PUBLIC TRANSPORT EFFICIENCY AND ANALYSIS

**Project Objective:**

The project aims to enhance public transport analysis on survey through educational campaigns, outreach programs, and digital media, fostering informed and engaged communities. By disseminating key information and encouraging active participation, we seek to empower individuals to make informed decisions and drive positive change in their communities.

**Design Thinking Process:**

**1. Empathize:**

* Empathizing with the public's awareness needs is essential for effective communication and engagement.
* By putting ourselves in their shoes, we can identify their priorities, fears, and aspirations, which in turn allows us to tailor our messaging to resonate with their emotions and values.

**2.Define:**

* Clearly define the problem: public awareness.
* Set specific goals and success criteria.

**3. Ideate:**

* Brainstorm potential solutions and approaches for survey 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 survey prediction system.
  + Integrate it with the people'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 public awareness:**

**1. Data Collection:**

* + Gather peoples data, including nasme,gender,age, 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 survey, e.g., working, product sale, customer support interactions.

**4. Model Selection:**

* + Choose machine learning algorithms suitable for survey 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 survey predictions to implement targeted retention strategies, such as personalized offers or improved customer service.

**11. Maintenance and Optimization:**

* + Regularly maintain and optimize the survey 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 public awareness are to:**

1. **Clarity**: Ensure that the objectives are clear and specific, leaving no room for ambiguity.

2. **Measurability**: Set quantifiable goals that can be measured to assess the campaign's success.

3. **Relevance**: Ensure that the objectives align with the overall mission and goals of the campaign.

4. **Target Audience**: Define the specific audience the campaign aims to reach and influence.

5. **Timeframe**: Establish a realistic timeline for achieving the objectives.

6. **Realistic Expectations**: Ensure that the objectives are attainable given the available resources and constraints.

7. **Alignment with Messaging**: Verify that the objectives are consistent with the message and theme of the campaign.

8. **Behavioural Change**: If applicable, consider how the campaign's objectives aim to change or influence public behaviour.

9. **Information Dissemination**: Assess whether the objectives focus on spreading key information or raising awareness about a particular issue.

10. **Impact Measurement**: Define how success will be measured, whether it's through surveys, website traffic, social media engagement, or other metrics.

11. **Integration with Channels**: Ensure that the objectives align with the communication channels and platforms being used.

12. **Stakeholder Involvement**: Consider how various stakeholders, including partners and influencers, may contribute to achieving the objectives.

13. **Adaptability**: Be prepared to adjust the objectives if the campaign evolves or if new information becomes available.

14. **Cost and Resource Management**: Ensure that the objectives are realistic within the campaign's budget and resource allocation.

15. **Evaluation**: Establish a process for ongoing evaluation to track progress toward achieving the objectives and make necessary adjustments.

**Data Visualization Using IBM Cognos:**

Using IBM Cognos for data visualization in public awareness campaigns can be a powerful tool to convey information effectively and engage your audience. Here's how you can utilize IBM Cognos for this purpose:

1. **Data Integration**: Integrate relevant data from various sources, such as surveys, government databases, or social media analytics, into IBM Cognos.

2. **Dashboard Creation**: Develop interactive and user-friendly dashboards that display key data points, trends, and insights.

3. **Customization**: Tailor the visualizations to cater to your specific public awareness objectives and target audience.

4. **Geospatial Visualizations**: Utilize maps and geospatial data to illustrate regional disparities or specific issues in different areas.

5. **Time-Series Analysis**: Create time-based visualizations to show the evolution of data over time, helping the audience understand trends and changes.

6. **Infographics**: Use charts, graphs, and infographics to simplify complex data and make it more accessible to the public.

7. **Interactive Elements**: Incorporate interactive features like drill-down options and filters to allow users to explore the data on their own.

8. **Storytelling**: Build a narrative around the data, explaining its relevance and impact in the context of your awareness campaign.

9. **Data Sharing**: Easily share visualizations through various channels, including websites, social media, or presentations.

10. **Mobile Compatibility**: Ensure that your visualizations are mobile-friendly, making them accessible to a wider audience.

11. **Data Security**: Implement security measures to protect sensitive data and ensure compliance with data protection regulations.

12. **Real-Time Data**: If applicable, integrate real-time data to keep the public informed about the latest developments.

13. **Data Insights**: Leverage Cognos' analytics capabilities to extract actionable insights from the data, guiding decision-making in your campaign.

14. **User Training**: Provide training to users or stakeholders on how to interact with and interpret the visualizations effectively.

15. **Evaluation**: Continuously monitor the performance and reception of your visualizations, making adjustments as necessary to enhance the impact of your public awareness efforts.

**Predictive Modelling for public awareness:**

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.

**Data Collection Process:**

1. **Define Objectives**: Clearly outline the objectives of your public awareness campaign to determine the type of data needed.

2. **Identify Target Audience**: Define the specific demographic and psychographic characteristics of the audience you want to reach.

3. **Select Data Sources**:

a. **Surveys and Questionnaires**: Design surveys to gather information directly from your audience. Use online or paper surveys, phone interviews, or in-person questionnaires

b. **Existing Data**: Utilize available data sources, such as government statistics, research studies, or industry reports, to complement your own data collection efforts.

c. **Social Media**: Analyze social media platforms to gather insights on audience sentiment, engagement, and trending topics.

4. **Data Collection Tools**:

a. **Online Survey Platforms**: Use tools like SurveyMonkey, Google Forms, or Qualtrics to create and distribute surveys.

b. **Social Media Analytics**: Platforms like Facebook Insights, Twitter Analytics, and Google Analytics can provide data on social media engagement and website traffic.

c. **Data Collection Apps**: Mobile apps can be used for on-site data collection, interviews, and observations.

5. **Data Privacy and Ethics**: Ensure compliance with data privacy regulations and ethical standards. Obtain informed consent where necessary and protect the anonymity of respondents.

6. **Data Analysis**: Process and analyze the collected data to extract actionable insights and identify trends.

**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):**

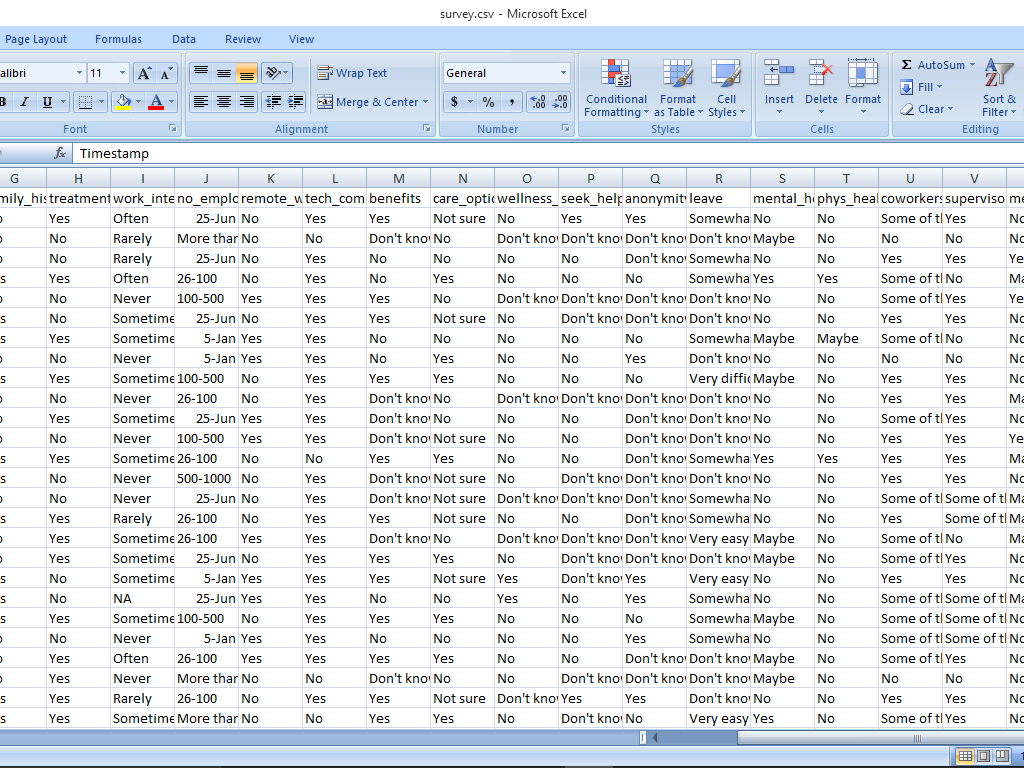
* If you plan to deploy the model, save it and build an API or interface for integration into your application.

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 PUBLIC HEALTH AWARENESS**

****

**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()

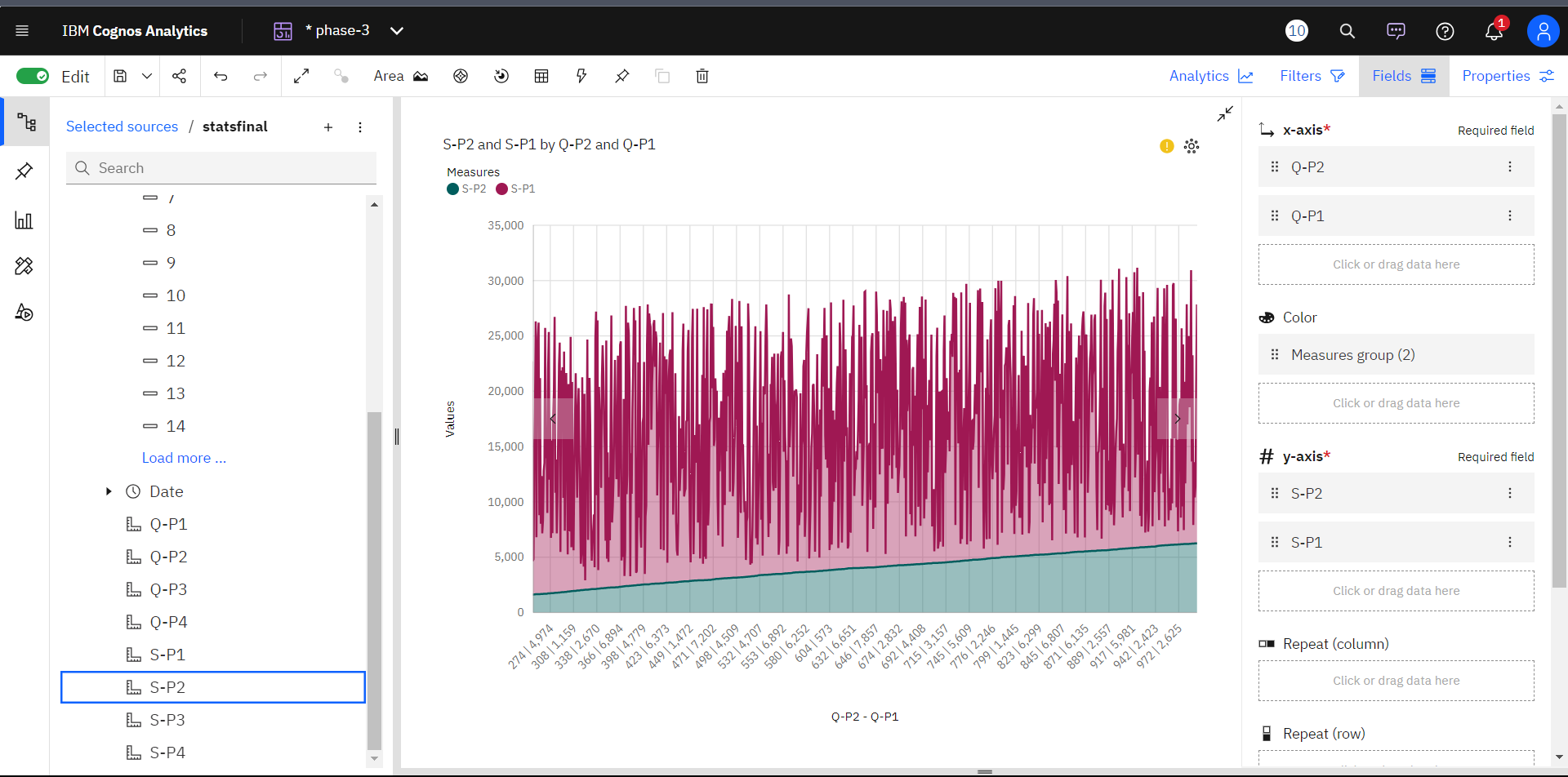
plt.xlabel('Predicted')

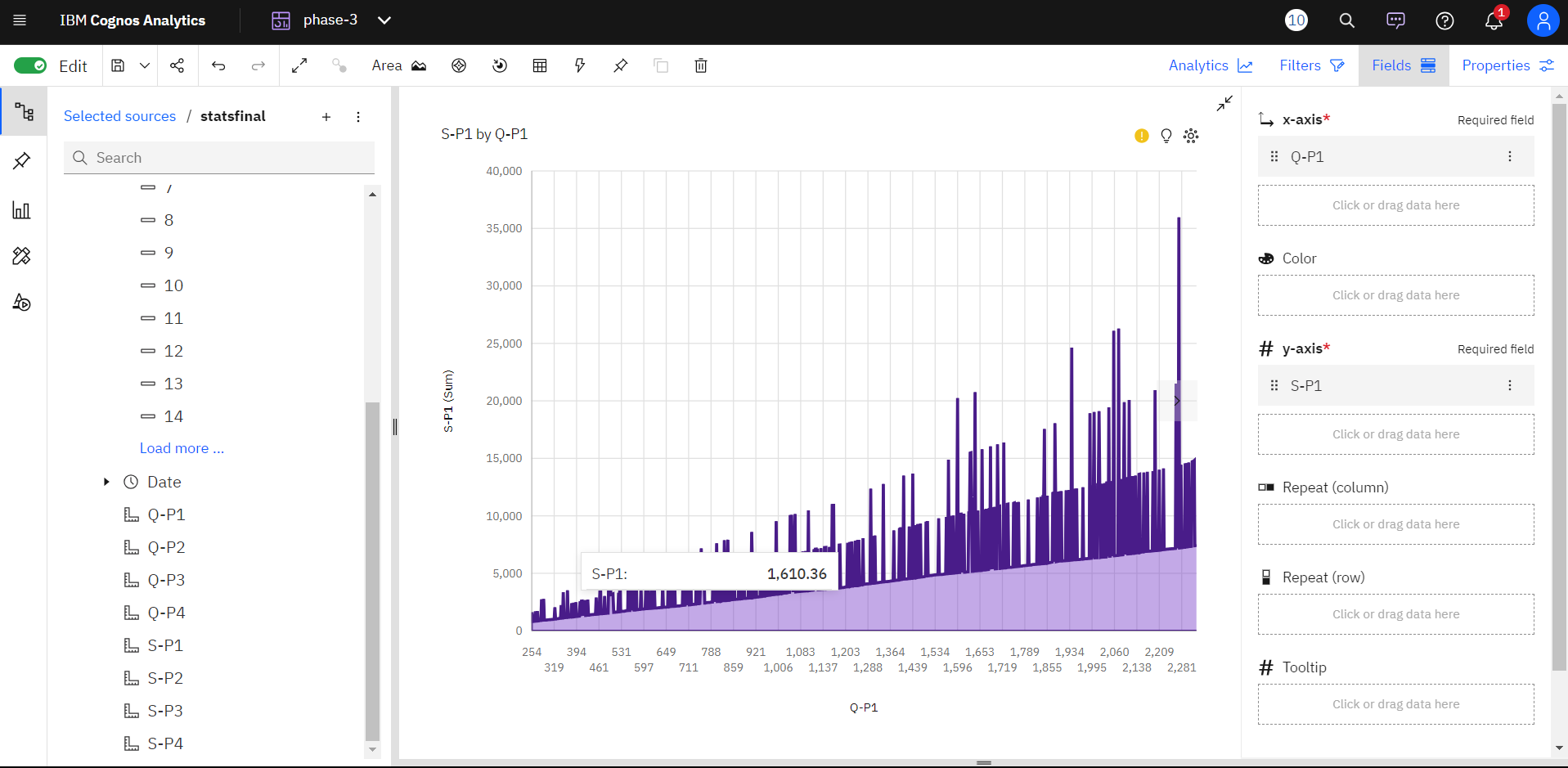
plt.ylabel('Actual')

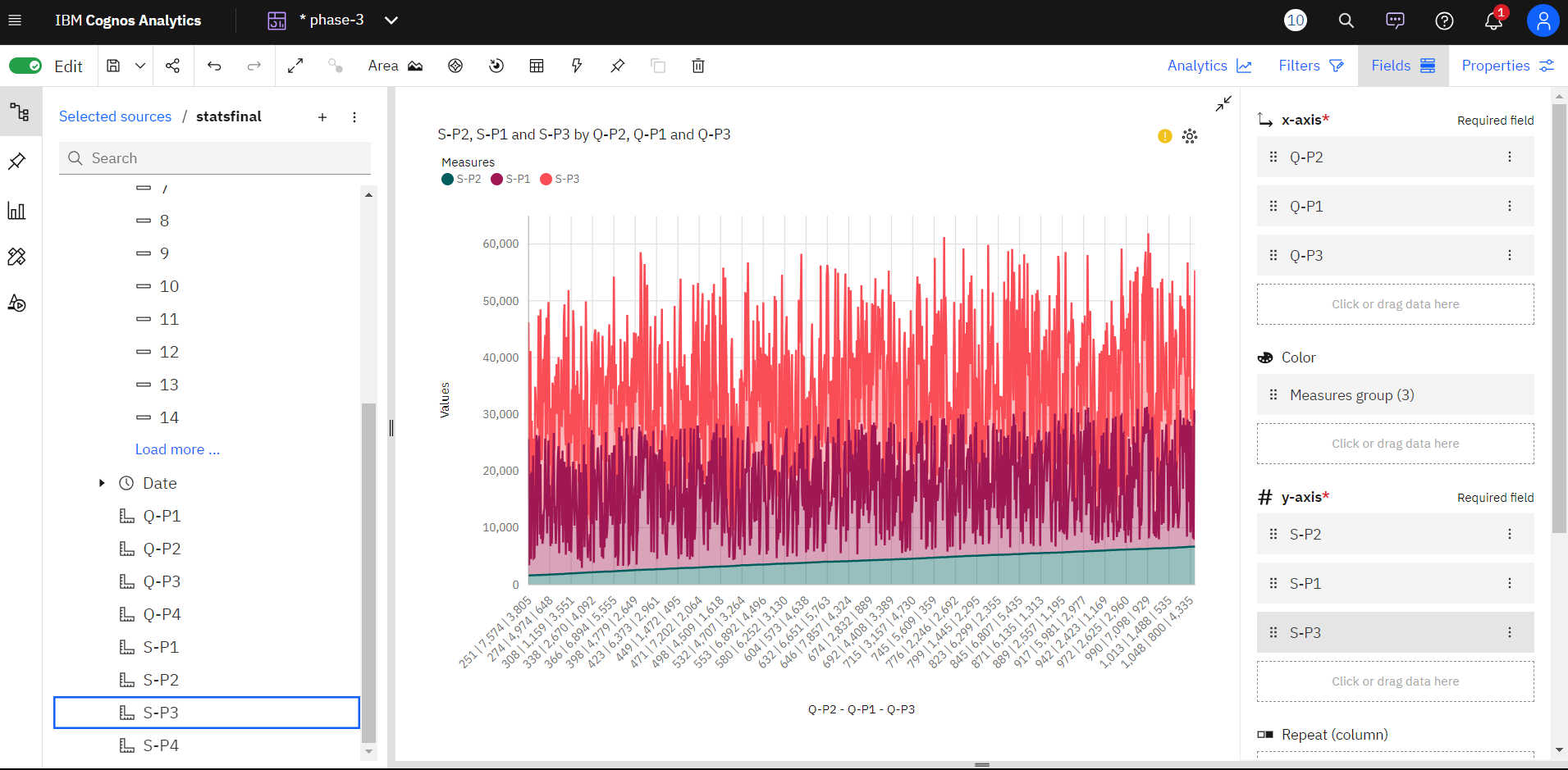
plt.title('Confusion Matrix')

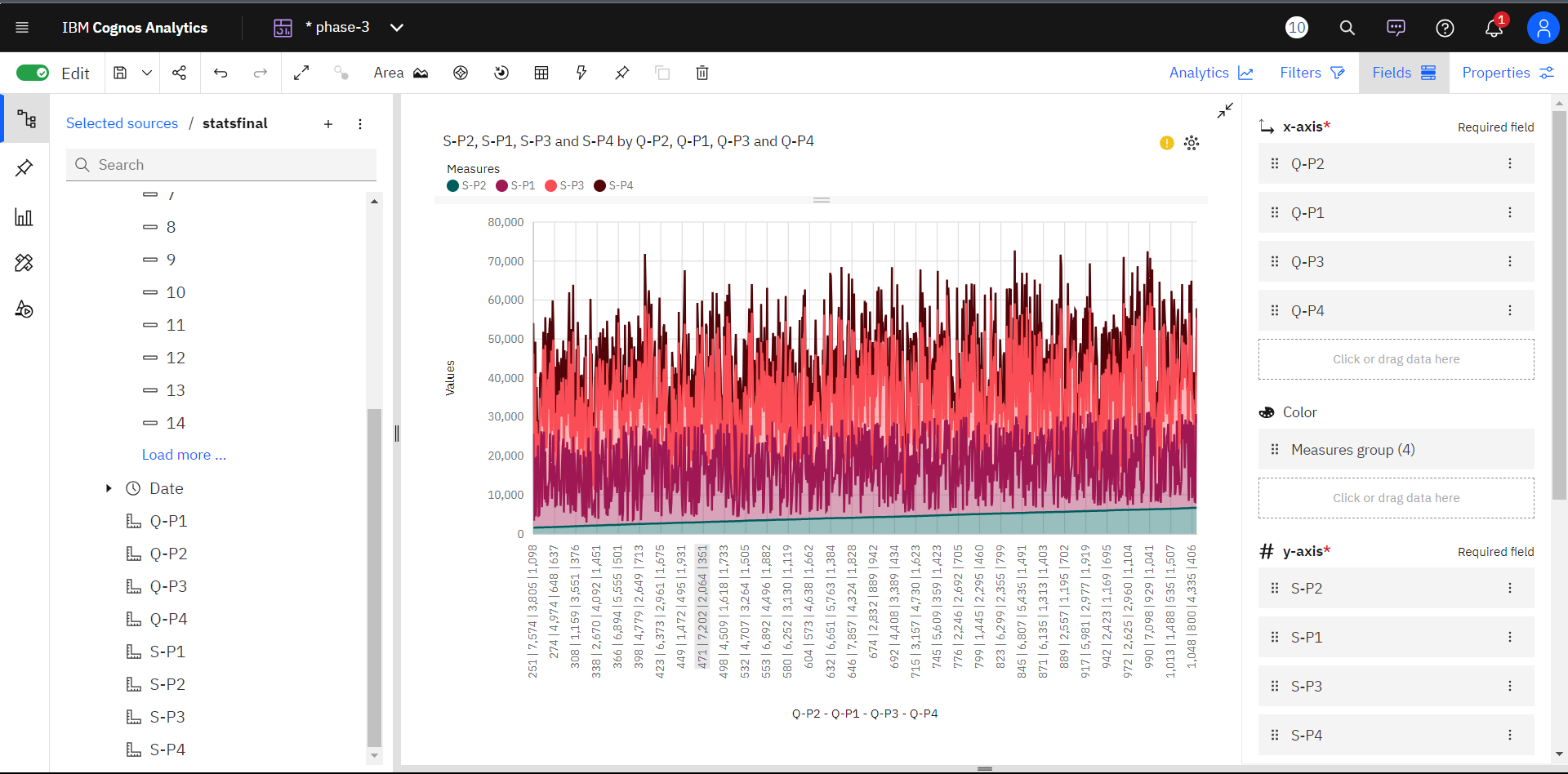
plt.show()

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



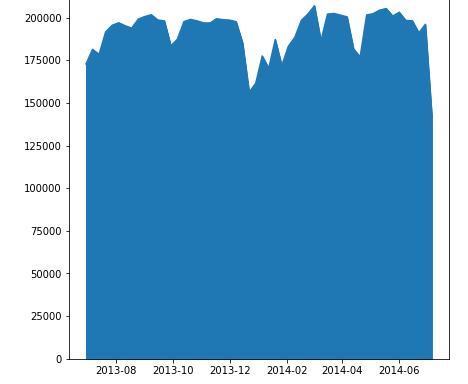


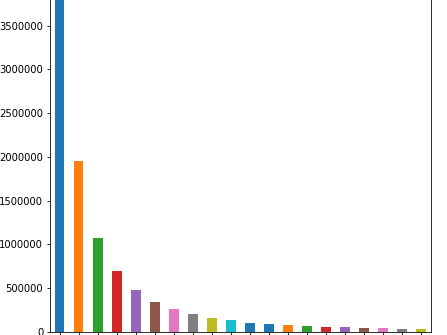


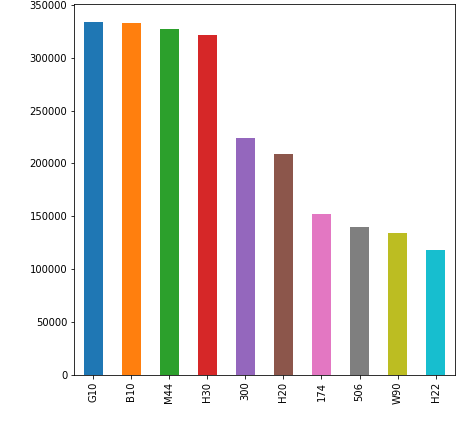


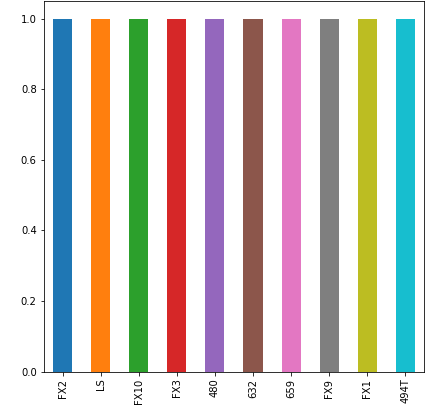
PROJECT REPORT: PUBLIC HEALTH AWARENESS

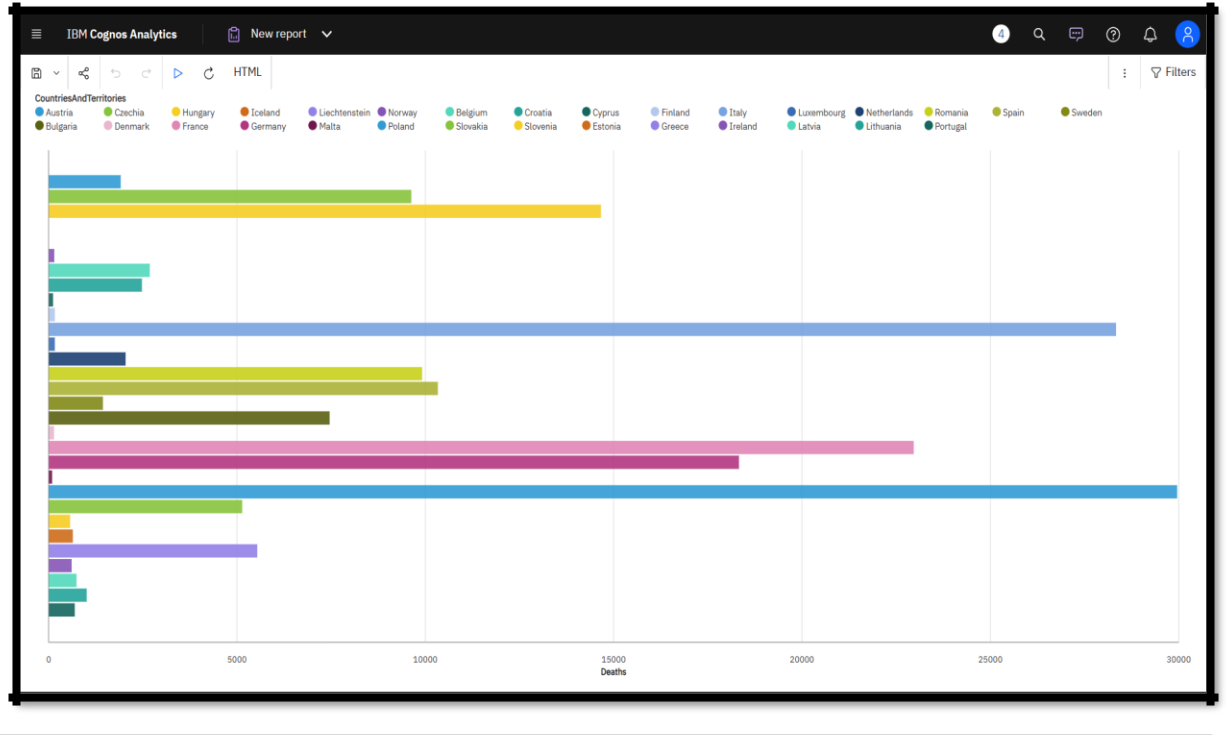
1. Number of people using telecommunication



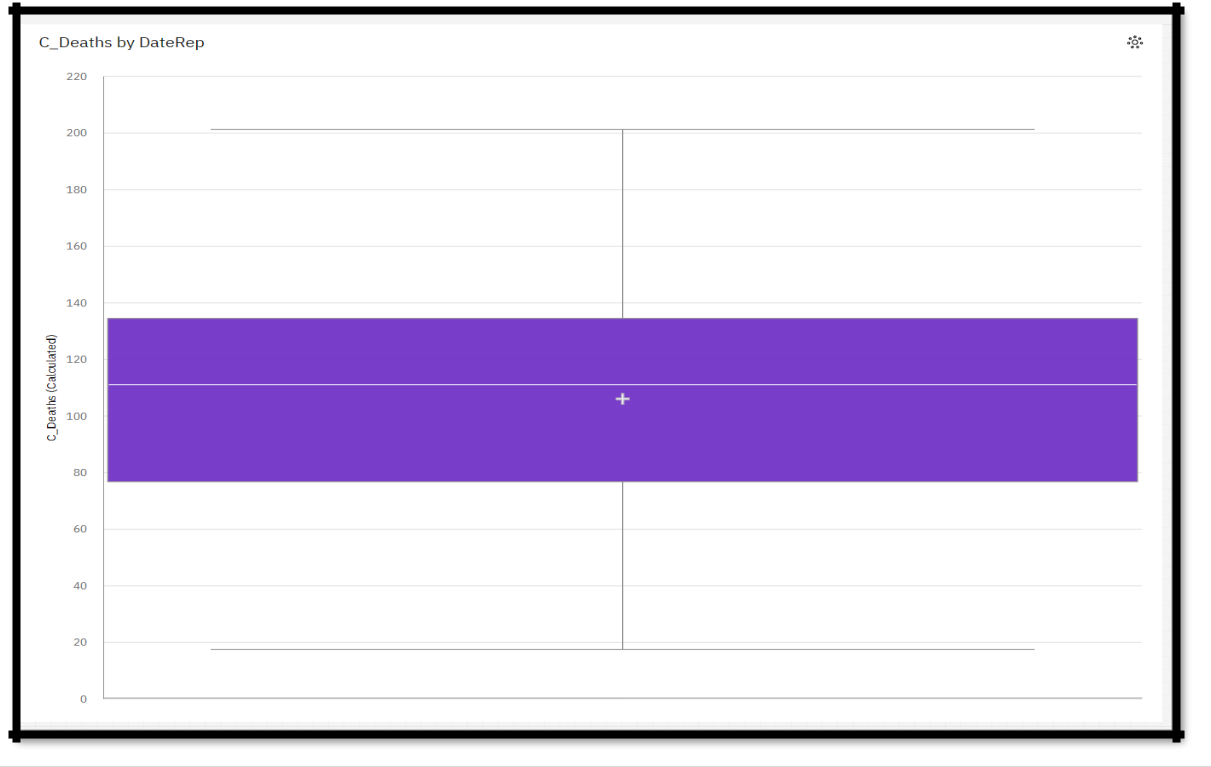


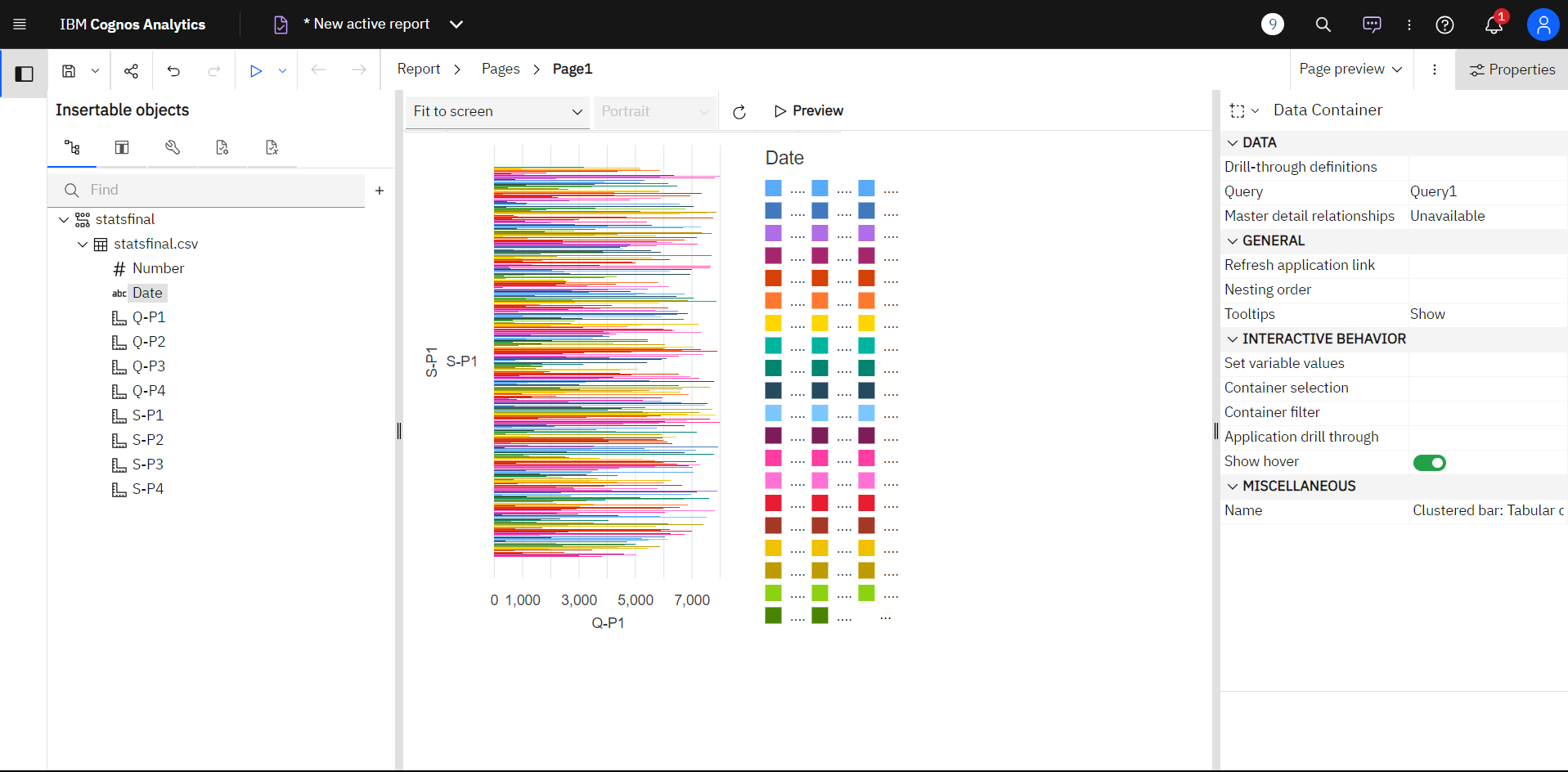




****

****

****



**Conclusions:**

The project data analysis using IBM Cognos provided valuable insights into the public health awareness. In conclusion, promoting public transport analysis is vital. Through education, outreach, and community engagement, we can empower individuals to make informed choices, improve well-being, and create healthier, more resilient communities.