
Luisa Decuir-Viruez

University of Kent
Department of Economics
Canterbury, Kent CT2 7NP
Tel. (44) 227 827679 Fax. (44) 227 827850
Email: ld24@kent.ac.uk

April, 2004

Preliminary

Keywords: Institutions, Regional Growth, Mexico, Cluster Analysis.

Abstract

This paper examines the role of institutional factors in the economic growth of Mexican regions. Assuming an institutional dimension of growth, we will assess the relationships between informal and formal institutional features, infrastructure conditions and growth. Following the approach of the document presented at the ERSA Conference (2003), we introduce the results obtained in the process of measuring the institutional factors for three different periods 1970-1985, 1985-1994 and 1994-2000. Cluster Analysis techniques were applied to a database of 13-14 variables; from we identified patterns of institutional characteristics for the Mexican regions and groups of states that share some features. The results are confronted with the evidence of growth for these periods, in order to identify if a particular regional institutional structure has led to higher or lower rates of growth.

I. The Institutional Dimension of Growth

One of the main inquiries in Economics is to explain the reasons of inequalities among societies, the causes of that division between rich and poor countries or regions. The question of unequal economic performance and territorial disparities has been unraveled through theories and models in which indicators of capital, investment, saving, productivity and some socioeconomic variables -such as education-, have been used. Within these studies, those that only emphasize on technological factors had led to a certain neglect of the role of social forces. But society matters, social and institutional features define the structure in which economic activity is realized. They influence the quality of investment, the level of technical efficiency and the ability of the regions and countries to assimilate technology from abroad.

The reasons for considering institutions in the determinants of growth can be found in empirical evidence analysis. For instance, the persistence of disparities in regions, such as Europe where mobility of economic and technological factors is greater, suggests a connection with local social structure that helps some societies to assimilate, replace and respond to challenges in a better way than others. Moreover, the successful results observed in regions where institutional elements seems to have an important role as a source of growth, have also opened a debate on the way regional policies must be. But precisely, the evidence of no-convergence among countries and regions and the studies of conditional convergence have triggered the challenge of assessing the role of institutions in growth. Also the connection among social indicator and institutions has shown that some of the indexes of social development are very useful for predicting subsequent growth (Temple and Johnson, 1988), because the social capability is reflected in those indicators.

In the literature of regional growth, some studies have supported that differences in social and institutional variables shape the growth rate per capita income of countries and consequently their convergence rates (Cofey and Polese, 1984; Hall and Jones, 1996; Helliwell and Putnam, 1995; Knack and Keefer, 1997; Putnam, 1993; Rupansingha, 2002; Temple and Johnson, 1998). These studies are examples of how
institutional features can explain growth in some regions, and added to other economic and social factors the results can elucidate the higher levels of efficiency and growth.

*Our proposal*

We start assuming that markets are socially constructed and economic life is rooted in networks of interpersonal relations; in which networks characteristics and properties - such as mutuality, trust and cooperation - are elements that can affect the economic outcomes. Secondly, the different actors-network can produce different forms of economic behaviour or decision making. Finally, the economy is shaped by collective forces that include formal institutions such as rules, laws, constitutions, property rights, as well as informal ones such as individual habits, groups’ routines, customs, traditions, social norms and values.

These collective forces are according to North (1991), *the institutions*: the game rules or the man-designed limits that determine the forms of social relationships and the incentives that modify human interaction. Institutions provide the framework of socioeconomic development, bring constraints and opportunities, influence all activities and define possibilities of growth. They reduce uncertainty in exchange, garner consensus and common arrangements and guide individual action.

In this research we define two kinds of institutional elements: a) *Soft Institutional Factors*: The characteristics of the networks of interpersonal relations that will be expressed by the following properties: mutuality, trust and cooperation. These also include individual habits, routines, customs, traditions, social norms and values. These are the socio-cultural characteristics in space, shared values and norms. b) *Hard Institutional Factors*: The long-lasting collective forces that shape the economy mainly identified with the formal institutional: rules, laws, constitutions, property rights.¹

In the first analysis that integrated institutional features (Barro, 1991; Easterly and Levine, 1997; Sachs and Warner, 1997), the interest was not to determine the contribution of these elements to growth. Later, as Janine Aron (2000) presents in an extensive comparative review, there is a series of recent works, which have tried to

¹ Because the term institutions could have a very broad meaning we focus the research of those institutions related to the economic strategies of local governments and the particular characteristics of their networks, such as civic behaviour, political preferences and polarization.
measure and assess the relationship of institutions and economic growth. This research has revised some of the works that had an explicit interest on institutional features classifying them into hard and soft institutional elements. In the next section we present the definition of the variables used for measuring these aspects (Section II), followed by the presentation of the results generated with Cluster Analysis and an analysis of the stability of the clusters in the section III. Section IV offers a balance of the main findings in terms of clusters, variables and growth. We have also included a brief note on some of the characteristics of Cluster analysis in the appendix.

II. Measuring Institutions and Economic Growth in Mexico

Having defined very briefly some of the features of the institutional approach, we explore the relation between institutional features and economic growth in Mexico. Particularly, our principal aim is to identify some institutional characteristics for Mexico that could be useful in the assessment of the regional economic growth. Can we explain the differences in regional growth using institutional features? Is there a simple positive connection between growth and institutions? Are there many institutional patterns and therefore different links with growth? Do less developed and rich regions have a particular set of institutional features? Have the institutional characteristics changed through time and space? How these changes are related to growth? These are the main questions of this research. Because the main purpose of this paper is to present the results generated by Cluster Analysis technique, the details of the methodology are included as an appendix. The next sections will focus on the definition of the variables and the results obtained.

Variables

Following the work of Putnam and Helliwell (1995), we include the density of associations, turnout in reference and the incidence of preference voting in order to evaluate some characteristics of the social capital of the regions. Similar to Rupansingha and Freshwater (2002) we also evaluate how associations (Olson-type and Putnam-type) affect regional growth, and how elements of ethnic diversity can also modified the performance of the regions.

---

For purposes of this analysis we present a classification slightly different from the one used in the previous paper (Decuir-Viruez, 2003). However, we use between 13 and 14 variables for the 32 observations (Mexican regions) in order to cover the main informal and formal institutional features. Assuming the new classification, the aspect we will evaluate are: a) Political Force of public institutions, b) Action Capacity of governments, c) Population Composition, d) Social Context, e) External links and d) Infrastructure Conditions. The aspects b and e are strongly related to the indicators of formal institutions, the rest (a, c, and d) give better information regarding the (social) informal links of the regions. The d) option refers to the physical condition of the regions. We briefly define the characteristics and sources in the following sections.

**Institutional Density**

This feature was including with the aim of observing the differences in the institutional stock of the regions, in one of its expressions, the importance of their local associations. This aspect is assessed as the number of associations per 10,000 habitants considering two types of associations, in accordance to their objectives: rent-seeking aims (Olson’s) or social network development (Putnam’s). Information was taken from the Economic Census of 1969, 1985 and 1994. The variables are identified as: (P0000) Institutional Density for Putnam’s type associations and (O0000) Institutional Density for Olson’s type associations.

**Political Force of public institutions**

Aspects of turnout and political preferences were also included as two aspects that can demonstrate the strength of political force of the local public institutions. Turnout is a proxy of the civic behaviour of the population, meanwhile voting preference could give an impression of which political option is more trusted by the people. The level of attendance used corresponds to the presidential elections in 1970, 1982 and 1994, published by Instituto Federal Electoral (IFE). The voting preference is the number of votes received by party 1 with respect to party 2. For these indicators we modified the concept of preference for trust. Trust in government refers to the ratio of votes received by the official party (PRI) over the votes received by opposition parties. For opposition

---

3 Putnam’s type associations include: the social, religious and handicapped assistance groups, civil organizations, social, recreational and sport associations. Olson’s type associations: entrepreneurial, commercial and production groups.

4 The identification of the variables is composed by one or two letters followed by two or four zeros. Two zeros correspond to an initial value and four zeros to the average values for a given period.
governments, we apply the same process with the ratio PAN/PRI votes. The information used covers the presidential elections in 1970, 1982 and 1994. The sources of this information were Instituto Federal Electoral (IFE) and Partido Acción Nacional (PAN) website. The variables are identified as follows: (TU00) Election turnout. (TG00 and TO00) Trust in Government (PRI) and Opposition, respectively.

**Action Capacity of governments**

The action capacity of the local governments has been identified with the size and management of Public Resources and also the degree of fiscal autonomy. This research proposes that a region which resources mainly come from central government –such as public expenditure and investment-, could be identified with a local government that has a low action capability. In contrast, it is expected that regions with less dependence of central government and more financial (fiscal) autonomy, will have a stronger capacity of action. This autonomy is given by taxes collection ability. In order to measure these aspects, the indicators proposed are:

(PE0000) Total Public Expenditure / GDP. Total expenditure of government.
(PE10000) Public Expenditure 1 / GDP. Expenditure of government -used for economic promotion and infrastructure purposes.
(PE20000) Public Expenditure 2 / GDP. Expenditure of government used for administrative objectives.
(PI0000) Public Investment / GDP. Total Public investment.
(TI0000) Tax Income/Total Local Public Income. Proportion of the local government income that is obtained through taxes;

With exception of TI indicator, all the variables are measured with respect to GDP for the periods 1970-2000, 1970-1985, 1985-1994 and 1994-2000. The source of these variables is INEGI.

**Population Composition**

The composition of the region defines an important feature, the degree of polarization modifies the social links and consequently it has an effect on in the informal and formal institutions. For this analysis we have considered the degree of polarization including the percentages of indigenous population and the size of the literate sector. These
indicators reflects the initial conditions of the population in each one of the periods, data is available from the population census for 1970, 1980 and 1990. The variables are defined as follows:

(IND00) Indigenous population. Percentage of population (5 years old and more) that speaks an ethnic language.

(LIT00) Literacy level. Percentage of population of 12 years old and more that is literate

Social Context
One of the characteristics selected for defining the social context of the regions was the insecurity level. This variable depicts the fragility of the social links which could impact the performance of the region. This indicator is given by the average of federal jurisdiction offences\(^5\) per 10,000 for the periods 1976-85, 1985-1992 and 1995-2000. This variable has been defined as (IS0000) for Insecurity Level.

External links
Despite the simple appearance that the ratio of Exports or FDI with respect to GDP can have, we decide to include both in order to have some indicator of the relationships of the regions with the abroad links. The changes suffered in the Mexican economy in the eighties justified to consider the modification of the external links. The variables are the average of exports and flows of FDI registered by states for the periods 1970-2000, 1970-1985, 1985-1994 and 1994-2000. FDI is only reported for the period 1994-2000. The identification of the variables is given by: (X0000) Exports/GDP and (FDI0000) Foreign Direct Investment/GDP. For this period we use data for exports from SIREM-Capem-Oxford, National Accounts, INEGI (1993-2000) and from the Secretary of Economy for FDI flows.

Infrastructure Conditions
Finally, we included the situation of infrastructure of the regions in order to have a more realistic scenario of the material conditions in which the economic activity of the regions is developed. The indicators are:

(U00) Urbanization. This is the percentage of population that lives in communities of more than 2500 habitants for the years 1970, 1980 and 1990.

\(^5\) The offences include all reported actions that violate health, security, communication, fiscal and private property laws.
(W00, E00 and D00). The availability of services of water, electricity and drainage, are the percentage of houses that count with these services for the years 1970, 1980 and 1990. Information was taken from Population Census by INEGI.

III. Cluster Analysis Results

In order to identify zones within Mexican territory with particular institutional characteristics, a series of clusters analysis were performed. This exercise was applied to the set of variables which were previously identified potential institutional features in different areas. The analysis divided the sample in three periods: 1970-1985, 1985-1994 and 1994-2000, it also included the estimation of clusters for the period 1970-2000.

In a first stage, the hierarchical cluster with Ward’s method and the squared Euclidean measure was used. The number of cluster estimated were defined by three steps, Firstly, a visual inspection of the dendrogram produced by hierarchical cluster analysis in order to have an impression of how many clusters (“the branches”) were generated given an arbitrary value for the fusion coefficient. The second step consisted in the construction of a scree test with the values of fusion coefficients and the number of clusters created at a particular coefficient. The final step involved the application of the Mojena criterion.

After obtaining the ideal number of clusters, the non-hierarchical method of K-means was applied on the 13-14 variables of each one of the four periods. In the following section, the main characteristics of the clusters are presented as well as their geographical localization.

Because one of the objectives of this research is to identify the links between these institutional features and growth, the results also include some indications of the rate of growth of the clusters members. Finally we add some comments with regards to the stability of the clusters membership and an overall balance of the results obtained in this stage of the research.

---

6 A more detailed explanation regarding cluster analysis is given in the appendix.
III.I Period 1970-2000

According to the Mojena criterion and the scree test for the fusion coefficients, seven is the ideal number of clusters for the K-means analysis for this period. Despite the use of Ward's method\(^7\); in the hierarchical method, the size of the clusters in the non-hierarchical method is not homogenous as we can see in the number of members each one has (See Figure 1).

The results show three special cases: Oaxaca (Cluster 1), Baja California (Cluster 2) and Distrito Federal (Cluster 6). Each one has very particular characteristics that make them unable to be clustered with the other cases.

For the Cluster 1, the size of the **government debt** and the **public expenditure** of Oaxaca are the characteristics that make it a special case among the thirty-two observations. The average of the Debt/GDP ratio is 5 times the national level for the period 1970-2000, meanwhile the Public Expenditure/GDP ratio is 3.9 times the national mean. (See Graph 1.1). These results suggest that the local government has been receiving **considerable resources from central government** throughout these years, at the same time such resources seem to be mainly used for **debt payment of this state**. Another flow of central government resources is the **public investment** that represents 0.7 times the mean. The role of export activity for the period 1970-2000 is below the average; therefore Oaxaca seems to have **been isolated of the export activity** of the country. In terms of the infrastructure provision, the initial situation in 1970 was also below the mean. The initial social links (1970) of this state were particularly defined by high polarization (given the level of indigenous population with the low levels of literacy) but at the same time, with positive levels of civic behaviour (turnout). Institutions, in terms of their densities, are below national means in 1971. According to

\(^7\) As it is mentioned in the definition of cluster analysis (See Appendix), Ward’s technique optimizes the minimum variance within clusters, which tends to create clusters of equal sizes and shapes.
the results of 1970, the political preferences favoured the PRI-government. Regardless of the disadvantages of **weak conditions of infrastructure and social and human capital**, Oaxaca managed to have a rate of growth of GDP above national for the overall period (1.9). After observing these results, could we expect that this growth is the result of the strong support received by central government in the form of **public expenditure and investment**? We will go back to this question at the end of the clusters analysis. As we will see in the following sections, Oaxaca’s situation has changed considerably throughout the three phases of analysis.

The second cluster: Baja California is characterized by high levels of insecurity (4.2 times the national mean) and Olson-type institutional density (3 times greater than national average). These two particular features show a controversial relationship: a very dangerous region with a good number of associations with economic-profit objectives. A third particular characteristic of this cluster is the initial strong preference for opposition parties, more than 2 times the national average (See Graph 1.2). In terms of the local government policies, the results show a **public expenditure** ratio slightly above the mean and a very diminutive positive difference of the investment used in economic promotion. The income collection capacity (Income/Tax Revenue Ratio) is below the mean.

The results illustrate a state with **a local government that limits its expenses, but still can not generate its own incomes**. Given the advantage of being located near to US markets, Baja California also had a positive level of export activity (1.4 times the mean). The initial conditions of the infrastructure were generally above the national standards, except the water service that was −1.5 times the mean availability. The initial social links of the cluster were characterized by good levels of literacy and low ethnicity, in a population with low participation in the elections but a high preference for opposition parties in 1970. The reality in terms of GDP per capita exposes a region with a rate of 0.9 percent, one of the lowest in the period 1970-2000. Contrasting with the conditions of Cluster 1, Baja California had greater advantages (**infrastructure, human capital, geographical position**) even though its insecurity levels. The growth of the region is one of the lowest for the period 1970-2000. A question arises here: is this rate related to a fragmented social capital in the region? We will analyses with more detailed the evolution of this region in the following sections.
As graph 1.3 shows, the capital city, Distrito Federal (Cluster 6) becomes a special case due to its notable high levels of exports and infrastructure accessibility, and also by its strong preferences for Opposition parties and high levels of Olson-type institutional density. Most of these indicators are at least three times above the national averages. Such conditions are expected for a region that concentrates the economic and political activities, therefore in this research; Distrito Federal is an outlier in most of the phases.
and generalizations could be puzzling. Its advantageous conditions have defined a high rate of growth (2.3 per cent) for the period 1970-2000.

Cluster 3 is composed by fifteen states located in most of the intermediate zone of the country. Their local government had in common low levels of public expenditure and investment, although the percentage used for administration is 0.1 times the national mean. The income generation capacity, measured by the income/tax revenue ratio is –0.1 times the mean. The exports indicator of the cluster is –0.4 times the average, this result suggests that exports were not crucial elements of the Gross Domestic Product of the members. In terms of infrastructure, the provision in 1970 of the availability services of electricity and drainage were 0.4 times and 0.2 times, respectively, below the standard. The initial social links of Cluster 3 had the following characteristics: scarcely levels of literacy, average ethnicity and turnout. Particularly, this cluster has a low level of insecurity. The institutional densities of these states were also under the national mean. The Olson and Putnam’s densities are lower (-0.5 and -0.2 times, respectively) than the national mean, as well as the preferences for PRI-government (-0.3). (See Graph 1.4). The rates of growth for GDP per capita of the members present a wide range of values. As we can see in graph 1.8, the cluster includes the states with the highest growth rates, but also three cases with the lowest rates. According to this result, most of the Mexican states did not present any of the following features: a strong central government influence through financial resources, a notable level of financial autonomy, an active export activity, a solid human capital, a high institutional density or a strong civic behaviour. They were “average states”, which growth results in very different rates of growth. However, these conditions are transformed throughout time and as we will see in the following sections their growth also changed.
Cluster 5 is composed by two states mainly characterized by high levels of **public expenditure and investment** (more than two times the mean). As graph 1.6 illustrates inside these flows, the size of the resources used in economic promotion and administration is 3 and 2.4 times the mean, respectively. The export activity is below the national levels; both entities have been segregated from this activity. Regardless of the planned resources for infrastructure development, the initial availability of water,
electricity and drainage services were below the standards. The initial institutional densities are also below the mean. The social links for these two states in 1970 were characterized by some degree of ethnicity and low literacy levels. Such population had an important participation in the elections of 1970 and its preferences were greatly supporting the PRI-governments, three times more than the national levels. A mixture of weak infrastructure, polarization, low institutional densities and important backing by central government supported by the population of the region correspond to two the lowest rate of growth of the period. The results in terms of growth per capita do not seem to match with the important financial resources received in this period.

Graph 1.5 Main Characteristics of Clusters.

The seven states that integrated the Cluster 7 have local governments that received flows of public expenditure and investment below the national mean, and also its uniqueness is the level of income-tax revenue above the mean. It is in fact the only cluster who has a positive ratio of income generation (0.7 times the national average), suggesting a certain level of financial autonomy. The second characteristic that distinguishes this cluster is the initial provision of infrastructure above the national standards (at least 0.8 times). At the same time, these states have initial level of literacy 0.8 times the average and low levels of indigenous population. (See Graph 1.7). The initial degree of institutional densities is slightly above the mean, but it is not one of the main characteristics of the group. Despite of the homogeneity of the population, the participation in election is below the mean and the group that decided to vote in 1970
had less preference for PRI-government with respect to the national mean. The main features of this cluster describe a region with some level of financial autonomy and more independence of central government resources; a good starting level of human capital and infrastructure; some initial degree of institutional density and a social capital composed by a no polarized population with low civic behaviour. These qualities correspond in terms of economic growth to rates above the national average.

Graph 1.6 Main Characteristics of Clusters.

In balance for the period 1970-2000, we observe that the clusters that have reached the highest rates of growth are related with cases that have important flows of public investment (Cluster 4). However, we have to consider that some of the members of these groups are the main oil-producers. In the other extreme, those clusters with the lowest rates of growth correspond to cases characterized by some positive levels of public expenditure, low exports and financial autonomy (Clusters 3 and 5). The Clusters 3 and 7 are two contrasting groups in the way the public finances are managed; the first has some positive level of public expenditure above the national mean but a negative level of financial autonomy. The cluster 7 has low levels of public expenditure and positive levels of income-tax revenue ratio. In terms of initial infrastructure accessibility, we detect cluster 7 in better conditions than cluster 3. Same situation is found for institutional densities. There is a stronger preference for opposition in cluster 7 than in 3 in the 1970 elections. The results also show that polarization of population
(ethnicity) is higher in cluster 3. Having these starting conditions in 1970, the results in terms of growth reveal that some members in both clusters reach similar high rates but the lowest rates are more likely to occur in Cluster 3, as can be seen in Graph 1.8. In the following section, we will expose the principal changes of these characteristics and how they related to the growth obtained.

**Graph 1.7 Main Characteristics of Clusters**

**Graph 1.8 Mexico: Growth Rates and Clusters. 1970-2000**
III. II Period 1970–1985

According to Mojena criterion the number of clusters used for this period is six, three of them resulted to be individual cases: Baja California, Distrito Federal and Quintana Roo. Similar to the whole period analysis, the size of the clusters is very different. The analysis generated two groups of five and one of nineteen observations.

The capital city, Distrito Federal, is again a particular case (Cluster 1), characterized by a local government that has a total public expenditure one time above national level, which is used in administrative and economic objectives but below national standards (see Graph 2.1). At the same time, public investment level is below national average but also the ratio of income-tax revenue. This last characteristic suggests that the capital city is not a financial autonomy state in the period 1970-1985. Export activity for this period is below the average. A particular feature of the cluster is the good provision of infrastructure (above 2.4 times national standards). Other special features are the type of initial social links it has, characterized by good level of literate population with low ethnicity with strong preferences for opposition parties (3.30 times the mean). It is also unique of this cluster, the positive level of Olson-type institutional density, -those associations devoted to rent-seeking activities-, which are 2.90 times national mean

The second outlier is Baja California, in Cluster 2, due to its high density of crimes per 10’000 habitants and the high level of Olson-type density (4.7 and 3 times national average, respectively). The management of public resources is defined by a high level of public expenditure in administrative objectives (above 2.60 times national average), but with levels below the mean for public investment and income-tax revenue ratio. With exception of water service, the rest of infrastructure indicators are above national standards. According to the levels of literacy and indigenous population is not a polarized zone; however there is a low participation in the elections but a high
preference for opposition parties. In 1970-1985 the performance of the national economy was considerable high, reaching an average of 2.9 per cent in per capita terms; however Distrito Federal and Baja California were below this rate (1.6 and 1.0 percent, respectively)/
Quintana Roo (Cluster 4) is characterized by an **impressive level of exports** (4.5 times national average) in the period 1970-1985. The local government of cluster 4 has levels of **public expenditure and investment above national average**, and also low levels of financial autonomy (see Graph 2.3). The initial availability of infrastructure is below national standards. The social links of this state correspond to a zone with high levels of polarization (ethnicity); in fact, it is the region with the greater proportion of indigenous population in the period 1970-1985. This population had an important participation in
1970 with a turnout level 2.70 times the mean and a preference for PRI-governments. At the same time, Putnam-type density is 3.20 times the national level and there is some degree of insecurity. The growth of GDP per capita for Cluster 4 was 2.9 per cent, equal to the average of this period.

Cluster 3 integrated by five members are characterized by a low civic behaviour of a population that according to graph 2.4 have less preference for PRI-governments. The public expenditure levels are slightly above national average, although financial autonomy is below it. The infrastructure conditions are relatively above national standards. Export activity is below national average; despite of their closeness to US markets as it is shown in figure 2. The social links of the cluster are defined by a nonpolarized population, particularly well-literate. The institutional densities have some positive levels with respect to the mean (above 0.50 times the mean). Three of the members of this cluster reported some of the lowest rate of growth in this period (Baja California Sur, Sonora and Nuevo Leon), although they are not negative rates. This group shows that despite of the geographical advantages and good infrastructure, the growth is slower than the national level in the period 1970-1985.

The five states that composed Cluster 5 are predominantly distinguished by a positive preference for PRI-government (1.80 times the national average) and also by some positive level of debt. The local governments of the member had slightly positive levels of public expenditure and low financial autonomy. However, the public investment is 1.20 times above national average. The initial conditions of infrastructure are below the mean, as well as the development of exports. In 1970, the population of this group was polarized, given the important proportion of indigenous people and the low level of literacy; however their civic behaviour was positive. The institutional densities of this period were both below average. These characteristics correspond to a cluster that had the highest rates of the period 1970-1985; it is a phase in which the less advantaged states are converging. This convergence seems to be supported by the flows of public investment received during 1970-1985.

---

8 Tabasco and Campeche had an extraordinary growth due to the oil production boom happened in this period.
In Cluster 6, the nineteen members show a positive level of financial autonomy with levels below average of public expenditure. As graph 2.6 shows, most of the indicators are below the mean. There is no evidence of important flows of public expenditure or investment going towards these states, neither of an important export activity. These
states have average infrastructure accessibility. The population of these zones is not mainly indigenous, but it does not show to be a high literate group. Institutional densities are below the mean. Such results suggests that most of the Mexican states were having these features during the period 1970-1985 resulting in most of them in intermediate rates of growth as we can see in graph 2.6.

The period 1970-1985 is a convergence phase, the states that reported the highest rates of growth are found in cluster 5, which main characteristics correspond to local governments supported by important levels of public expenditure and investment. The political preference of this region is strongly oriented towards PRI. The growth of this region in the period 1970-1985 was not related with export activity, institutional densities or infrastructure, indicators that are below national standards. In fact, for two of the members (Tabasco and Campeche) it was strictly related to the oil-production boom occurred in these years. Cluster 6 includes states with intermediate rates of growth which expose the development of some degree of financial autonomy, but the rest of the characteristics are slightly below national averages. The states that reported the lowest rate of growth correspond to regions that even with good initial infrastructure levels and institutional densities; have a positive level of public expenditure used for administrative objectives (cluster 2 and 3).
III. III Period 1985-1994

The number of clusters calculated for this period was five, resulting in only one special case: Oaxaca. The reason that made Oaxaca an outlier is given by the size of its **debt and public expenditure** compare with the national level (5.4 and 2.8 times, respectively).

The level of public investment in this zone is 0.9 times the mean. (See Graph 3.1). Population of Oaxaca in 1980 is polarized due to its level of indigenous population and low level of literacy. It lacks of a good infrastructure and the level of exports is also below national standards in 1980. Densities of institutions, turnout and preference for opposition are at least 1 time less than national mean.

For the period 1985-1994, the increasing tendency of Oaxaca’s GDP per capita growth collapsed and reported a negative rate of 1.6 per cent; its ranking was among the last eight regions. The particular situation of this cluster shows that between periods 1970-1985 and 1985-1994 the size of the debt increased considerably, meanwhile the polarization of population was intensified and the institutional and infrastructure levels became more distant from the national standards. The turnout level of the population in 1982 seems to have reduced as well as the preference for PRI-governments.

The twelve states that composed Cluster 1 have levels of total public expenditure and debt slightly below the mean, as well as public investment and income-tax revenue ratio (Graph 3.2). However, the proportion of the public expenditure used for economic promotion was slightly greater than the mean. Infrastructure provision is not at the national standard as graph 3.1 shows. The social links of the region share a particular low level of literacy and ethnicity. Institutional densities, turnout and preferences for opposition parties are below the mean. Ten of the members of this cluster were grouped in cluster 6 in the phase 1970-1985. According to the characteristics of the previous phase, the conditions of these members in 1980 seem to have abandoned the
development of a financial autonomy and the conditions of the infrastructure provision worsened. The period 1985-1994 was one of the most difficult for the Mexican economy, the rate of growth of GDP per capita was -0.8. The members of Cluster 1 were not exempted of this tendency; half of them reported rates of growth below the national average.

Graph 3.1 Main Characteristics of Clusters

1985-1994
Cluster 4: Oaxaca

Graph 3.2 Main Characteristics of Clusters

1985-1994
Cluster 1: Durango, Guanajuato, Guerrero, Hidalgo, Michoacán, Puebla, Querétaro, S.L.P., Sinaloa, Tlaxcala, Veracruz, Zacatecas
Cluster 2, formed by ten states, is distinguished by a good provision of infrastructure (1.2 times the national standard) and good level of literacy (0.8 times). A feature that differentiates this group is the level of exports (0.8 times). The performance of local government has the particularity of some level of financial autonomy (0.7 times), but also levels of public investment and expenditure below the mean (-0.3 and -0.6 times). The political preference tended to favour the opposition; however their participation in 1982 elections was below national level. In terms of institutional densities, Putnam-type is a little higher with respect to the mean (0.4). Five of the members of Cluster 2, used to be members of Cluster 6 in the period 1970-1985, with most of the indicators below the mean.

Comparing those results with the ones attained in the period 1985-1994, we observed that those members have improved in several areas: infrastructure, export activity, literacy and Putnam-type densities. The effort for a financial autonomy is still an important feature in the eighties, besides there is a strong definition of preferences for opposition parties. Eight of the members of this clusters had rates of growth of GDP per capita above national rate, and four of them were in the top five states (D.F. Aguascalientes, Chihuahua and Morelos).

Cluster 3 is defined by 7 states with positive levels of Putnam and Olson types institutional densities (1.5 and 0.8 times national mean), also by a preference for PRI-governments and some level of insecurity. The dimension of public resources (public expenditure) is slightly below the mean, but it is positive for public investment and income-tax revenue ratio. Export activity is not above the mean. Social links correspond to a region with some level of indigenous population but at the same time a positive level of literacy in the beginning of the eighties. Electricity service is the only indicator of infrastructure above national standards. The range of variation of the rate of growth of GDP is wide; it includes the highest and the lowest observations of the period (Quintana Roo and Campeche). Within this group there are five states growing at a higher rate than the national one. The element that gives the impression of being supporting the growth of the cluster is the public investment and some degree of financial autonomy, but there is no evidence of an important export sector.
Cluster 5 includes two states: Tabasco and Chiapas. In this period, the high level of public expenditure used in economic promotion and administration (2.7 and 1.8 times the national mean) is the characteristics that make these two observations a particular cluster. In the same way, public investment size is 2.3 times the mean. Even with this investment in the region, the infrastructure of the zone is still below the national mean.
standards in 1980. In the same way, export sector importance is below the mean of 1985-1994. The social links of the region are given by low literacy and high ethnicity, as well as low institutional densities. The preference for PRI-government is higher than national levels in 1982. Both states had the lowest rates of growth in the period 1985-1994. The strong support by central government to these states seems not to be correlated with a high growth rates for this period.

Graph 3.5 Main Characteristics of Clusters

Inside the members of Cluster 2 and 3 there are some of the cases with higher rates of growth of GDP per capita. These two cases show two types of strategies of economic development. The similarity among these two sets is the financial autonomy both are developing but that contrast with the level of public investment each one has. For Cluster 2, public investment is below the mean; meanwhile cluster 3 has slightly positive levels of this indicator. Another feature is the high institutional density for cluster 3, followed by cluster 2 in magnitude. Two particular differences are the preferences of a no-polarized population in cluster 2 for opposition parties; meanwhile a polarized group still prefers PRI-governments. A final common element that seems to support the performance of these two clusters is the provision of infrastructure. As can be seen in graph 3.5, cluster 5 is supported by important flows of public expenditure and investment, but it does not present a good infrastructure.
III. IV Period 1994-2000

The number of cluster estimated for this period were five, two of them are special cases: Tabasco (Cluster 4) and Oaxaca (Cluster 2). The characteristics that make Oaxaca an outlier observation are again the high difference of debt and public expenditure with respect to the national mean (5.4 and 3.8 times). Such values are inclusively higher that in the period 1985-1994. In the same way, the proportion of public investment 0.8 times the mean. Other features for the nineties are the high levels of ethnicity (2.6 times the mean), poor infrastructure and low export development and Foreign Direct investment. Although the low participation in elections, the population still prefer the PRI-governments. The institutional densities (Putnam and Olson types) are below the national levels, as well as the participation in elections. (See Graph 4.1). The rate of growth of Oaxaca for this period was slightly below the national rate (1.2 %). The results suggest that the policy of economic development of the zone has been
supported by public expenditure during the three phases, and by some decreasing levels of public investment after 1985. The constant feature along these three periods is the considerable level of debt of this state. The infrastructure conditions have not reached the national standards, meanwhile the polarization of the population is more intense and the densities of the institutions are below the mean. A revision of the growth of GDP per capita in these three phases shows that it was in the period 1970-1985 in which Oaxaca had the fastest rate, stage in which it received important flows of public investment.

The second outlier is Cluster 4, this case is distinguished by high positive levels of public expenditure and investment (1.9 and 3.9 times national level), however as graph 4.2 shows, the proportion of resources used in economic promotion is considerable bigger than the mean during the period 1994-2000. This result can be explained by the flows of resources that central government has used in the oil-production projects in this zone. According to the values of clusters centers, Tabasco’s public expenditure for economic promotion and infrastructure is 4.3 times the national mean. However, the availability of water and electricity is below national standards in 1990. Exports and Foreign direct investment flow do not have an important role in this case. The political preferences of the region are depicted by a strong preference for PRI-governments (3.9 times the national average). (See Graph 4.2). Contrasting with the considerable resources received from central government, the rate of growth for Tabasco in the period 1994-2000 has been -0.2. The main reason behind the development of this state is the oil-production boom, event that had its peak in the seventies. Most of the infrastructure and institutional conditions remained similar to the standards in 1970.

Cluster 1 composed by three entities, two of them with a rate of growth higher than national average for this period (Baja California and Nuevo Leon). As the graph 4.3 illustrates, the main characteristics that defined this group are high levels of Foreign Direct Investment, exports and infrastructure availability. The social links for 1990 are distinguished by low ethnicity and high literacy levels. For the same year, the institutional densities are slightly positive for Putnam-type and negative for Olson-type, although there is some level of insecurity. In terms of political characteristics, the preference for opposition is 1.1 times greater than the national average, and also there is a positive participation in elections of 1994. The indicators of the performance of local
government suggest levels of public expenditure below national mean and with an average degree of financial autonomy.

Graph 4.1 Main Characteristics of Clusters.

Graph 4.2 Main Characteristics of Clusters.
The nineteen members of Cluster 3 have in common positive levels of institutional densities and financial autonomy, as well as a good infrastructure provision above national standards. (Graph 4.4). The performance of their local government is identified by low levels of public expenditure and investment, contrasting with the positive level of income-tax ratio. The export activity is below the mean and the amount of FDI receive is equal to national level. The social links correspond to a region with no polarization, and the political preferences are defined towards opposition parties. The range of the rates of growth of the members is wide, from Quintana Roo with -1.4 percent to Queretaro’s rate of 3.9 percent. Twelve members of this cluster report rates of growth higher than the national average for this period.

In Cluster 5, there are nine states identified by levels below national average of infrastructure, literacy and financial autonomy. The local government performance has levels of public expenditure and investment slightly above the mean, particularly the expenditure is used for administrative objectives. The institutional densities are also below the mean of the period. There is some preference for PRI-government in the fairly polarized population (0.6 times above the mean). The growth of the member oscillates between 0.0 (Campeche) and 5.5 (Puebla ). Four of the observations reach the highest rate of growth, but not as high as in Cluster 3.
In balance, the highest rates of growth observed in this period correspond to the members in cluster 3, which are notable by the high levels of financial autonomy, institutional densities and infrastructure, but also by levels of public expenditure and investment below the mean. Most of the observations in cluster 3 have the highest rates of growth of the period; however, there are some members in cluster 3 that also have the
lowest rates. In particular, Quintana Roo, member of this group, has the most dramatic fall in terms of economic growth. The characteristics of Cluster 5 contrast with Cluster 3, the positive rate of growth observed for this period (although most of them below national mean) are identified with states that have important levels of public expenditure and investment. As graph 4.6 shows, the range of the rate of growth for cluster 3 is wider than cluster 5, but also it is evident that the higher rates are reached by the members of cluster 3.

**Graph 4.6 Mexico: Growth Rates and Clusters. 1994-2000**

In synthesis, the results generated by these series of cluster analysis show:

a) The high growth rates in the period 1970-1985 correspond to members that belong to clusters that received important levels of **public investment** and with some level of **financial autonomy**. The low rates belong to the members of a cluster that have some positive levels of **public expenditure**, mainly used in the administration of their government. Institutional densities are not strongly developed in the states with the highest rates of growth.

b) The high growth rates in the period 1985-1994 are linked to observation that are developing a **financial autonomy**, have good **infrastructure** and some level of
**institutional density.** However the size of public investment in these regions could be above or below the mean. The lowest rates correspond to regions that received important flows of **public investment** but that lack of **good infrastructure accessibility**.

c) In the period 1994-2000, the high rates of growth correspond to states with high levels of **financial autonomy, institutional densities and infrastructure**. The best rates do not belong to the states that export more or have received important resources of FDI. The regions with important levels of **public expenditure** seem to be the regions with slower growth.

d) An overall analysis of the period 1970-2000 illustrates that the members with the highest rates of growth of GDP belong to the clusters that: 1) have developed throughout the period some level of financial autonomy, 2) have a good infrastructure and human capital stock and 3) have some positive densities of institutions. The groups in which the size of public investment and expenditure is above the mean reported the lowest rates of growth for the whole period. Finally, there is a large group of observations with characteristics equal to or below the national average that did not reach any of the peak rates.

**III. V. How stable are the memberships of the clusters?**

A comparative analysis on the memberships for each period can illustrate how each one of the thirty-two states has performed. Some cases, sharing similar characteristics, move together during the whole period, but some other observations remain isolated or temporarily attached to bigger groups. The movements of the cases from one cluster to another can give some evidence of the evolution in the characteristics proposed at the beginning of this research. The changes could illustrate improvements or backward movements of the observations.

According to the changes in the membership in these thirty years, figure 5 shows the existence of three main groups of observations. The main group (left side) is integrated by nineteen states all identified by cluster 7, a group particularly characterized by some degree of financial autonomy. The middle group includes three clusters, which have in common the size of their public expenditure (above the mean), some degree of
infrastructure accessibility and institutional densities. As it was mentioned before, Baja California and Distrito Federal are independent cases due to their considerable levels of insecurity and infrastructure provision, respectively. The members of Cluster 4 and 5 (right side) shared the peculiarities of important levels of public expenditure and investment and the high degree ethnicity of their population. However, Quintana Roo becomes an isolated case due to the impressive export activity reported in this period.

Figure 5. Members of Clusters in the period 1970-1985.

In the period 1985-1994, the first change we observe in figure 6 is the separation of cluster 7 (1970-1985) in three sets: Clusters 1, 2, and 3. The second important movement is also the division of the middle group, whose members are attached to the clusters created in the main group. These divisions illustrates how members of cluster 1 became less able of generating a local income through taxes, the indicator of financial autonomy is smaller in this period for these set of observations. In contrast, Cluster 2 and 3 keep elevating this indicator and also reducing the size of their public expenditure. The elements that separated these two clusters are found in the contrasting political preferences, the support of public investment and better institutional densities in cluster 3, the less degree of ethnicity and the better infrastructure in cluster 2. From the middle group, we observe Durango moving to Cluster 1 in the period 1985-1994, suggesting a less level of financial autonomy and at the same a general decline in the values of the rest of the indicators. This same tendency is followed by Hidalgo. The
members of last group (left side) also modify their memberships. Campeche and Quintana Roo, despite of their high level of ethnicity, this group resulted to be clustered to group 3. This change is supported by the good levels of institutions, the positive levels support of public investment and at the same time, the reduction of public expenditure. Chiapas and Tabasco remained isolated due to their levels of public investment and expenditure, high ethnicity, weak infrastructure and low densities of institutions. Oaxaca is kept as independent cluster due to its level of public expenditure and particularly, the size of its debt.

The main changes in the last period (1994-2000) show how the main group of states of 1970-1985 is reintegrated; however there is a subgroup of states that seem to follow a diverse type of economic strategy. Cluster 1 from period 1985-1994, with indicators below the mean, is divided in two groups: the first one by Guanajuato, Queretaro, Tlaxcala and Sinaloa; the second by: Michoacan, Puebla, Guerrero, Veracruz and San Luis Potosi. The first group becomes part of Cluster 3, a change that implies that these states developed some degree of financial autonomy, reduced the sizes of public expenditure, improved infrastructure conditions and increased the institutional densities. A similar change must have occurred to the member of cluster 3 in 1985-1994: Nayarit, Tamaulipas, Yucatan, Baja California Sur, Colima and Durango. Additional to these
changes it seems that the preference for opposition in these regions also increased. Another important change is the separation of Nuevo Leon, Baja California and Distrito Federal due to their high levels of exports and infrastructure.

As we can see in figure 7, cluster 5 used to be part of the main group and for the period 1994-2000 it also includes three states from the group on the right: Campeche, Chiapas and Hidalgo. The common feature of this group is the positive levels of public expenditure and investment, with low financial independence, high degree of ethnicity and low conditions of infrastructure and institutional densities. The preference of this group tends to favour the PRI-governments. Tabasco could be part of this group, but due to its high level of public investment it becomes an independent observation. A similar situation happens with Oaxaca, which is maintained as a special case due to its level of debt. Meanwhile, Quintana Roo managed to be part of cluster 3.

According to the changes in the memberships, the states in the main group (left-side) in figure 5 could be considered the more stable clusters, in which we could identify at least four groups of states with similar characteristics changing together through time. The middle group is less stable at least each member has modified its membership once. The final group (right side) is not stable; half of the members have very particular characteristics that make them unable of being clustered with others.

**Figure 7. Members of Clusters in the period 1994-2000.**
It is important to mention that clusters 5 and 3 in 1994-2000 have a particular interest: they expose the most contrasting results in terms of public resources management and their related rates of growth.

**IV. Conclusions: Principal Findings from Cluster Analysis**

Assuming an institutional dimension on growth, we introduce a set of variables (13-14) that capture some institutional feature for the thirty two Mexican regions for the period 1970-2000. This data base was evaluated through the analysis of clusters in order to identify particular patterns of institutional characteristics, including formal and informal aspects. The clusters obtained were linked to the rates of growth reported for four periods 1970-2000, 1970-1985, 1985-1994 and 1994-2000, observing the following characteristics.

- In the period 1970-1985, the states that reported the highest rates of growth were members of cluster which had important flows of public investment. However these resources were directed to zones with high levels of ethnicity, low literacy, low conditions of infrastructure and institutions. It was the period in which the less advantaged states converged towards national standards. The second important observation on this period is the fact that most of the Mexican states exposed only a slightly positive degree of financial autonomy but that performed below in the rest of indicators. Their results in terms of growth correspond to intermediate levels. The cases with the lowest levels of growth belonged to member of clusters characterized by some level of public expenditure, without any relevance on the fact of having or not a good infrastructure or high institutional densities. It is also relevant to mention that during the period 1970-1985 the regions in the north border presented the slowest rates of growth, despite their geographical advantages and good conditions in terms of infrastructure, social and human capital. We can describe the performance of the less advantage regions –mostly in the south- : Campeche, Chiapas, Hidalgo, Oaxaca and Tabasco, like the **most active regions** of that period, given the speed of the rate of growth of their GDP per capita. On the other side, we can refer to the rest of the cases and particularly to the northern states as the **lethargic regions**, because despite their better conditions for the economic activity, the
main model ruling the economy was been debilitated and becoming not adequate for exploiting their potential.

- For the period 1985-1994, the highest rates of growth belong to two groups: (Cluster 2 and 3), each one with particular characteristics. One of the clusters kept a positive financial autonomy trajectory, and received fewer resources of public investment and expenditure. These regions became part of the export sector and already started with a good infrastructure, well-literate human sector and a high institutional density, particular Putnam-type. They defined their political preferences for the opposition. The second cluster showed a preferences PRI-government, had some institutions, some degree of ethnicity and also a literate population. This cluster is not part of the export sector that is being developed in the eighties. The intermediate rates correspond to members that are clustered in groups that use some positive levels of public expenditure for economic promotion, regions that are less financial autonomy but used to be in the previous period. They show infrastructure conditions to be worsened. The low rates are reported by regions which received important flows of public investment and expenditure; however they did not have any improvement in the rest of the conditions. The conditions of these zones show an increasing ethnicity level, difficulties in the local government such as financial debt and weaker infrastructure provision.

- For the most recent period (1994-2000), the highest rates of growth belong to members of clusters with high levels of financial autonomy, institutional densities and infrastructure. These regions also have low levels of public expenditure and investment. They are not the main regions connected to the flows of exports or FDI. Their preferences support the opposition parties. The next group in terms of growth is formed by cases that have slight support for central government through public expenditure and investment, and a level below the mean of financial autonomy. Their population —some how polarized— still prefer PRI-government. Meanwhile, the observations with the lowest rates are identified with regions with high levels of public investment and expenditure. The formula of public investment is ineffective in terms of growth promotion; there is no evidence of improvement in infrastructure or social
conditions, these circumstances make difficult to develop a strong institutional density. Despite their constant political preference, its civic behaviour is being reduced. The general dynamic of the country, exposes this last region as a *lethargic* type in which the conditions of 1970 seem not to have changed.

In terms of the variables used the main findings in this process were the following.

- The use of Public Investment and Expenditure were part of the regional development policies, which resulted successful for the southern regions in the period 1970-1985. Its effectiveness in terms of growth seems to have exhausted in the last two periods. These resources do not seem to be materialized in better conditions of infrastructure or human capital in these last regions.

- High Institutional densities seem to coincide with cases that have developed financial autonomy. The development of Putnam associations is more important in the eighties and Olson in the nineties. Meanwhile, regions with strong public expenditure and investment have densities below the national mean. Regions with low infrastructure have low institutional densities.

- High Turnout is found in the seventies and eighties in regions with polarized population, in contrast low turnout belongs to more literate population. In the nineties, the literate population participates more in elections and polarized regions reduces their turnout.

- The preferences for opposition correspond to regions with some degree of financial autonomy. Meanwhile, the preference for PRI-government belongs to regions that receive important flows of public investment or have an important size of the public expenditure. In the seventies high turnout coincided with high polarized regions that preferred PRI-governments.

- There is a particular positive relation among insecurity and Olson-type institutional density.

- Indigenous population is more persistent in southern regions.

- Literacy improves for some more financial autonomy regions in the eighties and nineties. It does not improve in the south (Oaxaca).

- Exports patterns are different in the three periods: the beneficiaries regions are different. Seventies is dominated by Quintana Roo’s exports, and the most
recent period corresponds to two urban agglomerations and one of the northern states.

- The infrastructure seems to have improved for the regions that developed some degree of financial autonomy. The conditions become weaker in the second period for those who are less financial independent.

The results suggest that there is a transformation on the management of the public resources in the last thirty years in most of the Mexican regions, which have some connections with the trends of growth. In contrast, exports and foreign direct investment links are not important determinants of the highest rates; however they tend to intensify differences among regions. Infrastructure is a key element in the success of the financial autonomous regions. This same variable exposes some important questions when we observe that huge levels of public resources have not been materialized in good levels of infrastructure or better conditions of the population. Finally, institutional densities become a strong indicator after the regions have developed some degree of financial autonomy and a positive growth. The role of Putnam and Olson types associations in seem to be overshadowed by the important force that the Mexican central government exerted in the last thirty years.

Finally in terms of the typology of the Mexican regions we observed in overall – considering all periods- four categories: regions with have developed a certain degree of financial autonomy and reduce their dependence of central government resources in the last thirty years, which have improved their material and institutional conditions of growth and have changed and reinforced their preferences for opposition parties. A second type is formed by regions that abandoned (or were unable of keeping) the financial autonomy, and still are supported with some level of public resources and not integrated to the exports and FDI flows. A third type of regions integrated to the external links and strongly supported by its infrastructure and geographical positions. A fourth type formed by regions strongly dependent of public resources but with material and institutional conditions that have not changed since 1970. For the period 1970-2000, the members with highest rates of growth of GDP per capita can be found in the first group; meanwhile the lowest rates are identified with the fourth type.
References

Appendix

I. Methodology: A Note on Cluster Analysis

General Characteristics

Cluster analysis, also called segmentation analysis or taxonomy analysis, refers to a complete series of quite divergent statistical methods which seek to identify homogenous subgroups (clusters) of cases in a population. This multivariable technique aims to group elements or variables trying to achieve the maximum homogeneity within a group (minimum variance) and the maximum heterogeneity between groups (maximum variance). This group of objects with a strong similarity among them are internally united, meanwhile, the different groups have a distinctive behaviour, which means, and each group is externally isolated of the others.

Cluster analysis is an exploratory technique because it does not use a statistical model for classifying. However, it is very useful for extracting information about a group of data without previous and explicit restrictions, and then it could be very practical as a tool for creating hypothesis about the topic. The obtained classification of the multivariable data results valuable for understanding the data and the original population they belong to. Aldenderfer (1984) formally classifies the four principal goals of cluster analysis in: a) development of a typology or classification; b) investigation of useful conceptual schemes for grouping entities; c) hypothesis generation through data exploration, and d) hypothesis testing or attempt to determine if types defined through other procedures are in fact present in a data set. Therefore, in our research we use cluster analysis with the aim of finding a typology of institutional features in the 32 Mexican regions and at the same time to test if there are relevant types defined, that in subsequent phases of the research could be use to test whether there is a connection with differences in economic growth.

According to Aldenderfer (1984) there are five basic steps that characterize the cluster analysis studies. The first one is the selection of a sample to be clustered. The second and probably the most important step is the variables selections in which the grouping will be based on, variables should describe the similarity between groups in relevant terms for the research questions. The selection could be base on previous research, theories or a new hypothesis to be tested. An important feature in this step is standardisation of the variables, a crucial step in this analysis, in order to avoid dispersion and increase similarity effects. The third step is the computation of the similarities among the entities. The fourth task if the creation of clusters and finally the validation of the resulting cluster solution.

What is a cluster?

---

9 This section is mainly based on the works developed by Aldenderfer and Blashfield, (1984), Everitt (1980) and Figueras (2001).
10 Different names are used for the object that is classified, such as entities, case, pattern, operational taxonomic unit (OTU). The aspects of the objects used for assessing their similarities are called variable, attribute, character or feature. Aldenderfer (1984, p.16)
Clusters are defined by Everitt(1980) as “continuous regions of a space containing a relatively high density of points, separated from other such regions by regions containing a relatively low density of points”. Sneath and Sokal(1963) characterized a cluster by the following most important properties:

1) Density, it is a relatively thick group of data points in the space.
2) Variance, it is the degree of dispersion of the points in the space
3) Dimension, it helps to measure cluster by radius.
4) Shape, It is a simply display of points in the space
5) Separation, a cluster needs to be separated from other cluster.

**Measures of similarity: Distances**

Because the clusters groups similar objects, it is important to evaluate the differences and similarities among objects. The similarity\(^{11}\) is a measure of correspondence among the elements that are going to be grouped. This is measure in terms of distance between a pair of elements. Those observations with the smallest distances are more similar between them than those with greater distances; therefore they will be grouped in the same cluster. The distances are presented in a table in which both rows and columns are the units of analysis and the cell entries are a measure of similarity or distance for any pair of cases. The distances measures that could be used are: correlation coefficients, distances measures as Euclidean distances, association measures and probabilistic similarity measures. We present some of the characteristics and examples of the first two types of measures.

\( a) \) **Pearson Correlation coefficient** measures the lineal association between two observations and it is defined as:

\[
 r_{jk} = \frac{\sum (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum (x_{ij} - \bar{x}_j)^2} \sqrt{\sum (x_{ik} - \bar{x}_k)^2}}
\]

where \( x_{ij} \) is the value of the variable \( i \) for the \( j \)-observation, \( \bar{x}_j \) is the mean of all values of the variable for \( j \)-observation and \( n \) is the number of variables. The value of the coefficients varies from -1 to +1, a value zero means no relationship between cases. Some limitations of this coefficient are its sensitivity to dispersion and elevation on profile data.

\( b) \) **Distance measures**

**Euclidean distance** is defined as the straight line distance between two points; it is estimated in the following way:

\[
 d_{ij} = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2}
\]

\(^{11}\) Besides the similarity or proximity measures, cluster analysis could also use dissimilarity or distance measures. The former measure the similarity degree between two objects, the greater the value is, the greater the similarity between them, the greater the probability that they are assigned to the same group. In contrast, the latter the greater the distances are, the more different are the two objects, consequently the less probability that they belong to the same group.
where \( d_{ij} \) is the distance between observations i and j, and \( x_{ik} \) is the value of the k-variable for the i-case.

This measured could be squared in order to place progressively greater weight on objects that are located further apart.

Other types of distances used are:

*City-block (Manhattan) distance*, defined as the average difference across dimensions, in other words, the rectilinear rout measured along parallels to X and Y axes. It is calculated using the following expression:

\[
d_{ij} = \sum_{k=1}^{p} |x_{ik} - x_{jk}|
\]

*Minkowski metrics* are a family of distance measurements which are generalized from the Euclidean distance formula, in which \( r \) is the Minkowski parameter:

\[
d_{ij} = \left( \sum_{k=1}^{p} |x_{ik} - x_{jk}|^{r} \right)^{1/r}
\]

*Mahalanobis D2* is a normalized distance in the normal distribution from the center to a point (X) in case of n dimensional normal distribution. \( d_{ij} = \left( X_i - X_j \right) \sum^{-1} \left( X_i - X_j \right) \)

where the \( \Sigma \) is the pooled within-groups variance-covariance matrix, and \( X_i \) and \( X_j \) are vectors of the values of the variables for observations i and j.

Euclidean and other distance metrics have the disadvantage of being affected by elevation differences in the data, in order to reduce the effect; the standardization process becomes appropriate in this step.

c) *Association coefficients*

This measure establishes similarity between observations described by binary variables. *Jaccard’s coefficient*, a matching coefficient, is defined as:

\[
S = \frac{(a)}{(a + b + c)}
\]

where S is the similarity between the two observations with values of 0 and 1. It ranges from 0 to 1.

**Types of Clustering**

According to Aldenderfer (1984:35) there are seven major families of clustering methods:

1) hierarchical agglomerative
2) hierarchical divisive
3) iterative partitioning
4) density search
5) factor analytic
6) clumping, and
7) graph theoretic

We will briefly present some characteristics of the first and third families.
The hierarchical agglomerative family

The hierarchical technique starts assuming each observation as a cluster and subsequently groups the most similar observations, those with the smallest distances among them.

Figure 1. Dendrogram

The easiest hierarchical agglomerative technique is single linkage, it starts searching for the two most similar entities in the matrix. These methods start from the similarity matrix and search for the most similar cases. The results of the hierarchical analysis usually include a tree diagram dendrogram” (Figure 1). The dendrogram is a graphical display of the hierarchical structure implied by the similarity matrix and clustered by the linkage rule (Aldenderfer, 1984:37). In which we observe that the first step or level all cases are seen as individual clusters and at the final step, all cases are merged into one large group. Each cluster is a member of a larger, more inclusive cluster at a higher level of similarity. One disadvantage of this method is that different solution could be generated simply by giving a different order to the data; therefore results are not stable when observations are eliminated from the analysis.

Within this technique besides single linkage, there are other rules of formation of clusters: complete linkage, average linkage and Ward’s method. As we mentioned before in the single linkage all cases are joined to existing clusters if at least one of the components of that cluster has the same level of similarity. The connections are based on single linkage. The advantage of this option is the low effects of data transformations in the results but a drawback is the tendency to form long clusters.

The complete linkage states a more rigorous inclusion rule in which any candidate for inclusion into an existing cluster must be within a certain level of similarity to all members of that cluster. This type of method leads to compact clusters that are less susceptible to outliers than in the single-linkage.
The average linkage computes an average of the similarity of an observation under consideration with all cases in the existing cluster and joins the case to that cluster if a given level of similarity is achieved using the average rule. Some variants calculate the similarity between the centroids of two merged clusters or the computation weights the results by the size of the cluster.

The Ward’s technique optimizes the minimum variance within clusters. The method joins the observations that have the minimum increase in the error sum of squares; it tends to create clusters of equal sizes and shapes. This method is less responsive to outliers; however a drawback is that clusters are ordered by their elevation.

The iterative partitioning family

The iterative partitioning methods start with an initial partition of the data into some specified number of clusters in which they compute the centroids of the clusters. Each data point is allocated to the cluster with the nearest centroid\(^{12}\). Subsequent, the new centroids are computed, clusters are not updated until there has been a complete pass through the data. The last two tasks: allocation and new centroids computation are repeated until no data point change clusters. The main disadvantage of this method is to discover the optimal partition of a data set.

Depending on the way in which cases are assigned to clusters there are two types of pass: k-means passes and hill climbing passes. The K-means cluster analysis tries to assign cluster to the nearest centroid. The passes can be either combinational or non-combinatorial. The first one requires the recalculation of the centroid after each change in its membership; the second one recalculates the cluster centroid only after an entire pass through the data is completed. K-means method could be exclusive or inclusive when they remove or include the case under consideration from the parent cluster.

Hill climbing pass moves cases on the basis of whether or not the proposed move optimizes a particular statistical criterion used. The statistical criteria could be: $trW$, $trW^{-1}B$, $det W$ and the largest eigenvalues of $W^{-1}B$, where $W$ refers to the pooled within-cluster covariance matrix and $B$ is the pooled between-cluster covariance matrix.

K-means cluster analysis

There is an initial partition of the data in which the k value is user-specified and centroids of clusters are computed. Initial clusters centers are chosen in a first pass of the data, then each additional iteration groups observations based on nearest Euclidian distance to the mean of the cluster. After all reassignments, the centroids of new clusters are computed. The process continues until cluster means do not shift more than a give cut-off value or the iteration limit is reached. The method works directly upon the raw data, it passes through the data more than one time, the clusters are not nested and do not overlap each other. As we mentioned before, the main drawback is to find out the optimal partition of the data.

---

\(^{12}\) Centroid is the weighted average location of one of the category of samples in ordination space.
**Number of clusters**

There is not an objective method for selecting the number of clusters, one first option is to calculate different cluster solutions and decide among the alternatives based on a previous criteria.

**Figure 2. Defining the number of cluster using a dendrogram**

Using the hierarchical analysis, the number of clusters could be determined using the dendrogram. This tree diagram shows the results of a cluster analysis, usually is depicted horizontally, with each row representing a case. Cases with high similarity are adjacent. Lines indicate the degree of similarity of dissimilarity between cases. After an inspection of the tree, the researcher does some “cuts”, this process makes possible to predict how many possible groups could be obtained. The “cuts” are defined by selecting a certain distance (height) of each observation. The branches obtained will group cases that similar to each other. In the example given in figure 2, we see how heights a, b and c produce different clusters (6, 4 and 3 groups, respectively). The main drawback of this procedure is the fact that it could be biased by the researcher opinion.

In the non-hierarchical analysis, the research could use a plot that compares the number of clusters in relation to the total variance of the groups, the variance among groups or the fusion values (at which cases merge to form a cluster). Similar to the “scree test” in Factor Analysis, the graphic point in which there is an evident change will indicate the appropriate number of clusters.

Another procedure to determine the number of clusters is to examine the fusion coefficients (distances) and identify a significant variation among them. This change would suggest that dissimilar clusters have merged. In figure 3, the fusion value for four clusters shows the most noticeable break of the series. The stopping rule number 1 by Mojena (1977) states that an optimal partition is selected when it satisfies the following inequality:

\[ z_{j+1} > z + k \sigma_z \]
where $Z$ is the value of the fusion coefficient, $z_{j+1}$ is the value of the coefficient at stage $j+1$ of the clustering process, $k$ is the standard deviate and $z$ and $s_z$ are the mean and the standard deviation of the fusion coefficient. There is more that one cluster if the inequality is satisfied. The standard deviate ($k$) is calculated as follows:

$$k_j = \frac{z_{j+1} - z}{s_z}$$

The estimation continues until the inequality is not satisfied. The rule suggests using the number of clusters given by the row in which this process is stopped.

**Figure 3. The scree test for Cluster analysis**

![Figure 3. The scree test for Cluster analysis](image)

**Some Problems in Cluster Analysis**

One of the potential problems in cluster analysis could be the presence of outliers in the variables; in such situation the research should decide whether to keep it or eliminate it from the data and restart a new classification. Another two aspects that could affect the performance of this technique are the overlapping of clusters and the choice of similarity measure used.

**The Validation stage**

Once the results have been interpreted, if it is possible, it is convenient to validate the results. Validation could be done internally or externally. The former criteria consists in divide the data in two groups and apply the cluster technique used originally. The results are compared in both groups. The external criteria imply to compare the results with external evaluations or previous research or applying the technique to a different sample.
## Appendix – Cluster Analysis


<table>
<thead>
<tr>
<th>Cluster</th>
<th>Final Cluster Centers</th>
<th>Cluster</th>
<th>Final Cluster Centers</th>
<th>Cluster</th>
<th>Final Cluster Centers</th>
<th>Cluster</th>
<th>Final Cluster Centers</th>
</tr>
</thead>
<tbody>
<tr>
<td>O71</td>
<td>2.90 3.00 0.70 -0.70 -0.30</td>
<td>O86</td>
<td>-0.40 -0.40 1.50 -1.20 -0.60</td>
<td>O94</td>
<td>-0.5 -1.2 0.3 0.0 -0.5</td>
<td>O71</td>
<td>-0.9 3.0 -0.5 0.6 -0.6</td>
</tr>
<tr>
<td>P71</td>
<td>-0.30 0.00 0.90 3.20 -0.80 -0.20</td>
<td>P86</td>
<td>-0.30 0.40 0.80 -1.70 -2.00</td>
<td>P94</td>
<td>0.3 -1.6 0.5 -1.4 -0.3</td>
<td>P71</td>
<td>-1.1 0.0 -0.2 1.7 -0.8</td>
</tr>
<tr>
<td>TU70</td>
<td>-0.10 -0.50 -1.20 2.70 1.20 -0.10</td>
<td>TU82</td>
<td>-0.10 -0.10 0.20 -0.50 0.50</td>
<td>TU94</td>
<td>0.8 -1.1 0.1 0.2 -0.4</td>
<td>TU70</td>
<td>1.0 -0.5 0.0 1.6 0.7</td>
</tr>
<tr>
<td>TO70</td>
<td>3.30 2.30 -0.10 -1.00 -1.00 0.00</td>
<td>TO82</td>
<td>-0.30 1.10 -0.60 -0.80 -1.00</td>
<td>TO94</td>
<td>1.1 -1.3 0.4 -1.9 -0.9</td>
<td>TO70</td>
<td>0.9 2.3 -0.1 -0.9 -1.1</td>
</tr>
<tr>
<td>TG70</td>
<td>-0.70 -0.70 -0.40 1.60 1.80 -0.40</td>
<td>TG82</td>
<td>-0.20 -0.60 0.80 0.20 1.30</td>
<td>TG94</td>
<td>0.7 1.2 -0.5 3.9 0.6</td>
<td>TG70</td>
<td>0.4 -0.7 -0.5 1.0 3.1</td>
</tr>
<tr>
<td>PE7085</td>
<td>1.10 2.60 0.60 1.30 0.30 -0.50</td>
<td>PE8594</td>
<td>-0.40 0.00 -0.20 2.80 1.90</td>
<td>PE9400</td>
<td>-0.4 3.8 -0.4 1.9 0.3</td>
<td>PE7000</td>
<td>3.9 0.8 -0.2 0.2 1.8</td>
</tr>
<tr>
<td>PE17085</td>
<td>-1.90 -0.10 10 2.80 0.30 -0.10</td>
<td>PE18594</td>
<td>0.10 -0.30 -0.40 -0.50 2.70</td>
<td>PE19400</td>
<td>1.0 -0.6 0.3 4.3 0.5</td>
<td>PE17000</td>
<td>0.6 -0.7 0.0 2.0 1.7</td>
</tr>
<tr>
<td>PE27085</td>
<td>-2.90 2.00 0.80 1.30 -0.20 -0.20</td>
<td>PE28594</td>
<td>0.00 -0.30 -0.20 0.00 1.80</td>
<td>PE29400</td>
<td>-0.6 0.1 -0.5 1.6 0.7</td>
<td>PE27000</td>
<td>0.2 -0.1 0.1 0.2 2.4</td>
</tr>
<tr>
<td>DEB7085</td>
<td>-0.70 -0.30 0.40 0.50 0.80 -0.30</td>
<td>DEB8594</td>
<td>-0.20 -0.20 -0.10 5.40 0.20</td>
<td>DEB9400</td>
<td>-0.3 5.4 -0.2 0.0 -0.2</td>
<td>DEB7000</td>
<td>5.4 -0.3 -0.2 0.1 0.4</td>
</tr>
<tr>
<td>PI7085</td>
<td>-0.50 -0.40 0.00 1.80 1.20 -0.40</td>
<td>PI8594</td>
<td>-0.20 -0.60 0.40 0.90 2.30</td>
<td>PI9400</td>
<td>0.5 0.8 -0.4 3.9 0.4</td>
<td>PI7000</td>
<td>0.7 -0.6 -0.2 1.5 2.0</td>
</tr>
<tr>
<td>TI7085</td>
<td>-2.90 -0.40 -0.40 -1.70 -0.40 0.30</td>
<td>TI8594</td>
<td>-0.50 0.70 0.10 -1.20 -1.00</td>
<td>TI9400</td>
<td>0.0 -1.2 0.4 -0.9 -0.6</td>
<td>TI7000</td>
<td>-0.8 -0.4 -0.1 1.0 -0.4</td>
</tr>
<tr>
<td>IS7085</td>
<td>0.20 4.70 0.40 1.10 -0.40 -0.40</td>
<td>IS8594</td>
<td>0.70 0.10 1.10 -0.40 -0.40</td>
<td>IS9400</td>
<td>0.6 -0.5 0.2 -0.4 -0.6</td>
<td>IS7000</td>
<td>-0.4 4.2 -0.5 0.9 -0.4</td>
</tr>
<tr>
<td>IN70</td>
<td>-0.60 -0.70 -0.60 2.90 0.90 -0.20</td>
<td>IN80</td>
<td>0.10 -0.40 0.50 2.40 0.50</td>
<td>IN90</td>
<td>0.6 2.0 -0.2 0.4 -0.4</td>
<td>IN7000</td>
<td>2.0 -0.7 0.0 1.1 0.3</td>
</tr>
<tr>
<td>LT70</td>
<td>1.50 1.20 1.00 0.00 -1.00 -0.10</td>
<td>LT80</td>
<td>-0.50 0.40 0.40 -2.00 -1.20</td>
<td>LT90</td>
<td>1.2 2.2 0.4 0.0 -1.0</td>
<td>LT70</td>
<td>-1.8 1.2 -0.5 0.4 -0.9</td>
</tr>
<tr>
<td>X7085</td>
<td>-0.40 0.40 -0.20 4.50 -0.10 -0.20</td>
<td>X8594</td>
<td>-0.40 0.80 -0.20 -0.50 0.50</td>
<td>X9400</td>
<td>2.1 -0.5 0.1 -0.1 -0.3</td>
<td>X7000</td>
<td>-0.4 1.4 -0.4 0.3 -0.5</td>
</tr>
<tr>
<td>W70</td>
<td>2.40 1.50 0.80 -0.60 -0.90 0.00</td>
<td>W80</td>
<td>-0.40 1.10 -0.10 -1.60 -1.70</td>
<td>W90</td>
<td>0.9 -1.8 0.6 -1.8 -1.0</td>
<td>W70</td>
<td>-1.0 1.5 -0.2 -0.5 -0.9</td>
</tr>
<tr>
<td>E70</td>
<td>3.30 0.70 0.40 -1.40 -0.80 0.00</td>
<td>E80</td>
<td>-0.50 0.30 0.20 -1.60 -1.40</td>
<td>E90</td>
<td>1.1 -1.5 0.4 -0.3 -1.0</td>
<td>E70</td>
<td>-1.3 0.7 -0.4 -0.9 -0.7</td>
</tr>
<tr>
<td>D70</td>
<td>2.60 1.60 0.40 -0.60 -1.10 0.00</td>
<td>D80</td>
<td>-0.60 1.20 -0.20 -1.80 -0.70</td>
<td>D90</td>
<td>1.4 -2.1 0.3 0.2 -0.9</td>
<td>D70</td>
<td>-1.7 1.6 -0.4 -0.3 -1.5</td>
</tr>
<tr>
<td>U70</td>
<td>2.40 1.70 0.40 -1.00 -1.00 0.00</td>
<td>U80</td>
<td>-0.70 1.00 0.40 -1.70 -1.50</td>
<td>U90</td>
<td>1.6 -1.8 0.4 -1.2 -0.9</td>
<td>U70</td>
<td>-1.5 1.7 -0.5 -1.3 2.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cluster</td>
<td>Growth Rate</td>
<td></td>
<td>Cluster</td>
<td>Growth Rate</td>
<td>Cluster</td>
<td>Growth Rate</td>
</tr>
<tr>
<td>Veracruz</td>
<td>3</td>
<td>0.2</td>
<td></td>
<td>Baja California Sur</td>
<td>3</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>México</td>
<td>3</td>
<td>0.3</td>
<td></td>
<td>Sonora</td>
<td>3</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>Nayarit</td>
<td>3</td>
<td>0.5</td>
<td></td>
<td>Baja California</td>
<td>2</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Tabasco</td>
<td>5</td>
<td>0.7</td>
<td></td>
<td>Sinaloa</td>
<td>6</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>Sinaloa</td>
<td>3</td>
<td>0.7</td>
<td></td>
<td>México</td>
<td>6</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>Chiapas</td>
<td>5</td>
<td>0.8</td>
<td></td>
<td>Veracruz</td>
<td>6</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>Sonora</td>
<td>7</td>
<td>0.8</td>
<td></td>
<td>Tamaulipas</td>
<td>6</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Baja California</td>
<td>2</td>
<td>0.9</td>
<td></td>
<td>Distrito Federal</td>
<td>1</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Baja California Sur</td>
<td>4</td>
<td>0.9</td>
<td></td>
<td>Nuevo León</td>
<td>3</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Jalisco</td>
<td>7</td>
<td>1.1</td>
<td></td>
<td>Guanajuato</td>
<td>6</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Guanajuato</td>
<td>3</td>
<td>1.1</td>
<td></td>
<td>Yucatán</td>
<td>6</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Guerrero</td>
<td>3</td>
<td>1.3</td>
<td></td>
<td>Coahuila de Zaragoza</td>
<td>6</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>Tamaulipas</td>
<td>7</td>
<td>1.4</td>
<td></td>
<td>Chihuahua</td>
<td>6</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>Morelos</td>
<td>7</td>
<td>1.4</td>
<td></td>
<td>Morelos</td>
<td>6</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Nuevo León</td>
<td>7</td>
<td>1.4</td>
<td></td>
<td>Jalisco</td>
<td>6</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Puebla</td>
<td>3</td>
<td>1.5</td>
<td></td>
<td>Nayarit</td>
<td>6</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>Zacatecas</td>
<td>3</td>
<td>1.5</td>
<td></td>
<td>Michoacán</td>
<td>6</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>Yucatán</td>
<td>3</td>
<td>1.6</td>
<td></td>
<td>Aguascalientes</td>
<td>6</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>Coahuila de Zaragoza</td>
<td>7</td>
<td>1.6</td>
<td></td>
<td>Puebla</td>
<td>6</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>Michoacán</td>
<td>3</td>
<td>1.6</td>
<td></td>
<td>Guerrero</td>
<td>6</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>Durango</td>
<td>3</td>
<td>1.7</td>
<td></td>
<td>Zacatecas</td>
<td>6</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>Hidalgo</td>
<td>3</td>
<td>1.8</td>
<td></td>
<td>Quintana Roo</td>
<td>4</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>Colima</td>
<td>7</td>
<td>1.8</td>
<td></td>
<td>San Luis Potosi</td>
<td>6</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Oaxaca</td>
<td>1</td>
<td>1.9</td>
<td></td>
<td>Colima</td>
<td>3</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>Tlaxcala</td>
<td>3</td>
<td>1.9</td>
<td></td>
<td>Durango</td>
<td>3</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>San Luis Potosi</td>
<td>3</td>
<td>2.0</td>
<td></td>
<td>Hidalgo</td>
<td>5</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>Distrito Federal</td>
<td>6</td>
<td>2.3</td>
<td></td>
<td>Querétaro</td>
<td>6</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>Chihuahua</td>
<td>7</td>
<td>2.5</td>
<td></td>
<td>Chiapas</td>
<td>5</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>Quintana Roo</td>
<td>4</td>
<td>2.6</td>
<td></td>
<td>Oaxaca</td>
<td>5</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>Aguascalientes</td>
<td>7</td>
<td>2.7</td>
<td></td>
<td>Tlaxcala</td>
<td>6</td>
<td>5.3</td>
<td></td>
</tr>
<tr>
<td>Querétaro</td>
<td>3</td>
<td>2.7</td>
<td></td>
<td>Tabasco</td>
<td>5</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Campeche</td>
<td>4</td>
<td>3.3</td>
<td></td>
<td>Campeche</td>
<td>5</td>
<td>16.5</td>
<td></td>
</tr>
</tbody>
</table>