

Progetto Basi di Dati

Catapano Pacifico e Giugliano Salvatore

Aprile-Maggio 2020

Tutto andrà bene. Così, senza nessuna ragione a sostenerlo, se non la ragione stessa. Mentre le strade si svuotano, l'angoscia ci riempie e i profeti di sventura scalzano gli arruffapopolo dai talk-show, mani anonime hanno cominciato a scrivere "Tutto andrà bene" sopra dei post-it che appiccicano sui muri, sugli alberi, sulle serrande dei negozi. La gente passa, li vede, scuote la testa e poi continua per la sua strada, un po' più leggera. Il web moltiplica l'effetto e rimbalza il messaggio dentro i telefonini. "Tutto andrà bene" era la frase che Gregory Peck sussurrava a Ingrid Bergman in certi filmi ingoiati durante l'infanzia. Sono le parole che mi disse la mamma lasciandomi per la prima volta davanti al portone della scuola. Niente di che, una piccola medicina dell'anima, un integratore emotivo per le difese immunitarie. Un placebo, forse, ma in certi momenti serve anche quello. "Tutto andrà bene" si coniuga al futuro, ma agisce subito. Sembra arrivato apposta per colmare il vuoto degli abbracci che non ci possiamo più dare. Si sa che l'idea è partita da una donna, una poetessa, talmente saggia da non far trapelare il suo nome, affinché tutti coloro che la imitano possano sentirsi parte dello stesso silenzioso esercito di assicuratori. Qualunque firma toglierebbe forza al messaggio. In fondo basta mettere quella di John Lennon: "Alla fine tutto andrà bene. E se non va bene, significa che non è ancora la fine".

- Massimo Gramellini, *Corriere della Sera*, Edizione del 07/03/2020

1 Descrizione delle specifiche di progetto

Si costruisca una base di dati contenente i valori dell'andamento dell'epidemia COVID-19 e la si rappresenti attraverso un'interfaccia utente.

2 Creazione della tabella master importando i file CSV in una tabella SQL

Accedendo al link allegato alla traccia di progetto (<https://github.com/pcm-dpc/COVID-19/tree/master/dati-province>) verrà visualizzato in formato CSV tutte le info relative ai contagi da Covid 19, giorno per giorno, per tutte le province italiane. I file registrano informazioni dal 24/02/2020 fino ad ora. Utilizzando la pagina web consigliata all'interno della traccia (<https://www.convertcsv.com/csv-to-sql.htm>), si ottiene una sintassi in forma tabellare SQL da copiare incollare all'interno di Oracle.

3 Verificare che la tabella segua uno schema in 3NF ed eventualmente decomporre definendo tutti i vincoli

Nella seconda parte abbiamo rielaborato la tabella master in modo da ridurla in 3NF. La tabella non originale risulta essere di prima forma normale, ma non di seconda. Quindi si procede con il processo di decomposizione prima in seconda e poi in terza forma normale. La chiave primaria all'interno della relazione risulta essere il sottoinsieme degli attributi (data, codiceProvincia)

I vincoli di FD presenti risultano essere:

- FD1 (data, codiceProvincia) -> Stato
- FD2 (data, codiceProvincia) -> codice_regione
- FD3 (data, codiceProvincia) -> denominazione_regione
- FD4 (data, codiceProvincia) -> denominazione_provincia
- FD5 (data, codiceProvincia) -> sigla_provincia
- FD6 (data, codiceProvincia) -> lat
- FD7 (data, codiceProvincia) -> lon
- FD8 (data, codiceProvincia) -> totale_casi
- FD9 (data, codiceProvincia) -> note_it
- FD10 (data, codiceProvincia) -> note_en

- FD11 (codice_Provincia) -> denominazione_provincia
- FD12 (codice_provincia) -> sigla_provincia
- FD13 (codice_provincia) -> lat
- FD14 (codice_provincia) -> lon
- FD15 (codice_regione) -> denominazione_regione
- FD16 (codice_regione) -> codice_provincia

4 Arricchimento dello schema mediante ALTER TABLE

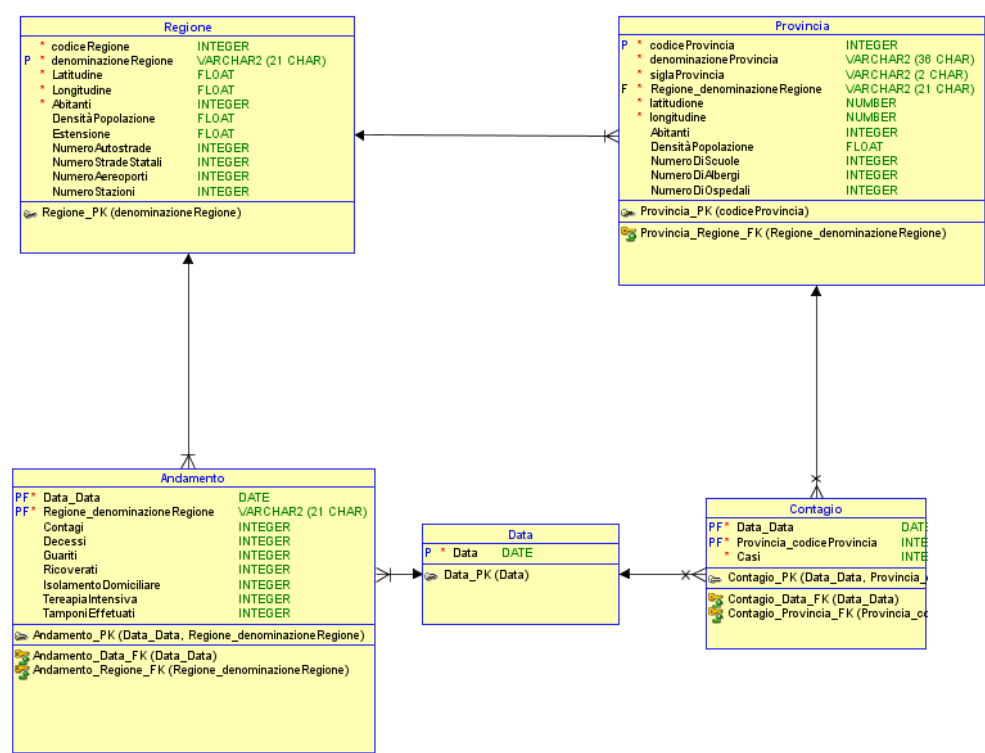
A. Aggiunta di morti, ricoveri e terapie intensive per regioni I dati catturati all’interno delle ALTER TABLE sono stati estratti dalla cartella di origine Github in COVID-19/dati_regioni

B. Numero di abitanti, la densità abitativa, superficie in km2, numero di autostrade e strade statali, numero di aeroporti e stazioni per regioni I dati catturati all’interno delle ALTER TABLE sono stati estratti da www.wikipedia.org cercando la pagina di ogni regione specifica, aggiungendo i dati manualmente nella tabella .

C. Numero il numero di abitanti, la densità abitativa, il numero di scuole, alberghi/strutture recettive, il numero di ospedali per provincia I dati catturati all’interno delle ALTER TABLE sono stati estratti da varie fonti:

- Numero abitanti, numero scuole per provincia: www.comuniecitta.it
- Numero ospedali per provincia: File allegato alla Cartella Onedrive di progetto all’interno della cartella ”Dati province”
- Numero strutture alberghiere per provincia: File allegato alla Cartella Onedrive di progetto all’interno della cartella ”Dati province”

5 Individuazione, attraverso un processo di reverse engineering, di un possibile schema E/R della base di dati



Il Modello E/R ivi rappresentato è composto da 3 entità e 3 relazioni.

Entità:

- Regione (codiceRegione, denominazioneRegione, latitudine, longitudine, abitanti, densitàPopolazione, numeroAutostrade, numeroSS, numeroAeroporti, numeroStazioni);
- Provincia (codiceProvincia, denominazioneProvincia, regioneAppartenenza: REGIONI, latitudine, longitudine, siglaProvincia, abitanti, densitàAbitativa, nrScuole, nrAlberghi, nrOspedali);

- DataCampione (data);

Associazioni:

- Andamento: Molti a molti tra Regioni e DataCampione. Gli attributi sono (dataAndamento:DATACAMPIONE, regione: REGIONI, contagi, decessi, guariti, ricoverati, isolamentoDomiciliare, terapiaIntensiva, totalePositivi, tamponiEffettuati);
- Contagio: Molti a molti tra Provincia e DataCampione. Gli attributi sono (data:DATACAMPIONE, provincia: PROVINCE, casi);
- appartenenzaRegione: Uno a Molti, seguendo il processo di traduzione delle associazioni è stata accorpata nell'entità province;

Per l'entità Regione è stato scelto come chiave primaria denominazioneRegione poichè codiceRegione risulta coincidere in P.A. Bolzano e P.A. Trento, le quali nonostante siano province, sono state considerate dal data source come regioni.

6 La specifica in SQL di una serie di query utili all'analisi dell'andamento del contagio

Sfruttando i meccanismi relativi alle viste virtuali, il tempo di esecuzione delle query viene ottimizzato. Ecco qui di seguito il codice SQL opportunamente commentato che compie queste operazioni:

https://communitystudentiunina-my.sharepoint.com/:u:/g/personal/salvatore_giugliano2_studenti_unina_it/EcS-D12Li8tNjrClT-z4Ya4Bhzq9sTT350QiVSX0cE7uqA?e=rdJJUr

7 Rappresentare i risultati attraverso un'opportuna interfaccia grafica

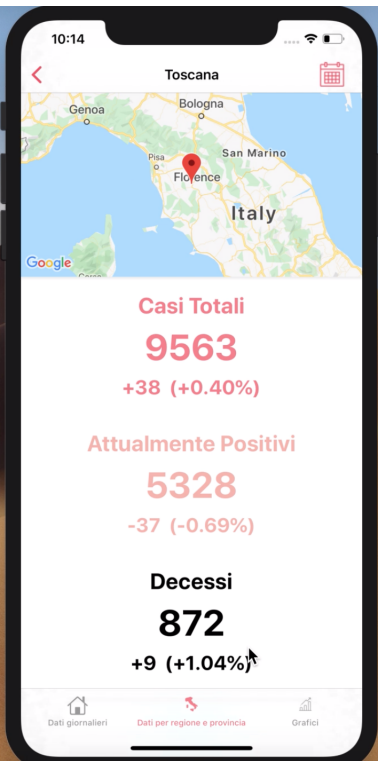
Si realizza un'applicazione nativa in linguaggio di programmazione Swift per sistemi Apple distribuzione Ios. L'app è diviso in tre sezioni; ciascuna caratterizzata dalla propria Tab Bar.

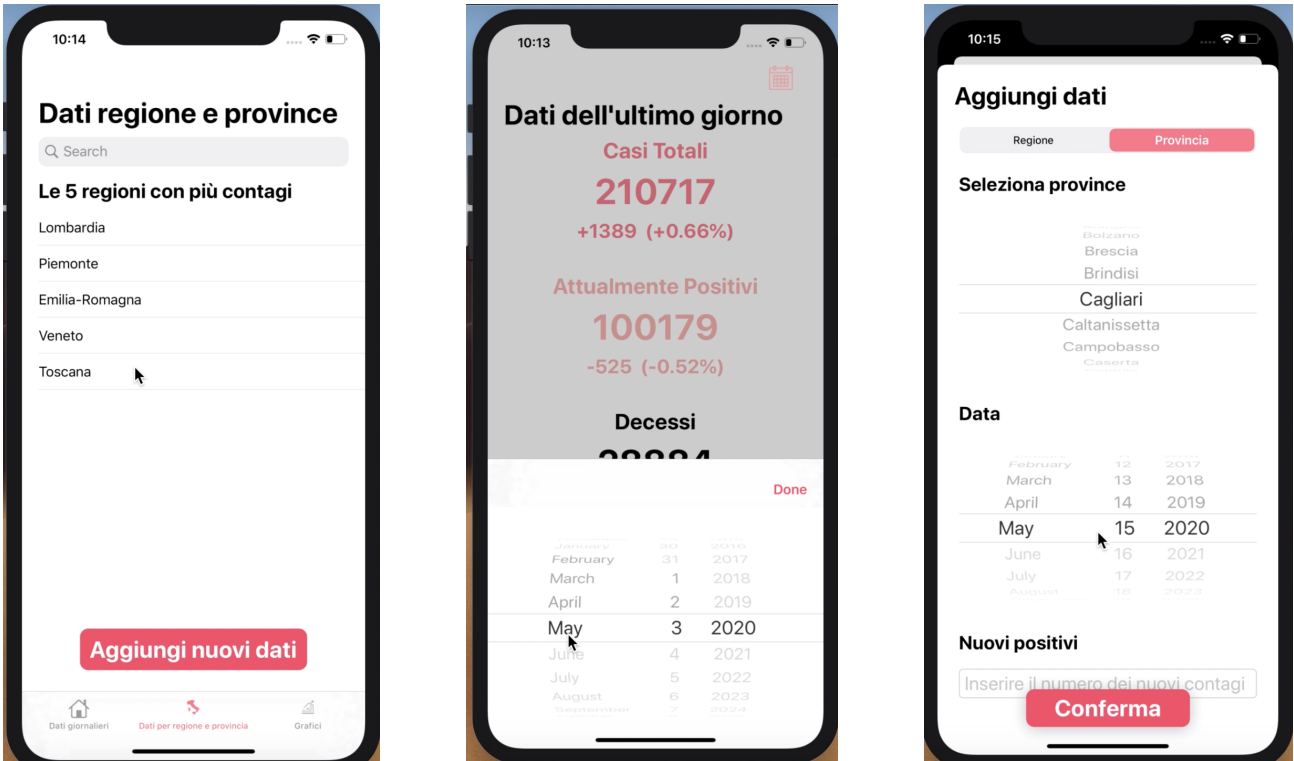
Nella sezione Home si riscontrano i dati relativi all'andamento giornaliero del contagio, o se esso non è presente nel database ci si riferisce all'ultimo giorno inserito nel dataset. Scorrendo i giorni nel calendario (icona in alto a destra) è possibile cambiare il giorno predefinito in uno ad esso antecedente; i dati si aggiorneranno automaticamente.

Nella sezione regioni e province è possibile visualizzare i dati per ogni zona geografica, effettuando una facile ricerca all'interno della Search Bar. La prima volta che si entra nella schermata verranno visualizzate le 5 regioni con più contagi. Tappando su ognuna di esse si potranno avere informazioni specifiche quali: dati dell'andamento epidemia, opportuni grafici, e visualizzazione della mappa integrata attraverso un API di Google Maps. Attraverso il bottone "Aggiungi dati" è possibile effettuare manualmente degli INSERT sul database, aggiornandolo permanentemente di conseguenza.

Nella sezione grafici verranno visualizzati gli andamenti dell'epidemia lungo tutto l'arco temporale considerato. La schermata è composta da tre grafici: uno per la variazione totale, uno per la variazione giornaliera, e infine l'ultimo per rapporto di tamponi su contagi.

Mostriamo ivi alcuni screen dimostrativi:





Si allega il link della repository scaricabile del progetto: <https://github.com/pacificocatapano/progettoBD.git>

8 La specifica in PL/SQL di una serie di procedura/trigger che possano consentire un’analisi più flessibile ed eventuali aggiornamenti automatici della base di dati qualora vogliano essere importati dati successivi al 03/5/2020

Link alla cartella del progetto:
https://communitystudentiunina-my.sharepoint.com/:u:/g/personal/salvatore_giugliano2_studenti_unina_it/ES4r8nCJjE9Iv5e8syVnDrIBR9mPsD0yLvFgmVVBMSPUJw?e=oFZMGy

Tutti i punti sopraelencati sono stati sviluppati in diversi file .sql di cui qui mostrati:
https://communitystudentiunina-my.sharepoint.com/:f:/g/personal/salvatore_giugliano2_studenti_unina_it/Er1YBnUb_UpIo-kJEtN1d0gBDh_Lk4VMLUhtuXNFAKCGdg?e=r4UU7a