

"تحصيل وتنظيف البيانات: دليل سهل لتحسين تصنيف محركات البحث"

المدة: 5 يوم

اللغة: ar

كود الكورس: PI2-108

هدف الكورس

عند إتمام هذه الدورة، سيكون المشاركون قادرين على:

- فهم كيفية الحصول على البيانات من منصات وقواعد بيانات مختلفة.
- تعلم تقنيات تنظيف وإعداد البيانات للتحليل.
- تطوير مهارات في تحديد ومعالجة البيانات المفقودة وغير المتسقة والمكررة.
- اكتساب الخبرة في استخدام أدوات مثل Excel وPython وR لتنظيف البيانات.
- تعلم تحويل البيانات الخام إلى صيغ منظمة ودقيقة وجاهزة للاستخدام.
- تطبيق أفضل الممارسات في إعداد البيانات لضمان مجموعات بيانات عالية الجودة وموثوقة.

الجمهور

هذه الدورة مخصصة لـ

- **محللي البيانات وعلماء البيانات** : المهنيين المسؤولين عن جمع وتنظيف البيانات للتحضير للتحليل وإعداد التقارير.
- **محترفي ذكاء الأعمال** : الأفراد الذين يعملون مع البيانات لدعم اتخاذ القرارات ويحتاجون إلى بيانات نظيفة ومنظمة للحصول على رؤى دقيقة.
- **الباحثين** : الذين يجمعون البيانات من الاستبيانات أو التجارب أو طرق البحث الأخرى ويجب عليهم ضمان دقتها وموثوقيتها.
- **محترفي تكنولوجيا المعلومات ومديري قواعد البيانات** : الأفراد المكلفين بدمج وصيانة وإدارة مجموعات البيانات الكبيرة في بيئات المؤسسات.
- **أي شخص يعمل مع البيانات** : الأفراد الذين يتعاملون مع البيانات يومياً ويرغبون في تحسين مهاراتهم في إعداد وتنظيف البيانات.

منهجية التدريب

يجمع هذا الدورة بين الرؤى النظرية والتجربة العملية لتزويد المشاركين بالمهارات الأساسية في جمع البيانات وتنقيتها. من خلال محاضرات يقودها مدربون، وتمارين عملية باستخدام مجموعات بيانات حقيقية، ومناقشات جماعية تفاعلية، سيستكشف المشاركون تقنيات جمع البيانات وتنقيتها وتحويلها. ستوفر الدروس التعليمية المعتمدة على الأدوات في إكسل وبايثون وR تدريبًا عمليًا، بينما ستتحدى دراسات الحالة المشاركين لتطوير حلول لمشكلات البيانات الواقعية. ستعزز التقييمات المنتظمة والتغذية الراجعة عملية التعلم وتضمن فهمًا شاملاً للمفاهيم الأساسية.

الملخص

تُعَد البيانات حجر الزاوية لاتخاذ القرارات المستنيرة في عالم اليوم الذي يعتمد على البيانات، ولكن غالبًا ما تكون البيانات الخام غير منظمة أو ناقصة أو غير مرتبة. تم تصميم دورة "الحصول على البيانات وتنظيفها" لمساعدة المهنيين على فهم العمليات الأساسية لجمع البيانات وتحضيرها وتنظيفها لضمان دقتها واكتمالها وجاهزيتها للتحليل. سواء كنت تعمل مع مجموعات بيانات كبيرة أو صغيرة ومركزة، فإن الحصول على بيانات نظيفة وموثوقة أمر حاسم لضمان استنتاجات ذات مغزى ونتائج دقيقة.

توفر هذه الدورة للمشاركين المعرفة والمهارات اللازمة للحصول على البيانات بفعالية من منصات متنوعة، بما في ذلك قواعد البيانات وواجهات برمجة التطبيقات واستخلاص البيانات من الويب والملفات المسطحة مثل CSV. أو Excel. سيتعلم المشاركون كيفية التعامل مع البيانات المفقودة، والتنسيقات غير الصحيحة، والتكرارات، والقيم الشاذة – وهي تحديات شائعة عند العمل مع البيانات الخام. بالإضافة إلى ذلك، ستغطي الدورة أفضل الممارسات في تنظيف البيانات وتحضيرها، بما في ذلك استخدام أدوات مثل Excel وSQL وPython وR.

بنهاية الدورة، سيكون المشاركون قادرين على جمع البيانات وتنظيفها بكفاءة، وتحويل البيانات الخام إلى صيغة قابلة للاستخدام يمكن الوثوق بها للتحليل. هذه الدورة مثالية لمحللي البيانات والباحثين والمتخصصين في استخبارات الأعمال وأي شخص يعمل مع مجموعات بيانات كبيرة أو معقدة.

محتوى الكورس والمخطط الزمني

Section 1: Introduction to Data Sourcing and Cleaning

- Overview of the Data Lifecycle: From Collection to Cleaning
- Why Clean Data Matters: The Importance of Accuracy and Consistency
- Common Data Issues: Missing Values, Duplicates, Outliers, and Inconsistencies

Section 2: Sourcing Data from Different Platforms

- Collecting Data from Databases Using SQL
- Extracting Data from APIs and Web Scraping Techniques
- Importing and Exporting Data from Excel, CSV, and Flat Files
- Handling Real-Time Data Streams and Integrating Data Sources

Section 3: Data Cleaning Fundamentals

- Handling Missing Data: Techniques for Imputation and Removal
- Identifying and Removing Duplicates
- Dealing with Outliers and Incorrect Formats
- Standardising Data for Consistency

Section 4: Data Transformation and Preparation

- Converting Unstructured Data into Structured Formats
- Transforming Data Using Excel Functions, Python, and R
- Aggregating, Merging, and Joining Datasets
- Normalising and Scaling Data for Analysis

Section 5: Tools and Techniques for Data Cleaning

- Cleaning Data in Excel: Functions and Tools for Data Validation
- Data Cleaning with Python: Pandas and Numpy Libraries
- Using R for Data Cleaning: dplyr and tidyr Libraries
- Automating Data Cleaning Tasks with Scripts and Macros

Section 6: Best Practices in Data Cleaning

- Creating Data Dictionaries and Documentation
- Ensuring Data Quality with Validation Rules
- Maintaining Data Integrity Throughout the Cleaning Process
- Continuous Monitoring and Iterative Data Cleaning

Section 7: ase Studies and Practical Applications

- Real-world Data Cleaning Examples from Various Industries
- Solving Complex Data Cleaning Challenges
- Applying Data Cleaning Techniques to Your Organisation’s Data

تفاصيل الشهادة

عند إتمام هذه الدورة التدريبية بنجاح، سيحصل المشاركون على شهادة إتمام التدريب من Holistique Training. وبالنسبة للذين يحضرون ويكملون الدورة التدريبية عبر الإنترنت، سيتم تزويدهم بشهادة إلكترونية (e-Certificate) من Holistique Training.

شهادات Holistique Training معتمدة من المجلس البريطاني للتقييم (BAC) وخدمة اعتماد التطوير المهني المستمر (CPD)، كما أنها معتمدة وفق معايير ISO 9001 وISO 21001 وISO 29993.

لهذه الدورة من خلال شهاداتنا، وستظهر هذه النقاط على شهادة إتمام (CPD) يتم منح نقاط التطوير المهني المستمر التدريب من Holistique Training. ووفقًا لمعايير خدمة اعتماد CPD، يتم منح نقطة واحدة عن كل ساعة CPD حضور في الدورة. ويمكن المطالبة بحد أقصى قدره 50 نقطة CPD لأي دورة واحدة نقدمها حاليًا.

التصنيفات

الذكاء الاصطناعي وإدارة البيانات, تطبيقات تكنولوجيا المعلومات والكمبيوتر, التكنولوجيا

مقالات ذات صلة



Data Cleaning Best Practices: Unlocking The Power Of Accurate Data

Learn the importance of data cleaning and its benefits in enabling accurate decision-making. Discover best practices to implement an effective data cleansing strategy and overcome common data issues.