

Visual Analytics
VAST 2021 Mini-Challenge 2

Chiara D'Anzi

Luglio 2022

Indice

1	Introduzione	3
2	Dati	3
2.1	Cleaning e particolarità	4
2.2	Utilizzo dei dati	4
3	Scelte progettuali e stilistiche	4
4	Riferimenti	5
5	Interattività	5
5.1	Spinbutton	5
5.2	Mappa	5
5.3	Lista delle località	6
5.4	Grafici	6

1 Introduzione

Il progetto consiste nel rappresentare, tramite un'applicazione web, le informazioni utili ad analizzare lo scenario descritto dalla Mini-Challenge 2 della VAST Challenge 2021, che ha come fine quello di osservare i movimenti dei dipendenti della GasTech di Abila.

Link per maggiori informazioni sulla Mini-Challenge: <https://vast-challenge.github.io/2021/MC2.html>.

Link all'applicazione web: <https://github.com/ChiaraDA/vast_mc2>.

2 Dati

Di seguito sono stati riportati i dataset di partenza e una riassuntiva descrizione degli attributi di ciascuno.

Car Assignments

Dati generali dei dipendenti della GasTech in relazione all'auto aziendale:

- LastName: cognome del dipendente
- FirstName: nome del dipendente
- CarID: identificatore dell'auto aziendale
- CurrentEmploymentType: tipo di occupazione
- CurrentEmploymentTitle: titolo dell'occupazione

Credit Card Data

Dati relativi alle transazioni delle carte di credito dei dipendenti:

- timestamp: data e orario (ore, minuti) della transazione
- location: località in cui è stata effettuata la transazione
- price: addebito
- last4ccnum: ultime quattro cifre della carta di credito

Loyalty Card Data

Dati relativi all'utilizzo delle carte fedeltà da parte dei dipendenti:

- timestamp: giorno dell'utilizzo
- location: località in cui è stata utilizzata
- price: addebito registrato
- loyaltynum: numero della carta fedeltà

GPS

Dati GPS di tracciamento delle auto aziendali:

- Timestamp: data e orario (ore, minuti, secondi) dello spostamento
- id: identificativo dell'auto
- lat: latitudine
- long: longitudine

Oltre ai dataset riportati sopra, erano presenti lo shapefile relativo alla mappa della città e una cartina raffigurante le varie attività del luogo.

2.1 Cleaning e particolarità

L'attributo *location* presente nei dataset sulle carte fedeltà e le carte di credito è stato modificato per la voce *Katerina's Café* per questioni di formattazione dei caratteri speciali.

Inoltre, all'interno del dataset *Car Assignment* gli identificativi dei camionisti non sono presenti.

2.2 Utilizzo dei dati

I dati sono stati utilizzati in modo tale da poter fare dei controlli incrociati tra le varie informazioni. In modo particolare, l'aspetto più rilevante riguarda i possibili collegamenti tra le operazioni eseguite con le carte e la posizione GPS dei dipendenti. Questi collegamenti permettono di ricavare informazioni completamente nuove, infatti rendono possibile l'identificazione del proprietario delle carte e, quindi, il tracciamento delle sue transazioni.

Un altro elemento importante è costituito da quali attività vengono maggiormente frequentate e quando, per cercare di capire se ci sono dei comportamenti abitudinali e, eventualmente, delle anomalie in essi.

Infine, allo stesso modo, avere un prospetto delle entrate delle attività può rappresentare un elemento utile alla formazione di un quadro generale e delle sue anomalie.

3 Scelte progettuali e stilistiche

Al fine di agevolare una lettura immediata delle informazioni i widget sono stati rappresentati in modo abbastanza compatto, cercando di non sprecare spazio senza, però, rimpicciolire troppo i vari elementi. Anche sotto il punto di vista responsive questo principio è stato rispettato.

L'elemento centrale della visualizzazione è la mappa della città di Abila. La mappa è costituita da tre "livelli": la mappa derivante dallo shapefile, la cartina di Abila e le traiettorie delle auto aziendali. Per gestire queste traiettorie è stato usato, anziché un elemento input di tipo range, uno spinbutton che permettesse di poter muoversi sulla linea temporale in modo, all'occorrenza, più preciso (utile anche per fare dei controlli sulle transazioni delle carte). Inizialmente lo step è minimo e avanza praticamente minuto per minuto. È possibile, però, muoversi più velocemente inserendo un valore arbitrario all'interno dell'input che è sopra lo spinbutton.

Le traiettorie sono state colorate con il blu per permettere maggiore contrasto con lo sfondo, dato che questo colore non è presente sulla cartina.

Il prospetto con i dati dei dipendenti selezionati sulla mappa, rappresentando dati molto collegati alla mappa, è stato accostato a quest'ultima per permettere una lettura sinottica.

All'interno del prospetto sui dati delle auto selezionate viene visualizzato il nome, il cognome e l'occupazione

del/i dipendente/i.

Nella parte successiva dell'applicazione web è possibile selezionare il nome dell'attività della quale si vuole approfondire la ricerca per visualizzare i dettagli sui metodi di pagamento utilizzati e un grafico a torta che ne mostra la distribuzione (entrambi relativi al giorno su cui è lo spinbutton).

Infine, sempre selezionando un'attività dalle lista a sinistra vengono rappresentate le informazioni settimanali su frequenza della località ed entrate grazie a un grafico a linea, in grado di mostrare le informazioni sulla linea temporale in modo efficace.

4 Riferimenti

Per la realizzazione del progetto sono stati presi in considerazione diversi riferimenti volti ad analizzare contesti di studio simili.

In primo luogo, sono stati un riferimento gli esempi di applicazione web realizzati per l'analoga VAST Challenge 2008. L'elemento della mappa e delle traiettorie in modo particolare.

Altri progetti utili sono stati i seguenti:

- Drug Overdose Deaths during COVID-19: utile soprattutto per la rappresentazione temporale dei dati.
- Environmental Justice Analysis: Contaminated Sites and Waste Disposal Facilities in Iowa: spunti per le rappresentazioni sulla mappa.

5 Interattività

Il progetto presenta degli elementi che interagiscono tra di loro in modo dinamico al cambiare delle informazioni in modo da rendere la ricerca più facile e la visualizzazione più interessante.

5.1 Spinbutton

Il primo elemento interattivo consiste nello spinbutton. Grazie ai bottoni di precedente e successivo presenti ai lati della casella è possibile navigare nella linea temporale entro la quale siamo in possesso dei dati (due settimane quindi).

Come detto in precedenza, al variare del valore arbitrario inserito nella cella precedente, lo step (1 di default) con cui avanzerà/decremerà il tempo sarà più alto o basso in base al valore inserito (il quale non può essere minore di uno o maggiore di 24713).

La data rappresentata nello spinbutton costituisce un filtro anche per le informazioni sulle carte che confluiscono nel prospetto sui dettagli di pagamento e la distribuzione dei metodi utilizzati all'interno, appunto, della giornata sulla quale si è fermato lo spinbutton.

Oltre a questo chiaramente la data e l'orario interagiscono con le traiettorie che cambiano in base a questi.

5.2 Mappa

La traiettorie presenti all'interno della mappa, se selezionate, interagiscono con il prospetto sulla destra il quale mostra le informazioni relative al proprietario della macchina aziendale. Una volta selezionate, il loro colore diventa rosso in modo da distinguerle dalle altre e poterle seguire senza confonderle.

5.3 Lista delle località

La lista delle località permette di concentrarsi su ogni singola attività filtrando i dati e mostrando solo ciò che riguarda la località selezionata.

In seguito alla selezione della località cambieranno tutti i grafici e il prospetto con i dettagli sui metodi di pagamento. Il prospetto mostrerà nella parte iniziale le informazioni generiche e in seguito la lista delle transazioni.

5.4 Grafici

Grazie alla libreria Plotly passando sopra i grafici con il cursore si possono visualizzare ulteriori informazioni: per quanto riguarda il grafico a torta, passando sopra le varie "fette" viene indicato il numero preciso di occorrenze per la categoria, mentre per i grafici a linea passando sui nodi è possibile visualizzare il numero preciso di dipendenti che hanno frequentato una determinata attività in un caso e l'ammontare delle entrate di un'attività nell'altro.