

GITHUB Token (bambam):

github_pat_11BOXC3SQ0J6XDiG6UfcSV_RP0bRT642zJOpYXqGVm0cFRzXxaLb2kb5S26Mun9QnIDD7ZTNLG
fo7OU7FU

https://books.adalab.es/materiales-modulo-0-d.a./Wqfv7ZYMdUmDap4wNqTK/recursos-adicionales/7-3-busqueda_de_soluciones

Objetivos generales

El objetivo principal de este curso es adquirir los recursos necesarios para ser autónomas en la extracción eficiente de información de bases de datos, al tiempo que adquieran conocimientos básicos sobre la creación, modificación y mantenimiento de bases de datos.

Además, conseguiremos la seguridad y comodidad al trabajar con Python para procesar, limpiar y extraer información relevante. También adquiriremos conocimientos para crear visualizaciones que complementen el storytelling de los proyectos. Asimismo, aprenderemos a utilizar el lenguaje técnico de manera fluida para una comunicación efectiva con otros analistas de datos.

Por último, nos podremos integrar fácilmente en equipos de trabajo siguiendo una filosofía ágil, utilizando marcos de trabajo como Scrum, y utilizando sistemas de control de versiones como Git.

Acerca del programa

Este repositorio recoge los materiales técnicos utilizados dentro del curso intensivo de Data Analytics impartido en Adalab. El curso está enfocado a mujeres sin conocimientos previos de programación.

La duración del programa es de 14 semanas. El bootcamp de análisis de datos consta de cuatro módulos diseñados para proporcionar a las estudiantes una base sólida en las habilidades y conceptos clave del análisis de datos. Empezaremos aprendiendo los fundamentos de Python, centrándonos en su aplicación en el análisis de datos. Además, aprenderemos a extraer datos de diversas fuentes utilizando Python y SQL, y adquiriremos las habilidades para limpiar y transformar datos antes de cargarlos en una estructura adecuada para el análisis. Por último aprendemos a como reportar de forma eficiente todos los datos utilizando herramientas de visualización especializadas.

Sobre Adalab

Adalab es una escuela especializada en formación digital para mujeres.

Trabajamos para formar y acompañar a mujeres que buscan un giro profesional adaptado a las nuevas necesidades digitales de las empresas.

A través de nuestro Curso Intensivo de Data Analytics las alumnas inician una carrera en tecnología, aportando diversidad al sector.

Visual Studio Code (también conocido como VS Code o Code a secas) es el editor de código con el que vamos desarrollar nuestras scripts y Jupyter notebooks haciendo uso del lenguaje de programación Python. Además dispone de un montón de extensiones y plugins que nos facilitan mucho la vida.

README

Si estás leyendo esto es porque has sido admitida en el bootcamp de Adalab y estás en el primer día de clases. ¡Qué emoción! 🎉

¡¡¡🎉 Enhorabuena futura analista de datos!!!

Estos son los materiales del curso de analista de datos de Adalab, están compuestos por cinco módulos:

Módulo 0 | ¡Hola, mundo!

En este módulo inicial, tendrás acceso a guías esenciales como Agile, Scrum y Pair Programming.

Prepararás tu equipo instalando las herramientas necesarias para el bootcamp y comenzarás con el pre-work, que incluye cursos sobre Jupyter Notebooks, Anaconda, la Terminal y la Línea de Comandos.

Además, te familiarizarás con los distintos roles en el campo de la analítica de datos para entender mejor las áreas y funciones del sector.

Módulo 1 | Python básico

En este módulo, aprenderás los fundamentos de Python y el control de versiones con Git. Te enfocarás en el desarrollar habilidades de programación y resolución de problemas utilizando Python.

Proyecto: Juegos Clásicos en Python Trabajaremos con Python para desarrollar la lógica de juegos clásicos como "Piedra, Papel o Tijera" y "Tic Tac Toe". Este proyecto nos permitirá familiarizarnos con Python, una herramienta clave para nuestro futuro como analistas de datos.

Módulo 2 | Extracción de datos (Python y SQL)

Aprenderás a extraer datos de diversas fuentes mediante técnicas de web scraping utilizando BeautifulSoup y Selenium. También comprenderás el funcionamiento de las APIs, incluyendo autenticación y autorización. Además, te familiarizarás con conceptos básicos de bases de datos y SQL, creando, consultando y gestionando datos en MySQL.

Proyecto MusicStream Explorarás cómo utilizar diversas herramientas y tecnologías para extraer información valiosa del mundo de la música, procesarla y almacenarla en una base de datos.

Módulo 3 | Transformando datos

Dominarás el uso de NumPy y Pandas para manipular y analizar datos estructurados. Crearás gráficos informativos con las librerías Matplotlib y Seaborn, diseñarán procesos ETL en Python. También aplicarás técnicas de análisis estadístico y realizarás pruebas A/B para tomar decisiones basadas en experimentación.

Proyecto: Optimización de Talento En el competitivo entorno empresarial actual, la retención de empleados y la satisfacción laboral son cruciales para el éxito. Desarrollarás un proyecto de análisis de datos y experimentación A/B con el objetivo de identificar factores clave que influyen en la satisfacción y retención de empleados. Presentaremos un análisis de datos, diseñarás un experimento A/B para probar hipótesis y proporcionarás información valiosa que apoye las decisiones estratégicas de la empresa.

Módulo 4 | Visualización de Datos con PowerBI y Tableau

En el último módulo, te harás con Tableau y Power BI, herramientas clave para la visualización de datos. Aprenderás a crear gráficos y paneles de control para comunicar datos de manera efectiva. También analizarás datos empresariales para extraer información valiosa y utilizarás estas herramientas para respaldar decisiones estratégicas basadas en datos.

Proyecto Data Insights: ETL y Visualización en Tableau/Power BI.

Aprenderás a explorar y limpiar datos, crear visualizaciones impactantes con Tableau o Power BI, y comunicar hallazgos efectivamente. Presentarás tus informes, mejorando tus habilidades en ETL y visualización.

PYTHON

Variable número entero = integers ; a = 3 int

Variable número decimal = float ; metros= 1.87 float

Variable texto = strings ; sacar el contenido de una variable = print()

Enunciado

print(a)

Variable "comida" que contenga "batata"

comida = "batata"

print(comida)

resultado: batata

PALABRA SIN COMILLAS EL ORDENADOR LO INTERPRETA COMO UNA VARIABLE Y BUSCARÁ EL CONTENIDO DE LA VARIABLE ENTRE TODAS LAS VARIABLES QUE HAN SIDO CREADAS EN ESTE JUPYTER

categoria1 = 'verde'

color_detalle = categoria1 + ' ' + 'oscuro'

print(color_detalle)

Resultado: verde oscuro

ó

print(categoria1 + ' oscuro')

Resultado: verde oscuro

ó

otra forma: separar valores con 1 coma (significa que pasamos a la función print dos argumentos distintos

print(categoria1, 'oscuro')

Resultado: verde oscuro

Hasta ahora hemos visto dos tipos de variables: números enteros, cadenas de caracteres o *strings*:

- Los números enteros en inglés se denominan **integer** ('int'), y son todos los números sin decimales tanto positivos como negativos.
- Los números decimales en inglés se denominan **float** ('float'), y son todos los números con decimales tanto positivos como negativos.
- Las cadenas de caracteres se llaman **string** ('str') en inglés y pueden delimitarse con comillas simples ' ' o dobles " " indistintamente.

C5 = int()

print(c5)

C5 = float()

print(c5)

C5 = int(b4 / b2)

print(c5)

C5 = float(b4 / b2)

print(c5)

```
C5 = str(b4 / b2)
print(c5)

type es cambiar el tipo de un dato
c5a = float(b4 / b2)
c5b = int(b4 / b2)
c5c = str(b4 / b2)
t = type(c5a)

print(t)
print(type(c5b))
print(type(c5c))
```

Para comprobar que una variable es del tipo que esperabas _: isisnstance()

Requiere dos argumentos, que se separan por una coma. Primero de la variable que quieres comprobar y segundo el tipo (float, int o str) - (decimal, números sin decimales, caracteres entre comillas) - la variable es sin comillas

ESTABLECER CONDICIONES:

Booleano (bool): si/no es true o false

NO A TODAS LAS FUNCIONES, LAS DIVISIONES NO ACEPTAN BOOL

Solo ejecutaremos un trozo de código si la condicion es verdadera, si es falso nos lo saltamos.

GitHub

Terminal — ls ("LS") (listar lo que tenemos dentro de la terminal en la que estamos)

— cd ("Change Directory") (nos movemos por la terminal, ponemos el nombre de donde queremos ir) cd modulo1 pair

cd .. (vamos un nivel hacia atrás)

Clonar un repositorio

git clone = copiar carpeta de git

MANTRA

git add . — decirle que eso ha cambiado en mi repositorio local, seguimiento que tenga en cuenta para subirlo

git commit — "foto" del momento como está el archivo

git push — empujo el repositorio a la nube

git pull — descargar repositorio de la nube