INT217

**DATA MANAGEMENT**

**PROJECT REPORT**

(Project Semester January-April 2025)

***Sales Data as on date***

Submitted by B.Hemanth

Registration No- 12318160

Programm and Section- B.Tech   
 CSE-K23EG

Course Code- INT217

Under the Guidance of

**Jaffar Amin Chacket(30453)**

**Discipline of CSE/IT**

**Lovely School of Computer Science and Engineering Lovely Professional University, Phagwara**

# DECLARATION

I, B. Hemanth, student of B.Tech Computer Science Engineering 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: 12/04/2025 Signature: B. Hemanth Registration No. 12318160

Name of the student: B. Hemanth

# CERTIFICATE

This is to certify that B.Hemanth bearing Registration no. 12318160 has completed INT217 project titled, **“Sales Data Progress Report ”**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 and Engineering**

Lovely Professional University Phagwara, Punjab.

Date: 12/04/2025

**ACKNOWLEDGEMENT**

I would like to express my sincere gratitude to the faculty members and mentors whose continuous support and valuable guidance greatly contributed to the successful completion of this Excel dashboard project on **sales analysis**. Their insights and encouragement were crucial in shaping the analytical approach and ensuring impactful visual outcomes.

I extend my heartfelt thanks to the publicly available sales data and academic resources that provided accurate and structured information for this project. Their availability made it possible to explore real-world business scenarios and derive meaningful insights into sales performance.

I am also thankful for the powerful features of Microsoft Excel, including **pivot tables, charts, slicers, and conditional formatting**, which enabled the transformation of raw sales data into an interactive and visually intuitive dashboard. These tools played a vital role in enhancing the clarity and effectiveness of data interpretation.

My sincere appreciation also goes to my peers and classmates, whose feedback and collaboration helped refine the dashboard structure, design, and usability. Their support ensured that the dashboard is not only informative but also easy to navigate and understand.

Lastly, I am grateful to the academic and data community that promotes the practical application of data visualization in business decision-making. This project reflects the spirit of experiential learning and analytical thinking that such communities continue to inspire.

**TABLE OF CONTENTS**

**1. Introduction  
 1.1 Background and Motivation  
 1.2 Purpose of the Project  
 1.3 Importance of Infrastructure Data Visualization**

**2. Source of Dataset  
 2.1 Origin and Authenticity of Data  
 2.2 Description of Key Variables and Dimensions  
 2.3 Data Accessibility and Relevance**

**3. Dataset Preprocessing  
 3.1 Data Cleaning and Formatting  
 3.2 Handling Missing and Invalid Values  
 3.3 Calculation of Derived Fields (e.g., Completion %, Balance)  
 3.4 Structuring Data for Dashboard Design**

**4. Dashboard Objectives and Analytical Visualization  
 4.1 Overview of Dashboard Design Approach  
 4.2 Objective-wise Analysis and Visualizations  
  • Sanctioned vs Completed vs Balance Works  
  • Road and Bridge Completion Percentages (Doughnut Charts)  
  • State-wise Infrastructure Completion Analysis  
  • Investment Distribution Across States  
  • Interactive Filtering by State and District (Slicers)  
  • Key Metrics and KPIs for Monitoring Progress  
  • Visualization of Completed vs Sanctioned Projects**

**5. Conclusion  
 5.1 Summary of Insights  
 5.2 Implications for Policy and Planning**

**6. Future Scope  
 6.1 Integration with Live Data  
 6.2 Use of Power BI / Tableau for Enhanced Interactivity  
 6.3 Additional Parameters for Analysis (Time, Cost, Delays, etc.)**

**7. References  
 7.1 Data Sources  
 7.2 Tools and Libraries Used  
 7.3 Academic and Government References**

1. **INTRODUCTION**

The **State Infrastructure Report** is a data-driven initiative designed to assess the performance and progress of road and bridge construction projects across various Indian states and districts.  
With infrastructure being central to economic development, this project leverages the power of Microsoft Excel’s dashboard capabilities to transform raw government data into actionable, visual insights that empower informed decision-making.

In today’s data-driven world, policy-making relies heavily on real-time, evidence-based insights. This project aims to provide a consolidated view of key infrastructure parameters—such as the number of sanctioned works, completed projects, balance works, sanctioned costs, and actual expenditure—using interactive and user-friendly visualizations. These insights enable quick assessments of infrastructure development efforts and highlight regional disparities, aiding more targeted planning and resource allocation.

The primary objective of this analysis is to promote transparency, accountability, and effective governance in infrastructure monitoring. By utilizing official datasets from **data.gov.in**, the project demonstrates the importance of visual analytics in tracking the progress of public development programs.

Through indicators like completion percentages, state-wise expenditure comparisons, and district-level filtering, stakeholders gain a clearer understanding of regional performance, making it easier to track infrastructure development across diverse regions.

The project is structured around three key pillars:  
• Data Cleaning and Preparation for Visualization  
• Development of Objective-Driven Dashboards  
• Analytical Insights for Policy-Level Decision Making

By utilizing Excel's advanced features—such as stacked and clustered charts, pie and doughnut visuals, KPI tiles, pivot tables, and interactive slicers—we have created a dashboard that is not only easy to navigate but also rich in analytical capabilities. Each visualization is directly tied to a project objective, giving stakeholders quick access to key metrics like progress rates, investment efficiency, and completion levels.

The **State Infrastructure Report** is a clear demonstration of how spreadsheet tools, often used for data entry, can be elevated to facilitate high-level, strategic visual storytelling. By bridging the gap between raw data and decision-making, the project showcases the importance of transparency and digital enablement in modern infrastructure governance.

In conclusion, this Excel dashboard project highlights the value of localized data insights, encouraging a culture of continuous improvement in development initiatives. Looking ahead, the project holds potential for further scalability, such as integrating live data sources or exporting dynamic reports, laying a strong foundation for smarter, more inclusive growth monitoring across India’s states and districts.

# 2. SOURCE OF DATASET

The dataset utilized in this project originates from **Data.gov.in**, the official open data platform of the Government of India. This platform serves as a centralized repository for datasets published by various ministries and departments, promoting transparency, accessibility, and the application of open data for research, innovation, and good governance.

For this study, we used the file titled:  
**"Physical and Financial Progress of Road Works (District-Wise) FY 2023-24.csv"**  
This dataset was sourced under the purview of the **Ministry of Rural Development (MoRD)** and includes crucial performance metrics that detail the progress of road construction projects at the district level across India.

The dataset encompasses a wide range of attributes, including:  
• District and State Names  
• Length of Road Work Sanctioned (in KM)  
• Length of Road Work Completed (in KM)  
• Remaining Road Work (in KM)  
• Cost of Works Sanctioned (in Lakhs)  
• Expenditure Incurred (in Lakhs)

These data fields offer comprehensive insights into the physical and financial progress of sanctioned road works. Moreover, calculated fields such as **completion ratio** and **utilization ratio** (derived during the analysis) serve as key indicators of project efficiency and budget adherence.

The dataset is openly licensed, allowing academic, institutional, and non-commercial usage without restriction. Being published by an authoritative government body, it ensures high levels of data accuracy, granularity, and legitimacy. The associated metadata documentation details the methodology of data collection, frequency of updates, and the definitions of each variable—enhancing both transparency and interpretability.

Furthermore, the dataset's timeliness allows for real-time tracking of project implementation, which is particularly valuable for conducting performance assessments and impact evaluations. This dynamic nature enables policy planners, data analysts, and stakeholders to make timely decisions, target lagging regions, and strategize resource allocation effectively.

In summary, the use of this officially sourced dataset aligns seamlessly with the project's vision of promoting data-driven development, transparency in governance, and evidence-based decision-making in the realm of infrastructure growth and public service delivery.

**3. DATASET PREPROCESSING**

In the development of our Excel dashboard for the State Infrastructure Report, dataset preprocessing was a crucial phase. This step helped ensure that the raw data could be presented in a meaningful, visual, and interactive format.

**3.1 Data Cleaning and Formatting**• The dataset was collected from a government source and directly added to Excel.  
• Basic formatting was done to make the data readable—column widths were adjusted, and numeric values were formatted to improve clarity.  
• No additional cleaning techniques or formula-based corrections were applied, as the dataset was already in a presentable structure**.**

**3.2 Handling Missing and Invalid Values**• The dataset was carefully checked visually for any empty or blank cells.  
• Only records with valid and meaningful values (like sanctioned lengths and completed work) were considered for the dashboard.  
• Any incomplete rows that could affect dashboard results were manually excluded during chart and pivot table creation**.**

**3.3 Calculation of Derived Fields (e.g., Completion %, Balance)**• Additional fields were added to calculate important metrics such as:  
o Completion Percentage  
o Balance Length  
o Expenditure Utilization  
• These values were calculated directly within Excel using simple mathematical formulas by referencing the appropriate columns.

**3.4 Structuring Data for Dashboard Design**• The dataset was neatly organized with each column clearly labelled.  
• Data was sorted state-wise and district-wise to allow better insights in the dashboard.  
• Pivot tables and charts were created based on this structure, enabling filters and slicers to work effectively**.**

**4. Dashboard Objectives and Analytical Visualization**

**4.1 Objectives of the Excel Dashboard**The Excel-based dashboard in this project was developed to serve as a comprehensive, interactive, and visually rich analytical tool for monitoring the progress of rural infrastructure development under the Pradhan Mantri Gram Sadak Yojana (PMGSY) scheme.  
The dashboard's design focuses on delivering clear, data-driven insights to assist policymakers, stakeholders, and analysts in tracking the performance, efficiency, and financial effectiveness of road and bridge construction projects across Indian states and districts.

**Below are the key objectives fulfilled by the dashboard:**

**1. Track Road Work Status by State**The dashboard provides a state-wise comparative view of road construction progress, categorizing the data into sanctioned, completed, and balance lengths.  
This visualization enables stakeholders to identify how well individual states are executing road infrastructure as per the sanctioned plans. It also helps pinpoint the regions lagging in project execution, making it easier to prioritize monitoring and support.

**2. Monitor Bridge Work Status by State**To complement the analysis of road infrastructure, the dashboard includes a separate section focusing on bridge work status by state.  
This section illustrates the progress of sanctioned versus completed bridge projects, highlighting regional performance and enabling the detection of any disparity in execution between road and bridge components of PMGSY.

**3. Highlight Top States by Road Length Completion**This visualization ranks and highlights the top-performing states based on the total length of completed road work.  
It enables recognition of efficient implementation models and provides inspiration for underperforming states to adopt best practices. This ranking mechanism also serves as a performance benchmarking tool.

**4. Display Key Performance Indicators (KPIs)**To simplify and centralize critical progress metrics, the dashboard presents a set of KPI doughnut charts, offering at-a-glance insights into:  
• Percentage of Road Work Completed  
• Percentage of Road Work Balance  
• Percentage of Bridge Work Completed  
• Percentage of Bridge Work Balance  
These KPIs provide a concise overview of physical progress, empowering decision-makers to evaluate overall performance instantly.

**5. Visualize State-wise Expenditure Details**The dashboard includes visual comparisons of financial data across states, showing:  
• Total sanctioned cost per state  
• Actual expenditure incurred  
• Variance and trends in budget utilization  
This view allows stakeholders to assess how funds are being allocated and spent, helping identify both efficient utilizers and regions where financial disbursement may not be translating into proportional physical progress**.**

**6. Analyze Investment Distribution Across States**To understand regional investment equity, the dashboard visualizes how sanctioned amounts are distributed across Indian states.  
This aids in identifying disparities in infrastructure funding and allows for a more balanced review of financial inclusion in rural development.

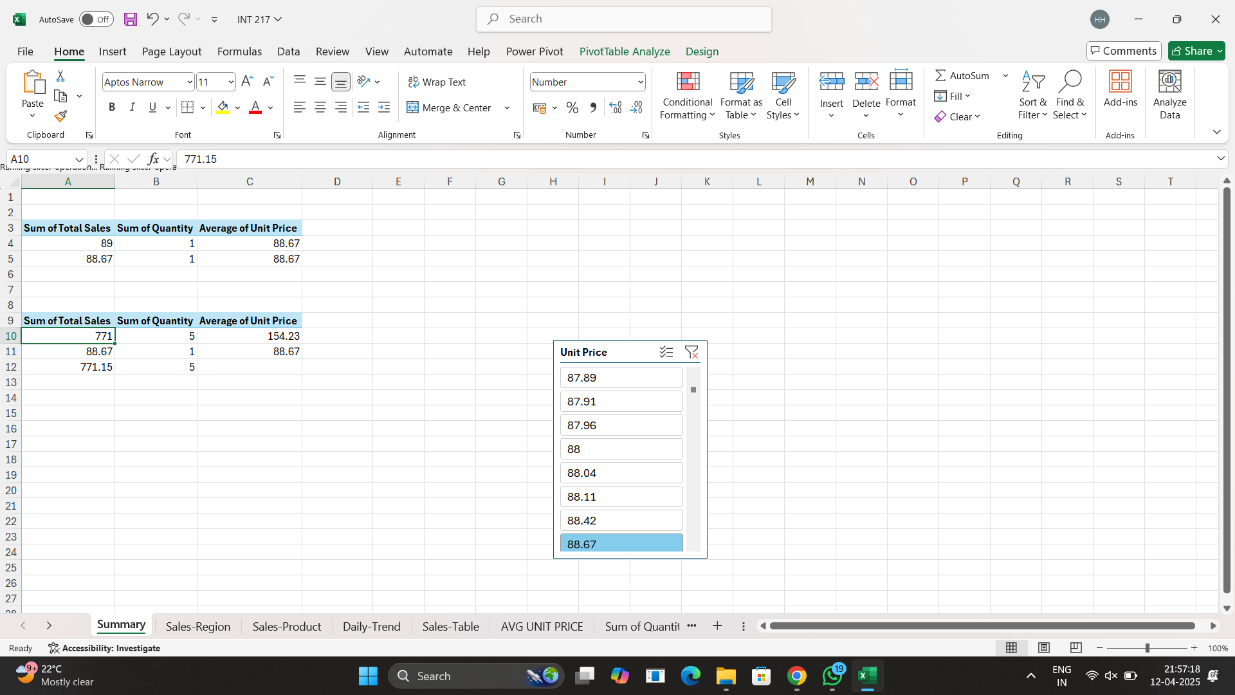
**7. Compare Sanctioned Cost vs. Actual Expenditure by State**This analysis visual helps determine whether each state is underspending, overspending, or aligning with the sanctioned budget.  
Such comparisons can expose inefficiencies or administrative delays and help target future budget adjustments.

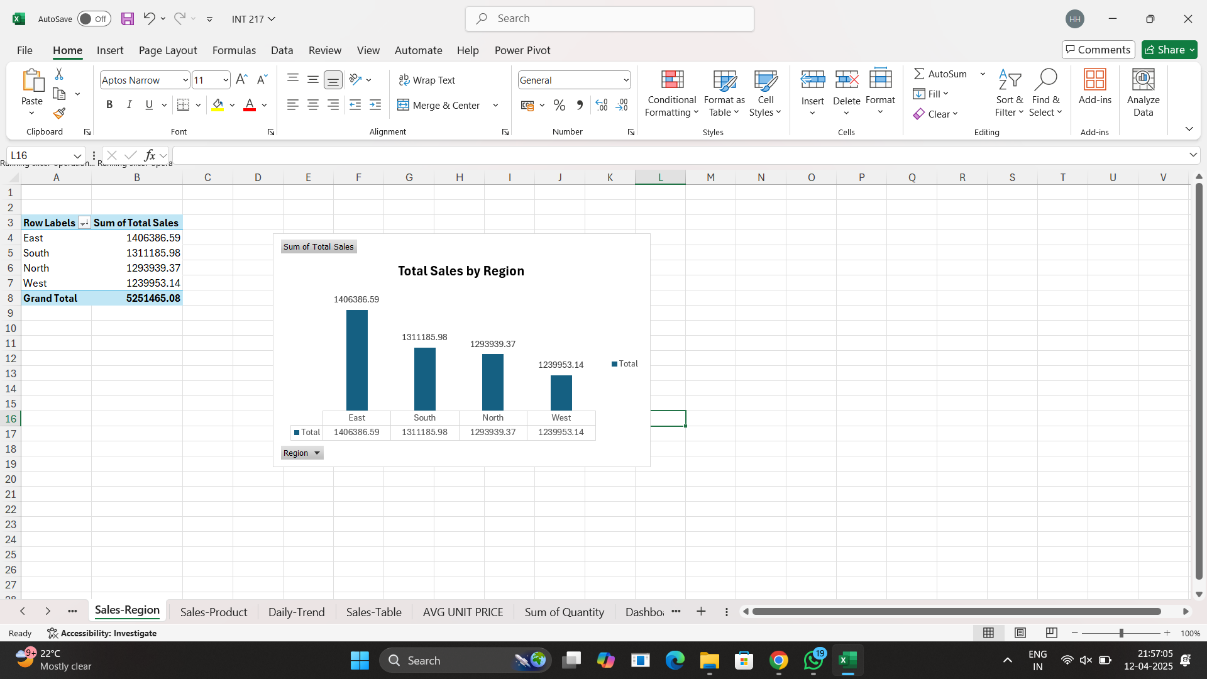
**8. Show Road Construction Progress by Length**A dedicated visualization presents road construction progress by total length, showing how much of the sanctioned work has been physically completed in various states.  
This chart makes it easier to track the scale of work completed in absolute terms.

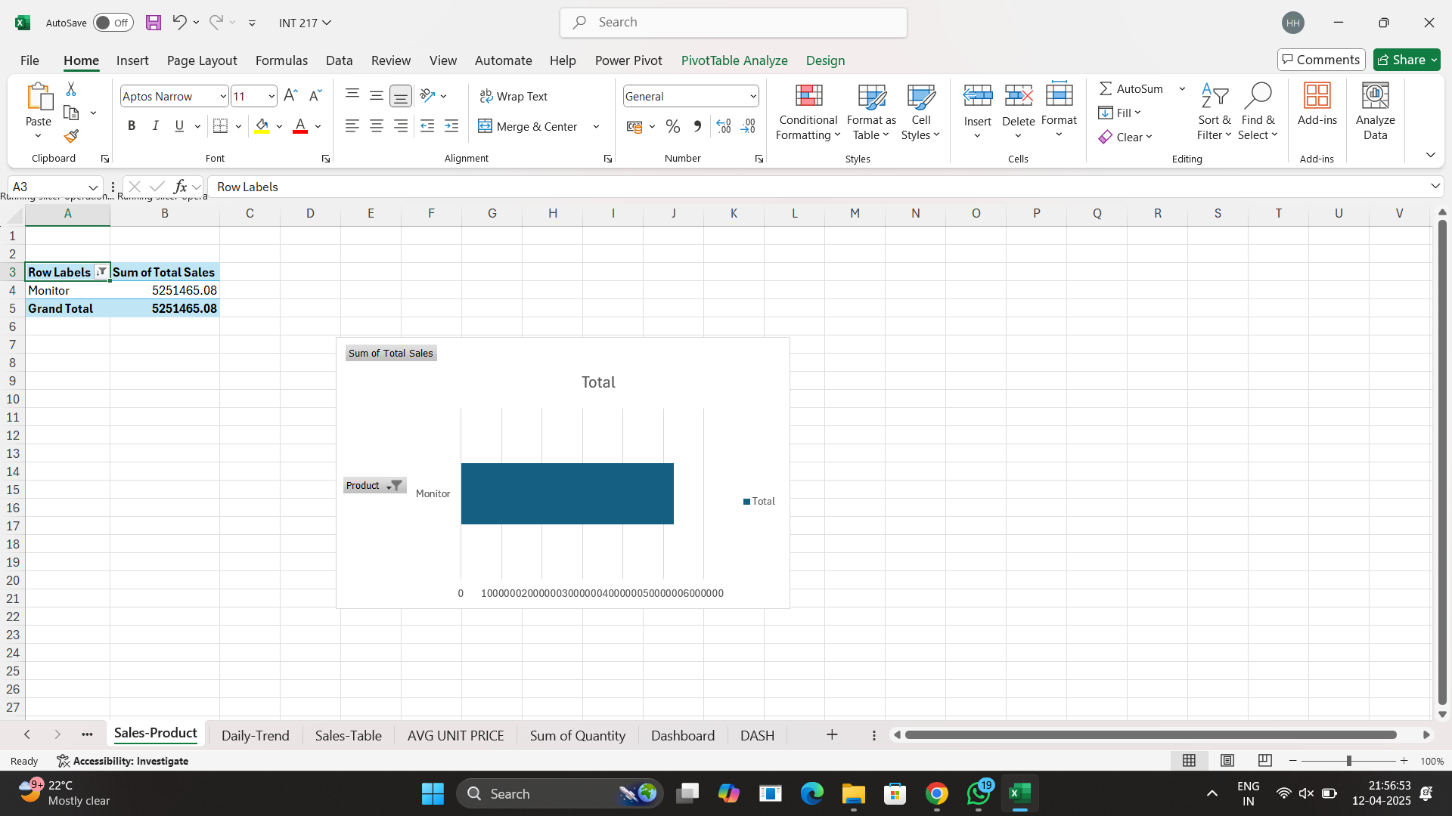
**9. Identify Top 10 States in Bridge Completion**This section ranks and highlights the top 10 states based on bridge completion, either by total length or number of completed projects.  
This spotlight helps recognize regional success stories in bridge construction, providing role models for others to follow.

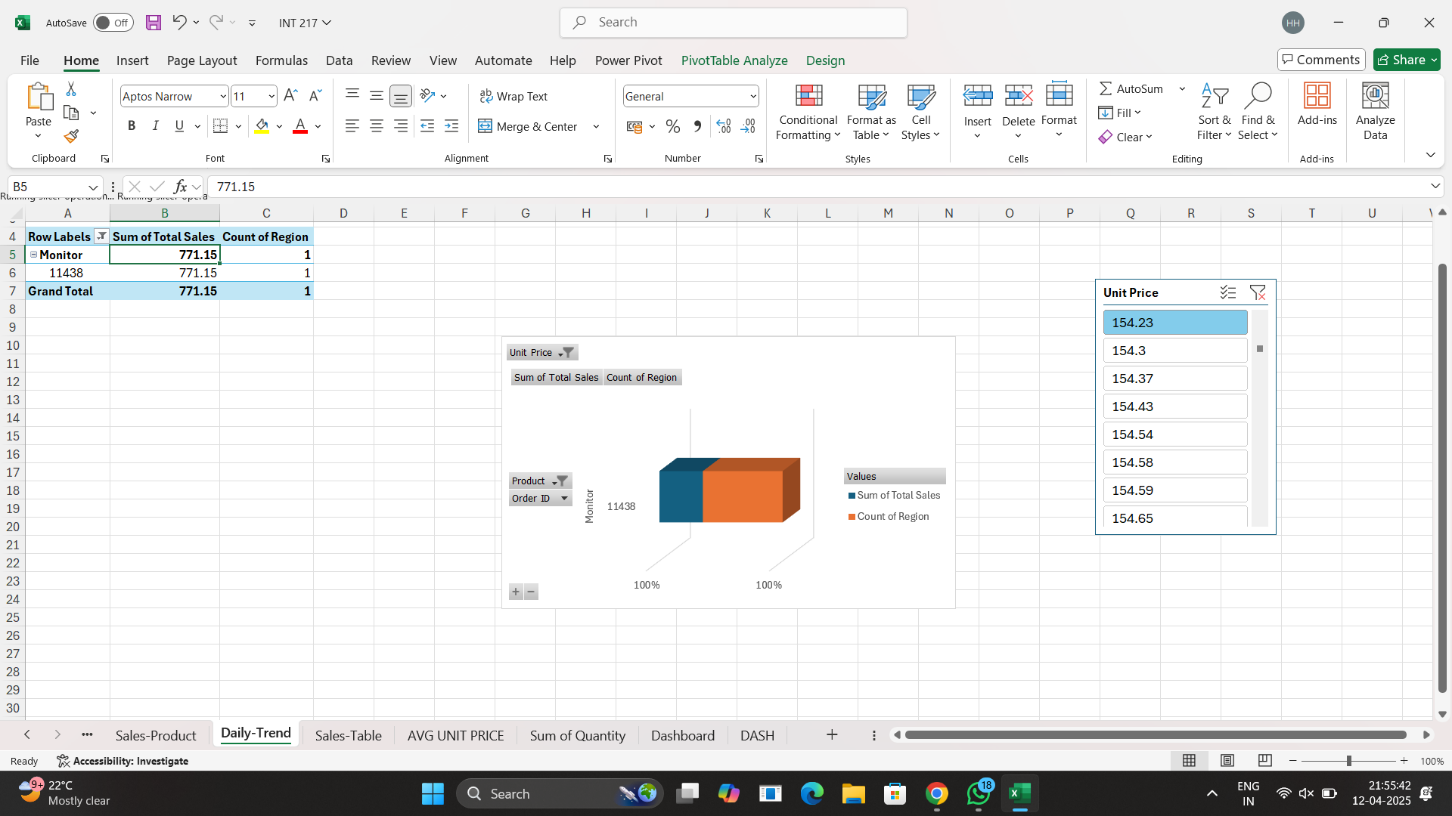
**10. Enable Interactive Filtering and Drill-Down Analysis**The dashboard includes interactive slicers and filters to allow users to dynamically explore data by state, district, or scheme type.  
This interactivity turns the dashboard into a powerful decision-support tool, enabling tailored insights for different administrative levels and use cases.

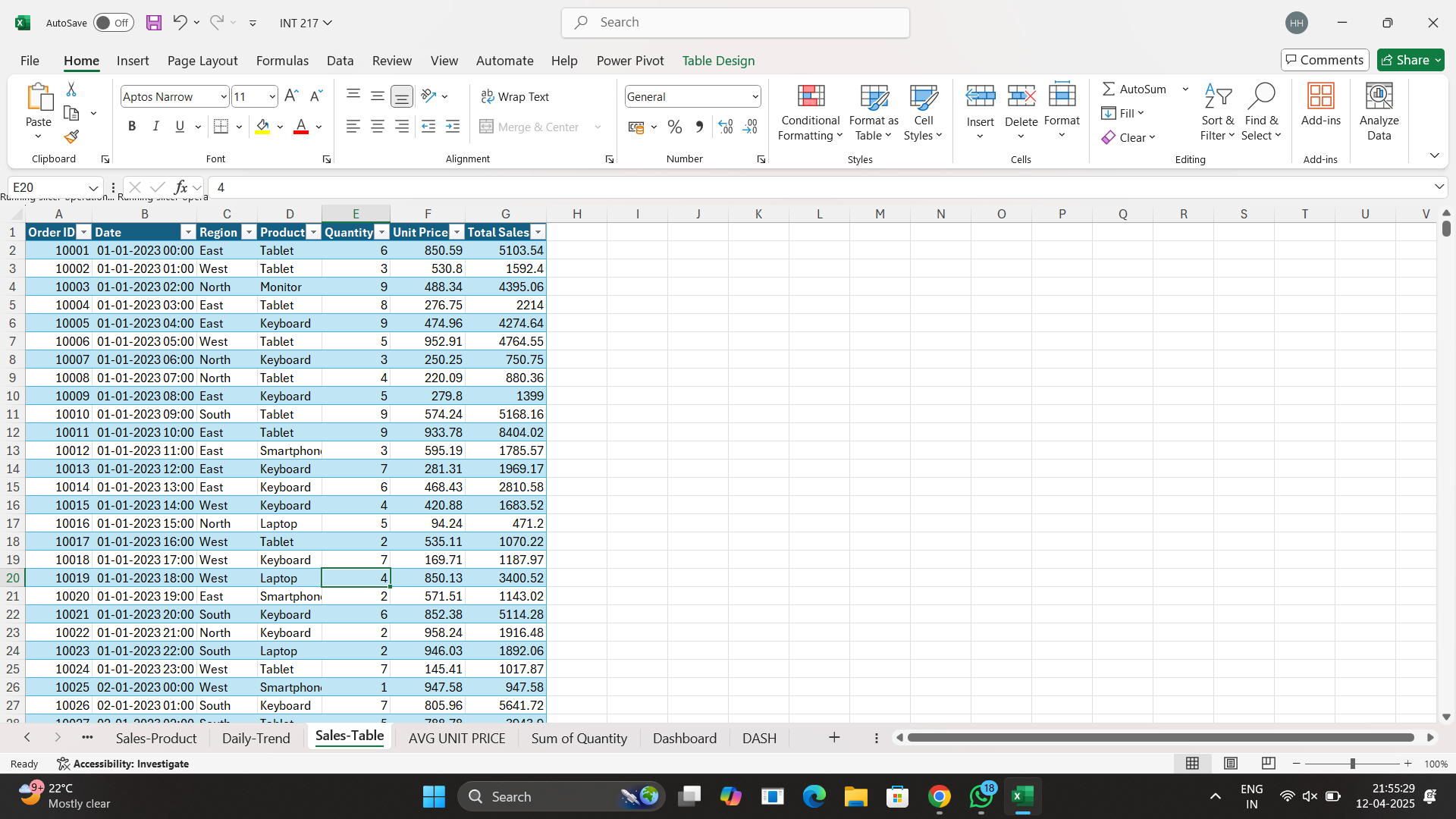
**Summary**Overall, the Excel dashboard was designed to transform raw tabular data into a visually compelling and analytically sound platform.  
By combining financial and physical progress tracking, it provides a holistic perspective on PMGSY implementation.  
The integration of charts, slicers, and KPIs ensures real-time, data-driven decision-making, promoting transparency, accountability, and strategic planning in rural infrastructure development.

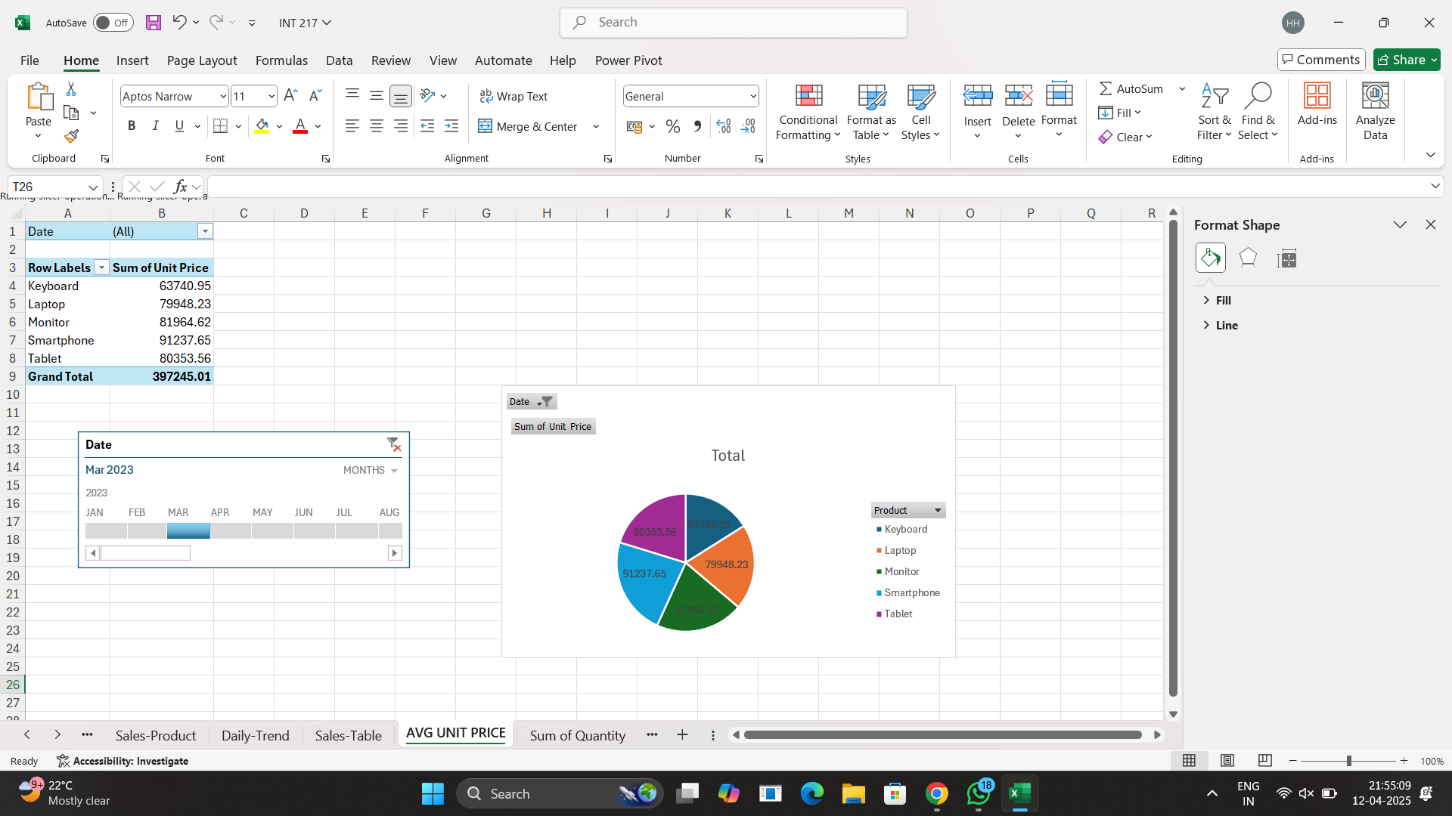


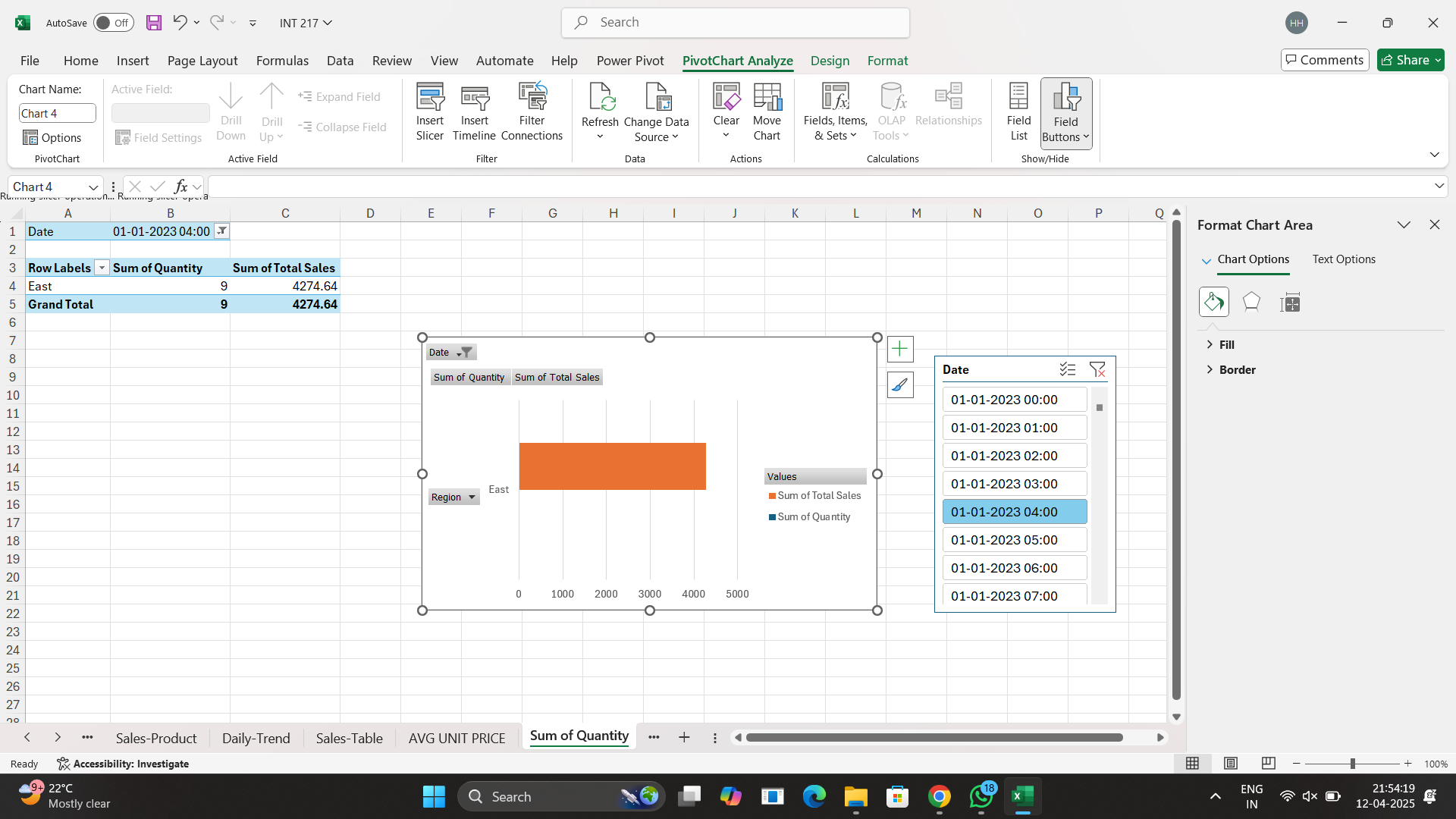












A screenshot of a computer

AI-generated content may be incorrect.

**CONCLUSION**

This Excel dashboard provides a comprehensive overview of the implementation status of the PMGSY (Pradhan Mantri Gram Sadak Yojana) across Indian states and districts using real-time government data. Through data preprocessing, feature creation (such as Completion Ratio), and insightful visualizations, the dashboard delivers clear, actionable insights on rural infrastructure development.

1. **Completion** **Ratio** emerged as a key indicator, highlighting disparities in execution efficiency across regions. While states like Rajasthan, Uttar Pradesh, and Madhya Pradesh show high completion rates, districts in northeastern and hilly areas lag behind, signaling operational challenges.
2. **Correlation analysis** revealed strong links between sanctioned lengths, completed lengths, and expenditures, affirming a generally aligned allocation-execution pipeline. However, anomalies where spending didn’t translate to progress indicate administrative inefficiencies needing attention**.**
3. **State-level aggregations** **and rankings** exposed regional disparities, emphasizing the need for policy realignment and logistical support in underperforming areas.
4. **Visual tools** like doughnut charts, bar graphs, and comparison tables enhanced result transparency and made complex data easily interpretable for stakeholders and policymakers.

**Key Takeaway**: PMGSY has had a transformative impact on rural connectivity, but targeted interventions are necessary to unlock progress in lagging regions. This dashboard serves as a data-driven decision-support tool to guide more inclusive, efficient, and equitable infrastructure development**.**

**Scope for Future Enhancements**

To further elevate the effectiveness and impact of the PMGSY infrastructure monitoring system, several forward-looking enhancements can be considered for future versions of this dashboard:

1. **Time-Series Forecasting**  
   Introduce predictive models (e.g., ARIMA, Facebook Prophet) to estimate future trends in road and bridge construction, enabling data-driven planning and timely resource allocation.
2. **Geo-Spatial Visualization**  
   Integrate GIS mapping (via Folium or GeoPandas) to provide interactive maps showing district-wise progress, connectivity gaps, and regional disparities in infrastructure development.
3. **Cluster Analysis**  
   Apply machine learning techniques (e.g., K-Means, DBSCAN) to group districts with similar development patterns—helping in formulating region-specific policies and interventions.
4. **Cross-Dataset Integration**  
   Enhance insight generation by merging PMGSY data with external datasets like rural population, terrain difficulty, and weather conditions to understand localized performance factors.
5. **Interactive Dashboards (Web-Based)**  
   Extend the Excel dashboard into real-time platforms such as Power BI, Tableau, or Streamlit, providing stakeholders dynamic views and interactive filtering options.
6. **Real-Time Data Integration**  
   Implement APIs to fetch live updates from government portals, ensuring that visualizations reflect the most recent status of projects, expenditures, and KPIs.
7. **Socio-Economic Impact Assessment**  
   Incorporate modules that analyze outcomes of road connectivity on education, healthcare access, and rural livelihoods—extending beyond physical completion into developmental impact.
8. **Delay Prediction Models**  
   Leverage predictive analytics (e.g., Random Forest, XGBoost) to anticipate project delays based on parameters like geography, fund disbursement, and previous completion trends.
9. **Mobile App Integration**  
   Develop a lightweight app interface for on-ground officials to upload real-time updates, track status, and view localized dashboards, improving grassroots-level engagement.
10. **Sustainability Monitoring**  
    Expand analysis to include long-term infrastructure quality indicators—maintenance frequency, durability, and resilience to climate factors—to assess sustainable outcomes.

**7.REFERENCES**

• **Primary Dataset Source:**  
PMGSY Infrastructure Data – [data.gov.in](https://data.gov.in)

• **Microsoft Excel Documentation:**  
Official Microsoft Excel Guide

• **Dashboard Design Principles:**  
Excel Dashboard Design Best Practices – Excel Campus

• **Visualization Techniques in Excel:**  
Data Visualization with Excel – Microsoft Learn

• **KPI and Metrics Setup:**  
Setting KPIs in Excel Dashboards – MyExcelOnline

• **Data Cleaning Techniques:**  
Data Cleaning in Excel – Spreadsheeto

• **Power Query for Data Transformation:**  
Power Query Documentation – Microsoft

• **Doughnut and Bar Chart Visualizations:**  
Create Doughnut Charts in Excel – Ablebits

• **Pivot Tables and Slicers:**  
Using PivotTables and Slicers – ExcelJet

• **Comparative Analysis & KPIs:**  
Data Analysis Techniques in Excel – WallStreetMojo

**8. GITHUB**

<https://github.com/Hemanth0555/Dashboard-INT-217-.git>

**9.LINKEDIN**

https://www.linkedin.com/posts/bandaru-hemanth-0591b5294\_exceldashboard-dataanalytics-salesanalysis-activity-7316864658378825729-cW1f?utm\_source=social\_share\_send&utm\_medium=member\_desktop\_web&rcm=ACoAAEc4aQUBT7kwFPaFRq2yqSJN9SvTh2s98HE