**PHASE 4 PROJECT SUBMISSION**

**PROJECT 1 - WEBSITE TRAFFIC ANALYSIS**

**TEAM MEMBERS:**

1. Nethaji K – 2021504531
2. Rohith M U – 2021504538
3. Sripushkalai S – 2021504547
4. Tharunsarwin U G – 2021504554

**PROBLEM DEFINITION:**

The project involves analyzing website traffic data to gain insights into user behavior, popular pages, and traffic sources. The goal is to help website owners enhance the user experience by understanding how visitors interact with the site. This project encompasses defining the analysis objectives, collecting website traffic data, using IBM Cognos for data visualization, and integrating Python code for advanced analysis.

**PHASE OBJECTIVE:**

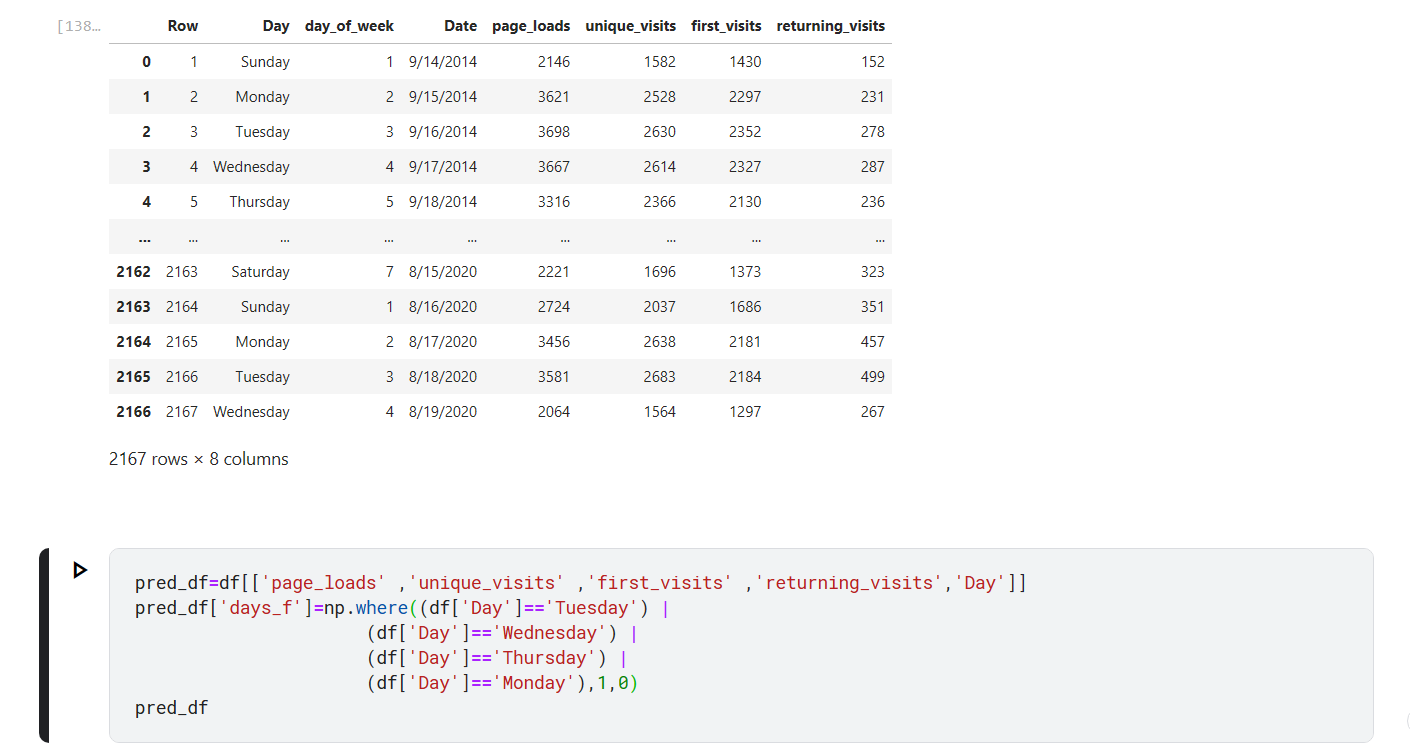
In this technology projects you will continue building your project by performing different analysis, model building and evaluation as per the project requirement. Perform different analysis and visualization using IBM Cognos.

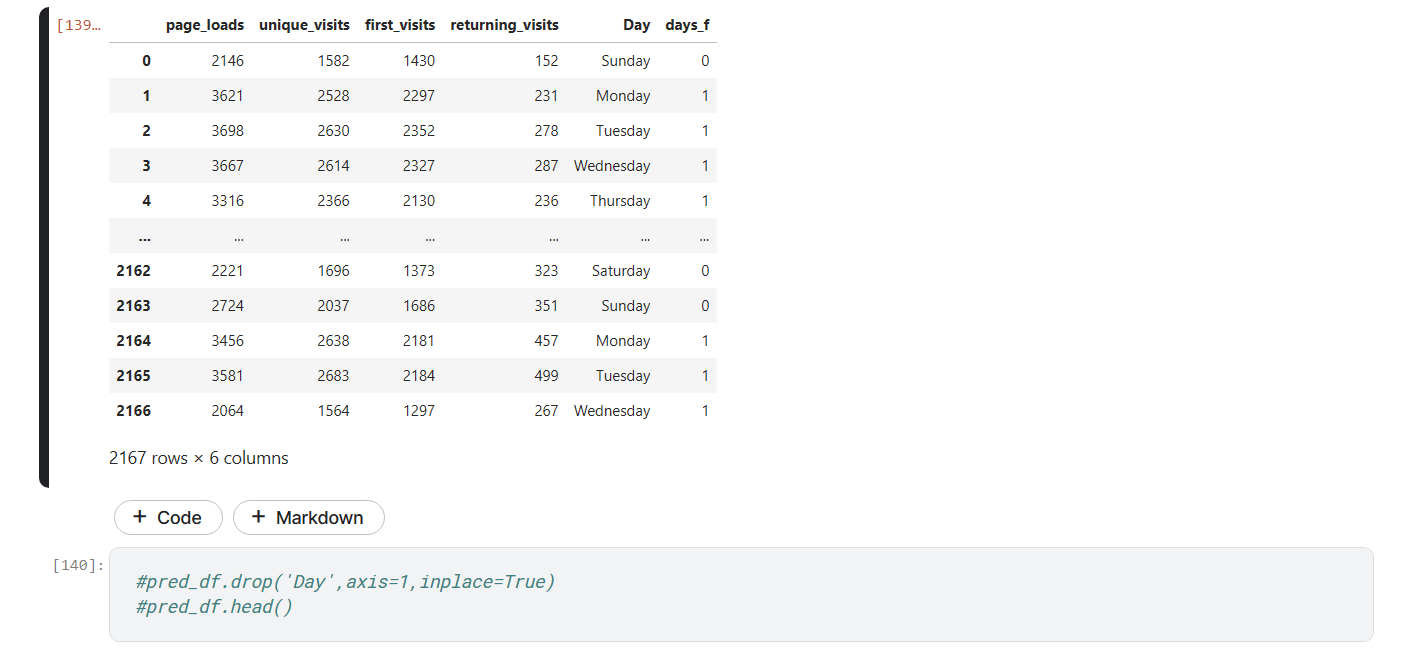
**DATASET LINK:**

<https://www.kaggle.com/datasets/bobnau/daily-website-visitors>

**PREDICTIVE MODEL:**







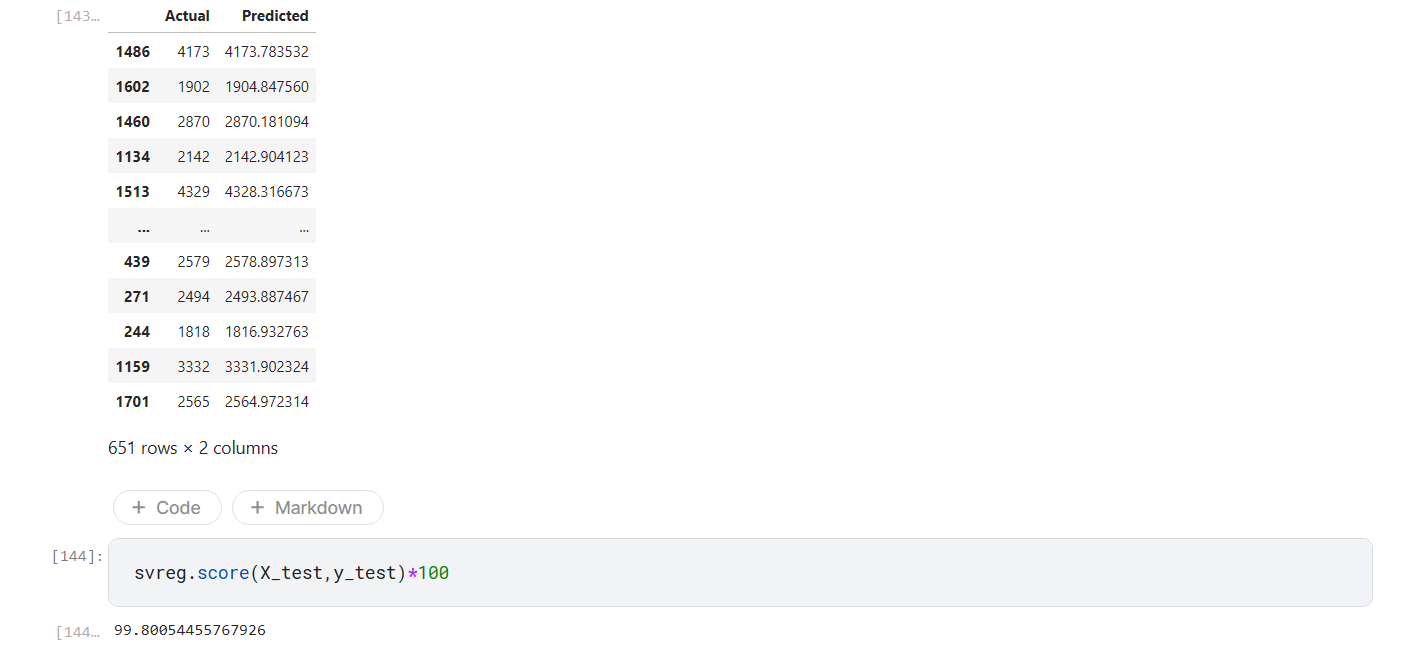
1. Linear Regression:





1. Support Vector Regression:





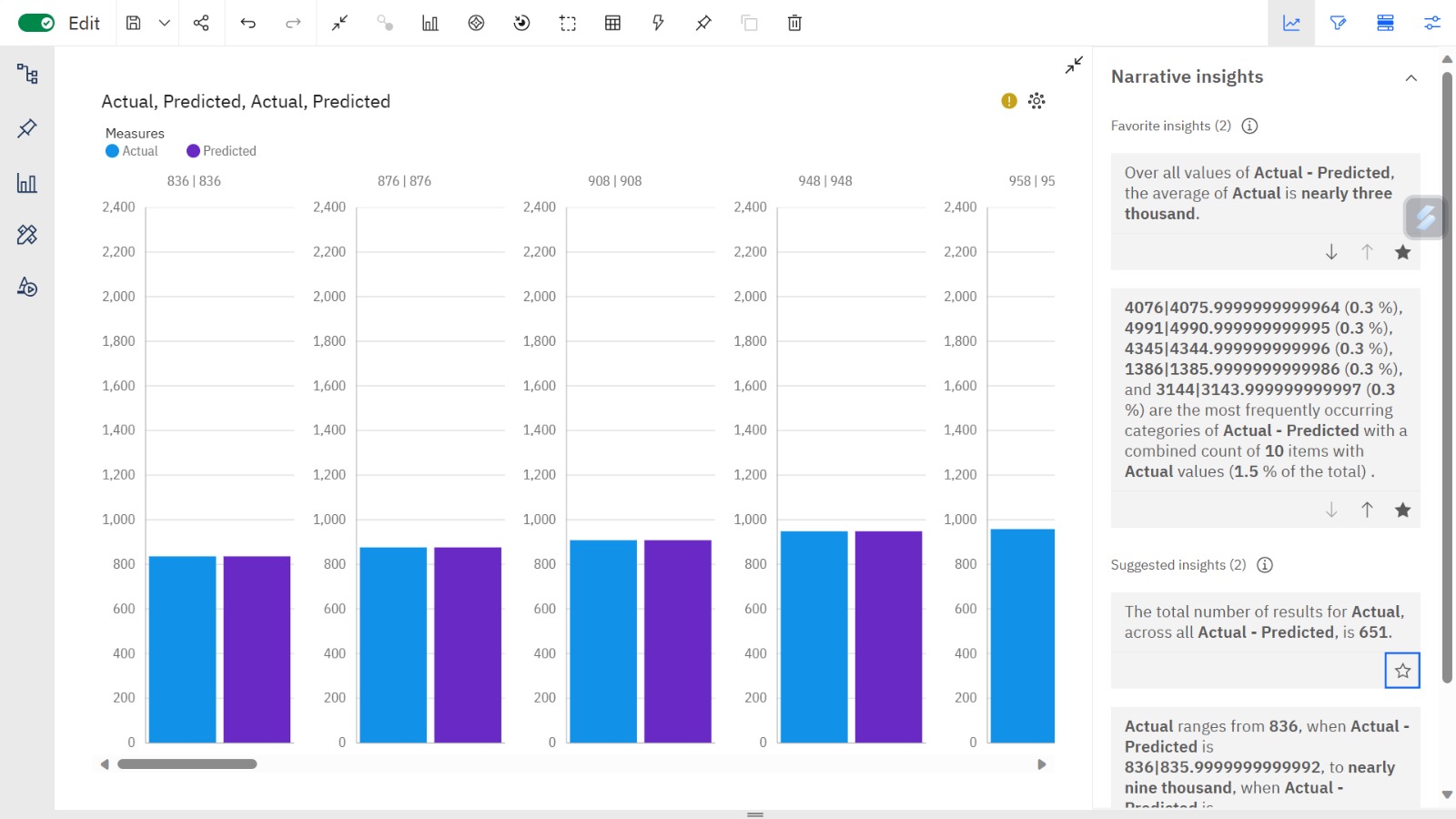
1. Decision Tree Regression:

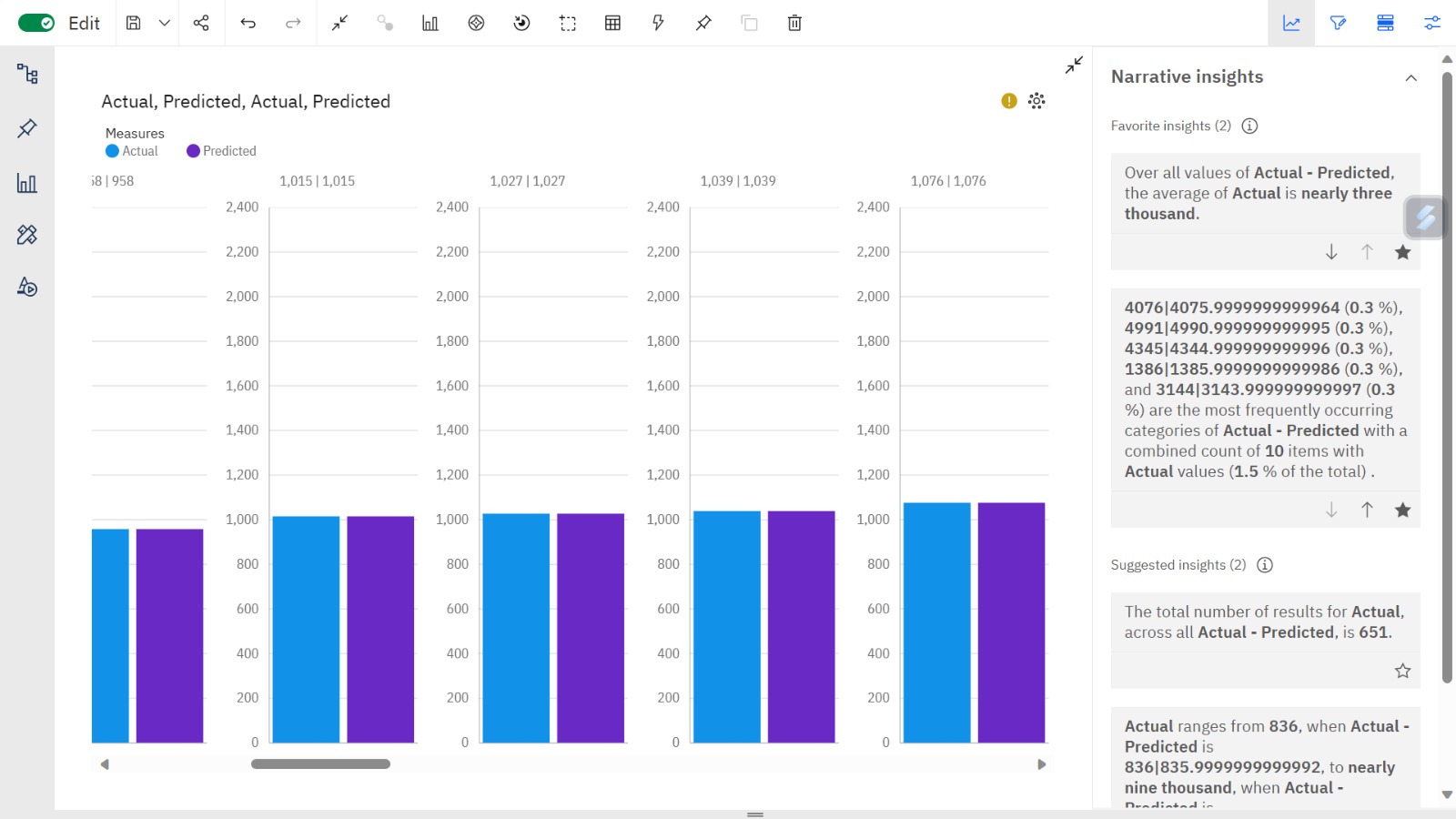


**COGNOS ANALYTICS:**

**Linear regression:**

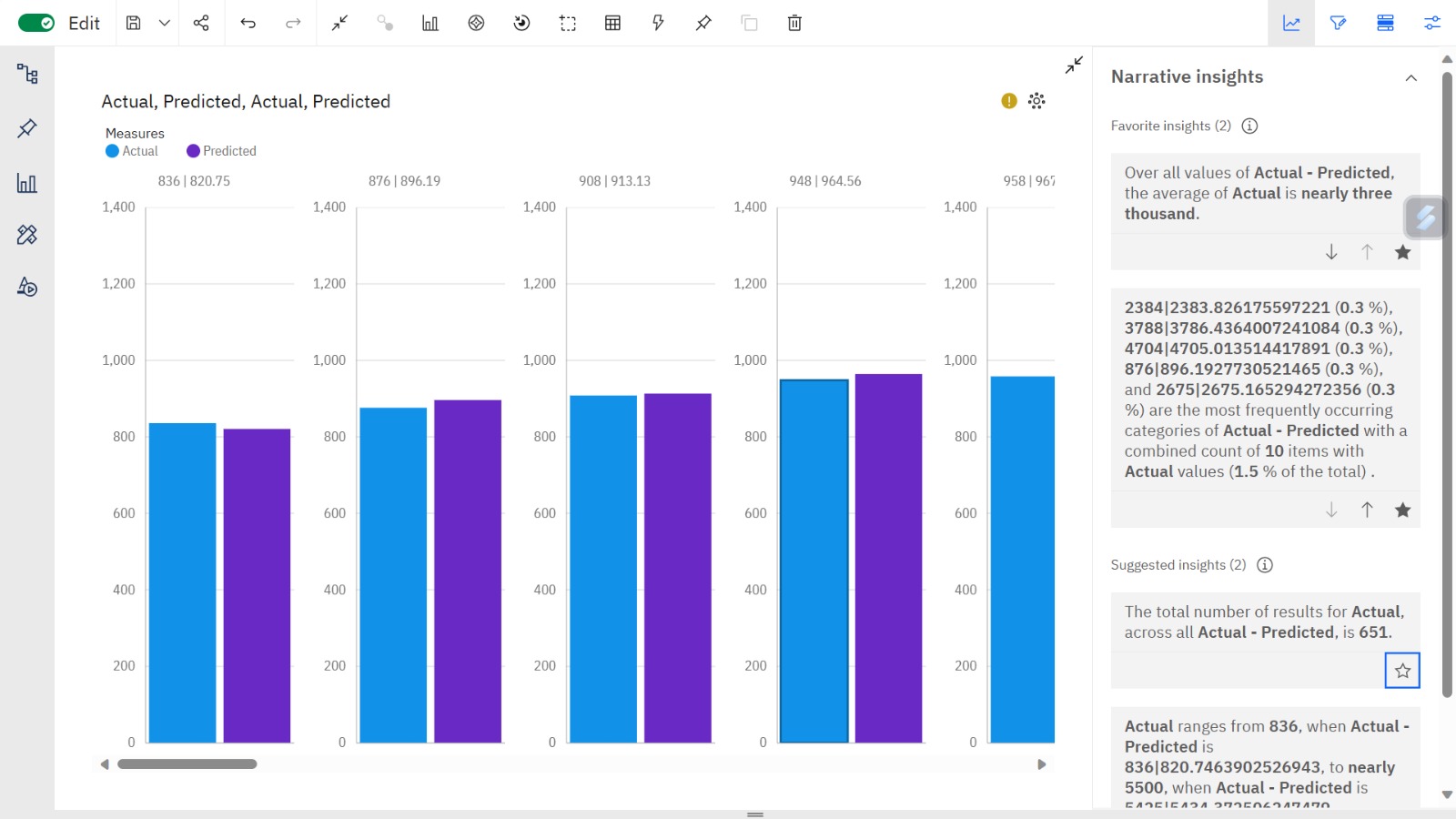
Linear regression plays an essential role in website traffic analysis for many reasons. By utilizing past data, it enables the prediction of future traffic. This not only assists in efficient resource planning, but also aids in the development of effective marketing strategies. The use of regression models aids in predicting valuable insights into website traffic trends and patterns, such as the influence of time of day, day of week, and marketing efforts. It is a valuable tool in evaluating the effects of A/B testing on traffic, detecting seasonal trends, and identifying anomalies and outliers. Linear regression is powerful, but it requires careful data selection and preprocessing for effective analysis.

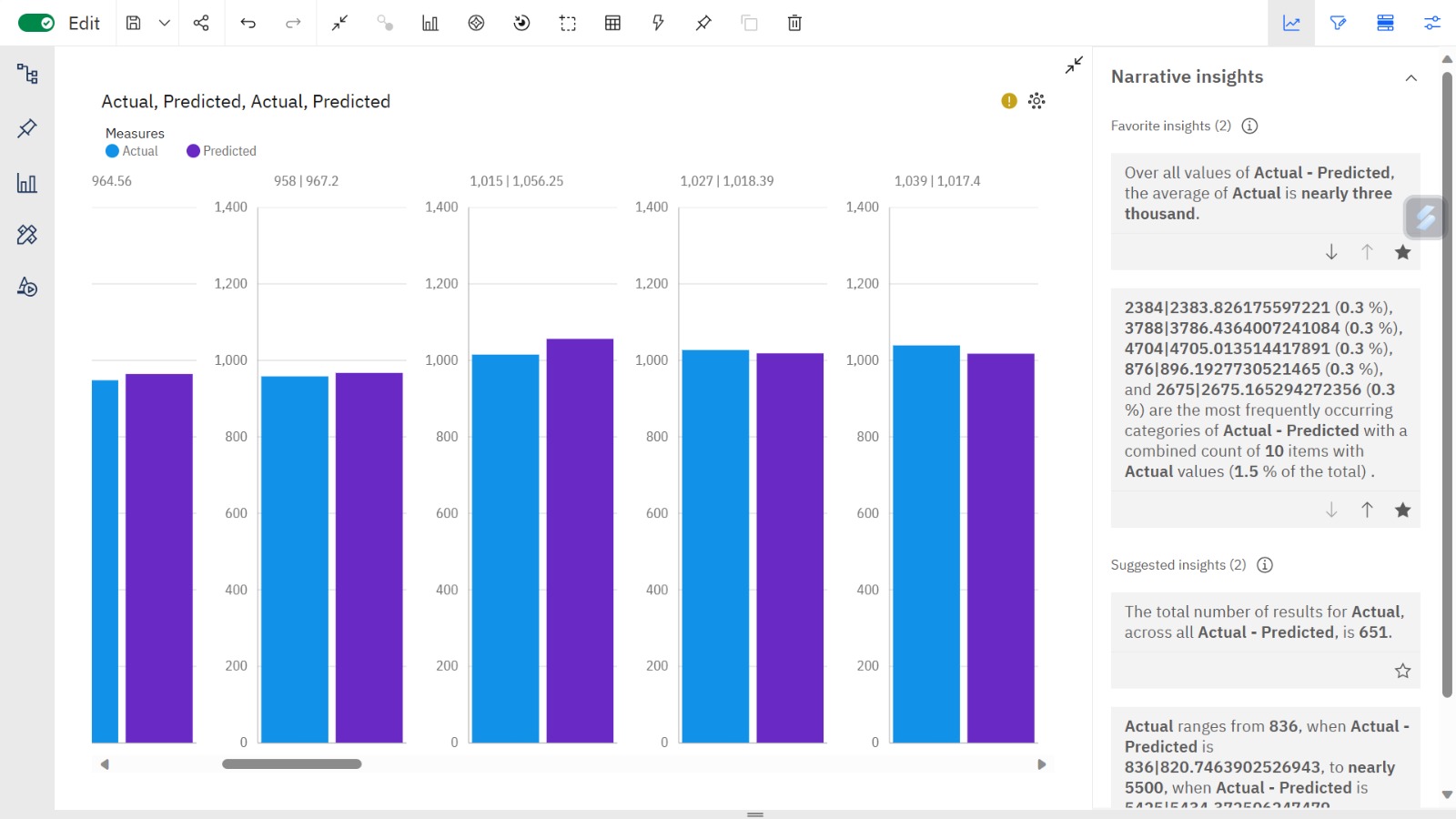




**Support Vector Regression:**

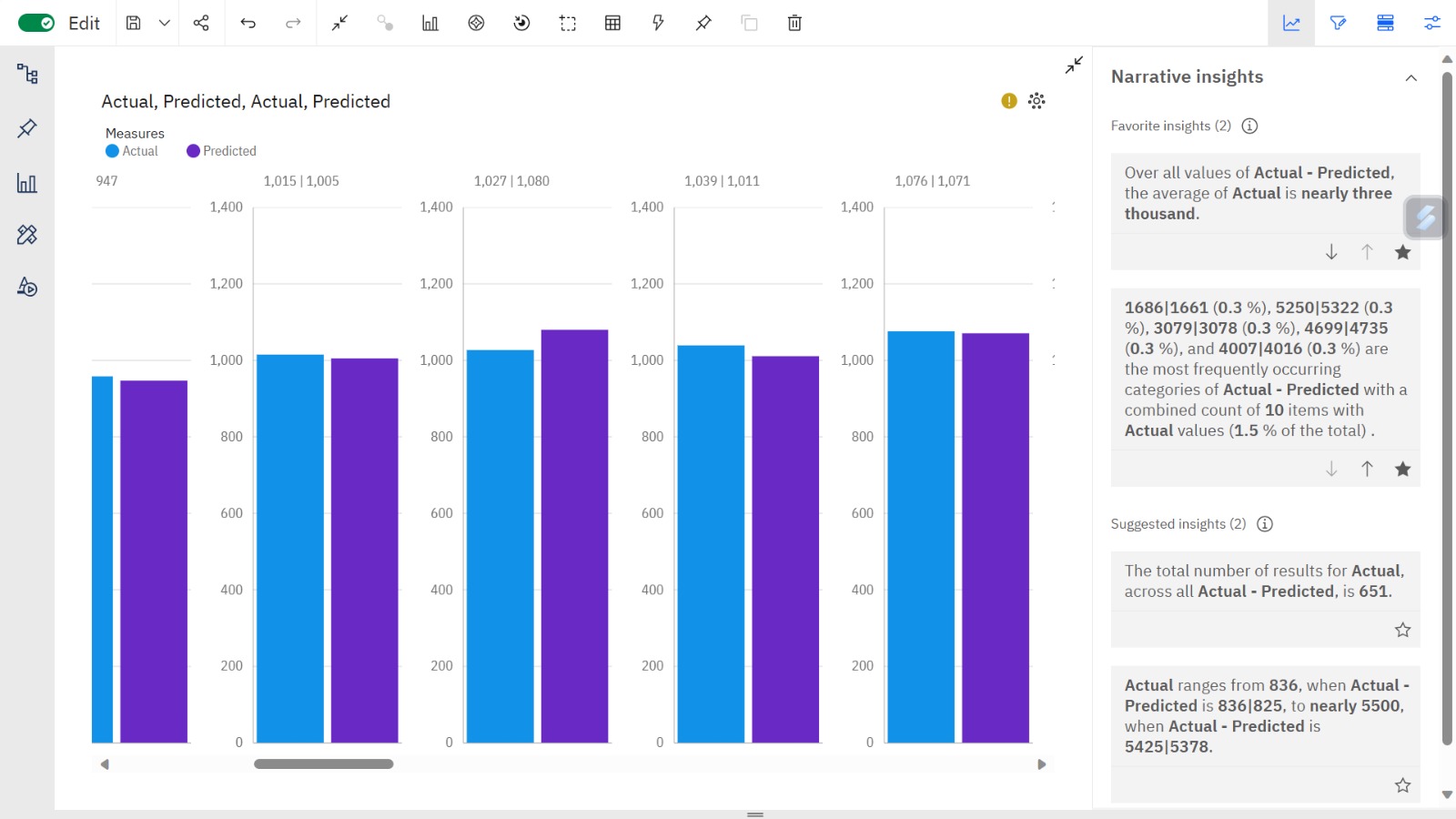
Support Vector Regression (SVR) is a powerful tool for website traffic analysis. Its exceptional ability to detect non-linear relationships and intricate patterns in data makes it the perfect choice for unraveling user behavior. SVR is robust to outliers, handles high-dimensional data effectively. This analytical tool offers customization through various kernel functions and can be applied to both regression and classification tasks, making it versatile.  SVR's adaptability to time series forecasting and ability to handle multicollinearity further amplify its effectiveness in website traffic analysis.

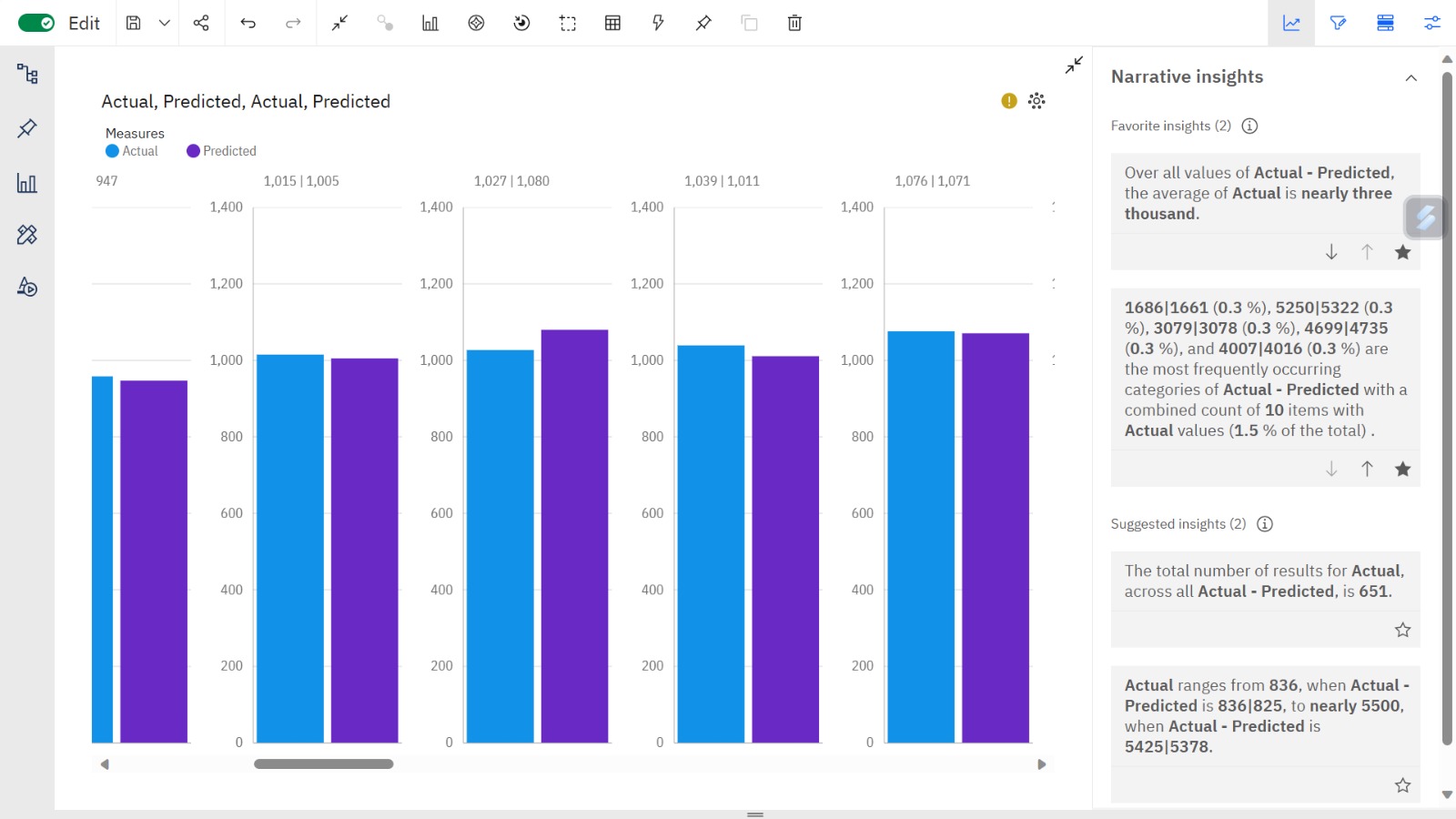




**Decision Tree Regression:**

Decision Tree Regression is a valuable asset when analyzing website traffic for several reasons. It excels at modeling complex, non-linear relationships between predictor variables and website traffic. Decision Trees handle both categorical and numerical variables, making them versatile for various types of website data. They are robust to outliers and can handle missing data. Decision Trees do not assume a linear relationship, providing flexibility in capturing different traffic patterns.





**CONCLUSION:**

Thus the predictive models were built and different visualizations were performed using IBM cognos.