

Esempio Modulo Statistica

May 4, 2022

1 Esempi di Utilizzo del MODULO STATISTICA della libreria IntelligenzaArtificiale

1.1 Installare la libreria

```
[ ]: #Per installare la libreria sul tuo computer puoi usare il comando :  
pip3 install intelligenzaartificiale  
  
#se invece utilizzi google colab puoi usare:  
!pip install intelligenzaartificiale
```

1.2 Importare la libreria

```
[1]: from intelligenzaartificiale import dataset as dt  
from intelligenzaartificiale import statistica as st
```

1.3 Importare o Caricare un Set di dati

```
[2]: il_mio_dataset = dt.leggi_csv("exams.csv")
```

Tempo impiegato per leggere il file: 0.004087924957275391

1.4 Statistiche su Valori Nulli o Corrotti

```
[3]: print(st.valori_nan(il_mio_dataset))  
print("\n ----- \n")  
print(st.percentuale_nan(il_mio_dataset))
```

gender	0
race/ethnicity	5
parental level of education	5
lunch	1
test preparation course	5
math score	1
reading score	2
writing score	0
dtype: int64	

```

-----
gender                                0.0
race/ethnicity                        0.5
parental level of education           0.5
lunch                                 0.1
test preparation course               0.5
math score                            0.1
reading score                         0.2
writing score                         0.0
dtype: float64

```

1.5 Statistiche su tutto il Set di dati

```
[4]: print(st.statistiche(il_mio_dataset))
```

	math score	reading score	writing score
count	999.000000	998.000000	1000.000000
mean	66.731732	69.225451	67.986000
std	15.234970	14.578954	15.344291
min	11.000000	23.000000	17.000000
25%	57.000000	59.000000	58.000000
50%	67.000000	69.000000	68.000000
75%	77.500000	80.000000	79.000000
max	100.000000	100.000000	100.000000

1.6 Statistiche su specifica colonna

```

[5]: #statistiche generiche su specifica colonna del dataset
print(st.statistiche_colonna(il_mio_dataset,"math score"))

print("\n ----- \n")

#media
print(st.media(il_mio_dataset,"math score"))

print("\n ----- \n")

#varianza
print(st.varianza_colonna(il_mio_dataset,"math score"))

print("\n ----- \n")

#quantili
print(st.quantile_25(il_mio_dataset,"math score"))
print(st.quantile_50(il_mio_dataset,"math score"))
print(st.quantile_75(il_mio_dataset,"math score"))

```

```
print("\n ----- \n")

#min e max
print(st.min(il_mio_dataset, "math score"))
print(st.max(il_mio_dataset, "math score"))
```

```
count    999.000000
mean     66.731732
std      15.234970
min      11.000000
25%      57.000000
50%      67.000000
75%      77.500000
max      100.000000
Name: math score, dtype: float64
```

```
-----

66.73173173173173

-----

232.10431272956322

-----

57.0
67.0
77.5

-----

11.0
100.0
```

1.7 Analizzare le correlazioni

```
[6]: #Correlazione tra i campi del dataset
print(st.correlazione(il_mio_dataset))

print("\n ----- \n")

#correlazione tra una colonna target e un'altra colonna
print(st.correlazione_radio(il_mio_dataset, "reading score", "math score"))

print("\n ----- \n")
```

```

#correlazione di Spearman tra una colonna target e un'altra colonna
print(st.correlazione_spearman(il_mio_dataset, "reading score", "math score"))

print("\n ----- \n")

#correlazione di Pearson tra una colonna target e un'altra colonna
print(st.correlazione_pearson(il_mio_dataset, "reading score", "math score"))

print("\n ----- \n")

#classifica correlazione tra una colonna target e un'altra colonna
print(st.classifica_correlazione_colonna(il_mio_dataset, "math score"))

```

	math score	reading score	writing score
math score	1.000000	0.822901	0.808995
reading score	0.822901	1.000000	0.952120
writing score	0.808995	0.952120	1.000000

```

-----
0.8229010198957898

```

```

-----
0.8067534176955062

```

```

-----
0.8229010198957898

```

```

-----

```

```

math score      1.000000
reading score    0.822901
writing score    0.808995
Name: math score, dtype: float64

```

2 CREARE UN REPORT AUTOMATIZZATO

```
[7]: st.report_dataset(il_mio_dataset)
```

```

Summarize dataset:  0%|          | 0/5 [00:00<?, ?it/s]
Generate report structure:  0%|          | 0/1 [00:00<?, ?it/s]
Render HTML:  0%|          | 0/1 [00:00<?, ?it/s]

```

```
Export report to file: 0%|          | 0/1 [00:00<?, ?it/s]
Report salvato in questa directory profile_report_pandas.html
<IPython.core.display.HTML object>
```

[7]:

```
[15]: #apri il tuo dataset sul web
      st.apri_dataframe_nel_browser(il_mio_dataset)
      #Ti consigliamo vivamente di provare questa funzione sul tuo set di dati

      ## verrà aperta in automatico una finestra nel tuo browser.
      ### Naviga tra le funzioni ti divertirai
```

2.1 Altre risorse

- [Documentazione Ufficiale](#)
- [Blog Ufficiale](#)
- [Corsi Gratis](#)
- [Ebook Gratis](#)
- [Progetti Python Open Source](#)
- [Dataset Pubblici](#)
- [Editor Python Online per il M.L.](#)

3 Per favore citaci se usi la Libreria.