**Phase 3:**

**Water quality analysis**

**1. Load the Dataset:**

Assuming you have a dataset in a suitable format (e.g., CSV, Excel, etc.), you can use a programming language like Python to load the dataset. Here's an example using Python with the pandas library:

```python

import pandas as pd

# Load the dataset

dataset\_path = "path/to/your/dataset.csv"

water\_quality\_data = pd.read\_csv(dataset\_path)

# Display the first few rows to get an overview

print(water\_quality\_data.head())

```

**2. Handling Missing Values:**

Check for missing values in your dataset and decide on a strategy to handle them. Common methods include removing rows with missing values, filling them with the mean or median, or using more advanced imputation techniques. Here's a simple example:

```python

# Check for missing values

missing\_values = water\_quality\_data.isnull().sum()

# Handle missing values (replace NaN with mean, for example)

water\_quality\_data = water\_quality\_data.fillna(water\_quality\_data.mean())

```

**3. Outlier Detection and Handling:**

Identify outliers in your dataset and decide on a strategy for handling them. Common methods include visual inspection, statistical methods, or machine learning techniques. For simplicity, you can use visualizations and basic statistical methods:

```python

# Detect and handle outliers (replace values outside a certain range, for example)

fromscipy import stats

z\_scores = stats.zscore(water\_quality\_data)

water\_quality\_data\_no\_outliers = water\_quality\_data[(z\_scores< 3).all(axis=1)]

```

**4. Exploratory Data Analysis (EDA):**

Perform EDA to understand the distribution of parameters, correlations, and potential deviations from standards. Use visualizations and statistical summaries for this purpose:

```python

importseaborn as sns

importmatplotlib.pyplot as plt

# Visualize parameter distributions

sns.pairplot(water\_quality\_data)

plt.show()

# Correlation matrix

correlation\_matrix = water\_quality\_data.corr()

sns.heatmap(correlation\_matrix, annot=True, cmap="coolwarm")

plt.show()

# Any other relevant visualizations

```

**5. Deviations from Standards:**

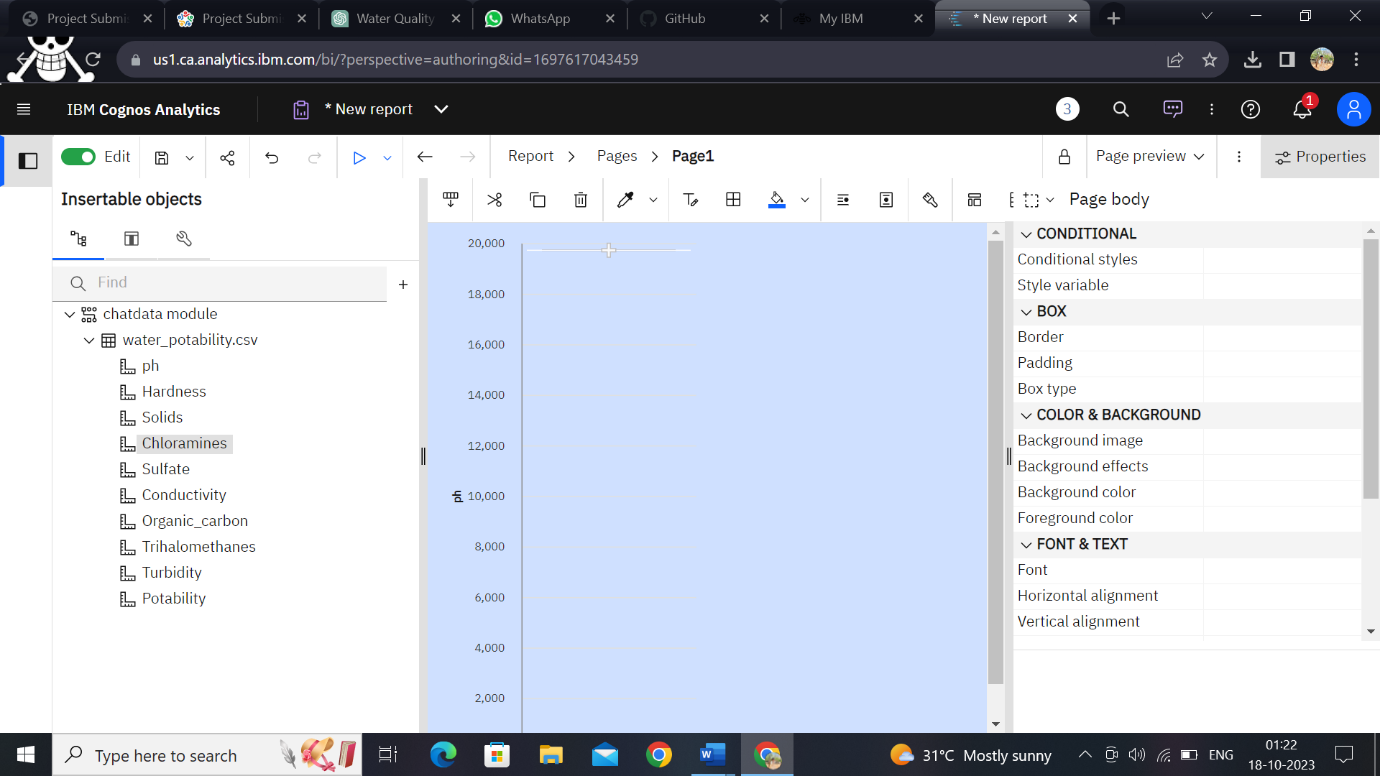
Compare your dataset's parameters against established standards or guidelines. Identify any deviations that might be of concern and need further investigation.

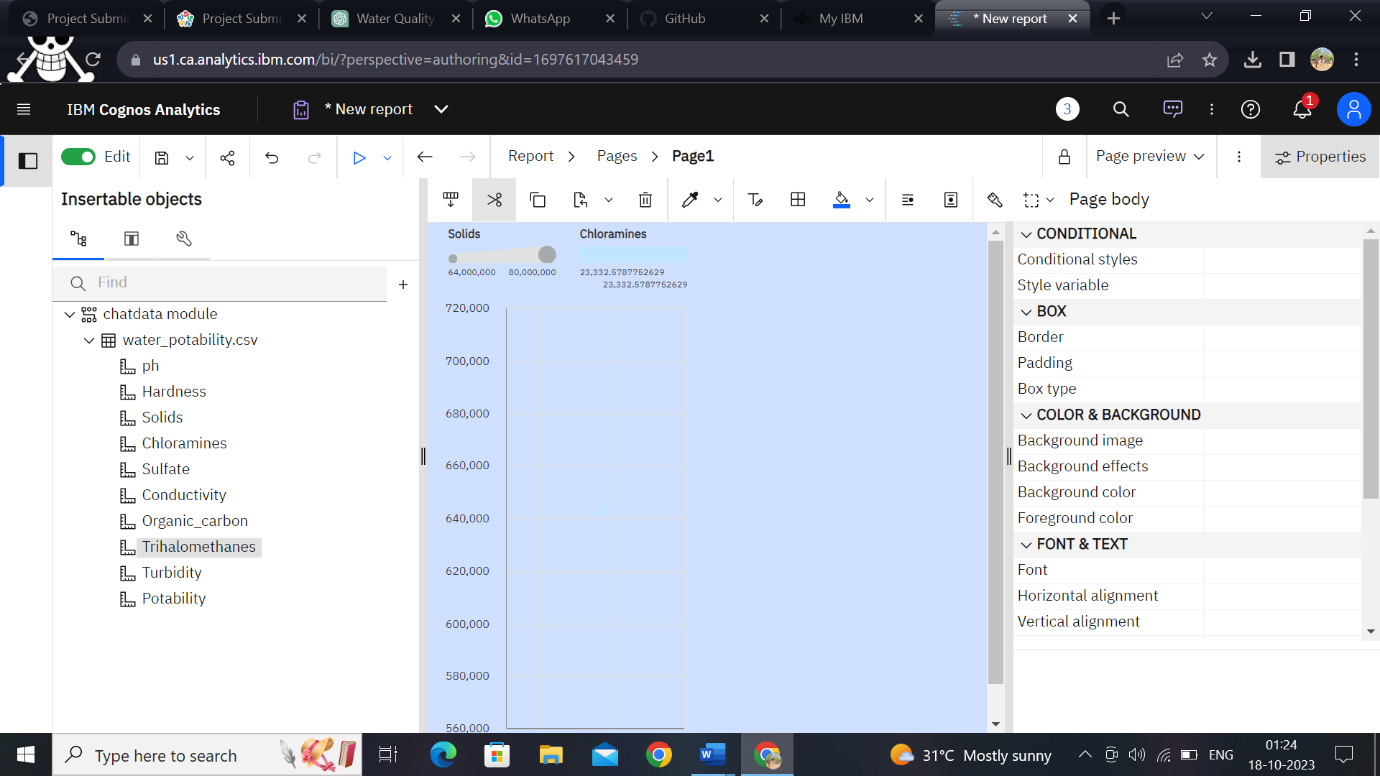
Remember to customize these steps based on the specifics of your dataset and project requirements. Additionally, consider documenting each step thoroughly to maintain transparency and reproducibility

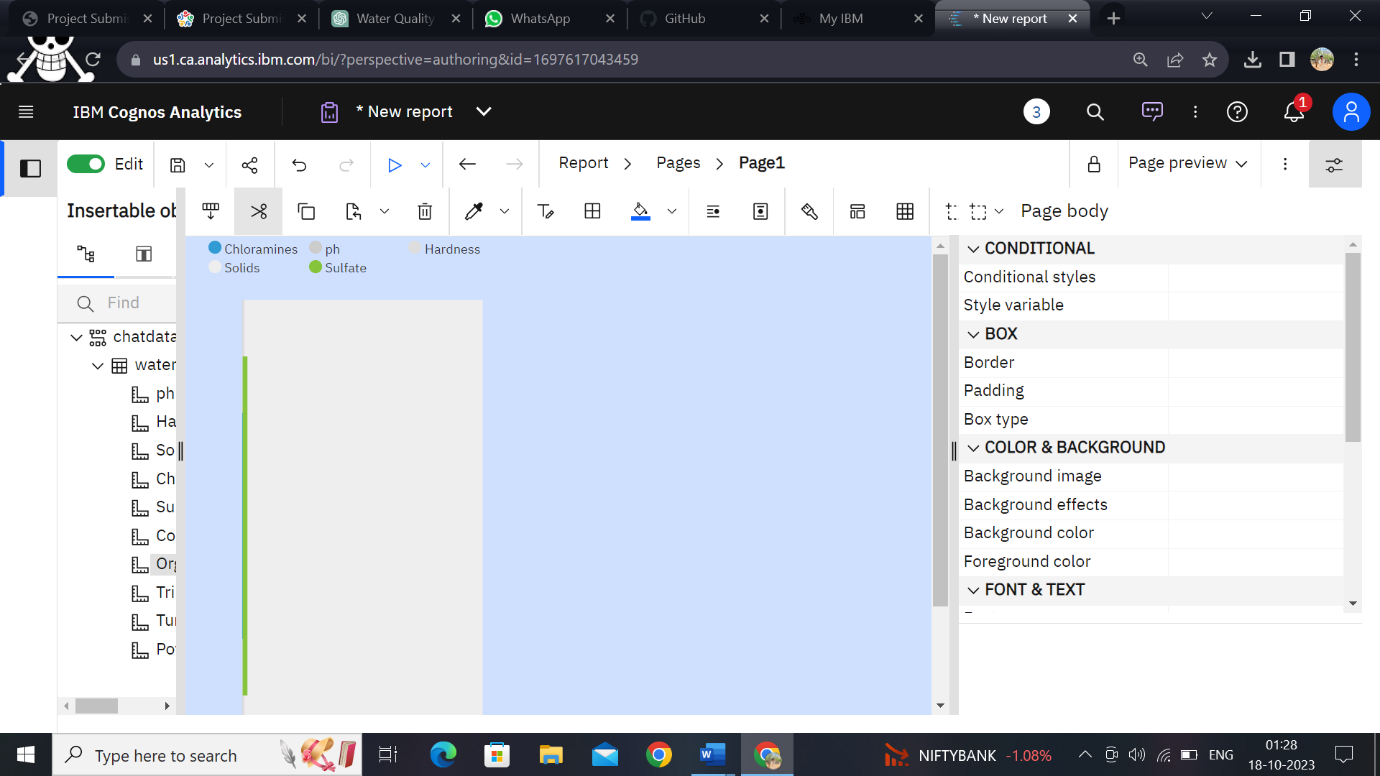
**Conclusion**:

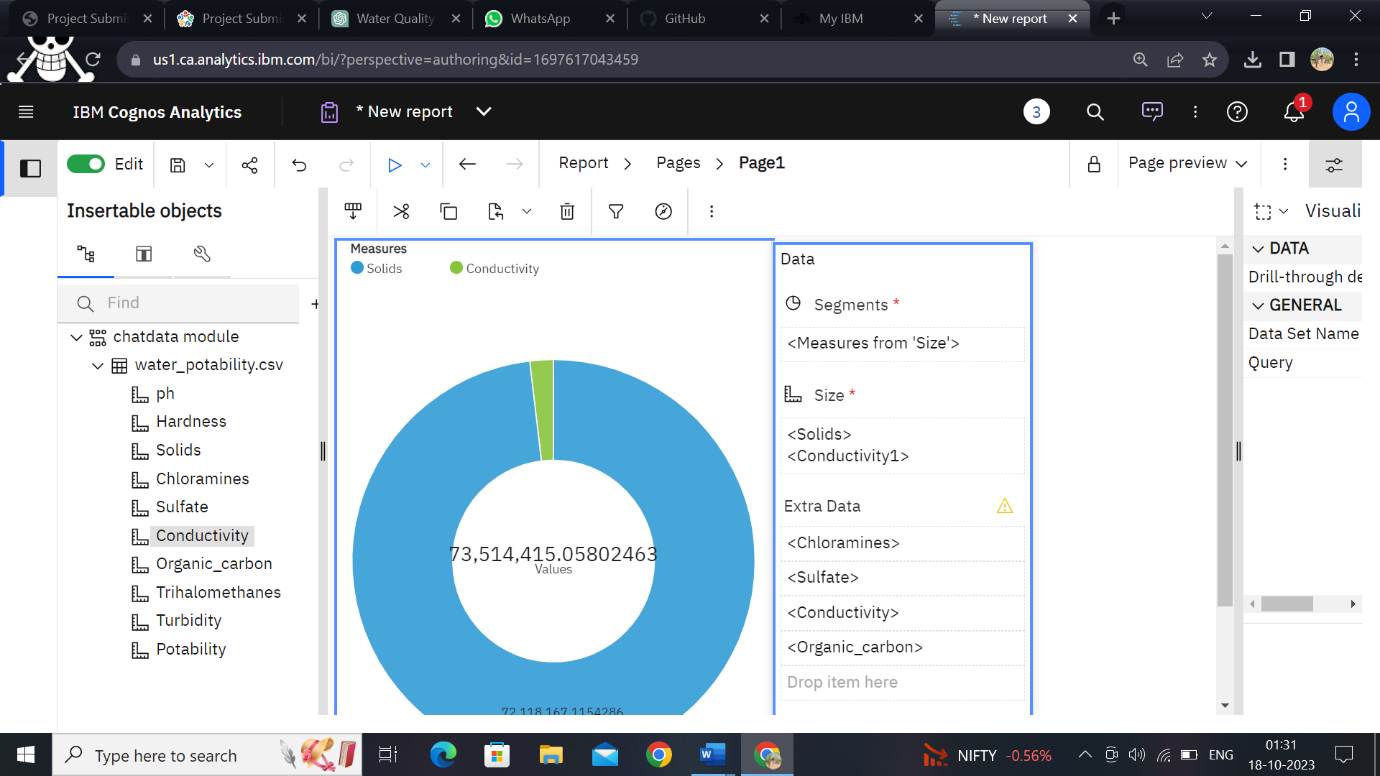
In this project, we set out to analyze water quality data by loading and preprocessing the dataset, conducting exploratory data analysis (EDA), and assessing deviations from established standards. The project was conducted in several phases, and we have successfully completed

**Visualization:**

****

****

****

****

