Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/SA6.20250196
Published: May 12, 2025

Determinants of State-Level Innovation: A Predictive Model for Patent Generation in Mexico's 32 States Over Two Decades

Juan Carlos Rodriguez Briones
Coordinator And Professor of The Executive and Top Management Master's Program
Universidad Iberoamericana Monterrey
Monterrey, Nuevo Leon, Mexico
jcrodriguez@strategicpractice.com.mx, juancarlos.rodriguez@c-iberomty.mx
https://orcid.org/0009-0004-0025-151X

María Guadalupe López Molina
Full-time professor in the science and engineering department
Universidad Iberoamericana Puebla
Puebla, Puebla, Mexico
musi.lopez@iberopuebla.mx

Abstract

This research presents a quantitative analysis of state-level competitiveness across Mexico's 32 states, based on data spanning 22 years. By exploring key economic, educational, and infrastructural variables, we develop a predictive model that estimates patent generation per 100,000 inhabitants, identifying the most influential factors driving innovation. The study employs exploratory data analysis, parametric statistics, and multiple linear regression to uncover insights into how regional competitiveness affects patent output, a proxy for innovation capacity. Our findings suggest that states with higher levels of academic performance, talent availability and economic diversification demonstrate significantly stronger innovation potential. To further clarify regional disparities, we propose a novel classification system that categorizes states into three tiers of innovation strength: Leading, Intermediate and Lagging. This classification provides policymakers with a tool to better understand innovation dynamics and tailor strategies for fostering technological growth at the state level. The results of this study are particularly relevant for regions seeking to enhance their innovation ecosystems by identifying and leveraging key drivers of patent generation. This model not only contributes to the existing literature on regional competitiveness and innovation but also offers practical insights for policymakers aiming to strengthen innovation capacity in emerging economies.

Keywords
Innovation drivers, Predictive Modeling, Patent Generation, Regional Competitiveness, Mexico.

1. Introduction
Innovation plays a critical role in driving economic growth and enhancing competitiveness, particularly in emerging economies where technological advancements are key to boosting productivity and global integration. As countries like Mexico seek to transition into knowledge-based economies, fostering innovation becomes a national priority.

Patents, often used as a proxy for innovation, are a tangible measure of technological progress. However, despite national efforts to enhance innovation capacity, significant disparities persist across Mexico's 32 states.

Some states show strong innovation performance, reflected in higher patent generation, while others lag behind due to regional imbalances in investment, human capital, and infrastructure. These disparities hinder the country's overall competitiveness and raise the question of what specific factors contribute to the uneven innovation landscape. Although policies at the national level have aimed to promote innovation, the lack of granular insights into state-level drivers has limited the effectiveness of these interventions. This uneven distribution of innovation potential poses a challenge to achieving balanced economic growth and inclusive development.

This research addresses this gap by conducting a comprehensive analysis of 22 years of data to identify the key factors influencing patent generation at the state level in Mexico. Through exploratory data analysis, parametric statistics, and multiple linear regression, we develop a predictive model to estimate patent output per 100,000 inhabitants. Additionally, the study introduces a classification system that categorizes states into high, moderate, and low innovation strength, providing policymakers with targeted insights to foster innovation across all regions. By understanding the specific determinants of state-level innovation, this research aims to contribute both to the academic discourse and to practical policy applications aimed at reducing regional disparities in innovation capacity. (10 font)

1.1 Objectives

This study aims to develop a predictive model that identifies the key factors influencing patent generation across Mexico's 32 states, providing insights into the regional dynamics of innovation and offering a classification system to differentiate states based on their innovation capacity. And the two specific objectives are, a) To analyze the relationship between state-level competitiveness variables and patent generation over the past 22 years, identifying the most significant predictors of innovation, and b) To create a classification framework that categorizes Mexican states offering policymakers a tool to design targeted innovation-boosting strategies.

2. Literature Review

The correlation between innovation and economic growth has been widely documented in the literature, with numerous studies demonstrating that technological advancements and innovative activities significantly contribute to a country's long-term economic development. According to Maradana et. al. (2017) innovation activities, such as patents filing and economic growth have a *"co-joined back-and-forth relationship between them",* as shown in his study of European countries, such relationship is particularly crucial for emerging economies, where enhancing innovation capabilities can drive structural transformations and reduce reliance on traditional, less competitive industries. In the context of Latin America, a region that exhibits a persistent gap in innovation performance compared to developed economies, Avila-Lopez, et. al. (2019) **,** found that, for countries of the largest economies in that region, such as México, innovation and economic growth have a bidirectional causality. When analyzing Latin America as a region, Gutiérrez Flores & Flores Pérez (2019) proposed that based on each countries characteristics, a series of factors such as technology adoption, research and development investment, intellectual property protection and talent development, have a different impact on each country's innovation performance.

Focusing in Mexico, the World Intellectual Property Organization (2023) in their Global Innovation Index, ranks the country in 58th place, out of 132 countries evaluated, that is good enough to be third place in the larin American region, behind Brazil and Chile. Specifically, Mexico is ranked 82nd in patents generation and 78th in tertiary education, both below average, and 12th in trade, diversification and market scale, those ranks depicts main weaknesses and strengths as a country regarding the most relevant factors that drives innovation.

While country information is relevant, this research strives to go deeper and understand innovation on a state level, as Unger (2018) shows, there is a strong and sustained evidence of a correlation between innovation and state level competitiveness in Mexico, even if the actual innovation efforts have been relatively modest across the country, Cabrero y de Los Cobos (2021) propose a classification for states, according to the proposed Innovation Capabilities model, in that research, based on economic, social and government variables, they propose a six level classification of states, going from Very Low to very High integrated innovation capabilities. Sánchez et. al. (2014) propose the existence of five different regions in Mexico, 1) non-metropolitan region with low innovative capacity, 2) industrial region with innovative potential, 3) small industrial regions, 4) metropolitan region with innovative capacity and

5)diversified industrial region, based on both, their geographical location and the level of three factors, regional resources for innovation, socioeconomical structure and productive structure,

This research seek to contextualize a relevant variable, in this case innovation in the Mexican states, measured in "Patents applied for per 100 thousand of the economically active population", which is treated as a proxy variable for the innovation performance, and finding the most relevant factors that drive it, this is relevant because, as proposed by Cabrero Mendoza (2023) there is a Virtuous Cycle of the Knowledge Economy, which starts with the investment, by various social actors, in science and the product of this generates knowledge, which is a public good, and the way in which this is incorporated into innovation processes is capable of generating wealth for society, allowing the cycle to be reinforced once again.

3. Methods

This study focuses on developing a predictive model that identifies the main factors that drive innovation at the state level in Mexico, using a quantitative approach, a three-step methodology is proposed, which is explained below.

Step 1. Data selection and processing. This process is required to clean and prepare the data, from a relevant and trusted source, for further analysis.

Step 2. Exploratory Data Analysis. Analysis of initial data, using graphical and analytical techniques to detect initial underlying patterns and trends that help describe the phenomenon under study.

Step 3. Regression analysis and hypothesis tests. Using statistical procedures of regression, correlation and hypothesis testing, the causal relationships between the dependent variable and the independent variables are identified, measuring their strength and direction.

3.1 Data selection and processing

The quantitative analysis conducted in this research used the State Competitiveness Index (Instituto Mexicano Para La Competitividad, 2022) **,** which is composed of 72 indicators and 10 sub-indexes, based on which the states are evaluated and classified into six levels of competitiveness. For this analysis, 18 indicators (variables) corresponding to 8 sub-indexes were pre-selected, using data from 2002 to 2016, 2021 and 2022, as this is the most recent information available now, it is worth mentioning that, at the date of writing this document, the data from 2017 to 2020 remains unpublished.

Variables belonging to 10 sub-indexes were used (Inclusive, prepared and healthy society; Stable and functional political system; Efficient and effective governments; Factor market; Stable economy; Precursors; Taking advantage of international relations; Innovation of economic sectors; Anchor variables), these 10 sub- indexes used are aligned with the proposal of Cabrero and De Los Cobos (2021)who in their study related to regional innovation systems in Mexico, take into account three dimensions of innovation:

Economic Innovation, focused on the levels of progress and development of the economy in the federal entities;
Government Innovation, focused on institutional, technical and administrative modernization actions; and
Social Innovation, focused on changes in social dynamics, focusing on professional and technical training and the creation of technological infrastructure, Table 1 shows how each of the selected indicators falls into these three categories.

Table 1. Classification of indicators used

Economic Innovation	**Government Innovation**	**Social Innovation**
GDP in high growth sectors	Perception of state corruption	Grade of education
Economic diversification	Government interaction via electronic means	Academic performance
Patents	Subnational regulatory improvement indicator	Labor training
GDP per capita		Internet access
Foreign direct investment		Talent
Export of goods		

The fourteen indicators used in the study are defined in detail in Table 2, including the full name of the indicator, the units in which it is expressed and the sub-index that it belongs to, an Excel database was built with them to standardize formats, organize the information and prepare it for later transfer to the Minitab statistical software for analysis and testing.

Table 2. Table of indicators used

No.	**sub-index**	**Is more better?**	**Indicator**	**Units**
28	Inclusive, prepared and healthy society	Yes	Scholarity	Average Years
30		Yes	Academic performance	Percentage of students at math performance level 3 and 4
41	Stable and functional political system	No	Perception of state corruption	Percentage of urban population over the age of 18 who consider corrupt practices in state government to be prevalent and very prevalent
50	Efficient and effective governments	Yes	Interaction with government via electronic means	Percentage of population aged 18 years and older who have had at least one interaction with government via electronic means
55		Yes	Subnational regulatory improvement indicator	Index (0-5)
66	Factor Market	Yes	Labor training	Economically active population that has received training as a percentage of the total economically active population
67	Stable Economy	Yes	GDP per capita	Pesos per person
68		Yes	GDP in high growth sectors	Millions of pesos at 2013 prices
77		Yes	Economic diversification	Number of sectors present in the economy
79	Precursors	Yes	Households with Internet access	Percentage of total households
92	Leveraging international relations	Yes	Foreign direct investment	Dollars per thousand of GDP
93		Yes	Export of goods	Percentage of GDP
97	Innovation of economic sectors	Yes	Patents	Patents applied for per 100,000 of the economically active population

100	Anchor	Yes	Talent	Percentage of the population aged 25 years or older with tertiary, teacher training or technical education

3.2 Exploratory Data Analysis

Counting 640 records over more than 20 years, the first analysis shows an average of 1.56 patents and a median of 0.91, as shown in Figure 1, shows both central tendency and dispersion, at the national level between 2001 and 2022.

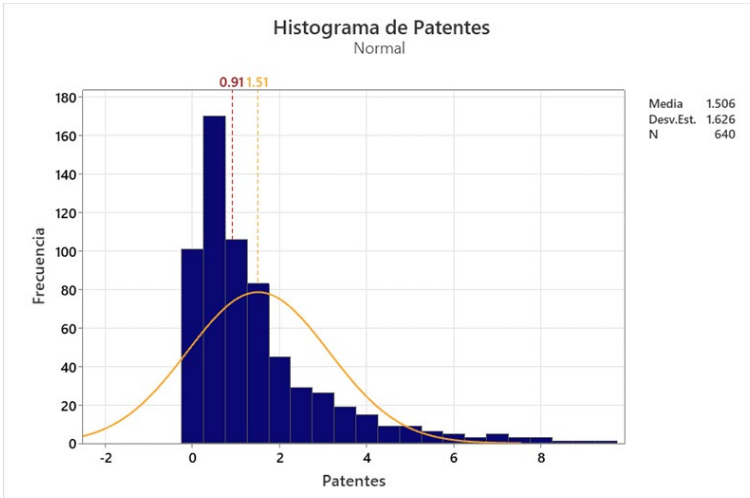

Figure 1. State level Patents, from 2001 to 2022

To continue with the exploratory data analysis, the trend over time of the variable of interest, Patents, is analyzed and complemented with the creation of a control chart of averages and standard deviations as shown in Figure 2, taking 32 as sample size, that is, for the average of each year, the patents generated by the 32 states are considered. When observing the behavior over the years, in addition to the upward trend in the generation of patents, only interrupted in recent years, patterns can be observed that denote the presence of special sources of variation, of particular relevance is that between 2001 and 2009, the yearly average of patents were below the central tendency line (1.506, the average of averages) and the continuous performance above said line and in an upward trend between 210 and 2017.

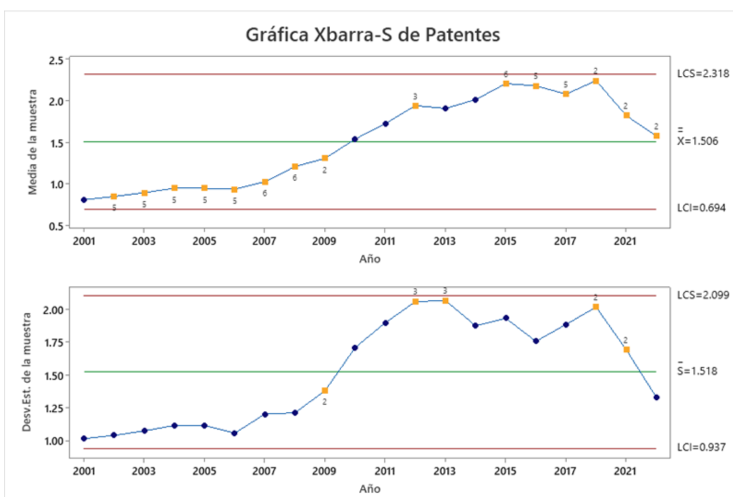

Figure 2. State level Patents by state, from 2001 to 2022

With the above mentioned trends and patterns, the hypothesis of the existence of "two processes", a "before and after" is proposed, observing that since that year, patents have been generated in Mexico consistently above the average of the last 20 years, it is relevant to run a hypothesis test to validate this assumption, therefore, a two-means hypothesis test was run, adding a qualitative variable to the database, the variable "New Process" which is a dichotomous variable, which accepts values of "YES", in case it is a year after 2011, "NO", for all values corresponding to previous years, the results are shown in Figure 3.

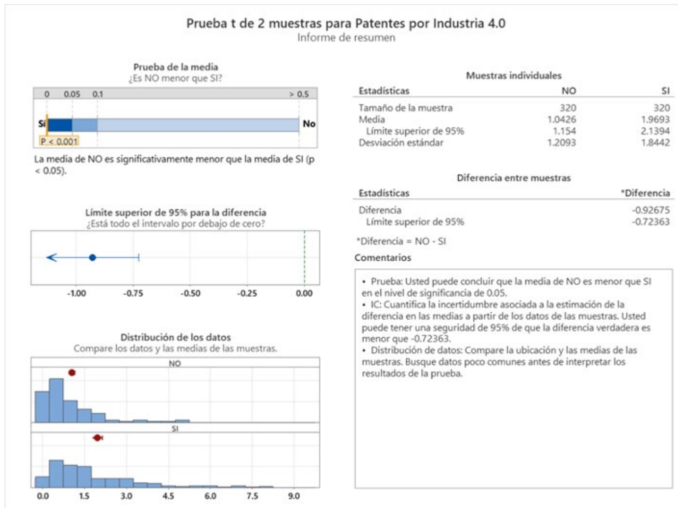

Figure 3. Two Sample t test for Patents before and after 2011 in Mexico

It is observed that the average goes from 1.04 patents at the national level before 2011 to 1.96 after said year, and with a 95% confidence level, when observing a p-value of less than 0.001, the difference is statistically significant, so it can be concluded that the hypothesis of the change in the Mexican environment, that "before and after" is confirmed as a valuable element to consider in the following analyses. Thanks to the large sample size, there is no risk that the normality in the distribution of the data could represent a problem in the conclusions of the study. In addition, the existence of "uncommon" data is observed, so it is suggested to continue investigating the cause of these, a hypothesis that will be taken up in the following analyses is the existence of some federal entities that stand out and generate most of the atypical data. A first step is to calculate the descriptive statistics of the variable Patents for all the states of the Mexican Republic, from 2001 to date, showing the results in Table 3.

Table 3. Descriptive statistics for Patents by State

State	**Total count**	**Mean**	**Std. Dev.**	**Median**
Aguascalientes	20	1.68	0.69	1.66
Baja California	20	0.82	0.57	0.77
Baja California Sur	20	0.76	0.58	0.57
Campeche	20	1.25	0.92	0.95
Chiapas	20	0.28	0.29	0.22
Chihuahua	20	1.55	0.53	1.57
Ciudad de México	**20**	**6.26**	**1.76**	**5.71**
Coahuila	20	2.53	1.08	2.23
Colima	20	1.33	0.81	1.08
Durango	20	0.58	0.33	0.54
Guanajuato	20	1.67	1.06	1.64
Guerrero	20	0.13	0.08	0.14
Hidalgo	20	1.16	0.99	0.84
Jalisco	**20**	**3.10**	**1.40**	**2.55**
México	20	1.05	0.24	1.00
Michoacán	20	0.46	0.22	0.46
Morelos	20	2.88	1.54	2.57
Nayarit	20	0.23	0.27	0.18
Nuevo León	**20**	**4.38**	**1.25**	**3.75**
Oaxaca	20	0.39	0.28	0.34
Puebla	20	1.90	1.24	1.98
Querétaro	**20**	**4.49**	**2.04**	**3.75**
Quintana Roo	20	0.57	0.50	0.45
San Luis Potosí	20	0.73	0.49	0.70
Sinaloa	20	1.38	1.03	1.27
Sonora	20	1.60	1.09	1.55
Tabasco	20	0.93	0.62	0.77
Tamaulipas	20	1.04	0.47	0.90
Tlaxcala	20	0.58	0.48	0.52
Veracruz	20	0.50	0.26	0.47
Yucatán	20	1.45	0.91	1.35
Zacatecas	20	0.55	0.29	0.70

Remembering that the total average is 1.51 patents per 100,000 inhabitants yearly, 4 entities stand out as they average a patent generation above twice the national average, they are Mexico City (6.25), Jalisco (3.09), Nuevo León (4.75) and Querétaro (4.49), this information raises an additional hypothesis, the existence of "*several Mexicos*" in innovation.

In accordance with the findings discussed, the following analyses only consider the records from 2011 onwards, as this is considered the "new process" to be analyzed, from a process perspective describing that the political and economic circumstances from that year onwards are significantly different in Mexico, compared to those of the previous decade. In analyzing the patents for this period, we begin with the graphical summary, which consolidates both the calculation of descriptive statistics, the graphical representation of the data and the estimation of confidence intervals, both for the mean and the median, which is shown in Figure 4.

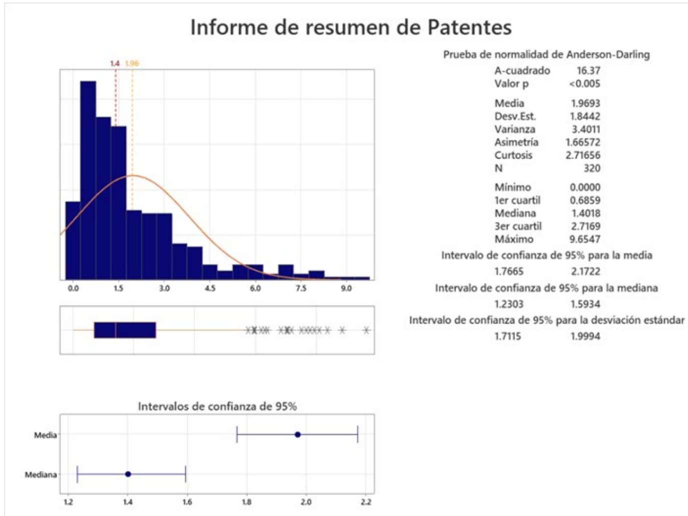

Figure 4. Graphical Summary for Patents from 2011 to 2022 in Mexico

When analyzing the data, the presence of outliers stands out, which are defined as those that are at a distance from the first quartile and the third quartile that is greater than 1. 5 times the interquartile range, these data belong, primarily, to the four states that were detected as those with the highest average number of patents in this sample from 2011 to 2022, which are Mexico City, Jalisco, Nuevo Leon and Queretaro, taking into account the median value (1.40) and the third quartile (2.46) a first classification of states can be made, as shown in the Figure 5.

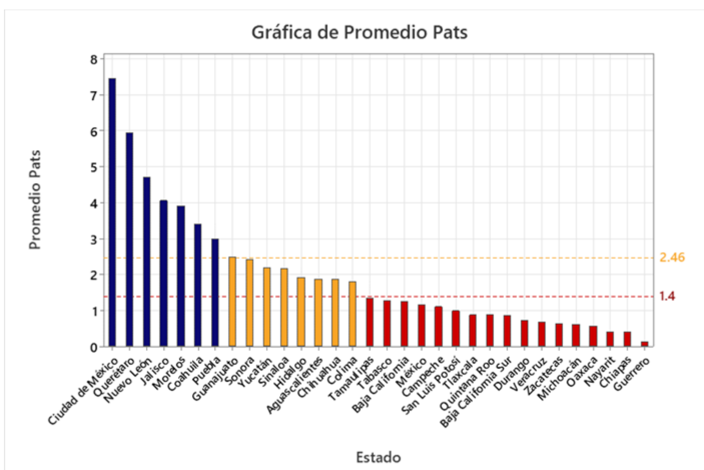

Figure 5. Bar chart for the average number of patents per state between 2011 and 2022

The graph shows the entities whose average is below the median (1.40) in red (which will be classified as "Lagging" in the rest of the analyses), while the states with an average between the median and the third quartile (2.46) are identified in orange (which will be classified as "Intermediate" in the rest of the analyses) and finally, states whose average is above the third quartile (which will be classified as "Leading" in the rest of the analyses) are identified in blue.

3.3 Regression analysis and hypothesis tests

To confirm the statistical validity of these three classifications, an ANOVA is performed to check if there is a statistically significant difference between the means of Patents for the three proposed families of states, which is shown in Figure 6.

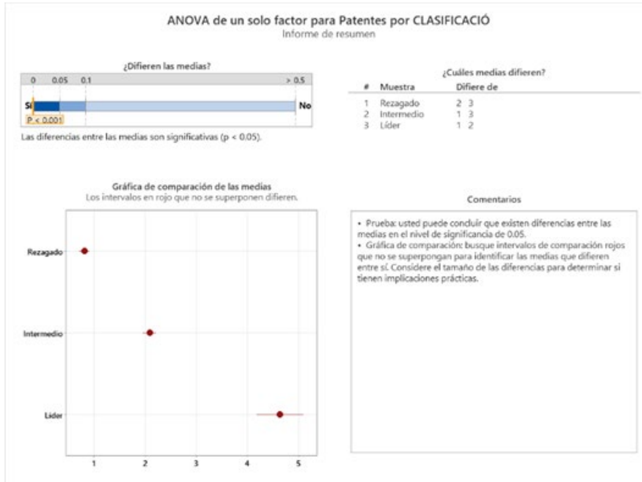

Figure 6. One way ANOVA results for Patents by proposed families, 2011 and 2022

A review of the results shows that, with a 95% confidence level, the averages of the three proposed classifications (Leader with mean of 4.6, Intermediate with mean of 2.1 and Lagging with mean of 0.8) differ among themselves and thanks to the size of the sample (340 data points) there is no problem of normality that affects the conclusions of the study. When observing the geographic distribution of the entities throughout the Mexican Republic, it is relevant how the seven entities identified as Leaders are dispersed in five different economic zones: Southcentral (Mexico City and Morelos), Northcentral (Querétaro), Northeast (Nuevo León and Coahuila), West (Jalisco), Noreste (Nuevo León and Coahuila), West (Jalisco), East (Puebla) which supports the hypothesis of the existence of "*Several Mexicos*" and also shows that geographic conditions are not the only determinant when trying to define the characteristics that make a federal entity an advanced place in terms of innovation.

Figure 7. Mexico's map, showing each state's innovation classification

To perform the initial correlation analysis between the variables used, a correlogram was created, as shown in Figure 8. This first analysis shows the direction (positive or negative) and intensity (strong or weak) of all the variables analyzed, including the calculation of the correlation coefficient for each of them. Remembering that the variable of interest (Patents) is the one we are interested in predicting, the first exercise is carried out to show the relationship of each of the dependent variables with the "Y" and with the rest of the "X" variables with each other.

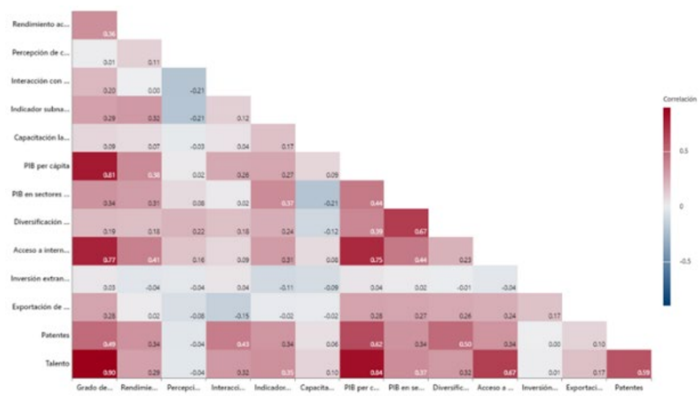

Figure 8. Correlogram of variables, from 2011 to 2022

Analyzing the correlation coefficients, three independent variables are detected with a strong positive correlation with the variable Patents, those are, Academic Performance (Percentage of students at math performance level 3 and 4), Economic Diversification (Number of sectors present in the economy) and Talent (Percentage of the population aged 25 years or older with tertiary, teacher training or technical education). Those are included in the multiple regression analysis, in addition the variable Classification is included, to consider the current state of progress that the federal entity has in the generation of patents. After performing the analysis, the regression model is obtained:

$$Patents = -3.992 + 0.752\ Academic\ Performance + 0.002150 Economic\ Diversification + 7.92\ Talent + 1.4341 Classification$$

When analyzing the coefficients, we observe that the four variables considered have a positive effect on the generation of patents, i.e., the more academic performance, economic diversification and talent in the states increase, the more patents are expected to be generated. It is highlighted that these four variables alone explain almost 70% of the changes in the dependent variable, according to the value of the estimated coefficient of determination (R-squared). In addition, a summary of the model is generated, shown in Table 4, from which we can see that three of the four variables have a p-value close to zero, which implies that their effect cannot be considered based on random causes, so a statistically significant causality relationship can be assumed between them and the dependent variable Patents.

Table 4. Coefficients and P Values for regression model

Coefficients		
Terms	Coeff	P Value
Constant	-3.992	0
Academic Performance	0.752	0.38
Economic Diversification	0.00215	0.009
Talent	7.92	0
Classification	1.4341	0

4. Conclusion

This study provides a comprehensive analysis of state-level innovation dynamics in Mexico, using a predictive model to identify key factors influencing patent generation across 22 years of data. The findings highlight the critical role of Academic Performance, Economic Diversification Talent in shaping regional innovation capabilities in Mexico. Furthermore, by categorizing states into Leading, Intermediate and Lagging states, the study offers a practical framework for policymakers to design tailored strategies that address regional disparities and foster balanced economic growth. These insights contribute to the broader understanding of how regional competitiveness factors drive innovation, with implications for emerging economies aiming to enhance their innovation ecosystems.

However, while the proposed model effectively predicts patent output, it should be considered an initial step towards a deeper understanding of state-level innovation. The dynamic nature of technological development and economic conditions requires continuous refinement of the model to maintain its predictive accuracy.

In order to build upon this research, we propose 3 follow-up steps. 1. Incorporate Additional Variables: Future research could expand the model by integrating more variables, such as intellectual property policies, digitalization levels or specific innovation programs available per state, to provide a more comprehensive understanding of factors influencing patent generation. 2. Conduct Longitudinal Case Studies: An in-depth exploration of states that have shown significant shifts in innovation capacity over time could reveal specific policies or initiatives that drive successful outcomes, offering valuable lessons for other regions. 3. Cross-Country Comparative Analysis: Expanding the study to compare Mexico's regional innovation dynamics with those of other emerging economies would allow for identifying common patterns or unique strategies that could be applied to improve innovation performance on a global scale.

References

Avila-Lopez, L. A., Lyu, C., & Lopez-Leyva, S, Innovation and growth: evidence from Latin American countries. *JOURNAL OF APPLIED ECONOMICS*, 287-303, 2019.

Cabrero Mendoza, E. (26 de Septiembre de , *Conferencia: Desafíos y oportunidades para la universidad en la era del conocimiento.* 2023. Obtenido de https://www.youtube.com/live/xr5arZwVimM?si=yWlxSyk2JyrwetjU&t=705

Cabrero Mendoza, E., & De Los Cobos, P., *Las capacidades de innovación regional como un elemento clave en la construcción de un Federalismo Responsable.* Ciudad de Mexico: CIDE.2021.

Gutiérrez Flores, L., & Flores Pérez, J, Factors which stimulate innovation activity in Latin America: a VECM approach. *Economía, Sociedad y Territorio*, 373-403, 2019.

Instituto Mexicano Para La Competitividad, *Índice de Competitividad Estatal* . Ciudad de México: IMCO.2022.

Maradana, R. P., Pradhan, R. .., Dash, S., Gaurav, K., Jayakumar, M., & Chatterjee, D Does innovation promote economic growth? Evidence from European countries. *Journal of Innovation and Entrepreneurship*, 2017. DOI 10.1186/s13731-016-0061-9.

Sanchez Tovar, Y., Garcia Fernandez, F., & Mendoza Flores, E, Determinantes de la capacidad de innovación regional en México. Una tipología de las regiones. *Región y sociedad*, 119-160.2014.

Unger, K. , Innovation, Competitiveness and Profitability in the Mexican Economy. *Gestión y Política Pública*, 3-37.2018.

World Intellectual Property Organization, *Global Innovation Index 2023: Innovation in the face of uncertainty.* Geneva: WIPO.2023.

Biographies

Juan Carlos Rodriguez Briones is the founder and managing partner of Strategic Practice Mexico, a boutique Consultancy Firm, focused on strategy and operational excellence, and serves as coordinator and professor of the Executive and Top Management Master's program at Universidad Iberoamericana Monterrey, in Monterrey, Mexico. Mr. Rodriguez holds a degree in Industrial and Systems Engineering and a Masters degree of Integrated Quality Management, both from Universidad de Monterrey. Mr. Rodriguez is an experienced executive in Strategic and Operational Excellence, with over 20 years of leadership roles in Quality, Lean, and Operations, in international companies, while successfully implementing projects in a wide array of sectors. He is currently pursuing a Doctorate in Business Administration at Universidad Iberoamericana, Puebla, his research focuses on strategy and innovation on small and medium enterprises. Additionally, the author is a member of the Strategic Management Society and was recognized as a distinguished professor by IEOM.

María Guadalupe López Molina is a full-time professor in the science and engineering department of the Universidad Iberoamericana Puebla in Puebla, Mexico. She holds a degree in Computer Systems Engineering from Universidad de las Américas Puebla, a Master's degree in Computer Science from Universidad Nacional Autónoma de México, a Master's degree in Higher Education from Universidad Iberoamericana Puebla and a PhD in Strategic Planning and Technology Management from Universidad Popular Autónoma del Estado de Puebla. She has worked in several projects mainly with small and medium sized companies, from these projects have resulted in several publications and presentations at conferences. Her interests include strategies for small and medium enterprises, lean enterprises, technological strategies and frugal innovation.