

Task-01



Create a bar chart or histogram to visualize the distribution of a categorical or continuous variable, such as the distribution of ages or genders in a population.

Sample Dataset :-

<https://data.worldbank.org/indicator/SP.POP.TOTL>

PRODIGY INFOTECH

Welcome to my submission for Task 1 of the Data Science Internship at Prodigy Infotech. In this task, I have performed Exploratory Data Analysis (EDA) on a dataset provided, focusing on creating a visualization to represent the distribution of a categorical or continuous variable.

Dataset

The dataset used for this task is worldpopulationdata. This dataset contains records of population from the year 2001 to 2022.

Tools and Libraries used

- Jupyter notebook
- Pandas
- Numpy
- Matplotlib & Seaborn for visualization

Exploratory Data Analysis (EDA)

During the EDA process, I performed the following steps:

1. **Data Cleaning:** Checked for missing values, duplicates, and outliers in the dataset and handled them accordingly.
2. **Visualization:** Created a bar chart, stacked chart to visualize the distribution of a categorical or continuous variable.

Conclusion

In conclusion, this EDA process provided valuable insights into the distribution of the selected variable in the dataset. This analysis lays the foundation for further exploration and modeling tasks in the data science workflow.

Thank you for reviewing my submission!