

# **DIRECCIÓN DE SÍNTESIS Y CUENTAS NACIONALES**

## **PRODUCTO INTERNO BRUTO -PIB- TRIMESTRAL**

**NOTA METOLÓGICA**

**Implementación fase 2 de la mejora integral en la medición de la  
actividad de construcción de edificaciones**

**Febrero de 2022**



**El futuro  
es de todos**

**Gobierno  
de Colombia**

## 1. Introducción

El Departamento Administrativo Nacional de Estadística - DANE, tiene entre sus funciones la elaboración de las cuentas nacionales del país, adoptando estándares internacionales e implementando en sus procesos las buenas prácticas en materia de producción estadística. En el marco del proceso de mejora continua, la Dirección de Síntesis y Cuentas Nacionales ha iniciado un proyecto que se orienta en la actualización de la medición de las actividades de la división 45 (Construcción de edificios) de la clasificación CIIU Rev. 4 A.C, y sus productos característicos de la división 531 (edificios residenciales y no residenciales) de la clasificación central de productos CPC. 2 A.C., Con el objetivo de crear una base de medición consistente con los actuales procesos productivos de la construcción de edificaciones.

De esta manera, la presente nota metodológica tiene como objeto exponer el proceso de la segunda fase de mejora integral del sector de edificaciones, específicamente en la actualización del modelo de conglomerados para la medición del resto país y las estimaciones para las construcciones sin licencias y rurales.

## 2. Contexto

La medición del valor agregado de la construcción de edificaciones se basa en el principio de causación, el cual consiste en registrar la producción a medida que avanzan las construcciones, realizando el seguimiento al avance del proceso constructivo del área total a construir de una edificación durante el tiempo de su ejecución.

Para el cálculo de la causación se cuenta con los siguientes componentes:

- **Componente observado.** Consiste en calcular la causación a través del avance de la obra mediante el capítulo constructivo en el que se encuentra la edificación y el grado de avance de este capítulo y la aplicación de los coeficientes de incidencia. A través del Censo de Edificaciones –CEED del DANE, se obtiene el registro observado de los metros cuadrados causados para las áreas de cobertura geográfica que hacen parte de dicha operación estadística.
- **Componente no observado.** Consiste en calcular indirectamente los metros cuadrados causados a través de un modelo de curvas de maduración para los municipios no cubiertos por el CEED, que registran información de aprobaciones en las Estadísticas de Licencias de Construcción –ELIC. No obstante, para los municipios que no se encuentran dentro de la cobertura de las operaciones estadísticas mencionadas, la Dirección de Síntesis y Cuentas Nacionales –DSCN implementa un modelo de conglomerados para imputar los metros cuadrados aprobados a partir de la información de la ELIC.

### 3. Modelo de Conglomerados

Para la medición de la construcción de edificaciones la fuente principal es el Censo de Edificaciones – CEED, el cual de acuerdo con su cobertura geográfica proporciona información observada al registrar los avances de las obras a partir de su inicio hasta su culminación. Es importante mencionar, que las cuentas nacionales al ser exhaustivas requieren de indicadores que le permitan alcanzar la cobertura nacional. Por lo tanto, para medir la producción del resto del país, es decir, para los municipios que no se encuentran en la cobertura del CEED, se realizan estimaciones indirectas a través de un modelo de curvas de maduración, en el cual se mide la duración promedio de las obras para ejecutar el total del área a construir, por lo que, permite calcular indirectamente el área causada, a partir de los metros cuadrados iniciados, los cuales se calculan aplicando rezagos de iniciación al área aprobada registrada en las Estadísticas de Licencias de Construcción –ELIC, para los municipios sin cobertura en el CEED.

Adicionalmente, para la medición del resto país, es decir, para los municipios que no hacen parte de la cobertura geográfica del CEED ni de la ELIC, se desarrolla un modelo estadístico de conglomerados, con el fin de distribuir geográficamente la información disponible de licenciamientos. Para ello, se crean grupos de municipios con licenciamiento (donantes) cuya información registrada en la ELIC se asigna a los grupos de municipios sin licenciamiento (receptores). Finalmente, una vez se imputan los metros cuadrados licenciados se pasan por el modelo las curvas de maduración para obtener el nivel de metros cuadrados efectivamente causados durante cada período.

Para la actualización del modelo de conglomerados se incorporaron los resultados del Censo Nacional de Población y Vivienda 2018, ya que, en el marco de los procesos de la Nueva Base 2015, solo estaba disponible la información del Censo General 2005.

Para la aplicación del modelo, se adoptaron una serie de procesos metodológicos propios de la estadística multivariada. En primera instancia, se realizó un análisis de componentes principales en el plano de las variables donde, de acuerdo con los criterios de representatividad dentro de los planos factoriales, se llevó a cabo una selección de estas. Estas variables se denominan variables sintomáticas y permiten caracterizar el comportamiento de los individuos, los cuales están representados por los 1.122 municipios.

En segundo lugar, esta vez en el plano de los individuos, se vuelve aplicar el modelo de componentes principales y se identifica la presencia de datos atípicos como es el caso de las ciudades principales, (ejemplos como Bogotá, Medellín, Cartagena, Barranquilla y Cali) para el caso de las edificaciones

residenciales al igual que para el ejercicio de las edificaciones no residenciales (son las mismas ciudades principales del caso residenciales incluyendo en este grupo a Bucaramanga).

Una vez se concluye el análisis de componentes principales, es decir, se han seleccionado las variables y los individuos con los que se requiere aplicar el modelo, se procede a la implementación del algoritmo de clasificación usando el paquete FactoClass{FactoClass} disponible en R. Este paquete presenta una combinación de métodos factoriales y de clasificación buscando agrupar los municipios en el territorio colombiano para el ejercicio.

**Figura 1.** Estrategia de Clasificación



<sup>1</sup> Si el número de "individuos" es muy grande, realizar un K-means de *preagrupamiento* en miles de clases, para luego realizar la clasificación jerárquica con el método de Ward.

**Fuente:** DANE, Cuentas nacionales

## 4. Otros elementos del componente no observado

Dentro de la medición de construcción de edificaciones se reconocen los procesos constructivos de las edificaciones sin licencia y las que se construyen en suelo rural. Para la actualización del primer elemento mencionado, se tomó como referencia observada información del Censo de Edificaciones –CEED, relacionada con las variables de formalidad de las obras, lo cual permitió identificar el comportamiento de este tipo de edificaciones y a su vez, actualizar los coeficientes de curvas de maduración utilizados para la estimación correspondiente.

Finalmente, para el segundo elemento se utilizó información de Estadísticas de Licencias de Construcción –ELIC con el objetivo de identificar a través de la variable *tipo licencia*, los metros cuadrados aprobados para suelo rural y a su vez, se establecieron los respectivos rangos de área aplicando un análisis de clústeres a partir de la metodología k-medias<sup>1</sup>. De igual manera, se realizó un análisis de información de la variable *Clase* registrada en el Censo de Edificaciones –CEED, para clasificar las construcciones rurales, lo cual permitió calcular rezagos de iniciación y actualizar los coeficientes de la curva de maduración para edificaciones rurales.

---

<sup>1</sup> Algoritmo no supervisión de machine learning cuyo objetivo es particionar n observaciones en k clústeres en los cuales cada una de sus observaciones pertenece al clúster con la media más cercana al centroide o centro del cluster.