

Contents

1	Il progetto	5
1.1	Le tappe del progetto	6
1.2	Definizione del contenuto delle prove standardizzate	6
1.2.1	Matematica	6
1.2.2	Italiano	7
1.3	Metodologia di analisi	8
1.4	Il pre-test	8
1.5	La prova definitiva somministrata	9
2	Caratteristiche socio demografiche	11
2.1	Nazionalità	11
2.2	Genere	11
2.3	Anno di nascita	11
2.4	Lingua Madre	14
3	Analisi dei risultati della prova di Matematica	16
3.1	Informazioni generali	16
3.2	Correlazioni tra le dimensioni valutate con la prova	18
3.3	Il contesto e la scuola	19
3.3.0.1	Analisi statistiche	19
3.4	Il grado di urbanizzazione del comune	20
3.4.0.1	Analisi statistiche	21
3.5	La classe	23
3.5.1	Dimensione della classe o numero di allievi	23
3.5.2	La pluriclasse	24
3.5.2.1	Analisi statistiche	26
3.6	L'allievo	26
3.6.1	La nazionalità	26
3.6.1.1	Analisi statistiche	26
3.6.2	La lingua madre	28

3.6.2.1	Analisi statistiche	29
3.6.3	Il Genere	30
3.6.3.1	Analisi statistiche	30
3.6.4	Origine Sociale	31
3.6.4.1	Analisi statistiche	33
3.7	Relazione tra risultati alle prove standardizzate e nota scolastica di Matematica	34
4	Analisi dei risultati della prova di Italiano	39
4.1	Informazioni generali	39
4.2	Correlazioni tra le dimensioni valutate con la prova	39
4.3	Il contesto e la scuola	40
4.3.0.1	Analisi statistiche	40
4.4	Il grado di urbanizzazione del comune	42
4.4.1	Analisi statistiche	43
4.5	La classe	44
4.5.1	Dimensione della classe o numero di allievi	44
4.5.2	La pluriclasse	44
4.5.2.1	Analisi statistiche	44
4.6	L'allievo	46
4.6.1	La nazionalità	46
4.6.1.1	Analisi statistiche	46
4.6.2	La Lingua madre	47
4.6.2.1	Analisi statistiche	47
4.6.3	Il Genere	49
4.6.3.1	Analisi statistiche	49
4.6.4	Origine Sociale	50
4.6.4.1	Analisi statistiche	51
4.7	Relazione tra risultati alle prove standardizzate e nota scolastica in Italiano	51

5	Analisi comparative - Matematica ed Italiano	55
5.0.1	Commenti generali	55
5.0.2	<i>Clusters</i> o di risultati	55
5.0.3	Relazione tra i risultati alla prova di italiano e matematica	55
5.0.4	Relazione tra i punteggi di matematica ed italiano in funzione della lingua madre	57
A	Appendice: Informazioni metodologiche	59
A.1	Test <i>t</i> di Student	59
A.1.1	Lettura del grafico	59
A.2	ANOVA (analisi di varianza)	60
A.2.1	Lettura del grafico	62
B	Appendice Matematica	64
B.1	Analisi di varianza (ANOVA) - Urbanizzazione	64
B.2	Test <i>t</i> di Student - La classe	64
B.3	Analisi di varianza (ANOVA) - Nazionalità	67
B.4	Test <i>t</i> di Student - La lingua madre	70
B.5	Test <i>t</i> di Student - Il genere	74
B.6	Analisi di varianza (ANOVA) - Origine sociale	77
C	Appendice Italiano	82
C.1	Analisi di varianza (ANOVA) - Urbanizzazione	82
C.2	Analisi di varianza (ANOVA) - Nazionalità	84
C.3	Test <i>t</i> di Student - La lingua madre	87
C.4	Test <i>t</i> di Student - Il genere	90
C.5	Analisi di varianza (ANOVA) - Origine sociale	93

title: | Prove standardizzate cantonali: Matematica e Italiano in V SE

```
author: - Emanuele Meier - Alberto Crescentini - Marina Pettignano date:
"2022-06-01" geometry: margin=3cm lang: it site: bookdown::bookdown_site
output: bookdown::gitbook documentclass: book bibliography: [refer-
ences.bibtex, packages.bib] biblio-style: apalike link-citations: yes github-repo:
rstudio/bookdown-demo
```

```
# fig_caption: yes
```

```
# includes:
```

```
# in_header: my_header.tex #description: "Esempio di report informatizzato
che si potrebbe pensare di presentare al cantone." —
```

1 Il progetto

Il presente rapporto fa seguito a tre rapporti analoghi prodotti negli anni. Il primo sulla base dei risultati alle prove di matematica in IV elementare [?], il secondo alle prove di matematica in V elementare [?] e il terzo alle prove di italiano in III elementare [?].

Per quanto riguarda le informazioni utilizzate in questo rapporto specifico il progetto che le ha prodotte nel suo insieme si è svolto, nelle sue varie parti, negli anni scolastici dal 2019 al 2021.

L'insieme delle valutazioni svolte tramite le prove standardizzate che sono state costruite nell'ambito di questo progetto ha il duplice obiettivo di fornire delle informazioni di monitoraggio del sistema educativo (oggetto specifico di questo rapporto) e di fornire ai docenti, direttori e ispettori delle informazioni di dettaglio relative all'andamento delle specifiche classi, rispettivamente sedi e ispettorati.

Lo svolgimento del progetto e la struttura di massima del rapporto riprendono quelli degli altri rapporti. La presentazione dei contenuti e le dimensioni teoriche sono molto simili e si rimanda per gli approfondimenti del caso ai documenti già esistenti; alcune parti che si riteneva potessero giovare alla lettura ancorché già esistenti sono state messe in appendice.

I primi mesi del progetto sono stati impiegati per il consolidamento e sviluppo della rete di collaborazione. Questa è stata composta dal gruppo di esperti del territorio (in questo caso il termine non fa riferimento esclusivo agli esperti disciplinari della Scuola Media bensì a persone che fossero portatori di un sapere e di una conoscenza utile alla riflessione su questo oggetto), individuati grazie alla collaborazione con la Sezione delle scuole comunali (SESCO), con la Sezione insegnamento medio (SIM) e con il Dipartimento formazione e apprendimento (DFA).

1.1 Le tappe del progetto



1.2 Definizione del contenuto delle prove standardizzate

La prima fase del progetto, sino ad agosto 2019, è stata dedicata alla definizione dei contenuti da testare e all'identificazione delle persone che avrebbero potuto sviluppare gli item. La scelta dei settori da investigare è fondamentale in quanto sarebbe estremamente difficoltoso valutare contemporaneamente tutte le competenze presenti in una disciplina insegnata e farlo in modo esatto.

Nell'elaborazione della prova sono stati presi in considerazione, attingendoli dal Piano di studio della scuola dell'obbligo ticinese [?] gli ambiti di competenza che si è deciso di valutare per la matematica e per l'italiano. Nello specifico, per la matematica i 3 ambiti di competenza che si è deciso di valutare sono: Numeri e calcolo, Grandezze e misure e Geometria. Per l'italiano sono invece stati valutati i 2 ambiti di competenza: Leggere e Scrivere.

1.2.1 Matematica

Dell'ambito Numeri e calcolo si è misurato il processo cognitivo Matematizzare e modellizzare. Dell'ambito Grandezze e Misure sono state misurate i processi cognitivi Eseguire e applicare e Sapere e riconoscere. Infine, dell'ambito Geometria è stata misurata il processo cognitivo Sapere e riconoscere e i processi cognitivi Matematizzare e modellizzare e Comunicare e argomentare.

Le sei dimensioni testate sono quindi:

1. Numeri e calcolo – Matematizzare e modellizzare (NC MT)
2. Grandezze e misure – Eseguire e applicare (GM EA)

3. Grandezze e misure – Sapere e riconoscere (GM SRD)
4. Geometria – Sapere e riconoscere (GEO SRD)
5. Geometria - Matematizzare e modellizzare (GEO MM)
6. Geometria - Comunicare e argomentare (GEO CA)

1.2.2 Italiano

Di entrambi gli ambiti (i.e., Leggere e Scrivere) si è testato il processo “Attivare”. Dell’ambito di competenza Leggere è stato trattato il processo Realizzare, mentre dell’ambito di competenza Scrivere si è anche misurato il processo Rivedere.

Gli item presenti nelle prove sono stati suddivisi in due parti:

- La Parte base (definita dall’acronimo PB), la quale racchiude esercizi che toccano l’ambito di competenza Leggere, valutandone gli Aspetti linguistici (processo Attivare), le Inferenze (processo Realizzare) e il Lessico (processo Attivare)
- La parte aggiuntiva, la quale racchiude esercizi di tre singole dimensioni: Punteggiatura (P), Ortografia (O) e Metacognizione (M). Gli item di Punteggiatura toccano l’ambito di competenza Leggere, valutandone il processo Attivare. Gli item di Ortografia toccano l’ambito di competenza Scrivere, valutandone i processi Attivare e Rivedere. Gli item di Metacognizione toccano i processi testati nelle altre aree, lavorando sulla consapevolezza e la riflessione.

Successivamente, sono stati sviluppati gli item per testare ognuna delle parti stabilite. Nel costruire gli item si è dovuto tener conto di alcune indicazioni. In primo luogo ogni item doveva essere quanto più possibile mono dimensionale. Il lavoro in questa fase è stato svolto in stretta collaborazione con gli esperti di Didattica della Matematica operanti all’interno del DECS e del DFA. Per poter misurare la capacità di discriminazione dell’item (si parla di capacità discriminativa relativamente al fatto che l’item riceva risposte corrette dagli allievi più abili e scorrette da quelli meno abili) e anche la sua coerenza con la dimensione che si desiderava valutare; (è infatti necessario che ogni item sia attinente, in maniera completa o quantomeno prevalente, a una e una sola dimensione). Questa caratteristica rende questi item in sé differenti da quelli che normalmente sono utilizzati dai docenti durante la loro attività professionale. In secondo luogo, si è dovuto creare un numero di item sovrabbondante rispetto all’uso finale. Si è dovuto infatti prevedere che successivamente alla prova campione (fase di pre-test) sarebbe stato eliminato almeno il 30% degli item prodotti.

1.3 Metodologia di analisi

Il modello di analisi che è stato utilizzato per la valutazione e la selezione degli item (nella fase di pre-test) e per la definizione dei punteggi degli allievi è legato alla Item Response Theory (IRT) ed è costruito al fine di avere delle misure precise di costrutti ben definiti e quanto più possibili unitari. I modelli IRT sono una valida alternativa ai modelli di misura basati sulla *Classic Test Theory* (CTT). I modelli IRT corrispondono ad un approccio psicometrico che considera che la risposta di un individuo ad un particolare item sia influenzato dalle qualità dell'item (i.e., difficoltà¹) e dalle caratteristiche individuali (i.e., competenza individuale sul costrutto latente misurato). In altri termini, i modelli IRT postulano che la probabilità di risposta corretta ad un item dipenda congiuntamente dall'abilità o competenza dell'individuo e dalla difficoltà dell'item².

Nello specifico le analisi si sono svolte all'interno del quadro concettuale del modello di Rasch a un parametro (i.e., 1PL) che fornisce un modello di misura appropriato al contesto di questo progetto e che permette di trattare degli item dicotomici (i.e., items la cui risposta può essere giusta vs. sbagliata). Tutte le analisi presenti in questo rapporto sono state effettuate tramite R [?]. Le analisi IRT sono state realizzate tramite la libreria TAM [?] mentre le varie operazioni di gestione dei *database* sono state realizzate con l'ausilio del corpo di librerie di tidyverse [?].

1.4 Il pre-test

Una volta prodotti gli item e sottoposti a verifica di contenuto con l'assistenza degli esperti, si è proceduto alla preparazione della prova per il pre-test. Questa prova aveva lo scopo di valutare la pertinenza degli item e di individuare quelli più efficaci a misurare e a discriminare. Ordinandoli per difficoltà crescente si dovrebbe trovare un numero inizialmente molto elevato di allievi che risponderà correttamente e questo numero dovrebbe ridursi al crescere della difficoltà degli item stessi. Se un item ad esempio riceverà un numero di risposte corrette elevato ma solo dagli allievi meno abili, questo sarà scartato, parimenti saranno eliminati gli item non discriminanti, quelli cioè ai quali tutti o nessuno avranno risposto in modo corretto. Queste procedure hanno infatti lo scopo di costruire delle scale valide non in termini astratti ma all'interno delle popolazioni reali. Gli item costruiti sono infatti coerenti con i contenuti presentati nei programmi della scuola dell'obbligo ticinese e la loro difficoltà è valutata rispetto agli allievi dello stesso sistema scolastico. Concretamente, sono stati realizzati dieci

¹Si deve sottolineare che in funzione del tipo di modello IRT utilizzato si possono anche considerare altri parametri degli items, come ad esempio il coefficiente di discriminazione, che possono influenzare la probabilità di fornire una risposta corretta.

²Ad esempio, un individuo con un alto livello di abilità matematica avrà più probabilità di rispondere correttamente a un item di matematica rispetto a un individuo con un basso livello di abilità matematica. Inoltre, a parità di abilità, la probabilità di rispondere correttamente ad un item semplice è più alta rispetto alla probabilità di rispondere ad un item difficile

differenti fascicoli, ciascuno di essi richiedeva un tempo di risposta di 45 minuti. Questa distribuzione apparentemente complessa era necessaria per garantire che ogni item fosse testato su almeno 300 allievi e ogni allievo venisse confrontato con 2 fascicoli, l'uno a distanza di una settimana dall'altro. La distanza di una settimana è stata ritenuta quella minima per poter ritenere l'effetto di apprendimento residuale. Ogni allievo ha quindi ricevuto due fascicoli diversi assegnati casualmente.

La somministrazione è stata curata da personale appositamente formato, il quale si è occupato di portare le prove nelle singole classi, far eseguire il lavoro agli allievi e recuperare poi i materiali distribuiti. Quest'ultima fase è estremamente rilevante in quanto gli esercizi proposti in questa fase dei lavori non hanno ancora subito alcun processo di validazione e non possono essere ritenuti efficaci alla valutazione delle competenze specifiche, non essendoci infatti valori che ne indichino l'efficacia o la difficoltà in alcun modo. Si deve anche sottolineare come un esercizio diffuso anticipatamente alla prova e/o in maniera non corretta (ad esempio tramite fotocopie del materiale) potrebbe rendere l'attività di valutazione non valida, in quanto introdurrebbe una condizione di non equità di fronte alla prova. Un problema più ampio legato alla disponibilità degli esercizi è quello definito in letteratura “*teaching for testing*” (Flukiger, 2004).

Le risposte degli allievi alle due prove (i.e., matematica ed italiano) sono poi state inserite in un archivio al fine di poter valutare la bontà metrica degli esercizi. I singoli esercizi, le scale e l'insieme degli esercizi sono stati quindi valutati utilizzando il modello di Rasch, al fine di capire come costruire le successive prove e quali esercizi conservare. Queste analisi hanno permesso di identificare gli items (i.e., domande) definitivi che possedevano le caratteristiche metriche desiderate. Per la matematica sono stati selezionati 144 items divisi nei 6 settori, mentre per l'italiano sono stati selezionati 228 items divisi nei 4 settori.

1.5 La prova definitiva somministrata

Gli items selezionati sono stati successivamente utilizzati per la concezione delle due prove definitive. Sia per matematica che per italiano, sono stati concepiti una serie di fascicoli cartacei contenenti gli items in ordine crescente di difficoltà³ Per matematica sono stati concepiti 6 fascicoli diversi contenenti ognuno items appartenenti a tutti i 6 settori identificati. Per italiano sono stati invece concepiti 2 fascicoli contenenti tutti e 4 i settori identificati. Come per la fase di pre-test si è deciso di far svolgere le attività di somministrazione esclusivamente a personale formato specificatamente, ciò anche per non portare un aggravio di lavoro supplementare agli insegnanti.

Le somministrazioni delle due prove di matematica ed italiano sono state effettuate in due sessioni distinte svoltesi a una settimana di distanza una dall'altra.

³La difficoltà degli items è stata valutata durante la fase del pre-test.

In entrambe le sessioni, gli studenti hanno risposto a un fascicolo di matematica e a un fascicolo di italiano ⁴.

Una volta raccolte le prove, gli item sono stati nuovamente sottoposti ad una analisi relativa alla identificazione della difficoltà rispetto alla popolazione degli allievi e, sulla base di questi valori, sono state successivamente svolte le analisi. Nel periodo estivo sono stati elaborati i rapporti delle singole classi e consegnati a docenti e direttori. Ogni docente ha ricevuto un rapporto relativo alla sua classe nel quale si mostrava per la classe e per ogni allievo relativamente ad ogni singolo settore il punteggio medio rispetto all'insieme della popolazione testata (media del Cantone). Analogamente ai direttori sono stati inviati i rapporti relativi all'istituto di competenza. Nel capitolo relativo alle analisi verranno presentati i dati raccolti secondo tre punti di osservazione differenti: il territorio, le classi e gli allievi.

La prova definitiva (di italiano e matematica) è stata somministrata a tutta la popolazione di allievi ticinesi di quinta elementare ($N = 3043$). Nel presente rapporto vengono presentati i risultati degli allievi che hanno partecipato ad entrambe le sessioni di somministrazione e risposto ad entrambi i fascicoli di matematica e italiano ($n = 2506$).

⁴L'ordine di presentazione dei fascicoli (matematica - italiano vs. italiano-matematica) è stato controbilanciato, di modo che se uno studente alla prima sessione ha risposto prima al fascicolo di matematica e poi a quello di italiano, alla seconda somministrazione ha risposto prima al fascicolo di italiano e poi a quello di matematica e viceversa. L'ordine di presentazione controbilanciato permette di controllare un eventuale effetto di fatica (Borgonovi, & Biecek, 2016).

2 Caratteristiche socio demografiche

Le tabelle ed i grafici proposti di seguito forniscono alcune informazioni sulla configurazione sociodemografica della popolazione degli allievi. Per quanto concerne i dati socio-demografici degli allievi, questi sono stati recuperati tramite il sistema GAGI – Gestione allievi e istituti.

2.1 Nazionalità

I dati relativi alla nazionalità degli allievi sono stati ricodificati in modo tale da mantenere una sola indicazione per ciascun allievo. Nel caso in cui fosse presente la nazionalità svizzera, accompagnata da un'altra nazionalità, l'allievo è stato considerato di nazionalità svizzera. Lo stesso principio è stato applicato per allievi con doppia nazionalità non svizzera (in questo caso ha prevalso la prima nazionalità indicata nei dati).

Table 1: Nazionalità svizzera, italiana o altra

Nazionalità	n	Percentuale
Svizzera	1798	71.75%
Altra	361	14.41%
Italiana	347	13.85%

Come mostrato dal grafico (1), la maggior parte degli allievi (72%) è di nazionalità svizzera. Si osserva inoltre una percentuale consistente di allievi provenienti dall'Italia (14%) o da altri Paesi (14%). I Paesi più frequentemente rappresentati sono mostrati in dettaglio nella tabella 2 e nel grafico 2 seguenti. Sono stati esclusi i Paesi di provenienza con effettivi inferiori a 5 allievi.

2.2 Genere

Per quanto concerne la ripartizione per genere, nei dati a nostra disposizione risulta una distribuzione equilibrata tra i due generi. Gli allievi costituiscono il 49.2% del campione, mentre le allieve il 50.2%⁵.

2.3 Anno di nascita

La grande maggioranza della popolazione testata è nata nel 2010 (81.5%). Al momento della somministrazione della prova (maggio 2021) la maggior parte dei

⁵Per uno degli allievi non è stato possibile risalire alle informazioni riguardanti il genere.

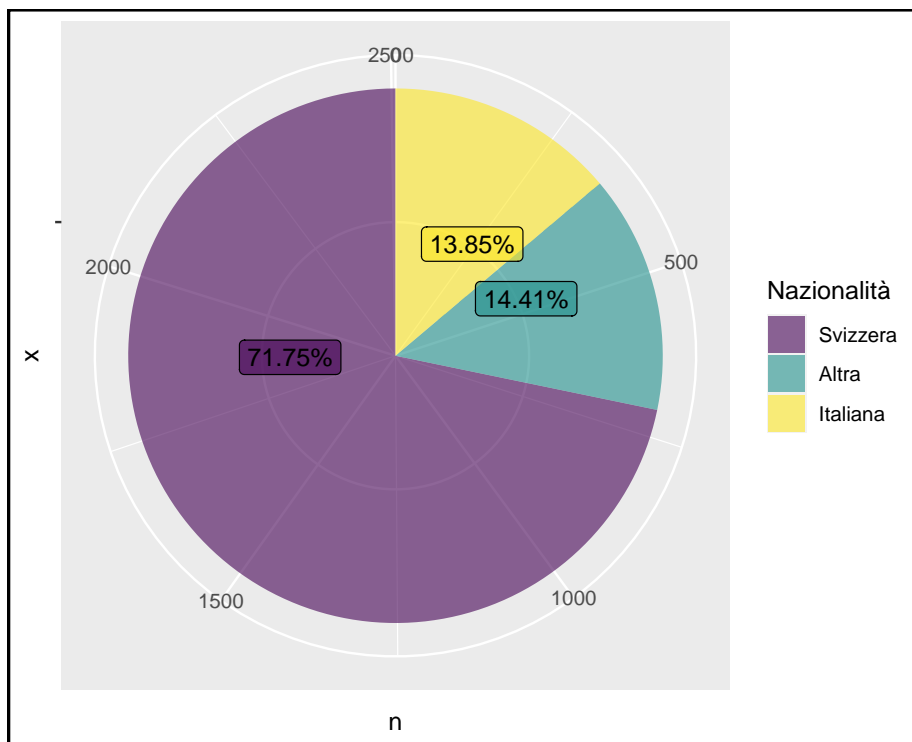


Figure 1: Nazionalità svizzera, italiana o altra

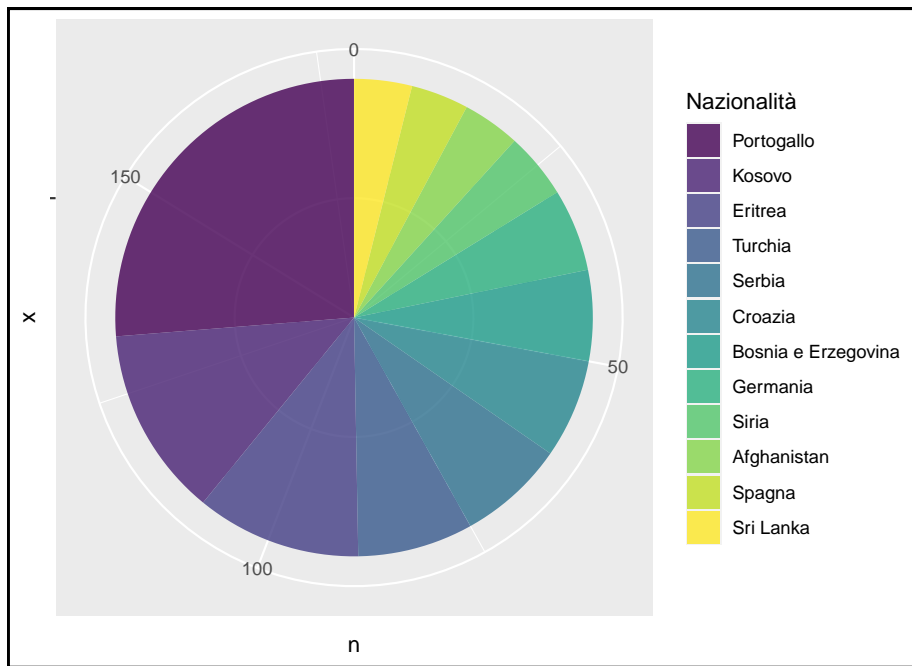


Figure 2: Nazionalità 'altra' (n>5)

Table 2: Nazionalità 'altra' (n>5)

Nazionalità 'altra' (n>5)	n	Percentuale
Portogallo	47	17.87%
Kosovo	23	8.75%
Eritrea	20	7.60%
Turchia	14	5.32%
Serbia	13	4.94%
Croazia	12	4.56%
Bosnia e Erzegovina	11	4.18%
Germania	10	3.80%
Siria	8	3.04%
Afghanistan	7	2.66%
Spagna	7	2.66%
Sri Lanka	7	2.66%

Table 3: Genere

Genere	n
F	1257
M	1248

soggetti aveva quindi un'età compresa tra i dieci e gli undici anni.

Table 4: Anno di nascita

Anno	n
2007	1
2008	13
2009	436
2010	1995
2011	2

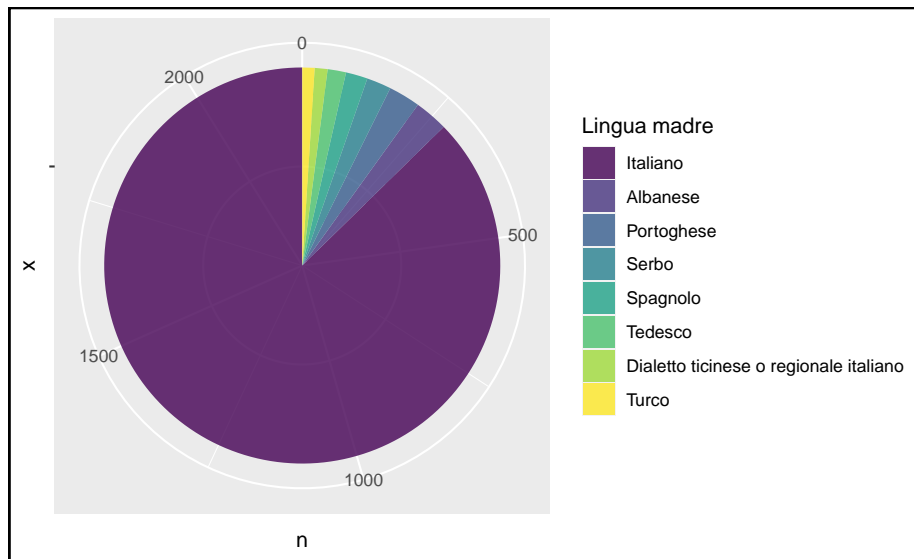
2.4 Lingua Madre

La maggior parte degli allievi è di lingua madre italiana (78%). Gli effettivi degli allievi con altre lingue madri è contenuto nel numero per ciascuna lingua considerata come mostrato nella tabella 5 e nel grafico ??.⁶

⁶Nella tabella e nel grafico sono indicate unicamente le lingue per i quali il numero di allievi è superiore a 20.

Table 5: Numero di allievi rispetto alla lingua madre ($n > 20$)

Lingua Madre	n
Italiano	1917
Albanese	60
Portoghese	57
Serbo	44
Spagnolo	39
Tedesco	33
Dialetto ticinese o regionale italiano	23
Turco	22



3 Analisi dei risultati della prova di Matematica

Per esporre i contenuti delle analisi si procederà con una logica a imbuto. Inizialmente saranno presentate quelle più generali poi si scenderà progressivamente nel dettaglio. Da un punto di vista di informazioni rispetto al sistema dopo i valori cantonali saranno considerati gli ispettorati, quindi alcune dimensioni territoriali e poi le classi. Nel capitolo successivo saranno approfondite le dimensioni relative alle caratteristiche sociodemografiche degli allievi.

Il livello dell'ispettorato è stato scelto anche per la funzione di vigilanza e di organizzazione che viene svolta dagli ispettorati e dal collegio degli ispettori. Il Cantone infatti ha compiti di istituzione e direzione della scuola svolti in collaborazione con i comuni. Il collegio degli ispettori e la Sezione delle scuole comunali (SESCO) sono, per le scuole comunali, gli organi di raccordo tra le scuole del territorio e il DECS (Legge della scuola, 1990).

3.1 Informazioni generali

Per avere una visione di insieme sono stati inizialmente analizzati i punteggi ottenuti nei singoli traguardi e la loro media. Questo permette un rapido raffronto tra le differenti parti che compongono la prova. Il punteggio nel traguardo è stato calcolato definendo per ogni item la difficoltà sul traguardo specifico e quindi sommandone i valori. Al fine di ottenere un valore che tenesse conto della difficoltà dei singoli item rispetto alla prova nel suo insieme è stato quindi fatto il medesimo calcolo tenendo però conto non più solo dei singoli traguardi ma della difficoltà complessiva della prova. Questo valore è stato definito “Matematica generale”. I risultati medi sono indicati nella tabella 6 e nel grafico 3.

Tutti i punteggi sono stati normalizzati in modo da assumere valori compresi tra 0 e 100. I punteggi non equivalgono però a percentuali corrispondenti al numero di esercizi svolti correttamente: ottenere 50 in un certo settore non significa infatti aver svolto correttamente il 50% degli item di quel traguardo di competenza. Per ciascun traguardo gli item sono stati infatti ponderati per il rispettivo coefficiente di difficoltà; questo è stato calcolato ponderando la quantità di risposte corrette allo specifico item. In pratica, chi ha svolto correttamente gli esercizi con elevato coefficiente di difficoltà ottiene un punteggio superiore a chi ha svolto un uguale numero di esercizi con coefficiente di difficoltà inferiore. All'interno dei fascicoli gli item utilizzati sono stati proposti in ordine crescente di difficoltà ed è quindi probabile che chi ha risposto correttamente agli item più difficili abbia risposto correttamente anche a quelli precedenti. Il valore “Matematica generale” è quindi quella variabile costituita dalla somma ponderata di risposte corrette rispetto a una scala da 0 a 100. Questo valore è stato calcolato per ogni allievo e può essere considerato come un indicatore sintetico della sua prestazione (Crescentini, Salvisberg, & Zanolla, 2014) alla prova. Il sistema con il quale è stato calcolato il valore di Matematica Generale è, rispetto ai precedenti rapporti, più raffinato e permette di valutare meglio

la competenza generale. Non si tratta infatti della media dei valori degli altri traguardi ma bensì è stato ricalcolato rispetto alla difficoltà di tutti gli item. Il valore medio per tutta la popolazione è pari a 51.68.

Table 6: Risultati medi cantonali

GEO_CA	45.42
GEO_MM	36.58
GEO_SRD	52.63
GM_EA	44.61
GM_SRD	59.45
NC_MT	50.81
Matematica_Generale	51.68

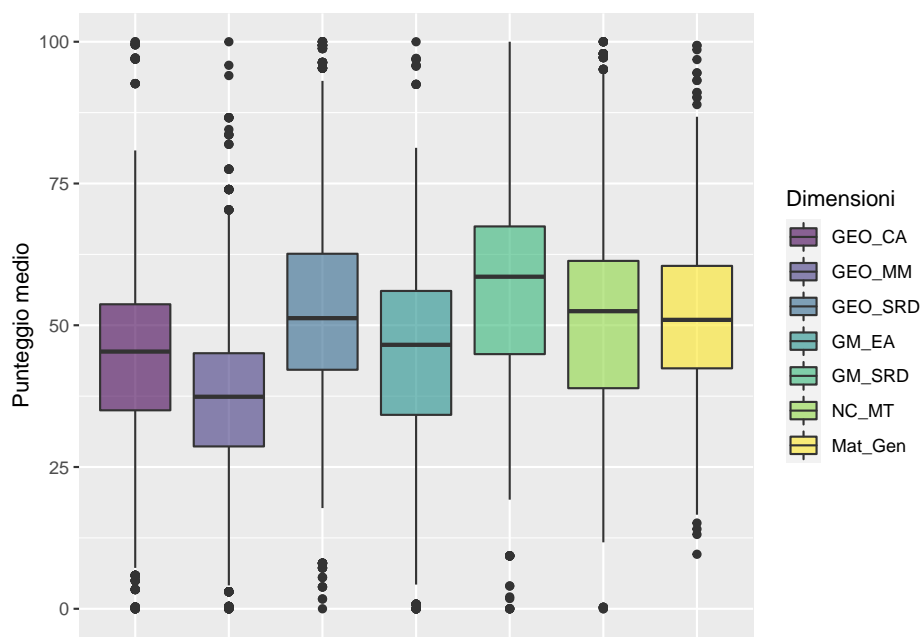


Figure 3: Risultati medi cantonali nelle varie dimensioni testate

Come illustrato dal grafico 3 il traguardo che ha raggiunto il valore medio più elevato è “Grandezze e misure - Sapere, riconoscere e descrivere” (GM SRD) mentre quello con il valore più basso è “Numeri e calcolo - Matematizzare e modellizzare” (NC MT). Si può notare come, all’interno dei due Ambiti di competenza scelti (Grandezze e misure e Numeri e calcolo) gli Aspetti di competenza relativi a Matematizzare e modellizzare risultino quelli apparentemente

più difficili e quelli relativi a Sapere, riconoscere e descrivere risultino i due con i risultati più elevati.

3.2 Correlazioni tra le dimensioni valutate con la prova

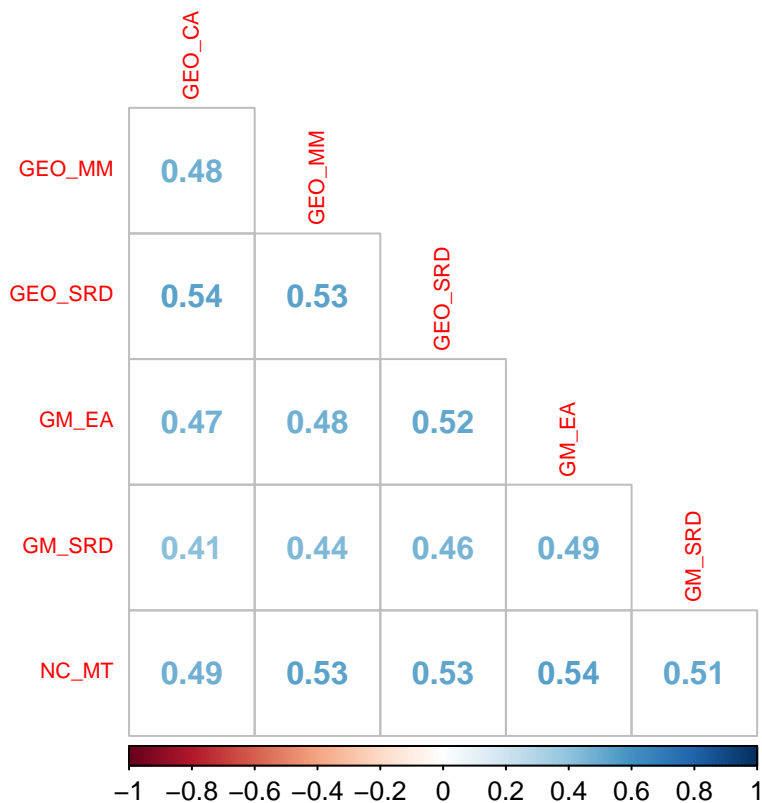


Figure 4: Correlazioni tra i risultati nelle varie dimensioni testate

Il grafico 4 mostra come i risultati nelle diverse dimensioni siano fortemente correlati tra di loro⁷. Come si può osservare i valori di correlazione sono tutti positivi e superiori a .41 ad indicare come più i risultati sono elevati in una dimensione più lo sono sulle altre. Questo risultato non è inatteso in quanto si tratta di dimensioni parte del costrutto più generale valutato che corrisponde alla matematica in generale. Si tratta però di una motivazione aggiuntiva alla

⁷Le correlazioni di Pearson sono generalmente utilizzate al fine di quantificare la covarianza (e.g., associazione) tra due o più variabili numeriche. I valori di una correlazione variano tra -1 e 1, dove -1 indica una perfetta associazione negativa, e 1 una perfetta associazione positiva. Al valore 0 è invece associata l'assenza di associazione, in questo caso le variabili sono indipendenti.

costruzione della dimensione Matematica generale che rappresentasse il punteggio globale del singolo allievo alla prova.

3.3 Il contesto e la scuola

Gli ispettorati sono le unità organizzative territoriali più ampie per quanto riguarda le scuole comunali. Il nostro territorio per quanto riguarda le scuole comunali è diviso in quattro aree. Da sud verso nord si trovano Mendrisiotto e Basso Ceresio (20.2%), Luganese (40.2%), Locarnese e Valli (15.5%) e Bellinzonese e Tre Valli (24%). La ricerca ha coinvolto 198 classi, tra pubbliche e private, di queste 54 sono pluriclassi.

Tutti gli ispettorati sono stati testati nel mese di maggio 2021.

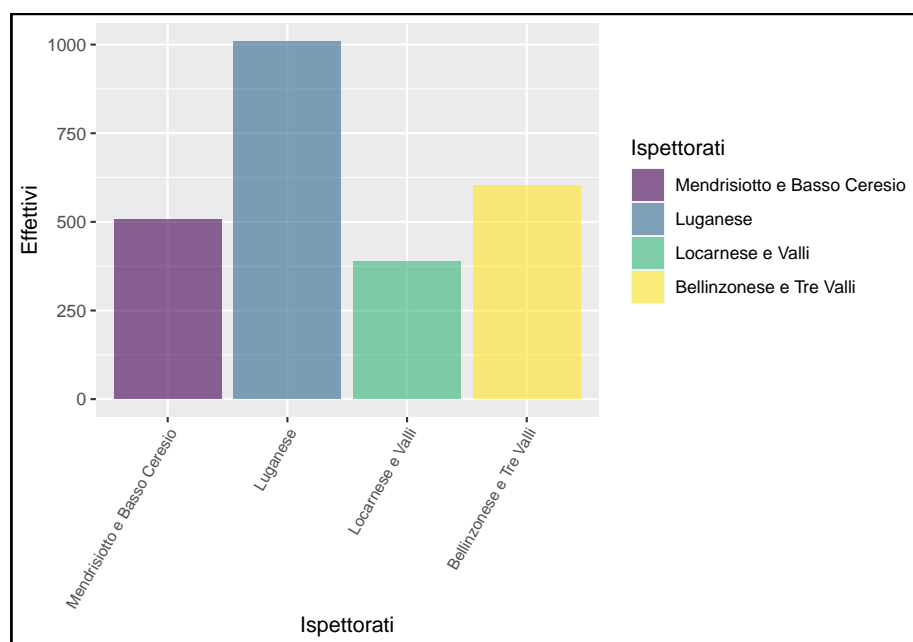


Figure 5: Numero di studenti nei 4 Ispettorati

Di seguito nella tabella 7 e nel grafico 6 troviamo i punteggi medi nei 4 ispettorati con i valori relativi alle diverse sottodimensioni analizzate.

3.3.0.1 Analisi statistiche Le differenze tra gli ispettorati non risultano essere statisticamente significative né a livello della dimensione “Matematica

Table 7: Risultati medi cantonali in funzione dell'ispettorato

Ispettorati	GEO_CA	GEO_MM	GEO_SRD	GM_EA	GM_SRD	NC_MT	Mat_Gen
Mendrisiotto e Basso Ceresio	45.63	35.93	52.43	45.64	58.51	49.63	51.24
Luganese	44.47	37.30	52.36	43.69	59.98	51.02	51.62
Locarnese e Valli	45.56	36.46	52.96	44.43	60.77	52.21	52.31
Bellinzonese e Tre Valli	46.75	36.01	53.05	45.40	58.48	50.56	51.75

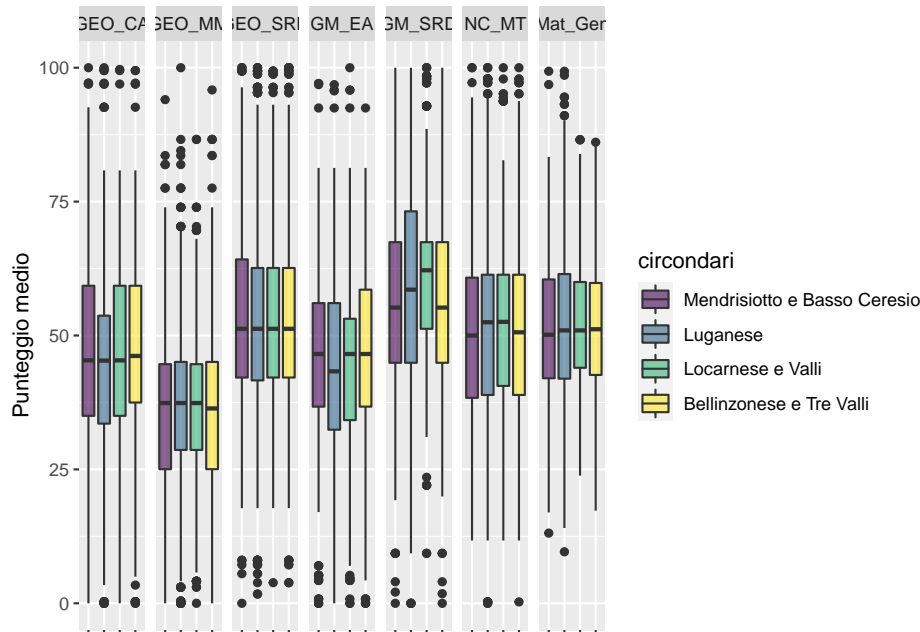


Figure 6: Risultati medi cantonali in funzione dell'ispettorato

generale”, né per le sottodimensioni valutate, analizzando l’insieme dei dati tramite l’analisi della varianza (ANOVA). Possiamo presupporre dunque una certa uniformità nelle competenze degli allievi appartenenti a ispettorati diversi.

3.4 Il grado di urbanizzazione del comune

Il livello organizzativo più direttamente visibile delle scuole elementari e dell’infanzia è il comune. Per lo studio dei comuni viene proposto dall’Ufficio Federale di Statistica un indice relativo alla dimensione del singolo comune. Per il Ticino vi sono tre gruppi: città nucleo; comuni della cintura; comuni rurali. Come si può osservare nel grafico 7, la maggior parte degli allievi si situa nei comuni della cintura (55.5%), seguiti dagli allievi dei comuni rurali (34.9%) ed infine quelli delle città nucleo (9.7%).

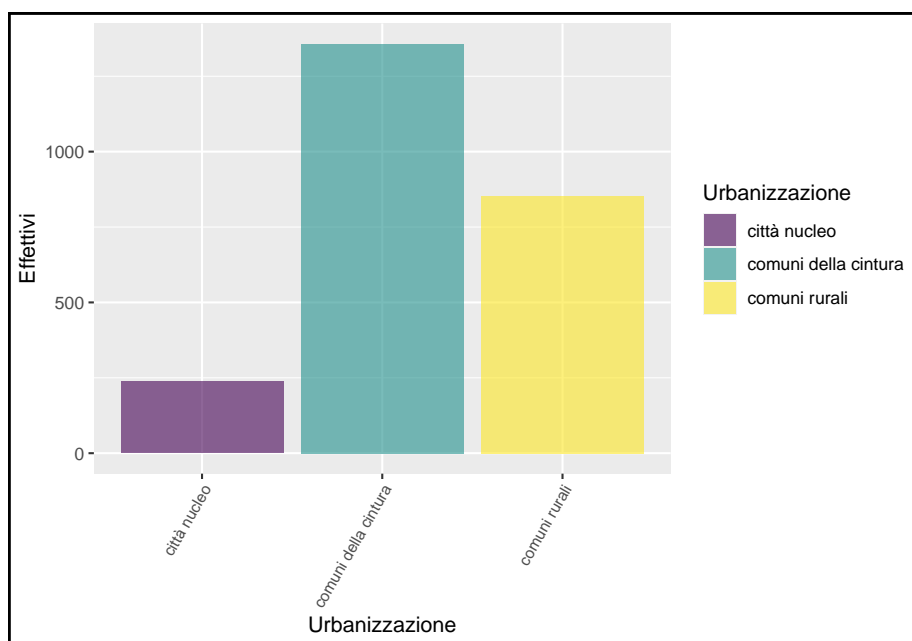


Figure 7: Numero di studenti in funzione del grado di urbanizzazione del comune

L'effetto del livello di urbanizzazione sugli apprendimenti scolastici è stato oggetto di studi rivelandosi una delle caratteristiche da considerare [?]. Nella tabella 8 e nel grafico 8 qui sotto vengono riportati i risultati di competenza generale e relativi alle diverse sottodimensioni secondo il grado di urbanizzazione del comune (città nucleo; comuni della cintura; comuni rurali).

Table 8: Risultati medi cantonali in funzione del grado di urbanizzazione del comune

Urbanizzazione	GEO_CA	GEO_MM	GEO_SRD	GM_EA	GM_SRD	NC_MT	Mat_Gen
città nucleo	40.27	37.58	52.23	43.40	57.94	50.63	50.54
comuni della cintura	45.68	36.33	52.55	44.56	59.42	50.82	51.65
comuni rurali	46.55	36.67	52.80	45.22	60.02	50.80	52.06

3.4.0.1 Analisi statistiche Le analisi di varianza realizzate non mettono in evidenza unicamente una differenza significativa ($F(2,2444)=12.75$, $p<.001$) unicamente per la dimensione Geometria – Comunicare e argomentare (GEO_CA). In questa dimensione i comuni rurali ($m = 46.55$) ed i comuni della cintura ($m = 45.68$) risultano avere prestazioni più elevate rispetto ai comuni delle città

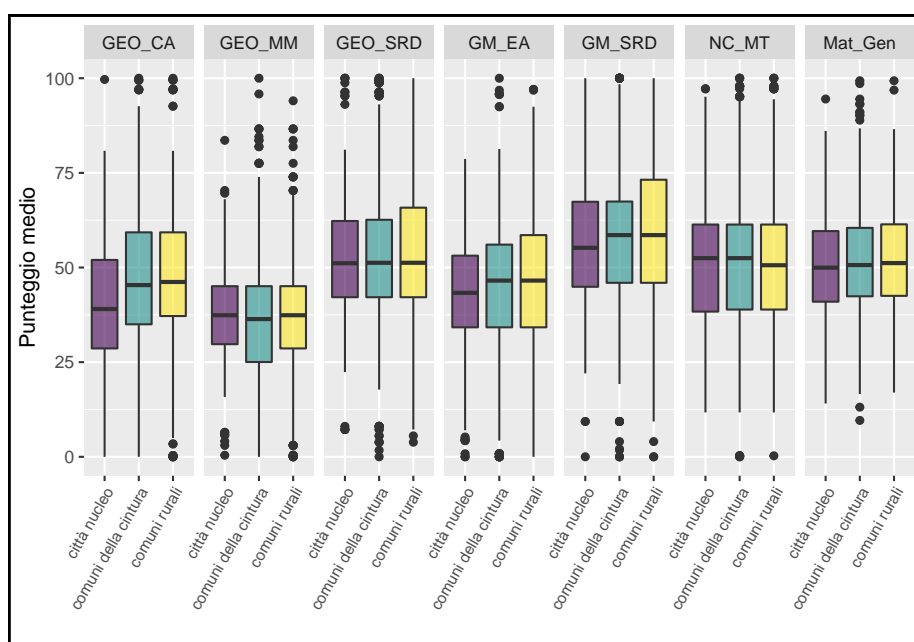


Figure 8: Risultati medi cantonali in funzione del grado di urbanizzazione del comune

nucleo ($m = 40.27$). I dettagli delle analisi realizzate possono essere consultate nell'Appendice Matematica: Analisi di varianza (ANOVA) - Urbanizzazione.

3.5 La classe

Nel valutare le prestazioni degli allievi le caratteristiche della classe sono ritenute estremamente rilevanti [?]. In coerenza con quanto considerato nei precedenti rapporti sono state considerate: numero di allievi per classe e il fatto che gli allievi testati fossero in mono o pluriclasse.

3.5.1 Dimensione della classe o numero di allievi

Il dibattito sull'effetto del numero di allievi per classe è, a livello internazionale, estremamente ricco e riguarda prevalentemente l'interazione che può avvenire nella classe tra docenti e allievi al ridurre gli effettivi (Crescentini & Galetta, 2014).

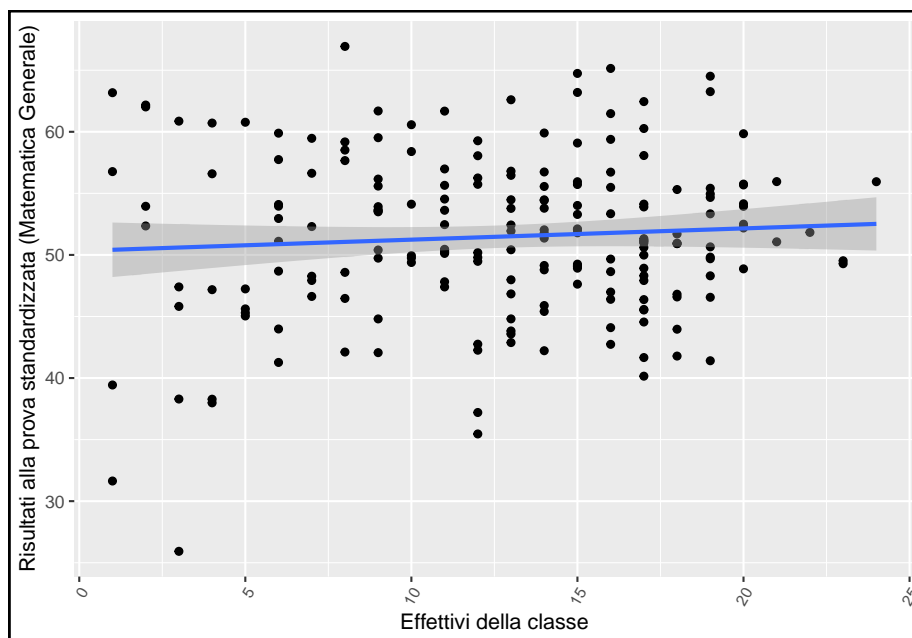


Figure 9: Relazione tra effettivi della classe e risultati alla prova standardizzate (Matematica Generale)

Per quanto riguarda i dati da noi analizzati, una breve analisi della relazione tra il numero di allievi per classe e il punteggio ottenuto nella prova standardizzata

indica come non sia possibile osservare nessuna tendenza esplicita (correlazione = 0.074) come rappresentato dalla linea blu orizzontale che rappresenta la correlazione praticamente nulla nel grafico 9.

Numerosità della classe e performance in Matematica non appaiono quindi collegati tra loro. Questo risultato è coerente con quanto osservato nei rapporti precedenti nei quali non si era rilevata alcuna differenza significativa collegata alla numerosità delle classi.

3.5.2 La pluriclasse

La pluriclasse è una classe formata da bambini che frequentano anni di corso differenti. Si tratta di una particolare organizzazione di classe della scuola primaria diffusa soprattutto nelle zone a bassa densità abitativa, laddove non si raggiunge il numero minimo di iscritti necessari per la formazione di classi omogenee. Gli studi sull'efficacia della pluriclasse e sugli effetti sugli allievi hanno iniziato a essere condotti negli anni '70 collegandosi alle riflessioni sulle scuole delle aree discoste negli Stati Uniti (Miller, 1988). Il tema è sfaccettato e vi sono ricerche che identificano sia vantaggi sia svantaggi.

Un elenco non esaustivo può prevedere tra i vantaggi: riduzione della disobbedienza, maggiore sviluppo linguistico [?]; migliori apprendimenti nella lettura e nelle capacità linguistiche [?]; migliore e più rapido sviluppo cognitivo [?]; migliori e più approfondite relazioni tra insegnanti e allievi e loro famiglie (Miller, 1994); maggiore tranquillità sin dal primo giorno [?]; maggiore possibilità di un apprendimento centrato sulla velocità del singolo studente [?]; il rapporto maggiormente individualizzato può portare a un atteggiamento degli studenti migliore nei confronti della scuola e delle relazioni con i pari [?]; le interazioni sociali e il lavoro collaborativo appaiono più sviluppati [?]; la differenziazione applicata aiuta gli allievi più dotati [?]; vengono sviluppate strategie di insegnamento maggiormente basate sulla collaborazione tra gli allievi [?]; lo sviluppo di strategie di assistenza tra gli allievi promuove la percezione che la differenza sia normale e che si sia parte di una comunità di apprendimento [?]; la presenza di allievi più maturi può essere di stimolo [?].

Tra gli svantaggi emergono invece: difficoltà maggiore per creare i materiali e i gruppi riportata dai docenti e mancanza di una formazione specifica [?]; difficoltà gestionali nell'organizzare le classi [?]; l'insegnamento è meno efficace e l'apprendimento più complesso [?]; la presenza di allievi più giovani può avere un effetto negativo sull'apprendimento [?].

Nella tabella 9 e nel grafico 10 vengono riportati i punteggi medi nella dimensione Matematica Generale e nelle altre dimensioni ottenuti dagli allievi delle monoclasse (classi tradizionali) e delle pluriclassi.

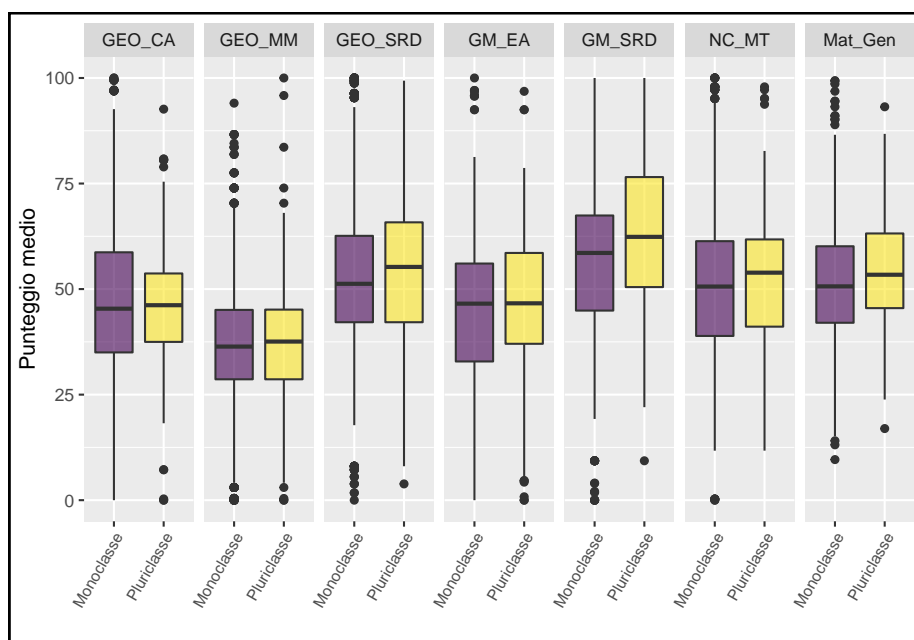


Figure 10: Punteggi medi degli allievi nelle diverse dimensioni in funzione del tipo di classe

Table 9: Punteggi medi degli allievi nelle diverse dimensioni in funzione del tipo di classe

Classe	GEO_CA	GEO_MM	GEO_SRD	GM_EA	GM_SRD	NC_MT	Mat_Gen
Monoclasse	45.291	36.421	52.442	44.244	59.183	50.460	51.438
Pluriclasse	46.561	38.017	54.304	47.814	61.747	53.885	53.846

3.5.2.1 Analisi statistiche Le analisi realizzate mettono in evidenza delle differenze significative nella competenza generale in matematica ($t(318.94) = -2.748$, $p = .006$) con gli allievi delle pluriclassi ($m = 53.85$) che ottengono risultati migliori degli allievi nelle monoclassi ($m = 51.44$).

Per quel che riguarda le singole dimensioni appaiono dei *pattern* di risultati simili per le dimensioni GM_EA, GM_SRD e NC_MT. Con gli allievi delle pluriclassi che ottengono risultati superiori a quelli delle monoclassi.

I dettagli delle analisi statistiche realizzate possono essere consultati nell'Appendice Matematica: Test t di Student - La classe.

3.6 L'allievo

Nel valutare le prestazioni in una prova è rilevante considerare le caratteristiche degli allievi e verificare se non possano esservi caratteristiche dell'allievo che abbiano un impatto sulla prestazione stessa. Questo oltre ad avere una rilevanza dal punto di vista dell'equità della prova potrebbe anche essere rilevante rispetto alla organizzazione in generale. Sono state quindi considerate caratteristiche socio demografiche delle quali verificare l'eventuale impatto: nazionalità e lingua madre; genere; origine sociale; età.

3.6.1 La nazionalità

I dati relativi alla nazionalità degli allievi sono stati ricodificati in modo tale da mantenere una sola indicazione per ciascun allievo. Nel caso in cui fosse presente la nazionalità svizzera, accompagnata da un'altra nazionalità, l'allievo è stato considerato di nazionalità svizzera. Lo stesso principio è stato applicato per allievi con doppia nazionalità non svizzera (in questo caso ha prevalso la prima nazionalità indicata nei dati).

La tabella 10 ed il grafico 11 riportano i punteggi medi ottenuti dagli allievi di nazionalità svizzera, italiana o altra alla dimensione di Matematica Generale e alle altre dimensioni testate.

3.6.1.1 Analisi statistiche Le analisi di varianza realizzate permettono di identificare delle differenze di media statisticamente significative per quel che riguarda i risultati in matematica generale ($F(2,2503) = 9.08$, $p < .001$) con

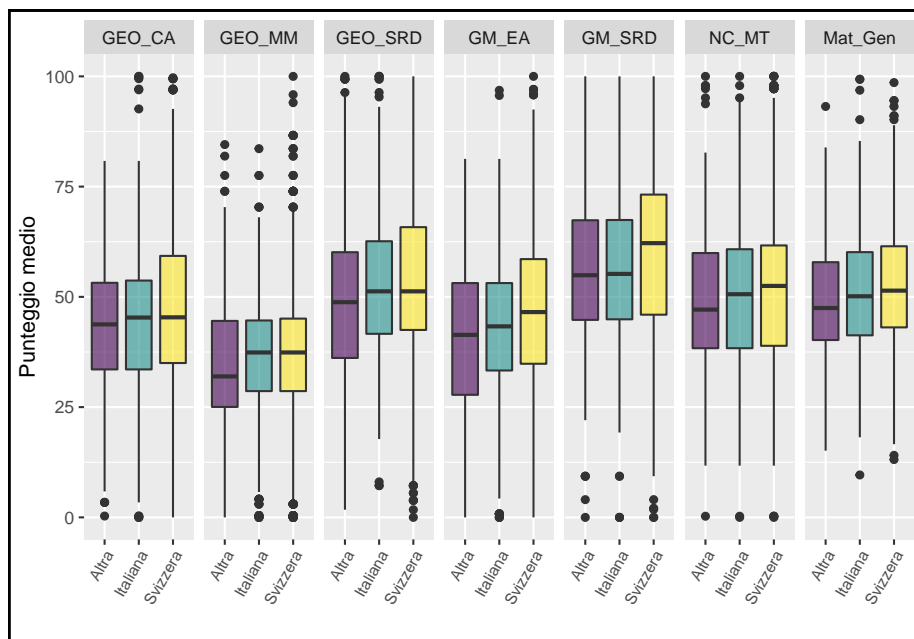


Figure 11: Punteggi medi degli allievi nelle diverse dimensioni in funzione della nazionalità

Table 10: Punteggi medi degli allievi nelle diverse dimensioni in funzione della nazionalità

Nazionalità	GEO_CA	GEO_MM	GEO_SRD	GM_EA	GM_SRD	NC_MT	Mat_Gen
Svizzera	46.02	37.03	53.32	45.57	60.22	51.33	52.35
Italiana	44.63	35.82	52.50	43.57	58.14	49.85	50.79
Altra	43.22	35.08	49.34	40.80	56.85	49.13	49.20

gli allievi di nazionalità svizzera ($m = 52.35$) che ottengono risultati più elevati rispetto agli allievi della categoria altre nazionalità ($m = 49.20$). Per contro non esistono differenze significative tra gli studenti di nazionalità svizzera e quelli di nazionalità italiana.

Per quel che concerne le singole dimensioni, a parte per la dimensione GEO_MM ($p = .053$) si osservano differenze statisticamente significative per tutte le dimensioni. Si osserva come gli allievi di nazionalità svizzera ottengano sistematicamente risultati migliori, seguiti dagli allievi di nazionalità italiana ed infine dagli allievi di altre nazionalità.

I dettagli delle analisi statistiche realizzate possono essere consultate nell'Appendice Matematica: Analisi di varianza (ANOVA) - Nazionalità.

3.6.2 La lingua madre

L'apprendimento della matematica è potenzialmente legato alle conoscenze linguistiche dell'allievo. A livello internazionale diverse ricerche empiriche hanno mostrato come la competenza linguistica possa impattare sull'apprendimento della matematica e sulle performance [??]. Si può dunque intuire che bambini con una seconda lingua diversa dall'italiano possano avere alcune difficoltà nell'apprendere nozioni matematiche complesse in una lingua seconda.

Anche alle nostre latitudini si sono avviate delle riflessioni in merito al legame tra la lingua italiana e l'apprendimento in matematica. Ne è un esempio il progetto Italmatica (si veda per approfondimento in merito al progetto il testo di ?).

In uno degli articoli legati a questo progetto, ? riportano i risultati della valutazione didattica della prova standardizzata di matematica somministrata nel maggio 2015 a tutti gli allievi di quinta elementare del Canton Ticino (la prova descritta in ?). Questi mostrano come diverse risposte sbagliate degli allievi siano legate a difficoltà di comprensione e di interpretazione linguistica del testo [??]. Ne consegue che la lingua italiana e la sua comprensione possono dunque avere un impatto sulla performance in matematica.

All'interno della tabella ??tab:linguaMadreDescr) e del grafico ??fig:linguaMadrePlot) si trovano i punteggi medi ottenuti dagli allievi nelle varie dimensioni testate ed in funzione della lingua madre (i.e, italiano vs. altra lingua).

Table 11: Punteggi medi degli allievi nelle diverse dimensioni ed in funzione della lingua madre

Lingua Madre	GEO_CA	GEO_MM	GEO_SRD	GM_EA	GM_SRD	NC_MT	Mat_Gen
Italiano	46.53	37.27	53.71	45.70	60.44	51.55	52.66
Altra lingua	41.82	34.34	49.12	41.07	56.22	48.39	48.51

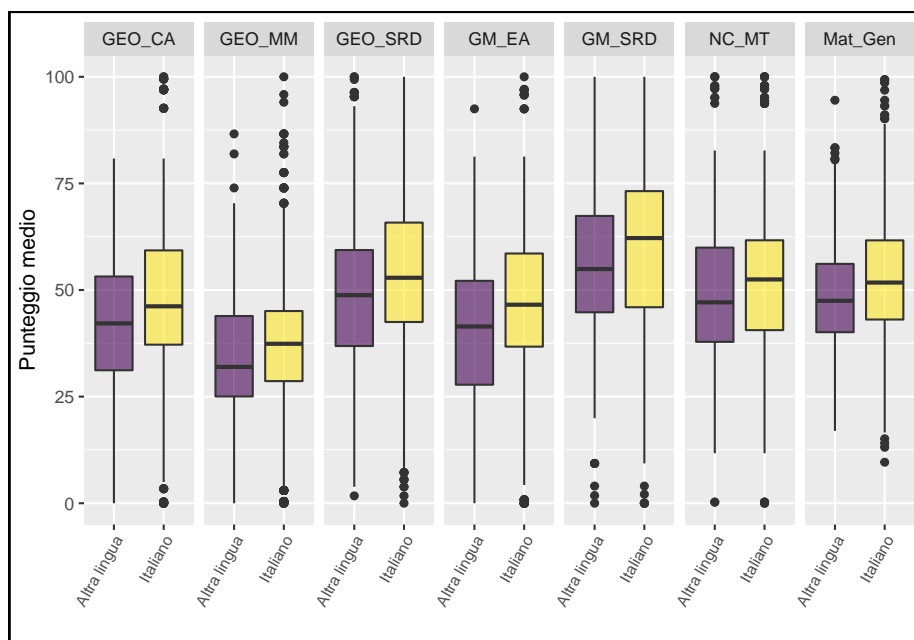


Figure 12: Punteggi medi degli allievi nelle diverse dimensioni in funzione della lingua madre

3.6.2.1 Analisi statistiche Per quel che riguarda i risultati in matematica generale, il test t di Student realizzato indica delle differenze statisticamente significative ($t(1037.1) = -6.803, p < .001$). Gli allievi di lingua madre italiana ottengono risultati migliori ($m = 52.66$) rispetto agli allievi di lingua madre diversa ($m = 48.51$).

Per quel che riguarda le singole dimensioni i *pattern* di risultati sono simili ed indicano in modo sistematico come gli allievi di lingua madre italiana ottengano risultati migliori rispetto a quelli di lingua madre diversa.

I dettagli delle analisi statistiche realizzate possono essere consultati nell'Appendice Matematica: Test t di Student - La lingua madre.

3.6.3 Il Genere

Come rilevato precedentemente [?] le differenze di genere possono avere un impatto sulla prestazione in considerazione sia della socializzazione primaria (in famiglia) sia di quella esterna alla scuola, sia in riferimento a eventuali effetti presenti a scuola. Riuscire a distinguere questi effetti va oltre le possibilità del presente lavoro ma è importante rendere attento il lettore in proposito.

Table 12: Punteggi medi degli allievi nelle diverse dimensioni in funzione del genere

Genere	GEO_CA	GEO_MM	GEO_SRD	GM_EA	GM_SRD	NC_MT	Mat_Gen
M	44.99	37.84	53.20	46.13	63.14	52.49	53.05
F	45.86	35.34	52.09	43.13	55.79	49.16	50.35

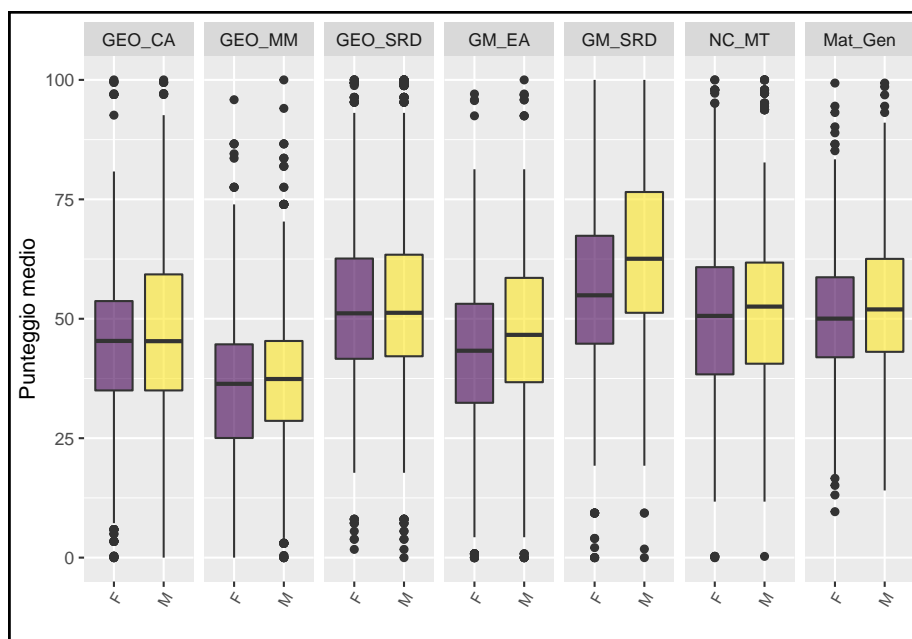


Figure 13: Punteggi medi degli allievi nelle diverse dimensioni in funzione del genere

3.6.3.1 Analisi statistiche Per quanto concerne la competenza generale in matematica le analisi realizzate mettono in evidenza delle differenze statisticamente significative ($t(2494.38) = -5.01, p < .001$) con i maschi ($m = 53.05$)

che presentano delle performance migliori rispetto alle femmine ($m = 50.35$). Questo risultato è in linea con quanto emerge solitamente a livello internazionale nei test di competenza di tipo standardizzato. In matematica i ragazzi ottengono infatti migliori risultati in matematica in tutti i livelli scolastici e nella maggioranza dei paesi (OECD 2016; Mullis et al. 2016 citati da ?, p. 81).

È interessante osservare come pure per quasi tutte le dimensioni (eccezion fatta per GEO_CA e GEO_SRD) i maschi ottengano risultati migliori rispetto alle femmine.

I dettagli delle analisi statistiche realizzate possono essere consultati nell'Appendice Matematica: Test t di Student - Il genere.

3.6.4 Origine Sociale

Come avviene nella maggior parte dei lavori sul tema anche nel presente documento saranno utilizzate le professioni dei genitori per categorizzare i figli con la consapevolezza che ci si riferisce alle condizioni di origine e che questo ha delle relazioni con le opportunità che possono essere offerte.

Nel presente rapporto l'operazione di categorizzazione sono state basate sullo *International Standard Classification of Occupation (ISCO)*⁸ che ha permesso di creare quattro categorie distinte:

- Colletti bianchi altamente qualificati (CBianchiAQ)
- Colletti bianchi scarsamente qualificati (CBianchiSQ)
- Colletti blu altamente qualificati (CBluAQ)
- Colletti blu scarsamente qualificati (CBluSQ)

L'origine sociale degli allievi viene spesso indicata come uno dei fattori rilevanti che influenzano le prestazioni scolastiche. Le ragioni sono molteplici e fanno riferimento principalmente alle aspettative della famiglia e alle risorse che sono messe a disposizione degli allievi sia in termini materiali sia in termini culturali nel senso più ampio del termine ivi compreso il capitale sociale. L'analisi condotta conferma che, in tutte le dimensioni, quanto più elevato è il retroterra socioculturale di un individuo, migliore sarà la sua prestazione. I figli di genitori appartenenti alla categoria “colletti bianchi altamente qualificati” riportano risultati significativamente migliori. Sia nelle prove standardizzate, sia nella valutazione da parte degli insegnanti, i figli dei gruppi socialmente più svantaggiati si posizionano peggio. Vale la pena osservare come tra la fascia sociale di origine più modesta e la fascia sociale di origine più elevata vi sia, nella valutazione di fine anno, una differenza media di mezzo punto come si può osservare dal grafico 14.

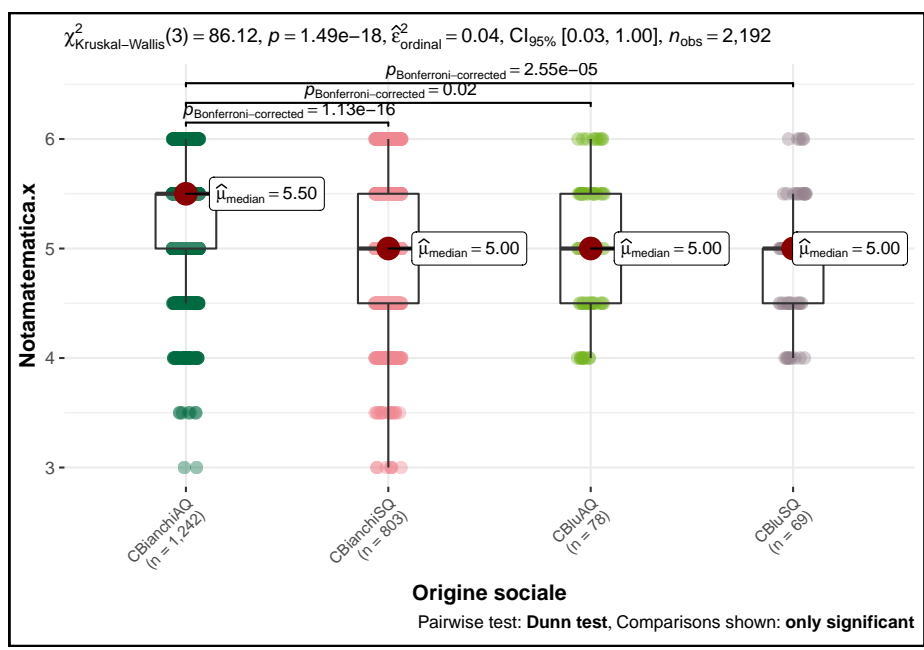


Figure 14: Risultati di fine anno (i.e., Note scolastiche) in Matematica in funzione dell'origine sociale

La tabella 13 ed il grafico 15 seguenti riportano i risultati ottenuti dagli allievi nelle varie dimensioni testate.

Table 13: Punteggi medi degli allievi nelle diverse dimensioni in funzione dell'origine sociale

Origine Sociale	GEO_CA	GEO_MM	GEO_SRD	GM_EA	GM_SRD	NC_MT	Mat_Gen
CBluSQ	40.93	31.70	47.07	36.48	54.24	46.32	46.29
CBluAQ	43.06	34.80	51.77	43.87	58.91	48.31	50.10
CBianchiSQ	43.18	34.27	49.82	42.37	56.41	48.01	48.97
CBianchiAQ	47.73	38.66	55.15	47.22	62.21	53.23	54.21

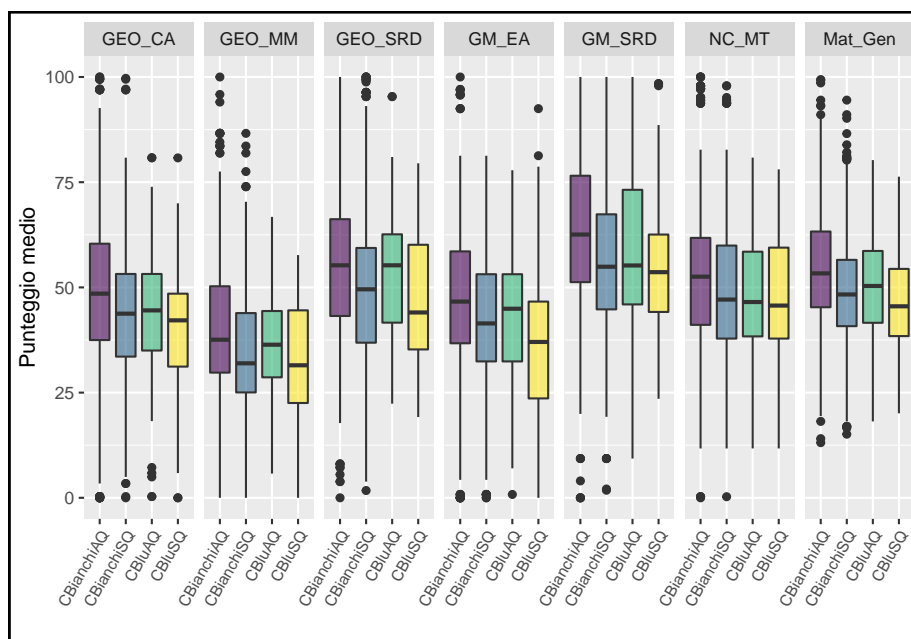


Figure 15: Punteggi medi degli allievi nelle diverse dimensioni in funzione dell'origine sociale

3.6.4.1 Analisi statistiche

Per quel che riguarda i risultati alle prove standardizzate l'analisi della varianza dei punteggi medi ottenuti nella dimensione

⁸<https://www.ilo.org/public/english/bureau/stat/isco/index.htm>

“Matematica generale” mostra delle differenze statisticamente significative in funzione dell’origine sociale degli allievi ($F(3,2188) = 30.21, p < .001$) con gli allievi di origine sociale “Colletti bianchi altamente qualificati” ($m = 54.21$) che ottengono risultati più elevati rispetto agli allievi di origine “Colletti bianchi scarsamente qualificati” ($m = 48.97$), rispetto agli allievi di origine “Colletti blu altamente qualificati” ($m = 50.10$) e rispetto agli allievi di origine “Colletti blu scarsamente qualificati” ($m = 46.29$). I confronti tra gli altri gruppi non sono statisticamente significativi.

Anche per quanto concerne le singole dimensioni testate, vi sono differenze significative rispetto alle performance degli allievi in funzione della loro origine sociale. Gli allievi di origine sociale più elevata hanno delle performance migliori in tutte le dimensioni testate.

I dettagli delle analisi statistiche realizzate possono essere consultati nell’Appendice Matematica: Test t di Student - L’origine sociale.

3.7 Relazione tra risultati alle prove standardizzate e nota scolastica di Matematica

Verrà ora approfondita la relazione tra i risultati ottenuti alle prove standardizzate e la nota finale in matematica degli allievi.

Si può osservare nel grafico 16 come sia il valore nella dimensione matematica generale, che quello delle singole dimensioni misurato dalle prove standardizzate correlino positivamente con i risultati scolastici degli allievi misurati dalle note scolastiche in matematica. Da notare come la correlazione più importante sia quella tra il livello di competenza generale in matematica e la nota di matematica ($r = .67$). Questo valore indica che circa il 45% della varianza dei valori è comune alle due misure (i.e., indice di una buona relazione tra i valori).

Nel grafico 17 possiamo osservare la relazione positiva tra i risultati alle prove e le note in matematica degli allievi. La linea blu nel grafico rappresenta la retta di regressione che esprime la relazione lineare tra queste due variabili. Come si può osservare, all’aumentare del punteggio alla prova standardizzata, corrisponde un aumento per quel che riguarda a nota scolastica in matematica ($r = .67$).

Proviamo ora ad analizzare se questa relazione sia differente per gli allievi che hanno ottenuto risultati più bassi alla prova standardizzata e quelli che hanno ottenuto i risultati più alti. Utilizzando il valore mediano della distribuzione dei risultati alla prova standardizzata creiamo dunque due “gruppi” di allievi.

Nel grafico @ref(fig:NoteMatePlot_LOW) sono rappresentate le relazioni tra i risultati alle prove e le note in matematica degli allievi in funzione dei risultati più o meno elevati alle prove. Per gli allievi che hanno ottenuto dei risultati

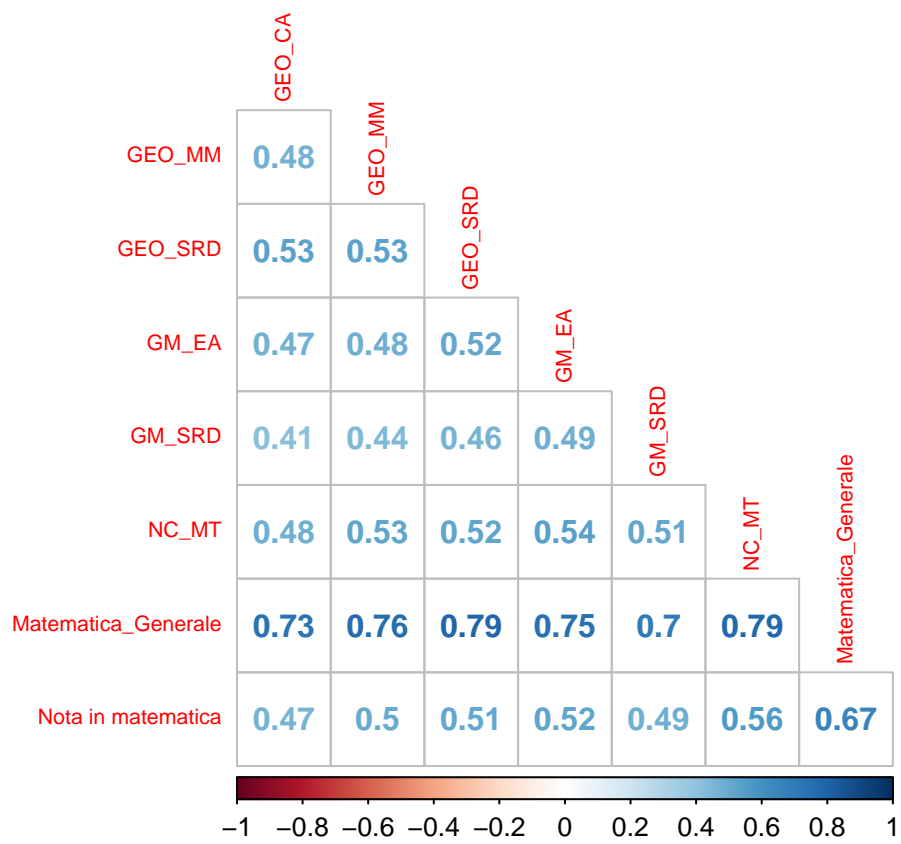


Figure 16: Correlazioni tra nota scolastica e prove standardizzate

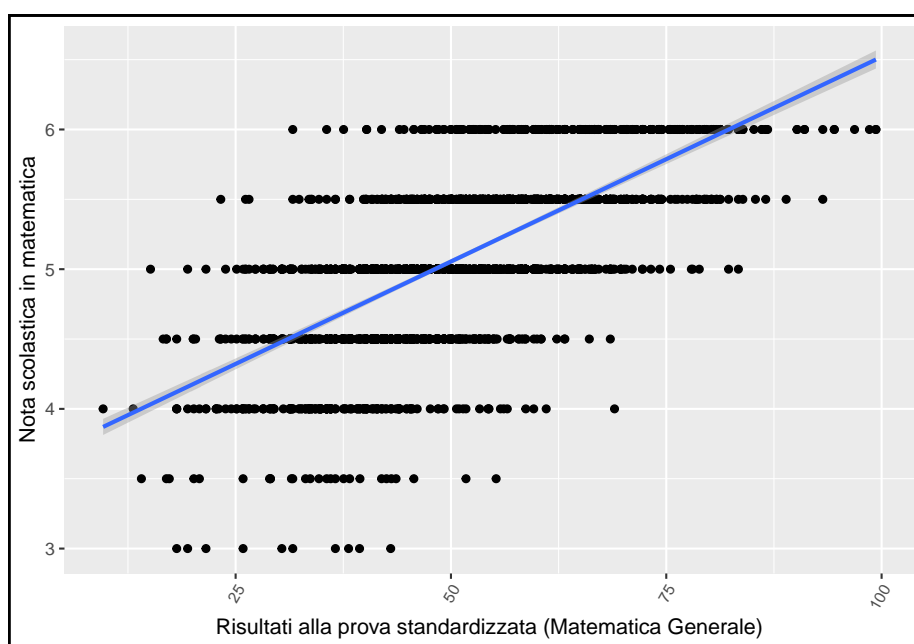


Figure 17: Relazione tra la nota scolastica in matematica e i risultati alle prove standardizzate

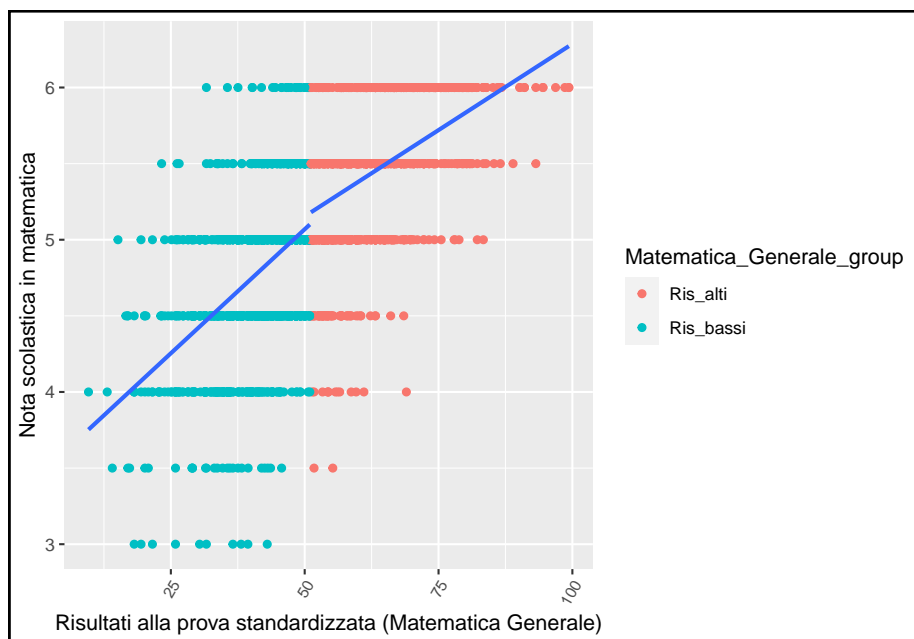


Figure 18: (#fig:NoteMatePlot_LOW)Relazione tra la nota scolastica in matematica e i risultati alle prove standardizzate in funzione dei risultati alle prove

inferiori alla mediana alle prove, la correlazione tra le due variabili è di .45, mentre è di .44 per gli allievi che hanno ottenuto dei risultati superiori alla mediana. Non è dunque possibile evidenziare dei *patterns* di risultati differenti in funzione di questi due gruppi.

4 Analisi dei risultati della prova di Italiano

4.1 Informazioni generali

Come indicato precedentemente (Capitolo 1) per l'italiano si è deciso di valutare i due ambiti di competenza Leggere e Scrivere. Ricordiamo che gli items presenti permettono di testare le seguenti dimensioni: Parte base (i.e., esercizi che toccano l'ambito di competenza Leggere, le Inferenze e il Lessico); Punteggiatura; Ortografia; Metacognizione.

Come per la matematica si è proceduto utilizzando una logica di analisi “ad imbuto” che permette di approfondire ed apprezzare le analisi con un livello crescente di dettaglio. Per quel che riguarda le informazioni contestuali (e.g., numero di allievi nei quattro ispettorati; grado di urbanizzazione, ecc.) il lettore può riferirsi al Capitolo 3] del presente rapporto.

Anche per le prove di italiano sono stati inizialmente analizzati i punteggi ottenuti nei singoli traguardi e la loro media. Anche in questo caso il punteggio nel traguardo è stato calcolato definendo per ogni item la difficoltà sul traguardo specifico e quindi sommandone i valori. Finalmente è stato calcolato un valore definito “Italiano Generale” che tenesse conto della difficoltà dei singoli item rispetto alla prova nel suo insieme. I risultati medi sono indicati nella tabella 14 e nel grafico 19.⁹

Table 14: Risultati medi cantonali

Parte_Base	47.51
Punteggiatura	59.38
Ortografia	41.22
Metacognizione	57.15
Italiano_Generale	49.17

Come illustrato nel grafico 19, la dimensione che presenta i risultati migliori sia “Punteggiatura” ($m = 59.38$), seguita da Metacognizione ($m = 57.15$), Parte Base ($m = 47.51$) ed infine Ortografia ($m = 41.21$).

4.2 Correlazioni tra le dimensioni valutate con la prova

Come per la matematica, si è valutata la correlazione tra le diverse dimensioni testate dalla prova. I valori di correlazione sono tra medi e forti. La dimensione legata alla padronanza della Punteggiatura è quella che presenta valori

⁹Così come effettuato precedentemente con la prova di matematica, i punteggi sono stati normalizzati in modo da assumere valori compresi tra 0 e 100.

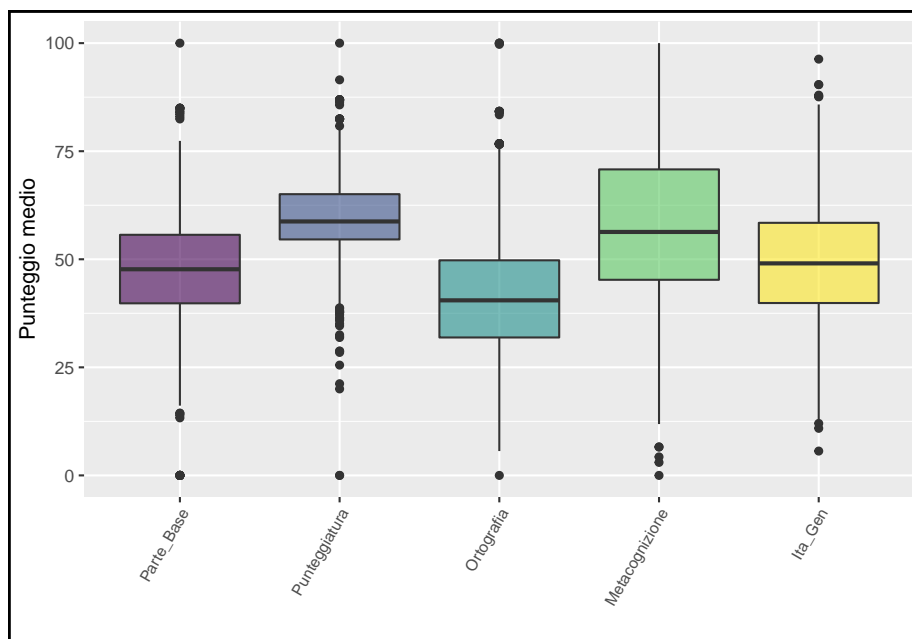


Figure 19: Risultati medi cantonali

di correlazione meno forti con le altre dimensioni testate. Da considerare che la suddetta dimensione è particolarmente circoscritta a una competenza molto puntuale e specifica.

Dal grafico 20 si può osservare la correlazione elevata tra le diverse dimensioni. Risultati questi che sottolineano che queste dimensioni siano parte del costrutto più generale valutato che corrisponde all'italiano in generale e che dunque giustifica la costituzione della dimensione Italiano generale che rappresentasse il punteggio globale del singolo allievo alla prova.

4.3 Il contesto e la scuola

Di seguito nella tabella 15 e nel grafico 21 si possono osservare i punteggi medi nei 4 ispettorati con i valori relativi alle diverse sottodimensioni analizzate.

4.3.0.1 Analisi statistiche Le analisi statistiche realizzate non evidenziano differenze statisticamente significative tra i quattro ispettorati. Così come già osservato per la matematica si può presupporre una certa uniformità nelle competenze degli allievi appartenenti a ispettorati diversi.

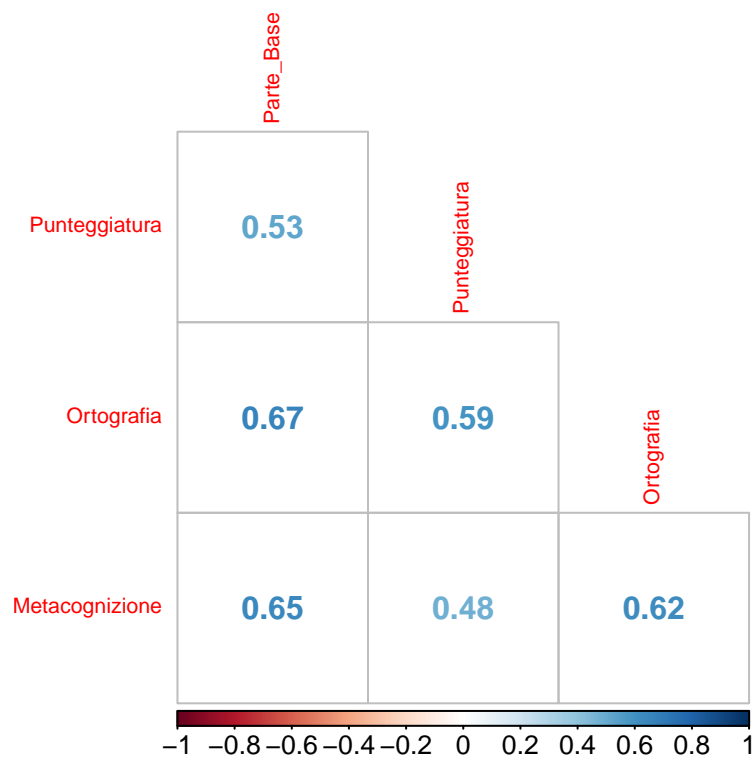


Figure 20: Correlazioni

Table 15: Risultati medi cantonali in funzione dell'ispettorato

Ispettorati	Parte_Base	Punteggiatura	Ortografia	Metacognizione	Ita_Gen
Mendrisiotto e Basso Ceresio	47.03	59.54	40.97	55.84	48.77
Luganese	47.60	59.59	41.57	58.53	49.59
Locarnese e Valli	47.57	58.33	39.67	55.85	48.02
Bellinzonese e Tre Valli	47.74	59.57	41.84	56.78	49.53

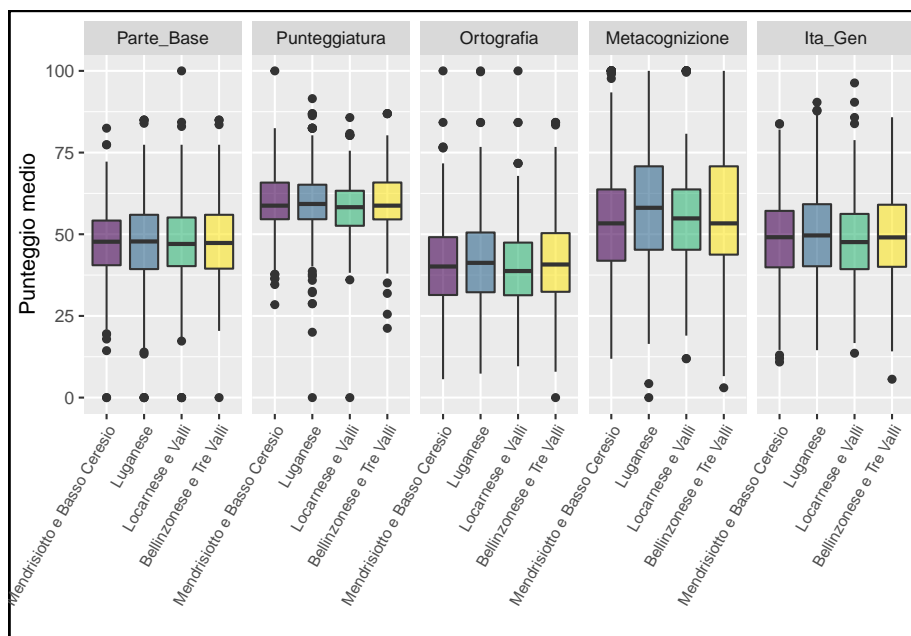


Figure 21: Risultati medi cantonali in funzione dell'ispettorato

4.4 Il grado di urbanizzazione del comune

Nella tabella 16 e nel grafico 22 si può prendere conoscenza dei risultati di competenza generale e relativi alle diverse sottodimensioni secondo il grado di urbanizzazione del comune (città nucleo; comuni della cintura; comuni rurali).

Table 16: Risultati medi cantonali in funzione del grado di urbanizzazione del comune

Urbanizzazione	Parte_Base	Punteggiatura	Ortografia	Metacognizione	Ita_Gen
città nucleo	46.29	59.98	42.02	57.93	49.27
comuni della cintura	47.24	58.99	40.48	56.19	48.47
comuni rurali	48.13	59.73	42.01	58.11	50.04

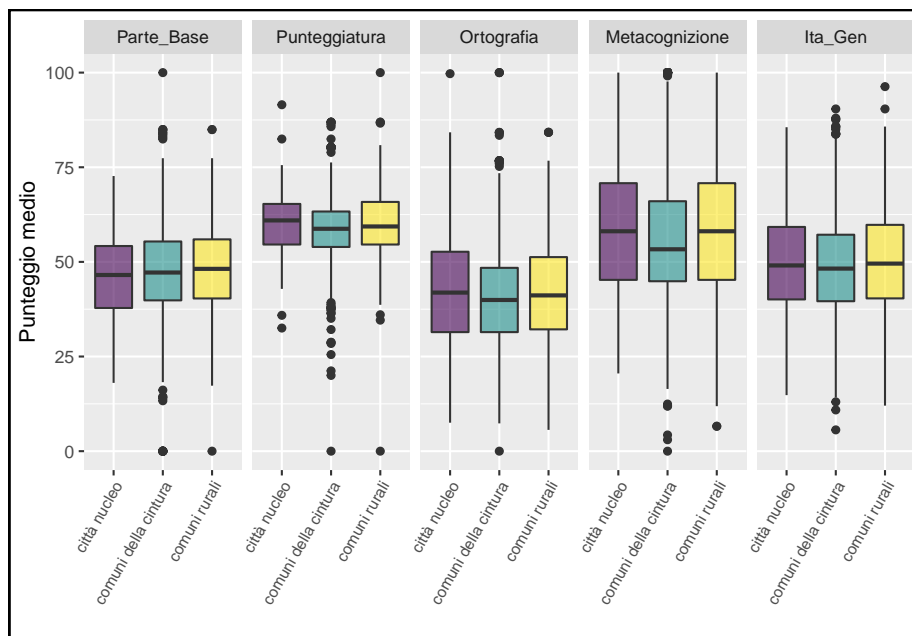


Figure 22: Risultati medi cantonali in funzione del grado di urbanizzazione del comune

4.4.1 Analisi statistiche

Le analisi realizzate mettono in evidenza differenze significative nella competenza generale in italiano ($F(2,244) = 3.42, p = .03$) in funzione dei 3 livelli di urbanizzazione. In particolare, gli allievi dei comuni rurali ($m = 50.04$) ottengono risultati migliori rispetto ai comuni della cintura ($m = 48.47$). Gli altri confronti non evidenziano differenze statisticamente significative.

Per quel che riguarda le singole dimensioni appaiono delle differenze statisticamente significative unicamente per le dimensioni di Metacognizione ed Ortografia. I risultati evidenziano dei *pattern* simili a quelli identificati per italiano generale con gli allievi dei comuni rurali che ottengono migliori risultati rispetto ai comuni della cintura. I dettagli delle analisi realizzate possono essere consultate nell'Appendice Italiano: Analisi di varianza (ANOVA) - Urbanizzazione.

4.5 La classe

4.5.1 Dimensione della classe o numero di allievi

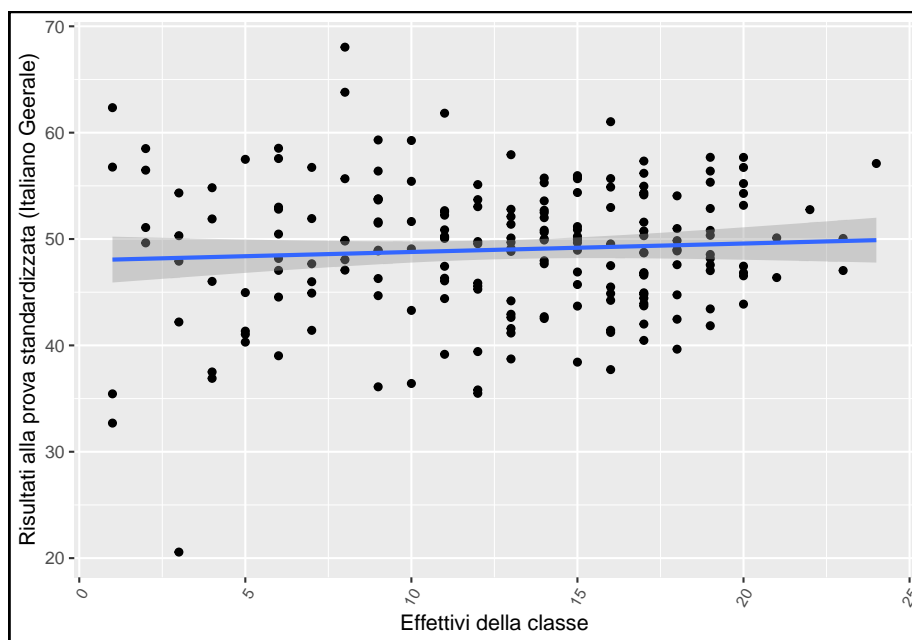


Figure 23: Relazione tra effettivi della classe e risultati alla prova standardizzate (Matematica Generale)

Come si è potuto osservare per la prova standardizzata di matematica, l'analisi della correlazione tra il numero di allievi per classe e il punteggio ottenuto nella prova standardizzata di italiano non evidenzia nessuna relazione (correlazione = .066), come rappresentato dalla linea blu nel grafico 23. Anche per italiano, la numerosità della classe non influenza i risultati alla prova standardizzata.

4.5.2 La pluriclasse

Nella tabella 17 e nel grafico 24 vengono riportati i punteggi medi nella dimensione Matematica Generale e nelle altre dimensioni ottenuti dagli allievi delle monoclasse (classi tradizionali) e delle pluriclassi.

4.5.2.1 Analisi statistiche Le analisi realizzate mostrano come non vi siano differenze statisticamente significative in funzione del tipo di classe sia

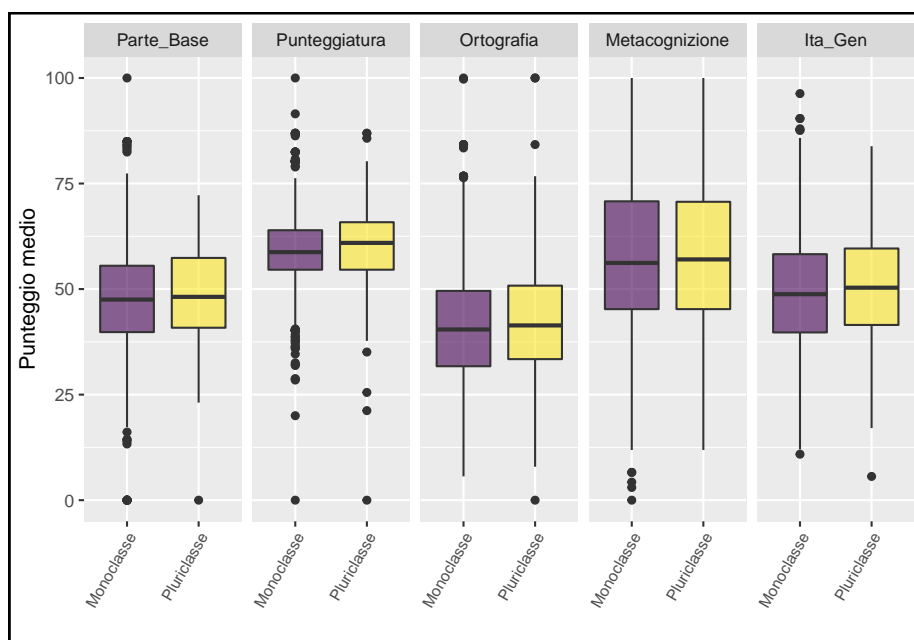


Figure 24: Punteggi medi degli allievi nelle diverse dimensioni in funzione del tipo di classe

Table 17: Punteggi medi degli allievi nelle diverse dimensioni in funzione del tipo di classe

Classe	Parte_Base	Punteggiatura	Ortografia	Metacognizione	Ita_Gen
Monoclasse	47.43	59.27	41.04	57.23	49.03
Pluriclasse	48.26	60.36	42.76	56.45	50.34

per quel che congerne i risultati alle singole dimensioni sia per quel che riguarda i risultati generali in italiano.

4.6 L'allievo

4.6.1 La nazionalità

La tabella 18 e il grafico 25 riportano i punteggi medi ottenuti dagli allievi di nazionalità svizzera, italiana o altra alla dimensione di Matematica Generale e alle altre dimensioni testate.

Table 18: Punteggi medi degli allievi nelle diverse dimensioni in funzione della nazionalità

Nazionalità	Parte_Base	Punteggiatura	Ortografia	Metacognizione	Ita_Gen
Svizzera	48.13	59.56	41.46	57.36	49.55
Italiana	47.66	59.63	42.79	58.60	50.49
Altra	44.32	58.27	38.53	54.73	46.00

4.6.1.1 Analisi statistiche Le analisi di varianza effettuate permettono di identificare delle differenze di media statisticamente significative in italiano generale, così come per ognuna delle dimensioni considerate. In modo sistematico le analisi evidenziano come gli allievi di nazionalità svizzera e gli allievi di nazionalità italiana ottengano risultati migliori rispetto agli allievi di altre nazionalità.¹⁰. Mentre non si rilevano però differenze statisticamente significative tra gli allievi di nazionalità svizzera e quelli di nazionalità italiana. I dettagli delle analisi realizzate possono essere consultati nell'Appendice Italiano: Analisi di varianza (ANOVA) - Nazionalità.

¹⁰Come si può evincere dalle analisi statistiche dettagliate presentate in appendice, per quel che riguarda la dimensione **Punteggiatura** non sussistono differenze tra gli allievi di origine italiana e gli allievi di altre nazionalità

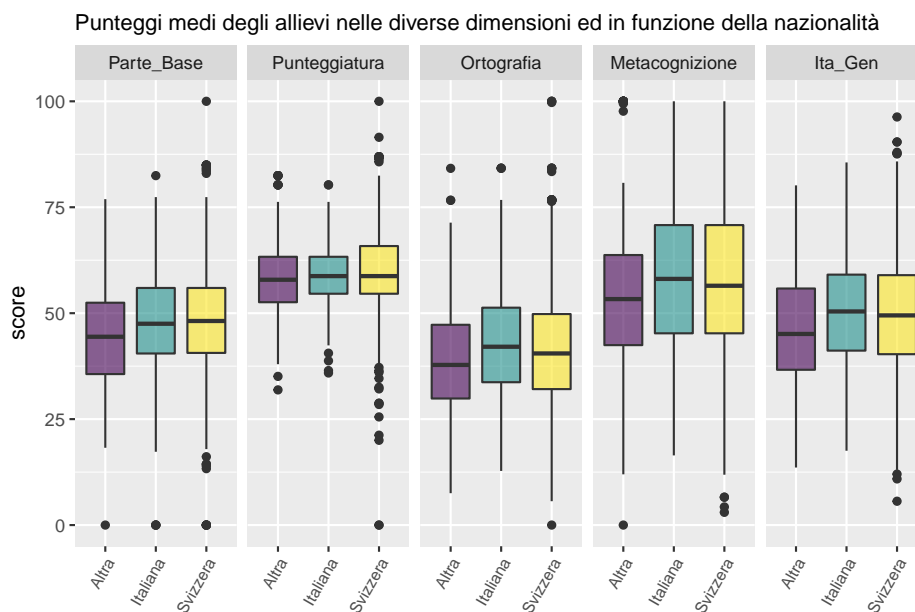


Figure 25: Punteggi medi degli allievi nelle diverse dimensioni in funzione della nazionalità

4.6.2 La Lingua madre

Nella tabella 19 e nel grafico 26 si possono visualizzare i punteggi medi ottenuti nella varie dimensioni ed in funzione della lingua madre degli allievi (i.e., italiano vs. altra lingua).

Table 19: Punteggi medi degli allievi nelle diverse dimensioni ed in funzione della lingua madre

Lingua Madre	Parte_Base	Punteggiatura	Ortografia	Metacognizione	Ita_Gen
Italiano	48.77	59.82	42.11	58.30	50.35
Altra lingua	43.44	57.96	38.33	53.43	45.31

4.6.2.1 Analisi statistiche In modo sistematico è possibile osservare come per ognuna delle dimensioni considerate, gli allievi di lingua madre italiana ottengano punteggi più elevati rispetto agli allievi con un'altra lingua madre. I dettagli delle analisi statistiche realizzate possono essere consultati nell'Appendice Italiano: Test t di Student - La lingua madre.

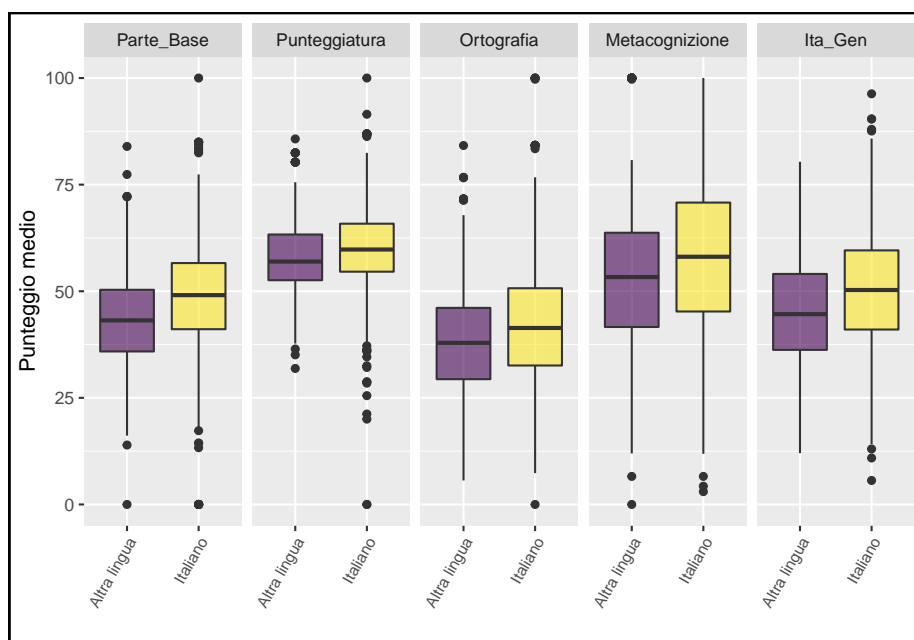


Figure 26: Punteggi medi degli allievi nelle diverse dimensioni in funzione della lingua madre

4.6.3 Il Genere

Nella tabella 20 e nel grafico 27 seguente riportano i risultati ottenuti in Italiano generale e nelle singole dimensioni in funzione del genere degli allievi.

Table 20: Punteggi medi degli allievi nelle diverse dimensioni in funzione del genere

Genere	Parte_Base	Punteggiatura	Ortografia	Metacognizione	Ita_Gen
M	46.51	58.26	39.64	55.87	47.52
F	48.51	60.51	42.79	58.43	50.81

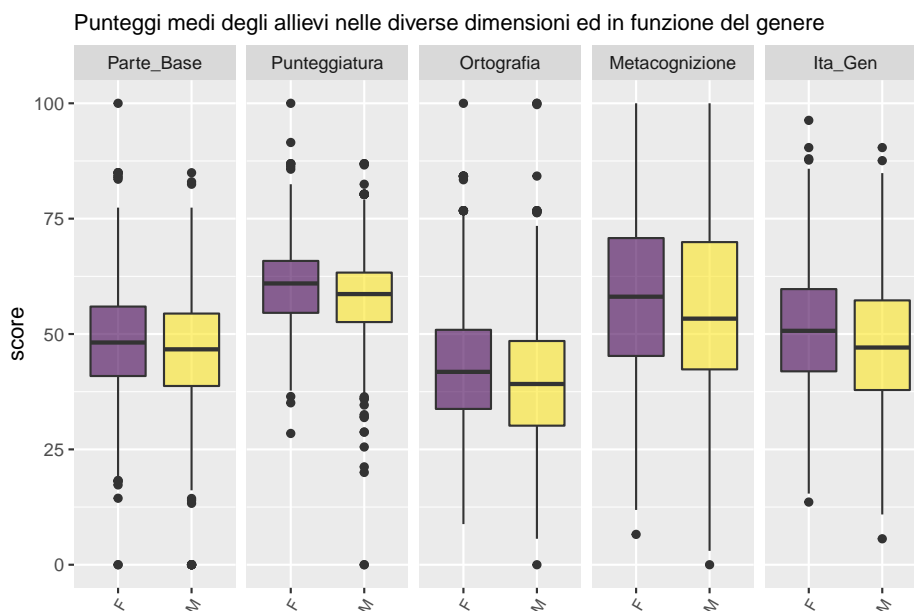
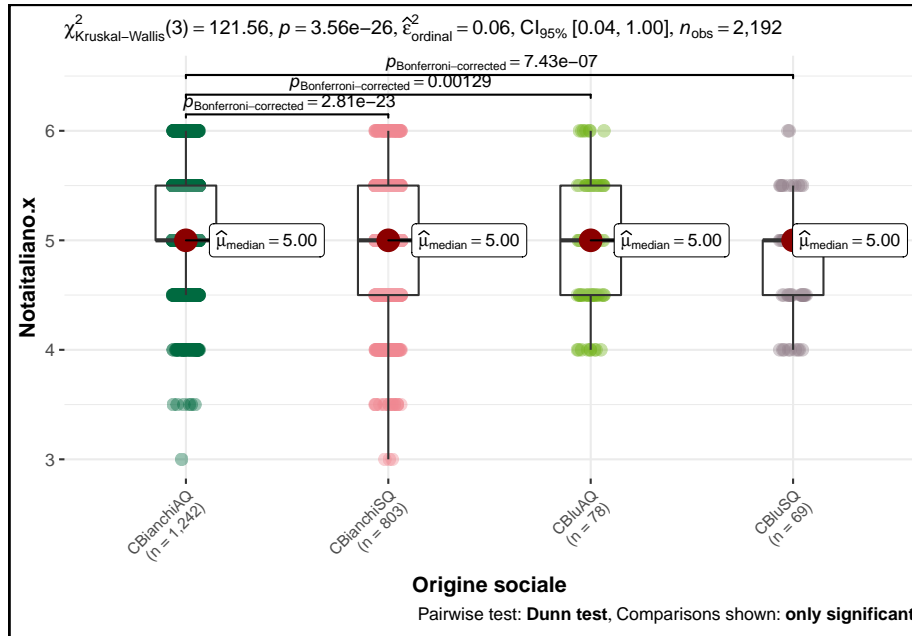


Figure 27: Punteggi medi degli allievi nelle diverse dimensioni in funzione del genere

4.6.3.1 Analisi statistiche Per quanto concerne la competenza generale in italiano, così come per le singole dimensioni appare in modo sistematico come siano le femmine ad ottenere risultati migliori rispetto ai maschi. Si ricorda come per le prove standardizzate di matematica i risultati mostrassero dei risultati opposti, con gli allievi maschi che ottengono risultati migliori rispetto alle femmine. I dettagli delle analisi statistiche realizzate possono essere consultati nell'Appendice Italiano: Test t di Student - Il genere.

4.6.4 Origine Sociale

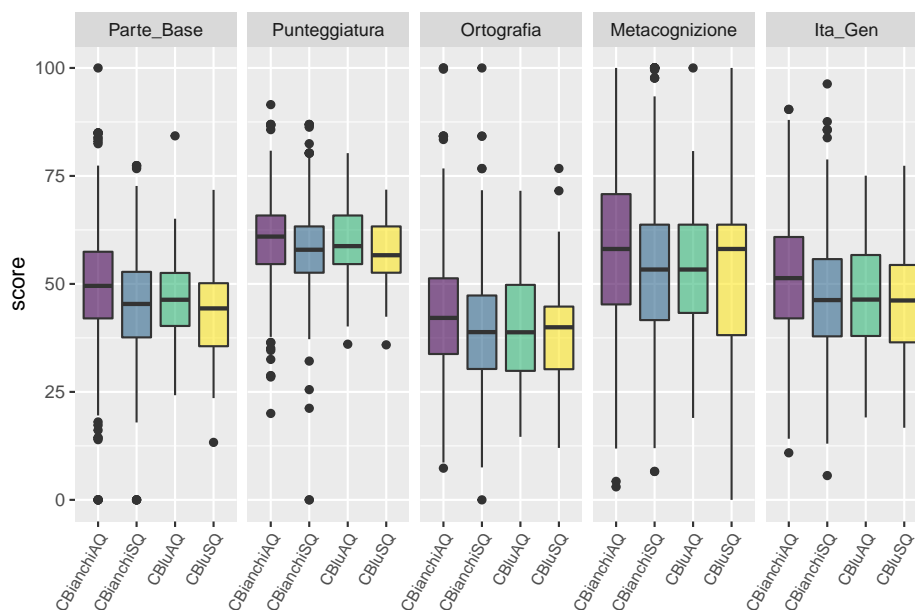
A differenza di quanto osservato per le note scolastiche in matematica (Capitolo 3.6.4), l'analisi condotta non permette di evidenziare differenze statisticamente significative rispetto alle note scolastiche in italiano in funzione dell'origine sociale degli allievi come illustrato dal grafico ??.



La tabella 21 ed il grafico ?? seguenti riportano i risultati ottenuti dagli allievi nelle varie dimensioni testate in funzione della loro origine sociale.

Table 21: Origine sociale e risultati alle prove

ISCO_Rec	Parte_Base	Punteggiatura	Ortografia	Metacognizione	Ita_Gen
CBluSQ	43.29	57.81	38.63	53.85	45.49
CBluAQ	46.23	58.86	39.54	53.83	47.07
CBianchiSQ	45.33	58.06	39.30	53.91	46.70
CBianchiAQ	49.44	60.49	42.91	59.61	51.34



4.6.4.1 Analisi statistiche Per quel che riguarda invece i risultati alle prove standardizzate, le analisi statistiche realizzate indicano come gli allievi di origine sociale “Colletti bianchi altamente qualificati” ottengono i risultati più elevati in tutte le dimensioni testate. È tuttavia interessante constatare come le differenze tra i vari gruppi presi “due a due” non siano sempre statisticamente significative come evidenziato dai dettagli delle analisi statistiche che possono essere consultati nell’Analisi di varianza (ANOVA) - Origine sociale.

4.7 Relazione tra risultati alle prove standardizzate e nota scolastica in Italiano

Verrà ora approfondita la relazione tra i risultati ottenuti alle prove standardizzate e la nota scolastica in italiano degli allievi.

Come si può osservare dal grafico 28, anche per quel che riguarda l’italiano si osserva come sia il livello di competenza generale in italiano, che quello delle singole dimensioni misurati dalle prove standardizzate correlino positivamente con i risultati scolastici degli allievi misurati dalle note scolastiche. Da sottolineare la correlazione molto elevata ($r = .94$) tra il livello di competenza generale in italiano e la dimensione di Ortografia ad indicare come l’ortografia sembri essere una componente molto importante per quel che concerne la valutazione delle competenze generali in italiano degli allievi.

Possiamo osservare nel grafico 29 la relazione positiva tra i risultati alle prove e le note in italiano degli allievi. La linea blu nel grafico rappresenta la retta di

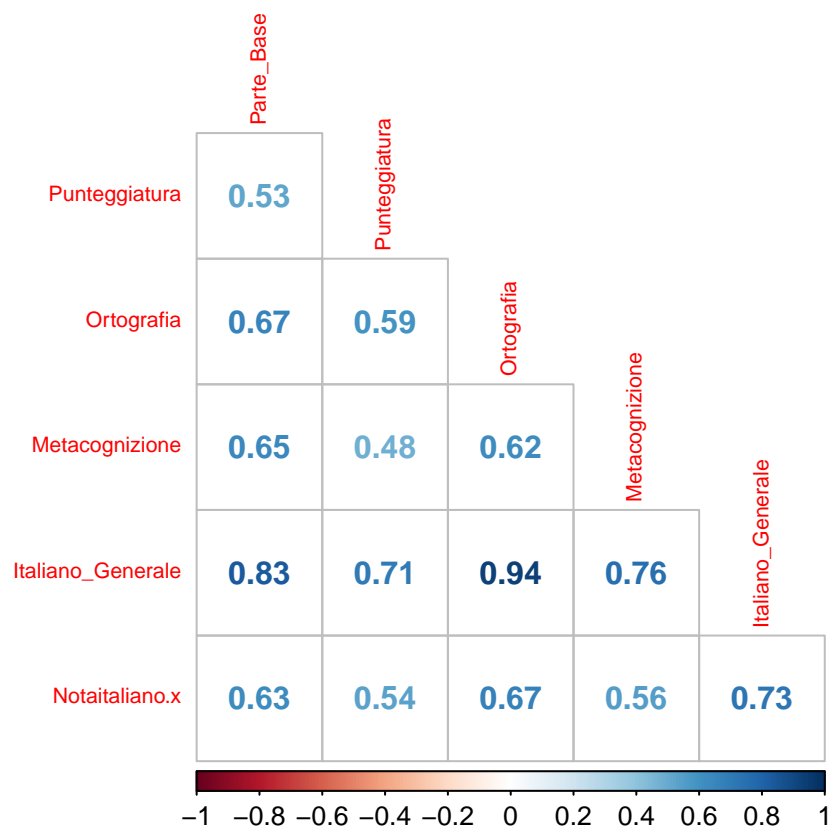


Figure 28: Correlazioni tra nota scolastica e prove standardizzate

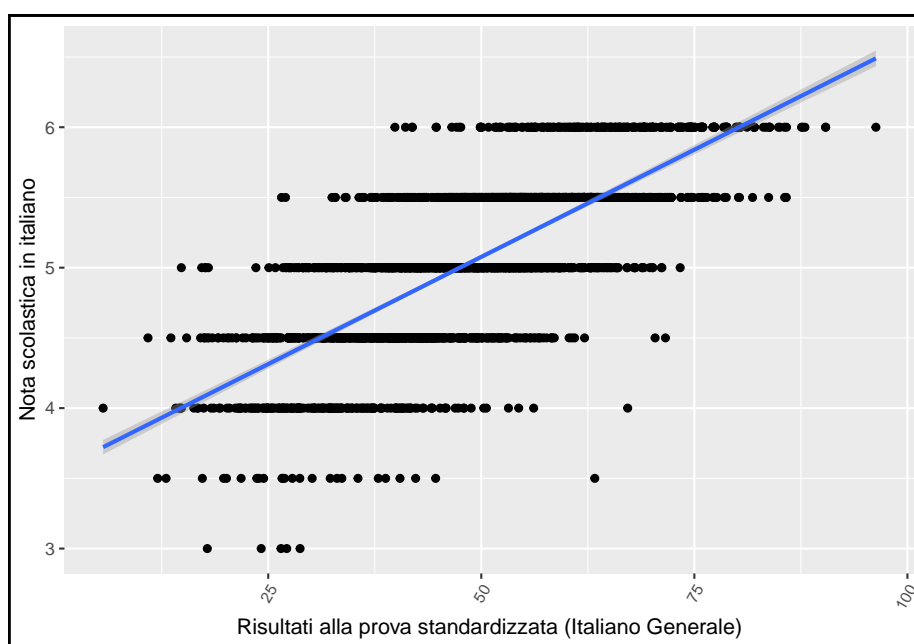


Figure 29: Relazione tra la nota scolastica in italiano e i risultati alle prove standardizzate

regressione che esprime la relazione lineare tra queste due variabili ad indicare la correlazione positiva di .73.

5 Analisi comparative - Matematica ed Italiano

In questa sezione conclusiva verranno presentate delle analisi che permettono di mettere in relazione i dati ottenuti dagli allievi nelle due discipline di italiano e matematica.

5.0.1 Commenti generali

È interessante osservare come i risultati già approfonditi e presentati all'interno del Capitolo 3 e del Capitolo 4 del presente rapporto siano molto simili tra di loro. Appare infatti come i risultati degli allievi nelle due discipline siano influenzati in modo analogo dalle variabili considerate (i.e., ispettorato; urbanizzazione; tipo di classe; caratteristiche socio-demografiche dell'allievo).

Unica differenza di rilievo quella concernente il genere dell'allievo. Si è potuto osservare come per la matematica fossero gli allievi di sesso maschile ad ottenere i risultati migliori (eccezion fatta per le dimensioni GEO_CA e GEO_SRD per le quali non vengono evidenziate differenze statisticamente significative), mentre per l'italiano sono le femmine ad ottenere i risultati migliori in tutte le dimensioni considerate.

5.0.2 Clusters o di risultati

Il grafico 30 indica le correlazioni misurate tra le varie dimensioni delle due discipline considerate.

Sulla base delle correlazioni (i.e., associazioni) tra i punteggi alle varie dimensioni misurate sia in italiano che matematica sono state realizzate delle analisi di *cluster* volte ad identificare dei raggruppamenti di variabili che permettano di massimizzarne l'omogeneità interna e che siano al contempo discriminanti tra di loro. Emergono così in modo chiaro due *clusters* o raggruppamenti di variabili che rimandano in modo netto alle due discipline. Troviamo difatti i due gruppi di variabili che riguardano l'italiano (in alto a sinistra) e alla matematica (in basso a destra).

5.0.3 Relazione tra i risultati alla prova di italiano e matematica

Il grafico seguente mostra in modo chiaro come i risultati ottenuti nelle due prove di italiano e matematica (operazionalizzati qui tramite i punteggi ottenuti nelle rispettive dimensioni generali) abbiano una forte correlazione positiva ($r = .69$) come rappresentato dalla linea blu (i.e., retta di regressione).

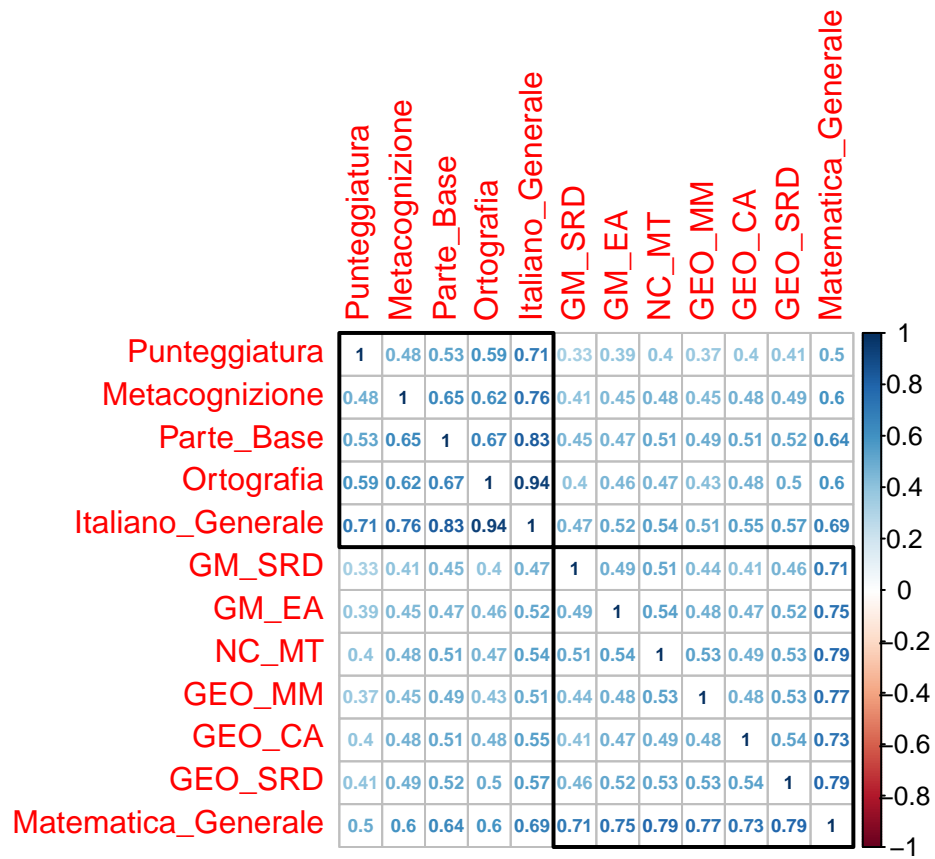


Figure 30: Correlazioni tra tutte le dimensioni testate

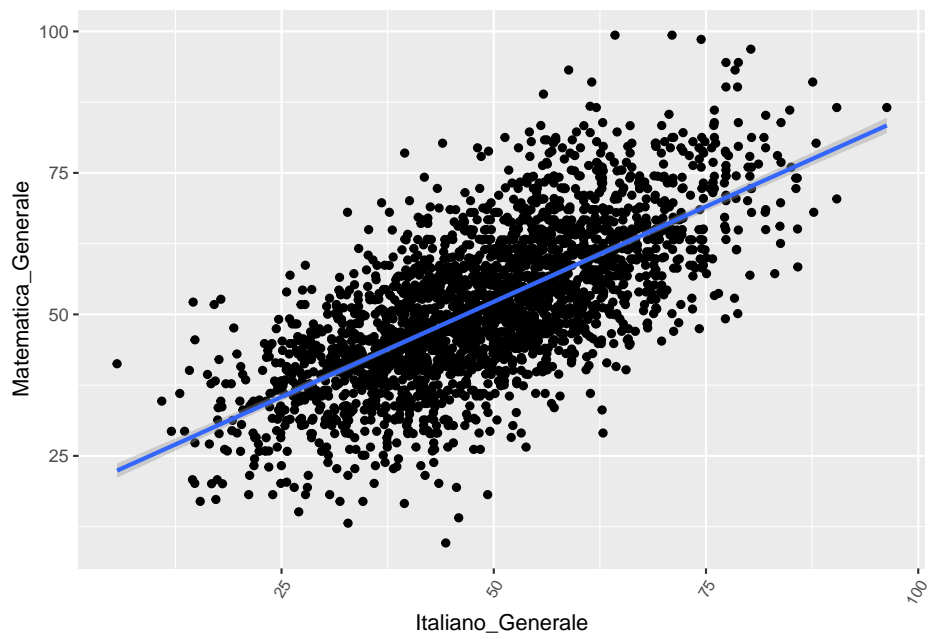


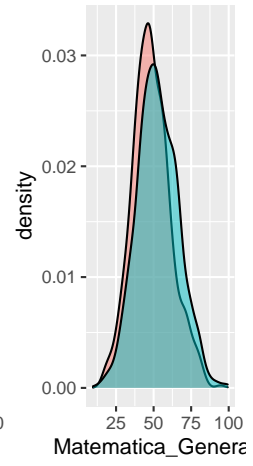
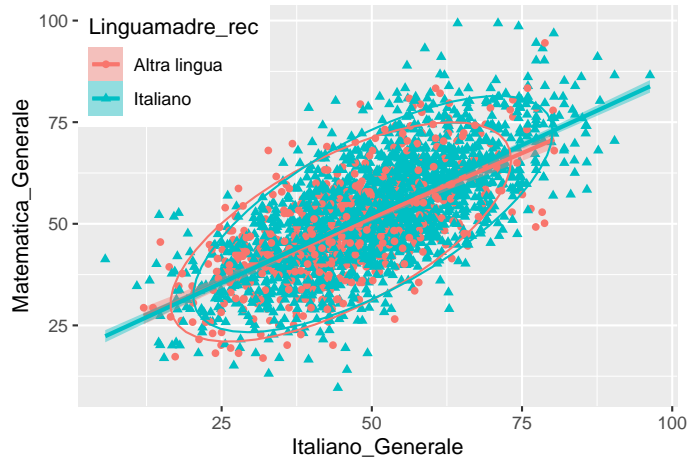
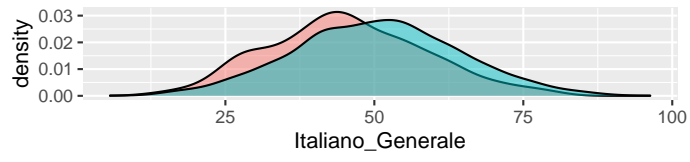
Figure 31: Relazione tra prove standardizzate in matematica ed italiano

5.0.4 Relazione tra i punteggi di matematica ed italiano in funzione della lingua madre

Come accennato in precedenza all'interno del Capitolo 3.6.2, la lingua madre è un potenziale fattore esplicativo delle differenze osservate. In particolare modo per quel che concerne il potenziale effetto delle competenze linguistiche sulle capacità di comprensione semantica di specifici esercizi di matematica.

Le analisi seguenti mostrano in modo chiaro come la relazione tra i risultati in italiano e in matematica sia leggermente più “positiva” per gli allievi di lingua madre italiana. Nel grafico si può osservare come la pendenza della linea verde (i.e., allievi di lingua madre italiana) sia leggermente più positiva di quella rossa (i.e., allievi di un'altra lingua).

Inoltre, come si può osservare dalle due distribuzioni appare come per gli studenti di lingua madre italiana i risultati siano leggermente più elevati sia in matematica che in italiano.



A Appendice: Informazioni metodologiche

In questa appendice sono presentati i dettagli delle analisi realizzate tramite delle apposite rappresentazioni grafiche.

Nello specifico, al fine di testare delle eventuali differenze a livello dei risultati medi ottenuti dagli allievi nelle prove di matematica ed italiano sono state realizzate due tipologie di analisi:

- test t di Student
- ANOVA (analisi di varianza)

In entrambi i casi, e per facilitare l'interpretazione delle analisi statistiche realizzate, i risultati delle stesse sono rappresentati tramite dei grafici che permettono di visualizzare e comprendere le differenze eventuali di media osservate. Di seguito verranno presentate succintamente le due tipologie di analisi realizzate e verranno fornite le informazioni necessarie ad una corretta lettura e interpretazione dei grafici.

A.1 Test t di Student

Il **test t di Student** a campioni indipendenti permette di testare se le differenze medie tra **due** gruppi di allievi (e.g., testare la differenza di risultati tra maschi vs. femmine). Il test statistico permette dunque di identificare delle differenze di media che non siano imputabili al campionamento effettuato¹¹.

A.1.1 Lettura del grafico

Nell'immagine 32 è possibile identificare i seguenti "parametri" di interesse: **Valori medi misurati; p value; taglia dell'effetto.**

- 1) **Valori medi misurati:** indicano i valori misurati nei vari gruppi.
- 2) **p value:** generalmente il risultato di un test statistico viene valutato rispetto al valore soglia $\alpha = .05$. Se il valore della *p value* è inferiore a $.05$, il test viene dichiarato statisticamente significativo. A questo proposito è importante sottolineare che **un risultato statisticamente significativo non implica che quanto osservato (in questo caso delle differenze**

¹¹Va sottolineato come, nonostante il fatto che si sia proceduto a somministrare le prove all'intera popolazione di allievi di quinta elementare, nel presente rapporto sono stati presentati ed analizzati unicamente i dati di un campione, ossia gli allievi che hanno partecipato ad entrambe le sessioni di somministrazione e risposto ad entrambi i fascicoli di matematica e italiano.

di media) sia “importante” (cf. punto seguente rispetto alla nozione di *effect size*).¹²

- 3) **Hedges g**: Per poter quantificare correttamente i risultati è dunque quanto meno necessario interessarsi a delle misure appropriate che permettano di quantificare l'importanza delle differenze osservate: la taglia dell'effetto. Esistono diverse tipologie di misure della taglia dell'effetto (che dipendono principalmente dal test statistico utilizzato), ed in questo caso si è optato per la misura chiamata *G di Hedges* che nell'ambito del test *t* di Student, permette di quantificare la differenza di media osservata.¹³

Il valore per il *G di Hedges* varia tra 0 e 1 e possono essere indicativamente interpretati come segue (Cohen, 1988):

- dei valori inferiori a .2 indicano delle minime differenze
- dei valori tra .2 e .5 delle differenze medie
- dei valori superiori a .8 delle differenze importanti

A.2 ANOVA (analisi di varianza)

L'analisi di varianza (ANOVA) viene comunemente utilizzata per testare la differenza di media tra due e più gruppi (e.g., per testare la differenza di risultati tra allievi dei comuni rurali, della cintura o nuclei). Il test statistico si compone di due parti distinte. La prima, comunemente chiamato *test omnibus* permette di testare la differenza di media a livello generale, senza però stabilire quale gruppo differisca da quale gruppo. La seconda, comunemente chiamato

¹²Un risultato statisticamente significativo (i.e., $p < .05$) indica unicamente che quanto osservato (o dei risultati ancora più estremi) abbiano meno del 5% di *chances* di “prodursi” nel caso in cui non esistessero reali differenze nella popolazione (i.e., ipotesi nulla). la *p value* corrisponde alla probabilità di ottenere un effetto almeno estremo come quello dei dati attuali, assumendo la verità dell'ipotesi nulla.

¹³È importante sottolineare come l'interpretazione di un test statistico, non possa realizzarsi unicamente tramite il meccanico utilizzo della *p value*. ? scrivono: “La significatività statistica non è equivalente alla significatività scientifica, umana o economica. Valori di *p* più piccoli non implicano necessariamente la presenza di effetti più grandi o più importanti, e valori di *p* più grandi non implicano una mancanza di importanza o anche una mancanza di effetto. Qualsiasi effetto, non importa quanto piccolo, può produrre un piccolo *p-value* se la dimensione del campione o la precisione della misurazione è abbastanza alta, e grandi effetti possono produrre *p-values* non impressionanti se la dimensione del campione è piccola o le misurazioni sono imprecise. Allo stesso modo, identici effetti stimati avranno diversi valori di *p* se la precisione delle stime differisce.” (traduzione libera).

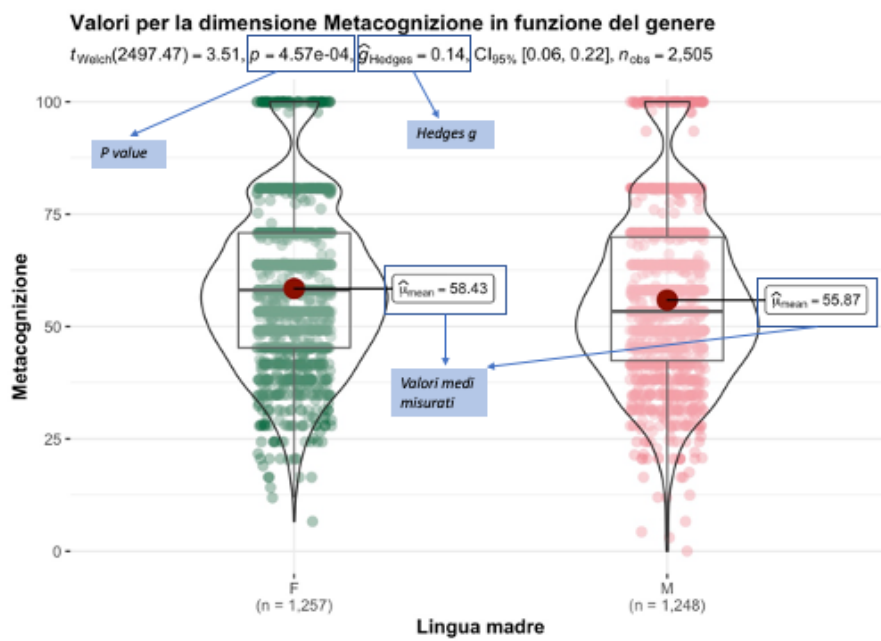


Figure 32: Esempio grafico test t di Student

*test post hoc*¹⁴ permette di testare ogni possibile confronto tra i gruppi presi “due a due” (e.g., comuni rurali vs. comuni della cintura; comuni della cintura vs nuclei; comuni rurali vs. nuclei).

A.2.1 Lettura del grafico

In modo simile a quanto descritto per il grafico relativo al test *t* di Student, nell’immagine 33 è possibile identificare i seguenti “parametri” di interesse: **Valori medi misurati; p value; taglia dell’effetto.**

Inoltre nel grafico è pure possibile identificare (cf. ellissi gialla) quali *test posthoc* siano statisticamente significativi. A titolo di esempio, nel grafico 33 si può osservare come sussistano delle differenze statisticamente significative tra i comuni della città nucleo ed i comuni della cintura, così come tra i comuni delle città nucleo ed i comuni rurali. Le differenze tra i comuni della cintura ed i comuni rurali non sono invece statisticamente significative.

Per quel che riguarda la **taglia dell’effetto**, nel caso dell’analisi di varianza corrisponde al parametro **eta parziale al quadrato** (η_p^2), che a discapito del nome indica semplicemente la percentuale di varianza dei valori misurati in funzione dall’appartenenza al gruppo. Cohen (1988) fornisce i seguenti valori indicativi che permettono di apprezzarne il significato.

- dei valori inferiori a .01 indicano delle minime differenze
- dei valori tra .01 e .06 delle differenze medie
- dei valori superiori a .14 delle differenze importanti

¹⁴Va rilevato come esistano diverse tipologie di test posthoc che rispondono ad esigenze specifiche. In questo caso si è optato per utilizzare un test di Bonferroni che permette di adattare la soglia di significatività statistica in base al numero di confronti realizzati.

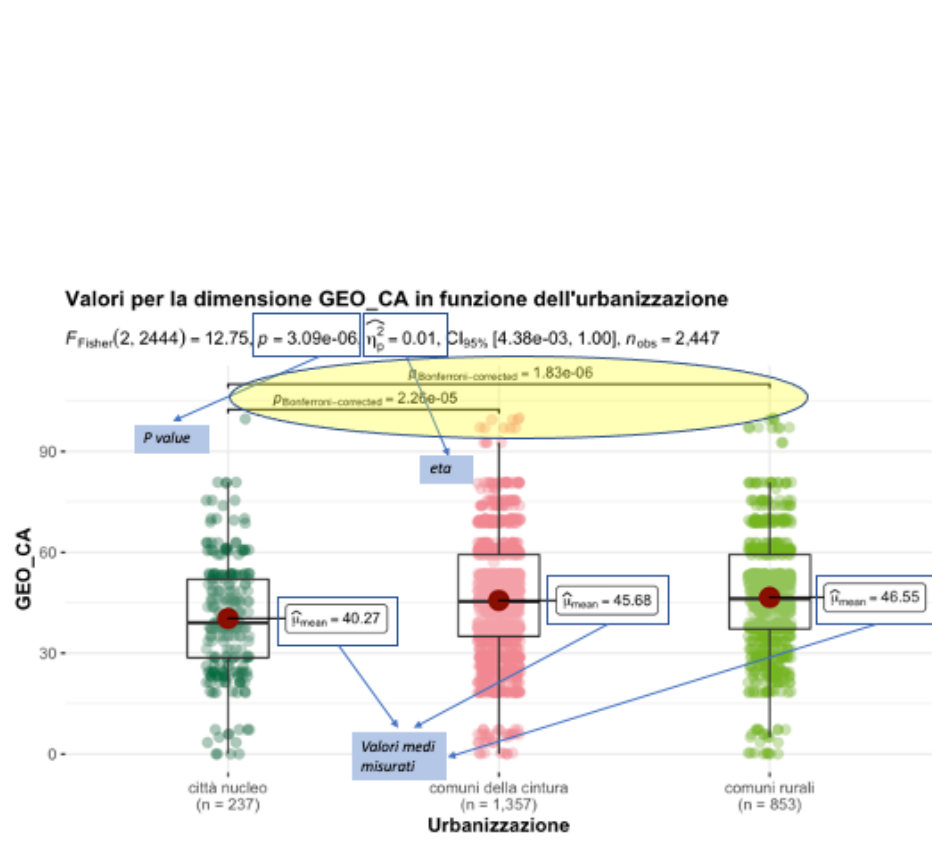


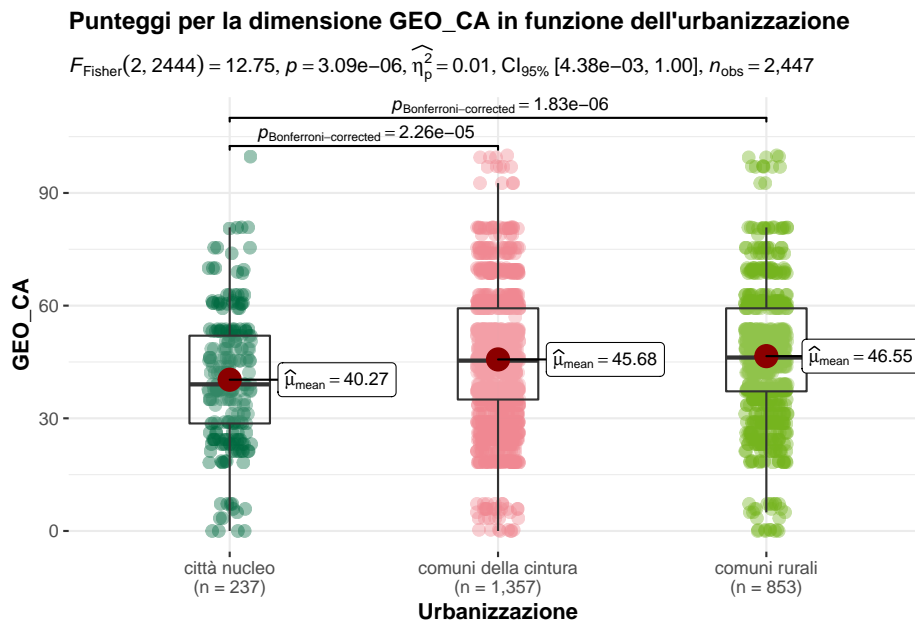
Figure 33: Esempio grafico ANOVA

B Appendice Matematica

Per le informazioni relative alla lettura e interpretazione dei seguenti grafici si invita il lettore a consultare l'appendice metodologica.

B.1 Analisi di varianza (ANOVA) - Urbanizzazione

Rispetto all'urbanizzazione dei comuni, appaiono delle differenze di risultati unicamente per la dimensione **GEO_CA**. I comuni rurali ($m = 46.55$) ed i comuni della cintura ($m = 45.68$) mostrano prestazioni più elevate rispetto ai comuni delle città nucleo ($m = 40.27$).



Ritornare al rapporto.

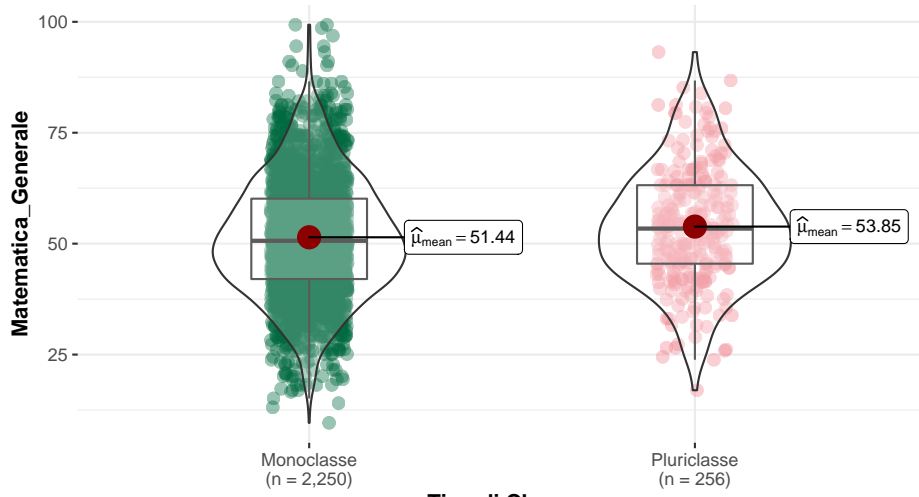
B.2 Test t di Student - La classe

Rispetto alla tipologia di classe (i.e., monoclasse vs. pluriclasse), laddove sono emerse delle differenze statisticamente significative (cf. grafici seguenti), queste mostrano in modo sistematico come gli allievi delle pluriclassi ottengano risultati migliori rispetto agli allievi delle monoclasse. Questi risultati vanno però

relativizzati in quanto le differenze sono minime, come attestato dai valori della taglia dell'effetto inferiori a .2 in valore assoluto.

Punteggi per Matematica generale in funzione del tipo di classe

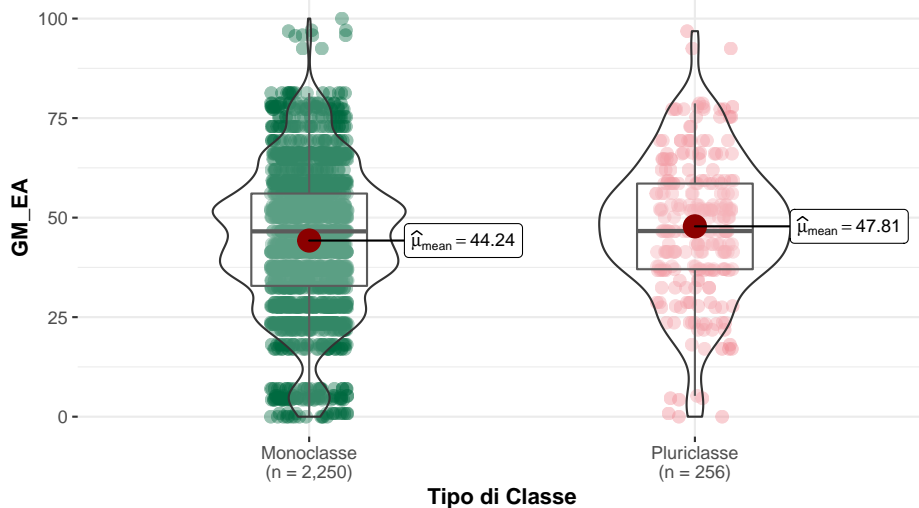
$t_{\text{Welch}}(318.94) = -2.75, p = 6.33e-03, \hat{g}_{\text{Hedges}} = -0.18, \text{CI}_{95\%} [-0.31, -0.05], n_{\text{obs}} = 2,506$



$\log_e(\text{BF}_{01}) = -0.97, \hat{\delta}_{\text{difference}}^{\text{posterior}} = -2.35, \text{CI}_{95\%}^{\text{HDI}} [-4.04, -0.63], r_{\text{Cauchy}}^{\text{JZS}} = 0.71$

Punteggi per la dimensione GM_EA in funzione del tipo di classe

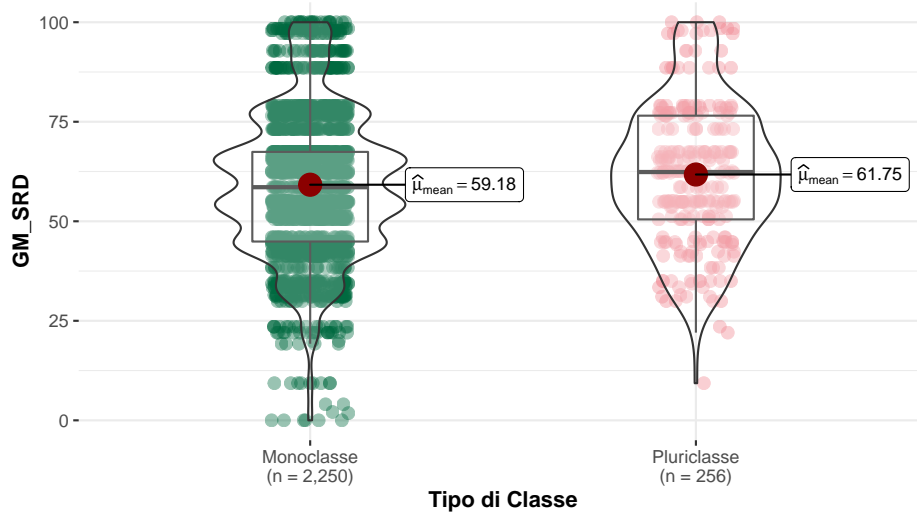
$t_{\text{Welch}}(319.7) = -3.13, p = 1.92e-03, \hat{g}_{\text{Hedges}} = -0.20, \text{CI}_{95\%} [-0.33, -0.07], n_{\text{obs}} = 2,506$



$\log_e(\text{BF}_{01}) = -1.98, \hat{\delta}_{\text{difference}}^{\text{posterior}} = -3.54, \text{CI}_{95\%}^{\text{HDI}} [-5.75, -1.21], r_{\text{Cauchy}}^{\text{JZS}} = 0.71$

Punteggi per la dimensione GM_SRD in funzione del tipo di classe

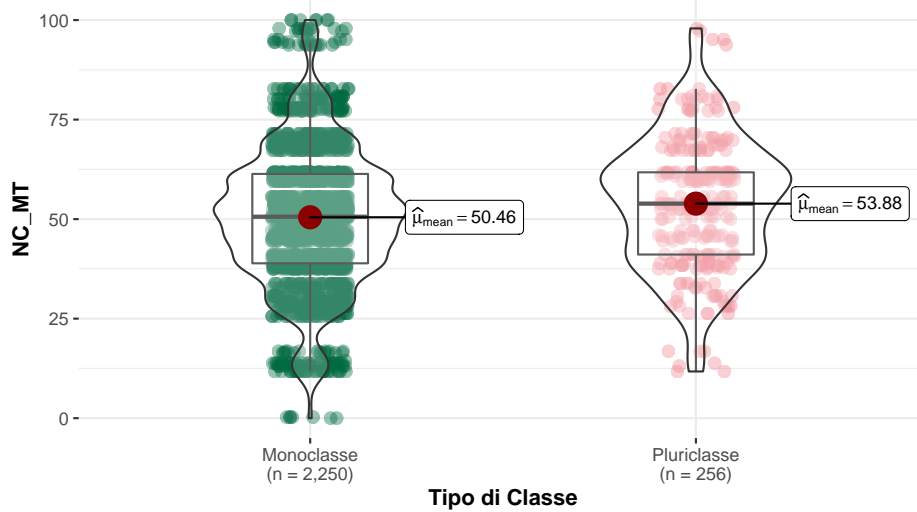
$t_{\text{Welch}}(317.25) = -2.20, p = 0.03, \hat{g}_{\text{Hedges}} = -0.14, \text{CI}_{95\%} [-0.27, -0.02], n_{\text{obs}} = 2,506$



$\log_e(\text{BF}_{01}) = 0.27, \hat{\delta}_{\text{difference}}^{\text{posterior}} = -2.63, \text{CI}_{95\%}^{\text{HDI}} [-4.99, -0.33], r_{\text{Cauchy}}^{\text{JZS}} = 0.71$

Punteggi per la dimensione NC_MT in funzione del tipo di classe

$t_{\text{Welch}}(319.86) = -3.13, p = 1.89\text{e-}03, \hat{g}_{\text{Hedges}} = -0.20, \text{CI}_{95\%} [-0.33, -0.08], n_{\text{obs}} = 2,506$



$\log_e(\text{BF}_{01}) = -1.99, \hat{\delta}_{\text{difference}}^{\text{posterior}} = -3.32, \text{CI}_{95\%}^{\text{HDI}} [-5.54, -1.16], r_{\text{Cauchy}}^{\text{JZS}} = 0.71$

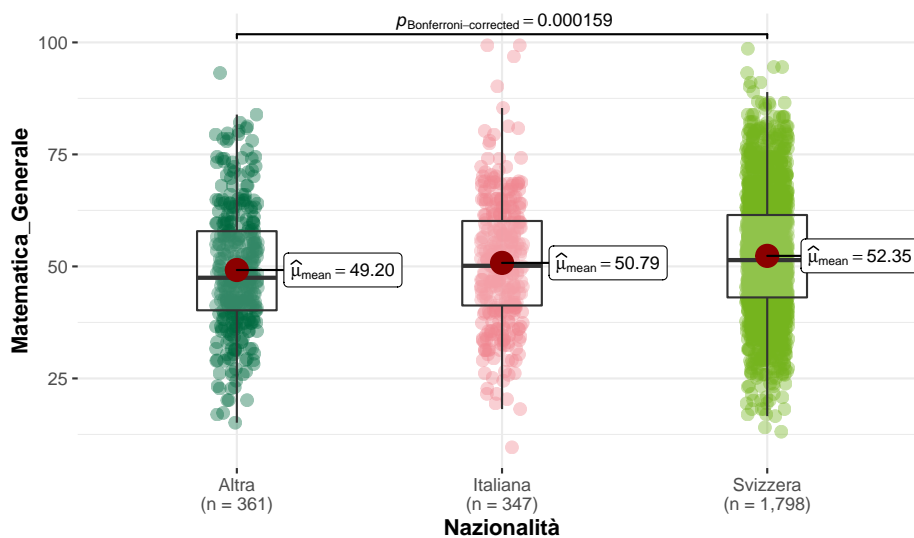
Ritornare al rapporto.

B.3 Analisi di varianza (ANOVA) - Nazionalità

Per quel che riguarda la nazionalità degli allievi, le analisi di varianza realizzate permettono di evidenziare delle differenze di risultati in alcune dimensioni (cf. grafici seguenti). Indipendentemente dalla dimensione misurata, vengono messe in evidenza in modo sistematico come gli allievi Svizzeri ottengano risultati migliori degli allievi di altre nazionalità (gli altri test *post hoc* non evidenziano differenze tra gli altri gruppi). Fa eccezione la dimensione GEO_SRD in cui si osservano delle differenze pure tra gli allievi di nazionalità italiana che ottengono risultati migliori che gli allievi di altre nazionalità. Inoltre, come si può osservare per la dimensione NC_MT, nonostante a livello di test globale (test *omnibus*) vengano rilevate delle differenze tra i tre gruppi, queste non emergono quando i gruppi vengono comparati “due a due” (i.e., test *post hoc*).

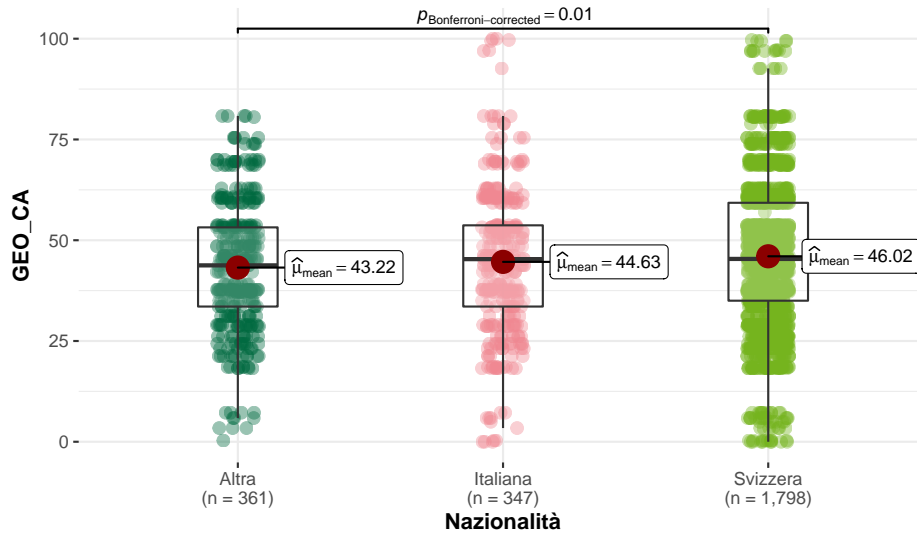
Punteggi per matematica generale in funzione della nazionalità

$$F_{\text{Fisher}}(2, 2503) = 9.08, p = 1.18e-04, \hat{\eta}_p^2 = 7.20e-03, \text{CI}_{95\%} [2.41e-03, 1.00], n_{\text{obs}} = 2,506$$



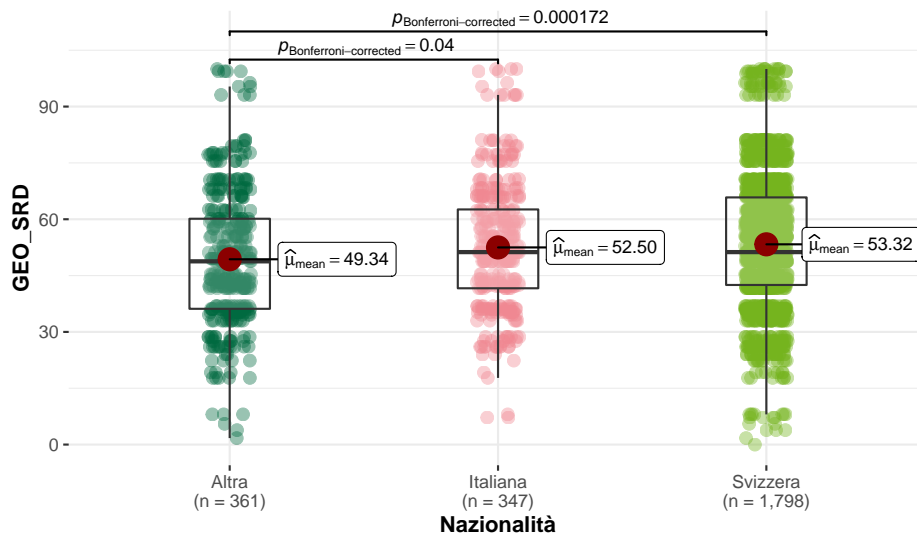
Punteggi per la dimensione GEO_CA in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 4.43, p = 0.01, \hat{\eta}_p^2 = 3.53e-03, CI_{95\%} [4.35e-04, 1.00], n_{\text{obs}} = 2,506$



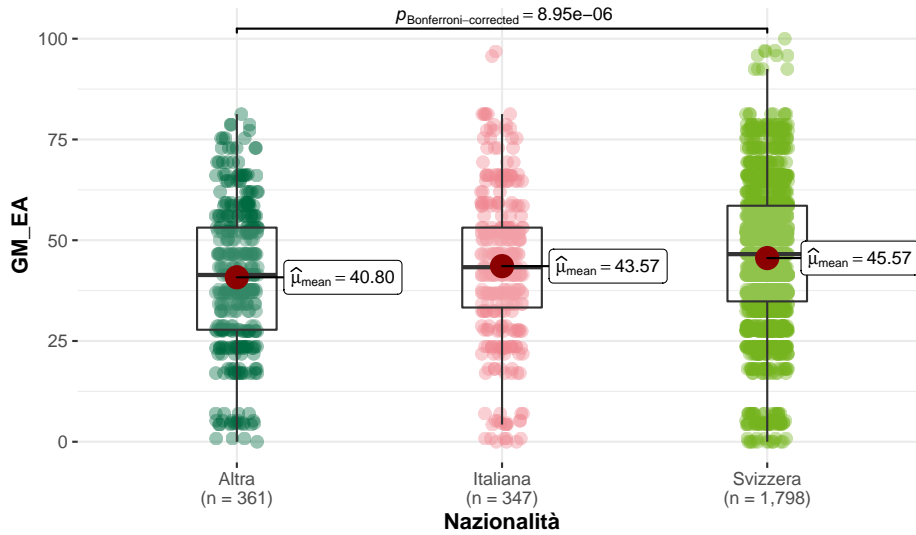
Punteggi per la dimensione GEO_SRD in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 8.13, p = 3.01e-04, \hat{\eta}_p^2 = 6.46e-03, CI_{95\%} [1.96e-03, 1.00], n_{\text{obs}} = 2,506$



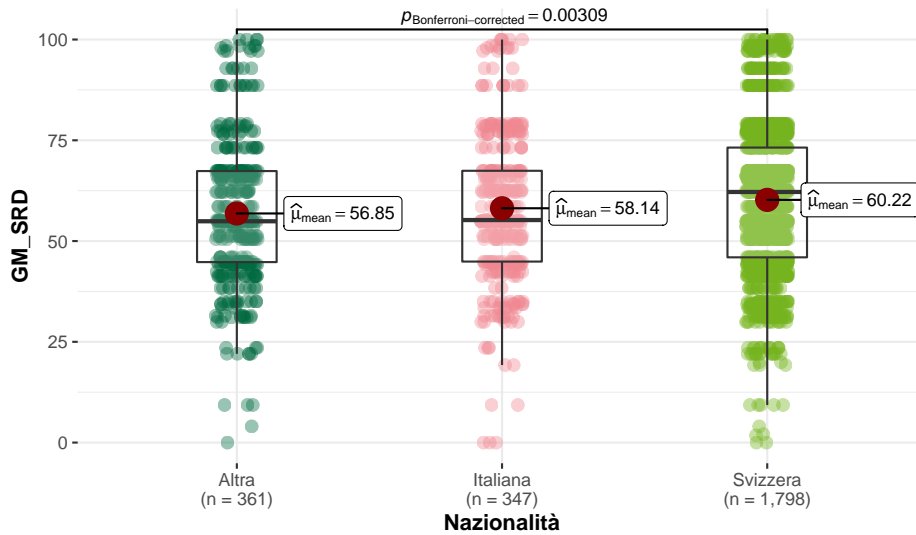
Punteggi per la dimensione GM_EA in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 11.66, p = 9.15e-06, \hat{\eta}_p^2 = 9.23e-03, \text{CI}_{95\%} [3.70e-03, 1.00], n_{\text{obs}} = 2,506$



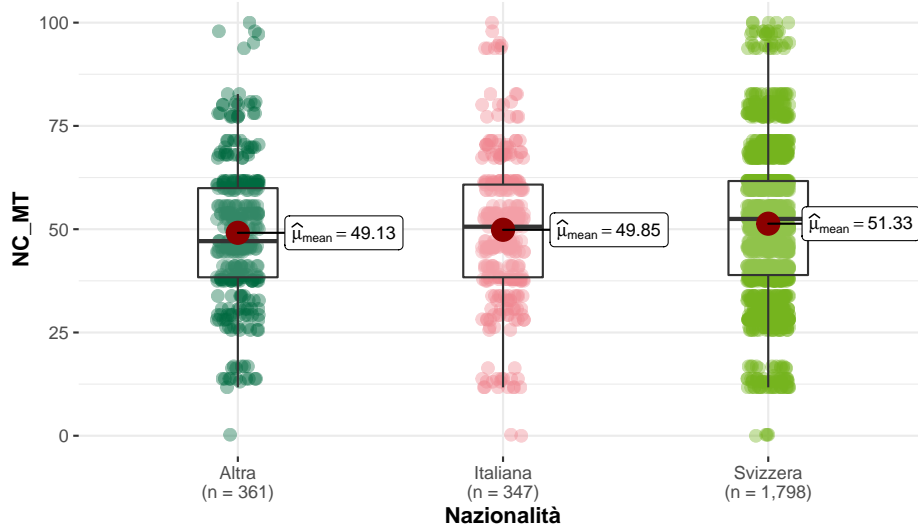
Punteggi per la dimensione GM_SRD in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 6.49, p = 1.54e-03, \hat{\eta}_p^2 = 5.16e-03, \text{CI}_{95\%} [1.23e-03, 1.00], n_{\text{obs}} = 2,506$



Punteggi per la dimensione NC_MT in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 3.18, p = 0.04, \hat{\eta}_p^2 = 2.53e-03, CI_{95\%} [4.85e-05, 1.00], n_{\text{obs}} = 2,506$



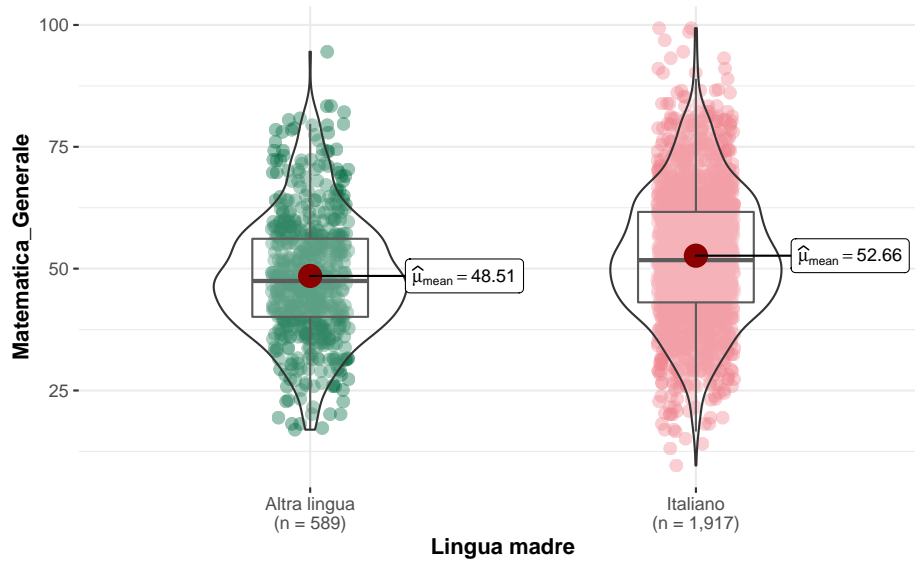
Ritornare al rapporto.

B.4 Test *t* di Student - La lingua madre

Come si può evincere dai seguenti grafici, emergono delle differenze statisticamente significative per tutte le dimensioni testate. Inoltre le differenze osservate possono essere quantificate come mediamente importanti come attestato dai valori della taglia dell'effetto $> .19$ in valore assoluto.

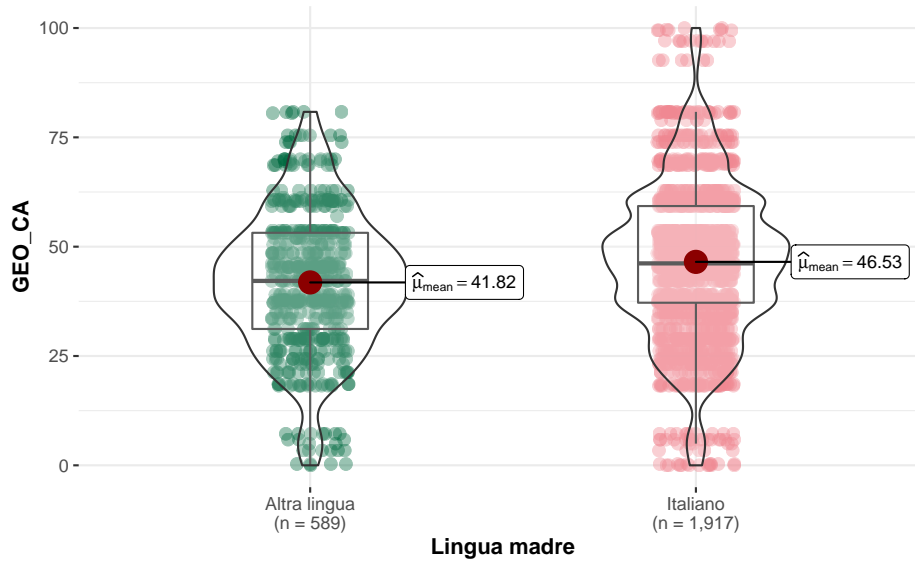
Punteggi per matematica generale in funzione della lingua madre

$t_{\text{Welch}}(1037.13) = -6.80, p = 1.73e-11, \hat{g}_{\text{Hedges}} = -0.31, CI_{95\%} [-0.41, -0.22], n_{\text{obs}} = 2,506$



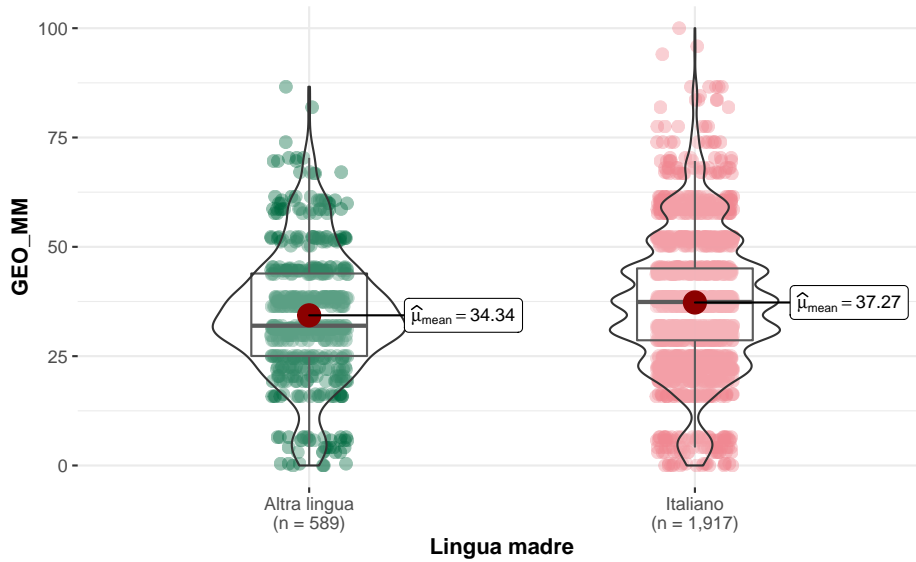
Punteggi per la dimensione GEO_CA in funzione della lingua madre

$t_{\text{Welch}}(1063.57) = -6.16, p = 1.01e-09, \hat{g}_{\text{Hedges}} = -0.28, CI_{95\%} [-0.37, -0.19], n_{\text{obs}} = 2,506$



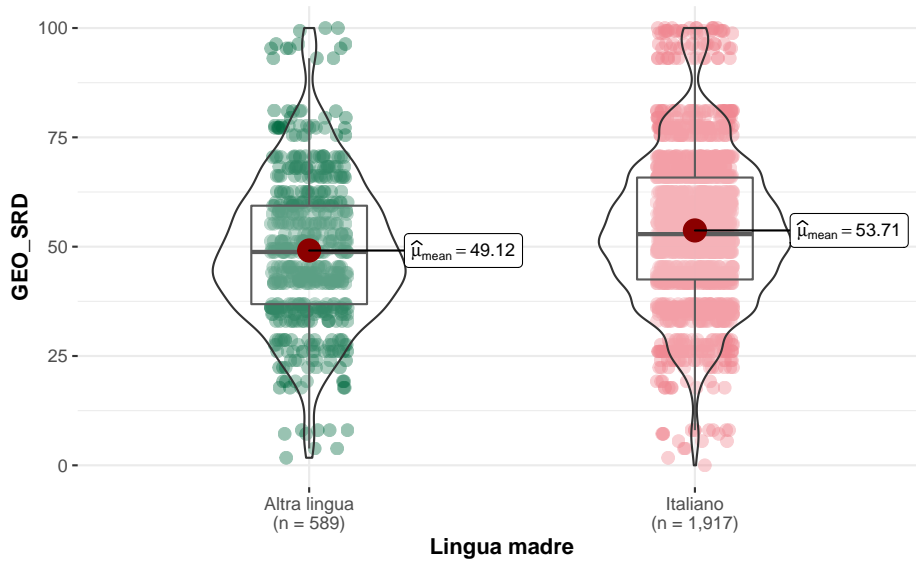
Punteggi per la dimensione GEO_MM in funzione della lingua madre

$t_{\text{Welch}}(1045.18) = -4.26, p = 2.22e-05, \hat{g}_{\text{Hedges}} = -0.20, \text{CI}_{95\%} [-0.29, -0.11], n_{\text{obs}} = 2,506$



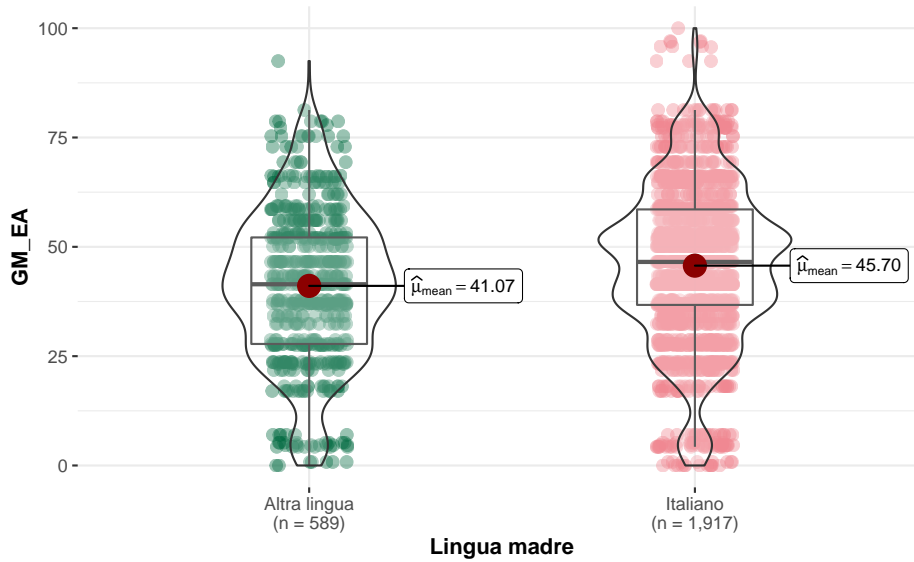
Punteggi per la dimensione GEO_SRD in funzione della lingua madre

$t_{\text{Welch}}(996.2) = -5.78, p = 9.79e-09, \hat{g}_{\text{Hedges}} = -0.27, \text{CI}_{95\%} [-0.36, -0.18], n_{\text{obs}} = 2,506$



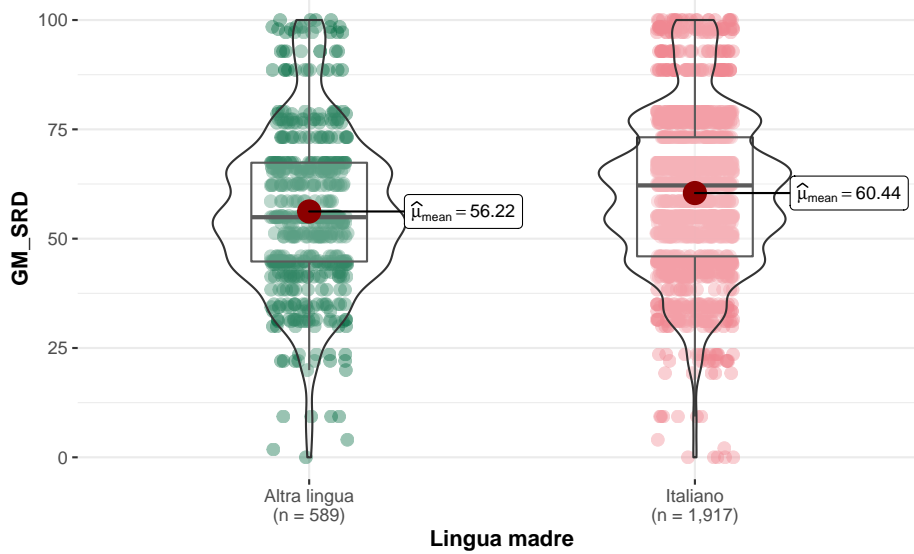
Punteggi per la dimensione GM_EA in funzione della lingua madre

$t_{\text{Welch}}(996.14) = -5.64, p = 2.25e-08, \hat{g}_{\text{Hedges}} = -0.26, \text{CI}_{95\%} [-0.36, -0.17], n_{\text{obs}} = 2,506$



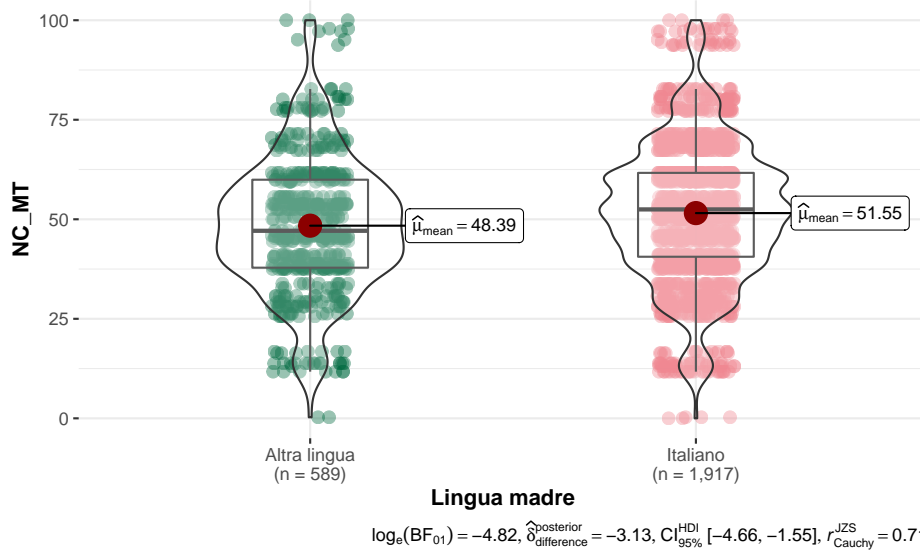
Punteggi per la dimensione GM_SRD in funzione della lingua madre

$t_{\text{Welch}}(986.38) = -5.08, p = 4.59e-07, \hat{g}_{\text{Hedges}} = -0.24, \text{CI}_{95\%} [-0.33, -0.15], n_{\text{obs}} = 2,506$



Punteggi per la dimensione NC_MT in funzione della lingua madre

$t_{\text{Welch}}(988.45) = -3.99, p = 7.08e-05, \hat{g}_{\text{Hedges}} = -0.19, \text{CI}_{95\%} [-0.28, -0.09], n_{\text{obs}} = 2,506$



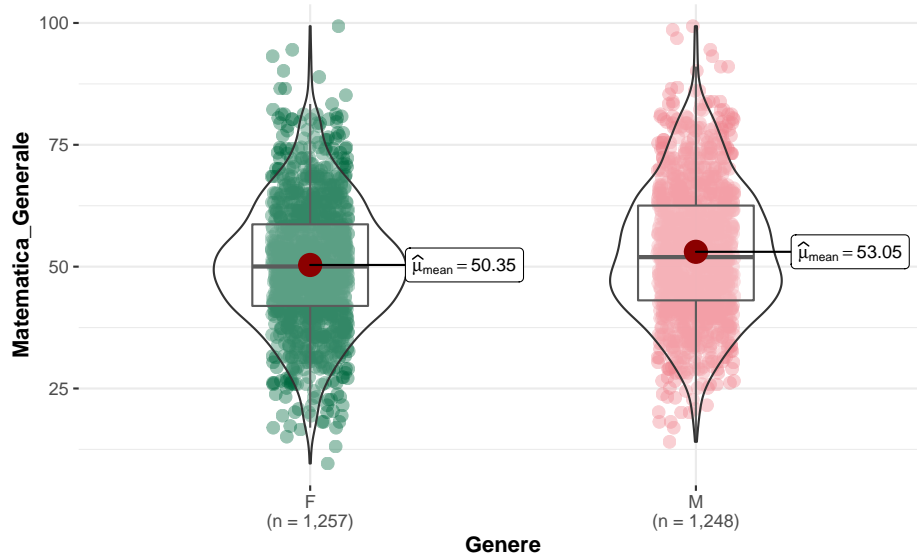
Ritornare al rapporto.

B.5 Test t di Student - Il genere

Dalle analisi realizzate, laddove sussistano delle differenze statisticamente significative, si può osservare come in modo sistematico i maschi ottengano dei punteggi più elevati rispetto alle femmine. Le differenze osservate sono di piccola entità, eccezion fatta per la dimensione **GM_SRD** in cui le differenze sono mediamente importanti con un valore della taglia d'effeto = .42 in valore assoluto.

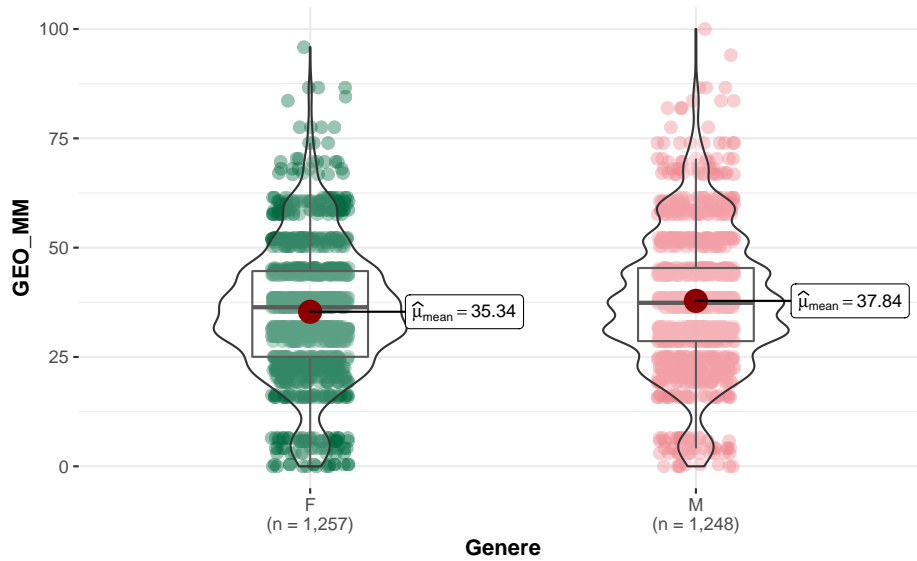
Punteggi per Matematica generale in funzione del genere

$t_{\text{Welch}}(2494.38) = -5.01, p = 5.80e-07, \hat{g}_{\text{Hedges}} = -0.20, CI_{95\%} [-0.28, -0.12], n_{\text{obs}} = 2,505$



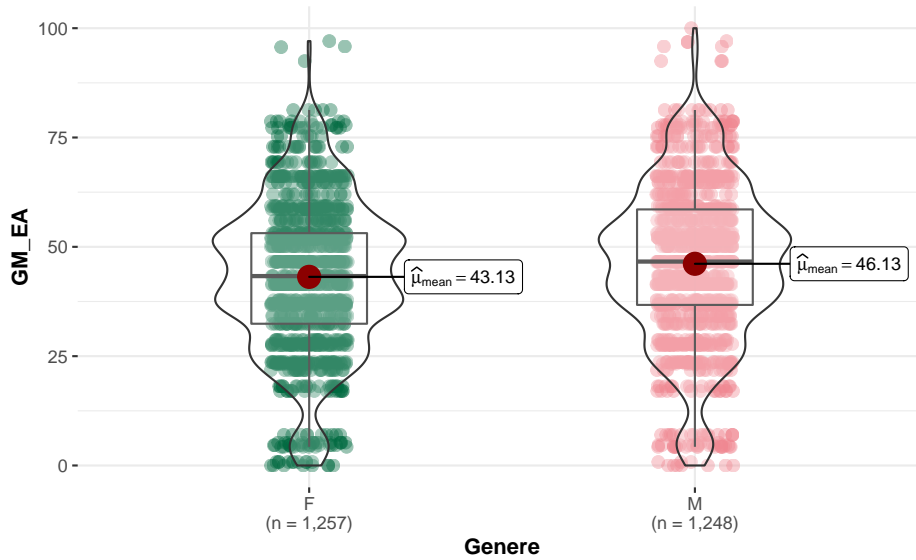
Punteggi per la dimensione GEO_MM in funzione del genere

$t_{\text{Welch}}(2494.82) = -4.09, p = 4.42e-05, \hat{g}_{\text{Hedges}} = -0.16, CI_{95\%} [-0.24, -0.09], n_{\text{obs}} = 2,505$



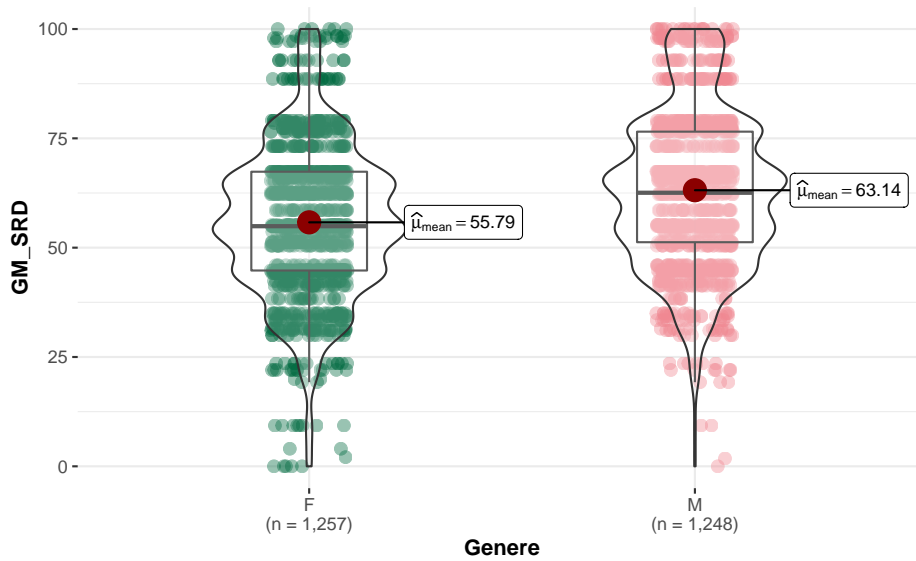
Punteggi per la dimensione GM_EA in funzione del genere

$t_{\text{Welch}}(2497.5) = -4.24$, $p = 2.35e-05$, $\hat{g}_{\text{Hedges}} = -0.17$, $CI_{95\%} [-0.25, -0.09]$, $n_{\text{obs}} = 2,505$



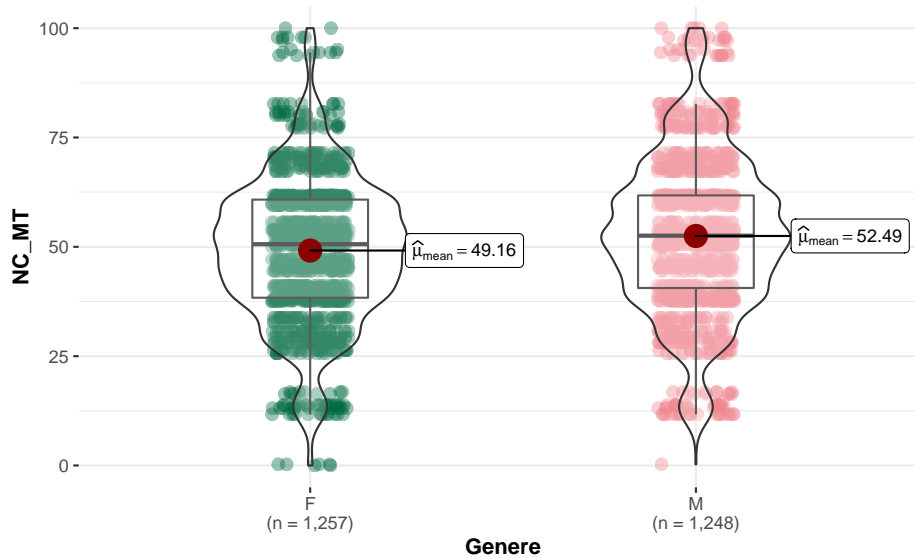
Punteggi per la dimensione GM_SRD in funzione del genere

$t_{\text{Welch}}(2499) = -10.53$, $p = 2.17e-25$, $\hat{g}_{\text{Hedges}} = -0.42$, $CI_{95\%} [-0.50, -0.34]$, $n_{\text{obs}} = 2,505$



Punteggi per la dimensione NC_MT in funzione del genere

$t_{\text{Welch}}(2486.15) = -4.92, p = 9.16e-07, \hat{g}_{\text{Hedges}} = -0.20, \text{CI}_{95\%} [-0.28, -0.12], n_{\text{obs}} = 2,505$



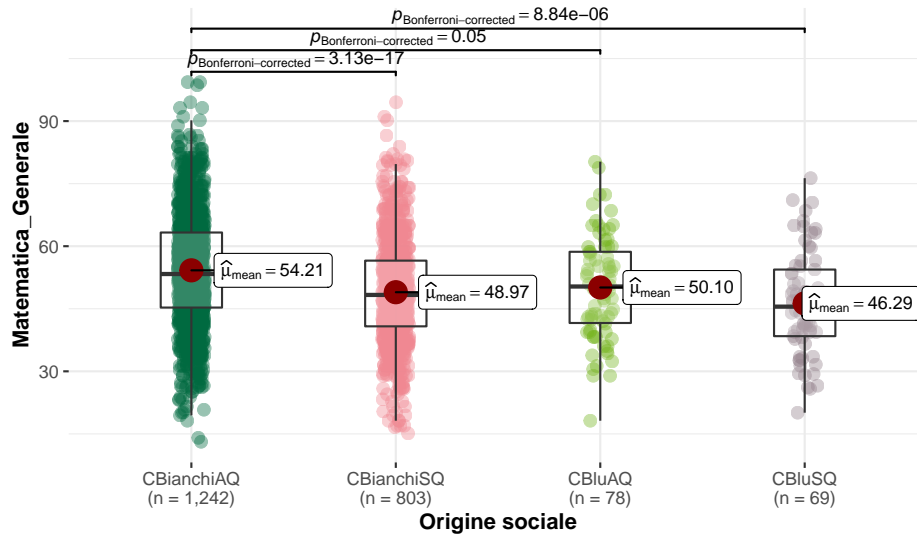
Ritornare al rapporto.

B.6 Analisi di varianza (ANOVA) - Origine sociale

Dalle analisi realizzate emerge in modo sistematico come gli allievi di origine sociale più elevata abbiano delle performance migliori. Nello specifico dai test *post hoc* realizzati si può osservare come gli allievi del gruppo “Colletti bianchi altamente qualificati” ottengano sempre dei risultati migliori rispetto ai “Colletti bianchi scarsamente qualificati” e rispetto ai “Colletti blu scarsamente qualificati”. Inoltre, per **matematica generale** appaiono delle differenze pure tra i “Colletti bianchi altamente qualificati” ed i “Colletti blu altamente qualificati”, e per la dimensione **GM_EA** risultano delle differenze tra i “Colletti bianchi scarsamente qualificati” ed i “Colletti blu scarsamente qualificati”. Da sottolineare come la misura della taglia dell’effetto evidenzi come le differenze medie osservate siano di media entità ($.02 < \eta_p^2 < .04$).

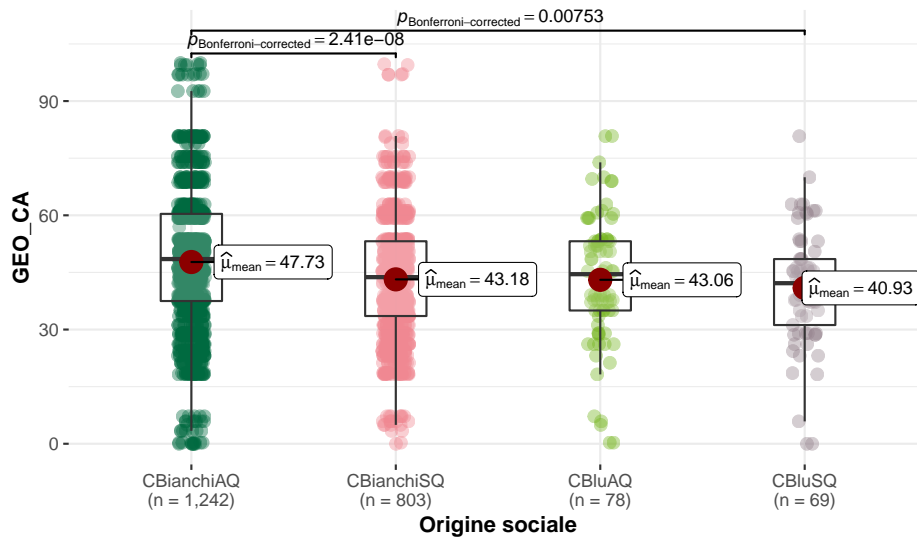
Punteggi per Matematica generale in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 30.21, p = 3.91e-19, \hat{\eta}_p^2 = 0.04, CI_{95\%} [0.03, 1.00], n_{\text{obs}} = 2,192$



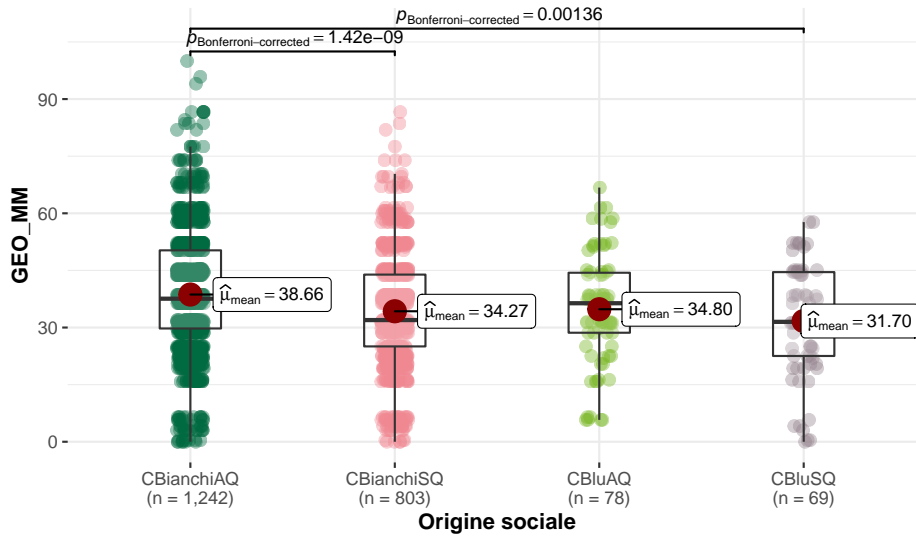
Punteggi per la dimensione GEO_CA in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 14.20, p = 3.62e-09, \hat{\eta}_p^2 = 0.02, CI_{95\%} [9.97e-03, 1.00], n_{\text{obs}} = 2,192$



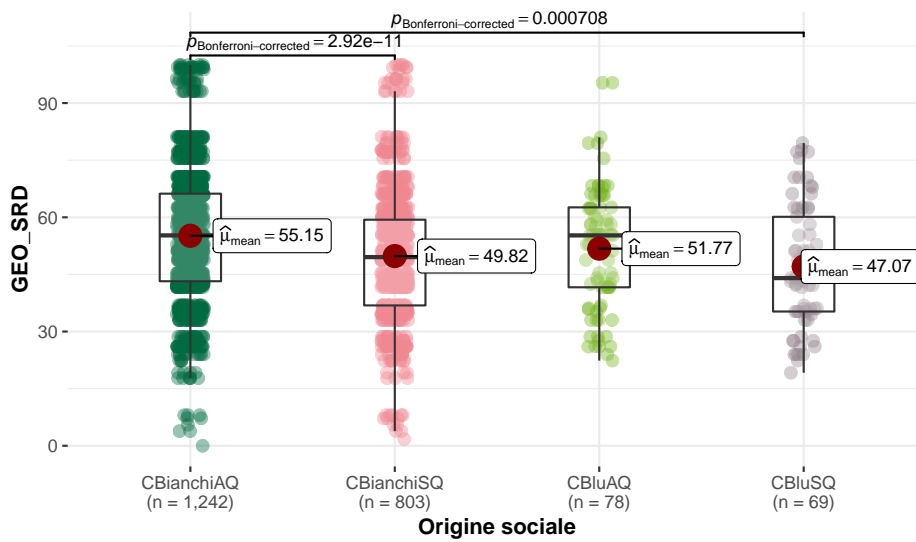
Punteggi per la dimensione GEO_MM in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 16.55, p = 1.24e-10, \hat{\eta}_p^2 = 0.02, CI_{95\%} [0.01, 1.00], n_{\text{obs}} = 2,192$



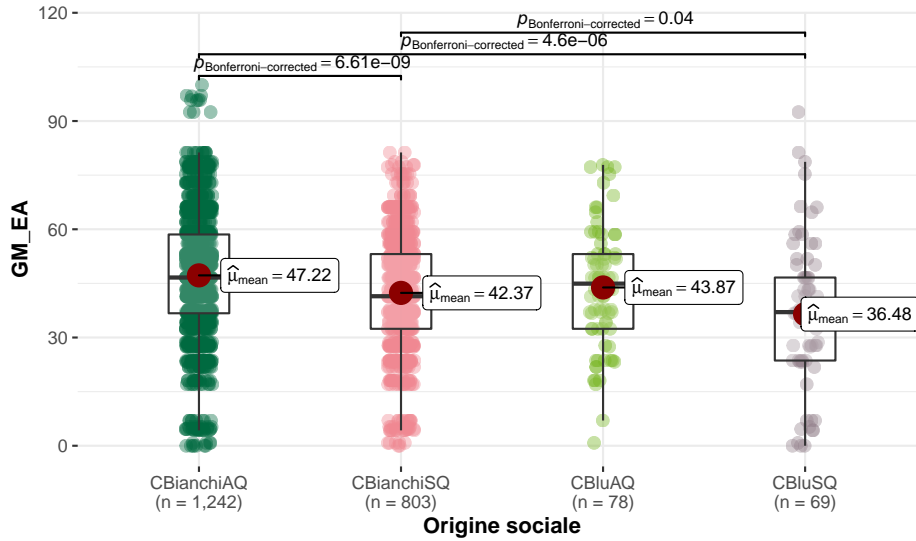
Punteggi per la dimensione GEO_SRD in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 18.98, p = 3.78e-12, \hat{\eta}_p^2 = 0.03, CI_{95\%} [0.01, 1.00], n_{\text{obs}} = 2,192$



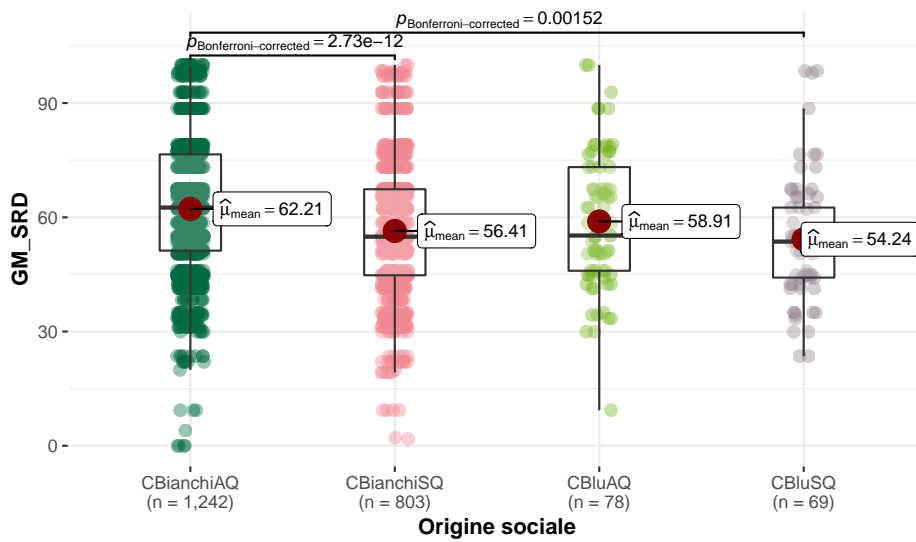
Punteggi per la dimensione GM_EA in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 18.26, p = 1.06e-11, \hat{\eta}_p^2 = 0.02, CI_{95\%} [0.01, 1.00], n_{\text{obs}} = 2,192$



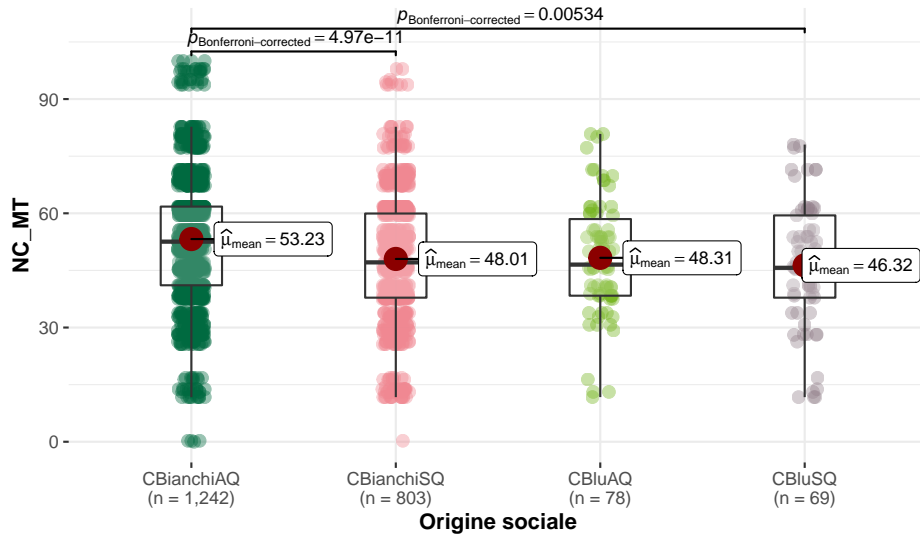
Punteggi per la dimensione GM_SRD in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 20.07, p = 7.90e-13, \hat{\eta}_p^2 = 0.03, CI_{95\%} [0.02, 1.00], n_{\text{obs}} = 2,192$



Punteggi per la dimensione NC_MT in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 18.26, p = 1.06e-11, \hat{\eta}_p^2 = 0.02, \text{CI}_{95\%} [0.01, 1.00], n_{\text{obs}} = 2,192$



Ritornare al rapporto.

C Appendice Italiano

Per le informazioni relative alla lettura e interpretazione dei seguenti grafici si invita il lettore a consultare l'appendice metodologica.

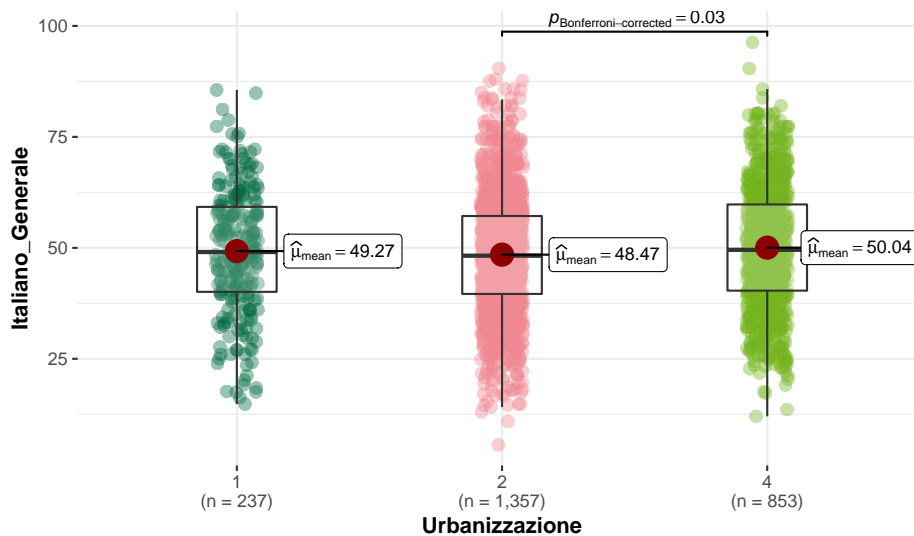
C.1 Analisi di varianza (ANOVA) - Urbanizzazione

Per quel che concerne l'urbanizzazione dei comuni, appaiono delle differenze di risultati unicamente per le dimensioni **Italiano Generale Metacognizione** e **Ortografia**. In entrambi i casi gli allievi dei comuni rurali ottengono prestazioni più elevate rispetto ai comuni della cintura. Gli altri confronti (i.e., tra comuni rurali vs. comuni delle città nucleo; comuni delle città nucleo vs. comuni della cintura) non evidenziano differenze.

Va però sottolineato come le differenze osservate siano di piccola entità come evidenziato dalla misura della taglia dell'effetto (i.e., $\eta_p^2 < .01$).

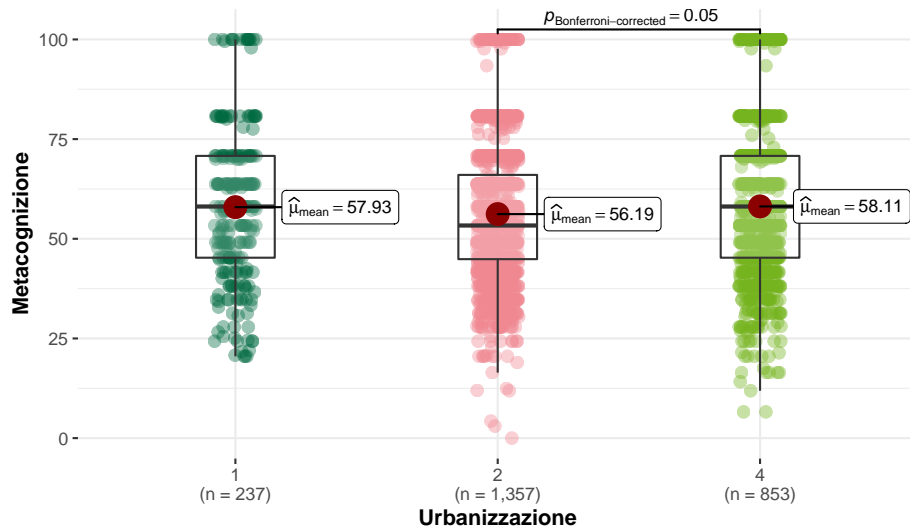
Punteggi per Italiano generale in funzione dell'urbanizzazione

$F_{\text{Fisher}}(2, 2444) = 3.42, p = 0.03, \widehat{\eta}_p^2 = 2.79e-03, CI_{95\%} [1.20e-04, 1.00], n_{\text{obs}} = 2,447$



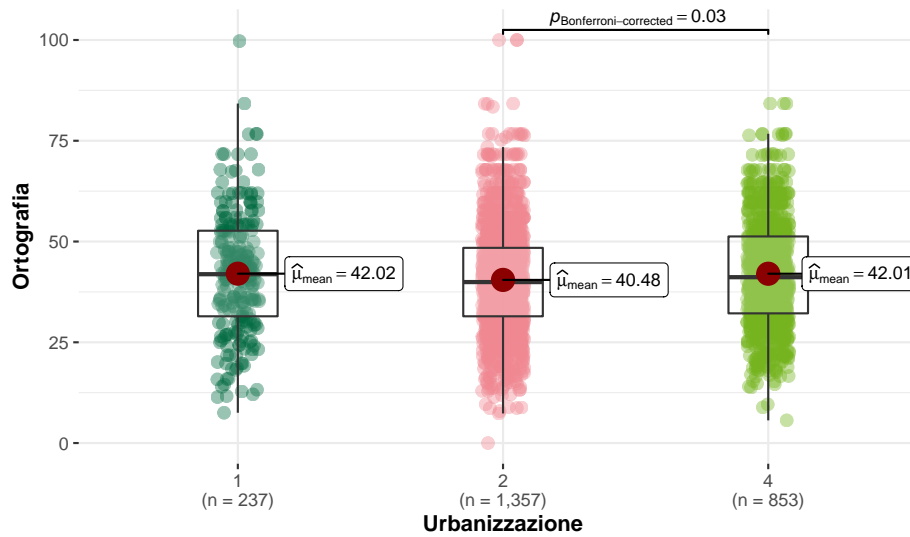
Punteggi per la dimensione Metacognizione in funzione dell'urbanizzazione

$F_{\text{Fisher}}(2, 2444) = 3.23, p = 0.04, \hat{\eta}_p^2 = 2.64e-03, CI_{95\%} [6.48e-05, 1.00], n_{\text{obs}} = 2,447$



Punteggi per la dimensione Ortografia in funzione dell'urbanizzazione

$F_{\text{Fisher}}(2, 2444) = 3.86, p = 0.02, \hat{\eta}_p^2 = 3.14e-03, CI_{95\%} [2.52e-04, 1.00], n_{\text{obs}} = 2,447$



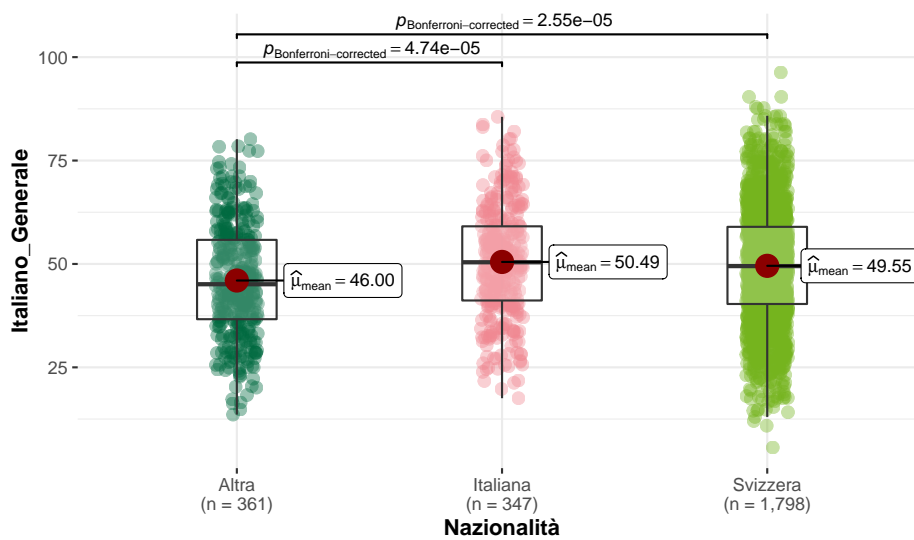
Ritornare al rapporto.

C.2 Analisi di varianza (ANOVA) - Nazionalità

Per quel che concerne la nazionalità degli allievi, le analisi di varianza effettuate permettono di mettere in evidenza delle differenze di risultati principalmente tra gli allievi di nazionalità svizzera o italiana e gli allievi di altre nazionalità (non sussistono invece differenze tra gli allievi di nazionalità italiana e di altre nazionalità). Si deve sottolineare come le differenze osservate siano di piccola entità come evidenziato dalla misura della taglia dell'effetto (i.e., $\eta_p^2 < .01$).

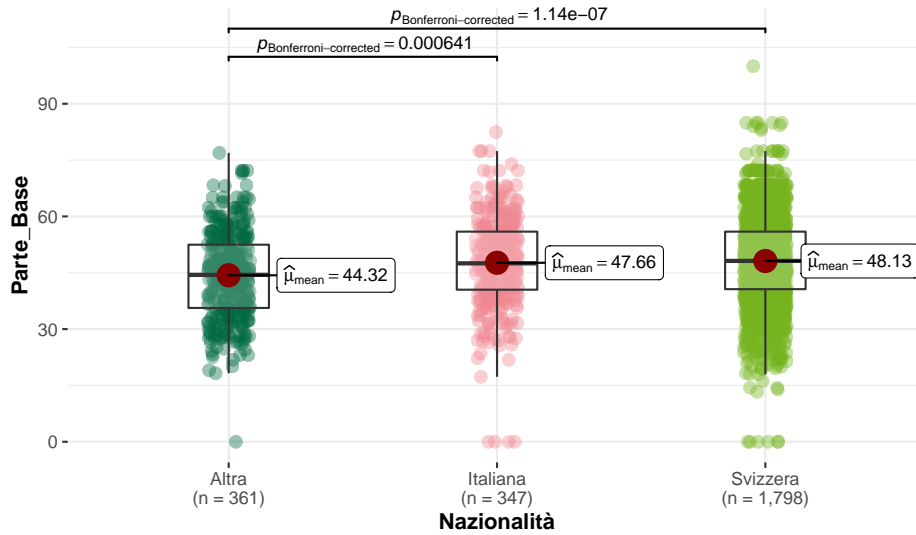
Punteggi per Italiano generale in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 11.79, p = 8.00\text{e-}06, \hat{\eta}_p^2 = 9.33\text{e-}03, \text{CI}_{95\%} [3.77\text{e-}03, 1.00], n_{\text{obs}} = 2,506$



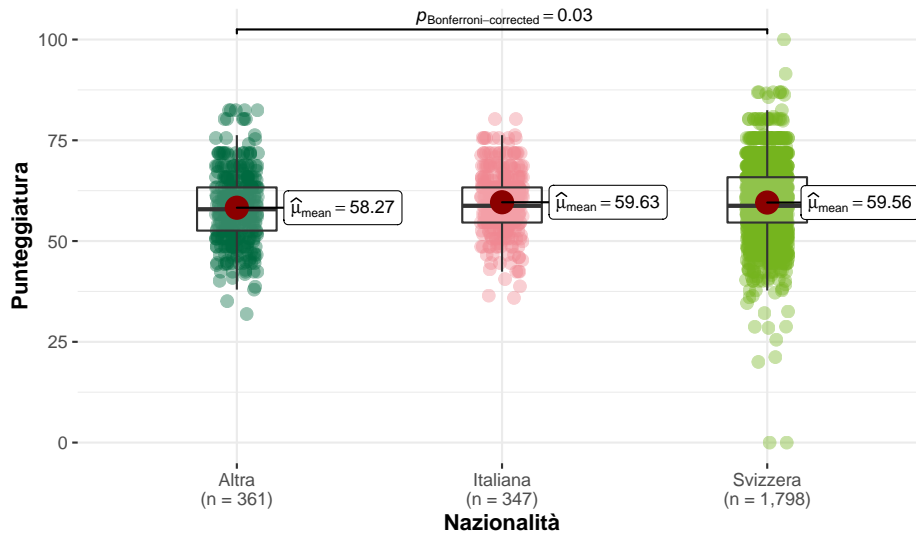
Punteggi per la dimensione Parte_base in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 15.24, p = 2.63e-07, \hat{\eta}_p^2 = 0.01, \text{CI}_{95\%} [5.62e-03, 1.00], n_{\text{obs}} = 2,506$



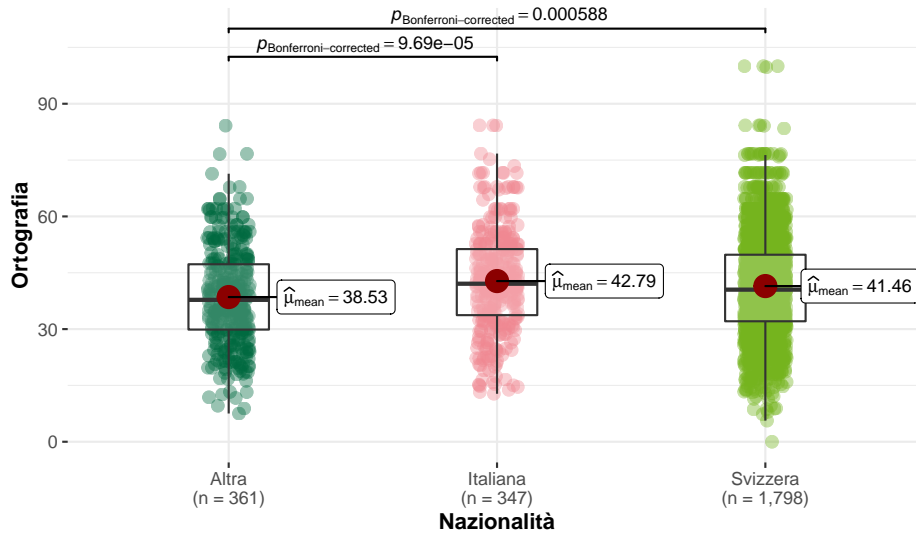
Punteggi per la dimensione Punteggiatura in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 3.47, p = 0.03, \hat{\eta}_p^2 = 2.77e-03, \text{CI}_{95\%} [1.32e-04, 1.00], n_{\text{obs}} = 2,506$



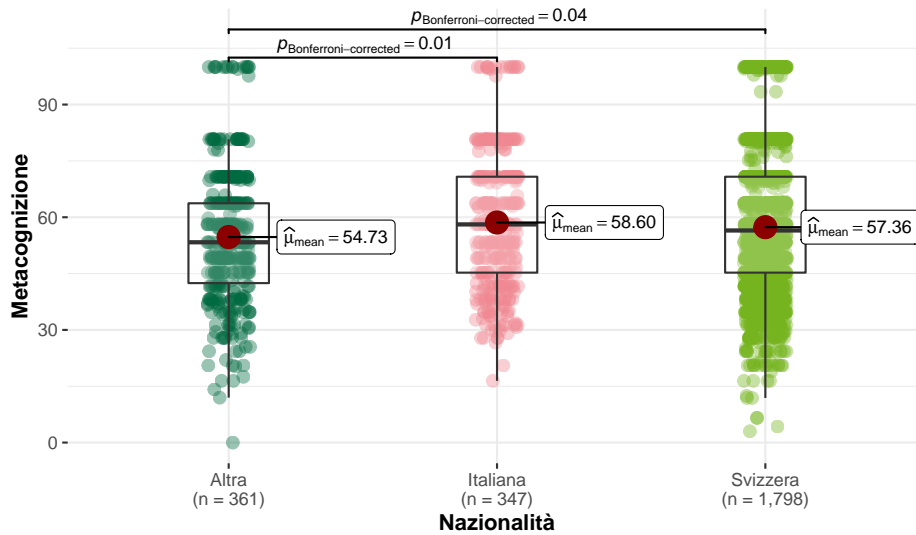
Punteggi per la dimensione Ortografia in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 9.64, p = 6.78e-05, \hat{\eta}_p^2 = 7.64e-03, CI_{95\%} [2.68e-03, 1.00], n_{\text{obs}} = 2,506$



Punteggi per la dimensione Metacognizione in funzione della nazionalità

$F_{\text{Fisher}}(2, 2503) = 4.39, p = 0.01, \hat{\eta}_p^2 = 3.49e-03, CI_{95\%} [4.20e-04, 1.00], n_{\text{obs}} = 2,506$



Ritornare al rapporto.

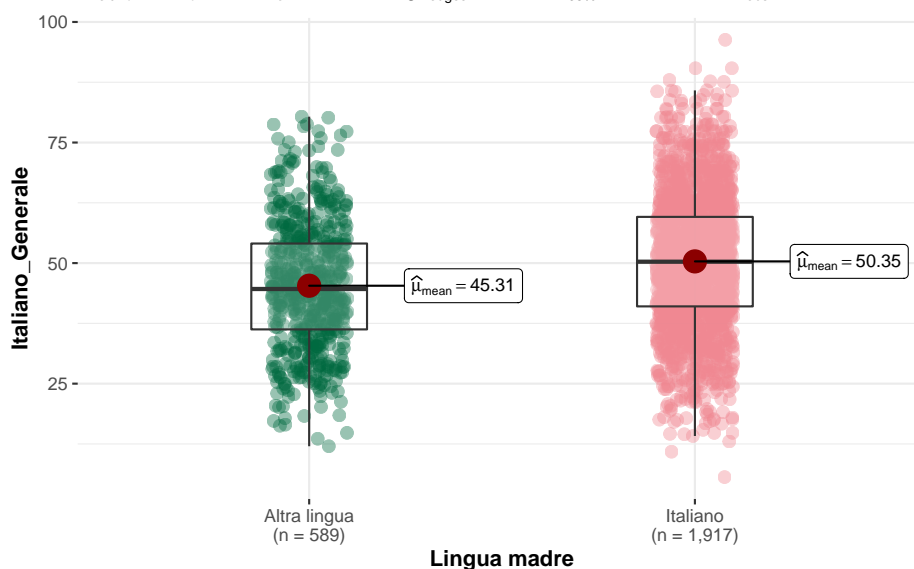
C.3 Test t di Student - La lingua madre

Per quel che riguarda i risultati in italiano generale, il test statistico evidenzia come gli allievi di lingua madre italiana ottengono risultati migliori ($m = 50.35$) rispetto agli allievi di un'altra lingua madre ($m = 45.31$). Anche i risultati relativi alle singole dimensioni evidenziano dei *pattern* di risultati simili che indicano in modo sistematico come gli allievi di lingua madre italiana ottengano risultati migliori rispetto a quelli di lingua madre diversa.

È inoltre importante rilevare come le differenze osservate siano importanti, soprattutto per la dimensione di **italiano generale** e per la dimensione denominata **Parte base** con dei valori di taglia dell'effetto di .37 e .47 rispettivamente.

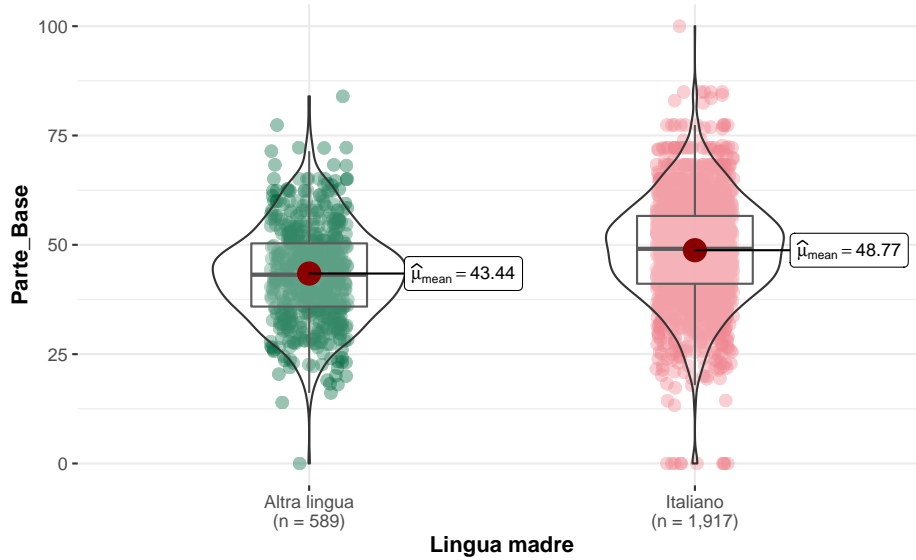
Punteggi per Italiano generale in funzione della lingua madre

$t_{\text{Welch}}(1021.3) = -8.04$, $p = 2.51e-15$, $\hat{g}_{\text{Hedges}} = -0.37$, $CI_{95\%} [-0.47, -0.28]$, $n_{\text{obs}} = 2,506$



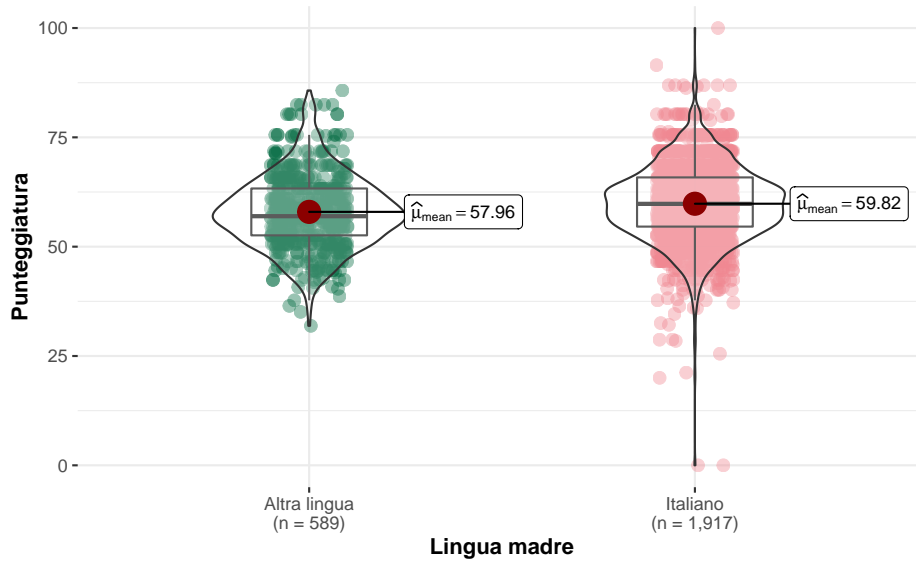
Punteggi per la dimensione Parte base in funzione della lingua madre

$t_{\text{Welch}}(1057.41) = -10.05, p = 8.97e-23, \hat{g}_{\text{Hedges}} = -0.46, \text{CI}_{95\%} [-0.55, -0.37], n_{\text{obs}} = 2,506$



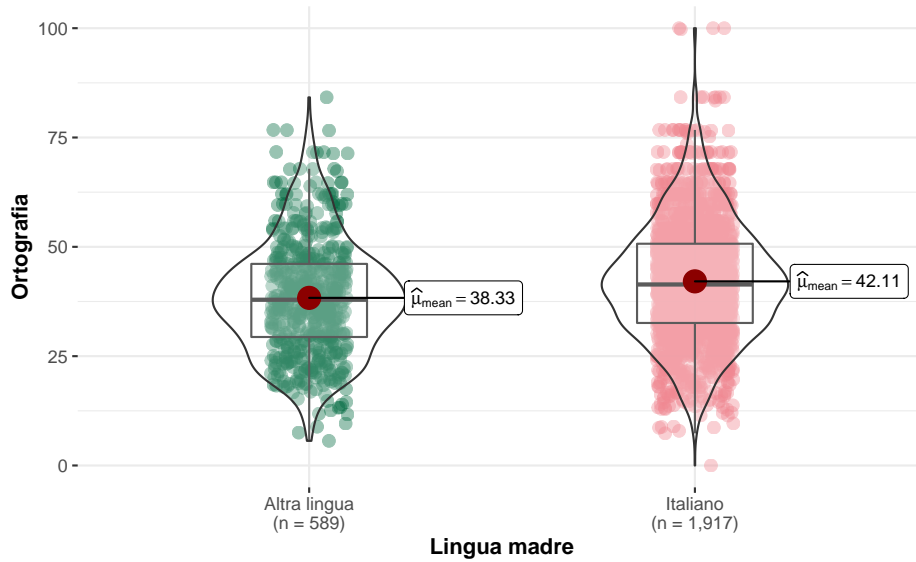
Punteggi per la dimensione Punteggiatura in funzione della lingua madre

$t_{\text{Welch}}(1010.8) = -4.66, p = 3.51e-06, \hat{g}_{\text{Hedges}} = -0.22, \text{CI}_{95\%} [-0.31, -0.13], n_{\text{obs}} = 2,506$



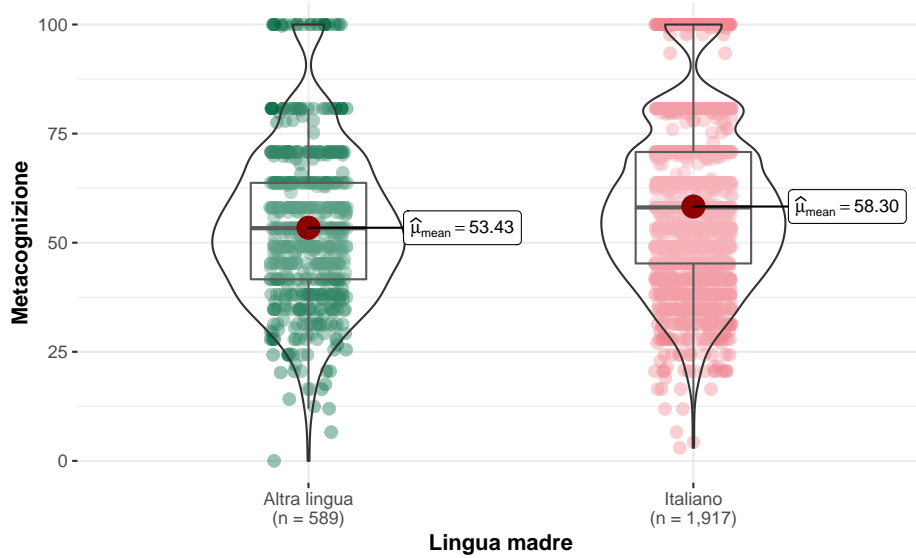
Punteggi per la dimensione Ortografia in funzione della lingua madre

$t_{\text{Welch}}(1018.06) = -6.07, p = 1.83e-09, \hat{g}_{\text{Hedges}} = -0.28, \text{CI}_{95\%} [-0.37, -0.19], n_{\text{obs}} = 2,506$



Punteggi per la dimensione Metacognizione in funzione della lingua madre

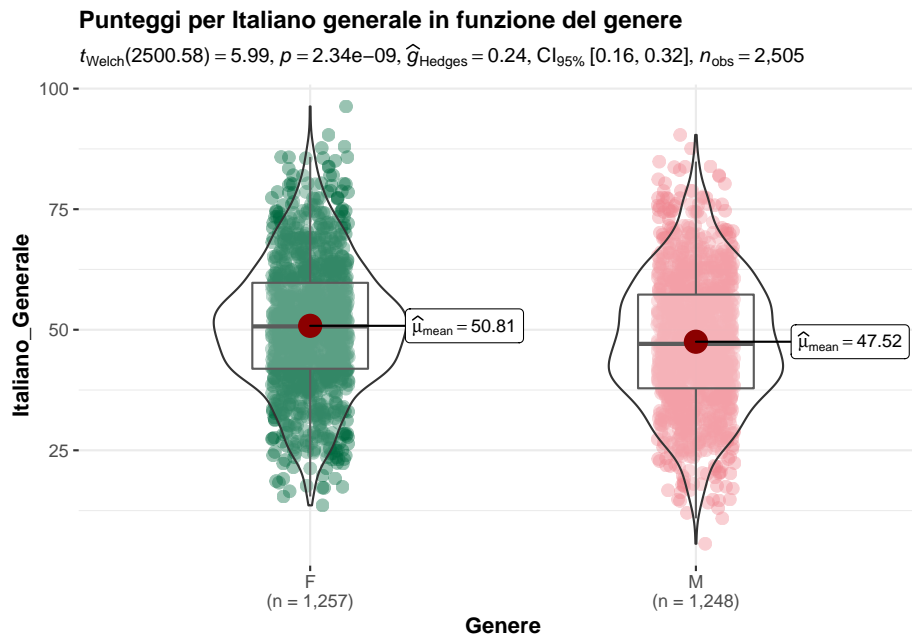
$t_{\text{Welch}}(1003.95) = -5.79, p = 9.31e-09, \hat{g}_{\text{Hedges}} = -0.27, \text{CI}_{95\%} [-0.36, -0.18], n_{\text{obs}} = 2,506$



Ritornare al rapporto.

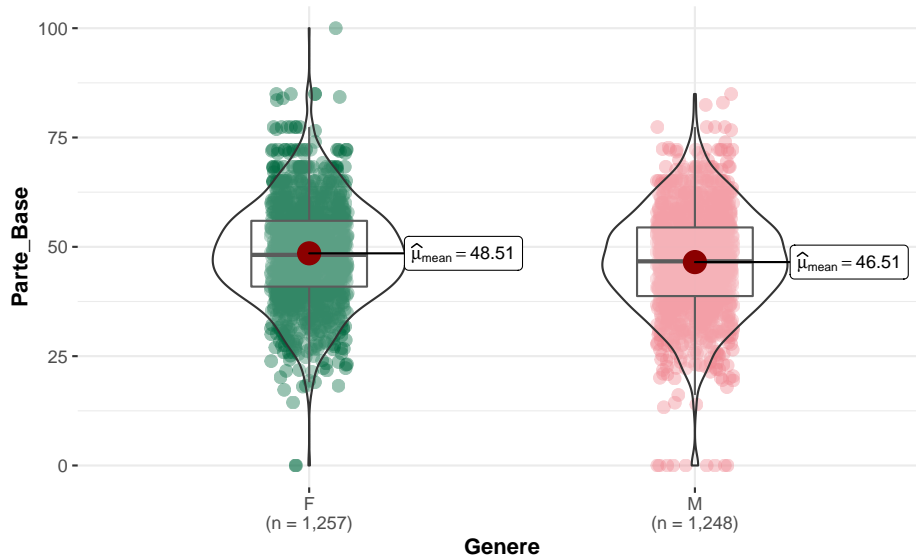
C.4 Test t di Student - Il genere

Le analisi realizzate mettono in evidenza delle differenze statisticamente significative in ognuna delle dimensioni testate, con le femmine che presentano delle performance migliori rispetto ai maschi. Rispetto ai valori della taglia dell'effetto, va sottolineato come le differenze siano contenute con dei valori inferiori a .26.



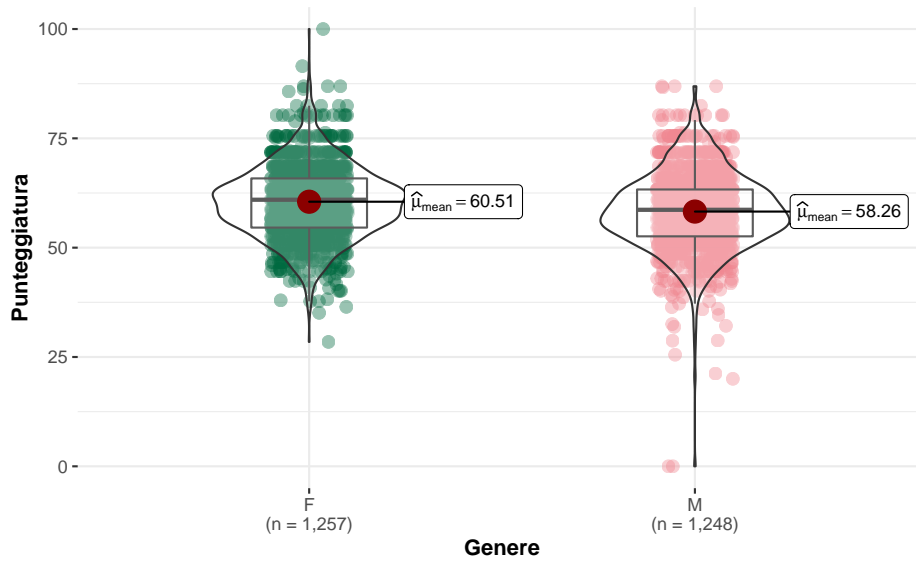
Punteggi per la dimensione Parte base in funzione del genere

$t_{\text{Welch}}(2499.64) = 4.18, p = 3.04e-05, \hat{g}_{\text{Hedges}} = 0.17, \text{CI}_{95\%} [0.09, 0.25], n_{\text{obs}} = 2,505$



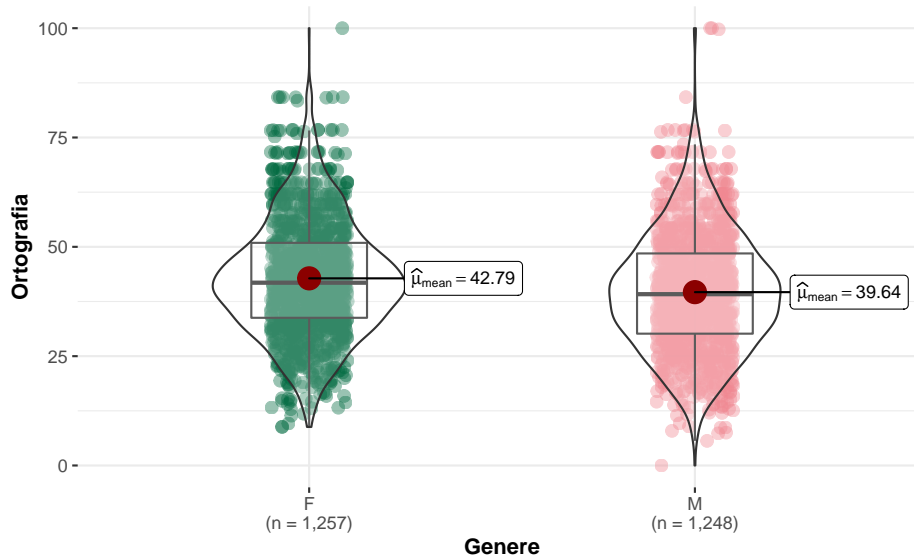
Punteggi per la dimensione Punteggiatura in funzione del genere

$t_{\text{Welch}}(2480.47) = 6.54, p = 7.65e-11, \hat{g}_{\text{Hedges}} = 0.26, \text{CI}_{95\%} [0.18, 0.34], n_{\text{obs}} = 2,505$



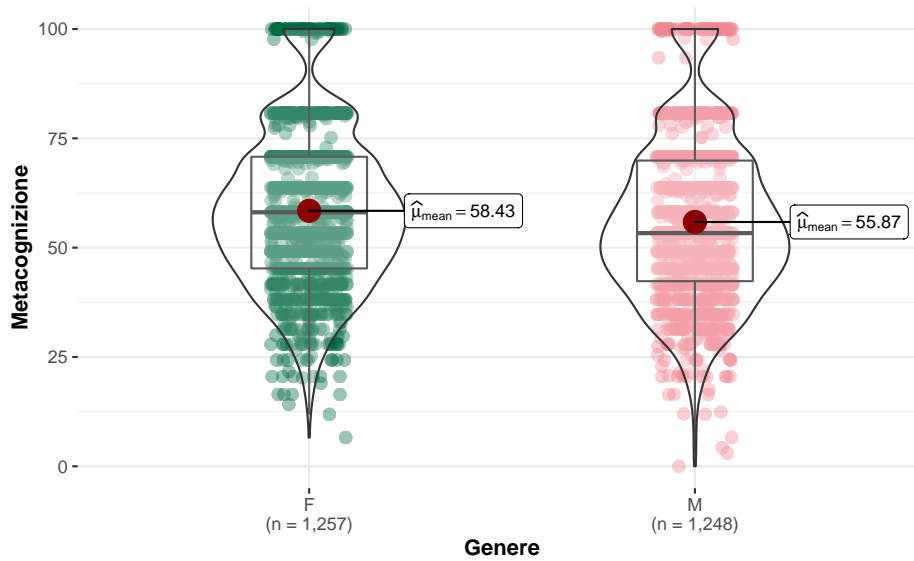
Punteggi per la dimensione Ortografia in funzione del genere

$t_{\text{Welch}}(2500.98) = 5.82, p = 6.78e-09, \hat{g}_{\text{Hedges}} = 0.23, CI_{95\%} [0.15, 0.31], n_{\text{obs}} = 2,505$



Punteggi per la dimensione Metacognizione in funzione del genere

$t_{\text{Welch}}(2497.47) = 3.51, p = 4.57e-04, \hat{g}_{\text{Hedges}} = 0.14, CI_{95\%} [0.06, 0.22], n_{\text{obs}} = 2,505$



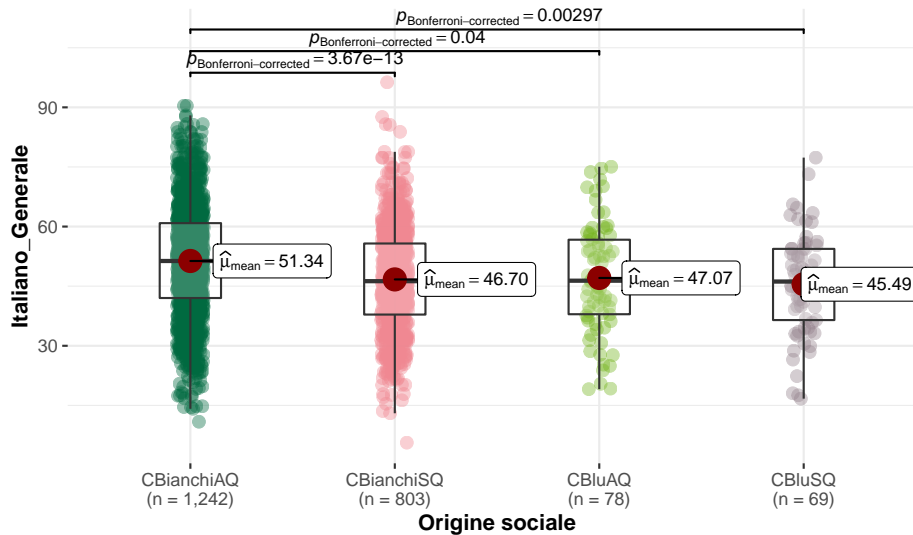
Ritornare al rapporto.

C.5 Analisi di varianza (ANOVA) - Origine sociale

Dalle analisi realizzate emerge in modo sistematico come gli allievi di origine sociale più elevata abbiano delle performance migliori. Nello specifico dai test *post hoc* realizzati si può osservare come gli allievi del gruppo “Colletti bianchi altamente qualificati” ottengano sempre dei risultati migliori rispetto ai “Colletti bianchi scarsamente qualificati”. Inoltre, per **italiano generale** appaiono delle differenze pure tra i “Colletti bianchi altamente qualificati” che ottengono dei risultati più elevati rispetto ai “Colletti blu altamente qualificati” ed ai “Colletti blu scarsamente qualificati”. Per la dimensione **Parte base** si evince come i “Colletti bianchi altamente qualificati” ottengano punteggi più elevati rispetto ai “Colletti bianchi scarsamente qualificati” e rispetto ai “Colletti blu scarsamente qualificati”. Per la dimensione **Metacognizione** i “Colletti bianchi altamente qualificati” ottengano punteggi più elevati rispetto ai “Colletti bianchi scarsamente qualificati” e rispetto ai “Colletti blu altamente qualificati”.

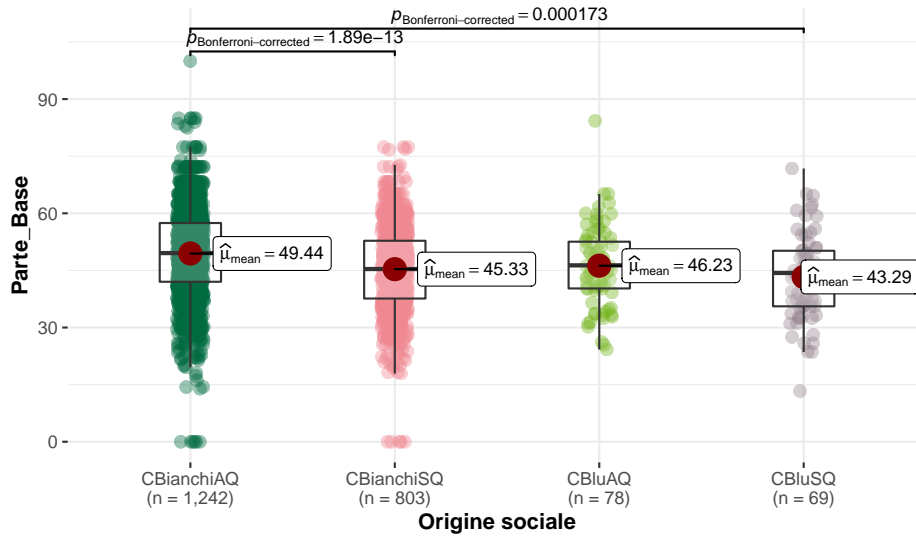
Punteggi per Italiano generale in funzione dell'origine sociale

$$F_{\text{Fisher}}(3, 2188) = 21.72, p = 7.31e-14, \hat{\eta}_p^2 = 0.03, \text{CI}_{95\%} [0.02, 1.00], n_{\text{obs}} = 2,192$$



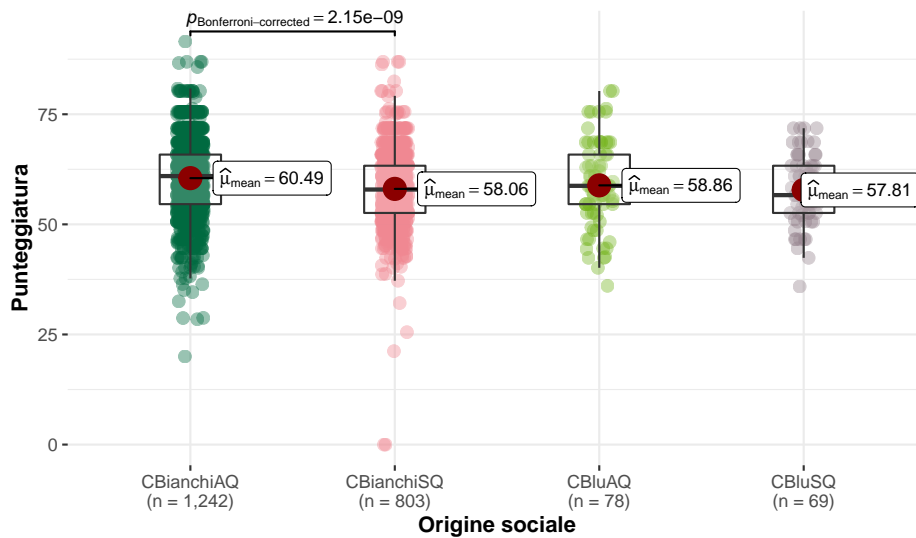
Punteggi per la dimensione Parte Base in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 23.10, p = 1.02e-14, \hat{\eta}_p^2 = 0.03, CI_{95\%} [0.02, 1.00], n_{\text{obs}} = 2,192$



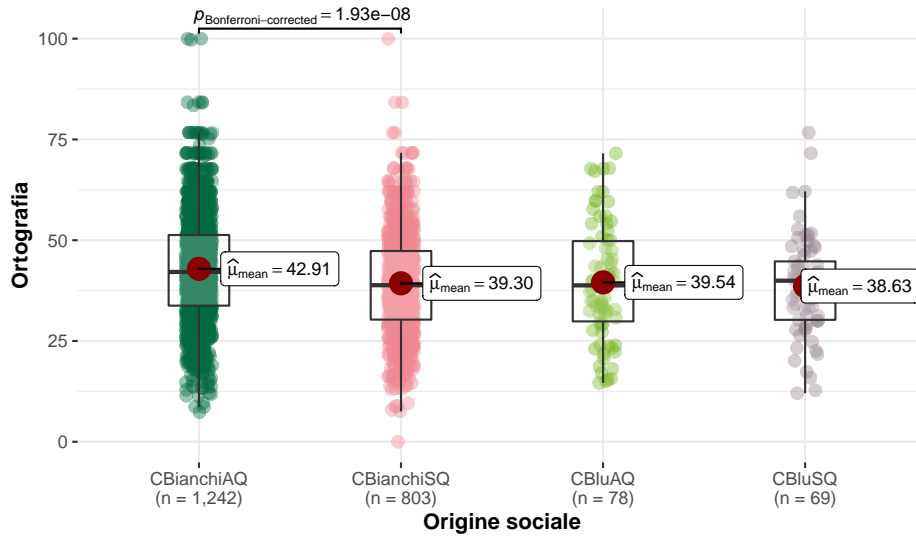
Punteggi per la dimensione Punteggiatura in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 14.28, p = 3.26e-09, \hat{\eta}_p^2 = 0.02, CI_{95\%} [0.01, 1.00], n_{\text{obs}} = 2,192$



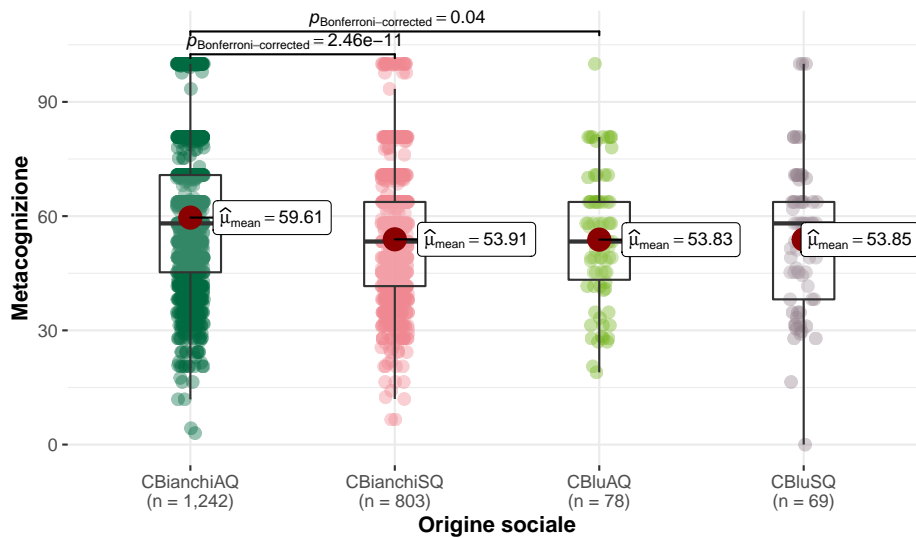
Punteggi per la dimensione Ortografia in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 13.27, p = 1.38e-08, \hat{\eta}_p^2 = 0.02, CI_{95\%} [9.04e-03, 1.00], n_{\text{obs}} = 2,192$



Punteggi per la dimensione Metacognizione in funzione dell'origine sociale

$F_{\text{Fisher}}(3, 2188) = 17.96, p = 1.64e-11, \hat{\eta}_p^2 = 0.02, CI_{95\%} [0.01, 1.00], n_{\text{obs}} = 2,192$



Ritornare al rapporto.