**UNIT – V: Case studies on using computational tools for data analytics**

**Case studies on Visualization with the help of Tools like Altair Tableau, Rapid miner, and MATLAB. [Access to open-source tools will be granted for practical application to work on cases studies]**

**Introduction to Computational Tools for Data Analytics**

**What are Computational Tools?**

Computational tools are software applications designed to process, analyze, and visualize data efficiently. These tools play a crucial role in data analytics by providing user-friendly interfaces, automation, and powerful algorithms to extract insights from data.

**Commonly Used Tools**

1. **Altair**
	* Description: Altair is a declarative statistical visualization library for Python, built on top of Vega and Vega-Lite visualization grammars. It emphasizes simplicity and concise syntax for creating interactive visualizations.
	* Key Features:
		+ Simple syntax for complex visualizations
		+ Seamless integration with pandas DataFrames
		+ Ideal for interactive visual analytics
2. **Tableau**
	* Description: Tableau is a powerful data visualization tool that helps users create interactive dashboards and reports. It is widely used in business intelligence and data analytics.
	* Key Features:
		+ Drag-and-drop interface for intuitive report building
		+ Advanced visualization options like heatmaps, treemaps, and scatter plots
		+ Real-time data analysis and integration with multiple data sources
3. **RapidMiner**
	* Description: RapidMiner is a data science platform that enables users to build predictive models through a visual workflow interface. It supports data preparation, machine learning, and model deployment.
	* Key Features:
		+ No-code/low-code interface for fast model development
		+ Extensive library of machine learning algorithms
		+ Strong integration with Python and R for custom solutions
4. **MATLAB**
	* Description: MATLAB is a high-performance language and environment for technical computing. It is widely used in engineering, data analysis, and scientific computing.
	* Key Features:
		+ Robust mathematical functions and toolboxes for specialized tasks
		+ Powerful visualization capabilities for plotting complex data
		+ Extensive support for automation and custom scripting

**Case Studies in Data Analytics**

**Case Study 1: Predicting Customer Churn with RapidMiner**

**Problem Statement: A telecom company aims to predict customer churn based on user behavior and subscription data.**

**Steps:**

* Data Preparation: Import customer data into RapidMiner.
* Data Cleaning: Handle missing values and normalize data.
* Model Building: Use decision trees and logistic regression models.
* Evaluation: Assess model accuracy using confusion matrix and ROC curves.

**Outcome:** The company identified key churn indicators, improving customer retention strategies by 15%.

**Case Study 2: Sentiment Analysis with Altair**

**Problem Statement: An e-commerce company wants to analyze customer feedback to improve product features.**

**Steps:**

* Data Import: Load reviews data in Altair.
* Data Cleaning: Remove stop words, punctuation, and irrelevant text.
* Visualization: Create bar charts for sentiment distribution and word clouds for keyword identification.

**Outcome:** Product teams gained actionable insights to improve user satisfaction.

**Case Study 3: Financial Forecasting with MATLAB**

**Problem Statement: A financial firm wants to predict stock market trends.**

**Steps:**

* Data Collection: Import stock market data from online sources.
* Data Processing: Apply time-series forecasting models.
* Visualization: Use MATLAB's plotting tools for trend analysis.

**Outcome:** Improved investment decisions with a 10% increase in portfolio returns.

**Case Study 4: Dashboard Development with Tableau**

**Problem Statement: A retail chain wants to monitor sales performance across regions.**

**Steps:**

* Data Integration: Connect Tableau to sales databases.
* Visualization: Create interactive dashboards to display KPIs such as revenue, product sales, and regional performance.
* Analysis: Use heatmaps and trend lines for insights.

**Outcome:** Enhanced decision-making through clear data representation, improving sales forecasting by 20%.

**Practical Implementation with Open-Source Tools**

**Available Open-Source Tools for Hands-on Practice**

* Orange: Ideal for machine learning and visual programming.
* KNIME: Powerful for data preparation, analysis, and visualization.
* Power BI (Free Version): Effective for creating dashboards and visual reports.

**Recommended Projects for Practice**

* Customer segmentation using clustering algorithms.
* Sales forecasting with time-series analysis.
* Social media sentiment analysis using NLP tools.
* Fraud detection using anomaly detection techniques.

**Conclusion**

By utilizing computational tools like Altair, Tableau, RapidMiner, and MATLAB, students can develop strong analytical skills while gaining hands-on experience through real-world case studies. Combining theoretical understanding with practical applications ensures effective learning outcomes.