A dataset to study equity in individual travel behavior and choices in Santiago, Chile

Niloofar Nalaee, Beatriz Mella Lira, Antonio Paez

# Abstract

The data set presented in this article would be advantageous to give intuition about the experience of daily travel and its psychological impact on travelers, ranging from positive feelings of enjoyment in some to the sensation of stress in many others. This dataset particularly examines the feelings of stress by users of active and motorized modes of transportation. Furthermore, it also investigates the importance that travelers attach to their feelings of stress. This allows us to explore the concept of “limited horizons”, the normalization of subpar experiences by those less able to adapt. This data provides information about commuters in terms of their individual characteristics, health-related factors while using transportation modes, feelings and emotions towards different modes of travel, own decisions, social interaction, attitudes towards nature and sustainability, shifts between modes, attitudes towards built environment, traveling to work. Data for the research are drawn from a survey conducted in Santiago, Chile, based on a quota-sampling method based on the information from Pre-Census of 2012, and in total, 451 persons validly completed the survey. Describe the statistics and conclude suggest that

## Key words

* Travel behavior
* Well-being
* Equity
* Transportation
* 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 conducted using questionnaire. The instrument contains descriptive data of respondents and 5-Likert scale questionnaire regarding most sections of the questionnaire |
| Data format | Raw Data |
| Parameters for data collection | The survey collected by 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 | State if data is with this article or in public repository. If public repository, please explicitly name repository and data identification number and provide a direct URL to data |

## Value of the data

Creating open data products can help accelerate research (Arribas-Bel et al. 2021). 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. Dataset 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. This dataset allows us to explore the concept of “limited horizons”, the normalization of subpar experiences by those less able to adapt, an advantageous resource for further research regarding transport inequalities, index of stress, travel behavior in the region or even as a representative for other areas with similar attributes. 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 tables as listed in [Table 2](#tbl-list-of-tables). Each table is thematic and includes information about the individual characteristics, health information, feelings and emotions information, reason and planning decision of respondents, social interaction of respondents, nature and sustainability, telecommunication and shifting, built environment and travel-work related information. As can be seen in data set, we have a wide variety of variables in each section.

Table 2: List of tables in data package

| Table | Theme |
| --- | --- |
| Santiago\_IC | A table with 23 variables that describe the individual characteristics of respondents |
| Santiago\_H | A table with 12 variables about health attributes of respondents |
| Santiago\_FE | A table with 22 variables with information about the feelings and emotions of respondents |
| Santiago\_RPD | A table with 23 variables that reveal information about the decision-making and planning characteristics of respondents |
| Santiago\_SI | A table with 4 variables about social interaction of respondents |
| Santiago\_NS | A table with 10 variables which indicate nature and sustainability factors from the viewpoint of respondents |
| Santiago\_IS | A table with 9 variables related to information and telecommunications and mode shifting of respondents |
| Santiago\_BE | A table with 22 variables about built environment at the place of residence of respondents |
| Santiago\_TW | A table with 7 variables regarding the commuting behavior of respondents |

Data collected regarding individual characteristics (see [Table 3](#X2eb449a934c2f8b9182ef914af54518f65500b3)) reveals characteristic variables like the information about the neighborhood of respondents. In this section most of them are complete but type of disability has numerous missing values and it seems like most people have not disability to specify its type. Also we have many missing responses for the commune of different activities like study or work. This is because people would more prefer one region for their regular trips than the second one(#origin).In terms of factorial variables like gender, age education it can be seen most of the variables are almost complete except mode 3. It looks like a large number of people tend to use one mode and less number of them use second or third mode of travel. It can be seen in numeric variables like number of adults and children in a family, weight and height that we have approximately complete answers. The only variables with considerable number of missing values, are weight and height which seems a few people did not perfectly response to them. These information give insights about different segments of gender, age and so on to organize planning decisions for them according to their requirements. Also this part would be advantageous for business developments to figure out marketing and apply strategies for future patterns based on consumer demands.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_IC |
| Number of rows | 451 |
| Number of columns | 23 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| character | 8 |
| factor | 10 |
| numeric | 5 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: character**

| skim\_variable | n\_missing | complete\_rate | min | max | empty | n\_unique | whitespace |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ORIGIN | 406 | 0.10 | 1 | 10 | 0 | 6 | 0 |
| 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\_COM\_TRAB\_1 | 18 | 0.96 | 4 | 16 | 0 | 30 | 0 |
| r0I\_COM\_TRAB\_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\_GENDER | 0 | 1.00 | FALSE | 2 | FEM: 242, MAL: 209 |
| r0D\_AGE | 1 | 1.00 | TRUE | 6 | 35t: 135, 18t: 134, 25t: 125, 55t: 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: 90, 238: 83, Les: 71, 155: 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 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | 0 | 1.00 | 226.00 | 130.34 | 1 | 113.5 | 226 | 338.50 | 451 | ▇▇▇▇▇ |
| r0E\_WEIGHT | 29 | 0.94 | 69.60 | 13.09 | 42 | 60.0 | 68 | 79.75 | 120 | ▃▇▅▂▁ |
| r0F\_HEIGHT | 20 | 0.96 | 167.21 | 9.46 | 146 | 160.0 | 167 | 174.00 | 197 | ▃▇▇▃▁ |
| r0K\_ADULTS | 5 | 0.99 | 2.81 | 1.29 | 1 | 2.0 | 3 | 4.00 | 7 | ▇▃▃▁▁ |
| r0L\_CHILDREN | 6 | 0.99 | 0.66 | 0.92 | 0 | 0.0 | 0 | 1.00 | 6 | ▇▂▁▁▁ |

Table 3: Variables regarding individual characteristics of respondents

The Health information (see [Table 4](#tbl-Health-Information-Descriptives)) indicates different factorial variables that are almost complete and organized categorically because of using Likert scale. In this table we have two part of variables: first the respondents’ assessment of each variable and second what level of importance do people assign to each variable. Speaking about stress most people experience average level of stress on regular trips and they feel it is very important for them. Most of people were neutral in demonstrating physical effort on their usual trips and they assign high level of importance to it. Many people mentioned poor level of proximity (proximity) to other transport users on their usual trips and felt it was very important in their viewpoint. In terms of environmental contamination most of people were very dissatisfied and they consider it is very important on their regular trips. About safety variable most respondents were unsatisfied with level of safety (not being assaulted or harassed) that they experience on their usual trips and declared it was very important in their trips. A large number of people specified very dissatisfying level of comfort throughout their trips and they assume it is very important when they are on their trips.

Health information could be useful in investigating affects of transportation policy decisions on public health.This information would be useful for planners and transportation experts to develop transport models used to identify how different policies or plans could affect active modes of travel and public transit. Having such a detailed data set would help them to adapt reform policies and management strategies in accordance with public health goals.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_H |
| 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 | MEDIUM: 145, LOW: 107, HIGH: 92, VERY HIGH: 54 |
| r1B\_EFFORT | 2 | 1.00 | TRUE | 5 | NEUTRAL: 152, POSITIVE: 117, VPOSITVE: 72, NEGATIVE: 67 |
| r1C\_PROXIMITY | 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\_STRESS | 6 | 0.99 | TRUE | 5 | VERY IMPORTANT: 243, IMPORTANT: 108, MODERATELY IMPORTANT: 57, SLIGHTLY IMPORTANT: 26 |
| r1GB\_EFFORT | 5 | 0.99 | TRUE | 5 | VERY IMPORTANT: 156, MODERATELY IMPORTANT: 121, IMPORTANT: 102, SLIGHTLY IMPORTANT: 46 |
| r1GC\_PROXIMITY | 7 | 0.98 | TRUE | 5 | VERY IMPORTANT: 169, IMPORTANT: 130, MODERATELY IMPORTANT: 97, SLIGHTLY IMPORTANT: 29 |
| r1GD\_CONTAM | 4 | 0.99 | TRUE | 5 | VERY IMPORTANT: 258, IMPORTANT: 86, MODERATELY IMPORTANT: 64, SLIGHTLY IMPORTANT: 23 |
| r1GE\_SAFETY | 5 | 0.99 | TRUE | 5 | VERY IMPORTANT: 311, IMPORTANT: 80, MODERATELY IMPORTANT: 35, SLIGHTLY IMPORTANT: 13 |
| r1GF\_COMFORT | 5 | 0.99 | TRUE | 5 | VERY IMPORTANT: 216, IMPORTANT: 111, MODERATELY IMPORTANT: 91, SLIGHTLY IMPORTANT: 15 |

Table 4: Variables regarding health information of respondents

First part of the data set of feelings and emotions (see [Table 5](#tbl-Feelings-Emotions-Descriptives)) reveals factorial variables that people consider them pertained to a specific mode of travel. In terms of freedom, health and social interactions most of people mention them related to walking as a transportation mode. Many people consider using bus connected to unsafety, poverty, unpunctuality, congestion and uncomfortale conditions variables.In terms of using car as the mode of travel, most of respondents identified it connected to functionality, safety, luxury and status. A large group of people declare that bike riding would be related to enjoyment, low-cost and environment care. The only efficiency variable is being connected to using metro as the mode of transportation in respondents’ opinions. In the second part of this table people evaluate the level of enjoy when they are traveling to their daily activities and most pf them uses average level for this variable and revealed the quality of their trips are very low which is not satisfying. Totally in this table we have half of respondents’ answer and the other half is missing values in all variables. feelings and emotions information can be used to mapping and understanding travel behavior and would lead to a more sustainable transportation network. By trying to identify a correlation between feelings, emotions and modes of transportation, planners would be able to persuade people to use more active modes of travel or public transport will be resulted in making our transportation system more sustainable.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_FE |
| 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 |
| --- | --- | --- | --- | --- | --- |
| r2AA\_FREEDOM | 218 | 0.52 | FALSE | 8 | WALK: 96, CAR: 79, BICYCLE: 43, MOTO: 8 |
| r2AB\_UNSAFETY | 217 | 0.52 | FALSE | 8 | BUS: 94, MOTO: 73, WALK: 22, BICYCLE: 21 |
| r2AC\_FUNCTIONALITY | 196 | 0.57 | FALSE | 8 | CAR: 90, MOTOO: 86, BICYCLE: 26, MOTO: 18 |
| r2AD\_ENJOYMENT | 212 | 0.53 | FALSE | 8 | BICYCLE: 118, CAR: 42, WALK: 42, MOTO: 15 |
| r2AE\_LOWCOST | 232 | 0.49 | FALSE | 8 | BICYCLE: 91, WALK: 84, BUS: 15, MOTO: 11 |
| r2AF\_POVERTY | 166 | 0.63 | FALSE | 8 | BUS: 139, WALK: 102, BICYCLE: 13, COLECTIVO: 10 |
| r2AG\_SAFETY | 142 | 0.69 | FALSE | 8 | CAR: 199, METRO: 62, TAXI: 16, WALK: 15 |
| r2AH\_WASTE\_OF\_TIME | 165 | 0.63 | FALSE | 8 | BUS: 222, METRO: 13, CAR: 12, WALK: 12 |
| r2AI\_UNPUNCTUALITY | 139 | 0.69 | FALSE | 7 | BUS: 263, METRO: 13, WALK: 11, CAR: 10 |
| r2AJ\_CONGEST | 233 | 0.48 | FALSE | 8 | BUS: 135, CAR: 52, METRO: 16, TAX: 5 |
| r2AK\_EFFICIENCY | 196 | 0.57 | FALSE | 8 | METRO: 94, CAR: 63, BICYCLE: 37, MOTO: 17 |
| r2AL\_LUXURY | 162 | 0.64 | FALSE | 7 | CAR: 200, TAXI: 66, BICYCLE: 8, MOTO: 6 |
| r2AM\_ENVIRONMENT | 251 | 0.44 | FALSE | 8 | BICYCLE: 121, WAL: 52, METRO: 16, CAR: 4 |
| r2AN\_HEALTH | 259 | 0.43 | FALSE | 7 | WAL: 94, BICYCLE: 76, CAR: 12, TAXI: 3 |
| r2AO\_INTSOCI | 224 | 0.50 | FALSE | 8 | WAL: 84, BUS: 41, METRO: 30, BICYCLE: 27 |
| r2AP\_UNCOMFT | 221 | 0.51 | FALSE | 8 | BUS: 134, METRO: 51, COLECTIVO: 15, MOTO: 14 |
| r2AQ\_HAPPINESS | 216 | 0.52 | FALSE | 8 | BICYCLE: 77, WALK: 77, CAR: 65, MOTO: 6 |
| r2AR\_STATUS | 173 | 0.62 | FALSE | 7 | CAR: 219, TAXI: 28, BICYCLE: 10, WALK: 10 |
| r2B\_DAILY\_ENJOY | 4 | 0.99 | TRUE | 5 | MEDIUM: 146, HIGH: 114, VERY High: 75, LOW: 72 |
| r2C\_IMP\_QUALITY | 6 | 0.99 | TRUE | 5 | VERY LOW: 276, LOW: 99, MEDIUM: 46, HIGH: 14 |
| r2D\_AFFECT | 38 | 0.92 | FALSE | 8 | ALL OF THEM: 136, TRAVEL TIME: 89, CROWDNESS OF PASSENGERS: 87, COMFORT ABSENCE: 38 |
| r2E\_FACILIT | 34 | 0.92 | FALSE | 8 | ALL OF THEM: 164, REDUCTION OF TIME TRAVEL: 82, LESS CROWDNESS OF PASSENGERS: 66, BETTER QULALITY ON STREETS: 36 |

Table 5: Variables regarding feelings and emotions of respondents

Table of decision-making and planning of respondents (see [Table 6](#X11a33f6ea4fb4ee37baff737c8ad81bc4a31c76)) includes different factorial variables using Likert scale and has two parts of an assessment of a variable and its importance from respondents’ viewpoints. Most of people assign level of good when they assess their access to employment opportunities through public transport and presume a good level about their access to public transport which allows them to access the employment they need. While people sometimes visit family and friends, do recreational, cultural and sport activities, and they assign a moderate importance to them, most of them often go for grocery/food shopping and social activities and consider it moderately important. In terms of options most of people assume it is very important to have several options in using different modes of transport and they assign very high when they consider quality of life depends on the access they currently have to public transport. Also most people highly think their quality of life would increase if they have better access to public transport.About the affordability and unaffordability of a mode of travel, a large group of people assign car and taxi to these variables, respectively. For different aspects of the public transport system most of people mention that it has very importance for them to improve access to offices and commercial areas, disponibility of different transport modes, comfort for the use of public transport and the incorporation of other modes to the fare system. Almost all the variables are fairly complete except being economic and uneconomic variables which nearly half of respondents reacted to them.

The most important role of the transportation network and public modes of travel is to provide people with access to different destination in order to travel for business, reuniting with other people, doing grocery shopping and so on. This data set gives us a wide range of variables helpful for identifying people’s travel pattern and what level of importance do people assign. So it would be useful for transport-related experts to have an insight about people movement to provide network system with a an appropriate level of performance to increase the efficiency and promote urban transportation system.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_RPD |
| Number of rows | 451 |
| Number of columns | 23 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| factor | 23 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: factor**

| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| --- | --- | --- | --- | --- | --- |
| r3A\_ACCESS | 11 | 0.98 | TRUE | 5 | GOOD: 186, FAIR: 91, VERY GOOD: 63, POOR: 57 |
| r3B\_ACC\_EM | 18 | 0.96 | TRUE | 5 | GOOD: 134, POOR: 108, VERY GOOD: 83, FAIR: 60 |
| r3CA\_FAM | 11 | 0.98 | TRUE | 5 | SOMWTIMEs: 117, OFTEN: 105, ALWAYS: 100, NEVER: 60 |
| r3CB\_REC | 15 | 0.97 | TRUE | 5 | SOMWTIMEs: 139, OFTEN: 103, ALWAYS: 72, RARELY: 65 |
| r3CC\_CUL | 14 | 0.97 | TRUE | 5 | SOMWTIMEs: 129, RARELY: 123, NEVER: 94, OFTEN: 65 |
| r3CD\_SPO | 15 | 0.97 | TRUE | 5 | SOMWTIMEs: 112, NEVER: 106, RARELY: 93, OFTEN: 75 |
| r3CE\_GROC | 14 | 0.97 | TRUE | 5 | OFTEN: 135, SOMWTIMEs: 112, ALWAYS: 83, RAR: 65 |
| r3CF\_SOC | 13 | 0.97 | TRUE | 5 | OFTEN: 146, SOMWTIMEs: 107, ALWAYS: 84, RAR: 55 |
| r3DA\_FAM | 45 | 0.90 | TRUE | 5 | MODERATELY IMPORTANT: 102, IMPORTANT: 85, NOT IMPORTANT: 79, VERY IMPORTANT: 74 |
| r3DB\_REC | 47 | 0.90 | TRUE | 5 | MODERATELY IMPORTANT: 121, IMPORTANT: 91, SLIGHTLY IMPORTANT: 67, NOT IMPORTANT: 64 |
| r3DC\_CUL | 52 | 0.88 | TRUE | 5 | MODERATELY IMPORTANT: 117, IMPORTANT: 87, SLIGHTLY IMPORTANT: 70, NOT IMPORTANT: 67 |
| r3DD\_SPO | 50 | 0.89 | TRUE | 5 | MODERATELY IMPORTANT: 111, NOT IMPORTANT: 85, IMPORTANT: 77, SLIGHTLY IMPORTANT: 65 |
| r3DE\_GROC | 46 | 0.90 | TRUE | 5 | MODERATELY IMPORTANT: 101, IMPORTANT: 91, NOT IMPORTANT: 75, VERY IMPORTANT: 70 |
| r3DF\_SOC | 47 | 0.90 | TRUE | 5 | MODERATELY IMPORTANTORTANT: 124, IMPORTANT: 89, NOT IMPORTANT: 64, SLIGHTLY IMPORTANT: 64 |
| r3E\_OPTIONS | 1 | 1.00 | TRUE | 5 | NOT IMPORTANT: 326, IMPORTANT: 75, MODERATELY IMPORTANT: 36, SLIGHTLY IMPORTANT: 8 |
| r3F\_ACCESS\_DEPENDENCY | 3 | 0.99 | TRUE | 5 | VERY HIGH: 184, HIGH: 128, MEDIUM: 100, BELOW AVERAGE: 21 |
| r3G\_QUALITY\_INCRS | 2 | 1.00 | TRUE | 5 | VERY HIGH: 267, HIGH: 97, MEDIUM: 54, BELOW AVERAGE: 18 |
| r3H\_ECON | 339 | 0.25 | FALSE | 8 | CAR: 30, ALL: 30, METRO: 18, BUS: 18 |
| r3I\_NOECON | 261 | 0.42 | FALSE | 9 | TAXI: 106, CAR: 39, MOTO: 29, COLECTIVO: 4 |
| r3JA\_OFIC | 20 | 0.96 | TRUE | 5 | NOT IMPORTANT: 195, IMPORTANT: 109, MODERATELY IMPORTANT: 82, SLIGHTLY IMPORTANT: 34 |
| r3JB\_MODES | 17 | 0.96 | TRUE | 5 | NOT IMPORTANT: 297, IMPORTANT: 91, MODERATELY IMPORTANT: 38, SLIGHTLY IMPORTANT: 6 |
| r3JC\_COMFORT | 14 | 0.97 | TRUE | 5 | NOT IMPORTANT: 300, IMPORTANT: 96, MODERATELY IMPORTANT: 34, SLIGHTLY IMPORTANT: 6 |
| r3JD\_OTHERS | 20 | 0.96 | TRUE | 5 | NOT IMPORTANT: 267, IMPORTANT: 82, MODERATELY IMPORTANT: 51, SLIGHTLY IMPORTANT: 16 |

Table 6: Variables regarding decision-making and planning of respondents

Table of social interaction (see [Table 7](#tbl-Social-Interaction-Descriptives)) disclose factorial variables and information using Likert scale to have an ordered data set. 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. We have a large number of missing values in mode variable because most of people said no when they were asked to specify the mode of travel if they ever felt discriminated and most of people who felt discrimination they connected it to using bus.

Information about discrimination while traveling by different modes of travel, would be a kind of racism and can affect people who can not afford private mode of travel. It may also lead to conflicting approach to cope with it as promoting private mode would be a danger to eviornment. So there might be a transport inequality needed to being tackled by urban planner, environmentalist or any other descion makers. Also getting familiar with the level of social interaction and its importance from respondents’ viewpoint will give an insight into certainty or uncertainty about the conditions among members of a social network, people’s tendency towards collaboration with others in different experiences like tourism networks and so on.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_SI |
| 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 7: Variables regarding social interactions of respondents

It can be seen there are different variables in nature and sustainability table (see [Table 8](#tbl-Nature-Nustainability-Descriptives)) more through using Likert scale to organize respondents’ answer to variables.We can see lots of missing values for changing the mode because most of people do not tend to change their main mode of travel when it comes to a climatic event like heavy rain or flood and for whom wants to change, using car has the most priority. Most of people have poor level of access to the currently available sustainable modes of transport (eg hybrid buses, electric cars, public bicycles) and they assign high level of importance to that. Again we have so many missing value in payment variable because almost half of people would be willing to spend more on transportation to gain access to more sustainable modes and they indicate 5-15% of their payments they tend to spend in this regard.In terms of level of importance of improving different aspects in public transport routes most respondents assume high level of importance to presence of trees, access to parks, access to sustainable transport modes and broaden supply of sustainable transport modes. Nature and sustainablity information will helps planner to provide people with a safe, convenient and evironment-friendly transportation system. This will also boost the quality of life, people’s mobility and goods and enhance economic growth through erfficient transpotation services.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_NS |
| Number of rows | 451 |
| Number of columns | 10 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| factor | 10 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: factor**

| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| --- | --- | --- | --- | --- | --- |
| r5A\_CHANGE | 6 | 0.99 | FALSE | 2 | NO: 256, YES: 189 |
| r5B\_CHANGE\_MODE | 284 | 0.37 | FALSE | 8 | CAR: 79, METRO: 34, TAXI: 27, COLECTIVO: 13 |
| r5C\_SUST | 7 | 0.98 | TRUE | 5 | POOR: 189, FAIR: 97, GOOD: 86, VERY GOOD: 42 |
| r5D\_IMP\_SUST | 6 | 0.99 | TRUE | 5 | VERY IMPORTANT: 195, IMPORTANT: 115, MODERATELY IMPORTANT: 86, NOT IMPORTANT: 26 |
| r5E\_PAYMENT | 6 | 0.99 | FALSE | 2 | YES: 228, NO: 217 |
| r5F\_PAYMENTS | 236 | 0.48 | TRUE | 2 | 5-15%: 144, 15-30%-: 71, 30% or more%: 0 |
| r5GA\_TREE | 16 | 0.96 | TRUE | 5 | VERY IMPORTANT: 236, IMPORTANT: 91, MODERATELY IMPORTANT: 74, NOT IMPORTANT: 19 |
| r5GB\_PARK | 12 | 0.97 | TRUE | 5 | VERY IMPORTANTORTANT: 254, IMPORTANT: 106, MODERATELY IMPORTANT: 58, SLIGHTLY IMPORTANT: 11 |
| r5GC\_MODE | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 236, IMPORTANT: 105, MODERATELY IMPORTANT: 63, SLIGHTLY IMPORTANT: 25 |
| r5GD\_MODE | 11 | 0.98 | TRUE | 5 | VERY IMPORTANT: 258, IMPORTANT: 102, MODERATELY IMPORTANT: 48, SLIGHTLY IMPORTANT: 22 |

Table 8: Variables regarding perspectives about nature and sustainability of respondents

The information, telecommunications and mode shifting of respondents (see [Table 9](#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 desiging 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.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_IS |
| 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 9: Variables regarding information and telecommunications and mode shifting of respondents

Built environment data set (see [Table 10](#tbl-Built-Environment-Descriptives)) gives intuition about how people evaluate and assign importance to facilities and infrastructures such as parking sites, quality of highways, spaces for pedestrian based on Likert scales and with mainly complete responses to the variables. Most of people consider a good level when they are asked to evaluate space for cars and amount of parking spaces near their home and felt a high level of importance. Lots of people assign a very good and good level to the quality of highways and both space for pedestrians and quality of sidewalks located near of their home, respectively and it is very important to them. Unfortunately while it is important for people to have clean bus stops, sitting areas in bus stops, weather protection in bus stops,amount of cycle ways, quality of cycle ways and bike sharing schemes, they felt poor level when they were asked to evaluate them.

Decision makers would use this information in different contexts such as planning for arrangement and appearance of the various physical elements in a transportation system. Some of them may focus on enriching the environment in a way that they can encourage people to use more active modes of travel especially walking to reach a destination. Some of them would contribute to climate change background and study human activity to identify how transportation system, constructions, land-use and so on could be destructive due to high green house gas emissions.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_BE |
| 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 GOODD: 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 10: Variables regarding the built environment at the place of residence of respondents

The data set of commuting behavior of respondents(see [Table 11](#tbl-Travel-work-Descriptives)) indicates several factorial variables mostly using Likert scales. While most of people were neutral when they think their access to the transport network has affected their chances of having a better job, others assign a goof level of job opportunities available in their commune of residence. In terms of the level of access to employment are people interested in having in their commune of residence most of them tend to have the highest level and most of them are very satisfied with their current job. Most of people spend more than one hour in their regular commuting and their frequent time of commuting for them is between 7:00 - 9:00. Also many people spend 35.000-75.000 monthly on their transportation expenditure.

commuting behavior of respondents will be helpful when it comes to accessibility and commuting time. This table provides useful information which can play a role in promoting efficient road network wich is not too much time consuming and also provide users with a efficacious network.

(a) Data summary

|  |  |
| --- | --- |
| Name | Santiago\_TW |
| 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 GOODD: 109, GOOD: 93, FAIR: 27 |
| r8D\_EMPLSATISF | 37 | 0.92 | TRUE | 5 | HIGH SATISFIED: 135, VHIGH SATISFIED: 130, MEDIUM SATISFIED: 116, LOW SATISFIED: 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 11: Variables regarding the commuting behavior of respondents

## 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

Arribas-Bel, Dani, Mark Green, Francisco Rowe, and Alex Singleton. 2021. “Open Data Products-a Framework for Creating Valuable Analysis Ready Data.” Journal Article. *Journal of Geographical Systems* 23 (4): 497–514. <https://doi.org/10.1007/s10109-021-00363-5>.