POC

Steps:

1. **Data Preparation**:
    - Load the generated dataset.
    - Preprocess the data to handle any missing values or outliers.
2. **Collaborative Filtering**:
    - Create a user-item matrix where the users are rows and the job titles (income sources) are columns.
    - Use matrix factorization techniques like Singular Value Decomposition (SVD) to decompose the user-item matrix.
    - Generate recommendations for users based on the decomposed matrices.
3. **Content-Based Filtering**:
    - Use the demographic information (age, city, state, etc.) to create user profiles.
    - Calculate the similarity between users based on their profiles.
    - Recommend jobs to users based on the similarity scores.
4. **Combining Both Approaches**:
    - Combine the results from collaborative filtering and content-based filtering to generate final recommendations.

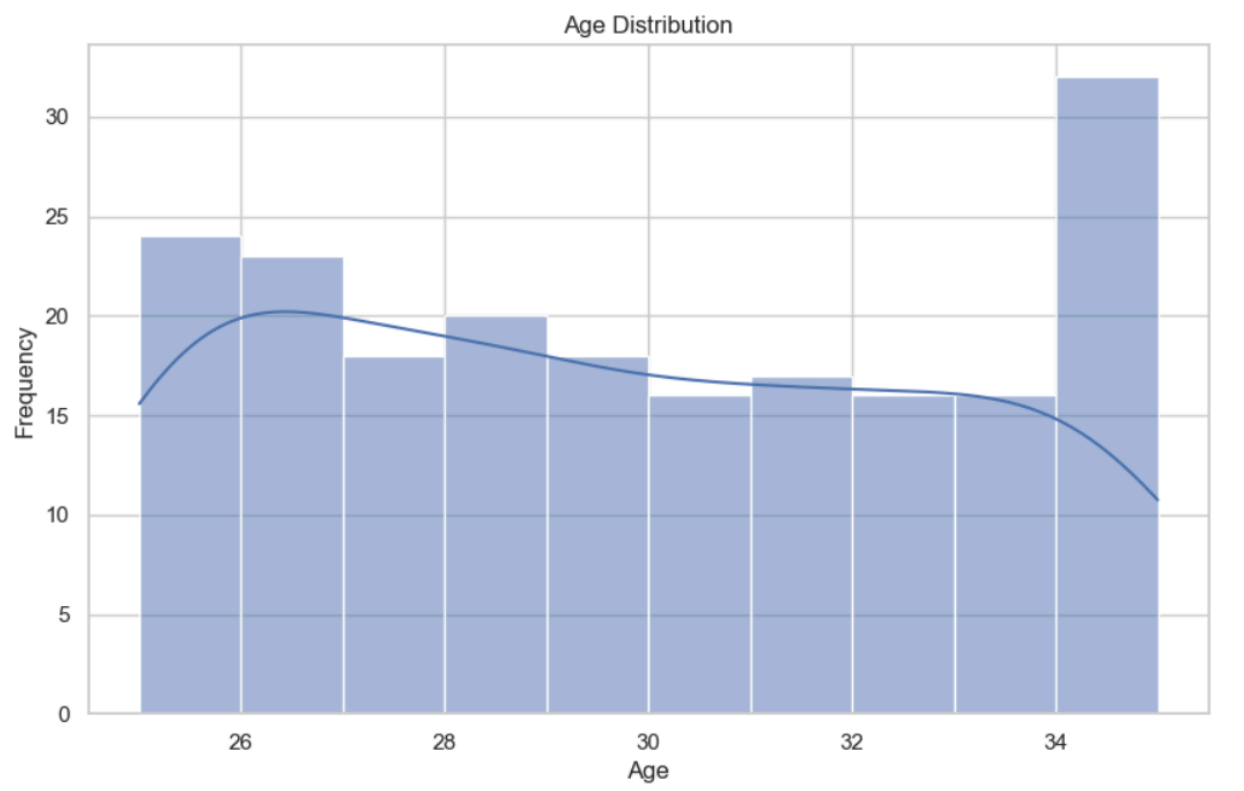
Data Visualization:

Data_Head: This shows the features and the type of dummy data generated.

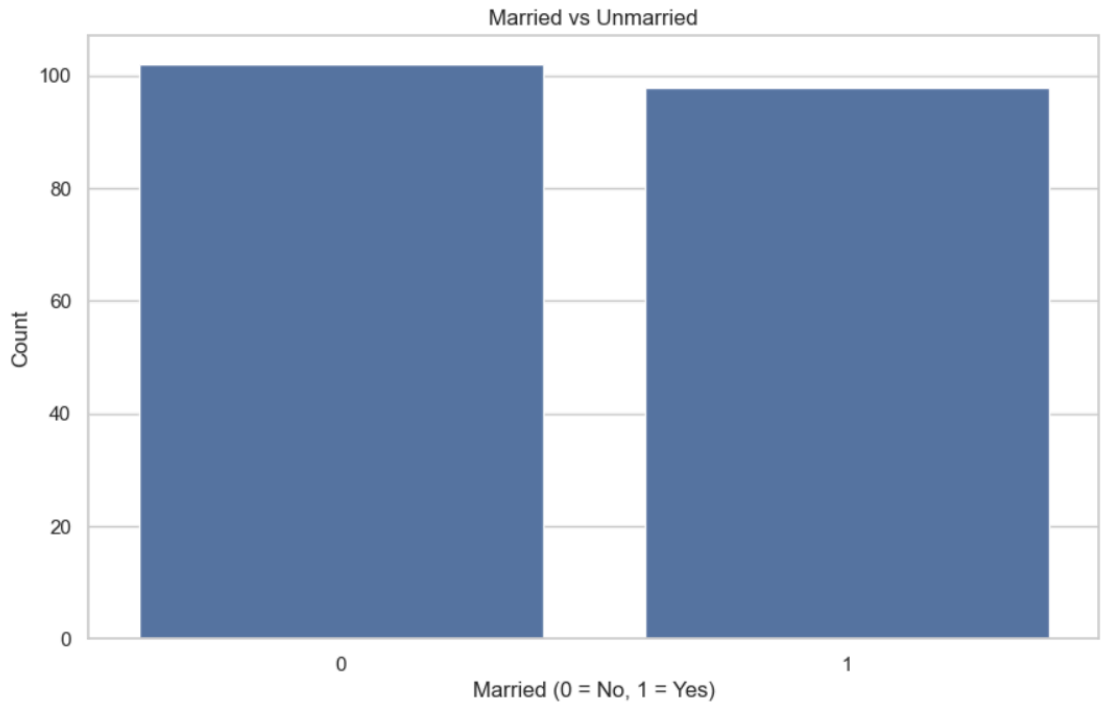
	User_id	User_name	DOB	PIN	City	State	Married	No_of_children	No_of_dependents	Source_of_income	Age
0	U100845	Arjun Patil	1994-07-13	826002	Dhanbad	Jharkhand	0	0	5	Software Engineer	29
1	U100879	Nisha Sengupta	1999-07-06	827002	Bokaro	Jharkhand	1	0	1	Software Engineer	25
2	U100997	Kiara Sethi	1996-09-14	831004	Jamshedpur	Jharkhand	0	0	1	Farmer	27
3	U100855	Vihaan Rao	1999-06-03	825304	Hazaribagh	Jharkhand	1	3	5	Engineer	25
4	U100988	Priya Chaudhary	1994-11-13	834002	Ranchi	Jharkhand	0	0	3	Manager	29

Age_Distriution:

```
Age
25    0.120
26    0.115
28    0.100
27    0.090
29    0.090
34    0.085
31    0.085
30    0.080
32    0.080
33    0.080
35    0.075
Name: proportion, dtype: float64
```

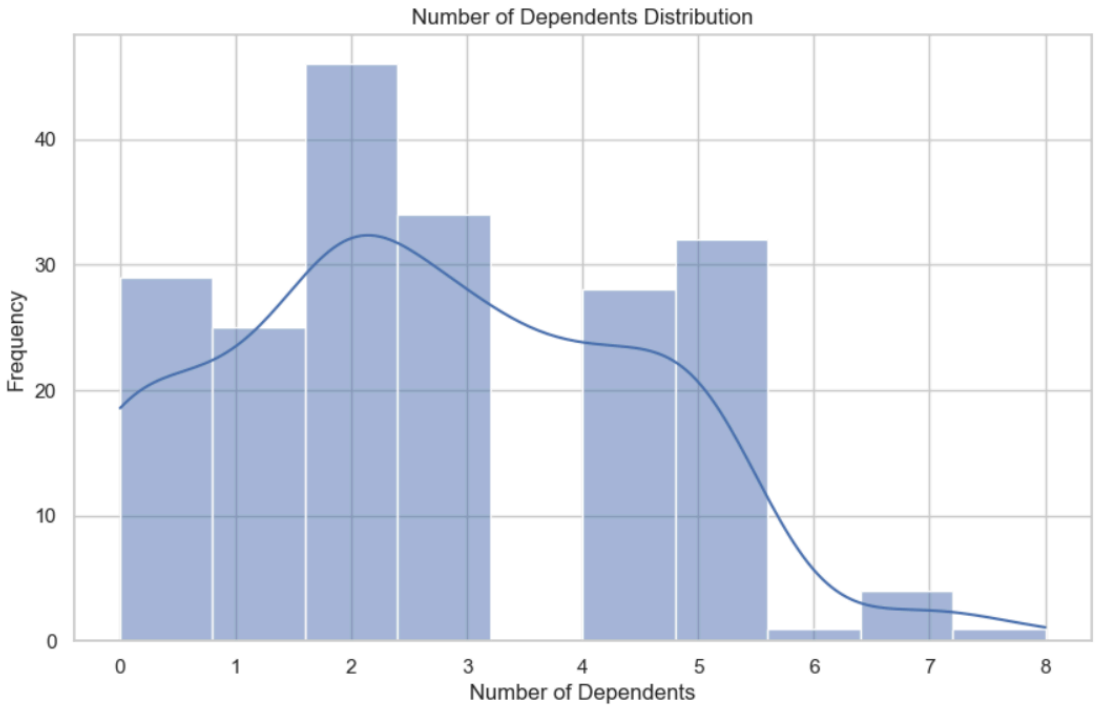


Married_status:



No. Of Dependents:

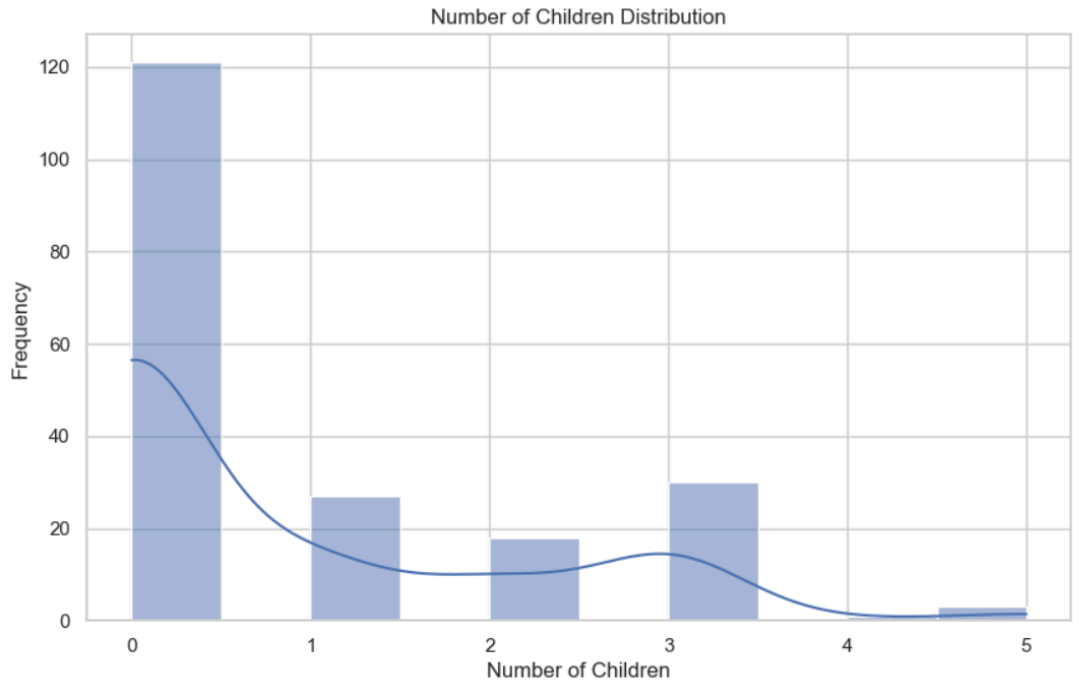
```
No_of_dependents
2    0.230
3    0.170
5    0.160
0    0.145
4    0.140
1    0.125
7    0.020
8    0.005
6    0.005
Name: proportion, dtype: float64
```



No. of Children:

This includes the values of the unmarried people which make the visualization inaccurate to understand.

```
No_of_children
0    0.605
3    0.150
1    0.135
2    0.090
5    0.015
4    0.005
Name: proportion, dtype: float64
```



No. of Children with respect to Married Status:

On an average, the married person have majority of 1,2 and 3 children.