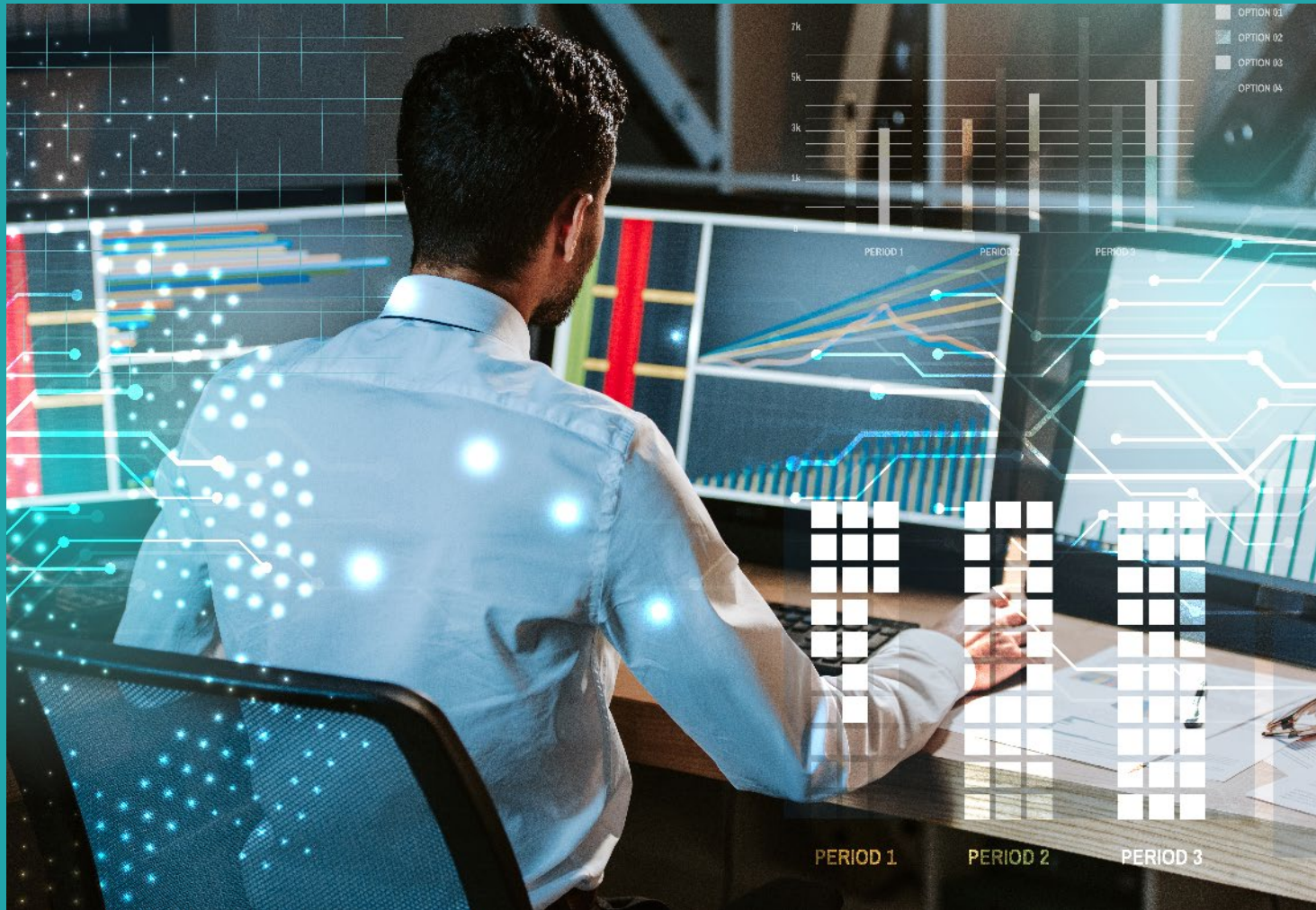


Introducción al Análisis de Datos para Políticas Públicas



Instituto de Investigación
y Educación Económica

1. Fundamentación

En los últimos años el volumen de datos en el mundo ha crecido de manera exponencial. Se estima que el 90% de los datos mundiales se ha creado en los últimos dos años y se predice un crecimiento de un 40% anual. A este ritmo, probablemente se supere la capacidad de almacenamiento. Por otra parte, la mejora de capacidad de las computadoras, la automatización de procesos y la generación de múltiples datos a través de la internet de las cosas (IOT por sus siglas en inglés) presentan múltiples desafíos a diversas disciplinas, tanto desde el punto de vista teórico como desde la aplicación de métodos específicos.

En lo que respecta a las ciencias sociales, la proliferación de datos ha ampliado las posibilidades de observación, experimentación y contrastación de hipótesis. La publicación y actualización sistemática de extensas bases de datos por parte de instituciones públicas y de organismos internacionales aporta el insumo necesario para potenciar las investigaciones. Al mismo tiempo, la capacidad de análisis de grandes volúmenes de datos se ve simplificada por la existencia de softwares de uso generalmente gratuito, la generación de bibliotecas de scripts o cadenas de comandos utilizables en la investigación, la estandarización de uso de indicadores y herramientas tales como los tableros de control.

Este curso representa un primer acercamiento para quienes deben realizar el análisis de grandes volúmenes de datos en sus investigaciones o en su ámbito profesional. Por su enfoque, está dirigido a quienes deseen profundizar en el impacto de políticas públicas de distintos ámbitos (educación, salud, producción, etc.).

II.- Objetivos

Luego de realizar el curso los asistentes serán capaces de:

- Manejar de forma más profesional los datos obtenidos de distintos tipos de fuentes
- Construir tableros de comando que permitan una comprensión rápida de los principales indicadores de sus áreas
- Elaborar potentes visualizaciones para evaluar con solidez el impacto de políticas públicas específicas
- Destacarse en su perfil profesional mediante la incorporación de habilidades para manejar las metodologías cuantitativas del futuro

II.- A quién está dirigido y conocimientos previos requeridos

Por sus contenidos y enfoque, el curso está diseñado para profesionales que estén interesados en contar con insumos para diseñar políticas públicas específicas y para evaluar convenientemente su impacto. Es por ello que será de interés para funcionarios públicos, investigadores, miembros de organizaciones del tercer sector, entre otros. Y, en términos más generales, a quienes deseen contar con habilidades que en los próximos años serán altamente valoradas por el mercado laboral.

En lo que respecta a conocimientos previos necesarios, es deseable que quienes participen cuenten con conocimientos básicos del uso de Excel.

Para un mejor aprovechamiento del curso, se requiere tener instalado Microsoft Excel 2016 o superior, u Office 365. Los usuarios de Mac podrán usar Excel, pero no Power Bi.

IV.- Cursado y metodología

El curso tiene una duración de 24 horas, distribuidas en 8 clases semanales de 3 horas cada una. La modalidad de cursado será por **streaming en vivo** los días jueves, de 6:30 a 9:30 p.m., hora argentina, comenzando el martes 4 de mayo y finalizando el 29 de junio de 2023.

La primera parte de cada clase estará dedicada a la revisión de los conceptos fundamentales asociados a cada tema, mientras que la segunda estará orientada a trabajar de modo práctico con bases de datos mediante la aplicación práctica de las diversas herramientas de análisis disponibles.

V.- Contenidos

Primera parte: Análisis de datos con Microsoft Excel

Clase 1. Conceptos generales en torno al análisis de datos

Data analysis versus Data science: diferencias y complementariedad. El proceso ETL. Tipos de bases de datos. MySQL. Storytelling: ¿qué queremos informar? Buenas prácticas de visualización. Microsoft Excel: trabajar con rangos y tablas. Autorrelleno. Atajos de teclado útiles para trabajar mejor.

Clase 2. Manejo de datos en Excel

Extracción de datos, formato CSV, separar texto en columnas. Tipos de datos: cualitativos y cuantitativos. Obtención de la información desde los datos, definición de insights. Tipos de datos: valores numéricos, texto y fórmulas. Inmovilizar paneles. Práctica: descargar un set de datos y aplicar el formato correcto.

Clase 3. Herramientas y Funcionalidades Útiles

Uso de filtros, buscar y reemplazar. Fórmulas y funciones: diferencias. Asistente de funciones: sumar, sumar.si, contar, contar.si, contara, función lógica Si. Referencia relativa y absoluta a una celda. Herramientas de análisis en Excel: estadística descriptiva, varianza, coeficiente de relación. Formatos condicionales.

Clase 4. Tablas dinámicas y macros

Tablas dinámicas: el procesamiento de grandes volúmenes de datos. Macros: automatizar Excel. ¿Un paso previo al RPA?. VBA, powerpivot y powermaps. Tableros de comando en Excel: ¿la manera más rápida y efectiva de resolver problemas?

Segunda parte: Análisis de datos con Power Bi

Clase 5. El paso de Excel a Power BI

Storytelling: ¿a quién queremos informar? Visualización de datos según los niveles jerárquicos. La transición de Excel a Power Bi, diferentes usos de cada uno, uso de Power Query. Diferentes lenguajes usados: VBA, lenguaje M, DAX. Versiones de Power Bi y el uso de datamarts. Ejemplos de usos. Otras herramientas de BI: Tableau y Google data studio.

Clase 6. El proceso de limpieza

El proceso de obtención, limpieza y transformación de datos. Selección de datasets de acceso público. Importancia de la limpieza. Normalizaciones. Breve introducción al lenguaje M.

Clase 7. El proceso de modelado de datos

Modelos de datos: copo de nieve vs estrella. Tipos de datos. Relaciones y campos clave, Jerarquías. Introducción a DAX. Diferencias entre DAX y Excel. Cálculos y expresiones DAX: Funciones comunes, columnas calculadas. Funciones de filtrado. Funciones de inteligencia de tiempo.

Clase 8. Visualización de datos

Definición de KPIs y creación de tableros de comando. Visuales predeterminadas. Galería de visuales. Mapas. Informes de alto impacto visual. Publicar y Compartir. Ejemplo práctico: trabajo con estadísticas de Covid durante la pandemia, descargados de <http://datos.salud.gob.ar/dataset>.

VI. Profesor

Hernán Aruj

Cuenta con un Máster en Economía y Desarrollo Industrial (Universidad Nacional de General Sarmiento) y con un Postgrado en Diseño y Evaluación de Políticas Públicas (Universidad Pompeu Fabra de Barcelona). Es Data Analyst (certificado por Digital House). Ha sido Auxiliar Docente en Macroeconomía y Política Económica en la Universidad de Buenos Aires. Actualmente es analista en el Ministerio de Seguridad de la Nación.

VII.- Certificado y condiciones de aprobación

El certificado de asistencia será otorgado por el **Instituto de Investigación y Educación Económica (I+E)**. Para acceder a él, los asistentes deberán:

Asistir al menos al 80% de las horas de cursado sincrónico
Presentar y aprobar un trabajo individual de aplicación. Las pautas para la realización del trabajo serán entregadas en el transcurso de la cuarta clase y la entrega final se hará durante la última clase. En escala de 1 a 10, el trabajo se aprobará con una calificación igual o mayor a 5.

Informes
contacto@ie.org.ar
(5411)4382-2332
www.ie.org.ar

