*data*

MDPI

*Data Descriptor*

# Sustainable Transportation Characteristics Diary—Example of Older (50+) Cyclists

**Sreten Jevremović [1], Carol Kachadoorian [2], Filip Arnaut [3], Aleksandra Kolarski [1,3] and Vladimir A. Srećković [1,3,*]**

[1] Scientific Society "Isaac Newton", Volgina 7, 11160 Belgrade, Serbia; jevremovic749@gmail.com (S.J.); kolarski@ipb.ac.rs (A.K.)
[2] dblTilde CORE, Inc., 105 Jacks Point Road, Oxford, MD 21654, USA; carol@dbltildecore.org
[3] Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia; arnaut@ipb.ac.rs
* Correspondence: vlada@ipb.ac.rs

**Abstract:** Cycling is a sustainable and healthy form of transportation that is gradually becoming the primary means of transportation over shorter distances in many countries. This paper describes the dataset used to determine the cycling characteristics of seniors in the USA and Canada. For these purposes, a specially created questionnaire was used in a survey conducted from August 2021 to July 2022. The questionnaire contained sections related to the general sociodemographic characteristics of the respondents, general characteristics of cycling (type of bicycle, cycle time, mileage, etc.), and specific characteristics of cycling (riding in night conditions, termination of cycling, motivating and demotivating factors for cycling, etc.). The total sample consisted of 5096 respondents (50+ years old). This database is particularly significant because it represents the first set of publicly available data related to the cycling characteristics of older adults. The database can be used by various researchers dealing with this topic, but also by the decision-makers who want to design a sustainable and accessible cycling infrastructure, respecting the requirements of this category of users. Finally, this dataset can serve as an adequate basis in the process of determining the specificities and understanding the needs of older cyclists in traffic.

**Dataset:** https://zenodo.org/doi/10.5281/zenodo.13908441, accessed on 9 October 2024.



**Keywords:** cycling; seniors; cycling infrastructure; cycling characteristics; traffic; older adults; dataset



**Citation:** Jevremović, S.; Kachadoorian, C.; Arnaut, F.; Kolarski, A.; Srećković, V.A. Sustainable Transportation Characteristics Diary—Example of Older (50+) Cyclists. *Data* **2024**, *9*, 123. https://doi.org/10.3390/data9110123

Academic Editor: Han Woo Park

Received: 2 September 2024
Revised: 9 October 2024
Accepted: 24 October 2024
Published: 25 October 2024



## 1. Summary

Cycling is gaining prevalence as a competitive mode of transportation in urban areas for shorter distances. In bicycle-centric nations, including the Netherlands, Denmark, and Germany, the proportion of bicycle usage is markedly elevated and exhibits a distinct upward trajectory relative to other countries [1]. However, a comparable tendency is evident in other nations, particularly across the American continent, where cycling is witnessing a revival post-COVID-19, even in relation to the pre-pandemic period. For instance, ridership in the USA has increased by a remarkable 37% since 2019. Every large city with a population of approximately 5 million or more experienced at least a 25% rise in biking activity during that period [2].

The popularity of cycling remains consistent with advancing age, particularly in highly developed nations like the Netherlands and Denmark. Senior bicycle users constitute a vulnerable demographic, necessitating particular considerations about interactions with other road users. In 2019, nearly half (47%) of cycling fatalities in the European Union involved individuals aged 65 or older [3], which illustrates some of the challenges associated with this mode of transportation, alongside its significant appeal.

This research presents a database with a sample of 5096 cyclists aged over 50 from the United States and Canada. The database was compiled with a specifically designed questionnaire over the timeframe from August 2021 to July 2022, encompassing the region of North America (USA and Canada). The survey consists of inquiries that document an individual's riding activities throughout their lifetime, in addition to their present cycling styles, habits, and preferences.

The questionnaire was promoted through many methods, including social media, websites, conference and webinar presentations, and email outreach utilizing a contact list compiled during the three years in which the surveys were conducted.

This research presents a dataset that continues a long-term investigation on the characteristics of senior cyclists in North America. The research was initially carried out in 2018, leading to the establishment of a multi-year best practice that resulted in several significant publications, as displayed in Table 1. The Year 5, 50+ Cycling Survey is now being implemented and is accessible via the following website: https://survey.alchemer.com/s3/7706636/50-Cycling-Survey-Year-5 (accessed on 9 October 2024).

**Table 1.** A summary of the most important publications.

| No. | Publication Name | Type | Reference |
|---|---|---|---|
| 1. | Cycling Past 50: A Closer Look Into the World of Older Cyclists (Year 3) | Report | [4] |
| 2. | Cycling Part 50: A Closer Look Into the World of Older Cyclists (Year 4) | Report | [5] |
| 3. | Ageing and Mobility: A Look at How Ageing Affects Driving and Cycling | Report/Article | [6] |
| 4. | It's About the Bike: What we Learned about eBikes from the 50+ Cycling Survey, Year 4 Survey | Report/Article | [7] |
| 5. | Cycling and Nature: The Connection between Cycling and Nature for Older Adults | Report/Article | [8] |

The primary advantage of this database is that it constitutes a freely accessible dataset on the riding characteristics of senior cyclists. This database serves as a great resource for decision-makers aiming to establish a unique, accessible, and sustainable cycling infrastructure while addressing the requirements of this vulnerable demographic. This dataset can provide a sufficient foundation for identifying the specificities and comprehending the demands of older bikers within the transportation and traffic systems of contemporary countries.

## 2. Data Description

The dataset comprises two sheets: the first, titled "Raw data", contains the data gathered during the survey, and the second, titled "Code book", provides details on all coded variables. If the question is absent from the second sheet, the responses are documented as text (sentences) rather than numerical values and can be located within the dataset.

The questionnaire had 26 questions presented via an algorithm, illustrated in Figure 1. All questions were of the closed-ended type, with responses selected from the provided options. Additionally, respondents could select "other" if their response was not included among the supplied options. In that instance, the respondent possessed the autonomy to compose the answer independently, as evidenced in the database.

Considering that the questionnaire was completed online, the initial page of the survey presented the study's objective and offered information on data protection, researcher contact details, and the anticipated duration for survey completion. The participants were advised that the survey was anonymous and voluntary, and that they could withdraw from the study at any time. Upon providing their informed consent to participate in the study, the respondents were forwarded to the initial set of questions.

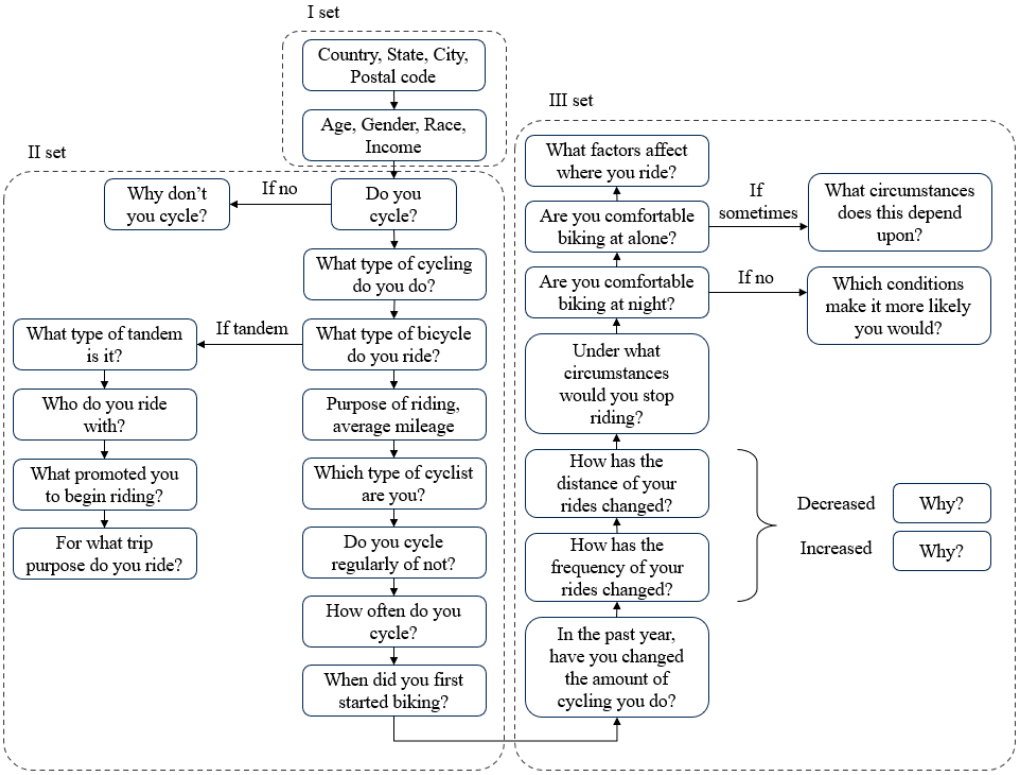


**Figure 1.** Algorithm of questions in the questionnaire.

Figure 1 illustrates that the questionnaire comprised three sets of questions:

Set I: General socio-demographic questions related to the place of residence and characteristics of the respondents such as gender, age, and economic status.

Set II: Questions such as the type of bicycle used, purpose of travel, mileage, and frequency of cycling.

Set III: Specifics related to the use of the bicycle, such as the conditions under which the bicycle would be used and the use of the bicycle at night or alone (without the presence of another person).

The primary objective of the questions structured in this manner and the organized questionnaire was to ascertain the cycling characteristics of older individuals, as well as to identify their preferences, such as situations in which they feel at ease or uncomfortable, their preferred times for cycling, and their inclination to use a bicycle for transportation, recreation, or other purposes.

The data were gathered and stored in an Excel database, where they were prepared and checked, which will be elaborated upon in the subsequent section.

The Excel file containing the data is structured such that each column header corresponds to a question or a component of a question from the questionnaire. The following is a concise overview of the data contained in the Excel file:

The initial four columns (A, B, C, and D) pertain to generic information regarding the respondents: country, abbreviated name, city, and postal code. No data coding occurred; all information was documented exactly as provided by the respondents.

Column E documents the responses to the inquiry "do you cycle?". The responses are encoded in binary (see the codebook).

Columns F-M provide the reasons for respondents' reluctance to cycle, with each column documenting the corresponding response. This question allows multiple answers.

Columns N-V correspond to the specific form of riding practiced by the respondents. Multiple answers could also be selected.

Columns W-AE denote the type of bicycle used.

Columns AF-AR delineate the attributes of the tandem bicycle: its purpose, rationale for use, and intended users.

Columns AS-AX shows the general socio-economic characteristics of the respondents, including age, gender, ethnicity, and income, as detailed in Table 2.

**Table 2.** Socio-demographic characteristics of the sample.

| Variable | *N* | [%] |
|---|---|---|
| **Country** | | |
| USA | 4324 | 84.8 |
| Canada | 772 | 15.2 |
| Age | | |
| 50–59 | 1401 | 27.5 |
| 60–69 | 2092 | 41.1 |
| 70–79 | 1025 | 20.1 |
| More than 80 | 249 | 4.9 |
| No answer | 329 | 6.4 |
| Gender | | |
| Female | 1665 | 32.7 |
| Male | 3142 | 61.6 |
| Other | 9 | 0.2 |
| No answer | 280 | 5.5 |
| Race | | |
| Asian | 128 | 2.5 |
| Native Hawaiian or Other Pacific Islander | 7 | 0.1 |
| Black/African American | 76 | 1.5 |
| White | 4328 | 84.9 |
| Hispanic/Latino | 64 | 1.3 |
| American Indian/Alaska Native | 11 | 0.2 |
| At least two | 39 | 0.8 |
| Other | 58 | 1.1 |
| Prefer not to answer | 385 | 7.6 |
| Annual Income | | |
| Less than USD 25,000 | 109 | 2.1 |
| USD 25,000 to USD 34,999 | 182 | 3.6 |
| USD 35,000 to USD 49,999 | 322 | 6.3 |
| USD 50,000 to USD 74,999 | 722 | 14.2 |
| USD 75,000 to USD 99,999 | 775 | 15.2 |
| USD 100,000 to USD 124,999 | 727 | 14.3 |
| USD 125,000 to USD 149,999 | 452 | 8.9 |
| USD 150,000 or more | 1272 | 25.0 |
| Prefer not to say | 535 | 10.4 |

Columns AY-BJ present data regarding the frequency of usage and the exact conditions under which the bicycle is used.

Columns BK-CQ correspond to a series of inquiries regarding the initiation of cycling, variations in the frequency and distance of cycling over the past year, and the rationale behind these alterations.

Columns CR-DD present information regarding the circumstances that would prompt respondents to resume cycling after having previously ceased.

Columns DE-DM present information regarding the user's solitary cycling habits and the settings under which such activity occurs.

The DN-EF columns present information regarding the causes and conditions that affect the user's cycling preferences.

Columns EG-GA present information regarding the purpose of travel and average mileage.

The final column, GB, presents information regarding the type of cyclist users identify themselves as.

## 3. Methods

The data included in this study were gathered in the USA and Canada from August 2021 to July 2022. Participants engaged in an online survey conducted via the Alchemer Survey platform, which was distributed over all of North America. The poll was disseminated through multiple methods, including social media, websites, conference and webinar

presentations, and email outreach utilizing a contact list compiled over the three years during which the surveys were conducted.

The survey was conducted in accordance with the Declaration of Helsinki of 1975, revised in 2008, and was approved by the Institutional Review Board of San Jose State University and the Board of Directors of dblTilde CORE, Inc. (No. 029/24).

By applying one of the most commonly used sampling methods, the snowball technique [9–11], participants were enlisted to disseminate the survey link across their networks. No inclusion criteria were established other than age, as the intended demographic for this dataset comprised riders aged 50 and above. Microsoft Excel was employed to recode, calculate, and analyze the obtained data as necessary.

### *3.1. Data Validation*

After data collection, validation and verification was carried out through several steps:

- Logical validation (consistency check)—ensuring that all answers make sense in relation to each other. For example, if a person indicated that the cycling distance decreased during the review period, and the average mileage for the same period is more than 40 but fewer than 50 miles, that might indicate an inconsistency. Such responses were thoroughly checked and excluded from the analysis if inconsistencies were found. We also removed responses registered from countries other than the USA and Canada. Also, we removed nearly incomplete responses, as well as other responses determined to be not suitable for inclusion.
- Range validation—verifying that numerical answers fall within the expected ranges. For example, if asking about age, the age range is between 50 and 100.
- Format validation—verifying that there is no placeholder text (e.g., "gfdgda", or "0000") in open-ended questions.
- Outlier Detection—for numerical data, we used min and max functions to identify responses that deviate significantly from the norm, mainly for the age question, since all of the other questions were prepared in such a way that the respondents could choose one of the offered answers.
- Cross-validation—checking for multiple submissions from the same respondent, and removing them. Also, for the sake of reliability, a brief cross-verification of the obtained results with existing and available socio-economic (e.g., income) and traffic data (for which it was possible to obtain the data, e.g., average cycling mileage, cycling purpose) at the state level was performed. Based on this comparison, the obtained data (in the survey) corresponds to the real data with a deviation of up to 5%.

In this way, 404 respondents were excluded from the sample. A summary of the final sample's ($N$ = 5096; MeanAge = 65, StDevAge = 33.2) socio-demographic characteristics is provided in Table 2.

### *3.2. Data Usability*

The provided dataset can be utilized in two principal ways.

Researchers can examine the various factors and characteristics of cycling among older users. It would be crucial to examine the reasons why respondents refrain from cycling and identify the necessary conditions to encourage these individuals to commence cycling [12–14]. Moreover, analyzing the sorts of bicycles utilized and their respective uses, particularly focusing on cargo and tandem bicycles, would be useful. Another approach is to examine the reasons for the alteration in cycling frequency (either a drop or increase) during the preceding period. An interesting direction for future research involves examining the characteristics of cycling under night conditions, along with the factors that influence bicycle usage in specific areas.

Decision makers: this dataset serves as a suitable foundation for the prospective planning of new infrastructure. For instance, the type of bicycle utilized can inform the planning of appropriate infrastructure for the specified categories [15]. Data on the predominant routes is required for this purpose. Conversely, alterations in the frequency and distance of cycling,

along with the motivations behind these changes [16], may indicate potential deficiencies in the current infrastructure, necessitating modifications to accommodate older users.

## 4. Limitations of the Study and Future Research

It is important to note some limitations and restrictions that should be considered by researchers who wish to use the current dataset. The survey was initially conducted online, potentially introducing sampling bias by excluding elderly persons without internet access. Future research of this nature, commencing with data collection in early 2025, also intends to conduct live surveys of older users to enhance the quality of the data acquired. Also, response bias should be taken into account due to the nature of data collection that involves surveying; that is, it should be taken into account that the answers are self-reported and may reflect a certain subjectivity on the matter. A concise cross-verification of the acquired results with available socio-economic (e.g., income) and traffic data (for which data were accessible, such as average mileage and cycling purposes) at the state level was conducted. The comparative analysis indicates that the survey data align with the actual data, exhibiting a variance of up to 5%.

The quality of the acquired data can be enhanced by gathering further information, including on health problems and environmental factors, which is scheduled for the next research phase and will be incorporated into subsequent questionnaires.

**Author Contributions:** Conceptualization, S.J., C.K., A.K., V.A.S. and F.A.; methodology, S.J.; validation, S.J. and C.K.; resources, A.K., V.A.S. and F.A.; data curation, S.J. and C.K.; writing—original draft preparation, S.J.; writing—review and editing, S.J., C.K., A.K., V.A.S. and F.A.; visualization, S.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The survey was conducted in accordance with the Declaration of Helsinki of 1975, revised in 2008, and was approved by the Institutional Review Board of San Jose State University and the Board of Directors of dblTilde CORE, Inc. (No. 029/24).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data described in this paper can be found at the following link: https://zenodo.org/doi/10.5281/zenodo.13908441 (accessed on 9 October 2024).

**Conflicts of Interest:** Author Carol Kachadoorian was employed by the company dblTilde CORE, Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Buehler, R.; Goel, R. Chapter Seven—A Global Overview of Cycling Trends. *Adv. Transp. Policy Plan.* **2022**, *10*, 137–158. [CrossRef]
2. Hurford, M. Good News, Everyone! People Are Riding More than Ever. Available online: https://www.bicycling.com/culture/a45325157/good-news-everyone-people-are-riding-more-than-ever/ (accessed on 9 September 2024).
3. European Commission Facts and Figures Cyclists. *European Road Safety Observatory*; European Commission Facts and Figures Cyclists: Brussels, Belgium, 2021.
4. Kachadoorian, C. *Cycling Past 50: A Closer Look into the World of Older Cyclists*; Mineta Transportation Institute: San Jose, CA, USA, 2021.
5. Kachadoorian, C. *Cycling Past 50: A Closer Look into the World of Older Cyclists, Year 4 Survey*; Mineta Transportation Institute: San Jose, CA, USA, 2023.
6. Kachadoorian, C. *Ageing and Mobility: A Look at How Ageing Impacts Driving and Cycling*; Mineta Transportation Institute: San Jose, CA, USA, 2023.
7. Kachadoorian, C. *It's About the Bike: What We Learned About EBikes from the 50+ Cycling Survey, Year 4*; Mineta Transportation Institute: San Jose, CA, USA, 2023.
8. Kachadoorian, C. *Cycling and Nature: A Look at the Connection Between Cycling and Nature for Older Adults*; Mineta Transportation Institute: San Jose, CA, USA, 2024.
9. Wright, R.; Stein, M. Snowball Sampling. *Encycl. Soc. Meas.* **2005**, *1*, 495–500. [CrossRef]
10. Stivala, A.; Koskinen, J.; Rolls, D.; Wang, P.; Robins, G. Snowball Sampling for Estimating Exponential Random Graph Models for Large Networks. *Soc. Netw.* **2016**, *47*, 167–188. [CrossRef]
11. Leighton, K.; Kardong-Edgren, S.; Schneidereith, T.; Foisy-Doll, C. Using Social Media and Snowball Sampling as an Alternative Recruitment Strategy for Research. *Clin. Simul. Nurs.* **2021**, *55*, 37–42. [CrossRef]

12. Tarkkala, K.; Vosough, S.; West, J.; Roncoli, C. Evaluating the influence of cyclists' route choices incorporation into travel demand modelling: A case study in greater Helsinki. *Trans. Res. Int. Pers.* **2024**, *27*, 101224. [CrossRef]
13. Uijtdewilligen, T.; Baran Ulak, M.; Jan Wijlhuizen, G.; Geurs, K. Effects of crowding on route preferences and perceived safety of urban cyclists in the Netherlands. Trans. *Res. Part A.* **2024**, *183*, 104030.
14. Zeng, Q.; Gong, Z.; Wu, S.; Zhuang, C.; Li, S. Measuring cyclists' subjective perceptions of the street riding environment using K-means SMOTE-RF model and street view imagery. *Int. Jour. App. Earth Obs. Geoinf.* **2024**, *128*, 103739. [CrossRef]
15. von Stülpnagel, R.; Rintelen, H. A matter of space and perspective—Cyclists', car drivers', and pedestrians' assumptions about subjective safety in shared traffic situations. *Transp. Res. Part A* **2024**, *179*, 103941. [CrossRef]
16. Hagemeister, C.; Bertram, L. Reported pushy driving against cyclists in Germany. *Jour. Saf. Res.* **2024**, *88*, 395–405. [CrossRef] [PubMed]