Course Name: Business Intelligence Lab Course Code:20CSP-421

EXPERIMENT:2.1

AIM: Implement OLAP (Online Analytical Processing) for Business Insights.

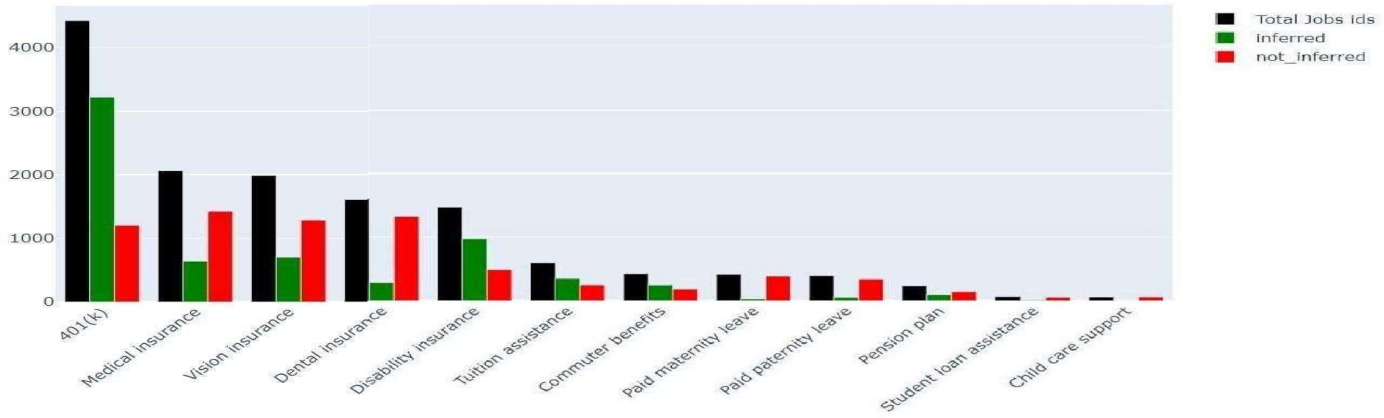
Objective: • The experiment involves setting up an OLAP system to analyse a dataset and extract meaningful insights.

- The dataset can be sourced from various data repositories within the organization, such as transactional databases or data warehouses.
- The OLAP system will enable users to explore the data from multiple dimensions or perspectives, facilitating deeper analysis and visualization

CODE:

Loading the Data from various sources

```
# Benfits offered for a particular job
benefits_df = pd.read_csv('benefits.csv')

# Description of companies
companies_df = pd.read_csv('companies.csv')

# Areas of working of companies
company_industries_df = pd.read_csv('company_industries.csv')

# Speicalites of a company
company_specialities_df = pd.read_csv('company_specialities.csv')

# Timestamps of employees & followers count of companies
employee_counts_df = pd.read_csv('employee_counts.csv')

# Job ID and Industry ID details
job_industries_df = pd.read_csv('job_industries.csv')

# Description of job posted
job_postings_df = pd.read_csv('job_postings.csv')

# Skills related to a job
job_skills_df = pd.read_csv('job_skills.csv')
```



Name: SUMIT KUMAR UID:20BCS8226

Course Name: Business Intelligence Lab Course Code:20CSP-421

TO GET INFO ABOUT JOB INFERRING

```
1 benefits_df.isnull().sum()
```

```
job_id      0
inferred    0
type        0
dtype: int64
```

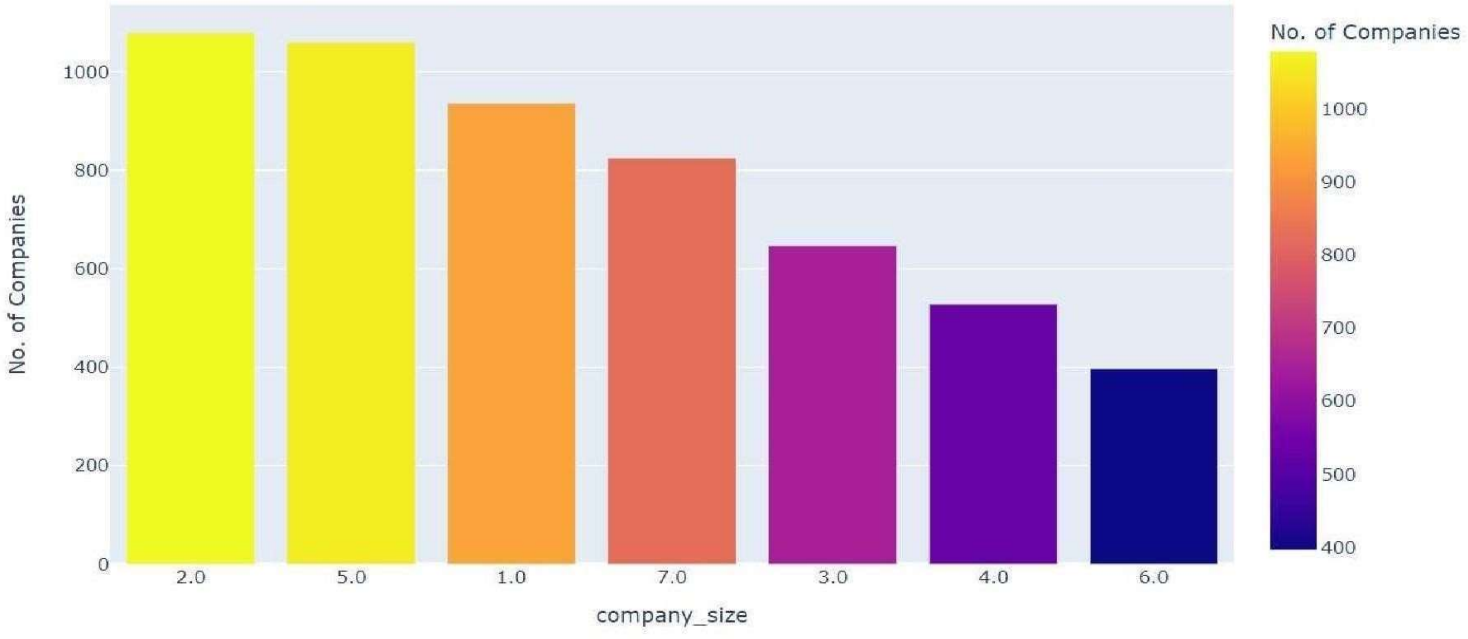
```
1 jobs_inferred_df = benefits_df[benefits_df['inferred']==1]
2 jobs_inferred_df.reset_index(inplace=True,drop=True)
3 jobs_inferred_df.head(3)
```

	job_id	inferred	type
0	3691763971	1	Dental insurance
1	3691763971	1	Disability insurance
2	3691763971	1	401(k)

How many jobs are inferred?

```
1 jobs_inferred_df['job_id'].nunique()
```

```
4100
```



Name: SUMIT KUMAR UID:20BCS8226

Course Name: Business Intelligence Lab Course Code:20CSP-421

What are the benefits offered by these jobs?

```
[ ]  1 jobs_inferred_df['type'].unique()
```

```
array(['Dental insurance', 'Disability insurance', '401(k)',
       'Medical insurance', 'Vision insurance', 'Pension plan',
       'Tuition assistance', 'Paid maternity leave',
       'Paid paternity leave', 'Commuter benefits',
       'Student loan assistance'], dtype=object)
```

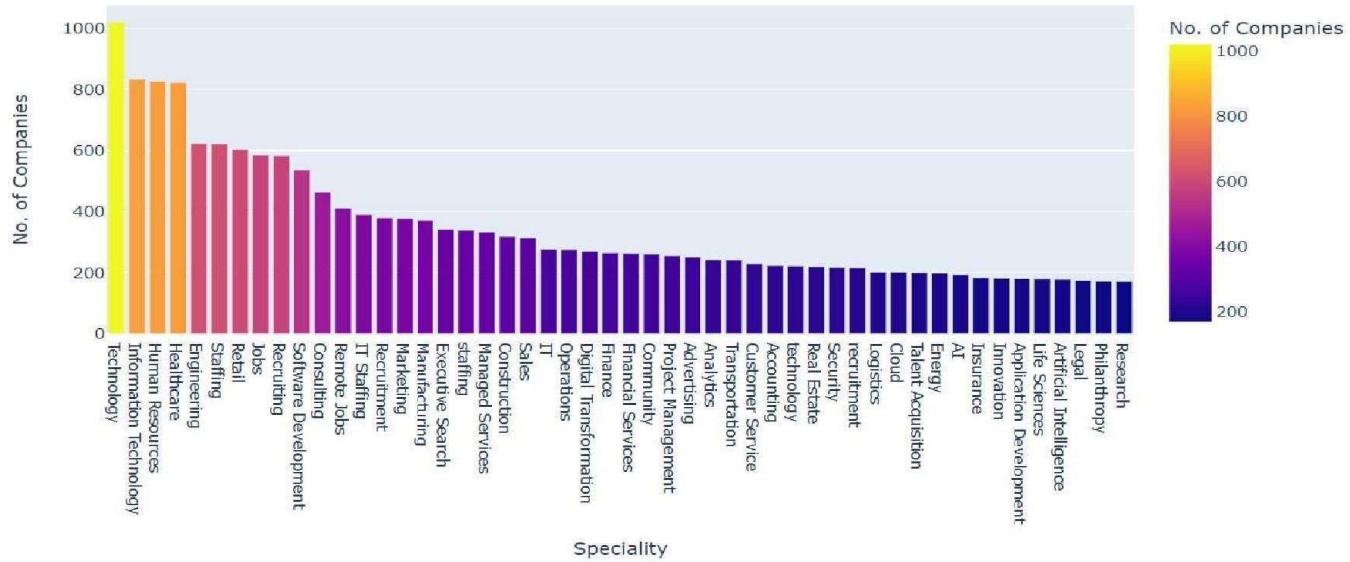
```
[ ]  1 jobs_inferred_df['type'].nunique()
```

```
11
```

Which Benefit type got more job infers?

```
1 jobs_inferred_df['type'].value_counts()
```

```
401(k)                    3221
Disability insurance       978
Vision insurance           703
Medical insurance          640
Tuition assistance         349
Dental insurance           283
Commuter benefits          248
Pension plan                93
Paid paternity leave        55
Paid maternity leave        29
Student loan assistance     11
Name: type, dtype: int64
```



Which Job gets most infered ?

or

Which job offers more benefits?

```
[ ]  1 jobs_inferred_df['job_id'].value_counts().head()
```

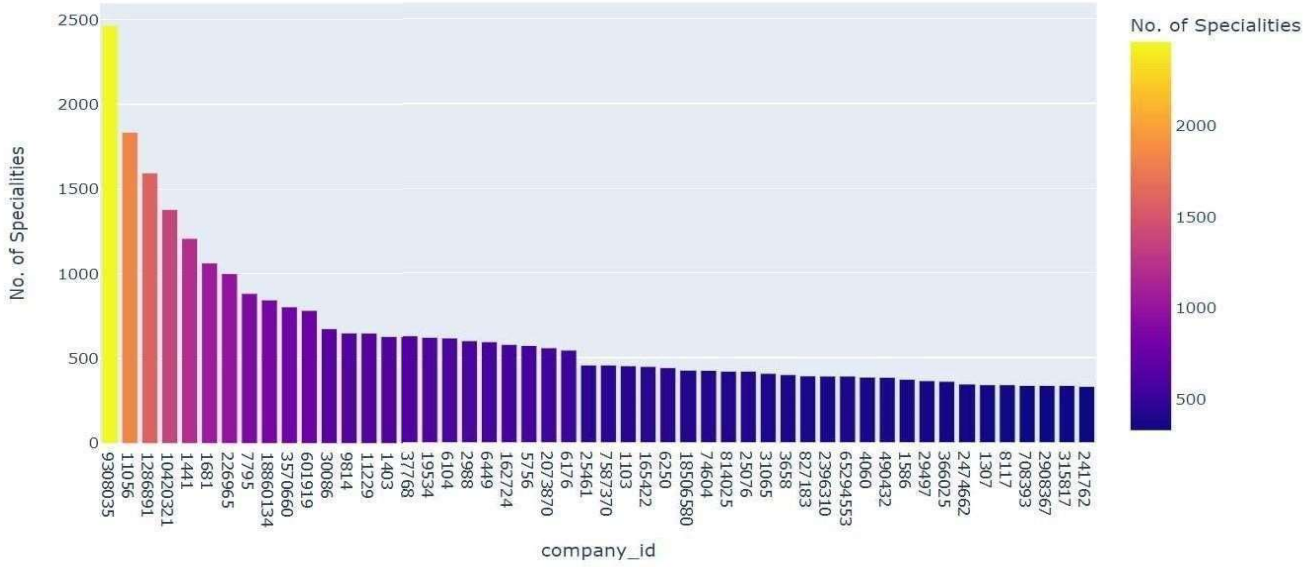
```
3690868934    7
3701305249    7
3699407132    6
3693041918    6
3697378492    6
Name: job_id, dtype: int64
```

What are those benefits of jobs which got inferred the most?

```
[ ]  1 jobs_inferred_df[jobs_inferred_df['job_id']==3690868934]['type'].values
```

```
array(['Medical insurance', 'Dental insurance', 'Vision insurance',
       '401(k)', 'Paid maternity leave', 'Paid paternity leave',
       'Commuter benefits'], dtype=object)
```

Course Name: Business Intelligence Lab Course Code:20CSP-421



3. Learning Outcomes:

- Understanding of OLAP: Participants will comprehend the concepts and principles of OLAP and its significance in business analytics and decision-making processes.

- Technical Skills: Participants will acquire hands-on experience in setting up and configuring an OLAP system, designing dimensional models, creating cubes, and loading data for analysis.

- Data Analysis Techniques: Participants will learn various OLAP analysis techniques, such as slicing and dicing, drill-down, and roll-up operations, to explore data from different dimensions and levels of detail.

- Business Insights: By analyzing the dataset using OLAP, participants will extract meaningful insights and understand the potential impact of such insights on business operations, performance, and strategic decision-making.