

Capitolo 21: Importazione e pulizia dei dati

1. Microsoft Excel - Importazione dei Dati

a. Importazione da un File

Excel consente di importare dati da una vasta gamma di fonti, sia file locali che servizi online. La funzione 'Ottieni e trasforma' (Power Query) facilita l'importazione e la trasformazione dei dati.

Percorso: 'Dati' > 'Ottieni dati' > 'Da file' > scegli il tipo di file.

b. Formati di File Supportati

i. Formati di file di fogli di calcolo

- .xlsx, .xls, .ods: supportano dati strutturati, formule, formattazioni.

ii. Formati di file di database

- .mdb, .accdb, .dbf, .sql: dati provenienti da database locali o server SQL.

iii. Formati di file di testo

- .csv, .txt, .prn: delimitati da virgole, tabulazioni o spazi.

iv. File HTML

- .html, .htm: importazione di tabelle contenute in pagine web.

v. File XML

- .xml: strutturati con tag che identificano i dati e i metadati.

vi. Immagini

- .jpg, .png, .bmp, .gif: non importabili come dati, ma inseribili come oggetti visivi.

c. Differenza tra Importare e Aprire

Aprire un file significa caricarlo direttamente come foglio Excel. Importare consente invece di prelevare dati da un'altra fonte e combinarli, trasformarli o aggiornarli dinamicamente tramite Power Query.

d. Importazione di un File di Testo

Percorso: 'Dati' > 'Da file di testo/CSV' > seleziona il file > scegli delimitatore > Carica o Trasforma.

- Opzioni avanzate: scelta della codifica, tipo di dati per ogni colonna, filtri preliminari.

e. Copiare e Incollare Dati

È possibile copiare dati da fonti esterne (browser, PDF, Word, ecc.) e incollarli in Excel:

- CTRL + C (copia) e CTRL + V (incolla)
- Incolla speciale: valori, formule, formattazioni, collegamenti dinamici

Esempi pratici

Esempio 1:

Importa un file CSV con l'elenco dei clienti e trasforma i nomi in maiuscolo con Power Query.

Esempio 2:

Importa un file XML con dati anagrafici e visualizza il contenuto in tabella Excel.

Esempio 3:

Copia da un sito web una tabella HTML e incollala direttamente in Excel, mantenendo la formattazione.

Esempio 4:

Importa una tabella da un database Access (.accdb) e crea una connessione per l'aggiornamento dati.

Esempio 5:

Inserisci un'immagine PNG come logo nel report e posizionala accanto alla tabella dati.

Esercizi

Esercizio 1:

Importa un file CSV con almeno 5 colonne e rimuovi le righe duplicate con Power Query.

Esercizio 2:

Apri un file Excel .ods e confronta i dati con un file .xlsx dello stesso contenuto.

Esercizio 3:

Copia una tabella da Word e incollala in Excel in formato solo valori.

Esercizio 4:

Importa un file di testo delimitato da tabulazioni e modifica il tipo dati della colonna 'Data'.

Esercizio 5:

Collega una tabella da un database SQL locale (se disponibile) e filtra solo i clienti attivi.

2. Microsoft Excel - Pulizia dei Dati

a. Rimozione di Righe Duplicati

Percorso: 'Dati' > 'Rimuovi duplicati'. Seleziona le colonne da analizzare per eliminare righe duplicate.

b. Identificazione dei Duplicati

Usa la formattazione condizionale: 'Home' > 'Formattazione condizionale' > 'Regole evidenziazione celle' > 'Valori duplicati'.

c. Divisione del Testo

Utilizza 'Testo in colonne': 'Dati' > 'Testo in colonne' > scegli delimitatore (es. virgola, spazio).

d. Utilizzo di Riempimento Automatico (Flash Fill)

Scrivi un pattern nella colonna accanto, poi premi CTRL + E oppure 'Dati' > 'Riempimento lampo'.

e. Modifica del Maiuscolo/Minuscolo

Funzioni:

- =MAIUSC (UPPER) – Tutto maiuscolo
- =MINUSC (LOWER) – Tutto minuscolo
- =MAIUSC.INIZ (PROPER) – Prima lettera maiuscola

f. Rimozione di Spazi Extra

Funzione: =ANNULLA.SPAZI (TRIM) – Rimuove gli spazi iniziali, finali e multipli interni.

g. Rimozione di Caratteri Strani

Funzione: =PULISCI (CLEAN) – Rimuove caratteri non stampabili.

h. Conversione e Classificazione Valori

- Conversione: =VALORE (VALUE) converte testo numerico in numero.
- Classificazione: usa =SE, =CERCA.VERT o Power Query per etichettare i dati.

i. Unione e Riordino Colonne

- Unione: =A2&B2 o =TESTO.UNISCI(" "; VERO; A2; B2)
- Riordino: seleziona colonna, usa 'Taglia' (CTRL+X) e 'Inserisci'.

j. Randomizzazione e Conversione Verticale/Orizzontale

- Randomizza: inserisci =CASUALE() e ordina i dati.
- Trasposizione: copia > incolla speciale > 'Trasponi'.

k. Estrazione e Confronto di Testo

- Estrai nome file da URL: =STRINGA.ESTRAI(A1; TROVA("/";A1;TROVA("/";A1)+1)+1; LUNGHEZZA(A1))
- Verifica presenza in elenco: =CONTA.SE(Intervallo; Valore)>0

l. Compilazione dei Vuoti in un Report

- Usa Power Query o riempi celle vuote con valore della cella sopra (con formule o GoTo Special).

m. Controllo Ortografico e Sostituzione Testo

- Ortografia: 'Revisione' > 'Controllo ortografia'
- Trova/Sostituisci: CTRL+T per 'Trova', CTRL+S per 'Sostituisci'

n. Aggiunta Testo e Correzione del Segno Negativo

- Aggiunta testo: ="Fattura: "&A2
- Correzione negativo a fine stringa: =-VALORE(SOSTITUISCI(A1;"-";""))

o. Checklist per la Pulizia Dati

1. Rimuovere duplicati
2. Controllare dati nulli

3. Uniformare il formato del testo
4. Correggere errori ortografici
5. Standardizzare valori e codifiche
6. Validare risultati con formule

Esempi pratici

Esempio 1:

Descrizione dell'esempio pratico 1 relativo alla pulizia dati in Excel.

Esempio 2:

Descrizione dell'esempio pratico 2 relativo alla pulizia dati in Excel.

Esempio 3:

Descrizione dell'esempio pratico 3 relativo alla pulizia dati in Excel.

Esempio 4:

Descrizione dell'esempio pratico 4 relativo alla pulizia dati in Excel.

Esempio 5:

Descrizione dell'esempio pratico 5 relativo alla pulizia dati in Excel.

Esercizi

Esercizio 1:

Descrizione dell'esercizio applicativo 1 per mettere in pratica la pulizia dei dati.

Esercizio 2:

Descrizione dell'esercizio applicativo 2 per mettere in pratica la pulizia dei dati.

Esercizio 3:

Descrizione dell'esercizio applicativo 3 per mettere in pratica la pulizia dei dati.

Esercizio 4:

Descrizione dell'esercizio applicativo 4 per mettere in pratica la pulizia dei dati.

Esercizio 5:

Descrizione dell'esercizio applicativo 5 per mettere in pratica la pulizia dei dati.

3. Microsoft Excel - Esportazione dei Dati

a. Esportazione in un File di Testo

Excel consente l'esportazione dei dati in formati di testo semplici, utili per l'interoperabilità con altri software. Percorso: 'File' > 'Salva con nome' > scegli il formato dal menu a discesa 'Tipo file'.

b. File CSV

Formato: Valori separati da virgole.

- Estensione: .csv
- Utilizzo: compatibile con molti database, strumenti di importazione e sistemi online.
- Attenzione: i campi contenenti virgole devono essere racchiusi tra virgolette.

c. File TXT

Formato: file di testo con dati separati da tabulazioni o spazi.

- Estensione: .txt
- Maggiore controllo su delimitatori personalizzati durante l'importazione.

d. File PRN

Formato: larghezza fissa per colonna.

- Estensione: .prn
- Ogni colonna ha una larghezza predefinita, ideale per stampa su sistemi legacy o terminali.

e. Esportazione in Altri Formati

Excel supporta l'esportazione in vari altri formati:

- PDF: tramite 'File' > 'Esporta' > 'Crea PDF/XPS'.
- HTML: salva come 'Pagina Web (.htm; .html)'.
- XML: per dati strutturati con tag personalizzati.
- XLSX/XLS: per compatibilità con versioni diverse di Excel.
- ODS: formato OpenDocument compatibile con LibreOffice e OpenOffice.

Esempi pratici

Esempio 1:

Esporta una tabella di fatture in formato CSV per inviarla al reparto contabilità.

Esempio 2:

Salva una tabella con descrizioni prodotto in formato TXT con separatore di tabulazione.

Esempio 3:

Esporta un report mensile in formato PDF per invio via email.

Esempio 4:

Genera un file PRN per importazione in un vecchio sistema gestionale.

Esempio 5:

Salva un foglio Excel come HTML per pubblicarlo su intranet aziendale.

Esercizi

Esercizio 1:

Esporta un elenco clienti in formato CSV, verificando che i nomi con virgole siano tra virgolette.

Esercizio 2:

Crea un file TXT con tabulazioni da una tabella Excel e aprilo in un editor di testo per controllare la struttura.

Esercizio 3:

Esporta un report sintetico in PDF con intestazioni e piè di pagina visibili.

Esercizio 4:

Salva un file in formato XML e aprilo per analizzare la struttura dei tag.

Esercizio 5:

Confronta un file CSV e un file TXT esportati dallo stesso foglio, analizzando le differenze nei delimitatori.