Commuting to college. An analysis of a suburban campus on the outskirts of Madrid

Mary Luz Mouronte-López[a,*], Alberto López[a]

[a]*Higher Polytechnic School, Francisco de Vitoria University, Madrid, Spain*

S1. Surveys

S1.1. Sample selection. Calculation of ideal size

For all analysed courses, both university population (N) and statistical sample are shown in Table S1. To determine an adequate (ideal) sample size (sz), a modification of the Cochran Formula Statistics How To [1] was applied. The following steps were carried out:

- Establish the margin of error (e).
- For a specific confidence level (α), determine the Z-score (z_α).
- For N, establish the estimated proportion (p) of the attribute that would be present in the population.

For a finite population (N), sz is defined as:

$$sz = \frac{N * z_\alpha^2 * p * (1-p)}{e^2 * (N-1) + (z_\alpha)^2 * p * (1-p)}$$

For the examined courses, for $p = 0.5$[1], $\alpha = 0.95$, and z_α=1.96, sz is shown in Table S1. Since the available sample is larger than sz, it is appropriate for this research.

*Corresponding author

Email address: maryluz.mouronte@ufv.es (Mary Luz Mouronte-López)

[1]Since p can take values in the range $[0, 1]$, $p*(1-p)$ is in the range $[0, 1]$. The value of p that maximises $p * (1-p)$ corresponds to $p = 0.5$. Therefore, if no information available exists to approximate p, then $p = 0.5$ can be utilised to produce the most conservative, or largest, sample size

Table S1: For all analysed courses, statistical sample & population (N)U. F. V. [2], U. F. V. [3], U. F. V. [4]. For $p = 0.5$, $\alpha = 0.95$, and z_α=1.96, size of the ideal sample (sz).

Year	Total number of individuals	population (N)	Ideal sample (sz)	Rounded ideal sample
2017-2018	3,992	12,293	981.95	982
2018-2019	2,532	13,511	989.06	989
2021-2022	3,250	17,780	1006.75	1,007

The adequacy of the sample size was also checked for all groups (individuals' profile and educational levels) of respondents to the survey. Identical values of p, α, and z_α were used in the verification. The following groups were considered:

- Individuals in undergraduate level.
- Individuals in graduate level.
- Individuals in other studies (vocational studies).
- Individuals in Le Cordon Blue.
- Individuals in administrative areas (Administrative staff).
- Teachers and researchers.

S1.2. Surveys of the 2017-2018 and 2018-2019 academic year

Only those used questions in this research have been included in this section, which were:

- Studies you are pursuing: (i) Degree, (ii) Postgraduate, (iii) Professional training, (iv) Le Cordon Bleu
- Postcode where you usually start your journey to university
- Community

- Which days do you travel to the university (you can indicate more than one option)?
  (i)Monday, (ii) Tuesday, (iii) Wednesday, (iv) Thursday, (v) Friday, (vi) Saturday, (vii) Sunday, (viii) All

- At what time do you travel to the university (you can indicate more than one option)?
  (i) Morning (between 07:00 and 14:00 hours), (ii) Afternoon (between 14:00 and 22:00 hours), (iii) Morning and afternoon (between 07:00 and 22:00 hours)

- Which are the transport modes that you usually use (more than one option is possible)?

  - Public transport (i) Bus, (ii) Subway, (iii) Commuters Train
  - University Transport
  - Private Car
  - Private Motorbike
  - Bicycle
  - Walking

- If you use a car to commute to university, which is its motorisation? (i) gasoline, (ii) diesel, (iii) hybrid, (iv) electric (v) unknown.

- If you use a car, how many people usually travel with you?
  (i) travelling alone, (ii) 1 person, (iii) 2 persons, (iv) 3 persons, (v) 4 persons, (vi) + 4 persons

- Approximate daily distance (km) used for university commuting (include one way only):
  (i) Less than 10 km, (ii) Between 10 and 30 km. (iii) Between 30 and 60 km. (iv) More than 60 km.

- Time taken to get to the university

  (i) Less than 15 minutes, (ii) Between 15 and 30 minutes, (iii) Between 30 and 60 minutes, (iv) More than 1 hour

- Is it comfortable for you to travel to the university (you can indicate more than one reason)?
  - Yes
  - No. Why? (i) Time, (ii) Distance , (iii) Traffic, (iv) Public Transport Insufficient, (v) Parking, (vi) Others
- Do you know and have you used platforms such as Blablacar/Emov/Bluemove/Car2go/Amovens...?

  (i) Yes, I know them but I have never used them, (ii) Yes, I know them and I have used them, (iii) I do not know them

S1.3. Survey 2021-2022 year

Only those used questions in this research have been included, they were:

- Which is your main profile? (i) Student, (ii) Campus administration and services staff
- Which is the postcode corresponding to the start of the usual journey to the university
- What days and times do you come to the university?
  - from Monday to Friday. Indicates at which time slot during the day (e.g. from 10 to 11 o'clock in the morning).
  - On Monday. Indicates at which time slot during the day (e.g. from 10 to 11 o'clock in the morning).

  - On Tuesday. Indicates at which time slot during the day (e.g. from 10 to 11 o'clock in the morning).
  - On Wednesday. Indicates at which time slot during the day (e.g. from 10 to 11 o'clock in the morning).
  - On Thursday. Indicates at which time slot during the day (e.g. from 10 to 11 o'clock in the morning).
  - On Friday. Indicates at which time slot during the day (e.g. from 10 to 11 o'clock in the morning).
  - on Weekends. Indicates at which time slot during the day (e.g. from 10 to 11 o'clock in the morning).

- How comfortable is it for you to travel to the university?. on a scale from 1 to 5. 1: very uncomfortable, 5: very comfortable.

- Which is the transport mode you usually use at every stage of the journey? Choose one in each step:
  - ON FOOT - Walking,
  - BICYCLE - Private electric bicycle,
  - BICYCLE - Private non-electric bicycle,
  - BUS - Alcorcón-University,
  - BUS - Aluche-University (561-561-A-561-B - AVANZA),
  - BUS - Aluche-University,
  - BUS - Boadilla-University (565),
  - BUS - Directa Moncloa-University (659 - AVANZA),
  - BUS - Shuttle University-MSI,
  - BUS - Las Rozas-University (with stops in Majadahonda),
  - BUS - Moncloa University with stops (659 - AVANZA),
  - BUS - Otras Bus Interurbano Público,

  - BUS - Other Private Intercity Bus,
  - BUS - Plaza Castilla-University,
  - BUS - Príncipe Pio-MSI,
  - BUS - Urban bus within a municipality other than Madrid,
  - BUS - Madrid EMT urban bus (within Madrid),
  - COMPANY CAR - Carsharing fleet (Zity/Sharenow/Free2move/Ubeeqo/Wible/Goto/Electricway/Bluemove or similar),
  - PRIVATE CAR - Private car as unorganised driver (own car),
  - PRIVATE CAR - Private car as an organised driver with an App platform (University car share/Blablacar/Socialcar/Amovens or similar)
  - PRIVATE CAR - Private car as an organised driver with an App platform (University car share/Blablacar/Socialcar/Amovens or similar),
  - PRIVATE CAR - Private car as an unorganised passenger (with a family member or friend but without an App platform or similar)
  - PRIVATE CAR - Private car as an organised passenger with an App platform (University care share/Blablacar/Socialcar/Amovens or similar),
  - METRO - Light rail/tram,
  - METRO - Metro,
  - MOTO - Private motorbike/moped,
  - OTHER - Other,
  - TAXI - Taxi,
  - TRAIN - Renfe Cercanías,
  - TRAIN - Rest of Renfe,
- Indicates the approximate time for each section in minutes.
  - 10 to 15 minutes

  - 15 to 20 minutes
  - 20 to 25 minutes
  - 25 to 30 minutes
  - 30 to 35 minutes
  - 35 to 40 minutes
  - 40 to 45 minutes
  - 45 to 50 minutes
  - 5 to 10 minutes
  - 50 to 55 minutes
  - 55 to 60 minutes
  - 60 to 70 minutes
  - 70 to 90 minutes
  - More than 90 minutes
  - Less than 5 minutes

- Indicate, if known, the motorization of the car you usually use when you travel to the campus (i) diesel, (ii) pure electric, (iii) gasoline, (iv) LPG, (v) plug-in hybrid, (vi) non-plug-in hybrid, (vii) unknown.

S2. Postcodes

- Postcodes corresponding to: Pozuelo de Alarcón (28223 and 28224), Boadilla del Monte (28660), Universidad (28015), Aravaca (28023), Moncloa (28008), Las Rozas de Madrid (28232), Galapagar, Torrelodones (28250), Chamberí - Rios Rosas (28003), Collado Villalba (28400) were among the most common postcodes for individuals who travelled to the Universidad Francisco de Victoria during the three academic years.

- 56 postcodes in which information on gross income in 2020 was available were: 28001 (Salamanca – Goya), 28002 (Chamartín), 28003 (Chamberí

– Rios Rosas), 28004 (Justicia), 28005 (Arganzuela), 28006 (Castellana), 28007 (Adelfas), 28008 (Moncloa), 28009 (Ibiza-Niño Jesús), 28010 (Almagro), 28011 (Latina), 28012 (Embajadores), 28013 (Centro), 28014 (Retiro), 28015 (Universidad), 28016 (Hispanoamerica – Costillares), 28017 (Ventas – Ciudad Lineal – La Elipa), 28018 (Palomeras), 28019 (Opañel – Comillas – San Isidro), 28020 (Valdeacederas), 28021 (Villaverde - S Cristóbal de Los Ángeles), 28022 (San Blas), 28023 (Aravaca), 28024 (Campamento), 28025 (Carabanchel – Abrantes), 28026 (Usera), 28027 (San Juan Bautista), 28028 (Guindalera – Fuente El Berro), 28029 (Tetuán – Barrio Del Pilar), 28030 (Moratalaz), 28031 (Villa Vallecas–Santa Eugenia), 28032 (Vicalvaro–Valdebernardo), 28033 (Hortaleza–Pinar Del Rey), 28034 (Tres Olivos), 28035 (Fuencarral), 28036 (Nueva España), 28037 (San Blas), 28038 (Pavones), 28039 (Bellas Vistas–Valdeacederas), 28040 (Ciudad Universitaria), 28041 (Orcasitas – San Fermín), 28042 (Barajas – La Alameda de Osuna), 28043 (Hortaleza – Canillas), 28044 (Cuatrovientos), 28045 (Delicias), 28046 (Castilla–Chamartín), 28047 (Aluche – Vista Alegre), 28048 (Fuencarral El Pardo), 28049 (El Goloso), 28050 (Las Tablas), 28051 (Villa de Vallecas), 28052 (El Cañaveral), 28053 (Entrevías – Puente de Vallecas), 28054 (Buenavista – La Fortuna), 28055 (Valdebebas - Valdefuentes).

S3. In-house software

Several programs in R language were developed in which the following features were implemented:

- Transformation of the 2021-2022 survey format to that of the surveys carried out in the 2017-2018 and 2018-2019 courses, which was necessary in order to compare the courses with each other.
- Examination of cumulative probability distributions and statistical quartiles of the analysed variables, as well as graphs plotting.

- Fitting cumulative probability distributions, through univariate distributional regression models, to functions based on the (Akaike Information Criterion (AIC) and Schwarz Bayesian criterion (SBC) metrics.
- Plotting of maps relating to the Madrid Community, showing for the individuals surveyed, the number of them residing in each postcode. Individuals using each mode of transport, as well as the distribution of stops by public transport mode are also displayed.

The following packages in R were utilised: `data.table`, `dplyr`, `dslabs`, `ggplot2`, `ggspatial`, `Hmisc`, `nortest`, `patchwork`, `raster`, `readr`, `readxl`, `sf`, `sm`, `sp`, `stats`, `stringr`, `tidyr`, `tidyverse`, `writexl`, `gamlss`, `NbClust`. `gamlss` in R was utilised Rigby & Stasinopoulos [5], Stasinopoulos et al. [6], Hastie & Tibshirani [7].

S4. Tables

Table S2: For undergraduate students, percentage of usage. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

	2017-2018 survey	2018-2019 survey
PT	49.61%	59.99%
UT	47.24%	53.21%
CAR	49.13%	38.86%
M	1.89%	1.59%
BI	0.66%	0.87%
W	4.76%	5.84%

Table S3: For undergraduate students, for the 2021-2022 course, in each step, percentage of use of each transport type. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

Transport Type	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
PT	29.68%	40.88%	22.07%	5.62%	0.47%	0.24%
UT	3.69%	4.06%	4.87%	0.66%	0.19%	0%
CAR	32,56%	1.80%	0.14%	0.05%	0%	0%
M	0.19%	0.14%	0.38%	0%	0%	0%
BI	0.05%	0.05%	0%	0%	0%	0%
W	31.95%	0.76%	0%	1.56%	0.66%	0.09%
Unfilled	1.89%	52.32%	69.57%	92.01%	98.63%	99.67%

Table S4: For graduate students, percentage of usage. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

	2017-2018 survey	2018-2019 survey
PT	23.81%	26.94%
UT	32.97%	42.47%
CAR	63.00%	52.51%
M	2.93%	1.83%
BI	0%	0.46%
W	0.73%	1.83%

Table S5: For graduate students, for the 2021-2022 course, in each step, percentage of use of each transport type. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

Transport Type	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
PT	15.58%	19.48%	11.69%	3.90%	0%	0%
UT	5.19%	1.30%	0%	1.30%	0%	0%
CAR	62.34%	0%	0%	0%	0%	0%
M	0%	0%	0%	0%	0%	0%
BI	0%	0%	0%	0%	0%	0%
W	14.29%	1.30%	1.30%	0%	0%	0%
Unfilled	2.60%	77.92%	88.31%	94.81%	100%	100%

Table S6: For vocational, percentage of usage. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

	2017-2018 survey	2018-2019 survey
PT	51.74%	46.13%
UT	40.93%	43.77%
CAR	53.28%	40.40%
M	2.70%	2.02%
BI	0.39%	0.67%
W	1.93%	2.36%

Table S7: For vocational studies, for the 2021-2022 course, in each step, percentage of use of each transport type. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

Transport Type	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
PT	35.83%	28.88%	14.97%	3.74%	0.53%	0%
UT	3.74%	1.07%	3.74%	0%	0%	0%
CAR	36,90%	2.67%	0.53%	0%	0%	0%
M	0,53%	0%	0%	0%	0%	0%
BI	0%	0%	0%	0%	0%	0%
W	18,18%	0.53%	1.60%	1.07%	0.53%	0%
Unfilled	4.81%	66.84%	79.14%	95.19%	98.93%	0%

Table S8: For Le Cordon Bleu, percentage of usage. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

	2017-2018 survey	2018-2019 survey
PT	24.94%	37.60%
UT	40.56%	36.00%
CAR	45.56%	44.80%
M	1.67%	1.60%
BI	0%	0%
W	1.11%	0.80%

Table S9: For Le Cordon Bleu, for the 2021-2022 course, in each step, percentage of use of each transport type. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

Transport Type	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
PT	13.04%	23.91%	6.52%	2.17%	0%	0%
UT	10.87%	6.52%	13.04%	0%	2.17%	0%
CAR	52.17%	0%	0%	0%	0%	0%
M	2.17%	0%	0%	0%	0%	0%
BI	0%	0%	0%	0%	0%	0%
W	21.74%	0%	2.17%	0%	0%	0%
Unfilled	%	69.57%	78.26%	97.83%	97.83%	100%

Table S10: For Administrative Staff, percentage of usage. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

	2017-2018 survey	2018-2019 survey
PT	4.35%	11.79%
UT	8.70%	14.63%
CAR	91.30%	84.96%
M	0%	2.85%
BI	0%	0%
W	0%	2.03%

Table S11: For Administrative Staff, for the 2021-2022 course, in each step, percentage of use of each transport type. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

Transport Type	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
PT	8.33%	9.56%	4.17%	1.47%	0.25%	0.25%
UT	0.74%	1.23%	0.98%	0.25%	0%	0%
CAR	79.41%	3.68%	0%	0%	0%	0%
M	1.96%	0.25%	0%	0%	0%	0%
BI	0%	0%	0%	0%	0%	0%
W	8.58%	0.25%	0.98%	0%	0%	0%
Unfilled	0.98%	85.05%	93.87%	98.28%	99.26	100%

Table S12: For Researchers & Teachers, for the 2021-2022 course, in each step, percentage of use of each transport type. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

Transport Type	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
PT	5.54%	5.54%	3.02%	0.25%	0%	0%
UT	1.01%	1.26%	0.25%	0%	0%	0%
CAR	84.13%	3.53%	0.25%	0%	0%	0%
M	2.27%	0.50%	0%	0%	0%	0%
BI	0.50%	0.25%	0%	0%	0%	0%
W	5.79%	0.25%	0.50%	0%	0%	0%
Unfilled	0.76%	88.66%	95.97%	95.97%	0%	0%

Table S13: For undergraduate students, similarity between cumulative probability distributions of transport utilisation ratios. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking. The obtained $p-value$ in the Kolmogorov Smirnov test is shown.

	2017-2018 survey						2018-2019 survey					
	PT	UT	CAR	B	BI	W	PT	UT	CAR	B	BI	W
PT	1	0.34	0	0	0	0	1	0.01	0	0	0	0
UT	0.34	1	0	0	0	0	0.01	1	0	0	0	0
CAR	0	0	1	0	0	0	0	0	1	0	0	0
M	0	0	0	1	0.97	0.15	0	0	0	1	1	1
BI	0	0	0	0.97	1	0	0	0	0	1	1	1
W	0	0	0	0.15	0	1	0	0	0	1	1	1

	2021-2022 survey					
	PT	UT	CAR	B	BI	W
PT	1	0	0	0	0	0
UT	0	1	0	0	0	0
CAR	0	0	1	0	0	0
M	0	0	0	1	1	0
BI	0	0	0	1	1	0
W	0	0	0	0	0	1

Table S14: For graduate students, similarity between cumulative probability distributions of transport utilisation ratios. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking. The obtained $p - value$ in the Kolmogorov Smirnov test is shown.

	2017-2018 survey						2018-2019 survey					
	PT	UT	CAR	B	BI	W	PT	UT	CAR	B	BI	W
PT	1	0.20	0	0	0	0	1	0.01	0	0	0	0
UT	0.20	1	0	0	0	0	0.01	1	0	0	0	0
CAR	0	0	1	0	0	0	0	0	1	0	0	0
M	0	0	0	1	1	1	0	0	0	1	1	1
BI	0	0	0	1	1	1	0	0	0	1	1	1
W	0	0	0	1	1	1	0	0	0	1	1	1

	2021-2022 survey					
	PT	UT	CAR	B	BI	W
PT	1	1	0	0	0	0.67
UT	1	1	0	0.02	0.03	0.91
CAR	0	0	1	0	0	0
B	0	0.03	0	1	1	0.31
BI	0	0.03	0	1	1	0.31
W	0.67	0.91	0	0.31	0.31	1

Table S15: For vocational training students, similarity between cumulative probability distributions of transport utilisation ratios. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking. The obtained $p-value$ in the Kolmogorov Smirnov test is shown.

	2017-2018 survey						2018-2019 survey					
	PT	UT	CAR	B	BI	W	PT	UT	CAR	B	BI	W
PT	1	0.10	0	0	0	0	1	1	0	0	0	0
UT	0.10	1	0	0	0	0	1	1	0	0	0	0
CAR	0	0	1	0	0	0	0	0	1	0	0	0
M	0	0	0	1	1	1	0	0	0	1	1	1
BI	0	0	0	1	1	1	0	0	0	1	1	1
W	0	0	0	1	1	1	0	0	0	1	1	1

	2021-2022 survey					
	PT	UT	CAR	B	BI	W
PT	1	0.25	0.01	0	0	0
UT	0.25	1	0	0	0	0.04
CAR	0.01	0	1	0	0	0
B	0	0	0	1	1	0
BI	0	0	0	1	1	0
W	0	0	0	1	1	0

Table S16: Le Cordon Bleu training students, similarity between cumulative probability distributions of transport utilisation ratios. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking. The obtained $p-value$ in the Kolmogorov Smirnov test is shown.

	2017-2018 survey						2018-2019 survey					
	PT	UT	CAR	B	BI	W	PT	UT	CAR	B	BI	W
PT	1	0.02	0	0	0	0	1	1	0	0	0	0
UT	0.02	1	0	0	0	0	1	1	0	0	0	0
CAR	0	0	1	0	0	0	0	0	1	0	0	0
M	0	0	0	1	1	1	0	0	0	1	1	1
BI	0	0	0	1	1	1	0	0	0	1	1	1
W	0	0	0	1	1	1	0	0	0	1	1	1

	2021-2022 survey					
	PT	UT	CAR	B	BI	W
PT	1	0	0	0	0	0
UT	0	1	0	0	0	0
CAR	0	0	1	0	0	0
B	0	0	0	1	1	0
BI	0	0	0	1	1	0
W	0	0	0	0	0	1

Table S17: Administrative staff, similarity between cumulative probability distributions of transport utilisation ratios. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking. The obtained $p - value$ in the Kolmogorov Smirnov test is shown.

	2017-2018 survey						2018-2019 survey					
	PT	UT	CAR	B	BI	W	PT	UT	CAR	B	BI	W
PT	1	1	0	1	1	1	1	1	0	0.28	0.07	0.19
UT	1	1	0	1	1	1	1	1	0	0.07	0.01	0.04
CAR	0	0	1	0	0	0	0	0	1	0	0	0
M	1	0	0	1	1	1	0.28		0.07	0	1	1
BI	1	0	0	1	1	1	0.07	0.01	0	1	1	1
W	1	1	0	1	1	1	0.19	0.04	0	1	1	1

	2021-2022 survey					
	PT	UT	CAR	B	BI	W
PT	1	0.59	0	0	0	0.71
UT	0.59	1	0	0.38	0.10	1
CAR	0	0	1	0	0	0
B	0	0.38	0	1	1	0.29
BI	0	0.10	0	1	1	0.07
W	0.71	1	0	0.29	0.07	1

Table S18: Researchers & teachers staff, similarity between cumulative probability distributions of transport utilisation ratios. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking. The obtained $p-value$ in the Kolmogorov Smirnov test is shown.

	2017-2018 survey						2018-2019 survey					
	PT	UT	CAR	B	BI	W	PT	UT	CAR	B	BI	W
PT	1	0.98	0	1	1	0.94	1	0.49	0	1	0.94	0.63
UT	0.98	1	0	0.73	0.61	0.28	0.49	1	0	0.26	0.06	0.02
CAR	0	0	1	0	0	0	0	0	1	0	0	0
M	1	0.73	0	1	1	1	1	0.26	0	1	1	1
BI	0.94	0.06	0	1	1	1	1	0.01	0	1	1	1
W	0.94	0.28	0	1	1	1	0.63	0.02	0	1	1	1

	2021-2022 survey					
	PT	UT	CAR	B	BI	W
PT	1	0	0	0	0	0
UT	0	1	0	0	0	0
CAR	0	0	1	0	0	0
B	0	0	0	1	1	0
BI	0	0	0	1	1	0
W	0	0	0	0	0	1

Table S19: For 2017-2018, 2018-2019, 2021-2022 survey, numbers of clusters, which were calculated utilising the following indexes KL: KL Krzanowski & Lai [8], CH: CH Caliński & Harabasz [9], HART: Hartigan Hartigan [10], CCC: CCC Sarle [11], SCO: Scott Scott & Symons [12], MAR: Marriot Marriot [13], TRC: TrCovW Milligan & Cooper [14] , TRA: TraceW Milligan & Cooper [14], FRI: Friedman Friedman & Rubin [15] , RUB: Rubin Friedman & Rubin [15], CIN: Cindex Hubert & Levin [16], DB: DB Davies & Bouldin [17], SIL: Silhouette Rousseeuw [18], DUD: Duda Duda & Hart [19], PSE: PseudoT2 Duda & Hart [19], BEA: Beale Beale [20], RAT: Ratkowsky Rousseeuw [18], BAL: Ball Ball & Hall [21], PTB: PtBiserial Milligan [22], Milligan [23], Milligan [24], FRE: Frey Frey & Van Groenewoud [25], MCC: McClain McClain & Rao [26], DUN: Dunn Dunn [27], HUB: Hubert Hubert & Levin [16], SDI: SDindex Halkidi et al. [28], DIN: Dindex Lebart et al. [29], SDbw Halkidi2.

2017-2018 survey

KL	CH	HAR	CCC	SCO	MAR
5	2	4	2	3	3
TRC	**TRA**	**FRI**	**RUB**	**CIN**	**DB**	**SIL**
3	5	3	5	3	5	4
DUD	**PSE**	**BEA**	**RAT**	**BAL**	**PTB**
4	4	2	5	3	5
FRE	**MCC**	**DUN**	**HUB**	**SDI**	**DIN**
2	2	0	3	0	5
SDB
5

2018-2019 survey

KL	CH	HAR	CCC	SCO	MAR
5	3	4	5	3	3
TRC	**TRA**	**FRI**	**RUB**	**CIN**	**DB**	**SIL**
3	3	3	3	3	3	5
DUD	**PSE**	**BEA**	**RAT**	**BAL**	**PTB**
3	3	3	5	3	5
FRE	**MCC**	**DUN**	**HUB**	**SDI**	**DIN**
1	2	2	0	5	0
SDB
3

2021-2022 survey

KL	CH	HAR	CCC	SCO	MAR
4	2	4	5	3	4
TRC	**TRA**	**FRI**	**RUB**	**CIN**	**DB**	**SIL**
4	4	5	4	5	2	5
DUD	**PSE**	**BEA**	**RAT**	**BAL**	**PTB**
4	4	4	5	3	5
FRE	**MCC**	**DUN**	**HUB**	**SDI**	**DIN**
1	5	2	0	2	0
SDB
4

Table S20: For 2017-2018, 2018-2019, 2021-2022 surveys, mean for each cluster and trip mode. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

			2017-2018 survey							2018-2019 survey			
C	PT	UT	CAR	M	BI	W	C	PT	UT	CAR	B	BI	W
			Mean							Mean			
1	0.42	0.42	0.11	0	0	0.04	1	0.56	0.27	0.09	0.02	0	0.05
2	0	0	0.98	0.02	0	0	2	0	0.89	0.10	0	0	0
							3	0	0	0.98	0.01	0	0

			2021-2022 survey			
C	PT	UT	CAR	B	BI	W
			Mean			
1	0	0	0.99	0	0	0.01
2	0.77	0	0.03	0	0	0.19
3	0.23	0.50	0.02	0.02	0	0.24

Table S21: For 2017-2018, 2018-2019, 2021-2022 surveys, median for each cluster. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

			2017-2018 survey							2018-2019 survey			
C	PT	UT	CAR	M	BI	W	C	PT	UT	CAR	B	BI	W
			Median							Median			
1	0	0	1	0	0	0	1	0.5	0.5	0	0	0	0
2	0.5	0.5	0	0	0	0	2	1	0	0	0	0	0
							3	0	0	1	0	0	0

			2021-2022 survey			
C	PT	UT	CAR	B	BI	W
			Median			
1	0	0	1	0	0	0
2	1	0	0	0	0	0
3	0.33	0.5	0	0	0	0

Table S22: For 2017-2018, 2018-2019, 2021-2022 surveys, mode for each cluster. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

		2017-2018 survey							2018-2019 survey				
C	PT	UT	CAR	M	BI	W	C	PT	UT	CAR	B	BI	W
			Mode							Mode			
1	0.42	0.42	0.11	0.01	0	0.041	1	0	0.89	0.10	0	0	0
2	0	0	0.98	0.02	0	0	2	0.56	0.27	0.092	0.02	0	0.05
							3	0	0	0.98	0.01	0	0

	2021-2022 survey					
C	PT	UT	CAR	B	BI	W
			Mode			
1	0	0	1	0	0	0
2	1	0	0	0	0	0
3	0	0.33	0	0	0	0

Table S23: For 2017-2018, 2018-2019, 2021-2022 surveys, standard deviation for each cluster. PT: public transport, UT: university transport, CAR: private car, M: motorcycle, BI: bike, W: Walking

2017-2018 survey							2018-2019 survey						
C	PT	UT	CAR	B	BI	W	C	PT	UT	CAR	B	BI	W
	standard deviation							standard deviation					
1	0.42	0.42	0.11	0	0	0.04	1	0	0.89	0.10	0	0	0
2	0	0	0.98	0.02	0	0	2	0	0	0.98	0.01	0	0.05
							3	0.56	0.27	0.09	0.02	0	0.05

2021-2022 survey						
C	PT	UT	CAR	B	BI	W
	Standard Deviation					
1	0.77	0	0.03	0	0	0.19
2	0	0	0.99	0	0	0.01
3	0.23	0.50	0.02	0.02	0	0.23

Table S24: For each transport mode in steps 1 and 2, for the variable transport mode by postcode, obtained $p - value$ in the Shapiro-Wilks test

step	Commuters trains	Light subway	Interurban bus	Urban bus	Madrid urban bus	subway
1	$5.2e-16$	$< 2.2e-16$	$8.6e-11$	$4.0e-16$	$6.2e-13$	$1.6e-13$
2	$1.5e-14$	$< 2.2e-16$	$6.2e-10$	$6.1e-15$	$4.0e-15$	$10e-12$

Table S25: For each type of public transport in steps 1 and 2, for the number of stops by postcode variable, obtained $p - value$ in the Shapiro-Wilks test

step	Commuters trains	Light subway	Interurban bus	Urban bus	Madrid urban bus	Subway
1	$< 2.2e-16$	$< 2.2e-16$	$< 2.2e-16$	$< 2.2e-16$	$< 2.2e-16$	$< 2.2e-16$
2	$< 2.2e-16$	$< 2.2e-16$	$< 2.2e-16$	$< 2.2e-16$	$< 2.2e-16$	$7.78e-14$

S5. Figures

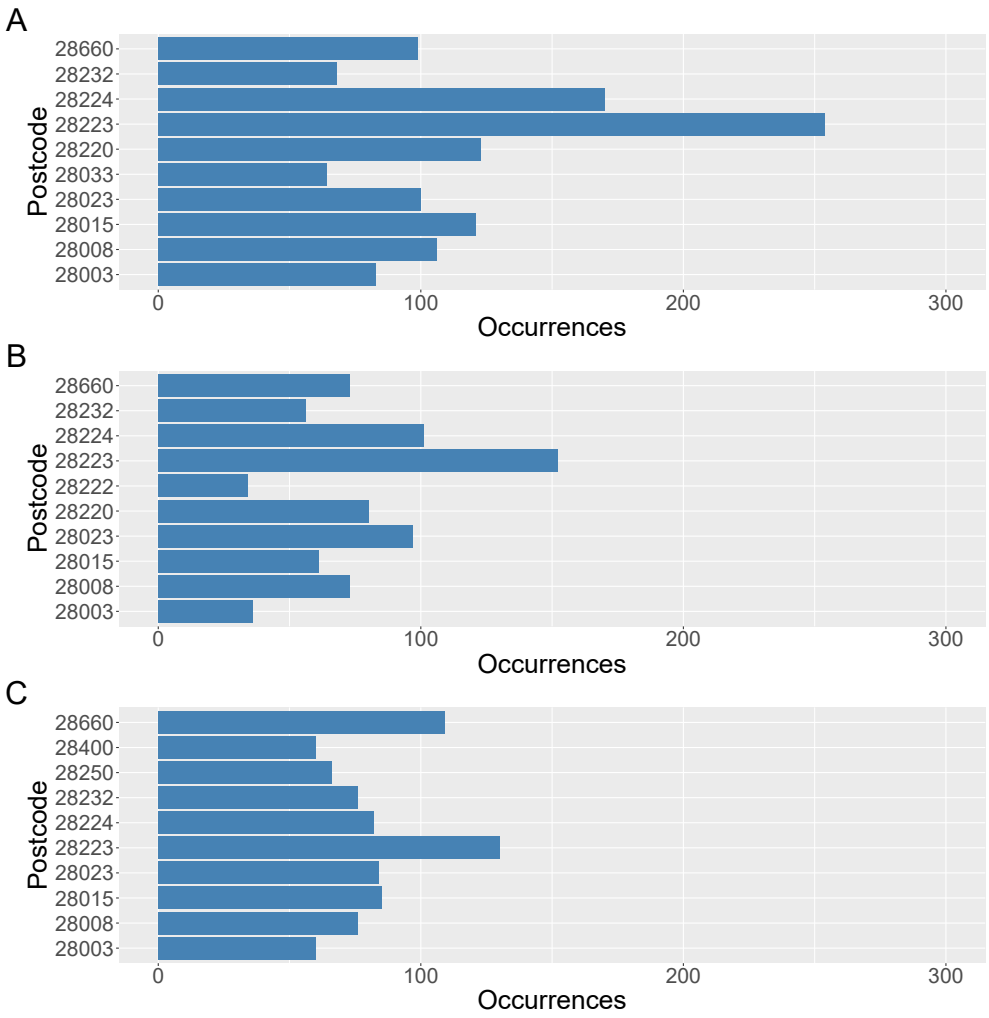

Figure S1: Most frequent postcodes in the 2017-2018, 2018-2019 and 2021-2022 courses .

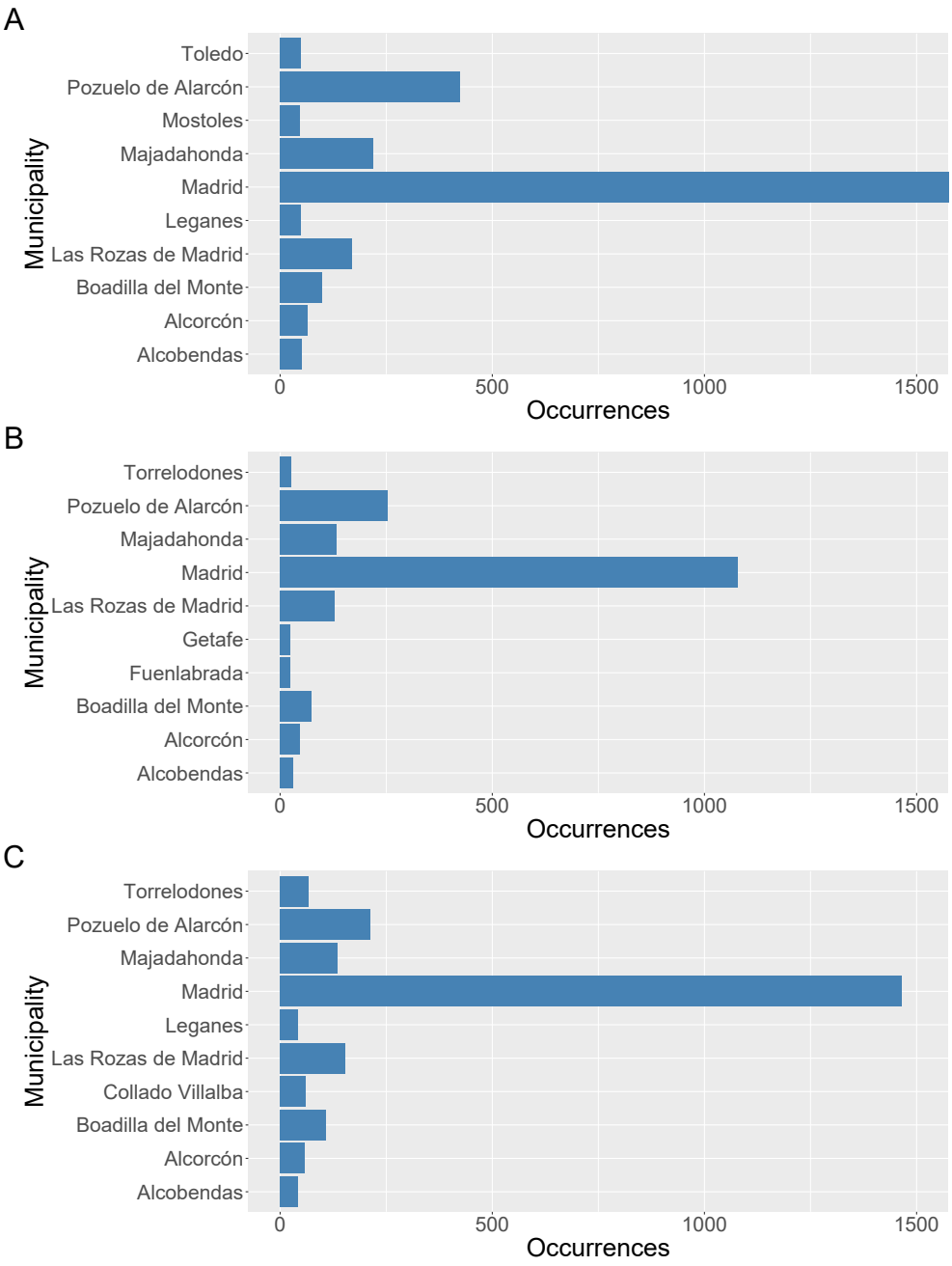

Figure S2: Most frequent municipalities in the 2017-2018, 2018-2019 and 2021-2022 courses .

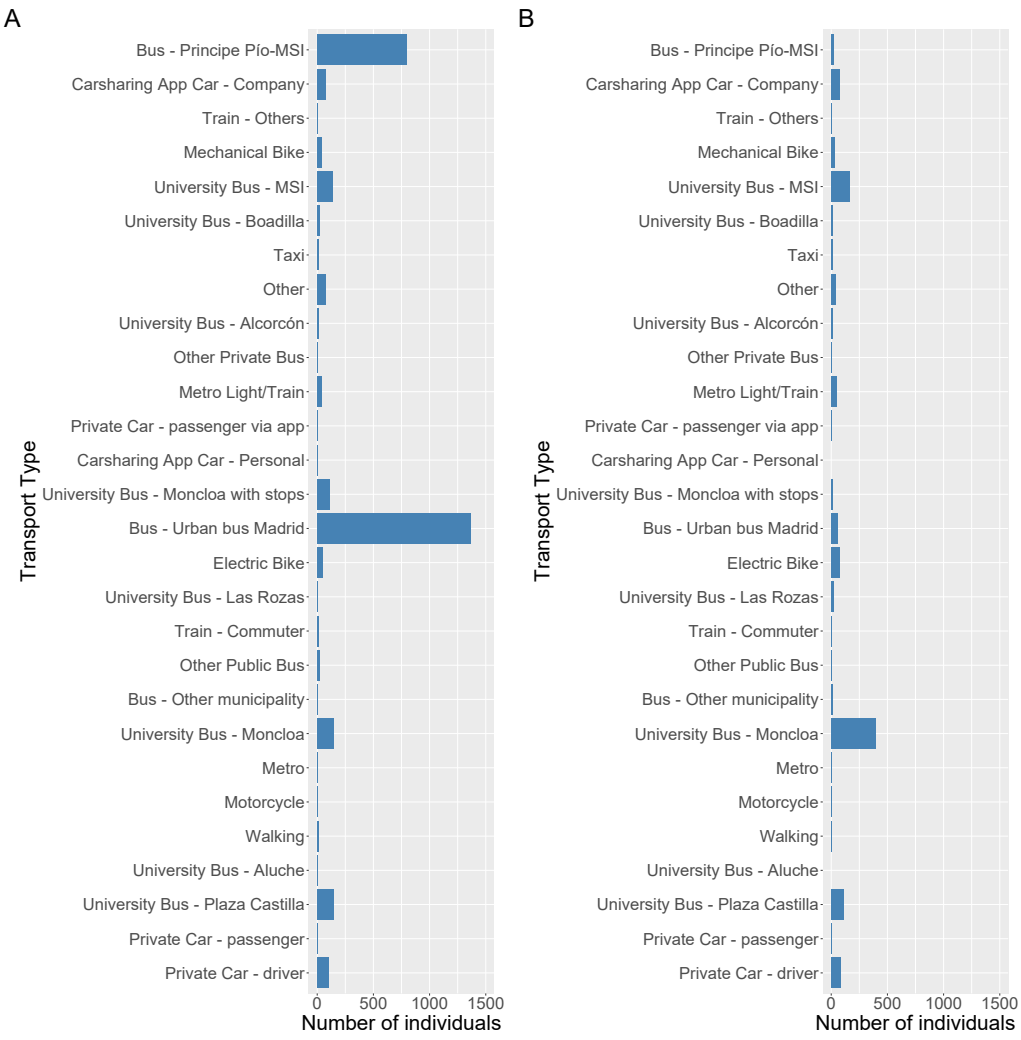

Figure S3: Histograms of type of transport in steps 1 and 2

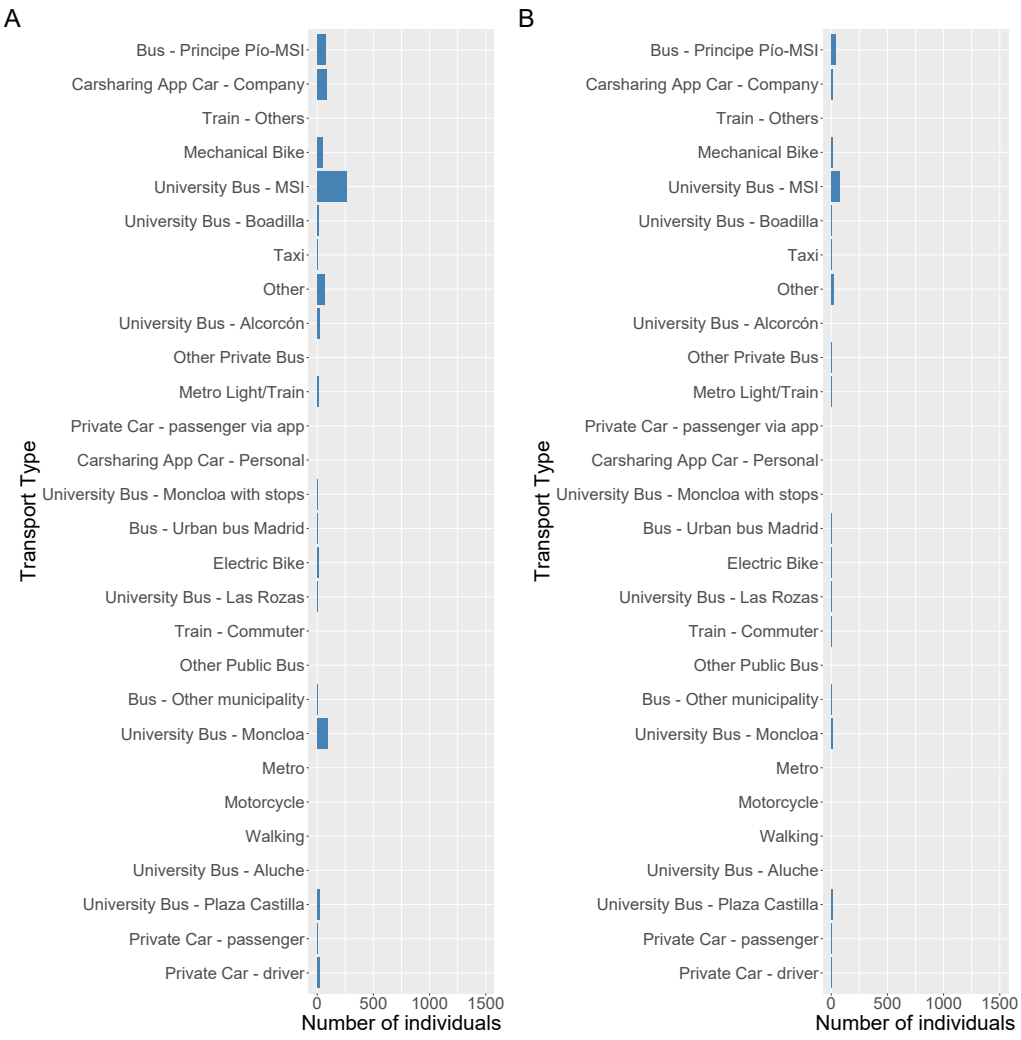

Figure S4: Histograms of type of transport in steps 3 and 4

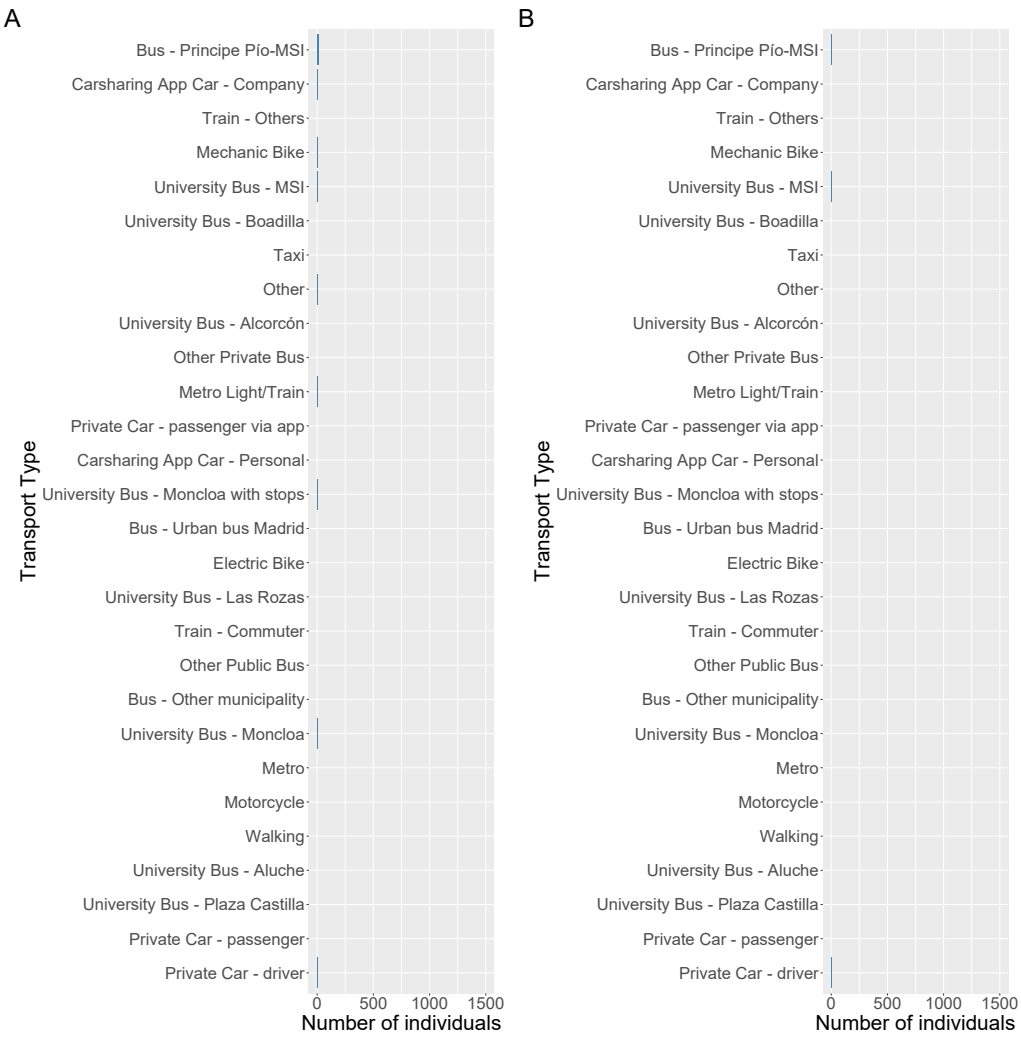

Figure S5: Histograms of type of transport in steps 5 and 6

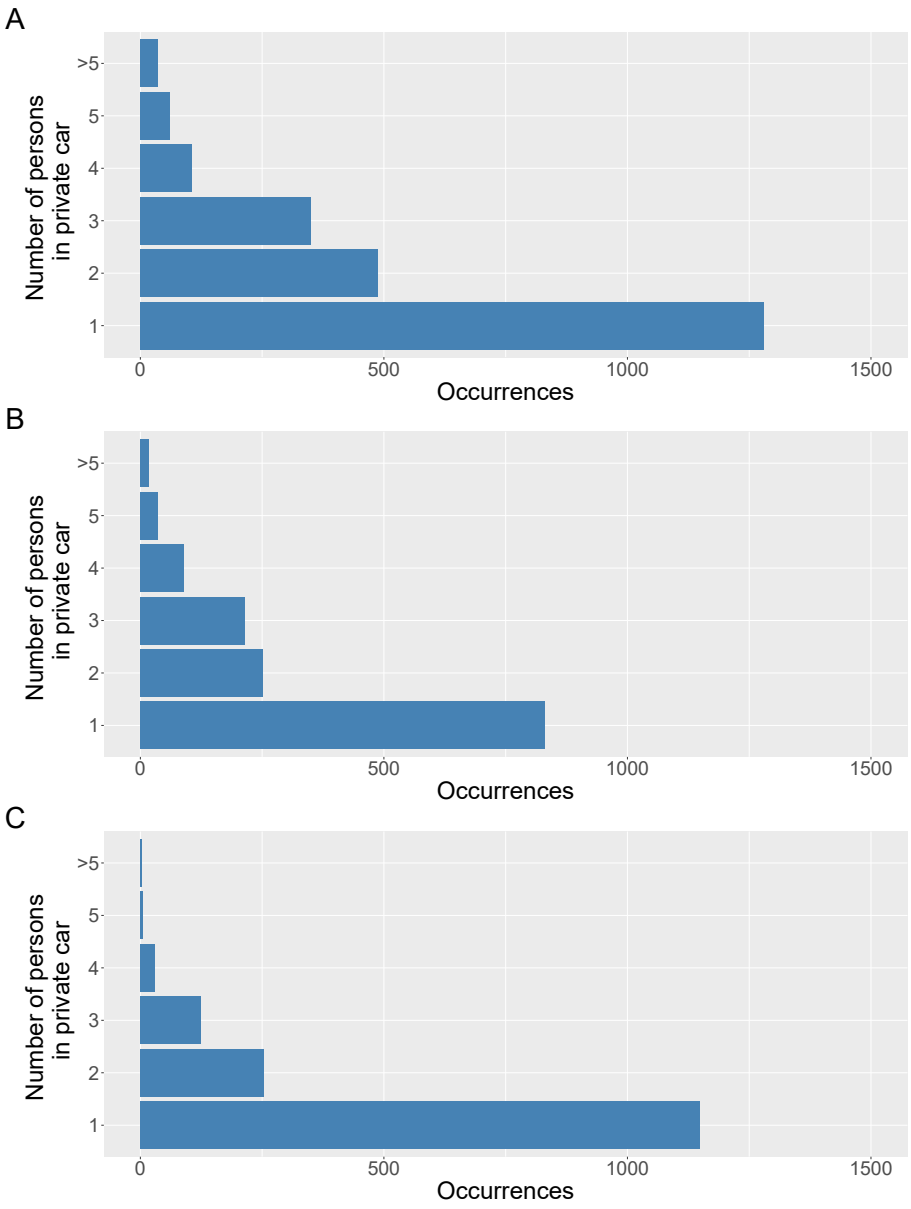

Figure S6: Histograms of number of persons travelling in a private car during a journey

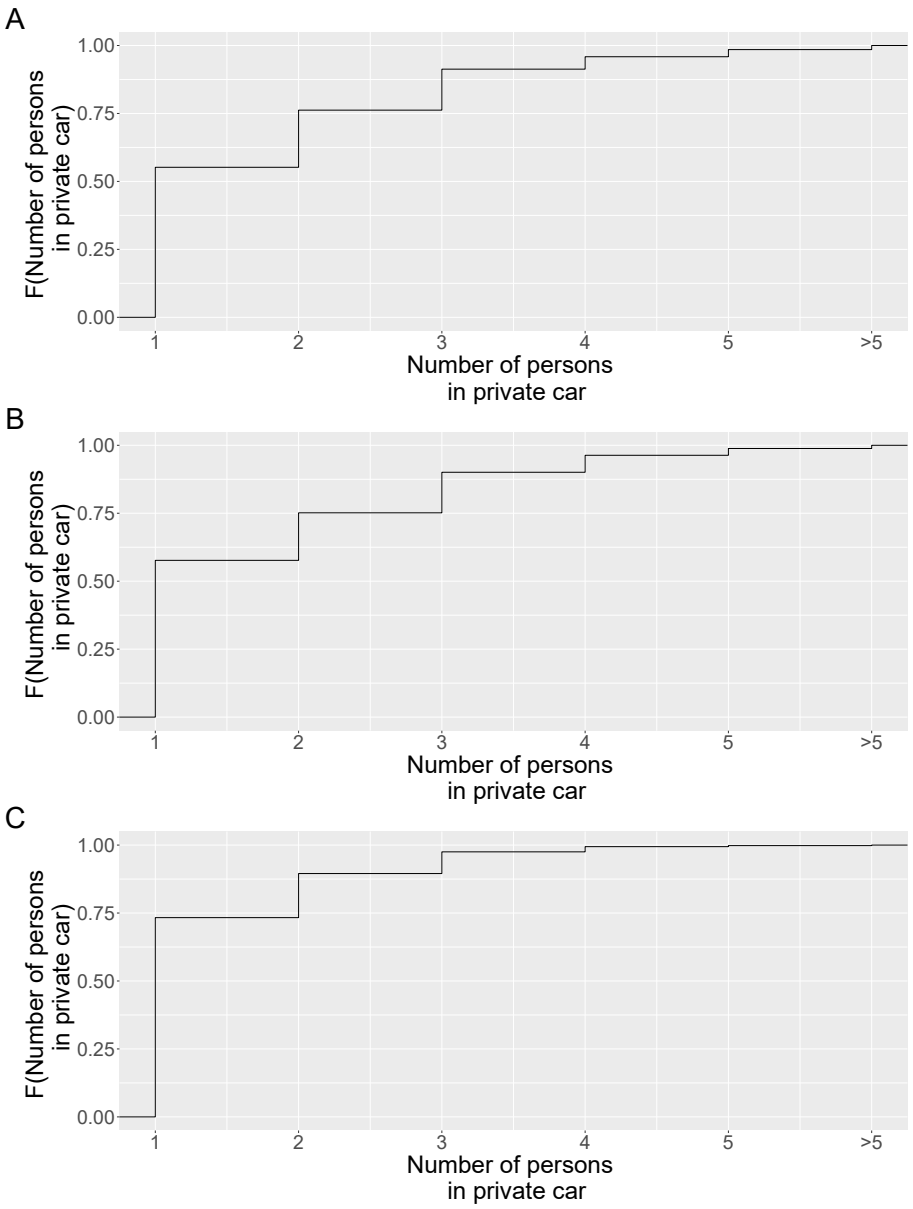

Figure S7: CDF corresponding number of persons travelling in a private car during a journey

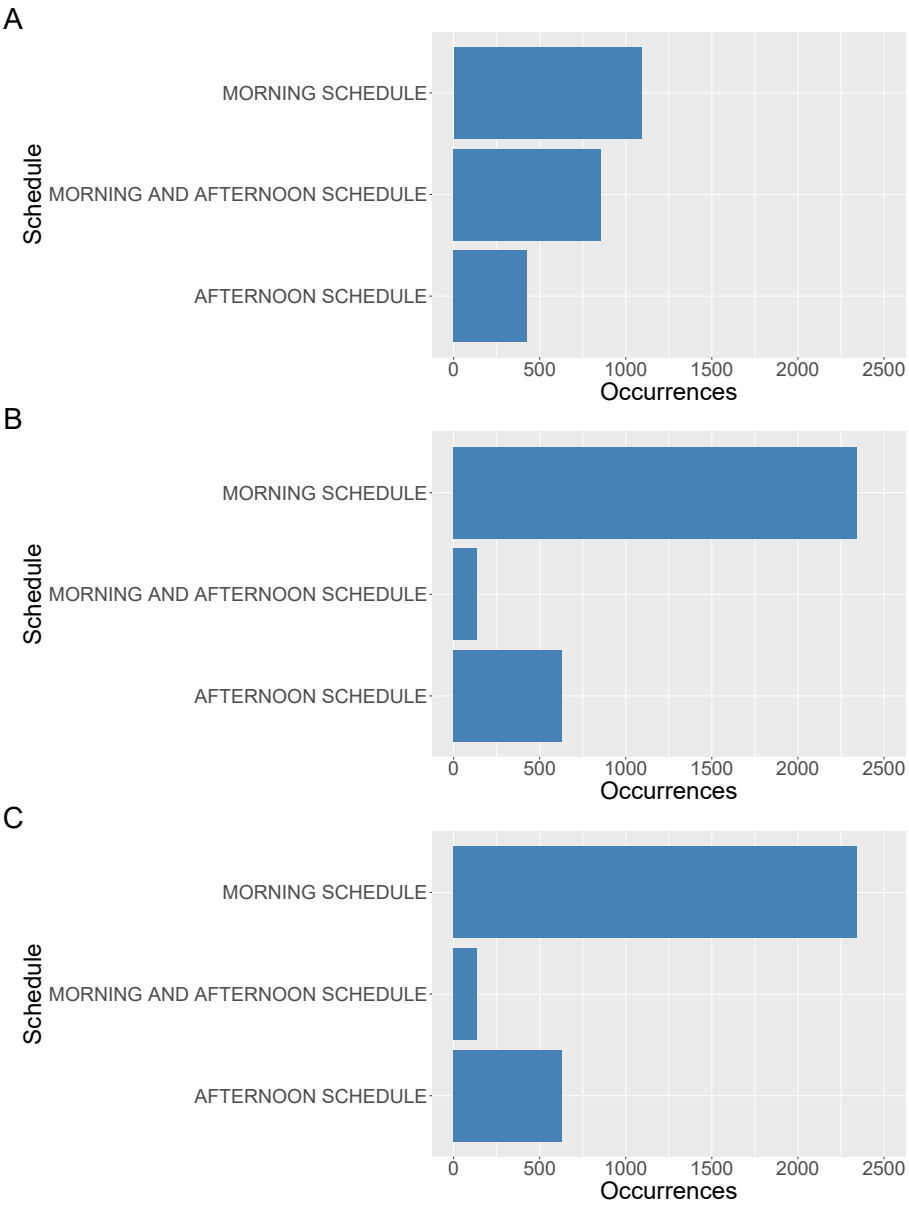

Figure S8: Histograms of schedules in which individuals stay in the university

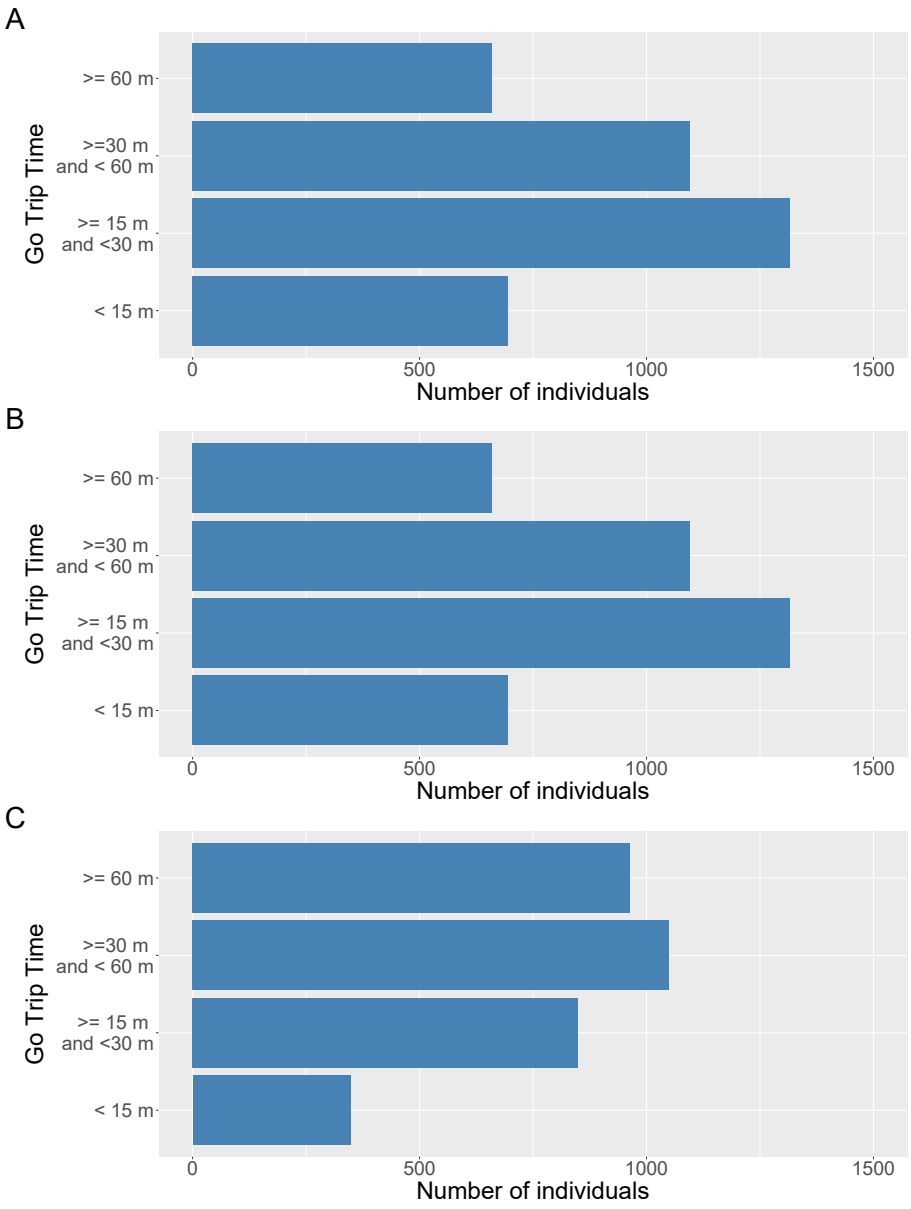

Figure S9: Histograms of trip times from origin in the (A) 2017-2018, (B) 2018-2019 and (C) 2021-2022 courses

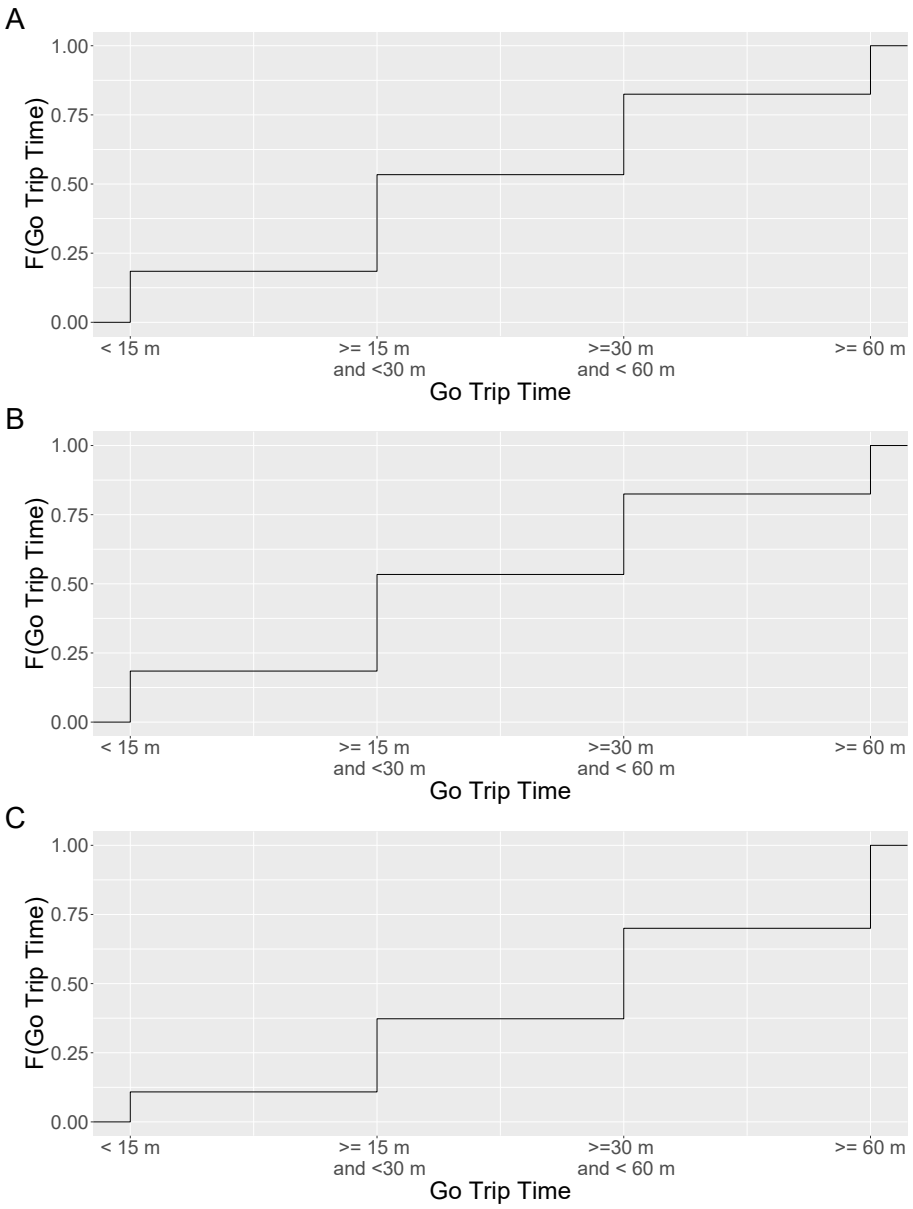

Figure S10: CDF of trip times from origin in the 2017-2018 (A), 2018-2019 (B) and 2021-2022 (C) courses

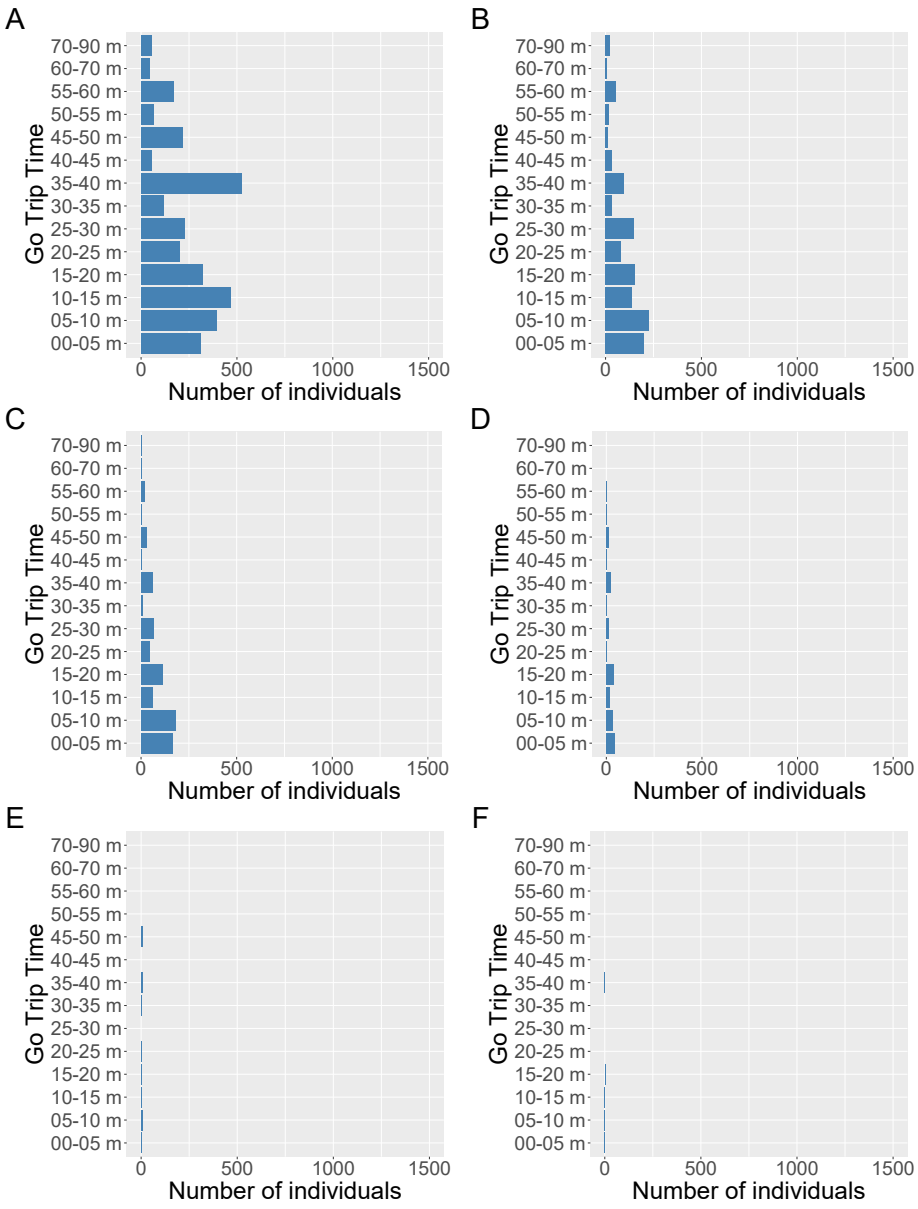

Figure S11: For 2021-2022, duration in each steps of the journey (A) step 1, (B) step 2, (C) step 3, (D) step 4, (E) step 5, (F) step 6.

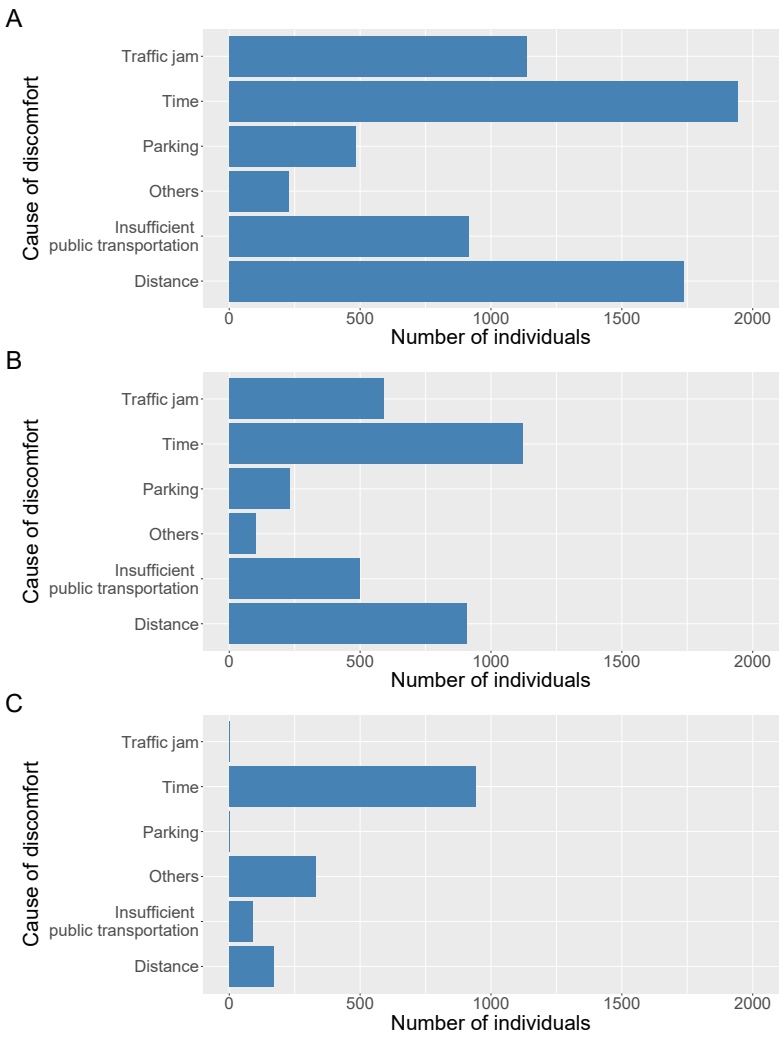

Figure S12: For 2017-2018 (A), 2018-2019 (B) and 2021-2022 (C) causes of discomfort

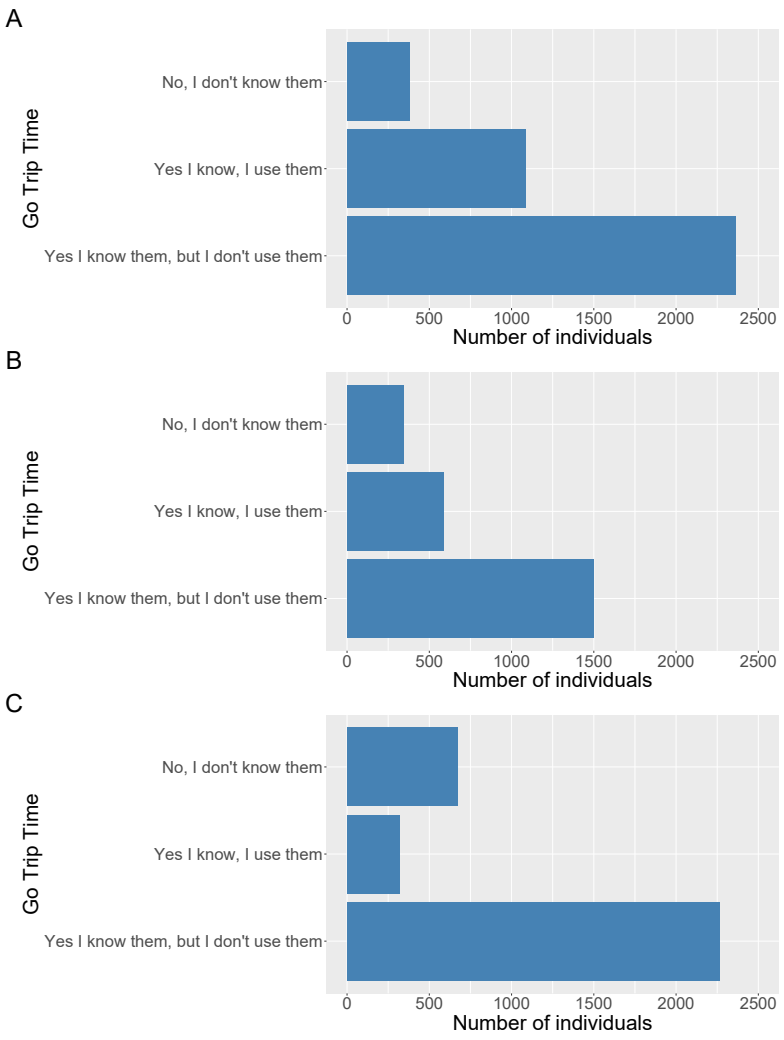

Figure S13: For 2017-2018 (A), 2018-2019 (B) and 2021-2022 (C) utilisation of mobile applications

Disclosure statement

The authors declare that they have no competing interest.

Funding

The authors received no financial support for publication of this article.

Data Availability

The anonymised data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors express gratitude to María Latasa Pérez, Carlota Gómez Rodríguez, and Javier Fuentes Lastau for his assistance in the statistical calculations and the maps construction.

This research was carried out as a result of the Project: PASCAL. A study of commuting to Francisco de Vitoria University, which was granted in the internal call for research projects in 2022 at the aforementioned university.

This research was partially supported by Telefónica Chair and AVANZA Chair at Francisco de Vitoria University.

References

[1] Statistics How To(w. d.). Statistics How To. https://www.statisticshowto.com/probability-and-statistics/find-sample-size/

[2] U. F. V.(w. da.). Memoria Académica UFV 2017-2018. Retrieved Feb 14, 2023, from: https://www.ufv.es/la-universidad/sobre-ufv/transparencia/memorias/

[3] U. F. V.(w. db.). Memoria Académica UFV 2018-2019. Retrieved Feb 14, 2023, from: https://www.ufv.es/la-universidad/sobre-ufv/transparencia/memorias/

[4] U. F. V.(w. dc.). Memoria Académica UFV 2021-2022. Retrieved Feb 14, 2023, from:https://www.ufv.es/la-universidad/sobre-ufv/transparencia/memorias/

[5] Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized Additive Models for Location, Scale and Shape. Journal of the Royal Statistical Society. Series C (Applied Statistics), 54(3), 507–554. Retrieved Feb 14, 2023, from: http://www.jstor.org/stable/3592732

[6] Stasinopoulos, M., Rigby, B., Voudouris, V., Akantziliotou, C., Enea M., & Kiose, D. (2022). Package 'gamlss'. Retrieved Jul 13, 2023, from: https://cran.r-project.org/web/packages/gamlss/gamlss.pdf

[7] Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. Statistical Science, 1(3), 297–310. Retrieved Feb 14, 2023, from: http://www.jstor.org/stable/2245459

[8] Krzanowski, W.J., & Lai, Y. T. (1985). A criterion for d-terming the number of groups in a data set using sum-of-squares clustering. Biometrics, 44, 23–34. https://doi.org/10.2307/2531893

[9] Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis, Communications in Statistics, 3(1), 1-27. https://doi.org/10.1080/03610927408827101

[10] Hartigan, J. A. (1975). Clustering Algorithms (Probability & Mathematical Statistics). Estados Unidos: John Wiley & Sons Inc.

[11] Sarle, W. S. (1983). Cubic clustering criterion (Tech. Rep. A-108). Research Triangle Park, NC: SAS Institute.

[12] Scott, A. J. & Symons, M. J. (1971). Clustering methods based on likelihood ratio criteria. Biometrics, 27, 387-397.

[13] Marriot, F. H. C. (1971). Practical problems in a method of cluster analysis. Biometrics, 27, 501-514.

[14] Milligan, G. W., & Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. Psychometrika, 50, 159-179.

[15] Friedman, H. R, & Rubin, J. (1967). On some invariant criteria for grouping data. Journal of the American Statistical Association, 62, 1159-1178.

[16] Hubert, L. J. & Levin, J. R. (1976). A general statistical framework for assessing categorical clustering in free recall. Psychological Bulletin, 83(6), 1072–1080. https://doi.org/10.1037/0033-2909.83.6.1072

[17] Davies, D., & Bouldin, D. (1979). A Cluster Separation Measure. Pattern Analysis and Machine Intelligence. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1(2), 224 - 227. https://doi.org/10.1109/TPAMI.1979.4766909

[18] Rousseeuw, P. J. (1987). Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. Journal of Computational & Applied Mathematics, 20, 53-65. Journal of Computational and Applied Mathematics, 20. 53-65 https://doi.org/10.1016/0377-0427(87)90125-7

[19] Duda, R. O. and Hart, P. E. (1973). Pattern Classification and Scene Analysis. Wiley, New York.

[20] Beale, E. M. L. (1969). Euclidean cluster analysis. A paper contributed to 37^{th} session of the International Statistical Institute.

[21] Ball, G. H., & Hall, D. H. (1965). ISODATA, a novel method of data analysis and pattern classification. Technical report, Stanford Research Institute. Recuperado el 26 de julio de 2022 de: https://apps.dtic.mil/sti/citations/AD0699616#:~:text=Abstract%3A-,ISODATA%2C%20a%20novel%20method%20of%20data%20analysis%20and%20pattern%20classification,calculations%20that%20the%20method%20uses

[22] Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika, 45, 325–342. https://doi.org/10.1007/BF02293907

[23] Milligan, G. W. (1981b). A Monte Carlo study of thirty internal criterion measures for cluster analysis. Psychometrika, 46, 187–199. https://doi.org/10.1007/BF02293899

[24] Milligan, G. W. (1981). A review of Monte Carlo tests of cluster analysis. Multivariate Behavioral Research, 16, 379–407. https://doi.org/10.1207/s15327906mbr1603_7

[25] Frey, T., & Van Groenewoud, H. (1972). A cluster analysis of the D2 matrix of white spruce stands in Saskatchewan based on the maximum-minimum principle. Journal of Ecology, 60(3), 873-886. https://doi.org/10.2307/2258571

[26] McClain, J. O., Rao, V. R. (1975). CLUSTISZ: A program to test for the quality of clustering of a set of objects. Journal of marketing research 12, 456–460.

[27] Dunn, J. C. (1974). Well separated clusters and optimal fuzzy partitions. Journal of Cybernetics, 4, 95–104

[28] Halkidi, M., Vazirgiannis, M., Batistakis, Y. (2000). Quality Scheme Assessment in the Clustering Process. In: Zighed, D.A., Komorowski, J., Żytkow, J. (eds) Principles of Data Mining and Knowledge Discovery. PKDD 2000. Lecture Notes in Computer Science book series (LNAI,1910). Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45372-5_26

[29] Lebart, L., Morineau, A. & Piron, M. (2000). Statistique exploratoire multidimensionnelle, Dunod, Paris

[30] Halkidi, M. et al. (2001). On Clustering Validation Techniques. Journal of Intelligent Information Systems 17, 107–145. https://doi.org/10.1023/A:1012801612483