A dataset to study transportation, residential context, and well-being in Santiago, Chile

Niloofar Nalaee, Beatriz Mella Lira, Antonio Paez

# Abstract

Large-scale travel surveys are invaluable sources of information to understand travel behavior and other aspects of the urban experience, such as residential context. By their nature, they often shy away from overloading respondents with additional questions. The data presented in this article result from an effort to purposefully collect data on various aspects of the experience of living and moving in a major city in the Global South. The data set contains essential socio-economic and demographic information about the respondents, as well as their built environment and behaviors commuting to work. In addition, the survey (conducted between DATE-DATE, 2016) includes information about the respondents’ feelings and emotions in relation to their commuting experience, the social experience of a variety of transportation modes, various self-assessed health questions, patterns of use of information and telecommunication technologies, and questions about sustainability and the environment. The survey was adopted a quota-sampling method that used the Pre-Census of 2012 as a frame, and in total includes 451 validated questionnaires.

## Key words

* Transportation
* Commuting to work
* Built environment
* Well-being
* Equity
* Santiago
* Chile

## Specifications Table

See [Table 1](#tbl-specifications) for details of the data set.

Table 1: Specifications table

| Items | Explanation |
| --- | --- |
| Subject area | Transportation, Geography, Public Health and Health Policy, Urban development |
| More specific subject area | Transport inequalities, Stress and limited horizons, Travel behavior, Global South |
| Type of data | R Data Package |
| How data was acquired | The survey was conducted using a (pen-and-paper??) questionnaire. The instrument contains some quantitative variables regarding the individual characteristics of respondents and mostly 5-point Likert scale responses in the rest of the questionnaire |
| Data format | Thematic tables and documentation in native R format. The thematic tables can be linked by means of a common ID field |
| Parameters for data collection | The survey was collected using a quota-sampling method based on the information from Pre-Census of 2012, and in total, 451 persons validly completed the survey and face-to-face in Santiago, Chile in 2016. The survey collected information on a wide range of travel-related issues (socio-demographics, health-related, perceptions and travel behavior, travel choices and planning, social interaction factors, built environment, among others) |
| Description of data collection | Data was acquired through the 5-Likert scale questionnaire regarding most sections of the questionnaire, using a face-to-face and quota-sampling method for individual characteristics |
| Data source location | Santiago, Chile |
| Data accessibility | https://paezha.github.io/bSantiago/ |

## Value of the data

A large body of the data has made inroads investigating psychological impact on travelers ranging from positive feelings of enjoyment in some to the sensation of stress in many others that can affect the effectiveness of policy measure (in the case of positive feelings) and are known to affect health outcomes (in the case of stress).This would be interesting for those with transport policies concerns as public feelings like happiness and satisfaction rating of travel mode switching has been studied previously (Abou-Zeid et al. 2012). Data set contribute to psychological impact on travelers both active and motorized modes of transportation to examines not only the feeling of stress, but also how these effects are experienced by travelers and investigates the importance that travelers attach to their feelings of stress which makes it valuable for researchers who focused on public sector development and health-related policies. It allows us to realize the concept of “limited horizons” and focus on developing people’s capability to select what they have their reason behind it. Furthermore, it would be advantageous for decision makers to evaluate their attempts to flourish well-being by expanding capabilities in further research regarding transport inequalities, index of stress, travel behavior in the region or even as a representative for other areas with similar attributes(Walker 2005). The dataset provides a wide range of travel-related issues such as socio-demographics, health-related, perceptions and travel behavior, travel choices and planning, social interaction factors, built environment, among others.

## Data

The data package described in this paper contains 9 data objects in tabular format, which are listed in [Table 2](#tbl-list-of-tables). Each object corresponds to a theme and includes information about the individual characteristics, travel-to-work related information, and perceptions of the built environment. Furthermore, other themes include social interactions by mode of transportation, mode shifting and use of information technologies, a number of self-assessed health responses, feelings and emotions as they relate to the commute experience, attitudes with respect to how well transportation systems satisfy the needs of respondents, and attitudes about nature and sustainability. As seen in [Table 2](#tbl-list-of-tables), the number of variables in these tables ranges from 5 to 24. The tables share a common unique identifier (ID) that can be used to link several tables in a consistent way for multi-theme analysis of the data.

Table 2: List of tables in data package

| Table | Theme |
| --- | --- |
| Santiago\_IC | A table with 22 variables that describe the individual characteristics of respondents |
| Santiago\_TW | A table with 8 variables regarding the commuting behavior of respondents |
| Santiago\_SI | A table with 5 variables about social interaction by mode of transportation |
| Santiago\_IS | A table with 10 variables related to mode-shifting and use of information technologies |
| Santiago\_BE | A table with 23 variables about built environment at the place of residence of respondents |
| Santiago\_H | A table with 13 variables about health attributes of respondents |
| Santiago\_FE | A table with 23 variables with information about the feelings and emotions of respondents |
| Santiago\_RPD | A table with 24 variables with information about attitudes with respect to the effectiveness of the transportation system |
| Santiago\_NS | A table with 11 attitudinal variables that refer to nature and sustainability aspects of city |

The theme of table Santiago\_IC is the individual characteristics of respondents and includes several essential socio-economic and demographic attributes. [Table 3](#X2eb449a934c2f8b9182ef914af54518f65500b3)) presents a summary of these variables, including information about the neighborhood of residence of respondents. All respondents were successfully geolocated to the level of Communes (large administrative partitions equivalent to municipalities) and Macro Zones (larger aggregations of communes used for descriptive purposes). The street address information is less complete. The place of work (geocoded at the level of Communes) is missing only in a small number of cases (complete rate 0.96) and few respondents report a *second* place of work/studies where they travel (r01\_WORK\_COM\_2).

As seen in the table, most respondents’ information is quite complete. Only 16 respondents reported facing some disability, and of these only 11 completed the section asking them about the type (r0O\_TYPE\_DIS). In terms of gender, the sample includes approximately 15.7% more women than men.

**BEA:** PLEASE COMMENT ON THE DISTRIBUTION OF AGES, EDUCATION, OCCUPATION, AND INCOME; ALSO, IS THE VARIABLE MONTHLY OR ANNUAL INCOME? HOW WHERE THE LEVELS/CUTS FOR INCOME SELECTED? ARE THE UNITS IN PESOS OR 1,000 PESOS?

**BEA:** r0P\_MODE\_1 IS PRIMARY MODE? AND r0Q\_MODE\_2 AND r0R\_MODE\_3 ARE SUBSEQUENT MODES FOR THE SAME TRIP? OR MODES USED IN TERMS OF DECREASING FREQUENCY?

Other variables regarding the respondents include their height (cm) and weight (kg), and the household context in terms of number of adults (between 1 and 7) and minors in the family (75% of respondents report no children, and of the remainder the number of children varies between 1 and 6).

These variables provide insights about different segments of the population. In terms of place of residence (**BEA:** ARE ALL COMMUNES URBAN? IS THERE AN URBAN-RURAL INTERFACE?).

(a) Data summary

|  |  |
| --- | --- |
| Name | select(Santiago\_IC, -ID) |
| Number of rows | 451 |
| Number of columns | 21 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| character | 7 |
| factor | 10 |
| numeric | 4 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: character**

| skim\_variable | n\_missing | complete\_rate | min | max | empty | n\_unique | whitespace |
| --- | --- | --- | --- | --- | --- | --- | --- |
| r0A\_COMMUNE | 0 | 1.00 | 4 | 16 | 0 | 43 | 0 |
| r0B\_MACR\_ZONE | 0 | 1.00 | 3 | 12 | 0 | 7 | 0 |
| r0B\_STREET\_1 | 38 | 0.92 | 4 | 24 | 0 | 200 | 0 |
| r0B\_STREET\_2 | 108 | 0.76 | 1 | 23 | 0 | 211 | 0 |
| r0I\_WORK\_COM\_1 | 18 | 0.96 | 4 | 16 | 0 | 30 | 0 |
| r0I\_WORK\_COM\_2 | 371 | 0.18 | 1 | 16 | 0 | 25 | 0 |
| r0O\_TYPE\_DIS | 440 | 0.02 | 5 | 19 | 0 | 10 | 0 |

**Variable type: factor**

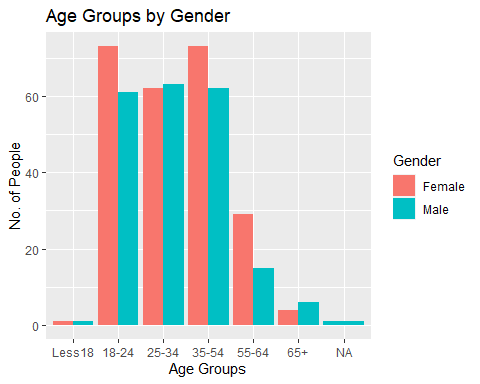
| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| --- | --- | --- | --- | --- | --- |
| r0C\_SEX | 0 | 1.00 | FALSE | 2 | WOM: 242, MAN: 209 |
| r0D\_AGE | 1 | 1.00 | TRUE | 6 | 35 to 54: 135, 18 to 24: 134, 25 to 34: 125, 55 to 64: 44 |
| r0G\_EDUCATION | 1 | 1.00 | TRUE | 5 | COL: 223, SEC: 104, PRO: 84, POS: 29 |
| r0H\_OCCUPATION | 1 | 1.00 | FALSE | 9 | FUL: 257, STU: 101, SEL: 26, HOM: 19 |
| r0J\_INCOME | 6 | 0.99 | TRUE | 6 | 977 to 1,550: 90, 2,380 or more: 83, Less than 423: 71, 1,550 to 2,380: 68 |
| r0N\_DRIVING\_LICENSE | 0 | 1.00 | FALSE | 2 | YES: 260, NO: 191 |
| r0M\_DISABILITY | 0 | 1.00 | FALSE | 2 | NO: 435, YES: 16 |
| r0P\_MODE1 | 0 | 1.00 | FALSE | 8 | MET: 160, CAR: 115, BUS: 110, WAL: 22 |
| r0Q\_MODE2 | 15 | 0.97 | FALSE | 8 | MET: 136, BUS: 131, CAR: 66, WAL: 32 |
| r0R\_MODE3 | 92 | 0.80 | FALSE | 8 | CAR: 74, WAL: 62, BUS: 59, TAX: 51 |

**Variable type: numeric**

| skim\_variable | n\_missing | complete\_rate | mean | sd | p0 | p25 | p50 | p75 | p100 | hist |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| r0E\_WEIGHT | 29 | 0.94 | 69.60 | 13.09 | 42 | 60 | 68 | 79.75 | 120 | ▃▇▅▂▁ |
| r0F\_HEIGHT | 20 | 0.96 | 167.21 | 9.46 | 146 | 160 | 167 | 174.00 | 197 | ▃▇▇▃▁ |
| r0K\_ADULTS | 5 | 0.99 | 2.81 | 1.29 | 1 | 2 | 3 | 4.00 | 7 | ▇▃▃▁▁ |
| r0L\_CHILDREN | 6 | 0.99 | 0.66 | 0.92 | 0 | 0 | 0 | 1.00 | 6 | ▇▂▁▁▁ |

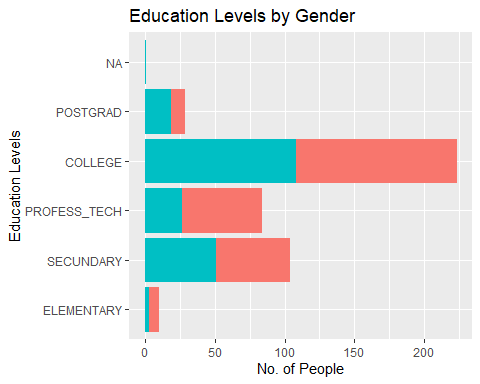
Table 3: Variables regarding individual characteristics of respondents

par(mfrow=c(1,2))  
  
# Age and Gender Bar Graph  
SantiagoSurvey2016 <- read\_excel(choose.files())  
  
ggplot(SantiagoSurvey2016,aes(x=factor(r0D\_EDAD), fill=factor(r0C\_GENERO)))+  
 geom\_bar(position="dodge")+  
 scale\_x\_discrete("Age Groups",   
 labels=c("A"="Less18","B"="18-24","C"="25-34",  
 "D"="35-54","E"="55-64","F"="65+"))+  
 scale\_fill\_discrete(name="Gender",  
 breaks=c("A", "B"),  
 labels=c("Female", "Male"))+  
 ggtitle("Age Groups by Gender")+  
 ylab('No. of People')

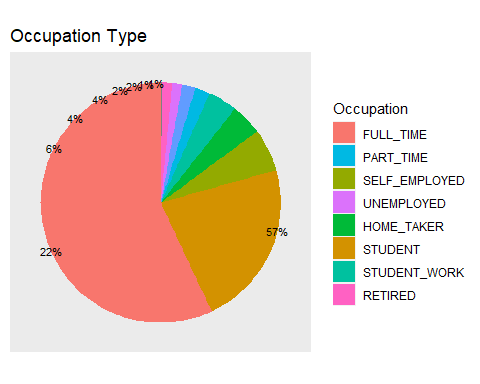


# Education bar graph  
qplot(as.factor(SantiagoSurvey2016$r0G\_EDUCACION), geom='bar', fill=as.factor(SantiagoSurvey2016$r0C\_GENERO))+  
 coord\_flip()+  
 scale\_x\_discrete("Education Levels",  
 labels=c("A"="ELEMENTARY","B"="SECUNDARY",  
 "C"="PROFESS\_TECH","D"="COLLEGE",  
 "E"="POSTGRAD"))+  
 scale\_fill\_discrete(name="Gender",  
 breaks=c("1", "2"),  
 labels=c("Male", "Female"))+  
 ggtitle("Education Levels by Gender")+ylab('No. of People')

Warning: `qplot()` was deprecated in ggplot2 3.4.0.



#Employment Pie Chart  
SantiagoSurvey2016$r0H\_OCUPACION <- reorder(SantiagoSurvey2016$r0H\_OCUPACION, X = SantiagoSurvey2016$r0H\_OCUPACION,   
 FUN = function(x) -length(x))  
  
at <- nrow(SantiagoSurvey2016) - as.numeric(cumsum(sort(table(SantiagoSurvey2016$r0H\_OCUPACION)))-  
 0.5\*sort(table(SantiagoSurvey2016$r0H\_OCUPACION)))  
  
label=paste0(round(sort(table(SantiagoSurvey2016$r0H\_OCUPACION))/  
 sum(table(SantiagoSurvey2016$r0H\_OCUPACION)),2) \* 100,"%")  
  
p <- ggplot(SantiagoSurvey2016,aes(x="", fill = r0H\_OCUPACION)) +  
 geom\_bar(width = 1) +  
 scale\_fill\_discrete(name="Occupation",  
 breaks=c("A","B","C", "D", "E", "F", "G", "H"),  
 labels=c("FULL\_TIME", "PART\_TIME","SELF\_EMPLOYED", "UNEMPLOYED", "HOME\_TAKER", "STUDENT", "STUDENT\_WORK", "RETIRED"))+  
 coord\_polar(theta="y") +  
 annotate(geom = "text", y = at, x = 1.5, label = label,size=3)+  
 theme(axis.text = element\_blank(),  
 axis.ticks = element\_blank(),  
 panel.grid = element\_blank(),  
 axis.title.x = element\_blank(),  
 axis.title.y = element\_blank())+  
 ggtitle("Occupation Type")  
print(p)



# df <- data.frame(matrix(ncol = 2, nrow = 6))   
# colnames(df)<-c('Income','Frequency')  
# df$r0J\_INGRESO<-c('Less423','423to639', '639to977','977to1550','1550to2380',  
# '2380more')  
#   
# for (i in 1:6){df[i,2]<-sum(SantiagoSurvey2016[,77+i])}  
#   
# df$r0J\_INGRESO<-factor(df$r0J\_INGRESO, levels=df$r0J\_INGRESO) #to prevent reordering of bars  
#   
#   
# ggplot(df,aes(x=r0J\_INGRESO, y=Frequency), fill=df$r0J\_INGRESO)+  
# geom\_bar(stat="identity")+  
# scale\_x\_discrete("Personal Income Groups")+  
# coord\_flip()+  
# ggtitle('Personal Income Distribution')+  
# ylab('No. of People')

The theme of the next table, Santiago\_TW, deals with commuting and work variables (see [Table 4](#tbl-Travel-work-Descriptives)) and consists of seven ordinal categorical variables (factors). Variable r8A\_ACCESSJOB refers to the impact that respondents feel the transportation network has affected their chances of having better jobs. The most common responses were SOME IMPACT and NO IMPACT, but we see that approximately 14.2% of respondents feel that the network has had a major impact. This variable tracks to some extent with the responses to job opportunities in the commune of residence (r8B\_JOBOPP), suggesting a possible correlation between local opportunities and the impact of the transportation network on job outcomes. When asked about their ideal level of accessibility in the commune of the residence (r8C\_ACC\_COM), a majority respondents opt for excellent and very good (**BEA:** I AM A LITTLE UNCLEAR ABOUT THE MEANING OF THE QUESTION THAT PRODUCES THIS VARIABLE).

In terms of the level of satisfaction with their current job, we see that almost 60% of respondents are at least highly satisfied. We also see that long commutes are frequent in this sample, with about one third or respondents spending 1 h or more travelling (**BEA:** ONE WAY OR TWO WAYS?) and about one quarter of respondents spending between 40 minutes and one hour in their daily commute. This distribution is noteworthy because time spent commuting has been recognized as a factor that can affect physical and mental health and well-being in particular in association with motorized transportation (Brutus, Javadian, and Panaccio 2017).

The most common time of the day for commuting is between 7 am and 9 am, but there are also 171 missing responses in this column, so not much can be read from it. Finally, we note that many people spend 35,000-75,000 (**BEA:** WHAT ARE THE UNITS OF THIS? IS IT 35 PESOS OR 35,000 PESOS) monthly on their transportation expenditure. (**BEA:** MAYBE COMMENT ON THE TYPICAL RATIO OF TRANSPORT EXPENDITURE TO INCOME).

(a) Data summary

|  |  |
| --- | --- |
| Name | select(Santiago\_TW, -ID) |
| Number of rows | 451 |
| Number of columns | 7 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| factor | 7 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: factor**

| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| --- | --- | --- | --- | --- | --- |
| r8A\_ACCESSJOB | 15 | 0.97 | TRUE | 5 | SOME IMPACT: 124, NO IMPACT : 123, MINOR IMPACT: 66, MAJOR IMPACT: 64 |
| r8B\_JOBOPP | 11 | 0.98 | TRUE | 5 | GOOD: 169, POOR: 86, FAIR: 85, VERY GOOD: 64 |
| r8C\_ACC\_COM | 15 | 0.97 | TRUE | 5 | EXCELLENT: 184, VERY GOOD: 109, GOOD: 93, FAIR: 27 |
| r8D\_EMPLSATISF | 37 | 0.92 | TRUE | 5 | HIGH SATISF: 135, VERY HIGH SATISF: 130, MEDIUM SATISF: 116, LOW SATISF: 20 |
| r8E\_TIMECOMMUT | 10 | 0.98 | TRUE | 4 | 1h and more : 133, 40-60 min: 113, 20-40 min: 103, 0-20 min: 92 |
| r8F\_SCHEDULE | 171 | 0.62 | TRUE | 6 | 7:00 - 9:00: 168, Others: 47, 9:00 - 13:00: 39, 18:00 - 21:00: 14 |
| r8G\_SPENDING | 7 | 0.98 | TRUE | 4 | 35,000-75,000: 199, LESS THAN 35,000: 166, 75,000-125,000: 46, MORE THAN 125,000: 33 |

Table 4: Variables regarding the commuting behavior of respondents

Table Santiago\_SI includes five variables that capture various aspects of social interaction while commuting (see [Table 5](#tbl-Social-Interaction-Descriptives)). Social interaction is a topic of interest for mode-related choices given earlier evidence that for some commuters privacy is an important consideration and/or a way to manage social stressors (Gardner and Abraham 2007; Paez and Whalen 2010; Lowe and Mosby 2016). When asked to rate the level of interaction with other people during their usual trips, respondents In terms of the level of interaction people have with others during their usual trips, most of them presume a good level and they moderately feel it is important when they consider the presence of other people during their usual travels (r4A\_INTERACC) a plurality of responses are “POOR” or “FAIR” (187) and only 73, that is 16.2% of respondents, rate their level of interaction as “GOOD” or “EXCELLENT” report poor or fair (there are 13 missing responses). With respect to the presence of other people during their travels, the responses tend to be somewhat more ambivalent, and the difference between those for whom this is less or more important is smaller (157 responses are “NOT” or only “SLIGHTLY IMPORTANT” whereas 132 responses are “IMPORTANT” or “VERY IMPORTANT”).

The next two variables in this table deal with feelings and the experience of discrimination: “have you felt discriminated against while travelling?” (r4C\_DISCRIM) and if so, “while using which mode of transportation” (r4D\_MODE). We can see in [Table 5](#tbl-Social-Interaction-Descriptives), that slightly fewer than one quarter of respondents (24.2%) report having felt discriminated while commuting (only 4 responses are missing), and of those close to half (44%) had that experience in public transportation (however, note that about 58% of those who felt discriminated against did not state the mode).

(a) Data summary

|  |  |
| --- | --- |
| Name | select(Santiago\_SI, -ID) |
| Number of rows | 451 |
| Number of columns | 4 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| factor | 4 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: factor**

| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| --- | --- | --- | --- | --- | --- |
| r4A\_INTERACC | 13 | 0.97 | TRUE | 5 | GOOD: 178, FAIR: 102, POOR: 85, VERY GOOD: 53 |
| r4B\_PERSON | 12 | 0.97 | TRUE | 5 | MODERATELY IMPORTANT: 150, NOT IMPORTANT: 88, IMPORTANT: 83, SLIGHTLY IMPORTANT: 69 |
| r4C\_DISCRIM | 4 | 0.99 | FALSE | 2 | NO: 338, YES: 109 |
| r4D\_MODE | 377 | 0.16 | FALSE | 7 | BUS: 33, METRO: 15, CAR: 8, TAXI: 7 |

Table 5: Variables regarding social interactions of respondents

**BEA:** Niloofar described [Table 6](#tbl-ITC-Shifting-Descriptives) as “telecommunication and shifting different modes” in the documentation, but I would like to ask for clarification, because it seems to me that the questions relate to *transfers*. For example, the question:

* Do you shift between transport modes on your usual trips?

Seems to refer to transfers between public transportation modes (e.g., metro to bus), and then the next questions refer to how long waits are, and ease of transferring.

* If YES in the previous question, how is the quality of these inerchanges?
* How would you assess your regular trips in terms of: Waiting times How would you assess your regular trips in terms of: Time of shifts
* How would you assess your regular trips in terms of: Travel time What main aspect makes your shifts difficult?

The last three questions relate to route planning tools, not ITC more generally (for instance, this is unrelated to working from home).

* Do you have access to technology tools such as smartphones and internet, with which you can view information on transportation services?
* How do you assess the level of technological information available to see transportation alternatives? (eg smartphone applications, internet, signs)
* How important is it for you to have access to technological information for your usual trips?

Please comment on these variables!

**THIS IS NILOOFAR’S TEXT:**

The information, telecommunications and mode shifting of respondents (see [Table 6](#tbl-ITC-Shifting-Descriptives)) reveals numerous factorial variables using Likert scale to identify the exact level of respondents’ viewpoints. We can see a considerable number of missing values in quality of changing mode of travel because of people’s decision on mostly saying yes to shift between transport modes on their usual trips and for those who change the quality of these inerchanges was good. Many people assign a good level when they are asked to assess the waiting times, time of shifts and total travel time in their regular trips. Most people have access to technology tools such as smartphones and internet, with which they can view information on transportation services and they assign a good level of technological information available to see transportation alternatives (eg smartphone applications, internet, signs) and most of them assume it is very important to have access to technological information for their usual trips.

This section will gives insights on designing a strong transportation network to facilitate people’s different activities such as work, school, grocery shopping and so forth. Also experts can make users informed about traffic circumstances, road information and cost of transportation by using various modes. Following this, car-use deduction has been turned into the political agenda around the world due to the increasing negative effects of motorized modes of travel on environment and public health. Therefore, improvements in public transportation sector are needed to achieve this goal regarding previous studies that has shown participants are more likely to use bus with increased frequencies, shorter travel time, and high access to bus stops(Ettema et al. 2011).

(a) Data summary

|  |  |
| --- | --- |
| Name | select(Santiago\_IS, -ID) |
| Number of rows | 451 |
| Number of columns | 9 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| factor | 9 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: factor**

| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| --- | --- | --- | --- | --- | --- |
| r6A\_SHIFT | 11 | 0.98 | FALSE | 2 | YES: 247, NO: 193 |
| r6B\_QUALITY\_SHIFT | 198 | 0.56 | TRUE | 5 | GOOD: 107, FAIR: 61, VERY GOOD: 38, POOR: 34 |
| r6CA\_WAITING | 41 | 0.91 | TRUE | 5 | GOOD: 124, POOR: 101, FAIR: 94, VERY GOOD: 61 |
| r6CB\_TIME\_SHIFT | 102 | 0.77 | TRUE | 5 | GOOD: 126, FAIR: 77, POOR: 62, VERY GOOD: 59 |
| r6CC\_TOTALTIME | 28 | 0.94 | TRUE | 5 | GOOD: 138, POOR: 113, FAIR: 75, VERY GOOD: 54 |
| r6D\_DIFFICULTY | 76 | 0.83 | FALSE | 7 | ALL THE PREVIOUS ONES: 107, TOO LONG SHIFTS: 89, UNCOMFORTABLE SHIFTING: 87, BAD INFRASTRUCTURE FOR WAITING TIMES: 33 |
| r6E\_TOOL | 8 | 0.98 | FALSE | 2 | YES: 372, NO: 71 |
| r6F\_INFO | 15 | 0.97 | TRUE | 5 | GOOD: 153, VERY GOOD: 118, FAIR: 63, EXCELLENT: 63 |
| r6G\_IMP\_INFO | 10 | 0.98 | TRUE | 5 | VERY IMPORTANT: 198, IMPORTANT: 122, MODERATELY IMPORTANT: 68, SLIGHTLY IMPORTANT: 33 |

Table 6: Variables regarding information and telecommunications and mode shifting of respondents

The theme of table Santiago\_BE is perceptions of the built environment. The way environments are perceived, and not just their objective attributes, has been shown to correlate with travel behavior (e.g., Jamal, Mohiuddin, and Paez 2020; Loo et al. 2015). Surveying perceptions is therefore a good remedy to the lack of canonical data sets in regions of the world where built environment attributes are not systematically collected.

There are in total 22 variables in this table, and their descriptive statistics appear in [Table 7](#tbl-Built-Environment-Descriptives). The variables are organized in pairs: one asks about the perception of an attribute and the second the importance of that attribute to the respondent. In this way, r7AA\_AUTOSPACE is about the perception of space for autos, and r7BA\_AUTOSPACE is about the importance that respondents assign to this attribute. Aspects of the built environment covered by this table, in addition to space for autos, are number of parking spaces, quality of highways (**BEA:** ROADS IN GENERAL OR HIGHWAYS SPECIFICALLY?), space for pedestrians, quality of sidewalks, cleanliness of bus stops and seating areas, protection from inclement weather at bus stops, amount and quality of cycleways, and bike sharing schemes.

The descriptive statistics suggest that most respondents have positive evaluations of space for cars and parking spaces, at the same time that they assign a high level of importance to these attributes. Respondents also show a slight tendency to assess positively the space for pedestrians and quality of sidewalks located near of their home, and these features are also regarded as very important or important by a majority of respondents. In contrast, although respondents feel strongly about the importance of facilities related to buses and cycling, the perception of these attributes tends to be poor.

(a) Data summary

|  |  |
| --- | --- |
| Name | select(Santiago\_BE, -ID) |
| Number of rows | 451 |
| Number of columns | 22 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| factor | 22 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: factor**

| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| --- | --- | --- | --- | --- | --- |
| r7AA\_AUTOSPACE | 10 | 0.98 | TRUE | 5 | GOOD: 127, VERY GOOD: 117, FAIR: 72, EXCELLENT: 67 |
| r7AB\_PARKING\_NUMB | 10 | 0.98 | TRUE | 5 | GOOD: 119, FAIR: 96, VERY GOOD: 88, POOR: 85 |
| r7AC\_QHIWAY | 11 | 0.98 | TRUE | 5 | VERY GOOD: 142, GOOD: 141, FAIR: 67, EXCELLENT: 49 |
| r7AD\_PEDESTRN | 10 | 0.98 | TRUE | 5 | GOOD: 141, VERY GOOD: 102, FAIR: 82, EXCELLENT: 68 |
| r7AE\_QSIDEWA | 9 | 0.98 | TRUE | 5 | GOOD: 119, FAIR: 110, VERY GOOD: 99, POOR: 67 |
| r7AF\_CLEAN\_STOP | 9 | 0.98 | TRUE | 5 | POOR: 129, FAIR: 108, GOOD: 102, VERY GOOD: 67 |
| r7AG\_SEAT | 9 | 0.98 | TRUE | 5 | POOR: 142, FAIR: 122, GOOD: 94, VERY GOOD: 53 |
| r7AH\_CLIMA | 10 | 0.98 | TRUE | 5 | POOR: 156, FAIR: 139, GOOD: 88, VERY GOOD: 36 |
| r7AI\_CICLEWA\_NUMB | 9 | 0.98 | TRUE | 5 | POOR: 189, FAIR: 96, GOOD: 84, VERY GOOD: 37 |
| r7AJ\_CICLEWA\_Q | 9 | 0.98 | TRUE | 5 | POOR: 171, FAIR: 96, GOOD: 96, VERY GOOD: 44 |
| r7AK\_BICSHARE | 9 | 0.98 | TRUE | 5 | POOR: 177, GOOD: 100, FAIR: 69, EXCELLENT: 49 |
| r7BA\_IMPAUTOSPACE | 12 | 0.97 | TRUE | 5 | VERY IMPORTANT: 140, MODERATELY IMPORTANT: 104, IMPORTANT: 95, SLIGHTLY IMPORTANT: 56 |
| r7BB\_IMPPARKING\_NUMB | 12 | 0.97 | TRUE | 5 | VERY IMPORTANT: 152, MODERATELY IMPORTANT: 102, IMPORTANT: 91, SLIGHTLY IMPORTANT: 51 |
| r7BC\_IMPQHIWAY | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 214, IMPORTANT: 100, MODERATELY IMPORTANT: 82, NOT IMPORTANT: 23 |
| r7BD\_IMPPEDESTRN | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 278, IMPORTANT: 103, MODERATELY IMPORTANT: 41, NOT IMPORTANT: 10 |
| r7BE\_IMPQSIDEWA | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 297, IMPORTANT: 86, MODERATELY IMPORTANT: 35, SLIGHTLY IMPORTANT: 12 |
| r7BF\_IMPCLEAN\_STOP | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 286, IMPORTANT: 92, MODERATELY IMPORTANT: 39, SLIGHTLY IMPORTANT: 13 |
| r7BG\_IMPSEAT | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 258, IMPORTANT: 92, MODERATELY IMPORTANT: 56, SLIGHTLY IMPORTANT: 19 |
| r7BH\_IMPCLIMA | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 296, IMPORTANT: 83, MODERATELY IMPORTANT: 34, SLIGHTLY IMPORTANT: 14 |
| r7BI\_IMPCICLEWA\_NUMB | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 296, IMPORTANT: 76, MODERATELY IMPORTANT: 42, SLIGHTLY IMPORTANT: 19 |
| r7BJ\_IMPCICLEWA\_Q | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 308, IMPORTANT: 67, MODERATELY IMPORTANT: 44, NOT IMPORTANT: 11 |
| r7BK\_IMPBICSHARE | 12 | 0.97 | TRUE | 5 | VERY IMPORTANT: 259, IMPORTANT: 78, MODERATELY IMPORTANT: 66, SLIGHTLY IMPORTANT: 22 |

Table 7: Variables regarding the built environment at the place of residence of respondents

Table Santiago\_H contains responses to twelve questions related to health. Health information could be useful in investigating affects of transportation policy decisions on public health. Also being aware of specific factors making modes stressful will help transportation and public health experts make commuting a safer, more enjoyable and less stressful activity. Consequently ,planners and transportation experts could mitigate the potentially serious health outcomes of a stressful commute (Legrain, Eluru, and El-Geneidy 2015). Having such a detailed data set and understanding the relationship between health and mode choice in commuting would help them to adopt reform policies and management strategies in accordance with healthy modes of travel as active and public transportation (Mattisson et al. 2018).

Similar to the set of questions about the built environment, these questions have two parts: first, the respndents were asked to assess their experience of some aspect of their commute, and then the importance they assign to this aspect of their commute. For example, variable r1A\_STRESS refers to the level of stress that respondents experience in their usual trips; its companion variable is r1GA\_IMPSTRESS, which refers to how important stress is to their experience. As seen in [Table 8](#tbl-Health-Information-Descriptives), a majority of respondents rate the stress in their commutes as medium or lower. This aspect of the experience is rated as very important or important by a majority of respondents. Other variables in this table contain information about the physical effort involved in their usual trips (r1B\_EFFORT and r1GB\_IMPEFFORT); proximity to other travelers in their usual trips (r1C\_PROX and r1GC\_IMPPROX); the environmental pollution they are exposed to in their usual trips (r1D\_CONTAM and r1GD\_IMPCONTAM); how safe they feel in their usual trips (r1E\_SAFETY and r1GE\_IMPSAFETY); and finally how comfortable their trips are (r1F\_COMFORT and r1GF\_IMPCOMFORT).

(a) Data summary

|  |  |
| --- | --- |
| Name | select(Santiago\_H, -ID) |
| Number of rows | 451 |
| Number of columns | 12 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| factor | 12 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: factor**

| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| --- | --- | --- | --- | --- | --- |
| r1A\_STRESS | 2 | 1.00 | TRUE | 5 | MOD: 145, LOW: 107, HIGH: 92, VERY HIGH: 54 |
| r1B\_EFFORT | 2 | 1.00 | TRUE | 5 | NEUTRAL: 152, POSITIVE: 117, VPOSITVE: 72, NEGATIVE: 67 |
| r1C\_PROX | 8 | 0.98 | TRUE | 5 | POOR: 130, GOOD: 103, FAIR: 95, EXCELLENT: 60 |
| r1D\_CONTAM | 4 | 0.99 | TRUE | 5 | VDISSATISFIEDISFIEDSATISFIEDISFIED: 149, DISSATISFIEDISFIED: 138, UNSURE: 107, SATISFIED: 33 |
| r1E\_SAFETY | 1 | 1.00 | TRUE | 5 | UNSURE: 121, DISSATISFIEDISFIED: 110, VDISSATISFIEDISFIEDSATISFIEDISFIED: 97, SATISFIED: 84 |
| r1F\_COMFORT | 1 | 1.00 | TRUE | 5 | VDISSATISFIEDISFIEDSATISFIEDISFIED: 111, DISSATISFIEDISFIED: 109, UNSURE: 103, SATISFIED: 73 |
| r1GA\_IMPSTRESS | 6 | 0.99 | TRUE | 5 | VERY IMPORTANT: 243, IMPORTANT: 108, MODERATELY IMPORTANT: 57, SLIGHTLY IMPORTANT: 26 |
| r1GB\_IMPEFFORT | 5 | 0.99 | TRUE | 5 | VERY IMPORTANT: 156, MODERATELY IMPORTANT: 121, IMPORTANT: 102, SLIGHTLY IMPORTANT: 46 |
| r1GC\_IMPPROX | 7 | 0.98 | TRUE | 5 | VERY IMPORTANT: 169, IMPORTANT: 130, MODERATELY IMPORTANT: 97, SLIGHTLY IMPORTANT: 29 |
| r1GD\_IMPCONTAM | 4 | 0.99 | TRUE | 5 | VERY IMPORTANT: 258, IMPORTANT: 86, MODERATELY IMPORTANT: 64, SLIGHTLY IMPORTANT: 23 |
| r1GE\_IMPSAFETY | 5 | 0.99 | TRUE | 5 | VERY IMPORTANT: 311, IMPORTANT: 80, MODERATELY IMPORTANT: 35, SLIGHTLY IMPORTANT: 13 |
| r1GF\_IMPCOMFORT | 5 | 0.99 | TRUE | 5 | VERY IMPORTANT: 216, IMPORTANT: 111, MODERATELY IMPORTANT: 91, SLIGHTLY IMPORTANT: 15 |

Table 8: Variables regarding health information of respondents

**NOTE FROM ANTONIO:** SEVERAL TABLES NEED MY ATTENTION BEFORE THEY ARE READY TO BE SEEN BY BEA

## Experimental Design, Materials and Methods

The study is based on a paper-based survey conducted face-to-face in Santiago in 2016. The survey collected information on a wide range of travel-related issues (socio-demographics, health-related, perceptions and travel behavior, travel choices and planning, social interaction factors, built environment, among others). The data collection considered a quota-sampling method based on the information from Pre-Census of 2012, and in total, 451 persons validly completed the survey. This paper considers the first part of the survey, with information about the basic socio-economic data, travel choices, activities and commuting information, and the question related to the levels of stress experienced in while traveling.

## Acknowledgements

## References

Abou-Zeid, Maya, Regina Witter, Michel Bierlaire, Vincent Kaufmann, and Moshe Ben-Akiva. 2012. “Happiness and Travel Mode Switching: Findings from a Swiss Public Transportation Experiment.” *Transport Policy* 19 (1): 93–104.

Brutus, Stéphane, Roshan Javadian, and Alexandra Joelle Panaccio. 2017. “Cycling, Car, or Public Transit: A Study of Stress and Mood Upon Arrival at Work.” *International Journal of Workplace Health Management*.

Ettema, Dick, Tommy Gärling, Lars Eriksson, Margareta Friman, Lars E Olsson, and Satoshi Fujii. 2011. “Satisfaction with Travel and Subjective Well-Being: Development and Test of a Measurement Tool.” *Transportation Research Part F: Traffic Psychology and Behaviour* 14 (3): 167–75.

Gardner, Benjamin, and Charles Abraham. 2007. “What Drives Car Use? A Grounded Theory Analysis of Commuters’ Reasons for Driving.” *Transportation Research Part F: Traffic Psychology and Behaviour* 10 (3): 187–200.

Jamal, Shaila, Hossain Mohiuddin, and Antonio Paez. 2020. “How Do the Perceptions of Neighborhood Conditions Impact Active Transportation? A Study in Rajshahi, Bangladesh.” Journal Article. *Transportation Research Part D: Transport and Environment* 87: 102525. https://doi.org/<https://doi.org/10.1016/j.trd.2020.102525>.

Legrain, Alexander, Naveen Eluru, and Ahmed M El-Geneidy. 2015. “Am Stressed, Must Travel: The Relationship Between Mode Choice and Commuting Stress.” *Transportation Research Part F: Traffic Psychology and Behaviour* 34: 141–51.

Loo, L. Y. L., J. Corcoran, D. Mateo-Babiano, and R. Zahnow. 2015. “Transport Mode Choice in South East Asia: Investigating the Relationship Between Transport Users’ Perception and Travel Behaviour in Johor Bahru, Malaysia.” Journal Article. *Journal of Transport Geography* 46: 99–111. <https://doi.org/10.1016/j.jtrangeo.2015.06.011>.

Lowe, Kate, and Kim Mosby. 2016. “The Conceptual Mismatch: A Qualitative Analysis of Transportation Costs and Stressors for Low-Income Adults.” *Transport Policy* 49: 1–8.

Mattisson, Kristoffer, Ahmed Osman Idris, Ellen Cromley, Carita Håkansson, Per-Olof Östergren, and Kristina Jakobsson. 2018. “Modelling the Association Between Health Indicators and Commute Mode Choice: A Cross-Sectional Study in Southern Sweden.” *Journal of Transport & Health* 11: 110–21.

Paez, A., and K. Whalen. 2010. “Enjoyment of Commute: A Comparison of Different Transportation Modes.” Journal Article. *Transportation Research Part a-Policy and Practice* 44 (7): 537–49. <https://doi.org/10.1016/j.tra.2010.04.003>.

Walker, Melanie. 2005. “Amartya Sen’s Capability Approach and Education.” *Educational Action Research* 13 (1): 103–10.