**CUSTOMER CHURN PREDICTION**

**Project Objective:**

The project's objective is to develop a customer churn prediction system for a business, helping it proactively identify and retain customers who are likely to leave. This can be achieved by using machine learning and data analytics to predict which customers are at risk of churning based on historical data.

**Design Thinking Process:**

**1. Empathize:**

* Understand the business and its customers' needs.
* Gather insights on why customers churn.

**2.Define:**

* Clearly define the problem: customer churn prediction.
* Set specific goals and success criteria.

**3. Ideate:**

* Brainstorm potential solutions and approaches for churn prediction.
* Consider different algorithms, data sources, and features.

**4. Prototype:**

* Create a basic model or proof of concept to test the chosen approach.
* Use a subset of historical data for initial testing.

**5. Test:**

* Evaluate the prototype's performance using metrics like accuracy, precision, recall, and F1-score.
* Gather feedback from stakeholders.

**6. Refine:**

* + Improve the model based on feedback and testing results.
  + Optimize features, algorithms, and hyperparameters.

**7. Develop:**

* + Build a full-scale churn prediction system.
  + Integrate it with the business's data infrastructure.

**8. Test Again:**

* + Conduct extensive testing and validation.
  + Ensure the system's reliability and accuracy.

**9.Deploy:**

* + Implement the system in the business's operations.
  + Provide necessary training to employees.

**10. Monitor:**

* + Continuously monitor the system's performance.
  + Make adjustments as needed to maintain accuracy.

**Development Phases for Customer Churn Prediction:**

**1. Data Collection:**

* + Gather historical customer data, including demographics, transaction history, support interactions, etc.
  + Ensure data quality and clean the dataset.

**2. Data Preprocessing:**

* + Handle missing values and outliers.
  + Encode categorical variables.
  + Split the data into training and testing sets.

**3. Feature Engineering:**

* + Create relevant features that can help predict churn, e.g., customer tenure, frequency of purchases, customer support interactions.

**4. Model Selection:**

* + Choose machine learning algorithms suitable for churn prediction, such as logistic regression, decision trees, random forests, or neural networks.

**5. Model Training:**

* + Train the selected model using the training dataset.
  + Tune hyperparameters for optimal performance.

**6. Model Evaluation:**

* + Assess the model's performance using appropriate evaluation metrics.
  + Conduct cross-validation to ensure generalizability.

**7. Model Deployment:**

* + Deploy the trained model in the business environment.
  + Set up data pipelines for real-time or batch prediction.

**8. Feedback Loop:**

* + Continuously gather new data for retraining.
  + Update the model periodically to adapt to changing customer behaviour.

**9. Visualization and Reporting:**

* + Create dashboards and reports to provide insights to stakeholders.
  + Monitor the model's predictions and their impact on customer retention.

**10. Business Strategy Implementation:**

* + Use the churn predictions to implement targeted retention strategies, such as personalized offers or improved customer service.

**11. Maintenance and Optimization:**

* + Regularly maintain and optimize the churn prediction system to ensure its effectiveness.
* This outline provides a structured approach to developing a customer churn prediction system, from design thinking to development and ongoing maintenance. The exact details and tools used may vary depending on the specific business and industry.

**Analysis Objectives:**

**The analysis objectives for customer churn prediction are to:**

1**. Identify Churn Risk:** Determine which customers are at risk of churning.

2**. Understand Churn Drivers:** Identify the factors and behaviours that contribute to customer churn.

3. **Develop Predictive Models:** Create accurate predictive models to forecast customer churn.

4. **Implement Retention Strategies:** Use insights from the analysis to develop and deploy retention strategies to reduce churn rates.

5. **Monitor and Adapt:** Continuously monitor and refine the churn prediction models and strategies.

**Data Collection Process:**

1. **Data Sources:** Gather data from various sources such as customer databases, sales records, customer support interactions, and marketing data.

2. **Data Integration:** Combine and clean the data to create a unified dataset for analysis.

3. **Data Sampling:** If the dataset is large, consider taking a random sample to speed up analysis and model development.

4. **Data Transformation:** Perform preprocessing tasks, including handling missing values, encoding categorical variables, and normalizing or scaling features.

5. **Feature Engineering:** Create relevant features such as customer tenure, purchase frequency, customer support interactions, and customer feedback scores.

**Data Visualization Using IBM Cognos:**

1. **Data Exploration:** Use IBM Cognos to explore the dataset by creating various visualizations like bar charts, histograms, and scatter plots to understand the data distribution and relationships between variables.

2. **Customer Churn Visualization:** Create visualizations to depict the churn rate over time, analyze patterns, and identify any seasonality or trends.

3. **Churn Driver Analysis:** Use Cognos to create visualizations that show the impact of various customer attributes and behaviours on churn. For example, create stacked bar charts or heatmaps to understand which factors correlate with churn.

4. **Geospatial Analysis:** Utilize IBM Cognos geospatial capabilities to map customer churn by location. This can reveal geographic patterns in churn.5.Dashboard Creation: Develop interactive dashboards in IBM Cognos to present the results of the analysis to stakeholders. These dashboards can include key performance indicators (KPIs) related to churn, trends, and predictive model performance.

**Predictive Modelling for Customer Churn Prediction:**

1. **Model Selection:** Choose appropriate machine learning algorithms for churn prediction, such as logistic regression, decision trees, random forests, or gradient boosting.

2. **Data Splitting:** Split the dataset into a training set and a testing set for model development and evaluation.

3. **Model Training:** Train the selected models on the training data using IBM Cognos or another preferred tool.

4. **Hyperparameter Tuning:** Optimize model hyperparameters through techniques like grid search or random search to improve model performance.

5. **Model Evaluation:** Assess model performance using metrics like accuracy, precision, recall, F1-score, and ROC AUC. Use cross-validation to ensure robustness.

6. **Model Deployment:** Deploy the trained model within IBM Cognos or integrate it with the business's data infrastructure for real-time predictions.

7. **Monitoring and Feedback Loop:** Continuously monitor the predictive model's performance and retrain it as new data becomes available. Update the model to adapt to changing customer behaviour.

8. **Strategy Implementation:** Utilize the predictions from the model to implement targeted retention strategies, such as sending personalized offers to at-risk customers.

9. **Reporting:** Create reports and dashboards in IBM Cognos to track the success of retention strategies and the impact on customer churn.

10. **Optimization:** Regularly refine and optimize the churn prediction model based on feedback and evolving business needs.

By following these steps, a business can effectively use IBM Cognos and predictive modelling to understand customer churn, develop strategies to retain customers, and continually improve its approach.

**Analysis Objectives:**

**The analysis objectives for customer churn prediction are to:**

**1. Identify Churn Risk:** Determine which customers are at risk of churning.

**2. Understand Churn Drivers:** Identify the factors and behaviours that contribute to customer churn.

**3. Develop Predictive Models:** Create accurate predictive models to forecast customer churn.

**4. Implement Retention Strategies:** Use insights from the analysis to develop and deploy retention strategies to reduce churn rates.

**5. Monitor and Adapt:** Continuously monitor and refine the churn prediction models and strategies.

**Data Collection Process:**

**1. Data Sources:** Gather data from various sources such as customer databases, sales records, customer support interactions, and marketing data.

**2. Data Integration:** Combine and clean the data to create a unified dataset for analysis.

**3. Data Sampling:** If the dataset is large, consider taking a random sample to speed up analysis and model development.

**4. Data Transformation:** Perform preprocessing tasks, including handling missing values, encoding categorical variables, and normalizing or scaling features.

**5. Feature Engineering:** Create relevant features such as customer tenure, purchase frequency, customer support interactions, and customer feedback scores.

**Data Visualization Using IBM Cognos:**

**1. Data Exploration:**

Use IBM Cognos to explore the dataset by creating various visualizations like bar charts, histograms, and scatter plots to understand the data distribution and relationships between variables.

**2. Customer Churn Visualization:**

Create visualizations to depict the churn rate over time, analyse patterns, and identify any seasonality or trends.

**3. Churn Driver Analysis:**

Use Cognos to create visualizations that show the impact of various customer.

Replicating analysis and generating visualizations using IBM Cognos and building a predictive model using Python involves a multi-step process. Here's a high-level overview of the steps:

**Using IBM Cognos for Analysis and Visualization:**

1. **Data Preparation:**

* + Start by importing your dataset into IBM Cognos.
  + Clean and preprocess the data as needed.

2. **Data Exploration:**

* Use Cognos' data exploration tools to understand your data.
* Create data visualizations like charts and graphs to explore relationships and patterns.

3. **Create Reports and Dashboards:**

* Build reports and dashboards in Cognos to present your findings.
* Add visualizations and interactivity for a comprehensive analysis.

4. **Analyse Data:**

* Use Cognos' analysis capabilities to perform statistical analyses or create custom calculations.

5. **Predictive Analysis (if needed):**

* Cognos has limited predictive modelling capabilities. For more advanced predictive modelling, it's best to switch to Python.

**Building Predictive Models Using Python:**

1. **Data Extraction:**

* Export the pre-processed data from Cognos to a format compatible with Python (e.g., CSV).

2. **Python Environment Setup:**

* Install Python and required libraries (e.g., Pandas, Scikit-learn, Matplotlib).

3. **Data Preparation:**

* Load the data in Python using Pandas.
* Perform feature engineering, handle missing values, and encode categorical variables.

4. **Split the Data:**

* Divide the data into training and testing sets.

5. **Model Selection and Training:**

* Choose an appropriate machine learning algorithm (e.g., regression, decision tree, random forest).
* Train the model on the training data.

6. **Model Evaluation:**

* Evaluate the model's performance on the testing data using appropriate metrics (e.g., accuracy, F1-score, ROC-AUC).

7. **Visualization:**

* Create visualizations to represent model performance (e.g., ROC curves, confusion matrices).

8. **Tuning and Optimization:**

* If necessary, fine-tune the model parameters to improve its performance.

9. **Deployment (if needed):**

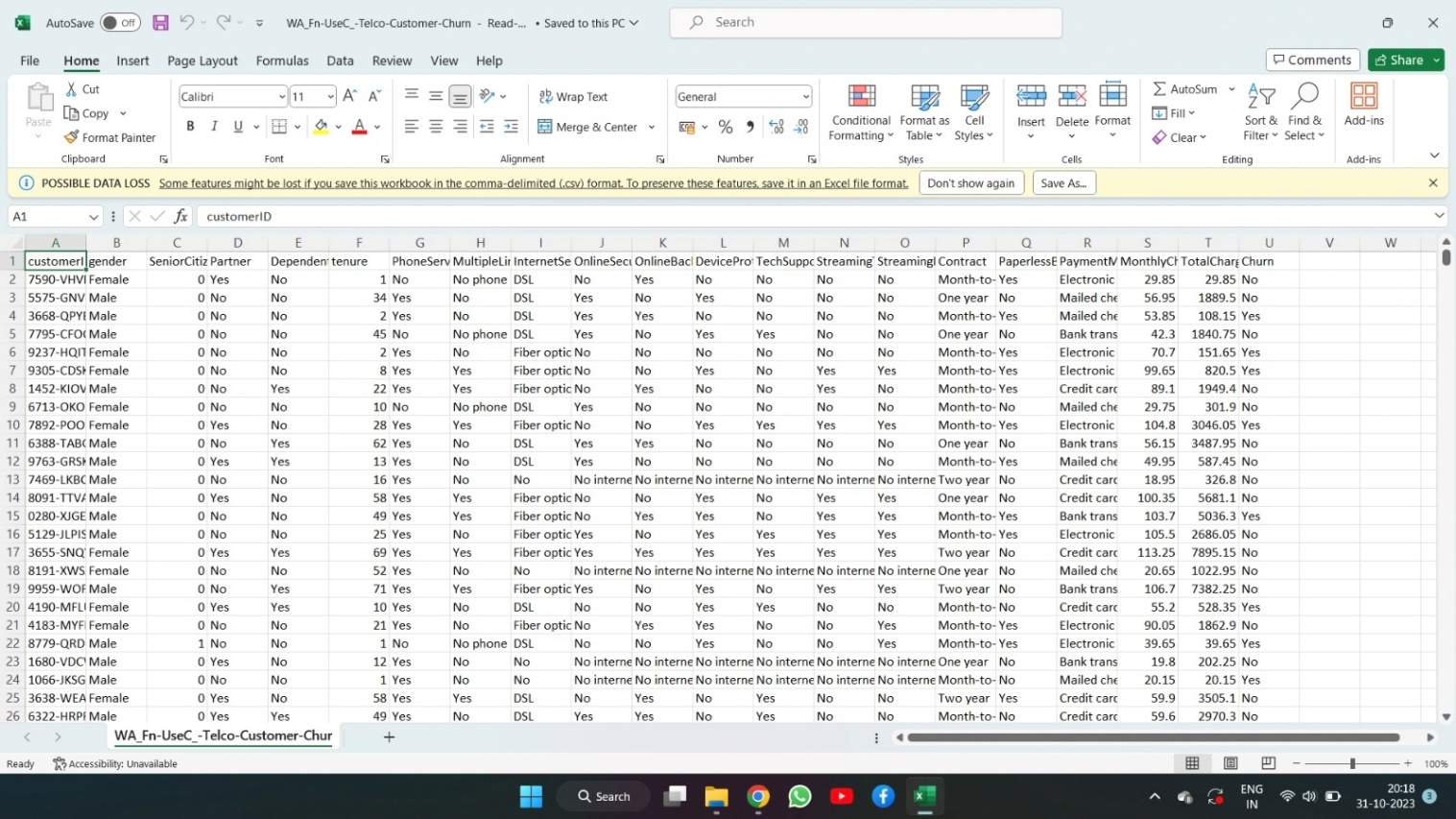
* If you plan to deploy the model, save it and build an API or interface for integration into your application.

10. **Documentation:**

* Document the entire process, including data preprocessing, model selection, and performance metrics.

Remember that the choice between using IBM Cognos and Python for analysis and visualization largely depends on your specific requirements and the complexity of your data. Cognos is great for business intelligence and reporting, while Python offers more flexibility for advanced data analysis and modelling. You can combine both tools as needed to leverage their respective strengths.

**DATA SET FOR CUSTOMER CHURN PREDICTION**



**PYTHON CODE FOR VISUALIZATION AND PREDICTIVE MODEL FOR THE DATA SET:**

python

# Import necessary libraries

import pandas as pd

import matplotlib.pyplot as plt

from sklearn.model\_selection import train\_test\_split

from sklearn.tree import DecisionTreeClassifier

from sklearn.metrics import accuracy\_score, confusion\_matrix

from sklearn.preprocessing import LabelEncoder

# Load your dataset

data = pd.read\_csv('your\_dataset.csv')

# Data Exploration and Visualization

# Example: Create a histogram of a numerical feature

plt.hist(data['Age'], bins=20)

plt.xlabel('Age')

plt.ylabel('Frequency')

plt.title('Age Distribution')

plt.show()

# Data Preprocessing

# Example: Encoding categorical variables

label\_encoder = LabelEncoder()

data['Category'] = label\_encoder.fit\_transform(data['Category'])

# Split the data into features and target

X = data.drop(columns=['TargetColumn']) # Replace 'TargetColumn' with your target variable

y = data['TargetColumn']

# Split the data into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Build a predictive model (Decision Tree classifier)

model = DecisionTreeClassifier()

model.fit(X\_train, y\_train)

# Make predictions on the test set

y\_pred = model.predict(X\_test)

# Model Evaluation

accuracy = accuracy\_score(y\_test, y\_pred)

confusion = confusion\_matrix(y\_test, y\_pred)

print(f'Accuracy: {accuracy}')

print('Confusion Matrix:')

print(confusion)

# Visualization of Model Results

# Example: Plot a confusion matrix

plt.imshow(confusion, cmap='Blues', interpolation='nearest')

plt.colorbar()

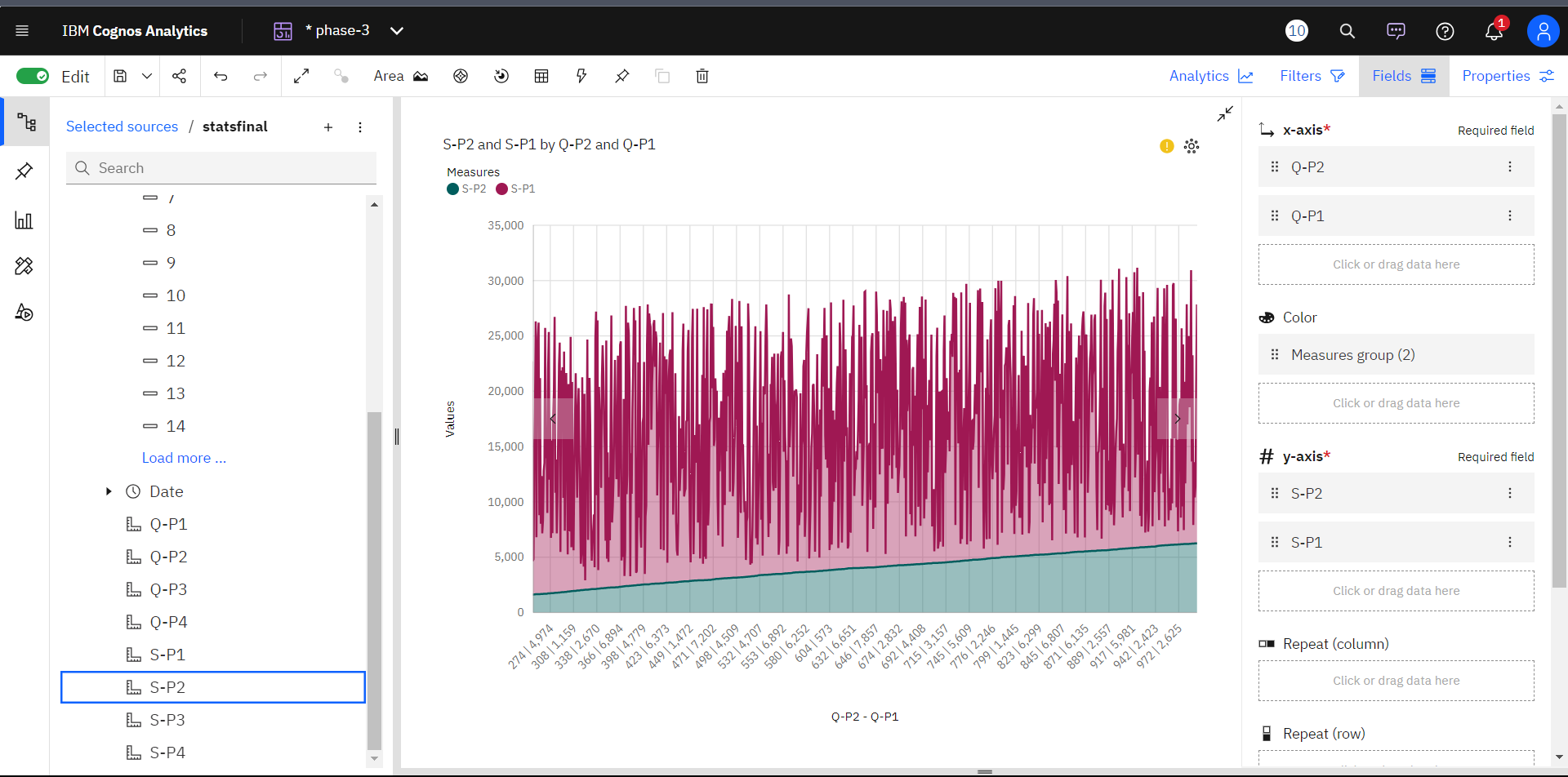
plt.xlabel('Predicted')

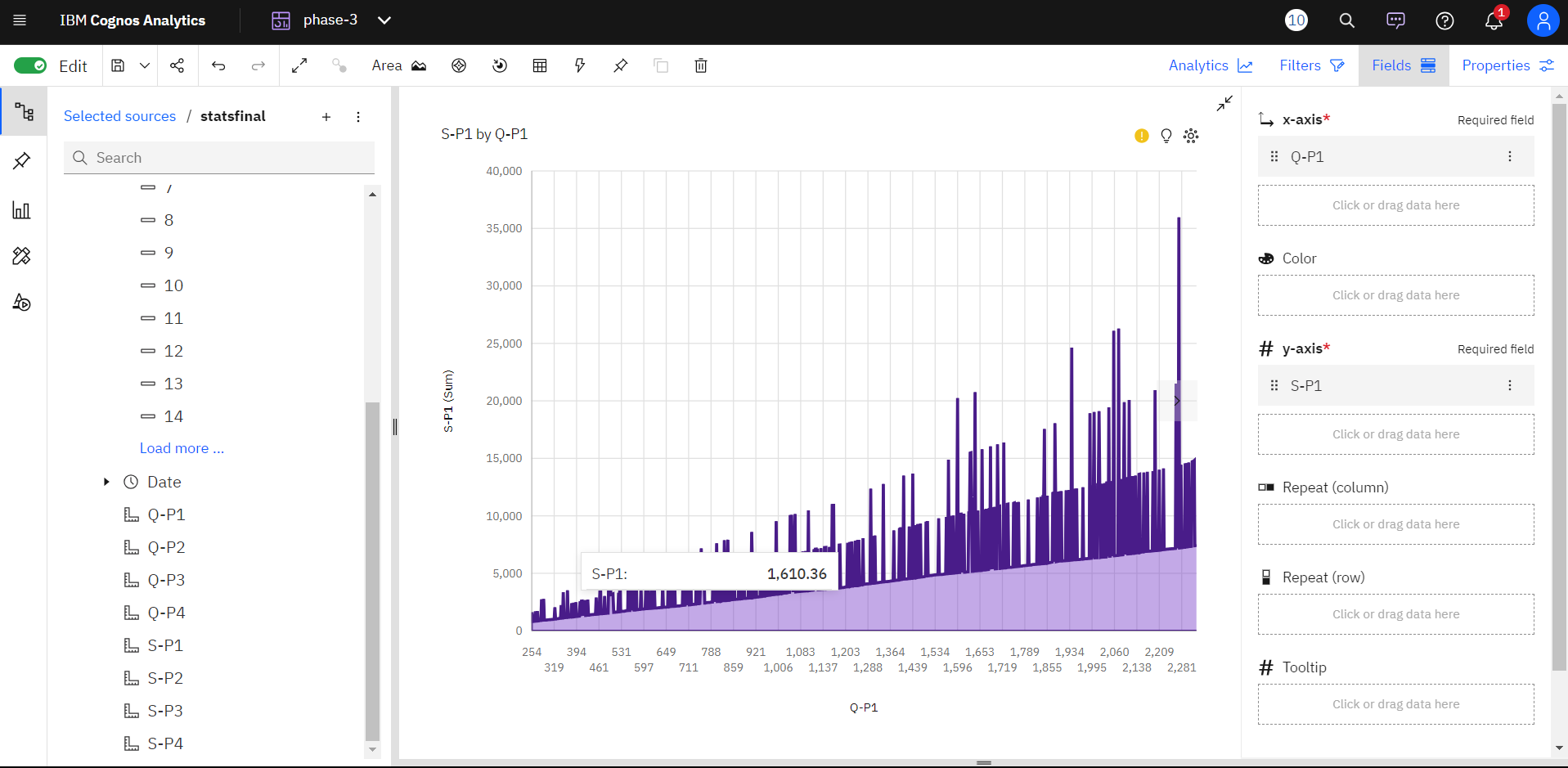
plt.ylabel('Actual')

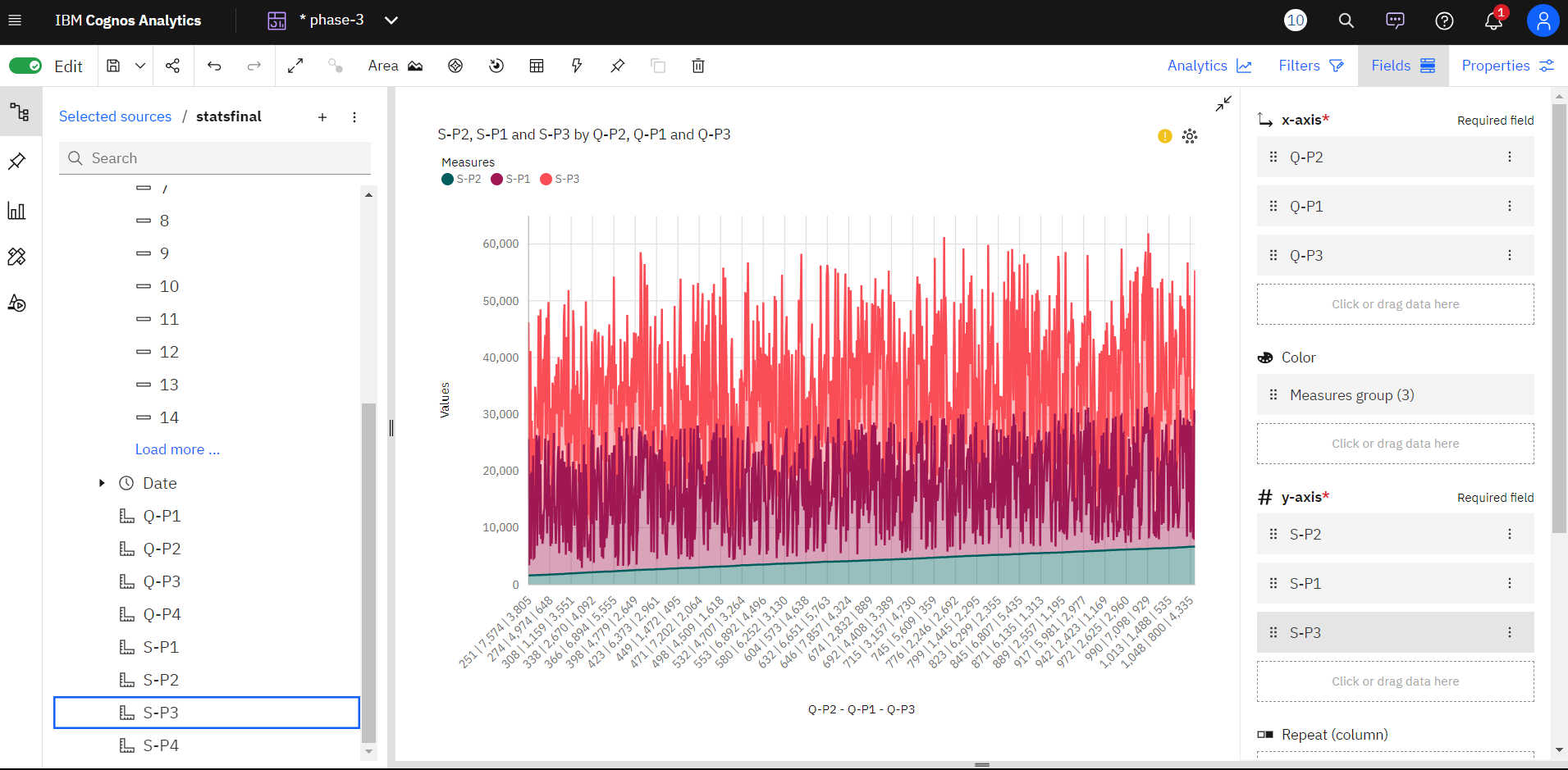
plt.title('Confusion Matrix')

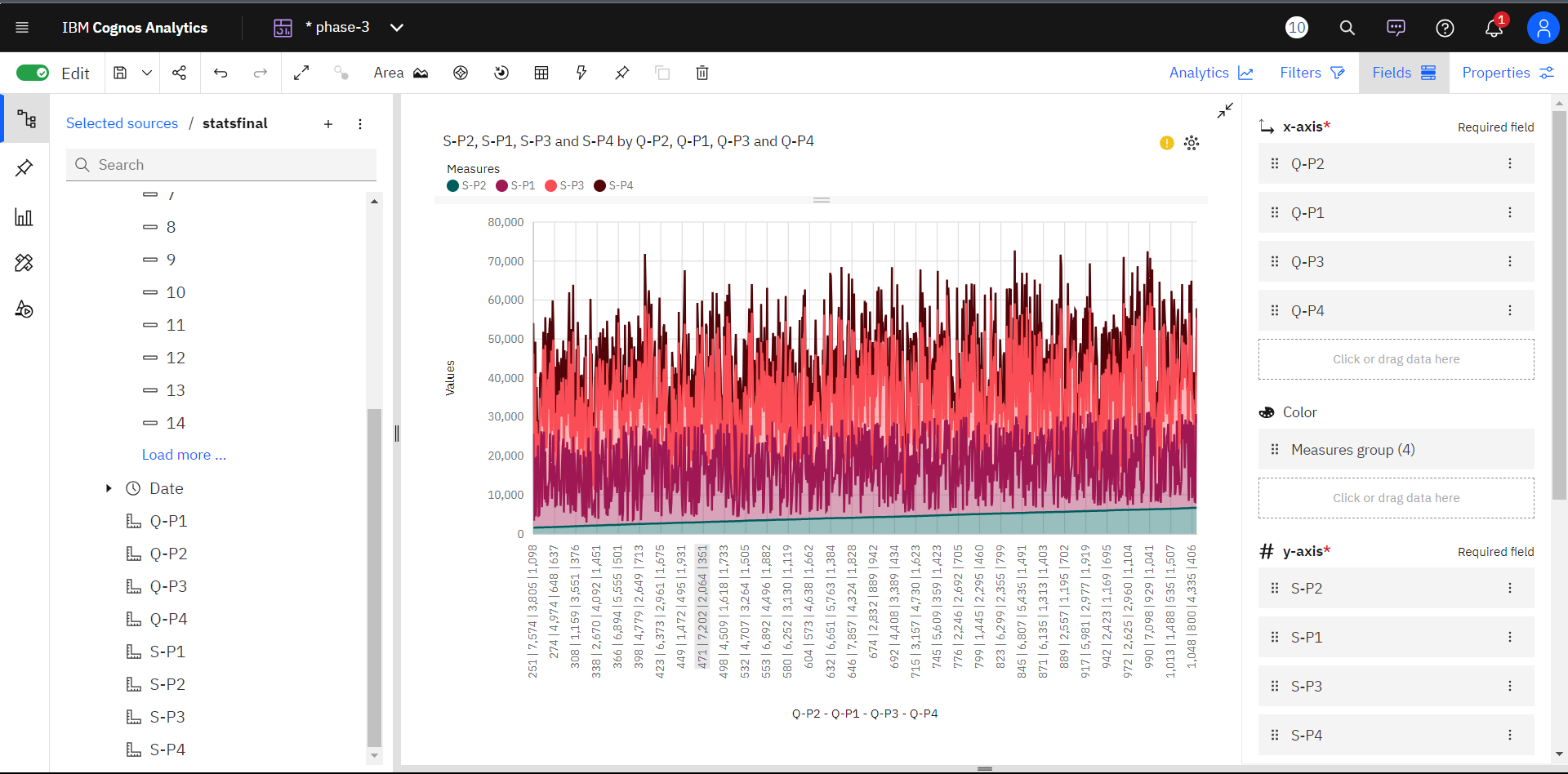
plt.show()

**VISUALIZATION USING IBM COGNOS AND PREDICTIVE MODELLING:**



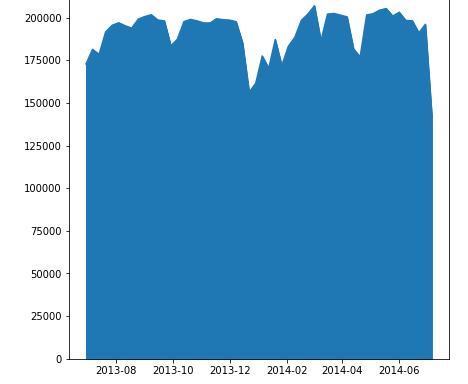


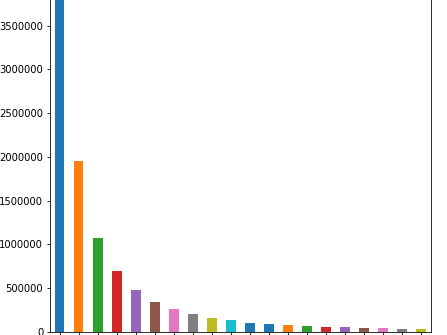


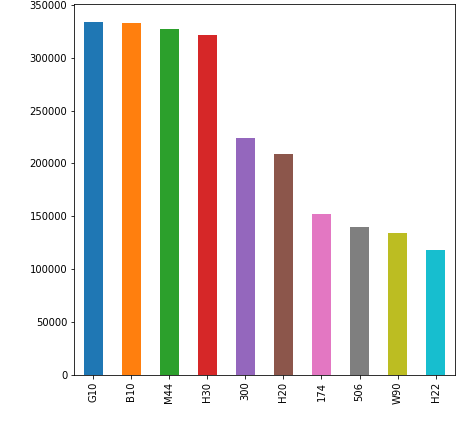


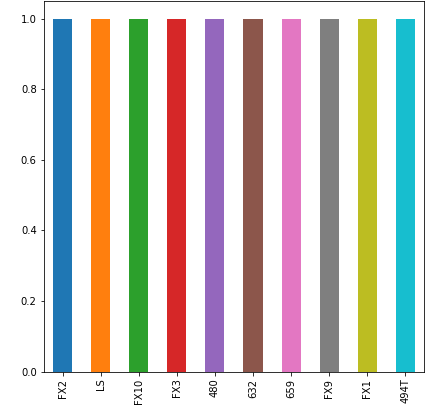
PROJECT REPORT: CUSTOMER CHURN PREDICTION

1. Number of people using telecommunication









**CONCLUSION:**

Customer churn prediction is a valuable tool for businesses to proactively identify and retain at-risk customers. By leveraging advanced analytics and machine learning models, companies can reduce customer churn, enhance customer satisfaction, and ultimately improve their bottom line. It is important for businesses to continuously refine and update their churn prediction models to stay ahead in the dynamic market landscape.