

Executive Summary

2024 Metro Water Recovery – Our Communities Dashboard: Statistical Analysis, Sentiment Analysis, and Community Awareness Survey

This executive summary provides the results of and insights into the statistical and sentiment analysis methodologies applied to the Post Rebranding Community Awareness Survey (PRCAS) conducted by Metro Water Recovery in late 2021. The survey, administered by the now acquired 9thWonder, aimed to assess community awareness in the Denver Metro Area.

Survey Demographics

The PRCAS gathered responses from 500 Colorado residents, focusing on non-demographic questions related to organizational and industry awareness, communication preferences, and more.

Confidence Interval Methodology

Using the Wilson Score Interval with Continuity Correction, and weighted adjustments to demographic discrepancies between the survey and reported Denver Metro demographics released by the US Census Bureau, a confidence interval was derived from respondent answers to PRCAS Question 16: “Prior to today, had you heard of Metro Water Recovery?” Results indicate that with 99% confidence the Denver Metro Area awareness of Metro Water Recovery is $14.88\% \pm 4.15\%$.

Sentiment Analysis Methodology

Survey responses to PRCAS Question 25: “Based on what you read or what you know, what 3 words come to mind when you hear Metro Water Recovery?” were used to perform the Sentiment Analysis reported in the Our Communities dashboard. Responses were categorized as Positive, Neutral, or Negative. Using a combination of the “Azure Machine Learning” add-on in Microsoft Excel and a manual review, the analysis revealed 427 positive, 29 neutral, and 44 negative sentiments among the 500 responses. These results indicate that Metro Water Recovery is generally viewed in a positive light.

Survey & Analysis Limitations

Statistical limitations include uncertainties about 9thWonder’s sampling methods, the assumption of random sampling for the confidence interval, a skewed binomial data distribution, a demographic weighting assumption that all demographics are equally influential, and whether the demographic weights are necessary at all.

Conclusion

Despite some limitations, the confidence interval and Sentiment Analysis offer valuable insights into a more nuanced understanding of community awareness and sentiment regarding Metro Water Recovery in the Denver Metro Area. Future analyses would benefit from a survey partner with transparent survey sampling methodologies, additional assessments of the significance of demographic variations, and an exploration of alternative statistical methodologies. In conclusion, while the analysis provides valuable insights, the highlighted limitations underscore the complexity

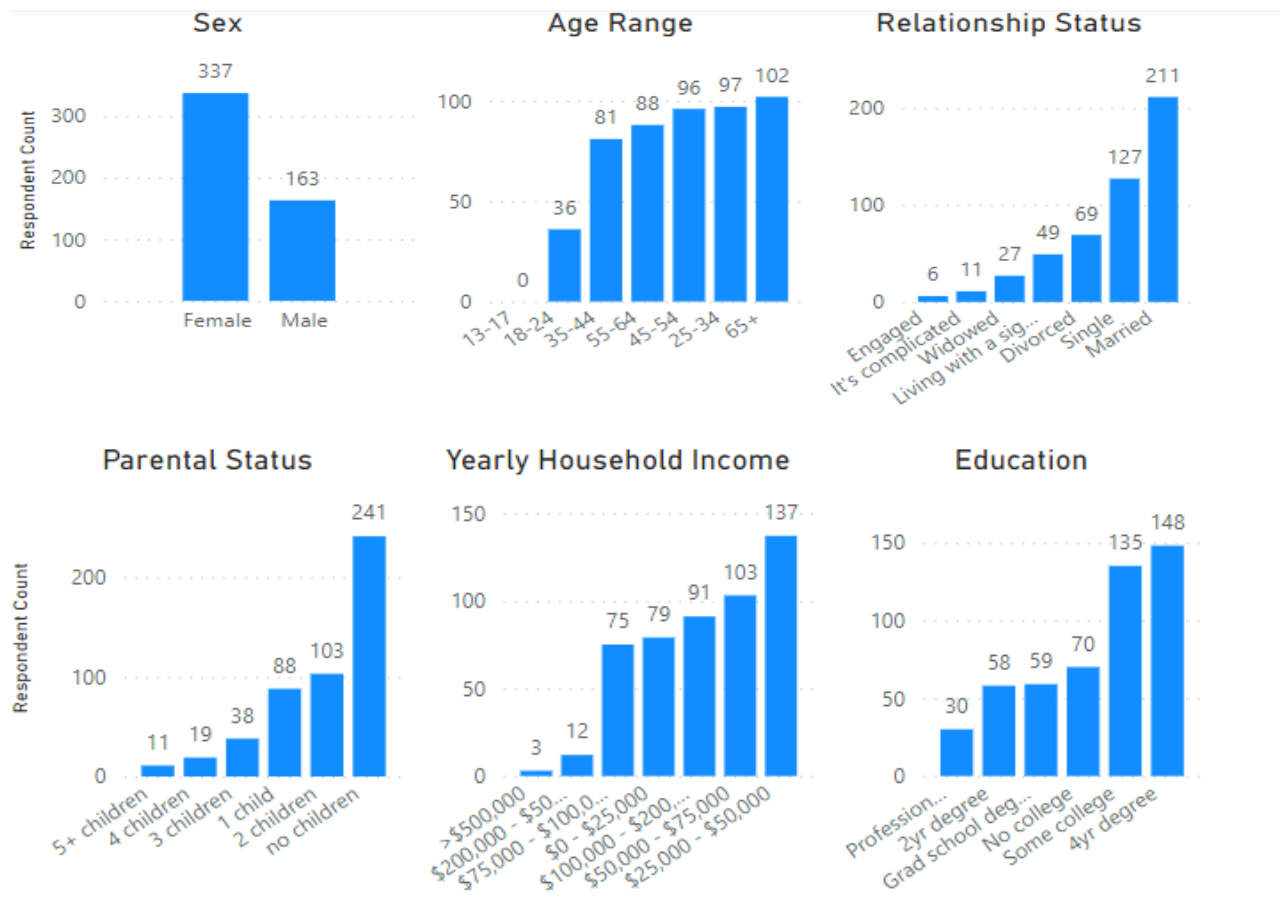
of deriving precise and universally applicable conclusions from survey data. Additional details can be found in the following pages.

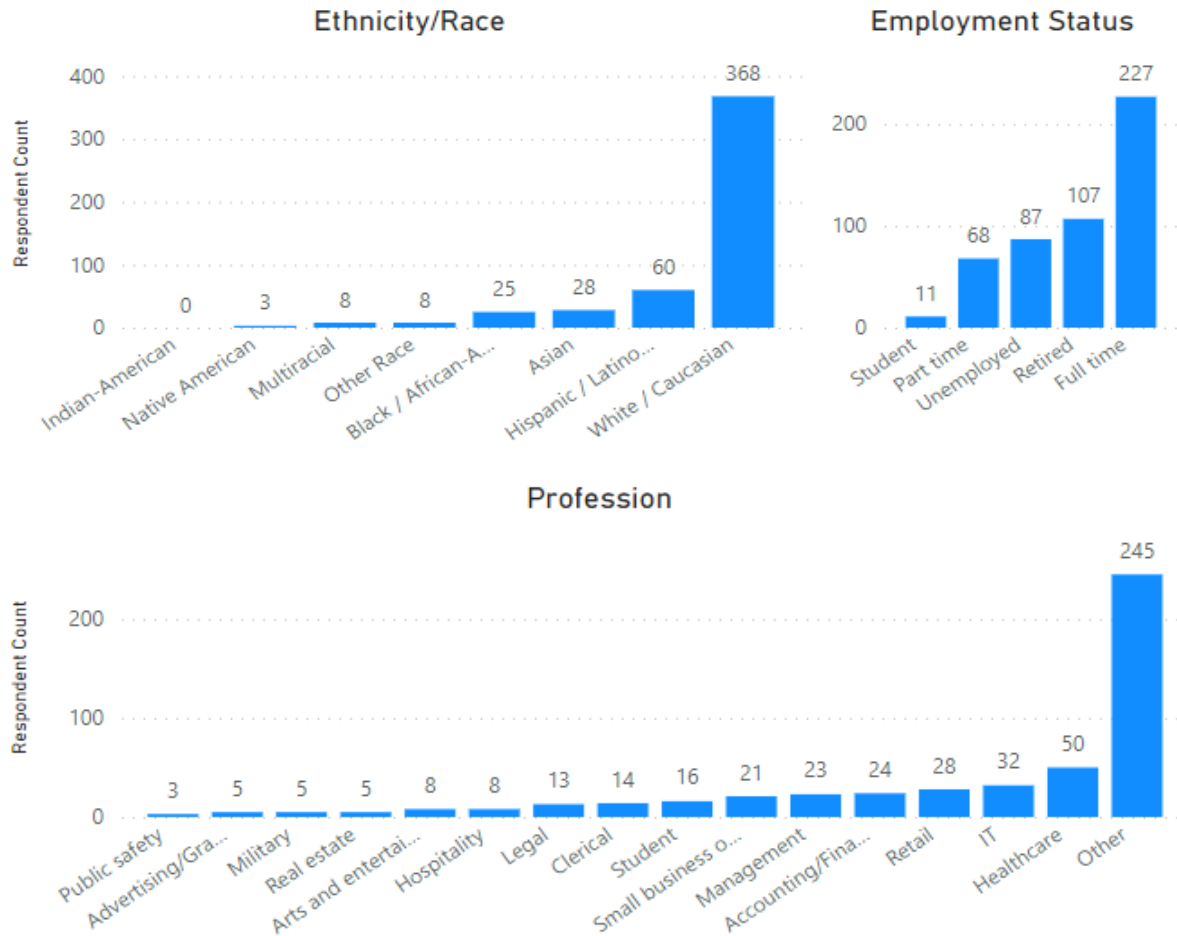
2024 Metro Water Recovery - Our Communities Dashboard – Statistical & Sentiment Analysis
Methodology and Community Awareness Survey Demographics

Community Awareness Survey Information & Demographics

Both the reported Confidence Interval and Sentiment Analysis were developed using the results of a Post Rebranding Community Awareness Survey (PRCAS) conducted by Metro Water Recovery late in 2021. Metro Water Recovery utilized a third-party marketing survey vendor called “9thWonder” to both develop and distribute the survey. However, as of 2023, 9thWonder was acquired by Gravity Global¹ and no longer exists. Survey results were then released to Metro Water Recovery in early 2022. It should be noted the survey contained a limitation that could have affected the analysis results. This limitation can be found in the Survey & Analysis Limitations Section of this document.

Overall, the survey consists of 500 respondents, all of whom are from Colorado; and a total of 29 non-demographic related questions. Every respondent reported from the Denver-Aurora Combined Statistical Area (CSA). Additional respondent demographics can be found in the visualizations below. Lastly, survey questions focused on topics ranging from organizational awareness to industry awareness to preferences around communications.





Confidence Interval Methodology

In short, leadership within the Strategy & Communications department at Metro Water Recovery posed the question, “Given the results of the Post Rebranding Community Awareness Survey, what proportion of our service area do we believe is actually aware of who we are?” In other words, could Metro Water Recovery make a statistically backed inference regarding what proportion of the Denver Metro Area is aware of them by utilizing the results from the PRCAS?

Proportional respondent awareness of Metro Water Recovery was deemed determinable from PRCAS question 16, which asks, “Prior to today, had you heard of Metro Water Recovery?” This question could take on two answers, ‘Yes’ or ‘No’, thus making inferences from it subject to a binomial distribution. All 500 respondents answered the question, resulting in 64 yeses and 436 noes.

It was then determined that the PRCAS demographics were not entirely representative of the expected proportions for the Denver Metro Area based on reported 2022 population demographics from the US Census Bureau² and 2022 American Community Survey³. These demographics included the following characteristics: Sex, Ethnicity, Household Income, Marital Status, Educational Attainment, and Employment Status. Thus, responses to PRCAS question 16 were weighted to account for discrepancies in these demographic features.

Each 'Yes' and 'No' carried the same numerical weight of 1. However, demographic weights were determined by dividing the expected proportion (i.e. the demographic proportion reported by either the US Census Bureau or American Community Survey) by the observed demographic proportion from the PRCAS. An overall weight for each respondent was then calculated by taking the arithmetic average of the six demographic class weights. This final weight was then multiplied by the numerical value of the 'Yes' or 'No'.

Lastly, the respondents were filtered on whether they initially answered, 'Yes' or 'No', and then the final weight value was summed and rounded to the next nearest whole number to determine shifts in the overall number of yeses and noes. This method resulted in a very minor shift to 63 yeses and 437 noes.

Using the weighted proportion of responses, a 99% Confidence Interval was then generated using the Wilson Score Interval with Continuity Correction. The Wilson Score Interval calculated a weighted proportional awareness for Metro Water Recovery in the Denver Metro Area of 14.88% ± 4.15%. In other words, based on the PRCAS responses to question 16, it is believed with 99% confidence that the actual awareness of Metro Water Recovery for the Denver Metro Area falls between 10.73% and 19.03%.

It should be noted, as with all statistical methods, there were limitations to the methodology outlined above. Details can be found in the Survey & Analysis Limitations section.

Sentiment Analysis Methodology

Sentiment Analysis is a conceptually simple technique used to determine if text responses about a topic would be considered positive, negative, or neutral. PRCAS Question 25, "Based on what you read or what you know, what 3 words come to mind when you hear Metro Water Recovery?", was deemed a good candidate for such an analysis. The initial analysis was performed using the "Azure Machine Learning" addon for Microsoft Excel. These sentiments were then reviewed for accuracy. Sentiment words were grouped in the following ways:

- **Positive Sentiment:** Environment (and derivatives of Environment such as Environmental), Water Treatment Program, Recycle, Clean, and Water. In general, if the response used words that demonstrated an awareness of who Metro Water Recovery is and what they do, the response was considered positive.
- **Neutral Sentiment:** Repeating Metro Water Recovery's name, replies that were clearly to just fill the space, or incoherent responses were considered neutral.
- **Negative Sentiment:** Boring, Gross, Nasty, Yuck, Vague, and any version of "I don't know", "I don't care", "Nothing", or "Unsure" were considered negative sentiments from an awareness perspective.

Of the 500 responses received to PRCAS Question 25, 427 were considered positive, 29 were considered neutral, and 44 were considered negative. These results indicate that Metro Water Watery is generally viewed in a positive way.

Survey & Analysis Limitations

- **Survey Limitations**
 - 9thWonder did not provide details about how survey participants were selected. Therefore, whether the sampling was random is questionable.
- **Confidence Interval Assumptions & Limitations**
 - Sampling was assumed to be random, but as outlined in the Survey Limitations, this assumption may not be valid.
 - Because the proportion of yes versus no responses is not 0.5, the data is skewed. This skew makes the data less subject to a normal distribution, thereby decreasing confidence that the actual population proportion of awareness is contained in the reported interval. The 99% Confidence Interval was selected to help account for this by providing a broader range of possible values. The Wilson Score Interval with Continuity Correction was also selected for its ability to better handle departures from normal distributions via adjustments to the normal approximation formula and the correction factor which makes interval estimates more conservative⁴.
 - Taking the arithmetic mean of the demographic weights assumes that each demographic is equally influential on the awareness response. An ANOVA confirms this assumption (i.e. there is no statistically significant difference in mean between demographics). However, this variable equality assumption may still not hold true because the ANOVA assumes the samples are randomly selected and independent. Due to the sampling discrepancy outlined above, not all required assumptions for an ANOVA are therefore met. More advanced statistical techniques would need to be deployed to further bolster confidence.
 - Lastly, applying weights to the awareness responses may not be necessary at all. The weights assume the PRCAS demographics are different enough from the expected Denver Metro Area Demographics to be statistically significant. More advanced statistical techniques would need to be employed to confirm the significant difference necessary to apply weights.

References

1. *9thwonder is now part of Gravity Global*. Gravity Global. (n.d.).
<https://www.gravityglobal.com/news/press/9thwonder-is-now-part-of-gravity-global>
2. United States Census Bureau. (n.d.). *Employment Status - Denver Metro Area*. Explore census data.
<https://data.census.gov/table?q=All+Congressional+Districts+%28116th+Congress%29+fully%2Fpartially+within+Denver-Aurora-Lakewood%2C+CO+Metro+Area&t=Employment+and+Labor+Force+Status&g=310XX00US19740&y=2022>
3. *Denver-aurora-lakewood, CO metro area*. Census Reporter. (n.d.).
<https://censusreporter.org/profiles/31000US19740-denver-aurora-lakewood-co-metro-area/>
4. Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158), 209.
<https://doi.org/10.2307/2276774>