St. Francis Institute of Technology, Mumbai-400 103

**Department Of Information Technology**

A.Y. 2024-2025

Class: TE-ITA/B, Semester: VI

Subject: **Business Intelligence Lab**

**Experiment – 10: Study and application of open source BI tool (Qlikview, Tableau, Pentaho, Rapid Miner)**

1. **Aim:** Study and application of open source BI tool (Qlikview, Tableau, Pentaho, Rapid Miner)
2. **Objectives:** After study of this experiment, the students will be able to know different BI Tools
3. **Outcomes:**

**CO6:**Apply BI to solve practical problems : Analyze the problem domain, use the data collected in enterprise apply the appropriate data mining technique, interpret and visualize the results and provide decision support

**Prerequisite:** Introduction to all Open Source BI tools.

**Requirements:** Personal Computer, Windows XP /Windows 7/8 operating system, Internet Connection, Microsoft Word.

1. **Theory:**

*Q.1 What is BI?*

Business Intelligence (BI) refers to the strategies and technologies used by enterprises for data analysis and management of business information.It provides historical, current, and predictive views of business operations, enabling organizations to make data-driven decisions. In essence, it transforms raw data into meaningful and actionable insights.

*Q.2 Need of BI*

**Improved Decision-Making:** BI provides accurate and timely information, enabling managers to make informed decisions.

**Enhanced Operational Efficiency:** BI helps identify bottlenecks, optimize processes, and improve overall efficiency.

**Competitive Advantage:** By analyzing market trends and customer behavior, businesses can gain a competitive edge.

**Revenue Growth:** BI helps identify new sales opportunities, improve customer retention, and increase profitability.

**Risk Management:** BI can help identify and mitigate potential risks by analyzing historical data and trends.

**Customer Insights:** Understand customer behavior, preferences, and needs.

*Q.3 Applications of BI*

**Sales Analysis:** Tracking sales performance, identifying top-performing products, and forecasting future sales.

**Marketing Analysis:** Measuring the effectiveness of marketing campaigns, analyzing customer segmentation, and identifying target audiences.

**Financial Analysis:** Monitoring financial performance, identifying trends, and forecasting future financial outcomes.

**Supply Chain Management:** Optimizing inventory levels, improving logistics, and reducing costs.

**Human Resources:** Analyzing employee performance, identifying training needs, and improving employee retention.

*Q.4 List Tools of BI*

Tableau

QlikView

Power BI

Pentaho

RapidMiner

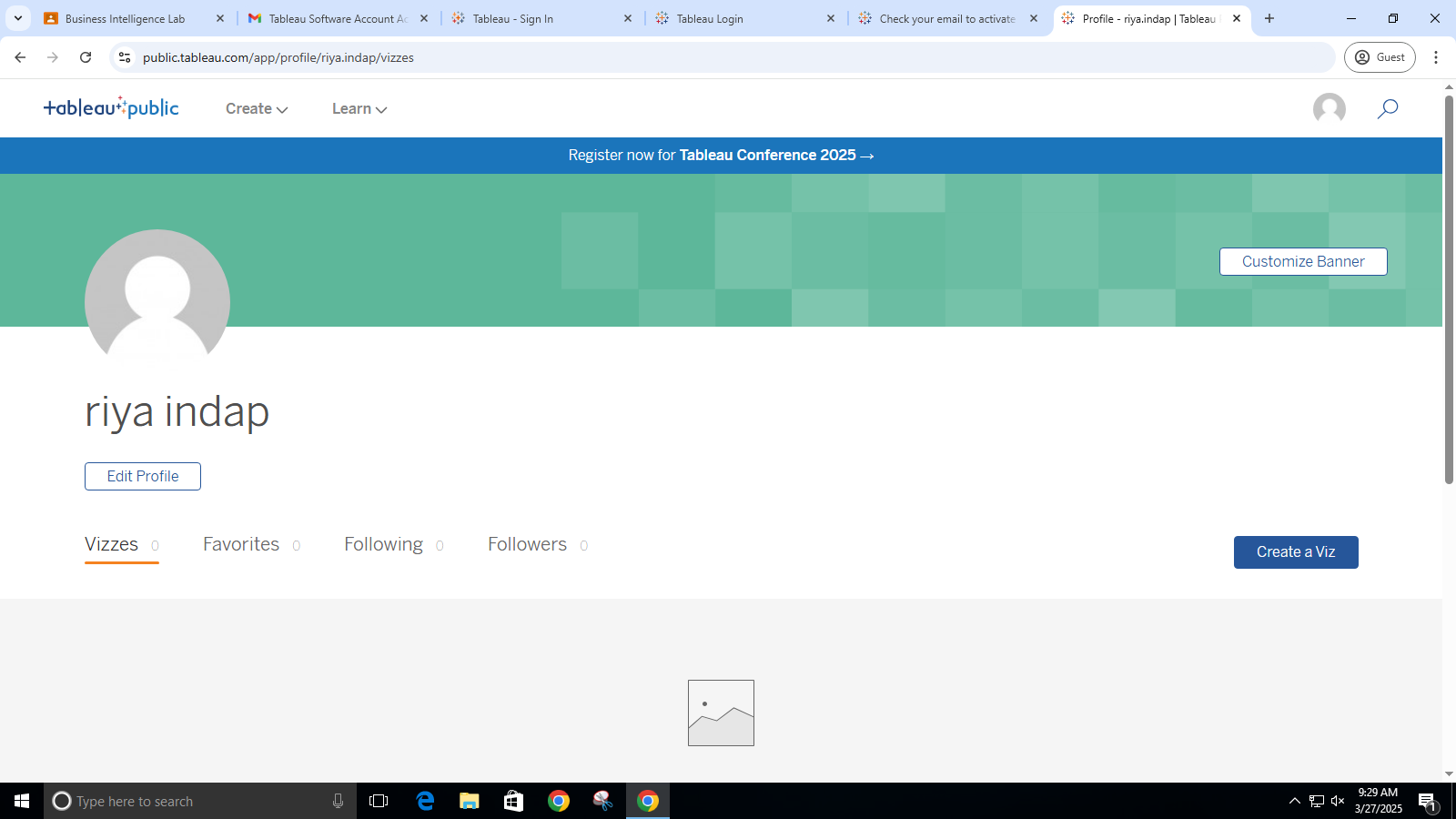
*Q.5 Elaborate on Pentaho, Rapid Miner, Qlikview, Tableau*

| **Feature/Aspect** | **Pentaho (Hitachi Vantara Pentaho Platform)** | **RapidMiner** | **QlikView** | **Tableau** |
| --- | --- | --- | --- | --- |
| **Features** | Data Integration (ETL), OLAP, Reporting, Dashboards, Data Mining, Big Data Integration, Open Source Core | Data Mining, Machine Learning, Predictive Analytics, Text Mining, Visual Workflow Design, Extensive Algorithm Library | Associative Data Model, In-Memory Processing, Interactive Dashboards, Self-Service BI, Data Visualization | Interactive Dashboards, Data Visualization, Drag-and-Drop Interface, Real-Time Analytics, Data Blending |
| **Data Set Required** | Relational Databases, Flat Files, Hadoop, NoSQL | CSV, Excel, Databases, Big Data Platforms | Databases, Spreadsheets, Web Services | Databases, Spreadsheets, Cloud Services, Big Data Platforms |
| **Working** | Platform for data integration and analytics; ETL, reporting, dashboards, advanced analytics | Visual environment for designing data mining workflows; drag-and-drop operators | Associative data model; in-memory data processing; interactive exploration | Connects to data sources; creates interactive dashboards and visualizations; visual data exploration |
| **Advantages** | Comprehensive suite, strong data integration, open-source flexibility, big data support | Powerful machine learning, user-friendly visual interface, extensive algorithm library, open-source and commercial versions | Fast performance, flexible data exploration, intuitive UI, associative engine | Powerful data visualization, user-friendly interface, fast performance, large community and support |
| **Limitations** | Steep learning curve, less intuitive UI compared to some tools | Performance issues with very large datasets, some advanced features require paid license | Can be expensive, scripting required for complex transformations | Can be expensive, data preparation can be challenging for complex datasets |
| **Applications** | Data warehousing, reporting, data mining, big data analysis | Predictive modeling, customer churn analysis, fraud detection, risk assessment | Sales analysis, financial reporting, supply chain analysis, customer analysis | Business reporting, sales analysis, marketing analysis, financial analysis, data discovery |

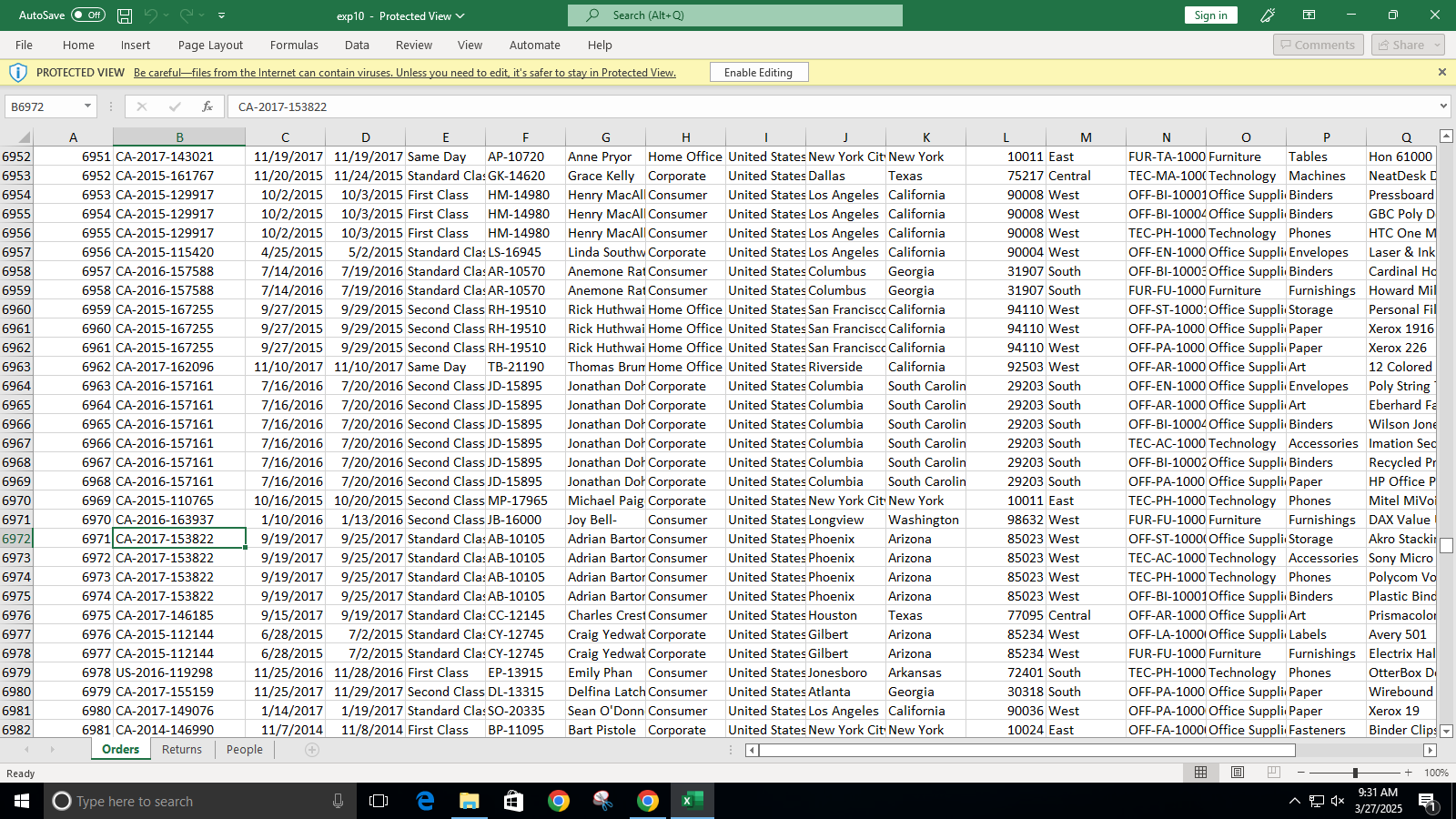
1. **Laboratory Exercise:** Attach screenshots for tableau and Qlikview using sample dataset
2. **Post-Experiments Exercise**
3. **Questions:**
   * Compare and Contrast between Qlikview and Tableau
4. **Conclusion:**
   * Summary of Experiment
   * Importance of Experiment
   * Application of Experiment

1. **Reference:** Business Intelligence: Data Mining and Optimization for Decision Making by Carlo Vercellis, Wiley India Publications

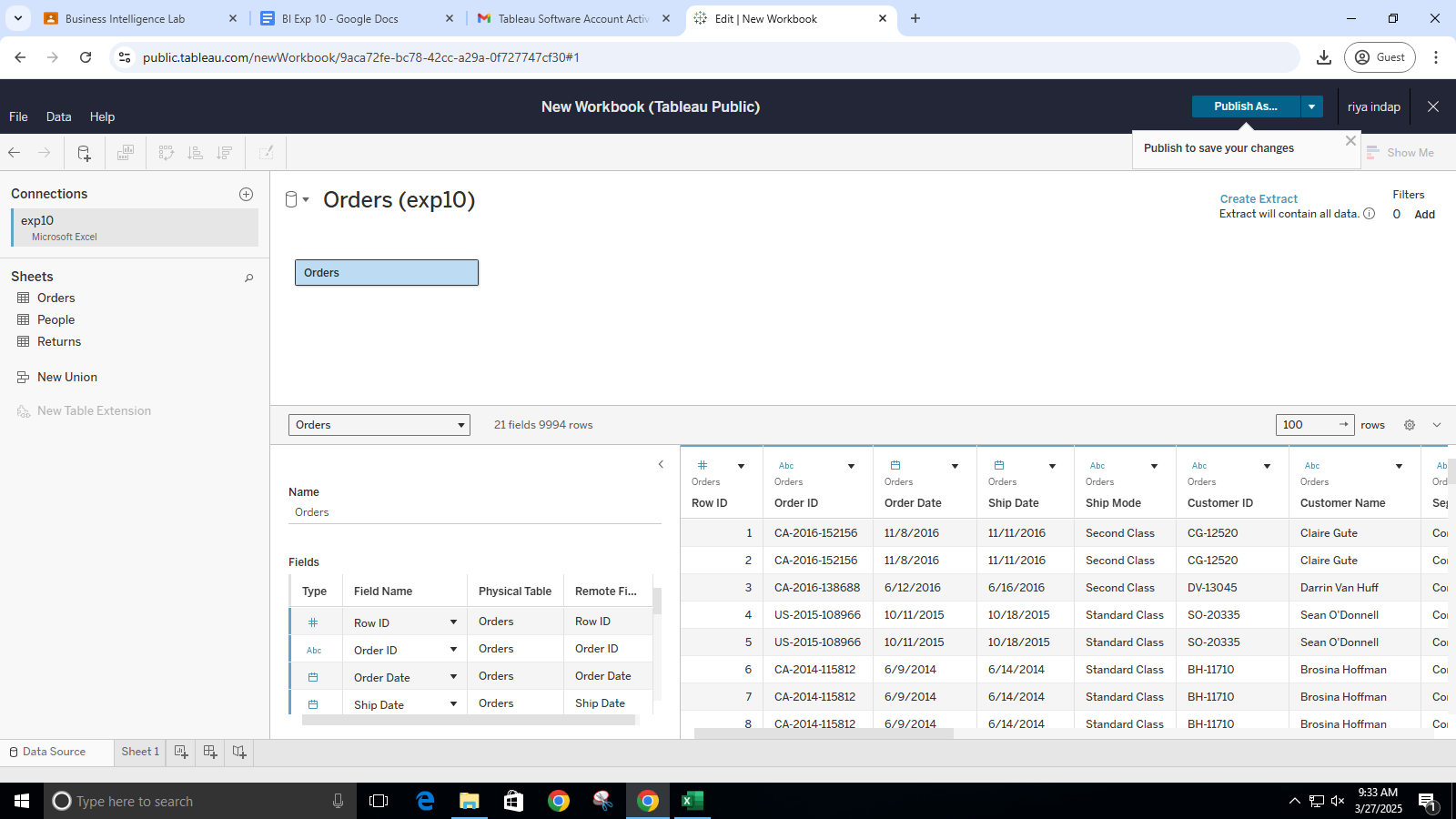
Created a Profile on Tableau



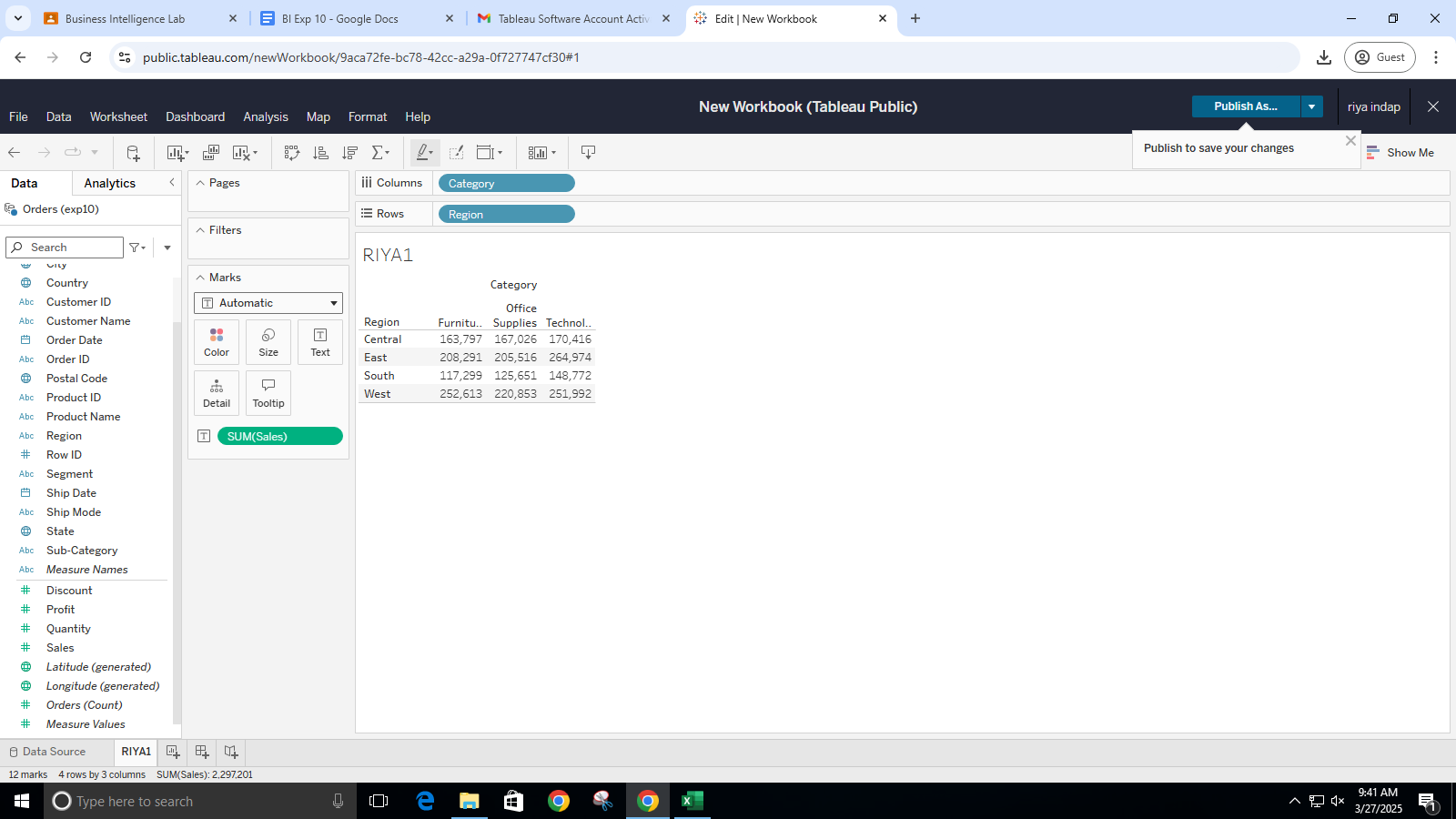
Downloading and loading dataset from tableau



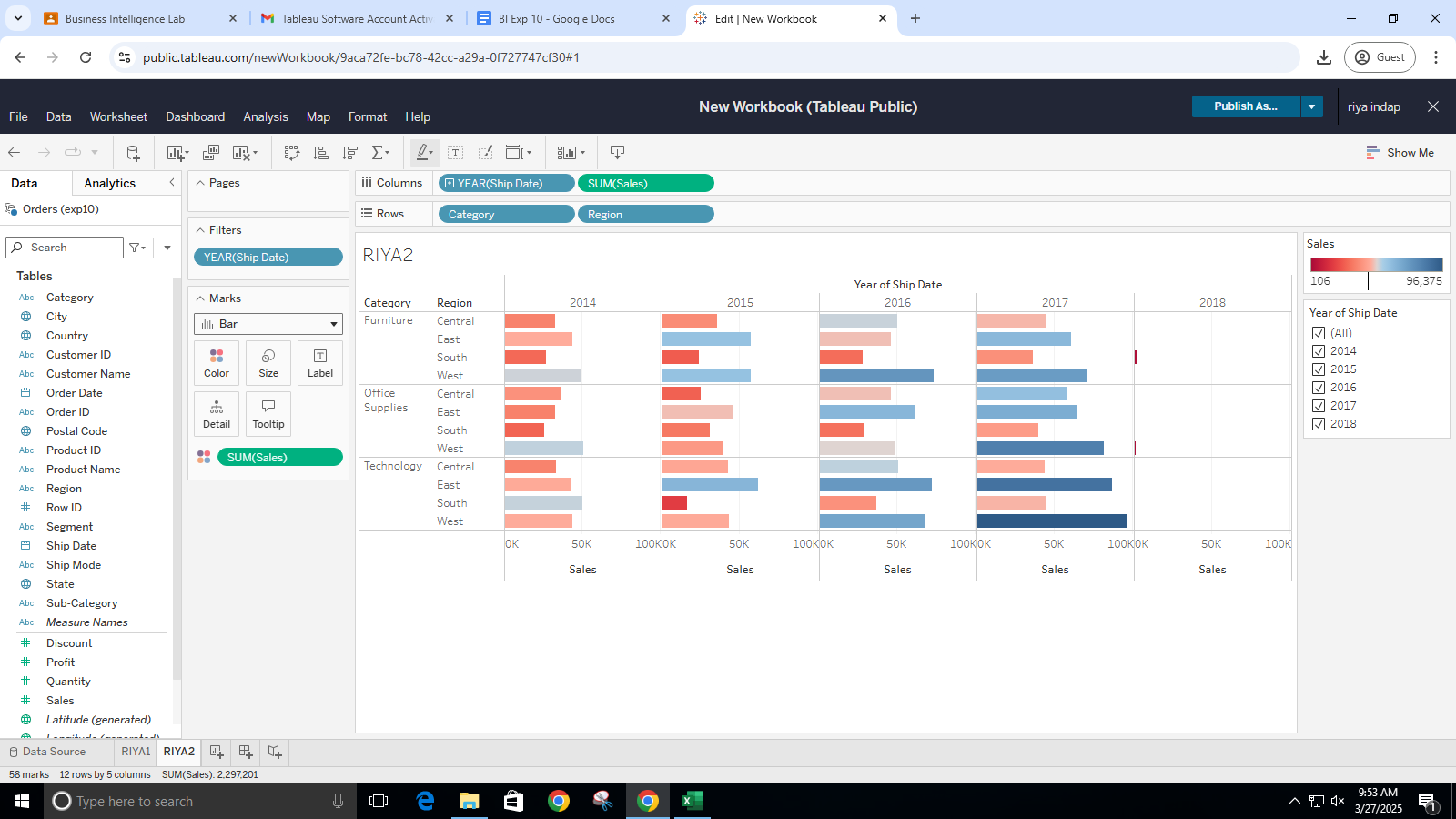
Loading the order sheet in Tableau



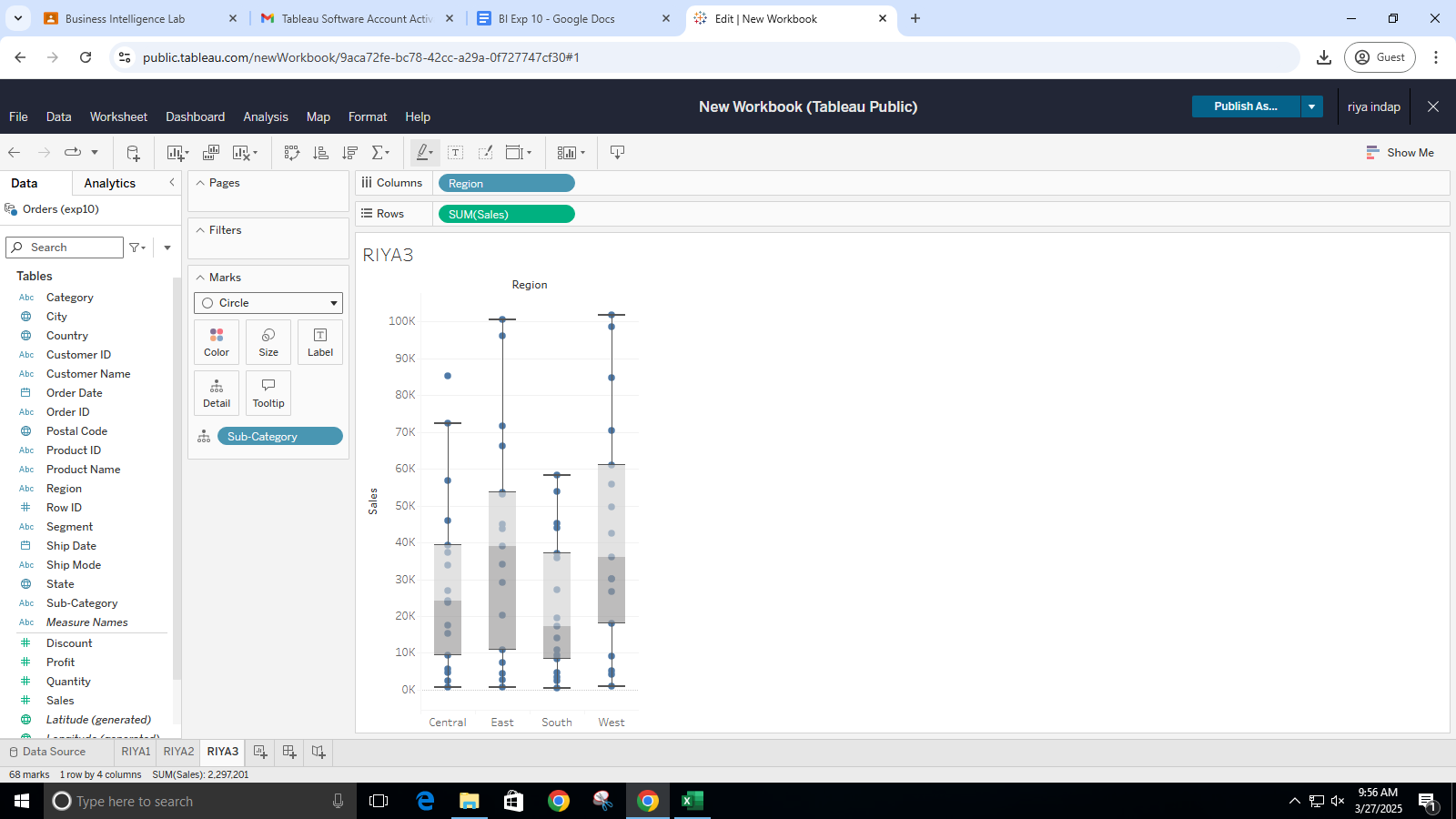
Creating a ‘RIYA1’ sheet having ‘category’ as columns and ‘regions’ as rows and applied ‘sales’ as measures which performed ‘sum’ operation on the sheet.



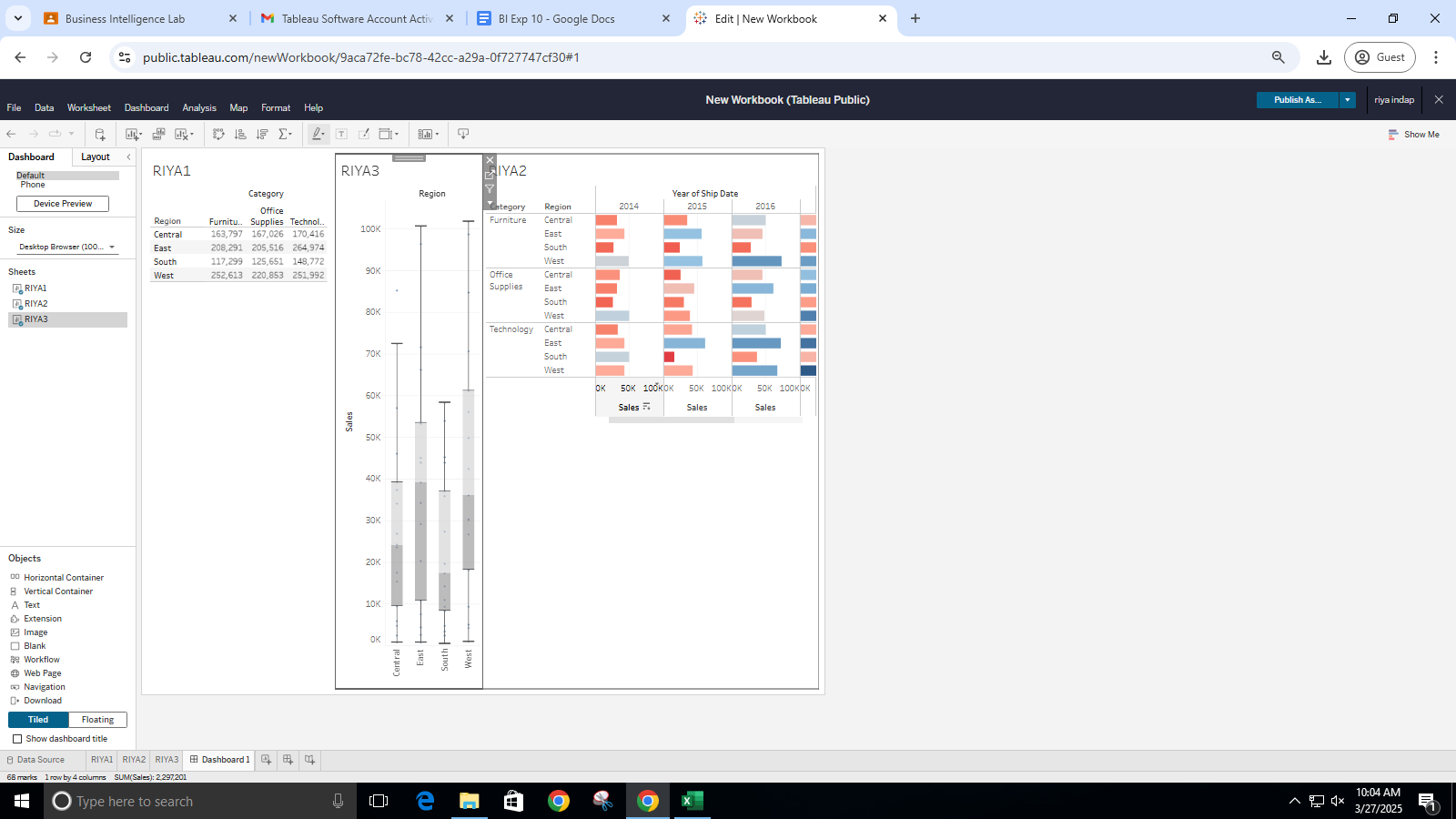
Creating a ‘RIYA2’ sheet having ‘YEAR(ship date)’ and ‘Sales’ measure as columns and ‘category’ and ‘regions’ as rows which gave us a visualization of the horizontal bar graph of the data.



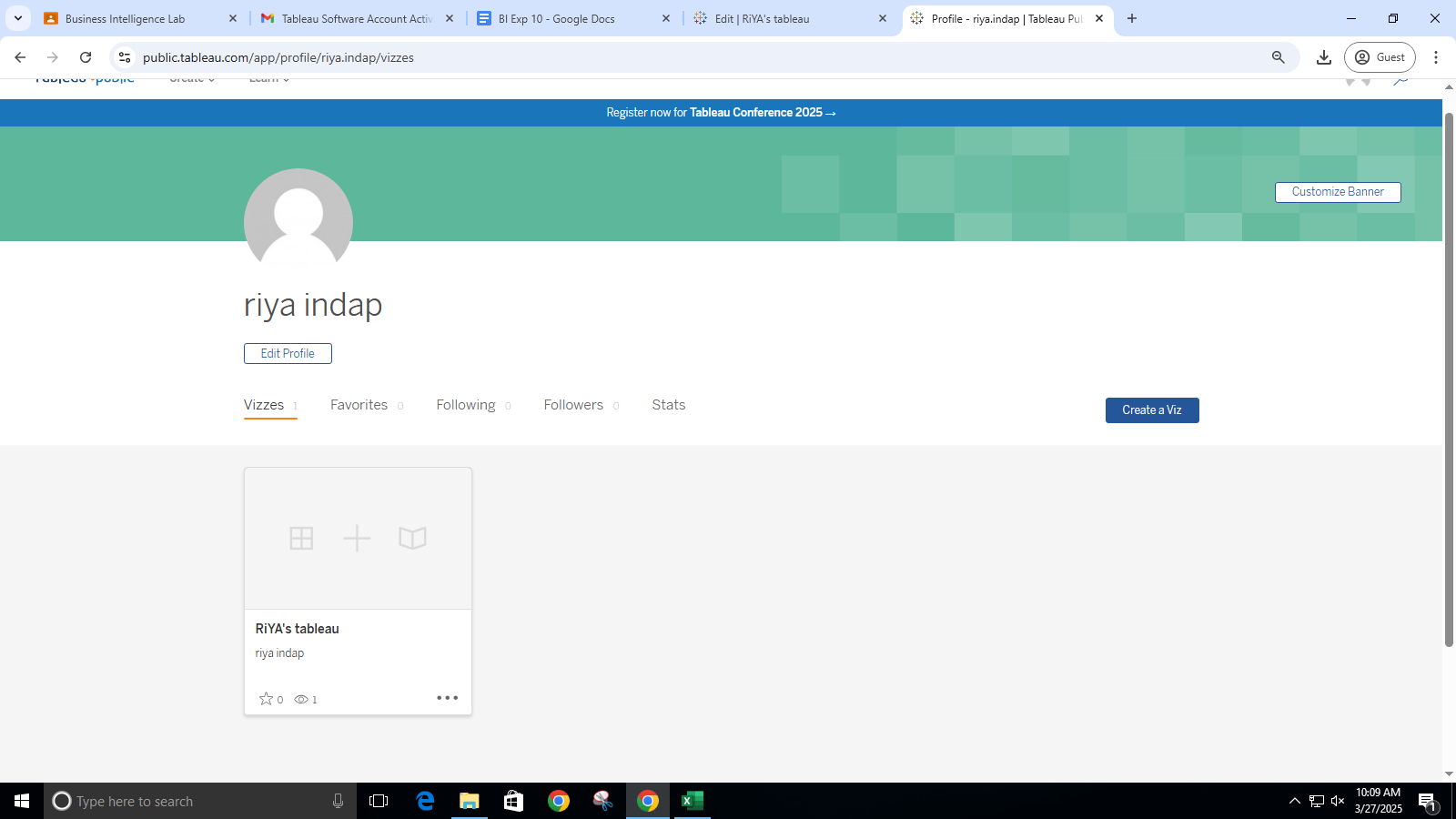
Creating a ‘RIYA3’ sheet having ’region’ as columns and ‘sum(sales)’ as rows which gave us a visualization of box-plot type of the data.



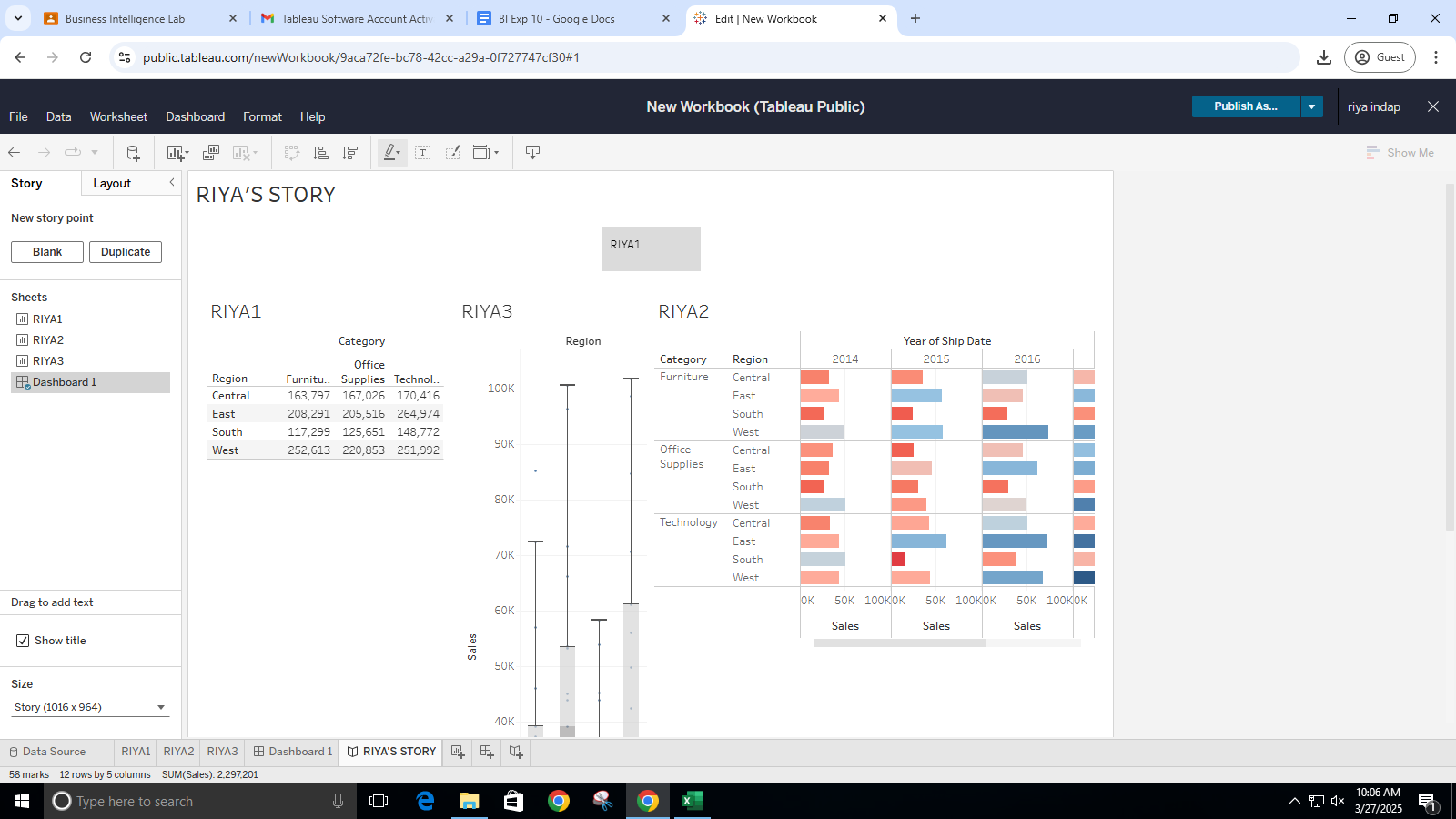
Creating dashboard with three sheets



Publishing Story



Creating story



QlikView

