

L'utilisation de graphiques dans le processus d'évaluation



Par George Canning, AACI, P.App, Canning Consultants Inc., London ON

Introduction

Les praticiens de l'immobilier utilisent plus d'aides visuelles que jamais auparavant pour décrire les données dans leurs rapports d'évaluation et de consultation. La connaissance des graphiques est indispensable à leur bonne utilisation.

Il existe sur le marché de nombreux programmes statistiques capables de produire un large éventail de graphiques, y compris une analyse de données perfectionnée comme SPSS, Mini Tab et 'R', pour n'en nommer que quelques-uns. Certains de ces programmes, notamment 'R', sont gratuits. L'utilisation de tous ces programmes statistiques exige des connaissances fondamentales en statistiques. Deux livres primés sur ce sujet sont Head First: *Data Analysis* et *Statistics* par Michael Milton et Dawn Griffiths.

Les graphiques sont très utiles pour dégager des modèles dans les données immobilières, mais très souvent seulement à la surface. Dans la plupart des cas, le vrai message des données est enfoui profondément dans les chiffres. Étant donné que les services d'évaluation et de consultation dans le domaine immobilier SONT ESSENTIELLEMENT UNE AFFAIRE D'ANALYSE DES DONNÉES, le conseil à donner à tout praticien de l'évaluation est d'entreprendre immédiatement le processus d'apprentissage des données. Des outils sont disponibles à des coûts relativement bas. Tout ce dont vous avez besoin, c'est le désir de commencer à explorer vos données.

Organiser vos données

Avant d'aborder la question des graphiques, il convient d'examiner les données immobilières et la façon de les organiser. Les données ne sont pas parfaites parce qu'elles ne se présentent pas dans un format que tout professionnel de l'immeuble peut suivre et comprendre. La première étape consiste donc à disposer les données sur un tableur. Peu importe que les données concernent

5 ventes ou 500 ventes, le processus est toujours le même. Tout cela fait partie de ce qu'on appelle le « nettoyage de vos données ». Au fil des ans, les évaluateurs collectent des données de vente de propriétés de types très divers. Ces données sont conservées dans des livres ou dans un document de traitement de texte. Dans la plupart des cas, les données deviennent redondantes dès que le rapport d'évaluation est envoyé au client. À moins que ces données soient transférées sur un tableur, leurs avantages graphiques et statistiques demeurent inexploités.

Un bon point de départ consiste à organiser les données par numéro de vente, date de vente, prix demandé, adresse, prix de vente (par unité de mesure), âge du bâtiment, superficie du terrain, zonage, façade du terrain, taux de capitalisation et taux GIG, qui sont les attributs de nom de propriété les plus courants. Il n'y a pas de règles absolues, sauf celle de consigner les informations de base pour chaque vente.

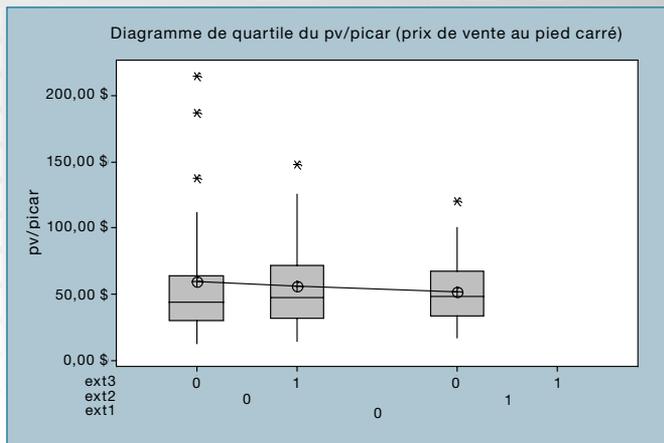
Les données de vente s'expriment généralement de deux façons : quantitativement et qualitativement, ce qui représente deux notions distinctes. Les données quantitatives concernent essentiellement la taille d'un bâtiment ou d'un terrain. Les données qualitatives sont une valeur attribuée à l'état d'un immeuble ou à son emplacement. Les données qualitatives peuvent être exprimées selon une échelle nominale telle 1 = passable, 3 = moyen et 5 = bon. On peut aussi utiliser une variable nominale ou discrète telle désignée plaine inondable = 1, non désignée plaine inondable = 0. On convertit les données qualitatives en chiffres ou notes parce que les ordinateurs peuvent traiter seulement un chiffre ou un chiffre assigné à une condition ou à un type. Le fait de convertir les données en un chiffre ne diminue aucunement l'importance de la variable, il établit simplement des catégories. Une note de 1 n'est pas 2 fois moins qu'une note de 3. C'est simplement une affaire d'appartenance.

Les données quantifiables sont plus pertinentes que les données qualifiables parce qu'on a affaire à un chiffre spécifique. L'utilisation des qualificatifs « passable » ou « bon » est subjective parce que chaque expert de l'immobilier peut avoir sa propre opinion quant à ce qui est passable ou bon. Oui, des notes peuvent être attribuées à ces mots aux fins de traitement informatique, mais ces notes ne sont pas aussi spécifiques qu'un chiffre donné. Cependant, on a besoin d'un système pour mesurer les variables de données qui ne sont pas absolues.

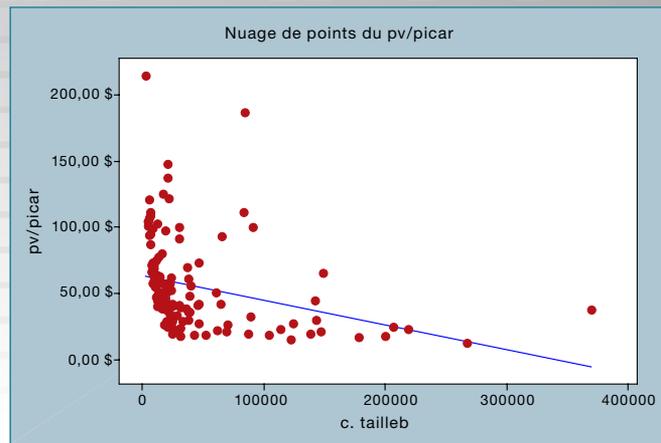
On fait souvent référence au prix moyen des données dans des documents d'évaluation. Cependant, une moyenne ne dit pas grand-chose parce qu'on ne connaît pas la dispersion autour de cette moyenne. La mesure qui nous précise cette dispersion est ce qu'on appelle l'écart-type, c'est-à-dire la distance par rapport à la moyenne. Plus la valeur de l'écart-type est petite, plus la distance par rapport à la moyenne est petite. Il peut être utile d'exprimer la tendance centrale au moyen du mode (c.-à-d., le chiffre le plus fréquent dans l'ensemble des données) ou de la médiane (c.-à-d., le chiffre au milieu de l'ensemble des données) des différences autour de la moyenne, mais leur valeur est limitée.

Les diagrammes de quartiles sont utiles quand on examine des moyennes entre groupes de données. Le graphique 1 est un diagramme de quartiles représentant trois différents types de revêtement extérieur de bâtiments industriels. Extérieur 1 regroupe les bâtiments dont le revêtement extérieur est en acier émaillé; extérieur 2 regroupe les bâtiments dont le revêtement extérieur est un mélange de maçonnerie et d'acier émaillé; extérieur 3 regroupe les bâtiments en maçonnerie. Le graphique 1 est une expression de trois différents types de finition extérieure utilisant des diagrammes de quartiles.

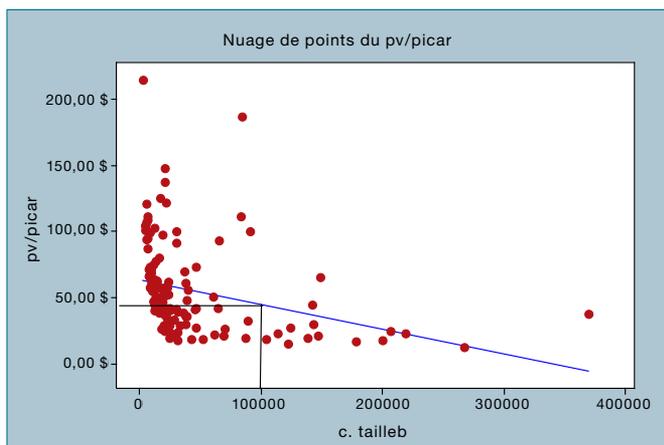
Le rectangle central, représentant un revêtement extérieur mixte, a un prix moyen de 60,00 \$. Le rectangle,



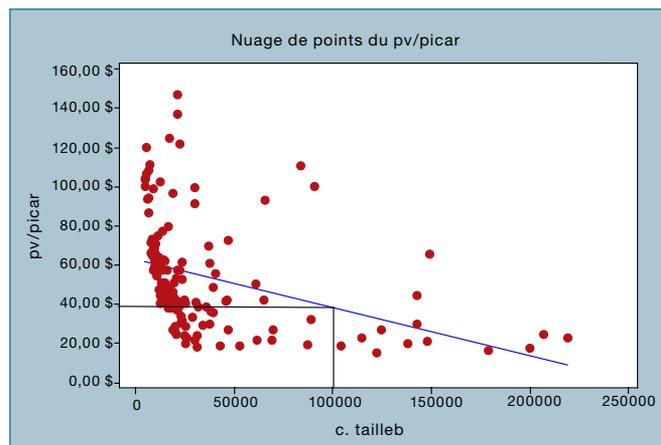
Graphique 1



Graphique 2



Graphique 3



Graphique 4

représentant un revêtement en maçonnerie, a un prix moyen de 62,50 \$. L'autre rectangle, représentant un revêtement en acier émaillé, a un prix moyen de 53,00 \$. L'utilisation de ces diagrammes de quartiles permet de conclure que les bâtiments à revêtement maçonnerie ou maçonnerie acier se vendent environ 8,25 \$ plus cher que ceux à revêtement en acier émaillé.

Ceci est un bon exemple des problèmes qui résultent de l'utilisation des moyennes seulement. On ne peut pas déclarer que la somme de 8,25 \$ est un chiffre fiable parce qu'on ne connaît pas le rôle statistique des autres variables dans l'ensemble des données. Dans le présent article, nous allons revenir sur cette question pour voir si la somme de 8,25 \$ a quelque validité.

Les diagrammes de quartiles peuvent nous en dire beaucoup sur la distribution. La ligne horizontale représente le prix médian des données, alors que le symbole relié à une ligne représente la moyenne. Quand la moyenne est au-dessus de la ligne de la médiane, on dit que les

données sont « désaxées vers la droite ». La ligne horizontale du fond est appelée Q1 et représente la valeur du milieu de la première moitié de la liste de données, alors que Q3 représente la valeur du milieu de la seconde moitié de la liste de données. Les moustaches représentent la largeur générale des données, alors que les étoiles situent les valeurs aberrantes ou au-delà de la gamme normale des données.

Pour réitérer ce que nous avons dit au sujet des graphiques, ce qui apparaît comme une tendance sur l'aspect visuel des données n'est pas nécessairement ce qui se produit réellement. Il en est ainsi parce que la plupart des données ne représentent que deux variables alors que d'autres variables peuvent intervenir dans l'explication du prix. Parfois, les données doivent être converties en format logarithmique. La raison étant qu'une variable donnée peut, par exemple, faire ressortir un modèle plus uniforme quand on convertit les données de pintes à litres, ou de litres à chopines. Les données ne sont pas modifiées, elles

sont tout simplement exprimées dans un format différent.

Nuages de points

Les nuages de points résultent tout simplement des données mesurées par deux points. Ils sont très utiles parce qu'ils permettent de visualiser très rapidement les tendances. Parfois, les tendances ne sont pas évidentes. Si c'est le cas, essayez de tourner le graphique sur le côté pour voir si une tendance se dessine dans cette perspective.

Les nuages de points sont constitués de points entrés sur un graphique à coordonnées X et Y.

Le graphique 2 est un nuage de points entre la taille d'un bâtiment industriel et le prix de vente au pied carré du bâtiment. Des économies d'échelles sont appliquées, selon lesquelles plus un immeuble est grand, plus son prix de vente au pied carré diminue. C'est un excellent renseignement à avoir à la lumière du processus d'ajustement de la MCD. On doit cependant procéder avec prudence. Ce graphique ne peut pas se substituer à la valeur.

Le graphique 3 est le même nuage de points servant à évaluer un bâtiment industriel de 100 000 pieds carrés. On trace une ligne vers le haut à partir de l'axe des X (à 100 000 pieds carrés) jusqu'à l'intersection de la ligne de régression simple (en bleu). On prolonge ensuite cette intersection vers la gauche jusqu'à l'axe des Y (prix de vente au pied carré du bâtiment). Selon cette illustration, le bâtiment de 100 000 pieds carrés vaudrait 49,00 \$ le pied carré.

Le problème vient du fait que la taille du bâtiment ne représente qu'une variable. Il y a d'autres variables qui pourraient influencer sur le prix de vente au pied carré d'un bâtiment. On doit aussi tenir compte des valeurs aberrantes qui peuvent avoir une influence sur la ligne de régression simple en bleu en tirant positivement ou négativement sur les données générales du nuage. En retirant tout simplement les valeurs aberrantes, on change la valeur du bâtiment de 100 000 pieds carrés, comme le montre le graphique 4. Maintenant, le bâtiment de 100 000 pieds carrés ne vaut que 40 \$ le pied carré.

La morale de cette histoire est qu'on ne doit pas utiliser les nuages de points pour déterminer la valeur. Utilisez-les plutôt comme aides visuels pour dégager des modèles de données. Pour utiliser les données aux fins d'évaluation, on doit établir un modèle de régression approprié.

Histogrammes

Les histogrammes sont des décomptes qui ont été placés dans un groupe commun. Ils sont très utiles, surtout si on insère une ligne représentant la forme des données. Les graphiques 5 et 6 représentent un histogramme des prix de vente. Le graphique 5 présente les données telles qu'elles apparaissent à l'état brut. Ces données s'étalent loin vers la droite. Cependant, une fois qu'elles ont été transformées (graphique 6), les données sont plus uniformes et plus proches d'une courbe normale. La transformation consistait à convertir en valeurs logarithmiques.

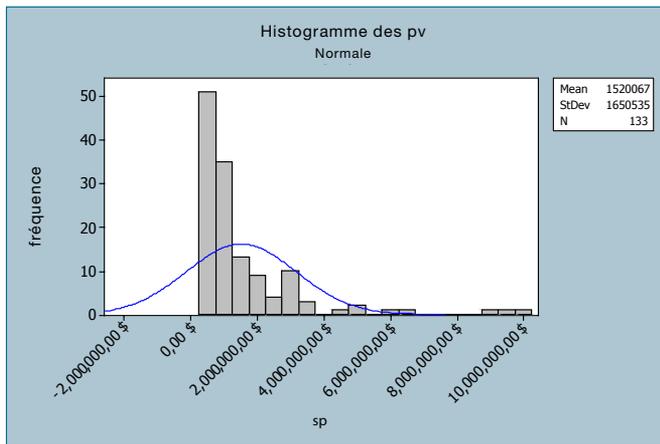
Le graphique 7 est un histogramme des tailles des bâtiments industriels. Que

voyons-nous? Un désaxement vers la droite. Après conversion en log, les tailles des bâtiments prennent une forme très proche d'une courbe normale, comme le montre le graphique 8.

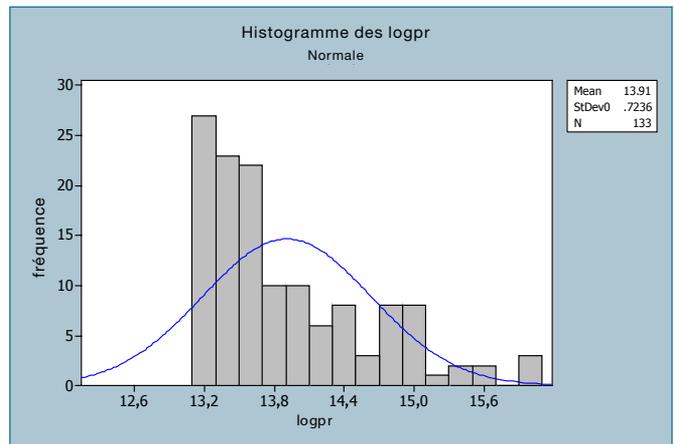
Les avantages de transformer les données se situent dans l'utilisation d'une analyse de régression multiple, puisque l'intention consiste à ajuster les données à un modèle linéaire et à retirer l'aberration des données. Plus les données se rapprochent d'une courbe normale, meilleurs les résultats seront dans le modèle de régression. La courbe normale fut conçue pour la première fois au 18e siècle par Abraham De Moivre, un statisticien et conseiller aux adeptes du jeu. La courbe normale ou courbe de Gauss présente d'étonnantes propriétés de prédiction qui bénéficient grandement aux joueurs de dés et de cartes.

Graphiques à barres

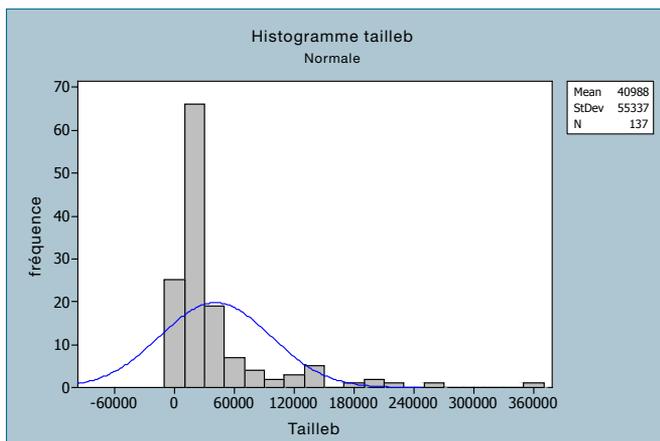
Les graphiques à barres sont des graphiques dont les barres peuvent être verticales ou horizontales. Le graphique



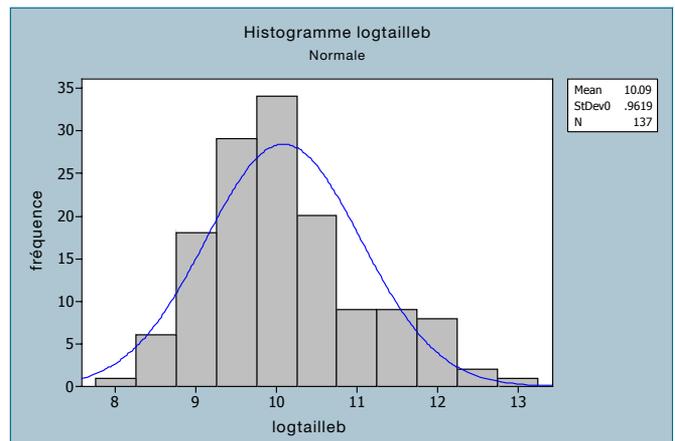
Graphique 5



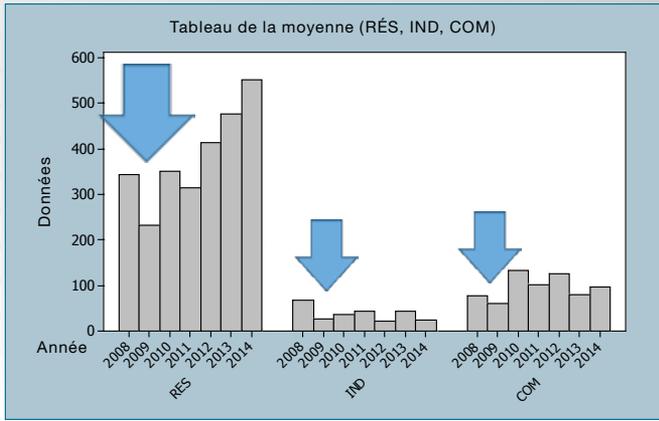
Graphique 6



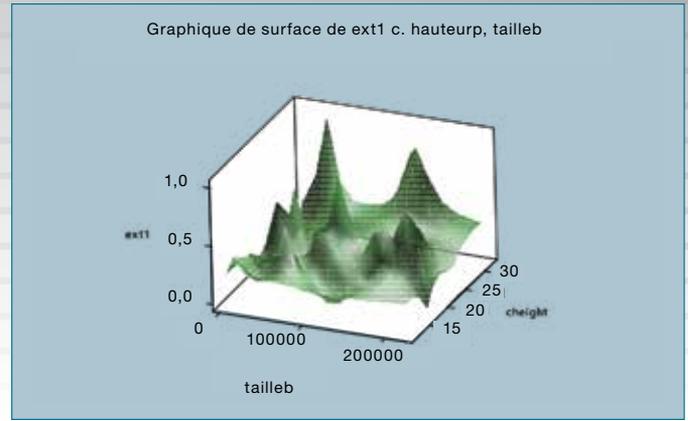
Graphique 7



Graphique 8



Graphique 9



Graphique 10

à barres est un moyen rapide et efficace de dégager des modèles dans les données. Le graphique 9 est un graphique à barres de la valeur, exprimée en dollars, des permis de construction des trois plus importantes catégories de bâtiments (c.-à-d., résidentiels, industriels et commerciaux).

À première vue, on voit un robuste marché d'aménagement résidentiel, une très faible valeur monétaire des permis de construction délivrés dans le secteur industriel et un secteur commercial et de vente au détail plus vigoureux.

Cependant, la valeur réelle dans le graphique est indiquée par les flèches bleues. Notez l'effet de 2008 (c.-à-d. l'effondrement du marché bancaire mondial) dans les trois catégories de bâtiments. C'est un renseignement important dans la préparation d'une évaluation rétrospective d'une propriété en 2009. Il est intéressant de noter que cette tendance s'est manifestée partout en Ontario, sauf dans la RGT (Région du grand Toronto).

Graphiques de surface

Les graphiques de surface sont des types de graphiques de haute technicité qui peuvent être utiles quand on étudie les relations entre diverses variables. Les graphiques de surface exigent qu'une variable soit catégorique et que les deux autres soient numériques.

Le graphique 10 est un graphique de surface de l'extérieur de bâtiments industriels par rapport à la taille des bâtiments et à la hauteur des plafonds. Il peut montrer les relations qui sont couvertes par chacune des variables.

Graphiques chronologiques

Le graphique chronologique illustré au graphique 11 montre les prix de vente au

piéd carré d'un bâtiment de 2005 à 2014 dans une ville donnée. L'index 1-130 ne représente pas des années, mais plutôt une période d'ensemble. Par exemple, 1-13 pourrait couvrir la période depuis le début des données, disons le 1er juin 2005 jusqu'au 17 avril 2006, date du début d'une autre période.

Les graphiques chronologiques sont une façon intéressante de démontrer la volatilité d'un type de données spécifiques. En examinant le graphique 11, quel modèle peut-on dégager des crêtes et creux des données? La réponse à cette question est qu'il y a des périodes de hausse des prix et des périodes de fléchissement du marché industriel. Ces tendances à la hausse et à la baisse ne reflètent pas seulement le temps, mais aussi d'autres variables comme la superficie du terrain, la taille du bâtiment et tout un ensemble d'attributs qui pourraient contribuer à expliquer les niveaux de prix. Ce phénomène est résolu par une ARM (analyse de régression multiple).

Faire rapport des tendances au moyen de graphiques

Les graphiques utilisés dans les rapports n'ont aucune valeur si on ne dégage pas ce qu'ils peuvent révéler au sujet des données. Les lecteurs des rapports ne sont pas des mentalistes. Ils se fient à l'évaluateur pour comprendre le comportement du marché. Les graphiques ne devraient pas figurer dans un rapport pour bien paraître, ils devraient y être pour informer. Si un graphique indique un marché stagnant, on devrait attirer l'attention d'un client qui s'apprête à entreprendre des négociations avec un acheteur.

Le graphique 12 est un nuage de points du nombre de nouvelles maisons construites dans une petite collectivité. Le graphique indique une tendance vers le bas

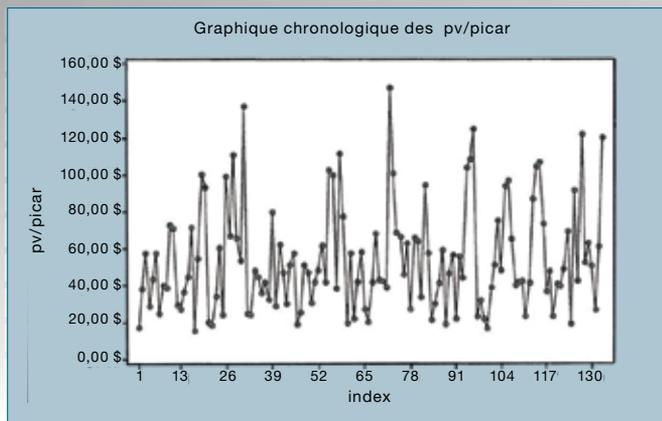
des nouvelles constructions résidentielles. Cependant, le graphique pourrait aussi faire ressortir le fait que la capacité du système d'égout était limitée en 2013 et 2014. Il pourrait souligner la demande refoulée de logements qui attend que la municipalité ajoute une nouvelle unité de traitement à son installation – précieuse information si une évaluation est effectuée sur un projet d'aménagement dans cette région.

Quand les graphiques ne disent pas tout au sujet des données

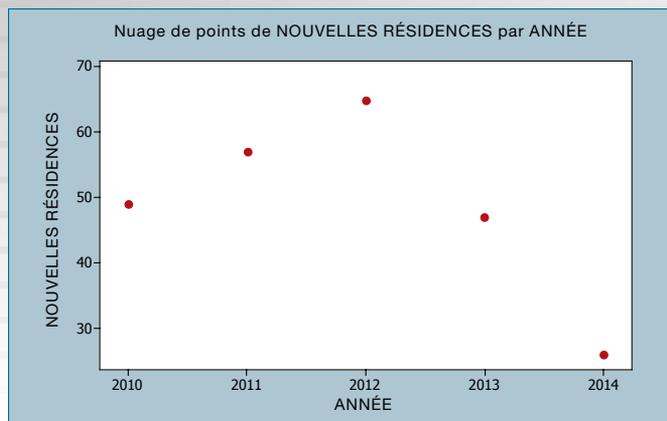
Les graphiques représentent généralement un espace bidimensionnel opposant une variable à une autre tel le prix par pied carré de logement. Les graphiques peuvent être révélateurs quand ils présentent une valeur nominale pour des aspects tels les valeurs des permis de construction, les taux de chômage et la quantité d'espace commercial absorbée dans une collectivité donnée. En pareils cas, l'information livrée par les graphiques va droit au but et est absolue.

Par contre, on ne peut se fier à l'information livrée par un graphique quand on utilise deux variables pour prédire une valeur. Le problème réside dans le fait que le secteur immobilier est généralement affecté par un bon nombre d'aspects/propriétés caractéristiques/variables. Ces variables peuvent même être en interaction les unes avec les autres. Comment peut-on alors savoir quelle variable est importante dans le processus d'évaluation pour expliquer les prix dans un ensemble de données déterminés? La seule façon de le savoir avec certitude consiste à soumettre un modèle de régression à un programme statistique.

Plus tôt dans cet article, nous avons démontré l'utilité d'un diagramme de quartiles et le graphique faisait ressortir une



Graphique 11



Graphique 12

différence dans les prix de vente des bâtiments industriels à finition de maçonnerie et ceux à revêtement d'acier émaillé.

Tableau 1 donne les résultats d'une ARM appliquée aux propriétés industrielles de London, en Ontario. Il n'est pas nécessaire d'expliquer tous les résultats, sauf pour ceux d'Extérieur 2 et d'Extérieur 3. Les résultats de ces variables sont en caractères gras pour mieux faire ressortir l'illustration.

Premièrement, qu'est-il arrivé à Extérieur 1? Rien, sauf que cette variable a été intégrée au programme Minitab pour la comparer aux deux autres. Les chiffres à l'extrême droite représentent les valeurs P ou les probabilités que les chiffres ou coefficients 0,05703 et -0,03458 ont été déterminés ou non de façon aléatoire ou que ces chiffres ou coefficients représentent une association avec les prix de vente des comparables. Étant donné que ces valeurs P sont très élevées, on ne peut s'y fier pour expliquer les différences de prix de vente des bâtiments industriels ayant différents types de finition extérieure. La deuxième colonne représente l'erreur-type. Ces chiffres sont plus grands que les coefficients, ce qui indique également qu'on ne peut se fier aux coefficients pour expliquer les différences de prix résultant des différents types de finition extérieure.

En définitive, le diagramme de quartiles a indiqué une différence de prix, mais en se fondant uniquement sur les prix de vente moyens de chaque catégorie de type de bâtiments industriels. Quand toutes les variables ont été considérées pour expliquer le prix, on a constaté que la différence entre Ext 2-3 n'était pas significative. Cela ne veut pas dire que Ext 1-3 ne devrait jamais être exclu de futures applications de régression quand de nouvelles données sont ajoutées. Tout ce que cet ensemble de données nous

Predictor	Coef	SE Coef	T	P
Constant	5,361	3,108	1,72	0,088
LotS Log	0,31513	0,06970	4,52	0,000
logbage	-0,36141	0,05084	-7,11	0,000
logcheight	0,2504	0,1688	1,48	0,141
logdmroad	-0,03883	0,02786	-1,39	0,166
log401	-0,17265	0,04883	-3,54	0,001
logumrate	-0,1422	0,2151	-0,66	0,510
logcdndol	0,0451	0,4986	0,09	0,928
logbuildsize	0,39254	0,06765	5,80	0,000
date	0,00014153	0,00003519	4,02	0,000
loc	-0,02633	0,04222	-0,62	0,534
fplain	-0,13077	0,09540	-1,37	0,173
zone	-0,05566	0,04453	-1,25	0,214
eland	-0,03115	0,07745	-0,40	0,688

ext3	0,05703	0,08798	0,65	0,518
ext2	-0,03458	0,09056	-0,38	0,703

S = 0,334323	R-Sq = 84,1 %	R-Sq(adj) = 81,7 %
--------------	---------------	--------------------

Tableau 1

permet de dire, c'est qu'on n'a trouvé aucune preuve que Ext 1-3 se reflète dans le prix de vente des données. Il se pourrait que l'ensemble du marché de l'immobilier industriel fût en période de dépression et que les acheteurs ne fussent pas prêts à payer la différence pour la finition extérieure. On ne sait tout simplement pas.

Conclusion

Les évaluateurs immobiliers qui désirent utiliser des graphiques dans un rapport devraient s'acquitter de leur « obligation de diligence » en abordant la question sous diverses optiques :

1. Est-ce que le graphique contribue, dans son ensemble, au processus d'évaluation?
2. Utilise-t-on le bon type de graphique?
3. Les données présentées sous forme de graphique sont-elles bien expliquées au lecteur?

4. La source des données présentées sous forme de graphique est-elle fiable?

Comme l'économiste américaine Emily Oster l'a si bien dit : « Les meilleurs moments se produisent quand les résultats d'un graphique ou d'une analyse statistique vous sautent aux yeux – quand vous réalisez que vous savez quelque chose que personne d'autre ne sait et que vous avez le plaisir de réfléchir sur la façon de leur communiquer ce que vous avez découvert. »

Bibliographie

Stephen Few : *Show Me the Numbers: Designing Tables and Graphs to Enlighten*
 Edward Tuffe : *The Visual Display of Quantitative Information*
 Naomi B. Robbins : *Creating More Effective Graphs* 