LA VERDAD CHRISTIAN COLLEGE, INC.
*Mac Arthur Highway, Sampaloc, Apalit, Pampanga*

# ANALYZING VARIABLES, DATA, PARTICIPANTS, AND QUESTIONNAIRE OF A STUDY

A

MIDTERM PROJECT

IN

## STATISTICAL ANALYSIS

## IN INFORMATION SYSTEM

Prepared by:

Reyes, Lorenz Genesis M.

BS in Information Systems 3

I. Research Title:

*"Does More Cycling Mean More Diversity in Cycling? "*

II. Respondents:

**People aged 16–74 in England and Wales who were employed at the time of the 2001 and 2011 UK Census** answered the question about their main commute mode, selecting options such as cycling for the longest part of their journey to work.

III. Institutions involved:

**Medical Research Council (MRC), through** the MRC Population Health Scientist Fellowship and MRC Centenary Award.

**National Institute for Health Research (NIHR)**, which funded a postdoctoral fellowship.

**Economic and Social Research Council (ESRC)**.

**NHS (National Health Service)**.

**Department for Health (UK)**.

IV. Year Published:

February 25, 2015

V. Abstract:

In low-cycling countries, cycling is not evenly distributed across genders and age groups. In the UK, men are twice as likely as women to cycle to work and cycling tends to be dominated by younger adults. By contrast, in higher cycling countries and cities, gender differences are low, absent, or in the opposite direction. Such places also lack the UK's steady decline in cycling among those aged over 35 years. Over the past fifteen years, some UK local areas have seen increases in cycling. This paper analyses data from the English and Welsh Census 2001 and 2011 to examine whether such increases are associated with greater diversity among cyclists. We find that in areas where cycling has increased, there has been no increase in the representation of females and a decrease in the representation of older adults. We discuss potential causes and policy implications. Importantly, simply increasing cycling modal share has not proved sufficient
to create an inclusive cycling culture. The UK's culturally specific factors limiting female take-up of cycling seem to remain in place, even where cycling has gone up. Creating a mass cycling culture may require deliberately targeting infrastructure and policies towards currently under-represented groups.

*CONTENTS*

**a. Identify the variables used in the study.**

**Cycling levels**: people that go to work and measured by percentage number level.

**Gender**: male and female cyclist representation.

**Age**: a particular group of cyclists that are present and part of the study "35 years and older."

**Local authority areas**: geographical locations in England and Wales where data was collected.

**b. Identify the qualitative and quantitative variables and indicate the highest level of measurement required in each.**

**Quantitative Variables**:

**Cycling levels**: ratio level (percentage of commuters that practice cycling to go to work).

**Age**: ratio level (measured in years, particularly among 35 years and older people.).

**Qualitative Variables**:

**Gender**: nominal level (male and female).

**Local authority areas**: nominal level (based on geographical location name).

**c. Justify why the used statistics (descriptive or inferential) is correct.**

This research study used both descriptive statistics (in order to summarize and describe the use and type of featured data, like average cycling levels by gender and age) and inferential statistics (that was, use and partake to help make conclusions about the broad population sets that were given and based on the sample data from the collected census). Descriptive statistics are mostly used in conveying the classification of cycling behaviors, and inferential statistics help and allow research to generalize the findings and see the relationship between cycling levels and demographic variables. This is the combination that is used in this research study.

**d. Determine the sources of data (primary and secondary) of the study.**

**Secondary Data**: The source of primary data for this research came from the English and Welsh census that happened to be collected in the time frame of 2001 to 2011. All of the included data in this source happened to be collected by the English government agency and was made to be used for research purposes.

**Primary Data**: any newly performed surveys or new interviews that were conducted in pairing with the collected census data; those things will not fit as the primary data source; however, this research study was based on the already provided text; in other words, the study primarily relies on the collected existing data status.

**e. Determine all the methods used in data collection. Briefly elaborate on how each method was done.**

**Census Data Collection**: These are the data that are gathered with the help of the national census, which is set to collect new demographic information every ten years. This was personally sent to households, those comprehensive surveys to the people across England and Wales, to capture various statistics, including commuting behavior.

**Statistical Analysis**: After the process of data collection, this study includes statistical methods that were used to track and analyze all of the occurring changes in cycling levels and demographic representation over time. This analysis commonly involves software tools that are used for data analysis to compute averages, percentages, and correlations.

**Comparative Regional Analysis:** This type of method was performed by comparing cycling data from different regions and local authorities that are part of England and Wales. By studying the given geographical differences, the researchers can identify different variations in cycling rates and demographic representation (gender and age) all across the areas with different cycling rules and policies.

**f. Are the objectives, statements of the problems, data collection method, tools, and analysis clearly defined?**

The objectives of this study aim to analyze whether increases in cycling are associated with greater diversity among cyclists; this is done specifically addressing gender and age inequities in participation. The statement of the problems tackles issues that are well articulated, and the use of Census data as a secondary source for data collection is clearly stated. Although no specific tools for analysis (such as software) are listed, the methodology suggests the use of statistical tools. This study shows an approach of using both descriptive and inferential statistics that will align to its objective in order to examine the trends and draw meaningful study conclusions.

**g. Examine the method(s) used in data collection. Know what made the researchers decide to use the chosen method?**

The methods used in data collection were utilized by the researchers from the **2001 and 2011 UK Census data**, which offered them large and reliable data sets from the entire population of England and Wales. This was the selected method because it gave a comprehensive census that collected data from millions of respondents, which allowed them to conduct a broad trend analysis over time. A longitudinal aspect of the census was conducted every ten years in order to provide an opportunity to compare more cycling participation across different demographics, making it ideal for studying changes in gender and age diversity among cyclists. The secondary source of data was also ensured to have a high level of consistency in data collection and allowed the researchers to save on the cost and time that were involved in primary data collection.

**h. Based on the Key Design Principles of a Good Questionnaire, explore the questionnaire and/or the interview questions used in the study, and tell (with proper narrative) whether the questionnaire and/or the interview questions were good or not.**

The census questionnaire that was used in the study is considered to be effective because of its key design principle. Although the exact questions are not provided on paper, we assure you that the national censuses are designed to adhere to strict standards for reliability and validity to ensure consistent responses across the population. The particular questions are mostly about those commuting modes, which are likely to be clear, straightforward, and well-structured, which reduces the possibility of confusion or misinterpretation. The closed-ended format (e.g., selecting "bicycle" as a commuting mode) will ensure objective data collection, which is crucial for large-scale studies. Also, based on the given high participation rates and the credibility of census data, we can reasonably assume that the questionnaire was designed to capture accurate and objective information without causing bias.

**i. Were the used of whether open-ended or closed-ended questions in the study reasonable? Do the researchers want to be subjective or objective in their data collection? Explain.**

In this study, the Census used closed-ended questions for information gathering about the participants commuting modes, like interviewing the respondents to give information about their primary mode of transportation from a list of options (including "bicycle"). This closed-ended format question is appropriate for collecting quantitative data as it helps to simplify the process of categorizing and analyzing responses. It also ensures that the data is easily comparable across different sets of respondents all over time, as everyone answers within the same predefined categories. This choice of closed-ended questions aligns with the study's objective to gather measurable data on cycling patterns rather than subjective opinions or open-ended descriptions, which would cause problems in the statistical analysis.

**j. How did the researchers decide about the number of participants in their study?**

The researchers decided the number of participants by including the entire population of England and Wales, but specifically, they focused on the working individuals who responded to the 2001 and 2011 censuses. Numbers of participants are not directly decided by the researchers but rather determined by the extent of the census's coverage. Census surveys aim to be extensive to capture data from all eligible individuals, making the dataset highly accurate. In 2011, for example, the census received a 94% response rate, indicating that the majority of the population was represented. The use of the full population data helps the researchers to make detailed analyses across various demographic groups without the need for sampling or limitations on participant numbers.

**k. What do you think were the considerations put in mind by the researchers in deciding which sampling method to use?**

Because the researchers' analysis depended on census data, there was no particular sampling method. However, because the census is designed to cover the entire population rather than just a subset, the study avoids frequent sampling difficulties like bias or underrepresentation. The researchers had access to data from all working-age people in England and Wales, giving them a very precise and comprehensive picture of cycling involvement. This strategy eliminates worries about whether the sample is representative of the community because the census strives for 100% inclusion, resulting in broad and diverse data across gender, age, and geographic region.

**l. By examining how the researchers did their sampling, do you think the sample is representative of the population? How did you say so?**

Yes, because the sample is very representative of the population, as it is based on data from the national census, which is intended to poll the whole population of England and Wales. The census data, which received an estimated 94% response rate in both 2001 and 2011, provides a nearly full picture of the population, encompassing all significant demographic groupings such as age, gender, and geographic location. The high sample size and broad coverage reduce the danger of bias or exclusion, ensuring that the results are representative of the entire population. Furthermore, the high response rate and broad involvement make the findings more generalizable and applicable to national policy conversations about cycling and transportation diversity.

**2001 and 2011 UK Census data *Questionnaires*:**

**How do you usually travel to work?**

**Options:** Car, Bus, Train, Bicycle, Walking, Motorcycle, Other

**How often do you use a bicycle to travel to work?**

**Options:** Daily, Weekly, Occasionally, Never

**What is the approximate distance of your commute to work (in kilometers/miles)?**

**Options:** Less than 1 mile, 1-3 miles, 3-5 miles, 5-10 miles, Over 10 miles

**What is your gender?**

**Options:** Male, Female, Other

**What is your age?**

**Options:** 16-24, 25-34, 35-44, 45-54, 55-64, 65+

**What factors prevent you from cycling to work?**

**Options**: Lack of safe routes, Weather conditions, Distance too far, Health issues, No access to a bicycle, Other (please specify)

What is the primary reason you cycle?

**Options**: Commuting to work, Leisure, Exercise, Running errands, Other (please specify)

Are you aware of dedicated cycling infrastructure in your area (e.g., bike lanes, cycle paths)?

**Options**: Yes, No

Have local government initiatives (e.g., cycling promotions, infrastructure improvements) influenced your decision to cycle?

**Options**: Very much, Somewhat, Not at all

How satisfied are you with the conditions for cycling in your area?

**Options**: Very satisfied, Satisfied, Neutral, Dissatisfied, Very dissatisfied