Overview of Tutorial Project

To answer the following questions a variety of visualizations were created. How does the incidence rate of breast cancer, vary across different countries? How does other demographic data impact the incidence rates of breast cancer? What other visualizations can be made to better understand breast cancer significance? While leveraging online resources and the data that is available three graphs were created. The data that was used was from SEER cancer statistics and the World Cancer Research fund. These data sets were used as they were available to the public but also can be leveraged in an academic setting such as in a committee meeting. Then a world graph of incidence rates of breast cancer was created to show that breast cancer is an issue worldwide. Furthermore, quantification of these incidence rates was made into a bar graph highlighting the countries in which they had the highest rates. Then a variety of bar graphs were created using other demographic data such as race, ethnicity, and age. With the assistance of online sources and papers found these visualizations answer the questions proposed. Something that was challenging was that sometimes quotations were not put where needed but then it was fixed. Additionally, since ggplot2 and other packages were used there are different ways to save the files which was challenging at first. In the future, different visualizations can be created using the base knowledge of R gained from this assignment. Furthermore, useful data sets were found for this assignment that will be helpful in future endeavors in creating PowerPoints for committee presentations and in lab meetings. Sources utilized for this project are cited on the following page.

Works Cited:

Bhatnagar, S. R. (2024, August 17). *Plot Cumulative Incidence and Survival Curves*. Plot cumulative incidence and survival curves. <https://cran.r-project.org/web/packages/casebase/vignettes/plotabsRisk.html>

*Breast cancer statistics*. World Cancer Research Fund. (2025, February 5). <https://www.wcrf.org/preventing-cancer/cancer-statistics/breast-cancer-statistics/>

Holtz, Y. (n.d.). *Help and inspiration for R charts*. The R Graph Gallery. https://r-graph-gallery.com/

liao, ling. (2025). Inequality in breast cancer: Global Statistics from 2022 to 2050. *The Breast*. <https://doi.org/10.2139/ssrn.4976619>

Nordmann E, McAleer P, Toivo W, Paterson H, DeBruine LM. Data Visualization Using R for Researchers Who Do Not Use R. Advances in Methods and Practices in Psychological Science. 2022;5(2). doi:10.1177/25152459221074654

Schneider, T. (2023, November 8). *Breast cancer subtypes analysis with R*. Medium. <https://medium.com/@taylorschneider_61189/breast-cancer-subtypes-analysis-with-r-e1d090ab151a>

SEER\*Explorer: An interactive website for SEER cancer statistics [Internet]. Surveillance Research Program, National Cancer Institute; 2025 Apr 16. [cited 2025 May 2]. Available from: <https://seer.cancer.gov/statistics-network/explorer/>. Data source(s): SEER Incidence Data, November 2024 Submission (1975-2022), [SEER 21 registries](https://seer.cancer.gov/registries/terms.html).

Zhang, Y., Ji, Y., Liu, S., Li, J., Wu, J., Jin, Q., Liu, X., Duan, H., Feng, Z., Liu, Y., Zhang, Y., Lyu, Z., Song, F., Song, F., Yang, L., Liu, H., & Huang, Y. (2025). Global burden of female breast cancer: New estimates in 2022, temporal trend and future projections up to 2050 based on the latest release from Globocan. *Journal of the National Cancer Center*. https://doi.org/10.1016/j.jncc.2025.02.002