



# ESTUDIO DEL MERCADO POTENCIAL A PARTIR DE LAS OCUPACIONES INFORMALES EN MÉXICO

Matilde Simental Herrera

Trainee Datos. Programa Be Bold

An abstract graphic consisting of numerous thin, overlapping, wavy lines that create a sense of movement and depth. The lines are light gray and flow from the bottom left towards the top right, forming a large, open, wave-like shape that occupies the right half of the page.

**Julio 2025**

# Índice

<b>1. Descripción del proyecto</b>	<b>2</b>
1.1. Fuentes de donde se obtuvo la información . . . . .	2
<b>2. Metodología</b>	<b>2</b>
2.1. Selección de Ocupaciones . . . . .	3
2.2. Explicación del código implementado . . . . .	4
<b>3. Anexos</b>	<b>8</b>

# 1. Descripción del proyecto

Segmentar las ocupaciones informales en México a nivel estatal y municipal para tener una dimensión del potencial de mercado para CrediClub. Utilizando la Encuesta Nacional de Ocupación y Empleo (ENOE) del INEGI al cuarto trimestre del 2024.

## 1.1. Fuentes de donde se obtuvo la información

A continuación se describen las fuentes de información así como una explicación de estas, todo el material utilizado se encuentra en el siguiente SharePoint [crediclubsa.sharepoint.com](https://crediclubsa.sharepoint.com). Se hace referencia a esta carpeta a lo largo de todo el documento bajo el nombre **carpeta**.

La información utilizada en este análisis fue obtenida del portal oficial del INEGI, específicamente del apartado correspondiente a la Encuesta Nacional de Ocupación y Empleo (ENOE):

[www.inegi.org.mx/enoe](https://www.inegi.org.mx/enoe).

A partir de los datos abiertos disponibles, se descargó la información correspondiente al cuarto trimestre de 2024. El archivo descargado, en formato .zip, contiene cinco tablas en formato .csv:

- ENOE\_VIVT424: Vivienda
- ENOE\_HOGT424: Hogar
- ENOE\_SDEMT424: Sociodemográfico
- ENOE\_COE1T424: Cuestionario de ocupación y empleo parte I
- ENOE\_COE2T424: Cuestionario de ocupación y empleo parte II

Para este análisis se utilizaron las bases de datos ENOE\_SDEMT424 y ENOE\_COE1T424. En el caso de ENOE\_COE1T424, se empleó la versión básica del cuestionario.

La descripción de cada campo contenido en estas tablas se encuentra disponible en el documento Encuesta Nacional de Ocupación y Empleo (ENOE). Estructura de la base de datos.pdf, ubicado en la **carpeta**.

## 2. Metodología

Con el objetivo de realizar una vinculación adecuada entre las distintas tablas utilizadas, se generó un identificador único (ID) para cada registro. Este identificador fue construido a partir de los campos que se muestran en la Figura II, y se incorporó como una nueva columna con el nombre ID.

Primeramente, de la tabla ENOE\_SDEMT424 se filtraron las personas empleadas informalmente por medio de los campos:

- emp\_ppal = 1: Empleo informal (1) o formal (2).
- t\_tra = 1: Número de trabajos (1 o 2)

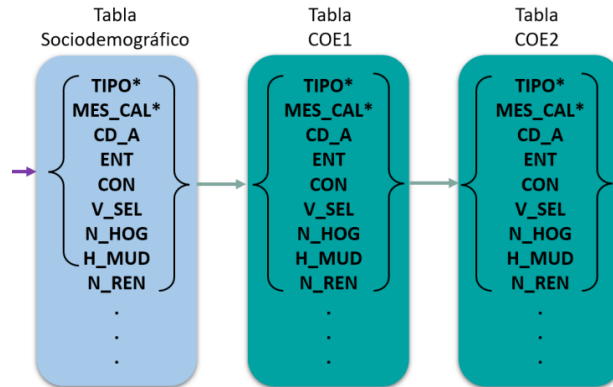


Figura 1: Captura obtenida del documento Encuesta Nacional de Ocupación y Empleo (ENOE). Estructura de la base de datos pg.4

- `fac_tri (sum)`: Ponderador trimestral<sup>1</sup>

Además, se consideraron las variables estado y municipio, correspondientes a las columnas `ent` y `mun`, respectivamente. Esta información permitió agrupar posteriormente las ocupaciones por estado y municipio.

## 2.1. Selección de Ocupaciones

De la tabla `ENOE_COE1T424` se extrajo la información correspondiente a las ocupaciones, registradas en la columna `P3`, la cual contiene claves de ocupación de cuatro dígitos. La descripción de todas las ocupaciones se basa en el Sistema Nacional de Clasificación de Ocupaciones (SINCO) 2019. Este documento, se encuentra bajo el nombre Sistema Nacional de Clasificación de Ocupaciones 2019. `SINCO.pdf`. Se recomienda consultar dicho documento a partir de la página 356.

En una primera iteración, se llevó a cabo una revisión exhaustiva del listado de ocupaciones del SINCO, agrupando aquellas que presentaban similitudes entre sí. Posteriormente, se seleccionaron las ocupaciones que son similares a las actividades desempeñadas por las y los clientes de CrediClub bajo criterio propio y de Carlos Calle.

Como resultado de este proceso, se identificaron 34 claves de ocupación, las cuales fueron clasificadas en 13 categorías, descritas en la Tabla<sup>1</sup>

Este listado de claves y descripción de ocupacion se encuentra en un archivo csv en la carpeta bajo el nombre `Listado.Ocupaciones.Seleccionadas`. Es importante mencionar que, en caso de retomar este proyecto en el futuro, el listado de ocupaciones deberá ser actualizado manualmente, añadiendo las nuevas claves junto con sus respectivas descripciones. Esto es necesario para asegurar la compatibilidad con el análisis implementado en el código.

En una segunda iteración, se realizó una agrupación más general de las ocupaciones, reduciendo su clasificación a ocho categorías, las cuales se enumeran a continuación:

- Trabajadores del hogar
- Producción de alimentos

<sup>1</sup> Al sumar esta columna se obtiene la cantidad aproximada de personas que representa cada registro.

Clave SINCO	Ocupación
5112, 5114	Vendedores de comida en establecimientos
9521	Vendedores ambulantes de comida
4224	Vendedores por catálogo
9512	Vendedores ambulantes de artículos diversos (juguetes, plantas, plumas, llaveros, radios, ropa, etcétera)
4213	Venta por teléfono o internet (negocio propio o no)
5211	Peluqueros, barberos, estilistas
5212	Maquillistas
9611, 5113	Trabajadores domésticos
9621	Trabajadores de limpieza en establecimientos o en la calle
9643	Lavanderos y planchadores domésticos
5221, 5222	Cuidadores de niños, personas con discapacidad y ancianos
7341, 7343, 7342, 7332	Elaboración de prendas y otros productos textiles (colchas, almohadas, sábanas, bolsas de tela, etc) en talleres o su casa
7311, 7312, 7313, 7351, 7352, 7353, 7411, 7611, 7614, 7322	Elaboración de accesorios y bisutería, manualidades, productos de cerámica, madera y cuero
7511, 7512, 7513, 7514, 7515, 7516	Transformación de materia prima en alimentos (Carnicero, tortillero, panadero, productos lácteos, dulces, etc)

Cuadro 1: Relación de claves SINCO y ocupaciones seleccionadas

- Vendedores ambulantes
- Venta por catálogo
- Venta en línea o por teléfono
- Venta de comida
- Peluqueros y maquillistas
- Artesanos

Esta segunda iteración se realizó dentro del código. Finalmente, se construyó una tabla que agrupa el número de personas que desempeñan cada una de estas ocupaciones a nivel municipal. El detalle de todo este procedimiento se presenta en la Sección [2.2](#)

## 2.2. Explicación del código implementado

El código utilizado para obtener las ocupaciones seleccionadas por municipio está hecho en Python, con el nombre `Ocupaciones_Municipal`, en la [carpeta](#). A continuación se explica por bloques la estructura del código.

En esta primera celda (Fig. [2](#)) se cargan los archivos del INEGI que se van a utilizar, así como el listado de ocupaciones (`LIST_OCUP`) con las claves y nombres de ocupación que fueron seleccionados en la Sección [2.1](#).

En la segunda celda se añade a cada tabla una columna `ID` con la que uniremos los datos de cada tabla.

```
import pandas as pd
import pandasql as psql # Instalar pandasql (pip install pandasql)
import numpy as np

COE1 = pd.read_csv('ENOE_COE1T424.csv')
SDEM = pd.read_csv('ENOE_SDEMT424.csv', encoding='latin-1')
LIST_OCUP = pd.read_csv('Listado_Ocupaciones_Seleccionadas.csv')
```

Figura 2: Primer bloque

```
# Añadir columna de ID a las tablas -----
COE1["ID"] = (
    COE1["tipo"].astype(str) +
    COE1["mes_cal"].astype(str) +
    COE1["cd_a"].astype(str) +
    COE1["ent"].astype(str) +
    COE1["con"].astype(str) +
    COE1["v_sel"].astype(str) +
    COE1["n_hog"].astype(str) +
    COE1["h_mud"].astype(str) +
    COE1["n_ren"].astype(str)
)
cols = ["ID"] + [col for col in COE1.columns if col != "ID"]
COE1 = COE1[cols]

# SDEM Crear ID
SDEM["ID"] = (
    SDEM["tipo"].astype(str) +
    SDEM["mes_cal"].astype(str) +
    SDEM["cd_a"].astype(str) +
    SDEM["ent"].astype(str) +
    SDEM["con"].astype(str) +
    SDEM["v_sel"].astype(str) +
    SDEM["n_hog"].astype(str) +
    SDEM["h_mud"].astype(str) +
    SDEM["n_ren"].astype(str)
)
cols = ["ID"] + [col for col in SDEM.columns if col != "ID"]
SDEM = SDEM[cols]
```

Figura 3: Segundo bloque

En la tercera celda (Fig. 4) se utiliza pandasql para filtrar las tablas más fácilmente, el query se puede consultar en el código. Esta información se guarda se guarda en Ocupaciones. En la cuarta celda (Fig. 5), las ocupaciones se

```
# Filtrado de datos
> query = """ ...

# Ejecutar el query
Ocupaciones = psql.sqldf(query, locals())
```

Figura 4: Tercer bloque

agrupan por clave de estado y municipio. Adicionalmente, se genera una clave única mediante la concatenación de ambas claves (ENTMUN) , con el fin de evitar ambigüedades en municipios que comparten nombre pero pertenecen a estados diferentes.

En la quinta celda se realiza una agrupación más general de las ocupaciones, lo que se mencionó anteriormente como la segunda iteración de agrupación. Aunque esta clasificación general puede definirse desde el inicio directamente en el archivo LIST\_OCUP . csv —lo cual resulta más claro y eficiente—, en esta ocasión no se implementó de esa forma por cuestiones de tiempo. Pero esta sería la alternativa más adecuada.

En la sexta celda (Fig. 7) se renombran las ocupaciones para acortar sus descripciones. Asimismo, se incorporan los nombres correspondientes del estado y del municipio, ya que hasta este punto dichas entidades solo habían sido identificadas mediante claves numéricas. Esta información se obtuvo del archivo Nom\_ESTMUN . csv, también extraído

```
# Agrupar por Estado, Municipio y Ocupacion, sumando fac_tri
ocupaciones_agrupadas = Ocupaciones.groupby(['Estado', 'Municipio', 'Ocupacion'], as_index=False)['fac_tri'].sum()

# Pivotar para que cada ocupación sea una columna
tabla_resultado = ocupaciones_agrupadas.pivot_table(
    index=['Estado', 'Municipio'],
    columns='Ocupacion',
    values='fac_tri',
    fill_value=0
).reset_index()

tabla_ocupaciones = tabla_resultado.sort_values(by=['Estado', 'Municipio']) # Ordenar por estado y municipio

# Crear clave de estado y municipio ENTNUM
tabla_ocupaciones['Estado'] = tabla_ocupaciones['Estado'].astype(str).str.zfill(2)
tabla_ocupaciones['Municipio'] = tabla_ocupaciones['Municipio'].astype(str).str.zfill(3)

tabla_ocupaciones['ENTNUM'] = tabla_ocupaciones['Estado'] + tabla_ocupaciones['Municipio']
display(tabla_ocupaciones)
```

Figura 5: Cuarto bloque

```
# Segunda iteración de agrupación de ocupaciones. Se agrupan en categorías más generales las ocupaciones enlistadas.

tabla_ocupaciones['Vendedores de comida'] = tabla_ocupaciones[
    ['Vendedores de comida en establecimientos',
     'Vendedores ambulantes de comida']
].sum(axis=1)

tabla_ocupaciones['Trabajadores domésticos'] = tabla_ocupaciones[
    ['Trabajadores domésticos',
     'Lavaderos y planchadores domésticos',
     'Trabajadores de limpieza en establecimientos o en la calle ',
     'Cuidadores de niños, personas con discapacidad y ancianos']
].sum(axis=1)

tabla_ocupaciones['Maquillistas y Peluqueros'] = tabla_ocupaciones[
    ['Maquillistas',
     'Peluqueros, barberos, estilistas']
].sum(axis=1)

tabla_ocupaciones['Elaboración de productos textiles y accesorios'] = tabla_ocupaciones[
    ['Elaboración de prendas y otros productos textiles (colchas, almohadas, sábanas, bolsas de tela, etc) en talleres o su casa ',
     'Elaboración de accesorios y bisutería, manualidades, productos de cerámica, madera y cuero.']
].sum(axis=1)

columnas_a_eliminar = [ # Eliminar solo las columnas que fueron usadas para agrupar
    "Vendedores de comida en establecimientos",
    "Vendedores ambulantes de comida",
    "Lavaderos y planchadores domésticos",
    "Trabajadores de limpieza en establecimientos o en la calle ",
    "Maquillistas",
    "Peluqueros, barberos, estilistas",
    "Elaboración de prendas y otros productos textiles (colchas, almohadas, sábanas, bolsas de tela, etc) en talleres o su casa ",
    "Elaboración de accesorios y bisutería, manualidades, productos de cerámica, madera y cuero.",
    "Cuidadores de niños, personas con discapacidad y ancianos"
]

tabla_ocupaciones.drop(columns=columnas_a_eliminar, inplace=True)
```

Figura 6: Quinto bloque

del INEGI<sup>2</sup>. Se puede encontrar en la **carpeta**

```
# Cambiar nombre a las columnas. Para tener nombres más sencillos.....
tabla_ocupaciones = tabla_ocupaciones.rename(columns={
    "Transformación de materia prima en alimentos (Carnicero, tortillero, panadero, productos lácteos, dulces, etc)": "Producción de alimentos",
    "Vendedores ambulantes de artículos diversos (juguetes, plantas, plumas, llaveros, radios, ropa, etcétera)": "Vendedores ambulantes",
    "Vendedores por catálogo": "Venta por catálogo",
    "Venta por teléfono o internet (negocio propio o no)": "Venta en línea",
    "Vendedores de comida": "Venta de comida",
    "Elaboración de productos textiles y accesorios": "Artesanos"
})

# Añadir el nombre del estado y municipio por su clave .....
ESTADOSMUN = pd.read_csv("nom_ESTMUN.csv")

tabla_ocupaciones = tabla_ocupaciones.merge(
    ESTADOSMUN[['ENTMUN', 'NOM_ENT', 'NOM_MUN']],
    on='ENTMUN',
    how='left'
)
```

Figura 7: Sexto bloque

En la séptima celda (Fig 8), se incorpora la cantidad de personas con ocupaciones informales en cada municipio. Aunque esta información también proviene de la tabla original del INEGI ENOE\_SDEMT424 , se almacenó en un archivo independiente, ya que su inclusión fue una decisión posterior.

<sup>2</sup>Cabe señalar que, según la documentación del INEGI, las claves de estado y municipio se mantienen consistentes en todos sus conjuntos de datos. Por tanto, cualquier archivo proveniente del INEGI que utilice estas claves será compatible con el catálogo de nombres correspondiente.

Los datos correspondientes a la población informal por municipio se encuentran en el archivo `PobInformal.csv`.

```
# Sumar el total de población empleada informalmente por municipio
PoblacionInformalNum = pd.read_csv("PobInformal.csv") # Cargar archivo en donde está el total de personas con ocupacion informal

PoblacionInformalNum['ent'] = PoblacionInformalNum['ent'].astype(str).str.zfill(2) # Renombrar las filas de claves de estado y municipio, y crear una nueva columna 'ENTMUN' para matchear
PoblacionInformalNum['mun'] = PoblacionInformalNum['mun'].astype(str).str.zfill(3)
PoblacionInformalNum['ENTMUN'] = PoblacionInformalNum['ent'] + PoblacionInformalNum['mun']

tabla_ocupaciones = tabla_ocupaciones.merge( # Matchear con tabla_ocupaciones
    PoblacionInformalNum[['ENTMUN', 'TrabajadoresInformales']],
    on='ENTMUN',
    how='left'
```

Figura 8: Séptimo bloque

Finalmente en la última celda (Fig. 9), se ordenan las columnas de la tabla. Y se muestra el resultado (Fig. 10).

```
# Ordenar columnas
tabla_ocupaciones = tabla_ocupaciones[
    [
        'ENTMUN', 'Estado', 'Municipio', 'NOM_ENT', 'NOM_MUN',
        'Trabajadores domésticos', 'Producción de alimentos',
        'Vendedores ambulantes', 'Vendedores por catálogo',
        'Venta en línea', 'Venta de comida',
        'Maquillistas y Peluqueros', 'Artesanos',
        'TrabajadoresInformales'
    ]
]

# Resultado
tabla_ocupaciones
```

Figura 9: Octavo bloque

NOM_ENT	NOM_MUN	Trabajadores domésticos	Producción de alimentos	Vendedores ambulantes	Vendedores por catálogo	Venta en línea	Venta de comida	Maquillistas y Peluqueros	Artesanos	TrabajadoresInformales
AGUASCALIENTES	AGUASCALIENTES	12919.0	2963.0	4458.0	2043.0	0.0	15358.0	5884.0	5377.0	169379
AGUASCALIENTES	EL LLANO	323.0	262.0	0.0	0.0	0.0	447.0	0.0	0.0	8849
AGUASCALIENTES	SAN FRANCISCO DE LOS ROMO	202.0	324.0	0.0	63.0	0.0	133.0	63.0	63.0	4173
AGUASCALIENTES	ASIENTOS	778.0	428.0	133.0	0.0	0.0	977.0	430.0	48.0	14893
AGUASCALIENTES	CALVILLO	693.0	820.0	92.0	0.0	0.0	749.0	172.0	612.0	13875
--	--	--	--	--	--	--	--	--	--	--
ZACATECAS	VILLA GONZALEZ ORTEGA	894.0	0.0	149.0	0.0	0.0	298.0	0.0	0.0	4768
ZACATECAS	VILLANUEVA	1179.0	410.0	0.0	0.0	0.0	0.0	188.0	205.0	6715
ZACATECAS	ZACATECAS	1925.0	640.0	945.0	355.0	0.0	2216.0	303.0	706.0	24630
ZACATECAS	TRANCOSO	0.0	407.0	0.0	0.0	0.0	2035.0	0.0	0.0	5698
ZACATECAS	CONCEPCION DEL ORO	0.0	0.0	224.0	0.0	0.0	171.0	0.0	0.0	2081

Figura 10: Tabla final



### **3. Anexos**

Puedes acceder a todo el material mencionado en el siguiente SharePoint:

[CrediClub Sharepoint Potencial Mercado.](#)