

Kod szkolenia: **SAS/DATAPROC**

Tytuł szkolenia: **Przetwarzanie danych w SAS**

Data Processing in SAS

Dni: 4

Opis:

Adresaci szkolenia

Szkolenie jest adresowane do osób chcących rozpocząć swoją przygodę z SAS, jak również osób posiadających już swoje pierwsze doświadczenia z tym narzędziem, które są zainteresowane usystematyzowaniem oraz wzbogaceniem posiadanej wiedzy w zakresie wykorzystania 4GL, SQL oraz makroprogramowania.

Cel szkolenia

Uczestnicy szkolenia opanują umiejętność wykorzystania SAS do analizy i przetwarzania danych zajmując się rozwiązywaniem praktycznych zagadnień z obszaru inżynierii cech (feature engineering), selekcji cech (feature selection), czyszczenia danych (data cleaning) oraz wstępnej analizy danych.

W szczególności:

- opanują składnię języka przetwarzania danych SAS (SAS 4GL).
- poznają podstawy implementacji SQL dostępnej w SAS.
- nauczą się tworzyć elastyczny kod z wykorzystaniem makroprogramowania.

Mocne strony szkolenia

Szkolenie obejmuje swoim zakresem spektrum najważniejszych, najczęściej wykorzystywanych zagadnień związanych z przetwarzaniem danych w SAS. Ich zrozumienie pozwoli uczestnikowi na rozpoczęcie samodzielnej, efektywnej pracy w tym środowisku.

Nauka prowadzona się będzie w oparciu o przykłady oraz zadania warsztatowe bazujące na dobrych praktykach związanych z przetwarzaniem oraz poznawaniem danych.

Wymagania

Ze szkoleniem nie wiążą się żadne specjalne wymagania. Natomiast znajomość podstaw pracy z danymi lub programowania może ułatwić zrozumienie treści przekazywanych podczas

szkolenia.

Specjalne wymagania techniczne

Ze szkoleniem nie wiążą się żadne specjalne wymagania techniczne.

Parametry szkolenia

4 * 8 godzin (4 * 7 netto) wykładów i warsztatów. Przewaga warsztatów skupionych na pracy z danymi.

Wielkość grupy maksymalnie 8-10 osób.

Program szkolenia:

1. Wprowadzenie do SAS.
 - Charakterystyka systemu SAS.
 - Składowanie danych w SAS.
 - Filozofia programowania w SAS.
2. Podstawy programowania w SAS 4GL.
 - DATA STEP - podstawowa składnia.
 - Praca z operatorami dostępnymi w DATA STEP.
 - Program Data Vector (PDV).
 - Selekcja zmiennych - instrukcje KEEP i DROP.
 - Zmiana nazwy zmiennej - instrukcja RENAME.
 - Selekcja obserwacji - instrukcje WHERE i IF.
 - Przegląd funkcji dostępnych w DATA STEP.
 - PROC STEP - podstawowa składnia.
 - Diagnostyka zbioru danych - procedury PRINT i CONTENTS.
 - Sortowanie zbioru danych - procedura SORT.
3. Zaawansowane aspekty pracy z DATA STEP.
 - Wybrane opcje instrukcji SET.
 - Zmienne opóźnione - funkcje LAG i DIF.
 - Zmienne trwałe w DATA STEP - instrukcja RETAIN.
 - Konwersja i wizualizacja danych w SAS - formaty i informaty.
 - Instrukcje warunkowe IF-THEN-ELSE oraz SELECT-WHEN-OTHERWISE.
 - Instrukcje pętli DO, DO-WHILE, DO-UNTIL.
 - Zmienne tablicowe - instrukcja ARRAY.
 - Agregowanie danych - instrukcja BY.
 - Praca z zewnętrznymi źródłami danych - instrukcje INPUT i INFILE.
4. PROC STEP - praca z wybranymi procedurami SAS.
 - Transponowanie zbiorów danych - procedura TRANSPOSE.
 - Operacje na zbiorach danych - procedury APPEND, COPY, DELETE i COMPARE.
 - Wstępne analizowanie danych - procedury FREQ, MEANS i SUMMARY.
5. Łączenie zbiorów danych w DATA STEP.

- Pionowe łączenie tabel - instrukcja SET.
 - Poziome łączenie tabel - instrukcja MERGE.
6. Wybrane zagadnienia z SQL w SAS.
- Podstawy przetwarzania danych w SQL - instrukcje SELECT i CREATE.
 - Specyficzne aspekty implementacji SQL w SAS.
 - Łączenie zbiorów danych w SQL.
7. Makroprogramowanie.
- Makroprogramowanie, a SAS 4GL.
 - Uogólnianie kodu z wykorzystaniem makrozmiennych.
 - Tworzenie własnych makroprocedur.
 - Makroprogramowanie z wykorzystaniem wbudowanych funkcji.
 - Instrukcja warunkowa %if-%then-%else.
 - Pętle %do, %do-%while, %do-%until.

