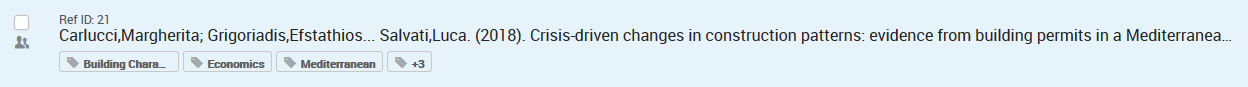
****

**Crisis-driven changes in construction patterns: evidence from building permits in a Mediterranean city**

* construction sector's response to the 2007-2008 recession based on the spatial analysis of 10 building activity indicators over a 25-year period (1990-2014) in Athens, Greece.
* Data were derived from a survey by the Hellenic Statistical Authority (ELSTAT) in co-operation with local municipalities.
* The collected data were analyzed using a three-part data mining framework:

(i) descriptive statistics

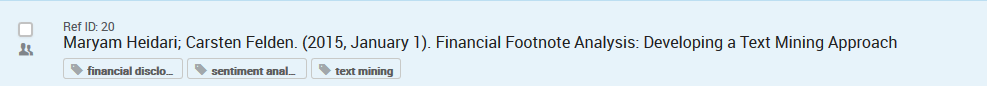
* The average number of floors per new building increased before the crisis, reaching a peak in the 2005-2009 period (2.5 floors) and decreasing moderately in recent years (2.3 floors).
* The average surface area per new building followed a similar pattern, decreasing by nearly 100 m2 per building in the crisis period.
* The highest number of building enlargements was observed in the first time interval, decreasing slightly afterwards. However, both the number of new floors in enlarged buildings and the average surface area of enlargements increased moderately over time, reaching a peak in 2005-2009,
* The percentage of small dwellings was stable (around 20%) up to 2000-2004, increasing rapidly thereafter.
* The ratio of building permits to resident population was quite stable (around 2 permits per 100 inhabitants), then declined 10-fold during the crisis. Taken together, descriptive statistics illustrate the impact of the 2007-2008 crisis on building activity with respect to previous expansion and recession cycles.

(ii) inferential approaches

* According to Wilcoxon tests, the average number of floors per new building is the only variable showing significant changes over the four decades studied.
* Seven of the 20 indicators differed significantly in distribution across the first (1990-1999) and second (1995-2004) decades, five in the third decade (2000-2009) and seven in the most recent decade (2005-2014).
* The percentage of small-size dwellings and the ratio of new building surface area to population increase showed the opposite spatial pattern, decreasing and increasing in concordance over time, respectively.

(iii) multivariate analysis.

* The MFA extracted four components explaining 57.5% of the total variance.
* Component 1 (27.4%) illustrates the mono-centric structure prevailing in Athens, being associated to 5 indicators in 1990-1994 and to 3 indicators in 2010-2014. Density of both new buildings and enlarged buildings, and the average number of floors per building showed positive and stable (or increasing) loadings to component 1 over the whole study period (Table 4).
* Component 2 (13.5%) was associated positively with the average surface area of new buildings over the entire study period. A negative loading was assigned to the density of new buildings for 2010-2014. Component 2 scores showed a moderate east-west gradient (particularly evident in the decade 1990-1999), and were correlated positively with average per-capita declared income, percentages of multiple-use buildings and of service/commercial buildings, population density and diversity in land-use.
* Component 3 (9.4%) was associated positively with the average surface area of enlarged buildings for all time intervals except 2010-2014. The percentage of small-size dwellings received a negative loading during the 2000-2004 'building boom'. High positive scores were concentrated in the western fringe of Athens.
* Positive loadings to component 4 (7.1%) were assigned to the ratio of new building surface area to absolute population increase and to the number of building permits released in the last decade. Component 4 score mapping illustrated a peri-urban spatial pattern consolidating over time, with higher scores observed in western and southeastern fringe municipalities.



What

* This study tries to make a contribution regarding financial analysts and any stakeholders who benefit from financial analysis by applying a text mining approach based on pre-defined classes in order to assist them in required financial information extraction from footnotes in a real time and automatic fashion.

How

* Concerning text mining application for financial footnotes analysis, we applied the Cross-Industry Standard Process (CRISP) process flow to define a complete lifecycle of the text mining workflow [7]. CRISP methodology is based on six phases providing a comprehensive coverage of all activities involved in data/text mining projects [21].
* The used database to get financial footnotes of companies' filings was Edgar online of U.S. SEC1. This database provides free public access to corporate information, allowing quick research on a company's financial information by reviewing registration statements and periodic reports filed on forms like 10-K (annual reports) and 10-Q (quarterly reports).

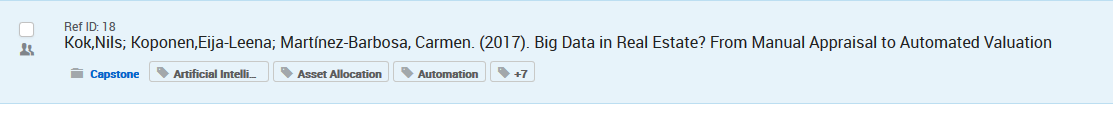
Conclusion

Despite significant developments in fields of financial analysis e.g. based on XBRL-formatted data, the textual parts of financial reports, which are critical for comprehensive financial analysis, are still dependent on time-consuming manual procedures.

Comparing to existing approaches in terms of financial footnote analysis, our preliminary results show that text mining could be an appropriate semi-automatic solution to facilitate manual analysis of unstructured parts of financial reports. As a matter of fact, the text mining approach helps users to access required soft information as a separate sentence based on each financial pre-defined category.

It is also of interest in this research to develop this solution by adding more capabilities thereby to map extracted sentences into related figures in financial statements. However, due to some limitations, it is not practically implemented, yet.

One of the limitations is that footnote parts are normally received by analysts or auditors as separate documents and are not attached to main financial statements. This makes it difficult to map financial sentences into figures in financial statements.



What/How

Real estate is the third-largest asset class for institutional investors, but determining the value of commercial real estate assets remains elusively hard. In this article, the authors provide a practical application of big data by employing a unique set of data on U.S.

In this study, we provide a practical application of*big data* in combination with sophisticated modeling techniques to develop an automated, machine-based valuation model for the commercial real estate sector. We focus on the multifamily sector, enabled by access to a dataset of some 54,000 U.S. multifamily assets. This dataset is enriched by a wide set of both standard demographic and economic measures and more modern, hyperlocal metrics, such as proximity to music events, bars and restaurants, and green space and local crime incidence.

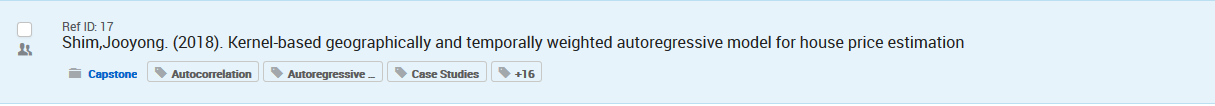
Conclusion

We find strong evidence on the superiority of automated valuation models (AVMs) over traditional appraisals--the median absolute error of the automated model we develop is 9%, which compares favorably against the accuracy of traditional appraisals, and the model can produce an instant value at every moment in time at a very low cost.

Previous research has showed that the difference between appraisals and transactions ranges from 10% to 15% (Cannon and Cole [2011]; MSCI [2016]). In addition an appraisal typically takes about three weeks to obtain and costs some USD 3,000-5,000 for the average commercial asset.

52% of automated valuations are within 10% of the actual transaction price, and 92% of automated valuations are within 30% of the actual transaction price.

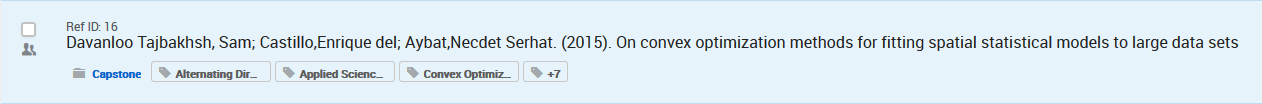
AVMs can provide an instant indication of property value, which saves significant time and resources for portfolios of both investors and lenders, as well as those interested in a single property.



<http://web.b.ebscohost.com.proxy.libraries.smu.edu/ehost/pdfviewer/pdfviewer?vid=1&sid=00811470-5484-478f-8bd4-abd153e2c311%40sessionmgr103>

This one was wayyyyyy too technical and super mathy so it was very difficult to understand/extract what they were doing.

Does have some cool graphs though that we might be able to use but overall its a lot.



<https://etda.libraries.psu.edu/files/final_submissions/11344>

A first goal of this dissertation is to propose an efficient algorithm that fits asecond-order stationary and isotropic GRF model to large data sets. The dominantfrequentist approach is Maximum Likelihood (ML) which requires non-convexoptimization.

The problem is aggravated in big data settings, given theO(n3)computational complexity of the ML method, wherendenotes the number ofspatio-temporal data points.

An extension of the proposed algorithm for prediction problemsunder a class of continuous, non-Gaussian data known to have a nonparanormaldistribution is studied.

GRF models are very popular for predictive modeling and also for classificationwhere the underlying data sets posses spatial/temporal dependencies. However,fitting these models is troublesome given the nonconvexity of their likelihood and the computational complexity of the Maximum Likelihood (ML) method especially in big n scenarios

The stage-one probleminvolves solving a weighted`1-norm penalized loglikelihood optimization where theweights are the pairwise distances of the spatial locations of observations. This101

problem is solved by the Alternating Direction Method of Multipliers (ADMM)which is shown to converge linearly given the structure of the problem. The secondstage involves a least-square problem which is solved sequentially. Furthermore,given an important class of exponential covariance functions, it is shown that thesecond stage problem is also convex. Finally, theoretical results provide convergencebounds for the optimal estimates of the first and second stage problems. Thesetheoretical results are supported with numerical simulation studies.