**PUBLIC TRANSPORT EFFICIENCY ANALYSIS**

**Introduction:**

In this analysis , we analyse the given CSV file with the power of IBM Cognos for insightful virtualization. Through data driven exploration is vital for promoting well-being and preventing diseases. Timely awareness campaigns can save lives and reduce healthcare burdens. Governments and organizations play a crucial role in fostering public health awareness.

**Step 1:**Define and analyze the dataset.

**Step 2:**Obtain the sales product analysis data.

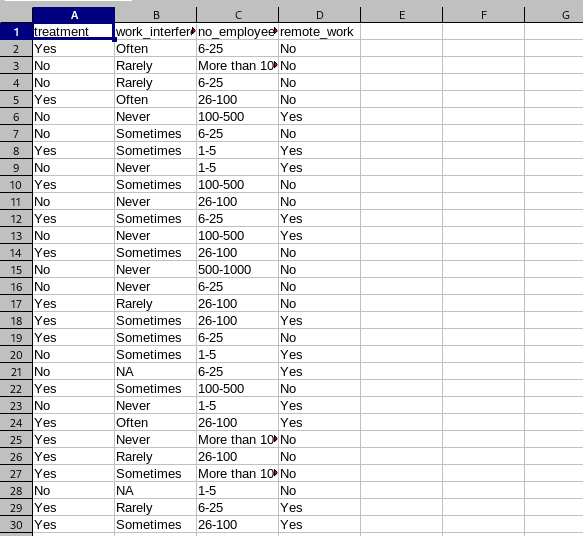
**Step 3:**Clean and Preprocess the give data.

**Step 4:** Load the data into IBM Cognos

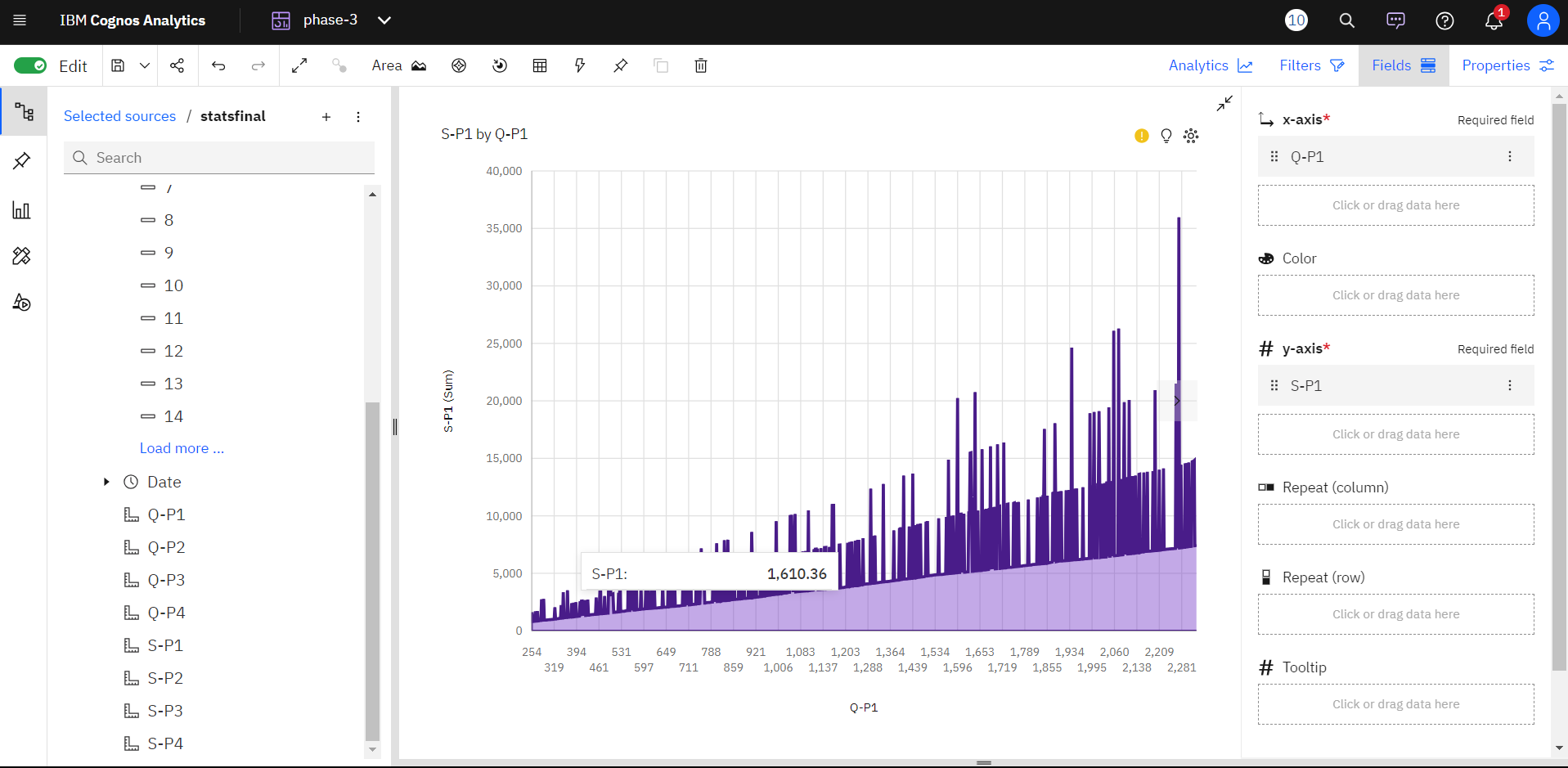
**Step 5:**Build the Visualizations.

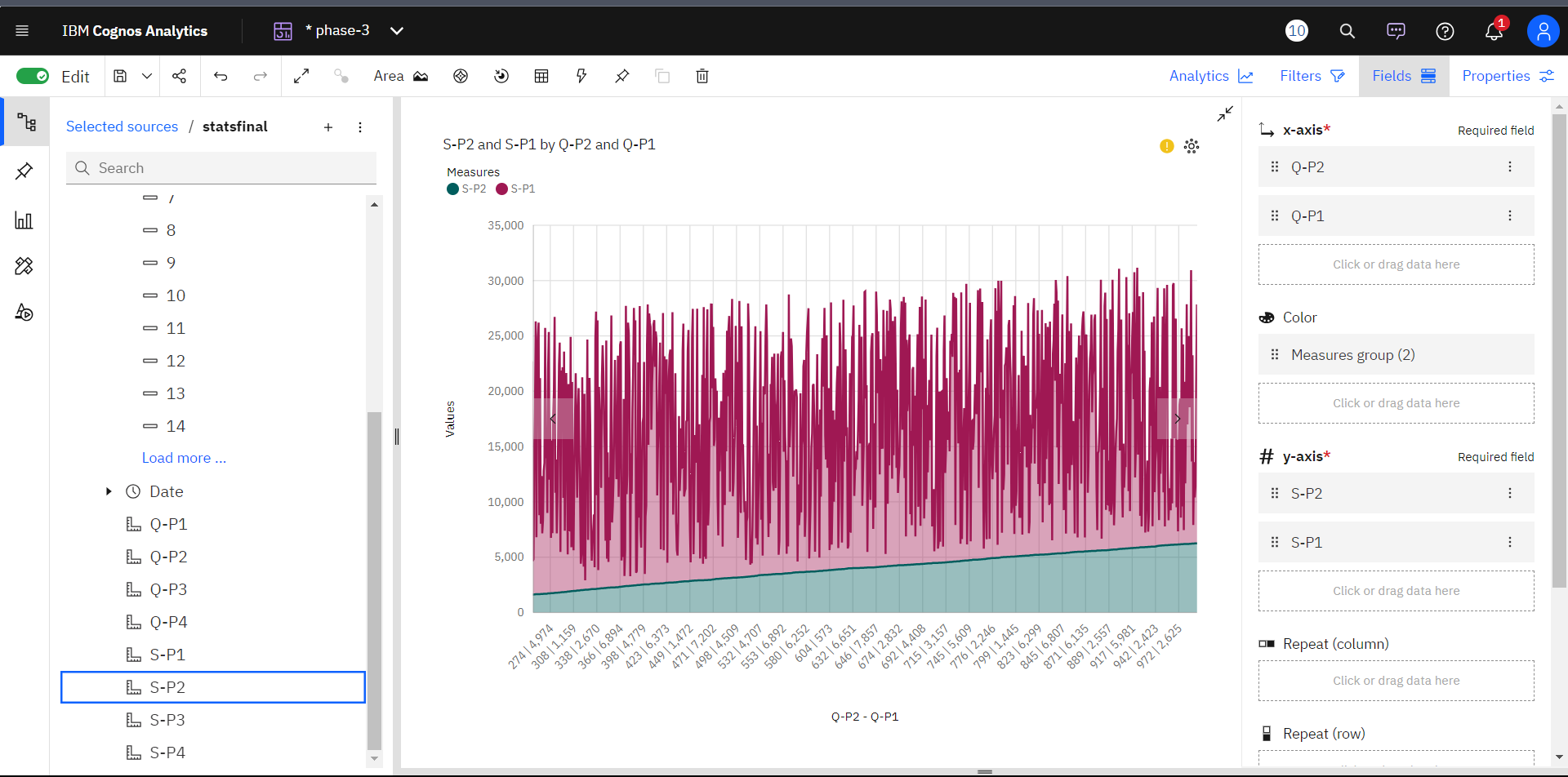
**Step 6:**Analyze and share the data.

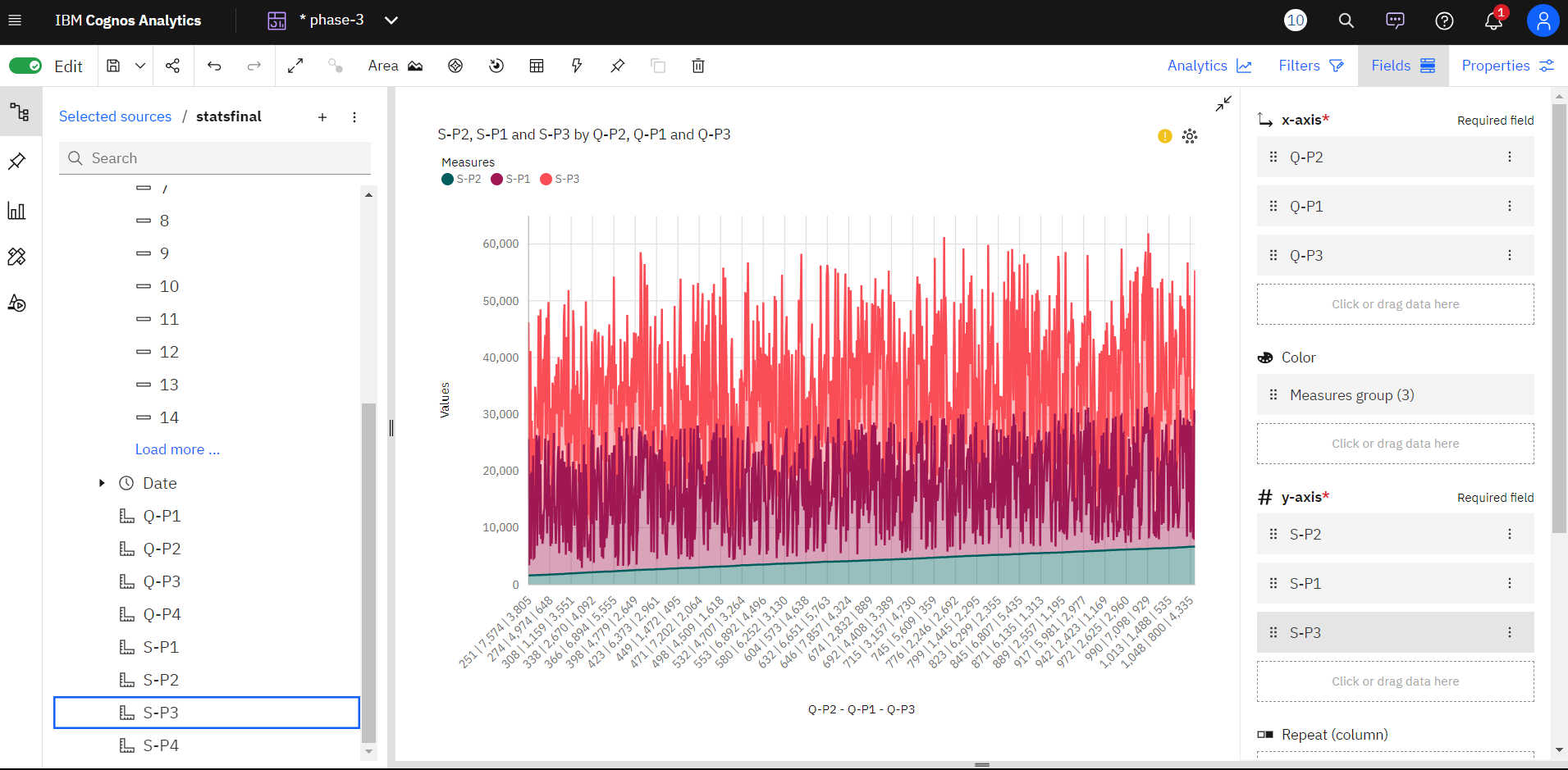
* In this session we have used Microsoft Excel to clean the data.

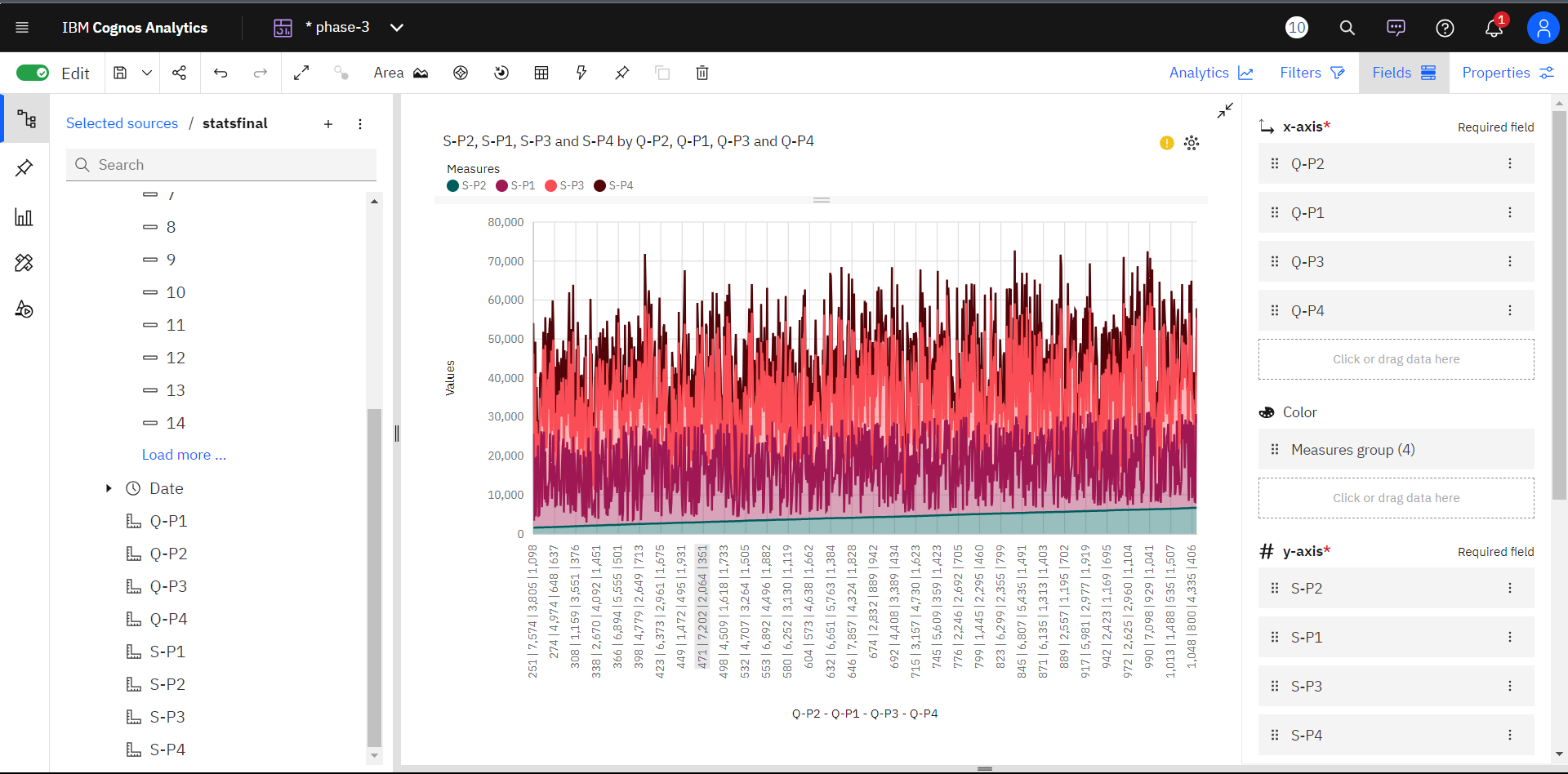


**Data Visualization And Analysis:**

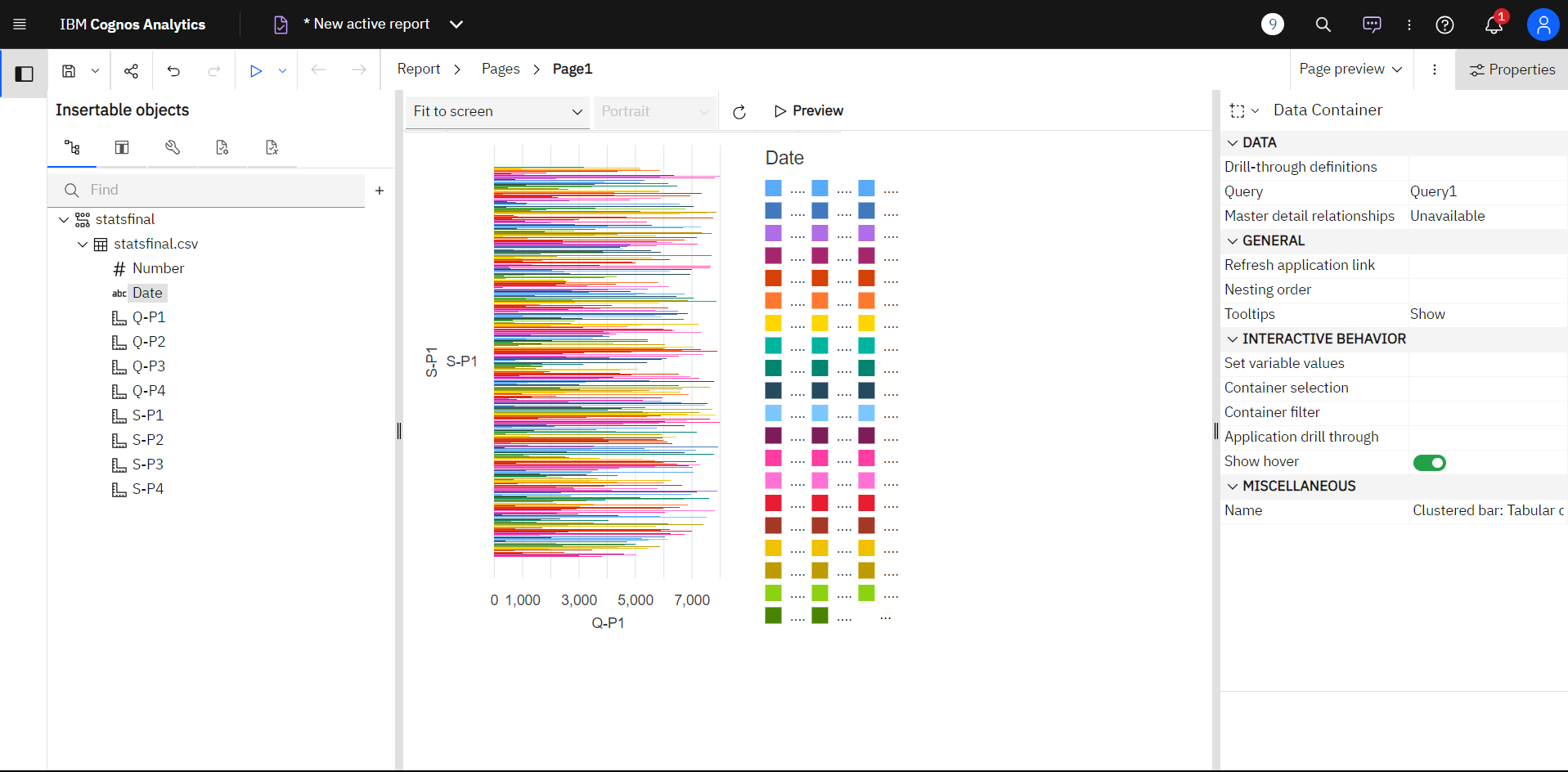








**Drilling Down For Deeper Insights**



**Code For The Data Cleaning And Preprocessing.**

import pandas as pd

import numpy as np

from scipy import stats

# Load the data

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

# Check for missing values

missing\_values = data.isnull().sum()

# Handle missing values by filling with the mean

data.fillna(data.mean(), inplace=True)

# Remove duplicate rows

data.drop\_duplicates(inplace=True)

# Convert data types if needed

data['column\_name'] = data['column\_name'].astype('desired\_dtype')

# Rename columns

data.rename(columns={'old\_name': 'new\_name'}, inplace=True)

# Clean text data

data['text\_column'] = data['text\_column'].str.strip().str.lower()

# Remove outliers using z-scores for a numerical column

z\_scores = np.abs(stats.zscore(data['numerical\_column']))

data = data[(z\_scores < 3)]

# Parse and format date/time columns

data['date\_column'] = pd.to\_datetime(data['date\_column'], format='%Y-%m-%d')

# Save cleaned data to a new CSV file

data.to\_csv('cleaned\_data.csv', index=False)

# Example analysis: Calculate the average age of survey respondents

average\_age = df['Age'].mean()

print("Average Age:", average\_age)

# Example analysis: Count the number of respondents in different age groups

age\_groups = pd.cut(df['Age'], bins=[0, 18, 35, 60, float('inf')], labels=['0-18', '19-35', '36-60', '60+'])

age\_group\_counts = age\_groups.value\_counts()

print("Age Group Counts:")

print(age\_group\_counts)

# Example analysis: Identify common health concerns

common\_health\_concerns = df['Health\_Concern'].value\_counts()

print("Common Health Concerns:")

print(common\_health\_concerns)

# Add more analysis as needed

# Save the results to a new CSV file

output\_file\_path = "/home/vignesh/Downloads/survey\_analysis.csv"

df.to\_csv(output\_file\_path, index=False)

print(f"Analysis results saved to {output\_file\_path}")

**conclusion:**

The use of IBM Cognos in Public transport analysis has proven instrumental in extracting actionable insights. By visualizing and interpreting the data, we have identified key trends, opportunities, and areas for improvement. This analysis equips us with the knowledge to make informed decisions, enhance sales strategies, and ultimately drive the success of our product in the market.