Appendix M

with accepted appraisal theory, an advantage to AVMs is the objectivity and efficiency of the resulting value estimates. Of course, sound judgment is required in model development and an appraiser should review the values produced by the model.

2.2 Purpose and Use of AVMs

2.2.1 General

AVMs are used to provide estimates of market value for a variety of public and private sector purposes. AVM estimates reflect a given time period and should be calibrated to produce market values as of a specific date. Although past market trends can be projected over a short time horizon, the credibility of appraisal estimates increasingly suffers as the projection is lengthened.

AVMs have the advantage of objectivity and consistency, reduced cost, and faster delivery time. It is important, however, that the AVM follow sound statistical and mathematical modeling practices and be tested for accuracy and uniformity before application. Section 8 discusses the important area of model testing and quality assurance and section 9 focuses on reporting of results.

2.2.2 Analysis of Impaired Properties

Properties subject to significant defects or that are affected by atypical circumstances impairing market value, including superadequacy or functional obsolescence, cannot be accurately modeled with an AVM. An appraiser may choose to apply the AVM to the property, but the defect or unique circumstance should be noted and a special adjustment made to compensate for the defect or special circumstance.

2.3 Steps in AVM Development and Application

The remaining portion of this section outlines the steps to take in development of an AVM. The following sections of this standard provide clarification and details concerning these steps and their application to particular property types.

2.3.1 Property Identification

The first step in any appraisal problem is to identify the property to be appraised. In developed economies, identification is normally straightforward, as maps, ownership records, property addresses, and legal descriptions will identify the property and owner. The appraisal assignment will usually require identifying physical characteristics and property rights to be valued as of the appraisal date. When applying an AVM to a particular property, improvements and renovations made before this date should be included in the appraisal; those made subsequent to the appraisal date should not.

The bundle of rights to be appraised generally includes the fee simple interest or full bundle of rights inherent in ownership of property. Nevertheless, the market analyst should make clear what rights are assumed and any limitations to full use or restrictions to transfer of the property.

2.3.2 Assumptions

The AVM supporting documentation should state all assumptions, special limiting conditions, extraordinary assumptions, and hypothetical conditions. A key assumption in many AVM applications concerns the assumed use of the property. Most real estate databases contain the actual use of property as of the inspection date. In some property tax systems, current use is stipulated as the basis for valuation. However, comparable market sales reflect the concept of highest and best (most probable) use. Market analysts and users of AVMs need to be aware of these subtleties.

Another key assumption relates to whether or not the fee simple bundle of rights is being appraised. This is generally the case for residential properties, but many commercial appraisals are made to estimate only the leased fee or leasehold interest when there is an existing lease (or leases) on the property.

Government appraisal agencies are responsible for collecting and maintaining property databases, although they often contract with private vendors for this purpose. Commercial AVM providers generally use data maintained by a government agency or third party service. In all cases, it is imperative that AVM market analysts test the reliability of the data and clearly state assumptions concerning its accuracy. If data important to value estimation are missing or the statistical process has shown the data to be inconsistent or unreliable, the AVM provider has a responsibility to not provide a potentially misleading value estimate to the intended user.

2.3.3 Data Management and Quality Analysis

The reliability of any appraisal depends on accurate data. Appraisal data fall into two general categories: property data and market data. Property data relate to location, land characteristics, and building features. Market data include sales, income, and cost information. Asking prices and independent appraisals can sometimes be used to supplement sparse sales data.

Computerized statistical tools used to develop AVMs afford the opportunity to screen data for missing or out-of-range occurrences and inconsistencies; examples include homes with more than two fireplaces or a bi-level home with no listed lower level living area.

Geographic information systems (GIS) can also help in data reviews. GIS software is used to maintain computerized maps and provide geographic representations of property attributes and features. It can be used to highlight properties with impossible, unlikely, or inconsistent data. For example, properties coded as being waterfront can be color-coded, displayed on a map, and reviewed for accuracy.

Only valid, open market sale and income data should be used in model development. (As mentioned, asking prices and independent appraisals can sometimes also be used to bolster sample sizes.)

Since the reliability of an AVM is dependent on the data from which it is generated, the integrity of the database should be monitored on a systematic and ongoing basis.

2.3.4 Model Specification

Model specification is the important process of determining the format (model structure) of the AVM. The market analyst must determine the type of model to be employed and specify the variables to be used in the model.

AVMs that employ property features, often characterized as "hedonic" models, can be categorized as additive, multiplicative, or hybrid models (see Section 3 on Specification of AVM Models). Market analysts must also determine the variables to be included in hedonic AVMs. These can represent property characteristics (e.g., square feet of living area and building age), location information, demographic data (e.g., income levels or school quality), or variables derived from property characteristics (e.g., the square root of lot size or living area multiplied by a quality index). The objective is always to include property features important in value determination and to capture actual market relationships. Skilled analysis is required to adequately specify an effective model structure.

Some models that are referred to as AVMs have only a time component; in other words, they merely track changes in property values over time. Where property characteristic information is unavailable or limited, these models can be used to trend a previous sale or value estimate to the target appraisal date.

2.3.5 Model Calibration

Calibration is the process of determining the coefficients in an AVM as well as which variables should be retained or deleted due to statistical insignificance. Several statistical tools can be used to calibrate AVM models (see Section 4 on Calibration Techniques). Proper use of these tools requires experience and training in statistical analysis and the software employed.

2.3.6 Model Testing and Quality Assurance

An AVM must be tested to ensure that it meets required accuracy standards before being deployed. This is accomplished through statistical diagnostics and a ratio study in which value estimates (e.g., estimated sale price or estimated rent) are compared to actual values (e.g., sale price or reported rent) for the same properties. GIS can be used to display color-coded ratios on maps and help spot groups of under- or over-valued properties. For more information, see Section 8 on Automated Valuation Model Testing and Quality Assurance. Before it is implemented, the AVM also should be tested on a holdout sample, which is a set of properties and their selling prices that were not used in the calibration process.

Properties with unusually large errors, termed "outliers," should be reviewed. It is likely that the sale price (or other value serving as the dependent variable in the model) is not representative, the data are partially incorrect, or the property exhibits atypical features that cannot be adequately accounted for in the model. Except where the data can be corrected, the property should be removed from the sample, and it and similar properties with similar features should not be valued by the AVM alone.

2.3.7 Model Application and Value Review

Once tested and validated, the AVM can be applied to estimate the value of other properties of the same type in the area or region where the model applies. These values should be reviewed for reasonableness and consistency with recent sales, either of the subject property itself or of similar properties in the same neighborhood or surrounding area, or where sales are not available, recent asking prices.

It is also good practice to systematically review the generated values for reasonableness and consistency with nearby properties in the same neighborhood. This affords the opportunity to ensure that the data are accurate, and to make individual adjustments to properties with unique features or that are subject to special influences, such as being located at a busy intersection or having a premium or obstructed view.

2.3.8 Stratification

Stratification is the process of grouping properties for modeling and analysis. Stratification begins with property type. Properties are delineated into generic use categories such as: single-family residential, condominium (if applicable), multi-family, commercial, and industrial. The number of property types will depend on the size and diversity of the geographic area being analyzed and the number of sales available within the proposed strata.

Residential properties in urban areas are generally stratified into "market areas." Market areas are broad, somewhat homogeneous geoeconomic areas that appeal to buyers in similar economic brackets. One AVM may be developed for each market area, or a regional model may be developed and individually calibrated for each market area. Location within the market area can be handled through neighborhood variables or other variables related to geographic location and desirability. Alternatively, a location value response surface analysis

The market analyst must be able to present the MRA results in an understandable and defensible format that appraisers and AVM clients can easily understand.

To avoid seriously violating assumption of linearity, additivity, and constant variance of the error term, the market analyst must consider the use of transforming variables or other calibration methods described in the standard. A multiplicative, nonlinear, or hybrid model structure is best for measuring interactive effects.

4.1.2 Diagnostic Measures of Goodness-of-Fit

Both the market analyst using regression and the user of AVM output must be aware of and understand how the various key statistical measures used in regression relate to the reliability of results. These statistics fall into two categories: overall measures that aid in the interpretation of model performance and individual variable measures that assist in the understanding of how well an individual variable performs in helping to estimate value, as well as keeping the standard error term to a minimum. Primary measures of goodness-of-fit for overall model performance are the coefficient of determination (R^2), standard error of the estimate (SEE), COV, and average percent error.

Goodness-of-fit measures for individual variables in a model are produced by most MRA software packages and include the coefficient of correlation (R), T-statistic, F-statistic, and beta coefficients. Each of these measures will provide information about an individual variable's linearity or importance of contribution toward improving predictive success, and relative importance, as variables are compared to each other.

(D'Agostino and Stephens 1986.)

When all the measures are used collectively, along with an understanding of data quality issues, those skilled in developing and using MRA can fully evaluate the credibility of the AVM estimates. Appraisers asked to review AVM results must understand the role that *goodness-of-fit* statistics play in evaluating AVM results. The application of AVM results to a single property may be better evaluated using historical market comparisons selected from a subset of data. Appraisers asked to review AVM results should review the Appraisal Standards Board's *USPAP* Standard and AO-18.

(Appraisal Foundation 2003, 46–56, 180–187; IAAO 1990; D'Agostino and Stephens 1986.)

4.1.3 MRA Software, Options and Techniques

MRA is the most widely used method for calibrating models. As such, the availability of MRA software provides users many choices. No one software package is deemed superior to another, as success using MRA is a combination of modeling skills and software familiarity. Variations of a selected MRA technique can be a decisive factor in selecting an MRA and statistical application package. Many MRA techniques have been adopted over the years to help regression take better advantage of its predictive powers. Stepwise, constrained, robust, ridge regression, and others are acceptable techniques used to improve predictive success. Many of the statistical software packages include variable selection routines that aid the market analyst in selection of significant variables.

4.1.4 MRA Strengths

1. Goodness-of-fit statistics—gives credence to the validity of results.
2. Software availability—many regression software products are available.
3. Widely-accepted calibration method.
4. Broad education network—MRA is taught at most colleges and universities around the world.
5. Credible values—in the hands of a skilled market analyst, MRA is proven to produce results that meet the test of model performance.

4.1.5 MRA Weaknesses

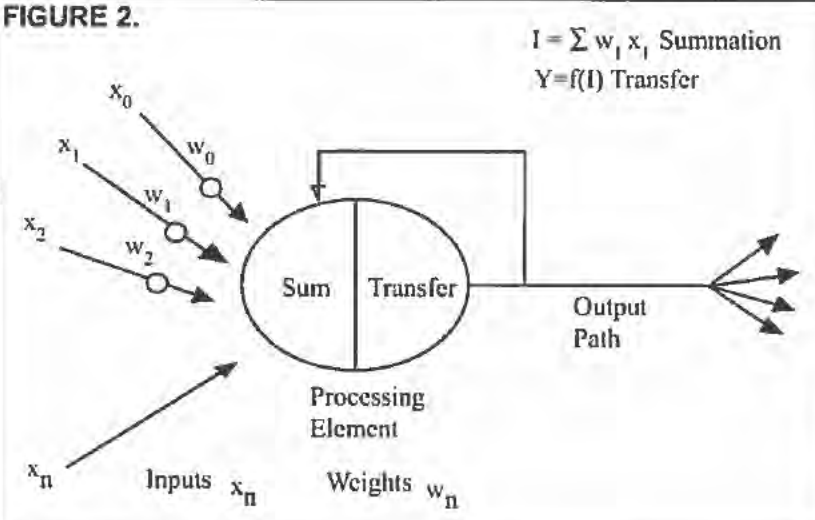
1. Requires a high level of statistical knowledge—market analysts must possess significant background in data analysis and statistical methods.
2. Predictive accuracy is restrained by assumptions.
3. Requires data sets that meet the test of sample size.
4. Interactive and nonlinear market trends are difficult to measure without transforming data.

4.2 Calibrating Using Adaptive Estimation Procedure (AEP)

Adaptive Estimation Procedure (AEP) is a calibration technique that was adapted to real estate value in the early 1980s. Also known as feedback, AEP is based on an engineering concept that relies on continual adjustment to coefficients as the calibration engine passes, or tracks, back and forth through the data until convergence, (minimum error is achieved) thus the feedback. For property valuation, the algorithm tracks the sale price as a moving target. It compares property characteristics as variables that measure the change in sale price, and calibrates a coefficient for each variable. The coefficients are used to estimate value that is then compared to sale price. A running tally is kept on the error term as the process continues. Figure 1 depicts the feedback loop.

AEP will make multiple passes through the sales file constantly adjusting coefficients before a final solution is reached. Success using AEP is dependent upon the

FIGURE 2.



4.4 Time Series Analysis

Time series analyses are a family of techniques that can be used to measure the cyclical movements, random variations, seasonal variations, and secular trends observed over a period of time. In property valuation, these analyses can be used to develop a multiplier or index factor to update existing appraised values or to adjust sales prices for individual properties to the valuation date. Since values can change at different rates in different markets, separate factors should be tested for each property type and market area.

Four methods used to develop time trend factors in the appraisal and assessment industries are: (1) value per-unit analysis, (2) re-sales analysis, (3) sales/assessment ratio trend analysis, and (4) inclusion of time variables in sales comparison models. These methods are summarized below (for a more detailed explanation and discussion, see *Mass Appraisal of Real Property* (Gloudemans 1999, 263-270).

Value per-unit analyses track changes in sale price per unit (e.g., per square foot for residential properties or per unit for apartments) over time. The method is easily understood and lends itself well to graphical representation, as well as to statistical modeling to extract the average rate of change. A downside is that the method does not account for the myriad of other value influences, such as age and construction quality, that impact per-unit values.

Re-sales analysis uses repeat sales occurring over a given time period. Price changes between sales are converted to monthly rates and an average (or median) rate of change is extracted. As can be imagined, the larger the number of repeat sales, the more reliable the estimated rate of change. The method can overestimate rates of change if repeat sales reflect substantial improvements (or other alterations) made to the property since the first sale.

Sales/assessment ratio trend analysis involves tracking changes in the ratio of sales prices to existing assessments made as of a common base date. Increases in the ratios indicate inflation and vice versa. The ratio also provides the index factor required to convert assessed value to a full value estimate. Like value per-unit analysis, the method lends itself well to graphical and statistical analysis. An advantage of the method is that assessments account for most value determinants and thus can isolate time trends better than the value per-unit method. The method assumes that the assessments share a common basis, and its reliability depends partly on the accuracy or uniformity of the assessments.

Time variables can be included directly into AVM models to capture the rate of price change over the period of analysis. This is usually the most accurate of the various methods. However, model developers must be careful that time variables are properly specified so that coefficients developed from the model reflect the desired valuation date.

Once a time trend is established, it can be used to adjust values to any point within the sales period.

Trend factors can be extrapolated for a short period beyond the sales period, but this must be done with caution and grows increasingly unreliable as the time frame is lengthened. If more than several months are involved, the first three methods can be used to calibrate the trend (one would not ordinarily develop time adjustments through use of a modeling approach without recalibrating the entire AVM model).

(The Appraisal Institute 2002, 291.)

4.5 Tax Assessed Value Model

Tax assessed value models derive an estimate of value by examining values attributed to properties by the local taxing authorities. As a matter of local law and custom, known as townhouses, row houses or zero-lot-lines, depending on geographic location throughout the world. All of these uses are residential in nature.

Valuing these various residential properties is somewhat similar to valuing detached single-family structures. All of the same principles apply and all can be modeled and valued using an AVM. In fact, because these properties exhibit a high degree of homogeneity compared to the detached single-family population, sales-based AVMs can produce values that are extremely reliable and accurate. The cost approach can also work well, in some cases, if adjusted to the market, but it is not appropriate for valuating condominium units because depreciated replacement cost will not properly reflect resale values. Data requirements for attached residences will not be the same as with detached residential properties. For example, floor level can be an important value determinant for condominiums, while lot size and yard improvements are irrelevant.

5.3 Two- to Four-Family Residential Property

Part of the residential housing market consists of structures built for the purpose of housing more than one family. Improvements designed to accommodate two, three, and four families within their own separate living areas are often referred to as small income-producing properties. A common theme among these property types is that the owner of the property may reside in one of the units. This concept, however, is not a requirement for classifying these structures in the market. Two-unit properties are more likely to be owner-occupied than four-unit properties. The concept to be recognized here is how such properties are treated in the marketplace, because that impacts their price and ultimately the value generated by any AVM. The ability to model the selling price of these small-income properties is reliant on what specific data is available, relating to number of units, age, condition, location and gross income. The motivation of buyers shifts when consideration is given to other property attributes that relate to producing rental income and not just owner occupancy. Direct market models, comparable sales models, and cost models are acceptable methods for valuing these small income-producing properties. With their income-producing potential, the income approach is also a model to be considered. With an adequate sample of gross income values for comparison to sale price, a model of GI * GIM will yield credible results where GI = Gross Income and GIM = Gross Income Multiplier (sale price/gross income). Some AVMs may even be set up to predict GI and the GIM. Each of these indicators can vary with size, age, location, style, and condition of a property.

5.4 Manufactured Housing

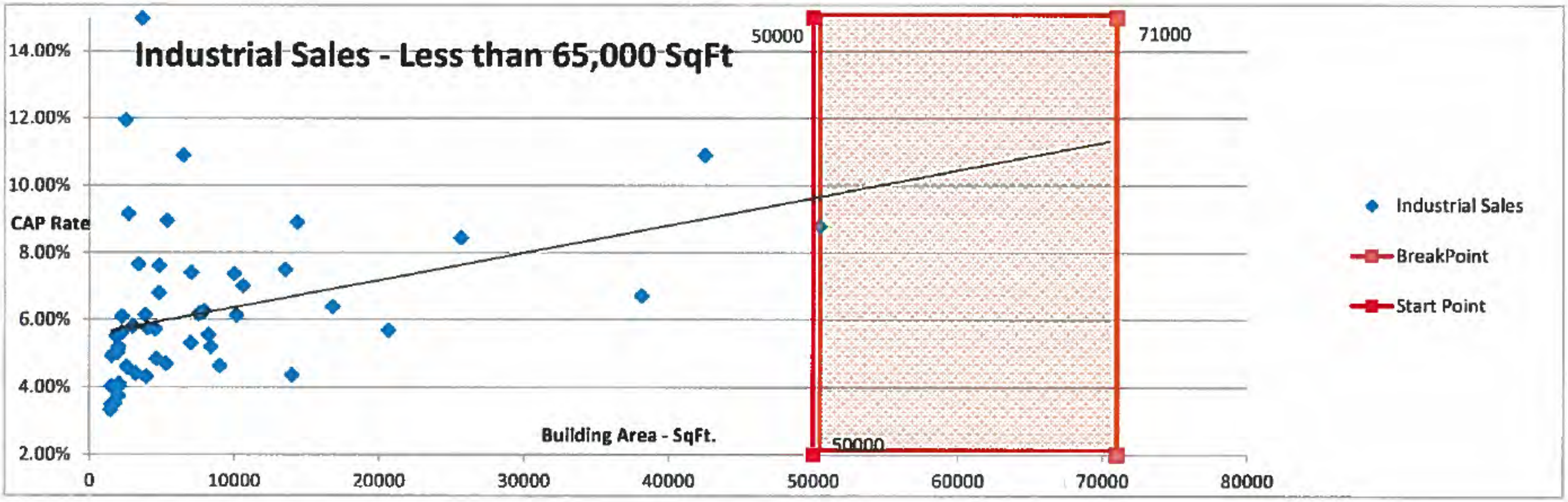
A manufactured home is a residential structure built in a factory. Construction standards for manufactured housing are controlled and monitored by the Department of Housing and Urban Development in the United States (HUD), and by the Canada Mortgage and Housing Corporation (CMHC) in Canada. While many manufactured homes are built with the same materials as site-built homes, the factory-controlled engineering process helps control cost and quality. The house can be financed as personal or real property on leased land, in a manufactured home community, or on a privately owned site. Buyers who desire to acquire land in conjunction with the home can finance the land and home together. Market conditions and trends will indicate how the manufactured homes compete in the market place. In some communities, zoning only allows manufactured homes in certain areas, confining the market area from which comparables can be derived. Once market conditions for a manufactured home are known, it can be modeled just like any other property type. Consistency is important when using an AVM for manufactured homes. Some manufactured homes are strictly treated in the market as mobile homes (i.e., personal property). An AVM developed to value manufactured homes as real property would give a false value in the case where the home was personal property, and vice versa. AVMs developed to value manufactured personal property homes cannot be used for homes classified as real property. Some manufactured homes compete in the market place with site-built homes. Where this is the case, it is possible that an AVM designed to value detached single-family structures will produce credible results, although the model should include a variable (or variables) to capture any differences between otherwise comparable manufactured and site-built homes.

5.5 Time Series Models for Residential Property

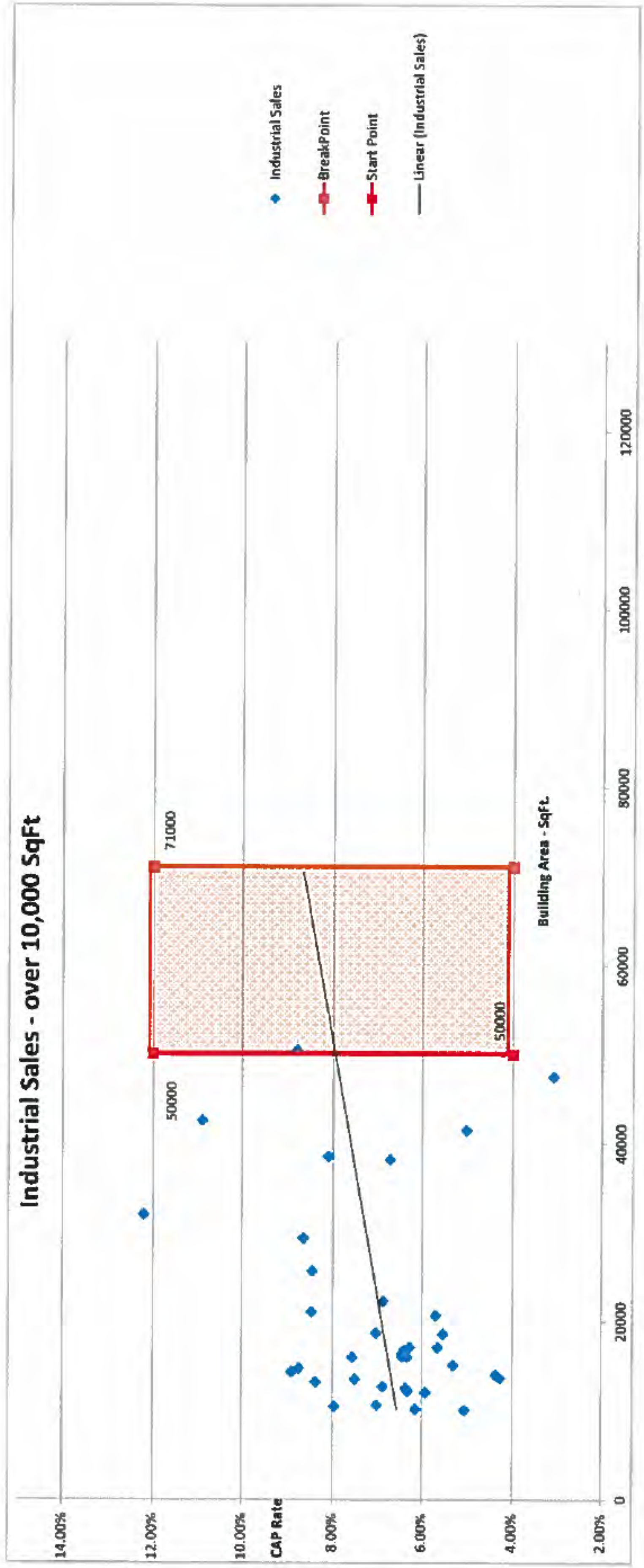
Indexed models relate to time-series analysis (see Section 4.4 on Time Series Analysis) as described earlier. Use of these models represents a common method of delivering quick automated value estimates. These models simply measure the average change in value over time and factor the value forward from a benchmark starting-place, such as the average value in a census block or market area. The accuracy of indexed models is inconsistent and less reliable than fully specified models. These models work best in areas of homogeneity where the range of value is close to the average value.

Indexing is a common method used to update cost tables to reflect current cost. As with market models, a benchmark in time is required as a starting point. Cost coefficients are then updated, using a single index factor representing the measurable change since the original cost coefficients were generated. One current method of indexing is to use an economic indicator such as the consumer price index (CPI). In the cost approach, indexed models have no way of adjusting values at the micro level for location and other market influences that impact value. Time adjustments may be developed from

Appendix N


Industrial Sales - Less than 65,000 SqFt
CAP Rate
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0
10000
20000
30000
40000
50000
60000
70000
80000
Building Area - SqFt.
50000
71000
50000
Industrial Sales
BreakPoint
Start Point

Appendix O



Industrial Sales - over 10,000 SqFt
CAP Rate
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0
20000
40000
60000
80000
100000
120000
Building Area - SqFt.
50000
71000
50000
Industrial Sales
BreakPoint
Start Point
Linear (Industrial Sales)

Appendix P

Exemption, Pseudo—Freedom from the property tax granted to property in recognition of the fact that it is taxed either directly or indirectly by other means; a reduction in the property tax base. Sometimes called "technical exemption." Note: For example, public utility property may be exempt from ad valorem property taxation but subject to a gross earnings tax in lieu thereof; corporation stock may be exempt in recognition of the taxation of the corporate assets.

Expense—A cost, or that portion of a cost, which, under accepted accounting procedures, is chargeable against income of the current year.

Expense Account—An accounting record maintained for recording particular expenses.

Expense Ratio—The ratio of expenses to gross income. A "typical" expense ratio is the relationship of normal expenses to effective gross income.

Expert Witness—One who is qualified to render expert testimony.

Exploratory Data Analysis—That part of concerned with reviewing the data set to isolate structures, uncover patterns, or reveal features that may improve the confirmatory analysis.

Exponent—A symbol usually written to the right and above an expression to indicate particular mathematical operations. For example, 6^2 means 6×6, or six squared. Fractional exponents indicate inverse operations; for example, an exponent of 1/2 signifies a square root. Exponents are also called powers. Valuation models make use of the following properties of exponents: A number raised to the exponent 0 is always 1.00; zero raised to any power is zero; any number raised to the power 1 is itself. Negative numbers cannot have exponents less than 1.

Expropriation—*See* **condemnation.**

External Diseconomies—Forces outside the activities of any single firm that cause resource prices to rise.

Statistics How To

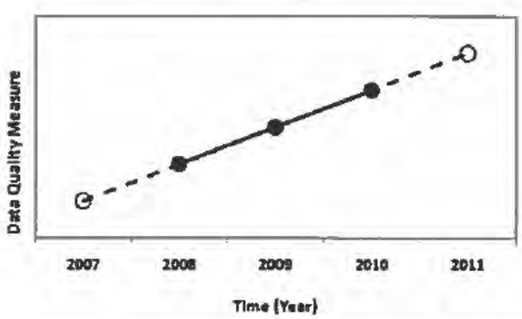
Statistics for the rest of us!

Extrapolation: What is it?

Statistics Definitions > Extrapolation

What is Extrapolation?

Extrapolation is a way to make guesses about the future or about some hypothetical situation based on data that you already know. **Interpolation** allows you to estimate *within* a data set; it's a tool to go beyond the data. It comes with a high degree of uncertainty. You're basically taking your "best guess" based on facts you already know.



The black line shows the data points. The dashed line shows a hypothetical extrapolation.

Real Life Uses

Find an article

Search

You [illegible] look forward to your monthly paycheck and you assume that you're going to get it based on known data (the fact that you've got paid monthly, on-time for the last year). But what if your company goes bankrupt? Or the market crashes? Or the bank mistakenly freezes your bank account? In this particular case, extrapolation has a fair amount of certainty (you're probably going to get your paycheck), but that isn't always the case.

Use in Statistics

Extrapolation can mean several things in statistics, but they all involve assumption and conjecture (extrapolation is far from an exact science!):

1. The extension of a statistical method where you assume similar methods will be used.
2. The projection, extension, or expansion of your known experience into an area that you do not know or that you haven't experienced yet.
3. The use of equations to fit data to a curve. You then use the equation to make conjectures. This is known as **curve fitting** or regression, which can get quite complex, with the use of tools like the Correlation Coefficient.

Other Practical Uses

Extrapolation is used in many scientific fields, like in chemistry and engineering where extrapolation is often necessary. For example, if you know the current voltages of a particular system, you can extrapolate that data to predict how the system might respond to higher voltages.

Cautions with Use

In general, you should only extrapolate for small amounts of data. For example, you might be able to rely on a steady paycheck coming in for a few months or years, but it probably wouldn't be a good idea to assume that same company is going to be still paying you 20 years down the road!

Extrapolation: What is it? was last modified: March 23rd, 2017 by Andale

By Andale | November 2, 2013 | Definitions | 1 Comment |

← Empirical Rule: What is it? Degrees of Freedom: What are they? →

One thought on "Extrapolation: What is it?"

1. ye ji
June 19, 2014 at 1:50 am

Thank you for explaining the concept of extrapolation!! It helped me a lot

Leave a Reply

Your email address will not be published. Required fields are marked *

Comment

Name *

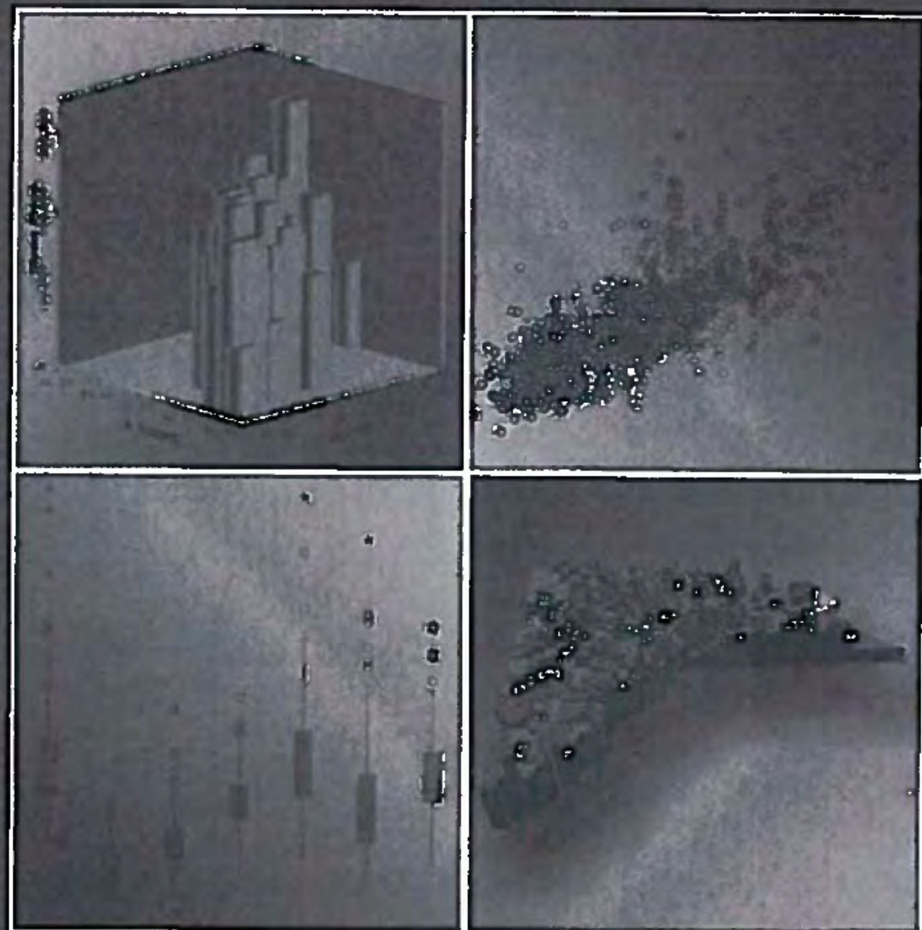
Email *

Website

Post Comment

Appendix Q

Fundamentals of Mass Appraisal



International Association of Assessing Officers

Fundamentals of Mass Appraisal

Robert Gloudemans
Richard Almy

International Association of Assessing Officers
Kansas City, Missouri

Spatial Data

Spatial data consist of maps, both hard copy and digital, that depict the location and geographic relationships among objects, namely, parcels and surrounding influences. These influences include political boundaries, market areas and neighborhoods, roads, rivers, lakes, greenbelts, mountains, and other relevant natural and man-made features. They can be displayed on maps and in geographic information systems (GIS). They can be turned on and off and represented by various colors or symbols. Maps that depict the size, degree, or relationship among objects in this matter are termed *thematic* maps.

Although spatial data are not used directly in modeling, they can be used to create variables used in modeling, for example, to determine the average elevation of parcels or to determine which properties abut or lie within a specified distance of a selected feature (e.g., lake, rivers, and major streets).

Market and property characteristics data maintained in traditional tabular format can be interfaced with spatial data and displayed thematically. For example, sales data can be linked to GIS data by parcel number and displayed in various colors depending on price. Estimated values and sales ratios can be similarly displayed.

Legal and Administrative Data

Legal and administrative data include legal description, property owner, billing address, assessment status, allowable exemptions, prior values, and appeals history. These data are not used in modeling but are needed to compute taxable values, notify property owners, and generate tax bills. How efficiently this information is maintained affects the cost and responsiveness of the assessment system.

Modern computer-assisted mass appraisal (CAMA) systems make data readily available to users, facilitate retrieval and integration of the data, and provide the public with convenient access to nonproprietary information.

Types of Variables

Independent variables used in model building represent or, more often, are based upon property characteristics. For modeling, it is helpful to categorize variables and the underlying data from which they were created as one of three types: (1) quantitative, (2) qualitative, or (3) binary. Different graphs and statistical analyses are appropriate for certain types of variables.

Quantitative

Quantitative data or
or counts, for examp
Quantitative data re
statistical analyses. T
modelers need to be
units. For example, t
fall as size increases a

During modelin
variables. For examp
ranges, as follows:

1 = 0–10 years
2 = 11–25 years
3 = 26–50 years
4 = more than

Qualitative

Qualitative data or
features or attribute
grade, condition,
normally assign c
data often are subj
programs explain

Wherever poss
and assigning vag
These categories
because they are
property values.

Binary Vari

Binary variables r
waterfront locatio
binary variables a

Binary varial
contribution of
other variables;
exterior from ext

Quantitative Variables

Quantitative data or variables, also known as *continuous*, are based on measurements or counts, for example, lot size, square feet of living area, or number of bathrooms. Quantitative data tend to be objective and verifiable and thus are well suited to statistical analyses. These data also are easily incorporated into models, although modelers need to be aware that value may not change in proportion to number of units. For example, the contribution of each additional square foot of living area may fall as size increases and a second fireplace may add less to value than a first fireplace.

During modeling, quantitative data sometimes are used to construct qualitative variables. For example, construction year could be converted to decades or to age ranges, as follows:

1 = 0–10 years
2 = 11–25 years
3 = 26–50 years
4 = more than 50 years.

Qualitative Variables

Qualitative data or variables, also termed *categorical* or *discrete*, represent property features or attributes that fall into predefined categories. Examples are construction grade, condition, building style, and type of heating or cooling. Data collectors normally assign codes for such features from a predefined list. Since qualitative data often are subjective, it is important that data collection manuals and training programs explain proper, consistent coding of such data.

Wherever possible, designers and users of CAMA systems should avoid creating and assigning vague or open-ended categories such as *other* or *to be determined.* These categories are exceedingly hard, if not impossible, to use in modeling, because they are unlikely to be bear any consistent, explainable relationship to property values.

Binary Variables

Binary variables represent the presence or absence of a specific feature, for example, waterfront location, corner lot, air conditioning, or swimming pool. In modeling, binary variables are usually coded 1 if present and 0 if not.

Binary variables provide an effective way for models to determine the contribution of property attributes to market value. They are often created from other variables; for example, a modeler could create a binary variable for brick exterior from exterior wall type (a qualitative variable).

The IAAO *Standard on Ratio Studies* (2010c) recommends the following standards for jurisdictions in which *current* market value is the legal basis of assessment. Individual jurisdictions may adopt tighter or alternative standards appropriate for their situation.

Appraisal Level

While the desired level of appraisal is 100 percent of market value, IAAO standards for measuring the level of appraisal allow a 10 percent variation. Based on the assumption that the law mandates appraisal at market value (before application of assessment ratios), this implies that the overall appraisal level should be between 0.90 and 1.10. The analyst can conclude that this standard has *not* been met when a 95 percent confidence interval (or other specified interval) about a chosen measure of central tendency fails to bracket either 0.90 or 1.10, or when a statistical test shows that the analyst can conclude with 95 percent confidence that the level of appraisal is not in this range.

The *window* of 10 percent about the market value standard provides a reasonable range in which measures of central tendency should fall in ad valorem mass appraisal. As long as this range is upheld and assessors are vigilant in reappraising property based on market value standards, property owners can use their appraisal as a reasonable indication of their property's true worth. Strict enforcement of a 100 percent standard is neither cost-effective nor practical. Such a policy could force assessors to make trivial annual adjustments in appraisals and encourage the loathsome and inequitable practice of sales chasing to achieve compliance. If strictly followed, such a policy would also force virtually half of properties to be appraised above market value, which tends to breed protracted and costly appeals and exacerbate inequities. The IAAO standard provides a reasonable, constructive, and cost-effective basis for ensuring that appraisals approximate market values. Of course, assessment officials can choose to enforce a more rigorous standard, such as 0.95 to 1.05, but a strict 100 percent standard is not recommended.

Appraisal Uniformity

As has been discussed, the three facets of appraisal uniformity are uniformity among strata, uniformity within a stratum, and value-related bias (regressivity and progressivity).

Uniformity among Strata

Each major stratum should be appraised within 5 percent of the overall level of appraisal for the jurisdiction. Thus, if the overall level is 0.900, each property class and area should be appraised between 0.855 and 0.945:

$0.900 - (0.05 \times 0.900) = 0.855;$

$0.900 + (0.05 \times 0.900) = 0.945.$

This aspect of appraisal performance is extremely important because it relates to systematic equity (or lack) thereof among major property groups.

Uniformity within a Stratum

IAAO recommends that the uniformity standard for a stratum vary with the type of property. This reflects differences in the difficulty of the appraisal task. IAAO recommends the following standards for the COD:

1. *Single-family residences.* CODs should generally be 15.0 or less and for newer and fairly homogeneous areas, 10.0 or less.
2. *Other residential property* (rural, seasonal, recreational, manufactured housing, 2- to 4-unit family housing). CODs generally should be 20.0 or less.
3. *Income-producing property.* CODs should be 20.0 or less, and in larger, urban jurisdictions, 15.0 or less.
4. *Vacant land and other property.* CODs should be 25 or less.
5. *Other real property and personal property.* Target CODs should reflect the nature of the properties, market conditions, and the availability of reliable market indicators.

While low CODs generally indicate good performance, note that the *Standard on Ratio Studies* (IAAO 2010c) regards a COD of 5.0 or less in any property category to be a possible indication of sales chasing or a nonrepresentative sample.

In general, low CODs are more difficult to achieve for heterogeneous property groups, very low- or high-value properties, and properties with older buildings or little market activity. Empirical work has shown that the ability to achieve good CODs is particularly affected by age of buildings. Difficult market conditions, in which prices are less rational and more difficult to predict, can also complicate achieving usual CODs. For example, while sales by financial institutions of previously repossessed property may constitute a significant portion of the market, use of such sales in ratio studies cannot be expected to produce CODs as good as when market conditions are more normal. Assessment officials may want to consider building such factors into performance measures or requirements.

For unique, very low-value or extremely high-value properties (outliers) and for transitional properties, even CODs of 20.0 may not be achievable. On the other hand, as indicated, CODs near zero are also not achievable and may indicate sales chasing.

Value-Related Bia

As stated in the *St*
0.98 and 1.03. As
for the measureme
graphs or formal s
in Appendix B) p
Calculated r
not proof, that
sample sizes are
of sampling erro
tests, should be

Evaluatio

The objective
population of
valuation mod
are used in th
difference in
however, sold
other criteria
and ratio sta
tendency are
very low CC
Fortuna
unsold prop
discussed i
(1) the tw
compariso
mass appra
If it is
central ten

Commu

Ratio st
based on
ratio stu

Chapter 10
Evaluating the Reliability of Ratio Study Statistics

This chapter focuses on statistical tests applicable to ratio studies and the formal evaluation of assessment performance. Assessors usually do not need to go beyond the ratio study analyses and statistics presented in Chapter 5, "Ratio Studies in Mass Appraisal," and supplemental feedback on valuation accuracy obtained through field reviews, the continuous monitoring of price trends, and informal inquiries and appeals. However, ratio study findings, particularly those of assessment oversight agencies, often have substantial fiscal implications, such as the distribution of school aid or equalization grants or the need to issue reappraisal or equalization orders. In such cases accurate estimates of appraisal level and performance are paramount. This chapter describes techniques for evaluating the reliability of computed ratio study statistics and making appropriate decisions based on ratio studies. It also describes methods for evaluating performance on unsold properties.

Confidence Intervals

As stated in Chapter 5, "Ratio Studies in Mass Appraisal," ratio study statistics can provide misleading indications of the true level and uniformity of appraisals when samples are inadequate. Fortunately, there are several methods for evaluating the reliability of a statistic. One method is to compute confidence intervals; another is to conduct an appropriate test, as discussed in the section on "Hypothesis Testing," below.

A *confidence interval* is an estimate of the range of values in which a population parameter lies depending on the analyst's specified confidence level. As discussed below and as with point estimates generally (the realized values of statistics or estimators from the sample in question), confidence intervals depend on the confidence level (confidence coefficient) chosen and on the sampling process. The *confidence level* is an expression of the probability that interval does in fact contain the parameter. The most commonly used confidence levels are 90, 95, and 99 percent. The higher the confidence level, the broader the range of the confidence interval. A 95 percent

confidence level is well suited to most ratio study analyses, although ot
can also be appropriate depending on the application at hand. In s
confidence interval provides an indication of the margin of error in a s

At any level of confidence, the width of the confidence interval is a fu
the sample size and the distribution of the ratios. Larger samples and lo
sures of dispersion are associated with tighter confidence intervals, as illu
Table 5-4 in Chapter 5.

To understand the role of confidence intervals, it is important to recall
ference between *statistics* (such as the mean and standard deviation) and pa
Statistics are calculated from samples and serve as *point estimates* of correspo
population parameters. The true value of the parameters is unknown and n
estimated. Confidence intervals quantify the range in which the analyst ca
clude that population parameters lie with a stated level of confidence.

Confidence intervals can be calculated about the mean, median, and w
mean ratios. The formulas assume that sales randomly represent the pop
of parcels. If sales are concentrated in areas in which appraisal levels are un
low or high or if sold and unsold properties are appraised differently, meas
central tendency and confidence intervals will not be representative.

Mean Confidence Interval

A confidence interval for the mean assumes a normal distribution. The formu

$$CI(\overline{AIS}) = \overline{AIS} \pm (t)(s) \div \sqrt{n},$$

where

$CI(\overline{AIS})$ = the abbreviation for the confidence interval about the mean
t = the t-value corresponding to the desired confidence level and sample
(see Appendix A)
s = the standard deviation
n = the sample size.

Note that the t-value is based on *degrees of freedom*, which, in this case, is the samp
size less 1. Since confidence intervals are two-tailed, use the t-value correspond
to a two-tailed test and the desired confidence level. For a 95 percent confiden
level and 60 degrees of freedom, the t-value is 2.00. For very large samples, th
t-value is 1.96 (equivalent to a z-value).

As an example, assume that a sample of 25 sales yields a mean sales ratio o
1.014 and a standard deviation of 0.132. The 95 percent confidence interval is

$$1.014 \pm (2.064 \times 0.132) \div \sqrt{25} = 1.014 \pm 0.054.$$

Thus, the lower confi
confidence limit is 1.01
Although not neces
to reflect the size of th

$$\sqrt{\frac{N-n}{N-1}},$$

where
N = the size of th
n = the sample s
Assume there we
correction factor is

$$\sqrt{\frac{1000-25}{1000-1}} = \ldots$$

and the confidence

1.014 ± (0.988

The adjusted conf
0.053) = 1.067, re
relative to the pop

Median Co

Unlike the mea
assumption of a
found by arrayir
to the lower anc
(j) that the ana
and upper confi

j = (1.96 ×

when n is odd,

j = (1.96 ×

when n is even
Consider
Round this n
even, the low
with the 20t

hough other levels
ıd. In summary, a
or in a statistic.
rval is a function of
les and lower mea-
als, as illustrated in

ant to recall the dif-
ion) and *parameters*.
tes of corresponding
known and must be
the analyst can con-
fidence.
edian, and weighted
sent the population
l levels are unusually
ferently, measures of
itative.

ation. The formula is
(1)

bout the mean
level and sample size

this case, is the sim
t-value correspond
95 percent confid
ery large sampl

s a mean sales
onfidence inter

Thus, the lower confidence limit is 1.014 − 0.054, or 0.960, and the upper confidence limit is 1.014 + 0.054, or 1.068.

Although not necessary in ratio studies, the confidence interval can be adjusted to reflect the size of the sample relative to the population. The correction factor is

$$\sqrt{\frac{N-n}{N-1}}, \qquad (2)$$

where

N = the size of the population

n = the sample size.

Assume there were 1,000 parcels in the population in the above example. The correction factor is

$$\sqrt{\frac{1000-25}{1000-1}} = \sqrt{\frac{975}{999}} = 0.988,$$

and the confidence interval could be refined as follows:

$$1.014 \pm (0.988 \times 0.054) = 1.014 \pm 0.053.$$

The adjusted confidence limits are 1.069 (1.014 − 0.053) = 0.961 and (1.014 + 0.053) = 1.067, respectively. The correction is minimal as long as samples are small relative to the population and can be ignored in most ratio studies.

Median Confidence Interval

Unlike the mean, the median confidence interval does not depend on the assumption of a normal distribution, nor is it as affected by outlier ratios. It is found by arraying the ratios and identifying the ranks of the ratios corresponding to the lower and upper confidence limits. The equation for the number of ratios (j) that the analyst must count up and down from the median to find the lower and upper confidence limits is

$$j = (1.96 \times \sqrt{n}) \div 2, \qquad (3)$$

when n is odd, and

$$j = (1.96 \times \sqrt{n}) \div 2 + 0.5, \qquad (4)$$

when n is even.

Consider the example of the 40 ratios in Table 10-1. The value of j is 6.70. Round this number to the next largest integer (7). Since the number of ratios is even, the lower confidence limit is found by counting *down* seven ratios beginning with the 20th-largest ratio (0.950) to obtain 0.890, and the upper limit is found

Appendix R

Google When to use a one-sample t-test

Sign in

All Videos Images News Maps More Search tools

About 13 200 000 results (0.04 seconds)

The one-sample t-test is used only for tests of the sample mean. Thus, our hypothesis tests whether the average of our sample (M) suggests that our students come from a population with a know mean (m) or whether it comes from a different population.

One-Sample t-Test
www.psychology.emory.edu/clinical/bliwise/Tutorials/TOM/meanstests/tone.htm

About this result • Feedback

One Sample T-Test - Statistics Solutions
www.statisticssolutions.com/manova-analysis-one-sample-t-test/
One sample t-test is a statistical procedure used to examine the mean difference between the sample and the known value of the population mean. In one sample t test, we know the population mean.

One-Sample t-Test
www.psychology.emory.edu/clinical/bliwise/Tutorials/TOM/meanstests/tone.htm
The one sample t test is used only for tests of the sample mean. Thus, our hypothesis tests whether the average of our sample (M) suggests that our students come from a population with a know mean (m) or whether it comes from a different population.
One Sample Z · T-test for Dependent Means · T-test for Independent Means

One-Sample T-Test using SPSS Statistics - Laerd Statistics
https://statistics.laerd.com/spss-tutorials/one-sample-t-test-using-spss-statistics.php
The one sample t test is used to determine whether a sample comes from a population with a specific mean ... This "quick start" guide shows you how to carry out a one sample t test using SPSS Statistics, as well as interpret and report the results from this test. ... Fortunately, when ...

QMSS e-Lessons | One-Sample T-Test
ccnmtl.columbia.edu/projects/qmss/the_ttest/onesample_ttest.html
To reiterate, the one-sample t-test compares the mean score of a sample to a ... the issues involved and demonstrate how to conduct a t-test using actual data.

Why should I use a 1-sample t-test? - Minitab
support.minitab.com/en-us/minitab/17/topic.../tests/tests-of.../why-use-1-sample-t/
To perform this test, select Stat > Basic Statistics > 1 Sample t. Use the 1 sample t test to estimate the mean of a population and compare it to a target or reference value when you do not know the ... Alternative hypothesis. Choose one ...

One Sample t-Test - YouTube

https://www.youtube.com/watch?v=VPd8DOL13Iw
Aug 6, 2010 - Uploaded by statisticslectures
One Sample t Test ... which statistical test to be used when you know both sample as well as ... This is y we ...

One Sample t Test - SPSS Tutorials - LibGuides at Kent State University
libguides.library.kent.edu/SPSS/OneSampletTest
4 days ago - The One Sample t Test determines whether the sample mean is ... data is significantly different than 66.5 inches using a one sample t test.

[PDF] One Sample T-test
lap.umd.edu/psyc200/handouts/psyc200_0810.pdf
Sample size can be small. The only difference between the z- and t-tests is that the t statistic estimates standard error by using the sample standard deviation ...

[PDF] One-Sample T-Test - NCSS.com
www.ncss.com/wp-content/themes/ncss/pdf/Procedures/.../One-Sample_T-Test.pdf
the one-sample t test is the best choice for assessing whether the measure of ... This section describes the assumptions that are made when you use one of ...

Independent One-Sample T-Test - Testing Samples against a Population
https://explorable.com/independent-one-sample-t-test

One-Sample T-Test using SPSS Statistics

Introduction

The one-sample t-test is used to determine whether a sample comes from a population with a specific mean. This population mean is not always known, but is sometimes hypothesized. For example, you want to show that a new teaching method for pupils struggling to learn English grammar can improve their grammar skills to the national average. Your sample would be pupils who received the new teaching method and your population mean would be the national average score. Alternately, you believe that doctors that work in Accident and Emergency (A & E) departments work 100 hour per week despite the dangers (e.g., tiredness) of working such long hours. You sample 1000 doctors in A & E departments and see if their hours differ from 100 hours.

This "quick start" guide shows you how to carry out a one-sample t-test using SPSS Statistics, as well as interpret and report the results from this test. However, before we introduce you to this procedure, you need to understand the different assumptions that your data must meet in order for a one-sample t-test to give you a valid result. We discuss these assumptions next.

One-Sample *t*-Test

Hypothesis

The one-sample *t*-test is used when we want to know whether our sample comes from a particular population but we do not have full population information available to us. For instance, we may want to know if a particular sample of college students is similar to or different from college students in general. The one-sample *t*-test is used only for tests of the sample mean. Thus, our hypothesis tests whether the average of our sample (M) suggests that our students come from a population with a know mean (μ) or whether it comes from a different population.

The statistical hypotheses for one-sample *t*-tests take one of the following forms, depending on whether your research hypothesis is directional or nondirectional. In the equations below μ_1 refers to the population from which the study sample was drawn; μ is replaced by the actual value of the population mean. The statistical hypotheses are identical to those used for one-sample *Z* tests.

$$H_0 : \mu_1 = \mu$$
$$H_A : \mu_1 \neq \mu$$

$$H_0 : \mu_1 \leq \mu$$
$$H_A : \mu_1 > \mu$$

$$H_0 : \mu_1 \geq \mu$$
$$H_A : \mu_1 < \mu$$

Study Design

The name of the one-sample *t*-test tells us the general research design of studies in which this statistic is selected to test hypotheses. We use the one-sample *t*-test when we collect data on a single sample drawn from a defined population. In this design, we have one group of subjects, collect data on these subjects and compare our sample statistic (M) to the population parameter (μ). The population parameter tells us what to expect if our sample came from that population. If our sample statistic is very different (beyond what we would expect from sampling error), then our statistical test allows us to conclude that our sample came from a different population. Again, in the one-sample *t*-test, we are comparing the mean (M) calculated on a single set of scores (one sample) to a known population mean (μ).

Available Information

The one-sample *t*-test compares a sample to a defined population. When we say "defined" population, we are saying that the parameters of the population are known. We typically define a population distribution in terms of central tendency and variability/dispersion. But, for a one-sample *t*-test, only the population μ is known. The one-sample *t*-test cannot be done if we do not have μ . The population s is not required for the one-sample *t*-test. All *t*-tests estimate the population standard deviation using sample data (S). Population means are available in the technical manuals of measurement instruments or in research publications. Population information for the attachment scales used in the class dataset is available in the articles on reserve.

Test Assumptions

All parametric statistics have a set of assumptions that must be met in order to properly use the statistics to test hypotheses. The assumptions of the one-sample *t*-test are listed below. These assumptions are identical to those of the one-sample ***Z*** test.

- Random sampling from a defined population
- Interval or ratio scale of measurement
- Population is normally distributed

When reading the psychological literature, we can find many studies in which all of these assumptions are violated. Random sampling is required for all statistical inference because it is based on probability. Random samples are difficult to find, however, and psychologists and researchers in other fields will use inferential statistics but discuss the sampling limitations in the article. We learned in our scale of measurement tutorial that psychologists will apply parametric statistics like the *t* test for dependent means on approximately interval scales even though the tests require interval or ratio data. This is an accepted practice in psychology and one that we use when we analyze our class data. Finally, the assumption of normal distribution in the population is considered "robust". This means that the the statistic has been shown to yield useful results even when the assumption is violated. The central limit theorem tells us that even if the population distribution is unknown, we know that the sampling distribution of the mean will be approximately normally distributed if the sample size is large. This helps to contribute to the *t*-test being robust for violations of normal distribution. There are conditions we may encounter when we should not use the t-test for dependent means. If we are conducting a directional test and our sample data are highly skewed, we should consider a nonparametric alternative.

Click here to review the one-sample *Z* test:	One-Sample *Z*
Click here to review the t-test for dependent means:	Dependent *t*
Click here to review the t-test for independent means:	Independent *t*
Click here to return to the main page:	Main Page

Appendix S

SECOND CANADIAN EDITION

Statistics
A First Course

Donald H. Sanders
Educational Consultant
Fort Worth, Texas

Robert K. Smidt
California Polytechnic State University
San Luis Obispo, California

Aminmohamed Adatia
University of Regina
Regina, Saskatchewan

Glenn A. Larson
University of Regina
Regina, Saskatchewan

McGraw-Hill Ryerson

Toronto Montréal Boston Burr Ridge, IL Dubuque, IA Madison, WI New York San Francisco
St. Louis Bangkok Bogotá Caracas Kuala Lumpur Lisbon London Madrid
Mexico City Milan New Delhi Santiago Seoul Singapore Sydney Taipei

Example 3.16 **AIDS in the Americas** In Example 3.12 (page 65), we found that the mean rate of incidence of AIDS per 100,000 in 1995 for the Americas was 13.8. Obtain the standard deviation of this data set.

◆ **Solution:**

Two possible ways of answering this question are to use either formula 3.5 or 3.6. But, examining Figures 3.15a and 3.15b, we find that MINITAB and the TI-83 calculator, in addition to calculating the measures of central tendency that we have previously considered, also compute several other items, including the standard deviation, denoted as "StDev" by MINITAB and as "Sx" by the TI-83 calculator. Both agree that for this data, $s = 23.92$. ◆

Interpreting the Standard Deviation

We know that dispersion is the amount of spread or scatter that occurs in a data set. If, for example, the values in the set are clustered tightly about their mean, the measured dispersion—in this case the standard deviation—is small. But if we have other data sets where the values become more and more scattered about their means, the standard deviations for those sets become larger and larger. To summarize, then, if a standard deviation is small, the items in the data set are bunched about their mean, and if the standard deviation is large, the data items are widely dispersed about their mean. To drive home this generalization in a more tangible way, let's first consider *Chebyshev's theorem*.

Russian mathematician P. P. Chebyshev has been dead for a century now, but his theorem lives on.

STATISTICS IN ACTION

Government Statistical Engines

Many government departments collect descriptive statistics that have multiple uses. Every five years, Statistics Canada tries to determine how many people there are in the provinces and in the country. These figures, along with facts about births and deaths that are generated elsewhere to advance public health, are raw materials used by demographers—the scientists who analyze the characteristics of human populations. Similarly, government-generated international trade data and weather facts support the development of new insights in economics and atmospheric sciences.

Chebyshev's Theorem

Chebyshev's theorem states that the proportion of *any* data set that lies within k standard deviations of the mean (where k is any positive number greater than or equal to 1) is *at least* $1 - (1/k^2)$.

Thus, if we substitute 2 for k in the theorem, we get $1 - (1/k^2) = 1 - (1/2^2) = 1 - (1/4) = 3/4$, or, in percentage form, $3/4 \times 100 = 75$ percent. This result means that *at least* 75 percent of the items in *any* data set (no matter how skewed it is) must lie within two standard deviations of the mean. And at least 88.9 percent $[1 - (1/3^2)$ or $8/9]$ of the items in *any* data set must fall within three standard deviations of the mean.

Chebyshev's theorem shows us how the standard deviation is related to the scatter of data items. But it tells us only the minimum percentage of items that must fall within given intervals in any data set. We've seen earlier (and in Figure 3.8 on page 53) though that many data sets have values that are found to be distributed or scattered about their means in reasonably symmetrical ways.

For bell-shaped distributions, known as normal distributions, the *empirical rule* applies and is of greater significance than Chebyshev's theorem.

Empirical Rule

The **empirical rule** for distributions that are generally bell-shaped or normal is that

- About 68 percent of all data items lie within one standard deviation of the mean ($\mu \pm 1\sigma$ or $\bar{x} \pm 1s$).
- About 95 percent of all data items lie within two standard deviations of the mean ($\mu \pm 2\sigma$ or $\bar{x} \pm 2s$).
- About 99.7 percent of all data items lie within three standard deviations of the mean ($\mu \pm 3\sigma$ or $\bar{x} \pm 3s$).

Let's look at an example of an application of this empirical rule. Suppose that many people are given a new type of IQ test, and the resulting raw scores are organized into a frequency distribution. A frequency polygon is prepared from the distribution and is found to be symmetrical in shape. The arithmetic mean of this mound-shaped distribution is 100 points, and the standard deviation is ten points. In this situation, the mean IQ score is directly under the peak of the curve, and the following relationships exist: (1) about 68 percent of the test scores fall within one standard deviation of the mean—that is, about 68 percent of the people have test scores between 90 and 110 points; (2) about 95 percent of the test scores fall within two standard deviations of the mean—that is, about 95 percent of those taking the test have scores between 80 and 120 points; and (3) virtually all (99.7 percent) of the test scores fall within three standard deviations of the mean (scores between 70 and 130). Figure 3.21 shows these relationships.

The relationships that exist between the mean and the standard deviation in a bell-shaped distribution may also be used for analysis purposes with distributions that are only approximately symmetrical. Let's return to our Slimline Fizzy Cola example and interpret the meaning of the standard deviation of 14.78 litres, since that distribution is approximately normal. We can conclude that about the middle two-thirds of the 50 employees sold syrup quantities between $\bar{x} \pm 1s$, that is, between 115.20 litres ± 14.78

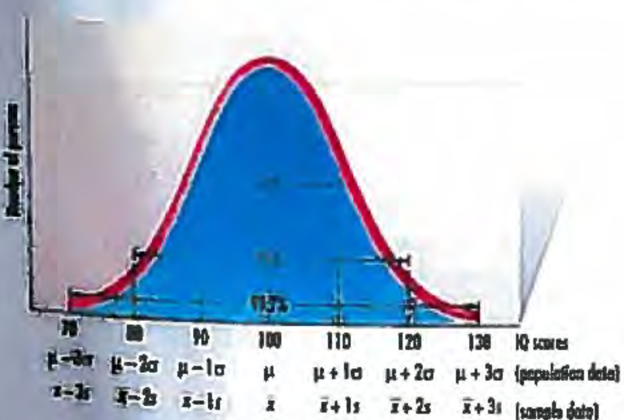


FIGURE 3.21 Illustration of the empirical rule

Appendix T

Standard on Ratio Studies

Approved April 2013

4.5 Sample Representativeness

In general, a ratio study is valid to the extent that the sample is sufficiently *representative* of the population.

The distribution of ratios in the population cannot be ascertained directly and appraisal accuracy can vary from property to property. By definition, a ratio study sample would be representative when the distribution of ratios of properties in the sample reflects the distribution of ratios of properties in the population. Representativeness is improved when the sample proportionately reflects major property characteristics present in the population of sold and unsold properties. As long as sold and unsold parcels are appraised in the same manner and the sample is otherwise representative, statistics calculated in a sales ratio study can be used to infer appraisal performance for unsold parcels.

However, if parcels that sell are selectively reappraised based on their sale prices and if such parcels are in the ratio study, uniformity inferences will not be accurate (appraisals appear more uniform than they are). In this situation, measures of appraisal level also will not be supportable unless similar unsold parcels are appraised by a model that produces the same overall percentage of market value (appraisal level) as on the parcels that sold (see Appendix E, "Sales Chasing Detection Techniques"). Assessing officials must incorporate a quality control program; including checks and audits of the data, to ensure that sold and unsold parcels are appraised at the same level.

Operationally, representativeness is improved when the following occur:

1. Appraisal procedures used to value the sample parcels are similar to procedures used to value the corresponding population
2. Accuracy of recorded property characteristics data for sold property does not differ substantially from that of unsold property,
3. Sample properties are not unduly concentrated in certain areas or types of property whose appraisal levels differ from the general level of appraisal in the population
4. Sales have been appropriately screened and validated (see Appendix A).

The first requirement generally is met unless sampled parcels are valued or updated differently from nonsampled parcels, or unless appraisals of sample parcels were done at a different time than appraisals of nonsampled parcels. For example, it is unlikely that the sample is representative of unsold parcels when the sample consists mostly of new construction, first-time sales of improved properties, condominium conversions, or newly platted lots.

The second requirement is met only if value-related property characteristics are updated uniformly for all property in a class as opposed to being updated only upon sale.

The third requirement relates to the extent to which appraisal performance for the sample reflects appraisal performance for the population.

The fourth requirement generally is met when the sales to be used in the sample are properly screened, adjusted if necessary, and validated.

4.6 Acquisition and Validation of Sales Data

Sales data are important in ratio studies and play a crucial role in any credible and efficient mass appraisal system. In some instances, it may be necessary to make adjustments to sales prices so they are more representative of the market. When there is more than one sale of the same property during a study period, only one of the transactions should be used in the ratio study. For guidelines on sales validation see Appendix A.

5. Ratio Study Statistics and Analyses

Once data have been properly collected, reviewed, assembled, and adjusted, outlier handling and statistical analysis can begin. This process involves the following steps.

1. A ratio should be calculated for each observation in the sample by dividing the appraised (or assessed) value by the sale price.
2. Graphs and exhibits can be developed that show the distribution of the ratios.
3. Exploratory data analysis, including outlier identification and screening, and tests of the hypotheses of normality may be conducted.
4. Ratio study statistics of both appraisal level and uniformity should be calculated.
5. Reliability measures should be calculated.

An example of a ratio study statistical analysis report is given in table 1-1.

5.1 Data Displays

Displays or exhibits that provide a profile or picture of ratio study data are useful for illustrating general patterns and trends, particularly to nonstatisticians. The particular form of the displays, as well as the data used (e.g., sales prices, sales ratios, and property characteristics) depends on the purposes of the particular display. Types of displays useful in ratio studies are arrays, frequency distributions, histograms, plots, and maps (Gloudemans 1999).

Graphic displays can be used to

- indicate whether a sample is sufficiently representative of the properties in a stratum
- indicate the degree of nonnormality in the distribution of ratios
- depict the overall level of appraisal

Table 1-1. Example of Ratio Study Statistical Analysis Data Analyzed

Rank of ratio of observation	Appraised value ($)	Sale Price ($)	Ratio (AV/SP)
1	48,000	138,000	0.348
2	28,800	59,250	0.486
3	78,400	157,500	0.498
4	39,840	74,400	0.535
5	68,160	114,900	0.593
6	94,400	159,000	0.594
7	67,200	111,900	0.601
8	56,960	93,000	0.612
9	87,200	138,720	0.629
10	38,240	59,700	0.641
11	96,320	146,400	0.658
12	67,680	99,000	0.684
13	32,960	47,400	0.695
14	50,560	70,500	0.717
15	61,360	78,000	0.787
16	47,360	60,000	0.789
17	58,080	69,000	0.842
18	47,040	55,500	0.848
19	136,000	154,500	0.880
20	103,200	109,500	0.942
21	59,040	60,000	0.984
22	168,000	168,000	1.000
23	128,000	124,500	1.028
24	132,000	127,500	1.035
25	160,000	150,000	1.067
26	160,000	141,000	1.135
27	200,000	171,900	1.163
28	184,000	157,500	1.168
29	160,000	129,600	1.235
30	157,200	126,000	1.248
31	99,200	77,700	1.277
32	200,000	153,000	1.307
33	64,000	48,750	1.313
34	192,000	144,000	1.333
35	190,400	141,000	1.350
36	65,440	48,000	1.363

Note: Due to rounding, totals may not add to match those on following table, which reports results of statistical analysis of above data.

Results of statistical analysis

Statistic	Result
Number of observations in sample	36
Total appraised value	$3,627,040
Total sale price	$3,964,620
Average appraised value	$100,751
Average sale price	$110,128
Mean ratio	0.900
Median ratio	0.864
Weighted mean ratio	0.915
Coefficient of dispersion (COD)	29.8%
Price-related differential (PRD)	0.98
Price-related bias (PRB) coefficient (t-value)	.232 (3.01)
95% median two-tailed confidence interval	(0.684, 1.067)
95% weighted mean two-tailed confidence interval	(0.806, 1.024)
Normal distribution of ratios (0.05 level of significance)	Reject—D'Agostino, Pearson K^2, and Shapiro-Wilk W
Date of analysis	9/99/9999
Category or class being analyzed	Residential

- depict the degree of uniformity
- depict the degree of value bias (regressivity or progressivity)
- compare the level of appraisal or degree of uniformity among strata
- detect outlier ratios
- identify specific opportunities to improve mass appraisal performance
- track performance measures over time

5.2 Outlier Ratios

Outlier ratios are very low or high ratios as compared with other ratios in the sample. The validity of ratio study statistics used to make inferences about population parameters could be compromised by the presence of outliers that distort the statistics computed from the sample. One extreme outlier can have a controlling influence over some statistical measures. However, some statistical measures, such as the median ratio, are resistant to the influence of outliers and trimming would not be required. Although the coefficient of dispersion (COD) is affected by extreme ratios, it is affected to a lesser extent than the coefficient of variation (COV) and the mean. The weighted mean and price-related differential (PRD) are sensitive to sales with high prices even if the ratios on higher priced sales do not appear unusual relative to other sales. Regression analysis, sometimes used in assessment ratio analyses (e.g., when ratios are regressed on sales prices or property characteristics, such as lot size or living area), is also affected by outliers: both ratio outliers and outliers based on the comparison characteristics (an excellent treatment of the assumptions made in regression and deviations from can be found in Cook, R.D. and Weisberg, S. 1982).

Outlier ratios can result from any of the following:

1. an erroneous sale price
2. a nonmarket sale
3. unusual market variability
4. a mismatch between the property sold and the property appraised
5. an error in the appraisal of an individual parcel
6. an error in the appraisal of a subgroup of parcels
7. any of a variety of transcription or data handling errors

In preparing any ratio study, outliers should be

1. identified
2. scrutinized to validate the information and correct errors
3. trimmed if necessary to improve sample representativeness

indicates, for example, that assessment ratios fall by 4.5% when values double and increase by 4.5% when values are halved. Like all regression coefficients, the statistical reliability of the PRB can be gauged by noting its *t*-value and related significance level, and by computing confidence intervals. In table 1-4 the PRB is -0.035 and is not statistically significant.

Unacceptable vertical inequities should be addressed through reappraisal or other corrective actions. In some cases, additional stratification can help isolate the problem. Measures of level computed for value strata should not be compared as a way of determining vertical inequity because of a boundary effect that is most pronounced in the highest and lowest strata (Schultz 1996).

5.7 Tests of Hypotheses

An appropriate test should be used whenever the purpose of a ratio study is implicitly or explicitly to test a hypothesis. A hypothesis is essentially a tentative answer to a question, such as, Are residential and commercial properties appraised at equal percentages of market value? A test is a statistical means of deciding whether the answer "yes" to such a question can be rejected at a given level of confidence. In this case, if the test leads to the conclusion that residential and commercial properties are not appraised at equal percentages of market value, some sort of corrective action on the part of assessing officials is clearly indicated.

Tests are available to determine whether the

- level of appraisal of a stratum fails to meet an established standard
- meaningful differences exist in the level of appraisal between two or more strata
- high-value properties are appraised at a different percentage of market value than low-value properties

Appropriate tests are listed in table 1-2 and discussed in Gloudemans (1999), *Property Appraisal and Assessment Administration* (IAAO 1990), and *Improving Real Property Assessment* (IAAO 1978, 137–54).

5.8 The Normal Distribution

Many conventional statistical methods assume the sample data conform to the shape of a bell curve, known as the normal (or Gaussian) distribution. Performance measures based on the mean or standard deviation can be misleading if the study sample does not meet the assumption of normality. As a first step in the analysis, the distribution of sample ratios should be examined to reveal the shape of the data and uncover any unusual features. Although ratio study samples typically do not conform to the normal distribution, graphical techniques and numerical tests can be used to explore the data thoroughly. Traditional choices are the binomial, chi-square, and Lilliefors tests. Newer and more powerful procedures are the Shapiro-Wilk W, the D'Agostino-Pearson K^2, and the Anderson-Darling A^2 tests (D'Agostino and Stephens 1986).

5.9 Parametric and Distribution-Free (Nonparametric) Statistics

For every problem that might be solved by using statistics, there is usually more than one measure or test. These measures and tests can be divided into two broad categories: parametric and distribution-free (nonparametric). Parametric statistics assume the population data conform to a known family of probability distributions (such as the normal distribution). When the mean, weighted mean, and standard deviation are used in this context, they tend to be more meaningful. Distribution-free statistics make less restrictive assumptions and do not require knowledge about the shape of the underlying population distribution. Given similar distribution of ratios in the underlying populations, distribution free tests, such as the Mann-Whitney test, can determine the likelihood that the level of assessment

Table 1-2. Tests of Hypotheses

Null Hypothesis	Nonparametric Test	Parametric Test
1. Ratios are normally distributed.	Shapiro-Wilk *W* test D'Agostino-Pearson K^2 test Anderson-Darling A^2 test Lillifores Test	N/A
2. The level of appraisal meets legal requirements.	Binomial test	*t*-test
3. Two property groups are appraised at equal percentages of market value.	Mann-Whitney test	*t*-test
4. Three or more property groups are appraised at equal percentages of market value.	Kruskal-Wallis test	Analysis of Variance
5. Low- or high-value properties are appraised at equal percentages of market value.	Spearman Rank test	PRB, correlation or regression analysis
6. Sold and unsold parcels are treated equally.	Mann-Whitney test	*t*-test

to determine whether it can be reasonably concluded that appraisal level differs from the established goal in a particular instance. Additionally, when uniformity measures show considerable variation between ratios, level measurements may be less meaningful.

9.1.1 Purpose of Level-of-Appraisal Standard

Jurisdictions that follow the IAAO recommendation of annual revaluations (*Standard on Property Tax Policy* [IAAO 2010] and *Standard on Mass Appraisal of Real Property* [IAAO 2013]) and comply with USPAP standard rules should be able to develop mass appraisal models that maintain an overall ratio level of 100 percent (or very near thereto). However, the local assessor may be compelled to follow reappraisal cycles defined by a legal authority or public policy that can extend beyond one year. During extended cycles the influence of inflation or deflation can shift the overall ratio.

The purpose of a performance standard that allows reasonable variation from 100 percent of market value is to recognize uncontrollable sampling error and the limiting conditions that may constrain the degree of accuracy that is possible and cost-effective within an assessment jurisdiction. Further, the effect of performance standards on local assessors must be considered in light of public policy and resources available.

9.1.2 Confidence Intervals in Conjunction with Performance Standards

The purpose of confidence intervals and similar statistical tests is to determine whether it can be reasonably concluded that the appraisal level differs from the established performance standard in a particular instance. A conclusion of noncompliance requires a high degree of confidence; thus, a 90 percent (two-tailed) or 95 percent (one-tailed) confidence level should be used, except for small or highly variable samples. The demonstration ratio study report in table 1-4 presents 95% two-tailed confidence interval estimates for the mean, median, and weighted mean ratio.

9.2 Appraisal Uniformity

Assuming the existence of an adequate and sufficiently representative sample, if the uniformity of appraisal is unacceptable, model recalibration and/or reappraisal should be undertaken. It is important to recognize that the COD is a point estimate and, especially for small samples, should not be accepted as proof of assessment uniformity problems. Proof can be provided by recognized statistical tests, including bootstrap confidence intervals.

In unusually homogeneous strata, low CODs can be anticipated. In all other cases, CODs less than 5 percent should be considered suspect and possibly indicative of nonrepresentative samples or selective reappraisal of selling parcels.

9.2.1 Uniformity among Strata

Although the goal is to achieve an overall level of appraisal equal to 100 percent of the legal requirement, ensuring uniformity in appraisal levels among strata also is important. The level of appraisal of each stratum (class, neighborhood, age group, market areas, and the like) should be within 5 percent of the overall level of appraisal of the jurisdiction. For example, if the overall level of appraisal of the jurisdiction is 1.00, but the appraisal

Table 1-4. Demonstration Ratio Study Report

Rank	Parcel #	Appraised value	Sale price*	Ratio	Statistic	Result
1	9	$87,200	138,720	0.629	Number (n)	17
2	10	38,240	59,700	0.641	Total appraised value	$1,455,330
3	11	96,320	146,400	0.658	Total sale price	$1,718,220
4	12	68,610	99,000	0.693	Avg appraised value	$85,608
5	13	32,960	47,400	0.695	Avg sale price	$101,072
6	14	50,560	70,500	0.717		
7	15	61,360	78,000	0.787	Mean ratio	0.827
8	16	47,360	60,000	0.789	Median ratio	0.820
9	17	56,580	69,000	0.820	Weighted mean ratio	0.847
10	18	47,040	55,500	0.848		
11	19	136,000	154,500	0.880	Coefficient of dispersion	14.5
12	20	98,000	109,500	0.895	Price-related differential	0.98
13	21	56,000	60,000	0.933	PRB	−0.035
14	22	159,100	168,000	0.947	PRB coefficient (*t*-value)	0.135 (2.4)
15	23	128,000	124,500	1.028		
16	24	132,000	127,500	1.035	95% conf. int. mean (two-tailed)	0.754 to 0.901
17	25	160,000	150,000	1.067	95% conf. int. median (two-tailed)	0.695 to 0.933
					95% conf. int. wtd. mean (two-tailed)	0.759 to 0.935

Date: 0/0/00. *No outlier trimming*

* *or adjusted sale price*

International Journal of Endocrinology Metabolism
KOWSAR Home | For Authors | Submit Online | Current Issue | Search

Int J Endocrinol Metab. 2012 Spring; 10(2): 486–489. PMCID: PMC3693611
Published online 2012 Apr 20. doi: 10.5812/ijem.3505

Normality Tests for Statistical Analysis: A Guide for Non-Statisticians

Asghar Ghasemi[1] and Saleh Zahediasl[1,*]

[1]Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, IR Iran
[*]Corresponding author: Saleh Zahediasl, Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, P.O. Box: 19395-4763, Tehran, IR Iran. Tel.: +98-2122409309, Fax: +98-2122402463, E-mail: zahedi@endocrine.ac.ir.

Received 2011 Nov 21; Revised 2012 Jan 21; Accepted 2012 Jan 28.

Abstract

Go to:

Statistical errors are common in scientific literature and about 50% of the published articles have at least one error. The assumption of normality needs to be checked for many statistical procedures, namely parametric tests, because their validity depends on it. The aim of this commentary is to overview checking for normality in statistical analysis using SPSS.

Keywords: Normality, Statistical Analysis

1. Background

Go to:

Statistical errors are common in scientific literature, and about 50% of the published articles have at least one error (1). Many of the statistical procedures including correlation, regression, t tests, and analysis of variance, namely parametric tests, are based on the assumption that the data follows a normal distribution or a Gaussian distribution (after Johann Karl Gauss, 1777–1855); that is, it is assumed that the populations from which the samples are taken are normally distributed (2-5). The assumption of normality is especially critical when constructing reference intervals for variables (6). Normality and other assumptions should be taken seriously, for when these assumptions do not hold, it is impossible to draw accurate and reliable conclusions about reality (2, 7).

With large enough sample sizes (> 30 or 40), the violation of the normality assumption should not cause major problems (4); this implies that we can use parametric procedures even when the data are not normally distributed (8). If we have samples consisting of hundreds of observations, we can ignore the distribution of the data (3). According to the central limit theorem, (a) if the sample data are approximately normal then the sampling distribution too will be normal; (b) in large samples (> 30 or 40), the sampling distribution tends to be normal, regardless of the shape of the data (2, 8); and (c) means of random samples from any distribution will themselves have normal distribution (3). Although true normality is considered to be a myth (8), we can look for normality visually by using normal plots (2, 3) or by significance tests, that is, comparing the sample distribution to a normal one (2, 3). It is important to ascertain whether data show a serious deviation from normality (8). The purpose of this report is to overview the procedures for checking normality in statistical analysis using SPSS.

2. Visual Methods

Go to:

Visual inspection of the distribution may be used for assessing normality, although this approach is usually unreliable and does not guarantee that the distribution is normal (2, 3, 7). However, when data are presented visually, readers of an article can judge the distribution assumption by themselves (9). The frequency distribution (histogram), stem-and-leaf plot, boxplot, P-P plot (probability-probability plot), and Q-Q plot (quantile-quantile plot) are used for checking normality visually (2). The frequency distribution that plots the observed values against their frequency, provides both a visual judgment about whether the distribution is bell shaped and insights about gaps in the data and outliers outlying values (10). The stem-and-leaf plot is a method similar to the histogram, although it retains information about the actual data values (8). The P-P plot plots the cumulative probability of a variable against the cumulative probability of a particular distribution (e.g., normal distribution). After data are ranked and sorted, the corresponding z-score is calculated for each rank as follows: $z = x - \bar{x} / s$. This is the expected value that the score should have in a normal distribution. The scores are then themselves converted to z-scores. The actual z-scores are plotted against the expected z-scores. If the data are normally distributed, the result would be a straight diagonal line (2). A Q-Q plot is very similar to the P-P plot except that it plots the quantiles (values that split a data set into equal portions) of the data set instead of every individual score in the data. Moreover, the Q-Q plots are easier to interpret in case of large sample sizes (2). The boxplot shows the median as a horizontal line inside the box and the interquartile range (range between the 25 th to 75 th percentiles) as the length of the box. The whiskers (line extending from the top and bottom of the box) represent the minimum and maximum values when they are within 1.5 times the interquartile range from either end of the box (10). Scores greater than 1.5 times the interquartile range are out of the boxplot and are considered as outliers, and those greater than 3 times the interquartile range are extreme outliers. A boxplot that is symmetric with the median line at approximately the center of the box and with symmetric whiskers that are slightly longer than the subsections of the center box suggests that the data may have come from a normal distribution (8).

3. Normality Tests

Go to:

The normality tests are supplementary to the graphical assessment of normality (8). The main tests for the assessment of normality are Kolmogorov-Smirnov (K-S) test (7), Lilliefors corrected K-S test (7, 10), Shapiro-Wilk test (7, 10), Anderson-Darling test (7), Cramer-von Mises test (7), D'Agostino skewness test (7), Anscombe-Glynn kurtosis test (7), D'Agostino-Pearson omnibus test (7), and the Jarque-Bera test (7). Among these, K-S is a much used test (11) and the K-S and Shapiro-Wilk tests can be conducted in the SPSS Explore procedure (Analyze → Descriptive Statistics → Explore → Plots → Normality plots with tests) (8).

The tests mentioned above compare the scores in the sample to a normally distributed set of scores with the same mean and standard deviation; the null hypothesis is that "sample distribution is normal." If the test is significant, the distribution is non-normal. For small sample sizes, normality tests have little power to reject the null hypothesis and therefore small samples most often pass normality tests (7). For large sample sizes, significant results would be derived even in the case of a small deviation from normality (2, 7), although this small deviation will not affect the results of a parametric test (7). The K-S test is an empirical distribution function (EDF) in which the theoretical cumulative distribution function of the test distribution is contrasted with the EDF of the data (7). A limitation of the K-S test is its high sensitivity to extreme values; the Lilliefors correction renders this test less conservative (10). It has been reported that the K-S test has low power and it should not be seriously considered for testing normality (11). Moreover, it is not recommended when parameters are estimated from the data, regardless of sample size (12).

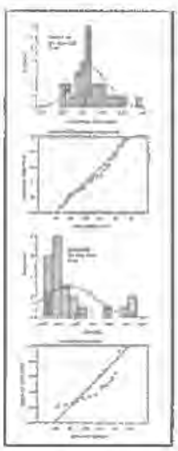
The Shapiro-Wilk test is based on the correlation between the data and the corresponding normal scores (10) and provides better power than the K-S test even after the Lilliefors correction (12). Power is the most frequent measure of the value of a test for normality—the ability to detect whether a sample comes from a non-normal distribution (11). Some researchers recommend the Shapiro-Wilk test as the best choice for testing the normality of data (11).

4. Testing Normality Using SPSS

Go to:

We consider two examples from previously published data: serum magnesium levels in 12–16 year old girls (with normal distribution, n = 30) (13) and serum thyroid stimulating hormone (TSH) levels in adult control subjects (with non-normal distribution, n = 24) (14). SPSS provides the K-S (with Lilliefors correction) and the Shapiro-Wilk normality tests and recommends these tests only for a sample size of less than 50 (8).

In Figure, both frequency distributions and P-P plots show that serum magnesium data follow a normal distribution while serum TSH levels do not. Results of K-S with Lilliefors correction and Shapiro-Wilk normality tests for serum magnesium and TSH levels are shown in Table. It is clear that for serum magnesium concentrations, both tests have a p-value greater than 0.05, which indicates normal distribution of data, while for serum TSH concentrations, data are not normally distributed as both p values are less than 0.05. Lack of symmetry (skewness) and pointiness (kurtosis) are two main ways in which a distribution can deviate from normal. The values for these parameters should be zero in a normal distribution. These values can be converted to a z-score as follows:



Figure

Histograms (Left) and P-P Plots (Right) for Serum Magnesium and TSH Levels

Table

Skewness, kurtosis, and Normality Tests for Serum Magnesium and TSH Levels Provided by SPSS

$Z_{Skewness} = Skewness\text{-}0 / SE_{Skewness}$ and $Z_{Kurtosis} = Kurtosis\text{-}0 / SE_{Kurtosis}$.

An absolute value of the score greater than 1.96 or lesser than -1.96 is significant at $P < 0.05$, while greater than 2.58 or lesser than -2.58 is significant at $P < 0.01$, and greater than 3.29 or lesser than -3.29 is significant at $P < 0.001$. In small samples, values greater or lesser than 1.96 are sufficient to establish normality of the data. However, in large samples (200 or more) with small standard errors, this criterion should be changed to ± 2.58 and in very large samples no criterion should be applied (that is, significance tests of skewness and kurtosis should not be used) (2). Results presented in Table indicate that parametric statistics should be used for serum magnesium data and non-parametric statistics should be used for serum TSH data.

5. Conclusions

Go to:

According to the available literature, assessing the normality assumption should be taken into account for using parametric statistical tests. It seems that the most popular test for normality, that is, the K-S test, should no longer be used owing to its low power. It is preferable that normality be assessed both visually and through normality tests, of which the Shapiro-Wilk test, provided by the SPSS software, is highly recommended. The normality assumption also needs to be considered for validation of data presented in the literature as it shows whether correct statistical tests have been used.

Acknowledgments

Go to:

The authors thank Ms. N. Shiva for critical editing of the manuscript for English grammar and syntax and Dr. F. Hosseinpanah for statistical comments.

Footnotes

Go to:

Implication for health policy/practice/research/medical education: Data presented in this article could help for the selection of appropriate statistical analyses based on the distribution of data.

Please cite this paper as: Ghasemi A, Zahediasl S. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. Int J Endocrinol Metab. 2012;10(2):486-9. DOI: 10.5812/ijem.3505

Financial Disclosure: None declared.

Funding/Support: None declared.

References

Go to:

1. Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society. Am J Physiol Endocrinol Metab. 2004;287(2):E189–91. doi: 10.1152/ajpendo.00213.2004. [PubMed] [Cross Ref]

2. Field A. Discovering statistics using SPSS. 3 ed. London: SAGE publications Ltd; 2009. p. 822.

3. Altman DG, Bland JM. Statistics notes: the normal distribution. Bmj. 1995;310(6975):298. [PMC free article] [PubMed]

4. Pallant J. SPSS survival manual, a step by step guide to data analysis using SPSS for windows. 3 ed. Sydney: McGraw Hill; 2007. pp. 179–200.

5. Driscoll P, Lecky F, Crosby M. An introduction to everyday statistics-1. J Accid Emerg Med. 2000;17(3):205–11. [PMC free article] [PubMed]

6. Royston P. Estimating departure from normality. Stat Med. 1991;10(8):1283–93. [PubMed]

7. Oztuna D, Elhan AH, Tuccar E. Investigation of four different normality tests in terms of type 1 error rate and power under different distributions. Turkish Journal of Medical Sciences. 2006;36(3):171–6.

8. Elliott AC, Woodward WA. Statistical analysis quick reference guidebook with SPSS examples. 1st ed. London: Sage Publications; 2007.

9. Altman DG, Bland JM. Detecting skewness from summary information. Bmj. 1996;313(7066):1200. [PMC free article] [PubMed]

10. Peat J, Barton B. Medical Statistics: A guide to data analysis and critical appraisal. Blackwell Publishing; 2005.

11. Thode HJ. Testing for normality. New York: Marcel Dekker; 2002.

12. Steinskog DJ. A cautionary note on the use of the Kolmogorov-Smirnov test for normality. American Meteor Soc. 2007;135:1151–7.

13. Ghasemi A, Syedmoradi L, Zahediasl S, Azizi F. Pediatric reference values for serum magnesium levels in Iranian subjects. Scand J Clin Lab Invest. 2010;70(6):415–20. doi: 10.3109/00365513.2010.504280. [PubMed] [Cross Ref]

14. Zahediasl S, Brojeni NK, Ghasemi A, Faraji F, Hedayati M, Azizi F. Alterations in osmotic fragility of the red blood cells in hypo and hyperthyroid patients. J Endocrinol Invest. 2009;32(1):28–32. [PubMed]

IBM SPSS Data Collection V6.0.1 documentation

Version 6.0.1

Welcome to the IBM SPSS Data Collection V6.0.1 documentation, where you can find information about how to install, maintain, and use the IBM SPSS Data Collection applications.

Getting started

- Product overview
- What's new
- Samples
- IBM SPSS Data Collection Desktop installation guide
- IBM SPSS Data Collection Server installation guide
- Release notes

Common tasks

- Author tutorials
- Professional tutorials
- Data Entry and Interviewer tutorials
- Paper tutorials
- Interviewer Server tutorials

Troubleshooting and support

More information

English

IBM Knowledge Center

Search Content Products Help

SPSS Data Collection

Select

Select a specific version or edition of SPSS Data Collection documentation from the drop-down menu.

Learn more

- Online catalog of software products: General information about software products
- Explore IBM Systems: General information about Systems products
- IBM Cloud Computing: Materials about the promise of cloud
- Redbooks: Technical publications by experts about hundreds of subjects
- developerWorks: A community site with a vast amount of information for IT professionals

Find support

- IBM Support Portal
- IBM Software Support
- IBM Client Success Portal: Support for cloud solutions
- IBM Passport Advantage: Support for licensed software

Stay current

Contact Privacy Terms of use Accessibility Feedback
Cookie preferences

English

Significance value

The significance value, or p value, is the probability that a result occurred by chance. The significance value is compared to a predetermined cutoff (the significance level) to determine whether a test is statistically significant. If the significance value is less than the significance level (by default, 0.05), the test is judged to be statistically significant.

The significance value does not indicate whether a result is practically significant. Effect size is another measure from a statistical test. It helps determine the practical significance. IBM® Watson Analytics™ uses both the significance value and the effect size to determine whether a result is important enough to display.

Parent topic:

Statistical terms

Setting the Significance Levels

Version 6.0.1

By default, the column proportions, column mean, net difference test, and paired preference tests are run at the 5% significance level. However, you can optionally run a test at another significance level, such as the 10% or 1% significance level.

You can also run the test at two significance levels on the same table. In the resulting table, the IDs of columns that are significant at the higher level appear in upper case, and those that are significant at the lower level appear in lower case.

You select this option using the *SigLevel* statistics property. For example:

```
TableDoc.Tables.MyTable.Statistics[0].Properties["SigLevel"] = 1
```

The Statistics and Statistic objects implement the mrScriptBasic dynamic property expansion feature. This means that an alternative way of writing this would be:

```
TableDoc.Tables.MyTable.Statistics.ColumnProportions.SigLevel = 1
```

See the topic Dynamic Property Expansion for more information.

To run the test at two significance levels, use the *SigLevelLow* statistics property to display the lower significance level. For example:

```
TableDoc.Tables.MyTable.Statistics.ColumnProportions.SigLevel = 1
TableDoc.Tables.MyTable.Statistics.ColumnProportions.SigLevelLow = 5
```

In the resulting table, the IDs of columns that are significant at the higher level appear in upper case, and those that are significant at the lower level appear in lower case.

Note: If you are using two levels of significance, ensure that the value of the *SigLevelLow* property is greater than that of the *SigLevel* property, as it represents a higher probability that the results are due to chance, and therefore a lower level of significance.

319

What value should I use for significance level?

Learn more about Minitab 17

Before you do a hypothesis test, you must choose a significance level for the test. Use the significance level to judge whether the test results are statistically significant. The significance level also determines the probability of error that is inherent in the test.

If the probability that an event occurs is less than α, the usual interpretation is that the event did not occur by chance. Formally, α is the maximum acceptable level of risk for rejecting a true null hypothesis (Type I error) and is expressed as a probability ranging between 0 and 1. The smaller the significance level, the less likely you are to make a Type I error, and the more likely you are to make a Type II error. Therefore, you should choose an alpha that balances these opposing risks of error based on their practical consequences in your specific situation.

Usually, a significance level (denoted as α or alpha) of 0.05 works well. A significance level of 0.05 indicates a 5% risk of concluding that a difference exists when there is no actual difference.

When to choose a larger alpha

Choose a larger alpha, such as 0.10, to be more certain that you will not miss detecting a difference that might exist.

For example, an engine manufacturer wants to compare the stability of new ball bearings with the current ones. If the new ball bearings are less stable, customers could have disastrous consequences. Therefore, they choose an α of 0.1 to be more certain that they will detect any possible difference in the stability.

When to choose a smaller alpha

Choose a smaller alpha, such as 0.01, to be more certain that you will only detect a difference that really does exist.

For example, a pharmaceutical company wants to be very certain before making an advertising claim that its new product significantly reduces symptoms. The company chooses an α of 0.001 to be sure that any significant difference in symptoms that they detect actually does exist.

Appendix U

Chapter 2. Frequencies

The Frequencies procedure provides statistics and graphical displays that are useful for describing many types of variables. The Frequencies procedure is a good place to start looking at your data.

For a frequency report and bar chart, you can arrange the distinct values in ascending or descending order, or you can order the categories by their frequencies. The frequencies report can be suppressed when a variable has many distinct values. You can label charts with frequencies (the default) or percentages.

Example. What is the distribution of a company's customers by industry type? From the output, you might learn that 37.5% of your customers are in government agencies, 24.9% are in corporations, 28.1% are in academic institutions, and 9.4% are in the healthcare industry. For continuous, quantitative data, such as sales revenue, you might learn that the average product sale is $3,576, with a standard deviation of $1,078.

Statistics and plots. Frequency counts, percentages, cumulative percentages, mean, median, mode, sum, standard deviation, variance, range, minimum and maximum values, standard error of the mean, skewness and kurtosis (both with standard errors), quartiles, user-specified percentiles, bar charts, pie charts, and histograms.

Frequencies Data Considerations

Data. Use numeric codes or strings to code categorical variables (nominal or ordinal level measurements).

Assumptions. The tabulations and percentages provide a useful description for data from any distribution, especially for variables with ordered or unordered categories. Most of the optional summary statistics, such as the mean and standard deviation, are based on normal theory and are appropriate for quantitative variables with symmetric distributions. Robust statistics, such as the median, quartiles, and percentiles, are appropriate for quantitative variables that may or may not meet the assumption of normality.

To Obtain Frequency Tables

1. From the menus choose:

   Analyze > Descriptive Statistics > Frequencies...
2. Select one or more categorical or quantitative variables.

Optionally, you can:
- Click **Statistics** for descriptive statistics for quantitative variables.
- Click **Charts** for bar charts, pie charts, and histograms.
- Click **Format** for the order in which results are displayed.

Frequencies Statistics

Percentile Values. Values of a quantitative variable that divide the ordered data into groups so that a certain percentage is above and another percentage is below. Quartiles (the 25th, 50th, and 75th percentiles) divide the observations into four groups of equal size. If you want an equal number of groups other than four, select **Cut points for n equal groups.** You can also specify individual percentiles (for example, the 95th percentile, the value below which 95% of the observations fall).

Chapter 4. Explore

The Explore procedure produces summary statistics and graphical displays, either for all of your cases or separately for groups of cases. There are many reasons for using the Explore procedure--data screening, outlier identification, description, assumption checking, and characterizing differences among subpopulations (groups of cases). Data screening may show that you have unusual values, extreme values, gaps in the data, or other peculiarities. Exploring the data can help to determine whether the statistical techniques that you are considering for data analysis are appropriate. The exploration may indicate that you need to transform the data if the technique requires a normal distribution. Or you may decide that you need nonparametric tests.

Example. Look at the distribution of maze-learning times for rats under four different reinforcement schedules. For each of the four groups, you can see if the distribution of times is approximately normal and whether the four variances are equal. You can also identify the cases with the five largest and five smallest times. The boxplots and stem-and-leaf plots graphically summarize the distribution of learning times for each of the groups.

Statistics and plots. Mean, median, 5% trimmed mean, standard error, variance, standard deviation, minimum, maximum, range, interquartile range, skewness and kurtosis and their standard errors, confidence interval for the mean (and specified confidence level), percentiles, Huber's M-estimator, Andrews' wave estimator, Hampel's redescending M-estimator, Tukey's biweight estimator, the five largest and five smallest values, the Kolmogorov-Smirnov statistic with a Lilliefors significance level for testing normality, and the Shapiro-Wilk statistic. Boxplots, stem-and-leaf plots, histograms, normality plots, and spread-versus-level plots with Levene tests and transformations.

Explore Data Considerations

Data. The Explore procedure can be used for quantitative variables (interval- or ratio-level measurements). A factor variable (used to break the data into groups of cases) should have a reasonable number of distinct values (categories). These values may be short string or numeric. The case label variable, used to label outliers in boxplots, can be short string, long string (first 15 bytes), or numeric.

Assumptions. The distribution of your data does not have to be symmetric or normal.

To Explore Your Data

1. From the menus choose:

   Analyze > Descriptive Statistics > Explore...
2. Select one or more dependent variables.

Optionally, you can:

- Select one or more factor variables, whose values will define groups of cases.
- Select an identification variable to label cases.
- Click **Statistics** for robust estimators, outliers, percentiles, and frequency tables.
- Click **Plots** for histograms, normal probability plots and tests, and spread-versus-level plots with Levene's statistics.
- Click **Options** for the treatment of missing values.

Explore Statistics

Descriptives. These measures of central tendency and dispersion are displayed by default. Measures of central tendency indicate the location of the distribution; they include the mean, median, and 5% trimmed mean. Measures of dispersion show the dissimilarity of the values; these include standard error, variance, standard deviation, minimum, maximum, range, and interquartile range. The descriptive statistics also include measures of the shape of the distribution; skewness and kurtosis are displayed with their standard errors. The 95% level confidence interval for the mean is also displayed; you can specify a different confidence level.

M-estimators. Robust alternatives to the sample mean and median for estimating the location. The estimators calculated differ in the weights they apply to cases. Huber's M-estimator, Andrews' wave estimator, Hampel's redescending M-estimator, and Tukey's biweight estimator are displayed.

Outliers. Displays the five largest and five smallest values with case labels.

Percentiles. Displays the values for the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles.

Explore Plots

Boxplots. These alternatives control the display of boxplots when you have more than one dependent variable. **Factor levels together** generates a separate display for each dependent variable. Within a display, boxplots are shown for each of the groups defined by a factor variable. **Dependents together** generates a separate display for each group defined by a factor variable. Within a display, boxplots are shown side by side for each dependent variable. This display is particularly useful when the different variables represent a single characteristic measured at different times.

Descriptive. The Descriptive group allows you to choose stem-and-leaf plots and histograms.

Normality plots with tests. Displays normal probability and detrended normal probability plots. The Kolmogorov-Smirnov statistic, with a Lilliefors significance level for testing normality, is displayed. If non-integer weights are specified, the Shapiro-Wilk statistic is calculated when the weighted sample size lies between 3 and 50. For no weights or integer weights, the statistic is calculated when the weighted sample size lies between 3 and 5,000.

Spread vs. Level with Levene Test. Controls data transformation for spread-versus-level plots. For all spread-versus-level plots, the slope of the regression line and Levene's robust tests for homogeneity of variance are displayed. If you select a transformation, Levene's tests are based on the transformed data. If no factor variable is selected, spread-versus-level plots are not produced. **Power estimation** produces a plot of the natural logs of the interquartile ranges against the natural logs of the medians for all cells, as well as an estimate of the power transformation for achieving equal variances in the cells. A spread-versus-level plot helps to determine the power for a transformation to stabilize (make more equal) variances across groups. **Transformed** allows you to select one of the power alternatives, perhaps following the recommendation from power estimation, and produces plots of transformed data. The interquartile range and median of the transformed data are plotted. **Untransformed** produces plots of the raw data. This is equivalent to a transformation with a power of 1.

Explore Power Transformations

These are the power transformations for spread-versus-level plots. To transform data, you must select a power for the transformation. You can choose one of the following alternatives:

- **Natural log.** Natural log transformation. This is the default.
- **1/square root.** For each data value, the reciprocal of the square root is calculated.
- **Reciprocal.** The reciprocal of each data value is calculated.
- **Square root.** The square root of each data value is calculated.

Explore

```
A1 - 37 Sales: Greater than 10000sqft and less than 65000sqft
```

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Total Net Area	37	100.0%	0	0.0%	37	100.0%

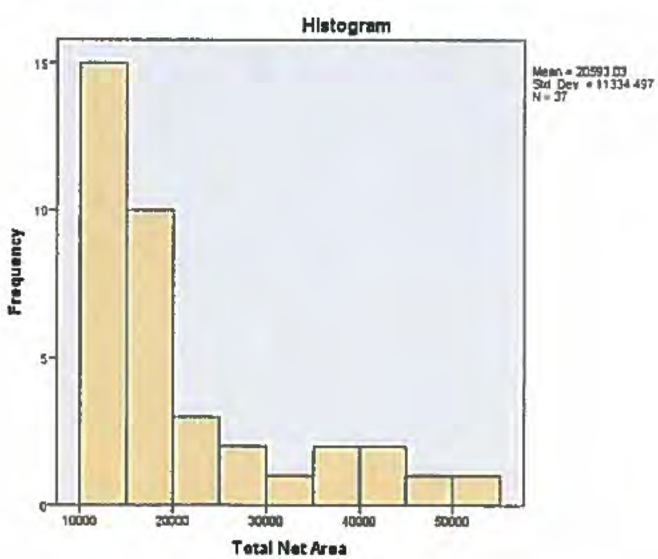
Descriptives

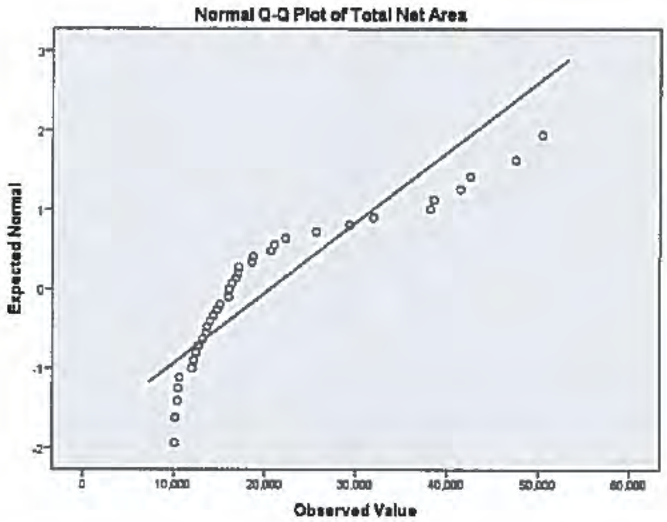
			Statistic	Std. Error
Total Net Area	Mean		20593.03	1863.380
	95% Confidence Interval for Mean	Lower Bound	16813.92	
		Upper Bound	24372.14	
	5% Trimmed Mean		19591.19	
	Median		16050.00	
	Variance		128470820.900	
	Std. Deviation		11334.497	
	Minimum		10057	
	Maximum		50462	
	Range		40405	
	Interquartile Range		11033	
	Skewness		1.394	.388
	Kurtosis		.818	.759

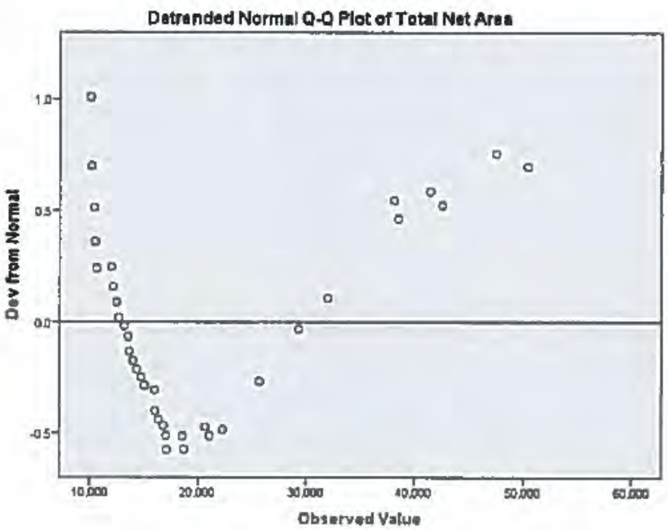
Tests of Normality

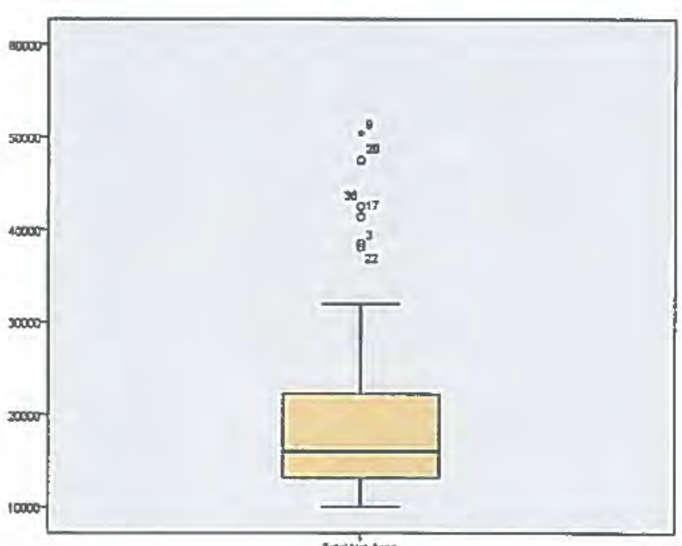
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Total Net Area	.243	37	.000	.795	37	.000

a. Lilliefors Significance Correction




Normal Q-Q Plot of Total Net Area
Expected Normal
Observed Value
3
2
1
0
-1
-2
0
10,000
20,000
30,000
40,000
50,000
60,000


Detrended Normal Q-Q Plot of Total Net Area
Dev from Normal
Observed Value
1.0
0.5
0.0
-0.5
10,000
20,000
30,000
40,000
50,000
60,000


60000
50000
40000
30000
20000
10000
9
20
36
17
3
22
Total Net Area

Appendix V

PRE 1970 AGE SUBSTRATIFICATION

MARK WHITNEY

RECORD OF DECISION

CITY OF SASKATOON, BOARD OF REVISION

APPEAL NO.: 46-2014	**ROLL NO.:** 514805350

RESPONDENT: City of Saskatoon

In the matter of an appeal to the City of Saskatoon, Board of Revision by:

APPELLANT: Altus Group Limited on behalf of
Northstar Innovative Developments

respecting the assessment of: 611 Avenue O South

Legal Description:	**Parcel No. 119886227**
Civic Address:	**611 Avenue O South**

for the year 2014

BEFORE Adrian Deschamps, Panel Chair
Dave Gabruch, Board Chair
Asit Sarkar, Member

Appeared for the Appellant Garry Coleman, Altus Group
Jesse Faith, Altus Group

Appeared for the Respondent Travis Horne, City of Saskatoon (Advocate)
Amy Huang, City of Saskatoon

The appeal was heard in the Council Chambers, City Hall, in the City of Saskatoon on May 1, 2014.

PRELIMINARY ISSUES:

The Panel Chair reviewed the hearing proceedings with the Appellant and Respondent.

The parties were advised that the proceedings were being recorded for the purposes of the Board and the Panel Clerk. The Chair introduced the Board members and the Panel Clerk.

Affirmation of truth statements were orally administered to all participants engaged in submitting evidence and testimony.

Mr. Faith stated that Grounds for this appeal relates to the issues dealt with in Appeal 44-2014 – therefore, no new testimony will be provided with regard to this ground. Instead, he requested that evidence and arguments provided in Appeal 44-2014 be carried forward to this appeal.

Mr. Horne indicated that the City's position is similar to what was presented in Appeal 44-2014 – therefore, the respondent will agree that evidence and arguments in Appeal 44-2014 be forwarded to this appeal.

GROUNDS AND ISSUES:

The Grounds and Facts are presented as outlined in A.1, as follows:

The assessment valuation is in excess and should be lower to reflect market value.

Ground 1: The assessor erred in the calculation of the capitalization rate used to determine the property assessment. Supporting facts are:

a. The capitalization (CAP) rate of 7.95% currently applied to the property is too low, and in error.
b. The sales used to determine the current CAP rate that are not comparable to the subject are as follows: 626 Weldon Ave, 608/616 Duchess St, 518 44th St E, 509 44th St E, 2206 Speers Ave, 2236 Avenue C N, 2225 Hanselman Ct, 2409 Thayer Ave, 307 Ontario Ave, 1939 Avenue B N, 318 Avenue K S.
c. The median sale price per square foot of properties built 1970 and prior is approximately half of the median price per square foot of properties with an effective year built lower than 1970.
d. The subject's year built is 1964.
e. 608-616 Duchess Street is a retail building that was used by the assessor to develop the 7.95% cap rate. As this building is retail, it should be removed from the warehouse cap rate analysis.

Ground 2 – The current stratification for warehouses that are less than 34,150 square feet and have a site coverage of greater than 0.47% is in error. Supporting facts are:

a. The current stratification does not accurately represent the realities of the marketplace.

b. The sales used by the assessor to develop the 7.95% cap rate indicate that properties built 1970 and prior have a median cap rate that is significantly higher than the current cap used to determine the estimated market value.

c. Re-stratifying the sales by effective year built results in a significantly more accurate market estimate and is supported by a lower COD for both new stratifications.

EXHIBITS:

The following were entered into the record:

A.1 - Notice of Appeal from Altus Group, received February 7, 2014

A.2 – Written Submission from Altus Group, (for use in 44, 46, 47, and 48-2014 only), received April 8, 2014

R.1 - Assessment Report, Warehouse & Automotive Response (for use in 44, 46, and 47-2014 only), received April 22, 2014

R.2 – 2014 General Law and Legislation Brief, (for use in 44, 46, and 47-2014 only), received April 22, 2014

FACTS

Salient Facts:

Property Description:

Roll Number: 514805350
Civic Address: 611 Avenue O South
Legal Description: Parcel: 119886227; 119886238; Plan H771 Block 7 Lot 15-16
Land: Lot/Parcel Size:
Building:
Zoning:
Predominant Property Type: Industrial, Flex Bld, Single Storey
Geographic Name
Geographic Code
Current Assessment **$286,900**

Current Taxable Assessment **$286,900**
Total Assessment **$286,900**
Percentage of Assessment **100%**

APPELLANT'S EVIDENCE & ARGUMENTS

As agreed by both the appellant and the respondent, the following has been carried forward from Appeal 44-2014.

APPELLANT'S EVIDENCE & ARGUMENTS

Mr. Garry Coleman presented no new evidence on behalf of the Appellant except to draw attention to the following:

- *The assessment of this property was appealed in 2013 (Appeal 392-2013). The Board of Revision sustained the appeal and directed that a cap rate of 08.94% be applied to the subject property.*
- *All other issues were denied.*
- *City of Saskatoon appealed the Board decision to the Saskatchewan Municipal Board. That appeal is still to be decided.*

1) *The Assessor grouped 19 sales together to develop the 7.95% warehouse cap rate. The main issue before the Panel is to determine if the sale of 608-616 Duchess Street (Duchess Street Property) is a warehouse property that can be included in the determination of warehouse cap rate of 7.95%.*

2) *The Duchess Street property is designed as a retail strip commercial property and acts like a neighbourhood shopping centre. Historically, the tenants in this property included retail establishments, professional services providers and non-profit organizations.*

3) *The Marshall and Swift costing manual describes warehouses as buildings "designed primarily for storage". The Duchess Street property is neither designed primarily for storage, nor used primarily for storage purposes.*

4) *The Market Value Assessment Handbook guides the assessor in establishing a cap rate by observing the "Economic conditions, competition and expected changes in competition, location, property age, property condition and property design." The design of the Duchess Street property is that of a retail commercial strip building.*

5) *It is the appellant's contention that when the Duchess Street property sale is excluded from the cap rate determination process, the new cap rate will be 8.94%. This is the cap rate that should be applied in the assessment of the subject property.*

ASSESSOR'S EVIDENCE & ARGUMENTS

The Respondent was represented by Mr. Travis Horne of the City of Saskatoon Assessment Branch. Mr. Horne agreed that the grounds of appeal and the evidence for this appeal were identical to those in Appeal 392-2013 and he did not speak further to the evidence that is on record. The following is a summary of his testimony, evidence and arguments:

1) *The Duchess Street property is located in a commercial/industrial area and is zoned MX1 whose purpose is to facilitate reinvestment in core neighbourhoods and industrial areas of the city by encouraging mixed uses in new development as well as promoting rehabilitation of existing structures. The MX1 district is intended to facilitate a broad range of compatible, industrial, institutional, cultural, and residential uses, including live/work units.*

2) *The Duchess Street property, upon inspection, was found to have the elements of a warehouse design such as masonry bearing walls, concrete floor, 12-foot ceiling and overhead door entries. Historically, this property has always been assessed as a warehouse.*

3) *Although the Duchess Street property has a variety of tenants, one of them – Medi-Chair – occupies the largest area (65%) of the building and 53% of the space occupied by this tenant is storage warehouse.*

4) *Although the Appellant may have a different interpretation of the Marshall-Swift manual, as long as the City applies it consistently, there is no error. The reference to SMB decision in Cadillac Fairview Corporation was used to justify the Appellant's discretion in including the sale of the Duchess Street property in determining the warehouse cap rate.*

Mr. Faith presented the following evidence with regard to Ground 2.

1) *Reference was made to BofR decision in 16-2013 (Exhibit A2, page 13). In this decision, the Board found it appropriate for stratification based on age of the property in order to determine the market value of a similar property. The Board noted that the model applied by the Assessor did not capture the obsolescence for age. Specifically, Board came to the following conclusion with regard to obsolescence – There is evidence of variances in the pre-1970 sales and there may be greater variances demonstrated if there were sales available of much older buildings (Exhibit A2, Appendix M, Page 196). Based on this, the Appellant argued that based on recalculation of the CAP rate after taking into account two age groups (1970 and before; post-1970), a CAP rate of 13.48% was obtained for sales of property with effective ages prior to 1970 (Exhibit A2, Appendix O, page 180). It was also argued that there was improvement in CODs for both the groups – from 34.9 for the unified group, 18.28 for 1970 and prior group, and 14.73 for post-1970 group (Exhibit A2, Appendix O, page 180).*

2) *The Appellant further argued that restratification based on the effective age of the property could be justified by using the Mann-Whitney Test. The results of the Mann-Whitney test were presented in Exhibit A2, Appendix P, page 182. It concluded that two groups were deemed to be independent samples.*

3) *The Appellant then argued that the CAP rate for the subject property should be set at 13.48% and this would result in a reduced value of assessment from $1,417,739 to $836,129. (Exhibit A2, paragraph 38, page 13). An alternative would be to apply an obsolescence of 48.71% based upon the variance of the median sale price per*

square foot, resulting in an assessed value of $727,158 (Exhibit A2, paragraph 39, pages 13-14).

4) *Upon cross-examination by the Assessor, the following additional points were made:*
   a. *19 sales were used in calculating the Mann-Whitney test. Of those, 7 sales were from the pre-1970 group and 12 from the post-1970 group.*
   b. *Mann-Whitney test was employed to provide a measure of statistical significance of the stratification based on effective age of the property. In determining the test of significance, p-value was used.*
   c. *Break-points other than 1970 were tested and not found to be statistically significant.*
   d. *Break-points were not tested for citywide sales data.*

ASSESSOR'S EVIDENCE & ARGUMENTS:

Mr. Chad Nunweiler from the City of Saskatoon's Assessment Department made the following points with regard to Ground 2:

1) *The Appellant's proposed stratification is contrary to the Mass Market Appraisal process. The Effective Age of a building is a citywide phenomenon and cannot be applied to just one market grouping, otherwise, this approach would be known as "sales chasing" (Exhibit R1, p,6). Sales chasing occurs when sales are used in a biased manner to achieve favourable results.*

2) *The Appellant's choice of year 1970 as an age break in justifying an Effective Age variable for Group 3 sales (subject property group) has no supported rationale.*

3) *As shown in Exhibit R1, page 17, three age stratifications (1985 or earlier; 1986-2007; and 2008 and later) were used in the development of the rent model. The adjustments for these stratifications reflected the market, as evidenced by the median ratios of market rent to contract rent (MCR). As all were close to 1.00, it would indicate that there was no bias in the rent model with respect to age.*

4) *When the results of the age stratification bias test was subjected to Mann-Whitney test, the conclusion was that the sales of properties with pre- and post-1977 effective year built did not provide any statistically significant difference. Therefore, it was concluded that there should not be an age break for warehouse cap rates.*

5) *The market valuation standard in the mass appraisal system must reflect typical market conditions for similar properties. Therefore, neither an Effective Age Break nor an obsolescence in age can be applied in the Warehouse Cap Rates to only one sub group, such as properties in Group 3. For equity to be achieved, these variables must be applied and tested to the entire City warehouse population.*

6) *Upon cross examination by the Appellant, it was noted that data for bias testing represented 104 citywide sales of warehouse properties. These were then sub-categorized into 4 groups based on site coverage. Another point coming out from*

cross examination was that 1977 break point was used for testing based on the graphical representation of sales.

7) *Upon questioning from the Panel, the point was made that the reason for choosing 1977 as the breakpoint to conduct the Mann-Whitney Test was because that is where data trends were evident. It was also indicated that possibility of 1970 as the breakpoint was tested, but no trend was evident at that breakpoint.*

Ms. Amy Huang was the next witness for the Assessor. Ms. Huang's evidence dealt primarily with the use of Mann-Whitney test, and the following points were made:

1) *In order to draw conclusions from Mann-Whitney test results, attention must be paid to the number of data points, and the correct use of statistical measure has to be made.*

2) *When conclusions on Mann-Whitney test are drawn using p-value, each sample size has to be higher than 9, as per IAAO guidelines. If the number of data points is less than 9, then the use of p-value is not appropriate and conclusions drawn need to be based on u-statistic value.*

3) *In determining the rationale for a citywide Market CAP rate, the first task is to see which variables are appropriate for any stratification. For that task, size and site coverage were tested and used as stratification criteria. Effective Year Built was also tested by using pre- and post-1977 as break points. But it was not found statistically significant based on Mann-Whitney test. As such, it was not used as a variable for Cap rate determination.*

4) *Although Effective Year built may be found to be important for a subgroup of warehouse property by using Mann-Whitney Test, the mandate for the Assessor is to set assessment according to the standards of mass appraisal. Therefore, the appropriate use of Mann-Whitney test is only for citywide sales of warehouse properties rather than for a subgroup of sales based on Effective Year Built.*

5) *The Appellant's evidence on Mann-Whitney test did not have the required minimum of 9 data points in each group; therefore, it was not appropriate to draw conclusions based on p-statistic value. In any case, even when p-value was used for the Effective Year Built variable, no statistically significant difference was found between two age groups.*

6) *While IAAO standard recommends the use of p-value when there is a minimum of 9 data points, there was no specific recommendation as to which value (p or u) should be used where there are less than 9 data points.*

7) *Upon cross-examination by the Appellant, it was noted that sales of warehouses that may have been utilized for some retail activity were not excluded in the determination of Cap rate. It was further noted that a particular variable must first be determined to be significant before doing further subgroupings based on this variables and attempting to alter the market cap rate.*

RULES, STATUTES, PRECEDENTS

In the general course of its deliberations, the panebal was guided by the principles expressed in Sections 164 and 165 of *The Cities Act*, the Market Value Assessment in Saskatchewan Handbook for non-regulated properties, and the Saskatchewan Assessment Agency Manual for regulated properties.

The relevant sections of *The Cities Act* are as follows:

- Section 165(2) provides that property is to be valued as of the "base date", which has been established by the Saskatchewan Assessment Management Agency (SAMA) as being January 1, 2011. In determining property value, all facts, conditions and circumstances that are required to be taken into account are to be applied as if they had existed on that base date.
- Section 165(3) directs that equity is the dominant and controlling factor in the assessment of property. Section 165(4) directs that equity in regulated property assessments is achieved by applying the regulated property assessment valuation standard uniformly and fairly. Section 165(5) states that equity in non-regulated property assessments is achieved by applying the market valuation standard so that the assessments bear a fair and just proportion to the market value of similar properties as of the applicable base date. If, as a general matter, the same methodology has been employed in the valuation of the property in question as has been employed in the valuation of other such properties in the municipality, then there is no basis, in general, for varying the valuation on appeal.
- It must be noted this is a "mass assessment" system, not an individualized appraisal system.

ANALYSIS AND CONCLUSIONS

As the evidence and arguments in this Appeal are identical to those in Appeal 44-2014, the Panel has used identical analysis and arrived at identical conclusions. These are provided below:

The Panel was presented with two grounds to consider. With respect to Ground 1, evidence consists of what was carried forward, with agreement of both the Assesso and the Appellant, from Appeal 45-2014. With respect to Ground 2, new evidence was presented.

Ground 1

Ground 1 is similar to that of Appeal 45-2014. In it, the principal issue is same as was presented in Appeal 392-2013. In arriving at the conclusion on Appeal 45-2014, the Board of Revision panel has considered what was presented in Appeal 392-2013 and the additional information presented in Appeal 45-2014. The issue before this panel is whether the information provided would justify arriving at a conclusion different from that in Appeal 45-2014. In reviewing the evidence presented, it is noted that evidence in the current appeal points to identical issues, identical facts and identical arguments. The

Panel has noted the point made by the Respondent that the Duchess Street property has the elements of a warehouse because of the large warehouse at the back, with overhead doors, and the largest tenant's use of considerable space for storage.

In considering the evidence of both parties, the Panel found that in reality, the majority of the tenants are retail or retail related. One tenant's use of some of the property for storage purposes is primarily to support the retail function of their business. As noted by the Panel in Appeal 392-2013, Medi-Chair business requires extended warehouse space due to the nature of their business of selling large bulky goods. The Panel in that appeal commented: "Medi-Chair was not a distributor and required the warehouse as such. The warehouse was exclusive to support the retail function." Further, the Panel in Appeal 392-2013 noted that there was no evidence presented by the Respondent that the Duchess Street property as a distribution warehouse.

In conclusion, from the evidence on record, this panel has not been persuaded to arrive at a decision different from what was reached in Appeal 45-2014.

Ground 2

With respect to Ground 2, the principal issue concerns the justification for further stratification of warehouse sales data by considering the Effective Year Built. Specifically, the Appellant asks for a third stratification variable, i.e., Effective Year Built, in addition to the two already adopted by the Assessor, i.e., size and site coverage. In arriving at conclusions regarding this ground, the Panel has considered the following:

1) *The Appellant provided justification for its request for further stratification from the Board decision in Appeal 16-2013. The Appellant maintained that it was actually following the direction from the decision in Appeal 16-2013. The Appellant's case is based on the statistical significance of the stratification by age, as measured by the Mann-Whitney Test.*

2) *The Assessor addressed the issue of stratification by age by saying that age was not found to be a statistically significant variable when all the citywide sales of warehouse properties were considered. The Assessor's main contention has been that establishing a market cap rate by introducing a new variable for a sub-group of sales for similar properties is contrary to the mandate of mass market valuation standard and would therefore be against the direction of the Saskatchewan Municipal Board. The appropriateness of a stratification variable must be tested by considering all sales of similar properties.*

3) *The Panel has not been presented with any evidence that contradicts the Assessor's position that property age (Effective Year Built) was not found to have statistical significance when all citywide sales of similar properties were considered. What the Appellant has presented is a justification for age stratification only for a portion of all citywide sales of warehouses. In particular, the Appellant's arguments are only with respect to Market Group 3 properties consisting of 19 sales that are less than 34,150 square feet and have site coverage greater than 47%. The Appellant has used the Mann-Whitney test as the basis for its justification for two breakpoints based on age for these properties.*

4) *In concluding on the issue of whether the Assessor has met the expectations of the Mass Appraisal process, the Panel has relied on the evidence of the Assessor that when all citywide sales were considered, age was not found to be a variable statistically significant enough for stratification. The Panel was also provided the evidence that while the graphical representation of all sales did provide a breakpoint at 1977 for age, the two resulting groups were not found to be statistically significant. Thus, the Panel was presented with the evidence that the Assessor provided valid statistical measure to justify the exclusion of age variable from the stratification. Therefore, the Panel was not persuaded that the Assessor erred by not including age as a separate stratification variable.*

5) *The Panel also took note of the evidence provided by the Appellant that at least one group of properties, i.e., Market Group 3 (group in which the subject property belonged) justified further stratification by age. In analyzing the evidence on this argument, the Panel considered two points:*

   a. *Has the Appellant used the statistically measure appropriately to support its ground? The Panel was presented with evidence that the Appellant had inadequate data points in its statistical measures, and therefore, the conclusions drawn are not supportable. The Panel was not presented with evidence to counter this position.*

   b. *Even if the debate on the appropriateness of the use of statistical measure is to be ignored, is subgrouping by a variable that was not found significant in the context of all citywide sales appropriate? The Panel was not provided evidence to justify such a substratification. Even though the Appellant referred to previous Board decisions (Appeal16-2013 and Appeal 24-2010), the Panel was not shown how their conclusions supported the Appellant's contention. In Appeal16-2013, the Panel did not address the issue of stratification as it was not a ground of appeal. The Panel in that case did find evidence of variances in sale prices because of the obsolescence factor but that has not been one of the points raised in this appeal. In Appeal 24-2010, the Board addressed issues related to the number of data points, but it did not address the issue of inclusion of a variable for further stratification when it was not found relevant in framing a model using all citywide sales. Therefore, the Panel has not persuaded that the Assessor erred by not including an age variable while assessing a subgroup of sales of warehouses.*

   c. *There was a residual argument made by the Appellant that since the identification of variables for stratification purpose was an iterative process (e.g., initial split of all sales was by size, then was by site coverage), age of the property could have been the next variable considered. However, the Appellant did not provide their own analytical framework to demonstrate how this would have resulted in age being a variable considered in the market cap determination. In making this point, the Panel took note of the fact that the Appellant's evidence dealt primarily with one sales group of 19 sales out of a total of 104 citywide sales. As such, this argument was also* *not able to persuade the Panel to arrive at a conclusion that the Assessor erred in valuing the property for assessment purposes.*

DECISION

The Panel's decision is identical to that in Appeal 44-2014 and provided below:

With respect to Ground 1,the panel rules that sale of the property 608-616 Duchess Street is not to be included in the determination of the warehouse cap rate applicable to the subject property. When the sale of this property is excluded, the resulting cap rate will be 8.94%. This is the rate that should be applied in the assessment of the subject property.

Ground 2 of the Appeal is dismissed.

The filing fee is refunded.

DATED AT SASKATOON, SASKATCHEWAN, THIS 25th DAY OF June, 2014.

CITY OF SASKATOON BOARD OF REVISION

Dave Gabruch, Chair

Asit Sarkar, Member

We concur

Adrian Deschamps, Panel Chair

Dave Gabruch, Member

City of Saskatoon

Warehouse & Automotive Response

2014 Assessment

LEAD APPEAL NO:	44-2014
ROLL NUMBER:	455007990
PROPERTY OWNER(S):	Various
APPELLANT/AGENT:	Altus Group Limited
HEARING DATE:	28/04/2014
CARRY FORWARD APPEAL NO(s):	44-2014,46-2014,47-2014

Contents

Executive Summary

An appeal has been filed with the Saskatoon Board of Revision against the assessed value of 525 43rd Street East. The subject property is assessed as an Industrial Flex Warehouse with 17,400 square feet of leasable area and has an effective year built of 1968.

Following are the appellant's grounds with the Assessment Branch's response to the grounds:

Ground 1: The assessor erred in the calculation of the capitalization rate used to determine the property assessment.

a) The capitalization (CAP) rate of 7.95% currently applied to the property is too low, and in error.

b) The sales used to determine current CAP rate that are not comparable to the subject are as follow:

626 Weldon Ave
608/616 Duchess St
518 44th St E
509 44th St E
2206 Speers Ave
2206 Speers Ave
2236Ave C N
2225 Hanselman Crt
2409 Thayer Ave
307 Ontario Ave
1939 Ave B N
318 Ave K S

c) The median sale price per square foot of properties built 1970 and prior is approximately half of the median price per square foot of properties with an effective year built newer than 1970.

d) The subject's year built is 1964.

e) 608-616 Duchess Street is a retail building that was used by the assessor to develop the 7.95% cap rate. As this building is retail, it should be removed from the warehouse cap rate analysis.

Ground 2: The current stratification for warehouses that are less than 34,150 square feet and have a site coverage of greater than 0.47% is in error.

a) The current stratification does not accurately represent the realities of the market place.

b) The sales used by the assessor to develop the 7.95% CAP rate indicate that properties built 1970 and prior have a median CAP rate that is significantly higher than the current CAP used to determine the estimated market value.

c) Re-stratifying the sales by effective year built results in a significantly more accurate market estimate and is supported by a lower COD for both new stratifications.

Assessment Response to Ground 1: *Currently Appealed to the SMB*

The subject property is assessed using the Warehouse cap rate of 7.95% which was calculated from the warehouse sales that were analyzed and verified. The sale of 608/616 Duchess Street is a warehouse sale that was used to calculate the 7.95% cap rate. This is not a retail building as the appellant suggests.

The property is located in a commercial/industrial area and is zoned MX1 whose purpose is to facilitate reinvestment in core neighbourhoods and industrial areas of the city by encouraging mixed uses in new development, as well as promoting the rehabilitation of existing structures. The MX1 District is intended to facilitate a broad range of compatible commercial, industrial, institutional, cultural, and residential uses, including live/work units.

The 608/616 Duchess Street sale was analyzed and verified as a valid warehouse sale based upon a September 2011 inspection. The building is coded as an MS-453 Industrial Flex building with 10,706 square feet, masonry bearing walls, concrete floor, a 12- foot ceiling and overhead door entries. It has an effective year built of 1974 and is of good quality. It has office and store-front finish in the front portion with the remaining rear area being warehouse space. The exterior is concrete block (70%) and brick with block back-up (30%). The heating consists of forced air (100%) and air conditioning (46% - none in warehouse areas). Historically, this property has always been assessed as a warehouse.

To classify properties into the specific categories of Warehouse, Retail, and Office, in the City of Saskatoon, the Assessment Branch relies on the information provided by the Marshall and Swift Manual and the Market Value Assessment in Saskatchewan Handbook.

The following are the definitions pertaining to the appellant's grounds. See Appendix F for further information.

Market Value Assessment in Saskatchewan Handbook:

Warehouse Valuation Guide – The primary functions of a warehouse are to store, mix, consolidate, and distribute raw materials, goods, and/or finished products. Warehouses can provide a number of these functions or can be designed for one specific use. Typically, warehouses are an integral part of a manufacturing or retailing operation or act as trans-shipment points for goods and materials. Although there is a wide variety of uses and styles of warehouses, these buildings are generally uncomplicated structures that can be adapted to a number of commercial and many light industrial uses.

Marshall & Swift Manual definition:

Industrial Flex Buildings -- are the modern multi-tenant loft structures, typically of low-rise construction. The lower qualities are purely light industrial buildings having minimal subdivisions and finish per shop space user with overhead door entries. The better qualities have fully finished customer service areas with storefront entries. Display-office areas in the higher qualities have finished floors and ceiling with good restroom facilities.

This property is considered to be a Class C based on having masonry bearing walls. It is considered to be a good quality when the interior finish consists of finished floors, ceilings and display rooms, with some extras. The Mechanicals consists of fluorescent lights, adequate restrooms and plumbing. The heating system is forced air unit and air conditioning.

The definition and physical characteristics of a Class C, good quality industrial flex building as defined by the Marshall & Swift manual meet the Duchess Street property's physical characteristics – overhead door entries, fully finished customer service areas with storefront entries, display office areas with finished floors and ceiling with good restroom facilities.

608/616 Duchess Street has been classified correctly as an Industrial Flex building based on the 2011 inspection, the Marshall & Swift definition, and the Assessment Handbook.

Assessment Response to Ground 2: *New Grounds*
The appellant states that the current stratification for warehouses that are less than 34,150 square feet and have site coverage greater than 47% is in error. This market stratification is known as Market Group 3 which consists of 19 sales which are dispersed throughout the city. The appellant has suggested that this grouping should be further stratified into 2 more groups separated by Effective Age, those that have ages older than 1970 (Pre 1970) and those with ages younger than 1970 (Post 1970).

The problem with this proposed stratification is that the Effective Age of a building is not specific to a single warehouse market sub group such as the Group 3 (subject property group). The Effective Age of a building is a city wide phenomenon and cannot be applied to just one market grouping, otherwise this approach would be known as "Sales Chasing". Sales chasing occurs when sales are used in a bias manor to achieve favourable results. In this case, the appellant has chosen to only use the 19 warehouse sales from Group 3 for the Effective Age variable, and has further chosen the year 1970 as an age break without providing any supported rationale. This is an example of sales chasing, and it is an unethical to practise for appraisers because it is not a true indication of market conditions.

If the warehouse sales are to be properly analyzed they must be analyzed City wide using all 104 Warehouse sales, otherwise the testing will be become bias towards to Group 3 and not truly reflect market conditions.

The Assessment Branch has considered the Effective Age variable for warehouse properties in its statistical testing. Appendix G shows the results of the City wide sales of 104 warehouse properties. As the graph shows, a slight trend emerges around the Effective Year built of 1977, suggesting the possibility of an age break (not 1970 as proposed by the appellant). However when testing the 1977 age break for significance, the Mann-Whitney Test is Accepted with a p-value of 0.46 (much greater than 0.05) meaning there is no significance to suggest independent age samples. This result proves that there should not be an age break for warehouse cap rates. It also proves that obsolescence for age would in fact be captured in the warehouse model, because similar to the Effective Age phenomenon, obsolescence for age would also have to be applied to the City wide warehouse sales. Therefore, we can confirm that because no Effective Age break was found to be significant, no obsolescence for age would be significant as well. It is also important to note, that if any obsolescence was present it would be captured by the Age Variable in the Rent Model. The Box Plot on page 17 shows that no bias was found for buildings that have Effective Ages older than 1985. Three age stratifications were used in the development of the rent model as the median ratios indicated by the mid-points are all close to 1.00.

The Market Valuation Standard in the mass appraisal system must "reflect typical market conditions for similar properties", therefore neither an Effective Age Break nor an obsolescence in age can be applied in the Warehouse Cap Rates to only one sub group because it would selective to the Group 3. For equity to be achieved, these variables must be applied and tested to the entire City warehouse population, and as shown in Appendix G, no significance for an age break in Warehouse Cap Rates was found. Nonetheless the Warehouse model still captures age through its rent model.

If the appellant is suggesting that they are simply making a modeling correction to obtain better statistics when there is no logical support of this, they are chasing statistics. In effect, appraisal logic is being ignored versus what is evident in the bigger picture - all in favour of simple statistics. If there is no logical reason that these sub-groups suggested by the appellant should behave different than our universe of evidence, then it is our position that this is statistics chasing. There is more to mass appraisal and market value than simply obtaining better statistics.

Argument

The Appellant relies on the guidance provided in the Handbook to suggest that the classification of 608/616 Duchess Street as a warehouse is incorrect. The City uses the M&S Manual to classify property. This practice has been accepted by the AAC:

- *Various c/o Altus Group v. Saskatoon (City)* (SMB 2009-0130 et al)
  - As it relates to the agent's position that the sale of 107 Gropper Crescent should not be included in the sales analysis, the assessor argued that the Board did not err in accepting his reliance on the low-rise building descriptors provided by Marshall. He noted that all four units included within 107 Gropper Crescent are at or above grade level and accordingly, has been categorized as a low-rise apartment building pursuant to Marshall's guidelines. He submitted that there is no requirement in the descriptors for a common entranceway and as such, existence of this feature is not a determining factor. According to the assessor, all low-rise apartment buildings have been categorized using these guidelines. [23]
  - The assessor submitted that the issue relating to the inclusion of 107 Gropper Crescent in the sales analysis is similar to an issue recently revisited by the Saskatchewan Court of Appeal (the Court) in its decision for *959360 Alberta Inc. v. City of Regina*, supra. He submitted that the following comment by the Court at paragraph 42 of this decision maintains the assessor's discretion to make reasonable choices in determining assessed values:

    > "In circumstances where the law does not dictate a particular outcome, discretion is the exercise of a decision-making power to choose one reasonable, just, fair or equitable outcome over another." [24]

- *Various c/o Altus Group v. Saskatoon (City)* (SMB 2010-0077 et al)
  - For the following reasons, the Committee does not accept the agent's position that the sale of 107 Gropper Crescent is not comparable to the low-rise sales included in the assessor's 'West WC-North Low-Rise" grouping (the subject neighbourhood). Firstly, the Committee heard the assessor's position that the following description as provided at page 1 of section 12 of Marshall has been used to classify all low-rise apartments in the city:

    > "Multiple residences, often referred to as garden apartments, are buildings of three or fewer stories, in which each unit has a kitchen and bath, and which are designed for other than transient occupancy. Priced per building, costs include common areas such as lobbies, hallways, laundry, recreation, etc." [10]

The AAC goes on to recognize the Assessors discretion when determining comparability:

- *Various c/o Altus Group v. Saskatoon (City)* (SMB 2010-0077 et al)
  - In deciding the comparability issue the Committee is aware of the following guidance provided by Cadillac Fairview Corporation noted above:

    > "[36] The word comparable is not defined in the manual! nor are the words compare, comparative, comparison or similar. We must take them to have their ordinary dictionary meaning, subject, of course, to the context in which they are used, that is, the surrounding words and the manual as a whole. The immediate context is that the purpose of the comparison to buildings that have been sold is to measure market influences on the value of any building, and to use the comparison in the determination of a MAF. Subject to this context, which will be considered in detail with the appellants I grounds of appeal, the use of words of such broad and general meaning confer upon the assessor a broad discretion in respect of determining whether buildings are comparable within the meaning of the manual.
    >
    > ...
    >
    > [42] There can be no question that the grouping of buildings chosen by the assessor, enclosed shopping centres, consists of comparable or similar buildings within the meaning of the manual. The criteria used by the assessor to arrive at the grouping were all relevant to comparability: size, age, specific use, zoning, geographic distribution, and market dominance. While a grouping which included Midtown with all other downtown commercial retail buildings (or other possible groupings) might also be said to consist of comparable buildings within the meaning of the manual by reason of having in common that they are commercial retail use buildings, the choice amongst these possible groupings was clearly left to the discretion of the assessor." **(Emphasis added)**

Although the Appellant may have a different interpretation of the M&S Manual as long as the City applies it consistently there is no error. The M&S Manual does not have the force of law and therefore cannot be reviewed as though it was.

The Appellant has provided no evidence to indicate that the classification of the Duchess sale was in any way *inconsistent* with the classification of the other sales or inventory.

Summary of Salient Facts

Appeal #	**44-2014**	**46-2014**	**47-2014**
Lead Appeal	44-2014	44-2014	44-2014
Type of Appeal	Regular	Regular	Regular
Roll	455007990	514805350	485004500
Address	525 43rd St. E	611 Ave O S	625 1st Ave N
Legal Description	118998310, 118998309, 118998523	119886227, 119886238	120321717, 120171653, 120171642, 120171631,
Land Size	34,821 ft2	6,018 ft2	28,498 ft2
Building Size	17,400 ft2	4,062 ft2	12,788 ft2
Effective Age	1968	1963	1950
Zoning	IL1	IL1	IL1
Predominant Property Type	3701 – Industrial Flex Building	3701 – Industrial Flex Building	3701 – Industrial Flex Building
Geographic Name	North Industrial	West Industrial	1st Ave North of 25th Street & City Park
Geographic Code	30023	30029	30001
Current Assessment	$1,417,700	$286,900	$1,276,300
Current Taxable Assessment	$1,417,700	$286,900	$1,276,300
Total Assessment	$1,417,700	$286,900	$1,276,300
Percentage of Assessment	100%	100%	100%

Summary of Salient Adjustments

Appeal #	44-2014	46-2014	47-2014
Lead Appeal	44-2014	44-2014	44-2014
Roll	455007990	514805350	485004500
Address	525 43rd St. E	611 Ave O S	625 1st Ave N
Base Rent	$6.92	$6.92	$6.92
Rental Market Area	1 (n/a)	4 ($ -0.88/sf)	3 ($1.49/sf)
Vacancy	2%	2%	3%
Occupancy Cost	$5.20	$5.20	$5.20
Structural Allowance	$0.20	$0.20	$0.20
Cap Rate	7.95%	7.95%	7.95%

Valuation Approach

The appraisal method employed for warehouses and automotive properties is the direct capitalization of market net operating income. Direct capitalization is widely used in mass appraisal and achieves good results while being relatively straightforward.

The analysis starts with estimating the market rents and vacancies for each property. Typical rental agreements for warehouse and automotive properties are 'net'. With net rents, the tenant is responsible for paying all the costs associated with occupying the property such as property taxes, insurance, utilities, routine maintenance, property management, etc.

The only expenses that are not passed on to the tenant are the costs associated with periodic replacement of major building components (such as roof cover or replacement of heating equipment) and costs that cannot be passed on to tenants due to vacancy. In valuation terminology, the first is a "structural allowance", and the second is a "non-recoverable expense".

Once market rents, vacancies, and expenses are determined, a net operating income (NOI) is derived for each property. The NOI is then compared to sales prices, and the sale price is expressed as a percentage of the net operating income.

Each valid sale is analyzed in this manner; then market groups defined and a typical capitalization rate is determined for each market segment. Because market rents, vacancies, and expenses are an integral part of the process, capitalization rates may vary when different market rents, vacancies, or expenses are employed.

Excluded from the analysis are partially completed buildings and those where there is a significant amount of deferred maintenance. There are a relatively large number of sales, so it is more practical to focus the analysis on those sales that are least likely to provide a distorted indication of capitalization rates.

Once typical capitalization rates are determined, they are used to value warehouse or automotive property where realistic market rents, vacancies, and expenses can be estimated.

Using direct capitalization rates in mass appraisal is straightforward when net rents are typical in the marketplace. When net rents are used, the influence of expenses in estimating net operating income is small. Direct capitalization rates are part of the common "language" of commercial real estate and reflect the rates of return negotiated by buyers and sellers. Ultimately, it is sales transactions that indicate capitalization rates.

Rent

Market rents in Saskatoon are usually negotiated on a per square foot per year basis. Our rental surveys ask property owners and managers to report on the amount of rentable space. For warehouses, the amount of space reported is almost always equal to the gross area of the building(s).

Property owners and managers were asked to provide rental information for the years 2008, 2009, and 2010. The data was analyzed using multiple regression analysis (MRA). When sample sizes are relatively large, MRA is the most commonly used analytical tool in the mass appraisal of real estate. MRA is a statistical technique that allows the user to predict one value (rent, etc) from the known values of other multiple variables simultaneously, such as varying contract rents, age, size, etc.

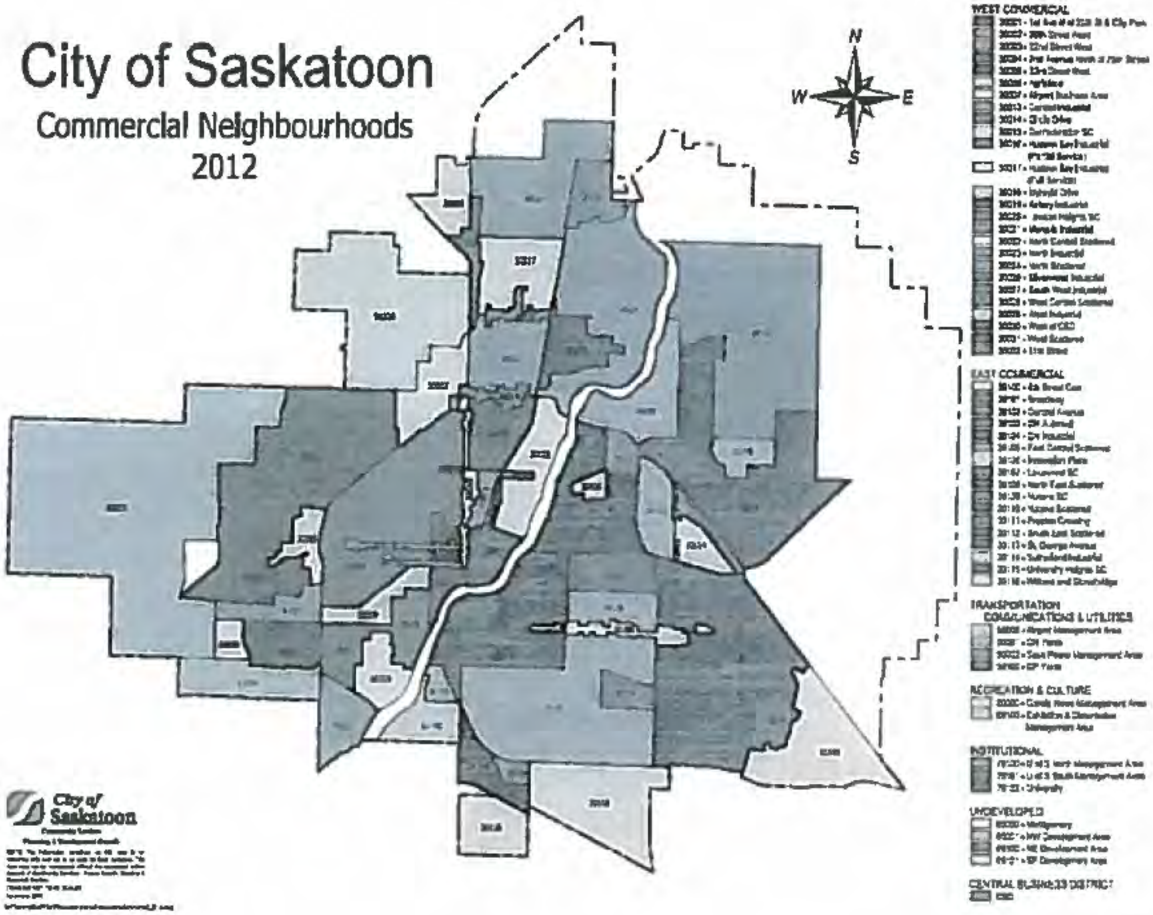
The data was tested for time trends over three years (2008, 2009, 2010) resulting in no significantly measurable changes. It can, therefore, be concluded that the rental data for all three years (2008-2010) is representative of the commercial rental market as of January 1, 2011 (the base year).

Properties with general use 3300 (Automotive) that are located on arterial roads or in commercial retail neighbourhoods were excluded from this warehouse-auto rent model analysis. They were excluded primarily because they compete in the retail market.

A total of 906 net rents were used for analysis which included warehouse-type properties and automotive-type properties located in industrial areas. Rents used for analysis were based on a lease start date between January 1, 2008 and December 31, 2010.

In order to reflect typical warehouse properties, analysis was done by excluding the following:

- reported rent where lease area was smaller than 850 square foot
- lease start date not between 2008 and 2010
- owner occupied
- gross rent, semi-gross rent
- commercial condo


City of Saskatoon
Commercial Neighbourhoods
2012
N
W
E
S
WEST COMMERCIAL
EAST COMMERCIAL
TRANSPORTATION COMMUNICATIONS & UTILITIES
RECREATION & CULTURE
INSTITUTIONAL
UNDEVELOPED
CENTRAL BUSINESS DISTRICT
City of Saskatoon

MRA Rent Model Summary

WAREHOUSE_AUTO RENT MODEL		
Assessment Year	2013	
	$/ft²	
Base Rent	6.92	

	Adjustment Amount	Count
Market Area 1 (30023 30016 30017 30114 30032 30006 30007 30020 30021 30024 30026) Base	0.00	727
Market Area 3 (30001,30003,30005, 30018,30102,30110, 30008, 30012)	1.49	35
Market Area 4 (30002,30013,30028, 30030, 30027,30029, 30104, 30103, 30113, 50002)	-0.88	61
Site Coverage < = 80%	0.00	889
Age <1985	0.00	731
WH/Flex, WH/Office, Combination, Fitness club, Automotive	0.00	846
Leasable Area < 20,000 ft²	0.00	879

Ratio Statistics for Market Rent / Contract Rent Summary

Number of rents	906
Mean	1.08
Median	1.04

Note the median ratio of the contract rent to the market rent is 1.04, indicating that the rental model successfully reflects typical market rents.

Graphical Analysis

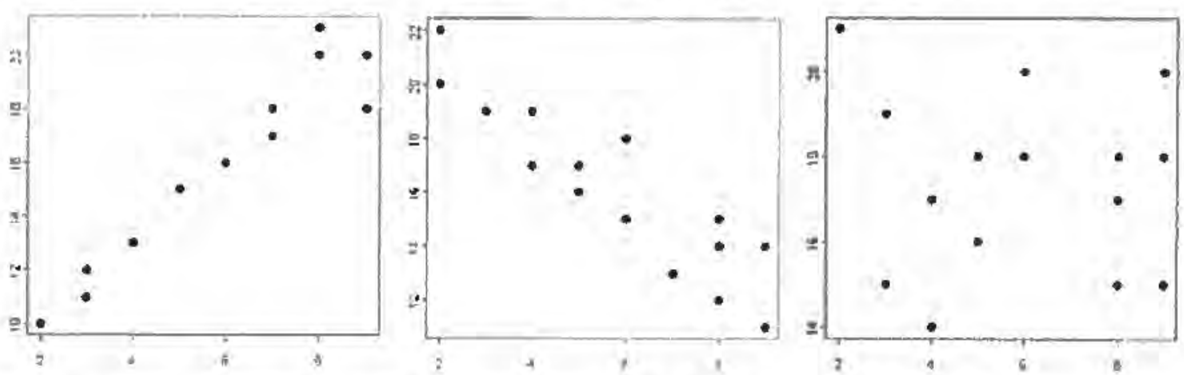
Scatter Plot Description

A bias test is used to check the quality of the regression model. The scatter plot graphically displays the market to contract rent ratio (MCR) of each contract rent in the dataset. The MCR compares the rent predicted by the model (Market Rent) with the contract rent. MCRs are calculated for each contract rent in the dataset and ideally, the MCRs should fall around 1.00.

In this case, the purpose of the bias test is to show how the model represents a particular variable or, property characteristic, by demonstrating if there is over prediction (positive correlation of MCRs) or, under prediction (negative correlation of MCRs). A non bias result is when the MCRs do not show an increasing or decreasing pattern (no correlation; that is, the data points are around 1.00 in the scatter plot.

Scatter plots show graphical relationships between continuous variables, which may show relationships and trends. The relationship between the two variables is called their correlation. The closer the data points come when plotted to making a straight line, the higher the correlation between the two variables.

Positive Correlation **Negative Correlation** **No Correlation**



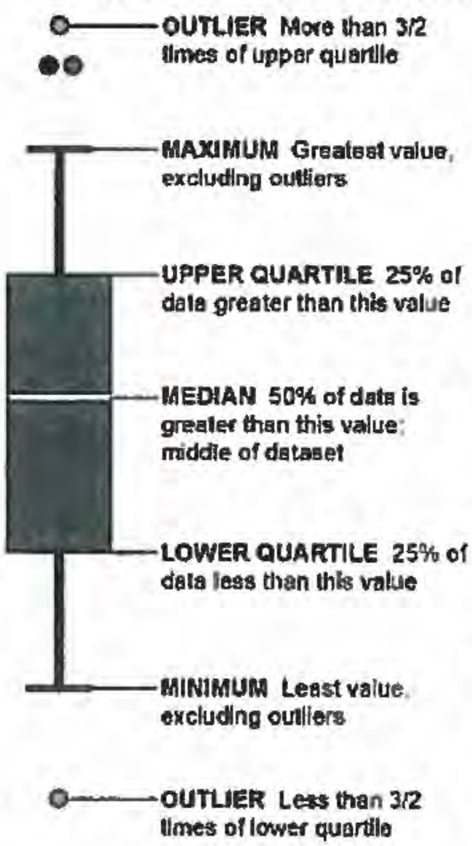
Scatter plots showing a pattern indicates that there is a relationship between the variables. Scatterplots can be used to visualize the relationship between the MCRs. The MCR is on the vertical axis, while the variables examined are on the horizontal axis. The line at 1.00 relative to each point displays the variability of the model application.

Box Plot Description

A bias test is used to check the quality of the regression model. The box plot graphically summarizes the market rent to contract rent ratio (MCR) of each contract and market rent in the dataset. The MCR compares the rent predicted by the model (Market Rent) with the contract rent. MCRs are calculated for each contract rent in the group and ideally, the median ratio is close to 1.00.

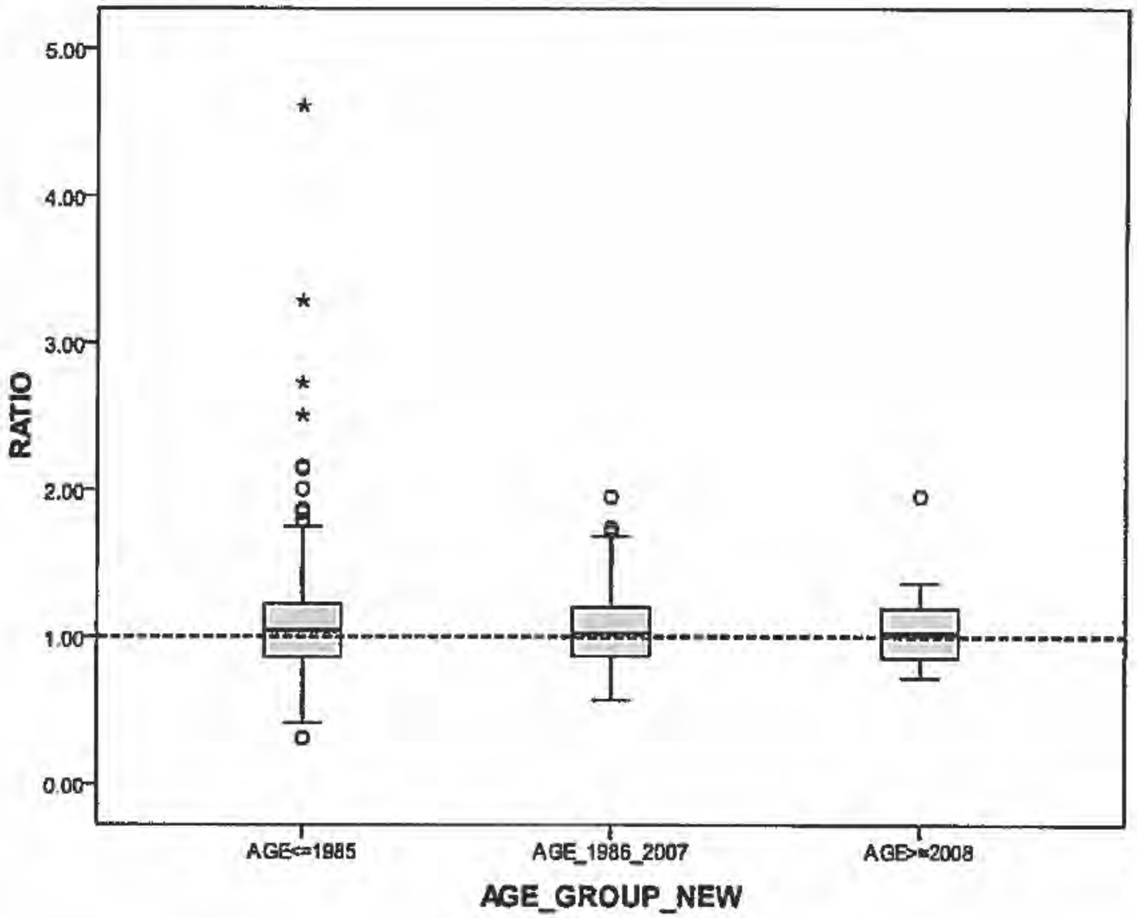
In this case, the purpose of the bias test is to show how the model represents a particular variable or, property characteristic, by demonstrating if there is over prediction (high median MCR) or, under prediction (low median MCR). A non-bias situation exists where the median MCR (middle line in the boxes) is at or near 1.00.

Box plots are helpful when reviewing non-continuous data (data stratified into specific categories). The MCR is on the vertical axis, while the variables examined are on the horizontal axis. Box plots allow for the graphical display of several statistical measures. The components of the box plots in this document display:



- The lower boundary indicates the 25th percentile, which is the value where 25 percent of the data is lower.
- The upper boundary indicates the 75th percentile, which represents the value that 75 percent of the data falls below.
- The area between the upper and lower boundary gives an indication of the spread of the middle 50 percent of the data.
- The line in the box indicates the median.
- The whiskers of the box plot are the vertical lines of the plot extending from the box, and indicate the minimum and maximum values that are not outliers in the dataset.
- Symbols are used to label outliers. The outliers are cases with the values between 1.5 and 3 box lengths from the 75th percentile or 25th percentile, while the extreme values are cases with the values more than 3 box lengths from the 75th percentile or 25th percentile.

Rent – Effective Age



Note: RATIO= Market rent / Contract rent

Three age stratifications were used in the development of the rent model. The adjustments for these stratifications reflect the market, as the median ratios indicated by the mid-points are all close to 1.00. This indicates that there is no bias in the rent model for these three age stratifications.

Vacancy

When investors purchase property, they gather information on historical and current trends and make judgements about the short to mid-term expectations. Because vacancies have a direct impact on the bottom line, vacancies are a key component of these considerations. A prudent buyer will not project near zero vacancies into the future, as it is realistic to expect that very low vacancies will spur increases in supply which will ultimately increase vacancies.

Similarly, if there is a general trend towards lowering vacancy rates, a purchaser will not base a purchase primarily on unusually high vacancies at any point in time. The above considerations essentially describe the process of stabilization. This is an area of valuation judgement.

The Assessment Branch annually sends requests for information from commercial property owners and managers. These queries include data on tenants, rents, property income and expenses, and vacancies. Both actual annual vacancies and vacant space, as of December, are asked for.

In 2007, 2008, and 2009, fully owner occupied properties were excluded from the survey. In 2010, all properties regardless of occupant type were canvassed. The response rate from our 2010 property survey by location was 93%. The response rates are very high, therefore the results are reliable.

Vacancies are employed in the income approach in order to arrive at a stabilized current estimate of net operating income. The assessor should not be overly optimistic or pessimistic in estimating a stabilized vacancy rate. It is generally acknowledged that unusually low vacancies will encourage development, and this will increase rates.

Accordingly, a minimum vacancy percentage of 2% will be used. The stabilized vacancy rates will be the greater of the average of the December 2009 and 2010 City rates, or 2%. The amounts are rounded. The stabilized rate used in the income approach is as follows:

AO Nbhd	Vacancy
30005, 30006, 30014, 30016, 30017, 30020, 30021, **30023**, 30024, 30026, 30027, 30028, **30029**, 30030, 30100, 30102, 30103, 30106, 30108, 30112, 50002	2%
30001, 30004, 30007, 30013, 30105, 30109, 30110, 30113, 50000, 70102	3%

Expenses

With net rents, the tenant is responsible for paying all the costs of occupying the property. The only expenses that are not passed on to the tenant are costs associated with periodic replacement of major building components (such as roof cover or replacement of heating equipment) and costs that cannot be passed on to tenants due to vacancy. In valuation terminology, the first is a "structural allowance", and the second is a "non-recoverable expense".

Structural Allowance

A structural allowance is normally provided in real estate valuations to account for periodic replacement of major building components. Considering the life span and cost of roof cover and heating equipment, an allowance of $0.20/ft^2 per year is applied.

Non-Recoverables

Tenants in warehouse and automotive properties typically directly pay for utility, grounds, and building maintenance costs. This means that property owners and managers of rented properties have very limited information on those expenses. In order to obtain reliable data on the full range of costs, owner occupants were surveyed for their 2010 annual operating expenses. Non-recoverable expenses may also be known as the occupancy cost.

A total of 434 properties reported a 2010 expense amount. In order to ensure that actual costs were not understated, only data where the stated amount was greater than zero was included in the analysis. The non-recoverable or occupancy costs for this property type are $5.20/ft^2.

Sales

Once market rents, vacancies, and expenses are determined, a net operating income (NOI) is derived for each property based on this analysis. The NOI is then compared to sales prices, and the sale price is expressed as a percentage of the net operating income.

Each valid sale is analyzed in this manner, then market groups defined, and a typical capitalization rate determined for each market segment. Because market rents, vacancies, and expenses are an integral part of the process, capitalization rates may vary when different market rents, vacancies, or expenses are employed.

There were a total of 104 warehouse sales that occurred between 2008 and 2010 that were used in a capitalization rate analysis. The 104 sales represent approximately 8% of the inventory.

Median Assessment to Sale Ratio (ASR)

An Assessment to Sale Ratio is a calculation comparing the assessment to the sale price for a particular property.

Assessment ÷ Sale Price = Assessment to Sale Ratio

A Median Assessment to Sale Ratio is the median Assessment to Sale Ratio found within a group. It is the middle value of the ratios when arrayed in order of magnitude. It divides the ratios into two equal groups, and is therefore only minutely affected by extreme ratios. The closer this value is to 1.00, the better. The assessment to sales ratio (ASR) study is a common statistical measure used to review the results of mass appraisal.

The goal is to achieve an ASR at or near 1.00. The IAAO Standard on ASRs is a range from 0.90 to 1.10. If the ASR falls within that range the ASR is acceptable and the goal is achieved.

If a grouping has produced an acceptable ASR close to 1.0 the same results may not necessarily be achieved by stratifying that same group differently.

Below is an example of a land market area in the City of Saskatoon for 2014 indicating the original group with an ASR of 1.00. If the grouping were restratified into two groupings based on a reasonable expectation on the influence of size, the applied base land rate and the resulting ASRs would change. The resulting ASR for the larger group of properties achieves an ASR of 1.11 which is outside IAAO standards. These results indicate that the original analysis of a single grouping is correct and achieves a better result than an alternate grouping.

Sale I.D.	Size	SP/ft^2	Asmnt/ft^2	ASR
A	5,435	101.19	60.80	0.60
B	23,530	71.61	60.80	0.85
C	*5,042*	*62.88*	60.80	0.97
D	*17,872*	*58.75*	60.80	1.03
E	17,495	30.87	60.80	1.97
F	25,278	28.48	60.80	2.13

Original above Median ASR 1.00

Sale I.D.	Size	SP/ft^2	Asmnt/ft^2	ASR
F	25,278	28.48	44.81	1.57
E	17,495	30.87	44.81	1.45
D	17,872	58.75	44.81	0.76
B	23,530	71.61	44.81	0.63

Larger group above Median ASR 1.11

Sale I.D.	Size	SP/ft^2	Asmnt/ft^2	ASR
C	5,042	62.88	82.035	1.3
A	5,435	101.19	82.035	0.81

Smaller group above Median ASR 1.06

Results Testing

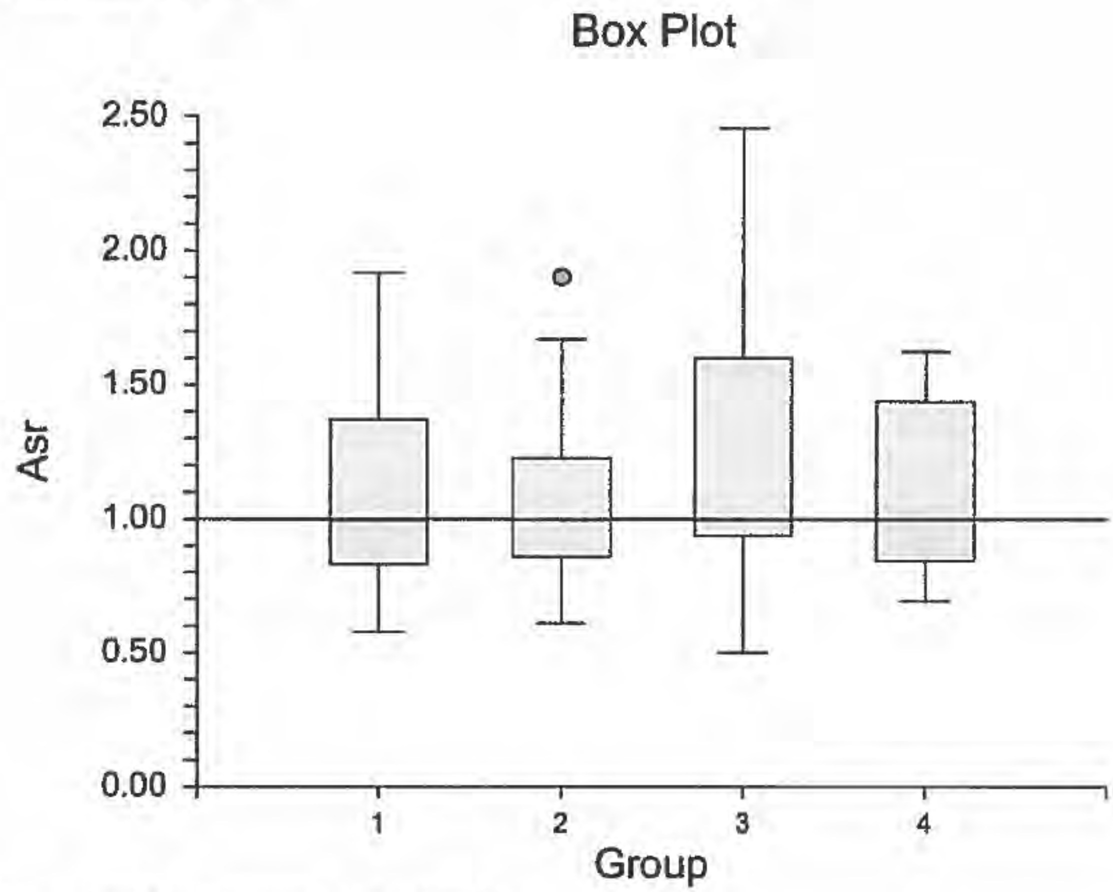
Bias testing was conducted by graphing ASRs versus different factors such as time of sale, size, age, etc. in order to check for biases. The results were satisfactory.

A scatter plot can be used as a test when the data is continuous while box plots can be used to test finite groupings of data.

The sales were stratified into four groups depending on their size and site coverage. The stratification below was applied to the subject property.

Group	# Sales	Site Coverage	Building Size	Median Cap	ASR
3	19	Greater than 0.47	<34,150	7.95%	.999
All	104	N/A	N/A	N/A	.999

Group Bias Test



Note: Group 1 Site Coverage <0.18, <34,150 ft^2

Group 2 Site Coverage 0.18 – 0.47, <34,150 ft^2

Group 3 Site Coverage >0.47, <34,150 ft^2

Group 4 >34,150 ft^2

The box plot above indicates no bias due to the city wide stratification by site coverage and size. All four groups have a mid-point close to a median ratio of 1.00 indicating no bias, and therefore reflects the market.

Certification

I certify, to the best of my knowledge and belief; that:

- The statements of fact contained in this report are true and correct;
- I have no present or prospective interest in the property that is the subject of this report and,
- The valuations and conclusions comply with the directions of *The Cities Act*, *The Assessment Management Agency Act* and the Saskatchewan Assessment Manual.

for / Les Smith, City Assessor

Appeal No.: 2015-0034; 0036;0037;0038; 0039; 0040;
0041;0043;0044; 0045;0046;0047 and 0048
YORKT-504903900; YORKT-515052500; YORKT-515053000;
YORKT-514901050; YORKT-484912650; YORKT-515054000;
YORKT-515057100; YORKT-505109000; YORKT-514802000;
YORKT-515101200; YORKT-505200050; YORKT-505201000;
YORKT-514900250; YORKT-505103000; YORKT-505104000.

SASKATCHEWAN MUNICIPAL BOARD
ASSESSMENT APPEALS COMMITTEE

Between:

McDonald's Restaurants of Canada Ltd. et al
C/O ALTUS GROUP

APPELLANT

- and -

CITY OF YORKTON
THE SASKATCHEWAN ASSESSMENT
MANAGEMENT AGENCY

RESPONDENT

WRITTEN SUBMISSION ON BEHALF OF THE RESPONDENTS

Saskatchewan Assessment Management Agency
45B Palliser Way
Yorkton, Saskatchewan S3N 4C5

Wallis Test is defined as "A non parametric test of then null hypothesis that three or more groups or subpopulations are equal centered. The test can be used to determine whether three or more property groups have equal appraisal levels. When only two groups are being compared, the appropriate test is the Mann-Whitney test."

[120] In statistics, the M-W test is used to compare the medians between two independent groups. If two property groups are being compared, a significant value of U (which is below 0.05 at 95% confidence level) could indicate that median levels differ significantly between the two groups.

[121] The K-W test is an extension of the M-W test to allow comparison of more than two independent groups. For example, if four property groups are being compared, a significant value of H (which is below 0.05 at 95% of confidence level) could indicate, at one extreme, that median levels differ between two of the groups, or, at the other extreme, that the medians of all four groups are significantly different. If hypothesis testing reject the null hypothesis, it does not necessarily mean that all compared groups have significantly different medians and should be all stratified or analyzed separately.

[122] In mass appraisal the M-W and K-W tests are excellent nonparametric tests to determine if property groups are appraised at the same percentage of market value. As found in the assessment textbook, *Fundamentals of Mass Appraisal (2011), IAAO, Chapter 10 Evaluating the Reliability of Ratio Study Statistics, Section - Testing Horizontal Equity, p.*379-385 in Appendix F, these tests are recommended to be used on ratio studies to compare the medians of assessment to sale ratio of different property groups. It is not typical that these tests are to be used for assessment model specification or model building.

[123] Assessment appraisers should not solely rely on or use the M-W, K-W, including Analysis of Variance (ANOVA) and t-test for model stratifications or model building.

[124] First of all, these tests deny any trend or spline relationships that are appropriate for quality class linearization, effective age adjustment or time adjustment. The trend and spline relationship can be easily detected by scatter plots, box plot or properly formulated regression model.

the underlying data. Also called "distribution-free statistic."" - *Mass Appraisal of Real Property, IAAO, p. 384.*

[153] The warehouse group is entered in the regression model as base occupancy type and Year Group 1 is entered in the regression model as base year group. The independent variables of office group and year group 3 are the significant variables affecting the variation of cap rate. The significance level of office group and year group 3 are below 0.05.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.130	.011		11.942	.000
	office_grp	-.044	.012	-.512	-3.648	.001
	retail_grp	-.022	.012	-.273	-1.752	.087
	restau_grp	-.001	.018	-.004	-.028	.978
	yrgrp_2	-.014	.011	-.196	-1.291	.204
	yrgrp_3	-.040	.015	-.370	-2.677	.011

a. Dependent Variable: Cap_Rate

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.115	.006		20.763	.000
	office_grp	-.037	.011	-.436	-3.415	.001
	yrgrp_3	-.034	.014	-.309	-2.423	.019

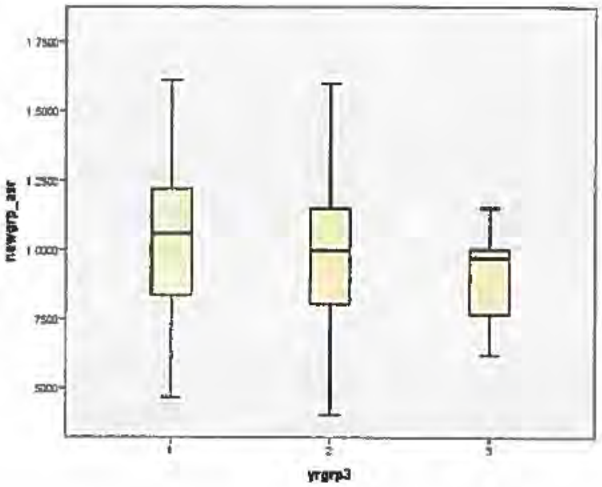
a. Dependent Variable: Cap_Rate

[154] The multiple regression analysis with the two significant variables noted above can provide a predication of the cap rates of the 3 cap rate groups (office group, non office group with year built of 1984 and older, and non office group with year built of 1985 and newer) based on minimizing the model residuals.

[155] Since there is only **one** office sale in the year group 3, it could be combined with the other office sales as a separate group, then a split by year for the other occupancy types. The assessment model would be as shown below:

	Count	Median Cap Rate	Median ASR	ASR COD
Office	11	7.10%	1	17.0%
Non Office (1984 and older)	32	11.60%	1	24.8%
Non Office (1985 and newer)	5	8.50%	1	15.4%
Overall	48		1	22.0%

[156] With the application of the above 3 cap rate groups (office, non-office 1984 and older and non-office 1985 and newer) stratification, the K-W test supports that the median ASR of the 3 years groups are now equal.



[157] K-W test of NEWGRP_ASR by 3 Year Groups:

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of newgrp_asr is the same across categories of yrgrp3.	Independent-Samples Kruskal-Wallis Test	.811	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[158] With the application of the above 3 cap rate group stratification, the K-W test supports that the median ASR of the occupancy types are now equal.