

Laboratorio 2 - Misure Ripetute

Jacopo D'alto, Sonia Spinelli

November 2023

INTRODUZIONE

Il dataset preso in considerazione raccoglie dati sulle pulsazioni cardiache di 30 individui. La popolazione è divisa in base al regime alimentare, poco grasso o grasso, e al modo con cui esegue esercizio ginnico, da ferma, camminando o correndo. Queste rilevazioni vengono, inoltre, ripetute in tre momenti differenti: un minuto, un quarto d'ora e mezz'ora dopo l'inizio dell'attività fisica.

ANALISI DESCRITTIVA

La variabile di interesse è **PULSE**, misura quantitativa che conta le pulsazioni cardiache in battiti al minuto. In generale, un adulto, non sottoposto a sforzo, registra dai 60 ai 100 battiti al minuto.

Le altre variabili che caratterizzano il dataset sono qualitative:

- **DIET** indica il regime alimentare (1 = poco grasso, 2 = grasso).
- **EXERTYPE** rappresenta la modalità con cui si svolge esercizio fisico (1 = da fermo, 2 = camminando, 3 = correndo).
- **TIME** indica il momento di misurazione dall'inizio dell'attività (1 = dopo un minuto, 2 = dopo 15 minuti, 3 = dopo 30 minuti).

Si osserva che ogni variabile categorica partiziona la popolazione in modo bilanciato per ciascun livello.

Variabile PULSE						
TIME	Taglia	Minimo	Media	Massimo	Range	Deviazione
dopo 1 minuto	30	80	93.133	103	23	6.152
dopo 15 minuti	30	82	101.533	135	53	14.564
dopo 30 minuti	30	83	104.433	150	67	18.876

Tabella 1: Summary della variabile **PULSE** al variare di **TIME**

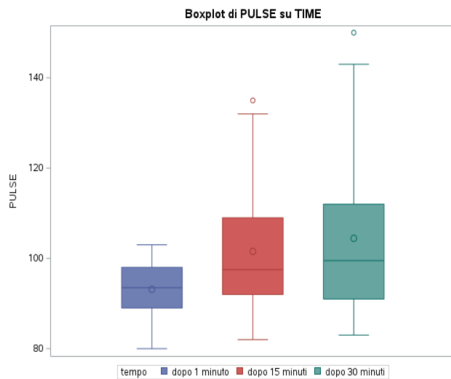


Figura 1:

Boxplot di PULSE al variare di TIME.

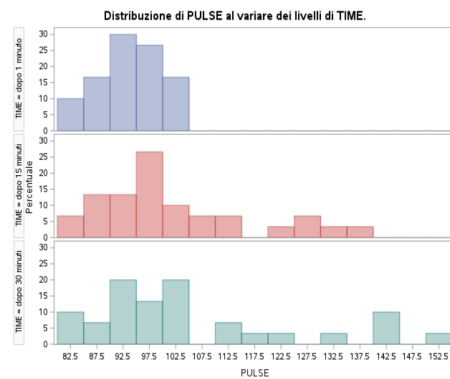


Figura 2:

Boxplot di PULSE al variare di TIME.

Inizialmente si studia la distribuzione delle pulsazioni nel tempo, come nelle *Figure 1 e 2* e nella *Tabella 1*. Dopo un minuto di attività i valori sono concentrati verso il basso, dopo 15 minuti tendono a disperdersi e dopo 30 minuti i dati sono i più dispersi e con valori più alti. Si decide di modificare il dataset, considerando separatamente le pulsazioni nei tre tempi per ciascun individuo, e si osserva il comportamento delle tre nuove variabili in base alla dieta e alla modalità di esercizio fisico:

- PULSE1: pulsazioni dopo un minuto
- PULSE2: pulsazioni dopo 15 minuti
- PULSE3: pulsazioni dopo 30 minuti

Come si osserva dalle *Figure 3 e 4* e dalla *Tabella 2*, le pulsazioni dopo un minuto sono tendenzialmente più alte quando l'individuo segue una dieta "grassa" e analogamente, nelle *Figure 5 e 6* e nella *Tabella 3*, sembra che più la modalità di attività fisica sia intensa più i battiti tendono a concentrarsi in valori più alti. In generale, la dispersione dei dati decresce all'aumentare dell'intensità dell'esercizio.

Variabile PULSE1						
DIET	Taglia	Minimo	Media	Massimo	Range	Deviazione
poco grasso	15	80	91.067	98	18	5.147
grasso	15	83	95.200	103	20	6.538

Tabella 2: Summary della variabile PULSE1 al variare di DIET

Variabile PULSE1						
EXERTYPE	Taglia	Minimo	Media	Massimo	Range	Deviazione
da fermo	10	80	90.200	100	20	6.545
camminando	10	84	93.100	103	19	6.297
correndo	10	87	96.100	103	16	4.483

Tabella 3: Summary della variabile PULSE1 al variare di EXERTYPE

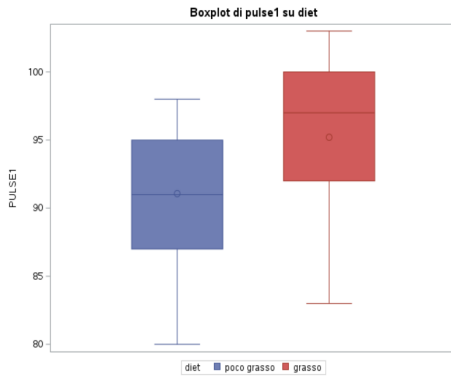


Figura 3 :
Boxplot di PULSE1 al variare di DIET.

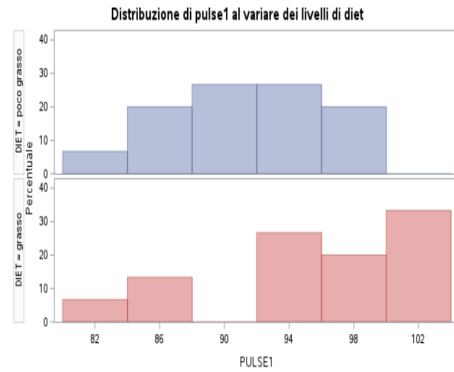


Figura 4:
Istogramma di PULSE1 al variare di DIET.

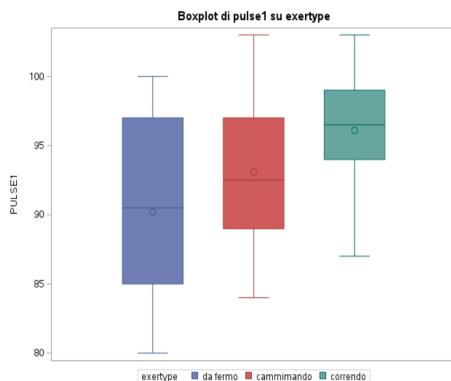


Figura 5:
Boxplot di PULSE1 al variare di EXERTYPE.

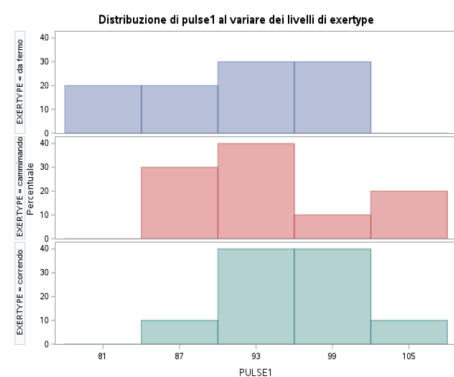


Figura 6:
Istogramma di PULSE1 al variare di EXERTYPE.

Dopo 15 minuti dall'inizio dell'attività, si nota dalle *Figure 7 e 8* che le prime metà dei dati delle due diete sono pressochè simili, infatti le mediane sono molto vicine, ma si accentua la differenza nelle seconde metà, dove per un

regime alimentare "grasso" si ha una dispersione maggiore. Dalle *Figure 9* e *10* si deduce che in base al tipo di esercizio la distribuzione è diversa: in particolare, se l'individuo corre i valori delle pulsazioni sono tendenzialmente più elevati e più dispersi rispetto ai battiti di chi cammina o sta fermo, che tra loro hanno simile concentrazione e una leggera differenza di valori.

Variabile PULSE2						
DIET	Taglia	Minimo	Media	Massimo	Range	Deviazione
poco grasso	15	82	98.000	132	50	12.236
grasso	15	83	105.067	135	52	16.2149

Tabella 4: Summary della variabile PULSE2 al variare di DIET

Variabile PULSE2						
EXERTYPE	Taglia	Minimo	Media	Massimo	Range	Deviazione
da fermo	10	82	90.900	99	17	6.118
camminando	10	86	96.600	109	23	7.442
correndo	10	98	117.100	135	37	12.991

Tabella 5: Summary della variabile PULSE2 al variare di EXERTYPE

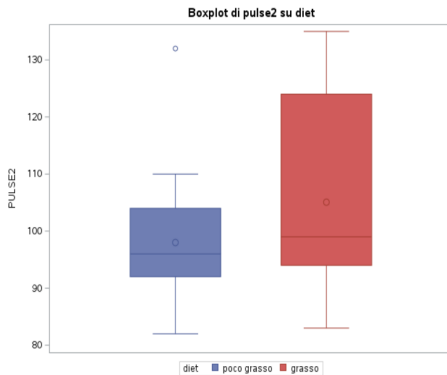


Figura 7:

Boxplot di PULSE2 al variare di DIET.

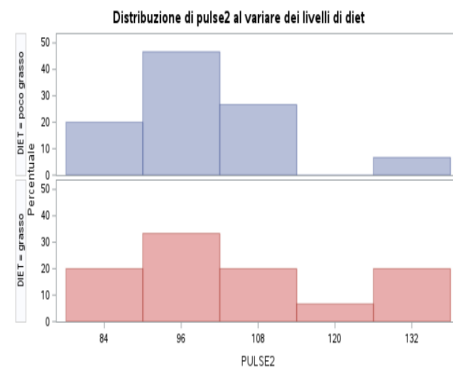


Figura 8:

Istogramma di PULSE2 al variare di DIET.

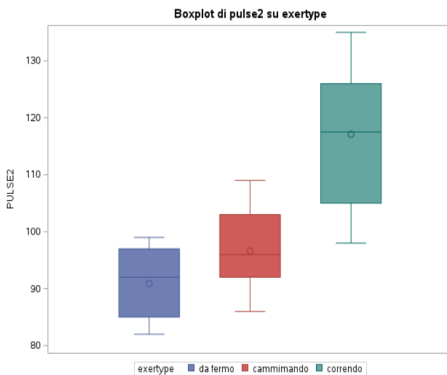


Figura 9:

Boxplot di PULSE2 al variare di EXERTYPE.

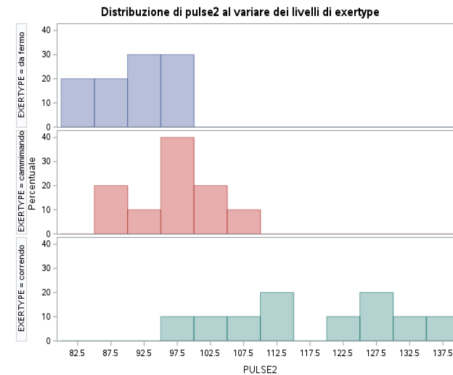


Figura 10:

Istogramma di PULSE2 al variare di EXERTYPE.

Dai grafici in *Figura 11* e *12*, si nota che la distribuzione dei valori dopo 30 minuti di esercizio non varia in modo significativo rispetto al tempo precedente, varia solo per un leggero aumento della dispersione nel livello poco grasso.

In base al tipo di esercizio, nelle *Figure 13* e *14* si accentua il fenomeno già osservato per le pulsazioni rilevate dopo 15 minuti.

Variabile PULSE3						
DIET	Taglia	Minimo	Media	Massimo	Range	Deviazione
poco grasso	15	83	98.800	120	37	11.416
grasso	15	84	110.067	150	66	23.233

Tabella 6: Summary della variabile PULSE3 al variare di DIET

Variabile PULSE3						
EXERTYPE	Taglia	Minimo	Media	Massimo	Range	Deviazione
da fermo	10	83	91.400	100	17	5.337
camminando	10	84	95.900	104	20	6.740
correndo	10	99	126.000	150	51	16.964

Tabella 7: Summary della variabile PULSE3 al variare di EXERTYPE

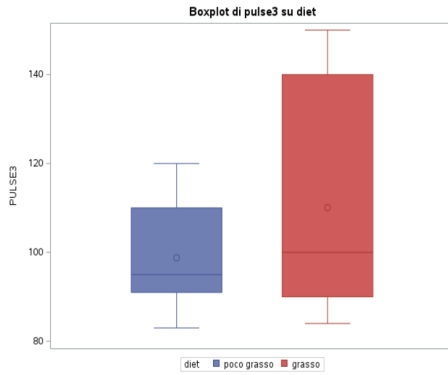


Figura 11:

Boxplot di PULSE3 al variare di DIET.

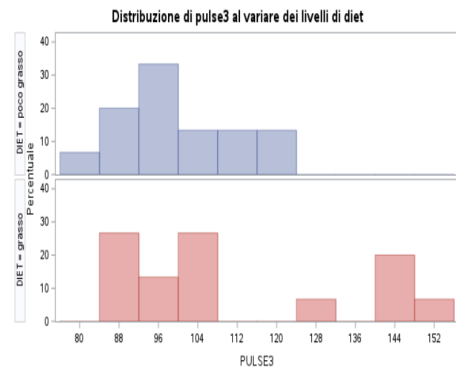


Figura 12:

Istogramma di PULSE3 al variare di DIET.

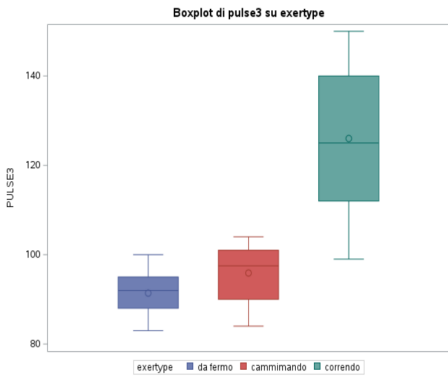


Figura 13:

Boxplot di PULSE3 al variare di EXERTYPE.

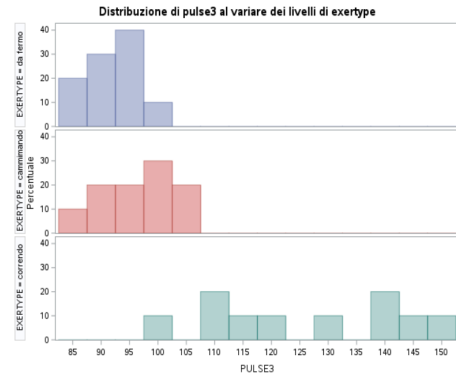


Figura 14:

Istogramma di PULSE3 al variare di EXERTYPE.

Da questo insieme di considerazioni, si potrebbe supporre che la modalità di attività fisica incida significativamente sui valori delle pulsazioni, mentre la dieta incida meno. Tuttavia, serve effettuare ulteriori verifiche per supportare o rifiutare questa ipotesi. Nello specifico, si effettua un'analisi della varianza a misure ripetute.

ANALISI DELLA VARIANZA DI MISURE RIPETUTE

Fissando il livello di significatività dei test d'ipotesi a 0.01, si effettua un'analisi della varianza a misure ripetute. Partendo dall'analisi univariata delle tre variabili risposta, si cerca un effetto globale delle variabili DIET ed EXERTYPE. Per PULSE1, come si nota nella Tabella 8, non si rifiuta l'ipotesi nulla, quindi la dieta e il tipo di esercizio non hanno effetto sulle pulsazioni misurate al primo minuto. A conferma di ciò, il test, riportato in Tabella 9, sull'incidenza dei diversi livelli dei due fattori non è rifiutato.

Tabella 8: Test d'ipotesi univariato per la variabile PULSE1

Origine	DF	SS Tipo I	Media quadratica	Valore F	Pr >F
DIET	1	128.1333333	128.1333333	4.19	0.0509
EXERTYPE	2	174.0666667	87.0333333	2.85	0.0763

Tabella 9: Test d'ipotesi univariato per la variabile PULSE1 con la distinzione in livelli

Parametro	Stima		Errore standard	Valore t	Pr >—t—
Intercept	88.134	B	2.01947780	43.64	<.0001
DIET grasso	4.134	B	2.01947780	2.05	0.0509
DIET poco grasso	0.000	B	.	.	.
EXERTYPE camminando	2.900	B	2.47334508	1.17	0.2516
EXERTYPE correndo	5.900	B	2.47334508	2.39	0.0246
EXERTYPE da fermo	0.000	B	.	.	.

Per PULSE2, il test non è rifiutato per la variabile DIET, ma lo è per la variabile EXERTYPE, come si nota nella Tabella 10. Ciò implica che la dieta non ha un effetto significativo sulle pulsazioni al quindicesimo minuto di attività, mentre il tipo di esercizio sì. Effettuando i test sui livelli della variabile EXERTYPE, si ottiene che questi hanno incidenza sui valori di PULSE2: col passare di 15 minuti, il tipo di esercizio comporta una variazione nelle pulsazioni.

Tabella 10: Test d'ipotesi univariato per la variabile PULSE2

Origine	DF	SS Tipo I -	Media quadratica	Valore F	Pr >F
DIET	1	374.533333	374.533333	4.92	0.0355
EXERTYPE	2	3797.266667	1898.633333	24.94	<.0001

Tabella 11: Test d'ipotesi univariato per la variabile PULSE2 con la distinzione in livelli

Parametro	Stima		Errore standard	Valore t	Pr >—t—
Intercept	87.367	B	3.18624179	27.42	<.0001
DIET grasso	7.0667	B	3.18624179	2.22	0.0355
DIET poco grasso	0.000	B	.	.	.
EXERTYPE camminando	5.700	B	3.90233329	1.46	0.1561
EXERTYPE correndo	26.200	B	3.90233329	6.71	<.0001
EXERTYPE da fermo	0.000	B	.	.	.

Per PULSE3 si rifiutano tutti i test nelle Tabelle 12 e 13: entrambi i fattori e i relativi livelli hanno un impatto sulle misurazioni.

Tabella 12: Test d'ipotesi univariato per la variabile PULSE3

Origine	DF	SS Tipo I -	Media quadratica	Valore F	Pr >F
DIET	1	952.033333	952.033333	10.75	0.0030
EXERTYPE	2	7078.066667	3539.033333	39.95	<.0001

Tabella 13: Test d'ipotesi univariato per la variabile PULSE3 con la distinzione in livelli

Parametro	Stima		Errore standard	Valore t	Pr >—t—
Intercept	85.767	B	3.43680432	24.96	<.0001
DIET grasso	11.267	B	3.43680432	3.28	0.0030
DIET poco grasso	0.000	B	.	.	.
EXERTYPE camminando	4.500	B	4.20920846	1.07	0.2949
EXERTYPE correndo	34.600	B	4.20920846	8.22	<.0001
EXERTYPE da fermo	0.000	B	.	.	.

Si passa ora a studiare l'analisi della varianza a misure ripetute e si sceglie di verificare l'invarianza della variabile risposta al tempo successivo, selezionando la matrice M come segue.

$$M = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$$

La scelta di questa parametrizzazione deriva dal fatto che le misurazioni sono effettuate in maniera ordinata rispetto al tempo. Perciò, sembra interessante valutare la differenza che occorre tra un tempo e il successivo.

Per l'ipotesi di nessun effetto del tempo in *Tabella 14*, si rifiuta rispetto a tutte le quattro statistiche test proposte dal software: i battiti, in generale, subiscono l'effetto del tempo.

Criteri di test MANOVA e statistiche F esatte per l'ipotesi di nessun effetto TIME H = Tipo III - Matrice SSCP per TIME					
Statistica	Valore	Valore F	DF num	DF den	Pr >F
Lambda di Wilks	0.34766180	23.45	2	25	<.0001
Traccia di Pillai	0.65233820	23.45	2	25	<.0001
Traccia Hotelling-Lawley	1.87635861	23.45	2	25	<.0001
Radice massima di Roy	1.87635861	23.45	2	25	<.0001

Tabella 14: test sull'effetto del tempo.

Complessivamente, nella *Tabella 15* non si rifiuta l'invarianza del tipo di dieta al passare del tempo sulle pulsazioni. Questo ha corrispondenza nei test univariati, dove si era concluso che il regime alimentare aveva un effetto significativo solo per le misurazioni dopo 30 minuti.

Criteri di test MANOVA e statistiche F esatte per l'ipotesi di nessun effetto TIME*DIET H = Tipo III - Matrice SSCP per TIME*DIET					
Statistica	Valore	Valore F	DF num	DF den	Pr >F
Lambda di Wilks	0.86022198	2.03	2	25	0.1523
Traccia di Pillai	0.13977802	2.03	2	25	0.1523
Traccia Hotelling-Lawley	0.16249064	2.03	2	25	0.1523
Radice massima di Roy	0.16249064	2.03	2	25	0.1523

Tabella 15: test sull'effetto del regime alimentare nel tempo.

Invece, nella *Tabella 16* si rifiuta l'ipotesi che il tipo di esercizio al variare del tempo non influenzi la frequenza cardiaca.

Criteri di test MANOVA e statistiche F esatte per l'ipotesi di nessun effetto TIME*EXERTYPE H = Tipo III - Matrice SSCP per TIME*EXERTYPE					
Statistica	Valore	Valore F	DF num	DF den	Pr >F
Lambda di Wilks	0.28927192	10.74	4	50	<.0001
Traccia di Pillai	0.71702894	7.27	4	52	0.0001
Traccia Hotelling-Lawley	2.43517324	15.04	4	28.992	<.0001
Radice massima di Roy	2.42619549	31.54	2	26	<.0001

Tabella 16: test sull'effetto della modalità di attività fisica nel tempo.

Considerando la media di tutte le pulsazioni in *Tabella 17*, si rifiuta l'ipotesi della non significatività di entrambe le covariate. Ciò potrebbe essere spiegato dal fatto che DIET influisce significativamente su PULSE3, che assume i valori più alti e, quindi, contribuisce maggiormente al valor medio.

Test di ipotesi per effetti tra soggetti					
Origine	DF	SS Tipo III -	Media quadratica	Valore F	Pr >F
DIET	1	1261.877778	1261.877778	11.31	0.0024
EXERTYPE	2	8326.066667	4163.033333	37.31	<.0001
Errore	26	2900.955556	111.575214		

Tabella 17: test univariato sulla media delle pulsazioni.

Distinguendo i test per i diversi tempi di misurazione in *Tabella 18*, si osserva che, fissato il tipo di esercizio, il variare del tempo ha impatto maggiore nei primi 15 minuti rispetto ai 15 successivi. La dieta, invece, non influisce in nessuno dei due casi.

Analisi della varianza di variabili contrasto											
Variabile del contrasto: time_1						Variabile del contrasto: time_2					
Origine	DF	SS Tipo III -	Media quadratica	Valore F	Pr >F	Origine	DF	SS Tipo III -	Media quadratica	Valore F	Pr >F
Mean	1	2116.800000	2116.800000	28.46	<.0001	Mean	1	252.300000	252.300000	2.92	0.0995
DIET	1	64.533333	64.533333	0.87	0.3602	DIET	1	132.300000	132.300000	1.53	0.2271
EXERTYPE	2	2420.600000	1210.300000	16.27	<.0001	EXERTYPE	2	547.200000	273.600000	3.17	0.0588
Errore	26	1934.066667	74.387179			Errore	26	2247.200000	86.430769		

Tabella 18: test per varaibili di contrasto.

Per effettuare i test univariati per effetti entro soggetti, è necessario verificare l'ipotesi di sfericità, che, per componenti ortogonali, non si rifiuta. Tuttavia, si osserva che per le variabili trasformate il *p-value*, seppur non alto, non è abbastanza basso. Di conseguenza, le conclusioni del test sono da interpretare con cautela. Nonostante ciò, la Tabella 20 conferma quanto già osservato: non si rifiuta per EXERTYPE e si rifiuta per DIET.

Test di sfericità				
Variabili	DF	Criterio di Mauchly	Chi-quadrato	Pr >ChiQuadr
Variate trasformate	2	0.8026721	5.4952234	0.0641
Componenti ortogonali	2	0.9867753	0.3328222	0.8467

Tabella 19: test di sfericità.

Test univariati di ipotesi per effetti entro soggetti						Pr corr >F	
Origine	DF	SS Tipo III -	Media quadratica	Valore F	Valore F	G - G	H-F-L
TIME	2	2066.600000	1033.300000	24.68	<.0001	<.0001	<.0001
TIME*DIET	2	192.822222	96.411111	2.30	0.1101	0.1109	0.1101
TIME*EXERTYPE	4	2723.333333	680.833333	16.26	<.0001	<.0001	<.0001
Errore(TIME)	52	2177.244444	41.870085				

Tabella 20: test univariati per effetti entro soggetti.

CONCLUSIONE

Dall'analisi finora condotta emergono le seguenti considerazioni sui contributi che il tipo di dieta e l'intensità dell'attività fisica hanno sulla frequenza cardiaca.

Il regime alimentare contribuisce solo nelle misurazioni dopo 30 minuti dall'inizio dell'attività. Tuttavia ciò non è abbastanza da renderlo un fattore globalmente incidente sul valore delle pulsazioni e si rifiuta l'idea che abbia un impatto sul ritmo cardiaco. Il tipo di esercizio fisico, invece, sebbene nella misurazione effettuata dopo un minuto non sembra essere discriminante, al passare del tempo pesa significativamente.

Si conclude che, in generale, la frequenza cardiaca al variare del tempo non è influenzata dalla dieta, ma dalla modalità di attività fisica che viene eseguita dall'individuo.