**SPECIFICHE FINALI ESTRAZIONE DATI COATTIVA**

I campi recuperati dall’estrazione per titolo di credito sono i seguenti:

1. **IdAnagrafica**: Id numerico per mascherare il codice fiscale del contribuente, serve per raggruppare i crediti nel caso il contribuente si fosse visto arrivare più titoli da pagare aggregati in un solo documento (stessa data notifica e stessa data di carico dei titoli). Più anagrafiche potrebbero avere lo stesso idAnagrafica, rappresentano tutte lo stesso contribuente ma possono presentare differenze nei campi indirizzo, cap, deceduto, pec ecc;
2. **DataNascita**: Data di nascita del contribuente ricavata dal codice fiscale;
3. **TipoPersonalità**: PF persona fisica(contribuente) PG persona giuridica (azienda);
4. **Telefono**: Presenza o assenza del numero di telefono. *Legato alla singola anagrafica*;
5. **IndirizzoResidenza**: indirizzo di residenza del contribuente. *Legato alla singola anagrafica*;
6. **Cap**: Cap ottenuto dall’indirizzo di residenza. *Legato alla singola anagrafica*;
7. **CittadinanzaItaliana**: Presenza o assenza della cittadinanza ottenuto dal codice fiscale (idAnagrafica);
8. **Deceduto**: Presenza o assenza di decesso per la persona fisica. *Legato alla singola anagrafica*;
9. **DecedutoDataInfo**: Data di registrazione nel database del decesso nel caso in cui questo fosse avvenuto. *Legato alla singola anagrafica*;
10. **Cessata**: Presenza o assenza di cessata attività per la persona giuridica;
11. **CessataDataInfo**: Data di registrazione nel database di cessata attività nel caso in cui questo fosse avvenuto. *Legato alla singola anagrafica*;
12. **PEC**: Presenza o assenza di una PEC verificata del contribuente. *Legato alla singola anagrafica*;
13. **DataCaricoTitolo**: Data in cui Abaco ha ottenuto il titolo di credito. Titoli di credito con la stessa data di carico appartengono allo stesso ruolo di Contarina;
14. **DataEmissioneTitolo**: Data in cui il titolo è stato emesso per il contribuente, utilizzato per calcolare la Vetustà del titolo;
15. **TipoCredito**: Tipo del titolo di credito;
16. **ValoreTitolo**: Valore monetario del titolo di credito;
17. **Pagato120Giorni**: Somma pagata dal debitore a 120 giorni dopo la data di primo sollecito per il titolo. Tale somma è probabilmente scorretta nel caso in cui più titoli con lo stesso idAnagrafica abbiano la stessa data di carico e stessa data di notifica, dato che il contribuente si è visto ricevere un unico documento da pagare e non i titoli singoli, la somma pagata è poi stata redistribuita in modo arbitraria tra tutti i titoli che andavano a comporla;
18. **DataPrimaNotifica**: Data in cui il contribuente è stato per la prima volta sollecitato al pagamento da Abaco, il campo Pagato120Giorni conta i pagamenti del debitore a partire da questo giorno;
19. **DataPagamentoTotale**: Data in cui è stata pagata la totalità della somma dovuta dal titolo;
20. **NumeroTitoliAperti**: Numero di titoli non pagati dal contribuente (stesso idAnagrafica) alla data di carico del titolo di credito, includendo anche titoli non provenienti da ruoli di Contarina;
21. **ImportoTitoliAperti**: Somma del valore dei titoli non pagati considerati nel campo precedente;
22. **NumeroTitoliSaldati**: Numero di titoli completamente saldati dal contribuente (stesso idAnagrafica) alla data di carico del titolo di credito, includendo anche titoli non provenienti da ruoli di Contarina;
23. **ImportoTitoliSaldati**: Somma del valore dei titoli saldati considerati nel campo precedente;
24. **NumeroTitoliRecenti**: Numero di titoli riferiti allo stesso contribuente (stesso idAnagrafica) con stessa data di carico del titolo. Non viene considerato il titolo stesso nel conteggio;
25. **TotaleTitoliRecenti**: Somma del valore dei titoli considerati nel campo NumeroTitoliRecenti.

**Note:**

- *Legato alla singola anagrafica:* Il campo può assumere valori diversi per stesso idAnagrafica a causa di anagrafiche duplicate nel database, non si ha modo di risalire a quale fosse il valore corretto alla data di carico del titolo, in questi casi va scelto un valore che sia uguale per tutti i titoli appartenenti allo stesso idAnagrafica e con stessa data carico e notifica.

**Pulizia dei dati:**

* **Verranno considerati solo i titoli con campo TipoPersonalità di valore ‘PF’**. *Vengono considerate solamente le persone fisiche in quanto quelle giuridiche presentano l’assenza di molti dati come data di nascita e cittadinanza italiana.*
* **Verranno eliminati i titoli con stesso valore idAnagrafica se ci fossero casi in cui, a quel valore di idAnagrafica siano associati valori di TipoPersonalità diversi in titoli diversi**. *Vengono eliminati i titoli riferiti a contribuenti che vengono a volte presentati come persone fisiche e a volte come aziende, questo perché è possibile che alcuni titoli facessero inizialmente riferimento ad un’anagrafica di tipo aziendale e poi siano stati passati, ad un tempo non ben precisato, alla persona fisica corrispondente all’azienda.*
* **Nel caso in cui titoli con stesso idAnagrafica presentino valori diversi nel campo DecedutoDataInfo, il valore di tale campo verrà impostato per tutti al valore minimo registrato nell’estrazione**. *Dato che tale campo è riferito all’anagrafica ed è stato calcolato basandosi su modifiche apportate all’entry corrispondente all’anagrafica sul database, è possibile che l’anagrafica sia stata duplicata in anni differenti, per cui la data di decesso corrisponde all’anno in cui è stato creato il duplicato. Siccome una stessa persona fisica non può resuscitare, la data di decesso più corretta per calcolare l’effettivo decesso alla data di carico di ogni titolo è il valore minore tra quelli presenti per lo stesso idAnagrafica.*
* **Nei titoli con stesso idAnagrafica, DataCaricoTitolo e DataPrimaNotifica, i campi IndirizzoResidenza, Cap e Telefono assumeranno come valore il valore di moda calcolato nel raggruppamento**. *I titoli che presentano valori uguali nei tre campi in prima riga verranno successivamente aggregati per calcolare l’effettivo pagato dal contribuente, dato che nella maggior parte dei casi (nonostante non possiamo averne la certezza) al contribuente è stato inviato un unico documento con un unico credito da pagare dal valore della somma dei titoli arrivati e inviati nello stesso momento. Siccome a questi possono essere collegate diverse anagrafiche e non si può sapere quale è l’anagrafica con data di aggiornamento più vicina alla data di caricamento dei titoli deve essere scelto un unico valore di Cap, Telefono e indirizzo con cui descrivere il contribuente.*
* **I contribuenti che presentano in ogni titolo a loro corrispondente (stesso idAnagrafica) il campo Deceduto con valore 1 ma con InfoDataDecesso sempre vuoto verranno considerati deceduti da sempre.**
* **Verranno esclusi i titoli che presentano Cap nullo e IndirizzoResidenza con valore ”SCONOSCIUTO ALL’ANAGRAFE”.**  *Si tratta di pochi titoli in cui non si sa proprio la residenza, e quindi il cap, del contribuente.*
* **Il valore del campo Cap nei titoli che presentano Cap nullo verrà sostituito con un valore unico significativo (ad esempio 0).** *Questo perché alcuni titoli presentano il cap nullo in quanto esteri, dunque facendo geocoding sull’indirizzo associato si può capire se sono effettivamente esteri e, nel caso, registrare l’informazione tramite una colonna aggiuntiva.*
* **Verranno eliminati i titoli che hanno lo stesso valore in simultanea sui campi idAnagrafica, DataCaricoTitolo e DataPrimaNotifica e valori diversi in o IndirizzoResidenza o Telefono o Deceduto.** *Possiamo trattare i casi in cui i 3 campi iniziali sono uguali su più titoli come un caso in cui al contribuente è stato chiesto di pagare il credito assieme, con un valore dato dalla somma dei titoli di credito. Il problema si presenta quando ci sono incongruenze nei dati anagrafici, segno che i titoli di credito POTREBBERO essere o stati mandati in un unico documento o in più documenti arrivati alla stessa data per la stessa persona. Dato che non si può risalire alla modalità di pagamento (singola o sdoppiata) in questi casi verranno eliminati i titoli (circa 1700 su 55000)*