**TITLE PAGE**

**DATA SCIENCE PROJECT**

**SUBMITTED BY: DIVYANATH MISHRA**

**REG. NO.: 12316777**

**SUBMITTED TO: Ms. Gargi Sharma (Faculty Coordinator)**

**SUBJECT: DATA SCIENCE TOOLBOX (PYTHON PROGRAMMING)**

**SUBJECT CODE: INT375**

**IN PARTIAL FULFILMENT FOR REQUIREMENTS OF THE AWARD OF THE DEGREE OF**

**COMPUTER SCIENCE AND ENGINEERING**



**CERTIFICATE**

This is to certify that DIVYANATH MISHRA bearing Registration no. 12316777 has completed INT375 project titled, **“SKILL BASED ASSIGNMENT”** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

**Signature and Name of the Supervisor**

**Designation of the Supervisor**

**School of Computer Science**

Lovely Professional University

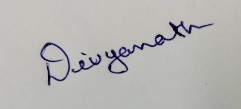
Phagwara, Punjab.

Date:

**DECLARATION**

I, DIVYANATH MISHRA, student of Data Science under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 10-04-2025



Signature

Registration No. 12316777

Name of the student : DIVYANATH MISHRA

**ACKNOWLEDGEMENT**

I would like to express my heartfelt gratitude to everyone who played a significant role in the successful completion of this data science project. First and foremost, I am deeply thankful to my mentors and academic guides whose constant encouragement, expert insights, and constructive feedback have been instrumental in shaping this project. Their guidance helped me navigate challenges and approach problems with a clear analytical perspective. I would also like to acknowledge the Government of India and various open data initiatives for providing access to rich and reliable datasets, without which this analysis would not have been possible. The availability of such data enabled a deeper understanding of the educational landscape in India and empowered this research with real-world relevance. I am equally grateful to my peers for the stimulating discussions and collaborative spirit, which enriched my learning experience. Finally, I extend my sincere appreciation to my family and friends, whose unwavering support, patience, and motivation provided the foundation I needed to stay focused and committed throughout this journey. This project has not only enhanced my technical skills but also deepened my awareness of the educational challenges and opportunities in India, and I am truly thankful for the opportunity to contribute to such a meaningful topic.

**TABLE OF CONTENT**

1. Introduction
2. Problem Statement
3. Objective
4. Source Of The Dataset
5. Code:
6. EDA
7. Objective one
8. Objective two
9. Objective three
10. Objective four
11. Objective five
12. Conclusion
13. Future Scope
14. References

**INTRODUCTION**

The objective of this project is to analyze funding allocation data with respect to geography and time using Python-based tools.

The goal is to uncover insights about regional funding disparities and budget distribution patterns, allowing for data-driven decisions in public planning.

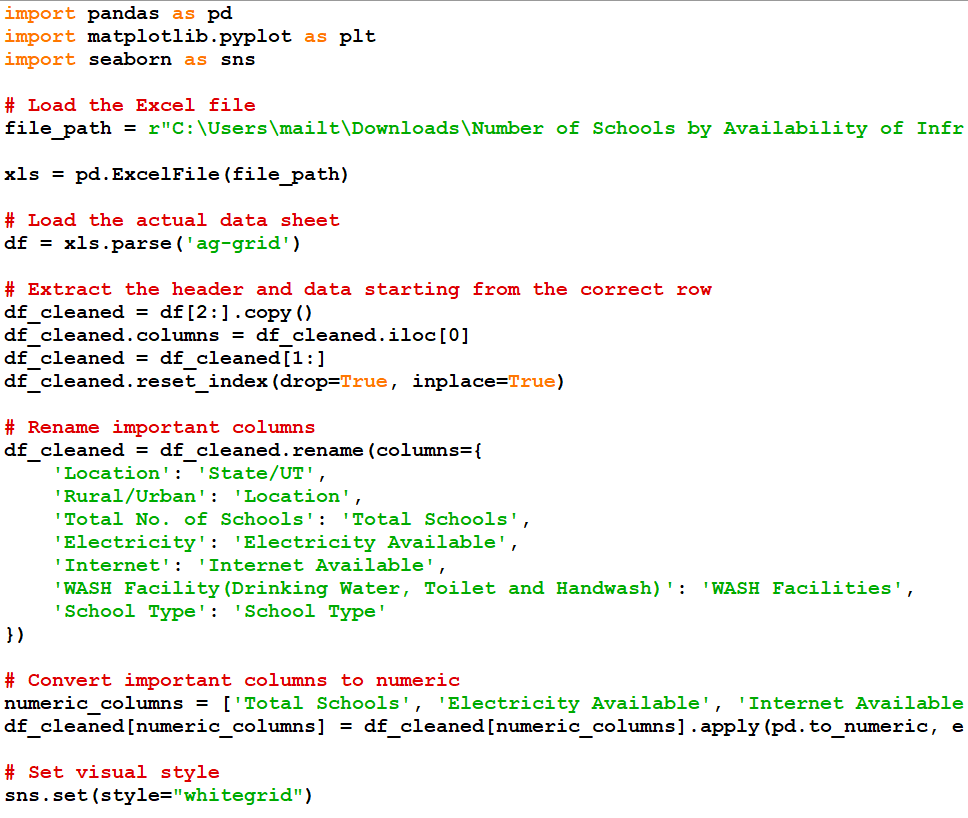
The analysis focuses on identifying trends, top states by funding, and regions that may be underfunded.

**SOURCE OF THE DATASET:**

<https://dashboard.udiseplus.gov.in/udiseplus-archive/#/reportDashboard/sReport>

NAME OF THE DATASET: Number of Schools by Availability of Infrastructure and Facilities, School Management and School Category

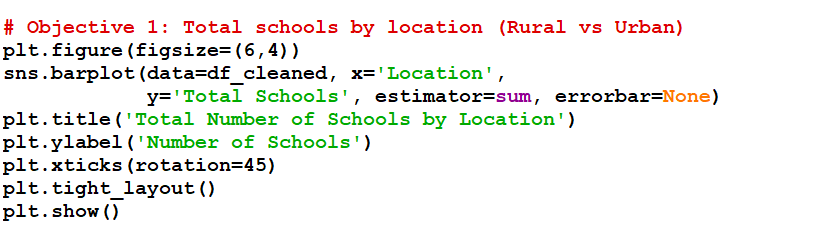
**EDA:**

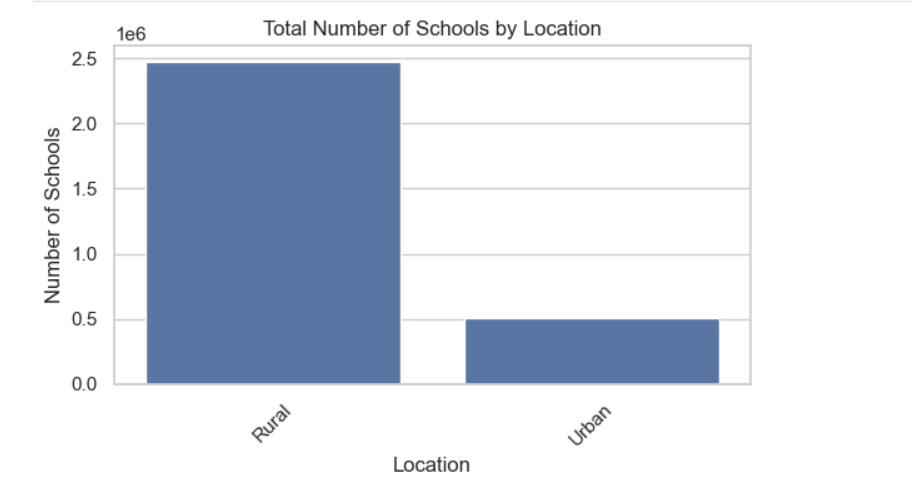


**OBJECTIVES:**

OBJECTIVE ONE:

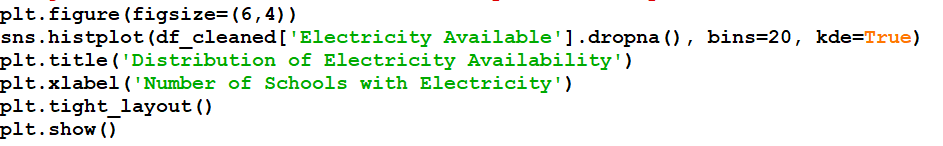
**Total schools by location (Rural vs Urban)**

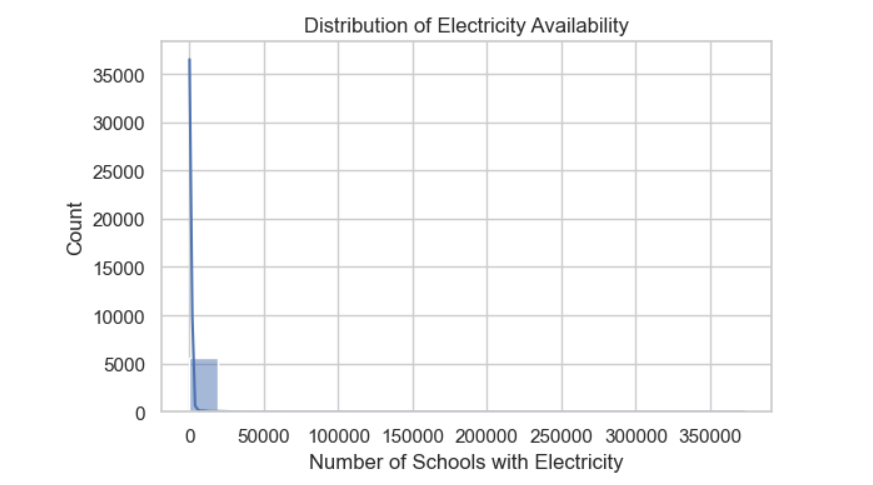




OBJECTIVE TWO:

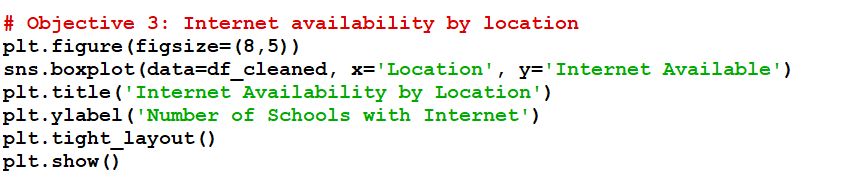
**Distribution of Electricity Availability**

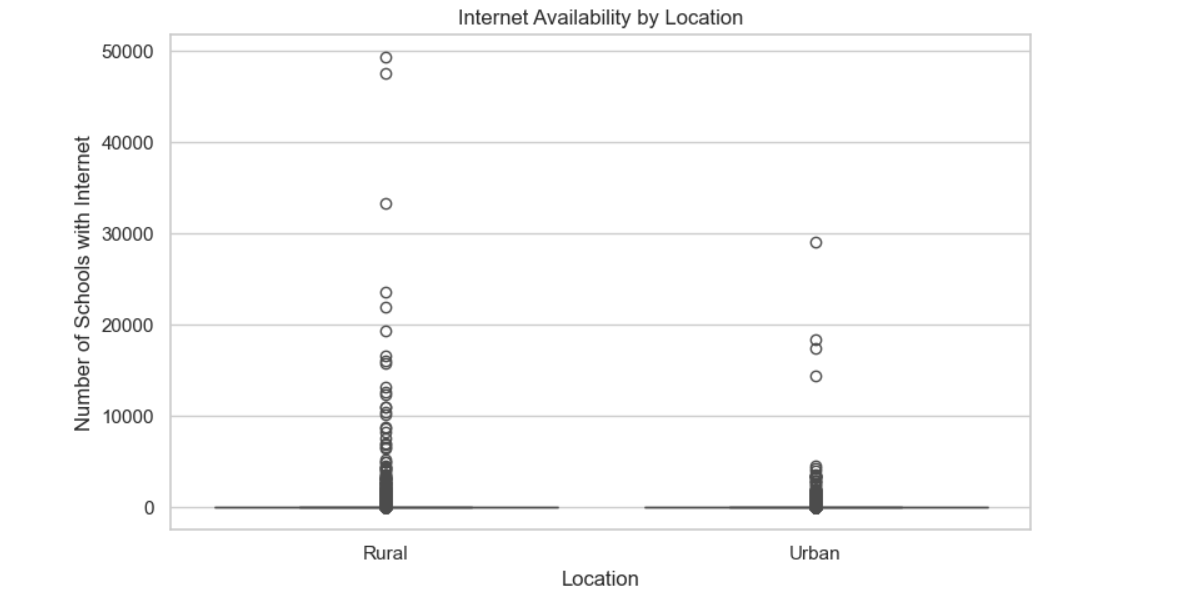


****

OBJECTIVE THREE:

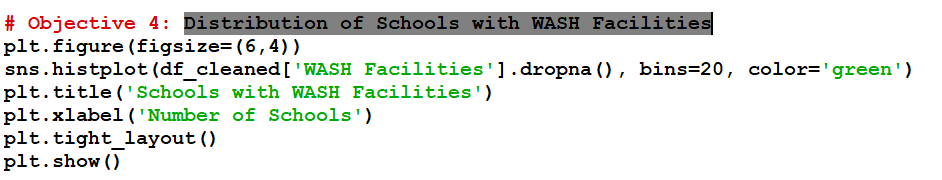
**Internet availability by location**

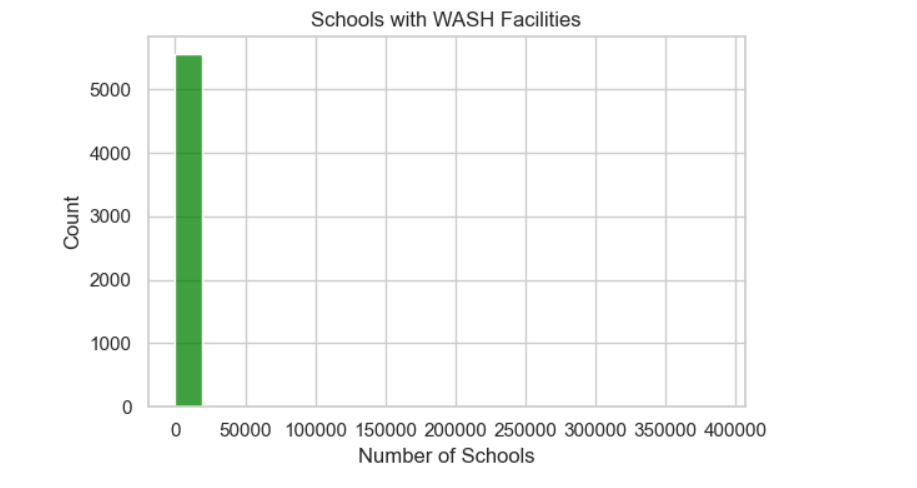




OBJECTIVE FOUR:

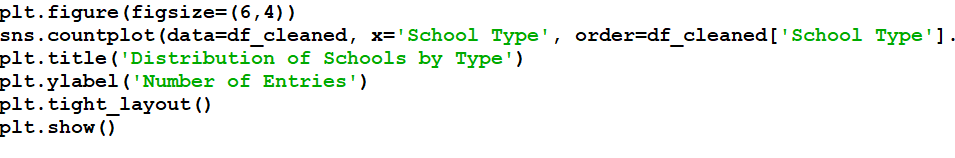
Distribution of Schools with WASH Facilities

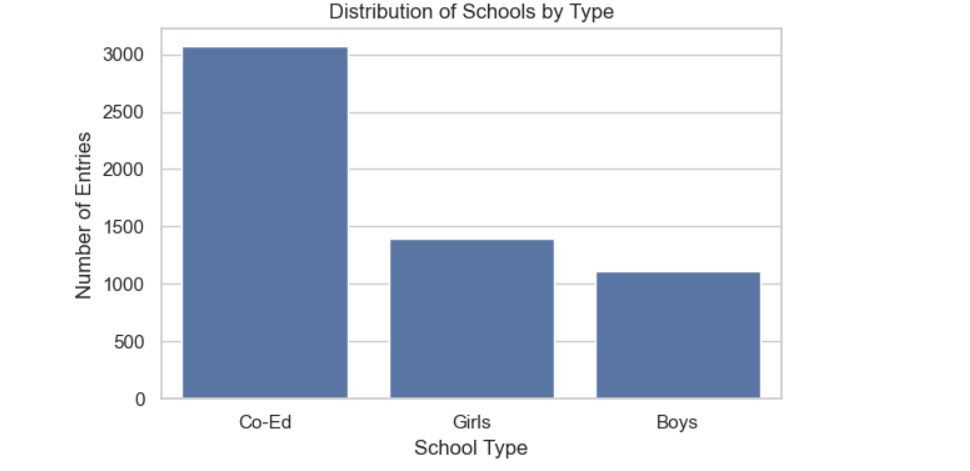




OBJECTIVE FIVE:

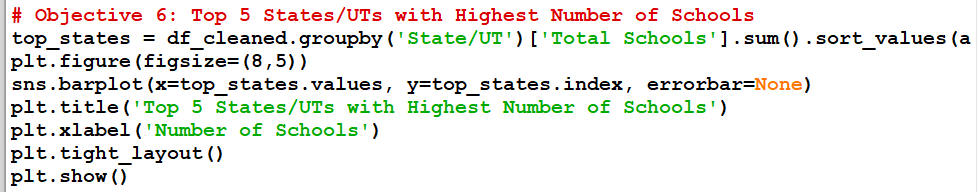
Distribution by School Type

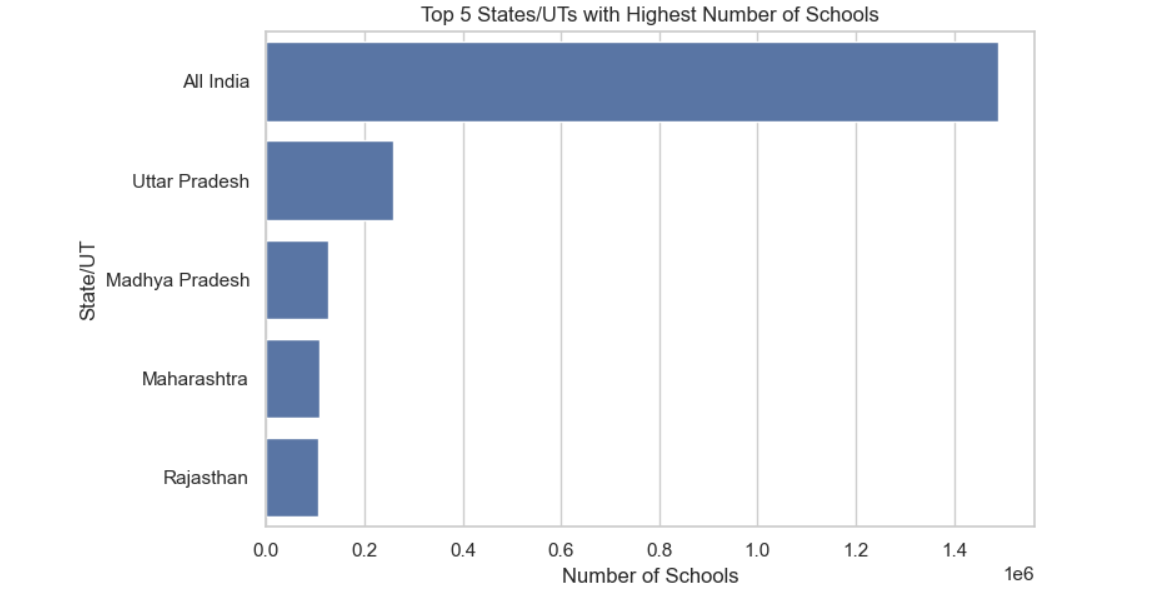




OBJECTIVE SIX:

Top 5 States/UTs with Highest Number of Schools





**CONCLUSION:**

The project revealed significant regional variation in government funding.

It emphasized the need for more balanced allocation models and highlighted data-driven strategies for identifying underserved regions.

The study also pointed to data integrity challenges due to outliers or gaps.

**FUTURE SCOPE:**

Extend dataset to include more years (2015–2024)

Integrate population and development indices to create need-based models

Deploy real-time dashboards for visualizing live funding patterns

Predictive modeling using machine learning to forecast future fund flows

**REFERENCES:**

1. <https://www.w3schools.com/python/numpy/numpy_random_seaborn.asp>
2. <https://www.w3schools.com/python/pandas/default.asp>
3. <https://www.w3schools.com/python/numpy/numpy_intro.asp>
4. <https://www.w3schools.com/python/matplotlib_intro.asp>
5. <https://chatgpt.com/>?
6. Data science from scratch first principles with python
7. Numpy, Pandas, Matplotlib, Seaborn Notes