

Comisión: Coderhouse #49100

Integrante: CORINTTI, Franco Ezequiel

Fecha de presentación: 21/03/2024

Tabla de Contenidos

Tabla de Versiones	3
Introducción.....	4
Descripción de la temática de los datos	4
Alcance.....	4
Hipótesis.....	4
Herramientas tecnológicas implementadas	5
Diagrama de Entidad-relación	5
Listado de tablas	6
Listado de columnas de cada tabla	8
Proceso de normalización	9
Modelo de datos relacional	9
Tabla de Medidas	10
Análisis General.....	11
Análisis de Ingresos	11
Análisis Educacional.....	11
Análisis Educacional.....	¡Error! Marcador no definido.

Tabla de Versiones

Version	Fecha
1.0	23/12/2023
2.0	22/01/2024
3.0	28/02/2024
Final	21/03/2024

Introducción

La Encuesta Anual de Hogares del año 2019 en la Ciudad Autónoma de Buenos Aires (CABA) proporciona una mirada detallada y significativa sobre la composición socioeconómica de los hogares en la región. Este estudio exhaustivo ofrece insights cruciales sobre diversos aspectos que abarcan desde la demografía hasta los niveles educativos y los ingresos familiares. A través de tres tablas de Excel que abordan información general, ingresos y educación, se revelan tendencias, desafíos y oportunidades que impactan directamente en el desarrollo y bienestar de la población porteña. Esta herramienta analítica se convierte en un recurso invaluable para formuladores de políticas, investigadores y aquellos interesados en comprender y abordar las dinámicas sociales y económicas de una de las principales urbes de América Latina.

Descripción de la temática de los datos

La base de datos trata sobre la encuesta anual de hogares es realizada por la DGEyC en todo el territorio de la Ciudad de Buenos Aires.

Alcance

Estratégico y Táctico:

Estratégico: Para responsables de políticas públicas y planificadores urbanos, el análisis puede proporcionar información valiosa para la toma de decisiones a largo plazo.

Táctico: Investigadores y analistas pueden utilizar los resultados para diseñar intervenciones específicas y programas de desarrollo.

Tipo de análisis:

Descriptivo: Analizar la distribución demográfica, niveles educativos, estado de salud y situación laboral en la Ciudad de Buenos Aires en 2019, utilizando gráficos y estadísticas para resumir patrones clave y características principales.

Diagnóstico: A través de un dashboard, se buscará explorar las relaciones entre nivel educativo, salud y situación laboral, identificando factores ocultos y diferencias demográficas.

Hipótesis

La hipótesis se plantea de la siguiente manera:

“Existe una relación significativa entre el nivel educativo, la salud y la situación laboral de los habitantes de la Ciudad de Buenos Aires en 2019.”

Examinar si las personas con niveles educativos más altos tienden a tener mejores condiciones de salud y si esto se traduce en una situación laboral más favorable. También explorar si existen

diferencias significativas en estas relaciones entre distintos grupos demográficos (por ejemplo, según la edad, género o nivel socioeconómico).

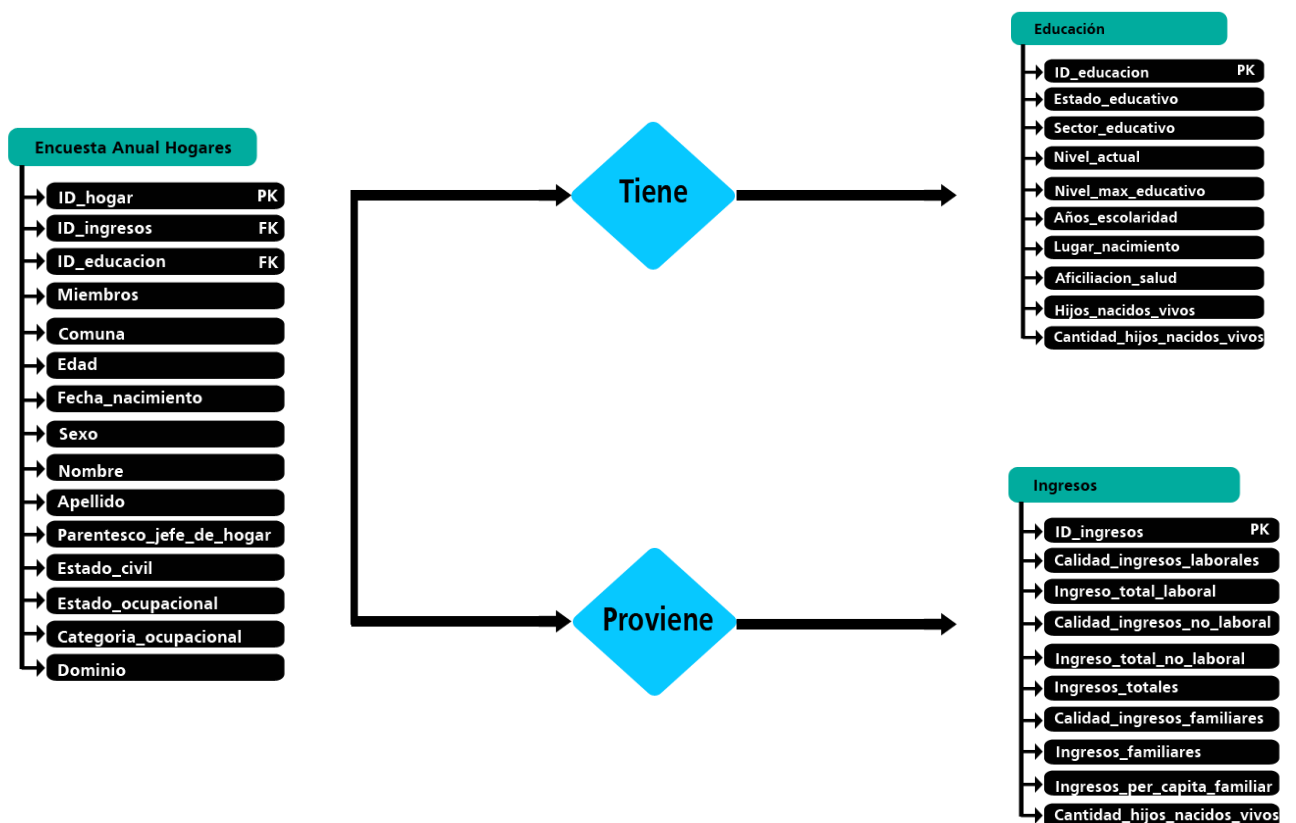
La hipótesis brinda flexibilidad para analizar una amplia gama de relaciones dentro del dataset, permitiéndote explorar la complejidad de cómo se entrelazan el nivel educativo, la salud y la situación laboral de los habitantes de la Ciudad de Buenos Aires en 2019.

Herramientas tecnológicas implementadas

Para la realización del proyecto final se utilizaron las siguientes herramientas:

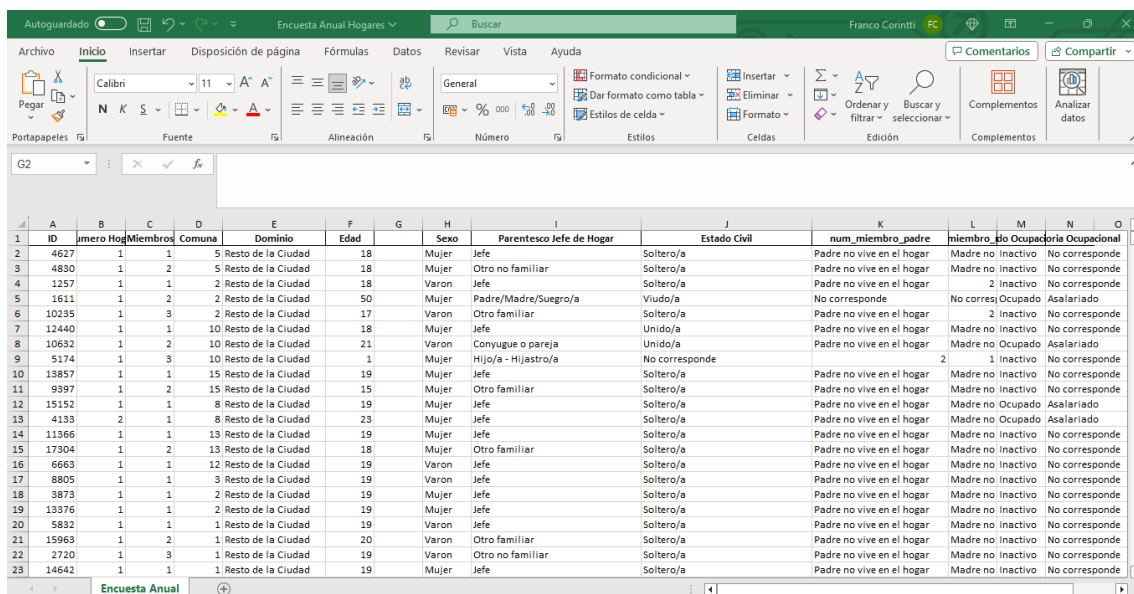
- Excel
- SQL Server
- Power BI Desktop

Diagrama de Entidad-relación



Listado de tablas

- **Tabla Encuesta Anual Hogares:** muestra los datos personales de los encuestados, la comuna a la que pertenecen, su estado civil y el estado ocupacional en el cual se encuentran.
 - ID_Hogar: PK.
 - ID_ingresos: FK.
 - ID_educacion: FK.
 - Miembros
 - Comuna
 - Dominio
 - Edad
 - Fecha_de_nacimiento
 - Sexo
 - Nombre
 - Apellido
 - Parentesco_jefe_de_hogar
 - Estado_civil
 - Estado_ocupacional
 - Categoria_ocupacional



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	ID	Hogar	Miembros	Comuna	Dominio	Edad	Sexo	Parentesco Jefe de Hogar	Estado Civil	num_miembro_padre	miembro_no ocupado	Ocupación			
1	4627	1	1	5	Resto de la Ciudad	18	Mujer	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
2	4890	1	2	5	Resto de la Ciudad	18	Mujer	Otro no familiar	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
3	1257	1	1	2	Resto de la Ciudad	18	Varon	Jefe	Soltero/a	Padre no vive en el hogar	2 Inactivo	No corresponde			
4	1611	1	2	2	Resto de la Ciudad	50	Mujer	Padre/Madre/Suegro/a	Viudo/a	No corresponde	No corresponde	Ocupado	Asalariado		
5	10235	1	3	2	Resto de la Ciudad	17	Varon	Otro familiar	Soltero/a	Padre no vive en el hogar	2 Inactivo	No corresponde			
6	12440	1	1	10	Resto de la Ciudad	18	Mujer	Jefe	Unido/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
7	10632	1	2	10	Resto de la Ciudad	21	Varon	Conyugue o pareja	Unido/a	Padre no vive en el hogar	Madre no Ocupado	Asalariado			
8	5174	1	3	10	Resto de la Ciudad	1	Mujer	Hijo/a - Hijastro/a	No corresponde	2	1 Inactivo	No corresponde			
9	13857	1	1	15	Resto de la Ciudad	19	Mujer	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
10	9997	1	2	15	Resto de la Ciudad	15	Mujer	Otro familiar	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
11	15152	1	1	8	Resto de la Ciudad	19	Mujer	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Ocupado	Asalariado			
12	4133	2	1	8	Resto de la Ciudad	23	Mujer	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Ocupado	Asalariado			
13	11366	1	1	13	Resto de la Ciudad	19	Mujer	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
14	17304	1	2	13	Resto de la Ciudad	18	Mujer	Otro familiar	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
15	6663	1	1	12	Resto de la Ciudad	19	Varon	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
16	8805	1	1	3	Resto de la Ciudad	19	Varon	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
17	3873	1	1	2	Resto de la Ciudad	19	Mujer	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
18	13376	1	1	2	Resto de la Ciudad	19	Mujer	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
19	5832	1	1	1	Resto de la Ciudad	19	Varon	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
20	15963	1	2	1	Resto de la Ciudad	20	Varon	Otro familiar	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
21	2720	1	3	1	Resto de la Ciudad	19	Varon	Otro no familiar	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			
22	14642	1	1	1	Resto de la Ciudad	19	Mujer	Jefe	Soltero/a	Padre no vive en el hogar	Madre no Inactivo	No corresponde			

- **Tabla Educación:** refleja los datos el nivel educativo de los encuestados, también otros datos socio culturales como el lugar de nacimiento, su afiliación a la salud y la cantidad de hijos.
 - ID_educacion: PK.
 - Estado_educativo
 - Sector_educativo
 - Nivel_actual
 - Nivel_max_educativo
 - Años_escolaridad
 - Lugar_nacimiento

Listado de columnas de cada tabla

Encuesta Anual Hogares		
Campo	Tipo de campo	Tipo de clave
ID_hogar	varchar(5)	PK - index
ID_ingresos	varchar(15)	FK
ID_educacion	varchar(5)	FK
Miembros	Int	-
Comuna	Int	-
Dominio	varchar(25)	-
Edad	Int	-
Fecha_de_nacimiento	Datetime	-
Sexo	varchar(5)	-
Nombre	varchar(10)	-
Apellido	varchar(10)	-
Parentesco_jefe_de_hogar	varchar(30)	-
Estado_civil	varchar(30)	-
Estado_ocupacional	varchar(10)	-
Categoria_ocupacional	varchar(30)	-

Educación		
Campo	Tipo de campo	Tipo de clave
ID_educacion	varchar(5)	PK - index
Estado_educativo	varchar(25)	-
Sector_educativo	varchar(15)	-
Nivel_actual	varchar(25)	-
Nivel_max_educativo	varchar(25)	-
Años_escolaridad	varchar(25)	-
Lugar_nacimiento	varchar(15)	-
Afiliacion_salud	varchar(25)	-
Hijos_nacidos_vivos	Int	-
Cantidad_hijos_nacidos_vivos	Int	-

Ingresos		
Campo	Tipo de campo	Tipo de clave
ID_ingresos	varchar(15)	PK - index
Calidad_ingresos_laborales	varchar(25)	-
Ingreso_total_laboral	Int	-
Calidad_ingresos_no_laborales	varchar(25)	-
Ingreso_total_no_laboral	Int	-
Ingresos_totales	Int	-
Calidad_ingresos_familiares	varchar(25)	-
Ingresos_familiares	Int	-
Ingreso_per_capita_familiar	Int	-

Proceso de normalización

La base de datos de obtuvo de <https://data.buenosaires.gob.ar/dataset/encuesta-anual-hogares>

El archivo originalmente estaba en formato CSV, el proceso se inicia por separar en columnas y a su vez, separar la información en distintas hojas (de la hoja matriz se desglosaron en 3 tablas). Se le aplicó el formato correspondiente a cada columna (fecha, contabilidad, texto, etc).

Los cambios ocurridos en función de a la primera entrega, de la tabla “Encuesta anual de hogares” se eliminaron columnas y se agregaron otras como la fecha de nacimiento, el nombre y apellido. También se movieron de la misma las últimas columnas a la tabla de “Educación” debido a que se consideró que la información era más relevante para ese tipo de tabla y para las relaciones que establecía.

Se generaron ID tanto para la tabla de “Educación” (ID numérico) como para la de “Ingresos” (ID alfanumérico), esto permitió aplicar el diagrama E-R de manera correcta.

Otro cambio fue en la tabla de “Educación” la cantidad de hijos vivos inicialmente figuraba como “No corresponde” y se cambió directamente por el número 0 (cero).

Modelo de datos relacional

Una vez cargadas las tablas previamente mencionadas al Power Bi, se realizaron transformaciones para poder trabajar con los datos de una manera más eficiente.

Se realizaron las siguientes transformaciones y creaciones:

- La aplicación de formatos en moneda a la tabla de “Ingresos”, tanto laborales como no laborales y familiares;
- Aplicación del formato fecha en la tabla de “Encuesta Anual Hogares” en las fechas de nacimiento debido a que había valores que los tomaba como número;

- Creación de columnas de rangos de edades en la tabla de “Encuesta Anual Hogares”, con el fin de poder graficar y contabilizar teniendo en cuenta la diversidad de datos que se manejan en cuanto a la edad;
- Selección y reemplazo de valores *null* en la tabla de “Educación”, en la columna de “Afiliacion_salud”, se le asignaron valores de manera que puedan ser representativos a la hora de asignar gráficos pertinentes;
- Creación de tabla calendario en función de la fecha de nacimiento. Esto se va a poder utilizar como filtro para definir un rango de edad específico;
- En función de la tabla calendario, se creó una tabla puente para poder solventar el problema de la relación de muchos a muchos.

A continuación, se deja una imagen con el diagrama de entidad relación creado en Power Bi:

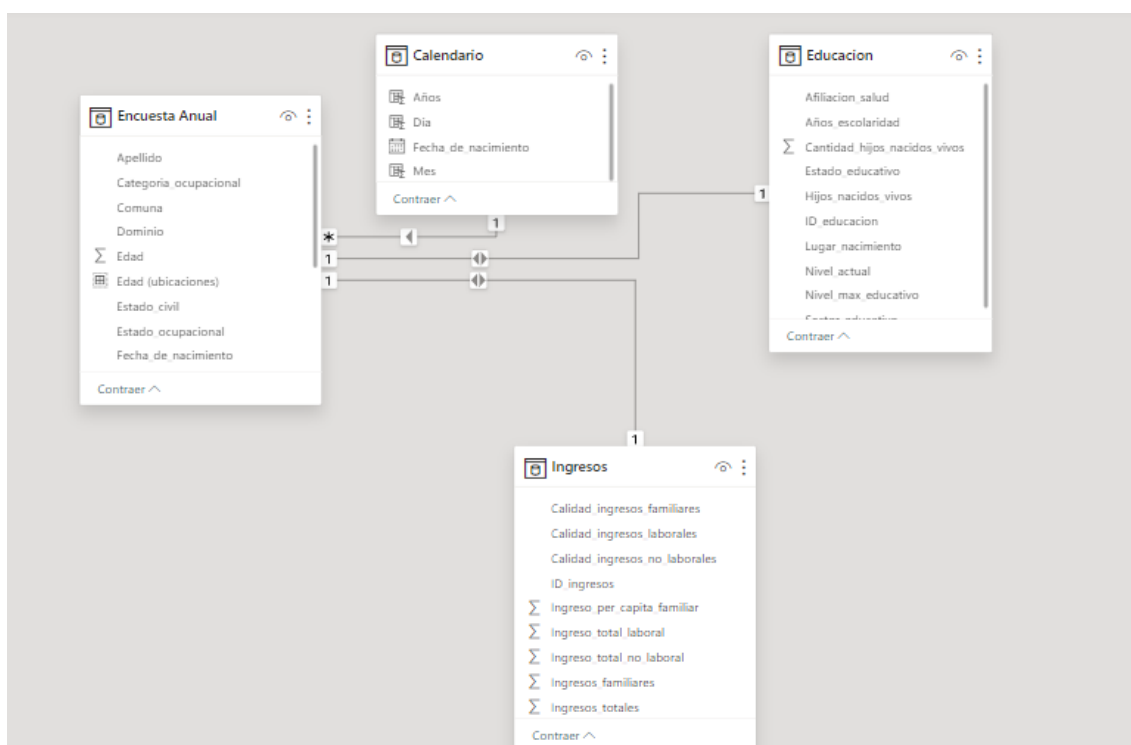


Tabla de Medidas

Las medidas creadas en la Tabla de Medidas son:

- 1) Cantidad de Encuestados = `COUNT('Encuesta Anual'[ID_Hogar])`
- 2) Edad Promedio = `AVERAGE('Encuesta Anual'[Edad])`
- 3) Promedio Cantidad Hijos = `AVERAGE(Educacion[Cantidad_hijos_nacidos_vivos])`
- 4) Total de Hijos Nacidos = `SUM(Educacion[Cantidad_hijos_nacidos_vivos])`
- 5) Total Ingresos Laborales = `SUM(Ingresos[Ingreso_total_laboral])`

6) Total Ingresos No Laborales = SUM(Ingresos[Ingreso_total_no_laboral])

7) Total Ingresos = [Total Ingresos Laborales]+[Total Ingresos No Laborales]

Análisis General

La primera solapa denominada "Encuesta Anual Hogares", constituye el núcleo de la recopilación de datos de la Encuesta Anual de Hogares del año 2019 en la Ciudad Autónoma de Buenos Aires (CABA). Esta tabla ofrece una visión detallada de los encuestados, presentando información personal clave, como nombres, apellidos, sexo, rango de edades y fecha de nacimiento. Además, proporciona datos relevantes sobre el contexto socioeconómico de los hogares, incluyendo la comuna a la que pertenecen, la estructura familiar, el estado civil y la situación ocupacional de los miembros del hogar. Cada entrada está identificada por un ID_Hogar único, y está vinculada a las otras tablas del conjunto de datos a través de claves foráneas (FK) que permiten un análisis integral de los datos.

Como segmentadores se utilizaron los siguientes:

- Comuna;
- Fecha de nacimiento (a partir de la tabla calendario);
- Sexo.

Análisis de Ingresos

La solapa "Ingresos" ofrece un análisis detallado de los ingresos de los encuestados, abarcando tanto las fuentes laborales como las no laborales, así como sus ingresos totales. Esta información crucial proporciona una visión amplia de la situación económica de los hogares encuestados, permitiendo entender tanto la contribución individual al ingreso familiar como la distribución de los ingresos dentro del hogar. La tabla está estructurada para identificar la calidad de los ingresos, su origen y su magnitud, arrojando luz sobre la diversidad y la dinámica de los ingresos en la población estudiada.

Como segmentadores se utilizaron los siguientes:

- Comuna;
- Sexo.

Análisis Educacional

La solapa "Educación" ofrece una visión detallada del nivel educativo de los encuestados, así como otros datos socio-culturales relevantes, como su lugar de nacimiento, afiliación a servicios de salud y la cantidad de hijos. Esta información es fundamental para comprender el perfil educativo de la población estudiada, incluyendo su trayectoria educativa y los niveles alcanzados. Además, los datos sobre lugar de nacimiento y afiliación a la salud proporcionan contexto adicional sobre el origen y el acceso a servicios básicos de la población encuestada. La inclusión de datos sobre hijos nacidos vivos añade un componente importante para comprender la dinámica familiar y su relación con la educación.

Como segmentadores se utilizaron los siguientes:

- Comuna;

Conclusión

Del análisis realizado sobre todos los gráficos puede respaldar la hipótesis propuesta. De total de encuestados (teniendo en cuenta todas las comunas) en cuanto a su categoría ocupacional en primer lugar figura como “no corresponde”, seguido en segundo lugar por la categoría de “asalariados” y así sucesivamente, esto nos indica que si bien una gran parte se encuentra asalariada el mayor porcentaje se desconocen los datos, pero dicho datos relacionados con los ingresos y la educación se verán a continuación. A su vez, determinándose en primer que el estado ocupacional se encuentra “Ocupado” seguido de los “Inactivos” correspondiente al rango de edades de entre 21-30 años y 31-40 años.

En cuanto a la distribución de ingresos, se observa que la mayor proporción del total de ingresos está compuesta por ingresos laborales, tanto en el sexo masculino como en el femenino. Este hallazgo sugiere que, independientemente del género, el trabajo sigue siendo la principal fuente de ingresos para los habitantes de la Ciudad de Buenos Aires en 2019. Este aspecto resalta la importancia de la actividad laboral como pilar fundamental en la economía de los hogares encuestados, subrayando la necesidad de políticas que promuevan el empleo y mejoren las condiciones laborales para garantizar la estabilidad financiera de la población.

El análisis de la variable educativa revela que un porcentaje significativo de los encuestados (69.16%) reporta que su sector educativo “no corresponde”, seguido por un 18.81% que asiste a instituciones educativas estatales. Aunque esta variable no muestra una correlación directa con la hipótesis inicial, es importante destacar que el acceso a la atención médica presenta una tendencia clara. La mayoría de los encuestados tienen acceso a la obra social, seguido por aquellos que solo tienen acceso al sistema de salud público. Además, se observa una concentración significativa de encuestados nacidos en la Ciudad de Buenos Aires, seguidos por aquellos nacidos en otras provincias, lo que puede tener implicaciones en la distribución y calidad de los servicios de salud y educativos en la región. Estos hallazgos resaltan la complejidad de las interacciones entre el nivel educativo, el acceso a la salud y otras variables demográficas, lo que subraya la importancia de un enfoque integral en el diseño de políticas que busquen mejorar el bienestar de la población.

Futuras Líneas

Considerando los resultados de este análisis, se pueden proponer varias iniciativas para abordar las brechas identificadas y promover el bienestar general de la población de la Ciudad de Buenos Aires.

- 1) Mejora de la calidad educativa: Implementar políticas que mejoren la calidad de la educación, asegurando un acceso equitativo y adecuado a recursos educativos y programas de apoyo.
- 2) Fortalecimiento del acceso a la salud: Desarrollar programas que mejoren el acceso a la atención médica, especialmente para aquellos que no cuentan con obra social o seguro médico privado, garantizando la disponibilidad de servicios de salud de calidad en todas las áreas de la ciudad.
- 3) Fomento del empleo y desarrollo económico: Diseñar estrategias para fomentar la creación de empleo y el desarrollo económico en la ciudad, especialmente en sectores que puedan beneficiar a personas con diferentes niveles educativos y habilidades laborales.

- 4) Investigación y seguimiento continuo: Realizar estudios adicionales para comprender mejor las interrelaciones entre la educación, la salud y la situación laboral, y evaluar el impacto de las iniciativas implementadas en la mejora del bienestar de la población.