

# **1. Purpose of the statistical system of “Real Estate Prices,” the publication objective**

Since 1997, the Ministry of Finance and the Czech Statistical Office have been cooperating on the formation of a system for monitoring of real estate prices in the Czech Republic. It was a milestone moment in the development when the Act No 151/1997 Sb, on Valuation of Property, was passed. In Section 33(3) thereof, a duty was put on financial authorities (internal revenue authorities) to transmit data contained in tax returns about prices found out in valuation of real estates and about prices agreed for those real estates in case of sale, to the Ministry of Finance and the Czech Statistical Office. The Act came into force on 1 January 1998.

The purpose of the system is to provide information on the distribution of the price level by real estate type, location, and by other decisive factor as well as on development of this distribution over time. The system should be reliable and up-to-date. The system purpose is not to replace so-called price maps, which local authorities form for their purposes yet to provide global information at the macroeconomic level.

This publication also does not strive to present such a system in full. Besides the providing of information on the real estate prices dependence on various determining factors, it also presents reasons for the selection of such factors in the given data source. One of the objectives is to find the maximum informative power of the current data status. The publication follows another one called “Prices of Types of Real Estate Monitored in 2016 - 2018”.

## **2. Data source, frequency**

Stamp duty land tax (SDLT) statements/returns (real estate tax returns) are the data source here. The statement must be submitted by new owners (purchasers) of real estates to relevant internal revenue authorities (having local jurisdiction) by the end of the third calendar month following the month, in which the ownership right was entered (the property was registered) in the Cadastre of Real Estate (land register). The CZSO shall receive these data encrypted once per month. The data have been collected since February 1999. They are delayed by 7 months, on average, after the date of sale or submission of the SDLT statement/return. The establishment of the database of price information at internal revenue authorities (financial authorities) started by the processing of data from submitted SDLT statements/returns in 1998 (according to the Decree No 279/1997 Sb).

Since 2014, an expert opinion is not required for SDLT statements/returns; therefore, only data from those real estate transfers are processed for which an expert opinion exists.

The main advantage of this administrative source is that it is based on real, actually paid prices (declared in SDLT returns). It is a nationwide and regular data flow on prices of transactions on the real estate market. Prospective objections (that a price declared for taxes may not always be identical with the price actually paid) almost lose their weight when it comes to a relative comparison of prices over time and also (in most of the cases) comparison of the location of real estates, because it can be assumed that this distortion is virtually the same in the given cases.

Internal revenue authorities enter selected data on sales of real estates from SDLT statements/returns into the database and the files created are then transmitted to the CZSO. These are types of real estates as follows: buildings and halls, family houses, recreational cottages and houses, recreational chalets and garden huts, garages, wells, flats (dwellings) and non-dwelling areas, building (p)lots, agricultural land, forest land, other (p)lots and land, and forest stands.

### **3. Variables monitored and derived, homogeneity**

Each (complete) real estate sold is usually formed of partial real estates of various types (e.g. buildings and building (p)lots). Location is reflected in the appraisal (assessed) value as calculated by the valuation surveyor according to the decree on property valuation. The appraisal (assessed) value of the whole, complete real estate is a sum of appraisal (assessed) values of partial real estates. The purchase price of a complete real estate is the recorded price, which the real estate was sold for. The purchase price of a partial real estate is subsequently derived proportionally to the portion of the appraisal (assessed) value of the partial real estate of the total appraisal (assessed) value. The unit price (appraisal/assessed value or purchase price) is the appropriate price for one measuring unit of the corresponding type of real estate (for 1 m<sup>2</sup> or 1 m<sup>3</sup>, etc.).

Starting from 2014, the valuation rules for valuation of partial real estates were slightly modified by the decree on property valuation.

The basic task is to determine the average purchase unit price of a certain type of real estate (also see 11.) in the area or band/zone limited by parameters set (for instance, Region, the extent of wear and tear, etc.). For this purpose, all realised sales of complete real estates in this area or band/zone, which contain a partial real estate of the surveyed type, are searched (found). Rules for the suitable sampling of real estates in the area or band/zone surveyed require homogeneity is defined as a portion (percentage) of the value of a partial real estate of the surveyed type in the whole sale of the given complete real estate.

The selection of real estates for the determination of the average purchase unit price of a certain type of real estate shall be carried out in three steps. First, only those complete real estates are selected, for which purchase price and appraisal (assessed) value are not extremely different. In the second step, only those complete real estates are taken out of this set, for which homogeneity of the type surveyed is higher than a fixed limit for this type of real estate. In the third step, improbable extremes, mostly generated by errors in data entering, are eliminated from the remaining data. Then the arithmetic average of the remaining purchase unit prices of the type of real estate surveyed in the given area is taken as the average purchase unit price of the given type of real estate in the area.

### **4. Explaining and explained variables, basic principles of their determination**

In tables that are the subject matter of this publication, the average purchase unit price is the priority explained variable everywhere. Then as a derivative, it is also the price index, which is the ratio of two such average prices for two different time periods. The subset of data, which are available in the aforementioned data sources, shall be selected as explaining variables. The determining factors shall be clearly territory and time aspect as a priority user's request for the real estate price statistics. Territorial breakdown inside the publication corresponds to the status as at 1 January 2020.

Further explaining determining factors must meet the condition that there is a statistically provable dependence of the explained variable on these factors and, simultaneously, the factors should be mutually independent, as much as possible. Failing to meet the latter condition would mean that adding of another explaining factor would not increase the informative power because it would not add a new dimension to the whole system of dependencies. It would only to some extent (namely to the extent of the dependence of the explaining factors) replicate the already known dependence of the explained variable (i.e. of the purchase unit price) on factors. However, it would reduce the number of price values of the explained variable for every possible combination of the explaining factors and thus it would decrease statistical reliability of the monitored dependence.

The extent of wear and tear (expressed as percentage - data from the assessor) was selected from explaining factors available. The dependence of the average sales unit price on the extent of wear and tear is very conclusive. Another explaining factor applied, which is virtually independent of all of the

previously mentioned, is the population size of the municipality (according to the population as at 1 January 2013), in which the surveyed real estate is located.

## 5. Selection of types of real estates

For this publication, only those real estates available in the source database were selected, for which there is a sufficient number of price data of the explained variable sorted by various combinations of explaining factors determined above. Moreover, the aforementioned limiting effect of homogeneity for certain types of real estates shall be taken into account. For instance, various types of land/(p)lots form mostly minority portions of prices in the sale of a complete real estate. Their prices are then more or less determined by other type of real estate, which they are sold with, e.g. a family house or a building. The applicable number of price data for such (p)lots is then substantially lower. Therefore, solely those types of real estates, where a relatively sufficient number of price data having homogeneity higher than the required limit, were selected for this publication as follows:

Order	Type of real estate	Band of the extent of wear (wear and tear by zone)	Minimum value of homogeneity
1	FH – single-dwelling family houses, of bricks	<0; 10>	60%
		(10; 50)	60%
		(50; 75)	70%
		(75; 100)	75%
2	FL – flats (dwellings)	<0; 100>	70%
3	MB – multi-dwelling buildings	<0; 100>	40%
4	BP – building (p)lots	X	10%

## 6. Frequency and dependence structure of the number of price data by determining factor

The selection of explaining factors for the unit price of a given type of real estate as well as the initial selection of real estate types itself is the outcome of relatively extensive numerical analyses. Some results are therefore presented in brief.

Based on the frequency of data about real estate sales in 2017 and 2018, a peak in the 4<sup>th</sup> quarter can be identified while another one can be traced to the 2<sup>nd</sup> quarter of 2019.

If we leave the above-mentioned seasonal development out of consideration, it is apparent that, out of the data that were available, only those from up until mid-2019 are fully usable. The amount of data for the following months decreases evenly. This is due to the fact that after the sale of real estate, it takes several months before the data are available in the database. The data for the period up until the end of 2018 can thus be considered practically final whereas those for 2019 are preliminary and it will be possible to make them slightly more precise later.

The frequency of family house sales per population size is above average particularly in the *Pardubický* Region, followed by the *Olomoucký* Region and the *Jihomoravský* Region; it is by far the lowest in Prague and low in the *Moravskoslezský* Region and the *Ústecký* Region. The situation is similar when it comes to building (p)lots – the highest relative activity is in the *Zlínský* Region, the *Jihočeský* Region, and the *Pardubický* Region while it is at its lowest in Prague, the *Moravskoslezský* Region, and the *Plzeňský* Region. Within the category of flats (dwellings), the highest number of sales per capita is in the *Karlovarský* Region, the *Ústecký* Region, and the *Plzeňský* Region while the lowest number is in the *Vysočina* Region, the *Královéhradecký* Region, and the *Středočeský* Region.

Statistically more conclusive is the number of real estates sold per capita by municipality population size. The relative activity with regard to family houses decreases with the rise in municipality population size while the trend is roughly opposite for flats (dwellings). Within multi-dwelling buildings, large (over 50 000 population) and mid-sized municipalities (10 000 – 49 999 population) dominate. The sales of building (p)lots are the highest in municipalities up to 1 999 population and the lowest in Prague.

## 7. Interdependence of determining factors

Further analyses studied mutual dependence of the following explaining factors:

- Period (usually the year of sale/purchase of real estate);
- Size category of municipality (4 categories + Prague);
- region (usually a Region, sometimes a District);
- Bands of the extent of wear (wear and tear by zone) (not applicable to building (p)lots, cumulated into 3–4 bands based on the type of real estate or, in a more fine-grained division, into 10% bands).

The results are as follows:

The structure of the number of real estates sold per 10% bands of the extent of wear (wear and tear by zone) is independent of the reference year. The structure of the number of real estates sold per Region is independent of the reference year (minor fluctuations in the case of multi-dwelling buildings). The structure of the number of real estates sold per size category of municipality is independent of the reference year (minor fluctuations in the case of multi-dwelling buildings). It is therefore not a major error to study some relations together for 2017–2019 aggregated data.

As regards the extent of wear and tear of family houses, there is a dependence on the municipality population size: the lower the population size, the higher the average wear and tear. The differences in the extent of wear and tear among Regions are hardly noticeable, with the exception of Prague whose average extent of wear and tear is about 18%. The *Středočeský* Region, which is next in order with its rate of about 27%, is already close to the countrywide average of 34%. It approximately holds that Regions with larger municipalities display lower rates of the wear and tear of family houses.

The overall average extent of wear and tear for flats (dwellings) is 9% and it is relatively even among Regions and different population sizes of municipalities with the highest rate being 13% in Prague.

For multi-dwelling buildings, there is a less conclusive dependence of the rising average extent of wear and tear and the decreasing population size of municipality (the countrywide average being at 60%). On average, this rate is relatively even among Regions in so far as it can be reliably stated considering the smaller amount of data.

The last apparent dependence is the share of larger or smaller municipalities in the analysed Regions. Leaving Prague out of account, the highest share of larger municipalities is in the *Moravskoslezský*

Region and the *Ústecký* Region. The most “rural” are the *Vysočina* Region, the *Pardubický* Region, the *Jihočeský* Region, and the *Plzeňský* Region.

## 8. Determination of bands of the extent of wear (wear and tear by zone)

To be able to classify real estates by wear and tear, bands of the extent of wear (wear and tear by zone) had to be determined, for which statistical dependencies of price on that factor (and in combination with other factors) could be monitored. Besides the breakdown into groups by 10% each, which is, however, too fine if combined with a breakdown by other factor and thus cause the population is too disintegrated, specific wider bands (zones) were determined for each type of real estate. The determination was based on the trend of average unit price of the given real estate type depending on the extent of its wear and tear. The groups were established the way that differences of unit prices inside are relatively low, differences among the groups are high, and none of the groups is too large or too small. This process resulted in the following bands of the extent of wear (wear and tear by zone) (in percentage) used in majority of analyses and tables of the publication:

Type of real estate	Band of the extent of wear (wear and tear by zone)			
	1	2	3	4
FH – family houses	⟨0; 10⟩	(10; 50)	(50; 75)	(75; 100)
FL – flats (dwellings)	⟨0; 5⟩	(5; 20)	(20; 45)	(45; 100)
MB – multi-dwelling buildings	⟨0; 25⟩	(25; 65)	(65; 100)	

For flats (dwellings), bands of the extent of wear (wear and tear by zone) have been preserved in the same form as they were used in the previous years, even though the precise extent of wear and tear of a flat (dwelling) has not been required for the needs of a stamp duty land tax (SDLT) statement/return (real estate tax return) since 2005. Therefore, it is missing in our source data. Only classification into one of four possible groups of technical wear and tear and moral wear (obsolescence) determined by the relevant decree of the Ministry of Finance of the Czech Republic is needed. However, those groups can still be considered as corresponding to the bands of the extent of wear (wear and tear by zone) determined by us.

## 9. The level of dependence of purchase unit prices on individual determining factors

The next step was to determine the relative level of dependence of the average unit price<sup>1</sup> on each of the chosen explaining factors. When studying this dependence, it is necessary to take into account a

<sup>1</sup> Price for 1 m<sup>2</sup> of building (p)lots and flats (dwellings) (floor area of a flat (dwelling)). Price for 1 m<sup>3</sup> of family houses and multi-dwelling buildings (architectural volume).

certain, albeit small, correlation of the explaining factors (see section 7). Examination of the influence of each of the factors, therefore, requires that all the others remain constant.

In the case of **family houses (FH)**, the highest dependence is, logically, on the level of wear and tear. The influence of municipality population size is less pronounced, that of the period is even weaker, and the lowest dependence is on the Region.

As regards **flats (dwellings) (FL)**, the highest dependence is currently on the period followed by the extent of wear and tear, municipality population size, and the Region. As for Regions, location in Prague exerts a dominant influence.

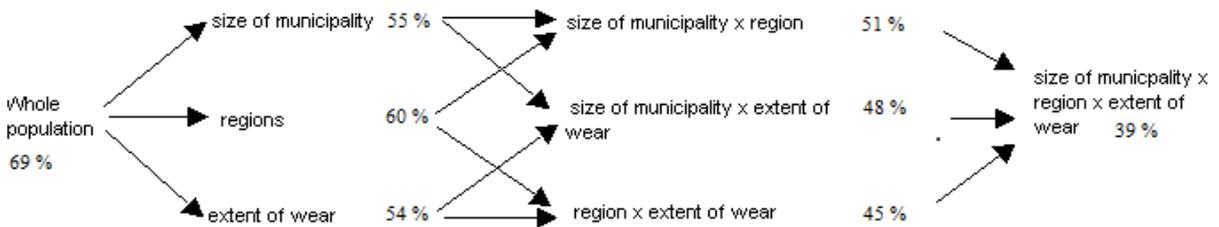
The extent of wear and tear and municipality population size have the highest influence in the case of **multi-dwelling buildings (MB)**, the period has a smaller influence, and the Region has the smallest influence.

In **building (p)lots (BP)**, there is a dominant dependence on municipality population size; in big cities, there is a considerable dependence on the location within a municipality. The dependence on the period is lower and the lowest influence is exerted by the Region. Out of all the real estate types, building (p)lots exhibit the greatest price variability.

### 10. Analysis by means of variation coefficient of groups by determining factor

Now, knowing the respective factors determining the unit price what remains is to determine their optimal combination for the tables in the publication. Subdividing the population by each factor much reduces the number of prices in each of the groups (cross-sections) formed this way. Therefore, effects of all the factors cannot be investigated concurrently to get representative results. Moreover, the four explaining variables established by us cannot by far explain all dependence of the unit price. It results from many other factors unknown to us. If a sufficient number of price data remains in each group (cross-section), formed by subdividing of the population by established factor, we may hope, that these effects, unknown to us, will partially eliminate one another.

Our optimising procedure consists in how fast the average variation coefficient (var. coef.; the ratio of standard deviation and average) of unit prices of the group decreases following each subsequent subdividing of the population by a further (newly added) explaining factor. The procedure is best depicted in the scheme of an analysis carried out for family houses (data for 2014-2016) – see below.



The scheme above shows, for example, that the most effective way is to subdivide the population into four groups by the extent of wear and tear, because it reduces the relative variance of prices from 69% to 54% and the generated groups are a quarter the amount in average. The subdividing procedure must stop in a certain moment when the number of price data is not representative in a certain portion of the groups. The aforementioned analysis was made for all of the 5 selected types of real estates (also garages were included up to 2010), while explaining factors (e.g. the Region and District for a region) were applied in different alternatives. This resulted in an optimum decision, which of the tables can be published.

## 11. Averaging unit prices

Further analyses were not directly related to the determination of published dependence types; they were more concerned with precision and marginal kinds of dependence. One example of this is the decision whether to use weighted or unweighted mean when determining the average unit price in each of the cross categories of determining factors. Weighted mean, where the weights are the sizes of real estates in units of measure, is equal to the so-called average value, i.e. the quotient of the sum of all the prices within a given category divided by the sum of the respective units of measure. Using weighted arithmetic mean could have a distorting effect on the comparison of average unit prices between individual categories if the composition of these categories differs significantly in terms of real estate size in units of measure. The comparison of both averaging methods, however, led to the finding that the resulting difference is relatively minor and can be disregarded.

## 12. Dependence of the other parameters

A condition for more fine-grained analyses is a sufficient amount of data in the population. These analyses were, therefore, mostly carried out only for family houses, flats (dwellings), and building (p)lots.

In particular, it was shown that the low dependence of the unit price (the price of a family house for 1 m<sup>3</sup>) on the size of a family house in m<sup>3</sup> does not in practice have an effect on statistics in this publication. At the same time, this “independence” was shown to hold true also for flats (dwellings) and building (p)lots, by which using of the unit price is justified.

For family houses, relatively precise dependence of the unit price on the extent of wear and tear can be demonstrated. Even though this dependence is not unequivocal and differs, albeit not significantly, according to the location of real estate (combination of the Region and the municipality population size), it can still be claimed that the following relationship holds true primarily for small and medium-sized municipalities. With the rate of wear and tear, the price of real estate decreases relatively quickly (linearly) to about 45% of the original price for the rate of wear and tear of about 30%. Then there is a break on the dependence curve, which goes down (once again in an approximately linear manner) in a much gentler slope to the price of about 11% of the original price for a family house with the rate of wear and tear of about 80%. This dependence allows us to carry out a hypothetical conversion of the price of a family house to a different, not too distant, level of the extent of wear and tear. Analogous dependence can be derived for the coefficient of building equipment.

## 13. Time dependence – price index

Besides the determination of the dependence of the average unit prices of real estates on various explaining factors (or, as the case may be, determination of absolute price levels by Region or locality), it is also important to determine the price development over time in the form of the price index. This index should be of the Laspeyres type with fixed base weights in order to be compatible with other price statistics. Most importantly, it should be adjusted for quality changes over time that means it should be a pure price index.

Because of the character of the price data acquired and due to the amount of the data, the shortest time unit possible is a quarter of a year. Then there is a specific issue occurring for building (p)lots, where there is enough data, yet their variability is by far the highest of all four types of real estates and the surveying of the (p)lot unit price in each quarter of the year keeping approximately the same quality of (p)lots is of extreme difficulty.

In order to meet the aforementioned requirement for a fixed quality of the type of real estates surveyed over time, a qualitative adjustment had to be applied. Because the index is based on the regional breakdown, that means observes the price development for all possible cross-combinations of Region x municipality population size, it is necessary to carry out a qualitative adjustment for differences in other known determining factors (size of a real estate) from one quarter of the year to another. In other words,

the combination of Region x municipality population size is fixed, what is changing is the period (in the quarter of a year interval) and the size of a real estate must be kept constant.

The method applied for family houses and flats (dwellings) uses regression analysis. The regression model assumes the price of the real estate surveyed depends on the period, the size category of a municipality, and the size of a real estate. The whole surveyed set of the real estate surveyed (e.g. family houses, aggregated data for 2015 - 2018) is subdivided into subgroups defined by a cross-combination of the municipality population size and the extent of wear and tear group. Then the aforementioned regression model is determined in every subgroup formed this way. It will provide a statistical dependence of the price on the "size" of the family houses in the subgroup. Then every price of the given real estate in this subgroup can be "converted" to an appropriate "standard" in this subgroup (i.e. one-floor house without a cellar), that means to carry out the required quantitative adjustment prior to the subsequent price comparison.

The model applied for family houses assumes a linear dependence of the logarithm of the unit price on the period, the size category of a municipality, and the "size" of family house (variables have discrete values). For flats (dwellings), the logarithm of the unit price depends on the period and on the fact whether the respective flat (dwelling) is in a bricked or a panel multi-dwelling building (variables have discrete values), and on its size (variables have continuous values). A quadratic form or any other non-linear form of explaining variables was excluded for statistical reasons. The necessary calculations were performed for multi-dwelling buildings yet it has come out that a better way to use prices is to keep them not adjusted by the aforementioned method, because regression models are not statistically provable and adjusted prices demonstrate even higher variability than the non-adjusted ones.

A special approach is used for building (p)lots because analogical explaining variables are not available. The main factor (and in fact the only one known) is the location. Therefore, the whole set of building (p)lots was subdivided into respective cadastral districts. Based on the data for 2017 – 2019, a "reference price" was determined for each cadastral district. Then the ratio of these prices for two different cadastral districts determines the "qualitative relation" of these two cadastral districts. Based on that relation, it was possible to make the qualitative adjustment. The method is relatively effective; variability of "adjusted prices" has been substantially reduced.

The year 2010 (whole-year period) was chosen as the price and index base period to construct the index. In 2013, real estate price indices were reviewed and new, more realistic weightings were created for all types of real estates monitored. Weights for the building (p)lot index and the multi-dwelling building index are determined on the basis of relative ratios of sums of absolute appraisal (assessed) values of real estates falling into the given categories for 2009 – 2011 (in order to get more robust weights). Weights for the aggregate index have been determined the same way. Weights for flats (dwellings) and family houses are solved as a combination of outputs from sums of absolute appraisal (assessed) values and data from the Population and Housing Census on the numbers of family houses and flats (dwellings) (these data are not available for other types of real estates). The weights for flats (dwellings) and family houses for Regions are taken as the relative shares of the volumes of real transactions.

## **14. Relationships of the appraisal (assessed) value and the purchase price**

For family houses and building (p)lots, dependence between the ratio of the purchase price and the appraisal (assessed) value on the one hand and some characteristics on the other was determined. The most conclusive dependence is between this ratio on the one hand and the absolute appraisal (assessed) value on the other hand (there is a relatively strong decreasing dependence (negative correlation) for lower appraisal (assessed) values while for higher appraisal (assessed) values, the ratio with a lower amplitude oscillates around 1). The dependence between the ratio and the extent of wear and tear is less conclusive. The extent of wear and tear and the size of a family house, however, correlate with the absolute price/purchase unit price/appraisal (assessed) value. Overall, for a family house with the highest price, it can be said that its purchase price on average almost equals its appraisal (assessed) value and that the lower the price of a family house is, the increasingly higher its purchase

price is compared to its appraisal (assessed) value. See also publication Table 1.5a, which presents this dependence as a function of the extent of wear and tear. There is a similar dependence, although statistically less conclusive, in building (p)lots.

Further, the distribution of the frequency of the ratio of the purchase price and the appraisal (assessed) value was examined for family houses, flats (dwellings), and building (p)lots. The peak of the frequency distribution lies at a point where the ratio of the purchase price and the appraisal (assessed) value equals 1 or in its immediate vicinity. Despite this similarity, a certain difference between the individual real estate types can be identified from this point of view.

Family houses can be considered as that type of real estate where purely market conditions are in force. For this reason, the corresponding division of the frequency (frequency distribution) of values of the ratio of the purchase price and the appraisal (assessed) value has only one sharp peak (in the above-mentioned value where the purchase price/the appraisal (assessed) value = 1). The frequency of sales where the purchase price is lower than the appraisal (assessed) value decreases very quickly; very few family houses were sold for less than 50% of the appraisal (assessed) value. There is a similar decrease in the frequency of cases when the purchase price is higher than the appraisal (assessed) value, but it is slower than in the opposite relation, particularly with the decrease in the municipality population size where it is not so rare to find that the purchase price is considerably higher than the appraisal (assessed) value.

Analogously, the division in flats (dwellings) has one sharp peak at a point when the ratio of the purchase price/appraisal (assessed) value equals 1, which is as narrow as in the case of family houses. However, the distribution varies considerably with different municipality population sizes. With the increase in the municipality population size, the share of sales that exceed the appraisal (assessed) value increases, too, which cannot be attributed to the sale of flats (dwellings) for, to a certain extent, “regulated” prices of originally state-owned (municipal) flats, because these were not part of the input data. There is a similar dependence between the ratio of the purchase price and the appraisal (assessed) value on the one hand and the extent of wear and tear on the other, even if it is not as conclusive as in the case of family houses.

The frequency distribution of this ratio for building (p)lots is more similar to family houses, i.e. one sharp peak at a point where the value of the ratio of the purchase price and the appraisal (assessed) value equals 1 (lower than in family houses). There are still relatively frequent sales for a price high above the appraisal (assessed) value (the relative frequency in this case is higher than in family houses). The larger the municipality population size, the (roughly) higher the frequency of sales below the appraisal (assessed) value (the maximum is for municipalities with a population above 50 000 population; there is a slight decrease for Prague).

## **15. Determination of the set of publication tables**

The analyses resulted in the determination of the extent of publication tables. One of the aims of this publication is also to establish limits of the informative power of the data investigated (from a perspective set in advance), i.e. determination of the real estate unit price by certain chosen factors. The maximum possible population is represented by the set of family houses, which is relatively large, relations are provable, and it is relatively homogeneous in terms of price. For other types of real estates, only a specific subset could be selected because the aforementioned characteristics are absent. Building (p)lots are a special case because the factor of the extent of wear and tear is missing there, while it is usually the most significant one for other types of real estates. Therefore, there are partially special modifications of otherwise standard types of tables given for building (p)lots.

The table called “Overview of tables in the publication” provides a complete overview. In its left section, one can find the table type marked by number, in the middle section there are determining factors given in tables, and the right section gives the number of the page, on which one can find the corresponding table for every type of real estate, if it exists.

Numbering of tables is coherent in the following sense (“?” stands here for the first digit, which denotes the type of the real estate to which the table pertains; the second digit signifies the type of the published table):

?-1. the factors are: Region and municipality population size – basic information – all types of real estates;

?-2. the factors are: Region and wear and tear;

?-3. a combination of factors from tables 1 and 2 – a maximum combination of factors – it is possible only for aggregated years 2017-2019;

?-4. more detailed regional breakdown by District (for breakdown of Prague by area see the table added to the end of the publication);

?-5. more detailed breakdown by the extent of wear and tear (wear and tear zone);

?-6. price index;

?-7. aggregate price index.

## **16. Informative power of tables, interpretation**

Each table includes the average purchase unit price of the given type of real estate for the breakdown by established factor. These are always simple arithmetic averages. The number of transfers, which the price was obtained from, is given as information on reliability and representativeness, whenever possible. In case there are less than three pieces of data, the price is not given yet the numbers are. Furthermore, the variation coefficient (var. coef.) of the set, from which the average price was calculated, is given as a measure of data accuracy and level of homogeneity of the given group, whenever possible. In some cases, quantiles are given as well. These are important namely for building (p)lots of high price variability. Average values of factors, which in the given table are not applied as determining ones (as the extent of wear and tear, for instance), or, as the case may be, other quantities (as the size of real estate given in appropriate measuring units, for example) are given in some cases, too. In every case, these are simple arithmetic means of the group.

Resulting values are to be interpreted with regards to all possible circumstances. It always has to be deduced how representative and precise a certain piece of data is. Due to a very detailed specification, the presented average unit price may be burdened by specifics of given sales of real estates that are unknown to us. On the contrary, general context and inherent laws are highly conclusive in tables and are relatively reliable.

## **17. Evaluation of the applicability and completeness of data**

Classification based on certain factors has exposed uneven data coverage from different points of view. One of the aims of this publication is to detect such unevenness, because the above-mentioned statistical system for real estate monitoring should continue to develop. Absence of data may be caused by the number of actual transactions being low for certain conditions set or it could be a question of uneven distribution of resources when capturing data. That is why no detailed conclusions on the distribution of real estate sales can be formulated yet, particularly from the regional perspective.