Liases Foras
Real Estate Rating & Research Pvt. Ltd.

Independent Non-broking
Real Estate Research Company

Housing Price Index (HPI) Computation Methodology

By

Liases Foras

TABLE OF CONTENTS

1 INTRODUCTION

Liases Foras, the only non-broking real estate research company in India, collects residential market data by conducting field/telephonic surveys of the primary residential projects through our dedicated survey team. Since 2003, Liases Foras has been collecting data through these primary survey techniques every quarter. It goes through rigorous checks and is verified using **mystery shopping** techniques. A dedicated team of analysts builds on this historical data to build the products and services tailored to meet the needs of businesses and institutions.

Information captured about the project includes,

- Essential attributes like Project Name, Location, City, RERA registration numbers, Geographical coordinates (latitude, longitude), etc.
- Builder quoted price - from the cost sheets and brochures.
- Total units and unit type (1BHK, 2BHK, Plot, Villa, etc.)
- The status of number of units available for sale (Unsold)
- The construction stage of each building/wing in the project
- Construction photographs

The project's quoted price is collected each quarter to create a consistent price trend for an individual project. Similarly, more than 18,000 such projects are being surveyed every quarter by Liases Foras across 60 cities in India. Overall Liases Foras has data of above 65,000 projects in its life time. Thus, the price collected at the most granular level helps us to derive the average price of any higher regional hierarchy such as pin-code, group of multiple pin-codes, sub-region, or district. Liases Foras adopts the "**repeat sales**" method to compute the price index for a project for the entire life span of its primary sales, which, on average, last for five years. The repeat-sales method assesses how house prices change over time by focusing on the sale prices of the same piece of real estate. Various housing price indexes have adopted the repeat-sales method to eliminate the problem of accounting for price differences in houses with varying characteristics.

2 HPI COMPUTATION

2.1 Scope

Liases Foras captures primary residential market data of India's top 60 Tier 1 and Tier 2 cities. These cities demarcate 80-90% of the builder market in the country. The data collected is at a project level and geo-coded to the project's location. This allows aggregating residential markets at various geographical levels such as pin-codes, sub-regions, regions, districts, etc. Since this data is updated at a regular interval, every quarter (3 months), a track record of price changes can be derived at various geographies by applying statistical analysis and developing a stable Housing Price Index.

Pin codes, one of the standard identifiers for location, are selected as the base geography for which the price index is developed. Given the 7000+ pin-codes of India won't have equity in marketable projects with a price, for such pin-codes where data is thin, the price index is derived using statistical methods and GIS technology *(ref sec 2.3.3)*.

For estimating the change in price every quarter, **repeat sales** method is adopted, where shift in price in each project in a quarter is averaged to derive the change in price for that geography. While this method has its limitations, such as there is no price trend as only a change in price is considered, it provides the most stable index trends and is independent of the type of supply in the region.

2.2 Definitions & Terminologies

2.2.1 Project

Project refers to a residential apartment project from an established developer. While the underlying data is diverse in its quality, where the project is divided into wings/sub-projects and further into unit types, such as 1BHK, 2BHK, etc., for the purposes of index preparation, the data is consolidated at the project level.

All project data is geocoded to its respective location. The geo-coordinates are used for aggregating prices at various boundaries.

Each project has a price, which is calculated as a simple average of various prices offered by the marketable supply/wings in the project. All prices considered are carpet prices. The source of the price data is from the RESSEX Residential Repository, maintained by Liases Foras. The data source is updated and validated every quarter via a primary survey conducted and managed by Liases Foras.

2.2.2 Boundary

A boundary could be any geographical boundary, such as pin codes, groups of pin codes, sub-regions, etc, on which the subject matter is discussed.

2.2.3 Pin-code

All pin-codes from the Department of Posts website (https://www.indiapost.gov.in/) are sourced/developed as GIS shapefiles. Project coordinates that fall under these pin-code shapes will be used for the calculation of the index for the respective pin-codes.

In addition to this, several pin codes are in use by HDFC but not found in the shapefile, these are represented as **missing pin-codes** for which index will be derived from the respective district-level index.

2.2.4 Sub-regions

Sub-region boundaries are Liases Foras designated boundaries for defining cities or breaking up of cities. While pin-codes are smaller than sub-regions in terms of area, pin-codes can't be combined to form sub-regions as the source and criteria for classification are independent for both types of boundaries. For example, in the figure below, the white dotted line represents boundary of the sub-region definition, and it is visible that the pin-code 400093 lies between two sub-regions.

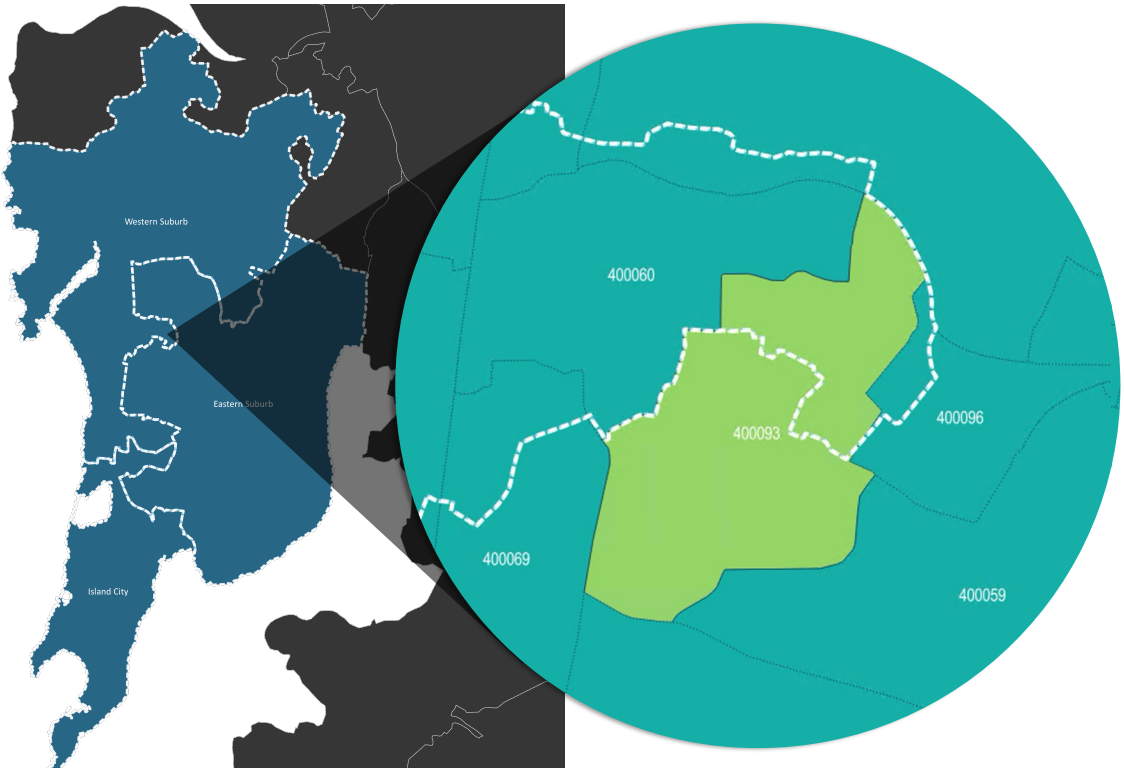


Figure 1 Sub-regions & Pin-codes

2.2.5 District

Latest district shapefiles available, sourced from public sources with definitions updated as of 2022, are used for the purpose of district-level index preparation. The same index is used to derive the price index for the **missing pin-codes**[ref 2.2.3]. The final shapefile for the district is modified with respect to the inputs from HDFC before use for index preparations. The modifications include delineating shapefile district boundary as per HDFC definitions of district, which is based on pin-codes and respective districts.

2.2.6 Cities Covered by Liases Foras

Liases Foras covers over 75 census cities (60 urban agglomerations), where residential builder market is prominent.

Table 1 List of cities covered by Liases Foras

MMR	Alibag Region	Agartala	Jaipur	Ranchi
	Greater Mumbai Region	Agra	Jamshedpur	Sangli
	Mira road - Virar Region	Ahmedabad	Jodhpur	Shimla
	Navi Mumbai Region	Alwar	Kanpur	Shirdi

	Neral - Karjat Region	Aurangabad	Karnal	Silvassa
	Palghar - Boisar Region	Bangalore	Kolkata	Solapur
	Panvel-Uran Region	Bhopal	Kota	Srinagar
	Pen Region	Bhubaneswar	Lonavala	Surat
	Shahpur - Titwala Region	Chandigarh	Lucknow	Thiruvananthapuram
	Thane Region	Chennai	Ludhiana	Tiruchirappalli
NCR	Delhi	Cochin	Madurai	Tiruppur
	Dharuhera-Bhiwadi	Coimbatore	Mangalore	Vadodara
	Faridabad	Daman	Meerut	Vapi
	Ghaziabad	Dehradun	Nagpur	Vijayawada
	Gurgaon	Dhanbad	Nashik	Vizag
	Noida & Greater Noida	Goa	Neemrana	
	Sonipat-Kundli	Guwahati	Patna	
PUNE	PCMC Region	Gwalior	Puducherry	
	Pune	Hyderabad	Raipur	
		Indore	Rajkot	

2.2.7 Actual Price Index

When the index is calculated by averaging change in price in projects within the considered boundary, the index is tagged as '*Actual*'.

2.2.8 Derived Price Index

When index is calculated by combining neighboring boundaries or by using a higher-level boundary, the index is tagged as *'Derived' (ref. sec. 2.3.3 for more info)*. When an index is derived, the boundary from which the index is derived is also mentioned separately. For example, if the index is derived by grouping pin-codes, it will be tagged as *'Derived from Adjacent Pincode Price Change'.*

Table 2 Possible tags for derived indices

Tag	Derivation
Actual	*Actual*
Derived	*Derived from Adjacent Pincode Price Change*
Derived	*Derived from SubRegion Price Change*
Derived	*Derived from District Price Change*
Derived	*Derived from Adjacent District Price Change*
Derived	*Derived from All India Price Change*

2.2.9 Marketable Project

A project being sold by a developer and has unsold units in the respective time frame considered is called a marketable project.

2.2.10 Booking Stop Project

A project which has stopped booking due to various reasons, such as litigations or construction has stalled etc. are booking stop projects.

2.3 Methodology

2.3.1 Repeat Sales Method

The repeat sales method tracks a project's price change across its lifetime, and the cumulative change will be the final price index for that project. Since Index is prepared for pin-codes, the average of all projects' price appreciation/depreciation is considered as the price appreciation/depreciation for the pin-code. Index is calculated as floating points hence not rounded to integers.

Table 3 Sample boundary index trend

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Project 1	-	7.00%	9.03%	-1.43%	5.01%	9.00%	12.99%	1.99%
Project 2	-	6.00%	-3.30%	-2.44%	5.82%	6.00%	7.00%	7.00%
Project 3	-	-6.00%	9.04%	-1.63%	6.00%	6.00%	12.00%	0.78%
Project 4	-	8.00%	1.01%	-4.17%	4.17%	2.46%	2.21%	0.86%
Project 5	-	8.00%	-2.78%	0.79%	2.90%	-1.29%	2.00%	-1.54%
Boundary	**-**	**4.60%**	**2.60%**	**-1.77%**	**4.78%**	**4.44%**	**7.24%**	**1.82%**
Price Index	**100**	**105**	**107**	**105**	**110**	**115**	**124**	**126**

2.3.2 Assumptions and Minimum Requirements

- Only marketable primary/builder apartment projects in the respective quarter are considered for index calculation.
- A minimum of 5 projects are required in any boundary (pin-codes, group of pin-codes, sub-regions, etc.) for calculating average price growth. This is to ensure that the average price change of the locality is diverse in its profile.
- A project with a quarterly price change of more than 15% (increase or decrease) and less than 3% of their supply as unsold is excluded from index calculation. This can mitigate sudden price shifts observed in near-to-sold-out projects due to market behaviour and practices.
- When a boundary, say a pin-code, doesn't meet the minimum project and price requirements, the adjacent boundary is combined to form a new boundary or a higher-level boundary like sub-region/district is selected for calculations.
- Booking stop[*ref. sec. 2.2.10*] projects are excluded from index preparation.
- When the price of projects is not available for the quarter, such projects are excluded due to lack of data points.

2.3.3 Deriving Price Index for boundary

Certain minimum requirements[ref. sec. 2.3.2] are applied to the boundary on which the index is calculated to ensure the price index holds up for stability and truthfulness.

When the minimum requirements are met, the average change in price in each project is calculated for the quarter and appended to the previous quarter's index for the same boundary. This index is tagged as *'actual'*.

Table 4 Index calculation sample

Q2 Change in Price		Q1 Index	New index	Q2 Index
Project 1	7.0%	**103**	$q_2 = q_1 (1+c)$	**110**
Project 2	6.0%	(q_1)		(q_2)
Project 3	6.0%			
Project 4	8.0%			
Project 5	8.0%			
Boundary (Avg. of Projects)	**7.0%** (c)			

If the minimum requirements are not met, the boundary is widened until the minimum requirement is met. This could be by grouping adjacent boundaries of the same type or selecting a higher-level boundary. For example,

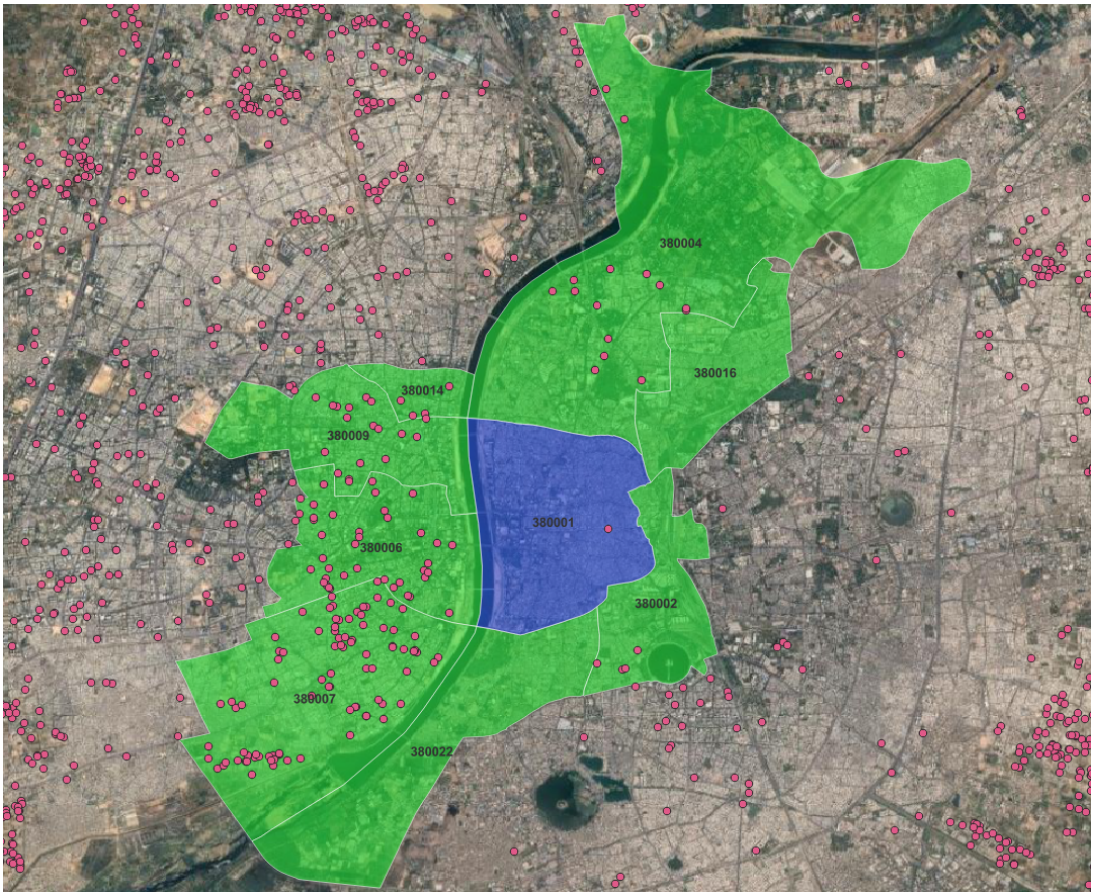


Figure 2 Pin-code with less than 5 projects and selection of adjacent pin-codes to widen boundary

The figure above shows pin code 380001 (blue), which has only one project; for such pin codes, adjacent pin codes (green) within the same district are grouped to create a larger boundary. If the group of pin codes doesn't meet the minimum criteria, then a higher-level boundary, such as a sub-region, is selected. If a pin-code overlaps multiple sub-regions, the sub-region with maximum overlap is considered for the respective pin-code in index derivation.

The following figure has pin-code 382426 (blue), which has no project, and even grouping the adjacent pin-codes (green) doesn't meet the minimum requirement. In such cases, the sub-region is considered as the source for calculating the price index. Since pin-codes do not add up to a sub-region*[ref sec 2.2.4]*, the respective sub-region is selected based on the maximum overlap of the pin code on the sub-region. When the sub-region selected doesn't meet the minimum criteria, the district is chosen for calculation.

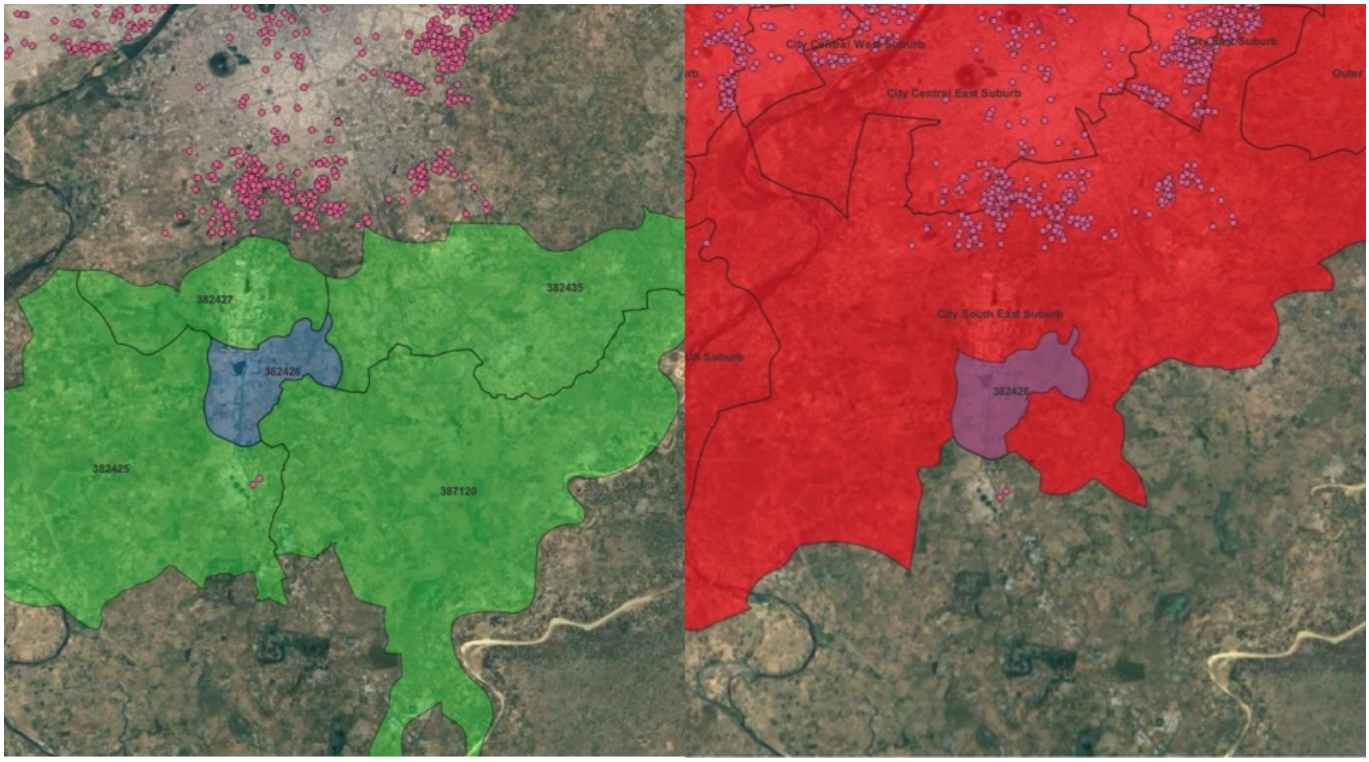


Figure 3 group of pin-code with less than 5 project & the sub-region where the pin-code has maximum overlap with.

If the districts don't meet the minimum criteria, adjacent districts within the same state are joined to make the new boundary. If even grouping the districts doesn't meet the minimum criteria, then the price is derived from Pan India data.

Whenever the price is derived by extending the boundary, the index calculated is tagged as 'derived.' The derivation level, whether from a group of pin codes or sub-regions, is also given separately for reference[ref. sec. 2.2.8].

2.4 Limitations

2.4.1 Price Trend & Index disparity

Since the index is calculated as a price change in projects, the average price trend will not match the index trend.

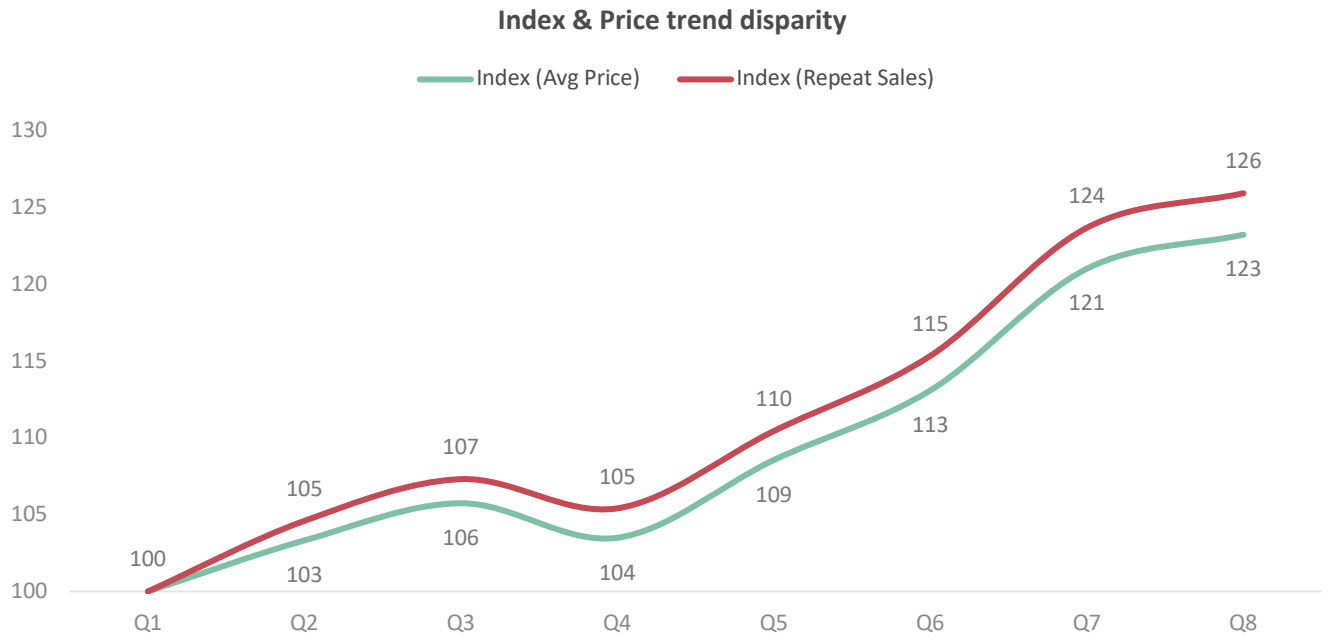
For example, below tables (Table 6 and 7) gives the hypothetical price trend of 5 projects and corresponding of change in price on QoQ. The trend of average price of each quarter, indexed to the initial price in Q1, will not match with index calculated using repeat sales method, hence price trend is not in alignment with the index.

Table 5 Sample Price Trend

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Project 1	3,000	3,210	3,500	3,450	3,623	3,949	4,462	4,551
Project 2	8,000	8,480	8,200	8,000	8,466	8,974	9,602	10,274
Project 3	12,000	11,280	12,300	12,100	12,826	13,596	15,227	15,346
Project 4	11,000	11,880	12,000	11,500	11,980	12,275	12,546	12,654
Project 5	6,000	6,480	6,300	6,350	6,534	6,450	6,579	6,478
Boundary Price (Avg)	**8,000**	**8,266**	**8,460**	**8,280**	**8,686**	**9,049**	**9,683**	**9,861**
Price Indexed to Price of Q1	**100**	**103**	**106**	**104**	**109**	**113**	**121**	**123**

Table 6 Trend of change in price

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Project 1	-	7.00%	9.03%	-1.43%	5.01%	9.00%	12.99%	1.99%
Project 2	-	6.00%	-3.30%	-2.44%	5.82%	6.00%	7.00%	7.00%
Project 3	-	-6.00%	9.04%	-1.63%	6.00%	6.00%	12.00%	0.78%
Project 4	-	8.00%	1.01%	-4.17%	4.17%	2.46%	2.21%	0.86%
Project 5	-	8.00%	-2.78%	0.79%	2.90%	-1.29%	2.00%	-1.54%
Avg Change in Price	**-**	**4.60%**	**2.60%**	**-1.77%**	**4.78%**	**4.44%**	**7.24%**	**1.82%**
Price Index	**100**	**105**	**107**	**105**	**110**	**115**	**124**	**126**


Index & Price trend disparity
Index (Avg Price)
Index (Repeat Sales)
130
125
120
115
110
105
100
Q1
Q2
Q3
Q4
Q5
Q6
Q7
Q8
100
105
107
105
110
115
124
126
103
106
104
109
113
121
123

3 FAQ

1. **How are sub-regions selected to impute pin-code values, for example, Chennai city pin-codes?**

Chennai District has multiple subregions. Each pin code has a different mapping of the subregion based on the largest overlapping area.

2. **Regarding derived values for the pin code, what is the preference from which boundary the index is derived?**

If the pin-code doesn't meet the minimum requirement to calculate the actual index, the next preference is grouping adjacent pin-codes to create a larger boundary. If that doesn't meet the criteria, the sub-region in which the pin code has maximum overlap is used. In case even that is not enough, the district level is considered.

3. **Are multiple pin-codes in the HDFC pin-code mapper missing in the Liases Foras pin-code repository?**

Liases Foras pin-code is derived from the Department of Posts website (https://www.indiapost.gov.in/). The shapefiles for the additional pin codes/missing pin-codes are unavailable. *(ref. sec. 2.2.3)*

For example, the pin-code 121011 is present in HDFC pin-code mapper as a part of Faridabad, while the same is not present in all India pin-code directory offered by the Department of Posts website.

4. **How the imputation are done for missing pin-codes mentioned in FAQ 3?**

For the missing pin codes, where adjacent boundaries can't be set since they are not part of the shapefile, the price is derived from the district level *(ref. sec. 2.3)*.

5. **The price trend in the RESSEX Residential dashboard doesn't tally with the index trend!**

The RESSEX dashboard for residential data shows the weighted average price at various boundaries, while the index is made with changes in price rather than the average price. Also notice that for index calculation several assumptions are taken which will filter some components which may be visible in RESSEX dashboard.

6. **Why carpet prices are chosen instead of saleable price in the exercise?**

Carpet area is the usable area in the apartment while saleable area could be built-up, super built-up etc. In general, the difference between carpet and saleable area is represented by Loading factor which is calculated as:

$$Loading = \frac{Saleable\ Area - Carpet\ Area}{Carpet\ Area}\%$$

What this indicates is that for the same carpet area, Developers can represent different price on saleable area by adding different loading factor.

For example:

Carpet Area (c)	Loading Factor (L)	Saleable Area $s = c*(1+L)$	Carpet Price C_p	Saleable Price $S_p = C_p / (1+L)$
800	40%	1,120	10,000	7,143
800	60%	1,280	10,000	6,250

In the table mentioned above, despite the carpet size being the same, the developer can quote a lesser saleable price (2nd case) by having a higher loading factor. The loading factor may include built-up area/super built-up area. Since price from different developers need to be comparable, carpet price is chosen to exclude effect from loading.