

# **RELAZIONE PROGETTO VISUAL ANALYTICS**

## **DINO FUN WORLD – VAST CHALLENGE 2015**

**Caterina Falchi Matr. 502145**

Il problema consisteva nell'analizzare i dati forniti per la Mini Challenge 1, ovvero i movimenti delle persone all'interno del parco Dino Fun World nei giorni di Venerdì, Sabato e Domenica.

### **PREPARAZIONE DEI DATI**

Il dataset era composto di 3 file CSV, uno per ogni giorno, contenenti record così formati: timestamp, id, tipo (check-in o movimento), coordinata x, coordinata y. I dati di ogni file sono stati importati in 3 tabelle su un db Sqlite: movFri, movSat e movSun. I dati sono stati poi elaborati con lo script Python Gruppi che ha permesso di creare la divisione delle persone in vari gruppi. In particolare sono stati individuati 968 gruppi per il Venerdì, 1162 per il Sabato e 2620 per la Domenica. Lo script funziona in questo modo:

- Legge il database inserendo i dati in una struttura creata appositamente
- Esegue l'algoritmo di raggruppamento. Scorre la struttura dati delle persone, ognuna viene confrontata con tutte le altre (che non sono già state inserite in un gruppo). Se le due persone hanno una corrispondenza del 70% dei movimenti, allora vengono inserite nello stesso gruppo. I parametri per la corrispondenza sono stati settati dopo aver fatto alcune prove.

I gruppi creati dallo script sono stati copiati in un file CSV e inseriti nelle tabelle groupFri, groupSat, groupSun.

Le istruzioni per la creazione del db sono contenute nel file instructions.txt.

### **IL PROGETTO**

Il progetto si compone di 4 pagine:

- Index
- Group Analysis
- CheckIn Movements
- Creighton Pavilion

**N.B.** Per consentire il corretto funzionamento del progetto è stata installata l'estensione di Chrome: Allow-Control-Allow-Origin: \*.

### **INDEX**

È la pagina principale del progetto. Contiene una breve descrizione del problema, e i link alle pagine di analisi.

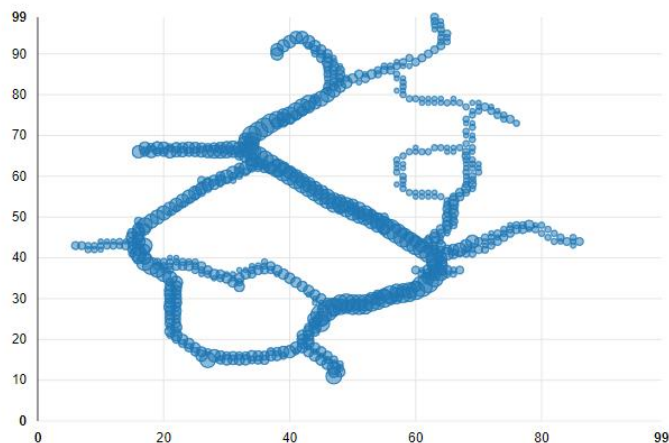
### **GROUP ANALYSIS**

È la pagina di analisi dei gruppi. Selezionando un giorno dalla toolbar appare l'elenco dei gruppi di persone per quel giorno. L'elenco mostra il nome del gruppo e il numero di persone che ne fanno parte. Selezionando dall'elenco un gruppo, appare a sinistra uno scatter plot dei punti del parco visitati dalle persone del gruppo. Per ogni punto, appare un cerchietto tanto più grande quante sono le volte che è stato visitato. Inoltre viene aggiornato lo spazio delle annotazioni con il nome del gruppo, quanti check-in e quanti movimenti hanno effettuato le persone che ne fanno parte.

Day: Friday

Group: 1 N. Check-in: 90 N. Movements: 5119

Day: **Friday** Saturday Sunday



Group 0	5
<b>Group 1</b>	<b>2</b>
Group 2	7
Group 3	3
Group 4	8
Group 5	7
Group 6	8
Group 7	4
Group 8	7
Group 9	4
Group 10	0

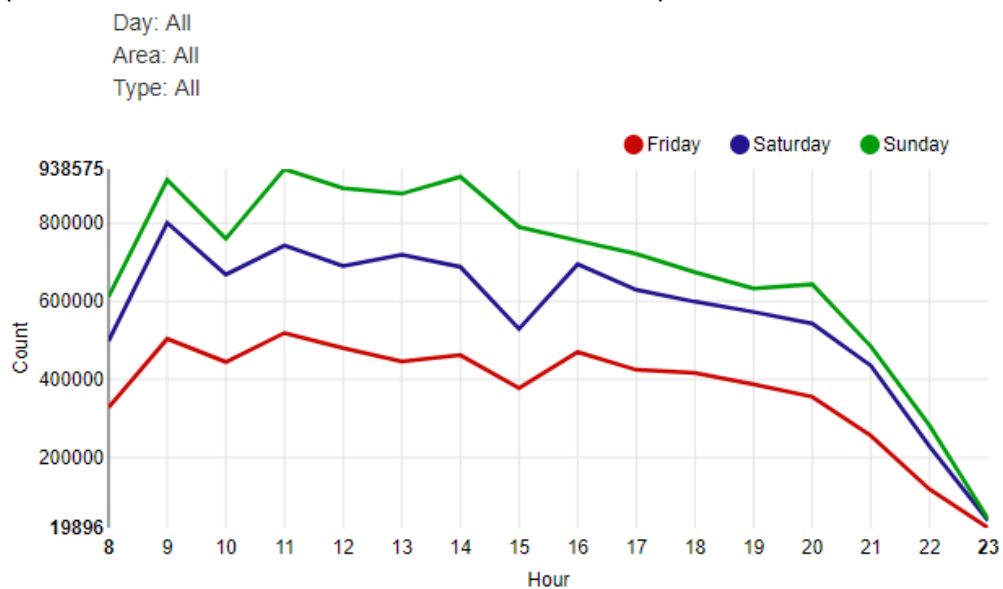
## CHECKIN MOVEMENTS

In questa pagina vengono analizzati i movimenti o i check-in o entrambi, per area e per giorno. Ci sono tre toolbar a disposizione:

1. Toolbar dei giorni
2. Toolbar dei tipi di movimento
3. Toolbar delle aree

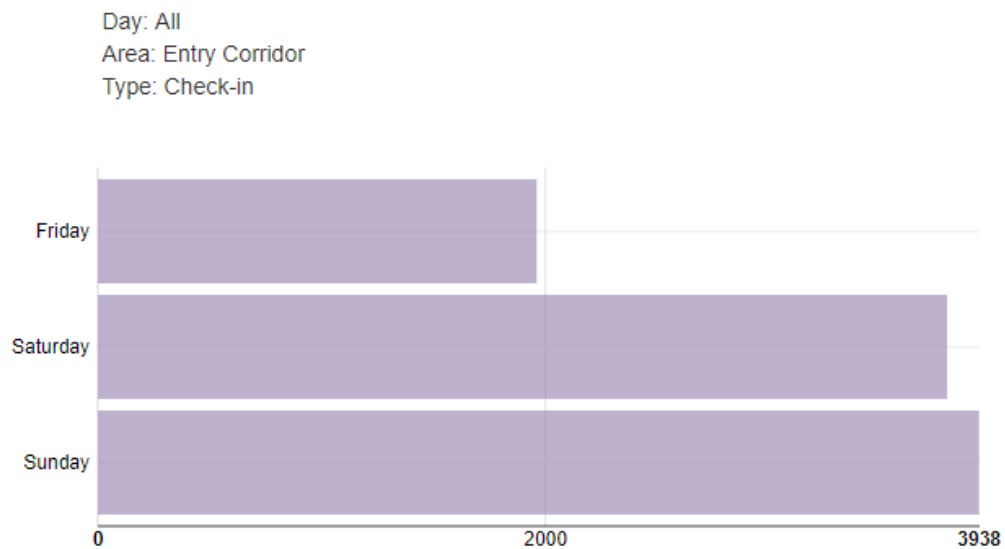
Abbiamo 4 combinazioni possibili di statistiche:

- Selezionando tutti i giorni (All) e tutte le aree (All) appare un lineChart che mostra per ogni ora del giorno quanti movimenti o check-in o entrambi sono stati compiuti:

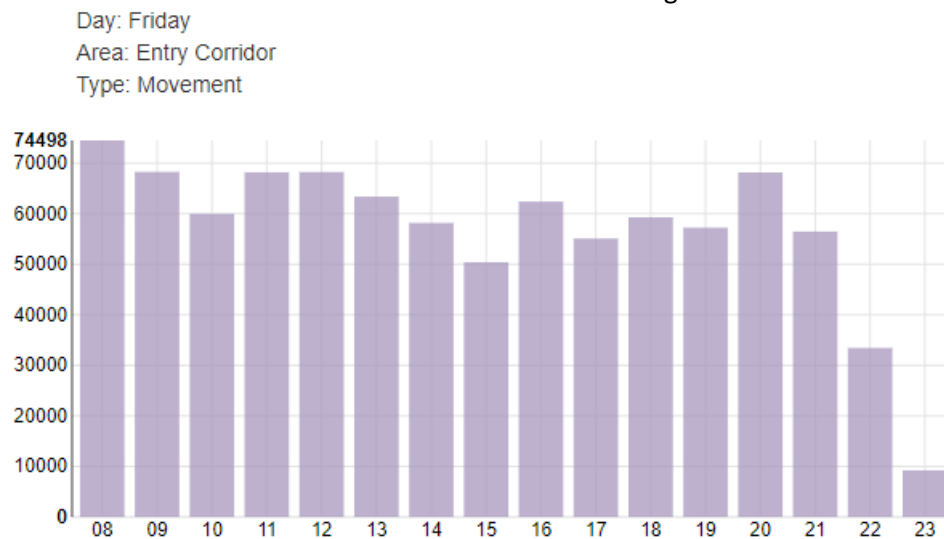


È possibile anche scegliere di vedere solo alcuni giorni, per mezzo dei pulsanti/legenda in alto a destra del grafico.

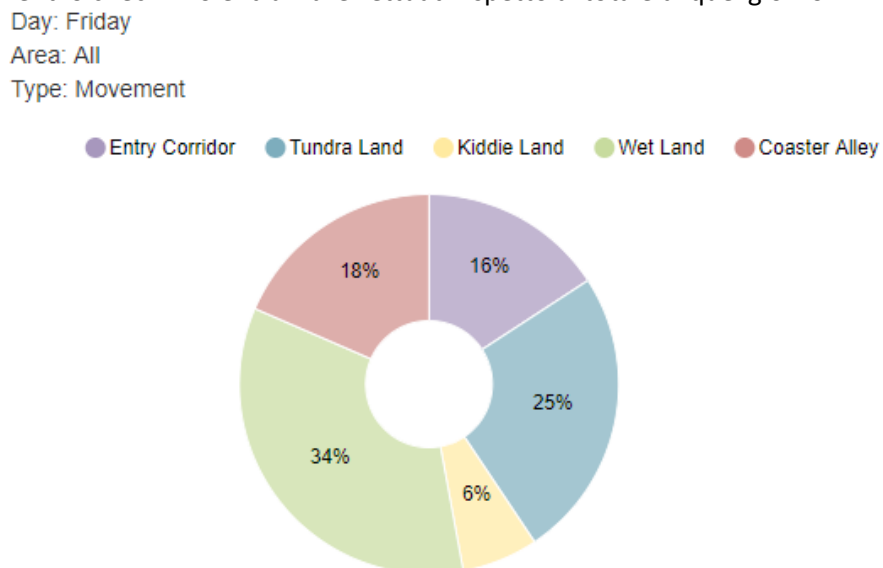
- Selezionando un'area, e lasciando tutti i giorni (All) appare invece un barChart orizzontale che mostra per ogni giorno i movimenti o i check-in o entrambi effettuati nell'area scelta.



- Scegliendo sia un'area che un giorno, appare un barChart verticale che mostra per ogni ora del giorno i movimenti o i check-in o entrambi effettuati nell'area scelta nel giorno scelto:



- Scegliendo invece tutte le aree (All) e un giorno, appare un pieChart che mostra per il giorno scelto la % di movimenti o check-in o entrambi effettuati rispetto al totale di quel giorno:

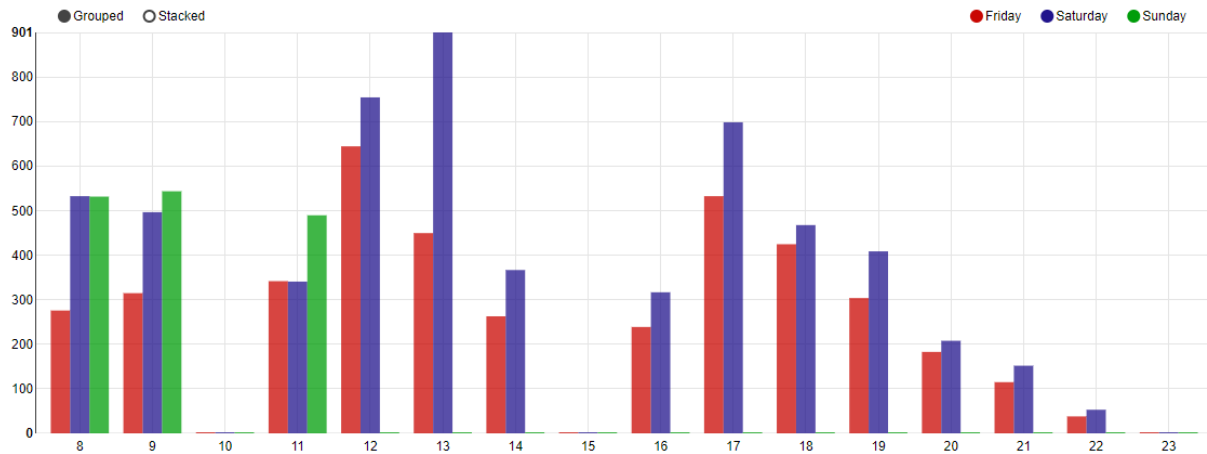


I colori dei grafici relativi alle aree riprendono quelli scelti per evidenziare le aree nella mappa.

**N.B:** Per la visualizzazione di questi grafici occorre aspettare qualche secondo il caricamento dei dati.

### CREIGHTON PAVILION

Questa pagina concentra l'analisi sull'attrazione 32, Creighton Pavilion, dove è avvenuto il crimine. Il grafico mostra i check-in avvenuti per giorno e per ora. È possibile scegliere anche una visualizzazione stacked.



Osserviamo che diversamente dagli altri giorni, la domenica registriamo check-in in questa attrazione solo fino alle 11. Con una semplice query sul database, è stato osservato che ogni giorno l'attrazione 32 chiude dalle 9.30 alle 11.30 e dalle 14.30 alle 16.30 perché in queste fasce orarie non vengono registrati check-in. Per questo possiamo dire che il crimine sia avvenuto dopo le 11.30 di Domenica, in particolare l'ultimo check-in è stato effettuato dall'id 38945 alle 11:59.