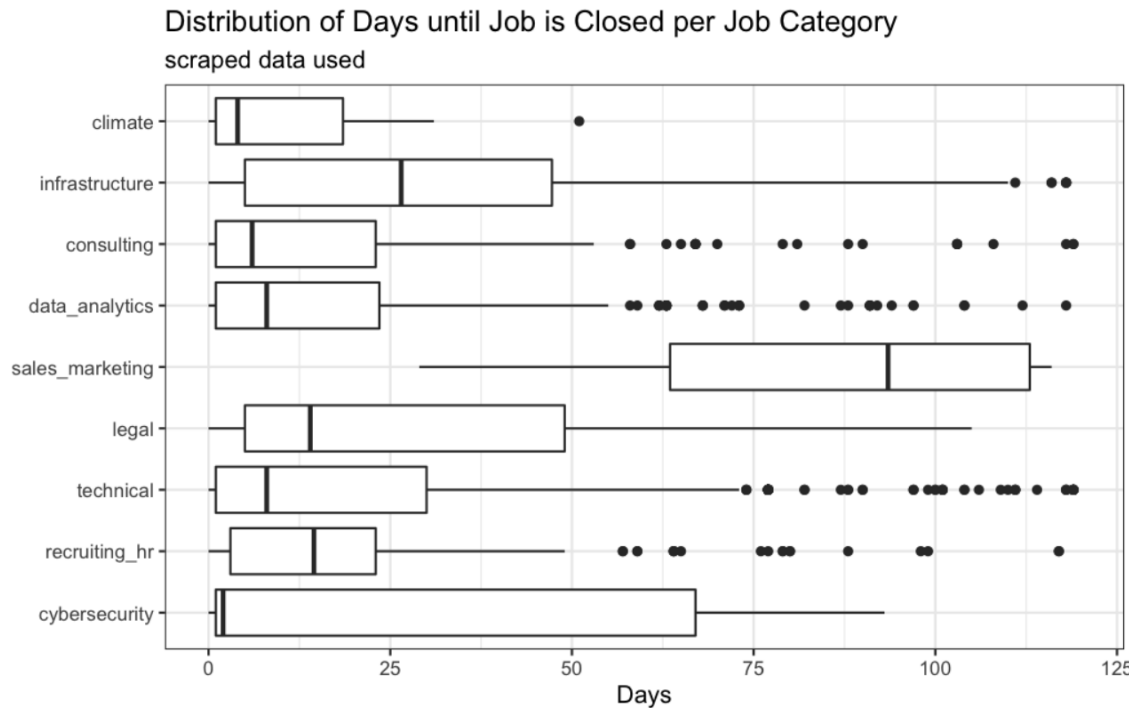
- Palantir has most job openings with longer open times
- PA Consulting and Palantir's job open time varies the most
- Job openings at National Grid UK takes short periods of time to be filled



# How long jobs take to be filled? – By Category

Distribution of Duration of Job Openings for each job category

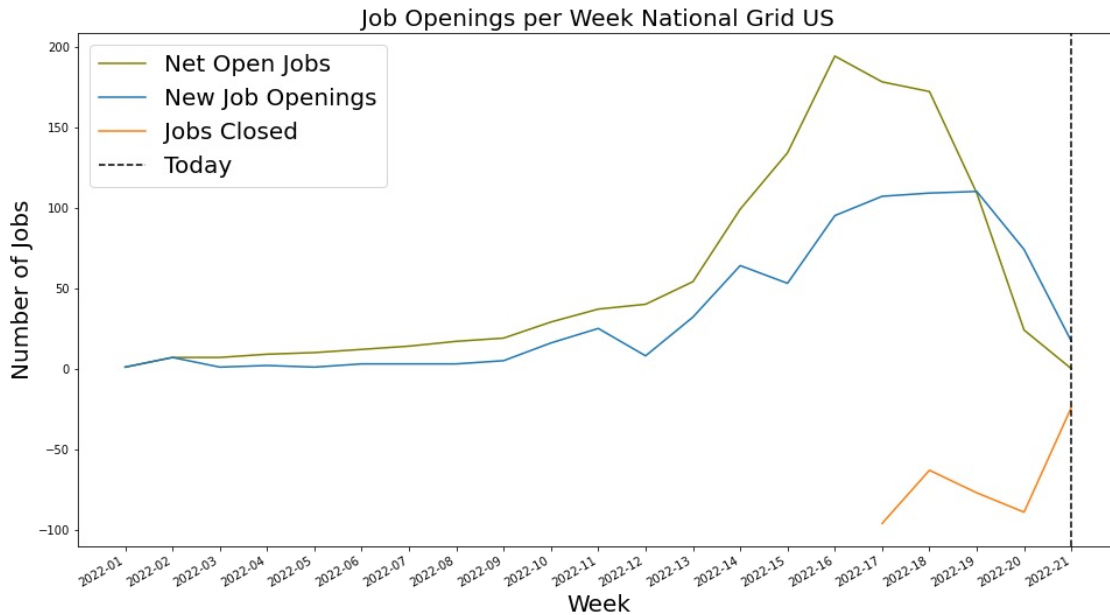
- With Palantir, Sales & Marketing role has most job openings with longer open times
- With Palantir, Data Analytics/Technical/Consulting role's job open time varies the most
- Job openings of climate/legal takes short periods of time to be filled



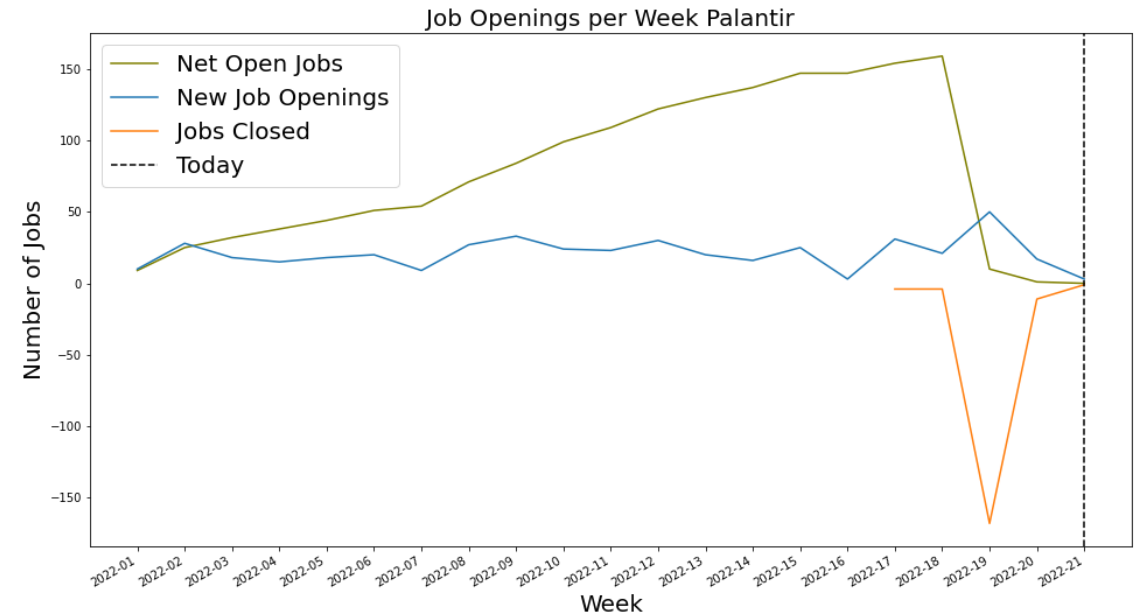
# Companies' Hiring Strategy & Dataset

Dataset shows growing hiring trend for National Grid and Palantir as mentioned in their strategy

## National Grid



## Palantir



### Strategies From Financial Statement

In 2021, employees/professionals' numbers will be **growing** because of design and implementation of new operation model

As Palantir continues to expand our business, industry verticals, and the breadth of our operations, Palantir will **hire additional employees** and **expand into new markets**

Source: National Grid plc Annual Report and Accounts 2020/21

Source: Palantir Quarterly Report ended March 31, 2022

## 2. Data Analysis and Recommendation

- I. Geographical analysis
- II. Industrial analysis
- III. Company analysis
- IV. Job Description Analysis**

Use NLP algorithms with OpenAI and Azure to Read **Faster**, Understand **Better**

### A Typical Full Job Description

#### Company Description

Almost identical within the same company.

**Subjectivity** makes it hard to compare them between companies. Descriptions written by companies themselves are typically cannot reflect satisfaction level of employees.

Employees care more about facts like company size, workplace environment, culture, etc.

#### Role Description

Key activities and accountabilities define a role in a company.

This section reveals most of the **business needs** of hiring companies.

Whether this section is in line with the expectation is critical to job seekers as well.

#### Qualifications

Required experiences and skills are also disclosed in this section.

This section can reflect **seniority** of the role and whether the company is strategically expanding.

Job seekers who are not qualified will tend not to apply.

#### Additional Information

Different companies would like to highlight different benefits for their employees.

This section is also **similar** between different jobs within a company, and usually not structured.

Key sections to describe a job



## Key Findings and Recommendations

- Data Processing:** The cleanness of texts can affect the performance of GPT-3 model on all platforms.  
**Pre-process** the texts before feeding them into the pre-trained model.
- Data Processing:** Neither key phrases nor the most frequent words in the description can be always the most important to job seekers.  
It's better to use **summarization for company description and key phrases extraction for role descriptions and qualifications** in order to maximize the simplicity and completeness.  
**Rule-based feature engineering methods** can be used to extract relevant information more precisely. (e.g., required years of experience)
- Modeling:** Parameters are important for performance. For example, parameter `maximum length` will truncate the full summarization accordingly rather than determine the degree of summarization. It can lead to severe incompleteness of contents.  
**Fine-tuning** is necessary before applying an algorithm for a specific job category or company so that comparisons are more reasonable.
- Modeling:** Azure's key phrases extraction is with relatively low level of comprehension, but it will retain the most information.  
Azure's key phrases extraction works better to generate keywords for role descriptions and qualifications.
- Cost and Compliance:** The pricing is shown in the table below. Davinci as the most complex one charges 75 times higher than Ada.  
With compliance, regional support, and enterprise-grade security services, **Azure is recommended** as the platform even for OpenAI APIs



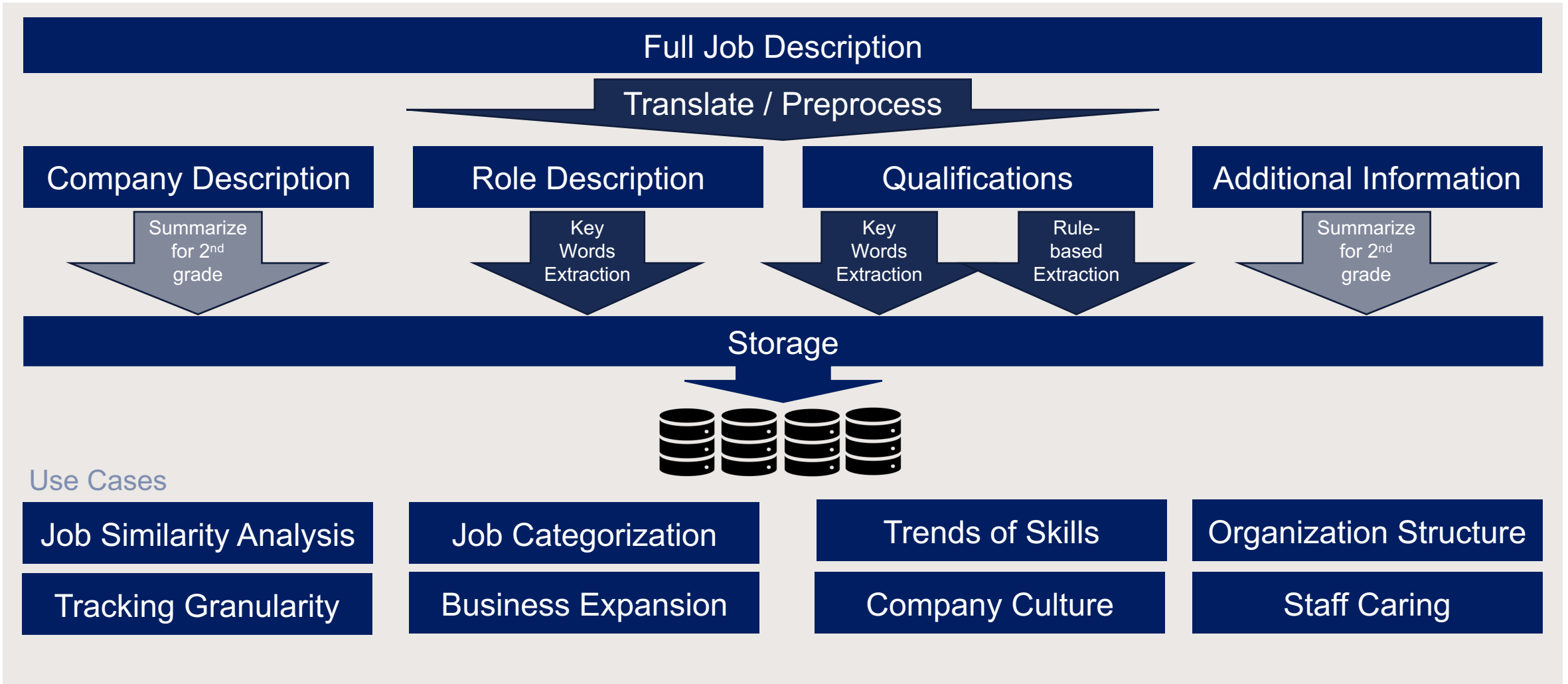
SAP, Maximo and Power Plant; Qualifications: ABET accredited Bachelor of Science Degree in Civil or Mechanical Engineering discipline required; 4+ years of experience in an engineering role required; 4+ years of AutoCAD experience required; EIT or PE certification preferred; Valid driver's license required; This position has a career path which provides for advancement opportunities within and across bands as you develop and evolve in the position; gaining experience, expertise and acquiring and applying technical skills. Candidates will be assessed and provided offers against the minimum qualifications of this role and their individual experience; National Grid is an equal opportunity employer that values a broad diversity of talent, knowledge, experience and expertise. We foster a culture of inclusion that drives employee engagement to deliver superior performance to

Platform	Simplicity (keywords)	Simplicity (summarize)	Completeness*	Ratio	Pricing
Azure	185 words out of 514	101 / 514	34 / 100	0.337	\$1 / 1K text records - (0.0M-0.5M) (\$0.0047 / 1K tokens) Degressive price to \$0.0012 / 1K tokens
Ada	216 / 514	84 / 514	45 / 100	0.536	\$0.0008 / 1K tokens
Babbage	2 / 514	153 / 514	76 / 100	0.497	\$0.0012 / 1K tokens
Curie	118 / 514	216 / 514	50 / 100	0.231	\$0.0060 / 1K tokens
Davinci	11 / 514	88 / 514	20 / 100	0.227	\$0.0600 / 1K tokens

\* A score card designed based on standard data scheme is used to evaluate completeness, according to <https://schema.org/JobPosting>  
london.edu

# Recommended Pipeline for Job Description Texts

Use NLP algorithms with OpenAI and Azure to Read **Faster**, Understand **Better**



# 3. Recommendations, Limitation and Conclusion

Project summary

Create website scrapers for the 6 companies

Dataset Overview



Time span:  
04/2022 - 05/2022



Size:  
~ 4200 job listings for 6 companies



Library / tools used:

- Scrapers (Python - BeautifulSoup, Selenium, etc.)
- Visualizations (Tableau, R)



Data Sources (Websites)



Dictionary

Column	Meaning	
ID	Unique ID generated by us to track each job openings	Existing features
Company	Company that posted this job	
...	...	
Job family	Classification of jobs according to Jacob's internal mapping	Feature to be developed
Category	Expertise or sub-industry	Newly developed features
Seniority	Rank or experience level	
Job description full	The combined descriptive text (e.g., role description, qualification)	

What are the limitations of this implementation and the next steps



## Limitations

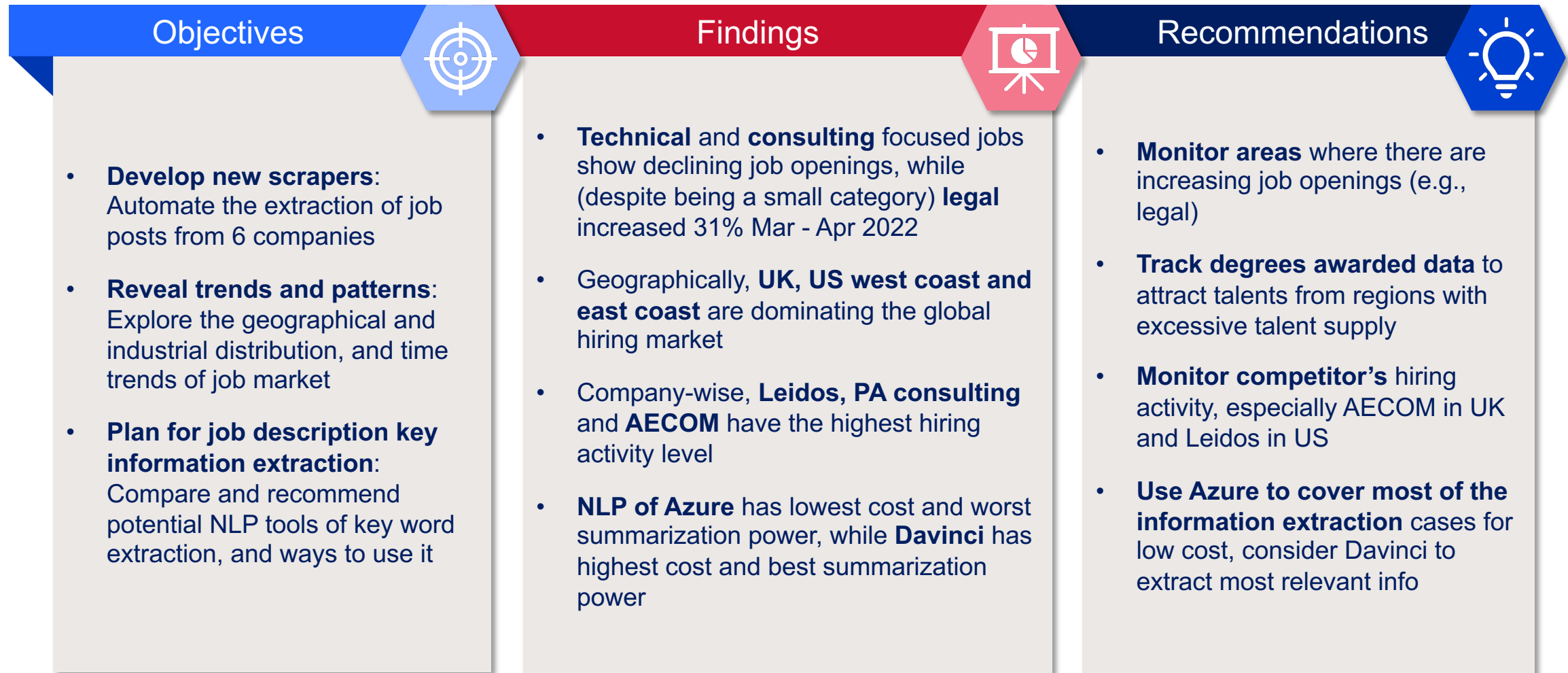
- Our analysis focuses on geography & area of expertise, **excluding salary and other perks**
- Some jobs have **no expiry date** (e.g., all posts in Palantir)
- **Missing data** in some fields (e.g., geographical level only to country/state, missing city)
- Limited analysis on 6 new companies because of **short scraping time**

## Next Steps

- **Integrate** new scrapers and datasets to existing dataset
  - **Add new features** (e.g., seniority) to existing dataset
  - **Add “Job family”** feature in the new dataset according to Jacob’s internal mapping
- **Deploy** language modules for executing keyword extraction from job description



Monitor metrics to respond to dynamic market



Monitor metrics to respond to dynamic market

## Recommended metrics to monitor

### Number of Open Jobs

Reveal the hiring activity, sudden changes in hiring activity.  
Can be filtered by Location, job category, Company  
Once enough scraping has been done, checks can be defined to flag deviations from default.

### How Long do Jobs take to close

This can reveal information about how quickly companies need to fill jobs and their ability to attract candidates.

### Talent Available vs Number of Jobs posted in an Area

Using degrees awarded data estimate the talent pool against the number of jobs posted in a location

### Skills trending within Job Descriptions

See the type of skills that are becoming sought after by extracting the keywords from Job Descriptions

### Monitor Job benefits (Salary, Num. Vacations days, Company perks)

Through NLP monitor the currently trending benefits companies offer and track how Jacobs compares. Identify the ones people care most about. (i.e. Remote working has become a trend after the pandemic)

# Thank you!

Group A11