

Table 1: Number of observations in each pre-defined category from Clark (1981)

Category	Nb observations	Percentage	Cumulative percentage	Nb subjects	Avg age
(-13,25]	4148	93.8673908	93.86739	536	56.12437
(25,40]	239	5.4084635	99.27585	91	71.85105
(40,55]	22	0.4978502	99.77370	14	75.69545
(56,70]	10	0.2262955	100.00000	1	70.18000

# Multivariate statistics: Assignment 1

**Team B:** Kendall Brown *r0773111*  
Stefan Velez *r0924289*

Raïsa Carmen *s0204278*  
Adhithya Unni Narayanan

## 1 Data trichotomization

To trichotomize the data, suitable cut-off points need to be found. The cutoff points are often chosen based on either expert knowledge or so as to optimize predictive power. An easy, often used method for dichotomization is a median-split since it assures that there are an equal amount of observation at either side of the cutoff value. Similarly, for trichotomization, we could aim for approximately 33.33% of the observations in each of the three categories. That would result in the following three categories: .

Alternatively, The Centers for Disease Control and Prevention distinguishes the following levels of hearing loss (“Centers for Disease Control and Prevention a Fact Sheet about Your Hearing Test,” n.d.), based on Clark (1981):

- $\leq 25$  dB: Normal hearing
- 26 - 40 dB: Mild hearing loss
- 41 - 55 dB: Moderate hearing loss
- 56 - 70 dB: Moderate / severe hearing loss
- 71 - 90 dB: Severe hearing loss
- $\geq 91$  dB: Profound hearingloss

Table 1 shows that, in this dataset, there is no one in the severe hearing loss categories and the large majority has normal hearing (93.867390812401).

It is quite common in literature to dichotomize hearing loss into normal hearing ( $\leq 25$  dB) and hearing loss ( $> 25$  dB) (see Garinis et al. 2017; Gallagher et al. 2019; Ju et al. 2022, for example). However, thichotomization is less common and it should be noted that it is generally not advised to discretize continuous data since some information is inevitably lost (Nelson et al. 2017; MacCallum et al. 2002).

## **2 Marginal model**

Q2

## **3 Random-effects model**

Q2

### **3.1 Empirical Bayes prediction**

Q3

## **4 Transition model**

Q4

## **5 Discussion**

### **Bibliography**

- “Centers for Disease Control and Prevention a Fact Sheet about Your Hearing Test.” n.d. <https://www.cdc.gov/niosh/mining/userfiles/works/pdfs/2008-102.pdf>.
- Clark, JG. 1981. “Uses and Abuses of Hearing Loss Classification.” *ASHA* 23 (7): 493–500.
- Gallagher, Nicola E., Chris C. Patterson, Charlotte E. Neville, John Yarnell, Yoav Ben-Shlomo, Anne Fehily, John E. Gallacher, Natalie Lynner, and Jayne V. Woodside. 2019. “Dietary Patterns and Hearing Loss in Older Men Enrolled in the Caerphilly Study.” *British Journal of Nutrition* 121 (8): 877–86. <https://doi.org/10.1017/S0007114519000175>.
- Garinis, Angela C, Campbell P Cross, Priya Srikanth, Kelly Carroll, M Patrick Feeney, Douglas H Keefe, Lisa L Hunter, et al. 2017. “The Cumulative Effects of Intravenous Antibiotic Treatments on Hearing in Patients with Cystic Fibrosis.” *Journal of Cystic Fibrosis* 16 (3): 401–9.
- Ju, Min Jae, Sung Kyun Park, Sun-Young Kim, and Yoon-Hyeong Choi. 2022. “Long-Term Exposure to Ambient Air Pollutants and Hearing Loss in Korean Adults.” *Science of The Total Environment* 820: 153124.
- MacCallum, Robert C, Shaobo Zhang, Kristopher J Preacher, and Derek D Rucker. 2002. “On the Practice of Dichotomization of Quantitative Variables.” *Psychological Methods* 7 (1): 19.
- Nelson, S. L. Prince, V. Ramakrishnan, P. J. Nietert, D. L. Kamen, P. S. Ramos, and B. J. Wolf. 2017. “An Evaluation of Common Methods for Dichotomization of Continuous Variables to Discriminate Disease Status.” *Communications in Statistics - Theory and Methods* 46 (21): 10823–34. <https://doi.org/10.1080/03610926.2016.1248783>.