



INFOTEC CENTRO DE INVESTIGACIÓN E  
INNOVACIÓN EN TECNOLOGÍAS DE LA  
INFORMACIÓN Y COMUNICACIÓN

DIRECCIÓN ADJUNTA DE INNOVACIÓN Y  
CONOCIMIENTO  
GERENCIA DE CAPITAL HUMANO  
POSGRADOS

# “Segmentación de Municipios de la Republica Mexicana”

(TIPO DE PROYECTO)

Que para obtener el grado de MAESTRO EN  
CIENCIA DE DATOS E INFORMACIÓN

Presenta:

**Pablo Gómez García**

Asesor:

**(Nombre y Apellidos)**

Ciudad de México, Mayo, 2023.

## **Autorización de impresión**

## Agradecimientos

## Tabla de contenido

Tabla de contenido.....	5
Introducción.....	1
Capítulo 1. Planteamiento de la Problemática .....	3
1.1 Objetivo .....	4
1.2 Alcance de la solución .....	4
Capítulo 2. Técnicas de modelación no supervisadas .....	2
2.1 K-Means .....	2
Capítulo 3. Metodología Aplicada al Problema .....	2
3.1 Procesamiento de los Datos.....	2
3.2 Análisis Exploratorio .....	2
3.3 Reducción de la Dimensionalidad .....	2
3.4 Modelado de los datos .....	2
3.5 Resultados Preliminares.....	2
Capítulo 4. Resultados y Conclusiones .....	9
Conclusiones.....	12
Bibliografía.....	13
ANEXO 1.....	15
Índice de términos.....	16

## Índice de figuras

## Índice de gráficos

## Índice de cuadros



## Siglas y abreviaturas

[illegible]

## Glosario

### “A”

**Aenean:** Sollicitudin sem lorem, nec tristique lacus auctor in. Etiam luctus purus non dui fringilla tempor. Suspendisse euismod urna id nisl cursus, nec tincidunt lacus sagittis. Sed mollis sem mattis ligula rutrum scelerisque. Ut mattis condimentum blandit. Curabitur ipsum mauris, ullamcorper non accumsan id, eleifend id velit. Nunc at semper massa, sit amet pellentesque metus. Pellentesque pulvinar eget mauris sit amet dapibus. Integer vel lorem ut sem pretium semper vel at est. Aenean vitae varius libero. Sed accumsan nisl eu nulla consectetur fringill.

### “B”

**Blandit:** Aenean laoreet ligula massa, ut varius lorem convallis ut. Integer at libero neque. Donec vestibulum neque in efficitur iaculis. Donec eros eros, porta suscipit auctor et, rutrum ut tortor.

### “C”

**Congue:** Quam nibh convallis mauris, et tincidunt massa dolor maximus mauris. Fusce pretium lectus vitae aliquet aliquet. Fusce elit ligula, cursus eu velit eu, maximus tempus lorem.



## Introducción

En la actualidad existen en México diversos estudios a nivel municipio, la gran mayoría realizados por el Instituto Nacional de Estadística y Geografía (INEGI), dichos estudios muestran características poblacionales de cada uno de ellos, también muestran estados de opinión, preferencias electorales, índices de Pobreza y Marginación etc.

Toda la información que se tiene acerca de los municipios es valiosa para distintos fines, sin embargo a la fecha hay muy pocos estudios que describen a los municipios por medio de varias de sus características, si es que existen tales estudios, lo hacen a nivel descriptivo simple.

Dentro de este proyecto planteamos la posibilidad de aplicar una técnica no supervisada para encontrar una segmentación muy característica de los municipios del País.

Tomando como punto de partida la información pública que existe de los municipios en diversos temas, planteamos la siguiente Hipótesis:

*“Existen variables entre los datos públicos de los municipios que permiten diferenciarlos en segmentos cuyos elementos que los conforman poseen características homogéneas entre sí pero heterogéneas con los elementos de otros segmentos”.*

Bajo el supuesto anterior la propuesta de este trabajo consiste en tomar la información pública de fuentes como el INEGI, la PGR, la CNBV entre otras más para construir un Modelo no supervisado con el que se pueda encontrar a los segmentos planteados en la hipótesis.

El importante enfatizar que el modelo propuesto no tiene como objetivo clasificar municipios que sean creados en el futuro, debido a que la creación de municipios no es algo que sucede a menudo, sin embargo busca mostrar de manera descriptiva el valor y aporte de cada segmento encontrado por el modelo.

Finalmente este trabajo muestra una serie de sugerencias sobre la aplicación de los resultados obtenidos.

# Capítulo 1

## Planteamiento de la Problemática



## Capítulo 1. Planteamiento de la Problemática

Como podemos ver en nuestro día a día, la información se genera en cantidades desproporcionadas por todos lados, se genera a través de las redes sociales, los sistemas de información tales como los sistemas bancarios, aeropuertos, los sistemas de cobro e inventarios de los centros comerciales por mencionar algunos. El gran reto de la ciencia de datos es dotar de utilidad para la toma de decisiones a dicha información.

Debe tenerse claro que la información por sí sola, no aporta más que un panorama muy general de lo que existe en todos esos sistemas. Por ello y para poder dar una propuesta de valor basada en el conocimiento se debe traducir dicha información por medio de la inteligencia a aplicaciones direccionadas a dar solución a un problema de negocio.

Derivado de lo anterior y siguiendo la misma directriz vamos a asumir que nuestro sistema de información es el Instituto Nacional de Estadística y Geografía (INEGI), que de acuerdo con lo citado en su página<sup>1</sup> se trata de *“Un organismo público autónomo responsable de normar y coordinar el Sistema Nacional de Información Estadística y Geográfica, así como de captar y difundir información de México en cuanto al territorio, los recursos, la población y economía, que permita dar a conocer las características de nuestro país y ayudar a la toma de decisiones”*. Hasta este punto sabemos que el INEGI provee de mucha información referente a temas del País México, particularmente a nivel de municipios. Ahora bien, esta información por si sola carece de inteligencia (definiremos el termino inteligencia como la capacidad de predecir eventos futuros a partir de información observada) por lo que no podría dar una respuesta concreta a una pregunta de negocio con enfoque predictivo.

Finalmente podemos apreciar que el problema se centra en poder tomar la información que existe y transformarla, combinarla y modelarla de una manera que pueda dar respuesta a la pregunta de negocio que planteamos en la introducción.

De una Manera más específica la nuestra pregunta de negocio es la siguiente:

*¿Existen variables entre los datos públicos de los municipios que permiten diferenciarlos en segmentos cuyos elementos que los conforman poseen características homogéneas entre sí pero heterogéneas con los elementos de otros segmentos?*

¿Por qué esta pregunta de negocio? La respuesta es porque se busca llevar más allá esta información, aportando un análisis más profundo que combine diversos

---

<sup>1</sup> [https://www.inegi.org.mx/inegi/quienes\\_somos.html](https://www.inegi.org.mx/inegi/quienes_somos.html)

tipos de datos acerca de los municipios de la República Mexicana con la finalidad de aportar conocimiento e inteligencia para la toma de decisiones.

Generalmente la información proporcionada por el INEGI y de más organismos que proporcionan información referente al país y a nivel regional, se usan para el lanzamiento de propuestas en beneficio de la población, tales como programas de alimentación, programas en beneficio del campo entre otras acciones más. Justo por ello la propuesta que se plantea en este trabajo busca ir más allá y por ello planteamos el objetivo del proyecto como se plantea en el siguiente apartado.

## **1.1 Objetivo**

Plantear una solución desde un enfoque Analítico basada en un Modelo no supervisado que logre segmentar a los municipios de la República Mexicana. Se pretende que dicha segmentación resulte de utilidad para estrategias de implementación de Políticas Públicas, de Marketing y para estrategias de Seguridad.

## **1.2 Alcance de la solución**

Plantear el alcance de la solución propuesta



# Capítulo 2

## Técnicas de modelación no supervisadas



## Capítulo 2. Técnicas de modelación no supervisadas

Aquí va el Desarrollo Teórico

### 2.1 K-Means



# Capítulo 3

## Metodología Aplicada al Problema



## Capítulo 3. Metodología Aplicada al Problema

Aquí va la explicación de cómo se aplicó la metodología

### 3.1 Procesamiento de los Datos

### 3.2 Análisis Exploratorio

### 3.3 Reducción de la Dimensionalidad

### 3.4 Modelado de los datos

### 3.5 Resultados Preliminares



# Capítulo 4

## Resultados y Conclusiones



## Capítulo 4. Resultados y Conclusiones

Aquí irán los Resultados obtenidos y la solución propuesta, así como las conclusiones finales del proyecto

# Conclusiones



## Conclusiones

Conclusiones Finales.

## Bibliografía



**ANEXOS**

## ANEXO 1

## Índice de términos

### “A”

Aliquam.....12

### “B”

Blandit.....3

### “C”

Consectetur.....7

### “D”

Donec.....12