FINAL PROJECT

Menganalisa pekerjaan dalam bidang data yang paling banyak diminati di negara Amerika dan Eropa

ANASTHASIA SUCI

Table of Content

DATA SET

DATA PREPROCESSING

VISUALIZATION

Table of Content

DATA SET

https://www.kaggle.com/datasets/hummaamqaasim/jobs-in-data/data

	work_year	job_title	job_category	salary_currency	salary	salary_in_usd	employee_residence	experience_level
0	2023	Data DevOps Engineer	Data Engineering	EUR	88000	95012	Germany	Mid-level
1	2023	Data Architect	Data Architecture and Modeling	USD	186000	186000	United States	Senior
2	2023	Data Architect	Data Architecture and Modeling	USD	81800	81800	United States	Senior
3	2023	Data Scientist	Data Science and Research	USD	212000	212000	United States	Senior
4	2023	Data Scientist	Data Science and Research	USD	93300	93300	United States	Senior

DATA PREPROCESSING

```
[ ]  Data columns (total 12 columns):
      #   Column              Non-Null Count   Dtype
     ---  ------              --------------   -----
      0   work_year           9355 non-null    int64
      1   job_title           9355 non-null    object
      2   job_category        9355 non-null    object
      3   salary_currency     9355 non-null    object
      4   salary              9355 non-null    int64
      5   salary_in_usd       9355 non-null    int64
      6   employee_residence  9355 non-null    object
      7   experience_level    9355 non-null    object
      8   employment_type     9355 non-null    object
      9   work_setting        9355 non-null    object
     10   company_location    9355 non-null    object
     11   company_size        9355 non-null    object
     dtypes: int64(3), object(9)
     memory usage: 877.2+ KB
```

```
    df.isnull().sum()

[→] work_year            0
    job_title            0
    job_category         0
    salary_currency      0
    salary               0
    salary_in_usd        0
    employee_residence   0
    experience_level     0
    employment_type      0
    work_setting         0
    company_location     0
    company_size         0
    dtype: int64
```
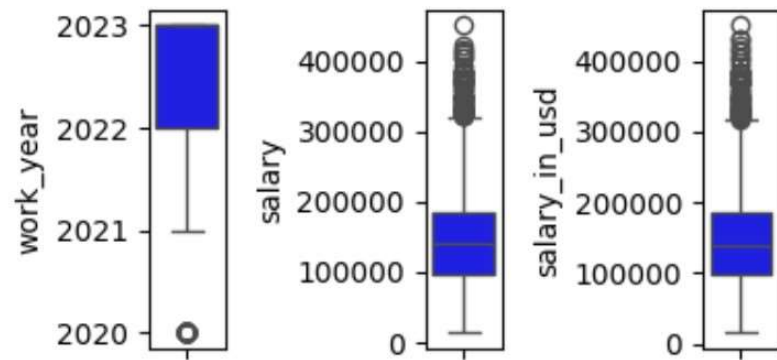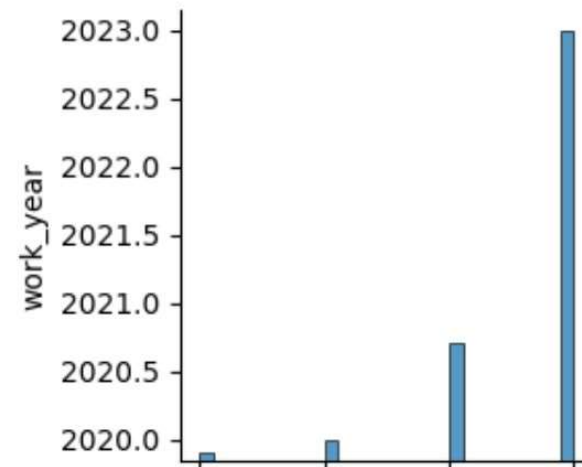
```
[9]  df.duplicated().sum()

     4014
```
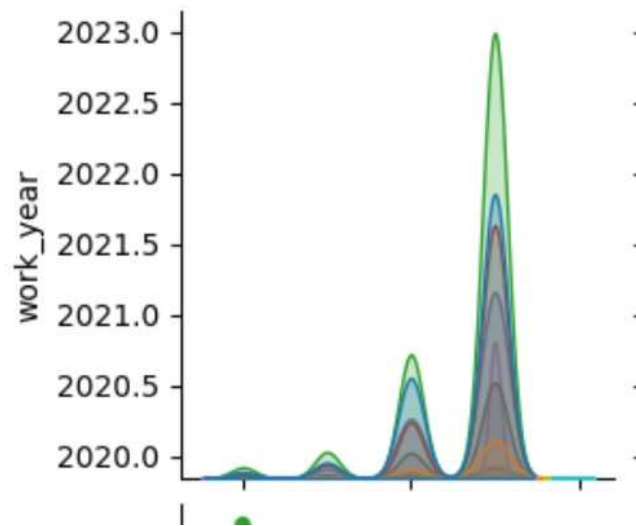
VISUALIZATION

BOXPLOT TO DETECT OUTLIER

INTERPRETATION

VISUALIZATION

INTERPRETATION

JOB CATEGORY

DATA ANALYST

What type of job_category is most popular with jobs in data in US and Europe?
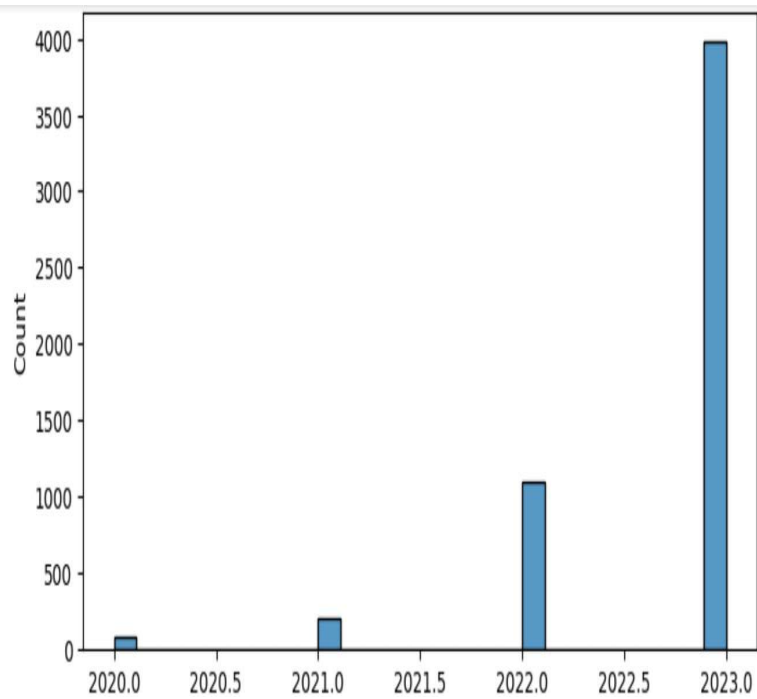
Job Category from Salary
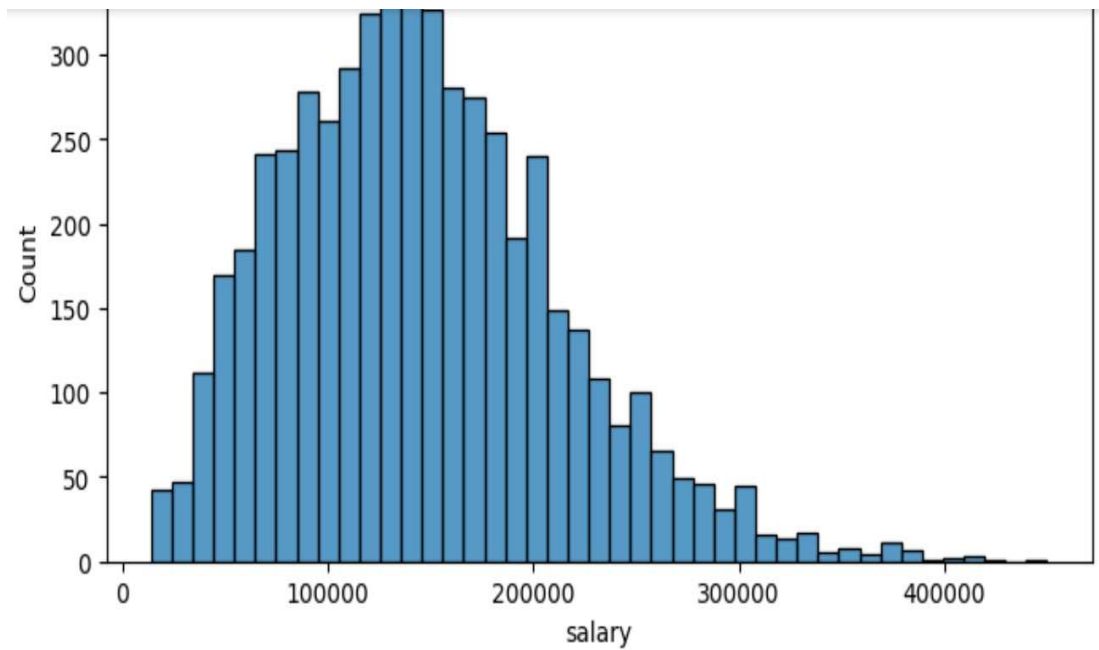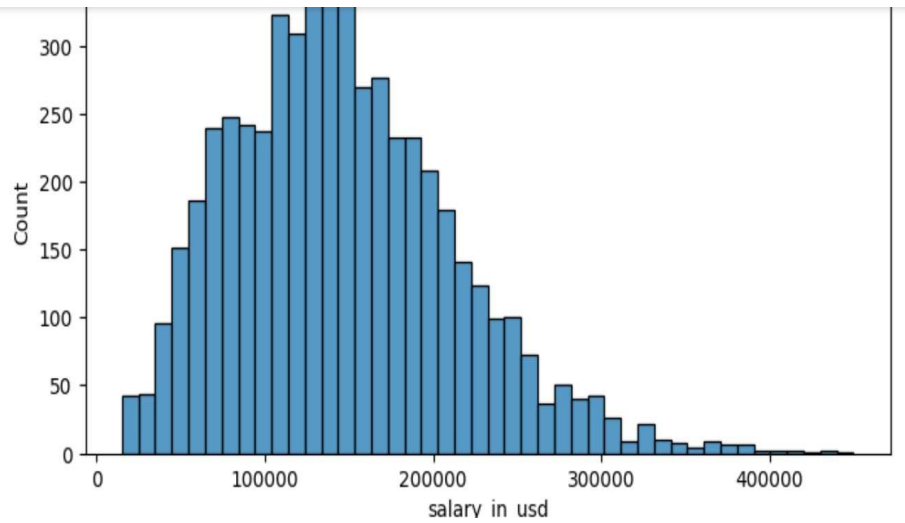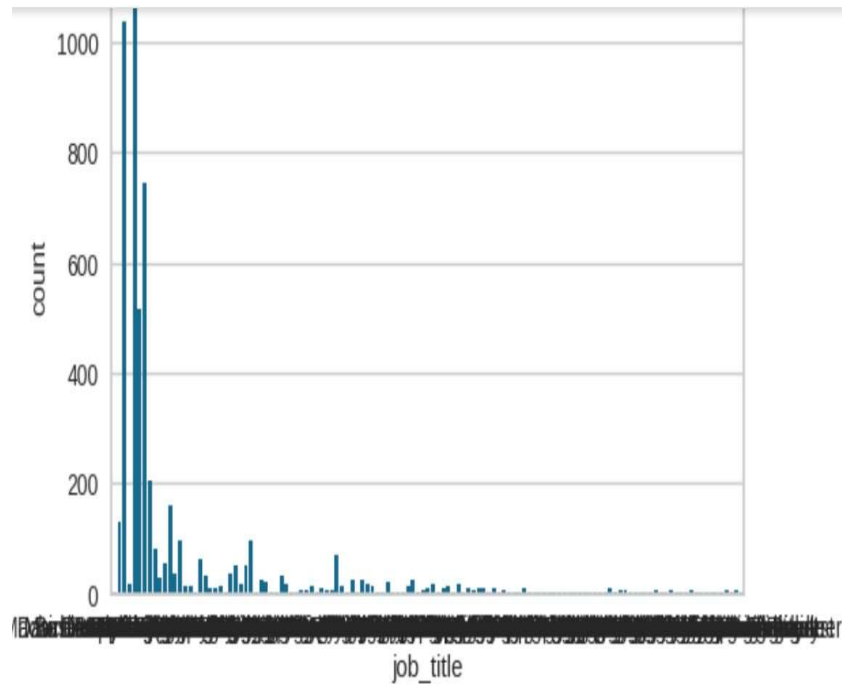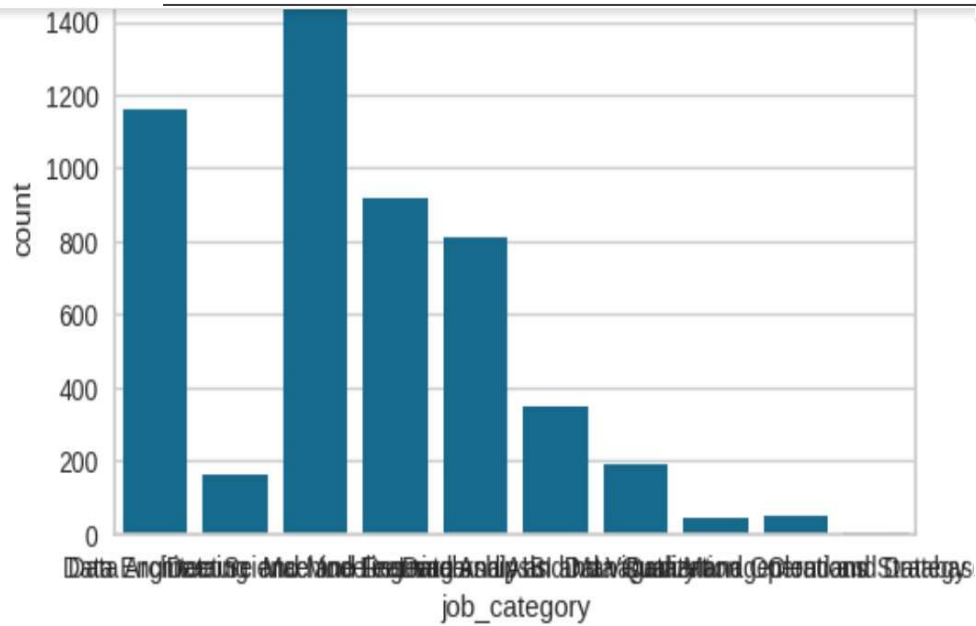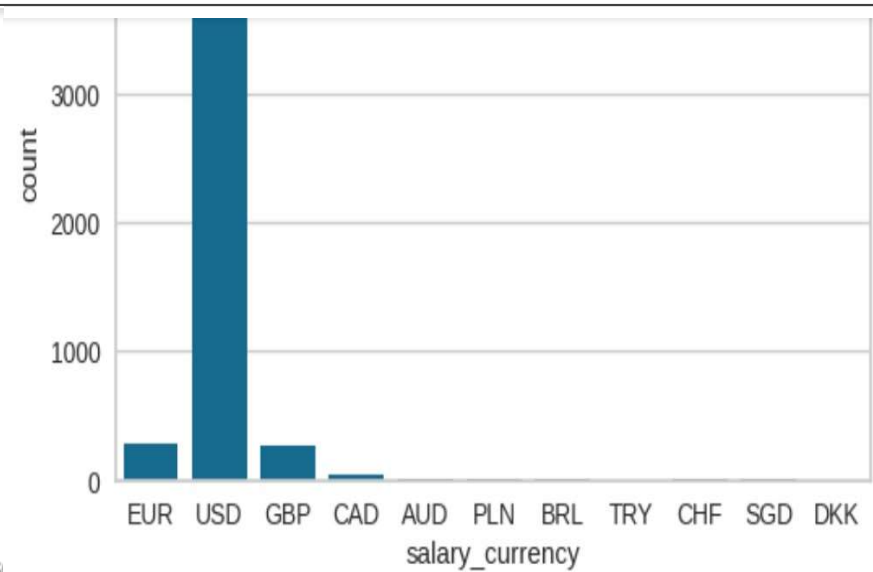
MACHINE LEARNING

CLUSTERING

CLUSTERING

MACHINE LEARNING

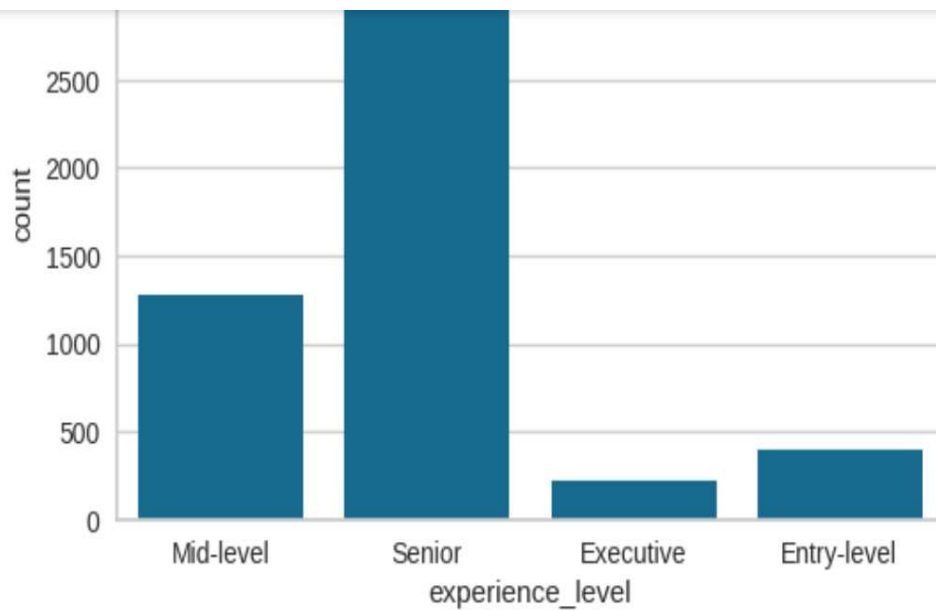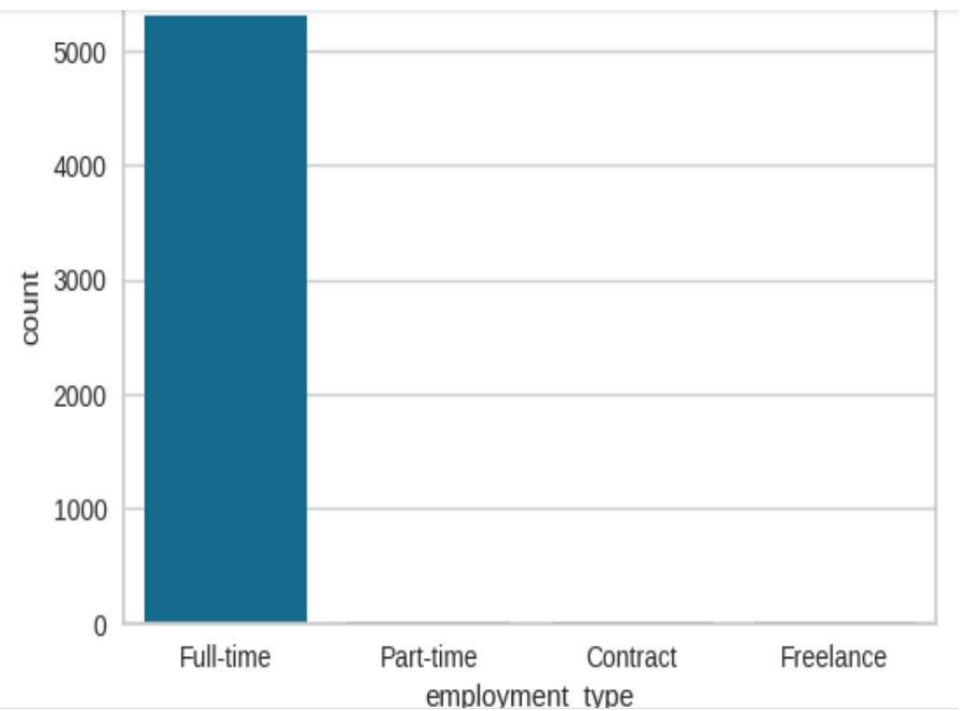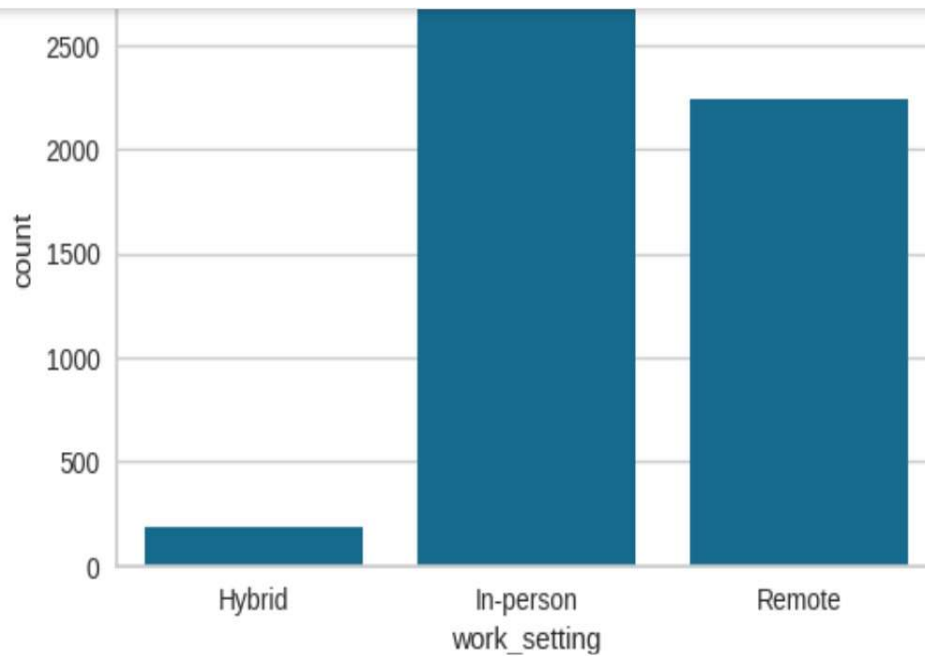CLUSTERING

CLUSTERING
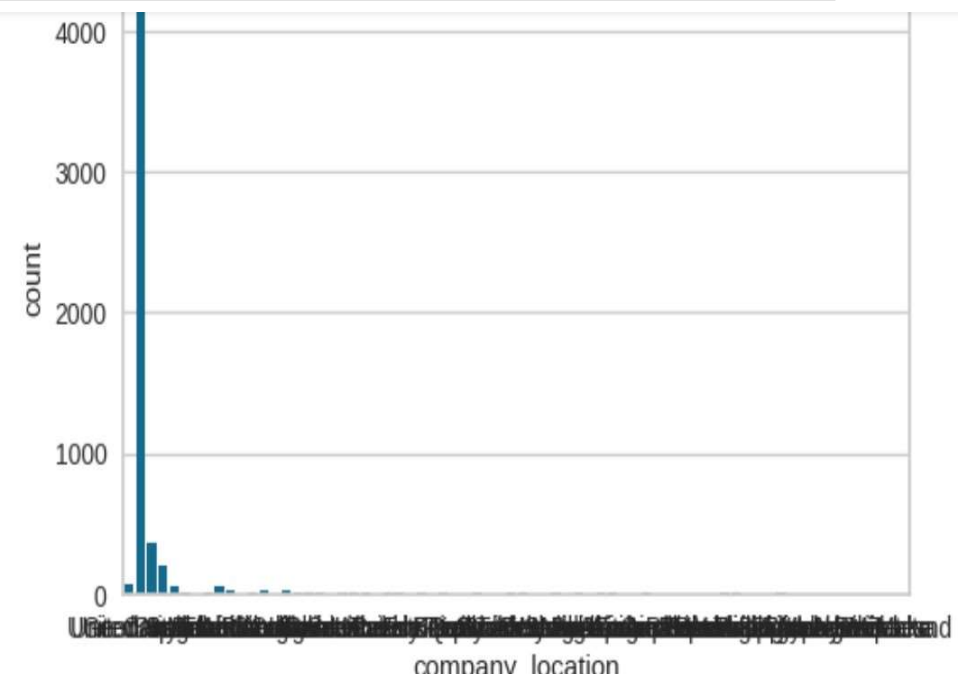
MACHINE LEARNING

CLUSTERING

CLUSTERING

MACHINE LEARNING

CLUSTERING   CLUSTERING

MACHINE LEARNING
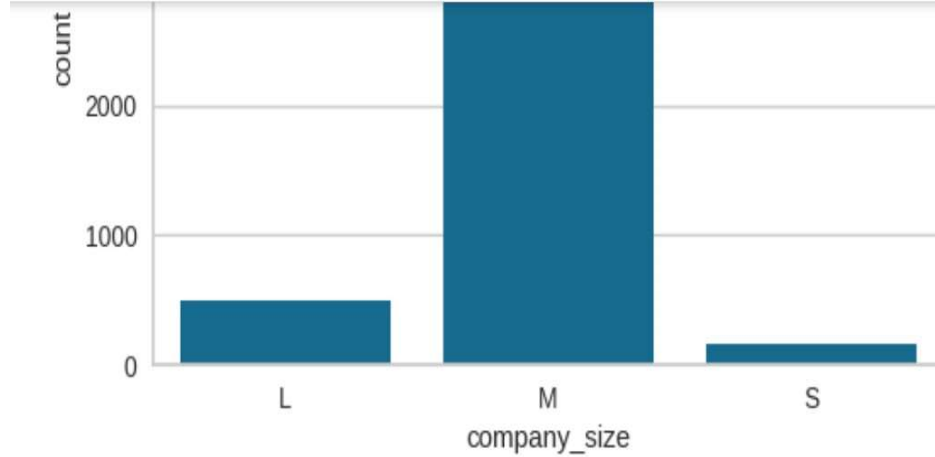
CLUSTERING

CLUSTERING

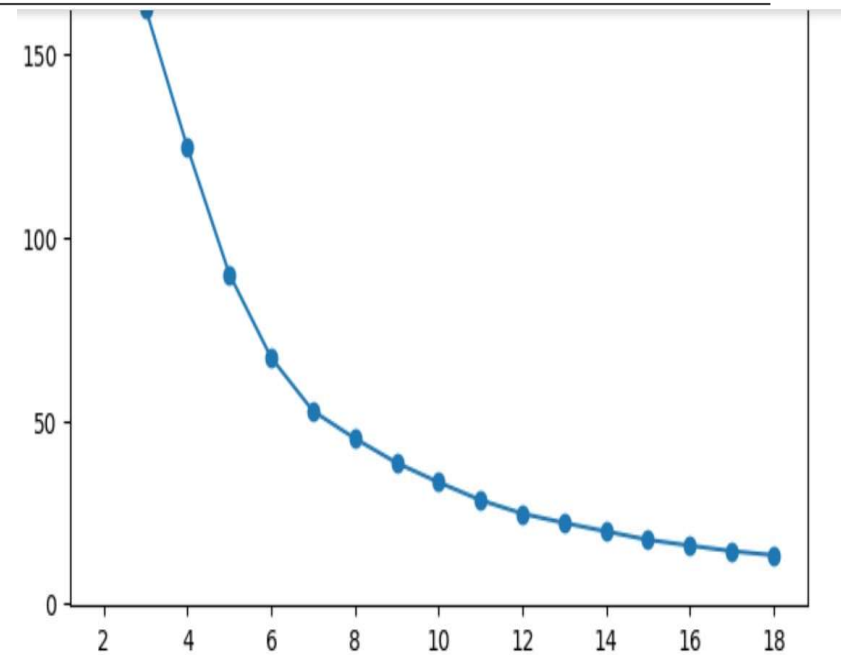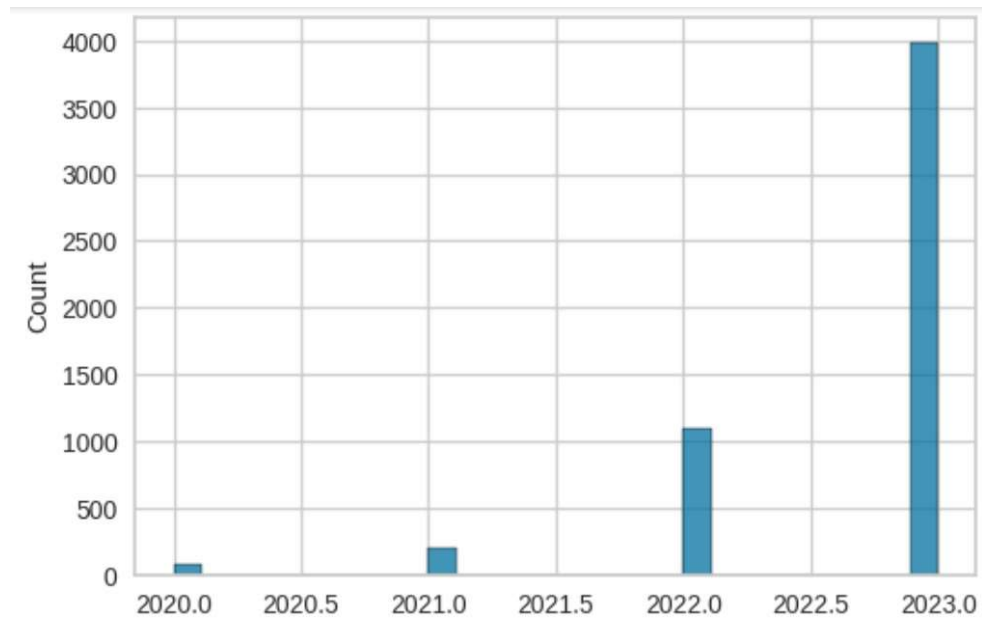# MACHINE LEARNING

CLUSTERING

CLUSTERING K-Means

MACHINE LEARNING

CLUSTERING JOB CATEGORY

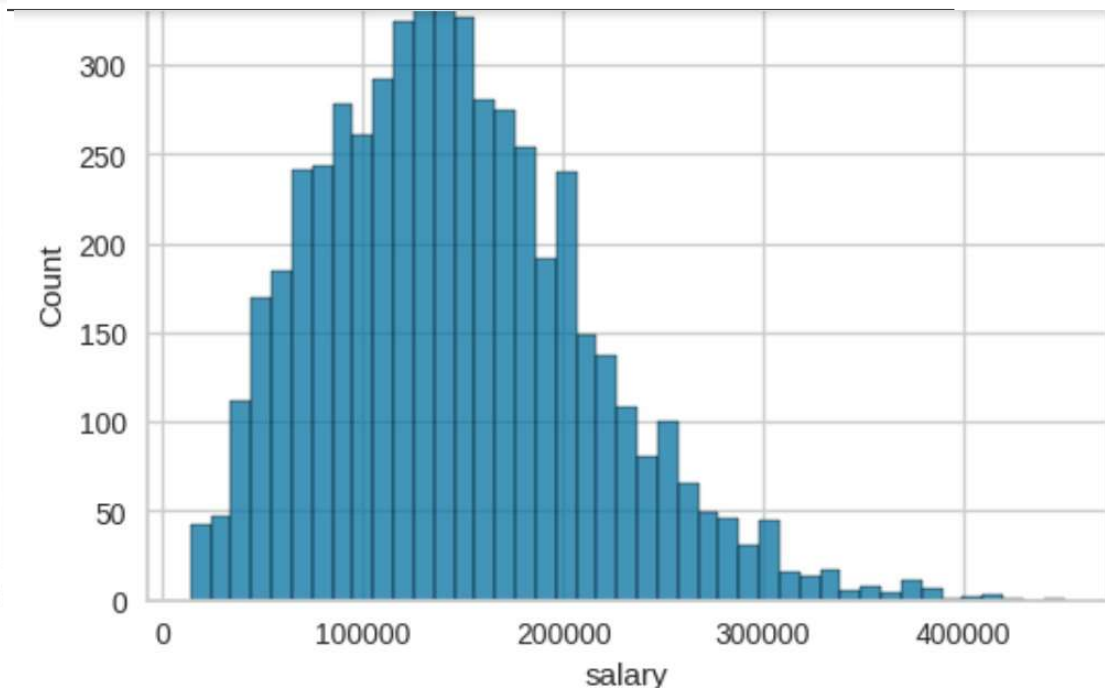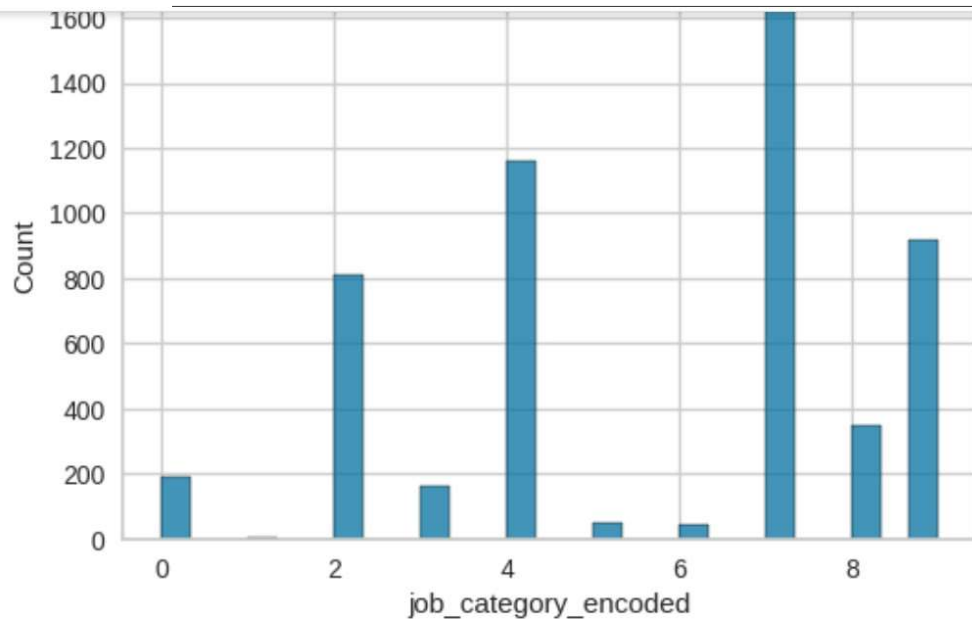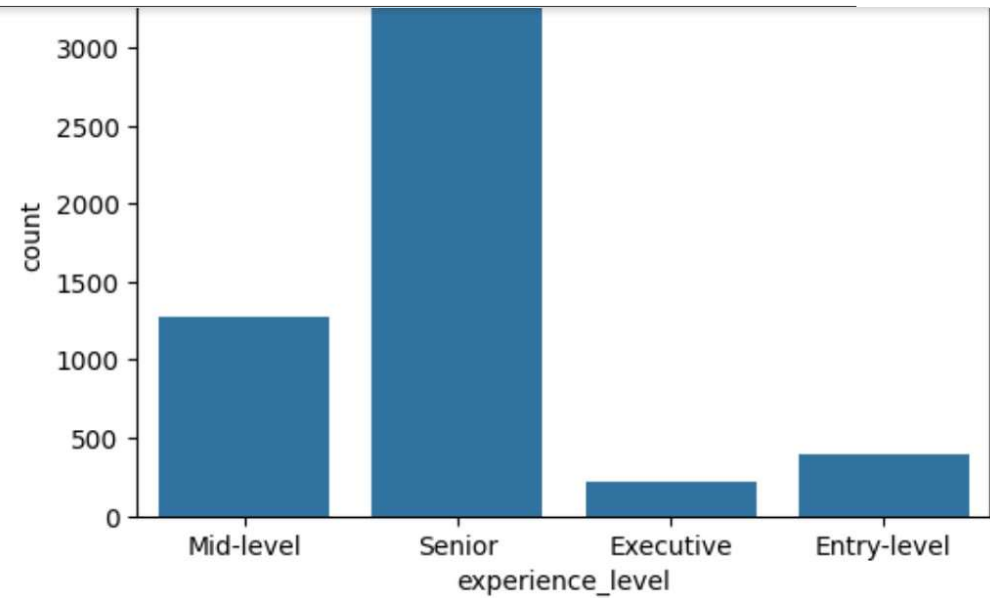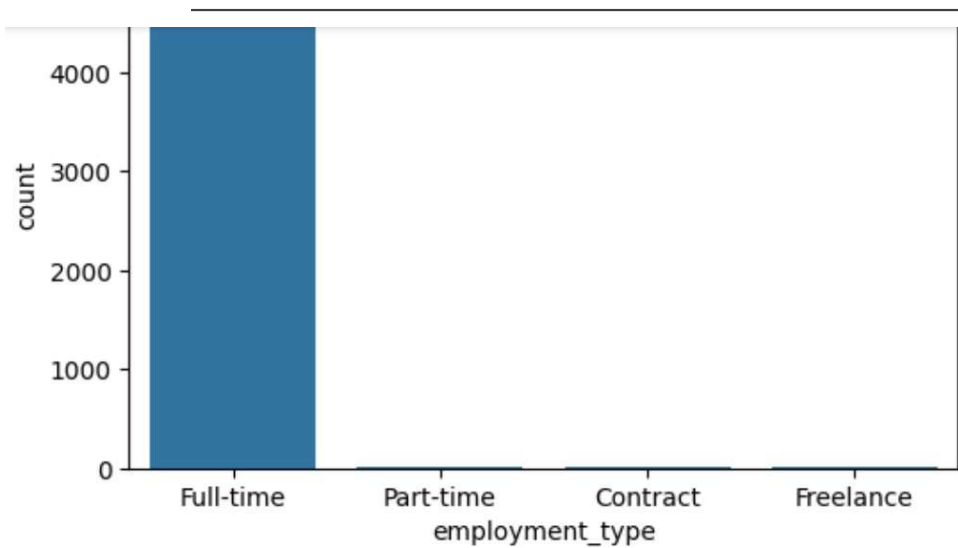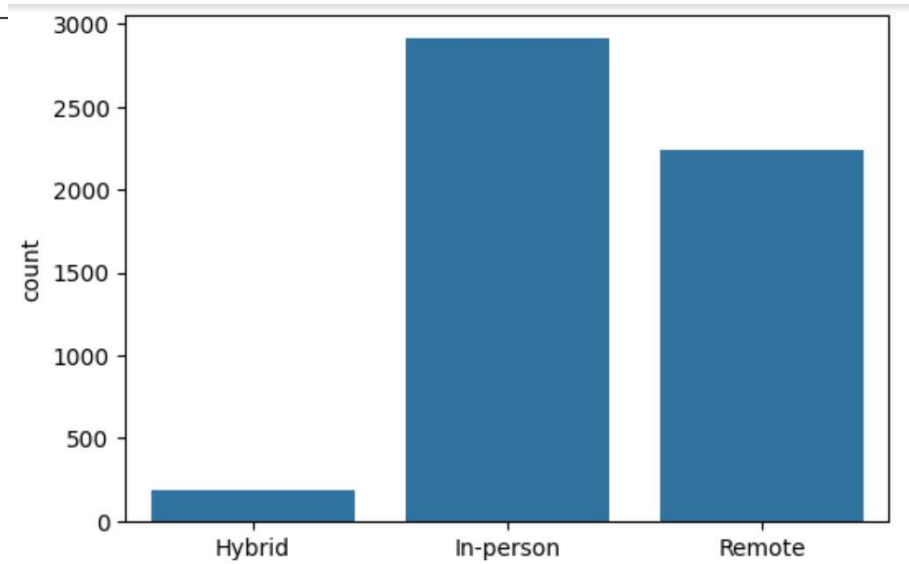## CLUSTERING JOB CATEGORY

MACHINE LEARNING

CLUSTERING JOB CATEGORY

CLUSTERING JOB CATEGORY
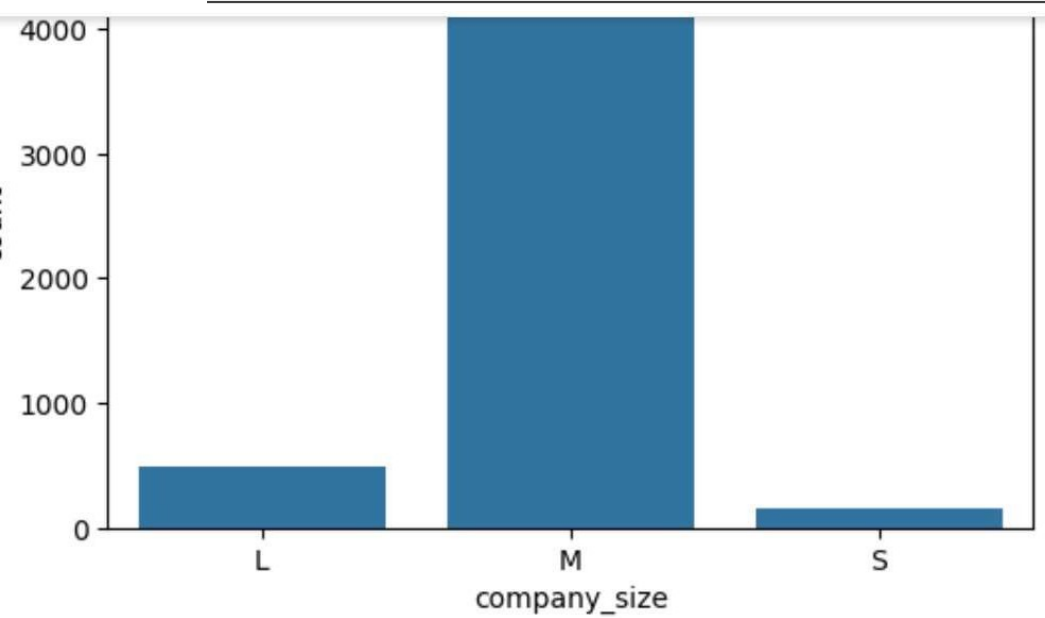
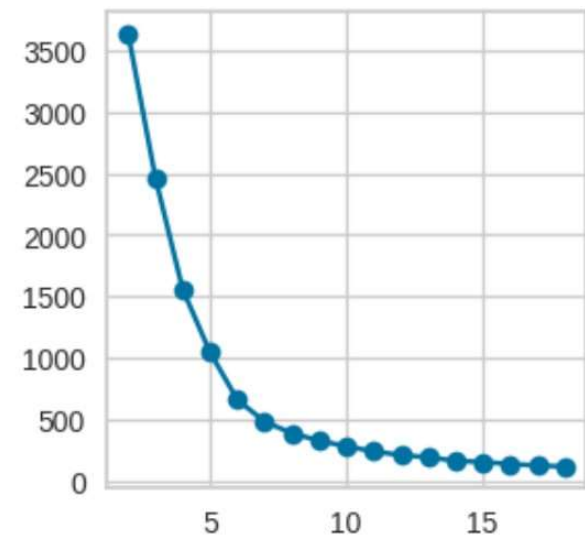# MACHINE LEARNING

CLUSTERING JOB CATEGORY

CLUSTERING JOB CATEGORY

MACHINE LEARNING

CLUSTERING JOB CATEGORY

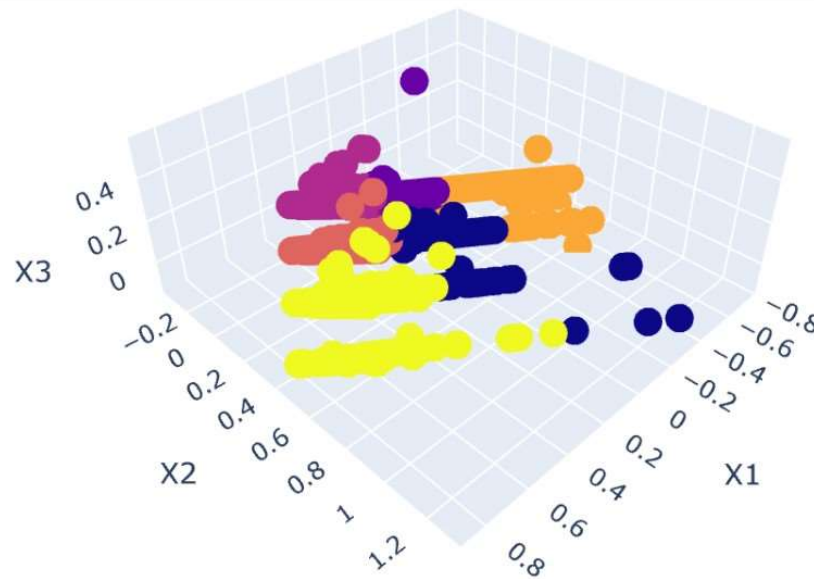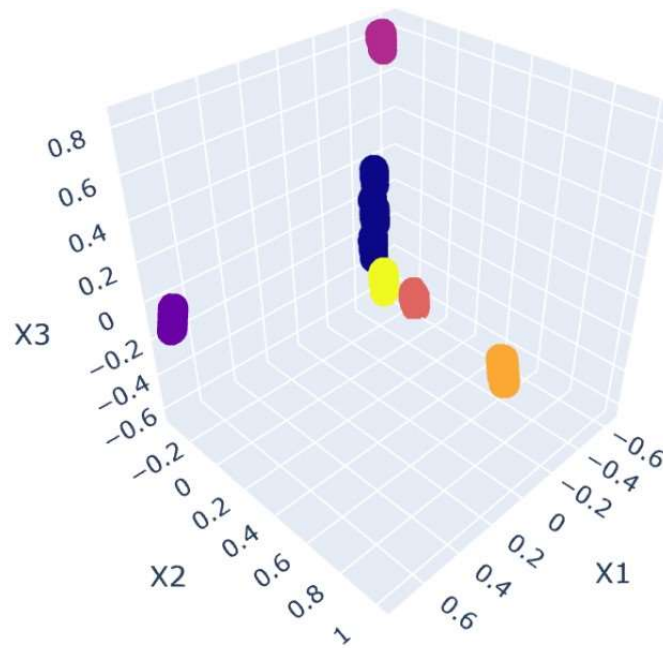CLUSTERING JOB CATEGORY

MODEL EVALUATION

Label 5 (kuning)
paling dominan

MODEL EVALUATION

Label 0 (biru)
paling dominan

INSIGHT & RECOMMENDATION

INSIGHT

Observation:

1. The highest rank in the job_title column is Data Engineer with 1100.
2. The highest rank in the job_category column is Data Science and Research with 1655.
3. The highest rank in the salary_currency column is USD with 4707.
4. The highest rank in the employee_residence column is United States with 4255.
5. The highest rank in the experience_level column is Senior with 3444.
6. The highest rank in the employment_type column is Full-time with 5296.
7. The highest rank in the work_setting column is In-person with 2913.
8. The highest rank in the company_location column is United States with 4301.
9. The highest rank in the company_size column is M with 4688.
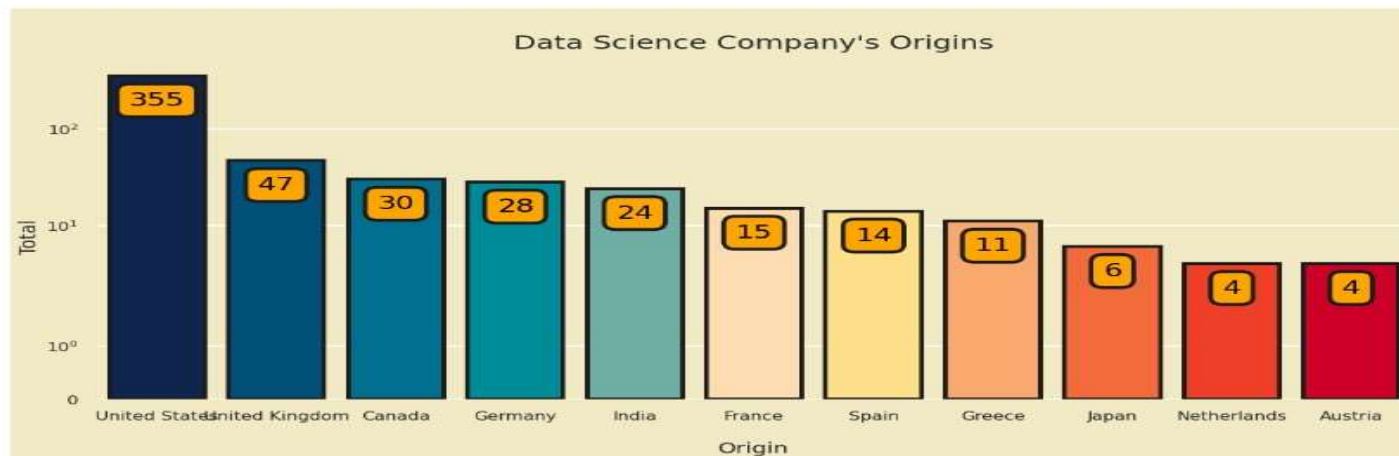
INSIGHT & RECOMMENDATION

RECOMMENDATION

ADDITIONAL INSIGHT

ADDITIONAL INSIGHT

ADDITIONAL INSIGHT

Analytics Insight

MARKETS ⌄

The Top 10 Data Science Jobs for the Newbies in 2023 are enlisted in this article

Data science is no longer a fancy term. Almost every sector that one can think of relies on data science in one form or the other. Today, this field has become so popular among individuals that every other person wants to make a career in data science. Data science is one of the most promising career opportunities that one can explore. On that note, let us have a look at the top 10 data science jobs for the newbies in 2023.

REFERENCES

- https://www.kaggle.com/datasets/hummaamqaasim/jobs-in-data/data

- https://www.kaggle.com/code/hasibalmuzdadid/data-science-jobs-salary-analysis-retro-vibe

- https://www.analyticsinsight.net/top-10-data-science-jobs-for-the-newbies-in-2023/