

Interaction Design Project 1

Visualization for Mumbai's Data

Part 1

By,
Akshay Kore
Interaction Design
IDC IIT Bombay
146330007

Guided by :
Prof. Venkatesh Rajamanickam
Information Design Lab, **IIT Bombay**



This Interaction Design Degree project titled '**Visualization for Mumbai's Data**' is done in partial fulfillment of the requirement for the degree of Master of Design in Interaction Design at Industrial Design Center, IIT Bombay.

It was the done with The **Information Design Lab** at IIT Bombay.

Duration : May to June, 2015

Declaration:

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all the principles of academic honesty and Integrity and have not misrepresented or fabricated or falsified any idea/data/fact/ source in my submission.

I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited or from whom proper permission has not been taken when needed.



Akshay Kore
Interaction Design IDC IIT Bombay
Roll No. 146330007

Mumbai, India
July 2015

Acknowledgement:

This project would not have been possible without the guidance and support of a number of individuals. I would like to thank Prof. Venkatesh Rajamanickam from the Information Design Lab at IIT Bombay for the opportunity to work on this project and for his support and guidance. I would also like to thank Prof. Arnab Jana from the Centre of Urban Science and Engineering (C-USE) at IIT Bombay for his valuable feedback during the course of the project.

I am highly indebted to the staff and officials at the Maharashtra Industrial Development Corporation (MIDC), Office of the Development Commissioner SEEPZ SEZ, Mumbai Metropolitan Region Development Authority (MMRDA), Maharashtra Housing and Development Authority (MHADA) and the Municipal Corporation of Greater Mumbai (MCGM) for providing data and access to maps. I am grateful to the staff at the Urban Design Research Institute (UDRI) for providing access to their library.

I am also thankful to IDC, IIT Bombay for giving me this opportunity and my faculty members for their support and encouragement.

Table of Contents

Abstract

8

Vast Quantities of Data

9

Scattered Datasets

12

Special Planning Areas

15

Existing Land Use

20

Visualization of the data

28

Abstract

Our governments collect vast amounts of data about us and our surroundings. A major part of this data is accessible in the public domain. This data shapes our cities, our villages. It is responsible for policies that affect us. And yet we know so little about it. This data tells stories of times gone by, it is a testament to our history. We can backtrace our paths and see where did we go wrong or better yet, where were we right.

A part of why we know so little is because the data does not speak to us nor can we make sense of numbers in tabular formats very easily. It does not come to us in a palatable form. We are visual beings by nature and it is much easier for us to see visuals; may it be a photograph or a graph. Those lines and pixels in two dimensional space tell us of a story that words alone could never describe that well.

Visualization of this vast amount of data becomes important in a way that it tells us much more than just numbers could ever tell us. It tells us about the rise and decline of our society, trends and maybe even help us predict what is to come. Maybe it could prevent a disaster or tell us about past ones so that we learn.

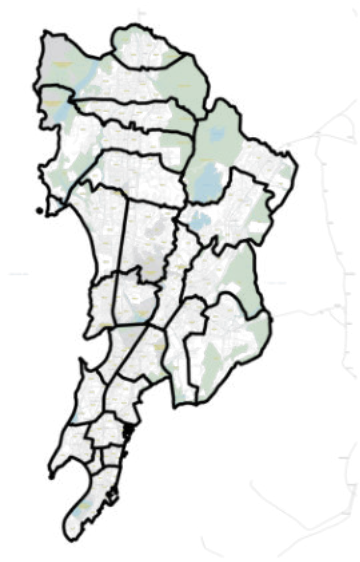


IMAGES SOURCE : GOOGLE IMAGES

Before telling of a story of our cities comes the knowing of our cities. Before the visualization comes the data. This project is the first part of the project titled 'Visualization for Mumbai's Data' and focuses on collection and partial analysis of the quantifiable data of the city of Mumbai. The overall project aims to set an example for presenting the data about our cities at large.

Vast Quantities of Data

Government bodies collect vast quantities of data by means of Census, Budget, Land



24 Wards

SOURCE : MCGM
MAP NOT TO SCALE

Use Studies, Topography studies etc. This data is significant due to the quantity and centrality of the data. It is also binding on the government to make most of this data available freely to its citizen.

Mumbai Data

The Municipal Corporation of Greater Mumbai (MCGM) has released a vast amount of data to the public. This data includes the existing land use, population and other relevant data sets. Also data about



151 Planning Sectors

SOURCE : MCGM
MAP NOT TO SCALE

governance, health, education, etc is made available by a number of other government bodies and NGOs. Part 1 of this project aims to look at the datasets and find out any insights or deficiencies in the data sets available.

This project tries to deal with the quantifiable information about the city's

infrastructure, population, growth and environment. Other datasets may be used to find interesting insights about the city and its divisions. The project tries to intentionally do away with the social and emotional aspects of the data.

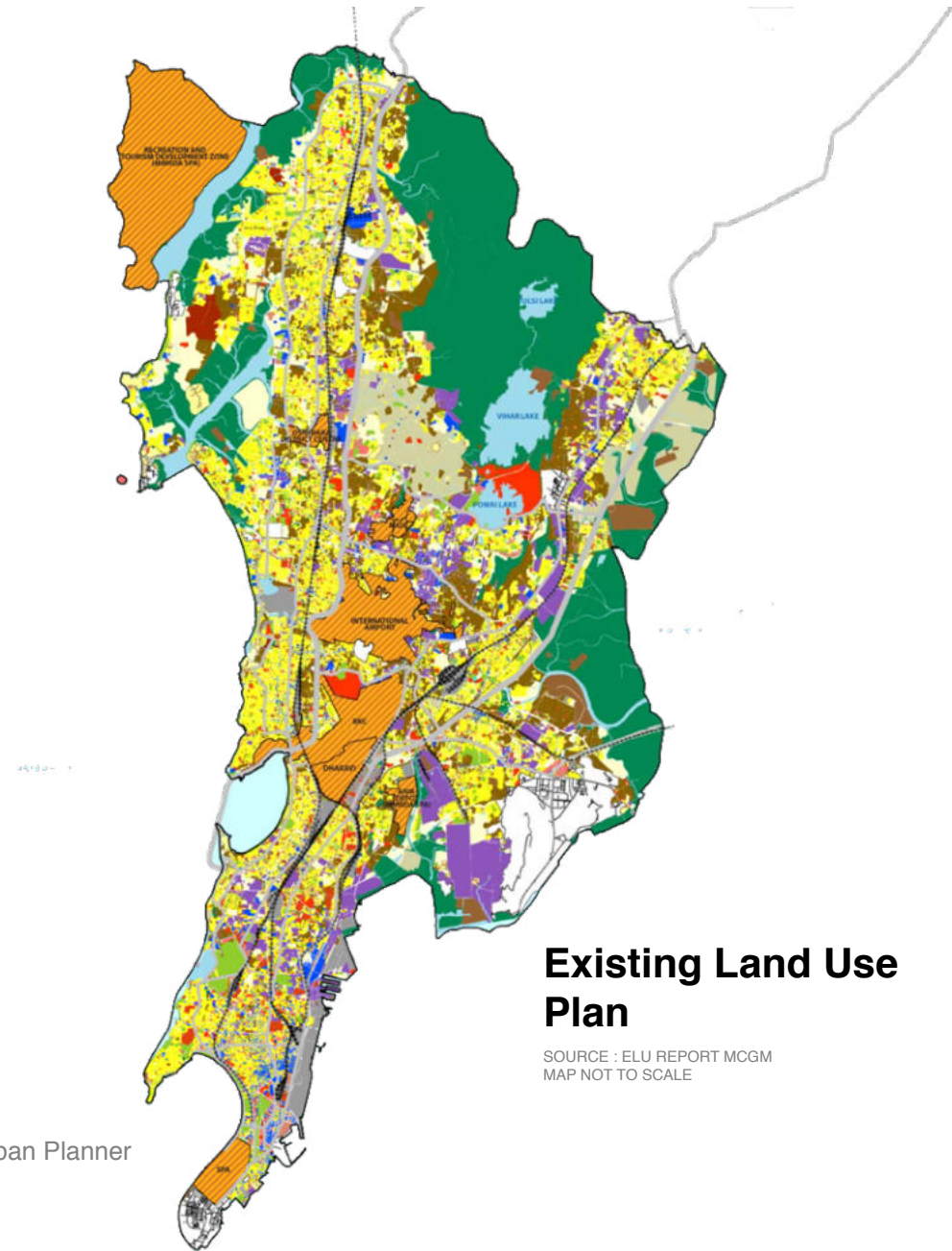
Note that the word infrastructure is used in its broad sense and includes residential, commercial, industrial infrastructure along with medical, educational, social amenities

and utilities. Environment includes water bodies, rivers, forests, national parks, open spaces, mangroves, hills, mud flats, etc.

Divisions

The City of Mumbai has approximately 1 lakh 27 thousand land parcels¹. For the purpose of ease of administration, the city is divided into 24 wards which are further subdivided into planning sectors. The City has a total of 151 Planning sectors.

The population of Greater Mumbai is 12.44 Million and is India's most populous city². Each ward constitutes of a population between 1,00,000 to 5,00,000 persons³.



¹ Talk on 22nd April 2015 in IIT Bombay by Mr. Vidyadhar Phatak, Urban Planner

² Census 2011

³ Preparatory Studies DP 2014-34 : MCGM

Scattered Datasets

Despite the availability of large quantities of data, the datasets are often scattered over multiple PDF documents. It becomes laborious to extract the data from these large number of PDF files.

The MCGM has released data for the Existing Land Use in Mumbai. This data mainly includes the quantity, location and category for different land parcels in the city. The land use categories are decided and assigned by the MCGM. However, it is very difficult to look at individual land use categories which helps in finding out deficiencies of that land use category in the city or a particular ward.

The ELU data includes tabular data of the land use in the city and its wards along with maps indicating location of land use categories in the wards and city. The adjoining figure is the land use map of the city.

GREATER MUMBAI – BROAD LAND USE CATEGORIES

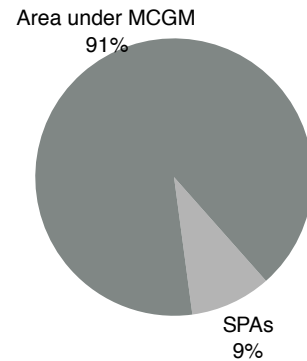
EXISTING LAND USE CATEGORIES#		ENTIRE CITY	ISLAND CITY										WESTERN SUBURBS										EASTERN SUBURBS						
			A	B	C	D	E	F/N	F/S	G/N	G/S	H/E	H/W	K/E	K/W	P/N	P/S	R/N	R/C	R/S	L	M/E	M/W	N	S	T			
Area (in ha)		41505.7	884.0	265.8	191.3	822.3	727.4	1108.5	978.8	573.1	929.4	657.9	811.2	1676.0	2403.5	4105.2	2471.7	1417.9	3362.3	1831.3	1408.3	3308.3	1740.4	2568.6	2975.2	4287.5			
Percentage (%)		100.0	2.1	0.6	0.5	2.0	1.8	2.7	2.4	1.4	2.2	1.6	2.0	4.0	5.8	9.9	6.0	3.4	8.1	4.4	3.4	8.0	4.2	6.2	7.2	10.3			
#		Ward Wise Area Distribution (in ha)																											
Residential	Area	10333.4	80.8	44.5	77.2	377.8	192.8	402.6	233.8	216.8	220.9	282.7	421.7	685.4	825.1	1027.2	499.1	416.8	626.4	614.0	544.9	571.1	391.9	550.1	622.6	406.9			
	%	24.9	0.8	0.4	0.7	3.7	1.9	3.9	2.3	2.1	2.1	2.7	4.1	6.6	8.0	9.9	4.8	4.0	6.1	5.9	5.3	5.5	3.8	5.3	6.0	3.9			
Commercial	Area	903.9	47.8	30.2	27.4	34.0	47.7	9.6	50.1	13.3	42.1	21.7	30.0	57.2	74.0	69.9	82.6	7.2	66.9	16.2	52.6	8.3	25.7	20.0	46.9	22.3			
	%	2.2	5.3	3.3	3.0	3.8	5.3	1.1	5.5	1.5	4.7	2.4	3.3	6.3	8.2	7.7	9.1	0.8	7.4	1.8	5.8	0.9	2.8	2.2	5.2	2.5			
Offices	Area	360.2	52.6	1.2	1.3	12.8	15.6	9.2	9.4	4.5	27.8	9.4	5.1	43.1	14.5	25.9	35.0	7.7	12.7	9.9	7.7	3.9	10.4	5.1	26.5	8.7			
	%	0.9	14.6	0.3	0.4	3.6	4.3	2.6	2.6	1.3	7.7	2.6	1.4	12.0	4.0	7.2	9.7	2.1	3.5	2.8	2.1	1.1	2.9	1.4	7.4	2.4			
Industrial	Area	2258.3	5.6	0.0	0.2	12.1	108.4	12.4	180.7	22.9	95.6	15.7	2.6	134.1	30.8	35.2	133.3	14.6	26.8	84.6	179.7	264.2	423.0	153.7	214.9	106.9			
	%	5.4	0.2	0.0	0.0	0.5	4.8	0.6	8.0	1.0	4.2	0.7	0.1	5.9	1.4	1.6	5.9	0.6	1.2	3.7	8.0	11.7	18.7	6.8	9.5	4.7			
Natural Areas and Open Spaces	Area	13075.5	78.2	2.0	13.5	151.0	42.8	222.4	57.2	47.7	209.0	31.5	66.1	97.1	560.8	1819.4	262.1	629.7	2050.6	603.9	126.3	615.9	392.7	1399.0	764.4	2832.0			
	%	31.5	0.6	0.0	0.1	1.2	0.3	1.7	0.4	0.4	1.6	0.2	0.5	0.7	4.3	13.9	2.0	4.8	15.7	4.6	1.0	4.7	3.0	10.7	5.8	21.7			
Educational Amenities	Area	852.3	20.1	3.2	2.7	17.8	18.3	39.9	26.7	18.3	10.9	97.6	22.5	50.1	58.5	33.0	33.0	13.3	20.1	16.5	34.9	23.3	16.6	25.2	232.4	17.6			
	%	2.1	2.4	0.4	0.3	2.1	2.2	4.7	3.1	2.1	1.3	11.4	2.6	5.9	6.9	3.9	3.9	1.6	2.4	1.9	4.1	2.7	1.9	3.0	27.3	2.1			
Medical Amenities	Area	316.8	10.9	2.2	5.4	32.2	46.1	18.3	41.9	5.7	19.4	2.8	9.6	16.4	20.4	10.7	8.3	4.9	5.9	4.1	7.5	9.9	5.6	10.3	7.9	10.2			
	%	0.8	3.4	0.7	1.7	10.2	14.5	5.8	13.2	1.8	6.1	0.9	3.0	5.2	6.4	3.4	2.6	1.5	1.9	1.3	2.4	3.1	1.8	3.3	2.5	3.2			
Social Amenities	Area	361.3	17.0	6.9	9.1	16.6	14.9	13.4	10.7	10.0	15.8	5.9	12.9	20.3	30.8	22.6	23.8	8.7	14.2	12.6	11.6	27.4	13.5	12.6	18.6	11.4			
	%	0.9	4.7	1.9	2.5	4.6	4.1	3.7	3.0	2.8	4.4	1.6	3.6	5.6	8.5	6.3	6.6	2.4	3.9	3.5	3.2	7.6	3.7	3.5	5.2	3.2			
Public Utilities and Facilities	Area	693.8	1.6	1.6	0.1	1.5	3.1	8.4	4.0	9.5	14.4	5.0	5.6	17.8	23.8	25.8	4.8	1.5	28.7	8.4	3.9	133.1	78.4	30.3	162.3	119.8			
	%	1.7	0.2	0.2	0.0	0.2	0.5	1.2	0.6	1.4	2.1	0.7	0.8	2.6	3.4	3.7	0.7	0.2	4.1	1.2	0.6	19.2	11.3	4.4	23.4	17.3			
Transport and Communication Facilities	Area	5319.5	280.8	168.5	50.7	144.9	203.9	223.2	283.7	198.3	162.7	117.4	171.0	257.6	409.6	353.1	275.1	147.5	265.6	188.0	245.0	243.9	210.5	239.1	272.7	207.0			
	%	12.8	5.3	3.2	1.0	2.7	3.8	4.2	5.3	3.7	3.1	2.2	3.2	4.8	7.7	6.6	5.2	2.8	5.0	3.5	4.6	4.6	4.0	4.5	5.1	3.9			
Urban Villages	Area	319.7	13.4	0.0	0.0	0.0	0.0	1.9	2.1	5.3	13.1	0.0	19.6	8.2	66.3	94.0	17.2	4.4	5.7	5.8	9.6	17.2	19.2	2.1	9.3	5.5			
	%	0.8	4.2	0.0	0.0	0.0	0.0	0.6	0.7	1.6	4.1	0.0	6.1	2.6	20.7	29.4	5.4	1.4	1.8	1.8	3.0	5.4	6.0	0.7	2.9	1.7			
Primary Activity	Area	1594.8	0.0	0.0	0.3	0.6	0.2	42.3	0.5	0.0	9.3	0.4	2.7	101.1	19.2	27.3	705.4	0.8	4.0	5.6	12.0	0.6	0.2	0.2	300.8	361.1			
	%	3.8	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0	0.6	0.0	0.2	6.3	1.2	1.7	44.2	0.1	0.3	0.4	0.8	0.0	0.0	0.0	18.9	22.6			
Unclassified	Area	1835.5	270.9	0.0	0.0	0.0	0.0	0.5	2.9	0.0	5.5	26.6	0.0	9.5	66.8	70.6	0.0	0.0	0.0	61.4	13.2	1179.5	15.9	18.3	94.0	0.0			
	%	4.4	14.8	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.3	1.4	0.0	0.5	3.6	3.8	0.0	0.0	0.0	3.3	0.7	64.3	0.9	1.0	5.1	0.0			
Vacant Land	Area	3280.6	4.1	5.5	3.4	20.9	33.4	104.5	75.1	20.7	82.8	41.3	41.8	178.0	203.0	490.5	391.9	160.8	234.7	200.2	159.2	210.0	136.7	102.4	201.8	178.0			
	%	7.9	0.1	0.2	0.1	0.6	1.0	3.2	2.3	0.6	2.5	1.3	1.3	5.4	6.2	15.0	11.9	4.9	7.2	6.1	4.9	6.4	4.2	3.1	6.2	5.4			

SOURCE : ELU 2012, SCE INDIA PVT LTD

However, the MCGM has provided data for 90.57% of the total area of Mumbai. The rest of the 9.43% area does not fall under the MCGMs jurisdiction. This results in scattered datasets and often incomplete information. Dividing the city areas is a good idea from the administrative point of view however, the boundaries are only a human representation and do not exist on ground. This division and lack of availability of data for certain parts of the city leads to difficulty in analysis for some wards and for the city overall.

These areas come under the jurisdiction of different government bodies which include the Maharashtra Industrial Development Corridor (MIDC), Mumbai Metropolitan Region Development Authority (MMRDA) and Maharashtra Housing and Development Authority (MHADA).

These areas have been termed at Special Planning Areas (SPAs) and the government bodies responsible for them release data in different formats like csv, PDF, doc etc. Even the coding of land use categories in these areas differs from what



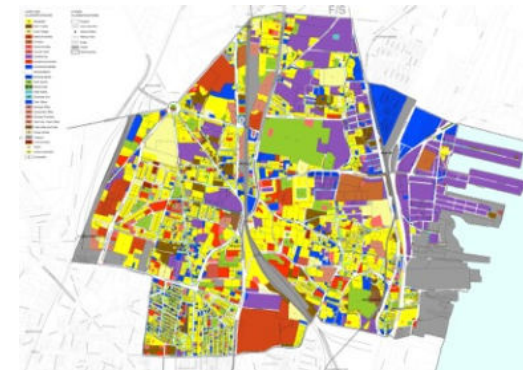
Area under MCGM

BASE MAP SOURCE : MCGM
MAP NOT TO SCALE



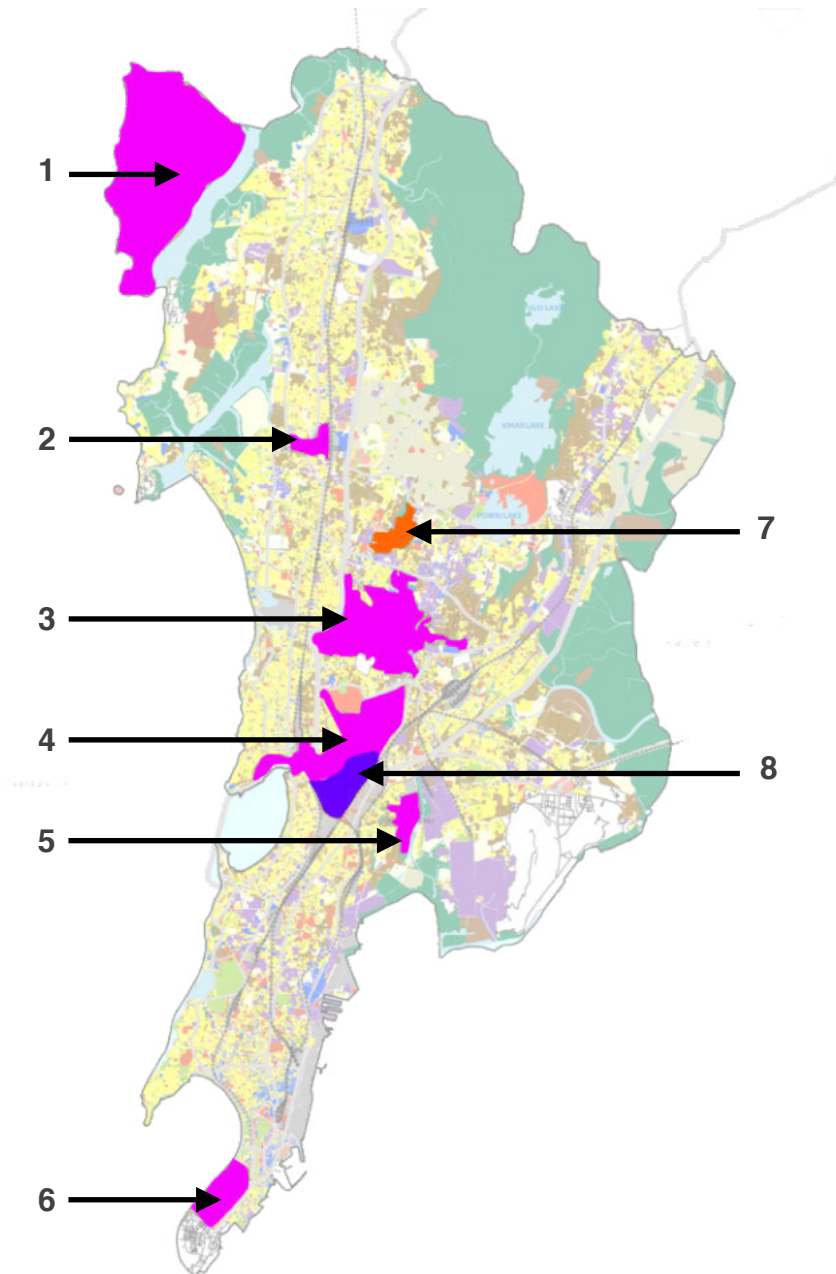
Dharavi

MAP SOURCE : MHADA
MAP NOT TO SCALE



Bandra-Kurla Complex

MAP SOURCE : MMRDA
MAP NOT TO SCALE



Area under SPAs

BASE MAP SOURCE : MCGM
MAP NOT TO SCALE

MMRDA

1. Gorai-Manori Tourism zone
2. Oshiwara District Centre (ODC)
3. Airport
4. Bandra Kurla Complex (BKC)
5. Wadala Truck Terminal
6. Backbay Reclamation Area

MIDC

7. Marol Industrial Area, SEEPZ SEZ

MHADA

8. Dharavi

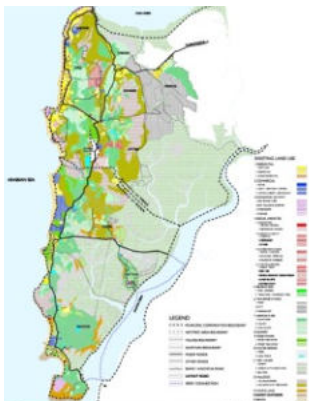
Special Planning Areas

Apart from the area under MCGM, the areas under the SPAs fall under the jurisdiction of MMRDA, MIDC and MHADA respectively.

SPAs under MMRDA

1. The Gorai-Manori Tourism Zone:

This region is located in the north-western part of Mumbai with large areas under mangroves, beaches, rocky outcrops and geological rock formations. Most notable parts of this SPA are the Gorai Beach, Essel World and the Global Vipassana Pagoda.



Gorai Manori Tourism Area

MAP SOURCE : MMRDA
MAP NOT TO SCALE

2. Oshiwara District Centre (ODC):

This region lies between Goregaon and Jogeshwari Railway stations on the western line and is primarily a commercial zone.



Oshiwara District Centre

MAP SOURCE : MMRDA
MAP NOT TO SCALE

3. International Airport SPA:



International Airport

MAP SOURCE : MMRDA
MAP NOT TO SCALE

4. Bandra Kurla Complex (BKC):

BKC is a predominantly commercial district and houses the Diamond bourse and offices of a number of national and international organizations.



5. Wadala Truck Terminal:

The Wadala Truck Terminal is developed by the MMRDA and will have a centralized facility for the transport of goods by road.

6. Backbay Reclamation Area

This region is located near the southern tip of Mumbai and was planned for reclamation in the pre-independence era.



Backbay Reclamation Areas

MAP SOURCE : MMRDA
MAP NOT TO SCALE

SPA under MIDC

Marol Industrial Area

This region is an industrial district with a large number of small and medium scale industries

SPA under MHADA

Dharavi

It is one of the largest slum areas in the world and houses residences and small scale industries for leather, pottery etc. It has one of the highest population densities in the world.



Dharavi

MAP SOURCE : MHADA
MAP NOT TO SCALE

Slum population accounts for 41.95% of the total population of the city¹. The Special Planning Areas such as Dharavi and Gorai-Manori Tourism Area house a sizable amount of population in the city and it becomes an impairment to analyzing or looking at the data of the city when the data of these SPAs is not readily available.

¹ MCGM Stakeholders consultation workshop, 2013

Despite the availability of so a large quantity data, it is very difficult for a person looking at the data to make any sense of it. It is due to the fact that most of the data is in a tabular format. One can infer little by looking at tables and maps with a long legend next to it.

GREATER MUMBAI - BROAD LAND USE CATEGORIES																									
REGIONAL LAND USE CATEGORIES	WARD CITY	WESTERN SUBURBS												EASTERN SUBURBS											
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Area (Sq Km)	4088.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Population (No)	4088.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Residential	Area	3889.4	95.1	44.1	23.2	27.8	10.3	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8
	Pop	3889.4	95.1	44.1	23.2	27.8	10.3	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8
	%	95.1	44.1	23.2	27.8	10.3	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	
	%	95.1	44.1	23.2	27.8	10.3	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	
	%	95.1	44.1	23.2	27.8	10.3	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	46.0	1.8	
Industrial	Area	209.6	5.1	2.6	3.1	1.0	2.1	0.8	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	
	Pop	209.6	5.1	2.6	3.1	1.0	2.1	0.8	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	
	%	5.1	2.6	3.1	1.0	2.1	0.8	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7		
	%	5.1	2.6	3.1	1.0	2.1	0.8	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7		
	%	5.1	2.6	3.1	1.0	2.1	0.8	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7	1.9	0.7		
Public Utilities and Facilities	Area	10.0	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
	Pop	10.0	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
	%	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		
	%	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		
	%	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		
Other	Area	10.0	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
	Pop	10.0	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
	%	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		
	%	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		
	%	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1		

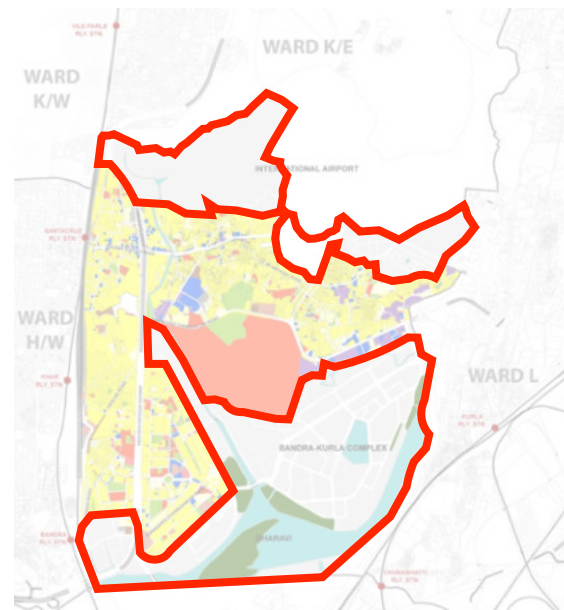
SOURCE : ELU 2012, SCE INDIA PVT LTD

Another problem with the data available is that it is spread out across multiple PDF files. There is little inference that can be gained from these files and it becomes much difficult to use the data in a constructive manner. overlapping datasets

Incomplete Datasets

Ward wise ELU data has been provided by the MCGM. However, this data is grossly incomplete for some wards, since parts of these wards fall under the SPAs, which is falls out of the MCGMs jurisdiction.

An example of this can be seen with ward H/E. A major part of the wards' area falls under the Airport and Bandra-Kurla Complex SPA. However, the ELU data that has been captured excludes these areas, which account for a sizable part of the ward's overall area. There are a number of examples like these in different wards.



WARD H/E : A LARGE PART OF THE WARD'S AREA FALLS UNDER SPA AND THEREFORE HAS BEEN EXCLUDED FROM THE WARD DATA.

SOURCE : ELU MAPS, MCGM
MAP NOT TO SCALE

It also becomes important to look at the data of the city in its entirety and not in silos.

1+1 = 3 in the case of data. Much value can be gained from data sets if they are looked at in context with each other. An affinity can be identified between different almost disparate data sets to gain new insights, generate new knowledge.

The data available can be either pointed out on a map or is a value that cannot be measured by its physical dimensions. There would also be data that falls into both these categories. It needs to be noted that all the data being used is quantifiable in nature i.e. we can define its position by means of a co-ordinate system or it is a number like the population of a ward. For this reason, it makes sense to divide the data type into spatial and non-spatial.

Spatial Data : It is data that can be pointed out on a map. It is situated in a space. The value of the data can be extracted from map, co-ordinates, etc in 2D space and 3D co-ordinates in a 3D space.

Non-Spatial Data : It is data that cannot be pointed out on a map, it exists as a number in time.

List of Available Datasets :

MCGM

Existing Land Use of all Wards

MMRDA (Land Use Data)

Gorai-Manori Tourism zone

Oshiwara District Centre (ODC)

Airport

Bandra Kurla Complex (BKC)

Wadala Truck Terminal

Backbay Reclamation Area

MIDC and SEEPZ (Land Use Data)

Marol Industrial Area, SEEPZ SEZ

MHADA (Land Use Data)

Dharavi

Housing

Cessed Buildings

Heritage Sites and Buildings

Ready Reckoner Rates (Housing prices)

Slums

Health

Hospitals in wards

Top 5 Diseases

Top 5 Sensitive Diseases

Education

Gender Ratio in Primary Education

Dropout numbers in Primary Eductaion

No. of teachers

Distribution of schools and colleges

Transportation

Existing transport networks

Proposed transport networks

Auto and Taxi stands

Peak time traffic roads

Heavy pedestrian congestion areas

Road Accident prone spots

Water and Sanitation

Waste management facilities

Settlements not served by sewer lines

Livelihood

Employment Rate

Open Spaces

Wardwise Open space data

Environment

Natural spaces and water bodies

Flood and Cyclone prone areas

Intertidal Zones

CRZ Boundaries

Wardwise Tree cover

Noise Pollution (Ambient noise)

Air Quality

Governance

Police Stations and Jurisdiction

Ward wise FIRs : Murders

Ward wise FIRs : Crimes against women

Ward wise FIRs : Rapes

Ward wise FIRs 2011 - 2014

Population

Ward wise population Data 2001-2011

Declining Growth rate of Population

Slum and Non-Slum population

Budget

Spatial Data	Non-Spatial Data
Existing Land Use*	Ready Reckoner rates
Cessed Buildings	Top 5 diseases
Heritage sites and buildings + ASI sites	Top 5 sensitive diseases
Slums	Gender ratio in Primary education
Existing and proposed transport networks	Dropout numbers in Primary education
Road accident prone spots	Number of teachers
Auto-rickshaw and Taxi stands	Peak time traffic records
Heavy Pedestrian Congestion areas	Employment rate
Settlements not served by sewer lines	Number of vehicles
Flood and Cyclone prone areas	Noise pollution

Spatial Data	Non-Spatial Data
Intertidal Zones	Air quality
CRZ boundaries	Ward wise FIRs
Ward wise tree cover	Ward wise population - Slum and Non-Slum
	Budget 2015-16

*The Existing Land Use data is a major dataset and includes data about :

- Medical Amenities
- Educational Amenities
- Open Space
- Natural Spaces and Water Bodies
- Residential areas
- Urban Villages
- Commercial Areas
- Offices
- Industrial Areas
- Social Amenities
- Public Utilities and Facilities
- Transport and Communication Facilities
- Primary Activity
- Vacant Land
- Unclassified Areas

The ELU data for all the wards has been taken from MCGM as well as from the SPA authorities which include MMRDA, MIDC and MHADA

Existing Land Use

Existing Land Use (ELU) is one of the major datasets and it includes information about the location and area of different land use categories. Each of these categories like Residential, Commercial, Industrial etc. has been further subcategorized into more than 140 subcategories.

The following table shows the different land use categories and sub-categories with their color code¹

R : Residential

R1 : Individual Housing

R1.1 : Primary Residential

R1C : Residential + Commercial

R2 : Apartments / Multifamily

R2.1 : Primary Residential

R2S : Residential + Shopping

R2C : Residential + Commercial

R3 : Government / Municipal Staff / Quarters / Housing

R3.1 : Primary Residential

R3S : Residential + Shopping

R4 : Chawls

R4.1 : Primary Residential

R4S : Residential + Shopping

R4C : Residential + Commercial

R4I : Residential + Industrial

R5 : Slums / Clusters

R5.1 : Primary Residential

R5C : Residential + Commercial

R5I : Residential + Industrial

C : Commercial

C1 : Retail Markets

C1.2 : Shopping Centre

C1.3 : Mall

C1.1 : Municipal Market

C1.4 : Informal Market

C1.5 : Weekly Market

C2 : Wholesale Markets

C3 : Hotels

C3.1 : Hotels / Boardings / Lodges

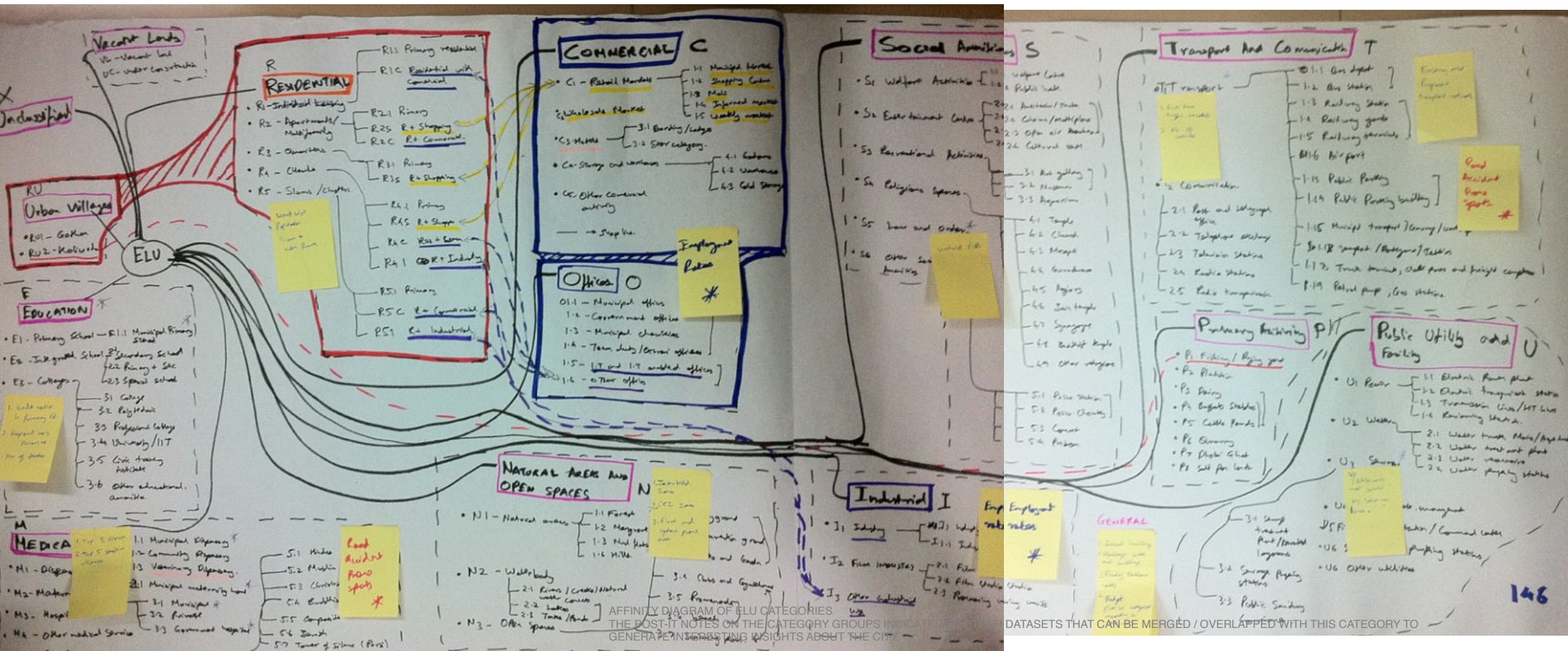
C3.2 : Star Category Hotels

¹ Preparatory Studies DP 2014-34 : MCGM

C4 : Storage and Warehouses	
C4.1 : Godowns	
C4.2 : Warehouses	
C4.3 : Cold Storages	
C5 : Other Commercial Activity	
---- Shop Line	
O : Offices	
O1.1 : Municipal Offices	
O1.2 : Government Offices	
O1.3 : Municipal Chowkies	
O1.4 : Town Duty / Octroi Offices	
O1.5 : I.T and I.T Enabled Offices	
O1.6 : Other Offices	
I : Industrial	
I1 : Industry	
I1.1 : Industrial Estate	
I2 : Film Industry	
I2.1 : Film City	

I2.2 : Film Studio	
I2.3 : Processing Units	
I3 : Other Industrial Use	
RU : Urban Villages	
RU1 : Gaothan	
RU2 : Koliwada	
P : Primary Activities	
P1 : Fishing Drying Yards	
P2 : Plantation	
P3 : Dairy	
P4 : Buffalo Stables	
P5 : Cattle Pounds	
P6 : Quarry	
P7 : Dhobi Ghat	
P8 : Salt Pan Lands	
N : Natural Areas and Open Spaces	
N1 : Natural Areas	
N1.1 : Forest	
N1.2 : Mangrove Forest	
N1.3 : Mud Flats	

N1.4 : Hills	
N2 : Water Body	
N2.1 : Rivers / Creeks / Natural Water Courses	
N2.2 : Lakes	
N2.3 : Tanks / Ponds	
N3 : Open Spaces	
N3.1 : Playground	
N3.2 : Recreation Ground	
N3.3 : Parks and Gardens	
N3.4 : Clubs and Gymkhana	
N3.5 : Promenade	
N3.6 : Beach	
N3.7 : Swimming Pool	
VL : Vacant Lands	
VL : Vacant Land	
UC : Under Construction	
X : Unclassified	
M : Medical Amenities	
M1 : Dispensary	
M1.1 : Municipal Dispensary	



M1.2 : Community Dispensary

M1.3 : Veterinary Dispensary

M2.1 : Municipal Maternity Home

M3 : Hospital

M3.1 : Municipal Hospital

M3.3 : Government Hospital

M3.2 : Private Hospital

M4 : Other Medical Services

M5 : Cemetery

M5.1 : Hindu Traditional or Electric

M5.2 : Muslim

M5.3 : Christian

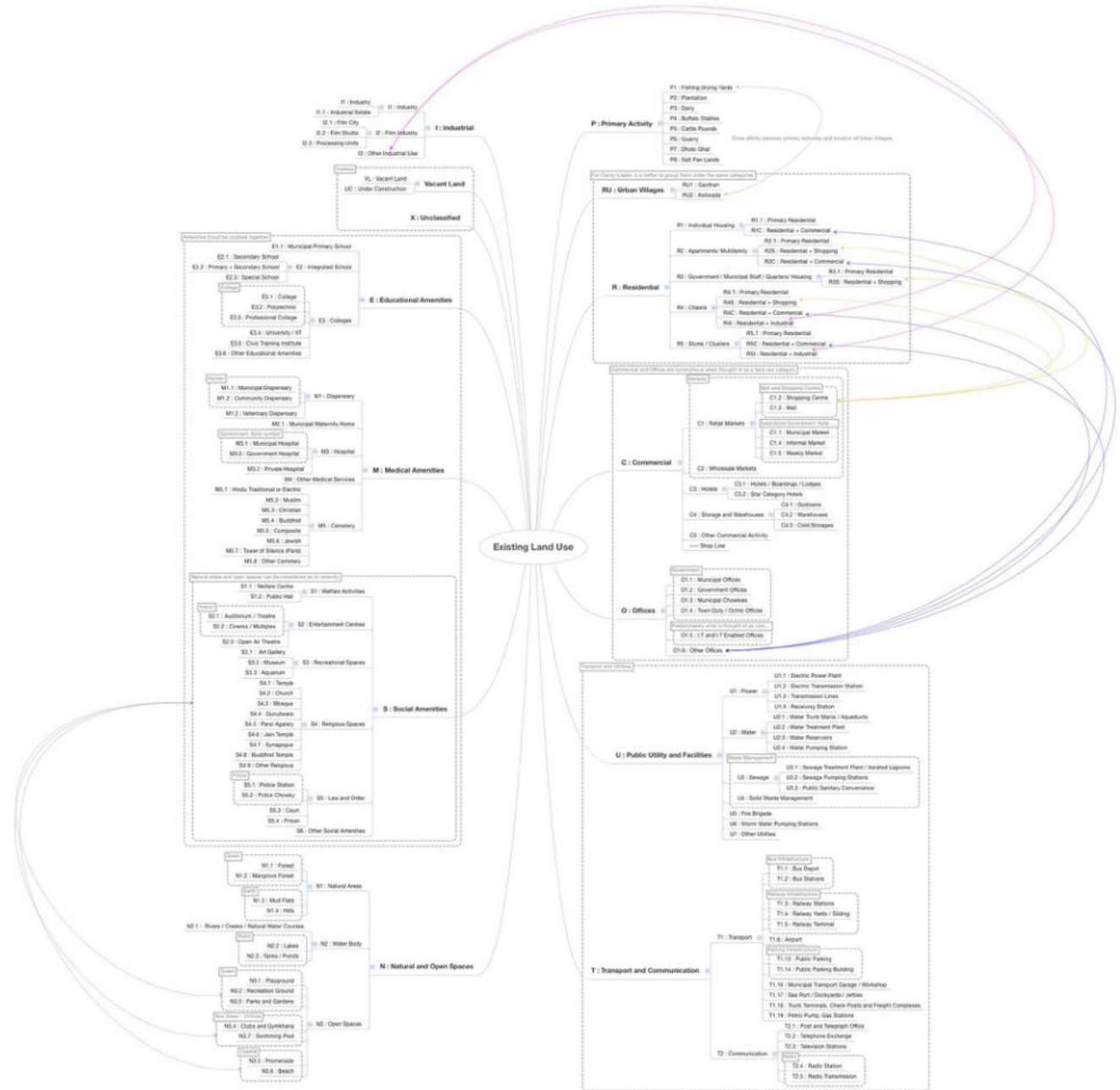
M5.4 : Buddhist

M5.5 : Composite

M5.6 : Jewish

M5.7 : Tower of Silence (Parsi)

M5.8 : Other Cemetery



M1.2 : Community Dispensary	
M1.3 : Veterinary Dispensary	
M2.1 : Municipal Maternity Home	
M3 : Hospital	
M3.1 : Municipal Hospital	
M3.3 : Government Hospital	
M3.2 : Private Hospital	
M4 : Other Medical Services	
M5 : Cemetery	
M5.1 : Hindu Traditional or Electric	
M5.2 : Muslim	
M5.3 : Christian	
M5.4 : Buddhist	
M5.5 : Composite	
M5.6 : Jewish	
M5.7 : Tower of Silence (Parsi)	
M5.8 : Other Cemetery	

E : Educational Amenities	
E1.1 : Municipal Primary School	
E2 : Integrated School	
E2.1 : Secondary School	
E2.2 : Primary + Secondary School	
E2.3 : Special School	
E3 : Colleges	
E3.1 : College	
E3.2 : Polytechnic	
E3.3 : Professional College	
E3.4 : University / IIT	
E3.5 : Civic Training Institute	
E3.6 : Other Educational Amenities	
S : Social Amenities	
S1 : Welfare Activities	
S1.1 : Welfare Centre	
S1.2 : Public Hall	
S2 : Entertainment Centres	
S2.1 : Auditorium / Theatre	

S2.2 : Cinema / Multiplex	
S2.3 : Open Air Theatre	
S3 : Recreational Spaces	
S3.1 : Art Gallery	
S3.2 : Museum	
S3.3 : Aquarium	
S4 : Religious Spaces	
S4.1 : Temple	
S4.2 : Church	
S4.3 : Mosque	
S4.4 : Gurudwara	
S4.5 : Parsi Agaiary	
S4.6 : Jain Temple	
S4.7 : Synagogue	
S4.8 : Buddhist Temple	
S4.9 : Other Religious	
S5 : Law and Order	
S5.1 : Police Station	
S5.2 : Police Chowky	
S5.3 : Court	
S5.4 : Prison	

T2 : Communication

T2.1 : Post and Telegraph Office

T2.2 : Telephone Exchange

T2.3 : Television Stations

T2.4 : Radio Station

T2.5 : Radio Transmission

U : Public Utility and Facility

U1 : Power

U1.1 : Electric Power Plant

U1.2 : Electric Transmission Station

U1.3 : Transmission Lines

U1.4 : Receiving Station

U2 : Water

U2.1 : Water Trunk Mains / Aqueducts

U2.2 : Water Treatment Plant

U2.3 : Water Reservoirs

U2.4 : Water Pumping Station

U3 : Sewage

U3.1 : Sewage Treatment Plant / Aerated Lagoons

U3.2 : Sewage Pumping Stations

U3.3 : Public Sanitary Convenience

U4 : Solid Waste Management

U5 : Fire Brigade

U6 : Storm Water Pumping Stations

U7 : Other Utilities

A diagram was made to see the affinity between different categories and sub-categories in the ELU. There are a number of instances where the subcategories in the same or different category have an affinity with each other.

Also post-it notes were pasted on top of category groups. These indicate the other datasets like crime, health, employment , etc., that can be merged or overlapped with these ELU categories to gain insights about the city.

Important instances of affinity between include the relationship between categories R1C, R2C, R4C and R5C with Other Offices category (O1.6) as well as the affinity between the Residential + Shopping Categories with Shopping Centers and Malls. Another interesting affinity was found between the Residential + Industrial Categories and the Other Industrial Use (I3). There is also a close affinity between Primary Activities (P) and the location of Urban Villages (RU). Also there is a relationship between Open Spaces and Social Amenities.

Re-categorization of the ELU

From the observation of the categorization of the ELU, it was seen that many categories and sub-categories have an affinity with each

other. Some sub-categories fit into two boxes.

There is a need to re-organize the categories. One need not do an overall organizational change since the data provided is linked to these category names. However, it was found that it is fine to group some of these categories and sub-categories together.

The Urban Villages Category (RU) and the Residential Category (R) could be clubbed together since both these area are predominantly residential in nature. Commercial (C) and Offices (O) could also fit into a broader category. All amenity categories like Educational (E), Medical (M) and Social (S) have a possibility of forming a larger group. At the same time there could be a broad category for Public Utilities (U) and Transportation and Communication (T) categories.

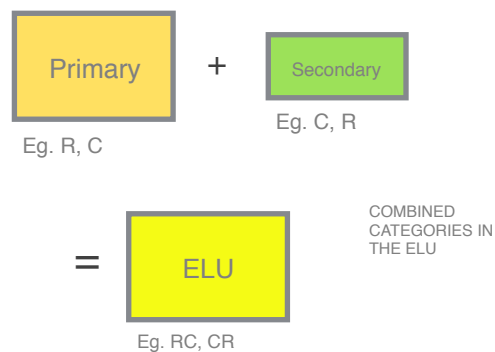
The organizing diagram on the page shows different land use categories and sub categories in the ELU. The colored lines indicate affinity between different sub-categories.

The boxes around multiple categories indicate that these can be merged into a single category and its content can be branched out.

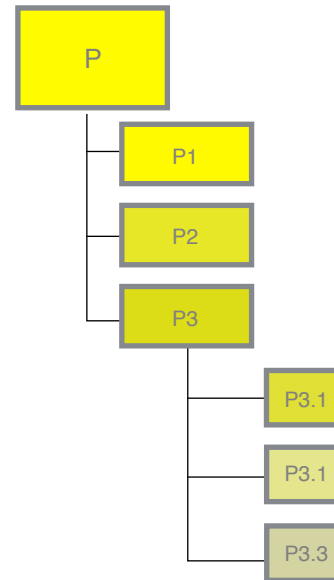
Visualization of the data

The Re-categorization of the ELU can have a much lesser cognitive load on the user if we build in redundancy and patterns in this re-organization.

The ELU mentions of sub-categories such as RC, CR, RS, etc. The first letter is the primary category and the next letter is the secondary category. In the example of 'RC'; 'R' indicates residential while 'C' indicates that the residential also has a commercial component in the land use. One can sometimes find a number between the two letters, which indicates a primary category number. For example, 'R2C' means the land use has residential as its primary component and commercial as the secondary



component, while the number 2 is associated by the primary category of apartments. This categorization although complex in the beginning is simple to understand if the underlying logic is known.



COLORATION FOR CATEGORIES IN THE ELU

For sub-categories in the same primary category, the difference earlier was shown by way of overlaying text on top of the colors. However this could be done away with by associating change in hue by change of category. For further sub-categories, the proposed solution could be a change in saturation.



EXAMPLE OF ICONOGRAPHY FOR THE ELU

Another way of building redundancy is by way of using appropriate iconography for the ELU categorizations. For combined categories like RC, CR, etc a similar approach like the for color can be followed. However, these icons may or may not be used directly on the map itself.

The project overall aims is to look at the open data about the city of Mumbai and present it in an easily understandable manner for its citizen without dumbing down the data. The objective is to do a visualization of data to show it in ways that the users, observers of the data can gain new insights about the city. The ELU data provides a majority of the dataset needed to visualize the data. However, additional data has been collected from various other government organizations such as the MMRDA, MIDC, SEEPZ and others mainly for land use data for the SPAs as well as NGOs and other organizations like Praja, The Urban Design and Research Institute (UDRI), Mashal, etc for other useful data like crime rate, education among other data sets.

“Re-describing the world is a necessary first step towards changing it¹.”
Salman Rushdie

The visualizations for the project can be categorized into broad categories as follows :

- Housing
- Livelihood
- Education
- Health
- Open Spaces

- Environment
- Crime
- Transportation
- Water and Sanitation

These categories are broad categories and mainly the available data from credible sources would be used to generate the visualizations. In some cases, only relevant data for a particular visualization may be used. A systemic design problem would be to create a system for visualization of the data to for reasons of scalability incase more diverse and relevant data becomes available.

It is also important to consider that the data should not be looked at in silos. For this reason, it is better to compare the data of the city either with other cities that are similar in nature or if it is the case of wards, looking at adjoining wards could be a solution. This helps in forming a judgement about the data in context with nearby areas. For example, the deficiency of Hospitals in a particular ward may be an alarming fact. Although if the adjoining ward has a large hospital in its boundaries; it could mean an excess of hospitals in the adjoining ward, thereby negating the deficiency in the ward being looked at.

It might not always be the case where the creator of the data visualization may know the intended use, or all the insights the visualizations shows. Too much data also is confusing. Giving the user a way to explore the datasets and find new insights is also a possible way to present the data. A system of layering datasets for multiple interpretations can be adopted.

Let the data speak.

“Simplicity of reading derives from the context of detailed and complex information, properly arranged. A most unconventional design strategy is revealed: to clarify, to detail².”
Edward Tufte

¹ Imaginary Homelands: Salman Rushdie

² Envisioning Information: Edward Tufte

