



Esercizi lezione 7

Esercizio

Sul database titanic, effettuiamo le seguenti analisi:

- Riusciamo a calcolare quanti ponti c'erano sulla nave (prima lettera dei dati nella colonna Cabin)?
- Visualizzare un grafico con il numero di passeggeri di ogni classe di imbarco (colonna PClass)
- Fare la stessa cosa per la colonna Survived
- Qual era la distribuzione delle tariffe (Fare)?
- Riusciamo a vedere la distribuzione delle età dei passeggeri rispetto alla classe di imbarco con un boxplot?
- Visualizziamo un boxplot rispetto alle colonne Fare e Survived; che cose ne deduciamo?

Esercizio

Scarichiamo il dataset stockdata.csv da

<https://github.com/plotly/datasets/blob/master/stockdata.csv> e:

- Visualizziamo l'andamento delle azioni della Apple ("AAPL")
- Visualizziamo l'andamento di tutte le azioni del dataset
- Visualizziamo l'andamento di Microsoft ("MSFT") e IBM ("IBM") negli anni 2012 e 2013
- Con un istogramma, troviamo il range di valori che le azioni di Microsoft hanno assunto più spesso

Esercizio

Scarichiamo il dataset `metal_bands_2017.csv` dalla piattaforma e:

- Con Seaborn visualizziamo l'istogramma del numero dei fan delle varie band; cosa si vede? Cose ne deduciamo?
- Delle band che suonano sia Progressive che Thrash, realizziamo un barplot che mostri quanti fan ci sono in totale rispetto al paese di origine (nota: in questo caso specifico, per accedere alla colonna "`style`" dovremo usare `.loc[:, "style"]` e non `.style` perché quest'ultimo è un metodo)
- Visualizziamo un istogramma delle date di formazione delle band (potremo dover convertire in qualche modo questo dato); si nota qualche trend?

Esercizio

Scarichiamo il dataset `elections.csv` da

<https://github.com/plotly/datasets/blob/master/election.csv>:

- Con un grafico a barre confrontiamo i voti totali presi dai tre candidati (come somma di tutti i distretti)
- Con un grafico a barre confrontiamo il numero di votanti per i primi dieci distretti

Esercizio

Scarichiamo i file `party_in_nyc.csv` dal dataset

<https://www.kaggle.com/datasets/somesnm/partynyc> relativo a quante volte è stata chiamata la polizia per delle feste eccessivamente rumorose:

- Realizziamo un grafico che mostra quante chiamate ci sono state per ogni città per tutti i tipi di locale
- Qual è il rettangolo di latitudine e longitudine considerato dal dataset?
- Quante segnalazioni sono state fatte per case private?
- Con Seaborn visualizziamo il conteggio delle segnalazioni avvenute in ogni quartiere; che succede se aggiungiamo il parametro `hue="Location Type"`?



GRAZIE
EPCODE