```python
# This Python 3 environment comes with many helpful analytics
libraries installed
# It is defined by the kaggle/python Docker image:
https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the read-only "../input/"
directory
# For example, running this (by clicking run or pressing Shift+Enter)
will list all files under the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/)
that gets preserved as output when you create a version using "Save &
Run All"
# You can also write temporary files to /kaggle/temp/, but they won't
be saved outside of the current session
```

```
/kaggle/input/data-science-job-postings-and-skills/job_summary.csv
/kaggle/input/data-science-job-postings-and-skills/job_skills.csv
/kaggle/input/data-science-job-postings-and-skills/job_postings.csv
```

```python
job_skills = pd.read_csv("/kaggle/input/data-science-job-postings-and-
skills/job_skills.csv")
job_summary = pd.read_csv("/kaggle/input/data-science-job-postings-
and-skills/job_summary.csv")
job_postings = pd.read_csv("/kaggle/input/data-science-job-postings-
and-skills/job_postings.csv")

job_skills.head()
```

```
                                          job_link  \
0  https://www.linkedin.com/jobs/view/senior-mach...
1  https://www.linkedin.com/jobs/view/principal-s...
2  https://www.linkedin.com/jobs/view/senior-etl-...
3  https://www.linkedin.com/jobs/view/senior-data...
4  https://www.linkedin.com/jobs/view/lead-data-e...


                                        job_skills
0  Machine Learning, Programming, Python, Scala, ...
1  C++, Python, PyTorch, TensorFlow, MXNet, CUDA,...
2  ETL, Data Integration, Data Transformation, Da...
3  Data Lakes, Data Bricks, Azure Data Factory Pi...
4  Java, Scala, Python, RDBMS, NoSQL, Redshift, S...
```

```
job_summary.head()

                                            job_link  \
0  https://www.linkedin.com/jobs/view/senior-mach...
1  https://www.linkedin.com/jobs/view/principal-s...
2  https://www.linkedin.com/jobs/view/senior-etl-...
3  https://www.linkedin.com/jobs/view/senior-data...
4  https://www.linkedin.com/jobs/view/lead-data-e...

                                         job_summary
0  Company Description\nJobs for Humanity is part...
1  Who We Are\nAurora (Nasdaq: AUR) is delivering...
2  Location: New York City, NY\nPosition Summary\...
3  Responsibilities:\nCandidate must have signifi...
4  Dice is the leading career destination for tec...

job_postings.head()

                                            job_link  \
0  https://www.linkedin.com/jobs/view/senior-mach...
1  https://www.linkedin.com/jobs/view/principal-s...
2  https://www.linkedin.com/jobs/view/senior-etl-...
3  https://www.linkedin.com/jobs/view/senior-data...
4  https://www.linkedin.com/jobs/view/lead-data-e...

              last_processed_time    last_status got_summary got_ner  \
0  2024-01-21 08:08:48.031964+00  Finished NER            t       t
1  2024-01-20 04:02:12.331406+00  Finished NER            t       t
2  2024-01-21 08:08:31.941595+00  Finished NER            t       t
3  2024-01-20 15:30:55.796572+00  Finished NER            t       t
4  2024-01-21 08:08:58.312124+00  Finished NER            t       t

  is_being_worked                                   job_title  \
0               f           Senior Machine Learning Engineer
1               f  Principal Software Engineer, ML Accelerators
2               f           Senior ETL Data Warehouse Specialist
3               f     Senior Data Warehouse Developer / Architect
4               f                          Lead Data Engineer

             company       job_location  first_seen search_city  \
0   Jobs for Humanity     New Haven, CT  2024-01-14  East Haven
1             Aurora  San Francisco, CA  2024-01-14   El Cerrito
2  Adame Services LLC       New York, NY  2024-01-14   Middletown
3    Morph Enterprise    Harrisburg, PA  2024-01-12     Lebanon
4               Dice          Plano, TX  2024-01-14     McKinney

   search_country               search_position  job_level job_type

0  United States  Agricultural-Research Engineer  Mid senior   Onsite

1  United States                  Set-Key Driver  Mid senior   Onsite
```

```
2   United States       Technical Support Specialist   Associate    Onsite

3   United States                          Architect  Mid senior    Onsite

4   United States       Maintenance Data Analyst  Mid senior    Onsite

job_skills_position = job_skills.merge(job_postings, on = 'job_link')
job_skills_position

                                                   job_link  \
0      https://www.linkedin.com/jobs/view/senior-mach...
1      https://www.linkedin.com/jobs/view/principal-s...
2      https://www.linkedin.com/jobs/view/senior-etl-...
3      https://www.linkedin.com/jobs/view/senior-data...
4      https://www.linkedin.com/jobs/view/lead-data-e...
...                                                  ...
12212  https://uk.linkedin.com/jobs/view/data-reporti...
12213  https://www.linkedin.com/jobs/view/corporate-a...
12214  https://www.linkedin.com/jobs/view/senior-data...
12215  https://www.linkedin.com/jobs/view/senior-data...
12216  https://www.linkedin.com/jobs/view/medical-tec...

                                                 job_skills  \
0      Machine Learning, Programming, Python, Scala, ...
1      C++, Python, PyTorch, TensorFlow, MXNet, CUDA,...
2      ETL, Data Integration, Data Transformation, Da...
3      Data Lakes, Data Bricks, Azure Data Factory Pi...
4      Java, Scala, Python, RDBMS, NoSQL, Redshift, S...
...                                                  ...
12212  Dashboard development, Reporting, Power BI, SQ...
12213  Investigation, Antimoney laundering, Fraud, Ba...
12214  Data Science, Quantitative Modeling, SQL, Data...
12215  Data Engineering, Data Quality, SQL, Python, T...
12216  Medical Technology, MLS, Microbiology, Clinica...

                   last_processed_time    last_status got_summary got_ner
\
0      2024-01-21 08:08:48.031964+00  Finished NER            t       t

1      2024-01-20 04:02:12.331406+00  Finished NER            t       t

2      2024-01-21 08:08:31.941595+00  Finished NER            t       t

3      2024-01-20 15:30:55.796572+00  Finished NER            t       t

4      2024-01-21 08:08:58.312124+00  Finished NER            t       t

...                              ...           ...          ...     ...

12212  2024-01-21 07:11:22.099082+00  Finished NER            t       t
```

```
12213   2024-01-19 15:10:41.177008+00   Finished NER              t        t

12214   2024-01-20 15:20:19.036168+00   Finished NER              t        t

12215   2024-01-19 23:25:28.107523+00   Finished NER              t        t

12216   2024-01-19 23:26:27.593516+00   Finished NER              t        t


       is_being_worked                                      job_title  \
0                    f              Senior Machine Learning Engineer
1                    f      Principal Software Engineer, ML Accelerators
2                    f            Senior ETL Data Warehouse Specialist
3                    f        Senior Data Warehouse Developer / Architect
4                    f                              Lead Data Engineer
...                ...                                            ...
12212                f      Data Reporting Manager, FOOTBALL ASSOCIATION
12213                f      Corporate AML Alert Investigation Specialist
12214                f                            Senior Data Scientist
12215                f                             Senior Data Engineer
12216                f                 Medical Technologist, MLS or MLT

                                     company                     job_location  \
0                            Jobs for Humanity                           New
Haven, CT
1                                       Aurora                           San
Francisco, CA
2                            Adame Services LLC                           New
York, NY
3                              Morph Enterprise
Harrisburg, PA
4                                         Dice
Plano, TX
...                                        ...
...
12212                            Guardian Jobs    Wembley, England, United
Kingdom
12213                     Glacier Bancorp, Inc.
Kalispell, MT
12214                                  Highnote                           San
Francisco, CA
12215    CompSource Mutual Insurance Company                        Oklahoma
City, OK
12216                    Community Health Systems
Mooresville, NC

        first_seen   search_city   search_country  \
0       2024-01-14    East Haven    United States
```

```
1       2024-01-14     El Cerrito     United States
2       2024-01-14     Middletown     United States
3       2024-01-12       Lebanon      United States
4       2024-01-14       McKinney     United States
...            ...           ...               ...
12212   2024-01-16   High Wycombe   United Kingdom
12213   2024-01-14       Montana      United States
12214   2024-01-16    San Rafael      United States
12215   2024-01-16       Arcadia      United States
12216   2024-01-14      Salisbury     United States

                          search_position      job_level  job_type
0       Agricultural-Research Engineer  Mid senior    Onsite
1                      Set-Key Driver  Mid senior    Onsite
2          Technical Support Specialist   Associate    Onsite
3                            Architect  Mid senior    Onsite
4            Maintenance Data Analyst  Mid senior    Onsite
...                              ...          ...       ...
12212          Manager Forms Analysis  Mid senior    Onsite
12213                          Teller  Mid senior    Onsite
12214                   Mathematician  Mid senior    Onsite
12215             Protection Engineer  Mid senior    Onsite
12216                       Biologist  Mid senior    Onsite

[12217 rows x 16 columns]
```

```python
print(job_skills_position['job_link'].head())
```

```
0    https://www.linkedin.com/jobs/view/senior-mach...
1    https://www.linkedin.com/jobs/view/principal-s...
2    https://www.linkedin.com/jobs/view/senior-etl-...
3    https://www.linkedin.com/jobs/view/senior-data...
4    https://www.linkedin.com/jobs/view/lead-data-e...
Name: job_link, dtype: object
```

```python
jobs =
job_skills_position['job_link'].str.split('view').str[1].str.strip('/'
)
jobs
```

```
0         senior-machine-learning-engineer-at-jobs-for-h...
1         principal-software-engineer-ml-accelerators-at...
2         senior-etl-data-warehouse-specialist-at-adame-...
3         senior-data-warehouse-developer-architect-at-m...
4                   lead-data-engineer-at-dice-3805948138
                           ...
12212    data-reporting-manager-football-association-at...
12213    corporate-aml-alert-investigation-specialist-a...
12214         senior-data-scientist-at-highnote-3793519077
12215    senior-data-engineer-at-compsource-mutual-insu...
```

```
12216    medical-technologist-mls-or-mlt-at-community-h...
Name: job_link, Length: 12217, dtype: object

ds_jobs  = job_skills_position[['job_skills','job_title']]
ds_jobs

                                              job_skills  \
0      Machine Learning, Programming, Python, Scala, ...
1      C++, Python, PyTorch, TensorFlow, MXNet, CUDA,...
2      ETL, Data Integration, Data Transformation, Da...
3      Data Lakes, Data Bricks, Azure Data Factory Pi...
4      Java, Scala, Python, RDBMS, NoSQL, Redshift, S...
...                                                  ...
12212  Dashboard development, Reporting, Power BI, SQ...
12213  Investigation, Antimoney laundering, Fraud, Ba...
12214  Data Science, Quantitative Modeling, SQL, Data...
12215  Data Engineering, Data Quality, SQL, Python, T...
12216  Medical Technology, MLS, Microbiology, Clinica...

                                          job_title
0              Senior Machine Learning Engineer
1      Principal Software Engineer, ML Accelerators
2             Senior ETL Data Warehouse Specialist
3       Senior Data Warehouse Developer / Architect
4                                  Lead Data Engineer
...                                              ...
12212  Data Reporting Manager, FOOTBALL ASSOCIATION
12213  Corporate AML Alert Investigation Specialist
12214                         Senior Data Scientist
12215                          Senior Data Engineer
12216              Medical Technologist, MLS or MLT

[12217 rows x 2 columns]

job_vals = ds_jobs['job_title'].value_counts()
job_vals = job_vals[job_vals>30]
job_vals

job_title
Senior Data Engineer
285
Senior Data Analyst
163
Data Engineer
149
Senior MLOps Engineer
138
Data Analyst
137
Data Scientist
```

128
Lead Data Engineer
123
Senior Data Scientist
119
Data Architect
111
Staff Machine Learning Engineer, Series A
101
Senior Machine Learning Engineer
59
Manager, Data Engineering
57
Manager, Data Loss Prevention (DLP) Engineer (Symantec)
57
Senior Data Engineer, Public Company
53
Senior Machine Learning Engineer - AI
53
Principal Associate, Data Loss Prevention (DLP) Engineer
53
Senior Database Administrator
49
Machine Learning Engineer
49
Lead Data Scientist
46
Senior MLS Engineer, Autonomous Driving Startup
45
Manager, Data Engineering (Python, AWS, Airflow, Snowflake) (Hybrid)
44
Senior Manager, Ontology and Data Modeling
41
Sr. Data Engineer
39
Database Administrator
37
Senior Manager, Machine Learning Engineering
36
Lead Machine Learning Engineer
35
Contractor - Epic Clarity Data Analyst - Revenue Cycle
34
Manager, Cyber Risk & Analysis (Machine Learning)
32
Contractor - Epic Clarity Data Analyst
32
Datacenter Technician

```
32
Name: count, dtype: int64

job_vals  = pd.DataFrame(job_vals)
job_vals.reset_index(inplace = True)
job_vals
```

	job_title	count
0	Senior Data Engineer	285
1	Senior Data Analyst	163
2	Data Engineer	149
3	Senior MLOps Engineer	138
4	Data Analyst	137
5	Data Scientist	128
6	Lead Data Engineer	123
7	Senior Data Scientist	119
8	Data Architect	111
9	Staff Machine Learning Engineer, Series A	101
10	Senior Machine Learning Engineer	59
11	Manager, Data Engineering	57
12	Manager, Data Loss Prevention (DLP) Engineer (...	57
13	Senior Data Engineer, Public Company	53
14	Senior Machine Learning Engineer - AI	53
15	Principal Associate, Data Loss Prevention (DLP...	53
16	Senior Database Administrator	49
17	Machine Learning Engineer	49
18	Lead Data Scientist	46
19	Senior MLS Engineer, Autonomous Driving Startup	45
20	Manager, Data Engineering (Python, AWS, Airflo...	44
21	Senior Manager, Ontology and Data Modeling	41
22	Sr. Data Engineer	39
23	Database Administrator	37
24	Senior Manager, Machine Learning Engineering	36
25	Lead Machine Learning Engineer	35
26	Contractor - Epic Clarity Data Analyst - Reven...	34
27	Manager, Cyber Risk & Analysis (Machine Learning)	32
28	Contractor - Epic Clarity Data Analyst	32
29	Datacenter Technician	32

```
job_title_list = list(job_vals['job_title'])
job_title_list
```

```
['Senior Data Engineer',
 'Senior Data Analyst',
 'Data Engineer',
 'Senior MLOps Engineer',
 'Data Analyst',
 'Data Scientist',
 'Lead Data Engineer',
 'Senior Data Scientist',
```

```python
 'Data Architect',
 'Staff Machine Learning Engineer, Series A',
 'Senior Machine Learning Engineer',
 'Manager, Data Engineering',
 'Manager, Data Loss Prevention (DLP) Engineer (Symantec)',
 'Senior Data Engineer, Public Company',
 'Senior Machine Learning Engineer - AI',
 'Principal Associate, Data Loss Prevention (DLP) Engineer',
 'Senior Database Administrator',
 'Machine Learning Engineer',
 'Lead Data Scientist',
 'Senior MLS Engineer, Autonomous Driving Startup',
 'Manager, Data Engineering (Python, AWS, Airflow, Snowflake)
(Hybrid)',
 'Senior Manager, Ontology and Data Modeling',
 'Sr. Data Engineer',
 'Database Administrator',
 'Senior Manager, Machine Learning Engineering',
 'Lead Machine Learning Engineer',
 'Contractor - Epic Clarity Data Analyst - Revenue Cycle',
 'Manager, Cyber Risk & Analysis (Machine Learning)',
 'Contractor - Epic Clarity Data Analyst',
 'Datacenter Technician']

job_title_list = ['Senior Data Engineer',
 'Senior Data Analyst',
 'Data Engineer',
 'Senior MLOps Engineer',
 'Data Analyst',
 'Data Scientist',
 'Lead Data Engineer',
 'Senior Data Scientist',
 'Data Architect',
 'Senior Machine Learning Engineer',
 'Manager, Data Engineering',
 'Senior Machine Learning Engineer - AI',
 'Principal Associate, Data Loss Prevention (DLP) Engineer',
 'Senior Database Administrator',
 'Machine Learning Engineer',
 'Lead Data Scientist',
 'Sr. Data Engineer',
 'Database Administrator',
 'Senior Manager, Machine Learning Engineering',
 'Lead Machine Learning Engineer',

 'Manager, Cyber Risk & Analysis (Machine Learning)',

 ]
```

```python
ds_jobs_cleaned = ds_jobs[ds_jobs['job_title'].isin(job_title_list) ]
ds_jobs_cleaned.reset_index(inplace = True,drop = True)
ds_jobs_cleaned
```

```
                                              job_skills  \
0     Machine Learning, Programming, Python, Scala, ...
1     Java, Scala, Python, RDBMS, NoSQL, Redshift, S...
2     Data Warehouse (DW), Extract/Transform/Load (E...
3     Machine Learning, Generative AI, Cloudbased se...
4     Data Loss Prevention (DLP), Cybersecurity, Web...
...                                                  ...
1893  Data Engineering, Technical Solutions, Agile D...
1894  Machine Learning, Generative AI, Cloud Risk Ma...
1895  Data Architecture, Data Engineering, Linux, Wi...
1896  Data Science, Quantitative Modeling, SQL, Data...
1897  Data Engineering, Data Quality, SQL, Python, T...

                                               job_title
0                      Senior Machine Learning Engineer
1                                     Lead Data Engineer
2                                   Senior Data Engineer
3        Manager, Cyber Risk & Analysis (Machine Learning)
4        Principal Associate, Data Loss Prevention (DLP...
...                                                  ...
1893                                  Lead Data Engineer
1894     Manager, Cyber Risk & Analysis (Machine Learning)
1895                                        Data Architect
1896                                 Senior Data Scientist
1897                                  Senior Data Engineer

[1898 rows x 2 columns]
```

```python
ds_jobs_cleaned.drop_duplicates()
```

```
                                              job_skills  \
0     Machine Learning, Programming, Python, Scala, ...
1     Java, Scala, Python, RDBMS, NoSQL, Redshift, S...
2     Data Warehouse (DW), Extract/Transform/Load (E...
3     Machine Learning, Generative AI, Cloudbased se...
4     Data Loss Prevention (DLP), Cybersecurity, Web...
...                                                  ...
1893  Data Engineering, Technical Solutions, Agile D...
1894  Machine Learning, Generative AI, Cloud Risk Ma...
1895  Data Architecture, Data Engineering, Linux, Wi...
1896  Data Science, Quantitative Modeling, SQL, Data...
1897  Data Engineering, Data Quality, SQL, Python, T...

                                               job_title
0                      Senior Machine Learning Engineer
1                                     Lead Data Engineer
```

```
2                              Senior Data Engineer
3     Manager, Cyber Risk & Analysis (Machine Learning)
4        Principal Associate, Data Loss Prevention (DLP...
...                                                    ...
1893                              Lead Data Engineer
1894  Manager, Cyber Risk & Analysis (Machine Learning)
1895                                   Data Architect
1896                             Senior Data Scientist
1897                             Senior Data Engineer

[1898 rows x 2 columns]
```

```python
ds_jobs_cleaned['job_skills'] = ds_jobs_cleaned['job_skills'].apply(
    lambda x: x.split(', ') if isinstance(x, str) else x
)



ds_jobs_exploded =
ds_jobs_cleaned.explode('job_skills').reset_index(drop = True)
ds_jobs_exploded.columns = ['job_skills','job_title']
ds_jobs_exploded
```

```
/tmp/ipykernel_30/703355473.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation:
https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#
returning-a-view-versus-a-copy
  ds_jobs_cleaned['job_skills'] = ds_jobs_cleaned['job_skills'].apply(
```

```
               job_skills                          job_title
0         Machine Learning  Senior Machine Learning Engineer
1              Programming  Senior Machine Learning Engineer
2                   Python  Senior Machine Learning Engineer
3                    Scala  Senior Machine Learning Engineer
4                     Java  Senior Machine Learning Engineer
...                    ...                               ...
49186             Snowflake             Senior Data Engineer
49187   Amazon Web Services             Senior Data Engineer
49188       Microsoft Azure             Senior Data Engineer
49189            SAFe Agile             Senior Data Engineer
49190                 SCRUM             Senior Data Engineer

[49191 rows x 2 columns]
```

```python
ds_jobs_exploded.drop_duplicates(inplace = True)

ds_jobs_exploded.reset_index(inplace = True, drop = True)
```

```
ds_jobs_exploded
```

```
                            job_skills
job_title
0                     Machine Learning  Senior Machine Learning
Engineer
1                          Programming  Senior Machine Learning
Engineer
2                               Python  Senior Machine Learning
Engineer
3                                Scala  Senior Machine Learning
Engineer
4                                 Java  Senior Machine Learning
Engineer
...                                ...
...
18845               Data Vault Express              Senior Data
Engineer
18846            Workers Compensation              Senior Data
Engineer
18847                       Insurance              Senior Data
Engineer
18848  CloudBased Data Warehouse Products          Senior Data
Engineer
18849                       SAFe Agile              Senior Data
Engineer

[18850 rows x 2 columns]
```

```python
skills_dict = {index: skill for index, skill in
enumerate(job_title_list)}
skills_dict
```

```
{0: 'Senior Data Engineer',
 1: 'Senior Data Analyst',
 2: 'Data Engineer',
 3: 'Senior MLOps Engineer',
 4: 'Data Analyst',
 5: 'Data Scientist',
 6: 'Lead Data Engineer',
 7: 'Senior Data Scientist',
 8: 'Data Architect',
 9: 'Senior Machine Learning Engineer',
 10: 'Manager, Data Engineering',
 11: 'Senior Machine Learning Engineer - AI',
 12: 'Principal Associate, Data Loss Prevention (DLP) Engineer',
 13: 'Senior Database Administrator',
 14: 'Machine Learning Engineer',
```

```
 15: 'Lead Data Scientist',
 16: 'Sr. Data Engineer',
 17: 'Database Administrator',
 18: 'Senior Manager, Machine Learning Engineering',
 19: 'Lead Machine Learning Engineer',
 20: 'Manager, Cyber Risk & Analysis (Machine Learning)'}

skills_df = pd.Series(skills_dict)
skills_df = pd.DataFrame(skills_df)
skills_df.reset_index(inplace = True)
skills_df.columns = ['Job_id','job_title']
skills_df['Job_id'] = skills_df['Job_id']+1
skills_df
```

```
    Job_id                                          job_title
0        1                               Senior Data Engineer
1        2                               Senior Data Analyst
2        3                                       Data Engineer
3        4                               Senior MLOps Engineer
4        5                                        Data Analyst
5        6                                       Data Scientist
6        7                                   Lead Data Engineer
7        8                               Senior Data Scientist
8        9                                       Data Architect
9       10                   Senior Machine Learning Engineer
10      11                           Manager, Data Engineering
11      12                Senior Machine Learning Engineer - AI
12      13  Principal Associate, Data Loss Prevention (DLP...
13      14                      Senior Database Administrator
14      15                          Machine Learning Engineer
15      16                                   Lead Data Scientist
16      17                                    Sr. Data Engineer
17      18                               Database Administrator
18      19       Senior Manager, Machine Learning Engineering
19      20                       Lead Machine Learning Engineer
20      21  Manager, Cyber Risk & Analysis (Machine Learning)
```

```
ds_jobs_exploded = ds_jobs_exploded.merge(skills_df,on='job_title')
ds_jobs_exploded
```

```
                             job_skills
job_title  \
0                        Machine Learning  Senior Machine Learning
Engineer
1                            Programming  Senior Machine Learning
Engineer
2                                  Python  Senior Machine Learning
Engineer
3                                   Scala  Senior Machine Learning
Engineer
```

```
4                                 Java  Senior Machine Learning
Engineer
...                                  ...
...
18845              Data Vault Express             Senior Data
Engineer
18846           Workers Compensation             Senior Data
Engineer
18847                      Insurance             Senior Data
Engineer
18848  CloudBased Data Warehouse Products        Senior Data
Engineer
18849                      SAFe Agile             Senior Data
Engineer

       Job_id
0          10
1          10
2          10
3          10
4          10
...       ...
18845       1
18846       1
18847       1
18848       1
18849       1

[18850 rows x 3 columns]

ds_jobs_exploded

                              job_skills
job_title  \
0                       Machine Learning  Senior Machine Learning
Engineer
1                            Programming  Senior Machine Learning
Engineer
2                                 Python  Senior Machine Learning
Engineer
3                                  Scala  Senior Machine Learning
Engineer
4                                   Java  Senior Machine Learning
Engineer
...                                  ...
...
18845              Data Vault Express             Senior Data
Engineer
18846           Workers Compensation             Senior Data
Engineer
```

```
18847                            Insurance          Senior Data
Engineer
18848  CloudBased Data Warehouse Products          Senior Data
Engineer
18849                            SAFe Agile         Senior Data
Engineer

       Job_id
0          10
1          10
2          10
3          10
4          10
...        ...
18845       1
18846       1
18847       1
18848       1
18849       1

[18850 rows x 3 columns]

ds_jobs_exploded.to_csv("DataScienceJobs.csv")
```