

## OEC020

### La Data Science con Excel

Durata: 3 gg

#### Descrizione

A giudizio pressoché unanime della comunità internazionale, Excel non è la prima scelta per la Data Science, tuttavia - se i partecipanti al corso hanno già consolidate esperienze in Excel base - esso può essere usato come un possibile step intermedio verso una più matura e completa adozione di una vera piattaforma di Data Science (come R). Sebbene infatti temi come Machine Learning , statistica, big data e grafica avanzata non siano seriamente e pienamente affrontabili con Excel, tuttavia alcune interessanti analisi (sopraindicate) possono essere condotte in Excel (2013+).

#### Prerequisiti

Il corso introduttivo sulla Data Science, una conoscenza di base di Excel

#### Contenuti

- Cos'è un modello.
- Outlier e dati mancanti.
- Modelli predittivi con Excel (regressione lineare e classificazione binaria).
- Creazione e valutazione dei modelli predittivi basati sui dati aziendali.
- Teoria dell'informazione e misurazione dell'entropia.
- Principali distribuzioni di probabilità degli output dei modelli.
- Overfitting e variabili omesse (i due principali problemi).
- Risolutore, Tabelle Dati, Analisi degli Scenari e Ricerca Obiettivo: a cosa servono e come si usano;

esempi.

- Analisi di sensitività.
- Simulazione di Montecarlo.
- Data Visualization ed Infografica in Excel: concetti, grafici avanzati e dashboard dinamiche.
- Punti di forza e limiti di Excel per la Data Science.
- Nuovi strumenti Excel per la Data Science: DB colonnare in-memory (Power Pivot), tool ETL da molti sorgenti incluso Hadoop (Power Query), data visualization (Power View e Power Map), add-in di Data Mining come front-end per SQL Server Analysis Services (SSAS, la componente SQL Server per la Data Analysis). [Visual Studio permette di accedere SSAS senza Excel ma è poco user-friendly]. VBA per Reti Neurali e simulazione di Monte Carlo.
- Machine Learning avanzato con Excel base (cenni).

Case study finali:

- Accettazione o rifiuto di una richiesta di carta di credito o di prestito;
- Costruzione di un modello per minimizzare il rischio di default ed un modello per massimizzare il profitto.

Focus sulla matematica, non su Excel in sé. Solo template Excel (no Power Pivot, VBA, macro, tabelle Pivot o altre funzionalità avanzate di excel)