

# ESERCITAZIONE LIBRERIE PYTHON: PANDAS

*Corso di Data Analytics*

25 Settembre 2020

1. Costruire un DataFrame Pandas contenente i seguenti dati relativi allo svolgimento delle edizioni dei mondiali di calcio dal 2000 ad oggi:

Anno	Sede	Vincitore	Numero Goal	Pubblico
2018	Russia	Francia	169	47371
2014	Brasile	Germania	171	53592
2010	Sudafrica	Spagna	145	49670
2006	Germania	Italia	147	52401
2002	Corea del Sud	Brasile	161	42268

Il campo pubblico è il numero medio di spettatori per partite.

2. Aggiungere una colonna con nome "Spettacolo\_OK", con valore booleano true/false se il numero di goal segnati in quella edizione è stato maggiore di 150.
3. Estrarre solo i nomi dei vincitori delle ultime tre edizioni.
4. Trasformare la colonna Numero Goal, calcolando il numero medio di goal per partita (anziché quello totale). In ogni edizione sono state giocate 64 partite.
5. Determinare la sede dell'edizione con numero medio minimo di goal segnati.
6. Costruire un secondo DataFrame contenente le seguenti informazioni sugli europei di calcio dal 2000 ad oggi:

Anno	Sede	Vincitore	Numero Goal	Pubblico
2016	Francia	Portogallo	108	47594
2012	Polonia	Spagna	76	46479
2008	Austria	Spagna	71	36383
2004	Portogallo	Grecia	77	37306
2000	Belgio	Francia	85	35220

Il campo pubblico è il numero medio di spettatori per partite. In ogni edizione sono state giocate 32 partite.

Restituire le squadre che vantano sia vittorie al campionato mondiale sia all'europeo.

7. Creare un nuovo DataFrame concatenando i due precedenti ed aggiungendo una colonna Tipologia, che vale 'MONDIALE' nel caso delle righe provenienti dal Dataframe al punto 1, 'EUROPEO' nel caso di righe provenienti dal punto 6.
8. Calcolare, per ogni Tipologia di competizione, la differenza tra l'edizione con più pubblico e quella con meno pubblico.