



UNIVERSITÀ DEGLI STUDI DI MILANO

CONCORSO PUBBLICO, PER TITOLI ED ESAMI, PER IL RECLUTAMENTO DI N. 1 UNITÀ DI PERSONALE AFFERENTE ALL'AREA DELLE ELEVATE PROFESSIONALITÀ - SETTORE SCIENTIFICO-TECNOLOGICO, CON RAPPORTO DI LAVORO SUBORDINATO A TEMPO DETERMINATO PRESSO L'UNIVERSITÀ DEGLI STUDI DI MILANO - DIPARTIMENTO DI SCIENZE CLINICHE E DI COMUNITÀ NELL'AMBITO DEL PROGETTO "DIPARTIMENTI DI ECCELLENZA 2023-2027" - CODICE 22508

La Commissione giudicatrice del concorso, nominata con Determina Direttoriale n. 19978 del 25/11/2024 modificata con Determina Direttoriale n. 21200 del 05/12/2024 e successivamente modificata con Determina Direttoriale n. n. 22051/2024 del 12.12.2024 composta da

Prof.ssa Marta Rossi	Presidente
Dott. Matteo Di Maso	Componente
Dott. Gianfranco Alicandro	Componente
Dott. Carlo Rinaldi	Segretario

comunica le tracce relative alla prima prova:

TRACCIA n. 1

1. Il/La candidato/a descriva i principali indicatori di tendenza centrale.
2. Il/La candidato/a descriva le caratteristiche principali delle diverse tipologie di variabili che possono essere contenute in un data set.

TRACCIA n. 2

1. Il/La candidato/a descriva i principali indicatori di dispersione.
2. Il/La candidato/a spieghi le modalità di gestione dei dati mancanti in un data set.

TRACCIA n. 3

1. Il/La candidato/a descriva le caratteristiche della statistica descrittiva e inferenziale evidenziandone le principali differenze.
2. Il/La candidato/a descriva i vari formati delle variabili di un dataset.

comunica le tracce relative alla seconda prova scritta a contenuto teorico - pratico:

TRACCIA n. 1

Nel file "ProvaPratica.xlsx" sono riportate le informazioni relative a 189 bambini nati al Baystate Medical Center, Springfield, Mass (1996). Le variabili incluse nel data set sono le seguenti:

- id identificativo del bambino;
- mwt peso della madre all'ultimo periodo mestruale espresso in libbre (lb);
- bwt peso del bambino alla nascita espresso in grammi (g).

Utilizzando un software per l'analisi dei dati a scelta del candidato/a,

1. Creare le seguenti variabili:

- mwt_kg peso della madre all'ultimo periodo mestruale espresso in kg (1 lb=0,454 kg);
 - mwt_kg_lt25per peso della madre all'ultimo periodo mestruale in kg inferiore al 25° percentile (dicotomica; 0: peso della madre all'ultimo periodo mestruale in kg \geq del 25° percentile, 1: peso della madre all'ultimo periodo mestruale in kg < del 25° percentile);
 - low_bwt bambino nato sottopeso (dicotomica 0: ≥ 2500 g; 1: peso < 2500 g);
- e riportare le prime 10 righe del dataset con tutte le variabili.

2. Costruire una tabella di contingenza a due vie che riporti la variabile dicotomica mwt_kg_lt25per (per riga) e la variabile dicotomica low_bwt (per colonna), specificando le frequenze assolute e le frequenze relative di riga.

3. Costruire un boxplot del peso del neonato alla nascita nelle madri con peso all'ultimo periodo mestruale in kg < del 25° percentile e un boxplot del peso del neonato alla nascita nelle madri con peso all'ultimo periodo mestruale in kg \geq del 25° percentile.

4. Il/La candidato/a spieghi l'operazione cross join tra due data set.



TRACCIA n. 2

Nel file "ProvaPratica.xlsx" sono riportate le informazioni relative a 189 bambini nati al Baystate Medical Center, Springfield, Mass (1996). Le variabili incluse nel data set sono le seguenti:

- id identificativo del bambino;
- mwt peso della madre all'ultimo periodo mestruale espresso in libbre (lb);
- bwt peso del bambino alla nascita espresso in grammi (g).

Utilizzando un software per l'analisi dei dati a scelta del candidato/a,

1. Creare le seguenti variabili:

- mwt_kg peso della madre all'ultimo periodo mestruale espresso in kg (1 lb=0,454 kg);
- mwt_kg_ltmed peso della madre all'ultimo periodo mestruale in kg inferiore alla mediana (dicotomica; 0: peso della madre all'ultimo periodo mestruale in kg \geq della mediana, 1: peso della madre all'ultimo periodo mestruale in kg $<$ della mediana);
- low_bwt bambino nato sottopeso (dicotomica 0: ≥ 2500 g; 1: peso < 2500 g);

e riportare le prime 10 righe del dataset con tutte le variabili.

2. Costruire una tabella di contingenza a due vie che riporti la variabile dicotomica mwt_kg_ltmed (per riga) e la variabile dicotomica low_bwt (per colonna), specificando le frequenze assolute e le frequenze relative di riga.

3. Costruire un boxplot del peso del neonato alla nascita nelle madri con peso all'ultimo periodo mestruale in kg $<$ della mediana e un boxplot del peso del neonato alla nascita nelle madri con peso all'ultimo periodo mestruale in kg \geq della mediana.

4. Il/La candidato/a spieghi la differenza tra le operazioni set e merge di due data set.

TRACCIA n. 3

Nel file "ProvaPratica.xlsx" sono riportate le informazioni relative a 189 bambini nati al Baystate Medical Center, Springfield, Mass (1996). Le variabili incluse nel data set sono le seguenti:

- id identificativo del bambino;
- mwt peso della madre all'ultimo periodo mestruale espresso in libbre (lb);
- bwt peso del bambino alla nascita espresso in grammi (g).

Utilizzando un software per l'analisi dei dati a scelta del candidato/a,

1. Creare le seguenti variabili:

- mwt_kg peso della madre all'ultimo periodo mestruale espresso in kg (1 lb=0,454 kg);
- mwt_kg_ltmed peso della madre all'ultimo periodo mestruale in kg inferiore alla media (dicotomica; 0: peso della madre all'ultimo periodo mestruale in kg \geq della media, 1: peso della madre all'ultimo periodo mestruale in kg $<$ della media);
- low_bwt bambino nato sottopeso (dicotomica 0: ≥ 2500 g; 1: peso < 2500 g);

e riportare le prime 10 righe del dataset con tutte le variabili.

2. Costruire una tabella di contingenza a due vie che riporti la variabile dicotomica mwt_kg_ltmed (per riga) e la variabile dicotomica low_bwt (per colonna), specificando le frequenze assolute e le frequenze relative di riga.

3. Costruire un boxplot del peso del neonato alla nascita nelle madri con peso all'ultimo periodo mestruale in kg $<$ della media e un boxplot del peso del neonato alla nascita nelle madri con peso all'ultimo periodo mestruale in kg \geq della media.

4. Il/La candidato/a spieghi l'operazione inner join tra due data set.

Milano, 17 dicembre 2024

La Commissione

Prof.ssa Marta Rossi Presidente

Dott. Matteo Di Maso Componente

Dott. Gianfranco Alicandro Componente

Dott. Carlo Rinaldi Segretario