The matrix followed across the documents was years and countries. We picked datasets that fit these conditions the best.

The chosen datasets:

1. Annual Freshwater Withdrawals: The dataset was chosen because it has most of the data for many countries and not much missing information.
2. Premature deaths due to UNSAFE\_WASH: This dataset has been included in our selection due to its commendable attributes, featuring restricted gaps in data and widespread availability across numerous countries. This dataset not only contributes valuable insights but also aligns with the common countries prevalent throughout the document.
3. Proportion of bodies of water with good ambient water quality: The dataset is in our selection despite having data for only one year. This decision was made due to its significance and adequacy in relation to the specified years and countries covered in the document.
4. Wastewater Discharges Per Year Per Country: This dataset is crucial in our selection. It has good coverage, not much missing data, and gives us a lot of useful information. We decided to pick common countries based on this dataset.

Datasets not used:

1. Generation and discharge of wastewater in volume: Even though it had a lot of information, it did not include the countries we commonly found in the document.
2. Mortality rate attributed to unsafe water, unsafe sanitation, and lack of hygiene (per 100,000 population): This dataset was left out because it had limited and not very good data. It only covered one country, and there were many missing values in the data.