

# Minicorso

## Word-processor, database, spreadsheet

di Andrea Saviano

### Parte 1

- Divieto d'Access ai non addetti ai lavori, premessa
- La gestione del reclamo... e io pago!
- Ta brutte e ta... belle
- Per centrare il bersaglio, più che Robin Hood... un little join
- 
- 
- 
- 

### Premessa



Il pacchetto **Ms Office** (ma anche Open Office, tanto non sponsorizzare Bill Gates), nonostante tutti i suoi limiti, può rivelarsi un interessante strumento di automazione d'ufficio, ad esempio: per l'archiviazione e l'analisi di un Sistema Gestione Qualità.

- **Ms Word** è un utile editore di testi in grado di generare dei modelli di documento univoci.
- **Ms Excel** permette di elaborare dati numerici sintetizzandoli in sofisticate analisi statistiche.
- **Ms Access** è un discreto database che si trova di frequente nelle aziende.

### Divieto d'Access ai non addetti ai lavori?

Capita spesso, quando si avvia Ms Access, di avere la strana impressione che non succeda nulla. La maggior parte delle persone, posta davanti a quella maschera bianca caratterizzata da pochi e misteriosi tasti, ha una frustrante sensazione d'impotenza, come se si trovasse davanti ad uno strumento difficilissimo da utilizzare.

In realtà è sufficiente un po' di pazienza (forse ben più che un po', lo ammetto) e d'interesse per i database per realizzare applicativi che sembrano realizzati da esperti professionisti.

Più in generale, la mia intenzione è quella di mostrare l'utilità del pacchetto Ms Office e in particolare di Ms Access per gestire professionalmente le informazioni cruciali di un'azienda.



La cosa più stupefacente di Ms Access è che, fino a certi livelli, non è strettamente necessario conoscere nessun linguaggio di programmazione e, grazie alla strutturazione di una base dati, Excel è in grado di risolvere agevolmente il suo compito di analisi e di sintesi...

### Database relazionali

Ms Access è un database relazionale. Ora, un database relazionale si presenta in forma tabellare in cui le righe rappresentano i record (dati) e le colonne rappresentano i campi (argomenti). Le operazioni di correlazione che si realizzano con questo tipo di database riguardano le tabelle e non i record individuali.

La norma fondamentale per stabilire relazioni tra tabelle, cioè tra contenitori di dati correlabili, è che il campo di collegamento non deve avere ripetizioni, ossia ogni record deve potere essere identificato in maniera univoca.

Il campo che permette l'identificazione di ogni record è detto **chiave primaria** e deve essere comune alle tabelle che s'intendono correlare. Ne deriva che, per concepire un database, sia indispensabile strutturarlo in modo che almeno un campo funga da chiave primaria (solitamente indicato con il prefisso ID davanti al nome campo).

I dati di una tabella per la quale è definita una chiave primaria sono ordinati automaticamente in base al valore di questo puntatore. Nonostante la chiave primaria si presenti come un indice (chiave di ricerca), occorre tuttavia premettere che in una medesima tabella possono essere presenti più indici, ma solo uno di loro può fungere da chiave primaria.

In particolare, la relazione tra due tabelle si basa sui concetti di chiave primaria e **chiave esterna**, dove la chiave primaria è un campo di una tabella A (che chiameremo **tabella primaria**) che contiene valori non ripetuti che identificano in maniera univoca ciascun record e la chiave esterna è un campo di una tabella B (che chiameremo **tabella secondaria**) che ha lo stesso tipo di dati della chiave primaria e che viene utilizzato per legare i dati della tabella B a quelli della tabella A.

### Speriamo non siano dati... per dispersi

Aperto una tabella in **modalità struttura dati** si visualizza la caratteristica dei campi che formano i record e si può notare come i **formati possibili per dati** siano:

- **testo**, una stringa con non più di 255 caratteri
- **memo**, un testo libero di non più di 255 caratteri
- **numerico**, un numero:
  - byte,
  - intero,
  - intero lungo,
  - precisione singola,
  - precisione doppia,
  - ID replica,
  - decimale;
- **data/ora**, un valore che può essere formattato per visualizzare data, ora o entrambi, i formati possono essere:
  - data generica,
  - data estesa,
  - data breve,
  - data in cifre,
  - ora estesa,
  - ora breve 12h,
  - ora breve 24h;
- **valuta**, un numero decimale con formato particolare per rappresentare importi in denaro:
  - numero generico,
  - valuta,
  - valuta predefinita (euro),
  - fisso,
  - standard,
  - percentuale,
  - nozione scientifica;

- **contatore**, un valore numerico particolare, utilizzato per individuare in modo univoco i record di una tabella:
  - intero lungo,
  - ID replica;
- **si/no**, un valore logico del tipo:
  - VERO/FALSO,
  - sì/no,
  - ON/OFF
- **oggetto OLE**, un elemento generato da un'altra applicazione software (ad esempio: un'immagine, un filmato, una traccia audio), associato al campo, ma residente in file separato;
- **collegamento ipertestuale**, una stringa di caratteri che individua un elemento interno o esterno collegato al record.

I campi memo o oggetto OLE non contengono dati effettivi, ma solamente un riferimento al documento dove l'informazione è contenuta.

### Prerogative dei campi memo

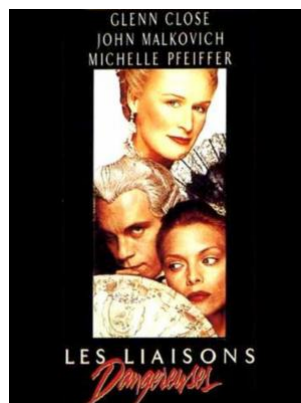
Rispetto ai campi testo, i campi memo sono molto più comodi per gestire delle informazioni lunghe e non strutturate all'interno di un record. Infatti, l'occupazione di un campo memo è molto ridotta, poiché il contenuto del campo viene salvato esternamente al data-base. Un altro indubbio vantaggio nella scrittura di note è che nei campi memo si usufruisce delle tipiche funzioni di un editor (taglia, copia e incolla).

### Indici

Poiché ogni indice può essere associato fino a 10 campi differenti, se si esamina la struttura di una tabella è possibile visualizzare in dettaglio ciò tramite una finestra che esamina:

- i **nomi dei campi associati** all'indice;
- il **criterio d'ordinamento** (crescente/decescente);
- le **proprietà** che caratterizzano l'indice:
  - **primario**, se si seleziona sì, quell'indice diventa una **chiave primaria**,
  - **univoco**, se si seleziona sì, ciascun valore dell'indice deve essere univoco,
  - **ignora nul**, se si seleziona sì, i record con valore **nul**<sup>1</sup> sono esclusi dall'indice.

### Relazioni... pericolose



Premesso che sono gli indici poi a gestire le relazioni, queste possono essere:

- **uno a uno** (1 a 1): a ciascun record di una tabella è associato un solo record di un'altra tabella;
- **uno a molti** (1 a  $\infty$ ): a ciascun record di una tabella sono associati più record di un'altra tabella;
- **molti a molti** ( $\infty$  a  $\infty$ ): a più record di una tabella sono associati più record di un'altra tabella.

Queste relazioni sono evidenziate nella finestra relazioni dai simboli 1 e  $\infty$ .

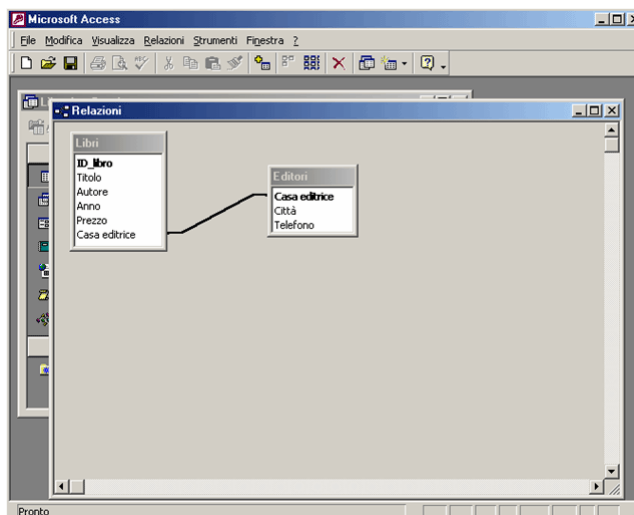
Per modificare il tipo di relazione basta selezionarla con un doppio click.

All'interno della finestra "modifica relazioni" si notano tre opzioni:

---

<sup>1</sup> Per valore **nul** s'intende un campo lasciato vuoto.

- **applica integrità referenziale**, se selezionata impone un vincolo di coerenza<sup>2</sup> tra due tabelle;
- **aggiorna campi correlati a catena**, dopo aver attivato l'opzione **applica integrità referenziale** è possibile, se lo si desidera, attivare anche la possibilità di aggiorna campi correlati a catena, selezionando tale opzione si fa in modo che una modifica effettuata nella tabella principale venga automaticamente riportata anche nella tabella correlata;
- **elimina record correlati a catena**, selezionando questa casella di controllo si fa in modo che l'eliminazione di un record dalla tabella principale elimini tutti i record correlati dalla tabella secondaria.



Per eliminare una relazione è sufficiente posizionarsi all'interno della **finestra relazioni**, su la relazione che unisce le due tabelle e utilizzare il tasto **Canc**.

### La gestione del reclamo... e io pago!

Per spiegare come utilizzare Ms Access partiamo dal primo cardine di ogni Sistema Gestione Qualità: il reclamo.



A prescindere dalla realtà aziendale che ci sta dietro, qualsiasi reclamo è sempre riportabile ad un **documento ufficiale** tra fornitore e cliente, ad esempio:

- **OFF** = richiesta d'offerta,
- **PRE** = preventivo,
- **ORD** = ordine,
- **DIS** = disegno,
- **FAT** = fattura,
- **DDT** = documento di carico e trasporto,

<sup>2</sup> Non è possibile eliminare un record nella tabella **1** senza eliminare tutti i record che sono collegati nella tabella  $\infty$ . In questo modo si mantengono integri i riferimenti. Il vincolo dell'integrità referenziale vale anche per la creazione di record: non si possono aggiungere record ad una tabella  $\infty$  se non associandoli ad un record esistente nella tabella **1**. Se non esiste, deve essere creato prima di aggiungere record nella tabella  $\infty$ .

- **NdC** = nota di credito,
- **RdC** = reclamo da cliente.

Ne deriva che, prima di strutturare il database, occorre esaminare quali siano in azienda i documenti ufficiali<sup>3</sup>.

Proseguendo, nella gestione di un reclamo si riconoscono sempre alcuni elementi comuni:

- data di apertura del reclamo;
- tipologia e n° di protocollo del documento attivo<sup>4</sup>;
- motivo del reclamo e dettagli sulle cause del reclamo;
- esito del reclamo<sup>5</sup> e avvio o meno di una non conformità;
- data di chiusura del reclamo<sup>6</sup>;
- costi sostenuti (azioni di tamponamento come: sconti, note di credito, fatture d'addebito);

pochi elementi, ma cruciali per permettere:

- tracciabilità,
- sintesi,
- analisi,

insomma, quei semplici tre elementi che dovrebbero caratterizzare qualsiasi Sistema Gestione Qualità.

### **Per gestire il reclamo... le do carta bianca**

Se si lavora in un'azienda in cui c'è un numero molto (ma molto) limitato di reclami<sup>7</sup>, anche un semplice modulo<sup>8</sup> Ms Word può bastare.



Se invece si lavora in un'azienda normale, che vuole comprendere e risolvere le problematiche (e i costi) connessi ai reclami, è necessario:

- **organizzare i dati** in più tabelle in modo che possano essere richiamati ed utilizzati efficacemente;
- **mettere in corrispondenza i dati** in modo da produrre statistiche e analisi;
- **condividere i dati** con gli altri utenti della rete aziendale;
- **eliminare la carta**, riducendo i tempi di lavoro.

A questo punto, appare evidente la necessità di ricorrere ad un database.

Ora, un database relazionale non è altro che un insieme di dati contenuti in una o più tabelle, le quali sono poi poste in interconnessione tra loro. Si comprende quindi che il numero delle tabelle da creare dipende essenzialmente da come s'intendono organizzare i dati.

Come dimostrato, appare evidente che la gestione del reclamo si possa suddividere in informazioni inerenti:

- la registrazione del reclamo e dei costi sostenuti;

<sup>3</sup> Solitamente... nessuno, perché alla gente non piace dover protocollare e archiviare i documenti, dopotutto ognuno si trova alla perfezione nel proprio disordine.

<sup>4</sup> Identifica in modo univoco il cliente e il prodotto/servizio implicato.

<sup>5</sup> In un Sistema Gestione Qualità c'è un filtro passa/non-passa per il reclamo ed è la sua ammissibilità, poiché se un Reclamo è ammissibile allora, automaticamente, si ha una Non Conformità accertata che è legata al Reclamo secondo la logica di tracciabilità padre-figlio.

<sup>6</sup> Insieme alla data d'apertura (differenza tra le due), definisce il tempo di risoluzione del Reclamo.

<sup>7</sup> Perché accada ciò, basta semplicemente non rispondere al telefono o, come capita spesso, non registrare alcun reclamo, perché come degli stupidi bambini, ci si vergogna.

<sup>8</sup> I moduli saranno affrontati più avanti, come forma d'output strutturato.

- il cliente e le sue aspettative;
- l'analisi del processo e le eventuali non conformità insorte.

Se si segue questo approccio, nascono almeno tre tabelle:

- una tabella per archiviare i reclami ed i dati di nota e costo, tabella **RECLAMI**;
- una tabella con le anagrafiche dei clienti, tabella **DOCUMENTI**;
- una tabella con l'analisi delle cause del reclamo, tabella **PROCESSI**.

## Ta brutte e ta... belle, ovvero vitto, alloggio, lavatura... e stiratura



Per ora, occupiamoci di semplici tabelle- dati in correlazione tra loro.

Innanzitutto, le tabelle non sono altro che degli insiemi di celle ordinate in colonne e righe, al cui interno vengono posti i dati caratterizzati da un certo formato.

Da quanto accennato sui formati, ne consegue che una tabella-dati non è un semplice insieme di cassetti in cui è possibile ammassare i dati, ma un luogo che si può compartimentare in modo adeguato, così da riporre in ordine le informazioni, facendo in modo che occupino il minor spazio possibile<sup>9</sup>.

Insomma, occorre realizzare un servizio “*all-inclusive*” del tipo: vitto, alloggio, lavatura... e stiratura.

Nella tabella principale i campi da inserire saranno:

- **ID reclamo**, tipo contatore e chiave primaria;
- **data apertura**, tipo data;
- **ID tipologia documento attivo**, tipo numerico;
- **riferimento documento attivo**, tipo numerico;
- **data documento attivo**, tipo data;
- **ID processo**, tipo numerico;
- **ID causa**, tipo numerico;
- **note**, tipo memo;
- **ammissibile**, sì/no;
- **data chiusura**, tipo data.
- **costo**, tipo numerico;

Per i campi di “tipo data”, può essere utile personalizzare il formato, specificando la proprietà **Maschera d’Input** così, quando s’inserirà una data sarà formattata automaticamente \_\_/\_\_/\_\_. Per i “tipo testo” può essere utile definire una lunghezza inferiore ai 50 caratteri standard e commisurata alla lunghezza effettiva dei possibili dati in ingresso.

La tabella DOCUMENTI si dovrà strutturare per effettuare una ricerca incrociata tra:

- tipologia del documento,
- numero di protocollo.

Nella tabella PROCESSI la chiave primaria corrisponderà ad un indice che identifica il processo e un indice che rimanda alla causa specifica.

Tutto ciò significa anche che non deve essere possibile inserire due clienti o due processi con lo stesso nome. Questi indici sono necessari per poter mettere in relazione i dati di queste due tabelle con quella principale.

<sup>9</sup> In informatica il concetto di “poco spazio” è spesso in similitudine con velocità d’accesso.

## Un servizio chiavi in mano

Tutti i database, per poter rintracciare le informazioni tra i vari dati, sono caratterizzati da delle chiavi d'identificazione, in modo da identificare in maniera univoca i dati, impedendo allo stesso tempo la possibilità d'inserire nella tabella dati ridondanti (duplicati non ammessi).

Vediamo quindi di comprendere meglio cosa sono le relazioni tra dati.

È intuitivo comprendere che un cliente può dar luogo a infiniti reclami e che non può esserci un reclamo senza che ci sia un cliente coinvolto.

Clienti e reclami sono quindi elementi in relazione fra loro e, in questo caso, la relazione è del tipo **1 a ∞** (uno a molti). Al fine di gestire agevolmente il dato, si può decidere la medesima cosa per le cause, cioè una categoria di reclamo può essere presente in più reclami, ma per facilitare le successive analisi si deve sempre aprire uno specifico reclamo per ogni differente tipologia di causa che si presenti. Da quest'ultima precisazione risulta evidente che un reclamo non deve contenere più motivazioni, ma conviene aprire più reclami distinti, senza vergognarsi del crescere del numero dei reclami, perché essi ci sono a prescindere dal fatto che qualcuno li registri o meno.

Emerge così anche la necessità di creare una quarta tabella secondaria: CLIENTI. Il campo principale sarà:

- **ID cliente** di tipo contatore,
- dimensione campo: intero lungo,
- indicizzato: sì, duplicati non ammessi.

Analogamente nella tabella PROCESSO, il campo principale sarà:

- **ID processo** anch'esso di tipo contatore,
- dimensione campo: intero lungo,
- indicizzato: sì, duplicati non ammessi.

Premendo, in entrambi i casi, il tasto destro del mouse nella modalità “visualizza struttura” sulla cella di selezione (quella a sinistra di colore grigio) è possibile selezionare dal pop-up l'opzione chiave primaria in modo da salvare l'integrità referenziale a cui s'è accennato.

Per quanto concerne la tabella CLIENTI, al suo interno si può creare tutto il necessario per definire l'anagrafica, cioè:

- **ID cliente**, contatore;
- **ragione sociale**, testo;
- **indirizzo** e numero civico, testo;
- **ID città** (da correlare ad un'altra tabella<sup>10</sup>), numero;
- **telefono**, testo;
- **fax**, testo;
- **sito web**, ipertesto.

Per quanto riguarda la tabella PROCESSI, si può prevedere l'aggancio ad un'altra tabella: CAUSE. Per ottenere tutto ciò è necessario prevedere un campo:

- **ID causa** di tipo numerico,
- dimensione campo: intero lungo,
- indicizzato: sì, duplicati ammessi.

In questo modo sarà possibile prevedere una tabella che contiene le sotto-famiglie come figli appartenenti alla medesima categoria padre, perché (ricordiamolo una volta ancora) lo scopo principale del database è quello di raggruppare i dati in categorie per favorire poi metodologie come:

- analisi di Pareto;
- analisi per stratificazione;
- analisi per correlazione.

## Qaerere licitum est, respondere... non semper

Una volta che si siano memorizzati i dati, occorre in qualche modo poterli analizzare. Per effettuare una qualsiasi analisi occorre effettuare delle richieste di dati a cui il database deve poter rispondere adeguatamente. Queste richieste, con tanto di risposte, vengono dette **interrogazioni** o **query**<sup>11</sup>.

<sup>10</sup> In questo modo elementi come il nome della località, del comune d'appartenenza, il codice d'avviamento postale, la nazionalità, il prefisso telefonico verranno di volta in volta immessi una sola volta, poi saranno sempre disponibili e selezionabili.

<sup>11</sup> Dal latino *qaerere* = chiedere, ricercare.

Quando si specificano più tabelle in una query, viene operato il **prodotto cartesiano** tra i record delle tabelle, cioè a ogni record della prima tabella, viene affiancato da ogni record della seconda, per cui se la prima tabella ha  $m$  record e la seconda ne ha  $n$ , vengono generati  $n \times m$  record.

Se s'introducono dei criteri di filtraggio sulle righe e sulle colonne d'entrambe le tabelle, vengono prodotti record costituiti unicamente dalle combinazioni dei record delle tabelle filtrate.

Uno dei filtri più comuni che si specifica è quello di chiedere che il campo di un record sia uguale ad un campo dell'altro.

L'operazione di filtrare su chiavi uguali viene chiamata **join**, dato che tale operazione è molto comune, esistono delle tecniche che ne ottimizzano il funzionamento, tali ottimizzazioni vengono attivate solamente se l'utente specifica in modo diverso le operazioni di join.

### **Per centrare il bersaglio, più che Robin Hood... un little join**



A questo punto si può creare la tabella CAUSE, nella quale si crea un elenco delle principali cause, rapportandole alla tabella PROCESSI mediante criteri d'appartenenza.

In maniera generica un'azienda che realizzi dei prodotti e ne commercializzi degli altri potrebbe suddividere i processi e le cause in questa maniera (NC è l'acronimo di Non Conforme):

- Prodotto NC:
  - colore o altri attributi qualitativi,
  - caratteristiche meccaniche implicite al tipo di lavorazione,
  - caratteristiche dimensionali implicite al tipo di lavorazione,
  - caratteristiche connesse a trattamenti chimico-fisici;
- Packaging NC:
  - etichetta,
  - confezione,
  - imballaggio interno,
  - altro motivo;
- Ordine NC:
  - errata conferma d'ordine,
  - prezzi sbagliati,
  - mancato annullamento,
  - dati offerta inesatti,
  - altro motivo;
- Quantità NC
  - eccedenza,
  - mancanza,
  - altro motivo;
- Trasporto NC:
  - rottura del prodotto,
  - manomissione dell'imballaggio,
  - errata consegna,
  - smarrimento del carico,
  - altro motivo;
- Materia prima NC:



- caratteristiche chimico-fisiche scadenti,
- additivazione carente,
- etichetta errata,
- altro motivo;
- Commercializzato NC:
  - prodotto scadente,
  - packaging imperfetto,
  - etichetta errata,
  - altro motivo;

### **L'unione (join) fa la forza, Luke!**

In Ms Access le operazioni di join si specificano trascinando il campo di una tabella su quello collegato di un'altra. I collegamenti vengono rappresentati mediante frecce che uniscono le due tabelle e, ovviamente, ci possono essere più tabelle implicate in più operazioni di join.

Anche per eliminare una relazione join è sufficiente posizionarsi all'interno della **finestra relazioni**, selezionare la relazione che unisce le due tabelle e utilizzare il tasto **Canc**.

Esistono tre tipi di join. Il **join simmetrico** è quello standard, in esso vengono considerati solamente i record di una tabella per i quali vi siano dei record corrispondenti (con la stessa chiave) nell'altra, se un record contiene nella sua chiave (interna od esterna) un valore che non ha corrispondente, questo viene scartato. Tuttavia, può accadere che si vogliano considerare anche dei record che non hanno un corrispondente specifico dall'altra parte che con un join simmetrico verrebbero scartati. Per questo esistono il **join interno** (o sinistro) e il **join esterno** (o destro):

- un **join interno** considera tutti i record della prima tabella, a cui collega (se possibile) quelli della seconda;
- un **join esterno** considera tutti i record della seconda, a cui associa (se possibile) quelli della prima.

Quindi i tre tipi differiscono su come operano, quando incontrano un campo contenente una chiave nulla.

Tornando al caso in esame, in un'azienda si può avere:

- un parco clienti, ma non è detto che tutti i clienti debbano necessariamente aver reclamato, quindi posso avere un collegamento tra tutti i clienti e i reclami che hanno effettuato (join interno);
- una serie di cause possibili, tra le quali anche quelle che non hanno dato luogo a reclami, e i clienti che eventualmente hanno effettuato i reclami (join esterno).