Internship Initial Status Report

##### **Name: Pasala Swathi Sri**

##### **Roll No: 24M11MC264**

##### **Batch: Batch 01**

##### **Duration: 19th May 2025 – Ongoing**

##### **Domain: Data Science & Big Data Analytics**

##### **Company: Blackbucks**

# 1. Weekly Learnings

## Week 1: (19th – 24th May 2025)

- Attended onboarding session on tools (Python, ChatGPT, Cloud Computing).  
- Completed Domain Test on statistics and Python basics.  
- Attempted MET (Mandatory Evaluation Test) covering logic & coding.  
- Outcome: Understood data science lifecycle; overview of Hadoop ecosystem; identified gaps in Python logic.

## Week 2: (26th – 31st May 2025)

- Participated in Placement Test-1 (Excel analytics & Python basics).  
- Live Sessions:  
 - 27th May: Pandas – filtering, grouping, merging, missing data handling.  
 - 30th May: Visualizations using Matplotlib & Seaborn.  
- Completed real-world dataset cleaning task using Pandas.  
- Created COVID-19 trend analysis using bar charts and line plots.  
- Outcome: Improved data wrangling and visualization skills.

## Week 3: (2nd – 6th June 2025)

- Attended live session on statistical analysis and probability.  
 - Topics: mean, median, mode, standard deviation, hypothesis testing.  
- Attempted quiz involving Python-based statistical measures.  
- Outcome: Applied statistics in Python; enhanced analytical interpretation of results.

# 2. Use Case Demonstration: Sales Insights

**Objective:**  
Analyze retail sales data to identify regional performance and peak seasons.  
  
**Techniques Used:**- Cleaned dataset using drop\_duplicates() and fillna() in Pandas.  
- Analyzed monthly revenue with groupby() function.  
- Created visualizations with Matplotlib to compare regional sales trends.  
  
**Outcome:**  
Discovered Region A contributed 40% to total sales. Recommended early stock replenishment for December.

# 3. Challenges & Improvements

|  |  |
| --- | --- |
| **Challenge** | **Solution** |
| Debugging Pandas logic | Used print(df.head()) and debugged stepwise |
| Choosing appropriate chart types | Referred to Seaborn documentation & examples |
| Interpreting statistical outputs | Practiced with datasets and reviewed core concepts |