**OSSERVAZIONI**

1. Ci sono un totale di 74.143 titoli unici, di questi 18.592 fanno riferimento a persone giuridiche
2. Ci sono un totale di 19.835 contribuenti unici, di questi 4296 sono persone giuridiche
3. 322 contribuenti persone fisiche sono senza data di nascita
4. Per 1986 contribuenti persone fisiche è presente la data in cui è stata registrata la morte
5. 91 Contribuenti mancano il campo cap perché l’indirizzo risulta sconosciuto o è estero
6. 443 contribuenti presentano diversi indirizzi nel tempo, dovuto alla presenza di anagrafiche duplicate, per un totale di 3039 titoli appartenenti a questi contribuenti
7. Ci sono, tra le persone fisiche 9 date uniche di carico titolo, segno che **il dataset include almeno 9 ruoli**
8. **158 contribuenti persone fisiche presentano l’attributo di cessata attività** , probabilemente questi sono gli unici casi in cui il titolo faceva riferimento ad una **persona giuridica** ma poi l’anagrafica è stata aggiornata. Inoltre **tutti (tranne 1) titolo hanno data cessata attività dopo la data di assunzione titolo**, indicando che **al momento dell’assunzione del titolo erano infatti azienda attive**
9. **Nessun contribuente/anagrafica ha diverse date di nascita per diversi titoli**, quindi se la **data di nascita è assente** su un titolo per quel contribuente è assumibile che in realtà il codice fiscale riportato nell’anagrafica fosse una partita iva e sono da rimuovere in quanto **persone giuridiche**
10. **Anagrafiche uguali con indirizzi differenti sono da vedersi come titoli differenti? Il numero di titoli recenti aperti non combacia**
11. **Calcolo dei titoli recenti / passati completamente sbagliato per anagrafiche che presentano indirizzi differenti in titoli differenti**
12. **Problema: ruoli che forse sono stati spezzati in 2, nonostante la data di carico dei titoli per lo stesso contribuente sia la stessa i titoli vengono divisi in sottogruppi,** esempio:anagrafica 19717
13. La **vetustà** (data assunzione-data scadenza) non sembra essere rilevante perché i titoli risultano praticamente sempre scaduti pochi giorni prima dell’assunzione, soprattutto **tutti i titoli di un contribuenti assunti nello stesso ruolo hanno tutti la stessa identica data di scadenza**
14. I **tipi di titolo** risultano 8 se si considerano solo le persone fisiche mentre risultano 9 guardando anche alle giuridiche (manca “**Servizi Sanitari**”). Negli ultimi 5 anni solo 70 titoli appartenenti a 50 anagrafiche erano di questo tipo.
15. **Solo 25 titoli hanno PEC presente, e risultano tutti non pagati**

Controllare anomalie nei dati:

1. Quanti contribuenti presentano più di un indirizzo?
2. Sono presenti date attività cessate su titoli riferiti a persone fisiche, da ignorare?
3. Eliminare i contribuenti con indirizzo sconosciuto?
4. Solo 1 contribuente presenta IndirizzoResidenza “SCONOSCIUTO ALL’ANAGRAFE SCONOSCIUTO ALL’ANAGRAFE”
5. 6 contribuenti presentano IndirizzoResidenza “ESTERO ESTERO EE”
6. Dunque dei 91 contribuenti senza cap 84 sono esteri e gli altri 7 anomali
7. Mettere a 0 il cap degli esteri per distinguerli? **COME DISTINGUERE INDIRIZZI ESTERI DA ITALIANI?**

ANOMALIE:

1. idAnagrafica 356: indirizzi diversi stesso ruolo, pesanti incongruenze su numero titoli aperti nello stesso ruolo e passati, tale numero è differente per titoli diversi nello stesso ruolo quando dovrebbe essere uguale
2. idAnagrafica 2471: per un ruolo ha conteggi diversi nei titoli recenti in base a dove ha l’indirizzo differente (“ruolo spezzato”) mentre in un altro ruolo nonostante ci siano indirizzi diversi il conteggio dei titoli recenti conta tutti e 3 i titoli appartenenti a quella data e non li separa.
3. idAnagrafica 19207: ha 3 cap differenti per un singolo ruolo, spezzato in 2 parti e non in 3

IDEE:

1. Campo nord centro e sud in base al cap
2. Nei casi in cui il ruolo è spezzato in due per uno stesso contribuente: ha senso tenerli separati al moemento dell’aggregazione? (si assume che il sollecito sia stato mandato in momenti diversi) o vanno aggregati? (si assume che il sollecito sia stato mandato allo stesso momento e sia corretto considerare il totale di pagamento per il contribuente anche se in alcuni casi i titoli di credito separati all’interno di uno stesso ruolo fanno riferimento a indirizzi diversi) **QUALE DELLE 2 SOLUZIONI?**

SOLUZIONI APPROVATE:

1. Trattare come persone giuridiche (quindi da escludere) i titoli che presentano anagrafiche senza la data di nascita (erano partite iva invece che codice fiscali)
2. Nel caso una identità abbia deceduto=1 in almeno un titolo bisogna impostare ad 1 tutti i campi deceduto dei titoli appartenenti a quella persona, verrà poi calcolato l’effettivo valore del decesso facendo la differenza tra data del titolo e data decesso (se positivo allora era già deceduto)

SOLUZIONI PROPOSTE:

1. Eliminare i titoli in cui, raggruppando per identità, data assunzione e numerotitoli recenti presentano un conteggio dei titoli raggruppati inferiore al numero titoli recenti.
2. Per i titoli in cui i ruoli si presentano spezzati (vedi IDEE punto 2) sembra che i 2 ruoli vengano segnalati differentemente, quindi è corretto mantenerli separati (dunque raggruppando per numero recenti, numero mancanti, data titolo e id anagrafica invece che solo gli ultimi 2 campi)
3. Per gli indirizzi: se su uno stesso ruolo ci sono più indirizzi mantengo quello che risulta più presente (in alternativa, ma decisamente più complicato, si può cercare di mantenere quello più aggiornato, guardando altri titoli relativi alla stessa anagrafica). Inoltre sarebbe possibile inserire una nuova colonna ad 1 se quel credito (aggregato di ruoli) presentava più indirizzi, è però sconsigliato dato che la maggioranza degli indirizzi multipli sono lo stesso indirizzo con uno spazio o punteggiatura aggiuntiva.

MODIFICHE DA FARE APPROVATE:

1. Fare +1 sui titoli recenti (si conta quindi anche il titolo attuale)
2. Droppare la casella importo titoli recenti (viene calcolata dalla somma del valore degli aggregati)
3. Droppare la casella datanotificatitolo (è sempre vuota)
4. Droppare righe con contribuente che presenta IndirizzoResidenza “SCONOSCIUTO ALL’ANAGRAFE SCONOSCIUTO ALL’ANAGRAFE” e senza cap
5. Droppare righe di contribuenti che erano aziende

SOLUZIONE RAPIDA PER AVERE UN MODELLO (ELIMINAZIONE DI TUTTE LE ANAGRAFI CON DATI AMBIGUI)

1. VIA QUELLE CHE ERANO AZIENDE
2. VIA QUELLE CHE HANNO PIU’ INDIRIZZI
3. VIA QUELLE CHE HANNO DECEDUTO DIFFERENTE