



Università degli Studi di Padova

Facoltà di Scienze Statistiche

Corso di laurea in Statistica e Tecnologie Informatiche

Uno studio preliminare dello stile di vita degli adolescenti

Relatore: Prof.ssa Laura Ventura

Correlatore: Prof. Attilio Carraro

Laureando: Nicola Bruni

Anno Accademico 2007/2008

Non posso certo dimenticare di ringraziare e stringere in un abbraccio tutti coloro che mi hanno accompagnato in questo rischioso percorso che ha segnato la mia vita. Ringrazio la mia famiglia per l'appoggio offertomi, Annapaola che è stata determinante per l'arrivo alla meta, tutti gli amici della Facoltà, Fiore, Fabrizio, Paolo, Andrea, gli amici di avventure di Piazzola, i rockers Giorgio, Maurizio e Ale. Un particolare ringraziamento alla prof.ssa Laura Ventura per la pazienza e la grande disponibilità.

INDICE DEI CONTENUTI

	Pag.
INTRODUZIONE	7
CAPITOLO 1 - Obesità: Questionario e dati	
1.1 Lo studio dell'obesità infantile	9
1.2 L'obesità in Italia	10
1.3 Questionario	17
1.4 Variabili	18
1.5 Dati mancanti	20
CAPITOLO 2 - Analisi esplorative	
2.1 Alcuni dati semplici sugli studenti	22
2.2 Aspetto genitori	26
2.3 Qualche domanda sullo sport	27
2.4 Alcune cose semplici sui genitori e sulla casa	33
2.5 Stato di provenienza dei genitori	35
2.6 Giudizio ultime pagelle	36
2.7 Con quale mezzo si va a scuola e quanto ci si impiega	39
2.8 Giudica te stesso	40
2.9 Qualcosa di più sull'alimentazione	41
2.10 Uso della televisione, computer o videogames	44
2.11 Qualcosa di difficile da rinunciare	53
2.12 Il Body Mass Index	53
2.13 Analisi della correlazione tra gli indici	55
CAPITOLO 3 - Analisi delle corrispondenze e analisi cluster	
3.1 Analisi delle corrispondenze	59
3.2 Variabili illustrative	61
3.3 Analisi cluster	65
CONCLUSIONI	70
APPENDICE	72
RIFERIMENTI BIBLIOGRAFICI	93

INTRODUZIONE

La tesi si prefigge di analizzare lo stile di vita (*lifestyle*) di 1005 adolescenti residenti in Veneto e Friuli Venezia Giulia, di età compresa tra gli 8,5 e gli 11 anni, attraverso lo studio di un questionario da loro compilato e che ne evidenzia vari aspetti della vita quotidiana.

Il questionario è parte di una ricerca che prende il nome di *Children's Lifestyle*, un progetto che comprende altri sette paesi della comunità europea ed è coordinato dall'Università di Colonia, in Germania. Per l'Italia il progetto è seguito dal Prof. A. Carraro, del Dipartimento di Scienze dell'Educazione dell'Università di Padova.

In particolare lo scopo di questa analisi è evidenziare quali possono essere le cause che portano ad una variazione di peso eccessiva nell'età adolescente. Negli ultimi anni si è, infatti, sentito spesso parlare di sovrappeso, una parola che inconsciamente viene associata alla fase adulta di un essere umano. Ritmi frenetici, lavori sedentari e cibi scadenti sono le principali cause di questa nuova patologia riconosciuta a livello mondiale. Molti paesi hanno visto crescere, tra la propria popolazione, molte nuove patologie e decessi legati al sovrappeso.

Ciò che non si tiene presente è che tale patologia non caratterizza solo la fase adulta della vita, ma si presenta fin dai primi anni di età con conseguenze non trascurabili nella vita adulta.

Paesi come Inghilterra e Francia si sono già attivati per contrastare tale fenomeno studiando nuove tasse che disincentivano l'acquisto dei cibi ritenuti troppo grassi, promuovendo pasti più leggeri nelle mense, diffondendo una campagna di prevenzione per sensibilizzare l'opinione pubblica e inserendo avvertimenti espliciti in tutte le confezioni di alimenti ricchi di grasso, sale e burro. Secondo l'Organizzazione mondiale della sanità (OMS) questa malattia non trasmissibile interessa nel mondo oltre 300 milioni di adulti ed è in costante aumento tra i bambini.

La ricerca e lo studio perciò costituiscono uno strumento per valutare gli stili di vita e i comportamenti degli adolescenti e un'utile fonte di informazioni per la comunità scientifica e per i responsabili delle politiche di promozione della salute.

Con la somministrazione di un questionario appositamente redatto, si è cercato di cogliere numerosi aspetti fisici, psicologici e sociali. Scopo della tesi è analizzare questi aspetti e delineare attraverso lo studio di particolari indicatori quali l'indice di obesità infantile (*BMI – Body Mass index*), l'indice di benessere familiare (*FAS – Family Affluence Scale*) e l'indice di utilizzo dei mezzi di intrattenimento (*MC - Media Consumption*), lo stile di vita degli adolescenti.

Il lavoro è suddiviso in vari capitoli.

Nel primo capitolo vengono presentati il questionario e le variabili rilevate, spiegandone la natura e procedendo al controllo di eventuali errori presenti nel dataset. Nel secondo capitolo,

viene presentata un'analisi esplorativa del dataset così corretto facendo uso prevalentemente di grafici, informazioni di sintesi, alcuni test per la per la verifica di ipotesi e tecniche bivariate. Nel terzo capitolo si utilizzano invece tecniche più avanzate quali l'analisi delle corrispondenze multiple e l'analisi cluster, utili per capire la struttura della dipendenza interna delle variabili e individuare l'esistenza di gruppi omogenei.

OBESITA': QUESTIONARIO E DATI

In questo capitolo si contestualizza in modo sintetico il problema dell'obesità nell'età infantile, per capire i motivi che hanno spinto allo studio di tale fenomeno. Si procede successivamente alla presentazione del questionario, focalizzando l'attenzione sulle varie parti che lo compongono e i vari aspetti della vita dell'adolescente a cui è rivolto. Per ultimo, si presentano tutte le variabili presenti nel questionario, specificandone la tipologia e controllando l'eventuale presenza di dati mancanti o errati.

1.1 LO STUDIO DELL'OBESITA' INFANTILE

Può essere utile dare una definizione di obesità, intesa come una condizione cronica in cui l'eccesso di tessuto adiposo è in grado di indurre un aumento di rischi per la salute (malattie cardiovascolari, pressione alta, diabete, ipercolesterolemia) e non meno importante a problemi relazionali e di qualità della vita. L'obesità risulta essere il disordine nutrizionale più frequente nei paesi sviluppati e in particolare l'obesità infantile è certamente uno dei problemi più frequenti in età pediatrica. L'obesità, in base ad un rapporto pubblicato dall' Ufficio Regionale Europeo e dall'Organizzazione Mondiale della Sanità (OMS), è definita come una vera e propria epidemia estesa a tutta l'Europa. Il trend più preoccupante riguarda soprattutto i più "piccoli", che si trovano a fare i conti con uno stile di vita profondamente modificatosi negli ultimi vent'anni. La tecnologia e l'abbondanza di cibo presenti giornalmente lasciano poco spazio al movimento spontaneo e ai giochi, affetti talvolta carenti rimpiazzati da cibo e televisione: comportamenti alimentari alterati e molto altro portano a quello che oggi viene definita come una malattia.

L'obesità infantile risulta complessa anche perché legata a diverse cause più o meno evidenti che interagiscono tra loro: cattiva ed eccessiva alimentazione, sedentarietà e fattori legati a patologie ereditarie e ambientali.

Ci sono numerose associazioni sia in ambito governativo che universitario che si propongono per lo studio di tale fenomeno, quali ad esempio l'Organizzazione Mondiale della Sanità (OMS), l'Ufficio Regionale Europeo (URE), citati in precedenza, oltre a enti statali specifici di ogni paese europeo come l'ISTAT (Istituto Nazionale di Statistica) e l' Istituto Superiore di Sanità. Oltre a queste

organizzazioni si affiancano altri istituti privati che svolgono ricerche su commissione presso l'opinione pubblica per ricerche sia sociali che di mercato come ad esempio la Doxa.

Un importante progetto creato specificatamente per approfondire le conoscenze sullo stato di salute e sugli stili di vita dei giovani in età scolare è lo studio HBSC (*Health Behaviour in School Aged Children*), messo a punto più di 20 anni fa e condotto da una rete internazionale di ricercatori in collaborazione sia con l'URE sia con l'OMS. Obiettivo del progetto è sviluppare una nuova visione dello stato di salute, del benessere psicofisico, degli stili di vita dei giovani e del contesto in cui vivono.

Dalla fondazione del progetto che vedeva solo ricercatori appartenenti a tre stati si è arrivati ad un crescente numero di paesi e regioni che hanno aderito a questo studio.

La ricerca HBSC prende in considerazione lo stato di salute degli adolescenti nel senso più ampio del termine monitorando molteplici caratteristiche che fanno parte integrante della vita quotidiana dei giovani e che presumibilmente si riflettono con l'andare del tempo in termini di "malattie". Le caratteristiche prese in esame sono la famiglia, la scuola, il gruppo degli amici, la realtà del luogo in cui vivono e crescono, tutti fattori che influenzano lo stato di salute e lo stile di vita.

1.2 L'OBESITA' IN ITALIA

Le ricerche sull'obesità, o più in generale sull'alimentazione e la sedentarietà dei giovani, sono da anni portate avanti per controllare il fenomeno e la sua crescita nelle varie regioni del mondo.

In questo paragrafo vengono richiamati alcuni studi relativi agli adolescenti residenti in Italia, rilevati ed elaborati nei primi anni del 2000, utili nei prossimi capitoli di questa tesi per un confronto con i nuovi dati analizzati.

Una delle ricerche effettuate da HBSC nell'anno 2001-2002 sull'attività fisica e la sedentarietà per un campione di adolescenti, appartenenti ad una fascia di età compresa tra gli 11 e 15 anni, viene così riassunto: per gli adolescenti italiani l'attività fisica praticata appare strettamente associata con la frequentazione di amici nel pomeriggio, dopo la scuola e presso associazioni sportive, piuttosto che come componente principale dello stile di vita. La sedentarietà appare poco influenzata dai determinanti sociali (percezione di sé, gruppo dei pari, contesto familiare, capitale sociale), mentre il livello culturale (in particolare del padre) sembra avere un centrale ruolo protettivo.

La suddivisione dei determinanti sociali e il loro significato è illustrato in Tabella 1.1, le differenze sociali agiscono in gran parte nel determinare o influenzare gli stili di vita, e attraverso essi, la salute. Per stile di vita si intende "un modo di vivere impostato secondo modelli di comportamento

identificabili, che sono frutto dell'azione reciproca delle caratteristiche proprie di un individuo, delle interazioni sociali con le condizioni di vita di carattere socioeconomico e ambientale.” (dal Glossario della promozione della salute, World Health Organization, 1998).

Da tale concetto si capisce come lo stile di vita sia un'importante area di intervento rispetto alla salute dei giovani per poter comprendere ed agire in modo appropriato.

Un altro fattore importante risulta l'attività fisica in quanto è un effetto preventivo e di protezione rispetto al rischio di obesità e sovrappeso e contribuisce ad un corretto ed armonioso sviluppo del corpo in termini di apparato muscolare (flessibilità, tonicità), scheletrico (postura, elasticità), metabolismo e pressione arteriosa. Inoltre sul piano psico-sociale da la capacità di socializzare e maggiore stima di sé. Le raccomandazioni OMS enfatizzano l'impatto sul benessere di una pratica quotidiana e regolare di attività fisica moderata con la cumulazione complessiva di 60 minuti minimo quotidiani di qualsiasi attività che faccia accelerare il battito cardiaco.

L'altro fattore considerato, la sedentarietà, fa correre il rischio di obesità e sovrappeso con la probabilità di protrarre uno stile di vita sedentario nell'età adulta comportando rischi ancora più seri, inoltre sul piano psico-sociale si hanno difficoltà di socializzazione, carente stima di sé, passività rispetto al proprio ambiente scolastico e agli amici.

Le modalità di misurazione del rischio di sedentarietà prendono in considerazione diversi comportamenti (tempo passato dai giovani davanti alla TV, computer, compiti). Secondo le più recenti indicazioni dell'OMS la misura che definisce il limite oltre al quale guardare la televisione diventa un indice di rischio di sedentarietà è 4 o più ore al giorno.

Di seguito sono riportati alcuni grafici di sintesi tratti da un'indagine HBSC (cfr. Report HBSC Italia), dell'attività fisica svolta durante la settimana, in confronto con la media europea, e della sedentarietà, misurata attraverso le ore passate a guardare la televisione. Interessanti sono anche i grafici in cui sono riportate le correlazioni tra i quattro determinanti sociali associati al livello di pratica sportiva e al rischio di sedentarietà, graficamente sono raggruppate considerando per ogni determinante sociale e gruppo d'età le correlazioni statisticamente significative rilevate all'interno di ciascuna categoria (Tabella 1.1).

Percezione di sé	<ul style="list-style-type: none"> • Soddisfazione per la propria vita • Immagine corporea • Percezione della propria salute
Gruppo dei pari	<ul style="list-style-type: none"> • Frequentazione amici dopo la scuola • Numero di amici veri (sesso opposto) • Difficoltà a fare nuove amicizie
Contesto familiare	<ul style="list-style-type: none"> • Titolo di studio del padre • Titolo di studio della madre • Benessere familiare (FAS)
Capitale sociale	<ul style="list-style-type: none"> • Partecipazione ad associazioni locali • Presenza di spazi per il tempo libero • La zona dove vivi è un bel posto?

Tabella 1.1: Determinanti sociali esaminati per l'attività fisica e la sedentarietà.

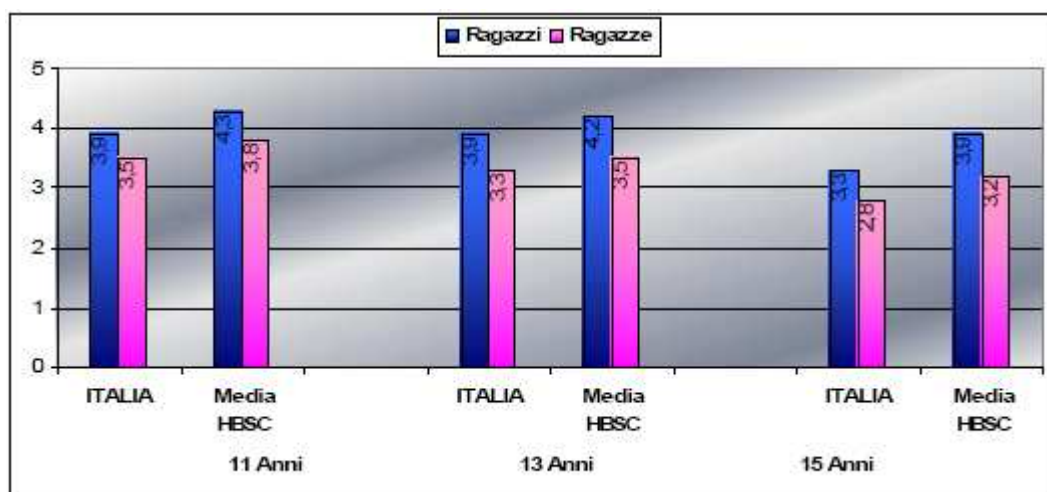


Figura 1.1: Numero medio di giorni alla settimana nei quali si praticano almeno 60 minuti di attività fisica: confronto tra i dati italiani e la media europea.

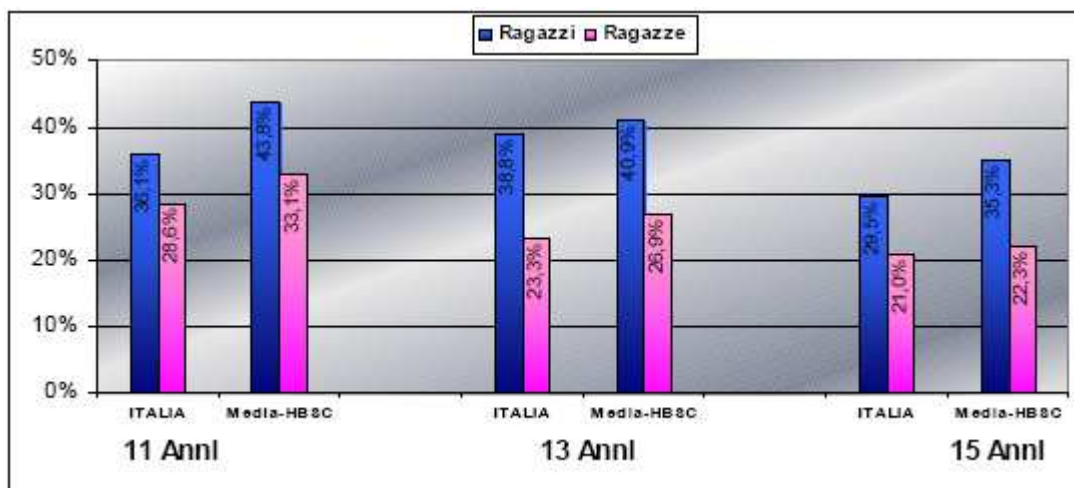


Figura 1.2: Percentuale di adolescenti in Italia che affermano di praticare almeno 60 minuti di sport per 5 o più giorni alla settimana, a confronto con al media europea.

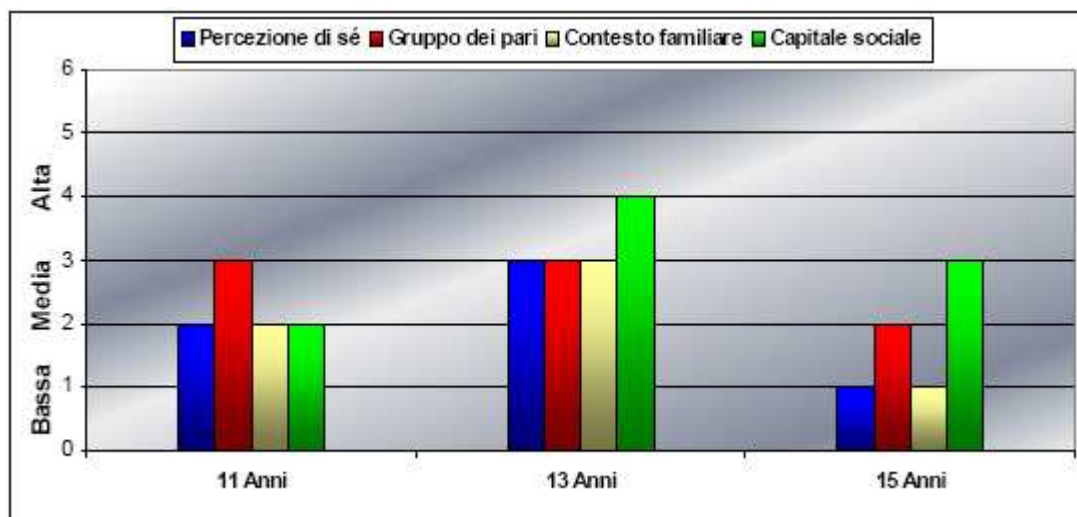


Figura 1.3: Rappresentazione delle correlazioni tra i 4 determinanti sociali e la pratica di un adeguato livello di attività fisica degli adolescenti.

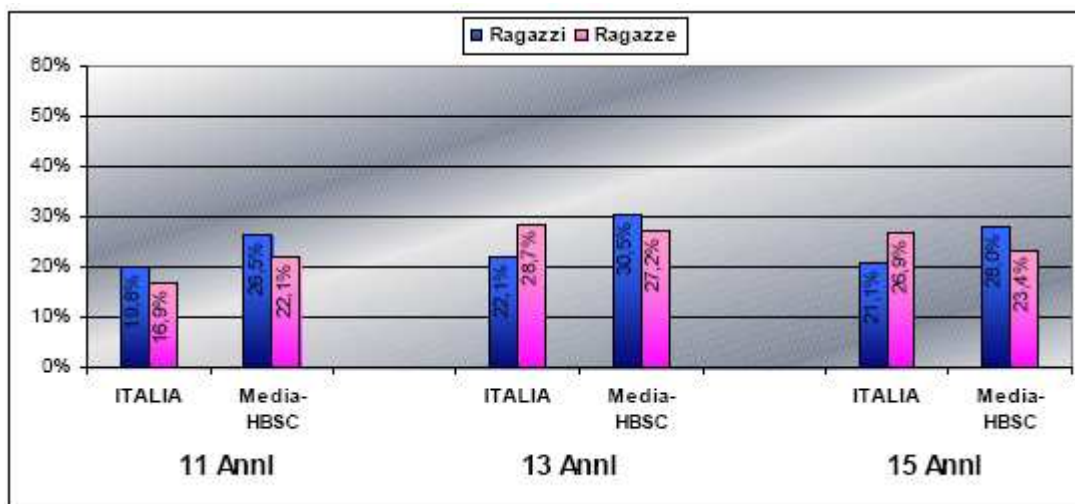


Figura 1.4: Percentuale di adolescenti in Italia che affermano di guardare la televisione almeno 4 ore al giorno (tranne il fine settimana), a confronto con la media europea.

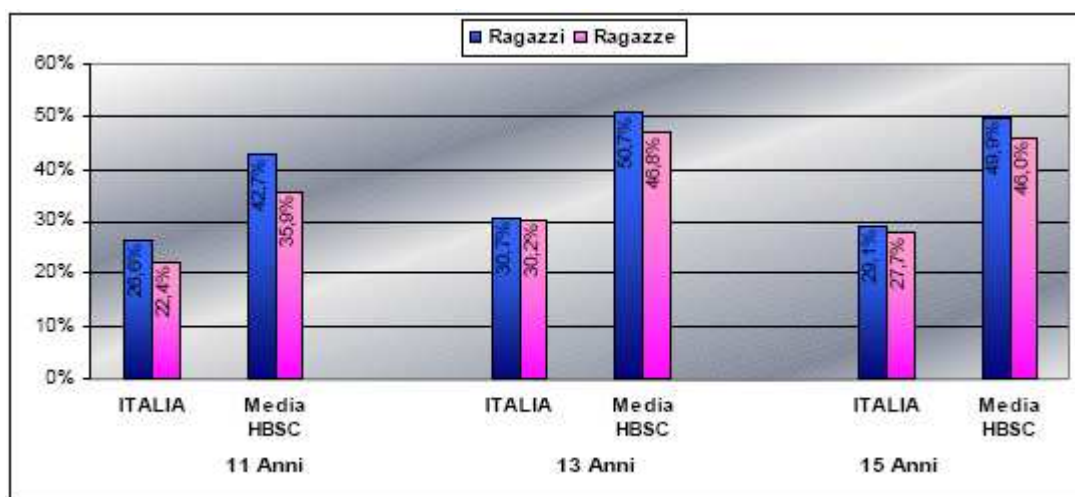


Figura 1.5: Percentuale di adolescenti in Italia che affermano di guardare la televisione almeno 4 ore al giorno nel fine settimana, a confronto con la media europea.

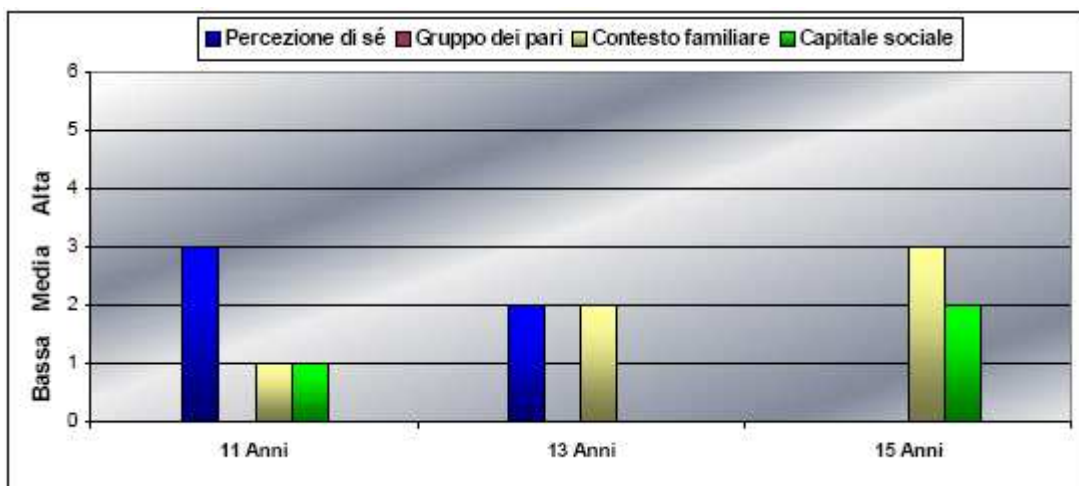


Figura 1.6: Rappresentazione delle correlazioni tra i 4 determinanti sociali e il rischio di sedentarietà per gli adolescenti in Italia.

Un'altra ricerca proposta dall'ISTAT, con i dati raccolti nei bienni 1999 - 2000 e 2001 - 2002, studia la percentuale di bambini e adolescenti obesi o in soprappeso in tutte le regioni italiane. Da tale ricerca risulta che la percentuale in Veneto è del 21,4% e in Friuli Venezia Giulia del 20,1% , con una media nazionale del 24,2% (vedi Tabella 1.2).

Percentuali di adolescenti obesi o in soprappeso			
Piemonte	17.1 %	Marche	25.8 %
Valle d'Aosta	14.3 %	Lazio	24.7 %
Lombardia	18.5 %	Abruzzo	27.0 %
Trentino Alto Adige	16.1 %	Molise	27.5 %
Veneto	21.4 %	Campania	36.0 %
Friuli Venezia Giulia	20.1 %	Puglia	26.0 %
Liguria	17.0 %	Basilicata	24.5 %
Emilia Romagna	22.7 %	Calabria	27.2 %
Toscana	17.0 %	Sicilia	26.8 %
Umbria	24.1 %	Sardegna	16.6 %
Media Italia		24.2 %	

Tabella 1.2: Percentuali degli adolescenti in soprappeso ed obesi in Italia divise per regioni.

In Figura 1.7 vengono mostrate le percentuali di giovani di 18 anni che praticano attività sportiva, suddivise per regioni italiane, si nota come le regioni con più le più basse percentuali sono quelle che riportano nella Tabella 1.2 le percentuali di adolescenti obesi o in soprappeso più alte della media italiana.

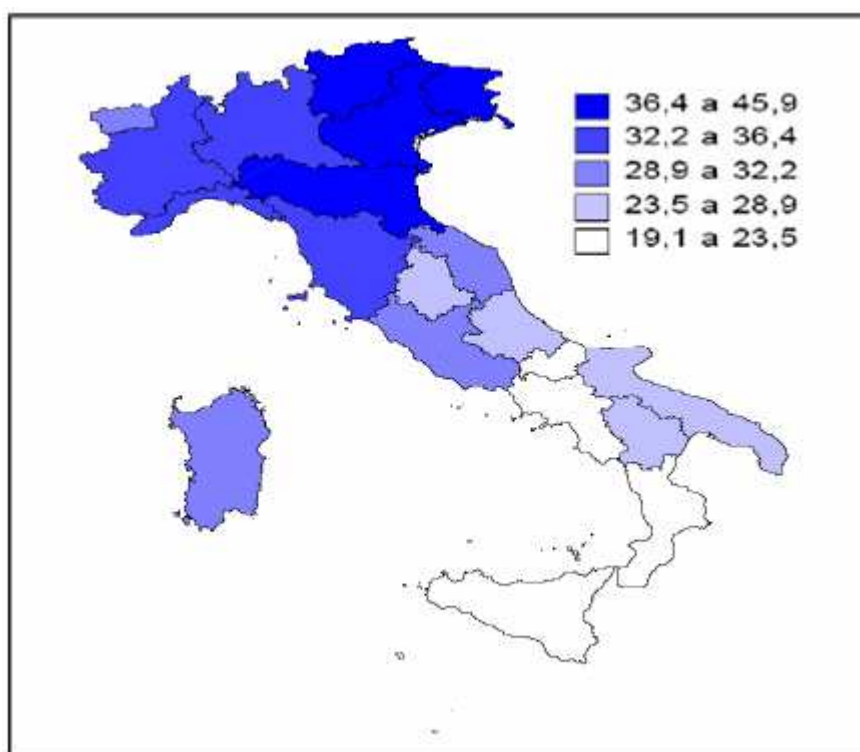


Figura 1.7: Percentuali di ragazzi all'età di 18 anni che praticano regolarmente attività sportiva in Italia, suddivise per regioni.

Sempre da uno studio svolto dall'ISTAT nel 2003, risulta che in Italia il soprappeso e l'obesità sono più diffusi in piccole città fino a 2000 abitanti (soprappeso: 35,6%, obesi: 9,1%) e in città da 2000 a 10000 abitanti (soprappeso: 35,1%, obesi: 10,5%), mentre in città più grandi il fenomeno è minore (soprappeso: 31,9%, obesi: 7,5%) .

Da uno studio della IOFT (International Association for the Study of Obesity) risulta che il 36% dei bambini di 9 anni di età in Italia risultano in soprappeso o obesi. Le percentuali invece ottenute da una ricerca svolta nel 2004 – 2005 nella regione Veneto, su un campione di 12853 bambini nati negli anni 1993 e 1995, sono riassunte in Tabella 1.3 .

		Maschi	Femmine
Anno 1993	Soprappeso	23.14 %	20.79 %
	Obesi	5.79 %	4.70 %
Anno 1995	Soprappeso	19.07 %	21.80 %
	Obesi	6.05 %	5.59 %

Tabella 1.3: Percentuali degli adolescenti in soprappeso ed obesi in Veneto.

1.3 QUESTIONARIO

Partendo dalle problematiche legate all'obesità e alla sua diffusione, sia in ambito europeo sia in ambito territoriale, si è deciso di studiare il fenomeno iniziando proprio dai più giovani, perché più esposti a fattori di rischio e perché più semplicemente, rappresentando il futuro, è giusto siano tutelati e non di meno si tenga conto dell'impatto futuro sulla spesa pubblica per curare le patologie derivate dall'obesità. In particolare il questionario, che viene utilizzato in questa tesi, è parte di una ricerca che prende il nome di *Children's Lifestyle*, un progetto che comprende altri sette paesi della comunità europea ed è coordinato dall'Università di Colonia, in Germania. Per l'Italia il progetto è seguito dal Prof. A. Carraro, del Dipartimento di Scienze dell'Educazione dell'Università di Padova.

Il questionario è stato somministrato ad un campione omogeneo di 1005 studenti di quarta elementare provenienti da Veneto e Friuli Venezia Giulia. I comuni interessati da tale progetto sono piccoli, con circa 3700 abitanti (con un solo caso di comune sopra i 20000 abitanti) e sparsi nelle province di Padova, Vicenza, Rovigo, Venezia, Pordenone e Udine. Per ogni questionario si hanno complessivamente 133 *items* a cui rispondere, che rappresentano informazioni sia psicologiche che fisiche dell'adolescente a cui è stato somministrato. Il questionario presenta una struttura semplice nella compilazione, con risposte suddivise su scala *Likert*, ordinata su una scala di lunghezza variabile in base alla necessità di giudizio. Solo le risposte sull'età, il peso, l'altezza e le ore percorse al giorno per il tragitto da casa a scuola, sono di tipo numerico.

Il questionario viene presentato allo studente e visionato dagli intervistatori presenti in classe, che sono in numero variabile da 3 a 4. Nella raccolta dei dati viene seguita una metodologia rigorosa,

proponendo a coloro che fungono da intervistatori delle regole da seguire. In tal senso gli intervistatori, oltre a presentare una spiegazione standardizzata del questionario, controllano anche la presenza di eventuali dati mancanti e procedono alla misura del peso e dell'altezza per ogni studente.

Il questionario completo è riportato nell'Appendice della tesi.

1.4 VARIABILI

Il *dataset* contiene le risposte di un campione di $n=1005$ unità statistiche suddiviso in parti uguali tra i due sessi. I dati raccolti si riferiscono ad adolescenti di età compresa tra gli 8,5 e gli 11 anni, ossia studenti delle scuole elementari di classe quarta. Le variabili complessivamente sono 133 e spaziano dal rappresentare dati oggettivi come peso, altezza ed età, a dati di tipo soggettivo. In particolare le variabili che compongono il dataset "*Project Lifestyle Analysis*" sono le seguenti:

Alcune semplici domande:

- S1 Sesso.
- S2_a Età espressa in mesi.
- S2_b Età espressa in anni.
- S3_a Statura in metri (dichiarata dall'alunno).
- S3_b Peso in kg (dichiarata dall'alunno).

Aspetto dei genitori:

- S4 Aspetto fisico del padre.
- S5 Aspetto fisico della madre.

Qualche domanda sullo sport:

- S6 Praticare sport per me è ...
- S7 Sei iscritto a una società sportiva?
- S8_a Lunedì pratico questa disciplina.
- S8_b Lunedì per quante ore?
- S9_a Martedì pratico questa disciplina.
- S9_b Martedì per quante ore?
- S10_a Mercoledì pratico questa disciplina.
- S10_b Mercoledì per quante ore?
- S11_a Giovedì pratico questa disciplina.
- S11_b Giovedì per quante ore?
- S12_a Venerdì pratico questa disciplina.
- S12_b Venerdì per quante ore?
- S13_a Sabato pratico questa disciplina.

- S13_b Sabato per quante ore?
- S14_a Domenica pratico questa disciplina.
- S14-b Domenica per quante ore?
- S15 Mio padre pratica sport.
- S16 Mia madre pratica sport.
- S17 Io mi scatenò e gioco all'aperto.

Alcune semplici domande sui genitori e sulla casa:

- S18 A casa ho una stanza solo per me.
- S19 Quante volte in vacanza negli ultimi 12 mesi?
- S20 La famiglia possiede un'auto?
- S21 Quanti computer possiede la famiglia ?

Provenienza dei genitori:

- S22 Mia madre proviene da..
- S 23 Mio padre proviene da ..

Giudizi ultima pagella:

- S24 Matematica: quale giudizio avevi nell'ultima pagella.
- S25 Italiano: quale giudizio avevi nell'ultima pagella.
- S26 Educazione motoria: quale giudizio avevi nell'ultima pagella.
- S27 Titolo di studio di uno dei genitori.

Il tragitto da casa a scuola:

- S28_a Con quale mezzo vai a scuola?
- S28_b Quante volte alla settimana?
- S28_c Quanti minuti per volta?
- S29_a In bici?
- S29_b Quante volte alla settimana?
- S29_c Quanti minuti per volta?
- S30_a Con il bus?
- S30_b Quante volte alla settimana?
- S30_c Quanti minuti per volta?
- S31_a A piedi e con il bus?
- S31_b Quante volte alla settimana?
- S31_c Quanti minuti per volta?
- S32_a Con l'auto?
- S32_b Quante volte alla settimana?
- S32_c Quanti minuti per volta?

Giudica te stesso:

- S33 - S65 Self Concept scale: valutare se stessi.

Qualcosa sull'alimentazione:

- S66_a - S72_a** Quale dei seguenti pasti assunti regolarmente nella settimana?
S66_b - S72_b Con chi assunti i pasti?
S73 - S93 Negli ultimi sei mesi con quale frequenza hai mangiato ...

Qualche domanda su televisione, computer e videogames:

- S94** Quanto guardi la TV da solo?
S95 Quanto guardi la TV con i tuoi genitori?
S96 Quanto guardi la TV con i tuoi amici?
S97 Quanto usi il computer da solo?
S98 Quanto usi il computer con i tuoi genitori?
S99 Quanto usi il computer con i tuoi amici?
S100 Quanto usi i videogames da solo?
S101 Quanto usi i videogames con i tuoi genitori?
S102 Quanto usi i videogames con i tuoi amici?
S103 Per quante ore al giorno guardi nel tuo tempo libero la TV (da Lunedì a Venerdì)?
S104 Per quante ore al giorno guardi nel tuo tempo libero la TV (fine settimana) ?
S105 Per quante ore al giorno usi il computer (da Lunedì a Venerdì)?
S106 Per quante ore al giorno usi il computer (fine settimana)?
S107 Per quante ore al giorno giochi con una consolle (da Lunedì a Venerdì)?
S108 Per quante ore al giorno giochi con una consolle (fine settimana)?
S109 Per me sarebbe veramente difficile rinunciare a ...

Misure:

- Peso** In chilogrammi (misurato dagli intervistatori).
Altezza In metri (misurato dagli intervistatori).

Oltre alle variabili del questionario, nei prossimi capitoli verranno analizzati anche alcuni indici, ricavati dai dati del questionario, quali:

- BMI** (*Body Mass Index*) o indice di massa corporea, utilizzato per misurare il grado di obesità;
MC (*Media Consumption*) un indice che misura il grado di utilizzo di apparati di intrattenimento;
FAS (*Family Affluence Scale*) un indice che misura lo status economico familiare.

1.5 DATI MANCANTI

Il dataset presenta fortunatamente pochi dati mancanti nonostante l'elevata numerosità. Anche per quanto concerne la presenza di valori anomali, questi sono pochi e sono stati facilmente individuati vista la natura delle variabili. Ad esempio, in una scala di tipo *Likert* da 1 a 4, la presenza di un 6 nel dataset è un errore.

Nonostante l'esiguo numero di errori e *missing data*, si deve procedere utilizzando qualche metodo per effettuare l'imputazione. Il metodo più semplice nei casi in cui i dati mancanti nelle variabili sono inferiori al 5%, consiste nel sostituirli con le medie, mediane o mode a seconda della natura delle variabili. Nel caso specifico, i dati mancanti o errati, relativamente alle risposte su scala *Likert*, sono stati sostituiti con le mediane dei valori osservati per la variabile presa in esame. Tale metodo risulta semplice e facilmente applicabile in molti casi.

Nel prossimo capitolo saranno effettuate le analisi esplorative e grafiche per le variabili del questionario rivedute e corrette, con il calcolo di alcune correlazioni tra variabili.

ANALISI ESPLORATIVE

In questo capitolo si presenta un'analisi esplorativa preliminare dei dati rilevati con il questionario. Tale analisi serve per focalizzare le caratteristiche dei dati in nostro possesso e, grazie ai grafici e all'uso di semplici test, per capire quali sono le variabili più interessanti e le eventuali relazioni tra esse. Il *dataset* studiato ha moltissime variabili e risulta indispensabile quindi riuscire a sintetizzarle. In questo capitolo si fa uso di metodi semplici ed intuitivi per riassumerle, utili successivamente a formulare delle ipotesi di lavoro che saranno considerate nei capitoli successivi. I prossimi paragrafi sono divisi per argomenti secondo lo schema proposto e seguito nel questionario.

2.1 ALCUNI DATI SEMPLICI SUGLI STUDENTI

Dei 1005 studenti, a cui è stato somministrato il questionario, il 49,9% sono di sesso maschile e il 50,4% di sesso femminile. L'età è compresa tra gli 8,5 e gli 11 anni, secondo la distribuzione riportata nella Figura 2.1. L'età media risulta pari a 9,65 anni, con deviazione standard pari a 0,35. Il questionario prende in esame studenti di classe quarta elementare, dai 9 ai 10 anni, ma dai dati analizzati risultano presenti anche adolescenti di età superiore, si può supporre che alcuni di essi, precisamente 8 di 10,5 anni e 4 di 11 anni, si trovano a dover frequentare una classe differente da quella prevista.

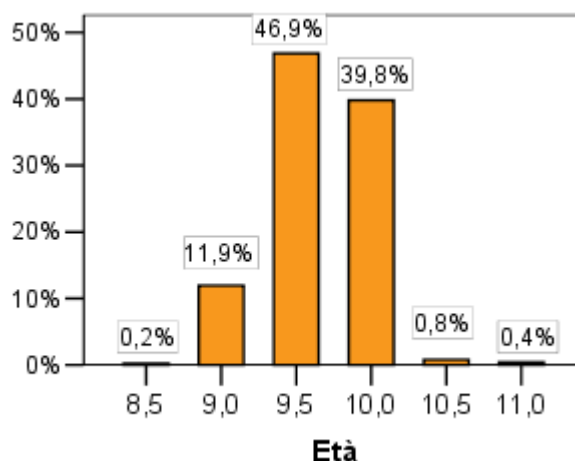


Figura 2.1: Istogramma della variabile età.

Le misurazioni del peso e dell'altezza degli studenti sono rappresentati in Figura 2.2 e in Figura 2.3 . Tali grafici sono la sintesi di due misurazioni: una di tipo soggettivo in cui gli studenti dovevano inserire nel questionario la propria altezza e peso, che di seguito è indicata con altezza e peso stimato, e una di tipo oggettivo, misurata effettivamente per ogni studente.

In prima analisi vengono confrontate tali misure allo scopo di vedere se gli studenti a cui è stato somministrato il questionario hanno una percezione reale del proprio fisico.

Per l'altezza si osserva che c'è una moderata correlazione tra altezza misurata e stimata (correlazione pari a 0,52). Un ulteriore test per verificare l'adattamento delle funzioni di ripartizione per le due misure è il test di Kolmogorov – Smirnov (Ercolani, Areni e Leone, 2002 Cap. 3.2.2), il cui p-value risulta 0,009, si rifiuta quindi l'ipotesi nulla di uguaglianza delle funzioni di ripartizione.

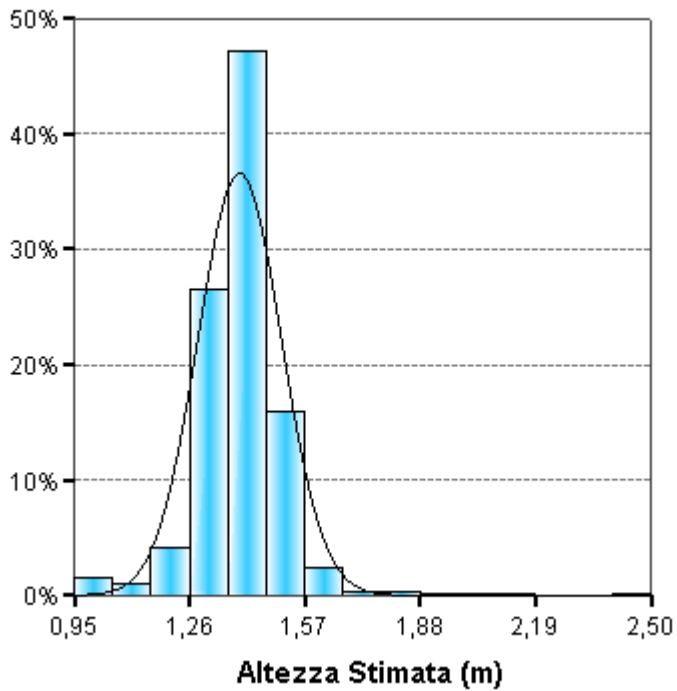
Interessante è il confronto con i grafici delle curve di crescita o percentili (World Health Organization, www.who.int/childgrowth) costruiti prendendo in esame migliaia di bambini in molte età diverse e usato in pediatria come valido strumento per valutare la crescita in peso ed altezza. Tale grafico si differenzia per età e per sesso: la crescita, infatti, è divisa fra maschi e femmine e per un'età che va da 0 a 2 anni e da 2 a 18 anni (vedi grafici riportati in Appendice).

I diagrammi percentili sono realizzati controllando il peso e l'altezza di migliaia di bambini di molte età diverse e poi dividendo le misure ottenute in modo che una proporzione definita dei bambini campione viene a trovarsi sopra e sotto le misure particolari in età specifiche.

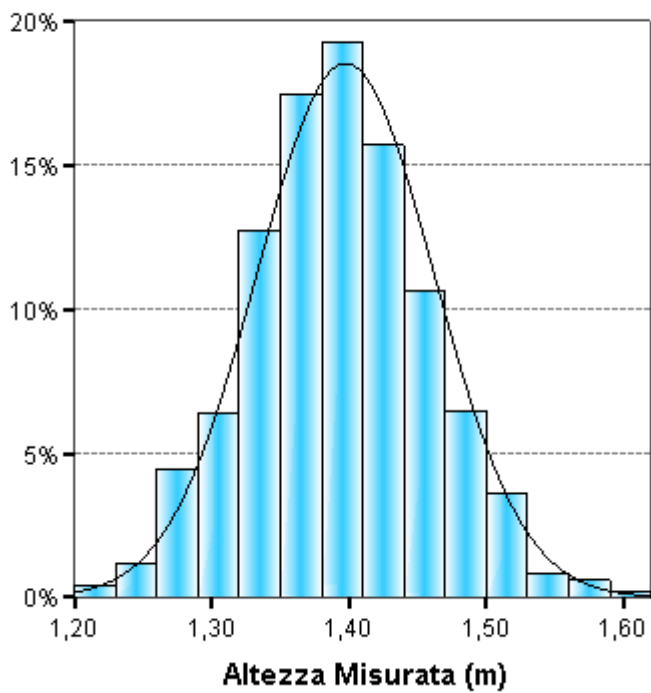
Il cinquantesimo centile è collocato in modo che il 50% dei bambini siano al di sopra e 50% al di sotto di esso in una data età.

Nel caso in questione utilizzando la curva di crescita dell'altezza si trova che per un bambino di 9,5 anni il cinquantesimo centile misura 1,35 metri per i maschi e anche per le femmine. La mediana dell'altezza misurata nel campione vale 1,39 metri per i maschi e 1,40 metri per le femmine ed è quindi di poco superiore alla curva di crescita attesa.

Per quanto riguarda le misure del peso stimato ed effettivo, la correlazione risulta più forte che per l'altezza (coefficiente pari a 0,86), mentre il test di Kolmogorov – Smirnov confuta l'ipotesi nulla con un p-value pari a 0. Si nota infatti come seppur le misure stimate non si discostano molto dall'effettivo peso, esse tendono ad essere leggermente sottostimate e meno precise. Dalla curva di crescita del peso risulta che per un'età di 9,5 anni il cinquantesimo centile misura 29 chilogrammi sia per i maschi che per le femmine. La mediana per il peso misurato del campione vale, rispettivamente, 36,2 e 35,7 per i maschi e per le femmine ed è anche in questo caso superiore al valore della curva di crescita.

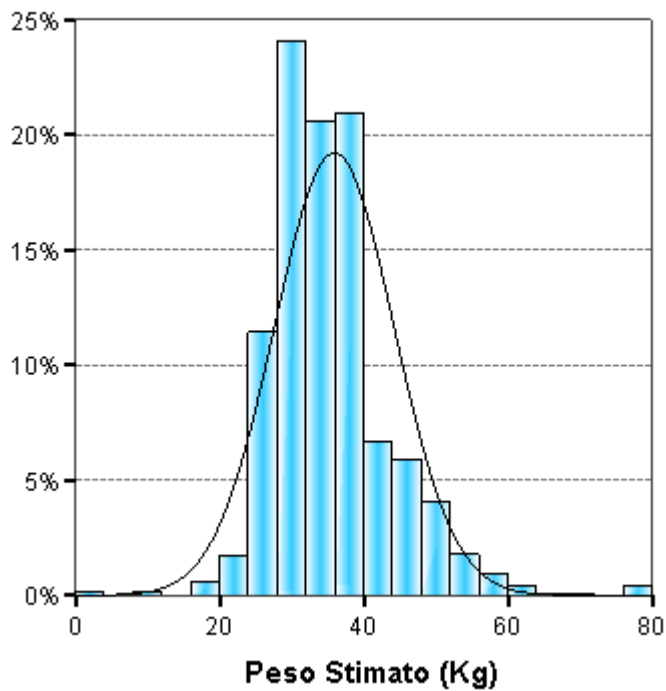


Statistiche descrittive Altezza Stimata	
Minimo	0.95
1° Quartile	1.35
Media (sd)	1.39 (0.004)
Mediana	1.40
3° Quartile	1.45
Massimo	2.50

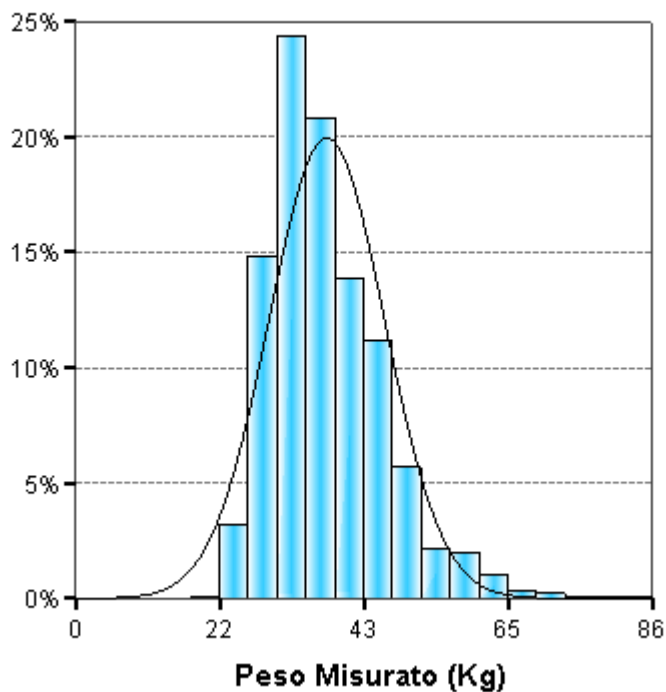


Statistiche descrittive Altezza Misurata	
Minimo	1.21
1° Quartile	1.35
Media (sd)	1.40 (0.002)
Mediana	1.40
3° Quartile	1.44
Massimo	1.62

Figura 2.2: Istogrammi frequenze altezza stimata e misurata e tabelle di sintesi statistiche descrittive.



Statistiche descrittive Peso Stimato	
Minimo	2
1° Quartile	30
Media (sd)	35.9 (0.26)
Mediana	35
3° Quartile	40
Massimo	80



Statistiche descrittive Peso Misurato	
Minimo	20.3
1° Quartile	31.4
Media (sd)	37.5 (0.27)
Mediana	35.9
3° Quartile	41.9
Massimo	86.2

Figura 2.3: Istogrammi frequenze peso stimato e misurato e tabelle di sintesi statistiche descrittive.

2.2 ASPETTO GENITORI

L'aspetto dei genitori viene valutato attraverso una scala *Likert* con un valore compreso da 1 a 7 in modo crescente, da molto magro a obeso. Per facilitare la scelta e la comprensione per gli studenti, nel questionario sono state inserite delle immagini relative all'aspetto. Di seguito vengono riportati i grafici con le percentuali delle risposte date e il boxplot riassuntivo (vedi Figura 2.4).

Sia per l'aspetto dei padri che delle madri i quartili sono rispettivamente 3 (sottopeso) e 4 (normopeso). La mediana, invece, per i padri risulta 4 (normopeso), mentre per le madri è 3 (sottopeso) con uno scarto interquartile per entrambi di 1.

Una considerazione va fatta sulla scala: come si nota dai grafici, le percentuali maggiori di risposta, sia per i padri che per le madri, sono concentrate in "sottopeso" e "normale". Infatti, se si osservano le immagini associate alle risposte 3 e 4 del questionario (vedi Appendice), sono molto simili e la distinzione non è netta. Le sette categorie potrebbero essere compattate in tre classi: sottopeso, normale e soprappeso.

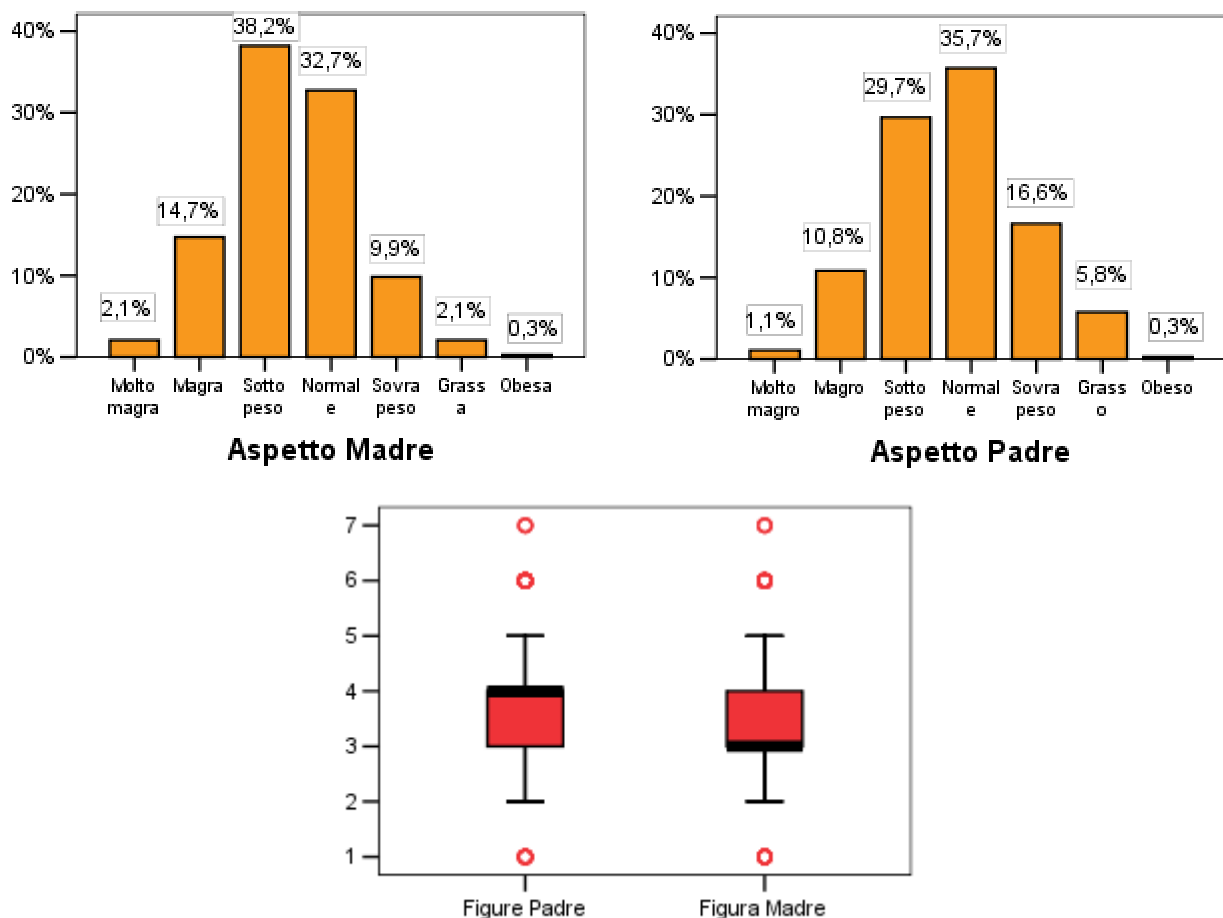


Figura 2.4: Istogrammi sull'aspetto della madre e del padre e boxplot.

2.3 QUALCHE DOMANDA SULLO SPORT

L'aspetto motorio legato ad attività fisiche é molto importante e nel questionario viene lasciato ampio spazio per rilevare le variabili legate allo sport. La prima domanda riguarda l'importanza data dagli studenti alla pratica sportiva ed emerge una netta risposta positiva, superiore al 50%. Tale variabile è misurata su scala *Likert* con valore crescente da 1 (nessuna importanza) a 10 (molto importante). La mediana risulta 10 (molto importante) con i quartili, rispettivamente, pari a 8 e 10 e scarto interquartile pari a 2. Un'ulteriore analisi fatta per distinzione per genere (vedi Figura 2.5) mostra come sia i maschi che le femmine reputino la pratica sportiva molto importante. Utile in questo caso risulta l'utilizzo del test di Wilcoxon (Ercolani, Areni e Leone, 2002 Cap. 3.3.2), un test non parametrico per verificare ipotesi su campioni correlati in cui si voglia testare la differenza tra due condizioni e basata sul confronto delle mediane (l'ipotesi nulla H_0 afferma l'assenza di differenze), il cui p-value risulta 0,2 avvalorando l'ipotesi di assenza di differenze per genere.

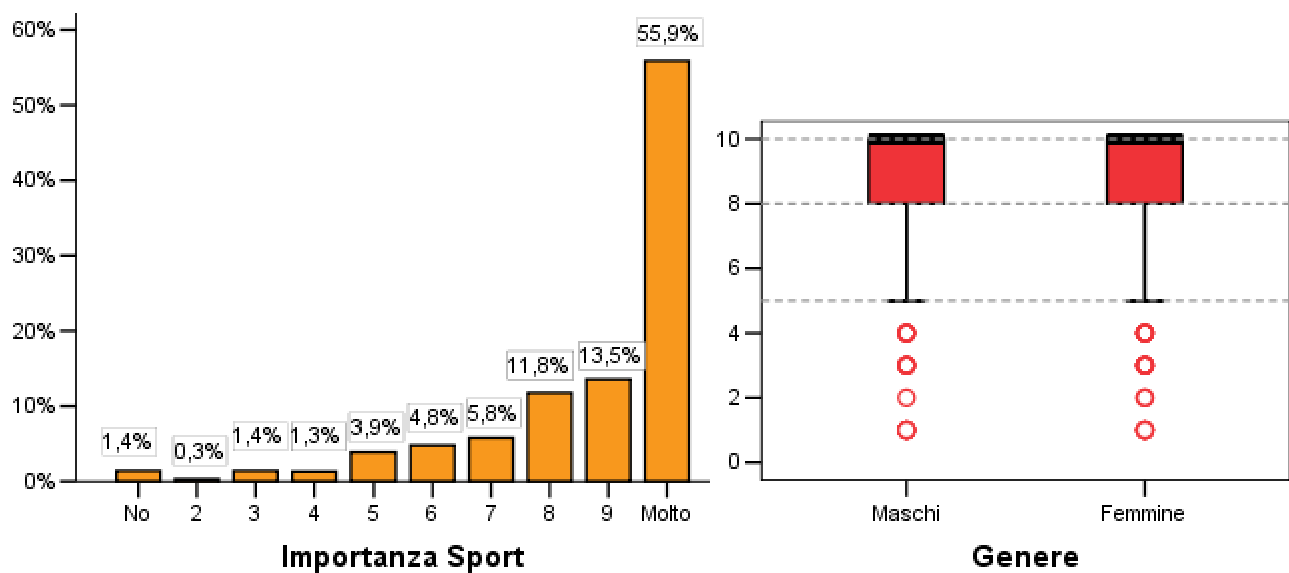


Figura 2.5: Istogramma con le percentuali di risposta sull'importanza dello sport e boxplot divisi per genere.

Successivamente sono stati indagati gli sport praticati durante i giorni della settimana e la frequenza, in ore, dedicata a tale attività. Il 78,3% degli intervistati risulta iscritto ad una società sportiva, mentre l'11,1% non lo è più e il 10,6% non lo è mai stato. Anche in questo caso nei due generi non c'è una sostanziale differenza, le risposte sono divise pressoché in eguale misura tra maschi e femmine (vedi Figura 2.6) e il p-value del test di Wilcoxon risulta pari a 0,1.

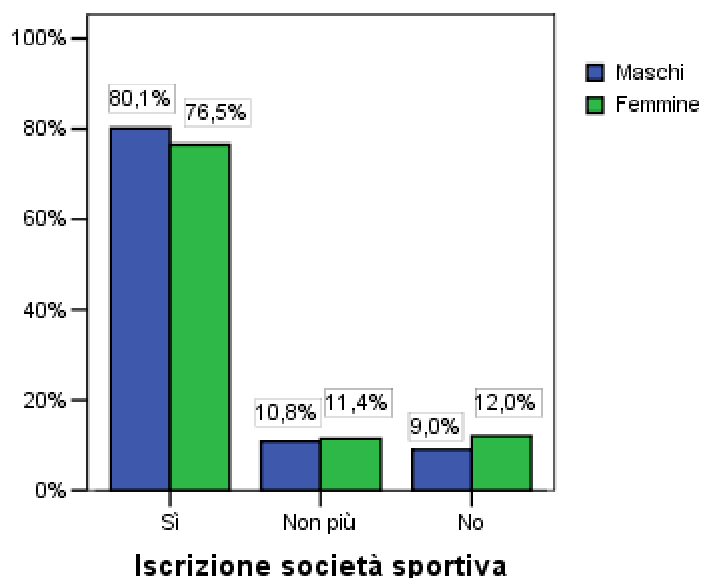


Figura 2.6: Istogramma con le percentuali di risposta all'iscrizione ad una società sportiva.

Tipo di sport praticati	Percentuale
Calcio	25,5%
Ginnastica - Danza	18,5%
Nuoto	15,1%
Pallacanestro	11,4%
Pallavolo	10,7%
Arti Marziali	3,5%
Corsa - Ciclismo	2,3%
Tennis	2,2%
Equitazione	1,5%
Altri Sport *	9,3%
Totale	100%

* Altri sport comprendono, in percentuale inferiore all'1%, sport quali: Tennis da tavolo, Pesca, Vela, Sport invernali, Scherma, Pattinaggio.

Tabella 2.1: Tabella di sintesi con le percentuali degli sport praticati.

Coloro che hanno risposto positivamente all'iscrizione ad una società sportiva, continuano il questionario elencando per ogni giorno della settimana (da lunedì a domenica), i tipi di sport praticati, scegliendoli da una lista codificata e riportando inoltre i minuti dedicati a tale sport. Nella Tabella 2.1 sono riassunti ed elencati gli sport più diffusi con la relativa percentuale dei rispondenti.

I minuti totali di pratica sportiva sono riassunti nella Figura 2.7. La mediana risulta di 180 minuti, e i valori si concentrano tra il primo quartile di 120 minuti e il terzo quartile pari a 286 minuti. Si possono notare nel boxplot numerosi *outliers* che si discostano dalla mediana in alcuni casi anche considerevolmente, dovuti forse ad un errore nell'imputazione o legati ad una pratica sportiva agonistica.

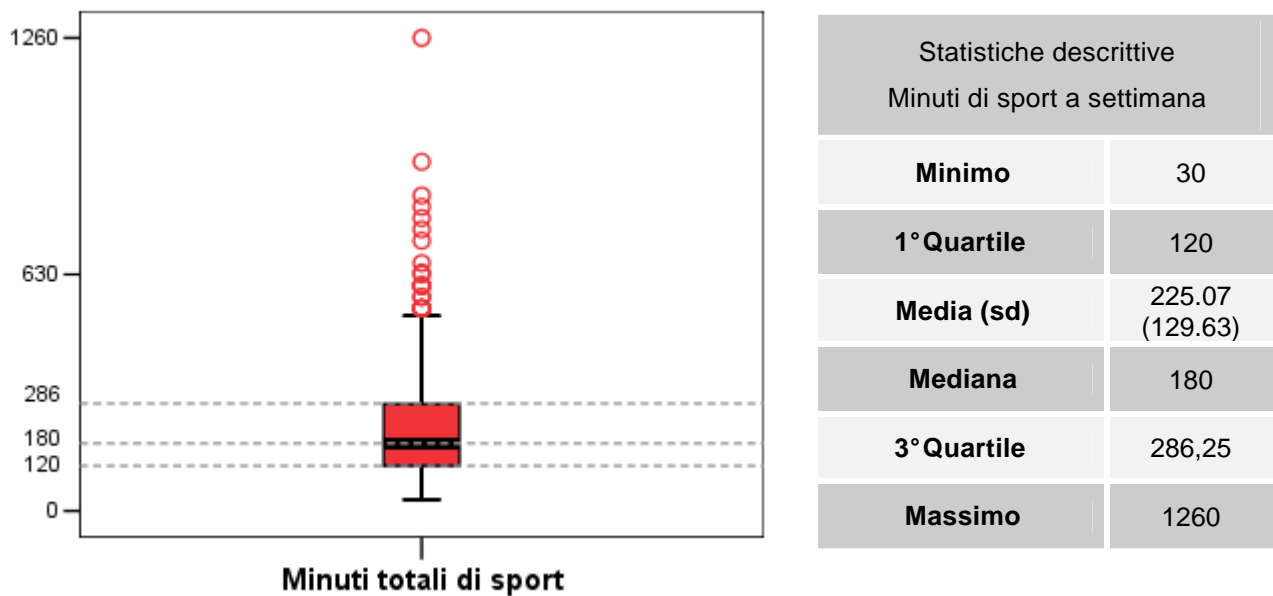


Figura 2.7: Boxplot dei minuti settimanali di pratica sportiva e tabella di sintesi.

Il tempo di allenamento per genere risulta maggiore per i maschi, con uno scarto interquartile pari a 180 minuti, contro i 120 minuti delle femmine (p-value test di Wilcoxon pari a 0). Dalla Figura 2.8 si nota una maggiore variabilità per i maschi, mentre per entrambi i sessi sono presenti numerosi valori che si discostano dalla media.

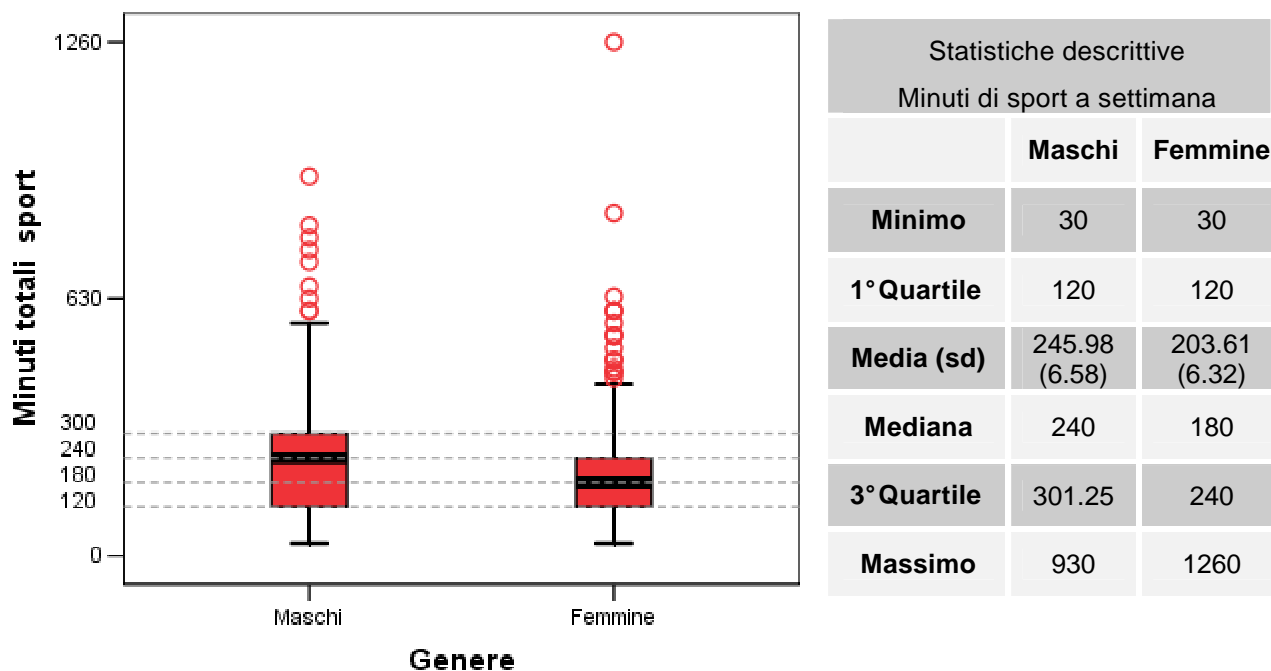


Figura 2.8: Boxplot dei minuti di pratica sportiva settimanale divisi per genere e tabella di sintesi per maschi e femmine.

Nella Figura 2.9 viene riportata la suddivisione dei minuti di attività per i giorni della settimana. Gli allenamenti si concentrano tra lunedì e venerdì, con una media costante per chi si allena che varia tra 26 e 36 minuti al giorno. La pratica diminuisce nei giorni del fine settimana con una media di 17 minuti di sabato e 8 minuti di domenica. Presumibilmente, soltanto chi pratica uno sport che prevede tornei continua l'attività fisica in questi giorni.

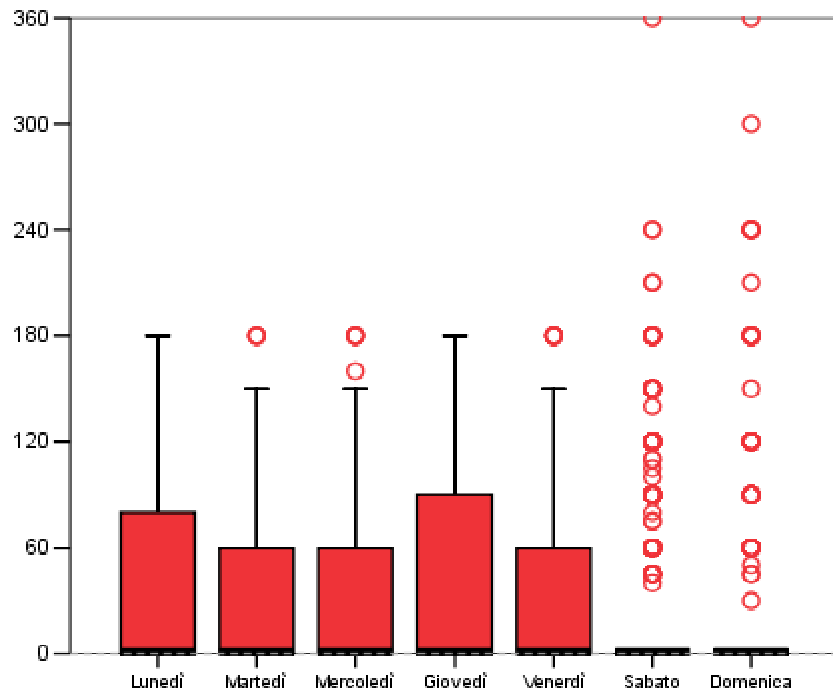


Figura 2.9: Boxplot dei minuti di pratica sportiva divisi per giorno.

Di seguito sono riportate le risposte sull'attività fisica dei genitori. Anche questa domanda é misurata su scala *Likert* che varia da 1 (tutti i giorni) a 5 (mai) e dalla Figura 2.10 si nota come più del 50% dei genitori non pratica mai sport. La mediana per entrambi i genitori risulta 5 (mai), mentre lo scarto interquartile risulta 1 per i padri e 2 per le madri. Nella Figura 2.11 vengono affiancati i due boxplot relativi alle risposte dei genitori, da cui si nota che le madri fanno un'attività fisica maggiore rispetto ai padri (test di Wilcoxon, p-value= 0.0002).

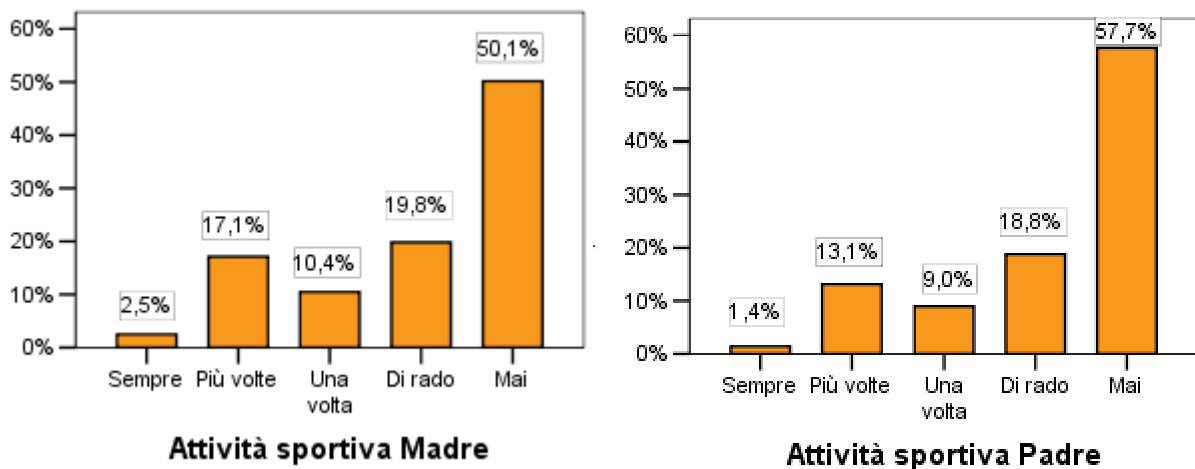


Figura 2.10: Istogrammi con le percentuali di pratica sportiva settimanale dei genitori.

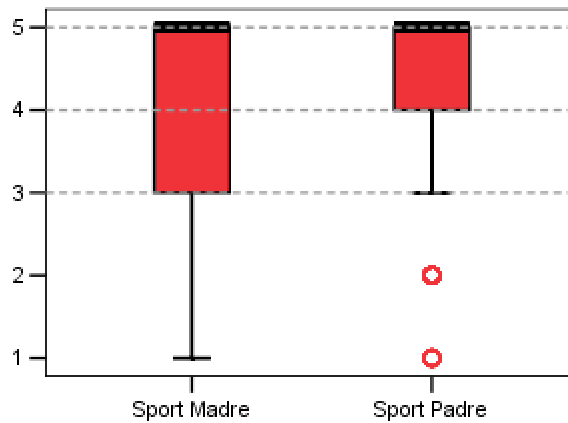


Figura 2.11: Boxplot pratica sportiva settimanale divisa tra padri e madri.

Oltre ad una attività fisica specifica legata all'iscrizione ad una società sportiva, si è indagato se nel tempo libero, nell'arco della settimana, i bambini giocano e si scatenano all'aperto. Anche questa domanda, come la precedente, è codificata nel questionario su scala *Likert* da 1 (tutti i giorni) a 5 (mai). Nella Figura 2.12 si osserva come per la maggior parte dei rispondenti il gioco è frequente quasi tutti i giorni. La mediana risulta 2 (più volte), con uno scarto interquartile pari a 1. Il test di Wilcoxon effettuato discriminando per genere riporta un p-value pari a 0,02, denotando una differenza tra le risposte date dai maschi e dalle femmine.

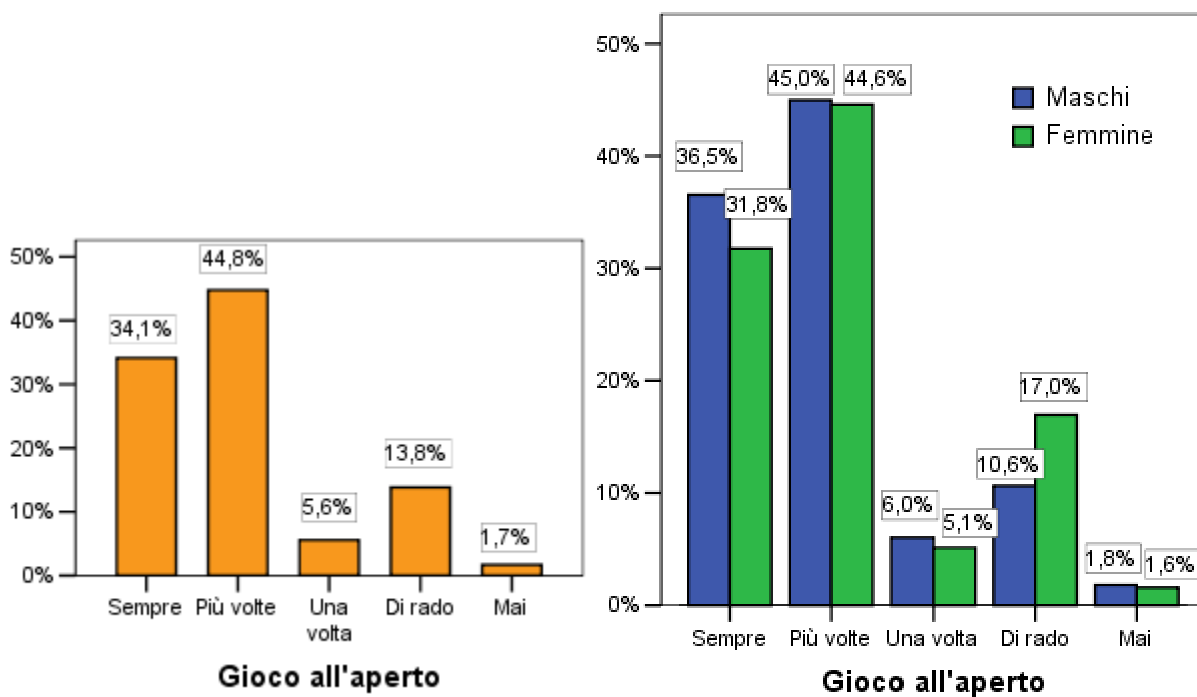


Figura 2.12: Istogrammi con le percentuali della frequenza settimanale di gioco all'aperto e suddivisione per genere.

2.4 ALCUNE COSE SEMPLICI SUI GENITORI E SULLA CASA

Questa parte del questionario viene utilizzata per acquisire informazioni legate in qualche modo alle possibilità economiche della famiglia. Viene richiesta la disponibilità di una stanza completamente personale, la frequenza delle vacanze, il numero di automobili e il numero dei personal computer a disposizione.

Queste informazioni vengono utilizzate per poi determinare un indice, denominato FAS (*Family Affluence Scale*, www.hbsc.org), uno strumento messo a punto dal gruppo di ricerca internazionale HBSC utile a dare una misura del benessere economico della famiglia attraverso la rilevazione della presenza di beni comuni.

La prima domanda richiede la disponibilità in casa di una stanza solo ad uso personale, a cui risponde positivamente il 51,7 % degli studenti, mostrando un certo bilanciamento nella risposta tra no e sì.

La seconda domanda riguarda il numero di vacanze durante l'arco di 12 mesi. Le possibili risposte sono codificate su scala *Likert*, che va da 1 (mai) a 4 (più di due volte). Come si nota dalla Figura 2.13 le risposte si distribuiscono per la maggior parte tra una volta e più di due volte all'anno.

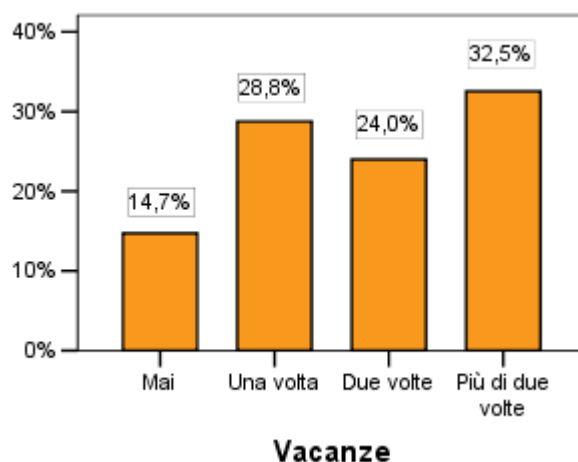


Figura 2.13: Istogramma con le percentuali di risposta sul numero di vacanze all'anno.

La terza e la quarta domanda fanno riferimento al numero di automobili e personal computer (pc) che sono disponibili in famiglia. Anche tali variabili sono codificate su scala *Likert*, che va da 1 (nessuna) a 3 (due o più) per il numero di automobili e da 1 (nessuno) a 4 (più di due) per il numero di computer. Il 61,7% delle famiglie ha due o più automobili, contro il 37,9% che ne ha una solamente. Invece, il 55,5% delle famiglie possiede un pc, contro l'11,3% che non ne ha e il 33,1% che ne ha due o più.

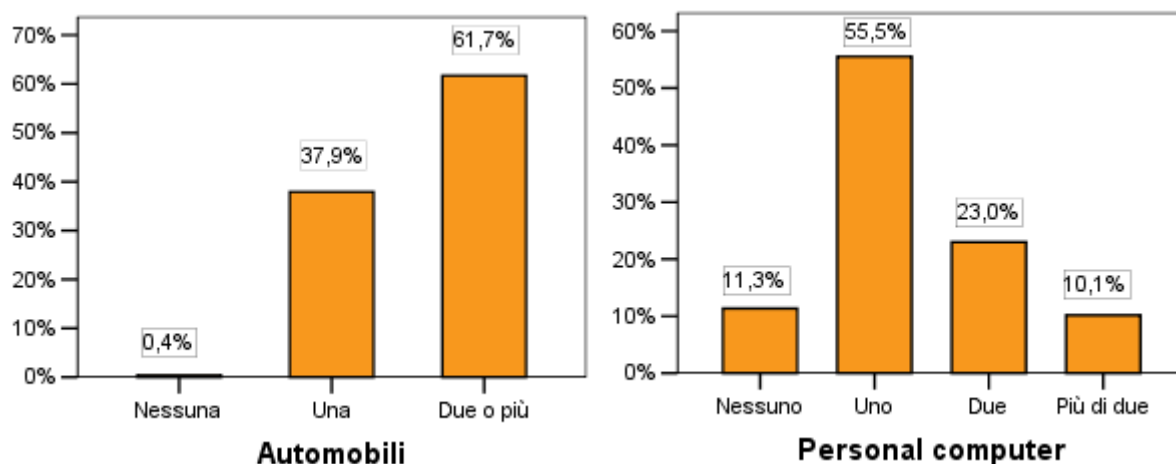


Figura 2.14: Istogrammi con le percentuali di risposta sul numero di automobili e pc in famiglia.

Le quattro domande precedenti servono a calcolare un indice (FAS - Family Affluence Scale) che viene qui di seguito riportato. Tale indice indica una scala con cui si misura lo status di benessere economico familiare.

Nella Figura 2.15 é riportato il barplot di sintesi dell'indice FAS: come si vede tale indice é suddiviso in una scala con 7 valori, con 1 valore minimo e 7 massimo. Come si nota nel boxplot di Figura 2.15, la mediana é pari a 5, con uno scarto interquartile pari a 2. I valori si concentrano, infatti, tra 4 e 6, che rappresentano un valore medio alto di status familiare. Compattando le 7 categorie in sole tre classi risulta che il 15,7% delle famiglie hanno uno status familiare basso, il 53,5% uno status medio e il 30,7% uno status alto.

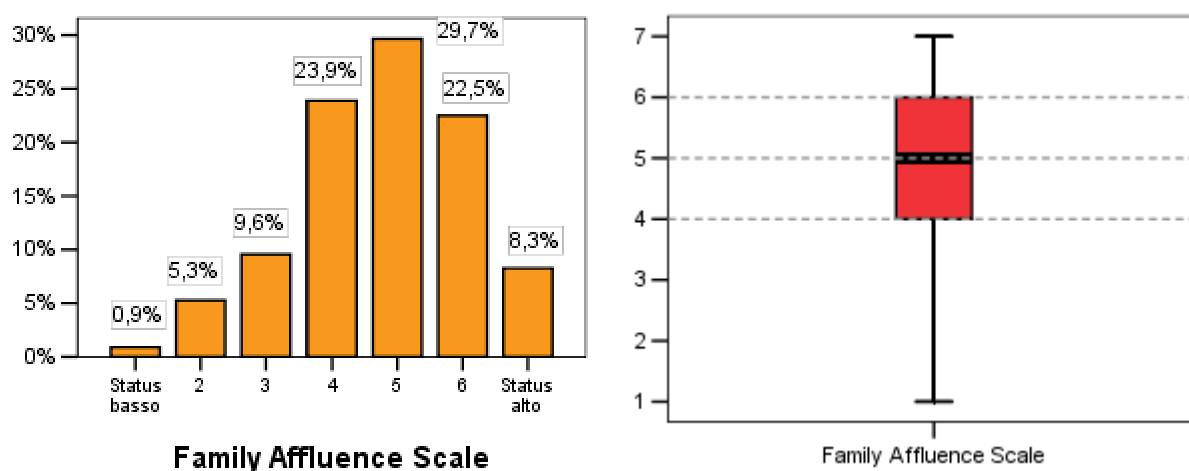


Figura 2.15: Istogramma con le percentuali indice FAS e boxplot relativo.

2.5 STATO DI PROVENIENZA DEI GENITORI

La provenienza dei genitori è stata rilevata dal questionario mediante l'inserimento di un codice che specificava il paese di origine nel caso non fosse italiano. Vista la numerosità dei dati si è provveduto a raggruppare l'insieme dei genitori come "entrambi italiani", "uno dei genitori è italiano" e "entrambi i genitori di altri paesi".

La provenienza dei genitori, separatamente per padre e madre, mostra come la maggior parte dei rispondenti ha sia la madre che il padre italiani. Più precisamente il 90,4% dei padri è italiano mentre nel caso delle madri tale percentuale si abbassa solo leggermente (87,2%).

Può essere interessante andare ad analizzare nel particolare il numero di coppie miste di genitori. Nella Figura 2.16 è riportato il barplot con le percentuali di genitori di diversa provenienza: la maggior parte delle famiglie è composta da genitori italiani, solo una piccola parte inferiore al 10% sono entrambi di altri paesi.

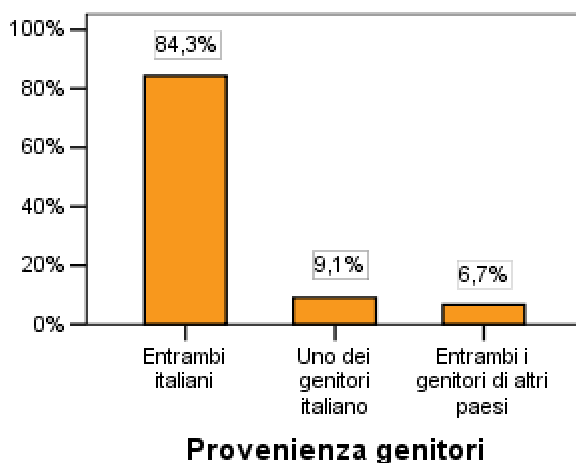


Figura 2.16: Istogramma con le percentuali di provenienza delle coppie di genitori.

2.6 GIUDIZIO ULTIME PAGELLE

In questa parte del questionario vengono raccolti dati relativi all'andamento scolastico di alcune materie, quali matematica, lingua italiana ed educazione fisica. Le risposte sono codificate su scala *Likert* che va da 1 (ottimo) a 5 (non sufficiente). Dalla Figura 2.17 si nota che per matematica e italiano i voti si distribuiscono in maniera simile, mentre per educazione fisica i voti si distribuiscono più uniformemente in tutto il range di voti. Più precisamente la mediana per matematica e italiano risulta uguale a distinto con uno scarto interquartile di 1, mentre per educazione fisica la mediana risulta sempre distinto ma con uno scarto interquartile di 2.

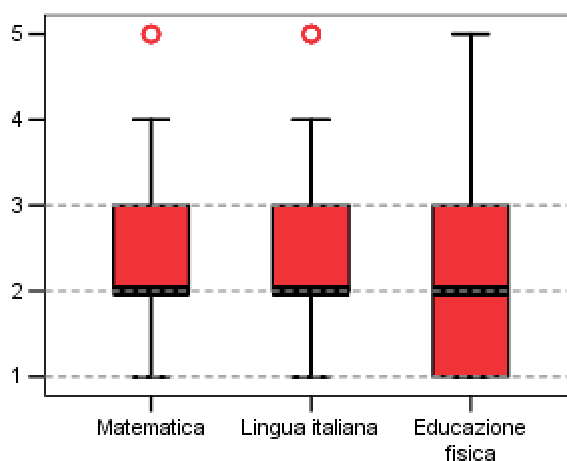


Figura 2.17: Boxplot giudizi di tre materie scolastiche nell'ultima pagella.

Una ulteriore analisi viene proposta suddividendo le varie materie e controllando la differenza dei voti per genere. Nelle Figure 2.18 e 2.19 sono riportati i boxplot relativi a tale analisi. Dal test di Wilcoxon effettuato per i voti delle pagelle differenziati per genere, risulta che per matematica non c'è effettivamente una differenza fra i sessi ($p\text{-value}=0,2$), mentre per italiano e educazione motoria c'è differenza fra maschi e femmine ($p\text{-value}=0,009$ e $p\text{-value}=0,005$ rispettivamente).

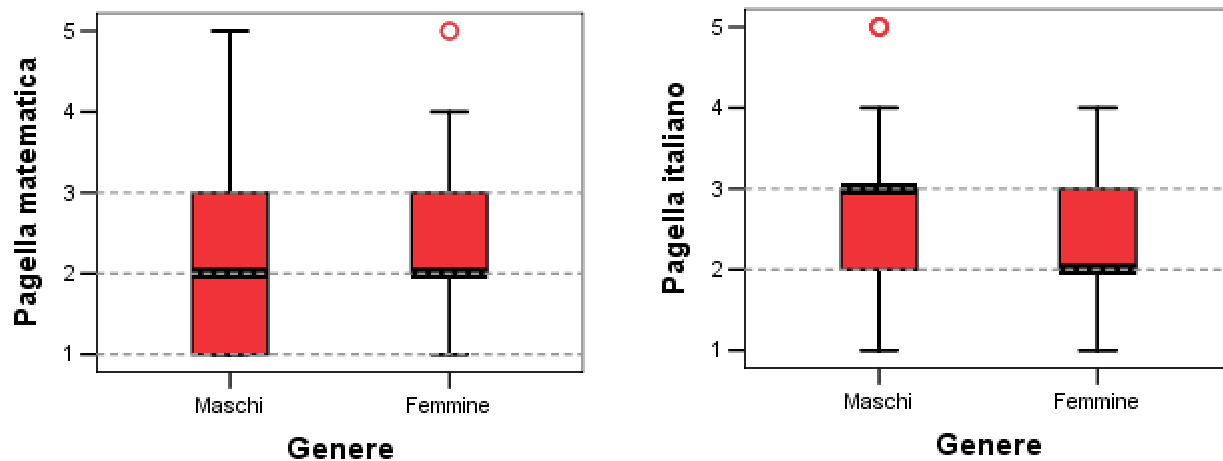


Figura 2.18: Boxplot giudizi ultima pagella di italiano e matematica distribuiti per genere.

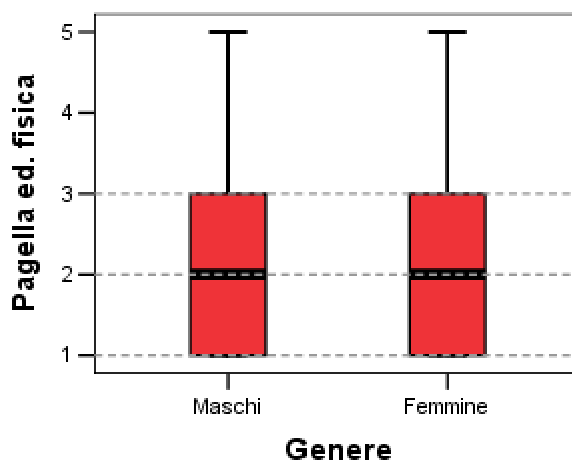


Figura 2.19: Boxplot giudizi ultima pagella di educazione fisica distribuiti per genere.

Per valutare se vi sono correlazioni fra il voto di una materia ed un'altra vengono riportati nella Tabella 2.2 i valori dei coefficienti di *Spearman* (Ercolani, Areni e Leone, 2002 Cap. 3.4.2).

La correlazione risulta significativa (livello 0,01) da cui si evince che vi é una correlazione marcata tra tutte le materie, in pratica coloro che hanno buoni voti in educazione fisica hanno anche un buon profitto in entrambi gli altri insegnamenti.

Correlazione di Spearman	Matematica	Italiano	Ed. fisica
Matematica	1	0,655	0,387
Italiano	0,655	1	0,368
Ed. fisica	0,387	0,368	1

Tabella 2.2: Correlazione di Spearman tra i voti per differenti materie.

La correlazione risulta significativa a livello 0,01.

Viene anche indagato il titolo di studio di uno dei genitori, non riportando nello specifico chi fra i due. Nella Figura 2.20 viene riportato il barplot con le percentuali del titolo di studio.

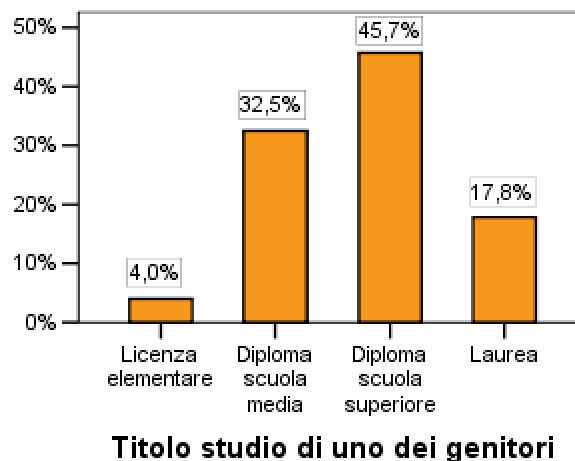


Figura 2.20: Istogramma con le percentuali del titolo di studio di uno dei genitori.

2.7 CON QUALE MEZZO SI VA A SCUOLA E QUANTO CI SI IMPIEGA

Un ulteriore modo per capire quanto la sedentarietà é rilevante in una giornata tipo di uno studente consiste nel valutare come ci si sposta per arrivare a scuola. Le possibili scelte nel questionario sono l'automobile, la bici, il bus, a piedi e un tragitto misto che prevede sia bus che a piedi. Per ogni mezzo scelto si devono riportare le volte alla settimana che viene utilizzato e i minuti di impiego. Nella Tabella 2.3 vengono riassunte le percentuali dell'utilizzo di un determinato mezzo di trasporto e le relative medie della frequenza di utilizzo settimanale e dei minuti per lo spostamento.

Mezzo Trasporto	Si %	Media Volte Settimana (sd)	Media Minuti per volta (sd)
Con l'Automobile	55,4%	4,75 (1,58)	5,88 (4,87)
A piedi	24,1%	4,29 (1,73)	7,38 (5,42)
Con il bus	20,7%	5,32 (0,99)	17,09 (11,97)
Con la bicicletta	15,2%	4,39 (1,76)	5,71 (3,92)
A piedi e con il bus	3,9%	5,13 (1,26)	14,08 (10,23)

Tabella 2.3: Frequenze di utilizzo dei mezzi per recarsi a scuola.

Il mezzo di trasporto più utilizzato risulta l'automobile, mentre la bicicletta ed il tragitto misto bus e a piedi hanno la percentuale più bassa. In tutti i casi la media di utilizzo é superiore alle 4 volte alla settimana: si tenga presente infatti che la frequenza massima risulta 6, che corrisponde ai giorni di scuola durante la settimana. La media di minuti per ogni mezzo invece varia di molto a seconda della scelta, il bus e il tragitto misto hanno una media di minuti circa tre volte superiore agli altri mezzi: ciò é coerente con il fatto che chi utilizza tale mezzo presumibilmente ha una distanza da scuola maggiore che negli altri casi. Ciò che si nota dai dati è che l'auto è il mezzo più usato nonostante il tragitto sia breve, in media inferiore ai sei minuti.

2.8 GIUDICA TE STESSO

Un consistente numero di domande é stato inserito nel questionario per dare un giudizio sul proprio aspetto fisico e sulle proprie capacità poste a confronto con gli altri. Nella Tabella 2.4 sono sintetizzate le affermazioni con le percentuali di risposta (in grassetto sono segnalate le mediane).

Affermazione	Per niente	Poco	Abbastanza	Molto
Mi piace il mio corpo così com'è	3,9	9,0	44,3	42,9
Sono molto bravo nello sport	2,8	7,3	48,8	41,2
Spesso dimentico quello che ho imparato a scuola	25,6	46,7	22,8	5,0
In confronto agli altri mi piaccio	7,1	15,2	49,7	28,1
Ho buoni motivi per essere orgoglioso di me stesso	3,7	10,2	45,6	40,5
Spesso sono triste senza motivo	41,1	35,1	15,8	8,0
Cambierei volentieri il mio aspetto	54,6	20,7	11,1	13,5
Facendo sport non mi accorgo di come passa il tempo	16,5	10,3	19,7	53,4
Faccio i compiti scolastici molto rapidamente	11,2	22,4	45,3	21,1
Complessivamente sono molto soddisfatto di me stesso	3,3	10,6	44,5	41,6
Nello sport imparo più velocemente degli altri della mia età	13,0	22,5	41,6	22,9
Certo di evitare le attività sportive	77,7	12,3	4,9	5,1
Sono molto contento del mio corpo	6,7	10,0	39,1	44,2
Vado molto bene a scuola	2,6	11,6	58,8	27,0
Non ho una grande stima di me stesso	40,4	24,9	23,7	11,0
Rido solo raramente	56,3	18,6	14,1	10,9
Molti ragazzi hanno un aspetto migliore del mio	12,1	34,1	36,6	17,1
Con lo sport dimentico tutto intorno a me	40,6	17,8	19,3	22,3
Imparo molto velocemente i nuovi esercizi sportivi	5,5	11,2	39,3	44,0
Credo di essere intelligente tanto quanto i ragazzi della mia età	6,8	16,2	40,8	36,2
Sono contento di me stesso	2,7	7,9	33,2	56,2
Durante il mio tempo libero preferirei fare sport più di ogni altra cosa	10,7	17,4	31,6	40,2
Mi sento veramente in forma	6,4	11,3	38,9	43,4
Spesso mi siedo e non vorrei fare nulla	37,4	31,2	17,2	14,1
Credo di essere bravo nello sport tanto quanto gli altri della mia età	7,9	10,9	37,4	43,8
Non sono bravo a scuola	49,4	25,9	19,5	5,3
Non mi piace il mio corpo	55,0	19,0	16,5	9,5
Quando gli altri si divertono io non posso ridere con loro	61,5	16,3	10,2	11,9
Mi sento impacciato nel mio corpo	66,4	19,4	9,7	4,6
Non sono bravo nello sport	61,7	19,7	11,4	7,2
Ho un aspetto migliore della maggior parte dei miei amici	22,8	33,0	33,5	10,6
Durante le lezioni mi é difficile trovare la risposta alle domande	29,3	38,7	23,9	8,2
Non c'è niente che mi piaccia davvero	72,8	15,0	6,7	5,5

Tabella 2.4: Percentuali di risposta alle domande per giudicare se stessi.

Come si nota dal tipo di affermazioni poste, il giudizio che ogni studente deve dare riguarda principalmente il proprio aspetto fisico, l'approccio con lo sport e la scuola.

In questo tipo di questionari, per evitare che durante le risposte si utilizzi una "scorciatoia di pensiero" e si scelga la risposta che occupa la stessa posizione nella scala, senza considerare il contenuto della domanda (*response effect*), si introducono nella batteria alcune domande di controllo a polarità invertita, in modo tale si può controllare la sincerità delle risposte.

Dalla percentuale di risposte alle affermazioni si nota come alla maggioranza degli studenti piace il proprio corpo e non lo cambierebbe e si ritengono bravi nello sport. Interessante è la percentuale di risposte all'affermazione: "spesso sono triste senza motivo" per cui la somma delle frequenze di poco e abbastanza supera il 40%. Stesso discorso per l'affermazione "non ho una grande stima di me stesso" in cui la somma di poco e abbastanza supera il 45%, in controtendenza rispetto ad altre affermazioni dal significato simile.

2.9 QUALCOSA DI PIU' SULL'ALIMENTAZIONE

Per capire le abitudini alimentari degli adolescenti, viene chiesto quali tipi di pasto sono consumati durante il giorno, specificando inoltre con chi sono assunti nell'arco temporale dal lunedì al venerdì. Le risposte in questo caso sono codificate come: 1 il pasto é consumato in famiglia, 2 con i fratelli, 3 da solo e 4 con gli amici.

I pasti sono: la prima colazione, consumata regolarmente da più del 92% degli studenti, la merenda (86%), il pranzo (99,7%), lo spuntino pomeridiano, ossia torte o dolci (60%), lo spuntino pomeridiano o frutta (63%), la cena (99,1%), lo spuntino di tarda sera (27%).

Le Figure 2.21 e 2.22 riassumono le percentuali di chi ha risposto positivamente al consumo regolare del pasto specificato. Ogni grafico riporta per ogni tipo di pasto, le frequenze rispetto a con chi viene consumato.

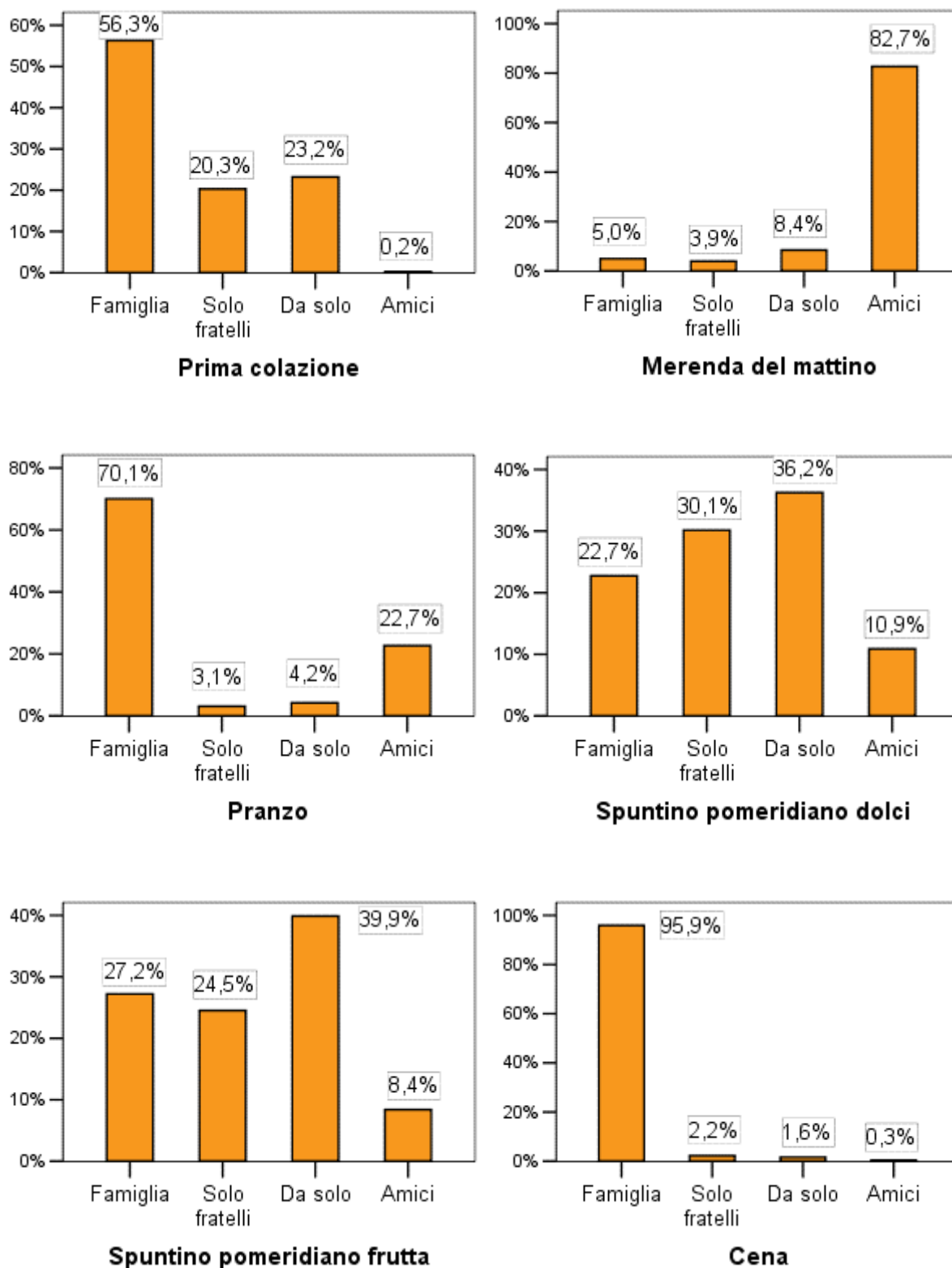


Figura 2.21: Istogrammi delle percentuali che indicano con chi si consumano i pasti durante la settimana da lunedì a venerdì.

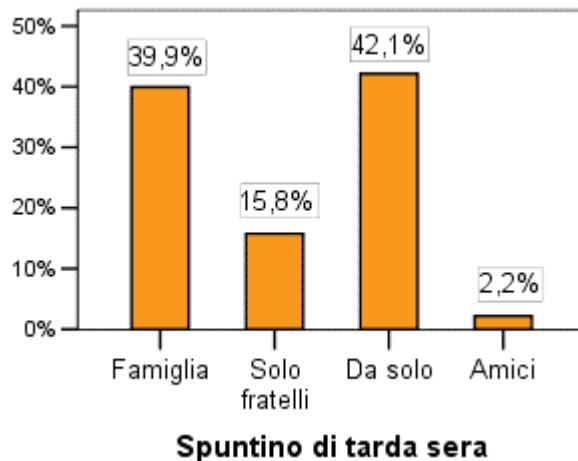


Figura 2.22: Istogramma delle percentuali che indicano con chi si consuma lo spuntino di tarda sera durante la settimana da lunedì a venerdì.

Dai grafici si nota come i pasti “principali”, quali prima colazione, pranzo e cena, sono consumati dalla quasi totalità dei rispondenti e in famiglia. Gli altri tipi di pasto, soprattutto gli spuntini e il dopo cena, si distribuiscono tra famiglia, fratelli e da solo. Unica eccezione é la merenda del mattino che viene consumata con gli amici: infatti, tale pasto corrisponde all’ora della merenda durante la scuola.

La tabella 2.5 riporta le percentuali di risposta sulla frequenza di consumo di alcune tipologie di cibo negli ultimi sei mesi (in grassetto sono segnalate le mediane).

Dalle distribuzioni di frequenza delle risposte si nota come la tipologia di cibi poco consumati siano: cornflakes e muesli (solitamente consumati a colazione), i pasticcini salati e gli hamburger (Big-Mac). La pasta e la carne invece sono i cibi consumati più volte alla settimana da più del 50% dei rispondenti. Per le bevande l’acqua risulta consumata dalla maggior parte degli adolescenti tutti i giorni, mentre, bevande come la cola più volte alla settimana.

Tipologia di cibo	Tutti i giorni	Più volte alla settimana	Una sola volta a settimana	Meno o mai
Pane fresco, pane ai cereali, pane nero	30,1	33,5	11,9	24,4
Pane conservato, pane da toast, cracker	15,6	39,2	21,2	24,0
Marmellata, miele, nutella	13,4	34,9	25,9	25,8
Cornflakes, corn pops	11,3	18,1	18,8	51,7
Muesli	4,0	7,3	6,9	81,9
Formaggio, ricotta, yogurt	18,0	39,1	21,6	21,3
Carne	17,9	60,5	17,6	4,0
Affettati, salame, mortadella, prosciutto	24,1	46,9	20,5	8,6
Frutta fresca	44,0	31,3	16,8	7,9
Verdura, insalata	36,8	32,4	13,6	17,1
Pasta, spaghetti	38,4	50,4	9,6	1,6
Patate, riso	10,6	41,6	37,5	10,2
Dolciumi, torte, biscotti	17,6	32,0	34,1	16,2
Pasticcini salati	4,1	14,2	27,2	54,5
Limonata, aranciata, coca cola	19,2	31,5	30,9	18,3
Patate fritte	7,3	25,3	44,9	22,6
Pesce, bocconcini di pesce	6,1	23,7	44,1	26,2
Pizza	6,4	25,0	58,9	9,8
Wurstel, salsicce	6,7	23,8	39,3	30,2
Hamburger, Big-Mac	3,7	9,1	21,7	65,6
Acqua, acqua minerale	88,4	7,2	2,4	2,1

Tabella 2.5: Percentuali di risposta sul consumo di varie tipologie di cibo.

2.10 USO DELLA TELEVISIONE, COMPUTER O VIDEOGAMES

La frequenza di utilizzo della televisione, del personal computer e dei videogames è una misura importante, soprattutto in un bambino, della sedentarietà ed è legata ad un aspetto sociale e comportamentale. In questa parte del questionario vengono raccolte delle variabili legate alla frequenza di utilizzo di tali apparati e con chi vengono utilizzati, se da solo, in famiglia o con amici. Le risposte sono sulla seguente scala: 1 significa “tutti i giorni”, 2 “più volte alla settimana”, 3 “una volta alla settimana”, 4 “più volte alla settimana”, 5 “una volta al mese o più raramente”, 6 “mai”. Di seguito sono riportati i barplot per le variabili in questione.

I tre barplot nella Figura 2.23 rappresentano la frequenza di utilizzo della televisione e con chi è vista. Le mediane per tali grafici possono venire così riassunte: la televisione viene vista da solo e con i genitori più volte alla settimana (p-value test di Wilcoxon pari a 0,79), mentre con gli amici raramente o una volta al mese. Nel boxplot di Figura 2.23 si nota come la televisione viene

prevalentemente guardata in famiglia o da solo con più frequenza, mentre con gli amici viene vista di rado.

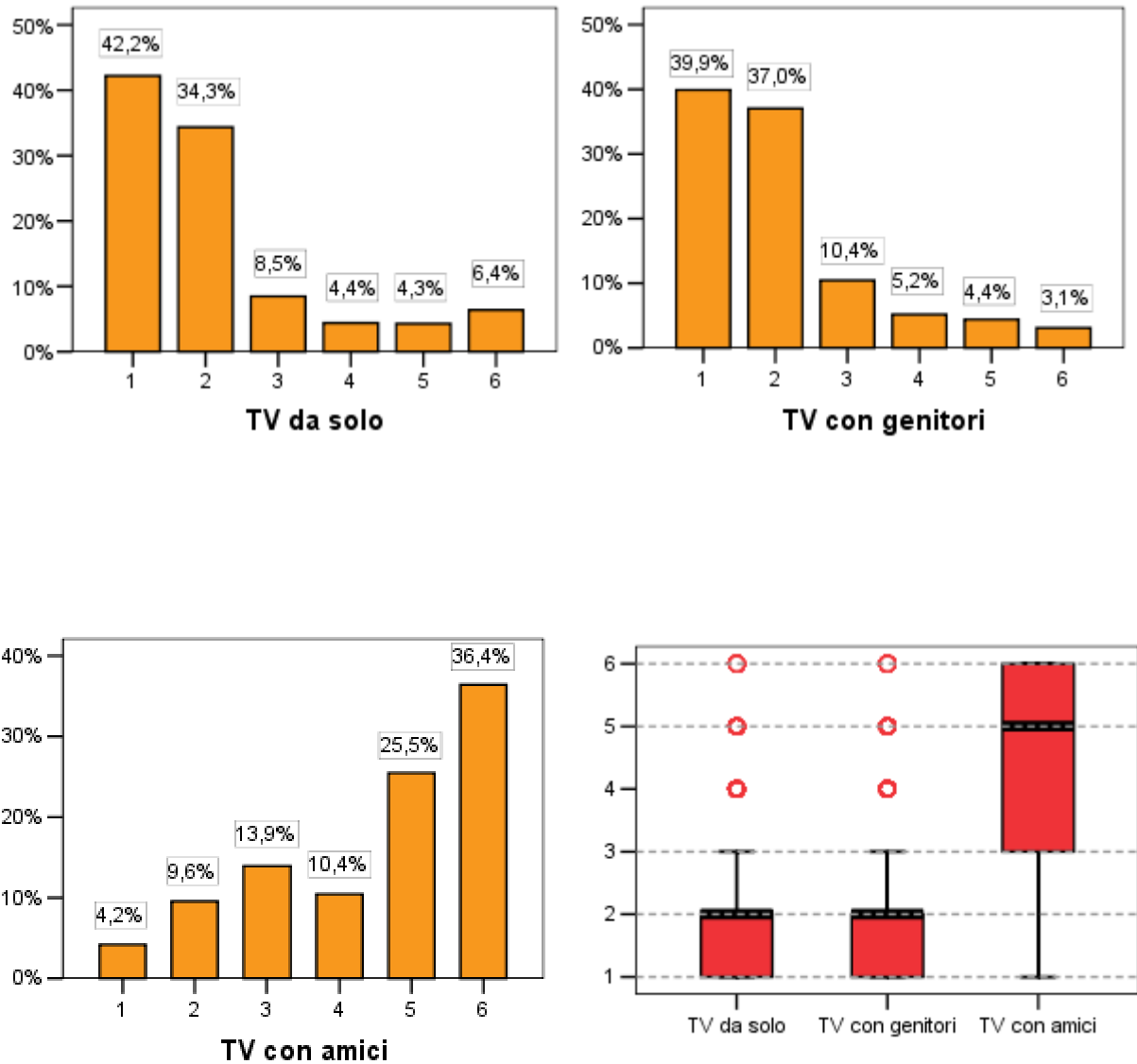


Figura 2.23: Istogrammi delle percentuali che indicano l'utilizzo della televisione e con chi viene vista e boxplot riassuntivo.

Stessa sintesi viene mostrata per l'uso del personal computer. Nei grafici della Figura 2.24, le mediane per ogni variabile sono: utilizzo del pc da solo una volta alla settimana, con i genitori mai, con gli amici raramente o una volta al mese. Dal boxplot di Figura 2.24 si nota come il computer non sia una fonte di interesse sia con gli amici che con i genitori (p-value test di Wicoxon pari a 0,52). Da soli, tuttavia, vi è un uso più frequente.

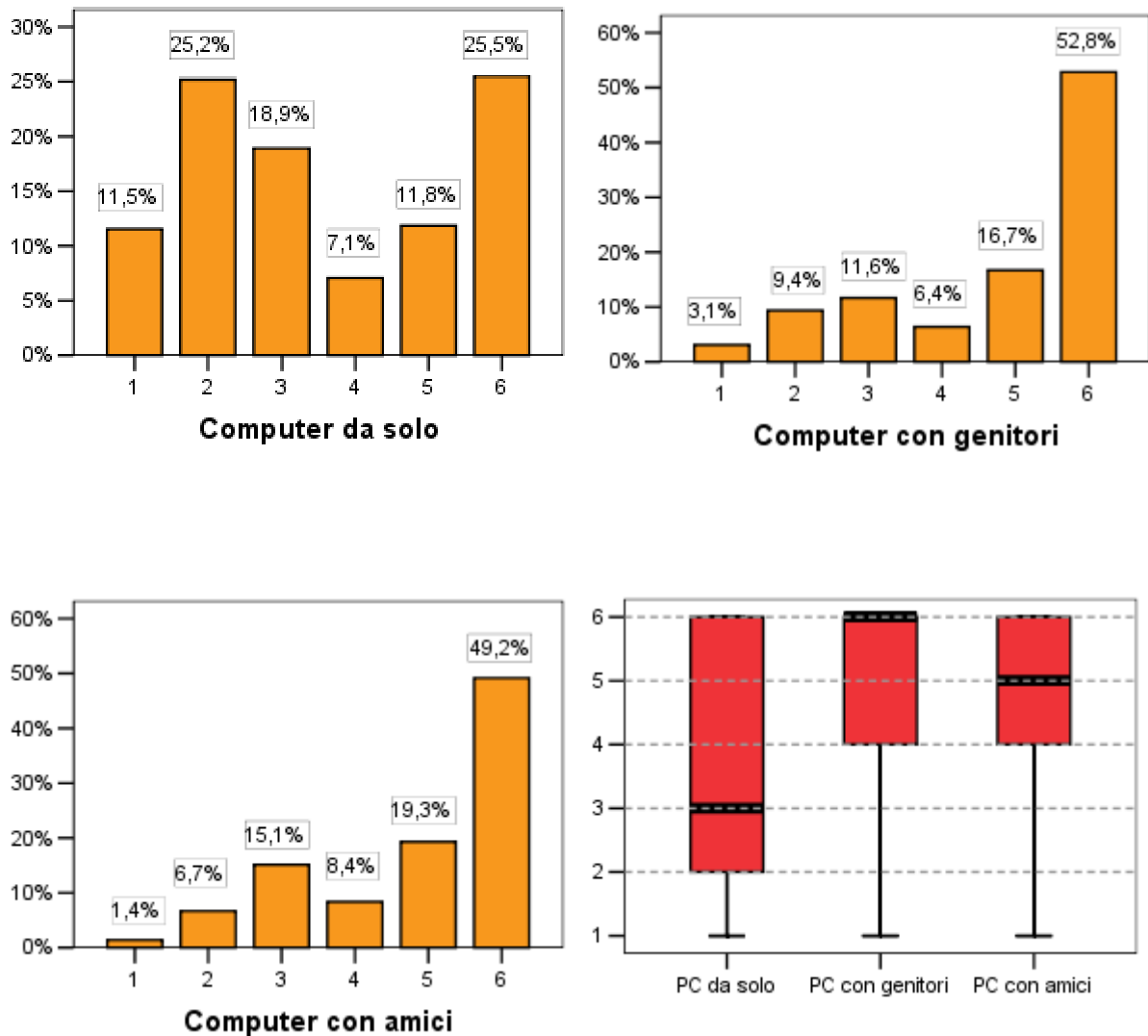


Figura 2.24: Istogrammi delle percentuali che indicano l'uso del computer e con chi viene utilizzato e boxplot riassuntivo.

Nella Figura 2.25 sono riportati i barplot relativi all'utilizzo dei videogames. Come si nota dalla distribuzione delle frequenze di utilizzo, i videogames, contrariamente a quanto ci si aspetta, sono raramente utilizzati. Le mediane in questo caso risultano: i videogames da soli sono utilizzati solo più volte al mese, con i genitori o amici invece mai. Nel boxplot, inoltre, si nota come l'utilizzo dei videogames é raro con amici o genitori, mentre da solo la frequenza di utilizzo si alza fino a più volte al mese. Il p-value del test di Wilcoxon effettuato fra le tre diverse variabili risulta in tutti i casi pari a 0, denotando una frequenza d'uso dei videogames diversa nei tre casi.

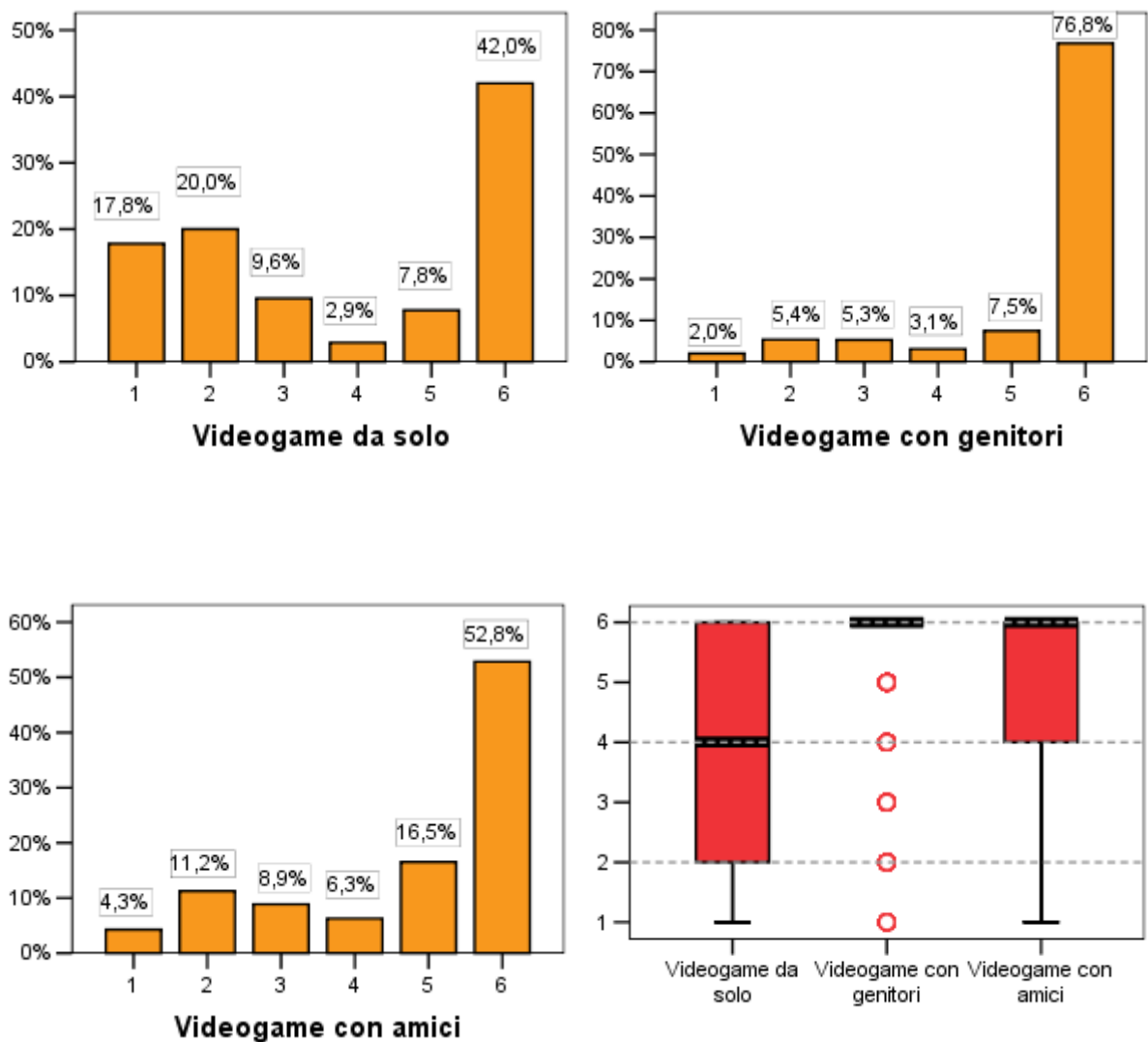


Figura 2.25: Istogrammi delle percentuali che indicano l'uso dei videogames e con chi vengono utilizzati e boxplot riassuntivo.

Le frequenze di utilizzo del computer e dei videogames da soli, si differenziano dalle altre per non essere uniformi ma sbilanciate agli estremi. Nella Figura 2.26 é riportato il boxplot relativo all'uso della televisione, del computer e dei videogames divisi per genere, nel caso "da solo". Per quanto riguarda la televisione, i boxplot mostrano una variabilità maggiore per le femmine rispetto ai maschi e la mediana in entrambi i casi risulta "più volte alla settimana" (p-value test di Wicoxon pari a 0,0001). Nel caso dell'uso del personal computer, la variabilità nei sessi é la stessa e la mediana risulta "una volta alla settimana" (p-value=0,13). Nel caso invece dell'uso dei videogames si hanno dei boxplot diversi: per i maschi la mediana risulta "più volte alla settimana", mentre per le femmine risulta "mai" (p-value=0).

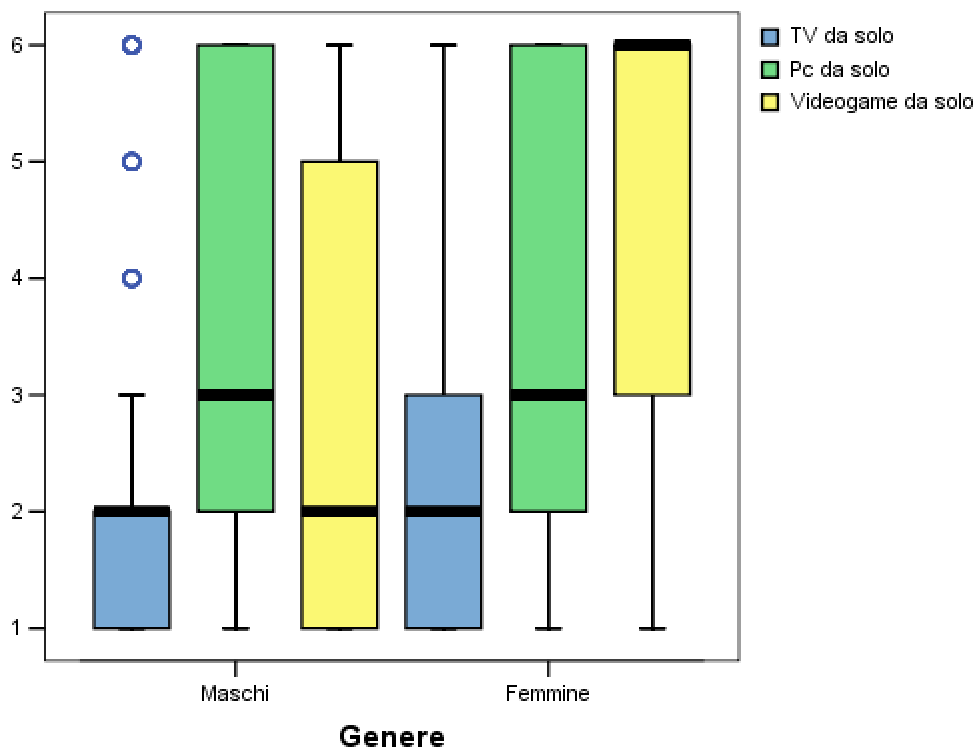


Figura 2.26: Boxplot frequenza di utilizzo Pc e Videogames suddivisi per genere.

Le variabili analizzate successivamente si riferiscono alle ore di utilizzo della televisione, computer e videogames nei giorni che vanno da lunedì a venerdì e nel fine settimana. Le risposte sono date su scala *Likert* e la scelta comprende: 1 “per niente”, 2 “circa trenta minuti al giorno”, 3 “circa un’ora al giorno”, 4 “circa due ore al giorno”, 5 “circa tre ore al giorno”, 6 “circa quattro ore al giorno”, 7 “circa cinque ore al giorno”, 8 “sei ore o più”.

Nelle Figure dalla 2.27 alla 2.32, che seguono, sono riportati i barplot con le percentuali di risposta e i boxplot di sintesi.

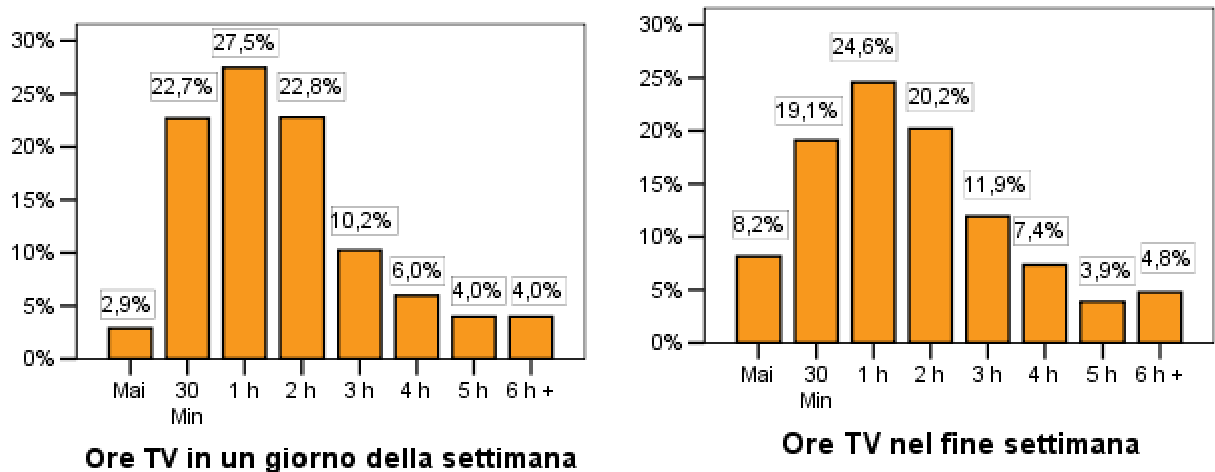


Figura 2.27: Istogrammi con le percentuali delle ore di utilizzo della televisione nei giorni della settimana e nel week-end.

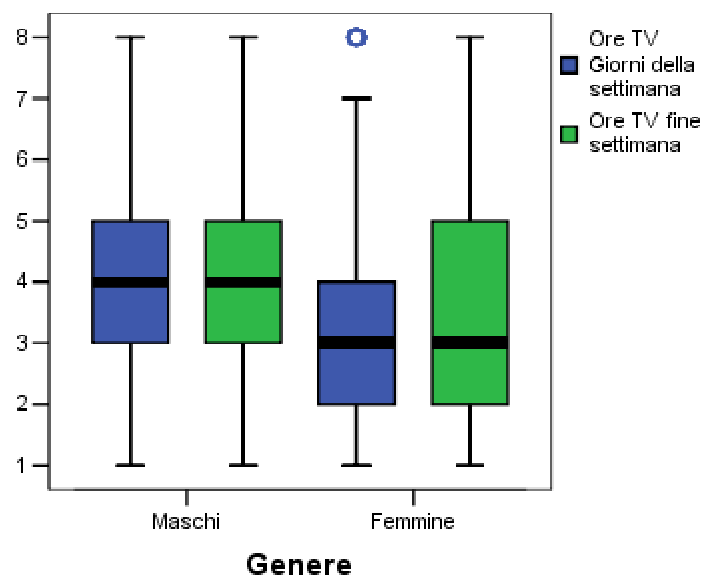


Figura 2.28: Boxplot delle ore di utilizzo della TV diviso per genere.

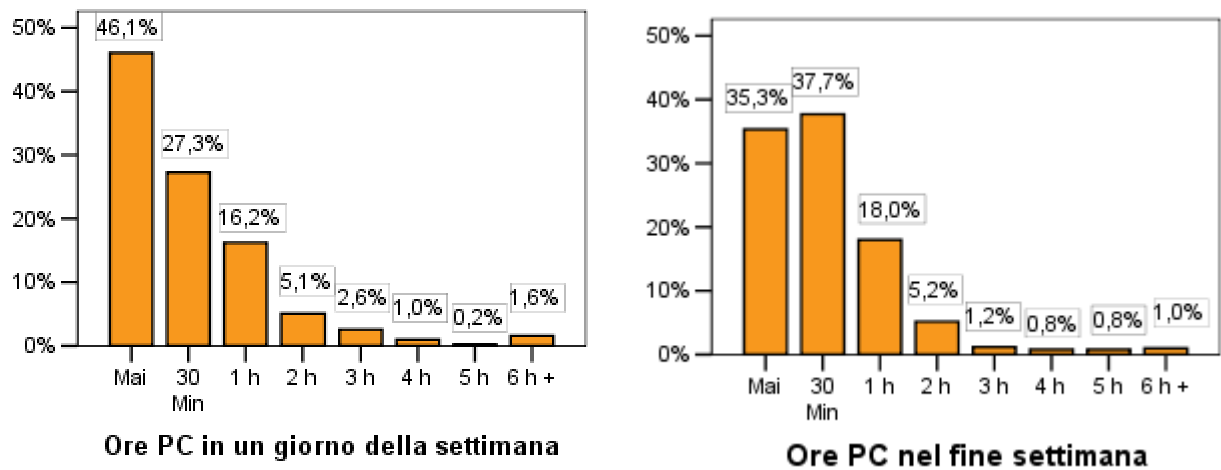


Figura 2.29: Istogrammi con le percentuali delle ore di utilizzo del computer nei giorni della settimana e nel week-end.

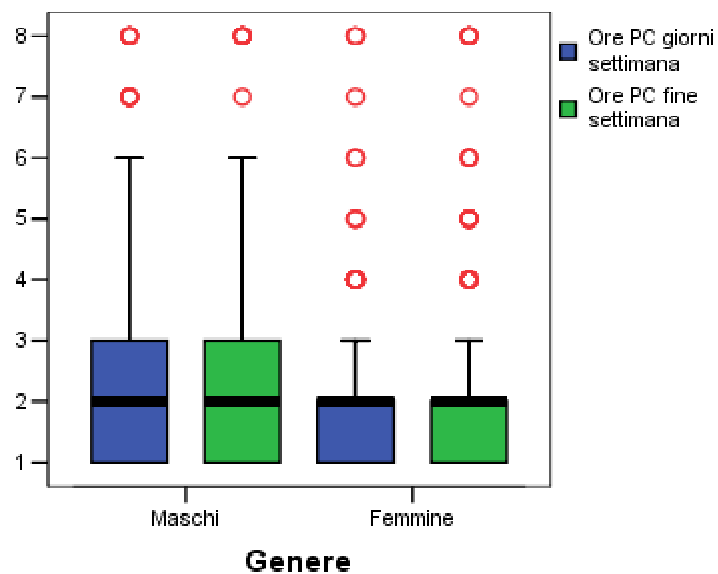


Figura 2.30: Boxplot delle ore di utilizzo del computer diviso per genere.

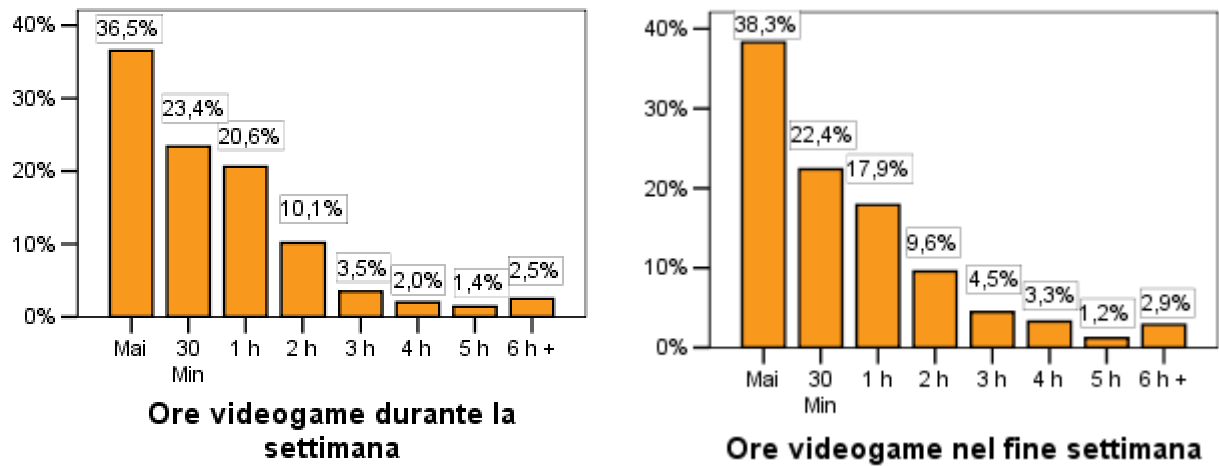


Figura 2.31: Istogrammi con le percentuali delle ore di utilizzo dei videogames nei giorni della settimana e nel week-end.

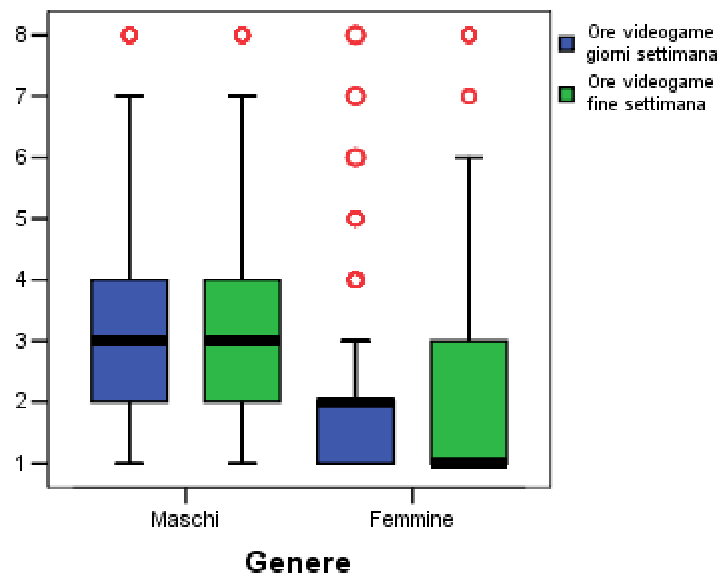


Figura 2.32: Boxplot delle ore di utilizzo dei videogames diviso per genere.

In sintesi, si può dire che la mediana per le ore dedicate alla TV, sia durante la settimana che nel fine settimana, è di circa 1 ora al giorno di utilizzo, mentre per il personal computer e i videogames corrisponde a circa 30 minuti al giorno di utilizzo.

I grafici mantengono una proporzione molto simile sia durante la settimana che nel fine settimana: le percentuali infatti non variano di molto sia nel caso della Tv che dei videogames, con un p-value del test di Wilcoxon rispettivamente di 0,96 e 0,87, mentre, nel caso del personal computer sembra esserci differenza (p-value = 0,002).

Nei boxplot si osserva come nel caso della TV, l'utilizzo per maschi e femmine risulta differente sia durante la settimana (p-value = 0,008) che nel fine settimana (p-value = 0,03), nel caso dell'uso del computer è simile durante la settimana (p-value=0,1) ma differisce nel week-end (p-value = 0,01), mentre nel caso del videogames l'utilizzo è molto diverso tra i sessi (p-value = 0) in entrambi i casi. Una sintesi fra tutti i grafici proposti in questo paragrafo e le variabili coinvolte è dato dall'indice *Media Consumption* (www.hbsc.org) riportato graficamente nella Figura 2.33, uno strumento messo a punto dal gruppo di ricerca internazionale HBSC con cui si dà una scala di valore sull'utilizzo più o meno intenso degli apparati di intrattenimento. Il barplot riporta nelle ascisse la scala di tale indice, dove 1 corrisponde ad un basso utilizzo, mentre 6 è il massimo utilizzo. Da tale indice si ricava che l'utilizzo di apparecchi da intrattenimento nel campione preso in esame non risulta molto intenso. Anzi la mediana risulta fissata ad un valore basso come 2.

Il boxplot nella Figura 2.33 riporta la suddivisione dei valori dell'indice media consumption per sesso: si nota come per i maschi la variabilità risulta maggiore rispetto alle femmine, la cui mediana si attesta a 1, il valore più basso della scala, denotando come siano i maschi ad utilizzare più intensamente gli apparati elettronici. Ciò è confermato dal test di Wilcoxon il cui p-value risulta 0.

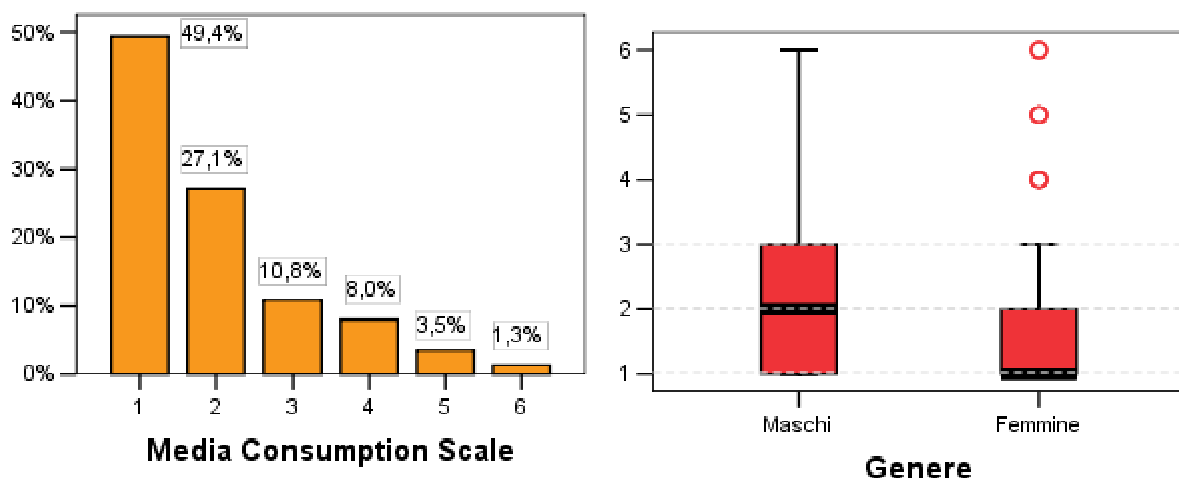


Figura 2.33: Istogramma con le percentuali dell'indice Media-Consumption e boxplot diviso per genere.

2.11 QUALCOSA DI DIFFICILE DA RINUNCIARE

A conclusione del questionario viene chiesto a cosa sarebbe difficile rinunciare tra televisione, personal computer e sport. Alla televisione hanno risposto il 28,3%, al computer il 10,6% mentre per lo sport il 61,1%. La Figura 2.34 mostra la percentuale dei rispondenti suddivisi per sesso e si nota una certa uniformità nelle scelte fra maschi e femmine (p-value=0,75).

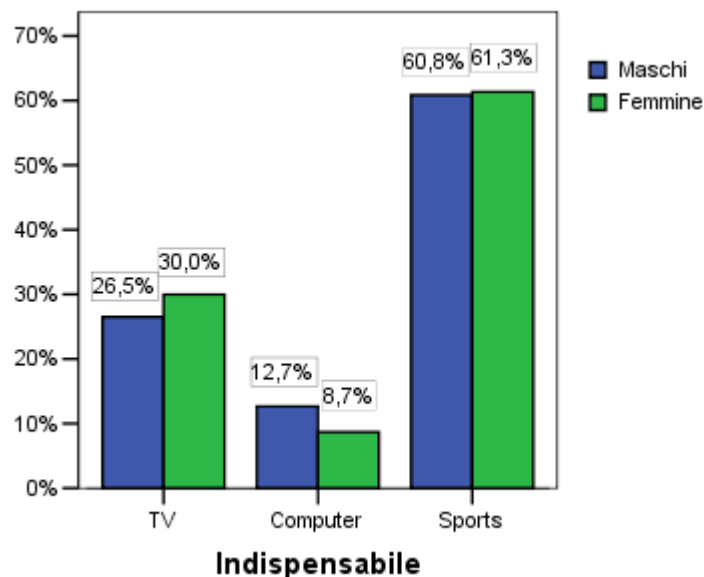


Figura 2.34: Istogramma con le percentuali di risposta suddivise per genere.

2.12 IL BODY MASS INDEX

Un indice importante che viene utilizzato spesso per misurare il grado di obesità è l'indice di massa corporea o *body mass index* (BMI). Tale indice, calcolato con il valore del peso e altezza della persona, fornisce un valore per indicare se si è in soprappeso o obesi.

Una premessa va fatta su tale indice. La definizione di obesità nel bambino o in un adolescente è più complessa rispetto all'adulto, il cui peso ideale è calcolato in base al BMI. In attesa di trovare dei parametri di riferimento più adeguati, il BMI è stato proposto anche per i più piccoli. Pertanto si definisce obeso un bambino il cui peso supera del 20% quello ideale; in soprappeso se lo supera dal 10 al 20%, oppure quando il suo BMI è maggiore del previsto.

Tale indice è riassunto in una scala riportata nella Tabella 2.6, i cui valori possono essere divisi in classi in modo indicativo.

BMI < 18,5	Sottopeso
18,5 < BMI < 25	Normopeso
25 < BMI < 30	Soprappeso
30 < BMI < 35	Obesità di I classe
35 < BMI < 40	Obesità di II classe
BMI > 40	Obesità di III classe

Tabella 2.6 : Indici di massa corporea e classi

Dai dati rilevati con l'ausilio del questionario é stato ricavato un barplot riassuntivo delle percentuali del BMI, visibile in Figura 2.35, diviso nelle sei classi principali. Tali classi sono calcolate tenendo conto dell'età, del sesso, del peso e dell'altezza di ogni adolescente.

Si nota come la maggior parte dei rispondenti ha un indice di massa corporea normale, pochi al di sotto di tale soglia, considerati magri, e un numero superiore al 20% nella soglia che va dal soprappeso all'obesità. Sempre in Figura 2.35 é riportato anche il boxplot relativo al BMI diviso per genere; le ordinate del grafico vanno dal valore 1 "molto magro", a 6 che identifica "molto obeso". Si nota che per entrambi i sessi la mediana si attesta a 3, cioè normopeso.

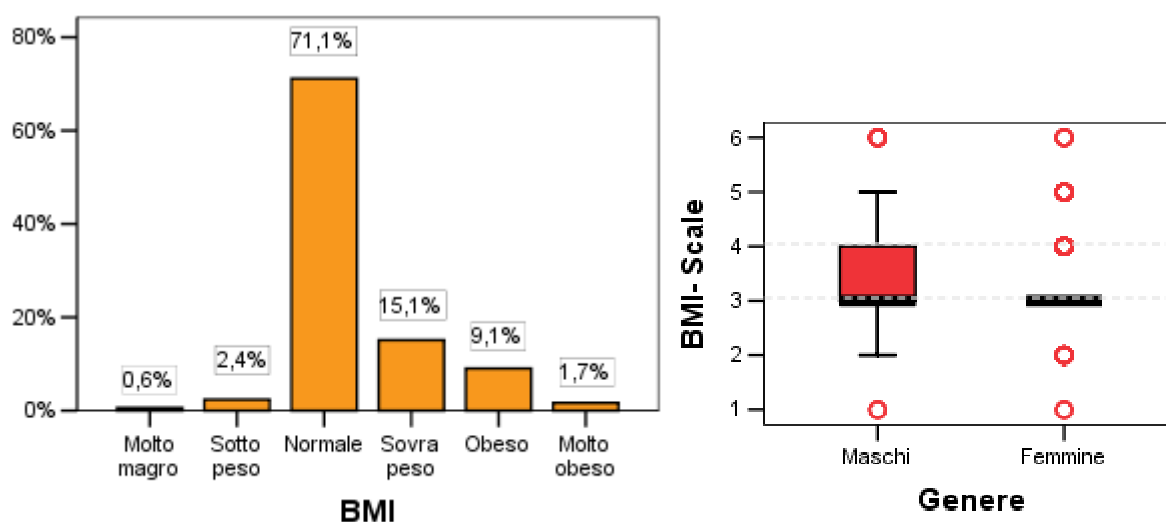


Figura 2.35: Istogramma indice BMI e boxplot diviso per genere.

L'andamento appena descritto si vede nella Figura 2.36, in cui è riportato il barplot delle percentuali del BMI suddiviso per genere. Si nota come l'indice associato ai maschi presenta, nel caso "obeso" e "molto obeso", una percentuale doppia di casi rispetto le femmine, mentre nel caso di "soprappeso" le percentuali sono le stesse, il test di Wilcoxon effettuato sul BMI discriminando per genere ne sottolinea la differenza con un p-value pari a 0,006.

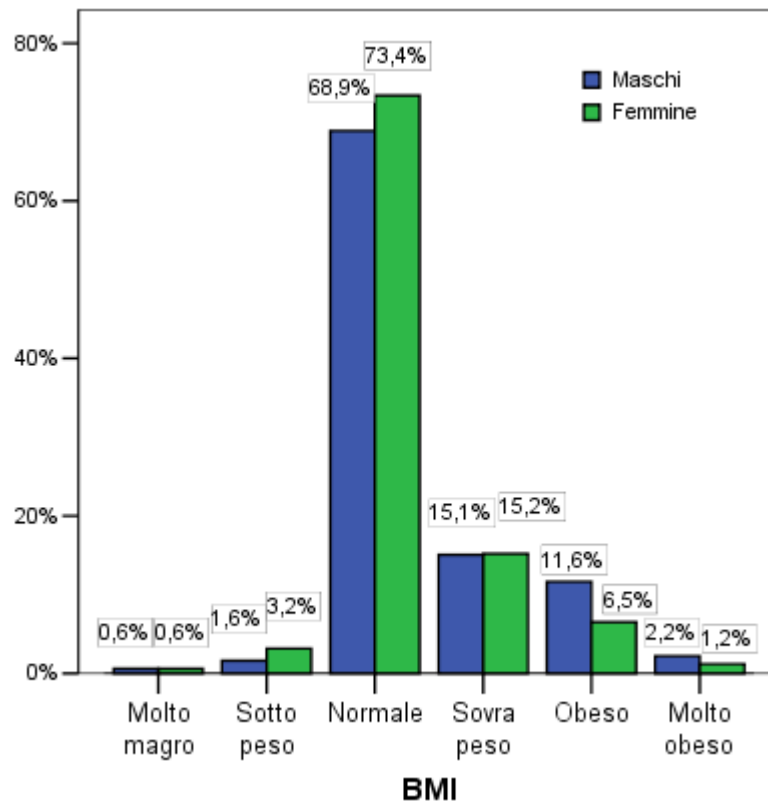


Figura 2.36: Istogramma delle percentuali indice BMI suddiviso per genere.

2.13 ANALISI DELLA CORRELAZIONE TRA GLI INDICI

Per analizzare se gli indici calcolati in precedenza quali FAS, MC e BMI, sono correlati tra loro si è provveduto ad utilizzare la correlazione di *Spearman*, inoltre gli indici sono stati opportunamente trasformati in una scala a tre valori ordinali.

Le Figure 2.37, 2.38 e 2.39 riportano gli istogrammi delle percentuali degli indici con la nuova scala, viene riportato inoltre per ognuno la suddivisione per genere. I p-value dei test di Wilcoxon effettuati per gli indici differenziati per genere risultano rispettivamente: per FAS p-value=0,69, per MC p-value=0 e per BMI p-value=0,2.

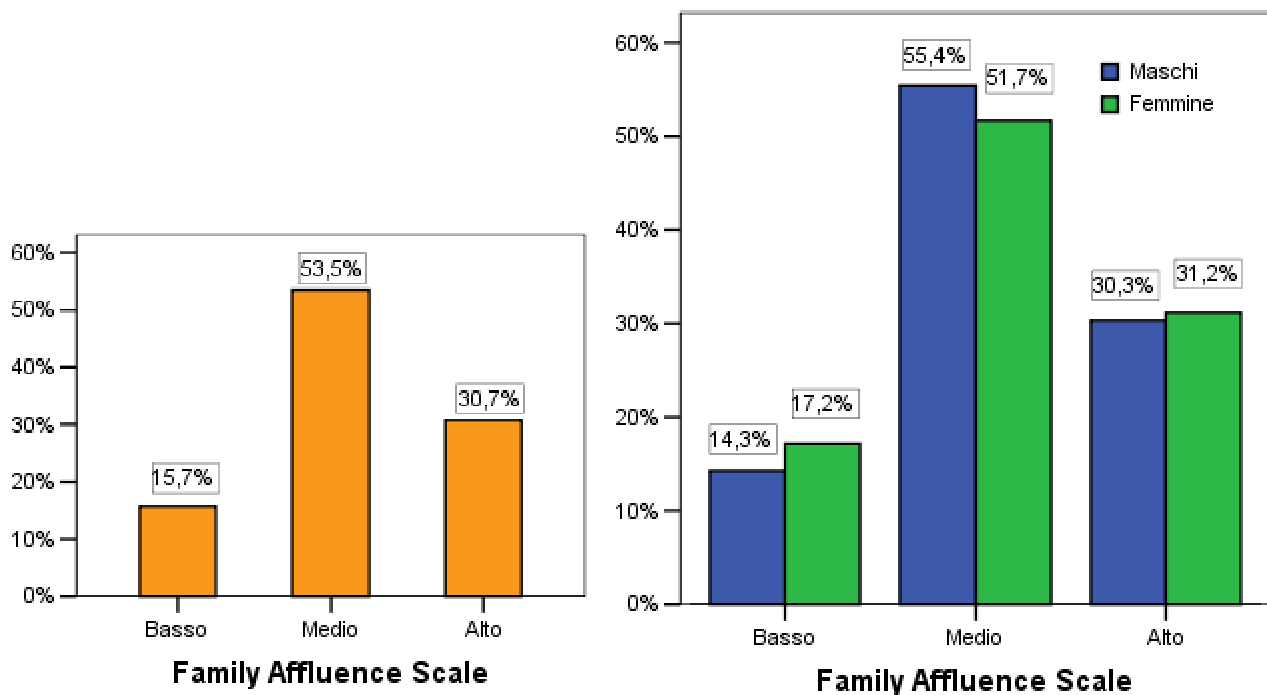


Figura 2.37: Istogrammi delle percentuali indice FAS e suddivisione per genere.

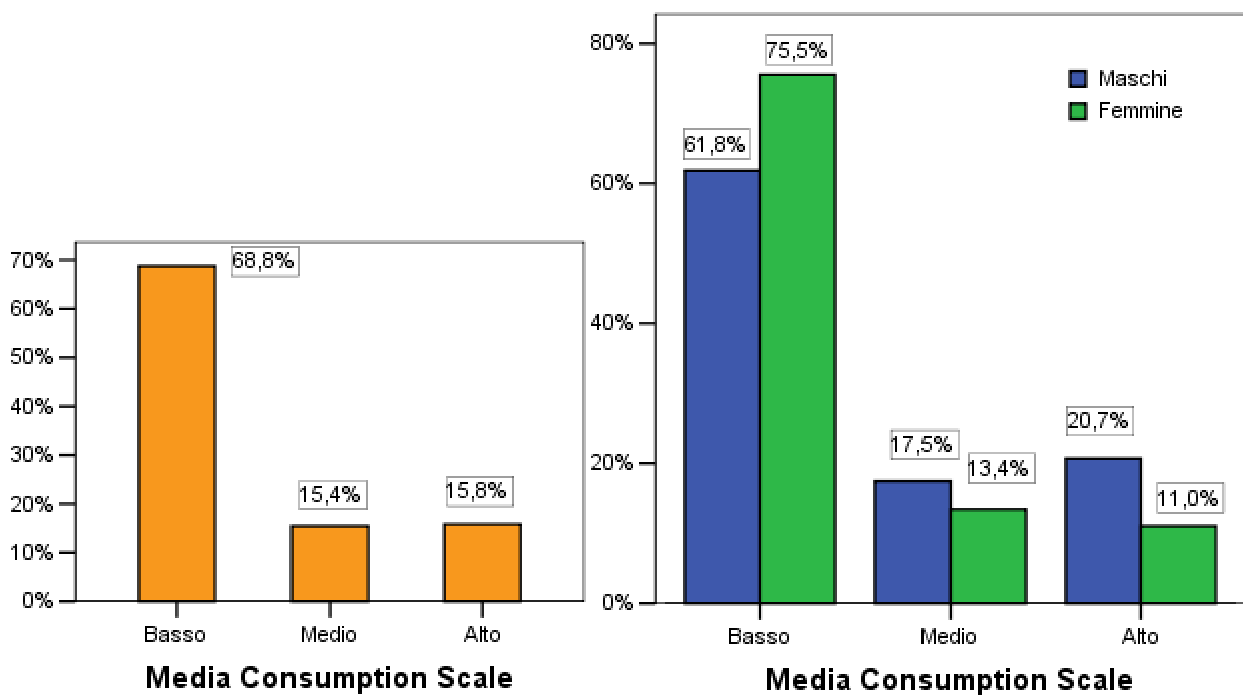


Figura 2.38: Istogrammi delle percentuali indice MC e suddivisione per genere.

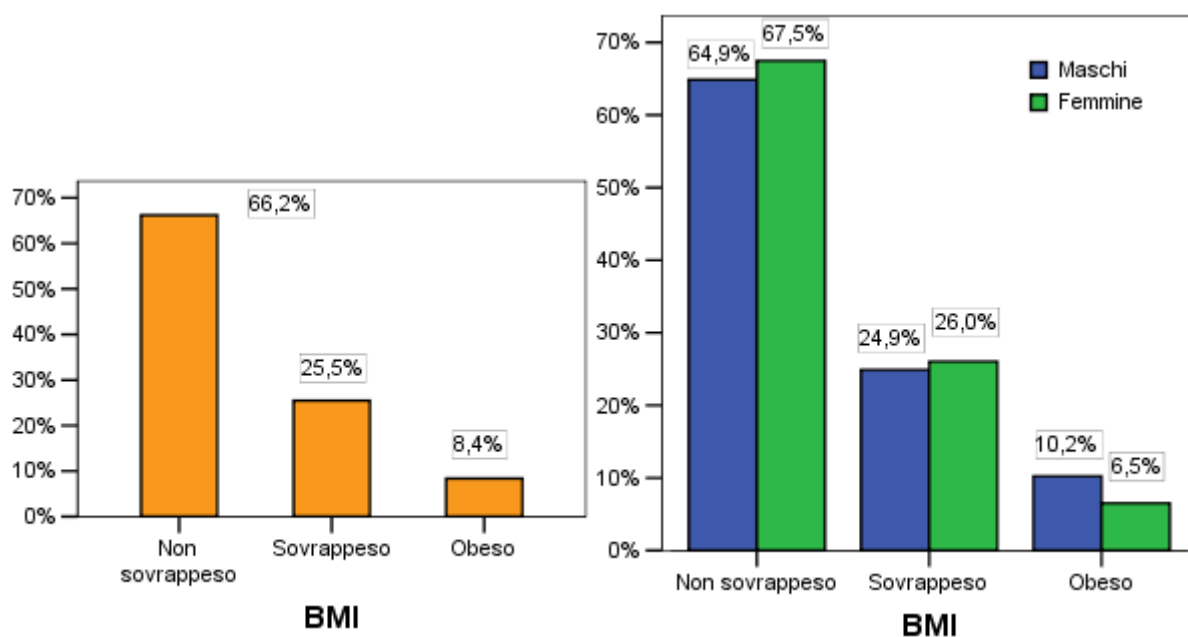


Figura 2.39: Istogrammi delle percentuali indice BMI e suddivisione per genere.

Di seguito in Tabella 2.7 sono riportati i valori della correlazione calcolata tra i vari indici.

Correlazione di Spearman	BMI	FAS	MC
BMI	1	-0,054	0,056
FAS	-0,054	1	0,064
MC	0,056	0,064	1

Tabella 2.7 : Correlazione di Spearman tra gli indici.

La correlazione risulta significativa a livello 0,05.

Dallo studio dei coefficienti si nota che la correlazione tra BMI-FAS e BMI-MC sia da considerarsi nulla (p-value test di correlazione pari a 0,08 in entrambi i casi), l'unica correlazione seppure poco marcata risulta essere fra l'indice FAS e l'indice MC (p-value del test di correlazione pari a 0,044). In tabella 2.8 e 2.9 vengono invece riportate le correlazioni fra gli indici discriminati per genere. Per quanto riguarda i maschi si nota che effettivamente c'è una correlazione fra i vari indici, negativa tra BMI e FAS, positiva negli altri casi. La correlazione negativa indica che per un maggior valore del BMI si ha una diminuzione dell'indice di benessere familiare (FAS), mentre negli altri casi con l'aumento dell'indice BMI si ha l'aumento dell'uso di apparecchi da

intrattenimento (indicato dall'indice MC) e con l'aumento dell'indice FAS si ha un aumento di MC. Questo ultimo punto può essere spiegato alla luce dei dati precedentemente studiati, in cui si evidenziava per i maschi l'uso degli apparecchi da intrattenimento come la TV, il computer e i videogames, che data la loro natura hanno un costo elevato che va rapportato quindi con il benessere familiare riassunto dal FAS.

Per le femmine invece la correlazione risulta assente e ciò è confermato dal p-value del test di correlazione che risulta in tutti i casi superiore a 0,6 confermando l'ipotesi nulla di incorrelazione.

MASCHI			
Correlazione di Spearman	BMI	FAS	MC
BMI	1	-0,129	0,076
FAS	-0,129	1	0,137
MC	0,076	0,137	1

Tabella 2.8 : Correlazione di Spearman tra gli indici per i maschi.

La correlazione risulta significativa a livello 0,05.

FEMMINE			
Correlazione di Spearman	BMI	FAS	MC
BMI	1	0,019	0,022
FAS	0,019	1	-0,019
MC	0,022	-0,019	1

Tabella 2.9 : Correlazione di Spearman tra gli indici per le femmine.

La correlazione risulta significativa a livello 0,05.

ANALISI DELLE CORRISPONDENZE E ANALISI CLUSTER

In questo capitolo verrà utilizzata l'analisi delle corrispondenze principali (ACOMU) utile per capire la struttura della dipendenza interna delle variabili di un dataset, tale tecnica si basa su una rappresentazione grafica delle modalità delle variabili. Successivamente si farà uso dell'analisi cluster utile per compattare l'insieme di unità statistiche studiate in gruppi.

3.1 ANALISI DELLE CORRISPONDENZE

L'analisi delle corrispondenze multiple è una tecnica utilizzata per analizzare dataset di grandi dimensioni, come nel seguente caso, dove sono presenti variabili di tipo qualitativo o variabili di tipo quantitativo opportunamente suddivise in classi. Nell'ACOMU si distinguono due tipi di variabili, variabili dette "attive" e variabili dette "illustrative" entrambe scelte da chi svolge l'analisi. Le variabili attive sono così chiamate perché vengono utilizzate direttamente per la formazione degli assi fattoriali che riassumono l'intreccio di interdipendenza tra le variabili stesse. Le variabili illustrative invece sono escluse dalla fase di estrazione dei fattori, ma vengono successivamente utilizzate proiettandole sugli assi fattoriali precedentemente individuati e utilizzate per l'interpretazione di questi ultimi.

In questo caso le variabili attive scelte sono gli indici BMI, FAS e MC che sintetizzano già a priori alcune caratteristiche degli adolescenti e altre due variabili quali l'importanza dello sport e i voti della pagella di educazione fisica, opportunamente codificate in una scala a tre valori (del tipo: alto, medio, basso) per evitare possibili problemi causati dal numero di modalità (effetto Guttman). Si è scelto inoltre un numero di fattori pari a 2, vista la bassa numerosità di variabili prese in esame e per poter rappresentare graficamente i risultati dell'analisi. In Figura 3.1 è riportato il grafico che si ottiene dall'analisi delle corrispondenze, le variabili prese in esame sono ordinabili per questo il grafico riporta le spezzate che uniscono i punti modalità e che permettono una migliore interpretazione degli assi fattoriali.

Analisi corrispondenze

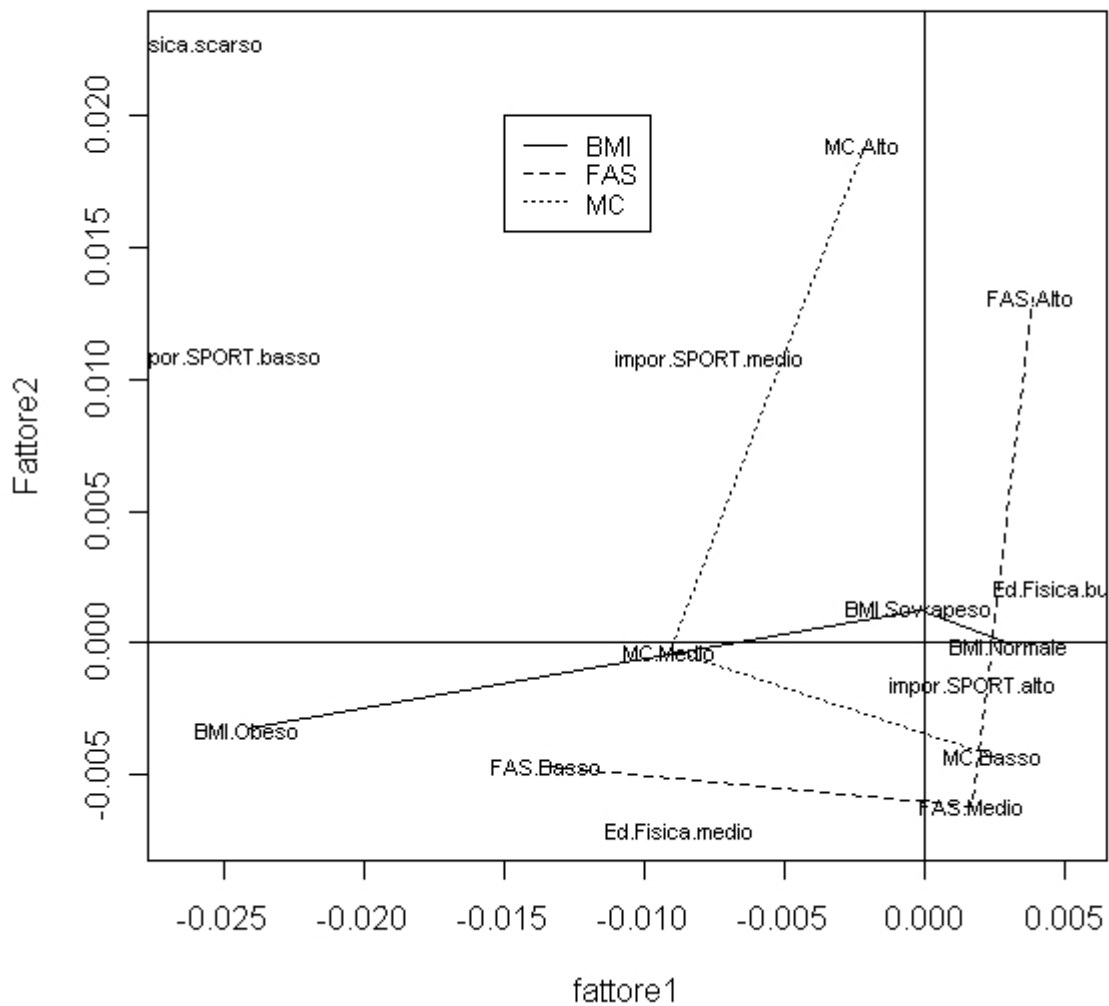


Figura 3.1: Grafico dell'analisi delle corrispondenze multiple.

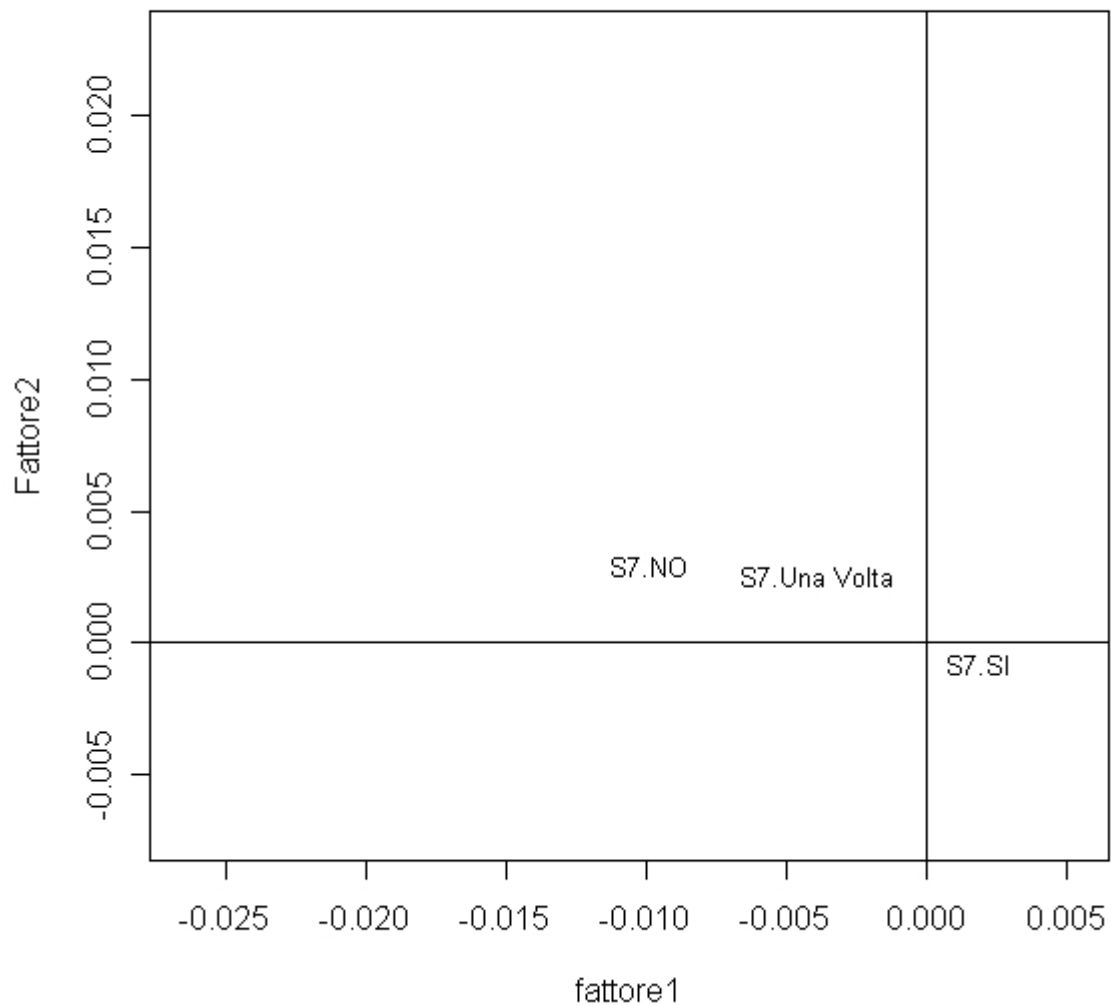
Dal grafico di Figura 3.1 si nota che a destra si concentrano le modalità “positive” quali un indice BMI normale, un buon voto in educazione fisica, un’alta importanza data allo sport, un FAS medio-alto e un MC basso. Nella parte sinistra del grafico invece sono raggruppate le modalità “negative” tra cui un indice BMI indicante obesità, nella parte centrale si hanno le modalità indicanti un valore medio. Da ciò si deduce che l’asse orizzontale, che individua il primo fattore, rappresenti la scala di misura delle variabili considerate. L’interpretazione del secondo fattore, individuato dall’asse verticale, risulta più difficile, si nota che nella parte superiore del grafico si concentrano le modalità elevate per gli indici MC e FAS e nel contempo le modalità più basse per l’importanza dello sport e il voto di educazione fisica, nella parte inferiore del grafico invece le modalità per le stesse variabili

assumono il valore opposto. Da tale comportamento si deduce che il secondo fattore è in qualche modo collegato alle “abitudini” legate al benessere familiare e all’uso di apparati di intrattenimento. Il grafico evidenzia inoltre un gruppo di modalità che si concentrano nella parte in basso a destra le quali sembrano individuare gli studenti più “attivi” con un indice BMI normale che usano poco gli apparecchi da intrattenimento e a cui piace lo sport.

3.2 VARIABILI ILLUSTRATIVE

Dopo aver svolto l’analisi delle corrispondenze si procede proiettando sugli assi fattoriali le variabili illustrative per migliorare l’interpretazione degli assi. Le variabili scelte come illustrative sono: “iscrizione ad una società sportiva”, “gioco e mi scatenano all’aperto” e “per me sarebbe veramente difficile rinunciare a”. Sono state analizzate anche altre variabili quali il sesso, il titolo di studio dei genitori e la provenienza dei genitori, ma non vengono riportati i grafici relativi perchè non sono risultate influenti nell’analisi in quanto le modalità si trovano molto vicine all’intersezione degli assi.

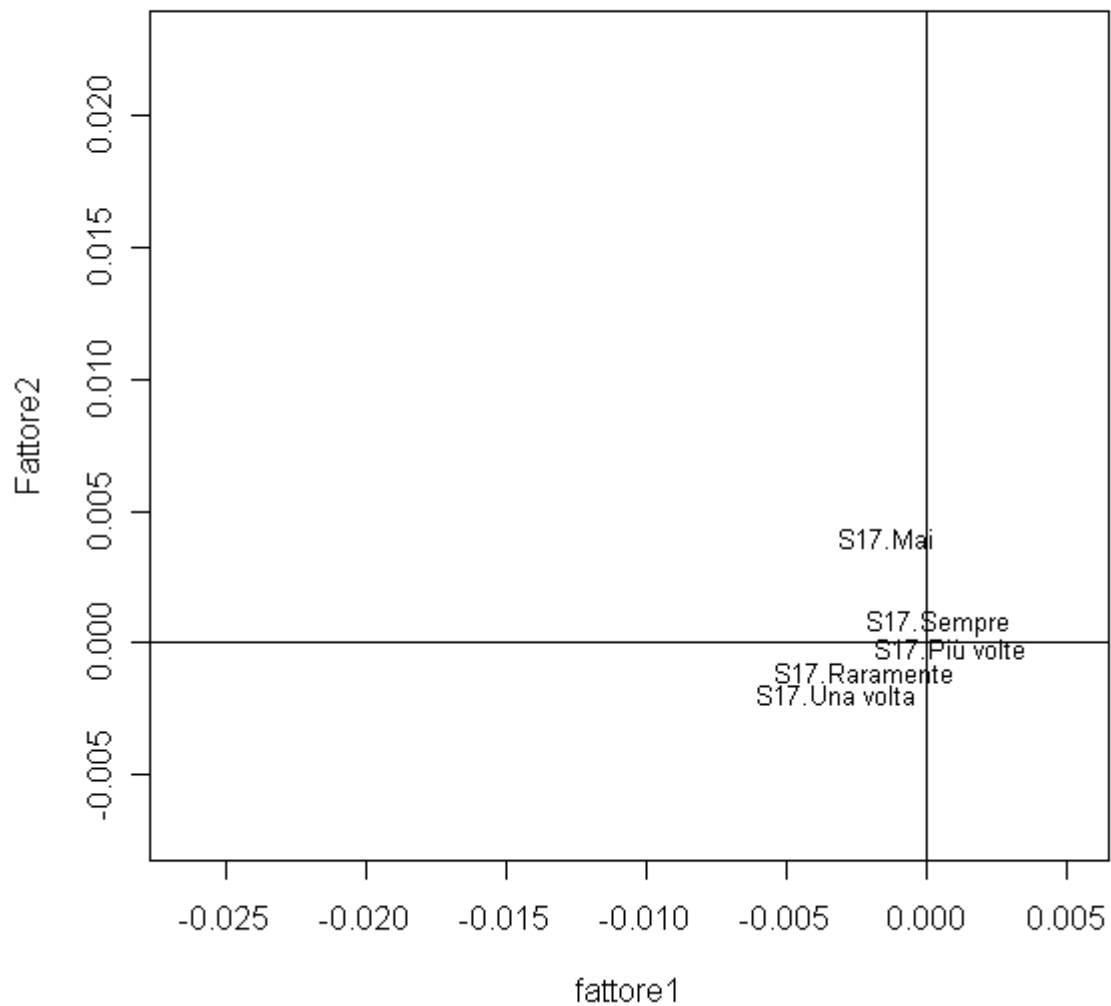
Illustrativa



*Figura 3.2: Proiezione sugli assi fattoriali della variabile illustrativa
“Iscrizione ad una società sportiva”.*

In Figura 3.2 è riportato il grafico della proiezione della variabile illustrativa “iscrizione ad una società sportiva” sugli assi fattoriali, si nota come le modalità sono allineate seguendo l’asse orizzontale che ne indica la scala di misura, inoltre le modalità “no” e “una volta” si trovano sopra l’asse orizzontale e indicano un’abitudine “sedentaria”.

Illustrativa



*Figura 3.3: Proiezione sugli assi fattoriali della variabile illustrativa
“Io mi scatenò e gioco all’aperto”.*

In Figura 3.3 è stata proiettata la variabile “io gioco e mi scatenò all’aperto”, si nota che la modalità “mai” è l’unica a fornire informazione perché le altre sono vicine all’origine degli assi, anche in questo caso chi non gioca all’aperto ha delle abitudini “sedentarie” legate ad un elevato uso degli apparecchi da intrattenimento e con un benessere familiare elevato.

Illustrativa

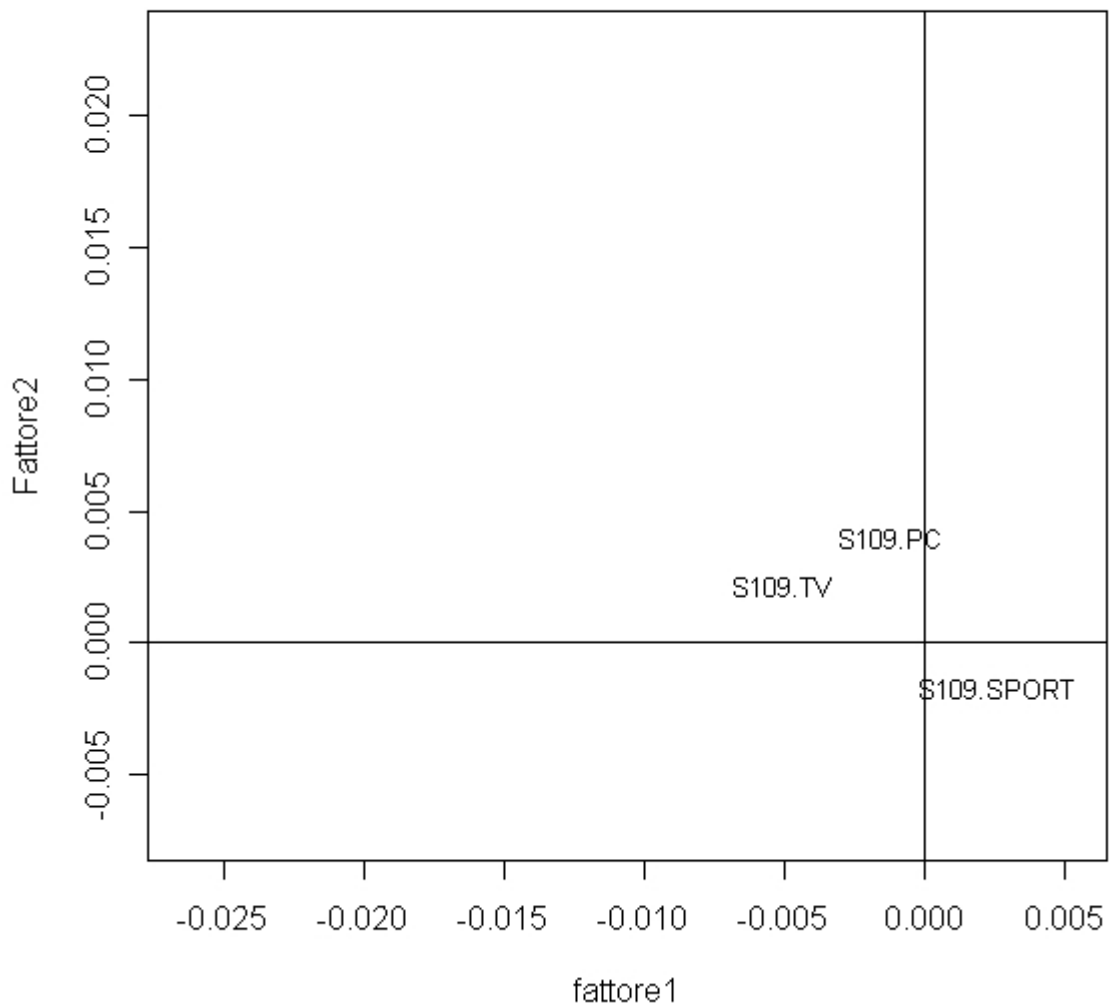


Figura 3.4: Proiezione sugli assi fattoriali della variabile illustrativa

“Per me sarebbe veramente difficile rinunciare a”.

In Figura 3.4 è stata proiettata la variabile “per me sarebbe veramente difficile rinunciare a”, si nota che chi risponde “sport” sono gli studenti più attivi cioè coloro che sono iscritti ad una società sportiva e fanno un basso uso degli apparati di intrattenimento. Coloro che rispondono “pc” e “tv” sono invece proprio coloro che ne fanno un largo uso e sono più “pigri”.

3.3 ANALISI CLUSTER

Lo scopo dell'analisi cluster è di riuscire a raggruppare n unità statistiche disponibili in k gruppi. I gruppi vengono generati in modo da massimizzare l'omogeneità interna e le differenze tra gli stessi. I metodi utilizzati per effettuare il raggruppamento si dividono in gerarchici e non gerarchici (Azzalini, A., Scarpa, B. (2004), "Analisi dei dati e Data Mining", Springer-Verlag, Milano).

L'analisi di seguito condotta fa uso di un metodo gerarchico agglomerativo di Ward la cui procedura di aggregazione si basa sulla minimizzazione della distanza tra centroidi, inoltre la funzione di distanza utilizzata è quella euclidea. In Figura 3.5 viene riportato il dendrogramma risultante dalla clusterizzazione degli indici BMI, FAS e MC.

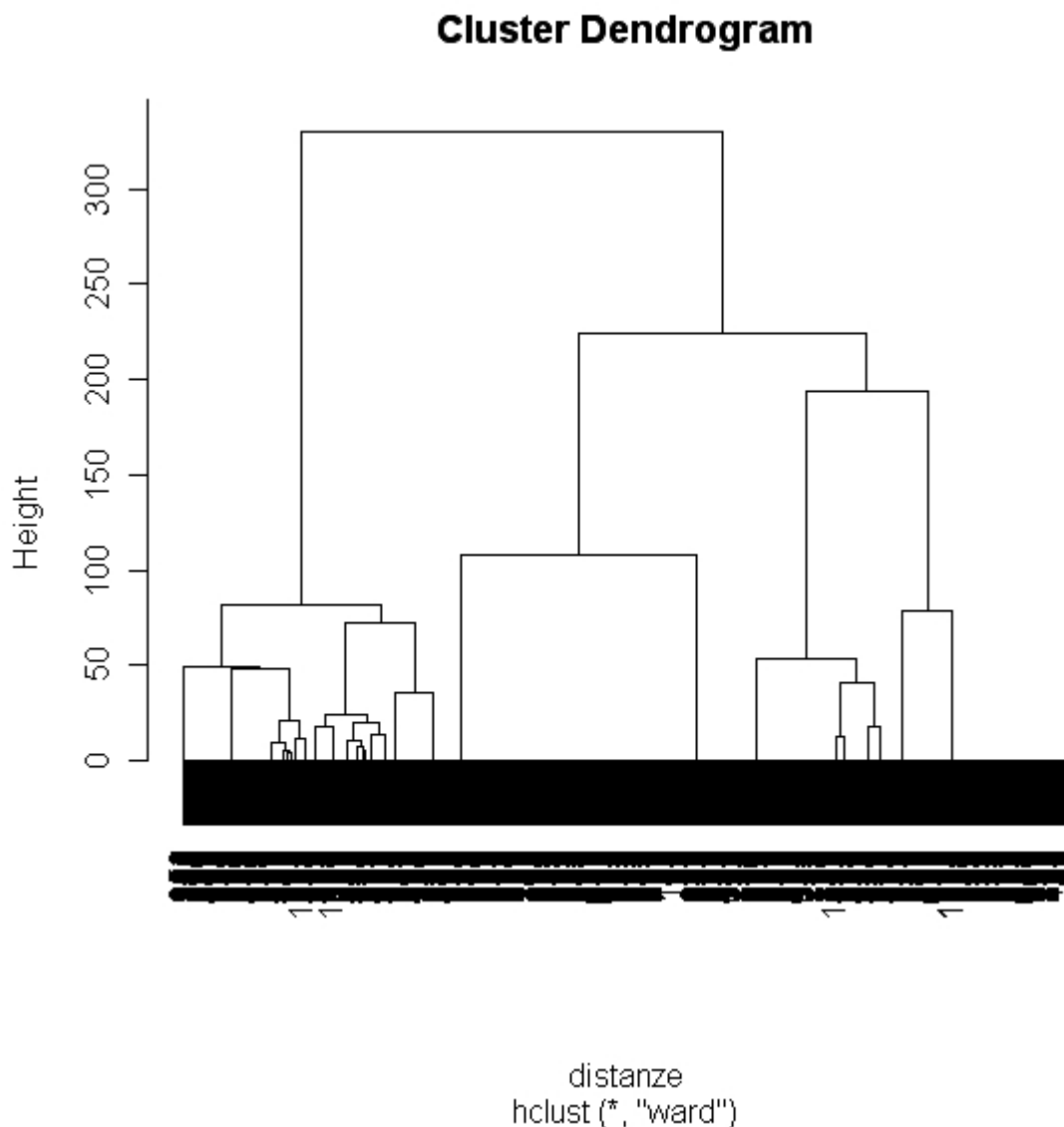


Figura 3.5: Albero di raggruppamento per gli indici.

Dal dendrogramma si nota che è possibile ottenere quattro sottogruppi, probabilmente legati tra loro da caratteristiche comuni. Il taglio viene effettuato tenendo conto della distanza indicata dalle aste verticali del dendrogramma e indicanti una dissimilarità marcata tra due gruppi.

La numerosità dei gruppi così formati è riportata in tabella 3.1.

Gruppi	1	2	3	4
Frequenze	335 (0,33%)	189 (0,19%)	314 (0,31%)	167 (0,17%)

Tabella 3.1: Frequenza delle distribuzioni dei gruppi.

Per capire le caratteristiche dei gruppi formati vengono riportati in Figura 3.7 i boxplot degli indici BMI, FAS e MC in funzione dei nuovi gruppi generati con l'analisi cluster.

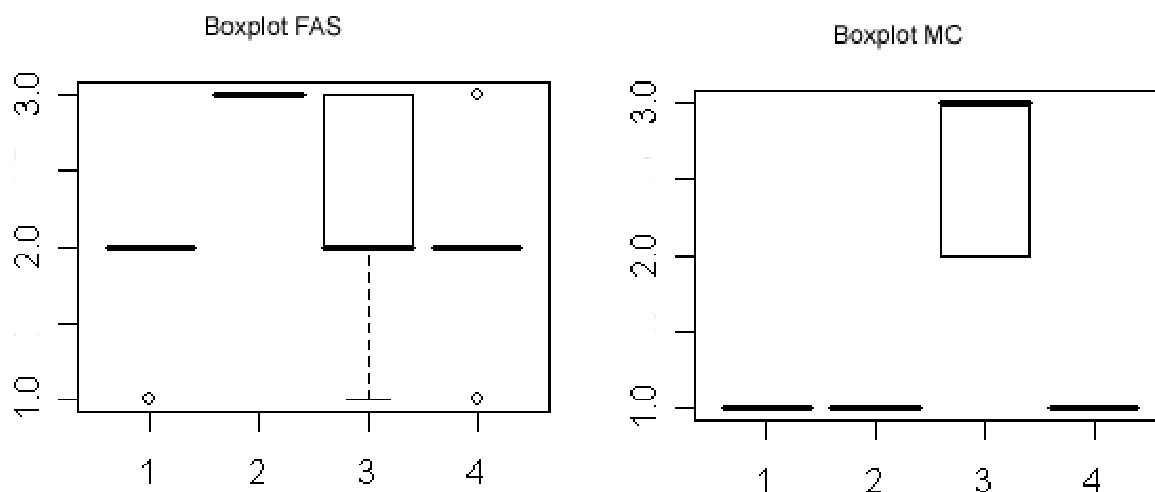


Figura 3.6: Boxplot indice FAS e MC nei gruppi.

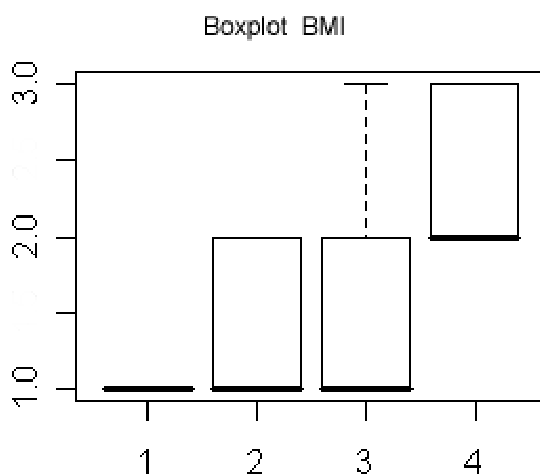


Figura 3.7: Boxplot indice BMI nei gruppi.

Dallo studio dei boxplot e delle Tabelle 3.2 e 3.3 si nota che i gruppi formati con l'analisi cluster possono essere così riassunti: il primo gruppo è composto da adolescenti con un peso forma normale, con un benessere familiare medio e che fanno scarso uso di apparecchi da intrattenimento. Il secondo gruppo è composto da adolescenti con un peso forma normale, benestanti e che fanno scarso uso di apparati di intrattenimento. Il terzo gruppo è in qualche modo il più omogeneo per gli indici BMI e FAS, ed è caratterizzato da quegli adolescenti che fanno maggior uso degli apparati di intrattenimento. L'ultimo gruppo è composto da quegli adolescenti in soprappeso e obesi.

Riassumendo ulteriormente si può indicare ciascun gruppo in base ad una caratteristica principale che li differenzia dagli altri: il primo gruppo è formato dagli adolescenti sportivi o più attivi fisicamente, il secondo dai più agiati, il terzo dai pigri o coloro che passano molte ore davanti un monitor, il quarto gruppo è formato da coloro che hanno problemi di peso.

	Gruppi			
	1	2	3	4
FAS Basso	68 (0,068%)	0	50 (0,05%)	40 (0,04%)
FAS Medio	267 (0,27%)	0	152 (0,15%)	119 (0,12%)
FAS Alto	0	189 (0,19%)	112 (0,11%)	8 (0,008%)

	Gruppi			
	1	2	3	4
MC Basso	335 (0,33%)	189 (0,19%)	0	167 (0,17%)
MC Medio	0	0	155 (0,15%)	0
MC Alto	0	0	159 (0,16%)	0

Tabella 3.2: Frequenza delle distribuzioni tra i gruppi dell'indice FAS e MC.

	Gruppi			
	1	2	3	4
BMI Normale	335 (0,33%)	133 (0,13%)	197 (0,20%)	0
BMI Soprappeso	0	56 (0,056%)	81 (0,081%)	119 (0,12%)
BMI Obeso	0	0	36 (0,036%)	48 (0,048%)

Tabella 3.3: Frequenza delle distribuzioni tra i gruppi dell'indice BMI.

CONCLUSIONI

Lo studio e l'analisi effettuati sui dati rilevati dal questionario sono finalizzati a dare una valutazione dello stile di vita degli adolescenti. I numerosi dati in esame e il numero di variabili associate hanno consigliato l'utilizzo di indici di sintesi, che riassumono aspetti legati alle abitudini, alla fisicità e al benessere familiare di ogni adolescente. Le variabili analizzate nel capitolo 2 danno un quadro completo e particolareggiato di moltissimi aspetti della vita quotidiana degli adolescenti, mostrandone le abitudini e le passioni. Proprio le abitudini nel corso dell'analisi sono diventate in qualche modo una chiave di lettura fondamentale per rispondere a quesiti quali il peso eccessivo. L'andare a scuola, i tipi di cibi consumati ogni giorno, lo sport praticato, le ore passate davanti ad un monitor, risultano i fattori determinanti su cui si è focalizzato lo studio. I maschi in particolare sono coloro che, nonostante una pratica sportiva maggiore rispetto alle femmine, hanno una percentuale di casi di sovrappeso maggiore. Dallo studio delle correlazioni tra i vari indici BMI, FAS e MC emerge una loro incorrelazione e ciò risulta in controtendenza rispetto a quello che ci si aspetterebbe, in particolare tra il BMI e le ore passate davanti alla tv (MC). Si è visto però che coloro che utilizzano maggiormente gli apparati di intrattenimento sono i maschi e in tal caso la correlazione tra gli indici risulta significativa. Lo studio delle corrispondenze multiple su tali indici ha ribadito che le abitudini sono uno dei fattori soggiacenti alla struttura dei dati.

Con l'analisi cluster successivamente si è provveduto a suddividere in gruppi il campione degli adolescenti, i gruppi così formati si possono distinguere tra sportivi, ricchi, sedentari e sovrappeso. La natura stessa del questionario e la sua complessità, per il numero di informazioni raccolte, ne fanno un vero e proprio tesoro, i cui dati possono venire ulteriormente utilizzati per provare a estrapolare nuove informazioni e nuove ipotesi di lavoro. In tal senso risulterebbe interessante poter confrontare il lavoro di analisi svolto sugli adolescenti italiani con un campione di adolescenti europei per capire se alcune caratteristiche dello stile di vita, sia fisiche che comportamentali, sono legate solo da abitudini in qualche modo soggettive o se, in un contesto sociale differente, tali caratteristiche cambiano radicalmente.

APPENDICE

Questionario

Per favore rispondi alle domande

Devi valutare se le frasi sono per te: *non giusta*, *poco giusta*, *abbastanza giusta*, *del tutto giusta*. Puoi fare solamente una crocetta per ogni domanda rimanendo dentro il cerchietto.

Un esempio su Pietro ed Anna:

Per Pietro è così: Pietro beve molto volentieri Coca Cola e perciò fa una croce alla voce "del tutto giusta". Se invece non mangia tanto volentieri gli spinaci fa una croce nella casella "poco giusta".

PIETRO

		non giusta	poco giusta	abba- stanza giusta	del tutto giusta
1. lo bevo molto volentieri Coca Cola	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
2. lo mangio molto volentieri gli spinaci	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Per Anna è così: Anna non può sopportare la Coca Cola e perciò fa una croce sulla casella "non giusta". Gli spinaci non sono il suo piatto preferito, ma lei prova a mangiarli molto volentieri e perciò fa una croce alla voce "abbastanza giusta".

ANNA

		non giusta	poco giusta	abba- stanza giusta	del tutto giusta
3. lo bevo molto volentieri Coca Cola	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. lo mangio molto volentieri gli spinaci	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Codice

--	--	--	--

Adesso vorremmo sapere da te alcune semplici cose

1. Io sono un

ragazzo



ragazza



2. Sono nato:

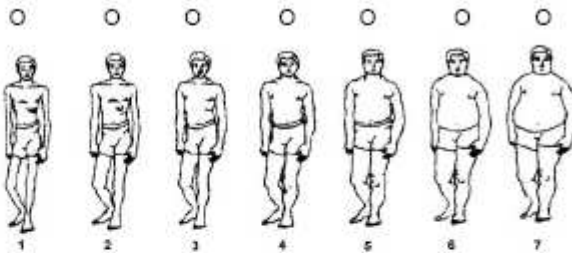
nel mese dell'anno

3. La mia statura (in
metri e centimetri) _____

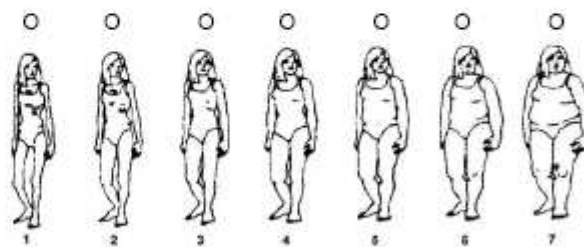
Il mio peso corporeo
(in chilogrammi) _____

Vorremmo sapere che aspetto hanno i tuoi genitori

4. Mio padre ha
all'incirca
questo aspetto



5. Mia madre ha
all'incirca
questo aspetto





Adesso abbiamo qualche domanda sullo sport

6. Praticare sport per me è ...

... non importante									... molto importante
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1	2	3	4	5	6	7	8	9	10

7. Sei iscritto ad una società sportiva Sì Non più No

Alla prossima domanda possono rispondere solo i ragazzi e le ragazze che sono iscritti ad una società sportiva e che hanno risposto all'ultima domanda con "Sì". Tutti gli altri possono andare alla pagina seguente alla domanda n°15.

In questi giorni della settimana io pratico nella società questa disciplina per quante ore
8. Lunedì	↳ _____	↳ circa _____ ore
9. Martedì	↳ _____	↳ circa _____ ore
10. Mercoledì	↳ _____	↳ circa _____ ore
11. Giovedì	↳ _____	↳ circa _____ ore
12. Venerdì	↳ _____	↳ circa _____ ore
13. Sabato	↳ _____	↳ circa _____ ore
14. Domenica	↳ _____	↳ circa _____ ore

Alcuni genitori praticano sport, altri no, com'è la situazione nella tua famiglia?

	Tutti i giorni	Più volte alla settimana	Una sola volta alla settimana	Raramente	Mai
15. Mio padre pratica Sport.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Mia madre pratica Sport.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Abbiamo ancora una domanda: con quale frequenza ti "scateni" e giochi all'aperto?

	Tutti i giorni	Più volte alla settimana	Una sola volta alla settimana	Raramente	Mai
17. Io mi "scateni" e gioco all'aperto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Vorremmo conoscere alcune semplici cose sui tuoi genitori e sulla tua casa

	No	SI			
18. A casa hai una stanza solo per te?	<input type="radio"/>	<input type="radio"/>			
	Mai	Una volta	Due volte	Più di due volte	
19. Quante volte sei andato in vacanza negli ultimi 12 mesi?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

	No	Si	Due o più
20. La tua famiglia possiede un'auto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Nessuno	Uno	Due	Più di due
21. Quanti computer possiede la tua famiglia	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Da quale stato provengono i tuoi genitori?

	Italia	Altro stato	
22. Mia madre proviene dall' ...	<input type="radio"/>	<input type="radio"/>	e precisamente <input type="text"/>
23. Mio padre proviene dall' ...	<input type="radio"/>	<input type="radio"/>	e precisamente <input type="text"/>



Quale giudizio avevi nell'ultima pagella?

	Ottimo	Distinto	Buono	Sufficiente	Non sufficiente
... Matematica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... Italiano	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... Educazione Motoria	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Licenza elementare	Diploma scuola media	Diploma scuola superiore	Laurea
27. Qual è il titolo di studio di tuo padre o di tua madre?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Come vai a scuola la mattina e di quanto tempo hai bisogno per il tragitto casa-scuola?

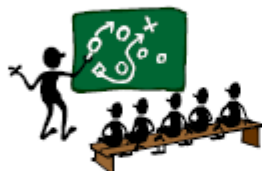


28. A piedi volte alla settimana minuti per volta
29. Con la bici volte alla settimana minuti per volta
30. Con il bus volte alla settimana minuti per volta
31. A piedi e con il bus volte alla settimana minuti per volta
32. Con l'auto volte alla settimana minuti per volta

Anche se qualche volta è difficile, tenta di giudicarti in base alle seguenti affermazioni

	Per niente	Poco	Abba- stanza	Molto
33. Mi piace il mio corpo così com'è	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34. Sono molto bravo nello sport	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35. Spesso dimentico quello che ho imparato a scuola	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36. In confronto agli altri mi piaccio	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Per niente	Poco	Abba- stanza	Molto
37. Ho buoni motivi per essere orgoglioso di me stesso	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38. Spesso sono triste senza motivo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39. Cambierei volentieri il mio aspetto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40. Facendo sport non mi accorgo di come passa il tempo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41. Faccio i compiti scolastici molto rapidamente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
42. Complessivamente sono molto soddisfatto di me stesso	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43. Nello sport imparo più velocemente degli altri della mia età	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44. Cerco di evitare le attività sportive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
45. Sono molto contento del mio corpo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
46. Vado molto bene a scuola	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
47. Non ho una grande stima di me stesso	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
48. Rido solo raramente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



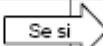
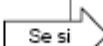
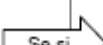
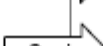

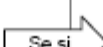

	Per niente	Poco	Abba- stanza	Molto
49. Molti ragazzi hanno un aspetto migliore del mio	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
50. Con lo sport dimentico tutto intorno a me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
51. Imparo molto velocemente i nuovi esercizi sportivi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Per niente	Poco	Abba- stanza	Molto
52. Credo di essere intelligente tanto quanto i ragazzi della mia età	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
53. Sono contento di me stesso	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
54. Durante il mio tempo libero preferirei fare sport più di ogni altra cosa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
55. Mi sento veramente in forma	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
56. Spesso mi siedo e non vorrei far nulla	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
57. Credo di essere bravo nello sport tanto quanto gli altri della mia età	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
58. Non sono bravo a scuola	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
59. Non mi piace il mio corpo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
60. Quando gli altri si divertono io non posso ridere con loro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
61. Mi sento impacciato nel mio corpo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
62. Non sono bravo nello sport	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
63. Ho un aspetto migliore della maggior parte dei miei amici	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
64. Durante le lezioni mi è difficile trovare la risposta alle domande	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
65. Non c'è niente che mi piaccia davvero	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Adesso vorremo sapere qualche cosa di più sulla tua alimentazione

Quale dei seguenti pasti assumi regolarmente (quasi ogni giorno) durante la settimana (da Lunedì a Venerdì)? Con chi consumi i pasti?

Nella settimana (da Lunedì a Venerdì)	regolarmente?		Se sì con chi per lo più?				
	no	si	Genitori, fratelli, nonna, nonno	Solo fratelli	Da solo	Amici	
66. Prima colazione	<input type="radio"/>	<input type="radio"/>	Se sì 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
67. Merenda del mattino	<input type="radio"/>	<input type="radio"/>	Se sì 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
68. Pranzo	<input type="radio"/>	<input type="radio"/>	Se sì 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
69. Spuntino pomeridiano (torta dolciumi)	<input type="radio"/>	<input type="radio"/>	Se sì 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
70. Spuntino pomeridiano (frutta)	<input type="radio"/>	<input type="radio"/>	Se sì 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
71. Cena	<input type="radio"/>	<input type="radio"/>	Se sì 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
72. Spuntino di tarda sera	<input type="radio"/>	<input type="radio"/>	Se sì 	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>




Negli ultimi sei mesi con quale frequenza hai mangiato i seguenti cibi?

	Tutti i giorni	Più volte alla settimana	Una sola volta alla settimana	Meno o mai
73. Pane fresco, pane ai cereali, pane nero	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
74. Pane conservato, pane da toast, cracker	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>


	Tutti i giorni	Più volte alla settimana	Una sola volta alla settimana	Meno o mai
75. Marmellata, miele, Nutella	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
76. Corn flakes, corn pops	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
77. Muesli	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
78. Formaggio, ricotta , yogurt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
79. Carne	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
80. Affettati, salame, mortadella, prosciutto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
81. Frutta fresca	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
82. Verdura, insalata	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
83. Pasta, spaghetti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
84. Patate, riso	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
85. Dolciumi, torte, biscotti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
86. Pasticcini salati	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
87. Limonata, aranciata, coca -cola	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
88. Patate fritte	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
89. Pesce, bocconcini di pesce	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
90. Pizza	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
91. Wurstel, salsicce	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
92. Hamburger, Big-Mac	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
93. Acqua, acqua minerale	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>




**Per favore rispondi ancora a qualche domanda su televisione,
computer o videogame**

Quanto guardi ... 

	Tutti i giorni	Più volte alla settimana	Una volta alla settimana	Più volte al mese	Una volta al mese o più raramente	Mai
94. ... la TV da solo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
95. ... la TV con i tuoi genitori	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
96. ... la TV con i tuoi amici	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Quanto usi ... 

	Tutti i giorni	Più volte alla settimana	Una volta alla settimana	Più volte al mese	Una volta al mese o più raramente	Mai
97. ... il computer da solo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
98. ... il computer con i tuoi genitori	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
99. ... il computer con i tuoi amici	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Quanto usi ... 

	Tutti i giorni	Più volte alla settimana	Una volta alla settimana	Più volte al mese Più volte al mese	Una volta al mese o più raramente	Mai
100. ... i videogame da solo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
101. ... i videogame con i tuoi genitori	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
102. ... i videogame con i tuoi amici	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Attualmente per quante ore al giorno guardi nel tuo tempo libero la TV? (compresi DVD o videocassette)



	Per niente	Circa 30 minuti al giorno	Circa 1 ora al giorno	Circa 2 ore al giorno	Circa 3 ore al giorno	Circa 4 ore al giorno	Circa 5 ore al giorno	Circa 6 ore al giorno o di più
103. ... in un giorno della settimana (da Lunedì a Venerdì) guardo la TV	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
104. ... nel fine settimana guardo la TV	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Per quante ore al giorno utilizzi abitualmente, nel tuo tempo libero, un computer (per giocare, scrivere mail, per fare qualcosa per la scuola, per chattare, per navigare in internet?)



	Per niente	Circa 30 minuti al giorno	Circa 1 ora al giorno	Circa 2 ore al giorno	Circa 3 ore al giorno	Circa 4 ore al giorno	Circa 5 ore al giorno	Circa 6 ore al giorno
105. ... in un giorno della settimana (Lunedì-Venerdì) utilizzo il computer per ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
106. ... nel fine settimana utilizzo il computer per ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Per quante ore al giorno giochi abitualmente con una console da gioco (Nintendo, Playstation, Xbox, Gameboy, ...)?

	Per niente	Circa 30 minuti al giorno	Circa 1 ora al giorno	Circa 2 ore al giorno	Circa 3 ore al giorno	Circa 4 ore al giorno	Circa 5 ore al giorno	Circa 6 ore al giorno
107. ... in un giorno della settimana (Lunedì-Venerdì) gioco ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
108. ...nel fine settimana gioco giornalmente ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	TV	Computer	Sport
109. Per me sarebbe veramente difficile rinunciare a ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Ce l'hai fatta !!!
Tantissime grazie per il tuo impegno.**



p.
h.
oss.

Istruzioni per la raccolta dati

La raccolta dei dati avviene nella classe, ogni studente è al suo posto normale; sono presenti 3 o 4 intervistatori per classe.

- Spiegazione standardizzata
- Misura di peso e altezza (se possibile fatte da una sola persona)
- Gli intervistatori sono disponibili per rispondere ad eventuali domande
- Alla fine: check dei questionari (non controllare le risposte per il loro contenuto, ma controllare se ci sono risposte mancanti)

Spiegazione standardizzata:

(Evitare artefatti degli intervistatori, non si deve improvvisare!)

- Noi siamo dell'Università di Padova, della SSIS del Veneto, dell'Università di Chieti, ...
- I nostri nomi sono:
- Siamo qui oggi perchè abbiamo alcune domande per voi.
- Non ci sono risposte giuste o sbagliate. Questo non è un test o una prova di verifica.
- Noi non sappiamo come voi conoscete questi argomenti e saperlo è molto importante per noi.
- Voi siete bambini, noi siamo adulti. Ci piacerebbe sapere come vi sentite, quanto sport fate, quanto guardate la televisione e cosa mangiate.
- Questa è come una vera intervista in cui voi siete le star o gli esperti che rispondono alle nostre domande.
- Vogliamo conoscere le vostre opinioni; per questo è importante rispondere alle domande con sincerità, scegliendo la risposta che corrisponde a quello che pensate.
- Ogni cosa sarà top-secret, nessuno sarà informato di come avete risposto, né i vostri insegnanti, né i vostri genitori o gli altri bambini della classe.

Distribuzione dei questionari

- Spiegare l'esempio di pagina 1.
- Spiegare bene come mettere le crocette.
- Da questo momento ciascuno compilerà il suo questionario da solo.

- Il tuo vicino di posto non può aiutarti, perchè non può sapere se “a te piacciono gli spinaci oppure no”!
- Non ci sono problemi se avete qualche domanda perchè non capite qualcosa. Alzate la mano. Qualcuno di noi verrà da voi. È per questo che noi siamo qui in (3-4) persone.
- Quando avete finito alzate la mano. Qualcuno di noi verrà a ritirare il vostro questionario (è a questo punto che gli intervistatori devono controllare se vi sono risposte mancanti!).

Alla fine:

- Vorremmo misurare la vostra altezza e il vostro peso.
 - Faremo questo in un angolo della classe, al riparo dallo sguardo degli altri
 - Per favore dovrete togliervi le scarpe e portare il questionario con voi.
 - Anche queste misure saranno top-secret, nessuno saprà nulla.
 - E perchè nessuno dei vostri compagni possa darvi fastidio non vi diremo i vostri valori.
- In questo modo nessuno potrà parlare degli altri perché le misure saranno sconosciute a tutti.

Compendio per gli aiuti standardizzati:

- Gli intervistatori avranno lo stesso numero di questionari (es. 5 intervistatori ciascuno con 6 questionari) e li distribuiranno alla classe.
 - Aspettare sino al silenzio.
 - Fare l'introduzione e dare le spiegazioni.
 - Se necessario, chiedere all'insegnante di restare nella classe, ma di *lasciare* gli alunni *soli* con le domande => questo permetterà un relazione paritaria tra intervistatori e intervistati.
 - Se i bambini sono troppo lenti, offrire loro un supporto individuale da parte di un intervistatore.
- La seguente procedura è raccomandata:
- leggi tu assieme al bambino la domanda, se necessario sottolinea le parti importanti della domanda attraverso l'intonazione, chiedi: “hai capito?”;
 - chiedi: “che cosa non hai capito?”;
 - chiedi: “hai capito adesso?”;
 - se una parola non fosse comprensibile cerca una parola sostitutiva, (per favore annota quale parola hai usato o usa una parola o una frase preparata precedentemente standardizzazione).
 - Se i bambini hanno problemi a valutare sé stessi, le seguenti procedure sono raccomandate:
 - prima di tutto accerta il segno della risposta (affermativo o negativo);

- poi accerta il valore della risposta, ad esempio se è un po' negativa oppure assolutamente negativa;
- se ci sono ancora problemi con la valutazione, l'alunno dovrebbe scegliere una valutazione mediana (anche se con una valutazione estrema, affermativa o negativa, non dovrebbero esserci particolari problemi).
- Se necessario, singoli alunni potranno avere più tempo per compilare il questionario, dopo la lezione e durante la pausa.

Grafici curve di crescita

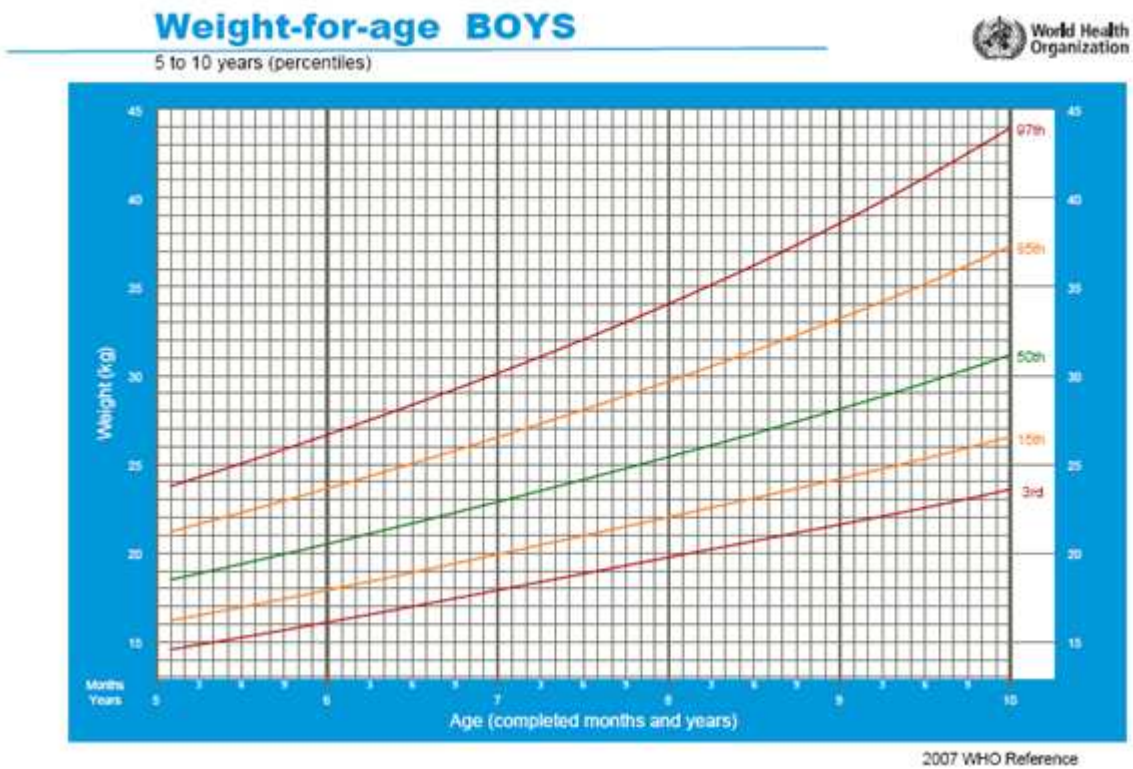
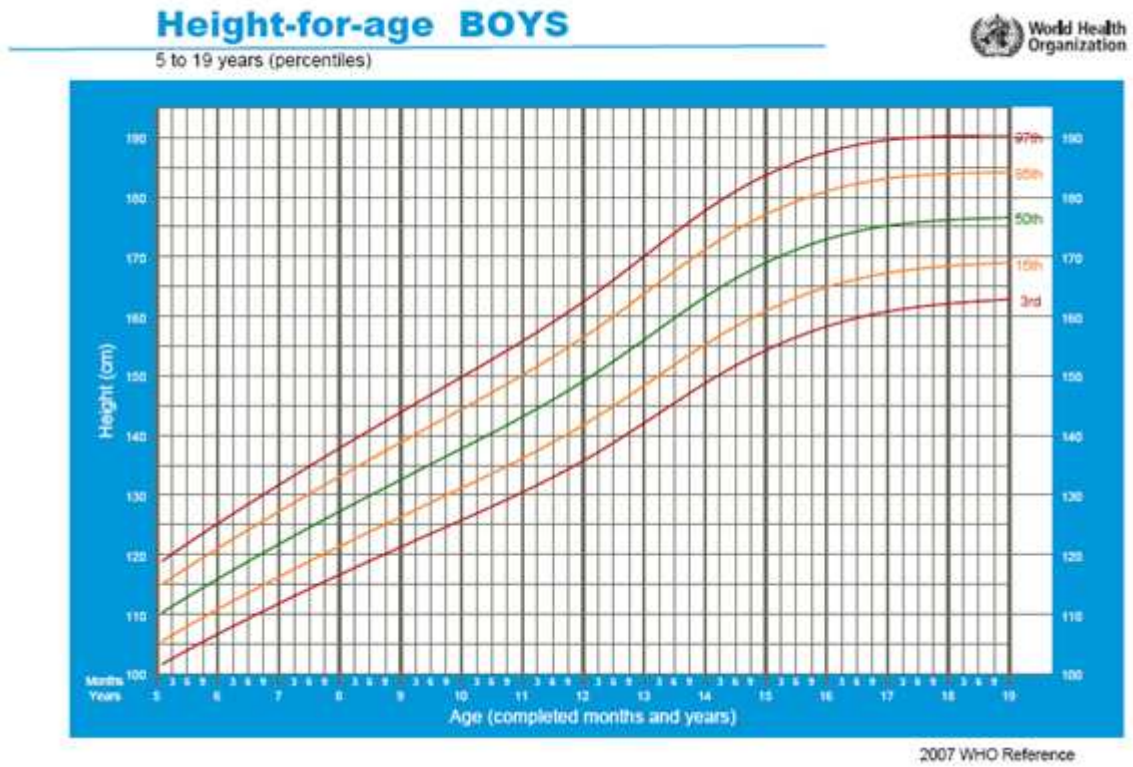
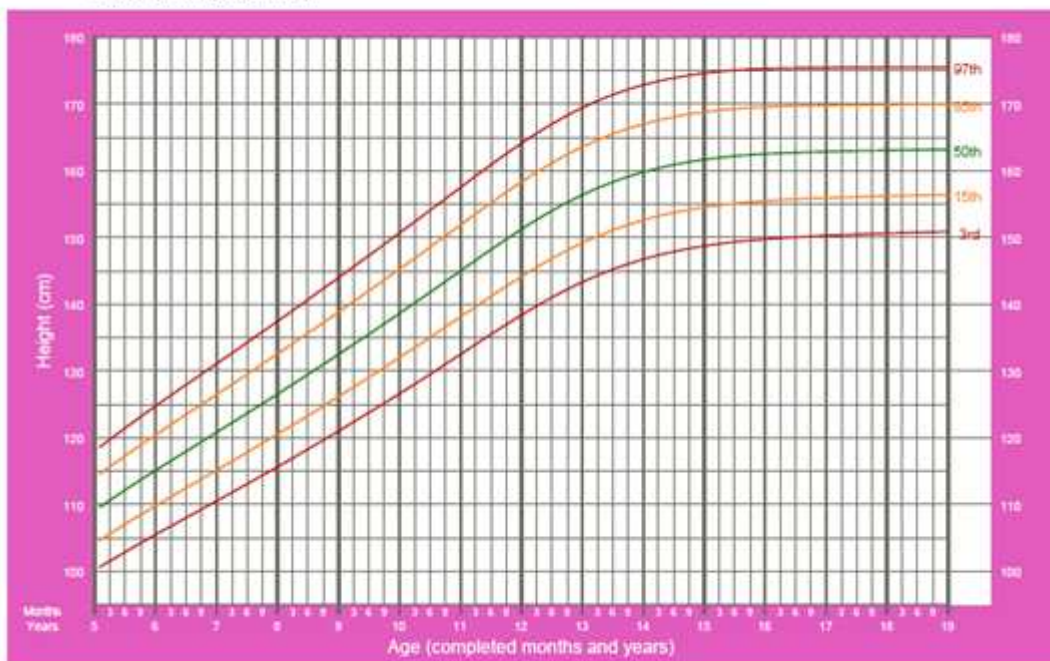


Figura 1: Percentili delle curve di crescita di altezza e peso per sesso maschile.

Height-for-age GIRLS

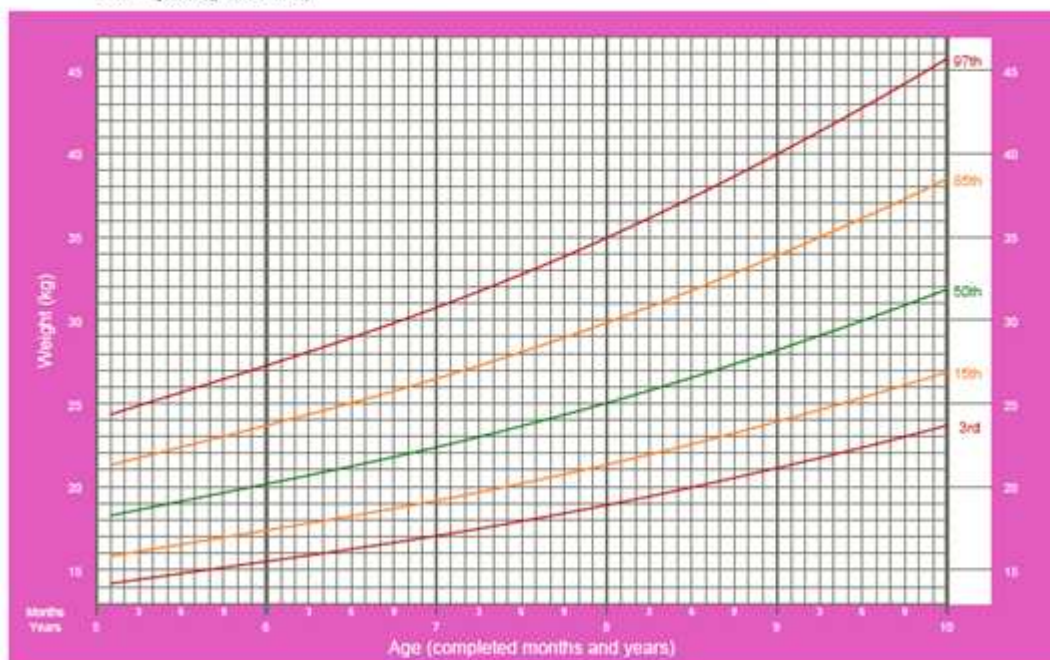
5 to 19 years (percentiles)



2007 WHO Reference

Weight-for-age GIRLS

5 to 10 years (percentiles)

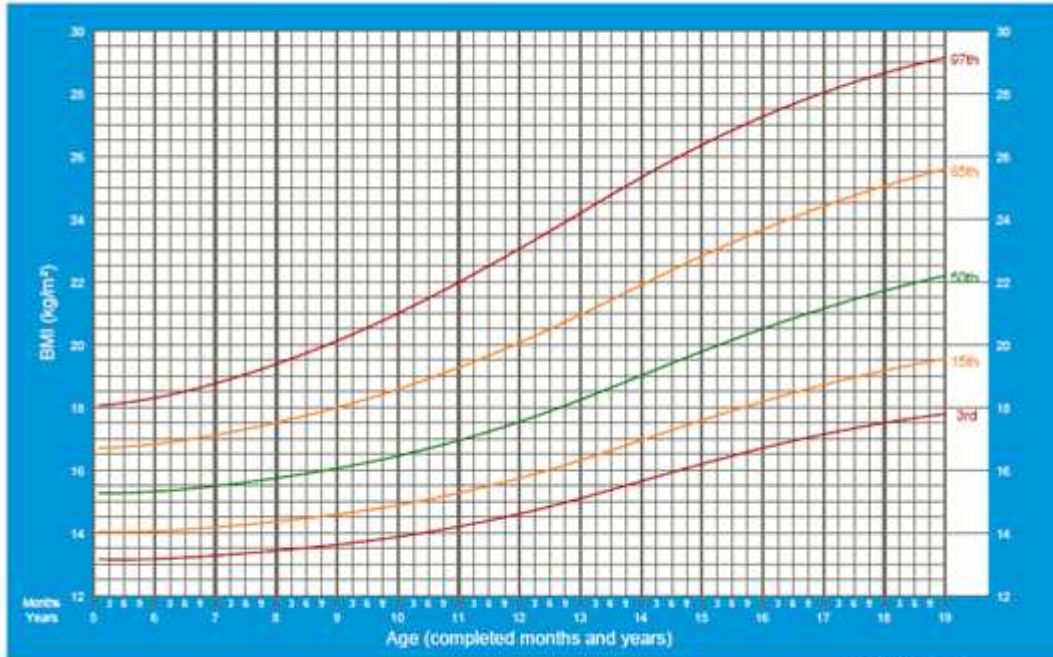


2007 WHO Reference

Figura 2: Percentili delle curve di crescita di altezza e peso per sesso femminile.

BMI-for-age BOYS

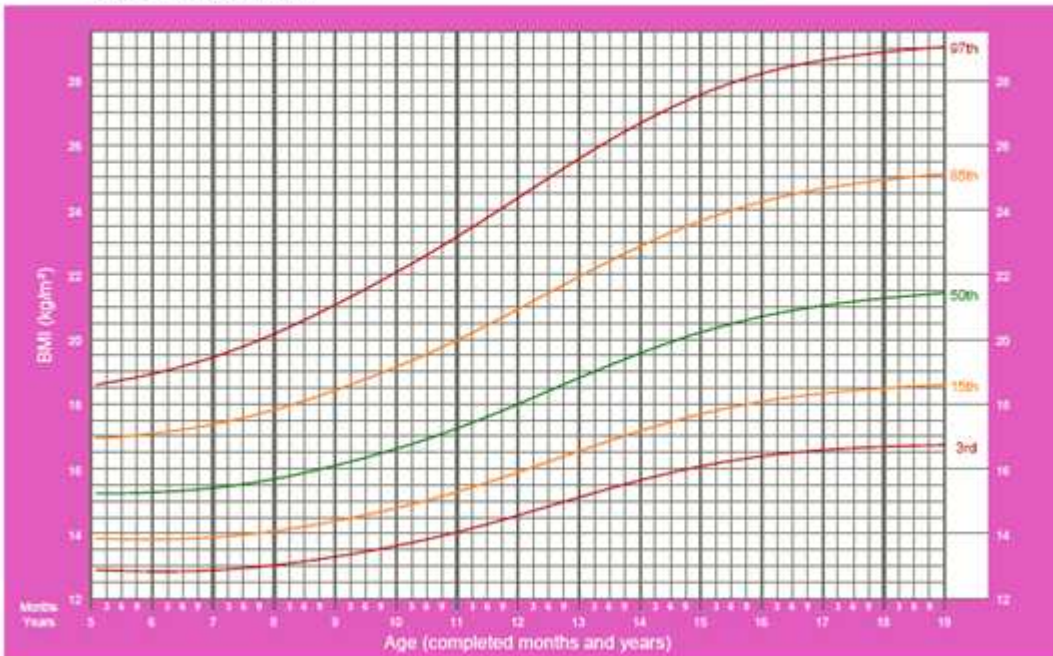
5 to 19 years (percentiles)



2007 WHO Reference

BMI-for-age GIRLS

5 to 19 years (percentiles)



2007 WHO Reference

Figura 3: Percentili delle curve di crescita del body mass index per maschi e femmine.

Comandi per analisi statistiche

Esempi comandi del software R utilizzati :

Creare un dataset da file:

```
dataset <- read.table ("percorso\\file.csv", header=T, sep=";")
```

Grafici a scatola e istogrammi:

```
>boxplot (dataset$variabile)
```

```
>boxplot (dataset$variabile ~ dataset$genere)
```

```
> barplot (table (dataset$variabile))
```

Fattorizzazione variabili qualitative:

```
>dataset$variabile <- factor (dataset$variabile)
```

Creazione di un dataframe:

```
>data <- data.frame (lista.variabili, labels=c("lista.etichette"))
```

Test di Kolmogorov-Smirnov :

```
>ks.test (dataset$variabile1 , dataset$variabile 2)
```

Correlazione tra variabili, test di correlazione e grafico:

```
>cor (dataset$variabile1, dataset$variabile2, method=c("spearman", "pearson"))
```

```
>cor.test (dataset$variabile1, dataset$variabile2, method=c("spearman", "pearson"))
```

```
>library (ellipse)
```

```
>plotcorr (cor (data, method=c("spearman", "pearson")))
```

Test di Wilcoxon:

```
>Wilcox.test (dataset$variabile1, dataset$variabile2)
```

```
> Wilcox.test (dataset$variabile ~ dataset$genere)
```

Analisi delle corrispondenze multiple:

```
>library (MASS)
```

```
>acomu <- mca (var.attive, nf=x, abbrev=FALSE)
```

```
>plot (acomu$cs, type="n")
>text (acomu$cs, rownames(acomu$cs), cex=0.8)
>lines (acomu$cs[riga:riga, fattore1] , [riga:riga, fattore2])
>plot (acomu$cs, type="n")
>illustrativa <- predict (acomu, newdata=as.data.frame(variabile), type="factor")
>text(illustrativa, rownames(illustrativa), cex=0.8)
```

Analisi Cluster:

```
>distanze <- dist (dati, method="euclidean")
>cluster <- hclust (distanze, method="ward")
>gruppi <- cutree (cluster, k=n)
```

Per richiamare la documentazione sui comandi in R digitare dal prompt dei comandi: *help()*.

Software

R - Project software for Statistical Computing.

SPSS – Analytical Software specializing in data mining, customer relationship management, business intelligent and data analysis.

Open Office – Multiplatform office suite open source.

RIFERIMENTI BIBLIOGRAFICI

Azzalini, A., Scarpa, B. (2004), “Analisi dei dati e Data Mining”, Springer-Verlag, Milano.

Corbetta, M. (1999), “Metodologia e tecniche della ricerca sociale”, Il Mulino, Bologna.

Ercolani, A., Areni, A., Leoni, L. (2007), “Statistica per la psicologia”, Il Mulino, Bologna.

Fabbris, L. (1990), “Analisi esplorative di dati multidimensionali”, Cleup Editore.

Masarotto, G., Iacus, S.M. (2003), “Laboratorio di statistica con R”, McGraw-Hill, Milano.

Pace, L., Salvan, A. (2001), “Introduzione alla statistica”, Cedam, Padova.

www.hbsc.org, sito Health Behaviour in School - aged Children.

www.ministerosalute.it, sito del ministero della salute.

www.who.int, sito World Health Organization.