|  |
| --- |
| **CICLO DE FORMACIÓN EN PYTHON. INTRODUCCIÓN A PYTHON CON ORIENTACIÓN AL ANÁLISIS DE DATOS** |

1. **Fundamentación**

La actividad se inscribe en la política de recursos humanos del INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS (INDEC), orientada a desarrollar programas de formación y actualización del personal profesional, técnico y operativo del Instituto acordes a los cambios tecnológicos y en los procesos de trabajo.

En su función de servir estadísticas públicas, las direcciones estadísticas del INDEC deben producir datos en procesos de manera periódica que abrevan de diversas fuentes además de sujetar esos datos a múltiples transformaciones con procesos que pueden ser sensiblemente complejos. Sobre ese carácter periódico de su labor se debe agregar que deben cumplir un marco metodológico replicable y comunicable que asegure la transparencia de los datos producidos.

Estas características descriptas de las tareas del instituto hacen idóneo que sus técnicos puedan incluir Python entre sus herramientas cotidianas de trabajo. Python facilita: a) la automatización de tareas repetitivas como descargas de gran cantidad de archivos o la repetición del mismo proceso en diferentes conjuntos de datos; b) la replicación de procesos utilizando programación con Python permite parametrizar los mismos para poder funcionar en nuevos contextos (por ejemplo, el caso de conjuntos de datos actualizados); c) la documentación de los procesos, favoreciendo la corrección entre pares y a la documentación de los procesos, permitiendo contrastar el proceso creado con el marco metodológico que debe seguir; y d) la comunicación y transparencia en el lenguaje de código abierto que utiliza Python implica que el mismo puede ser compartido y replicado por cualquier persona sin necesidad de pagar licencia, incluso cualquier usuario del instituto puede consumir librerías de Python que facilitan muchas tareas, actualizadas periódicamente y revisadas por la enorme comunidad de usuarios.

Dada la necesidad de actualización constante en cuanto a ciencias de datos respecta, siendo ello clave para la ejecución de las principales tareas desarrolladas por el organismo, se decide su inclusión en el INDEC como propuesta de capacitación institucional.

A partir de experiencias anteriores, se advirtió la necesidad de modificar la estrategia pedagógica de la capacitación institucional. Por ello se configurará la presente capacitación en formato de ciclo formativo de aprendizaje en Python consistente en un curso introductorio que se complementará con talleres de aplicación específica, siendo esos no correlativos.

1. **Contribución esperada**

Se espera que la organización logre agilizar la construcción de diseños metodológicos de los distintos operativos estadísticos nacionales y elaborar indicadores básicos e información de orden social y económico de manera más eficiente, mediante la utilización de las herramientas que ofrece Python para la toma de decisiones. Este curso permitirá obedecer a la necesidad de actualización constante en cuanto a ciencias de datos respecta, siendo ello clave para la ejecución de las principales tareas desarrolladas por el organismo.

1. **Perfil del participante**

Agentes desarrolladores, profesionales y/o técnicos que trabajan en operativos estadísticos como en la elaboración y presentación de la información.

Requisitos: utilización de bases de datos; análisis de datos estadísticos; tener instalado Python en el ordenador; conocimiento de inglés nivel intermedio.

1. **Objetivos**

Se espera que los participantes logren:

* Incorporar conocimientos básicos acerca de Python.
* Entender las ventajas y casos de uso de la automatización de un proceso.
* Tener la experiencia de desarrollar un proceso sencillo que permita la:
  + Lectura de datos de diversas fuentes
  + Exploración de los datos importados
  + Transformación de los datos con fines analíticos
  + Visualización de los datos y de relaciones entre sus variables
  + Elaboración de conclusiones en base al trabajo hecho
  + Facilitación de la comunicación de las conclusiones por los gráficos elaborados.

1. **Contenidos**

**Unidad 1 –  Introducción a Python**

* ¿Por qué Python?
* Comparación con Excel, semejanzas con R
* Utilidad
* Desambiguación: lenguaje, intérprete e interfaz gráfica
* Conceptos básicos
* Tipos de datos
* Expresiones y sentencias
* Operadores lógicos y aritméticos
* Métodos de strings
* Bucles for y while
* Ejercicios

**Unidad 2 –  Visualización de datos**

* Introducción a Jupyter Notebook
* Introducción a las librerías de Python
* Librerías para visualización en Python: Seaborn, Matplotlib y Plotly
* ¿Por qué Plotly?
* Propiedades y estructura de los gráficos de plotly
* Ejemplos

**Unidad 3 – Pandas (Data Wrangling) parte I**

* Concepto de Tidy Data
* Lectura de distintas fuentes: CSV, Excel, Web, Git
* Tipos de datos de Pandas
* Explorar Data Frame
* Renombrar, crear y eliminar columnas
* Subsets y Query con condiciones lógicas
* Merge, concat

**Unidad 4 – Pandas (Data Wrangling) parte II**

* Operaciones vectoriales (aritméticas, strings)
* Melt, Pivot y Transpose
* Group By
* Exportar en distintos formatos (Csv, Excel, Paquet, Pickle). Ventajas y desventajas. Excelwritter

**Unidad 5 – Pandas (Data Wrangling) parte III**

* Transformaciones avanzadas: apply
* Otros métodos de transformación: cumsum, pctchange, rank, etc.
* Integración: Data Wranlging – Visualización – Comunicación
* Introducción al ejercicio final

**Unidad 6 – Integración final: práctica y revisión**

* Consultas y devolución del trabajo final.
* Conclusiones y cierre.

1. **Estrategias metodológicas y Recursos didácticos**

**Estrategias metodológicas**

Las clases serán presenciales de modalidad teórico-práctica a dictarse en las aulas que cuentan con el equipo informático necesario. Se alternará entre una exposición participativa de los códigos utilizados en clase con la presentación de ejemplos. La práctica se basará en la resolución de actividades de manera individual y grupal con la posterior síntesis y puesta en común de los resultados. Se utilizará el *Campus* virtual INDEC en plataforma *Moodle* para consultas e intercambio de información; almacenamiento de los materiales teóricos y guía de ejercicios prácticos; actividades de evaluación por encuentro (cuestionarios de autoevaluación de resolución individual) y evaluación integradora final.

**Actividades de inicio o introductorias:**

Presentación del equipo docente, de los participantes y del curso (objetivos, contenidos, metodología y la evaluación). Aceptación de las condiciones de participación y su aprobación. Además, se presentará el “instructivo del *campus* virtual”.

**Actividades de desarrollo:**

Las actividades podrán comenzar con un repaso de los contenidos vistos la clase anterior y abordaje de dudas por parte del equipo docente, o con el planteo de una situación problemática para el abordaje de la temática y continuar con el desarrollo de los contenidos planificados para la clase. La misma será a través de la exposición dialogada dando uso del proyector multimedia para presentar los recursos didácticos preparados y seleccionados por el equipo docente. Y se espera que los participantes realicen las siguientes actividades de aprendizaje: análisis de casos, ejercicios, resolución de cuestionarios, entre otras, a ser resueltas de manera individual o grupal.

**Actividades de cierre o Integración:**

Para el cierre o integración del curso se realizará una evaluación final que integre los contenidos abordados en los módulos. Será resuelta de manera individual para ser corregida por el equipo docente. Dicho ejercicio consta de un análisis práctico con datos del instituto.

Recursos didácticos:

* Plataforma web del INDEC.
* Blogs, capítulos de libros, videos específicos.
* Códigos utilizados en clase.
* Proyector y pantalla

1. **Descripción de la modalidad**

Presencial.

1. **Bibliografía**

Buttu, M. (2016) “*El gran libro de Python”*. Editorial: Alfaomega Grupo Editor Argentino. Argentina.

Cane, A. (2019). “*Programación con Python”.* Editorial: Independently published. Argentina.

Ethan, W. (2019). “*Análisis de datos Python La guía definitiva para aprender A*”. Editorial: Independently published. Argentina.

Perez Castaño, A. (2016). “*Python fácil”*. Editorial: Alfaomega Grupo Editor Argentino. Argentina.

Raschka, V. y Mirjalili, S. (2019). “*Python machine learning”*. Editorial: Alfaomega Grupo Editor Argentino. Argentina.

Torres, J. (2020). “*Python deep learning”*. Editorial: Alfaomega Grupo Editor Argentino. Argentina.

PYTHON CRASH COURSE

1. **Evaluación de los aprendizajes**

Entendiendo que la evaluación es un proceso de recolección de información para emitir una valoración y tomar decisiones, destacamos dos tipos:

**Evaluación de proceso.** Los participantes realizarán actividades de aprendizajes de análisis de casos, ejercicios, resolución de cuestionarios, entre otras, a ser resueltas de manera individual o grupal. Actividades que tendrán una devolución por parte del docente la cual servirá de instrumento para la construcción del conocimiento. Ya que la orientación del curso es eminentemente práctica, los ejercicios siempre constarán de fragmentos de código capaz de ser ejecutados por el intérprete de Python el cuál devolverá un resultado o un error. Así podemos distinguir tres evaluaciones posibles para un ejercicio.

1. Satisfactorio: Es un ejercicio resuelto de manera tal que al ser ejecutado por Python arroje los resultados esperados.
2. Comprensión básica: En caso de que, al ejecutar el código, Python arroje un error o resultados incorrectos, el equipo docente asumirá la respuesta del participante como un pseudocódigo que puede expresar pasos correctos para resolver el ejercicio independientemente que no pueda ser ejecutado.
3. Incorrecto: Por último, si la ejecución de ese código no devuelve los resultados correctos, ni expresa entendimiento de los pasos para llegar a la solución, consideraremos que su respuesta es categóricamente errónea.

**Evaluación de producto.** Los participantes realizarán una evaluación final, que contemple los contenidos temáticos del curso. La misma será resuelta de manera individual para ser corregida por el equipo docente. Esta evaluación final consta de un análisis práctico con datos del instituto, en el que se integrarán los contenidos vistos en el curso. El participante deberá demostrar su conocimiento para:

* Leer de datos de las fuentes provistas.
* Explorar de los datos importados.
* Transformar de los datos con fines analíticos.
* Visualizar de los datos y relaciones entre sus variables.
* Elaborar de conclusiones en base al trabajo hecho.
* Comunicar efectivamente las conclusiones de las transformaciones y gráficos elaborados por sobre los datos originales.

Se aprobará la evaluación final según los siguientes criterios:

* Que demuestre haber utilizado de manera satisfactoria los contenidos vistos en el curso (30%).
* Que demuestre familiaridad con la herramienta (30%).
* Que haya llegado a la conclusión del análisis pretendido (40%).

1. **Instrumentos para la evaluación**

Se emplearán diversos instrumentos:

Instrumentos para la evaluación de los aprendizajes: se utilizará una “Grilla de Calificación y Asistencia”, donde se registrarán por cada participante las instancias de evaluación de proceso mencionadas anteriormente.

Para la evaluación final del curso, se administrará un instrumento de evaluación final individual del aprendizaje, consistente en la presentación de un proyecto puntual diseñado para que pueda aplicarse de manera colaborativa y que permita dar cuenta de la aplicación de los contenidos desarrollados en un caso de su entorno laboral.

Encuesta de satisfacción para la evaluación del curso. Se construirá de acuerdo con distintos aspectos pedagógico-didácticos: desempeño del/la docente, los contenidos temáticos, de desarrollo del curso. Así mismo se solicitará sugerencias y observaciones como propuesta de mejora. Se administrará al finalizar la cursada de forma anónima para conocer la opinión de los participantes.

1. **Requisitos de asistencia y de aprobación**

Los asistentes deberán acreditar:

* Asistencia mínima del 75% en las clases.
* Aprobación de la evaluación al finalizar el curso. Para aprobar el curso se debe obtener una nota mínima de siete (7) en una escala de 1 a 10.

1. **Duración de la actividad**

DOCE (12) horas

1. **Detalle de la duración**

Las clases se desarrollarán en SEIS (6) encuentros presenciales de DOS (2) horas de duración cada una.

1. **Cronograma y horario de realización**

Los encuentros serán distribuidos en clases los días lunes de 11:00 a 13:00 hs.

1. **Lugar de realización**

*Aula de capacitación. Av. Roca 609, piso 8.*

1. **Perfil del instructor**

Especialista en programación Python.

Economista especializado en el análisis del comercio internacional.

**Origen de la demanda**

Relevamiento de necesidades de capacitación 2023.