Aydogan, Y. (2023). The curse of sluggishness: rethinking firm entry and exit with machine learning. Gümüşhane Üniversitesi Sosyal Bilimler Dergisi, 14(3), 1036-1044.

The Curse of Sluggishness: Rethinking Firm Entry and Exit with Machine Learning

Ağırkanlılığın Uğursuzluğu: Firma Giriş ve Çıkışını Makine Öğrenmesi ile Yeniden Düşünme

Yigit Aydogan [1]

Abstract

The main contribution of this research lies in identifying a crucial insight: the slow growth of new firms in local economies may be attributed to a self-sustaining mechanism characterized by volatile influx of new firms. In other words, regions with lower long-term entry rates exhibit higher relative volatility in this aspect. A similar argument can be made for exit rates as well. To categorize spatial units in economics, Machine Learning algorithms can be utilized. In this study, Turkish cities were clustered based on firm dynamics data spanning from 2009 to 2020. Through the implementation of an Unsupervised Learning (k-means) algorithm, four clusters were identified based on entry rates, while six clusters were identified based on exit rates. This approach represents an improvement over traditional methods that often require extensive manual effort to incorporate numerous socioeconomic variables into a criterion. Furthermore, it helps reduce subjectivity inherent in such methods, which heavily rely on qualitative analyses. The proposed method empowers policymakers to obtain groupings that align with their economic objectives and foster policy success.

Keywords: *Firm Dynamics, Machine Learning, Unsupervised Learning, Clustering, k-means Algorithm*

Öz

Bu çalışmanın en önemli bulgusu çok önemli bir anlayışı tanımlama üzerinedir: yerel ekonomilerde yavaş yeni firma oluşumu, istikrarsız yeni firma girişleri açısından olası bir kendi kendini sürdürebilir mekanizma ile birliktedir. Başka bir deyişle, uzun vadede girişlerin az olduğu yerel ekonomiler, bu istatistik açısından daha fazla göreli dalgalanma yaşarlar. Aynı argüman çıkışlar için de geçerlidir. Ekonomide mekansal birimlerin sınıflandırılması, Makine Öğrenimi algoritmaları kullanılarak yapılabilir. Bu çalışmada, 2009-2020 yılları arasındaki firma dinamikleri verileri kullanılarak Türk şehirleri kümelendirilmiştir. Bir Denetimsiz Öğrenme (k-means) algoritması uygulanmasıyla, giriş temelli olarak dört ve çıkış temelli olarak altı küme belirlenmiştir. Önerilen model, birçok sosyoekonomik değişkeni bir kriterde birleştirme konusunda önemli çaba gerektiren geleneksel yöntemlere kıyasla bir iyileşme olarak görülebilir. Ayrıca, nitel analizlere yoğun şekilde dayanan bu tür yöntemlerin öznel olma halini azaltmaya yardımcı olabilir. Önerilen yöntem sayesinde, politika yapıcılar ekonomik hedeflerine uygun gruplamalar elde edebilir ve politikaların başarısını artırabilir.

Anahtar Kelimeler: *Firma Dinamikleri, Makine Öğrenmesi, Gözetimsiz Öğrenme, Kümeleme, k-means Algoritması*

Araştırma Makalesi [Research Paper]

JEL Codes: *C02, C61, L11, M13, R12*

Submitted: 08 / 07 / 2023
Accepted: 29 / 08 / 2023

[1] Assistant Professor, Kirklareli University, Kirklareli, Türkiye, yigit.aydogan@klu.edu.tr, Orcid No: https://orcid.org/0000-0002-1823-0352

Introduction

Economics is not a definitive science. Conventional toolkit of economists relies heavily on econometrics and simulations which demand accurate modeling of complex relationships and agents. Acceptable models in economics require rigorous modeling which build up on forgiving assumptions and immense data sets, if available. This way of doing economic research cannot match the speed of increasing problems of today's world. Policymakers do not have to, and perhaps need not to, figure out complete aspects of all economic units in order to propose solutions to the problems at hand. So, there should be a way of practicing economics just enough to meet the clients' needs. At this point, Machine Learning (ML) algorithms might step in and provide guidance without perfect sight just as needed.

Industrial dynamics is a relatively young field in economic research, labeled during the late 1980s (Carlsson, 1987). It builds upon the traditional industrial organization framework by incorporating spatial aspects of firms. In this line of research, localized firm entry, growth and survival has been inspected (Frenken et al., 2015).

Firms are one of the main actors in an economy, thus understanding how they enter and exit the market are important aspects of policymaking. In theory, most economic problems could be tackled by simple and broadly defined policies. However, defining the issue could be cumbersome in the first place. That might be the reason of why the profession has been stuck around developing practice recently rather than theory. Categorization of local economies is an important step both for economic research and policymaking which is not an easy task.

It has been accustomed in the economic literature to assess the determinants of firm entry or exit, mostly with a regional perspective. There are underlying driving forces affecting entry and exit which are ought to be captured individually, only then to be tackled. However, one can design a research framework that separates addressing the determinants from defining battlefronts. This approach originates from recognizing entry and exit as *complex phenomena*. Entry is a final decision triggered by many competing forces. Exit is also a fruit of many incidents and preferences. Therefore, one can see the occurrence of entry and exit being more than the sum of its parts. This paper finds its origin here: how can we handle entry and exit by capturing all their inherent characteristics, which are extensively surveyed in the literature, without picking just a few? The standard procedure to handle entry and exit would be collecting as many data as one can on their probable determinants, and try to minimize the weight of omitted ones. New approaches in the data science allow us to skip unraveling hidden aspects of data generating processes and extract useful stylized facts. Those stylized facts then lead the way for the researchers to build complex models of the economy. It provides insight for academic modelling purposes, and also for more direct policymaking.

This paper proposes a self-organizing application for grouping spatial metrics in economics using Unsupervised Learning (UL) algorithm. Cities in Turkey will be clustered based on their firm entry and exit data between 2009-2020. Due to the different underlying nature of firm entry and firm exit, the two phenomena has been treated separately rather than grouping cities by their total firm dynamics. The paper makes a unique contribution to the literature, reaching beyond spatial firm dynamics. Data handling in economic research could improve significantly with UL algorithms. Supervised Learning (SL) algorithms require assumptions about the labeling of the data. Unfortunately, there are many complex relationships among economic agents, many of which could (and perhaps need) not to be fully addressed. Letting the firms (or cities) self-organize (or letting the clusters emerge) might provide a better way to make predictions as exampled below.

1. Literature Review

The literature on determinants of regional differentiation of new firm formation booms after the seminal works of Krugman (1991; 1992) on “a new theory of economic geography”. It can be divided into two strands: the first body of research piled up during the 1990s, underlining determinants such as unemployment, population and financing (Audretsch & Fritsch, 1994; Davidsson et al., 1994; Garofoli, 1994; Guesnier, 1994; Hart & Gudgin, 1994; Keeble & Walker, 1994, Reynolds, 1994). The second group of research emerged after 2000, stressing that firm entry depends on high technology, differing among regions by levels of varying income, population and industrial density (Armington and Acs, 2002; Porter, 2003).

Examining the variation of regional entrepreneurship, Haveman (1993) proposes a regional economic structure that a favorable region for newly founded firms would attract more entry, hence, one might expect to find more stable entry rates in low entry regions over time. Moreover, Delgado et al. (2010) asserts that strong clusters, or linkages to strong clusters for regions provide higher entry rate.

Distributional analysis of firm dynamics has been popular for a long time since Jovanovic (1982) and Hopenhayn (1992). Evolution of the firm dynamics have been studied spatially for the USA (Black & Henderson, 1999), the Netherlands (Pellenbarg & van Steen, 2003), Germany (Duschl, 2016) and France (Arcuri et al., 2019) at the firm-level. Such studies have been based on traditional methods and disregarding aggregation issues.

On another narrow strand in the literature, there have been several attempts regarding utilization of ML algorithms predicting exit behavior (Barboza et al., 2017; Bargagli-Stoffi et al., 2020a) and performance of firms (Qiu et al., 2014; Miyakawa et al., 2017). van Witteloostuijn & Kolkman (2019) test the firm growth being random with ML perspective. Recently, SL algorithms has been implemented to the predicting firms' life-cycles (Bargagli-Stoffi et al., 2021). The power of UL has been promoted to the economists earlier (Kim et al., 2002; Prüfer & Prüfer, 2018; Athey, 2019; Athey & Imbens, 2019), yet it did not receive the expected high attention up until now.

2. Data and Methodology

In order to form economic clusters of Turkish cities based on firm behavior, there are limited options in terms of data. Firm entry and exit data in Turkey is being reported monthly by TOBB (The Union of Chambers and Commodity Exchanges of Turkey) since 2009. Data of entering and exiting firms in all 81 cities of Turkey is publicly available in the form of monthly press releases (TOBB, 2021). Therefore, the dataset contains firm entry and exits in 81 cities for 144 months, a total of 23,328 data points. The data points vary from zero to 20,216 firms (whether entry or exit) per month due to the significant differences in terms of city sizes. To comply with the machine learning literature, the data has been standardized using z-score technique as suggested in Mohamad & Usman (2013).

To group a multi-dimensional population (e.g. cities in Turkey) into a pre-determined number of clusters, k-means algorithm comes forward with ease-of-use (MacQueen, 1967). In this paper, k-means algorithm has been chosen over hierarchical clustering due to its flexibility with larger data sets and clearer output than dendrograms. The nature of city level firm data does not require arbitrary steps of moving forward along the cluster levels of the latter. A handful of fine-tuned clusters which could be obtained by k-means algorithm would be more than explanatory regarding the firm dynamics of the Turkish cities.

K-means algorithm can be stated as the minimization of the below squared-error function (i.e. distance from the center) where x_i is the cities' firm dynamics statistics in the related month, and c_j is the related cluster center:

$$J = \sum_{j=1}^{k} \sum_{i \in c_j} \left\| x_i - c_j \right\|^2$$

The k-means algorithm follows three steps, and repeats step 2 & 3 until the objective function is minimized (Chu et al., 2012):

1. Randomly choose k number of cluster centers at start.

2. Assign each data point to the cluster of the closest center.

3. Compute and re-define cluster means.

This method is particularly useful in economic research because it clusters data well without requiring guidance. Detailed specifications of k-means and alternative clustering methods can be found in Chapter 14 of Hastie et al. (2009).

As stated above, k-means algorithm requires a pre-determined number of clusters to work. Choosing the number of clusters can be justified by plotting different k-cluster's objective functions (i.e. within sum of squares [WSS]) and looking for a kink in the curve. The optimal number of clusters can be selected when an additional cluster center does not significantly reduce WSS (Makles, 2012). The method is also considered to be sensitive to the initial conditions (Bradley & Fayyad, 1998). To overcome this issue, k equal segments of the data has been used to initialize the cluster means at start.

3. Finding Optimal Number of Clusters

Providing accurate number of clusters to k-means algorithm is essential for obtaining better results. As mentioned earlier, WSS plot is a straightforward way of finding optimal number of k.

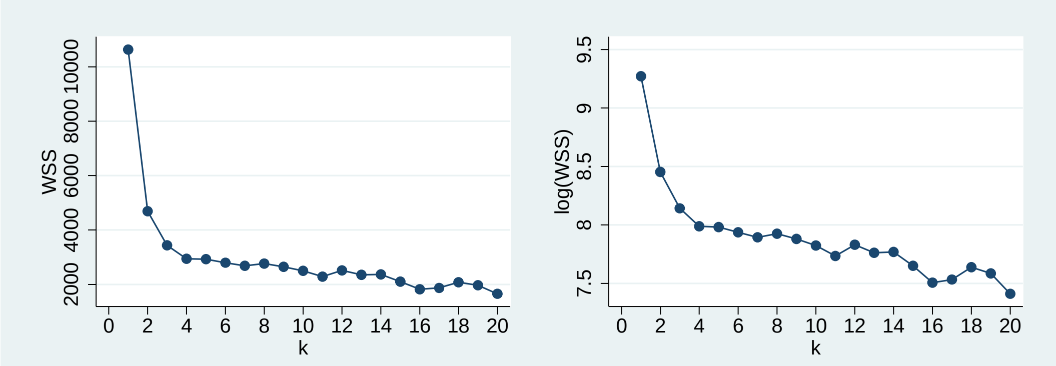

Figure 1. WSS Plot of Entry Statistics in Turkish Cities Between 2009-2020

Figure 1 depicts variation in WSS of clustering cities based on firm entry while the number of clusters (k) changes. While there is an unsurprising tendency towards higher numbers on the graph having lower WSS, the so-called kink is observed at k=4. It can be inferred from the plot that 4 clusters would provide satisfactory results for entry statistics.

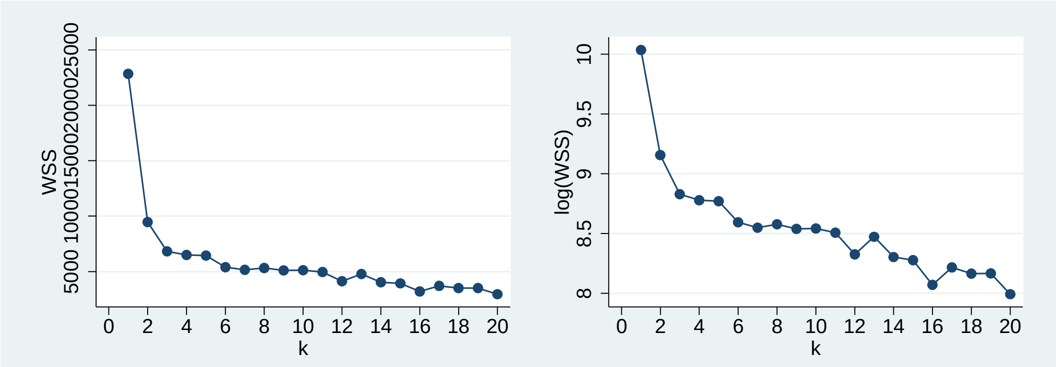

Figure 2. WSS Plot of Exit Statistics in Turkish Cities Between 2009-2020

WSS plot of pre-run clusters with exit data is depicted in Figure 2. The curve flattens at k=6 so the Turkish cities could be clustered in 6 different categories. One can argue that decrease in WSS while k increases is worth considering. In order to provide a general idea about the matter to ease policy-making, economic value of further clustering can be seen unnecessary.

4. Examining Proposed Clusters

The findings of clustering based on UL algorithm is presented in the following two tables, one for entry and the other for exit of firms. Table 1 provides summary statistics regarding 4 clusters of Turkish cities based on new firm entry to the market between 2009-2020. Cluster sizes vary from 7 to 50 cities[2]. Absence of clusters with a single member can be considered a good sign for the adequacy of the outcome of the algorithm.

Table 1. Clusters Within Firm Entry Data of Turkish Cities Between 2009-2020

Cluster ID	1	2	3	4
Cluster Size (# of Cities)	50	13	7	11
Mean of Monthly Entry	28.8	58.4	396.8	31.7
Mean of Monthly Std. Dev.	12.2	19.8	148	13.3

[2] Listings of cities based on entry and exit clusters are provided in the Appendix to this paper. For comparison, Table 5 provides the official classification by the government for investment incentives in Turkey.

Coefficient of Variation	0.425	0.34	0.37	0.419
Mean Entrants, per 100k Residents	3.24	3.48	6.91	3.86
Mean of Yearly Std. Dev. (of Entry per 100k Residents)	1.561	1.45	2.38	1.565
Coefficient of Variation (per 100k Residents)	0.48	0.42	0.34	0.41
Entrants, % of Total	27.0%	14.3%	52.2%	6.6%
Cluster Population, % of Turkey as of 2020	40.3%	14.5%	32.9%	12.2%

Note: Extra digits provided to avoid similarity due to rounding.

Table 1 presents proposed clusters of Turkish cities based on entry. In terms of new firm creation in cities, 4 clusters provide clear-cut categorizations. It seems that clusters 1 and 4 has close means and standard deviations. Coefficient of Variation (CV) of clusters 1, 4 are rather high despite being slower in terms of firm dynamics. It points out that sluggish entry of new firms have been accompanied with higher relative volatility in cluster 1.

Population based statistics in Table 1 portraits a clearer picture in terms of ordering clusters. While cities in cluster 1 has the lowest per capita (per 100k population) entry, they also suffer from the highest relative volatility by having the highest CV. CV of the clusters line up orderly as mean of per capita entry increases. It suggests that increasing per capita entry provides local economies a relatively stable influx of firms.

Table 2. Clusters Within Firm Exit Data of Turkish Cities Between 2009-2020

Cluster ID	1	2	3	4	5	6
Cluster Size (# of Cities)	10	45	3	5	14	4
Mean of Monthly Exit	13.3	3.3	212	12.8	6.1	4.3
Mean of Monthly Std. Dev.	7.3	2.6	103.3	8.5	4.2	3
Coefficient of Variation	0.55	0.77	0.49	0.66	0.68	0.71
Mean of Monthly Exit, per 100k Residents	0.74	0.47	2.35	0.85	0.57	0.45
Mean of Monthly Std. Dev. (of Exit per 100k Residents)	0.55	0.46	1.26	0.56	0.49	0.32
Coefficient of Variation (per 100k Residents)	0.74	0.99	0.54	0.65	0.87	0.72
Exits, % of Total	12.2%	13.9%	58.6%	5.9%	7.9%	1.6%
Cluster Population, % of Turkey as of 2020	14.6%	32.0%	24.9%	9.7%	13.9%	4.8%

Note: Extra digits provided to avoid similarity due to rounding.

Summary statistics of exit clusters are reported in Table 2. Cluster 2 has the lowest average exit of firms during the subjected period. Just like its counterpart in entry clusters, it has highest CV. Exit clusters are well ordered such that relative volatility (CV) decreases as average exits increase, too.

Calling the extreme values in exit statistics as good or bad for local economies might be troublesome. Since firms go out of business due to various reasons, the observed level of the number does not provide much insight on its own. For example, in an evolutionary perspective, varying number of closing firms from time to time, or place, could mean a vibrant economy with selection at work. Plus, entering new firms to the local economy shall create at least a little competition, and some firms should face challenges and create dynamism (i.e. go out-of-business).

Building on the framework, one can use the low per capita exit statistics to detect extremely dull economic environments. In this sense, the algorithm provides a finer filtering regarding the worst performing cities. Per capita exits in city clusters posted on Table 2 exhibits a close pattern. Here, cluster 6 maintains the lowest rank in terms of monthly exits but not the highest CV. Cluster 2 has the highest relative volatility of per capita exits. This anomaly can be attributed to the small size of the cluster 6, representing only 4.8% of population in Turkey and 1.6% of total exits during the subjected period. The algorithm singled out these 4 cities in a way that they represent the most ponderous local economies in Turkey in terms of selection. They all belong to the cluster 4 in terms of entry statistics, a group which has fine per capita entry performance but not as good in aggregate values as shown in Table 1.

Conclusion

In this paper, 4 entry-based and 6 exit-based clusters have been observed for Turkish cities using UL k-means algorithm. The cities in a cluster shares a similar pattern of firm dynamics between the period of 2009-2020. The findings of this paper can have large impact on solving some long-lived problems regarding policy effectiveness. Suggested groups are derived with an algorithm which reflects the full data generating process that leads to entry or exit. Therefore, proposed groups constitute a better tabulation, for instance, compared to Table 5 in the Appendix which is used for investment incentives planning by the Turkish government.

Entry based clusters show that more average new firms, whether aggregate or per capita, corresponds to lower volatility on influx of new firms to the local economy. Exit clusters also exhibit a similar pattern, except for cluster 6. Cities in exit cluster 6 represent the most plodder group in Turkish economy in terms of selection at work, by having the lowest per capita exit of firms alongside a lower variance and relatively higher per capita entry. This anomaly deserves further investigations as it might be caused from policies or subsidies disrupting market mechanisms that promote selection.

The policy-makers currently rely on straightforward methods to categorize cities in Turkey, just like the rest of the modern world. They are mainly based on geographical positions of the cities and have been justified by similarity of several socioeconomic variables. The proposed clusters are in completely different nature than the official divisions simply due the fact that the method at hand solely focuses on the local firm dynamics with no a priori directions. As a future study, smaller geographical units such as counties can be considered to decrease the scale and improve the robustness of clustering, if the data becomes available.

References

Acar, S., Kazancık, L. B., Meydan, M. C., & Işık, M. (2019). İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması SEGE-2017. Kalkınma Ajansları Genel Müdürlüğü Yayını, (3).

Arcuri, G., Brunetto, M., & Levratto, N. (2019). Spatial Patterns and Determinants of Firm Exit: An Empirical Analysis on France. The Annals of Regional Science, 62(1), 99-118. https://doi.org/10.1007/S00168-018-0887-0

Armington, C., & Acs, Z. J. (2002). The determinants of regional variation in new firm formation. *Regional studies*, *36*(1), 33-45.

Athey, S. (2019). The Impact of Machine Learning on Economics. A. Agrawal, J. Gans, & A. Goldfarb (Eds.) The Economics of Artificial Intelligence: An Agenda (507-547). University of Chicago Press. Https://Doi.Org/10.7208/9780226613475-023

Athey, S., & Imbens, G. W. (2019). Machine Learning Methods That Economists Should Know About. Annual Review of Economics, (11), 685-725. https://Doi.Org/10.1146/Annurev-Economics-080217-053433

Audretsch, D. B., & Fritsch, M. (1994). The geography of firm births in Germany. *Regional studies*, *28*(4), 359-365.

Barboza, F., Kimura, H., & Altman, E. (2017). Machine Learning Models and Bankruptcy Prediction. Expert Systems with Applications, (83), 405-417. https://Doi.Org/10.1016/J.Eswa.2017.04.006

Bargagli-Stoffi F.J., Niederreiter J., Riccaboni M. (2021) Supervised Learning for The Prediction of Firm Dynamics. In: Consoli S., Reforgiato Recupero D., Saisana M. (Eds) Data Science for Economics and Finance. Springer, Cham. https://Doi.Org/10.1007/978-3-030-66891-4_2

Bargagli-Stoffi, F. J., Riccaboni, M., & Rungi, A. (2020a). Machine Learning for Zombie Hunting. Firms' Failures and Financial Constraints. FEB Research Report Department of Economics DPS20. 06. http://Dx.Doi.Org/10.2139/Ssrn.3588410

Black, D., & Henderson, V. (1999). Spatial Evolution of Population and Industry In The United States. American Economic Review, 89(2), 321-327. https://Doi.Org/10.1257/Aer.89.2.321

Bradley, P. S., & Fayyad, U. M. (1998, July). Refining Initial Points for K-Means Clustering. ICML (98), 91-99.

Carlsson, B. (1987). Reflections on 'Industrial Dynamics': The Challenges Ahead. International Journal of Industrial Organization, 5(2), 135-148. https://Doi.Org/10.1016/S0167-7187(87)80016-4

Chu, H. J., Liau, C. J., Lin, C. H., & Su, B. S. (2012). Integration of Fuzzy Cluster Analysis and Kernel Density Estimation for Tracking Typhoon Trajectories in The Taiwan Region. Expert Systems with Applications, 39(10), 9451-9457. https://Doi.Org/10.1016/J.Eswa.2012.02.114

Davidsson, P., Lindmark, L., & Olofsson, C. (1994). New firm formation and regional development in Sweden. *Regional studies*, *28*(4), 395-410.

Frenken, K., Cefis, E., & Stam, E. (2015). Industrial Dynamics and Clusters: A Survey. Regional Studies, 49(1), 10-27. https://Doi.Org/10.1080/00343404.2014.904505

Garofoli, G. (1994). New firm formation and regional development: the Italian case. *Regional studies*, *28*(4), 381-393.

Guesnier, B. (1994). Regional variations in new firm formation in France. *Regional studies*, *28*(4), 347-358.

Hart, M., & Gudgin, G. (1994). Spatial variations in new firm formation in the Republic of Ireland, 1980–1990. *Regional Studies*, *28*(4), 367-380.

Haveman, H. A. (1993). Follow the leader: Mimetic isomorphism and entry into new markets. *Administrative science quarterly*, 593-627.

Hopenhayn, H. A. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. Econometrica: Journal of The Econometric Society, 1127-1150. https://doi.org/10.2307/2951541

Jovanovic, B. (1982). Selection and The Evolution of Industry. Econometrica: Journal of The Econometric Society, 649-670. https://doi.org/10.2307/1912606

Keeble, D., & Walker, S. (1994). New firms, small firms and dead firms: spatial patterns and determinants in the United Kingdom. *Regional studies*, *28*(4), 411-427.

Kim, Y., Street, W. N., & Menczer, F. (2002). Evolutionary Model Selection in Unsupervised Learning. Intelligent Data Analysis, 6(6), 531-556. https://doi.org/10.3233/IDA-2002-6605

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, *99*(3), 483-499.

Krugman, P. (1992). *Geography and trade*. MIT press.

Macqueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. In Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, 281–297.

Makles, A. (2012). Stata Tip 110: How to Get the Optimal K-Means Cluster Solution. The Stata Journal, 12(2), 347-351.

Miyakawa, D., Miyauchi, Y., & Perez, C. (2017). Forecasting Firm Performance with Machine Learning: Evidence from Japanese Firm-Level Data. Technical Report, Research Institute of Economy, Trade and Industry (RIETI).

Mohamad, I. B., & Usman, D. (2013). Standardization and Its Effects on K-Means Clustering Algorithm. Research Journal of Applied Sciences, Engineering and Technology, 6(17), 3299-3303. https://doi.org/10.19026/RJASET.6.3638

Pellenbarg, P. H., & Van Steen, P. J. (2003). Spatial Perspectives on Firm Dynamics in The Netherlands. Tijdschrift Voor Economische En Sociale Geografie, 94(5), 620-630. https://doi.org/10.1046/j.1467-9663.2003.00288.x

Porter, M. (2003). The Economic Performance of Regions. *Regional Studies: The Journal of the Regional Studies Association*, *37*(6-7), 545-546.

Prüfer, J., & Prüfer, P. (2018). Data Science for Institutional and Organizational Economics. In A Research Agenda for New Institutional Economics. Edward Elgar Publishing.

Qiu, X. Y., Srinivasan, P., & Hu, Y. (2014). Supervised Learning Models to Predict Firm Performance with Annual Reports: An Empirical Study. Journal of The Association for Information Science and Technology, 65(2), 400-413. https://doi.org/10.1002/asi.22983

Reynolds, P. (1994). Autonomous firm dynamics and economic growth in the United States, 1986–1990. *Regional Studies*, *28*(4), 429-442.

TOBB. (2021). Company Establishment and Liquidation Statistics [Database]. Available At: https://tobb.Org.Tr/Bilgierisimmudurlugu/Sayfalar/Eng/Kurulankapanansirketistatistikleri.Php

Van Witteloostuijn, A., & Kolkman, D. (2019). Is Firm Growth Random? A Machine Learning Perspective. Journal of Business Venturing Insights, 11, E00107. https://doi.org/10.1016/j.jbvi.2018.e00107

Appendix

Table 3. Listing of Turkish Cities Based on Firm Entry

Cluster 1			Cluster 2	Cluster 3	Cluster 4
Adana	Corum	Kirklareli	Ankara	Istanbul	Kayseri
Adiyaman	Denizli	Kirsehir	Malatya	Izmir	Manisa
Afyon	Diyarbakir	Kutahya	Kahramanmaras	Kocaeli	Sakarya
Agri	Edirne	Mardin	Ordu	Konya	Trabzon
Amasya	Elazig	Mus	Rize	Mugla	Sanliurfa
Antalya	Erzincan	Nevsehir	Sivas	Samsun	Van
Artvin	Erzurum	Nigde	Tokat	Tekirdag	Zonguldak
Aydin	Eskisehir	Siirt	Usak		Aksaray
Balikesir	Gaziantep	Sinop	Yozgat		Yalova
Bilecik	Giresun	Tunceli	Karaman		Osmaniye
Bingol	Gumushane	Bayburt	Kirikkale		Duzce
Bitlis	Hakkari	Sirnak	Batman		
Bolu	Hatay	Bartin	Karabuk		
Burdur	Isparta	Ardahan			
Bursa	Mersin	Igdir			
Canakkale	Kars	Kilis			
Cankiri	Kastamonu				

Table 4. Listing of Turkish Cities Based on Firm Exit

Cluster 1	Cluster 2			Cluster 3	Cluster 4	Cluster 5	Cluster 6
Ankara	Adana	Corum	Kars	Istanbul	Kayseri	Antalya	Manisa
Ordu	Adiyaman	Denizli	Kastamonu	Izmir	Kocaeli	Bursa	Trabzon

Sakarya	Afyon	Diyarbakir	Kirklareli	Mugla	Konya	Kutahya	Van
Sivas	Agri	Edirne	Kirsehir		Samsun	Malatya	Zonguldak
Sanliurfa	Amasya	Elazig	Mardin		Tekirdag	Kahramanmaras	
Usak	Artvin	Erzincan	Mus			Nevsehir	
Aksaray	Aydin	Erzurum	Siirt			Nigde	
Yalova	Balikesir	Eskisehir	Sinop			Rize	
Osmaniye	Bilecik	Gaziantep	Tunceli			Tokat	
Duzce	Bingol	Giresun	Bayburt			Yozgat	
	Bitlis	Gumushane	Sirnak			Karaman	
	Bolu	Hakkari	Bartin			Kirikkale	
	Burdur	Hatay	Ardahan			Batman	
	Canakkale	İsparta	Igdir			Karabuk	
	Cankiri	Mersin	Kilis				

Table 5. Official Classification by the Government for Investment Incentives in Turkey

Region 1	Region 2	Region 3	Region 4	Region 5	Region 6
Ankara	Aydin	Adana	Afyonkarahisar	Bayburt	Adiyaman
Antalya	Balikesir	Burdur	Aksaray	Cankiri	Agri
Bursa	Bilecik	Duzce	Amasya	Erzurum	Ardahan
Eskisehir	Bolu	Gaziantep	Artvin	Giresun	Batman
Istanbul	Canakkale	Karaman	Bartin	Gumushane	Bingol
Izmir	Denizli	Kirikkale	Çorum	Kahramanmaras	Bitlis
Kocaeli	Edirne	Kutahya	Elazig	Kilis	Diyarbakir
Mugla	Isparta	Mersin	Erzincan	Nigde	Hakkari
Tekirdag	Karabuk	Samsun	Hatay	Ordu	Igdir
	Kayseri	Trabzon	Kastamonu	Osmaniye	Kars
	Kirklareli	Rize	Kirsehir	Sinop	Mardin
	Konya	Usak	Malatya	Tokat	Mus
	Manisa	Zonguldak	Nevsehir	Tunceli	Siirt
	Sakarya		Sivas	Yozgat	Sanliurfa
	Yalova				Sirnak
					Van

Source: Acar et al., 2019