Programa Anual de Investigación del INEGI 2021

Contenido

Introducción:.....	4
Participantes en el ecosistema de investigación del INEGI*	4
Marco legal	10
El Programa Anual de Investigación del INEGI 2021 está integrado por:	10
Proyectos del PAINEGI 2021	12
Actualización cartográfica.....	13
1. Percepción remota para la estimación anual de la frontera agrícola nacional.	14
2. Implementación del modelo predictivo de desertificación en México.....	16
3. Detección de crecimientos urbanos y nuevas localidades con imágenes satelitales desde la ciencia de datos. (CGI).....	18
4. Identificación de panteones mediante imágenes satelitales. (DGEOSPJ).....	20
Coyuntura económica	22
5. Nowcasting para componentes mensuales de la demanda.....	23
6. Indicadores compuestos de alta frecuencia.....	25
7. Indicadores de la actividad económica generados a partir de información bancaria.....	27
Estructura económica.....	29
8. Indicadores de actividad económica municipal incluyendo mediciones satelitales de luminosidad nocturna.....	30
9. (7-2020) Aprovechamiento de los Censos Económicos y las Encuestas Especiales en Establecimientos para Contextualizar el Comportamiento de la Productividad en México.....	32
10. (10-2020) Modelos de Equilibrio General Computable para la Economía Mexicana.	34
11. (11-2020) Patrones de consumo de los hogares mexicanos.....	36
12. (9-2020) Construcción de una Matriz de Insumo-Producto Multi-regional de México año 2013.	38
13. Clasificación automática de unidades económicas de acuerdo con el apoyo que brindan a las mujeres. (Vicepresidencia a cargo del Subsistema Nacional de Información Geográfica, Medio Ambiente, Ordenamiento Territorial y Urbano).....	40
Ingresos	42
14. Valor imputado al uso de la vivienda por sus propietarios.....	43

15. Construcción del Modelo estadístico 2020 para la continuidad del MCS-ENIGH.....	45
16. Determinación de la subcobertura e imputación del ingreso corregido por MPVR en la ENIGH.....	46
Trabajo	48
17. Estructura y determinantes de la segregación ocupacional por tareas.	49
18. (29-2020) Culminación del modelo de experimentación e implementación de las recomendaciones de las CIET XIX y XX.	51
19. Preparación del documento de discusión de la XXI Conferencia Internacional de Estadísticos del Trabajo en materia de medición de la informalidad.	53
20. Codificación Automática de actividad económica y ocupación con Métodos de <i>Machine Learning</i> para proyectos estadísticos en hogares. Segunda etapa. (DGES).....	56
Bienestar	58
21. (12-2020) Uso de técnicas de BIG DATA a partir de información de Twitter para bienestar subjetivo y salud mental.	59
22. Actualización con CPV 2020 de la Identificación de nivel de marginación en asentamientos humanos mediante la integración y análisis datos estadísticos, geográficos y de percepción remota.	61
23. Indicadores sintéticos de salud física y mental con información longitudinal.	64
24. Sistema de indicadores del bienestar en México y sus entidades federativas en 2021.	66
Sociodemografía	68
25. (21-2020) Evolución de la consistencia de las estimaciones de fecundidad entre la EDER 2011 y la EDER 2017.	69
26. Corregir sesgos poblacionales de resultados obtenidos a partir de Twitter... 71	
27. Estratificador multivariado automático para el Censo de Población y Vivienda 2020.....	73
Recursos naturales.....	75
28. Análisis del consumo de agua en la industria manufacturera de México.....	76
Buscador de la página del INEGI.....	78
29. Buscador del Sitio del INEGI. (DGCSPIRI)	79

De conformidad con lo establecido en los artículos 29 fracción XI y 32 fracciones I, II III y V del REGLAMENTO INTERIOR DEL INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA, el Presidente de la Junta de Gobierno del INEGI somete a consideración de la Junta de Gobierno el **Programa Anual de Investigación del INEGI 2021**, para su aprobación.

Introducción:

La actividad económica y social está en permanente evolución; en esta era de profundas transformaciones que es preciso invertir tiempo y recursos para comprender dicho cambio mediante la generación de información sobre las diferentes dimensiones del cambio, a través de la realización de un metódico y constante trabajo de exploración de esa realidad, tomando ventaja de las nuevas técnicas y del avance del conocimiento científico. La tarea de investigación e innovación para la producción de información ha tenido una importancia creciente en los organismos responsables de la generación de la información estadística y geográfica, México no ha estado al margen de esta tendencia.

Para el desarrollo de sus actividades de investigación, el INEGI se vale del método científico con el propósito de generar conocimiento que le permita fortalecer sus capacidades para atender los objetivos que la Ley le impone. Así, con la generación y promoción de trabajos de investigación científica y tecnológica, el INEGI busca apuntalar sus capacidades para suministrar a la sociedad y al Estado Información de calidad, pertinente, veraz y oportuna, a efecto de coadyuvar al desarrollo nacional. Esto se hace mediante la generación de conocimiento sobre procesos y productos que atiende a tres preguntas básicas: 1) ¿Qué no hacemos que podríamos hacer para enriquecer la oferta de información estadística y geográfica oficial? 2) ¿Cómo podemos añadir valor a lo que ya hacemos? 3) ¿Cómo podemos hacer mejor lo que hacemos?

La atención de las necesidades de investigación del INEGI se hace con la participación de agentes y recursos, susceptibles de ser concebidos como un ecosistema, los cuales se listan a continuación:

Participantes en el ecosistema de investigación del INEGI*

Directos	Indirectos
<ul style="list-style-type: none"> • Investigadores permanentes. • Investigadores visitantes. • Convenios de colaboración nacionales e internacionales sobre temas de investigación. • Laboratorio de ciencia de datos. • Funcionarios de las Unidades Administrativas del INEGI que eventualmente realizan actividades de investigación. 	<ul style="list-style-type: none"> • Realidad Datos y Espacio. Revista Internacional de Estadística y Geografía. • Maestrías y diplomados. • Tesis de maestría. • Laboratorio de microdatos. • Hackatones y concursos de investigación. • Seminarios.

*/ El Fondo Sectorial CONACYT-INEGI, actualmente en proceso de extinción por mandato de ley, por medio del cual se financiaron 47 proyectos de investigación (10 aún en proceso de elaboración) en un lapso de 11 años; a partir de 2021 deja de reportarse como parte de los integrantes del ecosistema de investigación del INEGI dado que ya no será posible canalizar a través de él nuevas necesidades de investigación.

El marco legal del INEGI le permite disponer de instrumentos para desarrollar actividades de investigación tanto por cuenta propia como por vía de cooperación con terceros. El presente programa hace referencia a los proyectos de investigación a ser desarrollados por el INEGI.

Los trabajos de la Dirección General Adjunta de Investigación están dirigidos a apuntalar las capacidades del Instituto para suministrar a la sociedad y al Estado Información de calidad, pertinente, veraz y oportuna, a efecto de coadyuvar al desarrollo nacional. Dicho apuntalamiento, ofreciendo soluciones útiles a problemas reales, se hace evidente en ejemplos como los que se citan a continuación, elaborados o liberados durante 2019-2020.

Solicitante	Problema	Pregunta	Solución o Propuesta
Usuarios/ Presidente INEGI	El Indicador Global de la Actividad Económica (IGAE) aparece 56 días luego de cerrado el mes de referencia.	¿Cómo lograr mayor oportunidad en la medición del IGAE?	<ul style="list-style-type: none"> ▶ Indicador Oportuno de la Actividad Económica (IOAE).
Usuarios/ Presidente INEGI	La pandemia de 2020 afecta de manera diversa a sectores económicos y grupos sociales.	¿Cómo lograr una imagen oportuna y detallada del impacto económico y social de la pandemia?	<ul style="list-style-type: none"> ▶ Matrices de Contabilidad Social (MCS) ▶ Desagregación de IOAE con MCS
Dirección General de Estadísticas Sociodemográficas (DGES)	La codificación de actividad económica y ocupación es lenta y costosa.	¿Cómo clasificar de manera más rápida, precisa y barata?	<ul style="list-style-type: none"> ▶ Codificación automática con "machine learning".
Dirección General de Estadísticas Sociodemográficas (DGES)/ Organización Internacional del Trabajo (OIT)	El marco de referencia de la Encuesta Nacional de Ocupación y Empleo (ENOE) se está rezagando de los fenómenos laborales y las recomendaciones internacionales.	¿Cómo implementar las recomendaciones de la XIX y XX Conferencia Internacional de Estadísticos del Trabajo (CIET)?	<ul style="list-style-type: none"> ▶ Estrategia para implementación. ▶ Pruebas cognitivas. ▶ Pruebas de campo.

Solicitante	Problema	Pregunta	Solución o Propuesta
Dirección General de Estadísticas Económicas (DGEE)	No se cuenta con indicadores de actividad económica municipal.	¿Es posible generar indicadores de Actividad Económica Municipal consistentes con la información sectorial y estatal?	► Indicador de la Actividad Económica Municipal (IAEM)
Dirección General de Estadísticas Económicas (DGEE)	Series macro muy cortas.	¿Cómo ha evolucionado la economía en los últimos 40 años?	► Retropolación del Indicador Trimestral de la Actividad Económica Estatal (ITAEE) hasta 1980

Cabe señalar que la vocación de la Dirección General Adjunta de Investigación del INEGI es la investigación y no la producción de información; por lo tanto, cuando de sus actividades se desprenden proyectos o programas de generación de información, estos podrán ser incubados temporalmente en el área, normalmente en colaboración con otras áreas, pero eventualmente las tareas de generación regular de información deberán ser traspasada a alguna de las Unidades Administrativas del INEGI (o de las Unidades del Estado del Sistema Nacional de Información Estadística y Geográfica -SNIEG-) con vocación de generación de información estadística o geográfica. Así las cosas, por ejemplo, proyectos como la retropolación del ITAEE o las medidas de informalidad laboral de conformidad con la Matriz de Hussmanns, requirieron de una incubación muy breve por parte del área de investigación y desde un principio se reportaron al público como parte de la producción de línea de la DGEE y la DGEI, respectivamente. En contraste, la medición del bienestar subjetivo por medio del BIARE Básico o el Estado de Ánimo de los Tuiteros, han requerido de una incubación más prolongada, se han publicado en principio como estadísticas experimentales y están pendientes de ser transferidos a alguna Unidad Administrativa con vocación de producción de información. En todo caso, la finalidad es apuntalar el trabajo de las áreas encargadas de producir y difundir la información del INEGI y del SNIEG.

La actividad de investigación es por antonomasia colaborativa. Estar en condiciones de apuntalar las actividades del INEGI con propuestas innovadoras supone cercanía con los productores de información, pero también con la academia, el sector privado y agencias y expertos nacionales e internacionales. La Dirección General Adjunta de Investigación realiza seminarios académicos de manera continua en los cuales aborda temas de relevancia para las agendas de producción y difusión de información y de investigación del INEGI. En el siguiente cuadro se enlistan los 6 seminarios más recientemente realizados, indicando la procedencia de los participantes:¹

¹ <https://www.inegi.org.mx/eventos/seminarios/default.html>

Seminarios realizados más recientemente, indicando la procedencia de los conferencistas y organismos internacionales involucrados:

Fechas de realización	Nombre del seminario	Objetivo general	Total de ponentes nacionales	Total de ponentes del extranjero
Noviembre 3, 5 y 12, 2020	<p><i>Series of conferences: New data and methods for generating official information</i> https://www.inegi.org.mx/eventos/2020/newData/</p>	<p>Revisar las posibilidades de acceder a y explotar nuevas fuentes de información en posesión de repositorios privados de información, de los estados y/o de entidades gubernamentales, con una clara aplicación operacional, para generar información de interés público, en temas económicos y sociales.</p>		<p>6 Italia (2), Portugal, Canadá, Reino Unido y Alemania</p>
4, 11, 18, 25 de septiembre 2, 9 y 16 de octubre, 2020	<p>Mesas de diálogo: Transformaciones para la sostenibilidad https://www.inegi.org.mx/eventos/2020/transos/</p>	<p>Examinar los impactos ambientales de la actividad económica en México y discutir sus implicaciones para el bienestar de la población, así como ponderar las transformaciones necesarias para la sostenibilidad, identificando y valorando la información existente o necesaria para examinar la relación entre el medio ambiente y el bienestar de la población en México.</p>	20	<p>3 Estados Unidos (2)</p>

Fechas de realización	Nombre del seminario	Objetivo general	Total de ponentes nacionales	Total de ponentes del extranjero
Junio 26, julio 3, 10, 17, 24 y 31, agosto 6, 2020	<i>Towards a Comprehensive Measurement of Well-being: Series of invited lectures</i> https://www.inegi.org.mx/eventos/2020/towards/	Apoyar al grupo mexicano de expertos: Grupo Técnico de Expertos para la Medición del Ingreso y el Bienestar GTMIB, en su diseño de recomendaciones para una nueva encuesta nacional de hogares para medir el ingreso y el bienestar.		9 España, Nueva Zelanda (3), Italia (2), Reino Unido (2), República Checa
30 y 31 de octubre de 2019	Seminario internacional sobre Fechado de los Ciclos Económicos https://www.inegi.org.mx/eventos/2019/cide/	Analizar los datos y metodologías utilizadas por comités de fechado de ciclos económicos alrededor del mundo, así como alimentar la discusión de la pertinencia de que México cuente con un comité de este tipo.		7 Brasil (2) Canadá España Colombia Bélgica Estados Unidos
16 y 17 de mayo de 2019	Conferencia sobre grandes datos en Ciencias Sociales y Políticas Públicas https://www.inegi.org.mx/eventos/2019/bigdata/	Presentar trabajos relevantes en ciencias sociales y políticas públicas, basados en las nuevas fuentes de información, así como en metodologías de inteligencia artificial, para el uso de los datos y modelado.	5	8 Estados Unidos (2), España (3) Países Bajos, Canadá y Reino Unido
31 de enero y 1º. de febrero de 2019	<i>Seminar on Responsive and Adaptive Survey Design</i> https://www.inegi.org.mx/eventos/2019/seminar/	Presentar las posibilidades que las encuestas flexibles y adaptables ofrecen en un entorno de encuestas complejo y desafiante.	1	6 Estados Unidos (2), Países Bajos (2) y Reino Unido(2)

La actividad de investigación requiere también de un terreno fértil para que maximice su impacto y se beneficia de la formación de cuadros de alto nivel técnico-académico en las diferentes áreas del INEGI. Es por lo que también desde la Dirección General Adjunta de Investigación se impulsa la formación de este tipo de cuadros mediante recursos académicos como la Maestría en Análisis Estadístico y Computación, impartida por el Centro de Investigación en Matemáticas (CIMAT), que en 2021 iniciará los cursos de su quinta generación (la segunda con el nuevo programa más orientado a *ciencia de datos*).

En lo que corresponde específicamente al desarrollo de proyectos de investigación, se tiene con frecuencia participación activa en seminarios y foros internacionales en los que se recibe retroalimentación útil para enriquecer los proyectos y que provocan también nuevas ideas. Para citar algunos casos, investigadores del INEGI participan en el *High Level Expert Group for the Modernisation of Official Statistics* con algunos de los trabajos sobre *ciencia de datos aplicada* que forman parte del programa de investigación del INEGI. Asimismo, expertos del área en temas laborales participan activamente junto con la OIT en la implementación de los acuerdos de las Conferencias Internacionales sobre Estadísticas del Trabajo (CIET), además de que también participan propositivamente en las discusiones de las que derivan las emendaciones de las CIET actuales y venideras.

También se trabaja sobre diversos temas con expertos nacionales e internacionales del más alto nivel, incluyendo modelos de equilibrio general computable, matrices de contabilidad social (con Sherman Robinson, del *International Food Policy Research Institute* y el *Peterson Institute of International Economics*), migración México-EEUU y su efecto en el empleo y la inversión en ambos lados de la frontera (Michael Gove, de la *University of North Georgia*), impacto de la digitalización en la producción de información económica, corrección de la distribución del ingreso de la ENIGH con Nora Lustig (*Tulane University*), etcétera. Cabe mencionar también que la cooperación con las áreas estadísticas de la Organización para la Cooperación y el Desarrollo Económico (OCDE) y la Comisión Económica para América Latina y el Caribe (CEPAL) es frecuente. Detrás de la mayor parte de los trabajos de investigación que se generan existen artículos académicos de investigación que se publican en revistas especializadas *arbitradas* y libros, tanto del país como del extranjero. La revista científico-técnica del INEGI "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía" es también una plataforma de difusión de los trabajos de investigación promovidos desde el INEGI.

La cooperación para la investigación ha llevado a la conformación de grupos técnicos de expertos como el formado en cooperación con el Centro de Investigación y Docencia Económicas (CIDE) para evaluar la pertinencia de echar a andar un comité de fecho de ciclos económicos de México "Grupo Técnico de Expertos para el diseño de un Comité para el Fecho de los Ciclos de la Economía de México (GTDCFC)" y el formado en cooperación con la Universidad Iberoamericana (UIA) para diseñar desde cero una nueva encuesta social general "Grupo Técnico de Expertos para la Medición del Ingreso y el Bienestar". Estos grupos técnicos descansan en los esfuerzos articulados y enfocados a la solución de problemas específicos, de especialistas diversos y del más alto nivel.²

² [Propuesta para crear un comité de fecho de los ciclos económicos en México](http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/promo/GTDCFC_2020.pdf)

http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/promo/GTDCFC_2020.pdf

Marco legal

1. El artículo 3 de la Ley del Sistema Nacional de Información Estadística y Geográfica establece que "El Sistema Nacional de Información Estadística y Geográfica, tiene la finalidad de suministrar a la sociedad y al Estado Información de calidad, pertinente, veraz y oportuna, a efecto de coadyuvar al desarrollo nacional".
2. El Reglamento Interior del Instituto Nacional de Estadística y Geografía, establece en su artículo 29 que, entre las atribuciones específicas de la Dirección General de Integración, Análisis e Investigación, se encuentra la mencionada en su fracción XI consistente en "Dirigir la realización de investigaciones y estudios temáticos, sectoriales y regionales".
3. El mismo Reglamento Interior, establece en su artículo 32 que las atribuciones específicas de la Dirección General Adjunta de Investigación incluyen las de:
 - I. Promover la generación de proyectos de investigación en materia de información estadística y geográfica oficial, orientados a fortalecer las capacidades del Sistema;
 - II. Desarrollar proyectos de investigación en materia de información estadística y geográfica oficial, orientados a fortalecer las capacidades del Sistema;
 - III. Promover la incorporación de métodos y técnicas de ciencia de datos para la producción de información estadística y geográfica en coordinación con las Unidades Administrativas del Instituto, así como con las Unidades del Estado e
 - V. Integrar el Programa Anual de Investigación del Instituto y someterlo por conducto del Presidente a consideración de la Junta de Gobierno para su aprobación.

El Programa Anual de Investigación del INEGI 2021 está integrado por:

- a) Los proyectos a desarrollar por parte de investigadores y por el Laboratorio de Ciencia de Datos adscritos a la Dirección General Adjunta de Investigación de la Dirección General de Integración, Análisis e Investigación.
- b) Los proyectos a desarrollar de manera conjunta con personal de otras Unidades Administrativas del INEGI, así como los proyectos a desarrollar con especialistas independientes o adscritos a otras instituciones.

Dada la naturaleza de las actividades de investigación, algunos de los proyectos pueden tener una temporalidad multianual. Es así que en este Programa se retoman proyectos de investigación registrados desde el Programa Anual de Investigación del INEGI 2020, además de otros para iniciar en 2021.

El presente Programa incorpora temas de investigación propuestos por las diferentes Unidades Administrativas del INEGI y por la Dirección General Adjunta de Investigación. En total se presentan 29 proyectos, de los cuales uno es transversal, y el resto se puede agrupar de la siguiente manera según su afinidad de subsistema:

Subsistema	Número de proyectos
Economía	9
Sociodemografía	14
Geografía y medio ambiente	4
Gobierno, seguridad pública y justicia	1
Transversal	1
Total	29

Desde la perspectiva de la temática que se aborda, estos 29 proyectos están integrados por 4 de *actualización cartográfica* (todos nuevos), 4 más sobre *bienestar* (tres nuevos y uno que continúa), uno sobre el *buscador de la página del INEGI* (nuevo), 3 sobre *coyuntura económica* (nuevos), 6 sobre *estructura económica* (dos de ellos nuevos), 3 sobre *ingresos de los hogares* (nuevos), 1 sobre *recursos naturales* (nuevo), 3 sobre *socio-demografía* (dos nuevos) y 4 sobre *trabajo* (tres de ellos nuevos). En total se presentan 22 proyectos nuevos y 7 que continúan.

Dada la carga de trabajo implícita en este Programa y los perfiles profesionales de quienes participan en la Dirección General Adjunta de Investigación, dependiendo de las circunstancias, otras necesidades de investigación podrían ser consideradas para su desahogo a través de otros especialistas del INEGI o del resultado de cooperación con investigadores externos.

Proyectos del PAINEGI 2021

Actualización cartográfica

Nombre del Proyecto:	1. Percepción remota para la estimación anual de la frontera agrícola nacional.
Descripción general del problema a resolver:	<p>En México, el principal instrumento para poder estimar la producción Agrícola es el Censo Agropecuario que han venido llevando desde 1935 con una periodicidad cercana a los 10 años. En algunas ocasiones, estos ejercicios han sido suspendidos parcialmente por temas presupuestales. Una alternativa para reducir los costos y tener mediciones de mayor frecuencia ha sido la aplicación de la Encuesta Nacional Agropecuaria.</p> <p>Por otro lado, una de las primeras aplicaciones de percepción remota para la generación de estadísticas nacionales se realizó en Estados Unidos con imágenes generadas por los satélites de la National Aeronautics and Space Administration (NASA). Estas aplicaciones se vienen realizando desde mediados de los 70's y ha venido mejorando las estimaciones obtenidas.</p> <p>Actualmente, INEGI cuenta con los recursos de información, tecnológicos y humanos, para llevar a cabo este tipo de ejercicios de percepción remota y poder obtener estimaciones suficientemente precisas para estimar la producción nacional de los principales productos agrícolas.</p>
Importancia y utilidad del tema para el INEGI:	La Implementación de técnicas de aprendizaje computacional para el análisis de grandes volúmenes de datos satelitales Landsat, permita identificar la evolución de la frontera agrícola anual de los últimos 30 años. Lo cual podrá ser aprovechado para construir modelos de estimaciones de actividad agrícola para años en los que no se tienen censos agrícolas.
Información y fuentes:	<ul style="list-style-type: none"> • Imágenes contenidas en el Cubo de Datos Geoespaciales de México, las cuales cubren el periodo de 1984 hasta el año 2020. • Datos de Frontera Agrícola disponibles.
Hipótesis del trabajo:	El uso de imágenes de satélite en combinación con técnicas de ciencia de datos permitirá complementar los trabajos desarrollados por la Dirección General de Estadísticas Económicas y construir una serie de tiempo anual desde la cual se podrá conocer la evolución de la frontera agrícola en el territorio mexicano para al menos los últimos 30 años.
Objetivo general:	Construir una serie de tiempo anual de la evolución de la frontera agrícola en el territorio mexicano para al menos los últimos 30 años
Objetivos específicos:	<ol style="list-style-type: none"> 1. Desarrollo de metodología de Ciencia de datos. 2. Construcción de conjunto de entrenamiento. 3. Implementación y aplicación de algoritmo de aprendizaje computacional. 4. Determinación de la frontera agrícola. 5. Evaluación de los resultados obtenidos por el algoritmo. 6. Validación del modelo.

Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Mapas Nacionales de Frontera Agrícola para los continuos nacionales anuales disponibles en el Cubo de Datos Geoespaciales de México. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/). • Documento con reporte de investigación.
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas (DGEE). • Dirección General de Geografía y Medio Ambiente (DGGMA). • La Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT). • Secretaría de Agricultura y Desarrollo Rural (SADER).
Responsable de la ejecución del proyecto:	Elio Atenógenes Villaseñor García.
Observaciones:	Laboratorio de Ciencia de Datos - Elio Atenógenes Villaseñor García en colaboración con personal de la DGEE.
Subsistema:	Geografía y medio ambiente, ordenamiento territorial y urbano
Temática:	Actualización cartográfica
Técnicas:	Ciencia de datos
Fuente de datos:	No tradicional
Estatus:	NUEVO

Nombre del Proyecto:	2. Implementación del modelo predictivo de desertificación en México.
Descripción general del problema a resolver:	<p>Probablemente, el cambio climático es el mayor reto que actualmente confrontamos como humanidad. Medir y predecir los impactos de este fenómeno es de gran importancia. La cuantificación de los impactos sociales y económicos del cambio climático es un problema complejo y científicamente desafiante; al mismo tiempo, los impactos del cambio climático son cada vez más notorios.</p> <p>Este proyecto se enfoca en la construcción de un modelo basado en datos para el proceso de desertificación en el norte de México. Como una posible extensión, consideramos el modelado del impacto en otras áreas de interés, como temas relacionados con el desarrollo sostenible.</p> <p>Los modelos basados en datos parecen ser una alternativa factible para construir modelos, teniendo en cuenta que en este momento existen muchos datos brutos significativos que se pueden extraer de los sistemas de información y pueden ser útiles para considerar en la construcción de este tipo de modelos. Un ejemplo de estos datos son las imágenes satelitales (Yang et al. 2013).</p>
Importancia y utilidad del tema para el INEGI:	<p>La identificación de zonas con pérdida de vegetación será de utilidad para medir el impacto que tiene el cambio climático en zonas determinadas del territorio mexicano. Particularmente, para el INEGI será importante relacionar estas mediciones de cambio de vegetación con otros indicadores sociales para poder cuantificar el impacto que tiene el medio ambiente en el desarrollo social. Habiendo establecido dichas relaciones, el INEGI podrá ser un ente que se anticipe a cambios en las condiciones sociales con base en el cambio de tipo de suelo. Aunado a esto y sabiendo que el cambio climático y sus efectos guardan estrecha relación con los recursos naturales disponibles, el INEGI podrá tener estimaciones disponibles en zonas de interés y se podrá proveer de información más oportuna para la toma de decisiones y el aprovechamiento de los estos recursos naturales.</p>
Información y fuentes:	<ul style="list-style-type: none"> • Imágenes de percepción remota: <ul style="list-style-type: none"> • Geoeye • WorldView • Ortofotografía • Landsat • Cubo de Datos Geoespaciales de México. • Cartografía Física. • Cartografía Topográfica.
Hipótesis del trabajo:	El cambio climático está acelerando el proceso de desertificación en México.
Objetivo general:	Construir un modelo predictivo de desertificación en México vinculado con el proceso de cambio climático y basado en imágenes satelitales, registros de muestras de suelos y técnicas de aprendizaje computacional.

Objetivos específicos:	<ol style="list-style-type: none"> 1. Incorporar nuevos elementos de información. 2. Re-entrenar modelos de redes neuronales de aprendizaje profundo para identificar grado de desertificación a nivel de celdas de 600x600 mts. 3. Aplicar métodos de aprendizaje no supervisado para descubrir patrones de comportamiento en los conjuntos de series temporales. 4. Revisar el modelo de evolución de la desertificación. 5. Realizar predicciones actualizadas del grado de desertificación de la tierra para los siguientes 10 años.
Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Cubo de Datos Geoespacial con modelo predictivo integrado. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Geografía y Medio Ambiente (DGGMA). • Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT).
Responsable de la ejecución del proyecto:	Elio Atenógenes Villaseñor García.
Observaciones:	
Subsistema:	Geografía y medio ambiente, ordenamiento territorial y urbano
Temática:	Actualización cartográfica
Técnicas:	Ciencia de datos
Fuente de datos:	No tradicional
Estatus:	NUEVO

Nombre del Proyecto:	3. Detección de crecimientos urbanos y nuevas localidades con imágenes satelitales desde la ciencia de datos. (CGI)
Descripción general del problema a resolver:	La detección de crecimientos urbanos y nuevas localidades es una tarea complicada que requiere trabajo de campo y por lo tanto la fecha de actualización depende de los distintos operativos.
Importancia y utilidad del tema para el INEGI:	Implementación de técnicas de Aprendizaje Computacional para el análisis de Big Data que permita identificar crecimientos urbanos y nuevas localidades.
Información y fuentes:	<ul style="list-style-type: none"> • Imágenes contenidas en el Cubo de Datos Geoespaciales de México, las cuales cubren el periodo de 1984 hasta el año 2020. • Cartografía Censal.
Hipótesis del trabajo:	El uso de modelos de aprendizaje supervisado aplicados a imágenes de satélite permitirá encontrar vías alternativas y más eficientes que las actuales para identificar las áreas urbanas del país y su cambio a lo largo del tiempo.
Objetivo general:	Identificar crecimientos urbanos y aparición de nuevas localidades mediante técnicas de Aprendizaje Computacional utilizando imágenes satelitales.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Generar un algoritmo eficaz y eficiente de clasificación de imágenes de satélite para la detección de crecimientos urbanos y nuevas localidades. 2. Aplicar un algoritmo eficaz y eficiente de clasificación de imágenes de satélite para la detección de crecimientos urbanos y nuevas localidades.
Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Algoritmo de clasificación de imágenes. • Reporte de ejercicio de aplicación del algoritmo. • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas (DGEE). • Dirección General de Geografía y Medio Ambiente (DGGMA). • Investigadores y tomadores de decisiones interesados en estudiar la evolución del crecimiento de las zonas urbanas y rurales en México. • Coordinación General de Informática (CGI).

Responsable de la ejecución del proyecto:	Elio Atenógenes Villaseñor García.
Observaciones:	Laboratorio de Ciencia de Datos -Elio Atenógenes Villaseñor García en colaboración con la Coordinación General de Informática (CGI).
Subsistema:	Geografía, medio ambiente, ordenamiento territorial y urbano
Temática:	Actualización cartográfica
Técnicas:	Ciencia de datos
Fuente de datos:	No tradicional
Estatus:	NUEVO

Nombre del Proyecto:	4. Identificación de panteones mediante imágenes satelitales. (DGEESPJ)
Descripción general del problema a resolver:	A pesar de la vasta información con la que cuenta el INEGI sobre los panteones del país, por las características propias de los programas estadísticos, existen diversas limitaciones que impiden integrar en una sola base todos los datos que se tienen al respecto. Entre los principales retos se encuentran la falta de claves estandarizadas para estos establecimientos, las diferencias que existen entre los nombres registrados y la ubicación geográfica de cada uno de ellos.
Importancia y utilidad del tema para el INEGI:	<p>Conforme al artículo 115 de la Constitución Política de los Estados Unidos Mexicanos, los panteones son considerados como un servicio a cargo de las administraciones públicas de los municipios. Para dar cuenta de este servicio, el INEGI cuenta con cuatro fuentes de información estadística y geográfica: el Censo Nacional de Gobiernos Municipales y Demarcaciones Territoriales de la Ciudad de México (CNGMD), el Directorio Estadístico Nacional de Unidades Económicas (DENUE), las cartas topográficas 1:20 y 1:50, y más recientemente el proyecto especial denominado Recopilación de información de los Cementerios Públicos en las Zonas Metropolitanas del País (RICPZMP).</p> <p>A partir de estos programas estadísticos se puede entender la distribución y oferta de los panteones, las personas laborando en ellos, así como la caracterización del tipo de prestación de este servicio. En el caso de los panteones ubicados en los municipios que comprenden las zonas metropolitanas se cuenta con datos sobre la administración de los restos humanos, la disponibilidad de fosas comunes, los tipos de registro con los que cuentan, así como el grado de ocupación de la capacidad instalada.</p> <p>Los panteones son relevantes no solo por constituir un servicio público, sino porque en el contexto actual de seguridad, estos establecimientos constituyen un elemento fundamental en la agenda del Programa Nacional de Exhumaciones, para los servicios forenses del país, y en las políticas del Sistema Nacional de Búsqueda de Personas.</p>
Información y fuentes:	Con base en la información descrita en el apartado anterior, se sabe la ubicación de un gran conjunto de panteones. Estas ubicaciones serán referenciadas y caracterizadas con los datos disponibles en la plataforma de Google Earth Engine, específicamente las fuentes de imágenes satelitales LandSat8 y Sentinel.
Hipótesis del trabajo:	La actualización de la información sobre espacios con características específicas, como los panteones en México, se puede realizar de manera más rápida y eficiente que por las vías tradicionales mediante la incorporación de información de imágenes satelitales y el uso de técnicas de ciencia de datos.
Objetivo general:	Desarrollar una metodología para la identificación espacial de objetos temáticos diversos aplicada al caso especial de la identificación y ubicación de los panteones mediante el análisis de imágenes satelitales.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Desarrollar una metodología para la identificación espacial de objetos temáticos diversos. 2. Identificar los panteones mediante imágenes satelitales. 3. Relacionar estas ubicaciones con los datos correspondientes al programa estadístico de RICPZMP. 4. Generar claves estadísticas de los panteones del país a fin de construir un listado estructurado.

Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación que incluya: <ul style="list-style-type: none"> ○ Una capa de información geográfica con la unión de los datos que sean compatibles entre los programas estadísticos; ○ Un listado de panteones con claves geoestadísticas estructuradas. • Artículo con calidad suficiente para ser publicado en “Realidad, datos y espacio. Revista Internacional de Estadística y Geografía”, que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia (DGEESPJ).
Responsable de la ejecución del proyecto:	Elio Atenógenes Villaseñor García.
Observaciones:	Laboratorio de Ciencia de Datos -Elio Atenógenes Villaseñor García en colaboración con la Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia.
Subsistema:	Gobierno, Seguridad Pública e Impartición de Justicia
Temática:	Actualización cartográfica
Técnicas:	Ciencia de datos
Fuente de datos:	No tradicional
Estatus:	NUEVO

Coyuntura económica

Nombre del Proyecto:	5. Nowcasting para componentes mensuales de la demanda.
Descripción general del problema a resolver:	Como continuación del proyecto "Nowcasting: monitoreo en tiempo real de la actividad económica mexicana" que está enfocado a generar nowcasts del Indicador Global de la Actividad Económica y transversalmente, para las actividades secundarias y terciarias; con el fin de generar una mayor cantidad de estimaciones adelantadas y precisas que sean de utilidad para los tomadores de decisión, surge la necesidad de generar nowcasts para algunos componentes de la demanda como el Indicador mensual de consumo privado (consumo), Inversión fija bruta (inversión), exportaciones totales (exportaciones) e importaciones totales (importaciones).
Importancia y utilidad del tema para el INEGI:	Al ampliar la oferta de indicadores oportunos, el INEGI brinda información mensual, desagregada, precisa y anticipada, de algunos componentes de la demanda del Producto Interno Bruto, los cuales son útiles para la toma de decisiones informada.
Información y fuentes:	<ul style="list-style-type: none"> • Series de tiempo trimestrales y mensuales del Banco de Información Económica del INEGI. • Información económica y financiera reportada por el Banco de México • Fuentes de información no tradicionales como Google Trends, Twitter, noticias, etc. • Información relevante de otros países, la cual puede ser extraída de diversas fuentes como la OCDE, Banco Mundial, etc.
Hipótesis del trabajo:	La representación del consumo, inversión, exportaciones e importaciones mediante Modelo de Factores Dinámicos requiere en cada caso de especificaciones potencialmente diferentes de la utilizada para el Indicador Global de la Actividad Económica.
Objetivo general:	Generar nowcasts para el consumo, inversión, exportaciones e importaciones a través de metodologías estadísticas y econométricas válidas teórica y empíricamente.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Utilizar información de alta frecuencia, tradicional y no tradicional para generar los nowcasts. 2. Diferentes variables según la serie de tiempo objetivo 3. Estimaciones oportunas con hasta 35 días de anticipación respecto a las cifras oficiales. 4. Revisión de supuestos de los modelos de nowcasting y estimación de la incertidumbre asociada a la estimación. 5. Interpretación estructural de los resultados.
Tiempo para su terminación:	24 meses.

Productos esperados:	<ul style="list-style-type: none"> • Modelo de Factores Dinámicos específico para realizar nowcasting de consumo, inversión, exportaciones e importaciones programado en R. • Documentación técnica del trabajo realizado. • Artículo con calidad suficiente para ser publicado en “Realidad, datos y espacio. Revista Internacional de Estadística y Geografía”, que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas. • Banco de México. • Secretaría de Hacienda y Crédito Público. • Secretaría de Economía.
Responsable de la ejecución del proyecto:	Francisco de Jesús Corona Villavicencio.
Observaciones:	
Subsistema:	Económica
Temática:	Coyuntura económica
Técnicas:	Análisis de series de tiempo
Fuente de datos:	Mixto
Estatus:	NUEVO

Nombre del Proyecto:	6. Indicadores compuestos de alta frecuencia.
Descripción general del problema a resolver:	<p>Actualmente se cuenta con una gran cantidad de variables de alta frecuencia, tanto tradicional como no tradicional que da información en tiempo real de lo que acontece en diferentes temáticas, como económicas, financieras, sociales, de salud, etc., las cuales pueden o no, estar correlacionadas entre sí.</p> <p>Dicha información es proporcionada por diferentes fuentes y organismos, y actualmente, no existe ningún indicador o grupo de indicadores que resuma el comportamiento de las variables.</p> <p>En este proyecto se pretende dar solución a esta problemática y ofrecer al público en general, indicadores compuestos de alta frecuencia para temáticas económicas, financiera, sociales, de salud, etc., según sus características comunes que guarden dichas series de tiempo.</p>
Importancia y utilidad del tema para el INEGI:	Los tiempos de incertidumbre actuales hace clara la necesidad de contar con indicadores de alta frecuencia (por ejemplo, semanal) publicados oportunamente, los cuales concentren información de diferentes fuentes tradicionales y no tradicionales.
Información y fuentes:	<ul style="list-style-type: none"> • Sistemas de pagos del Banco de México. • tipo de cambio. • tasas de interés. • indicadores bursátiles. • indicadores de difusión. • redes sociales. • buscadores online. • indicadores económicos de coyuntura, entre otros.
Hipótesis del trabajo:	La creciente disponibilidad de datos de alta frecuencia permite generar indicadores de alta frecuencia y gran oportunidad que resulten de utilidad para el monitoreo de fenómenos económicos y sociales.
Objetivo general:	Generar indicadores compuestos de alta frecuencia para diversas temáticas, principalmente económica, financiera y de salud.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Indicadores con frecuencia semanal. 2. Los indicadores pueden compartir variables entre bloques de temáticas. 3. Validación de los supuestos que dan cabida a los indicadores. 4. Probar el carácter estructural de los indicadores en modelos de Nowcasting.
Tiempo para su terminación:	24 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Propuesta de indicadores de alta frecuencia a publicarse como estadística experimental.

	<ul style="list-style-type: none"> • Artículo con calidad suficiente para ser publicado en RDE, que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas. • Dirección General de Estadísticas Sociodemográficas. • Secretaría de Hacienda y Crédito Público. • Banco de México. • Secretaría de Economía.
Responsable de la ejecución del proyecto:	Francisco de Jesús Corona Villavicencio.
Observaciones:	
Subsistema:	Económica
Temática:	Coyuntura económica
Técnicas:	Análisis de series de tiempo
Fuente de datos:	Mixto
Estatus:	NUEVO

Nombre del Proyecto:	7. Indicadores de la actividad económica generados a partir de información bancaria.
Descripción general del problema a resolver:	Aprovechar, para la generación de estadísticas económicas y sociales, la información sobre transacciones que se registran a diario en terminales punto de venta (TPV), cajeros automáticos, comercio electrónico y dispersiones de nómina.
Importancia y utilidad del tema para el INEGI:	Explotar fuentes no tradicionales de información para generar estadísticas económicas y sociales complementarias. El trabajo supone el logro de asociaciones estratégicas con instituciones privadas y permitiría contar con referentes de la actividad económica de bajo costo, alta frecuencia y elevada granularidad.
Información y fuentes:	<ul style="list-style-type: none"> • Transacciones registradas por BBVA. • Transacciones registradas por Banorte. • Transacciones registradas por el Banco de México. • Transacciones registradas por banco Santander.
Hipótesis del trabajo:	La información generada por las instituciones bancarias es susceptible de convertirse en indicadores de estadísticas económicas y sociales que abonen a un seguimiento pormenorizado de aspectos relevantes de la actividad económica en el país, sus entidades federativas y municipios.
Objetivo general:	Emplear información bancaria para la generación de indicadores económicos y sociales con desagregaciones sectoriales y geográficas, que ofrezca mucho mayor frecuencia y oportunidad que los indicadores económicos actualmente generados por el INEGI.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Coordinar la generación y firma de convenios de colaboración. 2. Proponer indicadores y fichas técnicas para la homologación de la información que sea compartida con el INEGI. 3. Coordinar y revisar la entrega de información al INEGI. 4. Proponer y coordinar el diseño de una sección dentro de la página del INEGI para la publicación de la información. 5. Administrar y coordinar la publicación regular de la información.
Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Sitio de internet para publicación regular de la información • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).

Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas. • Dirección General de Estadísticas Sociodemográficas.
Responsable de la ejecución del proyecto:	José Alejandro Ruiz Sánchez.
Observaciones:	<p>Este proyecto sustituye y amplía el proyecto <<1.- Indicador adelantado de la Actividad Económica INEGI-BBVA>>, el cual se enfocaría solo en información de BBVA. Durante el primer semestre de 2020 se mantuvieron pláticas entre la DGIAI del INEGI y las instituciones bancarias Santander y Banorte, con las cuales se logró la firma de tres acuerdos en total. Los acuerdos establecen la transferencia de información hacia el INEGI para su publicación regular.</p> <p>A la fecha se tienen los tres convenios firmados; y un convenio modificatorio con BBVA para que pueda comenzar el intercambio de información.</p> <p>Comenzaremos pláticas con BANXICO para la firma del convenio con ellos.</p>
Subsistema:	Económica
Temática:	Coyuntura económica
Técnicas:	Visualización
Fuente de datos:	No tradicional
Estatus:	NUEVO

Estructura económica

Nombre del Proyecto:	8. Indicadores de actividad económica municipal incluyendo mediciones satelitales de luminosidad nocturna.
Descripción general del problema a resolver:	<p>Como continuación del proyecto de "Indicadores de actividad económica municipal II" donde se generan indicadores de actividad económica municipal basados en los Censos Económicos de 1993-2013 para todos los municipios de México, se pretende complementarlos y robustecerlos a través del uso de las mediciones satelitales de luminosidad nocturna que proporciona, por ejemplo, el departamento meteorológico satelital de la fuerza aérea de Estados Unidos.</p> <p>Con este proyecto se podrá contar con indicadores de actividad económica municipal más fiables a partir de 1993, los cuales puedan ser de utilidad para el público en general, por ejemplo, para la formulación de políticas públicas.</p>
Importancia y utilidad del tema para el INEGI:	Se amplía la oferta de información de coyuntura brindando al público en general, información económica municipal válida económica y estadísticamente usando información representativa geográficamente de los Censos Económicos e información de luminosidad nocturna, la cual sea útil a los usuarios para la toma de decisiones.
Información y fuentes:	<p>Además de la información de los Censos Económicos se explora información cartográfica y de imágenes satelitales:</p> <ul style="list-style-type: none"> • Estadística de luminosidad (HSW database). • Imágenes de visión nocturna (Earth Observation Group).
Hipótesis del trabajo:	La información cartográfica y de imágenes satelitales está correlacionada con los indicadores de actividad económica municipal extraídos mediante el uso de información de los Censos Económicos.
Objetivo general:	Generar indicadores de actividad económica municipal con periodicidad anual haciendo uso de información regionalmente representativa proporcionada por los micro datos de los Censos Económicos de 1994 a 2019 y también, de las mediciones satelitales de luminosidad nocturna. Lo anterior se realizará aprovechando la correlación existente entre actividad económica y luminosidad. Esto permitirá a su vez, actualizar los indicadores de actividad económica municipal en tiempo real, proporcionado a los usuarios indicadores económicos regionalmente desagregados y con una mayor oportunidad.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Indicadores de actividad económica municipal para grupos de sectores económicos con periodicidad anual a partir de 1993. 2. Validación de supuestos económicos y estadísticos como lo proponen Pérez y Corona (2020) y Guerrero y Mendoza (2019). 3. Publicar los indicadores como Estadística Experimental.
Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse

	en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas. • Secretaría de Economía. • Secretaría de Hacienda y Crédito Público. • Banco de México. • Gobiernos estatales y municipales.
Responsable de la ejecución del proyecto:	Francisco de Jesús Corona Villavicencio.
Observaciones:	
Subsistema:	Económica
Temática:	Estructura económica
Técnicas:	Análisis de series de tiempo
Fuente de datos:	Mixto
Estatus:	NUEVO

Nombre del Proyecto:	9. (7-2020) Aprovechamiento de los Censos Económicos y las Encuestas Especiales en Establecimientos para Contextualizar el Comportamiento de la Productividad en México.
Descripción general del problema a resolver:	Estudios recientes para México indican la relación entre el bajo crecimiento económico y la ineficiente asignación de recursos productivos. Entre estos estudios está el realizado por el área de investigación del INEGI sobre mediciones de la pérdida de productividad bajo la metodología desarrollada por Chang-Tai Hsieh y Peter Klenow. Sin embargo, poco se ha analizado sobre el contexto económico y social en el que se da la falta de productividad.
Importancia y utilidad del tema para el INEGI:	Al ligar registros contenidos en distintas fuentes de información el valor agregado de la información producida por el INEGI se incrementa considerablemente, y con ello su potencial analítico. Las encuestas especiales que realiza el INEGI contienen información que no es recopilada en los Censos Económicos y que sin embargo puede ayudar a explicar el comportamiento de temas contenidos en éstos. Este tipo de estudios contribuye a la generación de información de utilidad para la formulación y evaluación de políticas económicas que busquen aumentar la productividad a nivel nacional y regional, a través de un conocimiento más detallado de los fenómenos económicos.
Información y fuentes:	<ul style="list-style-type: none"> • "Esfuerzos mal recompensados. La elusiva búsqueda de la prosperidad en México", Santiago Levy. Banco Interamericano de Desarrollo, 2018. • https://publications.iadb.org/bitstream/handle/11319/8971/Esfuerzos-mal-recompensados-La-elusiva-busqueda-de-la-prosperidad-en-Mexico.pdf • "Misallocation and manufacturing TFP in China and India", Chang-Tai Hsieh & Peter J. Klenow. Quarterly Journal of Economics, Vol. CXXIV, Noviembre 2009. • http://klenow.com/MMTFP.pdf • Censos Económicos 2009, Censos Económicos 2014. • Encuestas Especiales: <ul style="list-style-type: none"> • Encuesta Nacional de Financiamiento de las Empresas. • Encuesta Nacional de Victimización de Empresas. • Encuesta Nacional de Calidad Regulatoria e Impacto Gubernamental en Empresas (ENCRIGE). • Encuesta Nacional sobre Productividad y Competitividad de las Micro, Pequeñas y Medianas Empresas (ENAPROCE).
Hipótesis del trabajo:	Las diferencias en la productividad de las unidades económicas al interior de una misma actividad económica están asociadas con temas de acceso al financiamiento (+), riesgo de ser objeto de acciones criminales (-) y la calidad del marco regulatorio (+), entre otros aspectos.
Objetivo general:	Combinar información de censos económicos y encuestas especiales en establecimientos para dilucidar algunas de las principales variables que explican o se correlacionan con las diferencias de productividad entre unidades económicas al interior de diferentes industrias.

Objetivos específicos:	<ol style="list-style-type: none"> 1. Ligar registros de encuestas especiales con registros de los Censos Económicos. 2. Formular e implementar modelos estadísticos que permitan identificar variables relacionadas con el comportamiento de la productividad y actividad económica al interior de diferentes industrias del país.
Tiempo para su terminación:	7 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Secretaría de Economía. • Dirección General de Estadísticas Económicas.
Responsable de la ejecución del proyecto:	José Alejandro Ruiz Sánchez.
Observaciones:	Este proyecto parte de los resultados del que formaba parte del PAINEGI 2018: "Productividad y Eficiencia en la asignación de los factores productivos en México a partir de los Censos Económicos 2009 y los Censos Económicos 2014".
Subsistema:	Económica
Temática:	Estructura económica
Técnicas:	Econometría
Fuente de datos:	Tradicional
Estatus:	CONTINÚA

Nombre del Proyecto	10. (10-2020) Modelos de Equilibrio General Computable para la Economía Mexicana.
Descripción general del problema a resolver:	Construir modelos suficientemente flexibles que permitan integrar amplios conjuntos de información para investigar asuntos económicos a nivel nacional y sectorial, enfatizando su interdependencia.
Importancia y utilidad del tema para el INEGI:	Una utilización intensiva de la base de información producida por el Instituto a nivel sectorial, es decir, Matrices de Insumo Producto, Cuadros de Oferta y Utilización, Matrices de Contabilidad Social, entre otros instrumentos estadísticos.
Información y fuentes:	<ul style="list-style-type: none"> • Matrices de Contabilidad Nacional. • Cuadros de Oferta y Utilización. • Matrices de Insumo-Producto. • Matrices de Contabilidad Social. • Encuesta Nacional de Ocupación y Empleo (ENOE). • Información adicional producida por el INEGI. • Bases de datos internacionales.
Hipótesis del trabajo:	El modelo de Equilibrio General Computable resultante permitirá simular de manera muy acertada el impacto de choques exógenos y de políticas públicas que incidan sobre variables clave de economía mexicana.
Objetivo general:	Desarrollar tres modelos de equilibrio general de la economía mexicana que responden a distintos enfoques de análisis y simulación de políticas.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Especificación de un modelo de Equilibrio General como un problema de optimización, cuyo énfasis es encontrar balances intersectoriales que lleven al máximo la producción y un mejor uso de los recursos primarios. 2. Especificación de un Modelo de Equilibrio General como el ensamble de sistemas de oferta y demanda con fundamentos microeconómicos, cuyo enfoque es conocido como <i>System Wide Approach to Microeconomics</i>. Esta especificación permite la solución de modelos de grandes dimensiones con detalles microeconómicos de los agentes de la economía. 3. Un modelo especificado con un enfoque macroeconómico de corte keynesiano, que permite la evaluación de macro-políticas a nivel sectorial y de equilibrio entre mercados de bienes, factores, hogares y sector externo. Actualmente son los modelos más empleados. A esta familia de modelos se les conoce como Modelos de Equilibrio General Computable.
Tiempo para su terminación:	5 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el

	artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas. • Banco de México. • Secretaría de Hacienda y Crédito Público. • Secretaría de Economía.
Responsable de la ejecución del proyecto:	José Valentín Solís y Arias.
Observaciones:	
Subsistema:	Económica
Temática:	Estructura económica
Técnicas:	Equilibrio General
Fuente de datos:	Tradicional
Estatus:	CONTINÚA

Nombre del Proyecto:	11. (11-2020) Patrones de consumo de los hogares mexicanos.
Descripción general del problema a resolver:	Investigar y ponderar los patrones de consumo de las familias mexicanas y aportar parámetros para su evaluación frente a variaciones de los ingresos de esas familias. Por otra parte, se investigarán las conductas de los hogares como respuesta a variaciones en los precios de los bienes que consumen.
Importancia y utilidad del tema para el INEGI:	Hay numerosos estudios en el país, sobre la distribución del ingreso entre familias; sin embargo, hay poco análisis de la demanda de bienes de consumo, a pesar de que se tiene una amplia información del consumo de los hogares a través de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). Esta investigación es una utilización extensiva de la información del gasto de los hogares.
Información y fuentes:	Encuestas Nacionales de Ingreso y Gasto de los Hogares, para los años 2016 y 2018. Índices de precios implícitos, Cuentas por Sectores Institucionales y las Matrices de Contabilidad Social.
Hipótesis del trabajo:	La conducta de los consumidores no necesariamente se refleja en agregaciones por deciles de ingresos. Los datos de los cuadros de Gasto de los hogares pueden homogeneizar la conducta de esos hogares de acuerdo con sus patrones de consumo. Para su evaluación debe estimarse un Sistema Completo de Ecuaciones simultáneas de Demanda.
Objetivo general:	Estimar, para los años 2016 y 2018, Sistemas de Ecuaciones de Demanda que cumplan con las restricciones impuestas por las modernas teorías del consumo.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Medir y comparar las conductas del consumo de los hogares de acuerdo con sus niveles de ingreso; precisión de bienes sustitutos y complementarios en la canasta de consumo de los hogares mexicanos.
Tiempo para su terminación:	6 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas. • Dirección General de Estadísticas Sociodemográficas. • Secretaría de Hacienda y Crédito Público. • Secretaría de Economía. • Secretaría del Trabajo y Previsión Social. • Universidades y Centros de Investigación.

Responsable de la ejecución del proyecto:	José Valentín Solís y Arias.
Observaciones:	
Subsistema:	Económica
Temática:	Estructura económica
Técnicas:	Equilibrio General
Fuente de datos:	Tradicional
Estatus:	CONTINÚA

Nombre del Proyecto:	12. (9-2020) Construcción de una Matriz de Insumo-Producto Multi-regional de México año 2013.
Descripción general del problema a resolver:	Por las diferencias económicas, políticas y sociales del país, es pertinente tener una Matriz de Insumo-Producto Multi-regional para estudiar los impactos económicos Inter e Intra-Regiones, y así contar con herramientas que ayuden a focalizar los impactos de la actividad económica en contextos geográficos.
Importancia y utilidad del tema para el INEGI:	La política industrial regional será uno de los principales temas del debate público en los próximos años. Actualmente, el Instituto no cuenta con la Matriz de Insumo-Producto necesaria para atender esta demanda de información.
Información y fuentes:	<ul style="list-style-type: none"> • Matriz de Insumo-Producto año 2013. • Sistema de Cuentas Nacionales de México. • Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2012. • Información adicional producida por el INEGI. • Regionalización de la Red "Crecimiento Económico y Economía Espacial" Constituida con la colaboración de la Universidad Autónoma de Nuevo León, El Colegio de México, Instituto Tecnológico y de Estudios Superiores de Monterrey, Universidad Autónoma de Coahuila.
Hipótesis del trabajo:	El grado de integración económica entre las entidades del centro del país es mayor que el de las del norte del país. Asimismo, el grado de integración de las entidades del sur y sureste es inferior al correspondiente del resto del país.
Objetivo general:	Generar dos Matrices de Insumo-Producto Multi-regional de México para el año 2013.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Construcción de Matrices de Insumo-Productos para las 32 entidades federativas para el año 2013. 2. Integración de las 32 matrices en una matriz Multi-regional.
Tiempo para su terminación:	4 meses.
Productos esperados:	<ul style="list-style-type: none"> • Matriz de Insumo-Producto Multi-regional. • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).

Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Económicas. • Banco de México. • Secretaría de Hacienda y Crédito Público. • Secretaría de Economía.
Responsable de la ejecución del proyecto:	José Valentín Solís y Arias.
Observaciones:	
Subsistema:	Económica
Temática:	Estructura económica
Técnicas:	Equilibrio General
Fuente de datos:	Tradicional
Estatus:	CONTINÚA

Nombre del Proyecto:	13. Clasificación automática de unidades económicas de acuerdo con el apoyo que brindan a las mujeres. (Vicepresidencia a cargo del Subsistema Nacional de Información Geográfica, Medio Ambiente, Ordenamiento Territorial y Urbano).
Descripción general del problema a resolver:	<p>Las mujeres dedican una parte importante de su tiempo al cuidado de otras personas, como parte de un trabajo generalmente no remunerado, no cuantificado y no reconocido. Sin embargo, una mayor participación femenina en la fuerza laboral remunerada beneficia sus propias condiciones de vida, las condiciones de vida de su familia y de toda la comunidad, permitiendo la reducción de brechas y un desarrollo económico más rápido.</p> <p>Ante esta realidad, este proyecto se plantea la siguiente pregunta ¿La proximidad de diferentes servicios facilita la integración de la mujer en el mercado laboral? Para responder a esta pregunta a nivel nacional, el primer desafío es hacer una clasificación de los servicios según el apoyo que brindan a las mujeres, para luego analizar la proximidad mediante la red vial, de este grupo de población a los diferentes tipos de servicios.</p>
Importancia y utilidad del tema para el INEGI:	Implementación de técnicas de Ciencia de Datos para el manejo y análisis de datos de texto no estructurado provenientes del Sistema de Clasificación Industrial de América del Norte (SCIAN) y del Directorio Estadístico Nacional de Unidades Económicas (DENUE) que permitan la clasificación de unidades económicas en categorías relacionadas con el apoyo que brindan a las mujeres.
Información y fuentes:	<ul style="list-style-type: none"> • Datos del DENUE y del SCIAN. • Bases de Datos etiquetadas manualmente por parte de expertos.
Hipótesis del trabajo:	El uso de técnicas de ciencia de datos y aprovechamiento de lenguaje natural permite la clasificación correcta, rápida y de bajo costo de unidades económicas con características particulares no necesariamente reflejadas en la clasificación SCIAN a seis dígitos pero recuperables mediante la explotación de elementos del Directorio Estadístico Nacional de Unidades Económicas.
Objetivo general:	Diseñar una metodología de ciencia de datos para clasificar automáticamente unidades económicas de acuerdo con el apoyo que brindan a las mujeres y grupos dependientes que ellas cuidan (infantes y adultos mayores). Utilizando información de texto proveniente de fuentes oficiales, como el Directorio Estadístico Nacional de Unidades Económicas (DENUE).
Objetivos específicos:	<ol style="list-style-type: none"> 1. Desarrollo de metodología de Ciencia de datos. 2. Construcción de conjunto de entrenamiento. 3. Implementación y aplicación de algoritmo de aprendizaje computacional. 4. Generación de la Clasificación global del DENUE. 5. Evaluación de los resultados obtenidos por el algoritmo. 6. Construcción del modelo para la caracterización del acceso a servicios con enfoque de género y su correlación con el acceso a mercado laboral. 7. Validación del modelo.

Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Listado de unidades económicas clasificadas de acuerdo con el servicio prestado. • Mapas nacionales que identifiquen la densidad de unidades económicas con productos y servicios orientados a solventar las necesidades de la mujer. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/). • Documento con reporte de investigación.
Destinatarios:	<p>Los beneficiarios potenciales serían, entre otros:</p> <ul style="list-style-type: none"> • Secretaría de Desarrollo Agrario, Territorial y Urbano. • Estados de la Federación. • Instituciones públicas y privadas que prestan un servicio de salud y cuidado, de docencia a nivel básico, así como instituciones de apoyo a las mujeres víctimas de violencia o de desarrollo económico. • Gobiernos locales. • Institutos nacionales de estadística y geografía de otros países.
Responsable de la ejecución del proyecto:	Elio Atenógenes Villaseñor García.
Observaciones:	Laboratorio de Ciencia de Datos -Elio Atenógenes Villaseñor García en colaboración con personal de la Vicepresidencia a cargo del Subsistema Nacional de Información Geográfica, Medio Ambiente, Ordenamiento Territorial y Urbano.
Subsistema:	Económica
Temática:	Estructura económica
Técnicas:	Ciencia de datos
Fuente de datos:	Tradicional
Estatus:	NUEVO

Ingresos

Nombre del Proyecto:	14. Valor imputado al uso de la vivienda por sus propietarios.
Descripción general del problema a resolver:	La principal fuente de ingreso no monetario de los hogares que reporta la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) corresponde a la renta imputada de las viviendas ocupadas por sus dueños. Actualmente la ENIGH se vale de la apreciación subjetiva de quien ocupa la vivienda asignarle el valor correspondiente. Es decir, al ocupante de la vivienda se le pregunta cuánto estaría dispuesto a pagar de renta por esa vivienda y ese es el valor que se utiliza para hacer la imputación. Sin embargo, esta no es la única ni necesariamente la mejor manera en que se puede hacer tal imputación. La literatura sobre el tema y las prácticas de los productores de información de otros países e incluso de otras áreas del INEGI, muestran que es posible usar modelos estadísticos que exploten la información sobre las viviendas captada por la propia encuesta en combinación con otros elementos de contexto a efecto de lograr estimaciones más objetivas. Los modelos hedónicos parecen especialmente promisorios. En estos modelos se parte de las viviendas para las que sí se paga renta y se estiman los parámetros que permiten el mejor ajuste entre la renta pagada y las características de las viviendas, las que usualmente se toman de la información proveída por la misma encuesta. Ejercicios preliminares desarrollados en la Dirección General Adjunta de Investigación (DGAJ) sugieren que el abanico de calidades de las viviendas propias es más amplio que el de las rentadas y que la estimación del coto de oportunidad de las viviendas ocupadas por sus dueños estaría condicionada por elementos del entorno que no capta directamente la encuesta, cuya incorporación en el modelo podría ayudar a mejorar la calidad de las imputaciones. En este ejercicio se experimentará con un conjunto de variables contenidas en la ENIGH y en otras fuentes para encontrar las más adecuadas que describan tanto a las viviendas rentadas como las no rentadas con el fin de lograr mejores imputaciones.
Importancia y utilidad del tema para el INEGI:	Ofrecer una alternativa al INEGI para reportar medidas más fidedignas de una de las principales fuentes de ingresos de los hogares, especialmente por la importancia de que los criterios de imputación sean objetivos.
Información y fuentes:	Variables diversas de la ENIGH, cartografía urbana, imágenes satelitales.
Hipótesis del trabajo:	El uso de modelos hedónicos permitirá contar con estimaciones de la renta imputada de las viviendas ocupadas por sus propietarios mejores que las que se producen y reportan actualmente en la ENIGH.
Objetivo general:	Desarrollar modelos estadísticos de la renta imputada de las viviendas ocupadas por sus dueños de manera que se tenga una alternativa a las estimaciones subjetivas que se utilizan actualmente en la ENIGH.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Revisar la literatura y las prácticas internacionales sobre imputación de los valores de las rentas de las viviendas ocupadas por sus dueños en las encuestas de ingresos de los hogares similares a la ENIGH. 2. Seleccionar la especificación de modelo hedónico más adecuada para la ENIGH, considerando variables internas y externas a la ENIGH. 3. Reportar las diferencias con respecto de la manera como actualmente se hacen las imputaciones de rentas en la ENIGH. 4. Hacer las recomendaciones pertinentes.
Tiempo para su terminación:	12 meses.

Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Producto de estadística experimental con estimaciones alternas del valor imputado de la vivienda propia. • Artículo con calidad suficiente para ser publicado en “Realidad, datos y espacio. Revista Internacional de Estadística y Geografía”, que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistacytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Productores y usuarios de la información de la ENIGH.
Responsable de la ejecución del proyecto:	Benito Durán Romo.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Ingresos
Técnicas:	Econometría
Fuente de datos:	Tradicional
Estatus:	NUEVO

Nombre del Proyecto:	15. Construcción del Modelo estadístico 2020 para la continuidad del MCS-ENIGH.
Descripción general del problema a resolver:	Generar desde el Modelo estadístico para la continuidad del MCS-ENIGH (MEC) los microdatos necesarios para el cálculo de la pobreza en 2020 del Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL).
Importancia y utilidad del tema para el INEGI:	El INEGI acordó con el CONEVAL la entrega de información que le permita a este último la medición consistente de la pobreza.
Información y fuentes:	<ul style="list-style-type: none"> • Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). • https://rde.inegi.org.mx/index.php/2018/08/27/modelo-estadistico-2016-la-continuidad-del-mcs-enigh/.
Hipótesis del trabajo:	Aplicando la metodología propuesta de ajuste de ingresos, se podrán generar datos para 2020 consistentes con los MCS-ENIGH del 2008 al 2014, MEC-2016 y MEC-2018.
Objetivo general:	Proveer al CONEVAL de los insumos necesarios para el cálculo de pobreza 2020.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Generar un algoritmo de ajuste en el vector de ingresos. 2. Albergar los resultados en la página web del Instituto.
Tiempo para su terminación:	8 meses.
Productos esperados:	<ul style="list-style-type: none"> • Algoritmo de Ajuste de Ingresos. • Documento metodológico. • Página Web.
Destinatarios:	<ul style="list-style-type: none"> • Consejo Nacional de Evaluación de la Política de Desarrollo Social.
Responsable de la ejecución del proyecto:	José Alejandro Ruiz Sánchez.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Ingresos
Técnicas:	Máximo Verosimilitud
Fuente de datos:	Tradicional
Estatus:	NUEVO

Nombre del Proyecto:	16. Determinación de la subcobertura e imputación del ingreso corregido por MPVR en la ENIGH.
Descripción general del problema a resolver:	La introducción de una nueva metodología para alcanzar una mejor aproximación a la distribución del ingreso de los hogares en México no tuvo entre sus propósitos la imputación de los ingresos subreportados, ni la determinación de la magnitud del ingreso asociado al truncamiento. A pesar de lo anterior ha permitido aproximar a nuestro juicio de una mejor manera la magnitud de la desigualdad de ingresos en nuestro país. Sin embargo, algunos usuarios de la información de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) la explotan a nivel de registro como insumo para las metodologías que les ayudan a cumplir sus fines. En consecuencia, si la información corregida por Máxima Pseudo Verosimilitud Restringida (MPVR) ha de ser aprovechada en sus tareas debe hacerse disponible de manera similar. Esta tarea se dificultaba pues se desconocían la proporción del ingreso total de los hogares efectivamente captado por la encuesta y, en consecuencia, la que está asociada al truncamiento. La propuesta de recurrir a versiones truncadas de los modelos utilizados parece permitir resolver este segundo problema y aportar elementos para la atención del primero.
Importancia y utilidad del tema para el INEGI:	Las encuestas por muestreo en hogares que tratan la variable ingreso cumplen sólo de manera parcial, en el mundo, el propósito de hacer disponible información precisa para la toma de decisiones. La ENIGH no es la excepción. Por ello, diversos investigadores han planteado la necesidad de corregir los valores declarados. Sin embargo, en ausencia de un procedimiento generalmente aceptado y a pesar de sus limitaciones, usuarios importantes de esta información han preferido usarla tal cual se registró en el cuestionario. Este trabajo de investigación permitirá al INEGI, en su caso, poner a disposición de tales usuarios información muestral de la que el efecto de las limitaciones ha sido reducido, y en un formato que satisfaga sus necesidades.
Información y fuentes:	Información referente al ingreso de las personas y de los hogares proveniente de la ENIGH, del Sistema de Cuentas Nacionales de México y del Servicio de Administración Tributaria.
Hipótesis del trabajo:	Al lograr identificar por separado la parte de la brecha entre el ingreso de los hogares reportado por la ENIGH y el reportado por las Cuentas Nacionales, según corresponda a truncamiento y sub-reporte, se hace posible no sólo un acercamiento más realista a la verdadera distribución del ingreso de los hogares mediante MPVR, sino que ello permite también ajustar por sub-reporte los ingresos reportados por la ENIGH para los hogares.
Objetivo general:	Hacer disponible información sobre el ingreso de los hogares a nivel de registro, corregida por el uso simultáneo de otras fuentes relevantes.

Objetivos específicos:	<ol style="list-style-type: none"> 1. Proponer un método de estimación para la magnitud del ingreso obtenido por hogares no incluidos en la muestra. 2. Consecuentemente, estimar el ingreso total de los hogares en la subpoblación representada por la muestra. 3. Aprovechando la información del punto anterior, corregir los ingresos reportados en la encuesta creando un nuevo vector de ingresos de los hogares.
Tiempo para su terminación:	18 meses.
Productos esperados:	<ul style="list-style-type: none"> • Reporte de estadística experimental que incluya: <ul style="list-style-type: none"> ○ Un nuevo vector de ingresos asociado a la muestra de la ENIGH. ○ Una estimación del ingreso total no reportado por la encuesta. • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Encargados de la planeación, el análisis y la difusión de resultados de la encuesta. • Usuarios tradicionales de la información de la Encuesta Nacional de Ingresos y Gastos de los Hogares.
Responsable de la ejecución del proyecto:	Víctor Alfredo Bustos y de la Tijera.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Ingresos
Técnicas:	Máximo Verosimilitud
Fuente de datos:	Mixto
Estatus:	NUEVO

Trabajo

Nombre del Proyecto:	17. Estructura y determinantes de la segregación ocupacional por tareas.
Descripción general del problema a resolver:	El cambio tecnológico está aumentando la demanda por los trabajadores que realizan tareas de naturaleza no rutinaria y reduciendo la demanda por los trabajadores que realizan tareas de naturaleza rutinaria. Esto es especialmente claro en el mercado laboral femenino. Resulta interesante entender qué condiciones socioeconómicas (edad, sexo, nivel educativo, estado civil, entre otras) y productivas (tamaño de la empresa donde se trabaja, sector rural o urbano, empresa pública o privada etc.) se asocian con la segregación ocupacional de los trabajadores, cuando las ocupaciones se definen a través del tipo de tareas que los trabajadores realizan.
Importancia y utilidad del tema para el INEGI:	El INEGI ha desarrollado en el pasado un marco para entender el efecto del avance tecnológico en el mercado laboral nacional. La literatura internacional afirma que el progreso de la tecnología afecta a los trabajadores en función del tipo de tareas que éstos realizan. Este trabajo busca dar continuidad a los esfuerzos para entender cuál es el efecto del avance tecnológico sobre la sociedad en México, y en especial sobre los trabajadores. Asimismo, se buscará generar una clasificación de las ocupaciones y de la población ocupada según el riesgo de desplazamiento a merced de la automatización. Esto posicionará al INEGI como generador de información relevante para la sociedad mexicana.
Información y fuentes:	Encuesta Nacional de Ocupación y Empleo (ENOE).
Hipótesis del trabajo:	Factores como la educación, el género o la región definen el tipo de tareas que los trabajadores realizan, lo que implica que la tecnología afecta de manera diferenciada a los trabajadores en función de estos factores. También, distintas ocupaciones son más o menos susceptibles a ser desplazadas por la automatización, en función del tipo de tareas que los trabajadores realizan.
Objetivo general:	Entender los determinantes de que un trabajador realice tareas de naturaleza no rutinaria, ya sea cognitivas o manuales, y de que realice tareas de naturaleza rutinaria (ya sea cognitivas o manuales). Esto con el objetivo de saber qué tipos de trabajadores son más susceptibles a sufrir los efectos laborales del avance tecnológico.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Conocer el efecto del avance tecnológico en el mercado laboral. 2. Entender los determinantes de la naturaleza de las tareas que realizan los trabajadores en México. 3. Hacer uso de las bases de datos del INEGI para entender mejor el funcionamiento del mercado de trabajo nacional. 4. Proponer un conjunto de indicadores derivado de información del INEGI que reporten la medida en que los trabajos son más o menos rutinarios, sean estas cognitivas o manuales. 5. Con base en lo reportado en el punto anterior, generar una clasificación de la población ocupada según el riesgo de desplazamiento a merced de la automatización.
Tiempo para su terminación:	10 meses.

Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Propuesta de reporte de estadísticas experimentales que considere: <ul style="list-style-type: none"> ○ Conjunto de indicadores derivado de información del INEGI que reporten la medida en que los trabajos son más o menos rutinarios, sean estas cognitivas o manuales. ○ Clasificación de las ocupaciones y de la población ocupada según el riesgo de desplazamiento a merced de la automatización. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Sociodemográficas. • Secretaría del Trabajo y Previsión Social. • Secretaría de Economía. • Instituto Mexicano del Seguro Social.
Responsable de la ejecución del proyecto:	Liliana Meza González.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Trabajo
Técnicas:	Econometría
Fuente de datos:	Tradicional
Estatus:	NUEVO

Nombre del Proyecto:	18. (29-2020) Culminación del modelo de experimentación e implementación de las recomendaciones de las CIET XIX y XX.
Descripción general del problema a resolver:	Se necesita actualizar los conceptos y los indicadores de fuerza de trabajo alineadas a las últimas recomendaciones internacionales de la Organización Internacional del Trabajo (OIT) y de la Organización para la Cooperación y el Desarrollo Económicos (OCDE).
Importancia y utilidad del tema para el INEGI:	Es fundamental retroalimentar a organismos internacionales sobre la viabilidad de los nuevos marcos conceptuales y co-diseñar un prototipo de encuesta de ocupación y empleo que explore formas no detectada de subutilización de fuerza de trabajo, así como nuevas modalidades de inserción en el mercado laboral que desafían a las encuestas y clasificaciones hasta ahora vigentes.
Información y fuentes:	<ul style="list-style-type: none"> • OIT, 19 CIET, Resolución II Resolución Sobre las Estadísticas de Trabajo, la Ocupación y la Subutilización de la Fuerza de Trabajo, (2013). • OIT, 20 CIET, Resolución I, Sobre las Estadísticas de las Relaciones de Trabajo (2018). • Documentos suplementarios emitidos sobre la OIT (<i>Room Documents</i>: contratistas dependientes, medición de desajuste de conocimientos y habilidades respecto a puestos de trabajo; actualización de la medición de la informalidad a la luz de la XIX y XX CIET, medición estadística del trabajo en cooperativas y medición estadística del trabajo forzado). • Informes sobre las pruebas realizadas entre 2017 y 2019.
Hipótesis del trabajo:	Las nuevas recomendaciones internacionales aprobadas por la OIT (Conferencias Internacionales del Trabajo) así como como los que promueve OCDE (Quality of Working Environment & Expert Group on Platform Work) son viables de concretar en un nuevo modelo de encuesta de fuerza de trabajo, sin embargo, ello requiere culminar una serie de pruebas, análisis de las mismas y creación de protocolos de transición, así como encontrar formas innovadoras de aproximarse a los conceptos en la captación no previstas por las nuevas recomendaciones internacionales.
Objetivo general:	Es fundamental retroalimentar a organismos internacionales sobre la viabilidad de los nuevos marcos conceptuales y co-diseñar un prototipo de encuesta de ocupación y empleo que explore formas no detectada de subutilización de fuerza de trabajo, así como nuevas modalidades de inserción en el mercado laboral que desafían a las encuestas y clasificaciones hasta ahora vigentes.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Las nuevas recomendaciones internacionales aprobadas por la OIT (Conferencias Internacionales del Trabajo) así como como los que promueve OCDE (Quality of Working Environment & Expert Group on Platform Work) son viables de concretar en un nuevo modelo de encuesta de fuerza de trabajo, sin embargo ello requiere culminar una serie de pruebas, análisis de las mismas y creación de protocolos de transición, así como encontrar formas innovadoras de aproximarse a los conceptos en la captación no previstas por las nuevas recomendaciones internacionales.

Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Reporte de investigación que incluya: <ul style="list-style-type: none"> ○ Informe y análisis sobre el ciclo de pruebas realizadas. ○ Un nuevo modelo de cuestionario de ocupación y empleo. ○ Documento sobre preguntas y respuestas de la transición de la ENOE a un nuevo modelo de encuesta laboral. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Sociodemográficas. • Organización Internacional del Trabajo. • Organización para la Cooperación y el Desarrollo Económicos.
Responsable de la ejecución del proyecto:	Rodrigo Tadeo Negrete Prieto.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Trabajo
Técnicas:	Pruebas cognitivas y de campo
Fuente de datos:	Tradicional
Estatus:	CONTINÚA

Nombre del Proyecto:	19. Preparación del documento de discusión de la XXI Conferencia Internacional de Estadísticos del Trabajo en materia de medición de la informalidad.
Descripción general del problema a resolver:	En los últimos años la comunidad de países ha venido adquiriendo experiencia en la medición de la informalidad; en paralelo nuevos conceptos del mercado de trabajo han surgido. Se necesita hacer un balance de la experiencia, actualizar conceptos y proponer esquemas de medición con base en las mejores prácticas que retroalimenten tanto a la XXI Conferencia Internacional de Estadísticos del Trabajo (CIET) como a la siguiente revisión del Sistema de Contabilidad Nacional, Conferencias ambas que tendrán lugar en 2023.
Importancia y utilidad del tema para el INEGI:	La medición del empleo y la economía informal distinguen al INEGI en América Latina y entre países emergentes. Evita que toda la discusión se centre en indicadores como el de desocupación o desempleo abierto y plantea una arena clara para el accionar de la política pública. El INEGI ha sido vanguardia en esto y debe seguir siéndolo; al mismo tiempo, ha de buscar que las actualizaciones y o modificaciones que se adopten sean debidamente fundamentadas y en ello tenga un peso específico la experiencia lograda por el Instituto. La temática en concreto incide de manera directa en la estadística laboral y repercute en el Sistema de Cuentas Nacionales de México.
Información y fuentes:	<ul style="list-style-type: none"> • Chen, Martha; Carré, Françoise (ed.) <i>The Informal Economy Revisited: examining the past, envisioning the future</i>. Routledge, London, 2020. • Chen, Martha; Vanek, Joann; Carr, Marilyn (ed.) <i>Mainstreaming Informal Employment and Gender in Poverty Reduction: a handbook for policymakers and other stakeholders</i>. London, Commonwealth Secretariat, 2004. • INEGI. Encuesta Nacional de Ocupación y Empleo 2005-2019. • _____. <i>Transición a un Nuevo Modelo de Encuesta Laboral en México: Informe de la Prueba Estadística ENOE 19/20 (pdf)</i>. 2020. • OECD. <i>Handbook for Measurement of the Non-Observed Economy</i>. Paris, 2002 • OIT. <i>Measuring Informality: A Statistical Manual on the Informal Sector and Informal Employment</i>. Geneve, 2013. • _____. Decimoquinta Conferencia Internacional de Estadísticos del Trabajo: Informe de la Conferencia, Ginebra 1993. • _____. Decimoséptima Conferencia Internacional de Estadísticos del Trabajo: Informe de la Conferencia, Ginebra 2003. • _____. Decimonovena Conferencia Internacional de Estadísticos del Trabajo: Informe de la Conferencia, Ginebra 2013. • _____. Vigésima Conferencia Internacional de Estadísticos del Trabajo: Informe de la Conferencia, Ginebra 2018. • Negrete, Rodrigo. <i>El concepto estadístico de informalidad y su integración bajo el esquema del Grupo de Delhi</i>. Realidad, Datos y Espacio, Vol. 2 Núm.3 septiembre-diciembre 2011. • _____. <i>Testing the Concepts of Dependent Contractors and Informality in Mexico</i>. Documento de discusión para la segunda reunión del grupo OIT

	para la revisión de los estándares de medición de la informalidad (pdf). 2020.
Hipótesis del trabajo:	Es posible alcanzar un nivel más alto de armonización de las estadísticas de informalidad al tiempo que generar un nuevo esquema integrador sobre las otras formas de trabajo y de inserción laboral definidas por la Organización Internacional del Trabajo (OIT) en las CIET 19 y 20, al tiempo que se establezca el lugar de la economía informal y sus vínculos con otros fenómenos que desafían a la contabilidad nacional. Para ello la experiencia acumulada por el INEGI, así como las pruebas de transición hacia un modelo de cuestionario ENOE 19/20 son fundamentales en proporcionar evidencia empírica y clarificar los puntos de discusión requeridos para una nueva resolución CIET en 2023.
Objetivo general:	Revisar los lineamientos de las CIET XV (1993) y XVII (2003) a la luz de la experiencia que se tiene en la materia, así como de los cambios que ha experimentado el mercado laboral y generar una propuesta actualizada de medición en lo conceptual y técnico.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Sintetizar y formular las mejores prácticas en materia de medición de la informalidad tanto desde el punto de vista laboral como el económico. 2. Fortalecer la discusión del grupo experto a partir de la experiencia mexicana. 3. Trascender el mero esquema de partición formal/informal por uno de gradientes de formalidad/informalidad. 4. Experimentar un modelo de indicadores al respecto. 5. Contribuir a la elaboración de los borradores de resolución de la XXI CIET.
Tiempo para su terminación:	36 meses.
Productos esperados:	<ul style="list-style-type: none"> • Reporte de investigación. • Paper y presentación para la tercera reunión del "Working Group for the Revision of the Standards for Statistics on Informality". • Propuesta de un marco integrador del empleo y la economía informal en el marco revisado de la Contabilidad Nacional. • Comentarios a toda la documentación que genere el grupo y subgrupos de trabajo en los que se participa. • Informe sobre los puntos de discusión del borrador de la resolución I de la 21 Conferencia Internacional de Estadísticos del Trabajo 2023.
Destinatarios:	<ul style="list-style-type: none"> • Organización Internacional del Trabajo (OIT). • Dirección General Adjunta de Encuestas Sociodemográficas. • Dirección General Adjunta de Cuentas Nacionales.
Responsable de la ejecución del proyecto:	Rodrigo Tadeo Negrete Prieto.

Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Trabajo
Técnicas:	Estadística descriptiva
Fuente de datos:	Tradicional
Estatus:	NUEVO

Nombre del Proyecto:	20. Codificación Automática de actividad económica y ocupación con Métodos de Machine Learning para proyectos estadísticos en hogares. Segunda etapa. (DGES)
Descripción general del problema a resolver:	A partir de los resultados positivos obtenidos en la primera etapa del proyecto de Codificación con algoritmos de Machine Learning, es pertinente transitar hacia su incorporación en los procesos productivos que llevan a cabo diariamente las oficinas de estadística. En este sentido, la segunda etapa del proyecto plantea avanzar en los siguientes aspectos: 1) evaluar los algoritmos en el estado del arte, 2) construir una base de datos tipo <i>ground truth</i> como insumo para las metodologías y algoritmos a evaluar, 3) Evaluar la incorporación dentro de un proceso productivo.
Importancia y utilidad del tema para el INEGI:	Este proyecto ayudará a investigar métodos, técnicas y procesos para la aplicación en fases productivas de metodologías de Procesamiento de Lenguaje Natural para tareas de Codificación que realiza el INEGI.
Información y fuentes:	https://statswiki.unece.org/display/ML/WP1+-+Pilot+Studies .
Hipótesis del trabajo:	El uso de métodos de inteligencia artificial permitirá mejorar la calidad de la codificación realizada actualmente, así como ahorrar tiempo y recursos económicos para esta fase dentro de las encuestas en hogares.
Objetivo general:	Automatizar procesos de codificación de actividad económica y ocupación de proyectos estadísticos en hogares mediante técnicas de Procesamiento de Lenguaje Natural.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Evaluar los algoritmos en el estado del arte. 2. Expertos temáticos y personal de la DGES construirán una base de datos tipo <i>ground truth</i> como insumo para los algoritmos de Machine Learning. 3. Evaluar la incorporación dentro de un proceso productivo.
Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento reporte de la investigación. • Algoritmo para el caso de uso que se llevará a cabo dentro de la DGES del INEGI. • Base de datos tipo <i>ground truth</i>. • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Sociodemográficas.

Responsable de la ejecución del proyecto:	José Alejandro Ruiz Sánchez.
Observaciones:	<p>En colaboración con el Laboratorio de Ciencia de Datos y la Dirección General de Estadísticas Sociodemográficas.</p> <p>El trabajo se desarrollará en el marco de las actividades del "Grupo de Aprendizaje de Máquina" de The United Nations Economic Commission for Europe y en colaboración entre la Dirección General de Integración, Análisis e Investigación y la Dirección General de Estadísticas Sociodemográficas.</p> <p>El trabajo es la continuación de la primera etapa del proyecto "Codificación Automática de actividad económica y ocupación con Métodos de Machine Learning para proyectos estadísticos en hogares".</p>
Subsistema:	Sociodemográfica
Temática:	Trabajo
Técnicas:	Ciencia de datos
Fuente de datos:	Tradicional
Estatus:	NUEVO

Bienestar

Nombre del Proyecto:	21. (12-2020) Uso de técnicas de BIG DATA a partir de información de Twitter para bienestar subjetivo y salud mental.
Descripción general del problema a resolver:	Extraer de los mensajes de Twitter elementos de información subyacentes que permitan generar indicadores del estado de ánimo e indicadores asociados con la prevalencia de condiciones de salud mental entre la población tuitera.
Importancia y utilidad del tema para el INEGI:	La disponibilidad de fuentes de información masiva y en tiempo real ha despertado el interés de diversas agencias nacionales y multinacionales productoras de estadísticas oficiales que consideran que dichas fuentes pueden contribuir a mejorar la oportunidad y la eficiencia con que producen estadística, aunque existe preocupación sobre la precisión de los resultados, así como sobre su representatividad. Resulta de interés para el INEGI participar activamente en este movimiento internacional con investigación propia que le ayude a estar mejor preparado para incorporar el enfoque de BIG DATA en su quehacer estadístico y geográfico. En particular, profundizar en la explotación de Twitter como fuente de información estadística geo-referenciada abre la puerta a nuevas maneras de atender las necesidades de información en el país.
Información y fuentes:	Twitter.
Hipótesis del trabajo:	La información que los usuarios de Twitter expresan en esa red social incorpora señales sobre su bienestar y salud mental que son susceptibles de ser extraídos mediante técnicas de ciencia de datos.
Objetivo general:	Generar indicadores de alta frecuencia de bienestar subjetivo y salud mental de la población tuitera.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Complementar el trabajo desarrollado hasta el momento para dar seguimiento al estado de ánimo de los tuiteros con refinamientos que permitan conocer los cocientes de positividad asociados con temas específicos. 2. Explorar el uso de Twitter para conocer sobre la salud mental de segmentos de la población tuitera, en especial mujeres adolescentes.
Tiempo para su terminación:	6 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Producto de estadística experimental con indicadores contruidos a partir de tuits. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Sociodemográficas.

	<ul style="list-style-type: none"> • Secretaría de Salud.
Responsable de la ejecución del proyecto:	Ana Miriam Romo Anaya.
Observaciones:	Trabajo colaborativo entre la Coordinación General de Informática con la participación de su personal (Juan Muñoz, Silvia Fraustro y Abel Coronado y parte de su grupo de colaboradores) es esencial para el desarrollo del proyecto.
Subsistema:	Sociodemográfica
Temática:	Bienestar
Técnicas:	Ciencia de datos
Fuente de datos:	No tradicional
Estatus:	CONTINÚA

Nombre del Proyecto:	22. Actualización con CPV 2020 de la Identificación de nivel de marginación en asentamientos humanos mediante la integración y análisis datos estadísticos, geográficos y de percepción remota.
Descripción general del problema a resolver:	<p>La dinámica en el crecimiento de las ciudades no siempre va acompañada de una debida planeación. En muchas ocasiones este crecimiento responde a la necesidad de poblaciones de escasos recursos que habitan zonas con poca o nula infraestructura urbana. Este fenómeno trae consigo la existencia de una población con alto nivel de marginación y de la cual no se tiene un registro estadístico actualizado continuamente.</p> <p>Existen índices de marginación como el que genera el Consejo Nacional de Población y el índice de Rezago social del Consejo Nacional de Evaluación de la Política de Desarrollo Social. Sin embargo, estos índices están basados en los Censos de Población y Vivienda y en la Encuesta Intercensal. Por lo que no se actualiza con la debida prontitud.</p>
Importancia y utilidad del tema para el INEGI:	<p>El combate a la pobreza (Objetivo 1) así como el desarrollo de asentamientos sostenibles (Objetivo 11) son parte importante de la agenda de los Objetivos de Desarrollo Sostenible (ODS). La estimación del grado de marginación en asentamientos humanos mediante el aprovechamiento de recursos de percepción remota y aprendizaje computacional podrá ser de gran utilidad para poder medir de manera oportuna indicadores tales como el 1.2.2 y el 11.1.3.</p> <p>La aplicación de percepción remota para la estimación de los ODS es una de las tendencias más innovadoras y con mayor potencial de los últimos años. Este tipo de aplicaciones posicionan al INEGI junto con las oficinas de estadística más avanzadas en el aprovechamiento de los avances tecnológicos y la ciencia de datos para la generación de información oportuna para la medición de los avances en los ODS.</p>
Información y fuentes:	<ul style="list-style-type: none"> • Censo de Población y Vivienda 2020 (CPV 2020). • Imágenes de percepción remota: <ul style="list-style-type: none"> • Geoeye • WorldView • Ortofotografía • Landsat • Cubo de Datos Geoespaciales de México. • Cartografía Geoestadística. • Cartografía Topográfica. • Variables geográficas (distancia a carreteras, instalaciones o edificios públicos). • Directorio Estadístico Nacional de Unidades Económicas (DENUE)/ El Registro Estadístico de Negocios de México (RENEM). • El Censo de Escuelas, Maestros y Alumnos de Educación Básica y Especial (CEMABE). • Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH). • Registros administrativos de consumo de energía eléctrica.

Hipótesis del trabajo:	Mediante la integración y análisis de fuentes tradicionales e imágenes satelitales de alta resolución es posible identificar asentamientos y medir el grado de marginación de las personas que los habitan.
Objetivo general:	Generar un indicador del grado de marginación actualizable anualmente, con una alta desagregación geográfica, y basado en datos estadísticos, geográficos y de percepción remota.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Integración de datos provenientes de fuentes tradicionales y diversas fuentes de percepción remota. 2. Optimización de algoritmo de aprendizaje computacional para la identificación de grado de marginalidad. 3. Aplicación de algoritmo para analizar la evolución del grado de marginalidad partiendo de los datos del censo 2010. 4. Realizar predicciones de 2011 a 2020. 5. Validación de los resultados obtenidos utilizando los datos del CPV 2020.
Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistasicytconacyt.mx/). • Mapas de la evolución de la marginalidad (2010-2020) con datos del Censo de Población y Vivienda 2020. Mapas de la evolución de la marginalidad (2010-2020).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Sociodemográficas. • Secretaría de Bienestar. • Consejo Nacional de Evaluación de la Política de Desarrollo Social.
Responsable de la ejecución del proyecto:	Elio Atenógenes Villaseñor García.
Observaciones:	El producto se encuentra dentro del marco del proyecto: Large-Scale High-Resolution Detection of Vulnerable Settlements que resultó favorecido en la convocatoria internacional emitida de manera conjunta por The Group on Earth Observations (GEO) y Google Earth Engine (GEE).
Subsistema:	Sociodemográfica
Temática:	Bienestar
Técnicas:	Ciencia de datos

Fuente de datos:	No tradicional
Estatus:	NUEVO

Nombre del Proyecto:	23. Indicadores sintéticos de salud física y mental con información longitudinal.
Descripción general del problema a resolver:	<p>Tanto la Encuesta Demográfica Retrospectiva (EDER) como la Encuesta Nacional sobre Salud y Envejecimiento en México (ENASEM) recogen información longitudinal que permite estudiar las relaciones entre características de origen, eventos, transiciones y trayectorias en la vida de las personas. En ambos casos se obtiene información relacionada con la salud física y mental a lo largo de varios años. En la EDER 2017 se compiló información de personas menores de 55 años de edad, y en la ENASEM 2018 se obtuvo información de individuos con 50 años y más, por lo que en ese sentido son complementarias.</p> <p>Hasta el momento, los indicadores de salud que ofrece el instituto a sus usuarios provienen de programas transversales y de estadísticas de morbilidad, pero es posible que del análisis de la información longitudinal pueda definirse otro tipo de indicador sintético para distintos grupos de población y regiones del país. Indicadores de este tipo son necesarios hoy en día para el diseño de políticas públicas.</p> <p>Se trata entonces de explotar la información longitudinal con miras al diseño de indicadores sintéticos de salud física y mental que mejoren la comprensión de los problemas de salud pública en México.</p>
Importancia y utilidad del tema para el INEGI:	<p>Actualmente, en el Catálogo Nacional de Indicadores hay 12 indicadores clave para el subtema salud. En el caso de los concernientes al estado de salud poblacional, ninguno se refiere a alguna enfermedad altamente prevalente o con tendencia creciente en los últimos años, de manera que se identifican oportunidades para la producción y suministro de información de calidad, pertinente, veraz y oportuna que coadyuve al desarrollo nacional.</p> <p>Esta investigación puede contribuir a la generación de indicadores clave en el tema de salud, de conformidad con el artículo 22 de la Ley del SNIEG (LSNIEG).</p>
Información y fuentes:	<ul style="list-style-type: none"> • Encuesta Demográfica Retrospectiva (EDER) 2017. • Encuesta Nacional sobre Salud y Envejecimiento en México (ENASEM) 2018.
Hipótesis del trabajo:	La explotación combinada de la EDER 2017 y la ENASEM 2018 permitirá identificar indicadores novedosos y útiles de salud física y mental.
Objetivo general:	Explotar la información longitudinal de la EDER 2017 y la ENASEM 2018 con el fin de proponer indicadores sintéticos de salud física y mental.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Identificar la tipología de trayectorias asociadas a la prevalencia de enfermedades físicas o mentales. 2. Explorar la diversidad regional de los indicadores.
Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Propuesta de reporte de estadísticas experimentales que considere un conjunto sintético de indicadores de salud física y mental para México. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en

	una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Subsistema Nacional de Información Demográfica y Social. • Dirección General de Estadísticas Sociodemográficas. • Secretaría de Salud. • Instituto Nacional de Salud Pública.
Responsable de la ejecución del proyecto:	Olinca Dessirée Páez Domínguez.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Bienestar
Técnicas:	Econometría
Fuente de datos:	Tradicional
Estatus:	NUEVO

Nombre del Proyecto:	24. Sistema de indicadores del bienestar en México y sus entidades federativas en 2021.
Descripción general del problema a resolver:	El levantamiento de la Encuesta Nacional de Bienestar Autorreportado (ENBIARE) 2021 brinda la oportunidad de contar con información por entidad federativa de bienestar subjetivo en México, la cual se puede combinar con información de otras fuentes para integrar un conjunto multidimensional de indicadores del bienestar que sirva como punto de partida para el monitoreo de la evolución del bienestar objetivo y subjetivo en México a lo largo de los años.
Importancia y utilidad del tema para el INEGI:	El INEGI ha establecido compromisos con la Organización para la Cooperación y el Desarrollo Económicos (OCDE) para generar indicadores sobre el progreso de las sociedades más allá de las cifras convencionales de crecimiento económico. Es fundamental posicionar ante la comunidad de usuarios una versión más amplia al respecto dentro del marco How's Life-WISE que ha establecido el organismo. Dicho marco a su vez permitirá evaluar áreas de oportunidad del INEGI en cuanto a la generación de fuentes de información e indicadores para estar a la par de lo que ofrecen los países miembros.
Información y fuentes:	ENBIARE 2021 más un conjunto de indicadores sociales, demográficos, medio ambientales y económicos por definir.
Hipótesis del trabajo:	Los resultados de la ENBIARE 2021 en combinación con otras fuentes de información disponible para México hacen posible combinar variables objetivas y subjetivas de diversas índoles a efecto de integrar un conjunto a la vez sintético y multidimensional del bienestar que sirva para monitorear la evolución del bienestar en México desde la perspectiva de lo que ocurre en sus 32 entidades federativas.
Objetivo general:	Explotar la información de la ENBIARE 2021, en combinación con un conjunto amplio de otras fuentes de datos para generar un sistema de indicadores de bienestar a nivel de entidad federativa inspirado en el marco de bienestar de la OCDE, pero que tome en consideración otros marcos de bienestar disponibles en el mundo actualmente y la información y experiencia disponible en México en la actualidad.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Partiendo del marco de bienestar de la OCDE, complementado de ser necesario con elementos de otros marcos de bienestar disponibles, construir el marco de bienestar sobre el cual se estará trabajando. 2. Identificar los indicadores del marco de bienestar para los que se cuenta con información disponible en México en la actualidad y aquellos indicadores para los que no se dispone de información. 3. Proponer vías de generación de la información faltante para que gradualmente pueda estar disponible en el futuro. 4. Partiendo de los elementos del marco de bienestar que puedan ser atendidos con la información disponible en México en 2021, alimentar todas las variables que sea posible del marco de bienestar por entidad federativa. 5. Reportar el sistema de indicadores de bienestar como estadística experimental. 6. Hacer un breve diagnóstico comparativo del bienestar en las 32 entidades federativas del país.

	<p>7. Seleccionar un conjunto de indicadores de bienestar que pueda ser reportado en una serie anual para una serie de al menos 10 años y repórtalo.</p> <p>8. Preparación de las salidas de información para la página web del INEGI respecto a la ENBIARE 2021</p>
Tiempo para su terminación:	24 meses a partir de que estén disponibles los resultados de la ENBIARE 2021.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Reporte del sistema de indicadores de bienestar en México como estadísticas experimentales. • Salidas de información para la página web del INEGI respecto a la ENBIARE 2021. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Sociodemográficas. • Secretaría de Bienestar. • Consejo Nacional de Evaluación de la Política de Desarrollo Social.
Responsable de la ejecución del proyecto:	Rodrigo Tadeo Negrete Prieto.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Bienestar
Técnicas:	Estadística descriptiva
Fuente de datos:	Tradicional
Estatus:	NUEVO

Sociodemografía

Nombre del Proyecto:	25. (21-2020) Evolución de la consistencia de las estimaciones de fecundidad entre la EDER 2011 y la EDER 2017.
Descripción general del problema a resolver:	<p>La expectativa de que la Encuesta Demográfica Retrospectiva (EDER) se levante con cierta periodicidad requiere de ejercicios analíticos que exploren la consistencia de las estimaciones y los patrones obtenidos a partir de levantamientos consecutivos de la encuesta con diseño estadístico y conceptual diferentes.</p> <p>Una temática clave para la que ha sido posible evaluar consistencia entre levantamientos de la EDER, y también respecto a otras fuentes de información transversales, es el estudio del cambio de la fecundidad en México. Con el levantamiento más reciente de la EDER en 2017 se obtuvo suficiente información para analizar, por primera vez, las transformaciones reproductivas en su diversidad geográfica, por lo que es posible evaluar la consistencia de las estimaciones y los patrones nacionales, al tiempo que se explora la variedad regional en el país.</p>
Importancia y utilidad del tema para el INEGI:	El Instituto requiere de elementos que permitan evaluar objetivamente la pertinencia de ajustes en el diseño estadístico y conceptual de próximos levantamientos.
Información y fuentes:	Encuesta Demográfica Retrospectiva, 2011 y 2017.
Hipótesis del trabajo:	Hay consistencia en las estimaciones y patrones nacionales sobre la fecundidad producidas por los dos levantamientos.
Objetivo general:	Evaluar la consistencia de las estimaciones nacionales de la fecundidad producidas con la información de la EDER 2017 respecto a las de la EDER 2011.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Actualizar las estimaciones del cambio y las diferencias reproductivas en México. 2. Explorar la diversidad regional de los patrones reproductivos.
Tiempo para su terminación:	3 meses.
Productos esperados:	<ul style="list-style-type: none"> • Publicación en obra colectiva basada en la EDER 2017. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistacytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Sociodemográficas (DGES). • Dirección General de Integración, Análisis e Investigación (DGIAI).

Responsable de la ejecución del proyecto:	Olinca Dessirée Páez Domínguez.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Sociodemografía
Técnicas:	Análisis de secuencias
Fuente de datos:	Tradicional
Estatus:	CONTINÚA

Nombre del Proyecto:	26. Corregir sesgos poblacionales de resultados obtenidos a partir de Twitter.
Descripción general del problema a resolver:	En México, como en otros países, la población de tuiteros activos exhibe características sociodemográficas diferentes a las de la población general. En efecto, la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH)-2018 nos permite afirmar que los tuiteros mexicanos tienen una edad promedio menor; o provienen principalmente de estratos socioeconómicos altos, contra lo que ocurre para la totalidad de la población. En consecuencia, los agregados estadísticos que se deriven de información extraída de sus cuentas podrán exhibir desviaciones más o menos importantes de los que se obtendrían a partir de muestras representativas. Aunque por otras razones las encuestas por muestreo pueden mostrar desviaciones semejantes, pero en este caso técnicas de reponderación han sido introducidas con el fin de corregir dichos efectos no deseados. Un ejemplo de tales técnicas es la que se conoce como post-estratificación. Se ha sugerido seguir un procedimiento similar para que resultados provenientes de redes sociales sean extendidos a la totalidad de la población.
Importancia y utilidad del tema para el INEGI:	Todo proyecto de producción de información estadística llevado a cabo por el INEGI hace explícito cada uno de los dominios de análisis; en particular, la mayor población a la que hacen referencia los resultados. El diseño de la muestra busca que esta sea representativa de aquella. Cuando se recurre a información geo-referenciada de redes sociales no queda claro que los resultados sean aplicables a poblaciones más amplias que las de los mismos usuarios de dichas redes, en cuyo caso las inferencias son limitadas. Su amplia y frecuente disponibilidad, el formato en que se encuentra almacenada, y el costo relativamente bajo de acceder a ella hacen atractiva a dicha información para complementar la producción de información estadística por medios tradicionales, sin embargo, en México gracias a ENDUTIHs recientes se cuenta con una descripción sociodemográfica de los usuarios de redes. Además de hacer explícitos los sesgos en ese sentido, lo que ya es útil, tal descripción puede ser aprovechada para post-estratificar la información obtenida de esas fuentes. Se busca que los resultados obtenidos de este modo sean representativos para la mayor población posible, deseablemente la misma de la encuesta. Bajo estas condiciones, el INEGI estará en condiciones de corregir los resultados obtenidos a partir de estas fuentes.
Información y fuentes:	Información sociodemográfica de los informantes de la ENDUTIH, identificados como usuarios de Twitter. Su nombre de usuario en esta red, voluntariamente aportado después de haberle explicado el propósito del estudio. Textos de los Tweets publicados por dichos usuarios.
Hipótesis del trabajo:	La publicación por parte del INEGI de indicadores que explotan información de redes sociales puede presentarse corregida de sus sesgos de representatividad respecto de la población en general mediante la aplicación de algoritmos de post estratificación de aprendizaje automático entrenados con la información combinada de la ENDUTIH y de la red social en cuestión.

Objetivo general:	Reducir sesgos en las estimaciones derivadas de la explotación de la información publicada en Twitter, debidos a las significativas diferencias entre las estructuras por edad, sexo, logro académico y nivel socioeconómico de la población general y la usuaria de redes, a través de la post-estratificación.
Objetivos específicos:	<ol style="list-style-type: none"> 1. El sexo, la edad, el estrato socioeconómico y los años de formación académica de los usuarios de redes sociales pueden ser inferidos con razonable precisión. 2. Las estructuras por grupo de edad, sexo, estrato socioeconómico y la formación académica de los tuiteros en un conjunto de datos pueden ser modificadas hasta hacerlas semejantes a las encontradas en población abierta. 3. Los resúmenes estadísticos obtenidos a partir de la muestra corregida pueden ser tomados como representativos para la población abierta. 4. Producir información estadística oportuna y de bajo costo a partir de publicaciones en Twitter.
Tiempo para su terminación:	24 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Encargados del levantamiento de información mediante encuestas por muestreo del INEGI.
Responsable de la ejecución del proyecto:	Víctor Alfredo Bustos y de la Tijera.
Observaciones:	Contingente a que suficientes informantes de la ENDUTIH acepten participar en el proyecto.
Subsistema:	Sociodemográfica
Temática:	Sociodemografía
Técnicas:	Post-estratificación
Fuente de datos:	No tradicional
Estatus:	NUEVO

Nombre del Proyecto:	27. Estratificador multivariado para el Censo de Población y Vivienda 2020.
Descripción general del problema a resolver:	La cantidad de información detrás de un Censo de Población y Vivienda como el que hizo el INEGI en 2020 puede resultar intimidante y hacer demasiado pequeño el número de sus usuarios intensivos. En este sentido, la disponibilidad de herramientas tecnológicas que faciliten la explotación del censo por parte de la ciudadanía abona a un mayor y más equitativo impacto de sus resultados. La tarea es entonces poner a disposición de los usuarios potenciales del Censo de Población y Vivienda 2020 una herramienta semi automatizada de estratificación de la que se puedan derivar agrupaciones con diferentes propósitos de la información censal e incluso la generación de productos derivados de utilidad pública.
Importancia y utilidad del tema para el INEGI:	Una vez realizado el Censo de Población y Vivienda 2020 es necesario favorecer su explotación adecuada por el mayor número posible de personas.
Información y fuentes:	Información sociodemográfica de los informantes de la ENDUTIH, identificados como usuarios de Twitter. Su nombre de usuario en esta red, voluntariamente aportado después de haberle explicado el propósito del estudio. Textos de los Tweets publicados por dichos usuarios.
Hipótesis del trabajo:	Un sistema de estratificación multivariado semi automatizado que haga uso de un menú suficiente de técnicas de estratificación y que ofrezca sugerencias sobre el número de estratos más conveniente para cada conjunto de variables a seleccionar, ayudará a que los resultados del Censo de Población y Vivienda 2020 sean de mayor impacto en la sociedad y permitirá generar productos derivados novedosos, útiles y adecuados a las necesidades de los usuarios.
Objetivo general:	Generar un sistema de estratificación multivariada semi automatizado que facilite la explotación de los resultados del Censo de Población y Vivienda 2020 mediante un menú suficiente de técnicas de estratificación y que ofrezca sugerencias sobre el número de estratos más conveniente para cada conjunto de variables a seleccionar.
Objetivos específicos:	<ol style="list-style-type: none"> 5. Recuperar y actualizar técnica y tecnológicamente el "Estratificador INEGI" generado originalmente para el Conteo de Población de 2010. 6. Vincularlo con los resultados del Censo de Población y Vivienda 2020. 7. Poner el estratificador a disposición del público por medio de la página del INEGI. 8. Generar explotaciones específicas de las capacidades de estratificación del sistema, comenzando por un mapeo de aspectos socioeconómicos o de bienestar en México.
Tiempo para su terminación:	12 meses.

Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Estratificador publicado en la página del INEGI. • Reporte gráfico por medio de mapas que reporten aspectos socioeconómicos y de bienestar relevantes derivados del Censo de Población y Vivienda 2020. • Artículo con calidad suficiente para ser publicado en “Realidad, datos y espacio. Revista Internacional de Estadística y Geografía”, que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Estadísticas Sociodemográficas. • Secretaría de Bienestar.
Responsable de la ejecución del proyecto:	Víctor Alfredo Bustos y de la Tijera.
Observaciones:	
Subsistema:	Sociodemográfica
Temática:	Sociodemografía
Técnicas:	Estratificación multivariada
Fuente de datos:	Tradicional
Estatus:	NUEVO

Recursos naturales

Nombre del Proyecto:	28. Análisis del consumo de agua en la industria manufacturera de México.
Descripción general del problema a resolver:	La preocupación por la escasez de agua, especialmente en algunos estados del país, es una constante en los últimos años. Determinar qué actividades económicas manufactureras consumen mayores cantidades de agua y en qué medida este recurso natural puede convertirse en un factor limitante del crecimiento económico y la reproducción social alertaría a los tomadores de decisiones para tomar medidas precautorias.
Importancia y utilidad del tema para el INEGI:	Explotar la información generada por el Instituto, su combinación con otras bases de datos externas para ser utilizadas en productos finales para la toma de decisión por usuarios interesados internos y externos. Además, el tema es de relevancia local, nacional e internacional ya que el agua es un bien escaso, difícilmente renovable afectado por el cambio climático global y su efecto local-regional. Por ser un bien relevante se ha considerado cotizarlo en el mercado bursátil.
Información y fuentes:	<ul style="list-style-type: none"> • Estadísticas e Indicadores de la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT). • Anuarios de la Comisión Nacional del Agua (CONAGUA). • Banco de Información Económica (BIE). • Sistema Estatal y Municipal de Base de Datos (SIMBAD) 2010. • Encuesta Mensual de la Industria Manufacturera (EMIM). • Cuentas nacionales y satélites. • Matriz de insumo producto (MIP) INEGI. • Indicadores del Banco de México. • Informes gubernamentales de los Estados. • Artículos y publicaciones nacionales e internacionales. • Matriz insumo-producto regional (MIPR) y • Matriz de contabilidad social regional (MCSR).
Hipótesis del trabajo:	El consumo de agua relacionado con las operaciones de unidades económicas manufactureras está asociado con el tamaño, el giro específico y la ubicación geográfica de esas unidades económicas, por lo que es susceptible de ser modelado fines de análisis de sostenibilidad del crecimiento.
Objetivo general:	Generar un modelo estadístico-matemático que estime el consumo de agua derivado de actividades manufactureras específicas en áreas geográficas específicas a efecto de generar estimaciones de sostenibilidad del crecimiento de esas actividades en relación con el mencionado recurso natural.
Objetivos específicos:	<ol style="list-style-type: none"> 1. Elegir la región o regiones geográficas de referencia y las actividades manufactureras a considerar. 2. Desarrollar el modelo estadístico-matemático que vincule consumo de agua con actividad manufacturera. 3. Establecer medidas de sostenibilidad del crecimiento de las actividades manufactureras consideradas en las regiones geográficas seleccionadas. 4. Reportar las medidas de sostenibilidad de crecimiento manufacturero encontradas.

Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en “Realidad, datos y espacio. Revista Internacional de Estadística y Geografía”, que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistascytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Geografía y Medio Ambiente. • Secretaría de Economía. • Secretaría de Medio Ambiente y Recursos Naturales.
Responsable de la ejecución del proyecto:	Dora Elena Ledesma Carrión.
Observaciones:	
Subsistema:	Geografía y medio ambiente, ordenamiento territorial y urbano
Temática:	Recursos naturales
Técnicas:	Análisis numérico
Fuente de datos:	Tradicional
Estatus:	NUEVO

Buscador de la página del INEGI

Nombre del Proyecto:	29. Buscador del Sitio del INEGI. (DGCSPIRI)
Descripción general del problema a resolver:	<p>En la encuesta de salida del Sitio del INEGI, durante 2020, el 54.4% de los usuarios mencionaron haber encontrado la información que estaban buscando. En 2017, cuando aún estaba disponible al público el sitio anterior que no contaba con un buscador, este indicador estaba en 49.1%. Por otro lado, el bounce rate tras realizar una consulta en el buscador fue de 68.8% en 2019. Estas son señales claras de las áreas de oportunidad que se tienen en materia de accesibilidad a la información.</p> <p>El buscador actual del Sitio utiliza un enfoque basado en palabras clave. Este tipo de buscadores son eficaces en los casos donde el usuario sabe exactamente cómo hacer la consulta para encontrar lo que busca. Por ejemplo, el buscador facilita al usuario encontrar el Uniform Resource Locator (URL) del Programa de Información que ya conoce al introducir su nombre, parte de éste o su acrónimo. Sin embargo, es común que las búsquedas expresadas en lenguaje natural no arrojen resultados.</p> <p>Por otro lado, el buscador actual no cuenta con un mecanismo de ordenamiento de los resultados de búsqueda en función de la relevancia para atender la necesidad de información. La mayoría de los buscadores entregan los artículos en donde puede encontrarse la información que satisfaga sus criterios de búsqueda ordenados por la probabilidad de que encuentren lo que están buscando. El buscador actual retorna al usuario listados exhaustivos y no ordenados de todos los registros que contienen los parámetros que proporcionó el usuario; dichos resultados a veces exceden los cientos de artículos dificultando nuevamente la accesibilidad de la información.</p> <p>Se han realizado acciones encaminadas a mejorar el desempeño del buscador actual. Entre estas acciones se encuentran: el almacenamiento sistemático de los procesos de búsqueda, el etiquetado semántico de una porción de las solicitudes y el uso de objetos de conocimiento para la descripción de los productos.</p> <p>Actualmente, las técnicas de procesamiento de lenguaje natural en conjunción con los métodos de aprendizaje computacional facilitan la caracterización y recuperación de información de toda índole. Por ejemplo, las recientes mejoras en el buscador Google se basan en la incorporación de redes neuronales y procesamiento de lenguaje natural (proyecto BERT).</p> <p>Las oficinas de estadística no son ajenas a estos avances. Existen ya proyectos muy avanzados en los que se incorporan motores de búsqueda que aprovechan el procesamiento de lenguaje natural. Los proyectos más avanzados han desarrollado robots (<i>Statbots</i>) que procesan requerimientos de información estadística, presentando resultados que se calculan al vuelo. Estos avances serán presentados por las oficinas de estadística de Canadá y Holanda en la siguiente edición del “<i>Workshop on the Modernization of Official Statistics</i>” organizado por The United Nations Economic Commission for Europe (UNECE).</p>

	<p>En general, las Oficinas de Estadística y la industria dedicada a facilitar la búsqueda de los usuarios están incorporando métodos de procesamiento de lenguaje natural y aprendizaje computacional para la caracterización semántica de los parámetros de búsqueda de los usuarios y la recuperación de información. Se considera que la aplicación de estas técnicas a la búsqueda de información estadística y geográfica está ya probada por otras oficinas nacionales de estadística y proporcionaría una mejora sustantiva en la accesibilidad de la información para los usuarios del Sitio del INEGI.</p>
<p>Importancia y utilidad del tema para el INEGI:</p>	<p>Durante 2019 el buscador del sitio del INEGI fue consultado 3,130,247 veces. Diariamente, tuvo entre 2,500 y 15,000 consultas. La tendencia de uso durante 2020 es similar, con un total hasta el 20 de noviembre de 2,336,080.</p> <p>El buscador es una herramienta básica para facilitar la accesibilidad de la información pues permite al usuario, independientemente de su capacidad técnica, detectar si la información que requiere se tiene en el INEGI y en dónde está.</p> <p>Desarrollar estos métodos son fundamentales para el INEGI por varias cuestiones: Primero, está el hecho de mejorar la atención que se brinda a los usuarios que a diario realizan consultas en el buscador mediante el análisis del texto a través de técnicas de Procesamiento de Lenguaje Natural e Inteligencia Artificial. También se destaca que la ordenación de los resultados es de alta relevancia pues se brindará un mejor servicio al priorizar los resultados más relevantes. Por otra parte, se resalta el hecho de que este tipo de técnicas de análisis de textos y ordenación de resultados podrían ser extrapolados a otros productos internos y externos de INEGI.</p>
<p>Información y fuentes:</p>	<p>Se necesitará información proveniente de todos los canales que arrojen la información digital que se analizará y procesará, entre estos:</p> <ul style="list-style-type: none"> • Consultas de los usuarios y, en su caso, respuesta que se les dio. • Información de centro de atención telefónica, correos, Customer Relationship Management, entre otros. • - SDMX (Intercambio de datos estadísticos y metadatos).
<p>Hipótesis del trabajo:</p>	<p>Mediante la aplicación de técnicas de Procesamiento de Lenguaje Natural e Inteligencia Artificial será posible mejorar la calidad de los resultados que entrega el Buscador del Sitio del INEGI, así como hacer más eficiente el ordenamiento de dichos resultados.</p>
<p>Objetivo general:</p>	<p>Desarrollar componente de procesamiento de lenguaje natural que pueda integrarse al buscador actual de la página del INEGI, dotándolo de una mayor capacidad de recuperación y que permita ordenar los resultados de las consultas de acuerdo a criterios de relevancia basados en el análisis del comportamiento de los usuarios.</p>
<p>Objetivos específicos:</p>	<ol style="list-style-type: none"> 1. Análisis y modelación de los datos recolectados de las consultas realizadas. 2. Construcción de conjuntos de entrenamiento que incorporen las caracterizaciones semánticas realizadas. 3. Incorporar la información de los objetos de conocimiento para caracterizar semánticamente los productos.

	<p>4. Construir algoritmo de recuperación y ordenamiento basado en criterios de relevancia.</p> <p>5. Evaluación de los resultados obtenidos.</p>
Tiempo para su terminación:	12 meses.
Productos esperados:	<ul style="list-style-type: none"> • Algoritmo productivo para transformar los parámetros de búsqueda establecidos introducidos por el usuario en el buscador del Sitio del INEGI en un listado de artículos de conocimiento del propio sitio en orden de la relevancia que les asigna el algoritmo para la recuperación de la información que da respuesta a dicha pregunta. • Documento con reporte de investigación. • Artículo con calidad suficiente para ser publicado en "Realidad, datos y espacio. Revista Internacional de Estadística y Geografía", que de acuerdo con la escala otorgada por el CONACYT, y dependiendo el área de orientación científica a la cual pertenece el artículo, pueda incluirse en una revista de clasificación al menos Q4 o equivalente (http://www.revistasicytconacyt.mx/).
Destinatarios:	<ul style="list-style-type: none"> • Dirección General de Comunicación, Servicio Público de Información y Relaciones Institucionales (DGCSPIRI). • Los beneficiarios últimos serán los usuarios del buscador del portal del INEGI.
Responsable de la ejecución del proyecto:	Elio Atenógenes Villaseñor García.
Observaciones:	<p>Laboratorio de Ciencia de Datos -Elio Atenógenes Villaseñor García en colaboración con la Dirección General de Comunicación, Servicio Público de Información y Relaciones Institucionales (DGCSPIRI).</p> <p>Se cuenta con una amplia cantidad de datos que permitirán el entrenamiento de algoritmos específicos:</p> <ul style="list-style-type: none"> • Consultas del buscador y comportamiento de los usuarios ante los resultados expuestos. • Solicitudes de información en todos los canales de atención, con información estructurada y la relación de las respuestas emitidas por los operadores de los servicios de atención a usuarios. • Bases de datos con tags específicos en todos los artículos de conocimiento contenidos en el Sitio del INEGI. <p>Se pretende que si el producto de datos puede ser mejorado al:</p> <ul style="list-style-type: none"> - Introducir variables adicionales en los tags de los artículos del Sitio o; - En los formatos de registro de las solicitudes-respuestas en los canales de atención a usuarios, <p>Se realizarían los ajustes necesarios a los procesos operativos.</p>

Subsistema:	Transversal
Temática:	Buscador de la página del INEGI
Técnicas:	Ciencia de datos
Fuente de datos:	Tradicional
Estatus:	NUEVO