Clustering Analysis in a Complex Real Estate Market: The Case of Ortigia (Italy)

Salvatore Giuffrida, Giovanna Ferluga, and Alberto Valenti

Department of Civil Engineering and Architecture, University of Catania
sgiuffrida@dica.unict.it,
gio.ferluga@virgilio.it,
albvlt79@gmail.com

Abstract. Ortigia, the historic center of Syracuse, is a complex urban entity characterized by high outer homogeneity and inner heterogeneity. The evolution of its real estate market during the last decade is somehow related to the global property market one. In addition its events are connected with the evolution of the exploiting policies still ongoing. The critical observations of its features aim at providing tools able to support the decisions about subsides and local property taxes. This study continues the observations we have carried out for five years, this time involving clustering analysis, a data mining technique able to recognize different submarkets, and suitable to make the valuation pattern fit to the different market areas. For each of the latter significant characteristics have been recognized with reference to the "monetary declination" of these particular capital assets.

Keywords: Imperfect real estate markets, mass-appraisals, clustering analysis, theory of capital, income method.

1 Introduction

The real estate of Ortigia, the historic center of Syracuse, is a complex urban entity characterized by high outer homogeneity and inner heterogeneity. The first is due to its geographical, landscape, architectural and cultural identity that makes it recognizable like a brand; the second is due to its location on an islet, and therefore to its need to concentrate in itself all the urban practical and symbolic functions and activities, so that it looks like a miniaturized complex and articulated context.

The huge amount of investments made during the last fifteen years in order to boost the general renovation policy for one of the main tourist target of tourism in the Mediterranean sea, has generated a bundle of positive economic and negative social and cultural externalities during a renovation process, still ongoing, dominated by the property market for better or worst.

The real estate market observations we have carried out for the last five years, reveal the sequence of about three different and recognizable phases during which the economic, financial and monetary characteristics of property have raised playing different roles. Furthermore, the surveys and the valuations we have carried out so far,

B. Murgante et al. (Eds.): ICCSA 2014, Part III, LNCS 8581, pp. 106–121, 2014.

highlight some criticalities that don't allow to consider Ortigia as a whole market and to use linear calculation tools, mostly because of the process that has been featuring it during the last decade.

At first, during its immature stage, this heterogeneous market has been attacked by a massive trading wave affected by the gap and inconsistency between local owners who were looking backward and foreign buyers who were looking forward. The former despised some functional features, the latter appreciated the symbolic and, above all, the monetary ones. Therefore, the first stage has been characterized by the massive sale of the best architectural heritage, mostly located along the waterfront or facing the most important streets and squares.

The second phase (2005-2009) has been characterized by the success of the brand of Ortigia that confirmed it as one of the most promising real estate markets; during this phase the prices reached the top. The professional intermediaries, who ruled the assessments and the bargaining, have played an important role in this process.

The third stage (since 2010 until now), started about one and half year after the start of the economic-financial crisis, and has been characterized by the lack of liquidity, due to the crisis of credit; trades decreased significantly and many of the former investors, who purchased during the first bullish phase are now trying to sell.

This sort of bubble [14] provided some positive and negative effects: the first ones are the general improvement of the physical and functional condition of this real estate, and the fair relationship between the quality of the location and the market prices; the second ones are the filtering of the local population and the concentration of the property because of some massive investments made by professional investors or groups of real estate investors.

The evolution of this eccentric market is still ongoing and allows us to apply a general approach aimed at recognizing and connecting the multiple relationships between *value* density and tensions of *prices*.

2 Materials: Ortigia and the Real Estate Market Observations

Ortigia is the historic center of Syracuse, situated in a 50 hectares islet connected to the mainland by three bridges. Although it has been for a long time a marginal and decaying quarter, mainly due to some adverse urban-environmental conditions, an improved awareness of its extraordinary heritage has deployed, since 1990 – the date of the most important Renovation Plan by Prof. Pagnano – a huge amount of human, political and economic resources that have been involved in its infrastructural renovation and in the enhancement of its cultural-historical identity. Some other enhancement plans and laws are: the Integrated Communitary Plan, "Urban Italy" including Syracuse in 1995, the "Progetti Sponda", financed by the Province and the Region ex lege n. 433/91, providing 15 mln €; the PIT – the Mediterranean "Environmental-Museum" providing further 16 mln € the special regional laws for Ortigia n. 70/76 and n. 34/85; the Urban Recovery Plan, 5,5 mln €, the Operating program FESR 2007-2013; the Sustainable Development Plan of Syracuse; the Strategic Plan Syracuse Renewal 2020, aiming at the land enhancement and valorisation with further 115 mln € to be devolved to the harbour renovation work in Public-Private Partnership.

Syracuse (Ortigia) and the Rocky Necropolis of Pantalica have been included in the Unesco World Heritage List since 2005 (Committee Decision 29COM 8B.41). Ortigia has become one of the most important economic mover for the city of Syracuse, and the undisputed brand of the whole Province, whose economic policy is aimed at improving tourism compensating some of the worst environmental criticalities caused by the former industrial development strategy; in fact, the huge chemical pole located in the last '70s along the northern waterfront of Syracuse has destroyed and marginalized so far a significant part of the environmental heritage of one of the most beautiful Mediterranean areas.

The real estate market of Ortigia took off in 2003, after a previous depression period when prices were comparable to the marginal areas' ones, despite the incomparable architecture and landscape of this location [6].

The current market situation, that does not follow the general national trend [16], is summarized by the result of the survey carried out in 2013 by collecting a sample of 96 properties for sale and 25 for rent; two databases have been compiled and each property (record) is described by 15 attributes (field) grouped in four types of characters (location, landscape, technologic, architectural-environmental). The fields are organized in a *work breakdown structure* in which each group is divided into a certain number of characters, then in sub-characters whose importance is measured by means of a weighed score system. The heterogeneity of this real estate requires both surface area (*sqm*) and number of rooms (*r*) to be considered as quantitative characters: in fact, sometimes ancient architectural typologies have wide rooms or they are not fairly usable, so that the effective utility do not correspond to the property area [4].

The sale prices overall range is 534-4,958 €/sq.m. and 14,194-148,750 €/r; in the different areas wide ranges have been registered as well: the waterfront properties with terrace reach the maximum value, while the properties with sea view range 735-3,129 €/sq.m. and 17,647-76,191 €/room; properties facing the main streets and squares (Cavour, Rome, Matteotti, Maestranza, Vittorio Veneto, Maniace etc.) range 1,100-3,409 €/sq.m. and 22,222-93,333 €/room; along the secondary streets and alleys (Amalfitania, Alagona, Crocifisso, Dione, Mirabella, Resalibera etc.) properties range 534-2,500 €/sq.m. and 14,194-85,714 €/room (Fig. 1). The wide price ranges reveal the heterogeneity of the real estate shown for each characteristic in the diagrams of the following of Fig. 1.

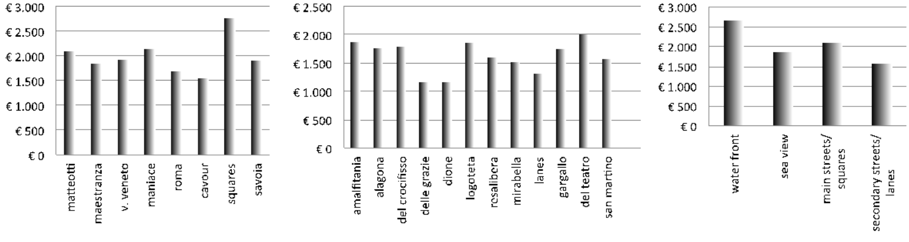


Fig. 1. Price ranges for each urban context

The overall location, landscape, technologic and architectural-environmental features of the properties, are synthesized in an appropriate analysis (Fig. 2).

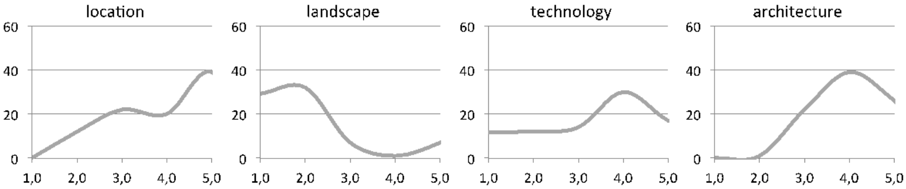


Fig. 2. Quality distribution within the sample

3 Methods and Procedures

3.1 Economic-Appraisal Features: Value and Valuation of Urban Real Estate

The study of this market has been carried out with the purpose of connecting and integrating architectural and economic features in the approach to real estate (re)production and management. In fact, the creation of value [2] in the historic centres has to be handled in order to reallocate the surplus that the property investment achieves as positive externalities due to the renovation public works.

Therefore, the main typical "three benefits" of property – 1. functional-symbolic qualities, 2. productiveness and 3. expectations – have been involved as general target of the complex behaviour of the investors, assuming that each of these benefits is considered by them as a motivation to invest or disinvest. Therefore, the three benefits can be considered located at different degrees of the goals/means scale: at the bottom we find the referential, practical and symbolic features; in the middle the economic-financial ones (costs and revenues of the investment); at the top the monetary ones, capital gains and implicit yield (Fig. 3).

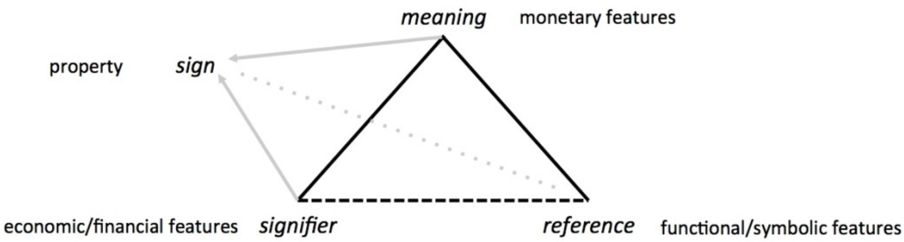


Fig. 3. The semiotic interpretation of a property and its benefits

In the proposed semiotic interpretation, properties, as signs, are the synthesis of value (as meaning) and productiveness (as signifier); the latter is the economic the economic phenomenon of the underlying physical object, the reference, that includes the bundle of functional and practical features of a property. The reference has a very weak relationship with the sign, so that the value of a property isn't strongly connected to its object features. This gap measures the eccentricity of a real estate market, that depends on how the bid-ask spread characterizes the different market segments.

Therefore, our research aims at representing on the one hand the ontology (how much the market is driven by speculation) and the phenomenology of this eccentricity (what it is like), on the other hand the possible segmentation of this property market.

The first question has been addressed by connecting some items of theory of capital with the analytic appraisal procedure, the second one by applying the clustering analysis, a data mining technique aimed at circumscribing the homogeneous submarket highlighting the three types of benefit.

The theory of capital, as addressed and defined by F. Rizzo [17], may be synthesized as follows. The income value V of an asset depends on forecasted income and capital gains: $V = k \pm |a|k$, where V is the income value, k is the supply value meant as the market current value, relevant for most operators, and |a| is a factor of increasing/decreasing in value.

The income approach can be expressed as $V = I_n 1/r$, where I_n is the Net Operating Income, and r is the cap rate. If the cap rate is considered a variable, it is: $r = r' \mp |a|r'$, or $r = r'(1 \mp |a|)$, where r' is the average, or market, cap rate. Therefore, $V = I_n/r'(1 \mp |a|)$.

The differences between the capitalisation value and the market value, and between the effective cap rate and the average cap rate, are the main concern of this theory, about which each asset is considered a singularity, and its value is the result of the combination of market rules and personal determination of the investor within a speculative market.

In fact, one of the main points of a monetary approach to economy is the inversion of the relationship between value and price: although it is generally believed that price depends on value, and the former is the monetary measure of the latter, in a monetary reality, that is the reality of the globalized financial capitalism, the price tensions modify our perception of value, so that the price becomes a sort of general and abstract value [19].

Assuming the market as a "semantic chain" [18], the value (meaning) of the assets (signs of the chain) arises mostly because they communicate by their monetary features – expectations and capital gains (that are goals) – and not by their economic or functional-symbolic characters (that mostly are means). Therefore the assets can be defined as money-goods, because they play the role of money: they measure, preserve and increase the value of the wealth that they (currently) are, and that they can (potentially) become. Each speculative transaction modifies the local property market (the semantic chain) both in prices (meanings) and in characteristics (signifiers), so that the expectations arise in advance of the current prices.

When the monetary perspectives (the capital gain expectations) prevail over the practical/symbolic functions, properties become treasures to be hoarded and the market becomes asymmetric and dynamic, for better and for worst.

The monetary characterization of a speculative asset market can be appraised as the difference between the fair and the expected cap ratio. According to Hicks [11] and

Rizzo [18], differently from Fisher [3] a capital asset is in *crescendo*, c, or in *diminuendo*, d: $c \rightarrow r' > r$; $d \rightarrow r' < r$ where r' is the cap rate of a standard stream of values, and r is the cap rate of the prospective, or expected, one.

Each specific cap rate describes how the related property plays the role of a "treasure". A treasure is hoarded by itself and doesn't provide any real income, but only a psychological one. The difference between the average (standard stream) rate and the specific (perspective stream) rate measures how much a property can be considered a treasure, that depends on the prevailing of the psychological income on the real one: the former can be considered the implicit liquidity, the latter the explicit one [18].

Therefore, we can measure the degree of hoarding of a real estate market and compare different markets from this point of view.

The inverse ratio of r' or r, $1/r'$ or $1/r$, is the average period of the standard or perspective stream, defined by Hicks as "the average length of time for which the various payments are deferred from the present, when the times of deferment are weighted by the discounted values of the payments" [11].

The average period measures the "average life" of the stream. Assuming a property as a particular stream, its average life measures its certainty degree as the number of "income-years" it is able to provide. The more r is low, the longer is the average life and vice-versa. In synthesis, we affirm that the renounce to a part of the probable income (related to the standard stream rate) prolongs the average life of the property whose cap rate is lower.

The main foundations of this theory is the well known Keynesian law of the inducement to investment, in which a difference between the supply price of an asset capital and its expected value calculated by discounting the perspective yields [12] is highlighted. The marginal efficiency e is the rate that equates the expected value to the supply price of the asset capital, so the marginal efficiency can be expressed by means of the expected rate: $e = i(1 \mp |a|)$, where e is the marginal efficiency and i the interest rate. The positive difference between e and i progressively activates successive amounts of investments. Therefore, *mutatis mutandis*, cap rate can be considered a sort of marginal efficiency in the property market.

A further representation of this theory can be addressed by considering the difference between optimistic and pessimistic approach about the same asset. A transaction implies that the bid price is higher than the demand price.

As a consequence, the value (bid price) can be expressed as the result of the double projection of k (*supply price* according to Keynesian address, *demand price* in financial current uses) forward and backward: $V = k(1 + r_l)^n/(1 + r_d)^n$ in which r_l is the expected (perspective) rate of increasing in value, r_s is the observed (at cost) discount ratio, and n is the time of this projection. The result of this projection depends on the characteristics of the two (dis)investors that influence these two rates. In general, $r_l > r_s \rightarrow V > k \rightarrow a > 0$ and viceversa. The relationship between increase in value rate and discount rate rules the result of the "negotiation adventure".

Such theory of capital represents the price tensions phenomena and the consequent bid-ask spread due to different expectations as measured by the two rates. During a bull trend, bullish operators expect capital gains, $V = k(1 + |a|)$, so that purchasers easily buy, owners hardly sell; during a bear trend, bearish operators expect capital losses, $V = k(1 - |a|)$, so that purchasers hardly buy, owners easily sell. The greater the price tension within a particular market, the greater the bid-demand spread.

Appraisal science mainly concerns about the fair value; according to a semiotic approach the fair value can be considered the conventional signification that rules a normal communicative system. When the communication system gets over-communicative, the intentional signification prevails, whereas, when the system becomes under-communicative, the real signification prevails; assuming market as a communicative system, the fair value dominates in a perfect market, whereas in a speculative market, *bid prices* prevail during a bullish trend, *demand prices* prevail during a bearish trend. Within an eccentric market the individual willingness/aversion "drags and drops" the prices and jerks the whole semantic chain by modifying the expectations that influence the cap rates. Several income method valuation tools have been recently provided following the increasing/decreasing logic, both in business [8], [9] and real estate market [21], [22], [15].

The real estate of Ortigia is characterized by the heterogeneity of capitalization ratios; the survey we propose aims at defining the characteristics that mostly influence it in the different submarkets [1]. The segments can be defined from physical, economic, monetary points of view [20]. The properties are substitute units in a physical sense if they have similar characteristics, in an economic sense if they have similar productiveness abilities, in a monetary sense if they have similar rate of capitalization classes [7].

The tool we propose in order to connect the urban, architectural and landscape with monetary characteristics is the clustering analysis referred to both the referential and the semantic characteristics of the properties.

3.2 Clustering Analysis. Theoretical Frame and the Real Estate Market

The statistical study of complex social systems, such as real estate markets, identifies the data mining techniques important applications such as the cluster analysis. The real estate market of Ortigia, because of its heterogeneity and articulation, and the large dimension of the examined dataset, is a quite interesting application context that needs some additional considerations that supplement the mere application of the chosen algorithm.

In order to better understand the theoretical framework of this work, it is convenient to make a brief digression about clustering theory [5].

Recently, it has been the tendency to treat inquiries in the real estate market using statistical techniques of clustering [10], [13].

Generally, the reason that leads to apply the clustering theory to a statistical population is grouping statistical units in subsets, called clusters, the most possible internally homogeneous and externally heterogeneous. This kind of groups can be realized through clustering algorithms.

In the following, with the term partition we will mean a family of subsets of the initial sample such that two of them are disjoint and their union is the entire sample.

A clustering algorithm is a succession of steps through which is made a succession of partitions is made starting from the partition of singletons and uniting at every step only two subsets in the previous partition, until to the partition is formed by the entire sample. Subsets in various partitions are called clusters. The aggregation is realized on the basis of a certain parameter, generally linked to a method and to a metric. Happening the aggregations of clusters, the initial sample becomes more and more compact. Therefore, through a clustering algorithm, a family of partitions called hierarchy is made, having as first the one formed by singletons and as last the one formed by only the entire sample. Moreover the previous partition has the same sets as the successive one, except for the ones aggregated at the previous step in a unique set in the successive one. This means that the previous partition is less fine than the successive one.

The following definitions formalize some of the above exposed concepts.

Def. ***1.*** *A partition P of a set U is a family of subsets U_i in U having the following properties*:

i. $\forall\, i,j\ (i \neq j \rightarrow U_i \cap U_j = \emptyset)$
ii. $\bigcup_i U_i = U$

Def. 2. *A partition P_2 is said less fine than partition P_1, symbolically $P_2 < P_1$ if:*
$\forall\, U \in P_2\ \ U = \bigcup_{i \in I} U_i$
where $\{U_i : i \in I\}$ is a family of sets taken from partition P_1*:* in other words, if every set in partition P_2 is union of sets in partition P_1.

Def. 3. *A hierarchy G is a family of partitions $P_1, P_2, \dots, P_n$, provided with a sorting having the following properties*:

i. P_1 is the partition of singletons
ii. P_n is the partition formed by only U
iii. $i < j \rightarrow P_i < P_j$

Def. 4. *A hierarchical aggregative algorithm is a proceeding such that in output provides a hierarchy, having accepted as input a population constituted by statistical units $u_1, u_2, \dots, u_n$, generally multivariate, i.e. with multidimensional vectors associated. These vectors are obtained from values assigned to variables $v_1, v_2, \dots, v_m$ accepted to describe the dataset.*

All the clustering algorithms have a common property, i.e. at every step, they aggregate among them the sets U_i, U_j minimizing a certain parameter, associated to a method. In the following table, there are the parameters associated to various methods:

Method	**Parameter**
Single linkage method	$min\,(D)$
Complete linkage method	$max(D)$
Average linkage method	$M(D)$
Centroids' method	$d(c_i, c_j)$

Where:

$$D = \left\{ d(ui, uj) \,\middle|\, ui \in U_i, uj \in U_j \right\} \tag{1}$$

$c_{i,} c_j$ are the centroids in the clusters U_i, U_j and d is a metric taken from various available d may be Euclidean metric, Manatthan metric or various other.

The algorithm used in the analysis is associated to the complete linkage method and to the Euclidean metric. So it predicts, at every step, the aggregation of two clusters of the previous partition, on the basis of the least Euclidean distance among the most distant elements.

Dendrogram. Once obtained the final hierarchy, it is possible to use a graphical object, called dendrogram, which allows us to globally visualize it, from the initial partition of singletons to the partition of the entire statistical population. In this graphic, there is an horizontal line where single elements in the sample are located. Climbing, at every level, there are the same sets of the last level with the addition of the one coming from the last aggregation. Therefore a specific partition of the hierarchy corresponds to each level.

R^2 and RMSSTD. After analyzing the dendrogram, the problem is choosing the level at which to cut the dendrogram. In other words, the problem is determining the best partition among the ones in the hierarchy, where in this context it better means the one maximizing heterogeneity between clusters and homogeneity within clusters.

To cut the dendrogram, there are two useful statistical indices that guide the choice of the best partition.

One of these indices is called R^2 and is defined as follows:

$$R^2 = 1 - \frac{W}{T} \quad (2)$$

Where:

$$W = \sum_{j,V} W_{jV} \qquad W_{jV} = \sum_{i \in I_V} (x_{ij} - \bar{x}_{jV})^2 \quad (3)$$

$$I_V = \{i : u_i \in U_V\} \qquad u_i \equiv (x_{i1}, x_{i2}, \dots, x_{im}) \quad (4)$$

Moreover:

$$T = W + B \quad (5)$$

Where:

$$B = \sum_{j,V} B_{jV} \qquad B_{jV} = \sum_V (\overline{x_j} - \overline{x_{jV}})^2 \quad (6)$$

W is known as within total variance while W_{jV} is the within variance of variable j inside cluster U_V. T is called total variance, and it is the sum of W and B, where B is the total between variance. $\overline{x_{jV}}$ is the average value of variable j relative to statistical units in the cluster U_V and $\overline{x_j}$ is the total average for variable j.

From definition of R^2, results that high values of this index indicate good partitions because these values correspond to small values of within variance W in relation to the total variance T.

The second index is known as *RMSSTD*. It is an acronym for *Root Mean Square Standard Deviation* and is defined as follows:

$$RMSSTD = \sqrt{\frac{W_V}{p(n_V - 1)}} \quad (7)$$

In this formula, p represents the number of variables investigated and n_V is the number of clusters in the current step. Moreover, in this formula, there is the within variance $\boldsymbol{W_V}$ relative to the cluster coming from the last aggregation, so it is necessary to know this value to calculate this index. Contrary to the first index, it is better having small values because RMSSTD grows up within variance of the last cluster so that great values indicate that the last cluster is very heterogeneous.

RMSSTD is not defined for trivial partition constituted by only the entire set because $n_V = 1$. Moreover it is not interesting for partition of singletons because it is the first partition so that there isn't a new entry inside it. In the cutting dendrogram, there is a simple rule predicting to stop with aggregations when this index has become much greater than the previous values.

Because it is scientifically meaningless establishing without any reference a good numeric boundary for values of R^2 and $RMSSTD$, to choose the best partition in the hierarchy arising, we will use an initial hypothesis about the maximum number of submarkets in the global market, considering only partitions with a number of clusters, corresponding to submarkets, lesser than a maximum value. This value is linked to the dimension of the global market.

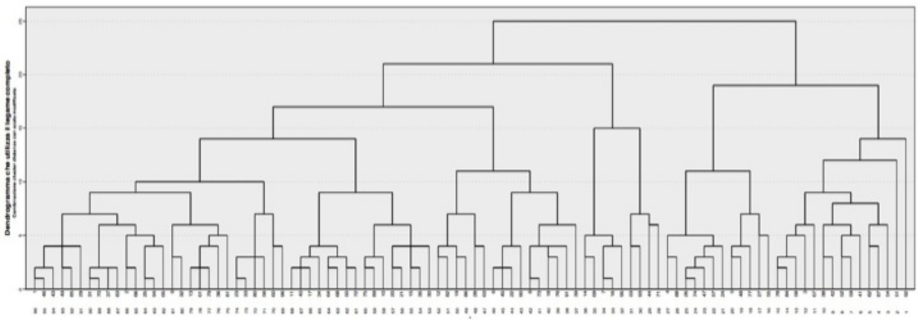


Fig. 4. The dendrogram

4 Applications and Results

The dendrogram (Fig. 4) provides a great number of possible bottom-up aggregations, so that some tests must be done in order to choose the right number of sub-markets. Therefore a spread-sheet model has been drawn up in order to represent the different segmentations. The maximum number of submarkets has been reasonably established on 8. The dendrogram shows the successive breakdowns so that the whole sample can be progressively cut (top-down) into two clusters (cut 2), three clusters (cut 3) and so on until the last cut that provides eight clusters. To do this, for each cut the first and the last identification numbers of the cluster are inserted in a triangular matrix (Fig. 5a) so that the eight segmentations are provided (Fig. 5c). The model calculates the indexes R2 and RMSSTTD (Fig. 5b) in order to choose the correct number of submarkets, in this case 5 because RMSSTD decreases more slowly the after fifth breakdown and no significant real estate differences have been recognized in successive groups. The location of the eight different segmentations (cuts) is shown in Fig. 6.

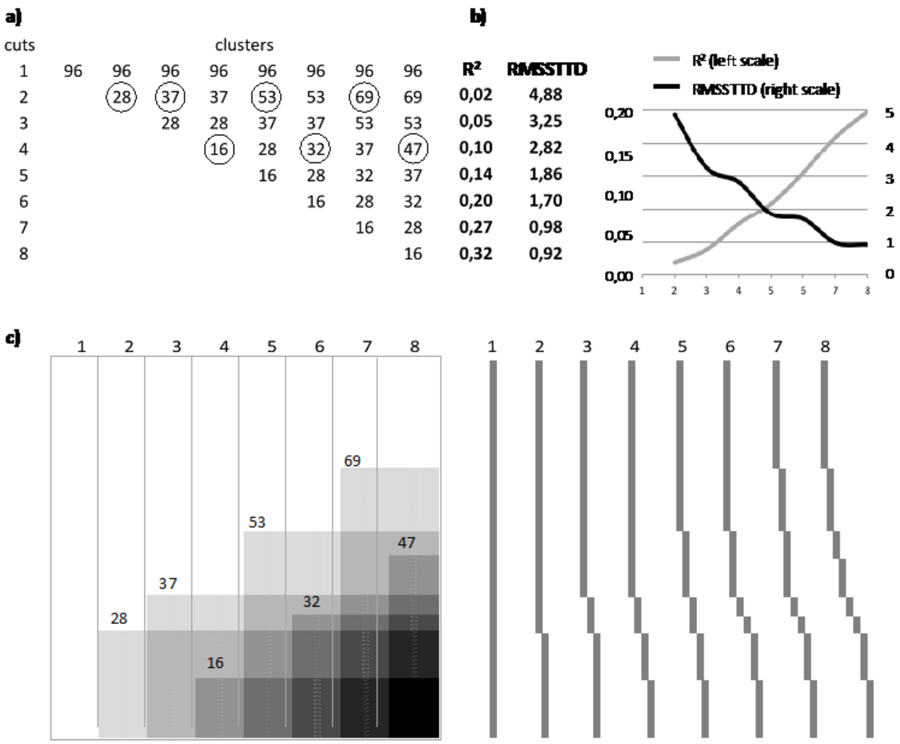


Fig. 5. Clustering application; a) each segmentation hypothesis (1 to 8) is defined by the two id. numbers shown in the column of the triangular matrix; the circle indicates the id. of each new cut; b) for each hypothesis the indexes R^2 and RMSSTTD are calculated: the graph shows their trend; c) the two histograms show the progressive clustering of each hypothesis: first column no segmentation, second column two clusters, the first from id 96 to 28, the second from 27 to 1; second column three clusters, 96-37, 36-28, 28-1, and so on; the scheme on the right differently shows the progressive clustering.

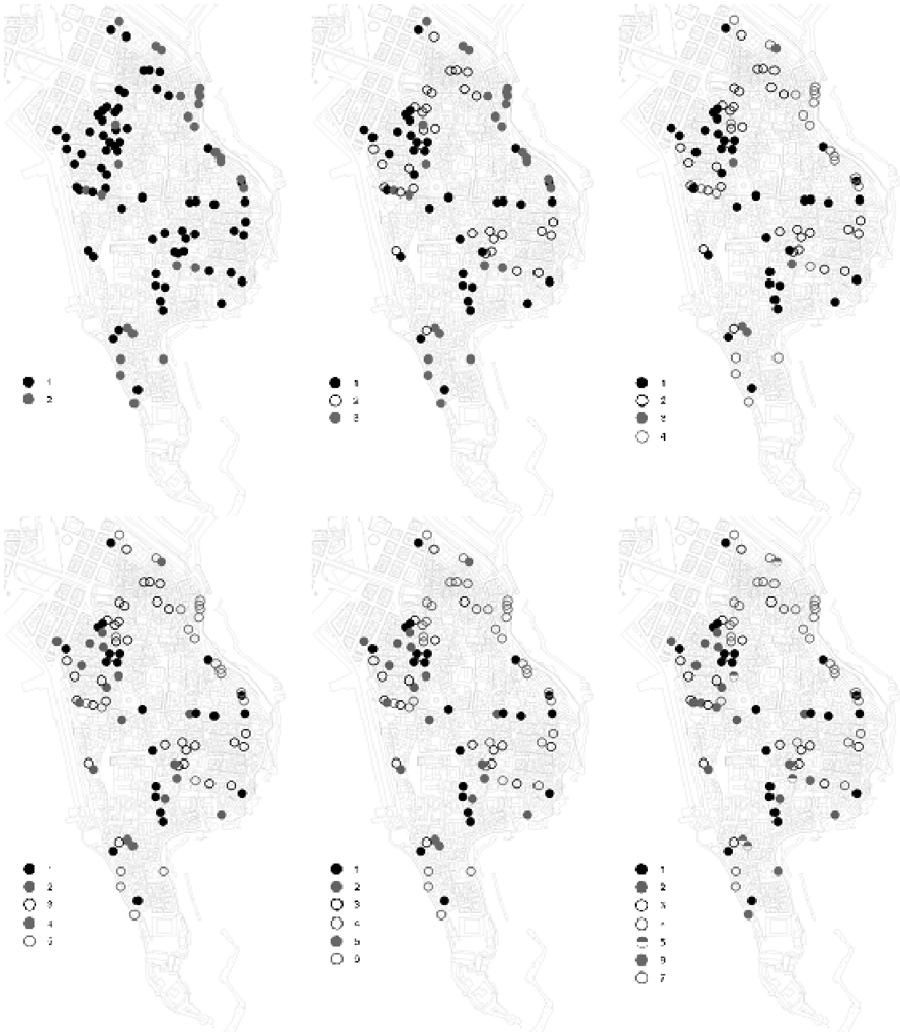


Fig. 6. Location of the different clusters for each cut (eighth cut omitted)

The next verification involves the price ranges within each cluster. This consistency analysis has been carried out by calculating the prices range (Fig. 7) and the logarithmic regression R2 index for each cluster of every cut (Tab. 1). Table one shows that after the fifth breakdown the groups are not significant (R2=1.00) in order to describe a different elasticity price/quality.

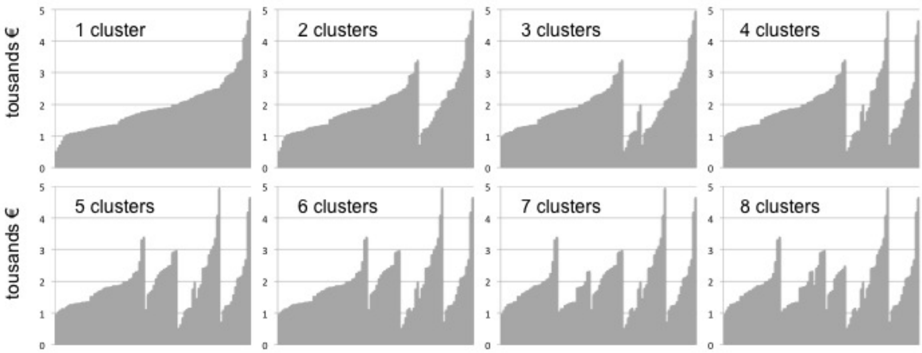


Fig. 7. Clustering price ranges for each cut (segmentation hypothesis)

Table 1. Statistic clustering consistency test: value of R^2 for each cluster of all cuts

	R^2 calculated by using room unit prices								R^2 calculated by using sq.m. unit prices							
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
cut 1	0,51	-	-	-	-	-	-	-	0,49	-	-	-	-	-	-	-
cut 2	0,50	0,48	-	-	-	-	-	-	0,47	0,47	-	-	-	-	-	-
cut 3	0,41	0,91	0,48	-	-	-	-	-	0,32	0,86	0,47	-	-	-	-	-
cut 4	0,41	0,91	0,36	0,56	-	-	-	-	0,32	0,86	0,47	0,46	-	-	-	-
cut 5	0,28	0,81	0,91	0,36	0,56	-	-	-	0,15	0,45	0,86	0,47	0,46	-	-	-
cut 6	0,28	0,81	1,00	1,00	0,36	0,56	-	-	0,15	0,45	1,00	1,00	0,47	0,46	-	-
cut 7	0,18	0,31	0,81	1,00	1,00	0,36	0,56	-	0,09	0,29	0,45	1,00	1,00	0,47	0,46	-
cut 8	0,18	0,31	1,00	0,83	1,00	1,00	0,36	0,56	0,09	0,29	0,70	0,56	1,00	1,00	0,47	0,46

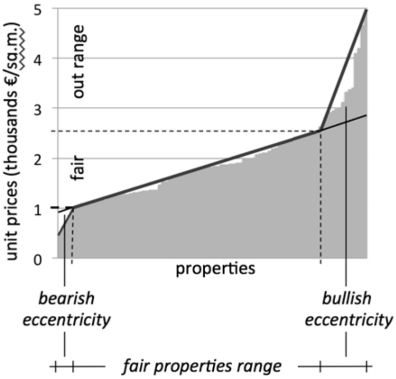
5 Discussions

According to the main concern of this study that involves the consistency of the monetary features, two final verifications have been carried out.

The first concerns the “bullish/bearish eccentricity” of the clusters given by multiplying the percentage of properties outside the “fair range” and the sum of the differences between the out-range and fair prices of the unit prices over and under the fair price level (Fig. 8). The more the clusters are bigger the more the indexes are significant.

The second verification concerns the cap rates. Basing of the rental survey, NOIs (I_n) have been calculated in order to find out the cap rate of each property for sale. The regression analysis carried out for each submarket allows to calculate the gross income (I_g) for each property, from which the managing expenses are deducted. According to the previous clustering, the comparison between the cap rates trend and the total quality index is displayed in Fig. 9 for the first five (the most significant) clustering hypotheses (the arrows indicate the dividing clusters that at each stage).

The clustering analysis provides some rational items for grouping the sample into five clusters. The different stages progressively distinguish ordinary and extraordinary property groups. By referring to the fifth clustering, we show the characteristics of the different groups in Fig. 10.



	1	2	3	4	5	6	7	8
1	*0,89*	*0,37*	*0,00*	*0,02*	*0,02*	*0,02*	*0,36*	*0,36*
	30,67	**3,10**	**7,32**	**7,32**	**3,38**	**3,78**	**7,59**	**1,22**
2		*0,09*	*6,39*	*6,39*	*0,54*	*0,54*	*0,19*	*0,19*
		9,82	**5,81**	**5,81**	**1,52**	**1,52**	**1,65**	**1,65**
3			*0,09*	*3,91*	*6,39*	*0,00*	*0,54*	*1,35*
			9,82	**6,36**	**5,81**	**0,00**	**1,52**	**0,00**
4				*0,09*	*3,91*	*0,00*	*0,00*	*0,74*
				8,63	**6,36**	**0,00**	**0,00**	**0,00**
5					*0,09*	*3,91*	*0,00*	*0,00*
					8,63	**6,53**	**0,00**	**0,00**
6						*0,09*	*0,00*	*0,00*
						8,63	**0,00**	**0,00**
7							*0,34*	*3,91*
							7,09	**6,36**
8								0,09
								8,63

bearish eccentricity
bullish eccentricity

Fig. 8. Bullish/bearish eccentricity

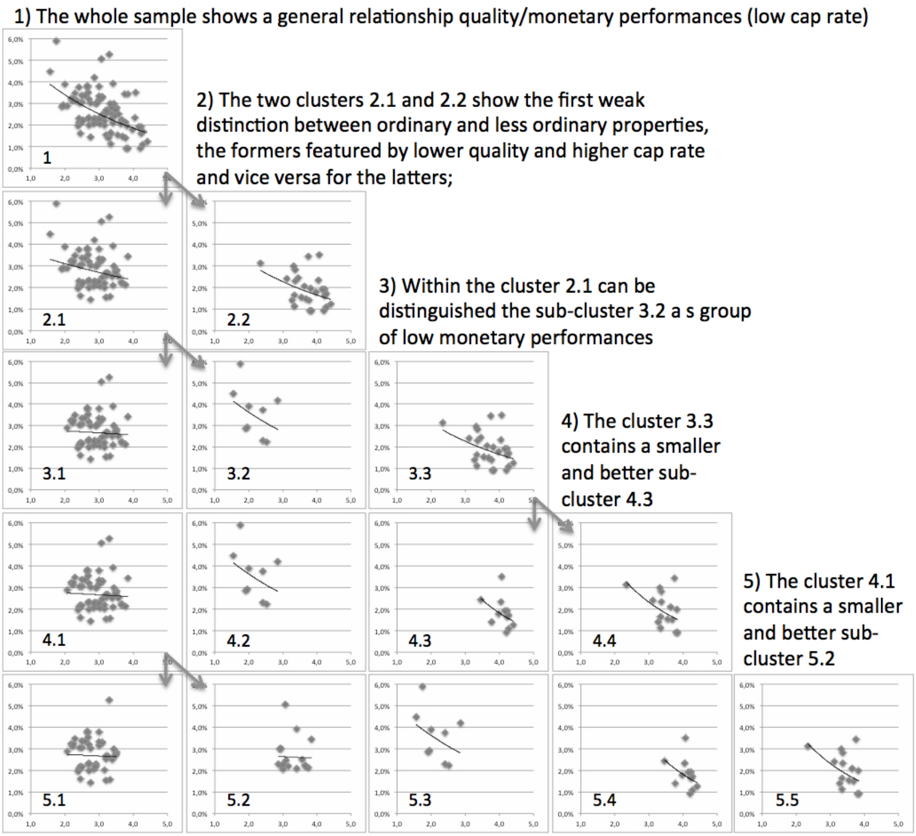


Fig. 9. Trend of the cap rate (y-axis, 0%-6%) compared to the total score index (x-axis, 1-5)

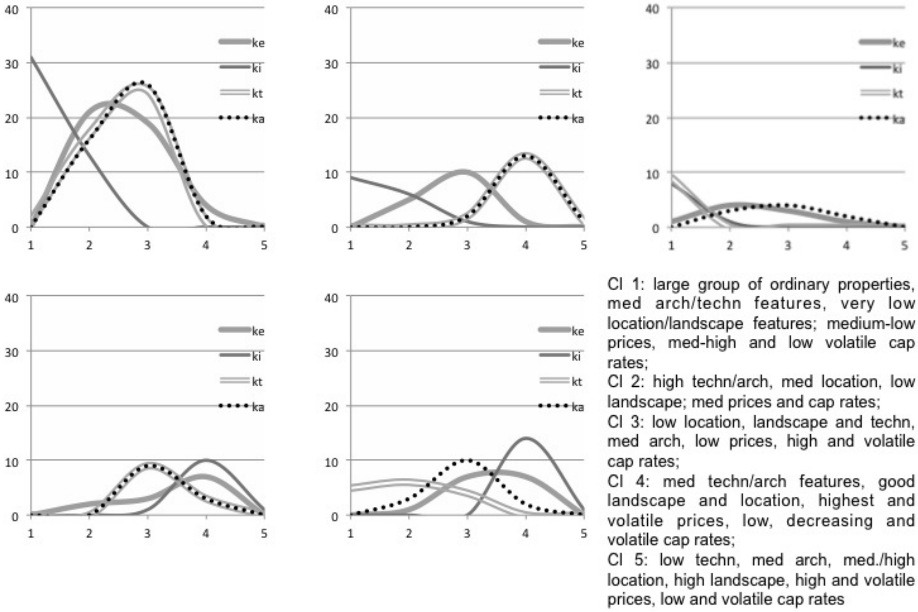


Fig. 10. Description of the different clusters

6 Conclusions

The clustering analysis can be considered an effective method for investigating the monetary characteristics of an atypical, heterogeneous and eccentric real estate market. Basing on a 96 properties sample, physical, symbolic and income features have been highlighted and connected in order to define their relationship with the monetary potential of this real estate market.

The significant eccentricity, that is the difference between fair and out-range prices, changes in each cluster and several clustering hypotheses have been carried out in order to find the best consistency.

The clustering analysis is suitable to complement the income method in mass appraisals for both ordinary estimation, like the cadastral ones, and extraordinary ones, as required within the urban equalisation processes in which the supposed enhancement encourages the expectation about capital gains.

Acknowledgements. Salvatore Giuffrida edited paragraphs 1, 3.1, 4, 5, 6, drew up the valuation model, performed the calculations on the basis of the clustering analysis, edited pictures 3, 5, 7-10 and table 1; Giovanna Ferluga edited paragraphs 2 carried out the whole market survey and edited pictures 1, 2, 6; Alberto Valenti edited paragraph 3.2 performed the cluster analysis and plotted the dendrograms (like Fig. 4) as the bases to the successive analyses.

References

1. Acciani, C., Gramazio, G.: L'Albero di Decisione quale nuovo possibile percorso valutativo. Aestimum 48, 19–38 (2006)
2. Donna, G.: La creazione di valore nella gestione dell'impresa. Carocci, Roma (1999)
3. Fisher, I.: The Theory of Interest. Macmillan, New York (1930)
4. Forte, C.: Elementi di estimo urbano. Etas Kompass, Milano (1968)
5. Fraley, C., Raftery, A.E.: Model-Based Clustering, Discriminant Analysis, and Density Estimation. Journal of the American Statistical Association 97(458) (2002)
6. Giuffrida, S., Martorina, L.: Densità di valore e tensione nei prezzi in un mercato immobiliare eccentrico. Il caso di Ortigia. Valori e Valutazioni 7, 145–162 (2011)
7. Grissom, T.V., Wang, K., Webb, J.R.: The Spatial Equilibrium of Intra Regional Rates of Return and The Implications for Real Estate Portfolio Diversification. Journal of Real Estate Research 7(1), 59–71 (1991), http://ideas.repec.org/s/jre/issued.html
8. Guatri, L.: Trattato sulla valutazione delle aziende. EGEA, Milano (1998)
9. Guatri, L.: Il giudizio integrato di valutazione. Dalle formule al processo valutativo. Università Bocconi Editore, Milano (2000)
10. Hepşen, A., Vatansever, M.: Using Hierarchical Clustering Algorithms for Turkish Residential Market. International Journal of Economics and Finance 4, 138–150 (2012), http://dx.doi.org/10.5539/ijef.v4n1p138
11. Hicks, J.R.: Valore e capitale. UTET, Torino (1959)
12. Keynes, J.M.: Teoria generale della moneta, dell'interesse e dell'occupazione. UTET, Torino (2001)
13. Maitra, R.: Volodymyr, Melnykov, Soumendra, N, L.: Bootstrapping for Significance of Compact Clusters in Multi-dimensional Datasets. Jasa (2012)
14. McDonalds, J., Stokes, H.: Monetary Policy and the Housing Bubble. Journal of Real Estate Finance and Economics (2011), doi: 10.1007/s11146-011-9329-9
15. Morano, N., Manganelli, B., Tajani, F.: La valutazione del rischio nell'analisi finanziaria di Ellwood per la stima indiretta di immobili urbani. Aestimum 55, 19–41 (2009)
16. Nomisma: III Rapporto 2008. Osservatorio sul mercato immobiliare, Bologna (2012)
17. Rizzo, F.: Analisi critica della teoria delle valutazioni. Seminario Economico-dell'Università di Catania, Catania (1977)
18. Rizzo, F.: Valore e Valutazioni. La scienza dell'economia o l'economia della scienza. FrancoAngeli, Milano (1999)
19. Rizzo, F.: Dalla rivoluzione keynesiana alla nuova economia. Dis-equilibrio, tras-informazione e coefficiente di capitalizzazione. FrancoAngeli, Milano (2002)
20. Simonotti, M.: La segmentazione del mercato immobiliare urbano per la stima degli immobili urbani. Atti del XXVIII Incontro di studio Ce.S.E.T., Roma (1998)
21. Simonotti, M.: Problemi di verifica del saggio di capitalizzazione. Estimo e territorio 12, 8–14 (2009)
22. Simonotti, M.: Ricerca del saggio di capitalizzazione nel mercato immobiliare. Aestimum 59, 171–180 (2011)