

# **Internship Report**

**Title: Data Analytics Internship at NullClass**

**Intern Name: Anshul Pundir**

**Internship Duration: 20th February 2025 – 20th  
March 2025**

**Organization: NullClass**

## 1. Introduction

This report provides an overview of my internship experience at NullClass, where I worked on data analytics tasks related to Google Play Store data. The internship aimed to enhance my skills in data visualization, sentiment analysis, and time-series forecasting using Python. Through this internship, I applied theoretical knowledge to real-world datasets and gained practical experience in handling and analyzing large data.

## 2. Background

NullClass provides industry-focused training programs and internships in data science, web development, and analytics. As part of this internship, I worked on real-time data analysis tasks, requiring me to extract insights from large datasets, create interactive visualizations, and apply advanced filtering techniques to present meaningful insights.

### 3. Learning Objectives

- Develop hands-on experience with data visualization techniques.
- Implement sentiment analysis using Python.
- Build interactive dashboards with Plotly and Matplotlib.
- Apply time-series analysis to identify trends in app installs.
- Enhance problem-solving and research skills by working independently.

## 4. Activities and Tasks Performed

During this internship, I worked on the following tasks:

### **Task 1: Sentiment Distribution Analysis**

- Visualized the sentiment distribution of user reviews using a stacked bar chart.
- Grouped ratings into three categories: 1-2 stars, 3-4 stars, and 4-5 stars.
- Filtered data to include only apps with more than 1,000 reviews.
- Focused on the top 5 app categories for meaningful insights.

### **Task 2: Interactive Choropleth Map**

- Built a global installs map using Plotly.
- Applied filters to show data only for the top 5 app categories.
- Highlighted app categories where installs exceeded 1 million.
- Ensured the visualization was visible only between 6 PM – 8 PM IST.

### **Task 3: Time Series Analysis**

- Created a line chart to visualize install trends over time.
- Highlighted periods where installs increased by more than 20% month-over-month.
- Included only apps with reviews greater than 500.
- Ensured the graph was visible only between 6 PM – 9 PM IST.

## 5. Skills and Competencies Developed

- Data Visualization: Gained proficiency in Matplotlib, Seaborn, and Plotly.
- Data Cleaning & Preprocessing: Used Pandas and NumPy to filter and clean datasets.
- Sentiment Analysis: Applied text processing techniques to categorize user reviews.
- Time-Series Analysis: Implemented trend detection and visual interpretation of app installs.
- Problem-Solving: Developed independent research skills to troubleshoot issues.

## 6. Challenges and Solutions

Challenge	Solution
Managing large datasets efficiently	Used Pandas optimizations and data filtering to handle large CSV files
Implementing time-based filtering in dashboards	Used Datetime functions to dynamically control display times
Finding appropriate data sources for analysis	Referred to Google Play Store datasets and research papers for validation
Debugging visualization issues	Used interactive debugging and community resources like Stack Overflow
Lack of country or state data for the Choropleth map.	Replaced it with a bar chart to visualize category-wise install distribution.



## 7. Outcomes and Impact

- Successfully built three interactive visualizations integrated into a dashboard.
- Improved data analysis and storytelling skills.
- Gained confidence in handling real-world data science projects.
- Strengthened my portfolio with practical data analytics projects.

## 8. Conclusion

This internship at NullClass provided me with valuable experience in data analytics and visualization. The hands-on tasks helped me apply theoretical knowledge to real-world datasets, improve problem-solving skills, and work independently on research-based projects. The experience has strengthened my foundation in data science and prepared me for future challenges in the field of analytics.

## 9. References & Submission Details

- GitHub Repository Link: [Your GitHub Repo URL]
- Live Dashboard Link: [Netlify/Vercel URL]
- Dataset Source: Google Play Store Analytics Dataset

This report is submitted as part of my internship completion requirements at NullClass.

Submitted by: Anshul Pundir

Date: [20/03/2025]