



VAST CHALLENGE 2014 - MC 2

CRISTIANO LANDI (559941)

VISUAL ANALYTICS

DATA SCIENCE & BUSINESS INFORMATICS

Anno Accademico 2020/2021

Indice

1	Descrizione dei dati e presentazione del pattern	2
1.1	Dati forniti	2
1.1.1	Preprocessing dei dati	2
1.1.2	Pattern da comunicare	3
2	Design	4
2.1	Design	4
2.1.1	Sezione 1: mappa	4
2.1.2	Sezione 2: selettore persone	4
2.1.3	Sezione 3: controlli temporali	6
2.1.4	Sezione 4: Transazioni	6
3	Casi d'uso	7
4	Conclusioni	9

1 Descrizione dei dati e presentazione del pattern

1.1 Dati forniti

Sono elencati in seguito i dati forniti, la loro semantica e le loro caratteristiche:

- *car-assignments.csv*: una lista di veicoli assegnati ai dipendenti. Per ogni dipendente è riportato, oltre che al nome e al cognome, la qualifica e il dipartimento. Si noti che tutti i dipendenti assunti con la qualifica di "Truck Driver" non riportano un'auto associata e che è presente una relazione 1:1 tra auto e dipendente.
- *gps.csv*: un file contenente, per ogni veicolo, un timestamp e la posizione in quel momento espressa in termini di latitudine e longitudine. Anche in questi dati si riscontrano veicoli tracciati, ma il quale id non è riportato nel file descritto al punto prima.
- *loyalty_data.csv*: file csv contenente i dati delle transazioni eseguite con la carta fedeltà. Per ogni transazione è riportata la data, il nome del negozio in cui è stata eseguita, il prezzo e il nome e cognome del proprietario della carta.
- *cc_data.csv*: file csv contenente i dati delle transazioni eseguite con la carta di credito. Per ogni transazione è riportata la data, l'ora, il nome del negozio in cui è stata eseguita, il prezzo e il nome e cognome del proprietario della carta.
- *Mappa*: per quanto riguarda la mappa di Abila sono stati forniti i files contenenti le strade sotto forma di shapefile e un'immagine JPG contenente i punti di interesse presenti nell'isola.

1.1.1 Preprocessing dei dati

Tutte le future elaborazioni sono state fatte nell'ottica di cercare di mantenere quante più informazioni possibili.

Mappa

Gli shapefiles sono stati convertiti tramite mapshaper in formato geoJSON, così da poter utilizzare le funzioni geografiche offerte da d3.

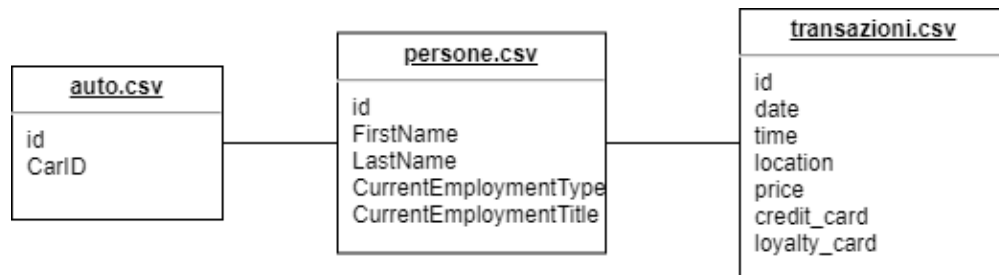


Figura 1.1

Dati delle persone

Tra i dati forniti, 3 file contengono dati sulle persone e 2 dati sulle auto. Per meglio visualizzare i dati con l'interfaccia web è stato deciso di creare un file contenente, per ogni persona presente in uno qualsiasi di questi file, un identificativo *id_D*, il ruolo e il dipartimento di appartenenza.

Ogni record in questo file appartiene ad una delle seguenti categorie: tutti i dati sono presenti, è disponibile soltanto il nome e cognome della persona (dati presenti nei file delle transazioni, ma non *car-assignments*), nessun dato disponibile (utenti fittizi creati per tenere traccia di tutti quei veicoli tracciati, ma non presenti in *car-assignments*)

Dati sulle transazioni

IPOTESI: la carta fedeltà non ha funzioni di credito o debito, le transazioni riportate servono soltanto a tenere traccia del denaro speso in quell'attività per usi futuri (tipo calcolo dei punti).

Si è deciso di unire i file in uno unico contenente i seguenti attributi: *id_P*, timestamp (data), location (negozio), prezzo, time (ore e minuti), *credit_card* (true se il pagamento è avvenuto tramite carta di credito) e *loyalty_card* (true se è stata usata una carta fedeltà).

Assegnamento auto

Per quanto riguarda l'assegnamento auto-dipendente è stato scelto di realizzare un file contenente l'id della persona e il rispettivo id dell'auto associata.

I nuovi file elaborati hanno la struttura riportata in Figura 1.1.

1.1.2 Pattern da comunicare

La visualizzazione dei dati dovrà permettere la selezione di uno o più utenti per visualizzarne gli spostamenti e le transazioni eseguite in un dato giorno. Dovrà inoltre essere possibile selezionare una fascia oraria all'interno del giorno selezionato.

2 Design

2.1 Design

La visualizzazione dai dati è realizzata in un'unica pagina.

È possibile individuare le seguenti quattro sezioni:

1. Mappa: visualizza la mappa di Abila e i percorsi effettuati dalle persone selezionate nell'intervallo orario selezionato
2. Selettore persone: permette di visualizzare le persone e i relativi dati e di selezionarne una o più. Permette inoltre di applicare dei filtri e di riordinare i dati per nome e titolo.
3. Controlli temporali: permette di selezionare il giorno e la fascia oraria di cui si desidera visualizzare i dati. Mette inoltre a disposizione una funzione per animare gli spostamenti degli utenti selezionati sulla mappa.
4. Transazioni: permette di visualizzare le transazioni eseguite dalle persone selezionate nel dato intervallo temporale.

2.1.1 Sezione 1: mappa

La mappa è implementata usando le funzionalità geografiche offerte dalla libreria *d3*.

La scelta di usare *d3* deriva dal fatto che librerie simili come Leaflet o l'uso di API verso servizi come Google Maps non permettono lo stesso livello di personalizzazione: con Leaflet non è infatti possibile usare un'immagine come quella fornita come layer base (essendo Abila una città di un'isola inventata, senza immagine le strade fornite risultano in mezzo al mare e sarebbe necessario mappare manualmente i punti di interesse in un file geoJSON), mentre con Google Maps invece sarebbe stato necessario eseguire chiamate ad un servizio esterno che avrebbero inficiato le performance dello strumento.

Questa sezione della pagina ha come unico scopo la visualizzazione della mappa e dei record gps risultati dopo l'applicazione dei filtri.

Gli spostamenti gps visualizzato sono convertiti in *LineString* per motivi prestazionali. Ogni percorso visualizzato è leggermente trasparente per permettere una chiara visualizzazione anche in caso di segnali sovrapposti.

2.1.2 Sezione 2: selettore persone

Questa sezione è implementata usando la libreria *vue-good-table*.

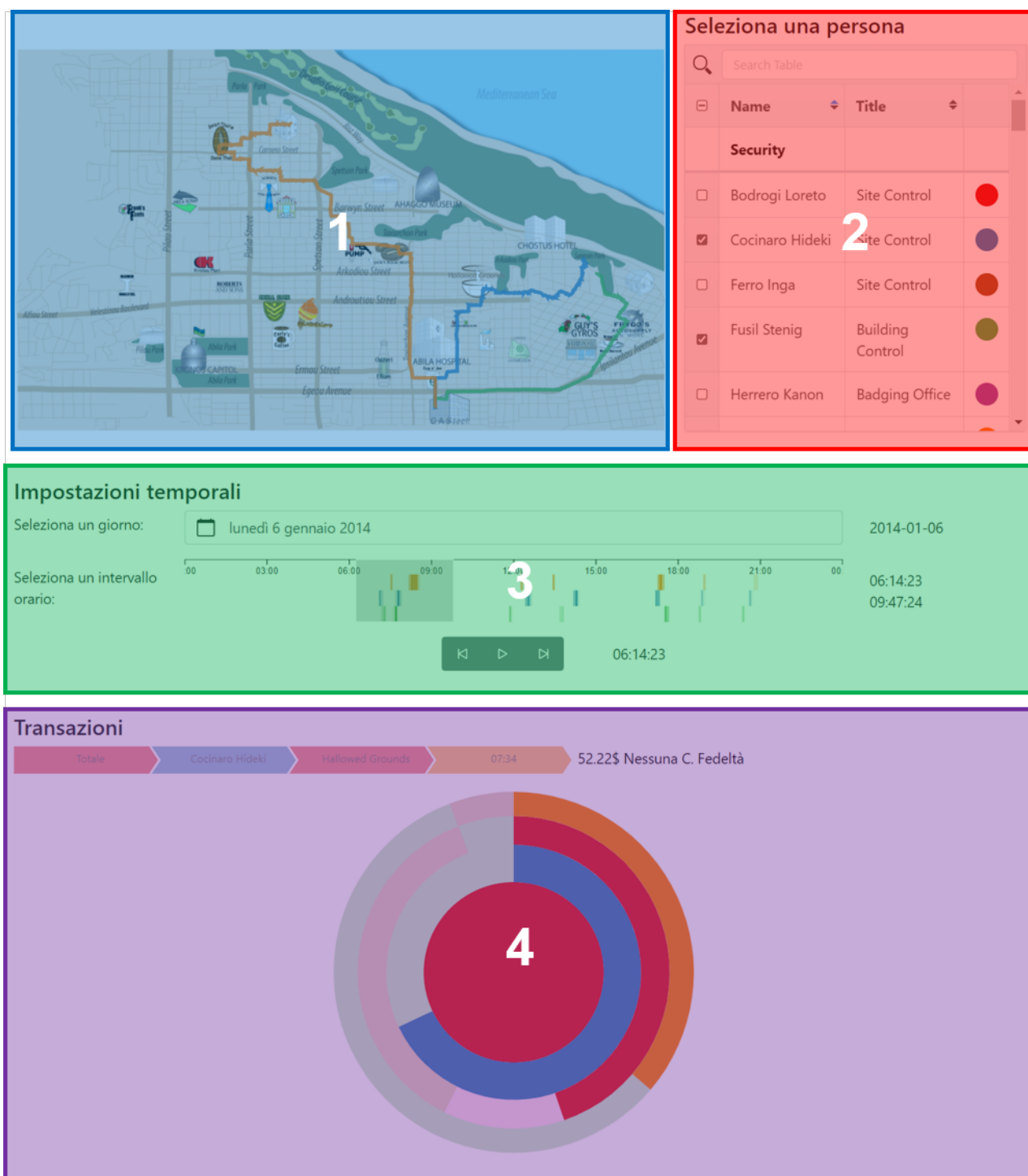


Figura 2.1: Sezioni della dashborad

La tabella visualizza il nome e il ruolo di ogni persona presente nei dati; inoltre è aggiunta un'ulteriore colonna per visualizzare il colore¹ assegnato all'individuo che sarà usato nelle altre sezioni della pagina. Le persone sono inoltre raggruppate per il dipartimento dell'azienda per il quale lavorano.

La tabella permette inoltre di selezionare, filtrare e ordinare i dati.

2.1.3 Sezione 3: controlli temporali

Questa sezione è formata da un selettore per la data alla quale si desidera restringere la selezione, implementata con il solo uso del tag *b-form-datepicker* da *BootstrapVue*, e da un selettore per la fascia oraria. Quest'ultimo è implementato in *d3* e permette di individuare più facilmente, durante il giorno selezionato, quando sono avvenuti gli spostamenti di una data persona (identificata dal colore) e restringere i dati da visualizzare ad una determinata fascia oraria (trascinamento per selezionare la fascia oraria, click per annullare la selezione). Questo componente poteva essere implementato anche istogramma con associato un range selector (ad esempio: Beautiful Histogram Range Slider Component), ma è stata scelta questa visualizzazione per permettere di visualizzare quale persona(tramite il colore) ha fatto determinati spostamenti.

Si è scelto inoltre di creare una piccola animazione che scorre l'intervallo temporale selezionato mostrando sulla mappa tutti gli spostamenti avvenuti nei 60 secondi precedenti.

2.1.4 Sezione 4: Transazioni

Infine è stato usato il componente *Vue.D3.Sunburst* per creare il grafico sunburst.

Questo grafico permette di visualizzare, oltre che alla quantità di denaro spesa, anche la persona che ha eseguito la transazione e il locale presso il quale è avvenuto il pagamento.

Con il passaggio del mouse sopra una qualsiasi sezione del grafico, è possibile visualizzare la quantità di denaro spesa aggregata per quel livello.

Nel caso di record troppo fitti è possibile cliccare su una qualsiasi parte del grafico per visualizzare il solo dato e i relativi figli.

¹La palette dei colori usata è quella suggerita da *colorbrewer* con la seguente configurazione: [link](#). È stato rimosso il colore FFFF33 (giallo) in quanto troppo chiaro.

3 Casi d'uso

Questo strumento è costruito per rispondere alle domande della mini challenge 2 della VA-ST challenge 2014. È possibile quindi usare questo strumento per individuare anomalie e identificare eventi o pattern presenti nei dati.

In seguito sono riportati alcuni esempi.

Teletrasporto

Selezionando "Calzas Axel" è possibile vedere come, alle 13 e 30 del 6 gennaio parta da un locale vicino al Coffee Shack e si fermi in una strada vicina (forse ha disattivato il dispositivo gps), per poi ripartire dopo le 17 dalla GASTech e fermarsi vicino a Carnero Street (Figura 3.1).

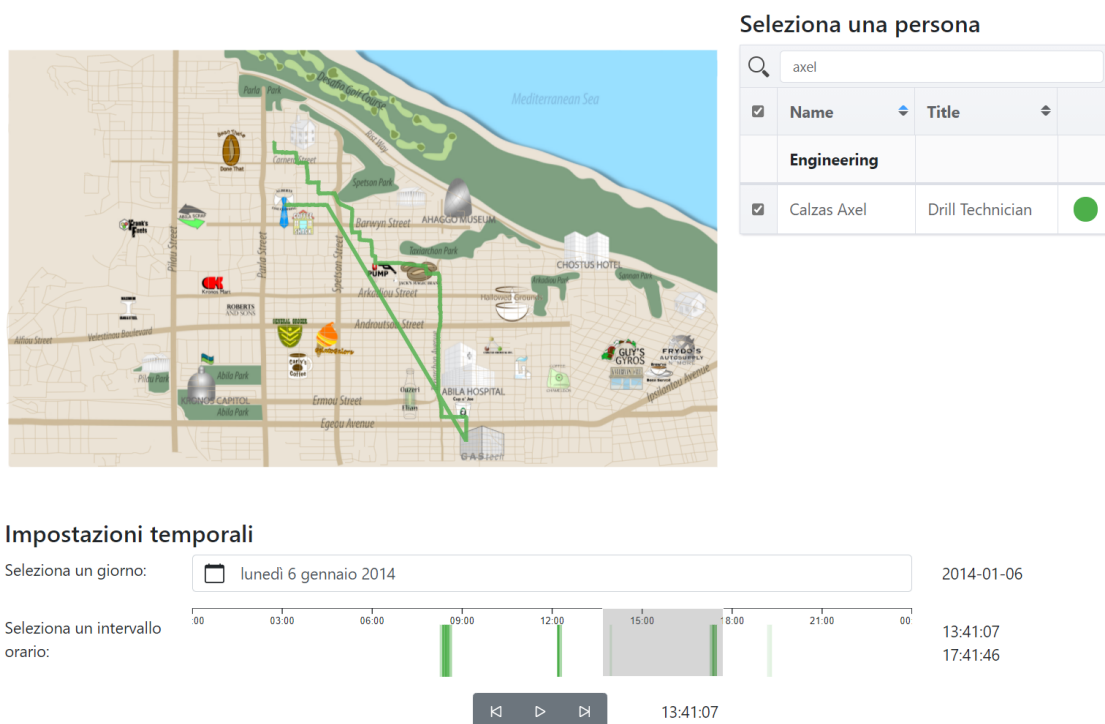


Figura 3.1: Uso dell'interfaccia per individuare gli spostamenti di Calzas Axel

Movimenti con pattern anomali

Selezionando il secondo utente "Sconosciuto" è possibile visualizzare (ad esempio il 15 gennaio), tramite la funzione di selezione della fascia orarie e l'uso dell'animazione, come vengano eseguiti spostamenti ripetuti tra l'Abila Hospital e una strada vicino a "Maximum Iron Steel".

Altri spostamenti anomali si possono individuare selezionando "Vann Isia" e notando come, a differenza di tutti gli altri giorni, il 6 e il 10 gennaio effettuati degli spostamenti alle 23, quando, solitamente, i suoi spostamenti si fermano intorno alle 21.

Malfunzionamenti

Selezionando "Orilla Elsa" è possibile vedere come, oltre che a piccoli "teletrasporti", la linea che identifica gli spostamenti sia molto ampia (dovuta a posizioni registrate a "zigzag", Figura 3.2). Questo potrebbe essere causato da un malfunzionamento del dispositivo gps o un posizionamento che ne inficia il segnale.

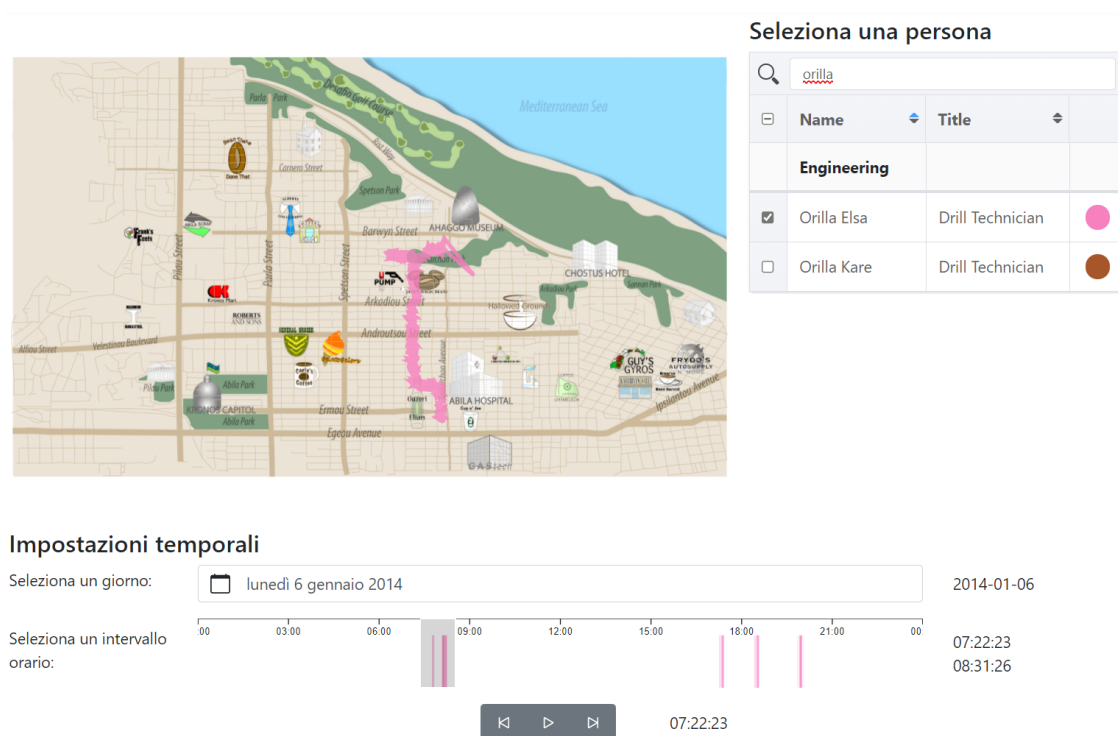


Figura 3.2: Uso dell'interfaccia per individuare gli spostamenti di Orilla Elsa

Completezza dei punti di interesse nella mappa

È possibile usare il sunburn e la mappa per individuare locali non presenti sulla mappa come punti di interesse. Selezionando "Calzas Axel" è possibile vedere come, alle 13 e 57 del 6 gennaio esegua una transazione al locale "Hippokampos" non riportato nella mappa fornita.

Correlazioni tra gruppi di persone e determinati luoghi

Selezionando tutti i dipendenti che lavorano nel reparto Security è possibile vedere come, ad esempio il 6 gennaio, la mattina siano tutti nella zona di Sannan Park. Inoltre, la sera dello stesso giorno, alcuni di loro si ritrovano a Guy's Gyros (oltre che dalla mappa, anche dalle transazioni eseguite) per cenare, probabilmente, insieme.

4 Conclusioni

Lo strumento permette di rispondere a tutte le domande richieste dal cliente. Gli unici strumenti simili esistenti che permettono, tramite interfaccia grafica, una visualizzazione e filtraggio analoghi, sono quelli realizzati in risposta a questa mini-challenge 2 della VAST challenge 2014.

Sono comunque elencate librerie e tecniche alternative utilizzabili per i singoli componenti della dashboard.