# Report cose fatte

Cosa ho fatto?

**Analisi dati demografici**

I dati sulla popolazione residente al 1 gennaio 2002 li ho scaricati dal sito dat.istat.it

Ho selezionato solo quelli del 2020, sia maschi che femmine,

Poi ho scaricato i dati per età sempre al 2020, così posso vedere solo quelli con più di x anni (ad esempio 65 oppure 80). L’interrogazione è salvata sul sito dati.istat.it

**22 marzo**

Ho aggiunto vari dati demografici importanti e ho scritto il codice per fare le mappe

Devo decidere lo stile migliore per le mappe

Altri stili di classificazione disponibili sono:

1. **"pretty"**: Questo stile genera intervalli di classificazione "carini" (cioè facilmente leggibili) per i dati.
2. **"equal"**: Questo stile divide i dati in intervalli di uguale ampiezza.
3. **"quantile"**: Questo stile divide i dati in intervalli di uguale numero di osservazioni.
4. **"jenks"**: Questo stile utilizza l'algoritmo di Jenks per trovare intervalli ottimali che minimizzano la variazione all'interno degli intervalli e massimizzano la variazione tra gli intervalli.
5. **"sd"**: Questo stile classifica i dati utilizzando intervalli basati sulla deviazione standard.
6. **"manual"**: Questo stile consente di specificare manualmente gli intervalli di classificazione.

Posso fare le mappe per i comuni piccoli <5000, medi 5000-60000 e grandi

Capire come analizzare gli “squilibri” tipo: tanti anziani, poca spesa. Forse la cosa migliore è percentuale di anziani sul totale della popolazione anziana della regione, e percentuale spesa sulla spesa totale, e poi vedere dove ci sono le differenze maggiori. Ad esempio se ad Ancona ci sono il 20% di anziani e si alloca il 30% di spesa sanitaria regionale c’è un divario.

# Cose da fare

## Stefano

Lavora alle funzioni per visualizzare le mappe, forse ad shyni app

## Mariagrazia

Aggiungere variabili alla mappa

* ~~Aggiungere il dato sulla percentuale di anziani (sia over 65 che over 80) sul totale degli anziani delle marche - CORREGGERE~~
* ~~Indice invecchiamento (numero over 65 / numero under 14)~~
* Aggiungere i dati su assistenza territoriale: tabelle excel nella cartella “data/socio-demo”
  + Domiciliare socio assistenziale
  + Domiciliare integrata
  + Strutture residenziale
* Per ciascun elemento bisogna mettere nella tabella
  1. Presenza servizio
  2. Numero utenti
  3. Spesa totale
  4. Spesa SSN
  5. Spesa utenti
  6. Spesa comune
  7. Per i dati dal 2 al 6 la percentuale del totale

Ho creato file ADI, SAD, RSA\_variabili spesa e l’ho messo nella cartella socio-demo in data.

Capire quale sia nei vari file il codice istat per poter fare il left joint.

--- FINE PRIMA PARTE – DATI ANAGRAFICI E SPESA SERVIZI TERRITORIALI ---

**Analisi dati SDO**

Abbiamo i dati per comune con i ricoveri ospedalieri da cui si evenice quante persone in un anno si sono ricoverate con alcune NCD rilevanti che potrebbero essere in maniera più efficiente sul territorio invece che in ospedale.

Questo ci permette di stimare il fabbisogno di questi servizi in maniera più accurata che la sola età utilizzata fino ad ora.

**Come lavorare i dati**

**MG**

* ~~Fare un raggruppamento per comune, con i ricoveri legati a ciascuna patologia, e attaccare alla tabella principale “db”~~
* ~~Calcolare anche lo stesso dato con~~
  1. ~~Una colonna per i ricoveri programmati~~
  2. ~~Un colonna per quelli da PS~~
* ~~Calcolare la percentuale di pazienti con classe età == 1 (< 65 anni) rispetto al totale di pazienti per ciascuna patologia, per capire se sono tutte patologie legate agli anziani. Per alcune patologie possiamo decidere in un secondo momento di considerare solo alcune classi di età~~ – Non sono riuscita a filtrare per ps = 0
* ~~mettere nel db reddito per fare una multivariabile (facile)~~
* mettere il livello di istruzione (più difficile): calcolare percentuale di licenza media in giù
* mettere gli stranieri (facile)
* Poi si fa la multivariabile anche con questi confounders e l'analisi dello scostamento si fa su questo modello
* ~~calcolare i ricoveri con quelle patologie solo per le classi di età 2 e 3 (anziani). Basta che fai il filtro prima di applicare quelle funzioni che hai scritto ieri. In questo modo togliamo dalla regressione i ricoveri associati a quelle patologie ma che non riguardano la popolazione anziana~~

**MG 26/05/2023**

* **Ho spostato nel codice 05b\_db\_mappe\_sdo le funzioni sulle sdo**

**STEFANO**

* Capisce quale è il modo migliore di pesare questi ricoveri per il loro burden sull’ospedale. Un modo possibile è vedere il DRG, oppure i giorni di ricoveri medi per quei codici.
* Sono più evitabili, e quindi più importanti, i ricoveri programmati o quelli in PS? Vedere con Claudia
* Quelli che si sono ricoverati con una diagnosi, a cui non è associato iun DRG, vuol dire che non sono stati rimborsati? Come è possibile? In altre parole, perché per alcune patologie Claudia non ha indicato nessun codice DRG?

~~Integrazioni da chiedere~~

* ~~Numero totali di ricoveri per riga~~
* ~~Se ha usato l’esenzione o solo DPR e DRG come gli abbiamo detto, visti alcuni numeri molto alti mi viene il dubbio che abbia usato l’esenzione~~
* ~~Farselo fare per il 2019 (senza covid)~~

La regressione percentuale ricoveri e percentuale anziani contiene l'elenco della differenza dal valore atteso. Se facciamo un dataframe per comune con questa differenza possiamo vedere i comuni più "sani" e meno sani dove fare più prevenzione