

Department of Civil Engineering/TRIPP  
Indian Institute of Technology Delhi

<http://tripp.iitd.ac.in>



Data Collection and Sampling issues  
-Preparing Low Carbon Mobility Plan

K. Ramachandra Rao, G. Tiwari, Sudeep Grover  
UNEP Workshop, 11-13 Apr 2012

# Outline

---

- ▶ Background
- ▶ Sampling
- ▶ Data collection
- ▶ Data needs for LCMP
  - ▶ City profile/structure
  - ▶ Transport network inventory
  - ▶ Public/para-transit system/NMT
  - ▶ Traffic conditions
  - ▶ Travel demand modelling
  - ▶ Environmental conditions
  - ▶ Alternative scenario development
- ▶ Case study –
  - ▶ Ludhiana: A choice experiment for Low Carbon Mobility options

# Background

---

- ▶ Data collection is one of the essential part of any planning exercise
- ▶ Most of the transportation planning models are data hungry
- ▶ Challenge lies in identifying the data needs and sampling sizes within the allocated budgets while preparing the CMP/LCMP
- ▶ Need to look at the data requirements and sampling strategies for LCMP
- ▶ Data collected for generating information would be used for decision making

# Sampling Plans...

---

- ▶ The selection of a proper sample is an obvious prerequisite to a sample survey
- ▶ A sample is defined to be a collection of units which is some part of a larger population and which is specially selected to represent the whole population
- ▶ Sampling is often done for reasons of **cost** and **practicality**
- ▶ A **sampling plan** is just a method or procedure for specifying how a sample will be taken from a population
- ▶ Methods of sampling:
  - ▶ Simple Random Sampling
  - ▶ Stratified Random Sampling
  - ▶ Variable fraction stratified random sampling
  - ▶ Cluster Sampling
  - ▶ Multi-stage sampling
  - ▶ Systematic sampling

# Sampling methods

00	01	02	03	04	05	06	07	08	09
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49
50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	99

A Population of 100 Sampling Units with Identifiers

00	01	02	03	04	05	06	07	08	09
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49
50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	99

A Simple Random Sample of 10 Sampling Units

Data collection and sampling issues

# Sampling methods

00	01	02	03	04	05	06	07	08	09
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49
50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	99

Female

Male

A Simple Random Sample from a Stratified Population

00	01	02	03	04	05	06	07	08	09
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49
50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	99

Female

Male

A Stratified Random Sample from a Stratified Population



# Sampling methods

Males 18 +	Adults 18 and Over			
	1	2	3	4
0	THE WOMAN	THE YOUNGER WOMAN	THE OLDEST WOMAN	THE SECOND YOUNGEST WOMAN
1	THE MAN	THE MAN	THE OLDER WOMAN	THE YOUNGEST WOMAN
2		THE OLDER MAN	THE OLDER MAN	THE OLDER MAN
3			THE OLDEST MAN	THE YOUNGEST MAN

EVEN

Males 18 +	Adults 18 and Over			
	1	2	3	4
0	THE WOMAN	THE OLDER WOMAN	THE OLDEST WOMAN	THE OLDEST WOMAN
1	THE MAN	THE WOMAN	THE YOUNGER WOMAN	THE SECOND YOUNGEST WOMAN
2		THE OLDER MAN	THE OLDER MAN	THE YOUNGEST MAN
3			THE SECOND OLDEST MAN	THE SECOND OLDEST MAN

Examples of Respondent Selection Grids

00	01	02	03	04	05	06	07	08	09
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49
50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	99

Female

Male

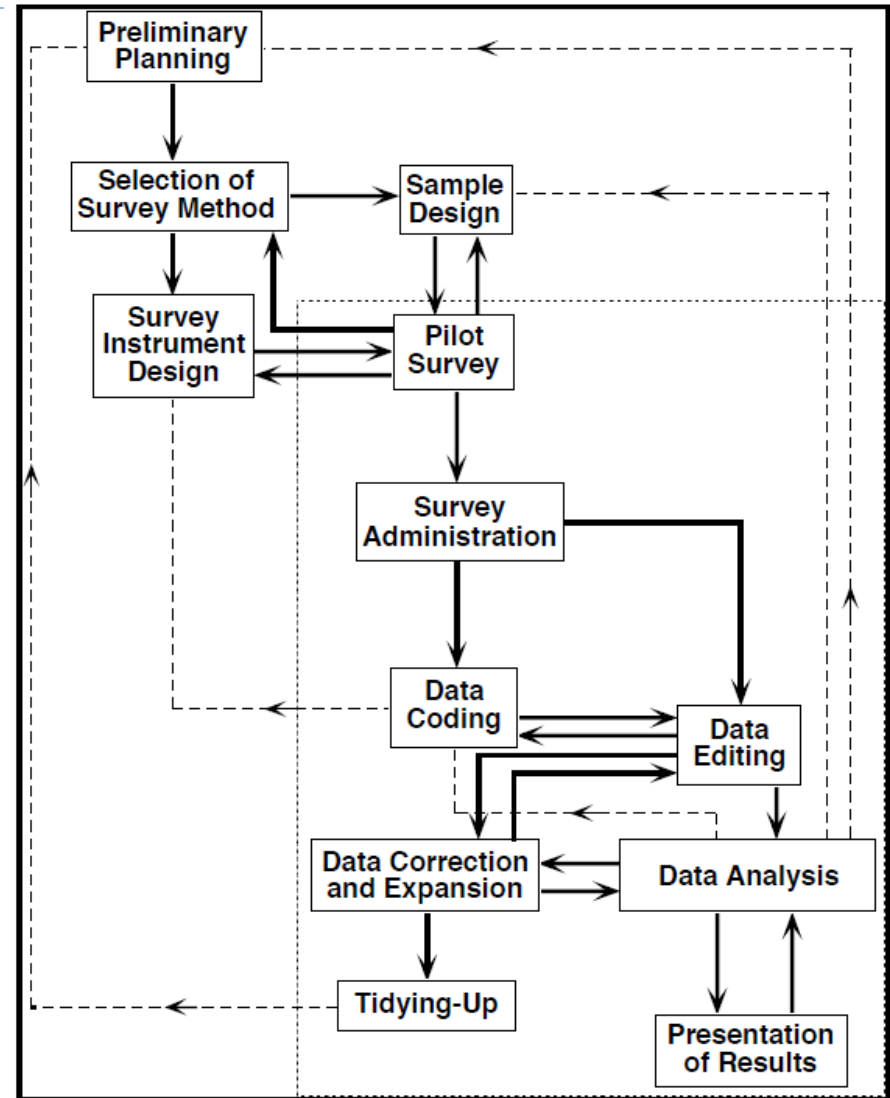
# Target population

---

- ▶ The target population is the complete group about which one would like to collect information.
- ▶ The elements of this group may be people, households, vehicles, geographical areas or any other discrete units
- ▶ The definition of the target population will, in many cases, follow directly from the objectives of the survey



# Data collection and survey process



# Data collection methods

---

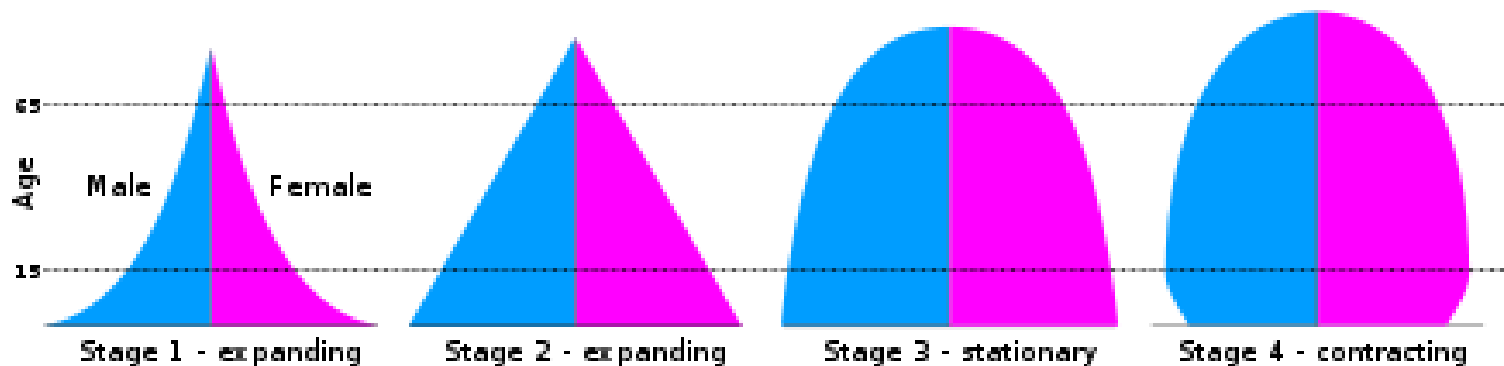
- ▶ Documentary searches
- ▶ Observational surveys
- ▶ Household self-completion surveys
- ▶ Telephone surveys
- ▶ Intercept surveys
- ▶ Household personal interview surveys
- ▶ Group surveys
- ▶ In-depth surveys
- ▶ IT – Wireless technology

# LCMP

Data required	Description	LCMP	CMP
Location	Geographical location	Y	Y
	<b>Climatic condition</b>	Y	
Land area	Total land area	Y	Y
	<b>Growth pattern</b>	Y	
	<b>Identification of notified areas</b>	Y	
Regional linkages		Y	Y
Demography	Population growth trend	Y	Y
	Number and size of HH	Y	Y
	<b>Age-sex pyramid</b>	Y	
Socio-economic data	<b>Population by social group*</b>	Y	
	Population by income/expenditure at zone	Y	Y
	Vehicle ownership by social group	Y	Y

# City profile: demography - sampling

- ▶ Age sex pyramid: A **population pyramid**, also called an **age structure diagram**, is a graphical illustration that shows the distribution of various age groups in a **population** which forms the shape of a **pyramid** when the population is growing



# City structure: mix intensity - sampling

---

Data required	Description	LCMP	CMP
Land use pattern	Map	Y	Y
Population density	Disaggregated at the level of study	Y	Y
<b>Mix intensity</b>	<b>Determined by job-housing ratio or balance in a zone or level of study</b>	Y	
Land consumption by land use		Y	Y

# Transport Network Inventory –Pedestrians - population

Data required	Description	LCMP	CMP
Footpath	Width	Y	Y
	Single sided/both side	Y	Y
	Continuity	Y	Y
	Encroachment by other activity/vehicles	Y	Y
	Pavement condition	Y	Y
	<b>Lighting</b>	Y	
	Clear markings	Y	Y
Intersections	Signalized crossing	Y	Y
	Level/raised crossing	Y	Y
	Traffic calming tools like speed breakers	Y	Y
	<b>Intermediate crossing sections</b>	Y	
	Crossing distance	Y	Y
Access	<b>Barrier free access to bus stops</b>	Y	
	<b>Barrier free access to footpaths</b>	Y	

# Transport Network Inventory – NMV - population

Data required	Description	LCMP	CMP
Lanes	Cycle lanes/tracks	Y	Y
	Width of cycle lanes/tracks	Y	Y
	Both sided/single sided	Y	Y
	Encroachment by other activity/vehicles	Y	Y
	<b>Lighting</b>	Y	
	Pavement condition	Y	Y
Intersection treatment	Signalized intersections	Y	Y
	Traffic calming tools	Y	Y
	<b>Traffic calming for access to properties</b>	Y	
Parking	<b>Number of parking</b>	Y	
	<b>Distance of parking from PT stop</b>	Y	
	<b>Parking charges</b>	Y	



# Transport Network Inventory – Para-transit Autos and Cycle Rickshaws - population

---

Data required	Description	LCMP	CMP
Para-transit (auto rickshaw and cycle rickshaw)	Restricted / non-restricted	Y	Y
	<b>Number of parking by specification of parking areas</b>	Y	
	<b>Distance of formal parking from bus stop</b>	Y	
	<b>Distance between parking stations</b>	Y	
	<b>Parking charges</b>	Y	



# Transport Network Inventory- Public Transport (Bus) - population

---

Data required	Description	LCMP	CMP
Infrastructure	<b>Dedicated bus lanes (type and description)</b>	Y	
	Average speed on different roads	Y	Y
Bus stop	Number of bus stops		Y
	Bus stop inventory	Y	Y
	<b>Average distance between bus stop</b>	Y	
	Location of bus stop with respect to junction	Y	Y
	Passenger amenities on bus stop	Y	Y
	Number of bus terminals		Y

# Transport Network Inventory- Personal Motorized Vehicles - population

Data required	Description	LCMP	CMP
Road infrastructure	ROW	Y	Y
	Number of lanes	Y	Y
	Average speed on different roads	Y	Y
	Road encroachment	Y	Y
	<b>Lighting</b>	Y	
Intersections	Type of intersection- signalized/un-signalized	Y	Y
	<b>Turning time from each direction</b>	Y	
	<b>Signal phasing</b>	Y	
	Existence of Area Traffic Control System	Y	Y
	Type of Signal Operation		Y
Parking	Ownership type		Y
	Parking regulations	Y	Y
	Number of parking	Y	Y
	Parking charges by mode	Y	Y

# Public Transport System- sample/population

Data required	Description	LCMP	CMP
Fleet usage detail	Type of ownership		Y
	Number of bus by type of bus (standard, mini, low floor), fuel used and age	Y	Y
	<b>Fleet utilization rate</b>	Y	
	Vehicular kilometer	Y	Y
	Percentage occupancy- peak hour and average	Y	Y
	Total passengers per day	Y	Y
Route detail	Route inventory	Y	Y
	Number of bus routes		Y
	Headway on different routes	Y	Y
	Average route speed	Y	Y
	Service reliability	Y	Y
Cost and fare	Operation cost per km	Y	Y
	Tax levied	Y	
	Fare structure	Y	Y
	<b>Revenue per km</b>	Y	
	Profit/loss	Y	Y

# Para-transit System- Auto, Cycle Rickshaw and Shared Auto - population

Data required	Description	LCMP	CMP
Fleet usage detail	Type of ownership	Y	Y
	Number of para-transit by type (shared, personal autos and cycle rickshaw), fuel used and age	Y	Y
	Vehicular kilometer	Y	Y
Route detail	<b>Route inventory for share auto</b>	Y	
	<b>Average waiting time for auto, cycle rickshaw and shared auto</b>	Y	
Cost and fare	Operation cost per km	Y	Y
	<b>Tax levied</b>	Y	
	Fare structure	Y	Y
	Revenue per km	Y	Y
	Profit/loss	Y	Y

# Traffic Study- Condition, Safety And Security - sample

---

Data required	Description	LCMP	CMP
Traffic count	Screen line by modes	Y	Y
	At intersection by modes	Y	Y
Delay and Queue length	Queue length by mode	Y	Y
	<b>Delay by mode</b>	Y	
	Travel speed by mode	Y	Y
Number and location of injury/fatalities on road	<b>By victim mode</b>	Y	
	<b>By impacting vehicle</b>	Y	
Reported crimes	<b>Disaggregated by mode</b>	Y	

# Modelling Travel Demand- Household Survey - sample

---

Data required	Description	LCMP	CMP
Personal information	Age	Y	Y
	Gender	Y	Y
	Occupation	Y	Y
	Monthly income	Y	Y
	Vehicle ownership and age of vehicle	Y	Y
	Monthly expenditure on transport	Y	Y
Transport infrastructure rating for different modes	Perception about Safety	Y	Y
	Perception about security	Y	Y
	Perception about comfort	Y	Y

# Modelling Travel Demand- Household Survey - sample

Data required	Description	LCMP	CMP	
Trip making information	Trip purpose	Y	Y	
	Trip origin and destination	Y	Y	
	<b>Travel distance</b>	Y		
	Mode used	Y	Y	
	<b>Access and egress mode</b>	Y		
	<b>Access and egress public transport stop</b>	Y		
	<b>Distance to access and egress public transport stop</b>	Y		
	<b>Travel time to access and egress</b>	Y		
	<b>Average waiting time to board Public transport</b>	Y		
	Total travel time	Y	Y	
	Total travel cost	Y	Y	
	<b>Average mileage if PMV used</b>	Y		
	<b>Fuel used</b>	Data collection and sampling issues	Y	
	<b>Reason for using the mode used</b>		Y	

# Environmental Condition & City Budget Analysis - sample

Data required	Description	LCMP	CMP
Air quality levels	NO <sub>x</sub> , CO <sub>2</sub> , CO, SO <sub>x</sub> concentration by location	Y	Y
Noise levels	By location	Y	Y
Investment trends in transport on different modes		Y	Y
Tax policies for different modes		Y	Y
Percentage of subsidy granted		Y	Y
Fuel price	Previous years trend	Y	Y



# Data for Developing Alternate Scenarios - sample

Data required	Description	LCMP	CMP
Socio-economic profile	Population growth	Y	Y
	<b>Economic growth- per capita income/household income</b>	Y	
	<b>Vehicle ownership trend</b>	Y	
	<b>Economic sector growth</b>	Y	
New development areas	Location	Y	Y
	Area	Y	Y
	<b>Land use plan</b>	Y	
	Number of residential population	Y	Y
	Number of Jobs	Y	Y
	Year of completion	Y	Y
Proposed projects	Type of project	Y	Y
	<b>Location</b>	Y	
	Project profile	Y	Y
	Purpose of project	Y	Y
	Current status	Y	Y
	Target year of completion	Y	Y
	<b>Likely impacts of project on travel pattern, traffic</b>	Y	

# Case study

---

- ▶ Ludhiana
- ▶ Sampling strategy: stratified random sampling
- ▶ Data collection method: personal- paper questionnaire
- ▶ User preferences to the Low carbon modes
  - ▶ Walk
  - ▶ Bicycle
  - ▶ Bus transit
  - ▶ Electric vehicles
    - ▶ Two-wheeler
    - ▶ Car
- ▶ Stated preference study

# Determining sample size: parameters

---

- ▶ **Margin of Error**: The margin of error is the amount of error that you can tolerate.
- ▶ If 90% of respondents answer yes, while 10% answer *no*, you may be able to tolerate a larger amount of error than if the respondents are split 50-50 or 45-55.
- ▶ Lower margin of error requires a larger sample size.
- ▶ **Confidence level**: The confidence level is the amount of uncertainty you can tolerate.
- ▶ Suppose that you have 20 yes-no questions in your survey. With a confidence level of 95%, you would expect that for one of the questions (1 in 20), the percentage of people who answer *yes would* be more than the margin of error away from the true answer.
- ▶ **Response percentage**: This is the percentage of people who give a particular answer to a question in a survey.
- ▶ If 98% of people select an answer there is not much room for error.
- ▶ However, if 45 or 55% select a particular answer, we don't have a clear majority and the chance for error gets bigger.

# Sample size for given population

Margin of Error (%)	Confidence level(%)	Response Ratio(%)	Sample Size
5	95	50	385
<b>6.68</b>	<b>95</b>	<b>50</b>	<b>216</b>
6	95	50	267
10	95	50	97
5	90	50	271
10	90	50	68
5	99	50	664
10	99	50	166
5	95	30	323
5	95	40	369
5	90	40	260

<http://www.raosoft.com/samplesize.html>  
<http://relevantinsights.com/research-tools>

# Population vs sample

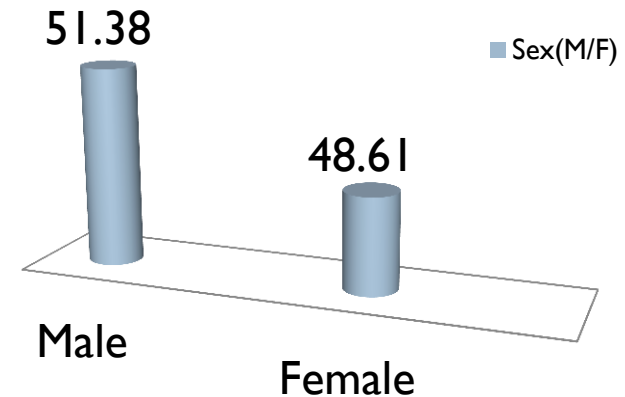
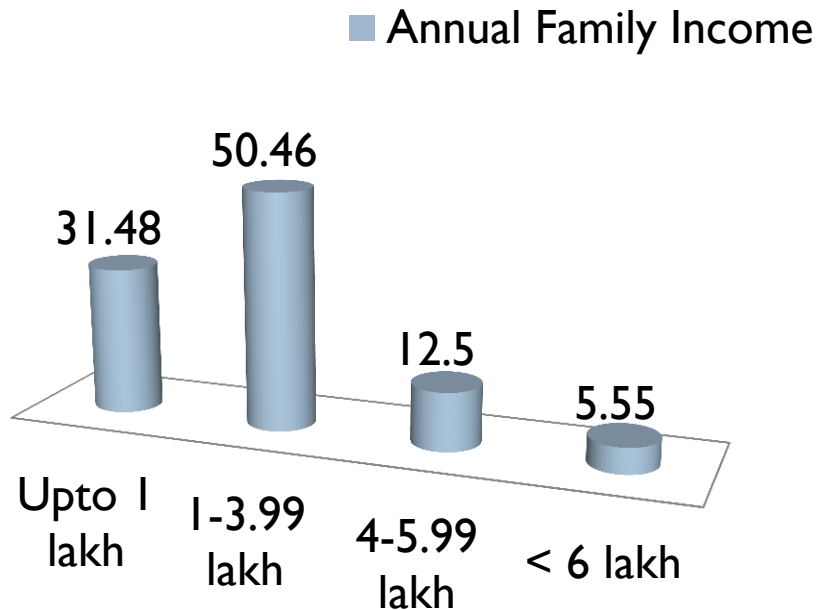
---

Parameter	Population (Census 2011)	Sample
SIZE (n)	<b>16,13,878</b>	216
SEX RATIO (%)	54.2 : 45.8	51.38 : 48.61
LITERACY	85.38% (city) , 82.5 % ( district)	Undergraduate 32.4 % Literate 100 %

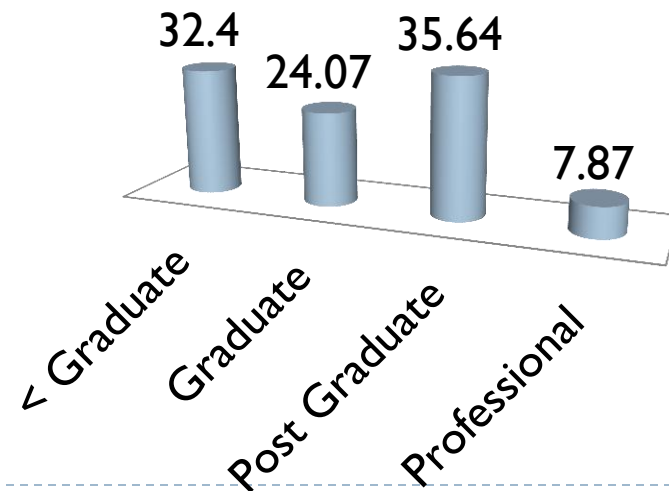
# Personal attributes

## Gender (%)

## Income Distribution (%)

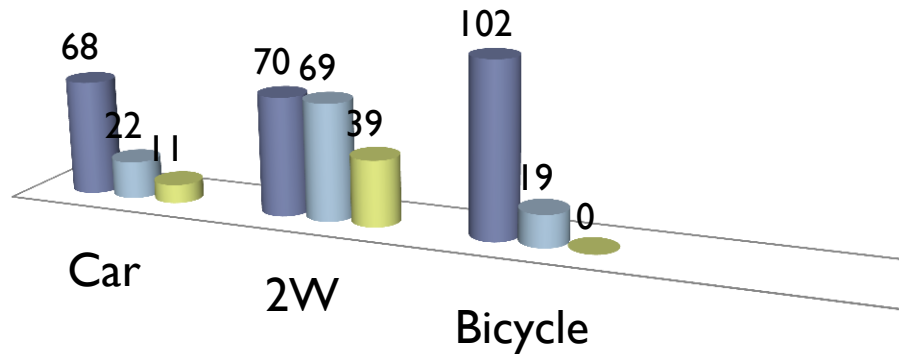


## Education (%)

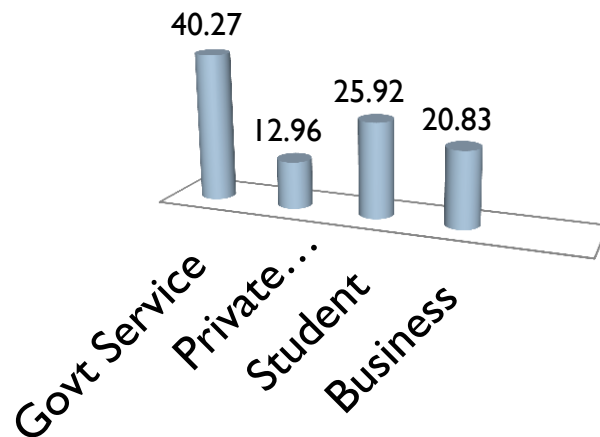


## Vehicle Ownership

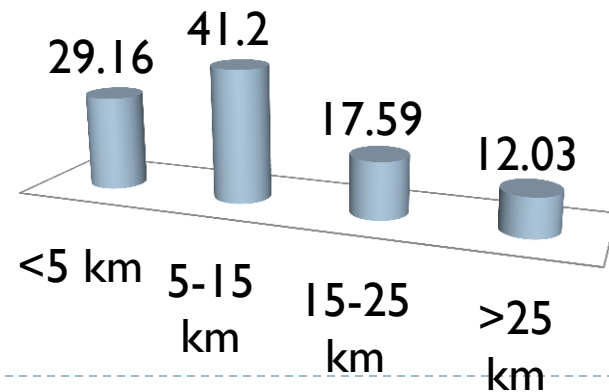
■ One ■ Two ■ Three



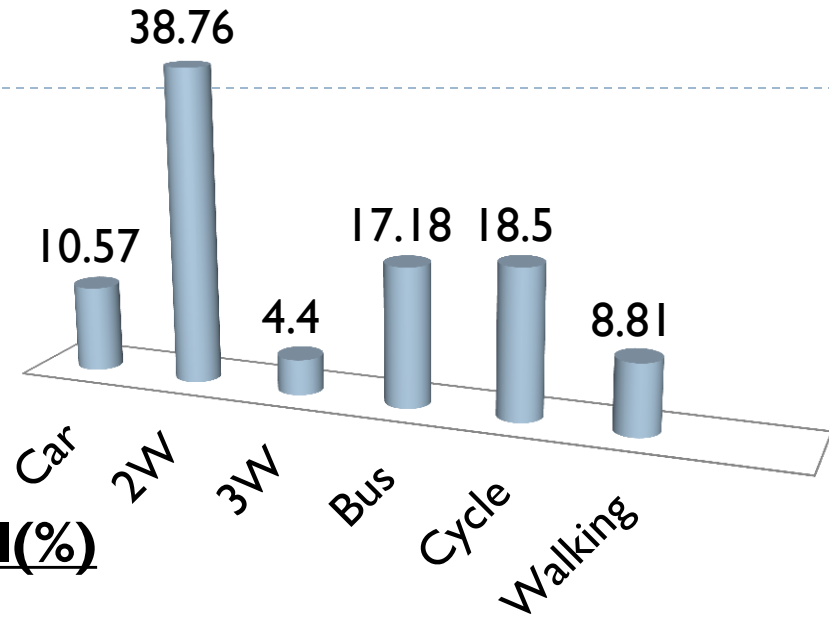
## Jobs Held (%)



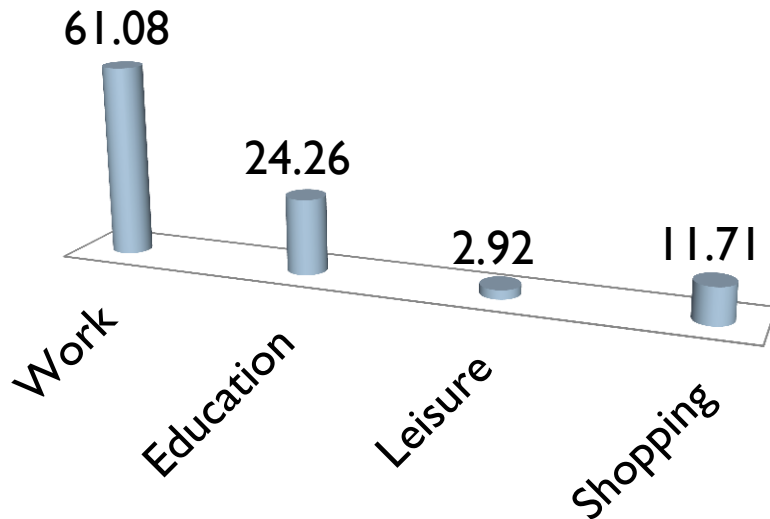
## Average daily Distance Commuted (%)



## Present Mode of travel(%)



## Primary Reason For Travel(%)





# Study design

---

## **Four choice cards for choosing desired mode.**

- ▶ **Card 1** Effect of providing independent lanes for buses, bicycles and walking
- ▶ **Card 2** Better transit policy in addition to the above
- ▶ **Card 3** Policy against cars in addition to the above
- ▶ **Card 4** IC engine vehicles vs. electric vehicles

# Summary of choices (%)

Option	Car	2W	Bus	Auto/ Taxi	Bicycle	Walk
<u>Present Mode</u> <u>Used</u>	10.57	38.76	17.18	4.4	18.50	8.81
Card 1 (Lanes for Bus,Cycle,Walk)	21.29	24.53	25.92	0.92	23.14	4.16
Card 2 (Card 1 + Better Buses)	18.51	22.68	41.67	0.92	12.5	3.7
Card 3 (Card 2 + Policy against Cars)	10.64	30.55	40.27	0.92	12.5	5.09

Option	Car	Electric Car	2W	Electric 2W
Card 4 (Electric Vehicles)	19.9	6.01	34.72	39.35

# References

---

- ▶ Hensher, D.A, Rose, J.M., Greene, W. H., (2005) *Applied Choice Analysis – A primer*, Cambridge University Press
- ▶ Keller, G. (2009) *Statistics for Management and Economics*, Eighth Edition, Cengage Learning
- ▶ Richardson, A.J., Ampt, E.S., Meyburg, A.H. (1995) *Survey Methods for Transport Planning*, 1<sup>st</sup> Ed, Eucalyptus press, Melbourne
- ▶ Ortuzar, J.D., Willumsen, L.G. (2011) *Modelling Transport*, 4<sup>th</sup> Ed, John Wiley, Chichester