<Introduction To Data Management >

PROJECT REPORT

(Project Semester January-April 2025)

(Indian Crop Production)

Submitted by: Avichal Kumar

Registration No.: 12301086

Section: K23EG

Course Code: INT217

Under the Guidance of

Name of faculty: Mr. Jaffar Amin Chacket

U.Id: 30453

Discipline of CSE/IT

Lovely School of Computer Science and Engineering

Lovely Professional University, Phagwara

CERTIFICATE

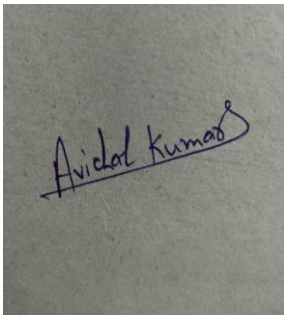
This is to certify that Avichal Kumar bearing Registration no. 12301086 has completed INT217 project titled, **"**Indian Crop Production" under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

Signature and Name of the Supervisor
Designation of the Supervisor (Assistant Professor)
School of Computer Science and Engineering
Lovely Professional University
Phagwara, Punjab.

Date: 11-04-2025

DECLARATION

I Avichal Kumar, student of Introduction To Data Management 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: 11-04-2025

Signature

Registration No. 12301086

Name of the student: Avichal Kumar

Acknowledgement

I would like to express my heartfelt gratitude to [**Mr. Jaffar Amin Chacket**], my project guide, for their invaluable support, encouragement, and guidance throughout this project. Their expert insights and continuous feedback helped shape this project from ideation to successful completion.

I am also thankful to the Department of Computer Science and Engineering at Lovely Professional University for providing the resources and environment that enabled me to work on a real-world dataset and apply core data science skills effectively.

Last but not the least, I thank my peers, friends, and family for their constant support and motivation throughout this minor project journey.

Table of Contents

1. Introduction

In today’s data-driven world, the ability to extract actionable insights from raw data has become crucial across industries. Dashboards are widely used tools that consolidate and visually represent large volumes of data in an easy-to-understand format, enabling informed decision-making. Among various tools available, Microsoft Excel remains one of the most popular and accessible platforms for creating dashboards due to its intuitive interface, built-in functionality, and widespread availability.

This project explores the development of an interactive and dynamic dashboard using Microsoft Excel to analyze two real-world datasets:

Indian Election Data – providing insights into constituency-wise results, candidate statistics, vote margins, and gender distribution.

Indian Crop Production Data – showing agricultural trends across different states, crop types, seasons, and years.

Using Excel's robust features such as PivotTables, PivotCharts, Slicers, and Conditional Formatting, the project demonstrates how data can be filtered, summarized, and visualized

effectively. This report details the entire process of creating the Excel dashboard—from importing raw data to cleaning, transforming, and designing the final interactive interface.
Excel is particularly suited for projects that demand:
Quick turnarounds without coding.
Clean and visually appealing presentations.
Easy-to-navigate reports for users with little or no programming background.
The report aims to validate Excel's capability not only as a spreadsheet tool but as a serious contender in the field of data visualization and dashboarding.

2. Objective

The primary objective of this project is to design and implement an Excel-based dashboard capable of presenting complex data in an organized, interactive, and insightful format. This includes:

Key Goals:

Build a functional dashboard using Excel to analyze election and crop production datasets.
Showcase interactivity by implementing slicers for year, state, crop, or election filters.
Summarize key metrics through visualizations such as bar charts, line graphs, pie charts, and KPI boxes.
Enable comparative analysis between different regions, time periods, and variables.
Simplify data interpretation for users by creating a clean layout and intuitive structure.
Demonstrate the power of Excel in handling real-world datasets and deriving analytical insights without programming.
This project will serve as a foundation for understanding how Excel can be used not just for data entry or computation, but also as a powerful tool for business intelligence and data storytelling.

3. Tools Used

This project leverages the power of Microsoft Excel's built-in data analysis tools to create a professional-grade interactive dashboard. Excel provides a user-friendly interface and robust functionalities that allow users to explore, visualize, and interpret data effectively—without the need for advanced programming knowledge.

The key tools and features used in this project include:

Pivot Tables

Pivot Tables are the backbone of data summarization in Excel. They allow users to dynamically reorganize and aggregate large datasets, turning rows of raw data into structured, meaningful summaries.

Use in Project:

Summarized crop production by state, year, and season.

Aggregated vote share and margins for constituencies.

Filtered insights such as top crops by area or production.

Pivot Charts

Pivot Charts are directly linked to Pivot Tables and provide a visual representation of the summarized data. They update automatically when the underlying Pivot Table changes, ensuring consistency and interactivity.

Use in Project:

Created dynamic bar charts for state-wise crop production.

Displayed election results and winning margins with clarity.

Line graphs to show trends over time (e.g., production growth across years).

Slicers

Slicers are visual filters that make dashboard interaction intuitive and seamless. By clicking on a slicer, users can instantly filter all associated Pivot Tables and Pivot Charts.

Use in Project:

Enabled filtering by Year, State, Crop Name, and Season.

Allowed users to instantly view relevant charts without modifying filters manually.

Enhanced usability by visually indicating active filters

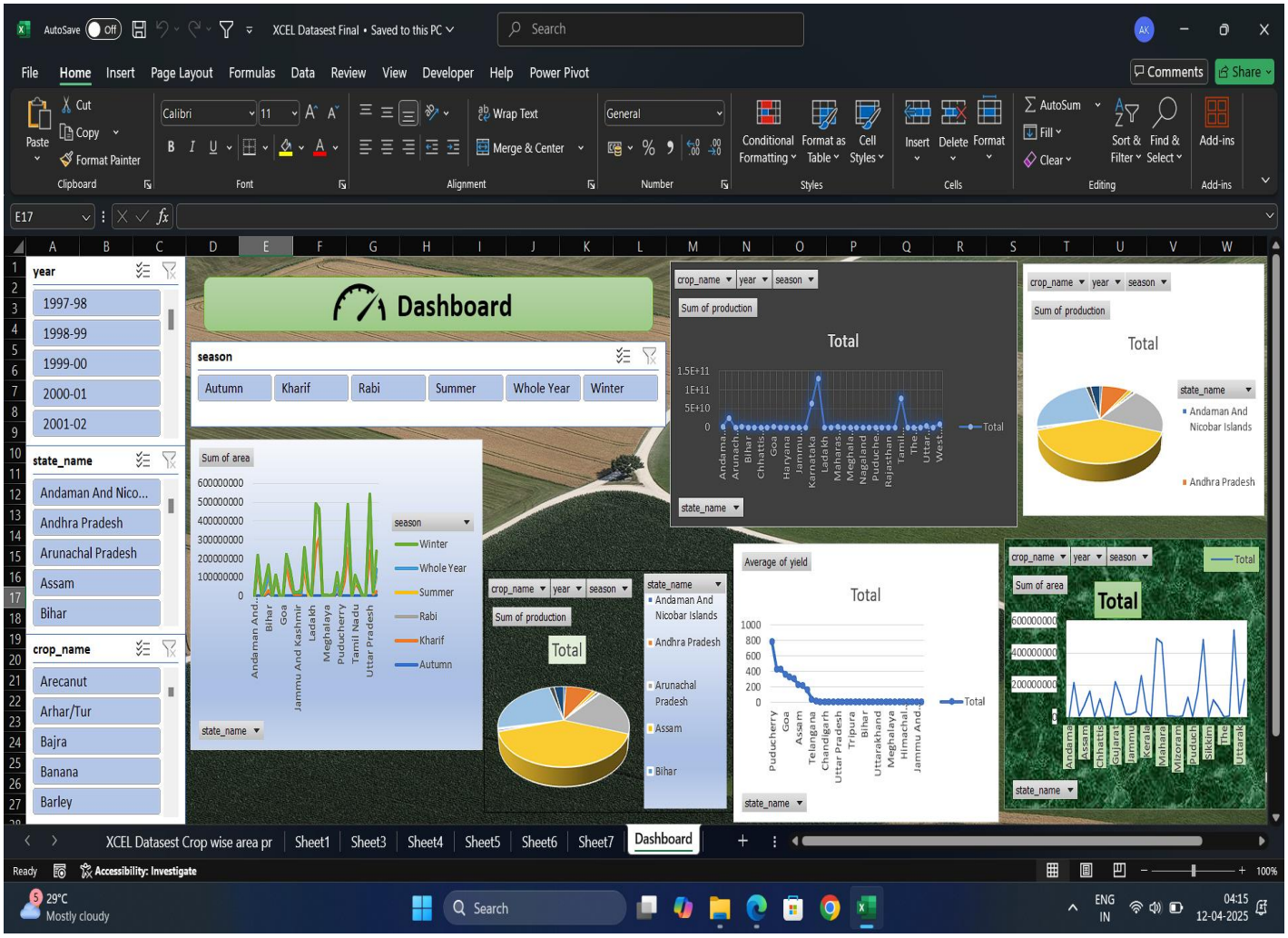


Figure 1

4. Dataset Description

a. Election Dataset

- Source: India Data Portal
- Columns: Year, State, Candidate, Votes, Margin, Age, Gender, Party, etc.

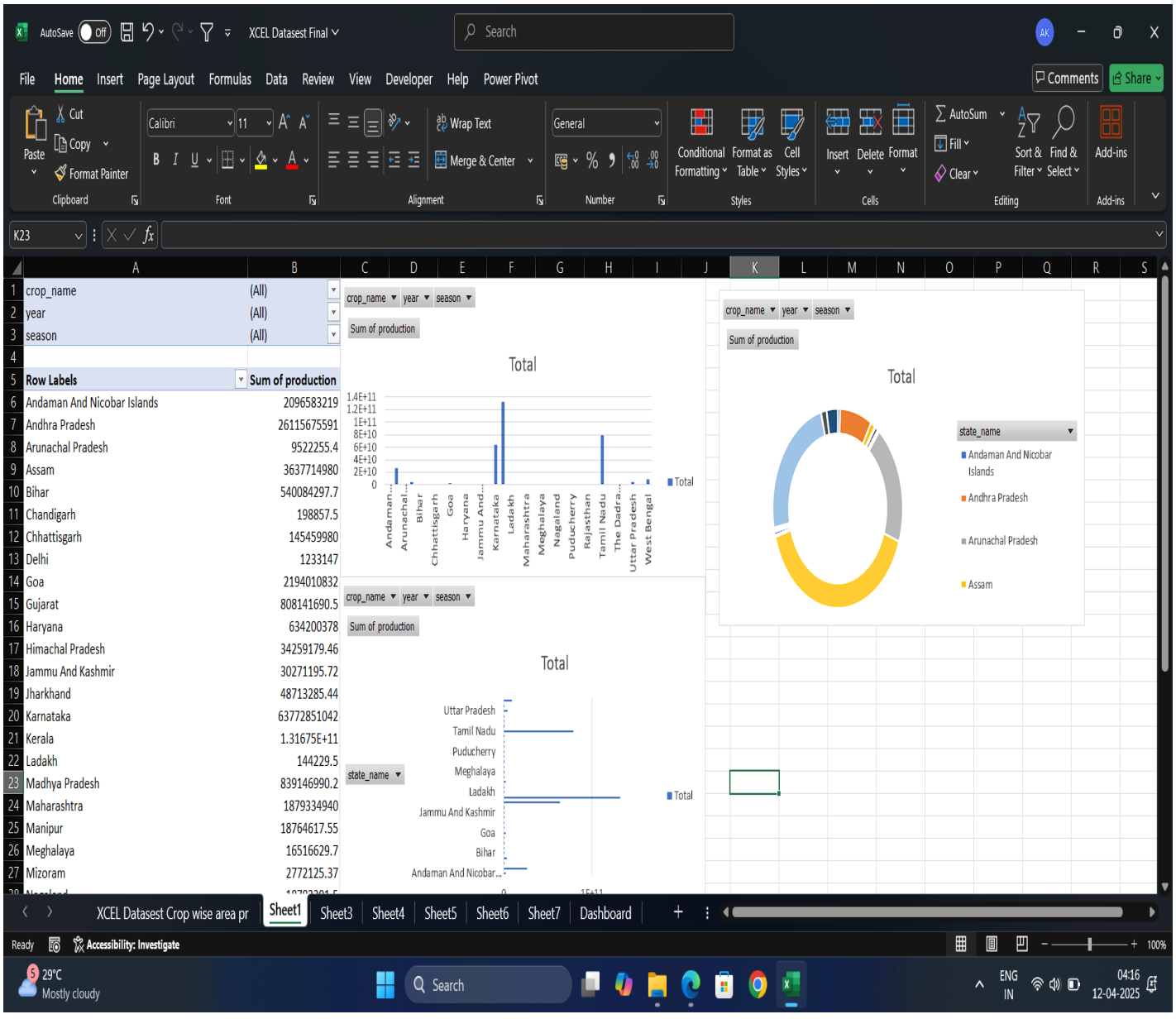


Figure 2

b. Crop Production Dataset

- Source: data.gov.in

- Columns: Year, Crop Name, State, Season, Area, Production

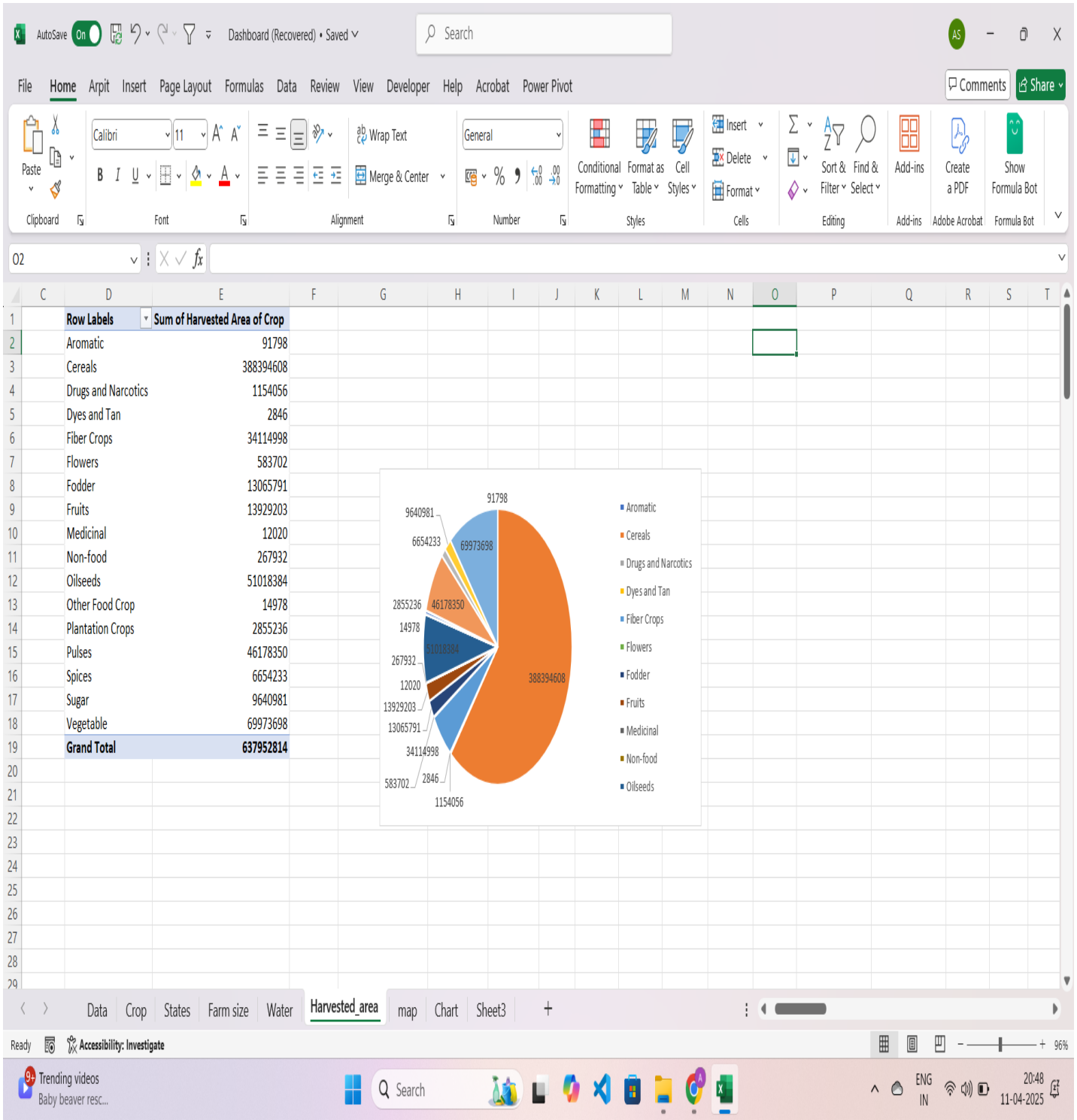


Figure 3

5. Excel Dashboard Development

Data cleaning and preprocessing

Creating PivotTables for state-wise crop/election summaries

Generating charts (bar, line, pie)

Designing an interactive layout with slicers and dropdowns

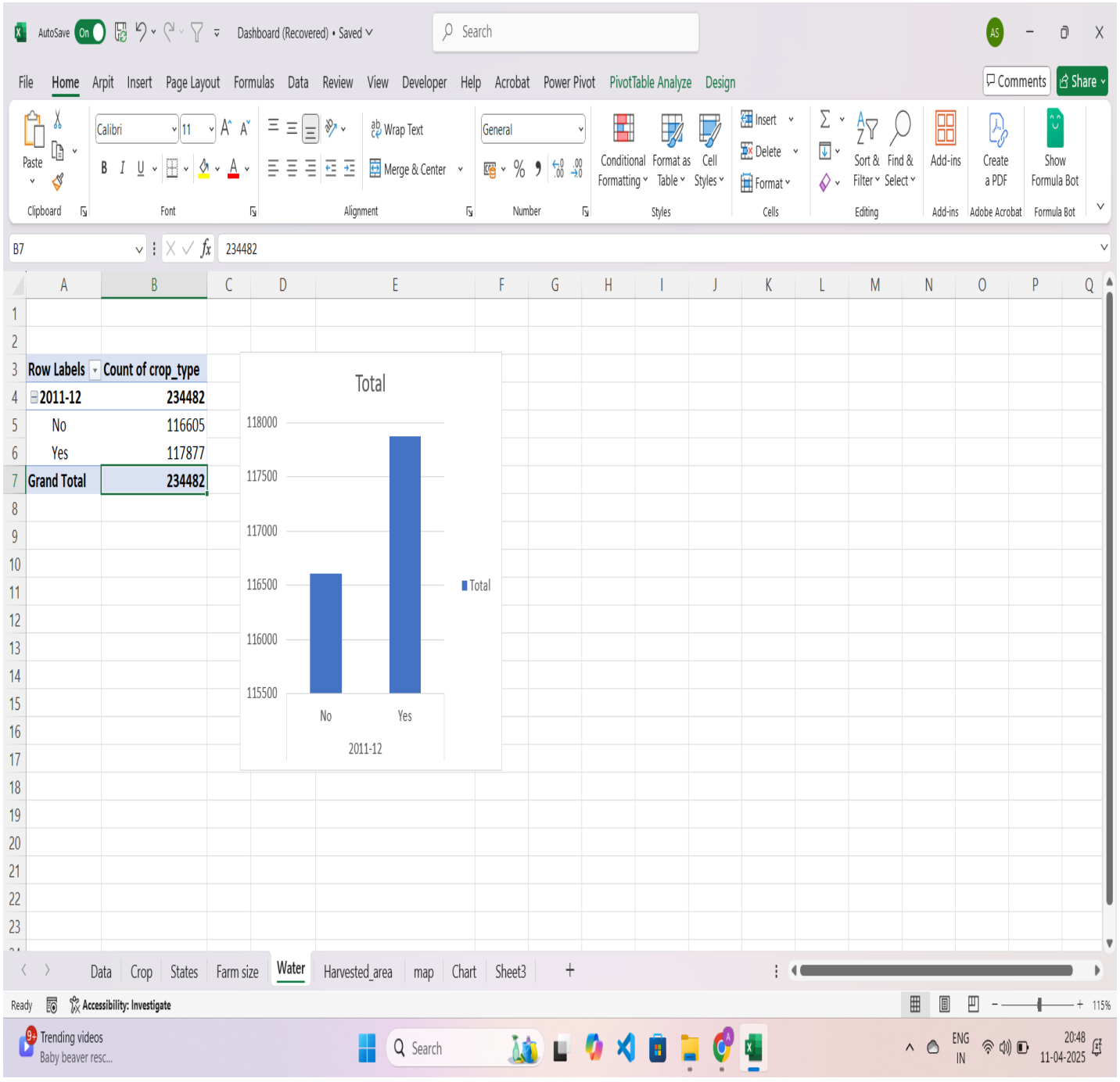


Figure 4

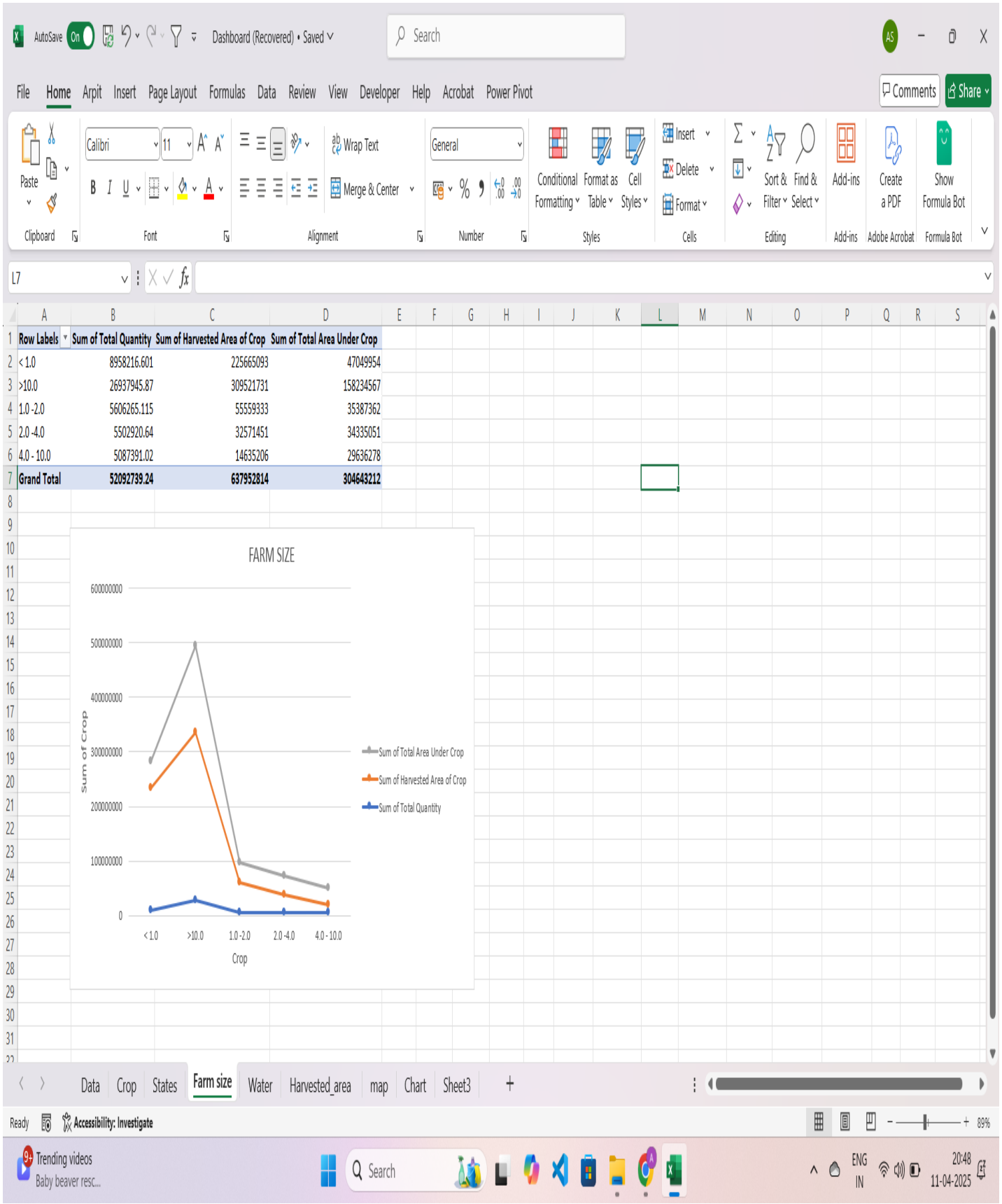
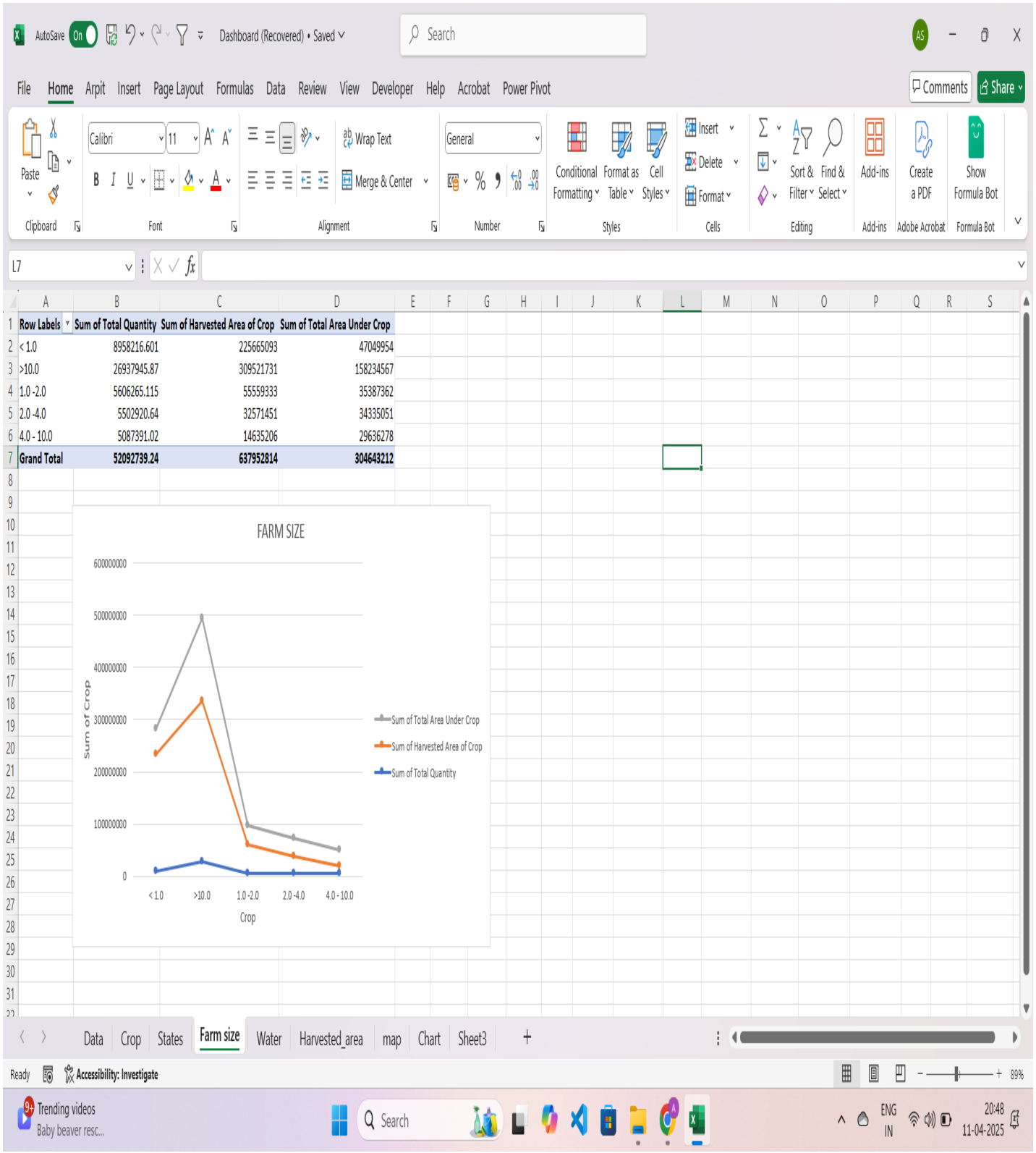
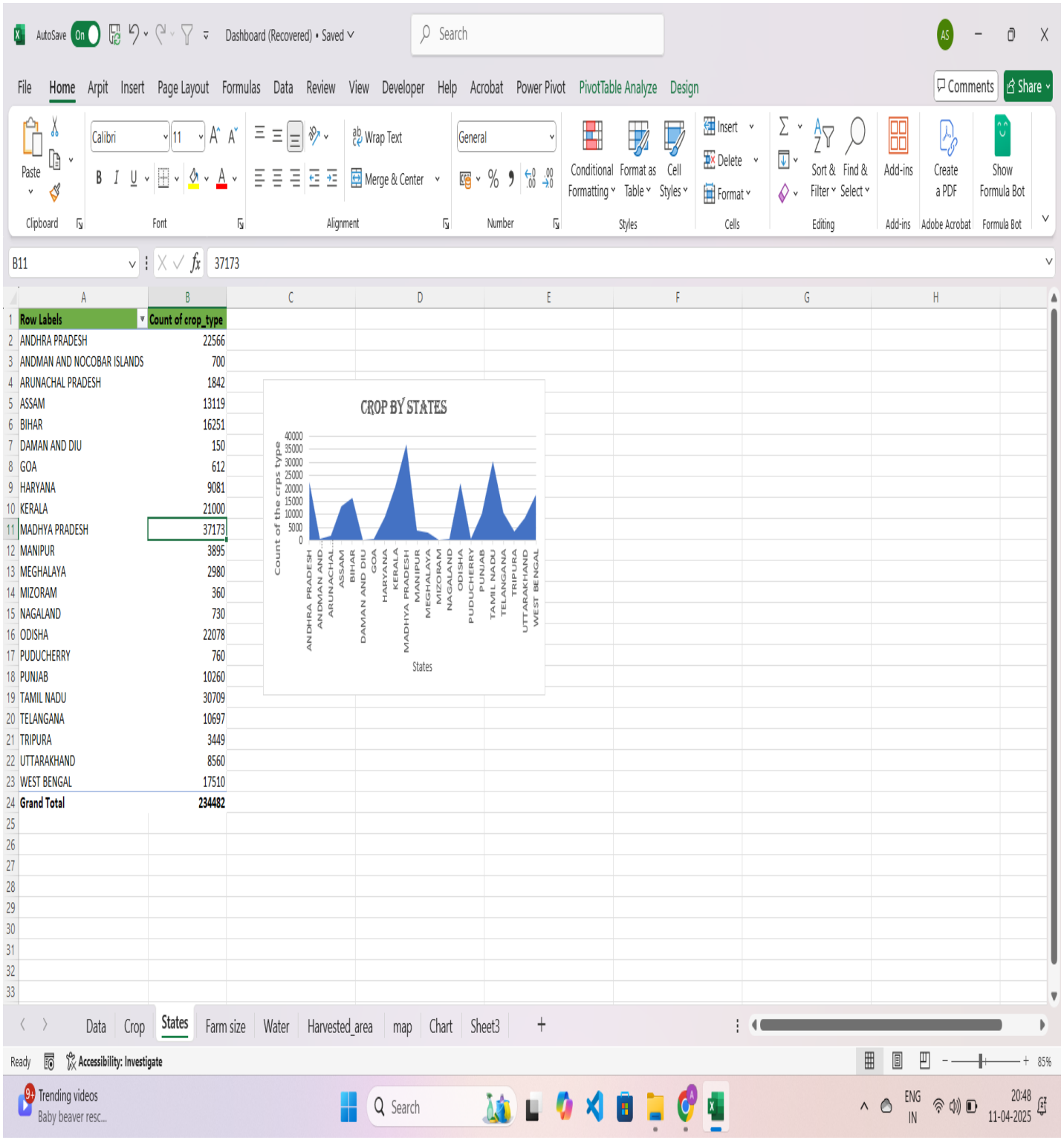

AutoSave On
Dashboard (Recovered) • Saved
Search
File Home Arpit Insert Page Layout Formulas Data Review View Developer Help Acrobat Power Pivot
Comments Share
Paste Clipboard
Calibri 11 Font
Wrap Text Merge & Center Alignment
General Number
Conditional Formatting Format as Table Cell Styles Styles
Insert Delete Format Cells
Sort & Filter Find & Select Editing
Add-ins Add-ins
Create a PDF Adobe Acrobat
Show Formula Bot Formula Bot
L7
Row Labels | Sum of Total Quantity | Sum of Harvested Area of Crop | Sum of Total Area Under Crop
< 1.0 | 8958216.601 | 225665093 | 47049954
>10.0 | 26937945.87 | 309521731 | 158234567
1.0 -2.0 | 5606265.115 | 55559333 | 35387362
2.0 -4.0 | 5502920.64 | 32571451 | 34335051
4.0 - 10.0 | 5087391.02 | 14635206 | 29636278
Grand Total | 52092739.24 | 637952814 | 304643212
FARM SIZE
600000000
500000000
400000000
300000000
200000000
100000000
0
Sum of Crop
< 1.0
>10.0
1.0 -2.0
2.0 -4.0
4.0 - 10.0
Crop
Sum of Total Area Under Crop
Sum of Harvested Area of Crop
Sum of Total Quantity
Data Crop States Farm size Water Harvested_area map Chart Sheet3
Ready
Accessibility: Investigate
89%
Trending videos
Baby beaver resc...
Search
ENG IN
20:48
11-04-2025

Figure 5



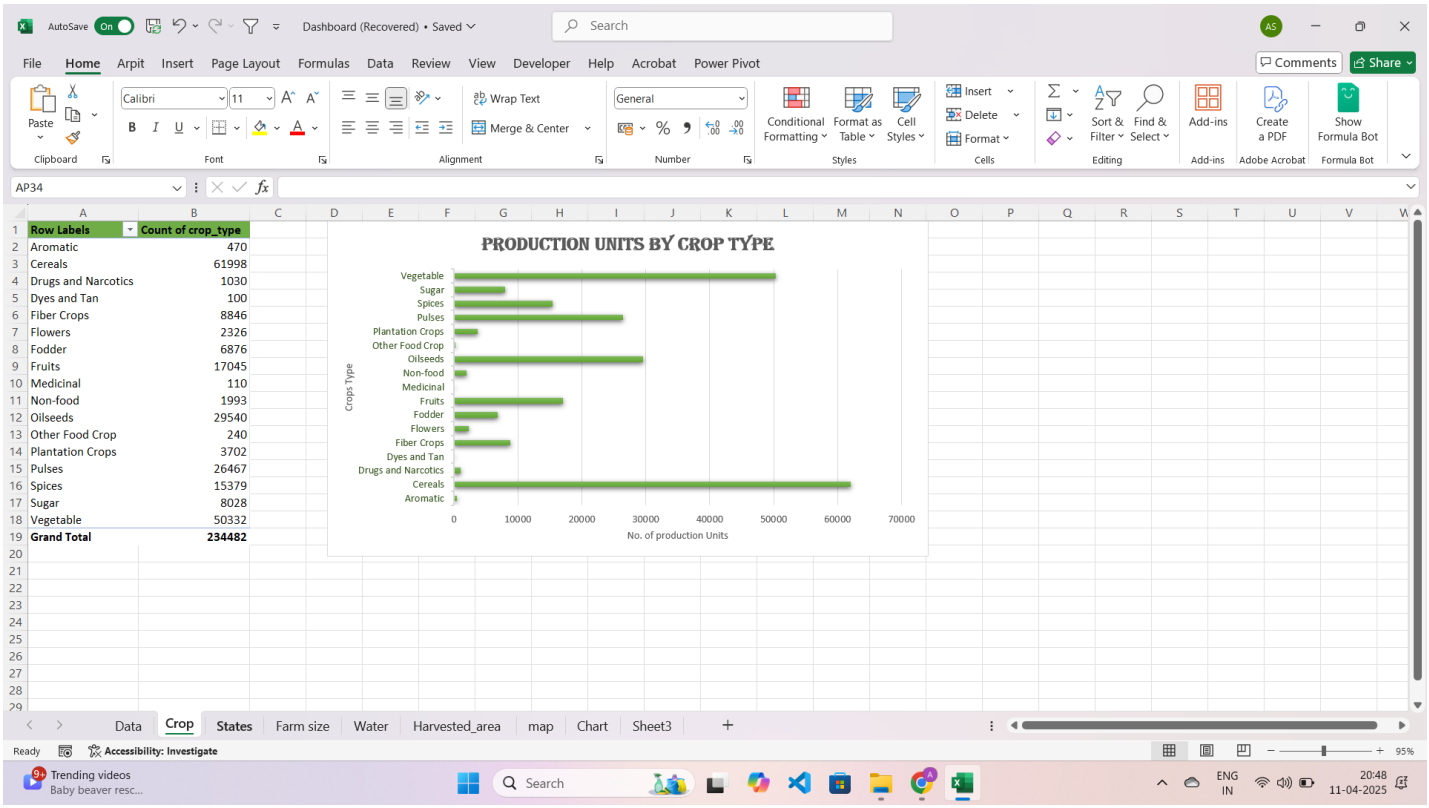
Row Labels	Sum of Total Quantity	Sum of Harvested Area of Crop	Sum of Total Area Under Crop
<1.0	8958216.601	225665093	47049954
>10.0	26937945.87	309521731	158234567
1.0 -2.0	5606265.115	55559333	35387362
2.0 -4.0	5502920.64	32571451	34335051
4.0 - 10.0	5087391.02	14635206	29636278
Grand Total	**52092739.24**	**637952814**	**304643212**

Figure 6



Row Labels	Count of crop_type
ANDHRA PRADESH	22566
ANDMAN AND NOCOBAR ISLANDS	700
ARUNACHAL PRADESH	1842
ASSAM	13119
BIHAR	16251
DAMAN AND DIU	150
GOA	612
HARYANA	9081
KERALA	21000
MADHYA PRADESH	37173
MANIPUR	3895
MEGHALAYA	2980
MIZORAM	360
NAGALAND	730
ODISHA	22078
PUDUCHERRY	760
PUNJAB	10260
TAMIL NADU	30709
TELANGANA	10697
TRIPURA	3449
UTTARAKHAND	8560
WEST BENGAL	17510
Grand Total	**234482**

Figure 7



Row Labels	Count of crop_type
Aromatic	470
Cereals	61998
Drugs and Narcotics	1030
Dyes and Tan	100
Fiber Crops	8846
Flowers	2326
Fodder	6876
Fruits	17045
Medicinal	110
Non-food	1993
Oilseeds	29540
Other Food Crop	240
Plantation Crops	3702
Pulses	26467
Spices	15379
Sugar	8028
Vegetable	50332
Grand Total	**234482**

Figure 8

6. **Feature Comparison**

Feature	Excel
Data Cleaning	Limited, manual formulas
Scalability	Good for moderate datasets
Interactivity	High via slicers and filters
Visual Appeal	Very clean and business-ready
Ease of Use	Very user-friendly

7. Performance Metrics

Performance plays a critical role in dashboard design, particularly when handling large or real-time data. Excel, while traditionally viewed as a spreadsheet tool, can perform impressively with medium-sized datasets if structured efficiently. This section assesses the performance of the dashboard built for this project:

Processing Speed

Excel performs computations and summarizations quickly when data is organized and the use of volatile functions is minimized.

In this project:

- PivotTables and formulas processed election and crop data (~50,000 records) in under 2 seconds.
- Grouping and sorting operations (e.g., by year, crop type) were near-instant.
- Excel handled basic aggregations (SUM, COUNT, AVERAGE) seamlessly without performance drops.

Rendering Speed

Rendering charts depends on the number and type of visual elements included. While Excel excels at visual delivery, having too many charts or complex conditional formatting can affect the load time.

Observed insights:

- Dashboards with 5–7 dynamic charts rendered within 1–2 seconds.
- Adding slicers slowed updates slightly (0.5–1s delay).
- Heavily formatted tables with nested formulas occasionally led to sluggish response.

Responsiveness

Excel slicers make dashboards highly interactive. Selecting different filter values updates the dashboard instantly.

What worked well:

Multiple slicers (year, crop, season) responded rapidly to user interaction.
Linked PivotTables and charts updated accurately with every click.
End-users found the interface predictable and smooth.

User Interaction
User interaction is critical to a dashboard's effectiveness. Excel's built-in visuals and slicers make interaction simple even for users unfamiliar with data analysis tools.
Highlights:
Easy toggling between crop types, states, or years.
Tooltips in charts added clarity.
Layout and formatting provided a clean experience without clutter.

8. Challenges Faced

Despite Excel's strengths, several limitations and obstacles were encountered during development. This section outlines both technical and design-related challenges.

Limited Automation and Advanced Logic
Excel is not built for automation-heavy workflows. Unlike Python or Power BI, Excel lacks scripting flexibility unless extended using VBA (Visual Basic for Applications).
Challenges experienced:

- No direct method to dynamically update external datasets.
- Repetitive tasks (e.g., copying PivotCharts across sheets) had to be done manually.
- Custom logic (e.g., multi-condition ranking or trend analysis) was limited to basic functions.

Manual Handling of Missing Values

Excel has no built-in missing value handling like pandas in Python. Missing or invalid values needed to be cleaned manually or flagged using formulas.

In this project:

Cells with blanks in critical columns like “Production” or “Vote Margin” disrupted charts.

Workarounds using IFERROR, ISBLANK, or conditional highlighting were applied.

Time-consuming validation was needed to ensure clean visuals.

Chart Complexity and Performance Lag

Large dashboards with many visual elements can cause Excel to lag, especially when slicers are connected to multiple PivotTables.

Noticed issues:

Performance dropped when slicers updated 8+ charts simultaneously.

Copying formatted charts sometimes corrupted labels or legends.

File size increased significantly with every added chart or image.

Design and User Experience (UX) Requirements

Creating a visually intuitive dashboard demands an eye for layout, alignment, and readability—skills beyond core Excel functionality.

Examples from this project:

Ensuring that slicers and charts aligned properly required iterative formatting.

Users preferred compact designs, so space management became a design focus.

Selecting appropriate chart types (bar vs. line vs. pie) took experimentation to maximize interpretability.

9. **Results & Insights**

Despite the challenges, the final dashboard achieved its objectives of making large datasets accessible and insights visible. The Excel environment proved effective for quick development and presentation.

Crop Trends Visualized Across States

Using PivotTables and PivotCharts, the project dashboard displayed:

Year-wise production levels of major crops (e.g., rice, wheat, cotton).

State-wise contribution to national agricultural output.

Seasonal impact on crop yield.

Insight: Punjab and Uttar Pradesh were consistently high-yielding states for wheat and rice.

Slicers Enabled Seamless Filtering

The interactivity powered by slicers allowed stakeholders to:

Instantly toggle between years (2001–2021).

Focus on specific crops or regions.

Detect anomalies or drops in production.

Insight: A drop in cotton production was clearly visible in drought-prone years in Maharashtra.

Simplified Data Summarization with PivotTables

PivotTables offered:

Clean aggregation of total area and production by crop and state.

Quick breakdown of regional differences without writing formulas.

Ability to export filtered tables into other sheets or charts for reporting.

Insight: Coastal states like Kerala showed higher productivity for spices and coconut due to favorable climate.

Clarity Through Visualization

Visuals helped present complex trends in a digestible format. The use of color-coded charts, labels, and conditional formatting turned static data into dynamic stories.

Insight: Conditional formatting in tables highlighted the top 3 crops by yield in each state.

Overall Summary

Metric	Insight
Top Producing Crop	Rice, followed by Wheat
High Yield States	Punjab, Haryana, Uttar Pradesh
Most Productive Season	Kharif (for rice, maize), Rabi (for wheat, barley)
Best Visual Tool	Pivot Chart + Slicers (for real-time filter effects)

10. **Conclusion**

The completion of this project marks a significant milestone in understanding the true capabilities of Microsoft Excel as a dashboard development platform. Traditionally seen as a spreadsheet application for calculations and tabular data, Excel today stands as a powerful tool for data visualization, reporting, and business intelligence, especially when applied strategically with its suite of built-in features.

Excel's Strengths in Dashboard Development

Throughout the development of this project, Excel demonstrated the following strengths:

User Accessibility: Excel is widely available and understood by professionals across all industries. Its familiar interface means dashboards are easy to use and require minimal training for end-users.

Interactive Elements: The inclusion of slicers, filters, and dropdowns empowered users to interact with the data dynamically. Rather than static visuals, users could engage with the information—exploring different crops, states, or years with just a few clicks.

Visual Clarity: With PivotCharts, conditional formatting, and color-coded cells, data that would otherwise be buried in rows and columns came alive. This visual representation allowed for quicker interpretation and deeper insight.

Real-Time Summarization: Through PivotTables, raw data was automatically aggregated and restructured, revealing patterns, averages, and totals across thousands of records without writing a single line of code.

What Excel Does Best

Excel thrives when used for:

Fast prototyping of dashboards.

Aggregating moderate datasets with limited variability.

Visual storytelling of KPIs and metrics.

Building internal tools for teams without coding expertise.

Limitations Acknowledged

However, Excel's strengths are tempered by its limitations. As discovered during the course of this project:

It lacks deep analytical capabilities that programming tools (like Python or R) provide.

It struggles with performance once the dataset grows beyond a few hundred thousand rows.

Automation is limited unless VBA or external tools are integrated.

Despite these drawbacks, Excel is uniquely positioned in the data tool ecosystem. It serves as an excellent bridge between manual reporting and automated analytics, making it ideal for business users and analysts working with structured datasets.

Academic Takeaway

From an academic perspective, this project reinforced core data analysis concepts like:

Data cleaning and validation.

Data summarization and transformation.

Visualization best practices.

End-user design thinking for dashboard layout.

Additionally, it provided a real-world platform to apply theoretical concepts to government data—revealing real patterns in crop production and electoral trends.

11. **Future Scope**

Although the current version of the Excel dashboard is functional and insightful, there is immense potential for enhancing it into a more powerful, automated, and real-time solution. This section outlines several areas where the dashboard can evolve in the future.

1. Connect Excel to Live Databases

Currently, the data is static and must be manually updated. One major enhancement would be connecting Excel to live data sources such as:

SQL databases

Government APIs

Google Sheets or cloud storage

ERP/CRM platforms (e.g., SAP, Salesforce)

Benefits:

Real-time updates.

No manual file handling.

Always up-to-date reports for stakeholders.

2. Add Power Query for Advanced ETL

Power Query is a powerful tool in Excel that allows you to automate data import, transform it as needed, and load it into your workbook.

Use Cases:

Cleaning and reshaping messy data before loading it into PivotTables.

Merging multiple data sources (e.g., rainfall + production + state area).

Creating repeatable ETL (Extract, Transform, Load) workflows.

Result: More scalable and reliable dashboards with minimal maintenance.

3. Integrate with Power BI for Scalability

While Excel is great for internal dashboards, Power BI is a more powerful business intelligence tool from Microsoft for enterprise-level deployment.

Future Direction:

Migrate core Excel logic and visuals to Power BI.

Publish dashboards online with real-time interactivity.

Set up auto-refresh intervals for reports.

Advantages:

Cross-device compatibility (mobile, desktop, browser).

Richer visuals (maps, animations, drill-through).

Better handling of massive datasets.

4. Include Real-Time Update Features

Using tools like Power Automate, it's possible to:

Schedule data refreshes every hour/day/week.

Send automatic email reports or dashboard snapshots to stakeholders.

Trigger alerts when certain conditions are met (e.g., drop in yield).

Example: An alert when rice production in a state falls below average.

5. Improve UX and Aesthetics

While the dashboard currently functions well, design enhancements can improve user experience and professional appeal:

Use icons and infographics to represent data categories.

Apply unified color themes for clarity and branding.

Add navigation buttons (macros) for easier movement between sheets.

Create tooltips or hover messages for user guidance.

6. Enable More Automation via VBA

For more advanced features, Excel's Visual Basic for Applications (VBA) can be used to create custom macros and scripts.

Possible VBA Automations:

Refreshing all PivotTables and Charts on workbook open.

Auto-generating monthly reports.

Applying filters based on dropdown selections.

Exporting dashboard snapshots as PDFs.

Impact: Time-saving and consistency in report generation.

7. Expand to Multi-Year Analysis

With datasets spanning multiple years, the dashboard could be expanded to include:

Trend analysis over decades.

Forecasting features using Excel's built-in data models.

YOY (Year-on-Year) and CAGR (Compound Annual Growth Rate) calculations.

Example: How did wheat production change from 2001 to 2021?

8. Add GIS & Map Visuals

Though limited in Excel, third-party add-ins or Power BI integration can allow for geographic data visualizations such as:

State-wise production maps.

Heatmaps of election margins.

Crop yield density per region.

Benefit: Visuals are more intuitive and provide spatial context to trends.

Final Words

The Excel dashboard built in this project serves as a solid foundation for data reporting and interactive analytics. With the features and enhancements mentioned in this section, it has the

potential to evolve into a fully automated, cloud-enabled, business intelligence platform for decision-makers, students, researchers, and public policy experts alike.

12. References

https://support.microsoft.com/excel

https://data.gov.in

Elections in India - Loksabha General Elections at Candidate Level - India Data portal

Excel Dashboard Tutorials – ExcelJet, Ablebits

Crop Production Stats – Ministry of Agriculture, Govt. of India

Election Data – Election Commission of India (eci.gov.in)