



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

Scuola di Scienze Matematiche, Fisiche e Naturali  
Corso di Laurea in Informatica

Tesi di Laurea

ANALISI E CARATTERIZZAZIONE  
DELL'ABBANDONO NEL CORSO DI LAUREA  
IN INFORMATICA

ANALYSIS AND DESCRIPTION OF THE  
STUDENT'S DROPOUT IN THE COMPUTER  
SCIENCE DEGREE COURSE

ALBERTO BILIOTTI

Relatrice: *Maria Cecilia Verri*  
Correlatrice: *Renza Campagni*

Anno Accademico 2021-2022



## INDICE

---

<b>1</b>	<b>Introduzione</b>	<b>9</b>
1.1	Obbiettivi	9
1.2	Testi di riferimento	9
<b>2</b>	<b>Dati analizzati</b>	<b>13</b>
2.1	Dati sullo stato d'iscrizione	14
<b>3</b>	<b>Operazioni preliminari e strumenti usati</b>	<b>17</b>
3.1	Stimare l'abbandono	17
3.2	Strumenti utilizzati	18
3.3	Caricamento e pulizia dei dati	18
3.4	Aggregazione degli esami svolti da uno studente	19
3.5	Rappresentazione ed organizzazione dei dati	20
3.6	Trovare gli studenti laureati	20
<b>4</b>	<b>Studenti inattivi</b>	<b>23</b>
4.1	Trovare gli studenti inattivi	23
4.2	Distinguere gli studenti inattivi in base ai crediti	24
<b>5</b>	<b>Confronto dei dati</b>	<b>25</b>
5.1	Il voto di maturità può avere un'influenza?	25
5.2	La distanza dal luogo di residenza influisce?	26
5.3	Il test d'ingresso può prevedere l'inattività?	27
5.4	Il genere ha una qualche influenza sull'inattività?	30
5.5	Ci sono scuole superiori "migliori"?	31
5.6	La media dei voti può influenzare l'inattività?	33
5.7	I crediti conseguiti al primo anno condizionano l'inattività?	35
<b>6</b>	<b>Analisi ed evoluzione temporale</b>	<b>39</b>

6.1	L'andamento dell'abbandono in base alla coorte di immatricolazione	39	
6.2	Quanto tempo impiega uno studente prima di diventare inattivo?	40	
6.3	Quali sono gli esami che gli studenti inattivi non riescono a superare?	42	
6.4	Evoluzione della scuola di provenienza degli studenti inattivi	43	
6.5	Andamento temporale degli studenti inattivi, attivi e laureati	45	
<b>7</b>	<b>Analisi e confronto dei dati aggiornati</b>		<b>53</b>
7.1	Correttezza della precedente stima	53	
7.2	Abbandoni nelle coorti di iscrizione 2018 e 2019	56	
7.3	Stato d'iscrizione di tutti gli studenti	59	
<b>8</b>	<b>Conclusione</b>		<b>61</b>
8.1	Analisi finale della valutazione al test d'ingresso	61	
8.2	Altri parametri da tenere in considerazione	63	
8.3	Considerazioni finali	63	

## ELENCO DELLE FIGURE

---

Figura 1	Diagramma Entità Relazione delle Tabelle degli studenti e degli esami superati	22
Figura 2	Proporzione studenti inattivi, attivi e laureati	23
Figura 3	Tipologie di studenti inattivi divisi per crediti	24
Figura 4	Distribuzione voto di maturità degli studenti iscritti e di quelli inattivi	25
Figura 5	Distribuzione voto di maturità degli studenti iscritti e di quelli inattivi	26
Figura 6	Distribuzione della provenienza degli studenti iscritti e di quelli inattivi	27
Figura 7	Distribuzione della provenienza degli studenti attivi	28
Figura 8	Distribuzione superamento prova d'ingresso degli studenti iscritti e di quelli inattivi	28
Figura 9	Distribuzione in base al superamento della prova d'ingresso degli studenti iscritti e di quelli inattivi delle coorti dal 2016 in su	29
Figura 10	Distribuzione in base al superamento della prova d'ingresso degli studenti attivi o laureati delle coorti dal 2016 in su	30
Figura 11	Distribuzione del genere negli studenti iscritti ed inattivi	31
Figura 12	Distribuzione del genere negli studenti attivi	31
Figura 13	Distribuzione del tipo di scuola superiore di provenienza negli studenti iscritti ed inattivi	32
Figura 14	Distribuzione del tipo di scuola superiore di provenienza negli studenti attivi o laureati	33
Figura 15	Distribuzione della media dei voti degli esami superati negli studenti iscritti ed inattivi	34
Figura 16	Distribuzione della media ponderata dei voti degli esami superati dagli studenti attivi o laureati	35

#### 4 Elenco delle figure

Figura 17	Crediti conseguiti al primo anno dagli studenti iscritti ed inattivi	37
Figura 18	Crediti conseguiti al primo anno da studenti attivi o laureati	37
Figura 19	Proporzione di studenti inattivi sul totale per coorte	40
Figura 20	Grafico riportante il numero di studenti che è diventato inattivo dopo un certo numero di anni	41
Figura 21	Periodo di tempo prima che uno studente che ha superato almeno un esame diventi inattivo espresso in anni	42
Figura 22	Esami del primo anno del Corso di Laurea superati da studenti inattivi	44
Figura 23	Scuola superiore di provenienza degli studenti inattivi per coorte	45
Figura 24	Scuola superiore di provenienza degli studenti per coorte	45
Figura 25	Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2010 anno per anno	47
Figura 26	Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2011 anno per anno	47
Figura 27	Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2012 anno per anno	48
Figura 28	Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2013 anno per anno	48
Figura 29	Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2014 anno per anno	49
Figura 30	Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2015 anno per anno	49
Figura 31	Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2016 anno per anno	50
Figura 32	Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2017 anno per anno	50
Figura 33	Andamento complessivo degli studenti attivi, inattivi e laureati degli studenti iscritti per anni dall'immatricolazione	51
Figura 34	Stato dell'iscrizione degli studenti valutati erroneamente come non inattivi	56
Figura 35	Stato dell'iscrizione degli studenti valutati erroneamente come inattivi	56

Figura 36	Stato dell'iscrizione degli studenti delle coorti 2018 e 2019 registrato nel 2021	57
Figura 37	Stato dell'iscrizione degli studenti iscritti nella coorte 2018	58
Figura 38	Stato dell'iscrizione degli studenti iscritti nella coorte 2019	58
Figura 39	Stato dell'iscrizione degli studenti iscritti di tutte le coorti (2010-2019)	59
Figura 40	Evoluzione della media delle valutazioni standardizzate conseguite da studenti inattivi in base alla coorte	62





*"Non mi preoccupo mai del futuro, arriva sempre abbastanza presto"*  
— *Albert Einstein*



## INTRODUZIONE

---

### 1.1 OBIETTIVI

L'obiettivo di questa analisi è lo studio e la caratterizzazione statistica dell'abbandono universitario da parte degli studenti del corso di Laurea triennale in Informatica: in particolare ci siamo interessati alla ricerca dei tratti che più significativamente caratterizzano i gruppi di coloro che si ritiene abbiano abbandonato il corso di Laurea da tutti gli studenti iscritti alle varie coorti che abbiamo preso in esame. Tali coorti di iscrizione sono quelle che vanno dal 2010 al 2019 e di cui possediamo i dati, che illustreremo nei capitoli seguenti, aggiornati al 2020. I dati che abbiamo usato in quest'analisi provengono dall'archivio e dall'anagrafica dei servizi online di ateneo. Inoltre sono stati analizzati anche i dati riguardanti lo stato della carriera degli studenti aggiornati al 2021. Tali dati però, come illustreremo meglio in seguito, sono arrivati a nostra disposizione quando l'analisi era ormai avviata da mesi, motivo per il quale il numero di analisi che abbiamo potuto condurre su questi dati "aggiornati" è limitato. Infine è necessario specificare che, non avendo a disposizione dati certi su quali studenti avessero effettivamente abbandonato gli studi, per poter essere in grado di studiare e caratterizzare gli studenti che hanno abbandonato, in questo lavoro è stato assunto che coloro che risultano come inattivi da due anni abbiano abbandonato il Corso di Laurea (assunzione che, come vedremo in seguito, non risulta essere troppo errata) ponendo quindi come improbabile l'eventualità che uno studente ritorni attivo dopo più di due anni di inattività.

### 1.2 TESTI DI RIFERIMENTO

È stato usato come riferimento di partenza lo studio "How Does Learning Analytics Contribute to Prevent Students'Dropout in Higher Education: A Systematic Literature Review" di Catarina Félix de Oliveira,

Sónia Rolland Sobral, Maria João Ferreira e Fernando Moreira [1]. Questo testo raccoglie un numero considerevole di studi riguardanti l'uso del "Learning Analytics" per l'analisi di diversi fenomeni e caratteristiche riguardanti gli studenti iscritti a corsi universitari di natura molto variegata (come università telematiche o corsi di studio umanistici). Nello specifico [1] si sofferma in maniera consistente sui precedenti studi riguardanti l'abbandono da parte degli studenti universitari, in particolare su come i "Learning Analytics" permettano di descrivere, caratterizzare ed infine anche predire le cause dell'abbandono universitario con lo scopo, come evidenziato anche nel titolo, di contribuire a ridurre tale fenomeno. Con "Learning Analytics" intendiamo diverse caratteristiche che possono essere proprie degli studenti come i dati "anagrafici" (età, genere, etnia, nazione di provenienza, impiego, stato coniugale, ma anche informazioni riguardanti i genitori come la professione da loro svolta o la condizione economica) oppure i dati più difficili da "quantificare" come l'impegno profuso nello studio, il livello di soddisfazione o lo stato di salute fisica e mentale uniti ai dati più prettamente "accademici" come i risultati nei diversi test di ingresso ai corsi di laurea, i crediti conseguiti, i corsi che lo studente sta frequentando, i voti conseguiti per ogni esame, gli esami non superati, le informazioni riguardo all'eventuale prova finale ed anche se lo studente abbia abbandonato gli studi. Oltre ai dati più strettamente legati ad ogni singolo studente nei vari studi presi in esame in [1] vengono usate anche caratteristiche "esterne" allo studente, che possono essere, ad esempio, le infrastrutture, i servizi, la posizione geografica in cui si trova l'università, la distanza tra i luoghi di residenza e l'università ed il supporto ricevuto dagli studenti attraverso borse di studio, corsi di recupero e tutor. Anche nel lavoro che andremo ad illustrare in seguito è stato fatto ricorso ad i "Learning Analytics" per poter analizzare e caratterizzare l'abbandono universitario, pur avendo però a disposizione, come illustrato in seguito, una quantità inferiore di dati forniti rispetto a quelli disponibili negli articoli citati nello studio. Questa mancanza è dovuta, oltre ai problemi riguardanti la privacy e la riservatezza dei dati, principalmente alle differenze di tipo burocratico, amministrativo ed organizzativo che caratterizzano i diversi atenei anche nella stessa nazione. L'esempio più lampante, che analizzeremo dettagliatamente in seguito, è quello riguardante l'abbandono stesso degli studenti universitari che nel nostro studio non può essere stabilito con un criterio "sicuro" ma deve necessariamente essere "approssimato" cosa che in altri studi dello studio viene del tutto ignorata in quanto condotta in atenei nei quali è possibile stabilire con certezza se uno studente

abbia abbandonato o meno. Tale approssimazione si è resa necessaria in quanto, a differenza di altre università o corsi di studio, il Corso di Laurea in Informatica presso l'Università di Firenze non prevede che uno studente possa rimanere escluso dall'università se non rispetta determinati parametri (che nelle altre università possono essere crediti, esami, media o anche il pagamento delle tasse d'iscrizione) motivo per il quale risulta impossibile stabilire con certezza quali e quanti studenti abbiano effettivamente abbandonato il Corso di Laurea senza avere a disposizione i dati sullo stato della carriera degli studenti (che sono arrivati in nostro possesso solo quando ormai il nostro lavoro era ormai quasi concluso). La nostra analisi si è quindi basata quasi esclusivamente sui dati "accademici" attraverso i quali abbiamo ricavato un'approssimazione dei possibili studenti che hanno abbandonato ed in base a tale approssimazione abbiamo analizzato e caratterizzato i principali attributi di tale gruppo in contrapposizione a quelli di tutti gli studenti iscritti al Corso di Laurea.

Un altro articolo, riguardante lo studio dei fattori che influenzano l'abbandono, che si è rilevato utile per condurre la nostra analisi è "Factors influencing university drop out rates" di Francisco Araque, Concepción Roldán, Alberto Salguer [2].

In particolare questo studio è stato usato come riferimento per la pulizia, il caricamento e la gestione dei dati relativi ad ogni studente anche se, invece di usare un Data Warehouse come nel lavoro sopra citato, abbiamo usato un più semplice database relazionale perché più semplice ed immediato da usare non avendo necessità di aggregare dati da fonti diverse come accaduto invece per il lavoro sopra citato.

Anche questo lavoro però, essendo svolto in un'università spagnola, non tiene conto dei dati che ci sono stati forniti, che non ci permettono di ricavare direttamente quali studenti abbiano effettivamente abbandonato. Per tentare di ovviare a tale problema sono stati consultati anche i lavori "Predicting Students' Dropout at University Using Artificial Neural Networks." di Anna Siri [3] e "Deep learning approach for predicting university dropout: a case study at Roma Tre University." di Francesco Agrusti, Mauro Mezzini e Gianmarco Bonavolontà [4] svolti rispettivamente presso l'Università di Genova e l'Università di Roma Tre. In particolare lo studio svolto presso l'Università di Genova è stato utile in quanto ha individuato il limite di due anni, che illustreremo meglio in seguito, per individuare gli studenti che si stima abbiano abbandonato anche se lo studio si riferisce a due anni in cui lo studente non rinnova la propria iscrizione non pagando le tasse (dato che inizialmente non era

in nostro possesso) mentre nel nostro studio abbiamo usato come soglia due anni dall'ultimo esame superato da uno studente non laureato. Tuttavia entrambi gli studi citati in precedenza hanno a disposizione molti più dati di partenza rispetto a quelli in nostro possesso; infatti questi due studi hanno a disposizione anche dati riguardanti la condizione psicologica dello studente, le condizioni familiari ed anche eventuali disabilità possedute dallo studente (inoltre lo studio condotto all'Università di Roma Tre ha già disponibile come dato l'informazione se lo studente abbia abbandonato o meno).

Alcune delle nostre analisi si sono basate inoltre su quelle effettuate presso l'Università di Bari nello studio "A Statistical Analysis of Factors Affecting Higher Education Dropouts" [5] di Paola Perchinunno Massimo Bilancia e Domenico Vitale, che hanno individuato come causa preponderante per l'abbandono universitario la difficoltà nel superare gli esami, soprattutto al primo anno, mentre in maniera decisamente inferiore anche l'età avanzata dello studente (dato di cui noi non siamo a conoscenza) ed il voto conseguito nell'esame di Maturità (che da noi è risultato, come vedremo in seguito, non rilevante). Questo studio ci ha permesso di individuare quindi le caratteristiche principali sulle quali indagare nella nostra analisi (anche quelle che si sono rilevate non essere influenti come, ad esempio, il genere), oltre ad essere uno studio svolto in un'Università italiana in cui però si conoscevano già gli studenti che hanno abbandonato gli studi.

Un altro contributo fondamentale, soprattutto per quanto riguarda la validazione dei nostri dati con lo stato della carriera dello studente ma anche per quanto concerne la pulizia ed il caricamento dei dati, viene dallo studio riportato nell'articolo "Dropout Situation of Business Computer Students di Praty" Nuankaew [6] nel quale vengono introdotti i concetti di "Falso positivo" e "Falso negativo" che riprenderemo in seguito nel Capitolo 7.

Infine, come ultimo testo, è stato consultato anche lo studio "Predicting Computer Engineering students' dropout in Cuban Higher Education with pre-enrollment and early performance data" [7] che, pur usando metodologie diverse dalle nostre, giunge a conclusioni simili alle nostre: ossia che è difficile, se non impossibile, usare i dati che derivano da esperienze precedenti all'iscrizione di uno studente o comunque che non abbiano relazione con la carriera universitaria per prevedere un possibile futuro abbandono mentre i dati relativi alla carriera universitaria stessa dello studente risultano essere molto più utili per il nostro scopo.

## DATI ANALIZZATI

---

I dati a nostra disposizione includono il resoconto degli esami superati dagli studenti immatricolati nel corso di Laurea triennale in Informatica nelle coorti che vanno dal 2010 al 2019 aggiornato alla fine del 2020 divisi per coorte di immatricolazione. Tale resoconto include sia informazioni sull'esame svolto (descrizione, data in cui l'esame è stato sostenuto, crediti, codice dell'esame oltre che la valutazione nel caso fosse prevista oppure il giudizio nel caso di esami che non richiedano una valutazione numerica) sia informazioni sullo studente che ha sostenuto l'esame (comune di residenza, provincia di residenza, genere, voto del diploma di scuola superiore, voto conseguito al test d'ingresso, scuola superiore di provenienza e se lo studente usufruisce dell'iscrizione part-time). Inoltre, per distinguere gli studenti all'interno di una stessa coorte è stato fornito anche un numero progressivo unico associato ad ogni studente. Tale numero progressivo è stato inserito allo scopo di rendere anonime le informazioni in modo da non poter risalire a chi corrisponda una determinata informazione e sostituisce, assieme alla coorte, l'identificativo unico rappresentato dal numero di matricola.

La Tabella 1, contenente i dati forniti riguardo gli esami superati da uno studente, con le variabili che fanno parte della chiave composta evidenziate in grassetto, può essere osservata nella prossima pagina. La chiave composta (costituita da tre variabili) è stata scelta in quanto prog e coorte servono a distinguere univocamente uno studente da un altro (in quanto prog è un valore incrementale univoco all'interno di ciascuna coorte) mentre il codice univoco del corso permette di distinguere un esame da un altro e siccome uno studente durante la sua carriera può superare un esame (senza considerare eventuali rifiuti di valutazioni ritenute non soddisfacenti) una sola volta non ci possono essere più codici identificativi di un esame superato uguali per lo stesso studente.

NOME VARIABILE	TIPO	DESCRIZIONE
<b>prog</b>	<b>int PK</b>	<b>Numero progressivo di riconoscimento per uno studente</b>
CORSO	text	Codice del Corso di Laurea
<b>code</b>	<b>text PK</b>	<b>Codice identificativo esame superato</b>
datae	date	Data di superamento esame
tipo_esame	text	Tipologia esame
voto_esame	int	Votazione ottenuta
giud_esame	text	Eventuale giudizio sull'esame se previsto
crediti	int	Crediti ottenuti dal superamento dell'esame
descrizione_esame	text	Nome dell'esame superato
<b>coorte</b>	<b>int PK</b>	<b>Coorte d'immatricolazione dello studente</b>

Tabella 1: Tabella dei dati riguardanti gli esami superati dagli studenti

### 2.1 DATI SULLO STATO D'ISCRIZIONE

Oltre ai dati citati in precedenza sono stati messi a nostra disposizione i dati riguardanti lo stato della carriera di ogni studente aggiornati al 2021. Tali dati non sono stati usati per la nostra analisi principale in quanto sono stati messi a nostra disposizione ad aprile 2022 quando la nostra analisi era partita da oltre due mesi e la gran parte del lavoro era già stata svolta. Per questo motivo, come vedremo in seguito, questi dati hanno avuto un impiego più ristretto rispetto ai dati citati nella scorsa sezione.

I dati sullo stato di iscrizioni comprendono tre variabili per ogni studente, due delle quali, prog e coorte, utili per poter distinguere a quale studente corrisponda un determinato stato d'iscrizione mentre la terza, lo stato della carriera, permette di capire se lo studente paga ancora regolarmente le tasse per l'iscrizione oppure se abbia deciso di cambiare Corso di Laurea o università oppure di abbandonare gli studi. Ci sono infatti sei possibili stati d'iscrizione per uno studente: ancora iscritto (che paga regolarmente le tasse), rinuncia (che quindi ha presentato una rinuncia formale agli studi), passato ad un altro Corso di Laurea, trasferito ad un altro ateneo, laureato ed infine mancata iscrizione nel caso lo studente risulti ancora iscritto ma senza pagare le tasse.

Abbiamo riportato la struttura di questi nuovi dati nella Tabella 2 in cui abbiamo evidenziato, in grassetto, le variabili che abbiamo scelto per comporre la chiave composta ossia l'identificativo prog e la coorte di



immatricolazione che permettono di distinguere ciascuno studente da un altro.

NOME VARIABILE	TIPO	DESCRIZIONE
<b>prog</b>	<b>int PK</b>	<b>Numero progressivo di riconoscimento dello studente</b>
<b>coorte</b>	<b>int PK</b>	<b>Coorte d'immatricolazione dello studente</b>
stato_carriera	text	Stato dell'iscrizione dello studente

Tabella 2: Tabella dei dati riguardanti lo stato della carriera di uno studente



## OPERAZIONI PRELIMINARI E STRUMENTI USATI

---

In questo capitolo andremo ad approfondire come abbiamo "stimato" gli abbandoni e le operazioni necessarie per ricavare gli studenti inattivi che stimiamo aver abbandonato il Corso di Laurea con i relativi strumenti che abbiamo utilizzato per raggiungere tale scopo.

### 3.1 STIMARE L'ABBANDONO

Anzitutto, come accennato in precedenza, si è reso necessario stimare l'abbandono in base ai dati in nostro possesso: infatti era impossibile, con i dati che ci erano stati forniti inizialmente, conoscere con certezza quali studenti abbiano effettivamente abbandonato poiché, a differenza di quanto accade in altri stati esteri, in Italia è possibile sospendere la propria carriera universitaria per un periodo di tempo indefinito per poi riprendere gli studi in qualsiasi momento. Ciò rende quindi difficile, se non consideriamo i nuovi dati riguardanti lo stato d'iscrizione (i quali non permettono comunque di riconoscere con assoluta certezza se uno studente abbia effettivamente abbandonato o meno), separare in maniera precisa gli studenti che hanno abbandonato definitivamente gli studi da quelli che hanno preso un periodo di pausa. per questo si è reso necessario un criterio di separazione arbitrario che è stato individuato nel tempo trascorso dall'ultimo esame superato da un determinato studente al momento in cui sono stati raccolti i dati. La soglia di tempo che definisce uno studente come probabile abbandono è stata quindi individuata come due anni dall'ultimo esame sostenuto nel caso di studente che non usufruisca dell'iscrizione part-time e tre anni per gli studenti part-time. Questa distinzione esclude gli studenti delle coorti 2018 e 2019 che non verranno quindi tenuti in considerazione in questa analisi in quanto immatricolati da troppo poco tempo. Inoltre tale distinzione comporta che la nostra stima degli abbandoni si concentri sugli studenti inattivi, in tale categoria però troviamo tre categorie separate di studenti:

coloro che hanno presentato una rinuncia ufficiale agli studi, coloro che pur non avendo presentato una rinuncia ufficiale non hanno comunque intenzione di riprendere gli studi ed infine come ultima categoria coloro che hanno sospeso il proprio percorso accademico per un periodo indefinito di tempo con l'intenzione però un giorno di riprendere la loro carriera. L'ultima categoria è quella per noi più problematica in quanto, nonostante lo studente risulti essere inattivo, è impossibile conoscere le loro intenzioni future, ma, come accennato nell'introduzione, tale eventualità è stata considerata come poco probabile poiché è difficile che uno studente, dopo due o più anni di inattività, possa riprendere gli studi.

### 3.2 STRUMENTI UTILIZZATI

Per gestire i dati, forniti sotto forma di file Excel, ci siamo serviti innanzitutto di una base di dati MySQL, in particolare abbiamo usato MySQL Workbench come client per eseguire le nostre operazioni. Inoltre, abbiamo usato il software Stata (versione 17) per l'analisi statistica oltre che per rappresentare graficamente i dati forniti, compito per il quale è stato sfruttato anche MATLAB. Infine, si è reso necessario l'utilizzo di un foglio elettronico per convertire il formato originale dei dati forniti (ODS) in uno "leggibile" da MySQL ovvero CSV per poter caricare i dati nella nostra base di dati. Tale base di dati, basata sul modello relazionale in quanto stiamo usando MySQL, è stata strutturata su due Tabelle di cui una contenente gli esami svolti da tutti gli studenti che è già stata illustrata in precedenza ed un'altra contenente tutti gli studenti immatricolati nelle coorti dal 2010 al 2019 che illustreremo in seguito. Oltre ciò abbiamo anche un numero considerevole di viste su queste due Tabelle che contengono informazioni sulle analisi riportate in seguito, esempi di tali viste possono essere gli studenti che abbiamo individuato come inattivi o quelli che sono riusciti a laurearsi.

### 3.3 CARICAMENTO E PULIZIA DEI DATI

Una volta ottenuti i dati si è reso necessario anzitutto il loro caricamento nella tabella della base di dati creata per ospitarli. Successivamente è stata effettuata anche un'operazione di controllo e di successiva pulizia dei dati a causa della presenza di valori accidentali o di metodi di valutazione diversi da quelli attuali nel caso del voto di maturità, che risultavano assurdi nel contesto della base di dati. Tali dati sono stati quindi sostitui-

ti con valori corretti (o standardizzati in centesimi nel caso del voto di maturità). Nella pulizia dei dati è stata tenuta di conto anche la presenza di valori nulli che sono stati sostituiti con valori di default ignorati oppure inseriti in categorie apposite nelle analisi che eseguiremo in seguito. Alcuni esempi di tale lavoro di pulizia sono i dati di residenza e scuola superiore di provenienza, per cui abbiamo creato due categorie apposite che illustreremo nei capitoli successivi. In particolare però il lavoro di pulizia dei dati è stato di fondamentale importanza per individuare gli studenti che non avevano superato nessun esame: tali studenti sono stati individuati grazie alla mancanza dei campi del codice univoco e della descrizione dell'esame superato che hanno permesso quindi la ricerca e l'individuazione di tali studenti.

### 3.4 AGGREGAZIONE DEGLI ESAMI SVOLTI DA UNO STUDENTE

Per poter avere una panoramica migliore sulla situazione di ogni singolo studente e per essere in grado di stabilire quali studenti siano inattivi, si è reso necessario aggregare gli esami svolti da ogni studente per ottenere così informazioni su quanti crediti uno studente abbia conseguito fino al 2020 e sulla data dell'ultimo esame superato che si aggiungono alle informazioni che già conosceamo in precedenza. Per fare ciò abbiamo sfruttato l'opzione `group by` di MySQL, che permette di raggruppare elementi di una Tabella in base a dei parametri specificati, che nel nostro caso sono stati il numero progressivo e la coorte di appartenenza di tale studente. Volendo quindi ottenere informazioni sui crediti totali ottenuti da uno studente e sulla data dell'ultimo esame superato abbiamo inserito come parametri di selezione della query, attraverso la quale abbiamo generato la vista contenente tali informazioni riguardo gli studenti, anche la somma dei crediti degli esami registrati (con la funzione `sum` di MySQL) e la data più recente (ottenuta con la funzione `max`, che applicata ad un raggruppamento di date restituisce quella più recente) di registrazione di un esame. La struttura della Tabella contenente gli studenti che abbiamo ottenuto è riportata nella Tabella 3.

Abbiamo scelto per comporre la nostra chiave composta `prog` e `coorte` in quanto, come già esposto in precedenza, permettono insieme di riconoscere univocamente ciascuno studente.

### 3.5 RAPPRESENTAZIONE ED ORGANIZZAZIONE DEI DATI

Una volta illustrati i dati in nostro possesso nelle Tabelle 1 e 3 possiamo andare ad illustrare il diagramma entità-relazione che è servito ad evidenziare la relazione che intercorre fra le due Tabelle che abbiamo costruito nelle sezioni precedenti. Tale diagramma è riportato nella Figura 1.

Possiamo notare come la relazione sia di uno a molti (con cardinalità minima 0) sia per gli studenti in quanto uno studente può aver superato da zero a molti esami nella propria carriera sia similmente anche per i vari esami che possono essere superati da uno o da più studenti (considerando anche gli esami a scelta).

### 3.6 TROVARE GLI STUDENTI LAUREATI

L'ultima informazione sugli studenti che ci resta da ottenere è quella riguardante gli studenti laureati. Tale informazione ci è infatti di fondamentale importanza in primo luogo per rapportarla al numero di abbandoni ma anche per avere un quadro più completo sull'evoluzione temporale delle varie coorti. Per ricavare tale informazioni è stato sufficiente raggruppare in una vista tutti gli studenti che risultano aver superato la prova finale di laurea. Il numero di studenti laureati nelle corti che stiamo analizzando aggiornato al 2020 risulta quindi essere di 154 studenti.

NOME VARIABILE	TIPO	DESCRIZIONE
<b>prog</b>	<b>int PK</b>	<b>Numero progressivo di riconoscimento</b>
sex	text	Genere dello studente
nazione	int	Nazione di provenienza
comres	text	Comune di residenza
pres	text	Provincia di Residenza
resid	text	Studente residente in Italia o meno
esami	text	Lista di esami superati dallo studente
part_time	int	Studente che usufruisce dell'iscrizione part-time o meno
data_ultimo_esame	date	Data di superamento dell'ultimo esame
crediti	int	Totale dei crediti maturati da uno studente
punti_mod_mat_base	int	Punti ottenuti al test d'ingresso
scuola_tipo	text	Scuola superiore di provenienza
voto_diploma	int	Votazione del diploma di scuola superiore
<b>coorte</b>	<b>int PK</b>	<b>Coorte d'immatricolazione dello studente</b>

Tabella 3: Tabella dei dati che ci siamo ricavati riguardanti gli studenti iscritti

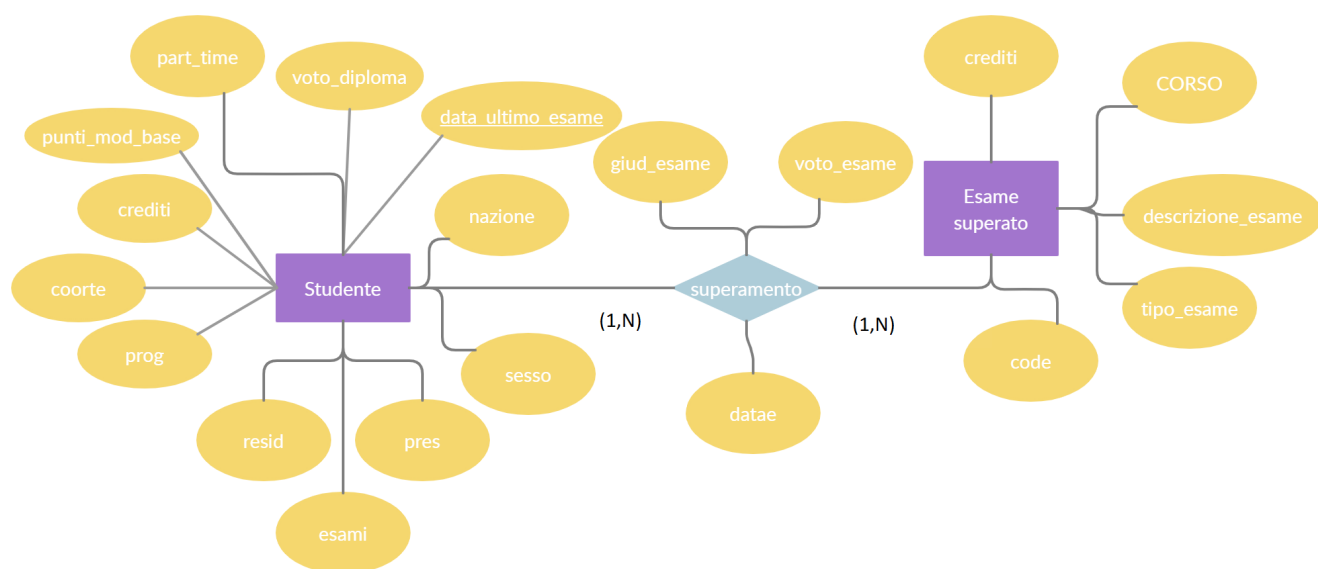


Figura 1: Diagramma Entità Relazione delle Tabelle degli studenti e degli esami superati



## STUDENTI INATTIVI

---

In questo capitolo ci occuperemo di trovare e di iniziare a caratterizzare gli studenti inattivi che abbiamo definito nel capitolo precedente dividendoli in base ai crediti che essi hanno già conseguito.

### 4.1 TROVARE GLI STUDENTI INATTIVI

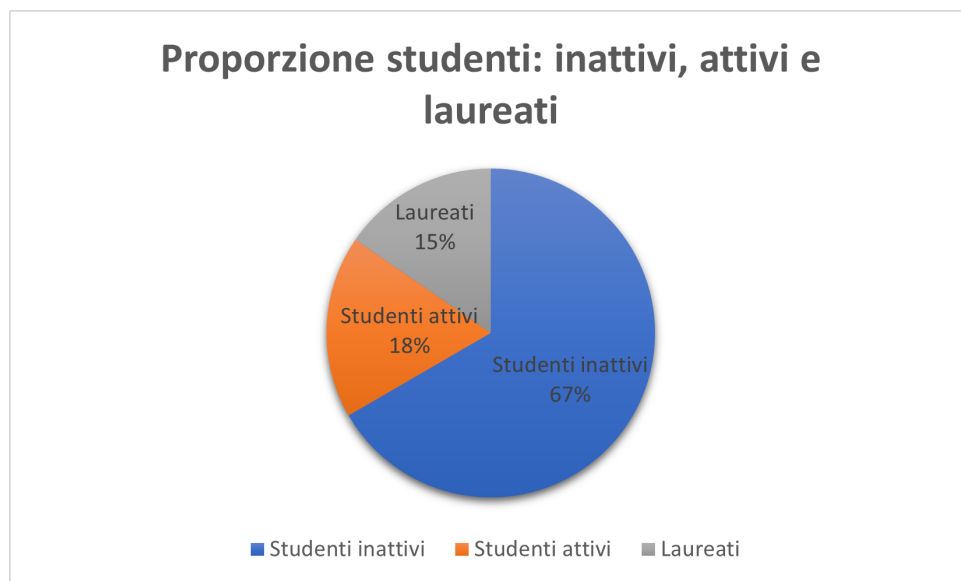


Figura 2: Proporzione studenti inattivi, attivi e laureati

Una volta che abbiamo ottenuto informazioni sugli studenti iscritti al Corso di Laurea ed una definizione di studente inattivo possiamo distinguere gli studenti inattivi da quelli che riteniamo essere ancora attivi. Per fare ciò è stato sufficiente fare una selezione su MySQL con le condizioni specificate in precedenza, ossia il fatto che per essere inattivo deve possedere meno di 165 crediti e non aver sostenuto nessun esame negli ultimi due anni (tre anni nel caso di studenti part-time), ed inserire i

risultati di tale operazione in una vista. La proporzione degli studenti che riteniamo aver abbandonato gli studi risulta quindi essere di circa due terzi del totale degli immatricolati nelle corti fino al 2017 ovvero gli studenti che abbiamo preso in considerazione per la nostra analisi. Inoltre, per completezza, abbiamo inserito nella proporzione anche gli studenti laureati (che abbiamo individuato raggruppando gli studenti che hanno superato la prova finale) che risultano essere circa il 17 per cento del totale degli studenti immatricolati dal 2010 al 2017. Il grafico con le proporzioni precedentemente specificate è visibile nella Figura 2.

#### 4.2 DISTINGUERE GLI STUDENTI INATTIVI IN BASE AI CREDITI

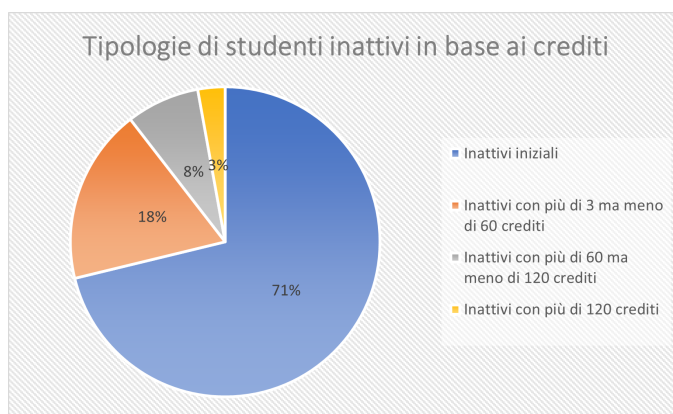


Figura 3: Tipologie di studenti inattivi divisi per crediti

Avendo individuato gli studenti inattivi, che supponiamo aver abbandonato gli studi, possiamo effettuare un'ulteriore suddivisione sulla base della quantità di crediti posseduti da uno studente. Definiamo come inattivi iniziali gli studenti che hanno conseguito tre crediti o meno (ovvero che al massimo hanno superato l'abilitazione d'Inglese), inattivi al primo anno se ne hanno conseguiti più di tre ma meno di 60, al secondo anno tra 60 e 120 e al terzo anno i restanti. Questa divisione è stata effettuata per mostrare a quale punto della carriera è più probabile che uno studente decida di lasciare il regolare percorso di studi. Come riportato nel grafico mostrato nella Figura 3 la maggior parte degli studenti che abbiamo individuato come inattivi non è stata in grado di conseguire più di tre crediti complessivi (parte blu del grafico) ed anche la maggior parte di coloro che hanno superato almeno un esame non sono andati poi oltre i sessanta crediti iniziali.

## CONFRONTO DEI DATI

In questo capitolo ci occupiamo di confrontare i dati degli studenti inattivi con tutti gli iscritti al Corso di Laurea di Informatica per riuscire a trovare qualche parametro che distingua gli studenti inattivi dalla popolazione generale degli studenti iscritti.

### 5.1 IL VOTO DI MATURITÀ PUÒ AVERE UN'INFLUENZA?

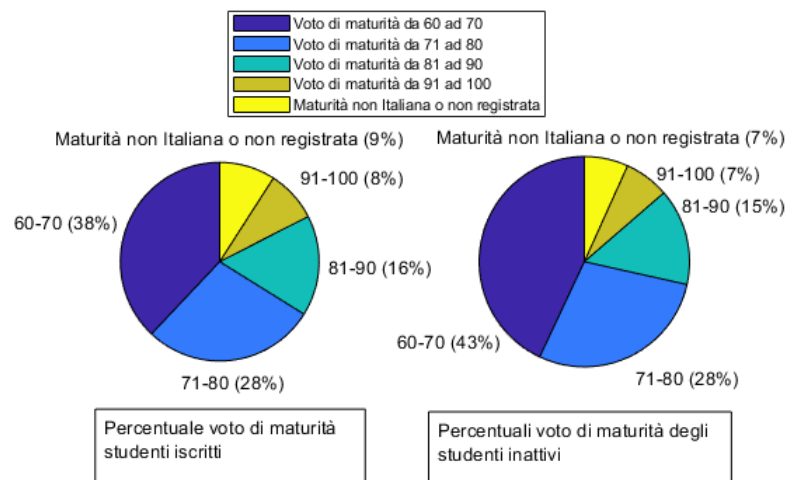


Figura 4: Distribuzione voto di maturità degli studenti iscritti e di quelli inattivi

Per rispondere a questo quesito è stato necessario raggruppare i voti di maturità in cinque categorie di cui quattro contengono i voti raggruppati per decine (60-70, 71-80, 81-90, 91-100) mentre l'ultima è stata usata per coloro che o non hanno registrato un diploma di maturità italiano o che, per qualche ragione, non hanno registrato un diploma di scuola superiore valido. Successivamente abbiamo messo a confronto i grafici

(riportati nella Figura 4) della proporzione dei voti di maturità tra tutti gli studenti e tra gli studenti inattivi. Ci siamo ricavati inoltre il grafico della Figura 5, riportante la distribuzione del voto di maturità tra gli studenti attivi o che hanno già conseguito la laurea. Da questi grafici possiamo notare che, seppur con qualche discrepanza (dovuta in parte anche alla presenza di diplomi non registrati o non italiani), i tre grafici non risultano essere troppo dissimili tra di loro anche se nel grafico degli attivi possiamo notare una minore presenza di votazioni tra il 60 ed il 70 rispetto agli inattivi mentre il resto delle categorie non sembra evidenziare differenze molto marcate, il che potrebbe indicare che un voto molto basso può rappresentare un lieve fattore di rischio verso il successo.

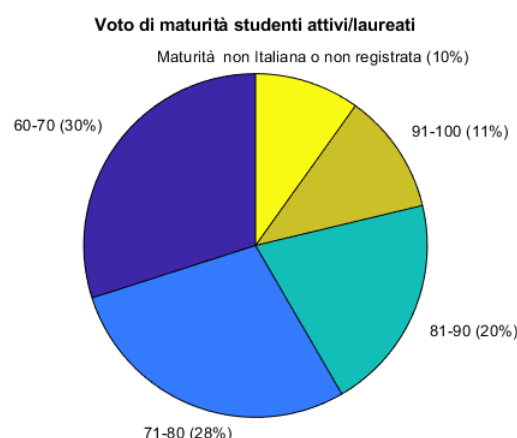


Figura 5: Distribuzione voto di maturità degli studenti iscritti e di quelli inattivi

## 5.2 LA DISTANZA DAL LUOGO DI RESIDENZA INFLUISCE?

Un'altra domanda che può essere interessante porsi è quanto la distanza del luogo di residenza dall'università influisca sull'inattività di uno studente, per rispondere a tale quesito dividiamo sia tutti gli studenti che quelli inattivi in quattro categorie: residenti nel territorio della provincia di Firenze, residenti nel resto della Toscana, nel resto d'Italia e quelli che sono residenti all'estero. Le proporzioni delle categorie sopracitate sono riportate nel grafico della Figura 6. Anche in questo caso è possibile notare come non ci siano differenze di notevole rilevanza nelle proporzioni di provenienza tra studenti ed il sottogruppo di studenti

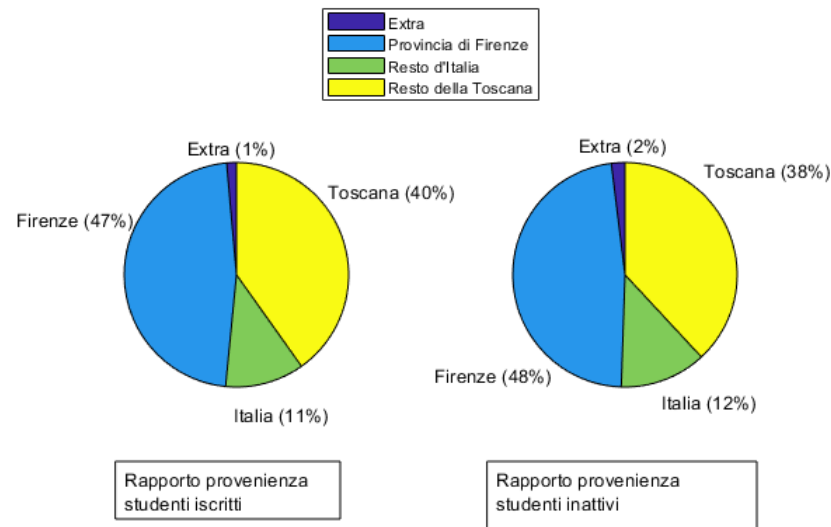


Figura 6: Distribuzione della provenienza degli studenti iscritti e di quelli inattivi

inattivi che abbiamo ipotizzato aver abbandonato gli studi. Inoltre anche analizzando la distribuzione del luogo di provenienza degli studenti che risultano essere attivi o laureati (Figura 7) non sembrano esserci differenze di particolare rilievo.

Il fatto che la distanza del luogo di residenza non influisca sulle possibilità di abbandonare il corso può essere utilizzata, nella fase di orientamento in ingresso, per convincere un eventuale studente fuori sede ad iscriversi al Corso di Laurea.

### 5.3 IL TEST D'INGRESSO PUÒ PREVEDERE L'INATTIVITÀ?

Gli studenti che si iscrivono al Corso di Laurea devono sostenere una prova di verifica delle conoscenze in ingresso. Il mancato superamento di tale prova comporta degli obblighi formativi aggiuntivi, possiamo voler quindi sapere se il superamento di questa prova possa essere un indicatore sulla futura inattività da parte di uno studente. Il problema però è che negli anni sono stati usati due diversi metodi di valutazione: fino al 2015 il test prevedeva un punto per risposta corretta e zero nel caso di risposta incorretta o mancante mentre dal 2016 ad oggi prevede sempre un punto per risposta corretta e zero per risposta mancante

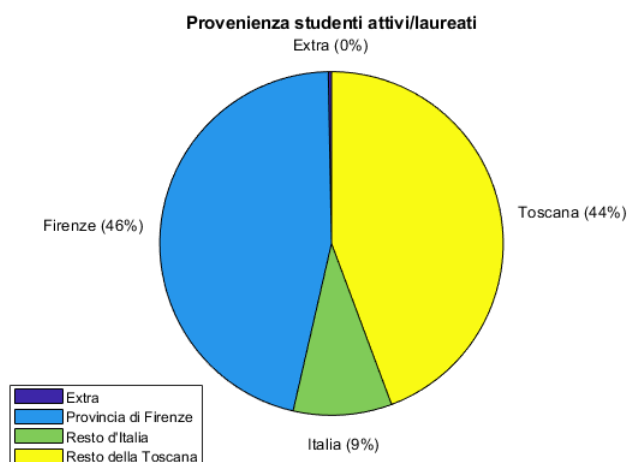


Figura 7: Distribuzione della provenienza degli studenti attivi

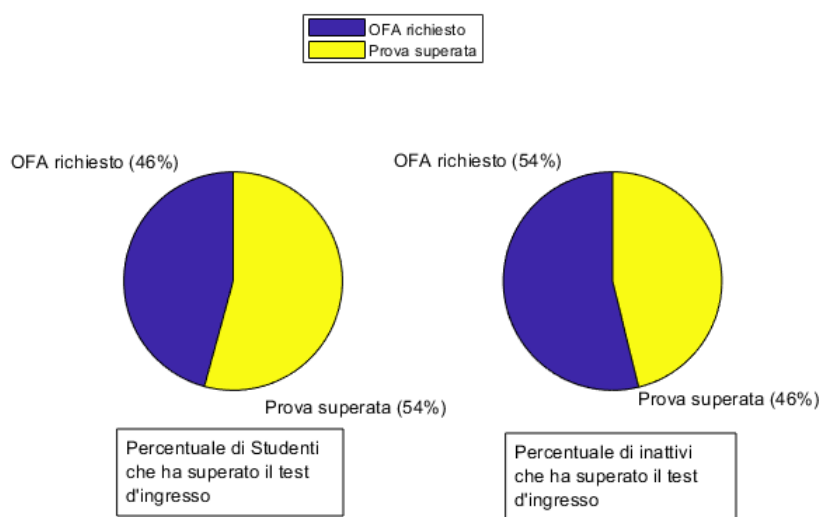


Figura 8: Distribuzione superamento prova d'ingresso degli studenti iscritti e di quelli inattivi

mentre per la risposta errata è prevista una penalizzazione di un quarto di punto. Per questo motivo è stato preferito considerare separatamente coloro che avevano sostenuto una tipologia di prova diversa. Una volta che abbiamo trovato coloro i quali hanno superato il test o meno è stato possibile riaggregare i dati per condurre quindi uno studio più genera-

le sugli studenti che hanno passato o meno il test a prescindere dalla tipologia di test effettuata. Le proporzioni trovate sono evidenziate nel grafico riportato nella Figura 8.

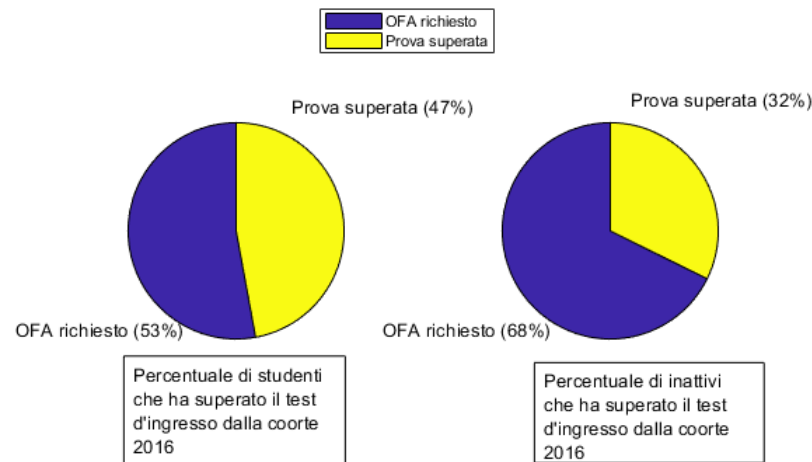


Figura 9: Distribuzione in base al superamento della prova d'ingresso degli studenti iscritti e di quelli inattivi delle coorti dal 2016 in su

Possiamo notare come le due proporzioni nel grafico abbiano una differenza relativamente marcata; anche mettendo a confronto le proporzioni con il comando "prtest" di Stata è possibile notare come le due proporzioni risultino significativamente diverse con una differenza di più del 8% tra i due gruppi. Possiamo poi, per completezza, voler osservare se il nuovo metodo di valutazione della prova preso singolarmente riesca a "predire" in maniera più accurata i possibili inattivi. Nella Figura 9 sono riportati i grafici delle proporzioni di superamento della prova per le coorti che hanno sostenuto la nuova prova.

Possiamo notare come il nuovo metodo di valutazione identifichi con maggior accuratezza i futuri inattivi rispetto alla proporzione aggregata dei grafici precedenti (con entrambi i metodi di valutazione), per cui possiamo affermare che il nuovo metodo sia in grado di predire con più accuratezza quali studenti diventeranno inattivi rispetto al precedente.

Un ulteriore conferma della validità della "nuova modalità" del test d'ingresso è data dal grafico della Figura 10, rappresentante la distribuzione del superamento del test d'ingresso nelle coorti che hanno sostenuto il test con la nuova modalità, nella quale notiamo che quasi due terzi degli studenti attivi o laureati hanno superato la prova mentre per gli inattivi era quasi esattamente l'opposto

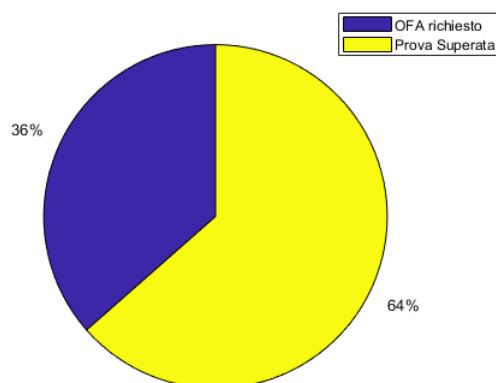


Figura 10: Distribuzione in base al superamento della prova d'ingresso degli studenti attivi o laureati delle coorti dal 2016 in su

#### 5.4 IL GENERE HA UNA QUALCHE INFLUENZA SULL'INATTIVITÀ?

Il Corso di Laurea in Informatica è frequentato in maggioranza da studenti di genere maschile (87%), potrebbe essere interessante sapere se questa caratteristica possa in qualche modo influenzare l'inattività delle studentesse di genere femminile che si trovano ad essere in minoranza. Al solito per verificare tale ipotesi riportiamo nella Figura 11 i grafici con le proporzioni del genere negli studenti iscritti e negli studenti che risultano essere inattivi.

Come possiamo notare anche ad occhio dal grafico le due proporzioni sono praticamente uguali e perciò dai dati in nostro possesso non è possibile attribuire al genere nessun effetto significativo sull'abbandono di uno studente, anzi possiamo notare come le studentesse tendano addirittura ad abbandonare meno rispetto ai loro colleghi di genere maschile. Possiamo ulteriormente convincerci di ciò sia con un test statistico sull'uguaglianza del tasso d'inattività fra studenti di genere maschile e femminile (il cui risultato non permette di confutare tale ipotesi) che analizzando separatamente la distribuzione, riportata nella Figura 12, del genere negli studenti attivi e laureati delle coorti analizzate in precedenza (2010-2017) in cui la percentuale di studentesse risulta essere addirittura maggiore rispetto alle due distribuzioni precedenti.



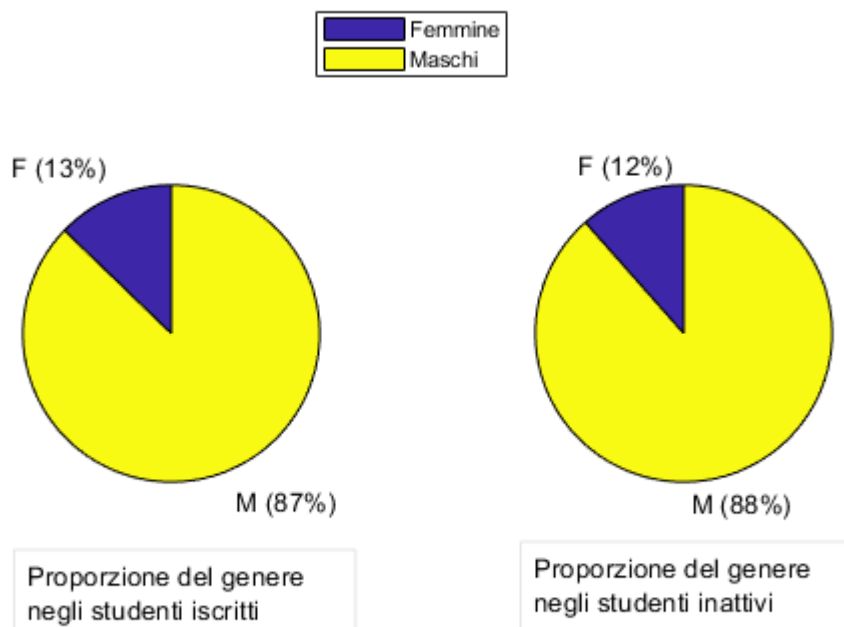


Figura 11: Distribuzione del genere negli studenti iscritti ed inattivi

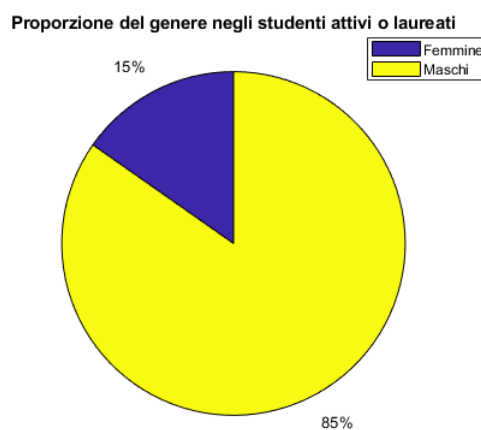


Figura 12: Distribuzione del genere negli studenti attivi

## 5.5 CI SONO SCUOLE SUPERIORI “MIGLIORI”?

Un altro confronto che potrebbe risultare di nostro interesse è quanto la scuola superiore di provenienza influisca sulle proporzioni dell’inat-

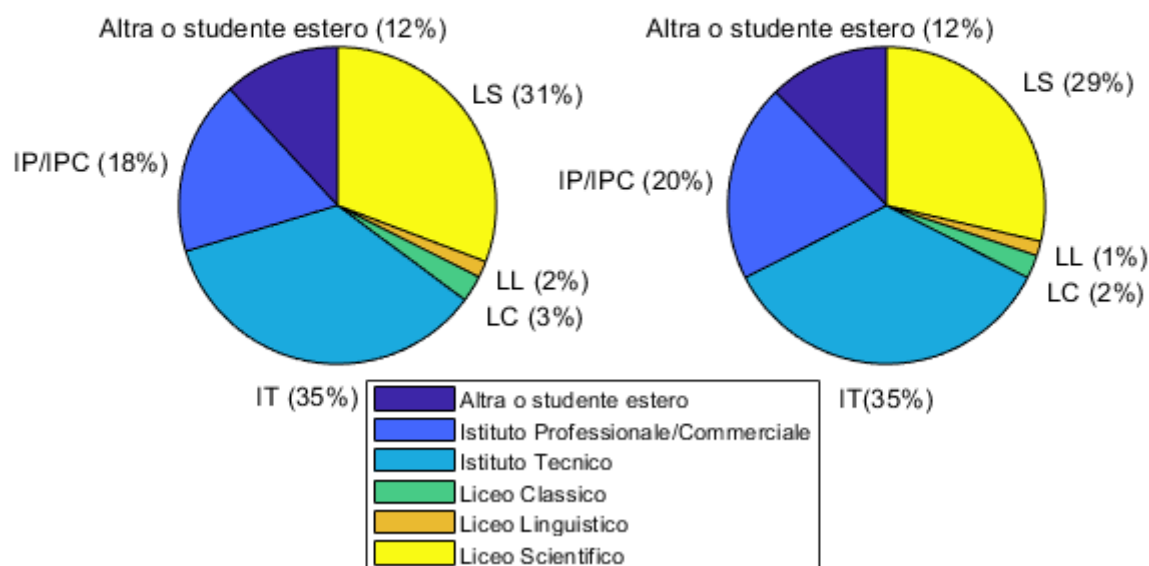


Figura 13: Distribuzione del tipo di scuola superiore di provenienza negli studenti iscritti ed inattivi

tività. Per verificare ciò dividiamo gli studenti in sei categorie in base alla scuola superiore di provenienza: Liceo Scientifico, Liceo Classico, Liceo Linguistico, Istituto Tecnico, Istituto Professionale o Commerciale e infine ne è stata creata un'ultima per contenere gli studenti che o provengono dall'estero o la cui scuola non appartiene alle categorie precedentemente illustrate. Queste distinzioni sono state effettuate sulla base delle tabelle fornite dal Ministero dell'Istruzione che non permettono di risalire alla scuola precisa di provenienza di uno studente (a meno che esso non provenga da un liceo classico, scientifico o linguistico), in quanto numerosi indirizzi (ad esempio tutti gli istituti tecnici) sono raggruppati sotto un'unica sigla, il che rende impossibile, ad esempio, osservare i dati separatamente per chi proviene da un Istituto Tecnico ad indirizzo informatico rispetto a chi proviene da un indirizzo generico. Nella Figura 13 possiamo osservare la distribuzione del tipo di scuola superiore di provenienza negli studenti iscritti ed inattivi mentre nella 14 possiamo osservare quella degli studenti classificati come attivi o laureati. Da quanto è possibile osservare comparando i grafici non vi è alcuna differenza significativa nelle tre proporzioni, se non forse per una maggiore

presenza di studenti provenienti da un liceo scientifico nella distribuzione degli studenti attivi o laureati.

In virtù di ciò possiamo affermare che la scuola di provenienza, in base ai dati forniti, non influenzi significativamente un'eventuale inattività dei futuri studenti. Non è possibile però osservare, come spiegato in precedenza, se all'interno delle varie categorie le proporzioni rimangono costanti o se ci siano delle variazioni significative all'interno di esse.

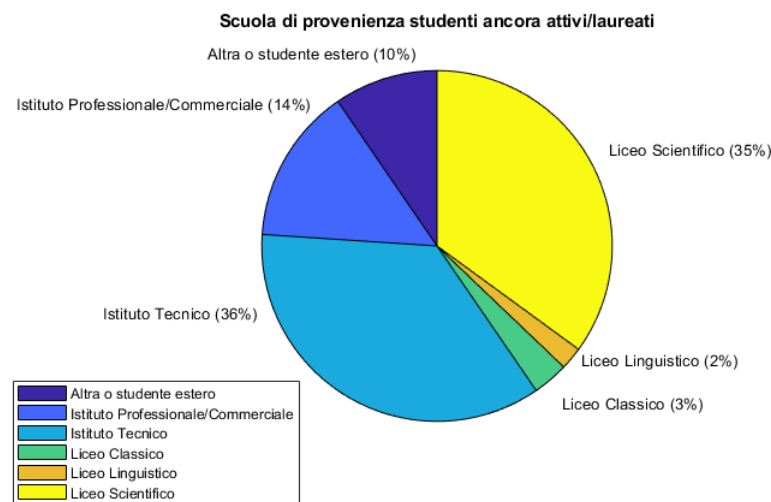


Figura 14: Distribuzione del tipo di scuola superiore di provenienza negli studenti attivi o laureati

## 5.6 LA MEDIA DEI VOTI PUÒ INFLUENZARE L'INATTIVITÀ?

Ora effettuiamo un'analisi differente rispetto alle precedenti andando ad osservare se esiste una differenza significativa nella media tra gli studenti inattivi, che riteniamo aver abbandonato, ed il resto degli studenti delle coorti precedenti al 2018 (anno di immatricolazione che abbiamo imposto come limite per l'inattività). L'obiettivo è verificare se una determinata media possa in qualche modo scoraggiare uno studente a proseguire gli studi. Per fare ciò abbia usato la funzione "ttesti" di Stata che ci ha permesso, inserendo il numero, la varianza e la media ponderata dei voti di entrambi i gruppi, che ci siamo ricavati precedentemente attraverso le funzioni "mean" e "std" di MySQL per la media e la deviazione standard (che non è altro che la radice quadrata della varianza), di eseguire un test statistico di uguaglianza delle medie. Per ricavarci inve-

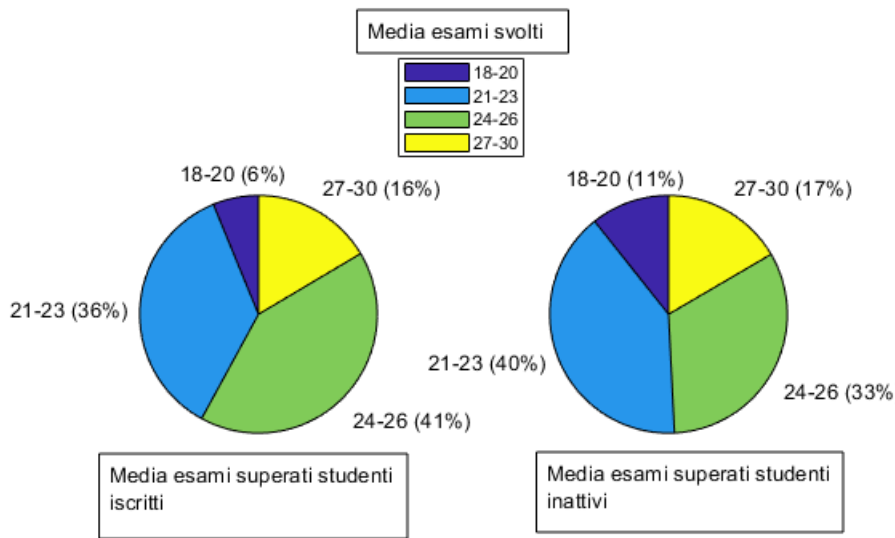


Figura 15: Distribuzione della media dei voti degli esami superati negli studenti iscritti ed inattivi

ce la media ponderata è stato necessario applicare la formula algebrica  $\frac{(\sum x_i * n_i)}{\sum n_i}$ , dove  $x_i$  indica la valutazione conseguita in un esame mentre  $n_i$  il numero di che crediti relativo all'esame, che abbiamo applicato ad una query di selezione degli esami che hanno ricevuto una valutazione.

I risultati mostrano una differenza nelle due medie ponderate di poco più di mezzo voto (0.59), che seppur piccola risulta altamente significativa, il che potrebbe indicare che le persone con una media più bassa tendono a diventare inattive con più frequenza ma potrebbe anche indicare il fatto che gli studenti nel proseguire i loro studi tendano ad aumentare la loro media col tempo (cosa che chi abbandona precocemente non può sperimentare). Quanto osservato in precedenza è confermato anche dai grafici che illustrano la differenza nelle proporzioni delle medie tra tutti gli studenti e gli inattivi riportati nella Figura 15. Possiamo notare come gli studenti che hanno una media compresa tra 18 e 23 siano in proporzione nettamente maggiore tra gli inattivi (più del 35%) rispetto a tutti gli altri studenti (in cui la proporzione è circa del 25%) il che conferma il test svolto in precedenza. Inoltre, anche analizzando la distribuzione della media ponderata dei voti degli esami superati dagli studenti attivi o laureati riportata nella Figura 16, possiamo ulteriormente affermare

come ci sia effettivamente una differenza nella media ponderata degli esami svolti da studenti attivi o laureati e da studenti che invece sono in seguito diventati inattivi

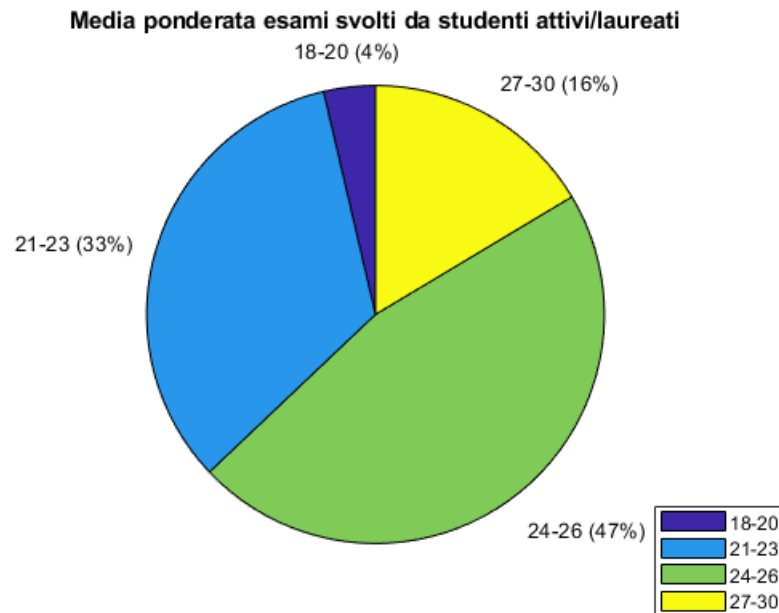


Figura 16: Distribuzione della media ponderata dei voti degli esami superati dagli studenti attivi o laureati

### 5.7 I CREDITI CONSEGUITI AL PRIMO ANNO CONDIZIONANO L'INATTIVITÀ?

In quest'ultimo confronto cerchiamo di stabilire se, come evidenziato nello studio svolto presso l'Università di Bari [5], il numero di crediti che uno studente consegue nel primo anno di studi possa in qualche modo influenzare un'eventuale futura inattività.

Abbiamo innanzitutto dovuto ricavare quali esami fossero stati svolti nel primo anno di studi: per fare ciò è bastato controllare quali esami uno studente avesse superato fino alla fine dell'anno solare successivo a quello di immatricolazione, infatti nel primo semestre di solito uno studente non sostiene esami se non quello di Inglese o l'eventuale prova d'ingresso che però non assegna crediti e quindi possiamo considerare come primo anno effettivo l'anno successivo a quello di immatricolazione. Per fare un esempio uno studente immatricolato nel 2019 difficilmente so-

sterà esami nel 2019 ma inizierà a sostenere esami dal 2020 per questo avrebbe poco senso considerare un anno a partire dall'anno effettivo di immatricolazione nella nostra analisi.

Detto ciò non resta altro che aggregare gli esami svolti nel primo anno per ogni studente e confrontare gli studenti inattivi con tutti gli studenti iscritti fino al 2018 con un test statistico sulla media dei crediti conseguiti. Il risultato di tale test indica come vi sia una differenza molto significativa tra le due medie di circa 9 crediti, il che conferma lo studio precedentemente citato. Inoltre questa differenza la ritroviamo anche nei valori mediani di entrambi i gruppi che sono di 15 crediti nel caso di studenti inattivi e di 27 per tutti gli studenti. Un'ultima conferma la possiamo avere confrontando i grafici delle figure 17 e 18 che evidenziano ulteriormente la differenza di proporzione nei crediti conseguiti nel primo anno tra studenti iscritti e inattivi. Oltre ciò possiamo osservare come gli studenti attivi differiscano in maniera consistente dalle due distribuzioni citate in precedenza in quanto la maggior parte degli studenti attivi o laureati hanno ottenuto un numero di crediti al primo anno compreso tra 31 e 45, mentre la maggior parte degli inattivi e degli studenti in generale ha ottenuto un numero di crediti compreso tra 1 e 15 crediti nel primo anno. Per questa analisi non abbiamo tenuto in considerazione gli studenti che non hanno superato nessun esame in quanto in primis avrebbero condizionato troppo l'analisi facendo risultare il distacco ancora più marcato ed in secondo luogo non avendo superato nessun esame al primo anno sarebbe risultato contraddittorio inserire anche questo gruppo nella nostra analisi.

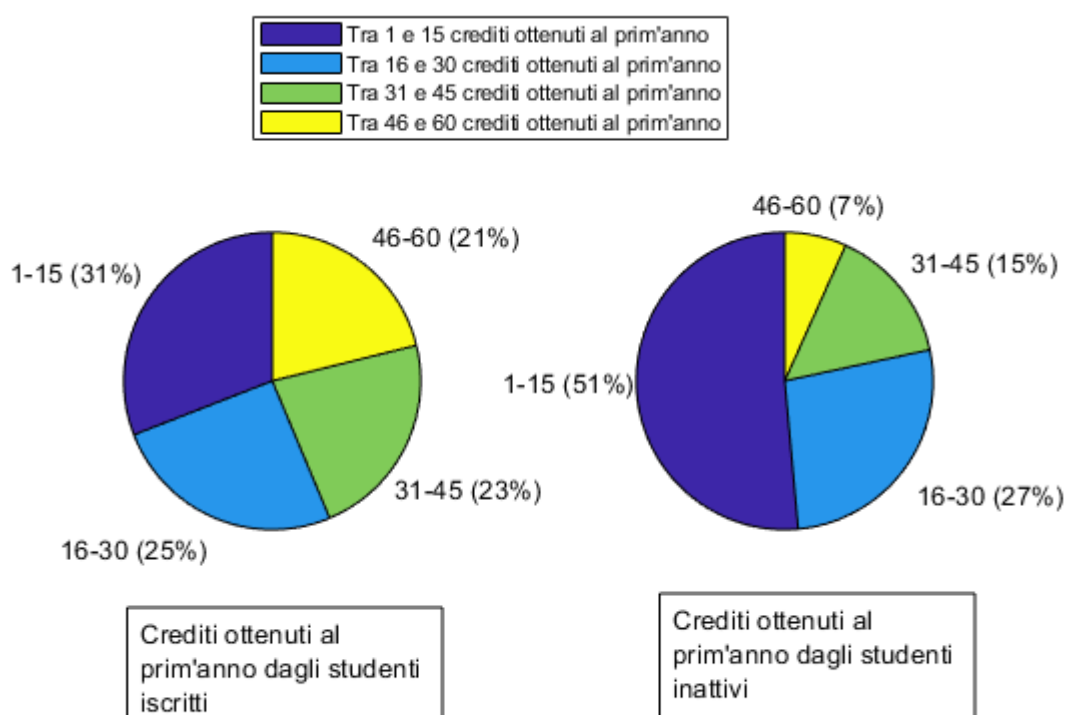


Figura 17: Crediti conseguiti al primo anno dagli studenti iscritti ed inattivi

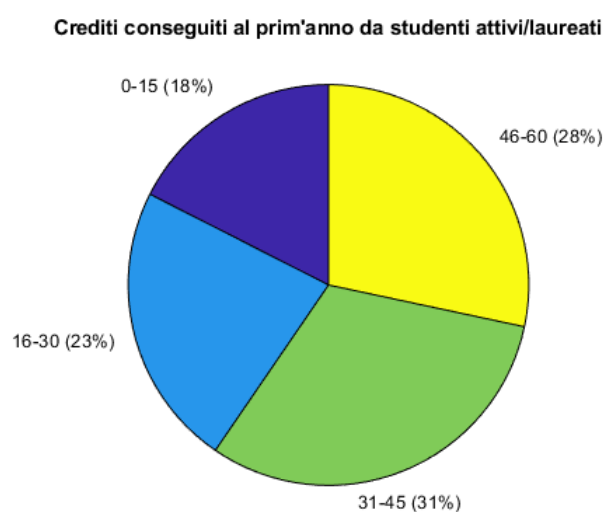


Figura 18: Crediti conseguiti al primo anno da studenti attivi o laureati





## ANALISI ED EVOLUZIONE TEMPORALE

---

In questo capitolo ci occupiamo di caratterizzare ed analizzare gli studenti inattivi in base a parametri di tipo sia "temporale" (come il tempo che uno studente impiega prima di diventare tale od il tasso di abbandono per coorte di immatricolazione) sia di carattere più prettamente accademico come gli esami che gli studenti inattivi non riescono a superare.

### 6.1 L'ANDAMENTO DELL'ABBANDONO IN BASE ALLA COORTE DI IMMATRICOLAZIONE

Un'altra analisi che potrebbe risultare interessante è quella di come gli inattivi si distribuiscano in base alla coorte di immatricolazione dello studente. Per osservare ciò ci siamo in primo luogo, per ciascuna coorte, ricavati il numero di studenti inattivi (tenendo conto che abbiamo individuato gli inattivi fino alla coorte del 2017 non potendo quindi analizzare le coorti successive), in secondo luogo abbiamo diviso tali valori per il numero totale di studenti iscritti in quella coorte. Di seguito, nella Figura 19, è stato quindi riportato il grafico con i tassi di inattività così ottenuti. Possiamo notare che, a differenza di quanto fosse logicamente attendibile, non abbiamo una discesa continua (dovuta al fatto che con il tempo si presume che sempre più persone diventino inattive), né tantomeno lineare, ma abbiamo un "picco" di inattivi che ha come massimo la coorte del 2014 e che vede come minimo la coorte del 2016 invece che quella più "logica" che sarebbe quella più recente ovvero quella del 2017. Spiegare questo fenomeno non è affatto semplice, infatti stabilire una relazione causale in questo caso è parecchio azzardato soprattutto per la presenza di dati a noi oscuri (che possono essere esterni all'università stessa).

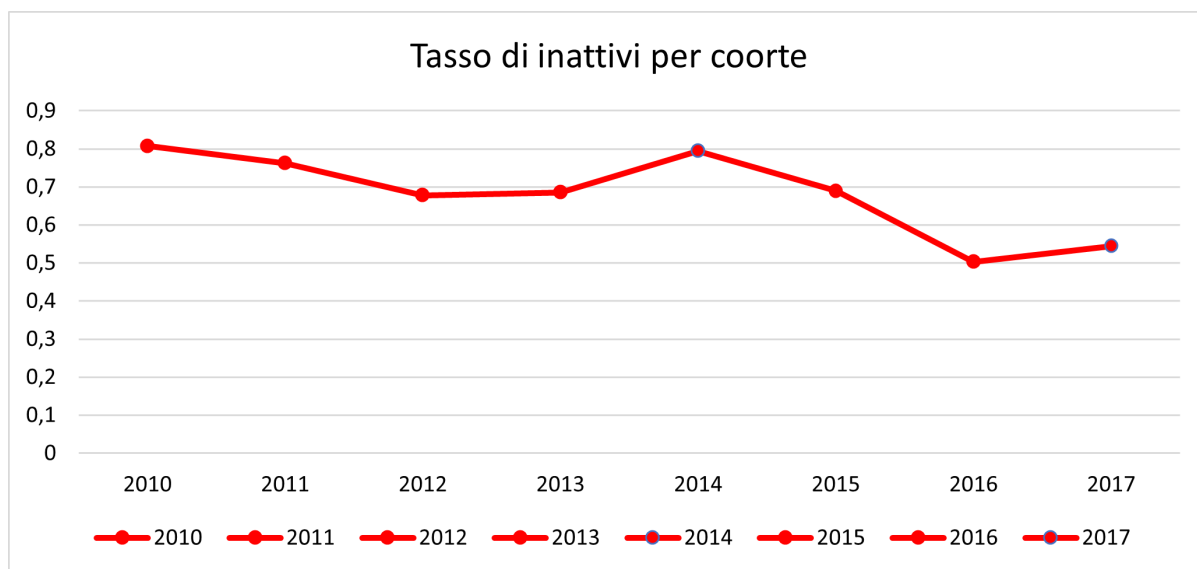


Figura 19: Proporzione di studenti inattivi sul totale per coorte

## 6.2 QUANTO TEMPO IMPIEGA UNO STUDENTE PRIMA DI DIVENTARE INATTIVO?

Per rispondere a tale quesito dobbiamo anzitutto definire cosa intendiamo per periodo prima che uno studente diventi inattivo. In questa analisi si è deciso di indicare con tale termine il numero di anni che intercorre tra l'anno di immatricolazione e l'anno in cui uno studente inattivo ha superato il suo ultimo esame (se non ne ha superato nessuno il periodo si considera uguale a zero). Questo implica che questa analisi può essere svolta solo sugli studenti che abbiamo individuato come inattivi. Per calcolare il periodo è stato necessario ricorrere alla funzione "year" di MySQL che permette di ricavare l'anno da una data passata in input per poter così ricavare l'anno in cui si presume che lo studente sia diventato inattivo dalla data dell'ultimo esame. A questo punto è bastato sottrarre per ciascuno studente la coorte d'immatricolazione dall'anno dell'ultimo esame superato per ottenere il valore desiderato. Una volta ottenuti i dati è stato possibile osservare che la media degli anni che uno studente "impiega" prima di diventare inattivo è di circa 0,74 anni (0,7406) con un intervallo di confidenza al 95% che va da 0,64 a 0,83 anni. Di seguito, nella Figura 20 è riportato anche l'istogramma contenente i valori appena analizzati. Come possiamo notare anche dal grafico la maggior parte degli studenti che abbandonano non "resiste" oltre il primo anno, ma questo era abbastanza prevedibile in quanto la maggior parte degli

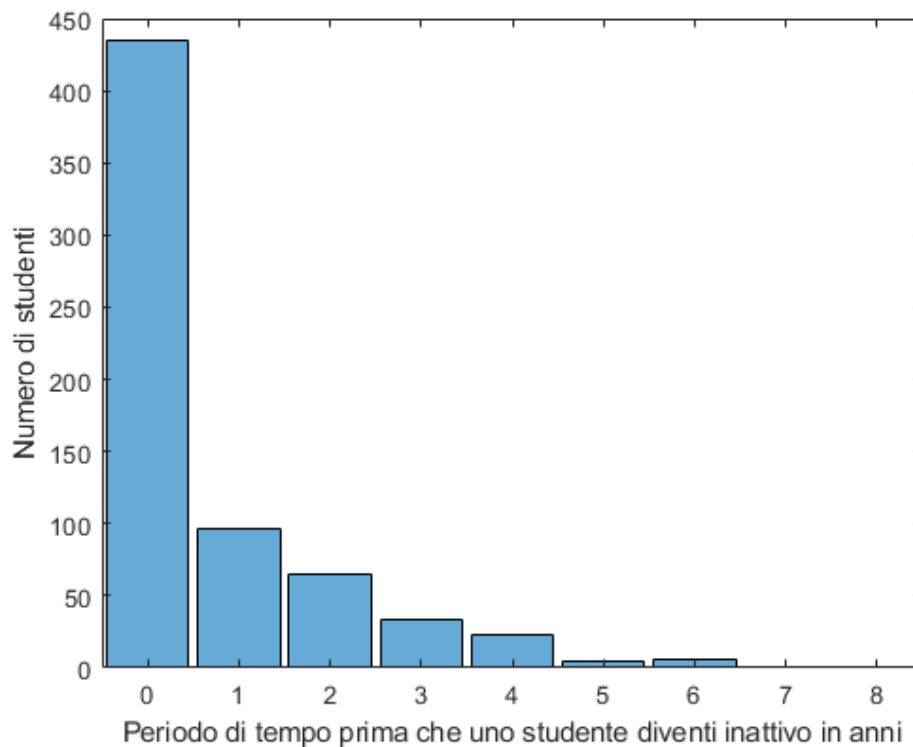


Figura 20: Grafico riportante il numero di studenti che è diventato inattivo dopo un certo numero di anni

studenti che abbandona non supera nemmeno un esame. Potrebbe essere quindi interessante analizzare gli studenti che invece sono riusciti a superare almeno un esame (esclusa la prova iniziale). Per fare ciò escludiamo dalla nostra analisi tutti coloro che non hanno conseguito nessun credito ed eseguiamo le stesse analisi svolte in precedenza. Questa volta la media del periodo dopo il quale uno studente abbandona risulta essere di più di due anni (2,06) con un intervallo di confidenza al 95% incluso tra 1.89 e 2.24 anni. Possiamo quindi osservare una discrepanza estremamente significativa fra la media del periodo di abbandono dei due gruppi, confermata anche dall'istogramma riportato di seguito nella Figura 21. Tale discrepanza è ulteriormente rilevabile osservando anche i valori mediani di entrambi i gruppi: zero anni se consideriamo coloro che non hanno superato nessun esame mentre due anni per coloro che invece sono riusciti a superare almeno un esame. Questa differenza così marcata potrebbe essere causata da diversi fattori: uno, come accennato in precedenza, è il fatto che la maggior parte degli studenti che

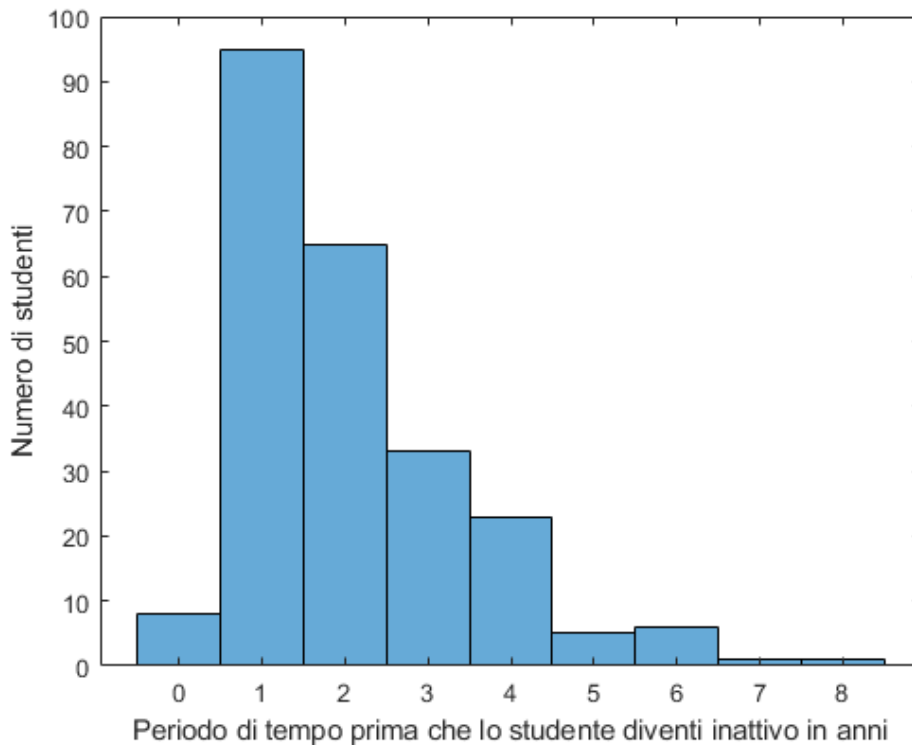


Figura 21: Periodo di tempo prima che uno studente che ha superato almeno un esame diventi inattivo espresso in anni

abbandona non abbia superato nemmeno un esame, un altro potrebbe essere il fatto che uno studente possa sentirsi maggiormente “motivato” dal superamento di un esame riducendo così la probabilità di diventare inattivo.

### 6.3 QUALI SONO GLI ESAMI CHE GLI STUDENTI INATTIVI NON RIESCONO A SUPERARE?

Per scoprire quali esami abbiano messo più in difficoltà gli studenti del corso di Laurea abbiamo osservato il numero di studenti che, dopo aver superato un determinato esame, hanno smesso di essere attivi. Per eseguire questa analisi è stato necessario ricorrere alla parola chiave “like” insieme al carattere jolly “%”, che, applicati ad una condizione su una stringa di una query MySQL, verificano se quella stringa contiene o meno i caratteri specificati all’interno dei due simboli %, questo per poter osservare quanti studenti abbiano superato determinati esami. Abbiamo

inoltre scelto di limitare gli esami su cui effettuare l'analisi a quelli inseriti nel primo anno del programma del Corso di Laurea in quanto, in primo luogo, nonostante i professori possano essere cambiati i corsi sono rimasti invariati sia per nome che per numero di crediti, il che permette di condurre un'analisi più "equa" poiché, nel proseguo del Corso di Laurea, gli esami differiscono sensibilmente in base alla coorte di immatricolazione (per non parlare poi degli esami a scelta del terzo anno in cui uno studente è libero di scegliere due corsi che abbiano attinenza con l'Informatica). In secondo luogo, molti esami del secondo e terzo anno sono bloccati da corsi del primo, il che renderebbe ancora più problematica un'analisi degli esami di altri anni oltre al primo. Ottenuti quindi, attraverso la funzione "count" di MySQL, il numero degli studenti inattivi che hanno superato un determinato esame possiamo andare a rappresentare i valori ottenuti attraverso un istogramma, riportato nella Figura 22. Per comprendere al meglio tale grafico riportiamo anche che il numero totale di studenti che sono diventati inattivi dopo aver superato almeno un esame è di 238 studenti. Possiamo quindi notare come gli esami più "problematici" risultino essere Architetture degli Elaboratori e soprattutto Matematica Discreta e Logica in quanto sono gli esami che gli inattivi tendono ad aver superato meno: infatti, solo circa il 25% degli studenti inattivi, che hanno superato almeno un esame, hanno superato almeno uno di questi due esami. Ovviamente va anche tenuto di conto che l'esame di Inglese richiede una preparazione nettamente diversa rispetto a quella necessaria per superare gli altri esami, per questo risulta così tanto preponderante nell'istogramma sopra riportato. Nonostante ciò, è comunque curioso notare come l'esame con meno crediti (certificazione di Inglese esclusa) ovvero Matematica Discreta e Logica sia anche quello che è stato passato da meno studenti inattivi. Questo fatto potrebbe essere spiegato sia dal fatto che gli studenti preferiscano concentrarsi sugli esami che forniscono un numero maggiore di crediti ma anche dal fatto che il non superamento dell'esame di Matematica Discreta e Logica non impedisca lo svolgimento degli esami dell'anno successivo a differenza degli altri esami del primo anno.

#### 6.4 EVOLUZIONE DELLA SCUOLA DI PROVENIENZA DEGLI STUDENTI INATTIVI

Nel corso degli ultimi anni il Corso di Laurea ha tentato, attraverso varie iniziative di orientamento in ingresso, di indirizzare gli studenti provenienti da scuole superiori che di solito non prenderebbero in considera-

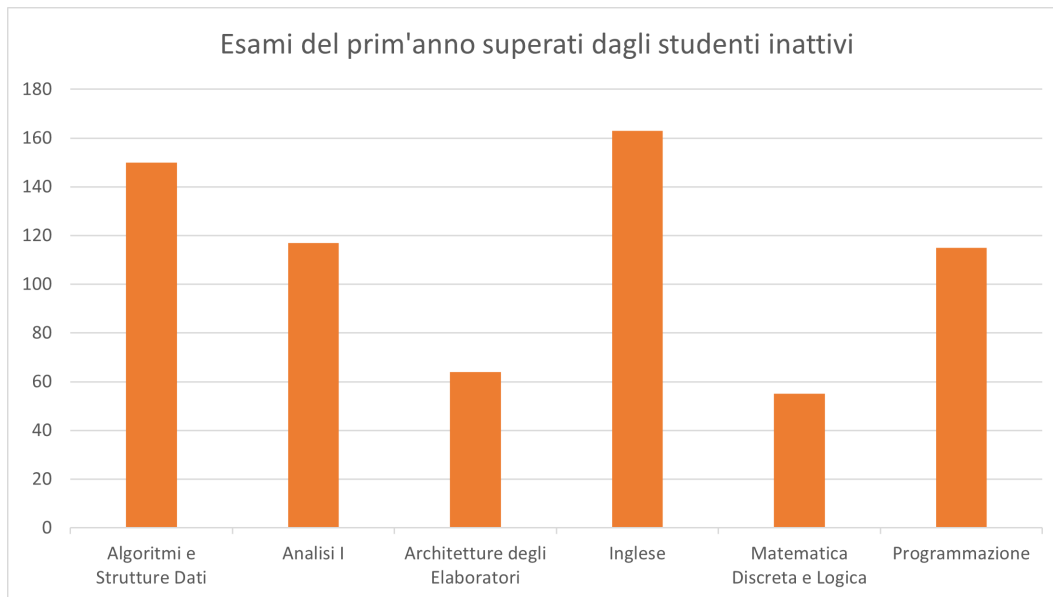


Figura 22: Esami del primo anno del Corso di Laurea superati da studenti inattivi

zione l'idea ad iscriversi al Corso. Potrebbe essere quindi interessante osservare se in qualche modo questi tentativi abbiano sortito un qualche effetto sugli studenti inattivi. Per verificare ciò è stato sufficiente raggruppare, attraverso l'opzione "group by" di MySQL, gli studenti per scuola di provenienza per ciascuna coorte di immatricolazione e contare il numero di studenti per ciascun gruppo di scuole grazie alla funzione "count". I raggruppamenti di scuole superiori sono gli stessi usati nell'analisi sugli inattivi per scuola superiore di provenienza. I risultati che abbiamo ottenuto sono stati quindi riportati nel grafico a barre impilate riportato nel grafico della Figura 23. Possiamo notare come negli ultimi tre anni si abbia una maggiore eterogeneità nella proporzione della scuola di provenienza, per rendersi conto di quanto questo fenomeno possa essere causato da una semplice differenza nella proporzione degli immatricolati alle varie coorti riportiamo anche tale proporzione nel grafico della Figura 24. Possiamo osservare come negli ultimi anni anche fra gli studenti iscritti vi sia molta più eterogeneità rispetto agli scorsi anni, per questo possiamo concludere che, pur incidendo sulla provenienza degli studenti iscritti, le iniziative di orientamento non sembrano avere effetto sulla scuola di provenienza degli studenti inattivi.

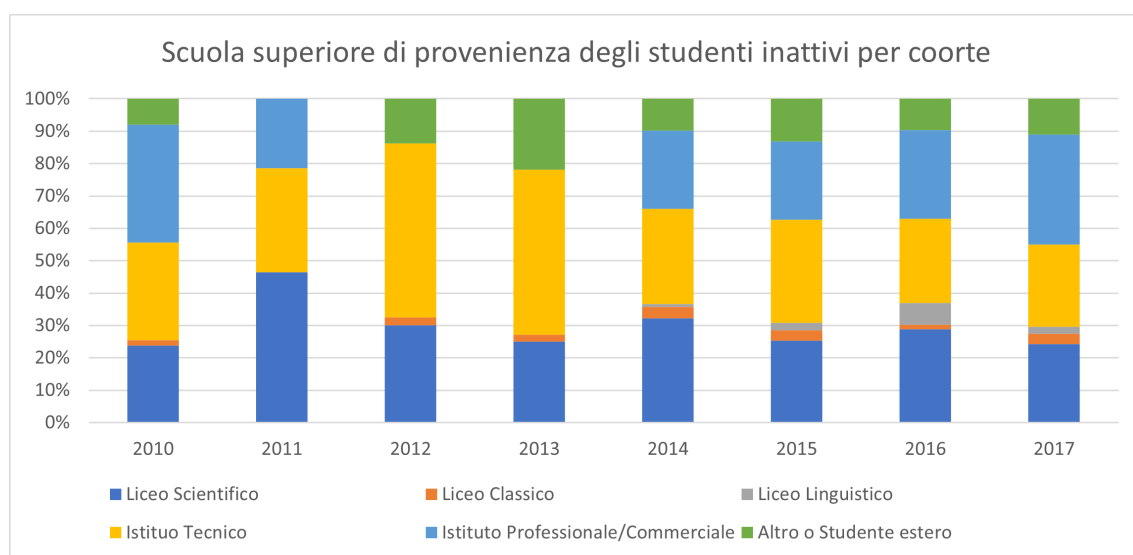


Figura 23: Scuola superiore di provenienza degli studenti inattivi per coorte

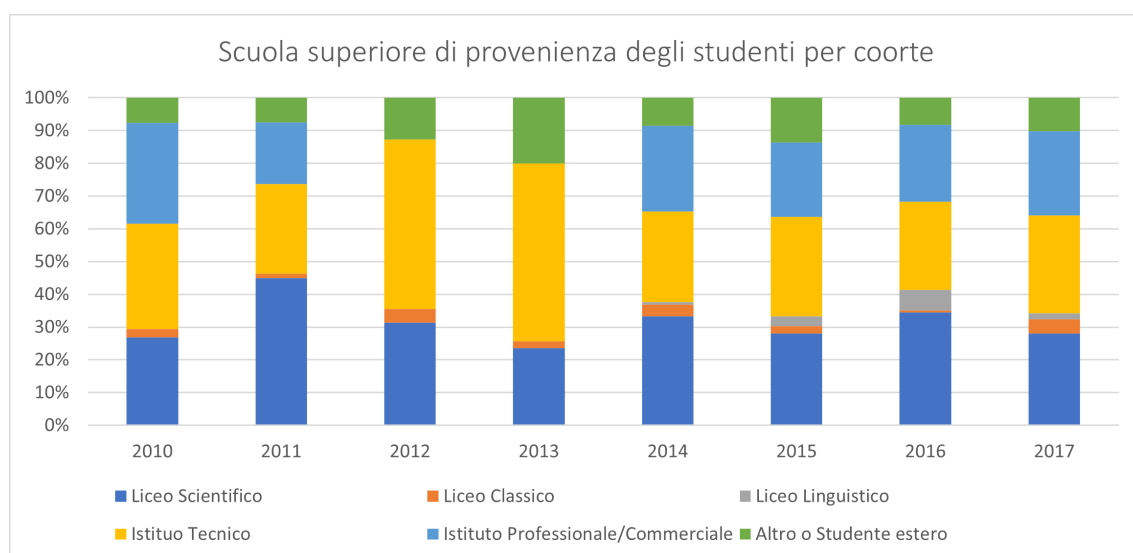


Figura 24: Scuola superiore di provenienza degli studenti per coorte

## 6.5 ANDAMENTO TEMPORALE DEGLI STUDENTI INATTIVI, ATTIVI E LAUREATI

Cerchiamo ora in questa sezione di ottenere un resoconto "temporale" sull'andamento delle varie coorti prese in esame ed in particolare tentiamo di entrare più nel dettaglio del grafico mostrato nella Figura 2 mostrando come questa distribuzione si è evoluta nel corso del tempo

per ogni singola coorte di iscrizione, andando quindi a suddividere anno per anno gli studenti che risultavano essere attivi da quelli inattivi. A questo scopo ci siamo serviti della suddivisione degli studenti inattivi, vista nella Sezione 6.2, in base al periodo trascorso prima di smettere di superare esami (ossia il periodo che intercorre tra l'anno d'immatricolazione e l'ultimo esame superato), con l'obiettivo di tentare di ordinare cronologicamente, per ogni anno di distanza dall'immatricolazione, gli studenti che diventano inattivi. Successivamente abbiamo dovuto suddividere anche i laureati (ovvero coloro che hanno superato l'esame finale) in base alla coorte di iscrizione e agli anni impiegati per laurearsi (ossia l'anno di laurea a cui abbiamo sottratto l'anno di immatricolazione). I risultati sono stati rappresentati con i grafici a barre impilate riportati nelle figure dalla 25 alla 31 (ovviamente abbiamo preso in considerazione le coorti fino al 2017 in quanto sono quelle fino a cui possiamo individuare studenti inattivi). Nei grafici abbiamo riportato solo gli anni in cui ci sono stati dei cambiamenti nelle proporzioni tra studenti attivi, inattivi e laureati iscritti alla coorte a cui fa riferimento il grafico (per questo, soprattutto nei grafici delle coorti più vecchie, i periodi più "recenti" non sono stati inseriti in quanto non è stato rilevato nessun cambiamento tra studenti attivi, inattivi e laureati). Per questo nella coorte del 2011 troviamo meno anni rispetto a quella del 2012, infatti nella coorte del 2012 esistono studenti laureati nel 2020, cosa che non è avvenuta per gli studenti delle coorti precedenti. Possiamo notare come l'incremento del numero inattivi risulti, per tutte le coorti, notevolmente più marcato nei primi tre anni, come evidenziato anche nello studio svolto presso l'Università di Roma Tre [4], per poi stabilizzarsi negli anni successivi, assumendo un comportamento simil-logaritmico con il numero degli inattivi che tende ad incrementare sempre di meno col passare degli anni.

Infine abbiamo voluto raggruppare tutte le coorti analizzate in un unico diagramma a barre impilate proporzionale ottenuto sommando, coorte per coorte, il numero degli studenti laureati, inattivi ed attivi dopo un determinato numero di anni dall'immatricolazione. Per questa caratteristica il diagramma a barre impilate risultante avrebbe avuto altezze delle barre differenti in quanto il numero di studenti iscritti dopo un certo numero di anni non può essere costante poiché uno studente non può essere inattivo, attivo o laureato da più tempo rispetto a quello da cui è immatricolato, quindi, per risolvere tale inconveniente, è stato scelto di utilizzare un diagramma proporzionale che ha cioè tutte le altezze complessive delle barre uguali, il che è utile per individuare la percentuale di studenti che anno per anno è rimasta inattiva o si è laureata rispetto



al totale degli studenti iscritti dopo un determinato numero di anni. Tale diagramma è rappresentato nella Figura 33.

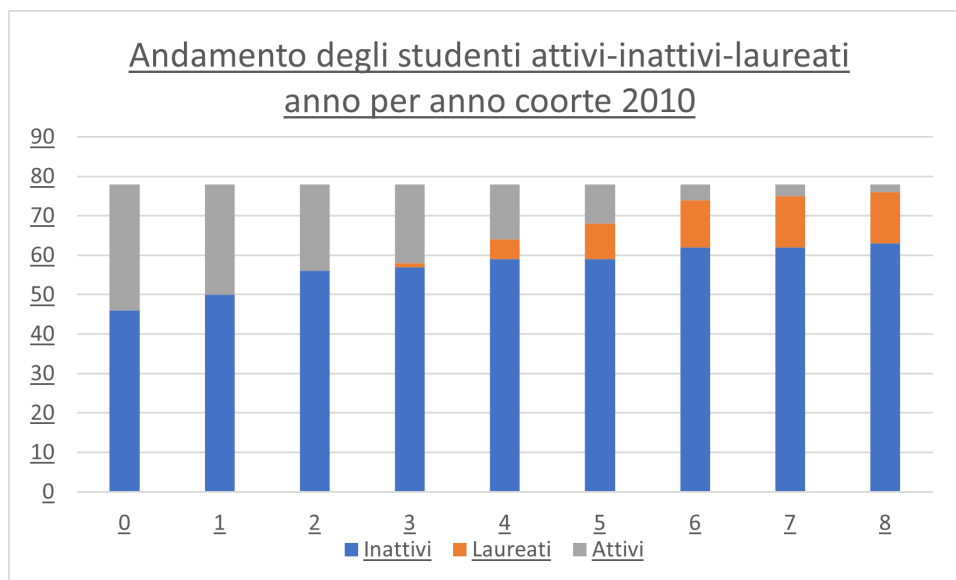


Figura 25: Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2010 anno per anno

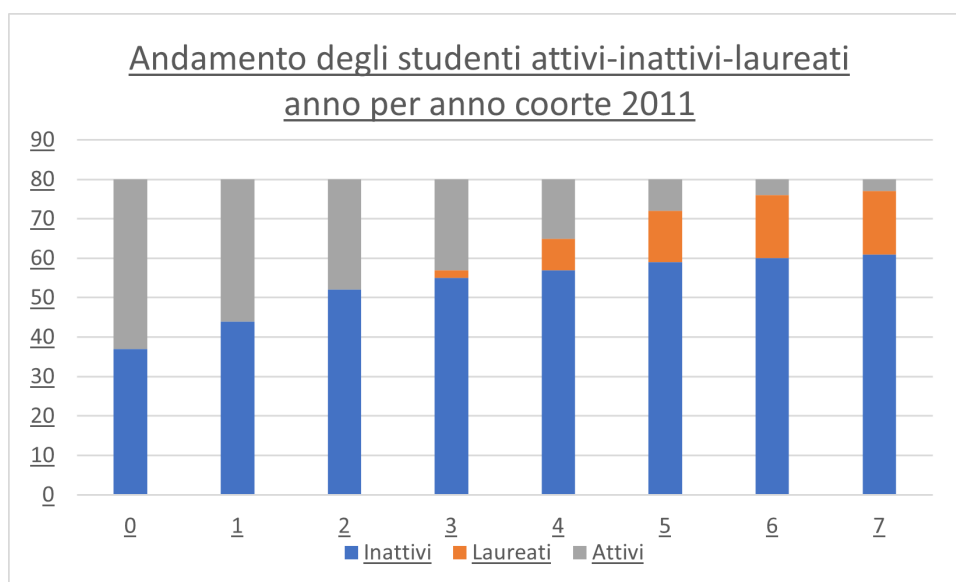


Figura 26: Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2011 anno per anno

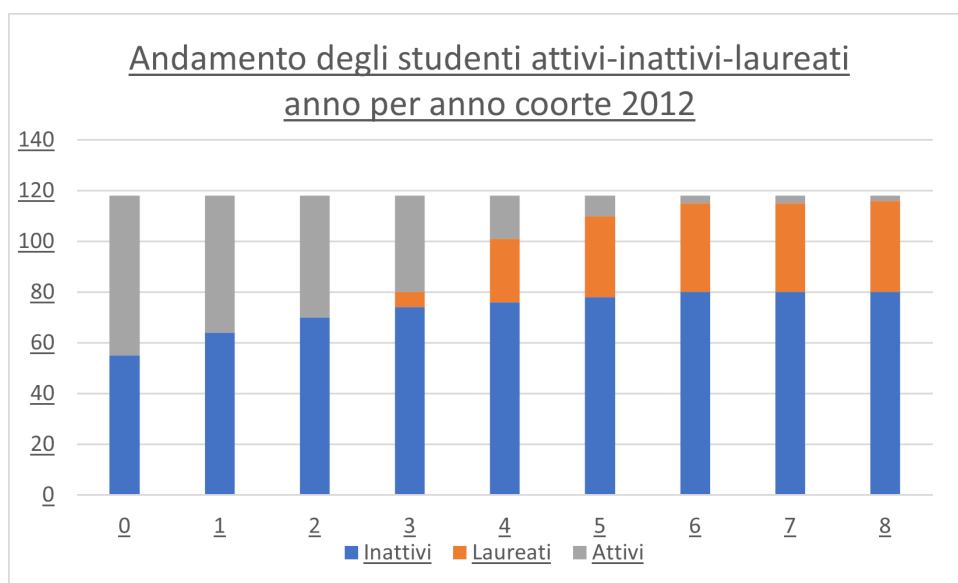


Figura 27: Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2012 anno per anno

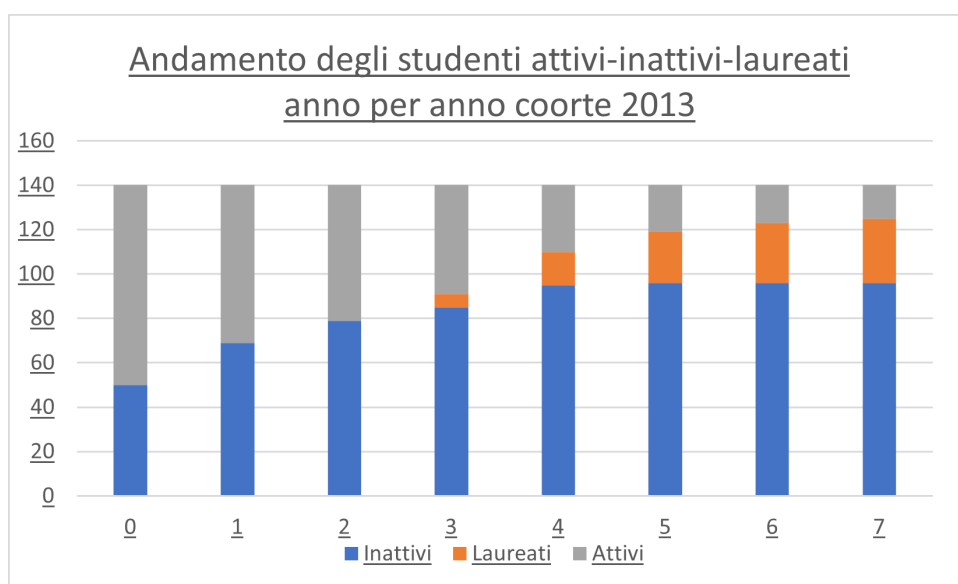


Figura 28: Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2013 anno per anno

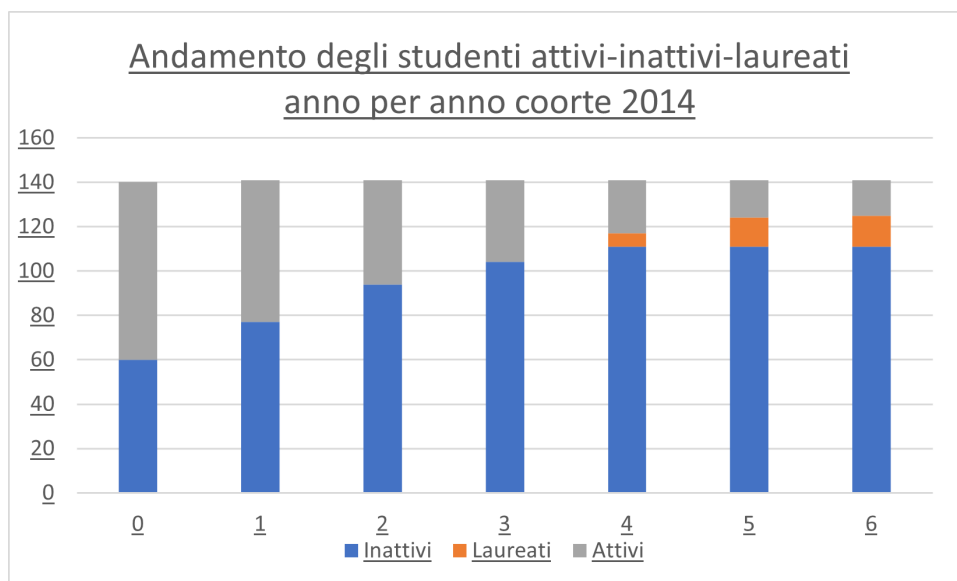


Figura 29: Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2014 anno per anno

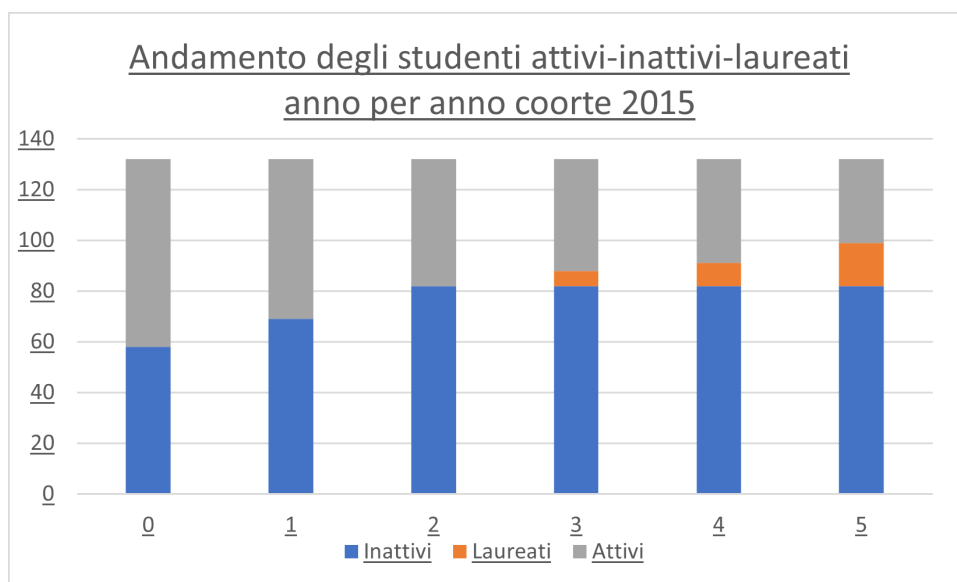


Figura 30: Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2015 anno per anno

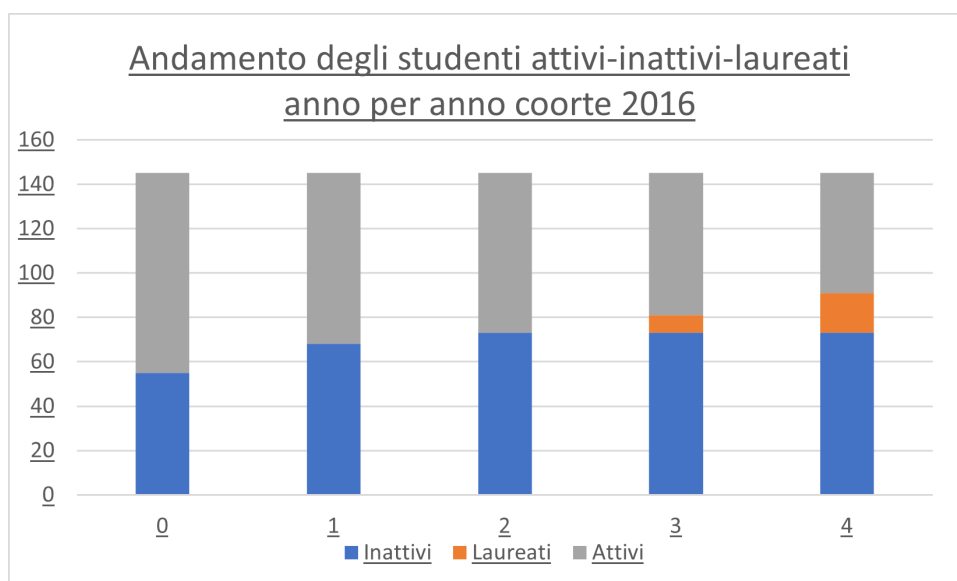


Figura 31: Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2016 anno per anno

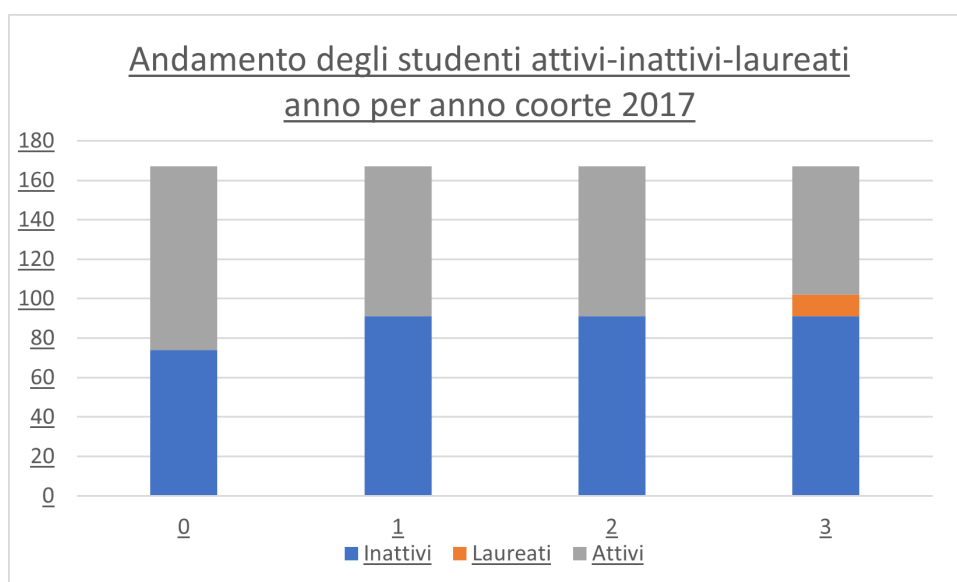


Figura 32: Andamento temporale degli studenti attivi, inattivi e laureati iscritti alla corte 2017 anno per anno

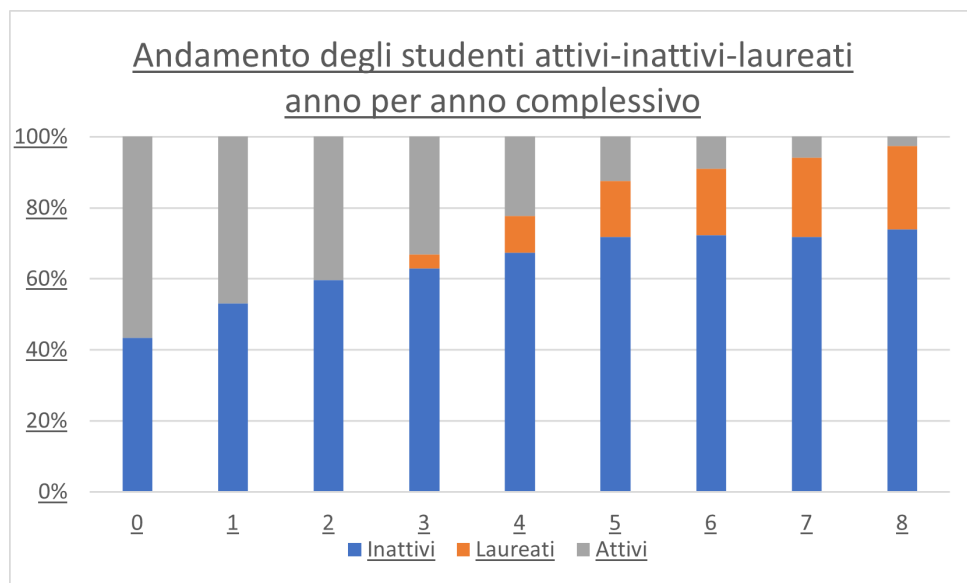


Figura 33: Andamento complessivo degli studenti attivi, inattivi e laureati degli studenti iscritti per anni dall'immatricolazione



## ANALISI E CONFRONTO DEI DATI AGGIORNATI

---

In questo capitolo analizzeremo i nuovi dati che, come accennato in precedenza, sono arrivati in nostro possesso ad aprile 2022. Essendo però la nostra analisi iniziata a febbraio dello stesso anno, l'indagine era già ampiamente sviluppata, motivo per il quale abbiamo deciso di usare questi nuovi dati che abbiamo ricevuto non come base per la nostra analisi ma bensì per verificare la correttezza dell'analisi già svolta in precedenza. Oltre a ciò, essendo i nuovi dati aggiornati al 2021, possiamo usare tali informazioni per osservare l'andamento degli abbandoni anche tra le coorti successive a quella del 2017, che era il limite che avevamo posto nella nostra precedente analisi, in quanto nei nuovi dati è riportato lo stato di iscrizione di tutti gli studenti delle corti fino al 2019. Una cosa che va però tenuta in considerazione è che tali dati saranno, almeno in parte, influenzati dalla pandemia dovuta al COVID-19 che ha stravolto anche l'Università causando disagi agli studenti che potrebbero aver deciso di abbandonare gli studi.

### 7.1 CORRETTEZZA DELLA PRECEDENTE STIMA

Per verificare la correttezza dell'analisi svolta in precedenza per prima cosa andiamo ad associare a ciascuno studente il proprio stato di iscrizione attraverso l'uso di un join basato sul numero identificativo prog e la coorte di iscrizione tra la tabella degli studenti e quella dello stato di iscrizione dello studente. Una volta ottenuto quindi per ogni studente il suo stato di iscrizione, aggiornato al 2021, è stato quindi possibile verificare l'accuratezza della nostra analisi svolta nei capitoli precedenti andando ad osservare sia quali studenti individuati come inattivi abbiano continuato a pagare le tasse o si siano laureati sia gli studenti classificati come attivi che risultano invece avere effettuato un passaggio ad un altro Corso di Laurea, un trasferimento verso un altro ateneo, una rinuncia formale agli studi o che non abbiano rinnovato l'iscrizione non

pagando le tasse. Non resta quindi che contare il numero, grazie alla funzione "count" di MySQL, di studenti precedentemente individuati. Il risultato di tale conteggio è che il numero di studenti che abbiamo stimato incorrettamente risulta essere di 72 studenti pari a circa il 7% degli studenti delle coorti prese in esame dal 2010 al 2017. Di questi 72 studenti valutati incorrettamente 31 appartengono al gruppo di coloro che non risultavano essere inattivi mentre in realtà lo erano, che corrispondono ad i "falsi negativi" (prendendo in prestito la terminologia usata nell'articolo [6]), mentre gli altri 41 studenti appartengono al gruppo di coloro che avevamo stimato aver abbandonato il Corso di Laurea ma che in realtà risultano ancora attivi, che corrispondono ad il gruppo dei "falsi positivi".

Tali risultati sono estremamente soddisfacenti, infatti il tasso di correttezza generale risulta essere pari al 93%, mentre i tassi di precisione positiva e negativa (con tasso di precisione positiva intendiamo il rapporto tra il numero di studenti evidenziati erroneamente come inattivi che in realtà non lo sono ed il totale degli studenti evidenziati come inattivi mentre per il tasso di precisione negativa consideriamo il numero di studenti considerati come attivi che in realtà non lo risultano sul totale degli studenti classificati come attivi) sono pari rispettivamente al 6% e 9%. Inoltre va considerato anche il fatto che, come accennato in precedenza, c'è stata di mezzo anche la pandemia da Covid-19, che ha avuto importanti ripercussioni sulla didattica e sulla vita stessa degli studenti, che potrebbe aver influenzato molti studenti portandoli ad abbandonare prematuramente il Corso di Laurea. Va tenuto comunque in considerazione che alcuni studenti potrebbero invece aver beneficiato degli strumenti e dei vantaggi offerti dalla didattica a distanza facendo così in modo che qualche studente, che in condizioni normali avrebbe abbandonato il Corso di Laurea, sia invece rimasto attivo.

Abbiamo infine deciso di rappresentare gli studenti che abbiamo stimato in maniera errata su un grafico a barre dividendo ciascuno dei due sottogruppi individuati in precedenza (ovvero coloro che avevamo incorrettamente definito come inattivi e coloro che in maniera errata avevamo definito come attivi) in base allo stato dell'iscrizione. Tali grafici sono riportati nelle figure 34 e 35. Possiamo notare come la maggior parte degli studenti che abbiamo stimato in maniera errata sono gli inattivi che però continuano a pagare le tasse universitarie, probabilmente con la speranza di poter riprendere gli studi in futuro. Il motivo invece per il quale alcuni studenti sono stati classificati come inattivi quando in realtà risultano come laureati è dovuto, probabilmente, al fatto di aver impostato



una soglia di crediti conseguiti per non essere classificati come inattivi troppo alta (era stata impostata a 165 crediti), classificando quindi come inattivi anche studenti che in realtà si stavano per laureare (senza aver comunque superato nessun esame in due anni).

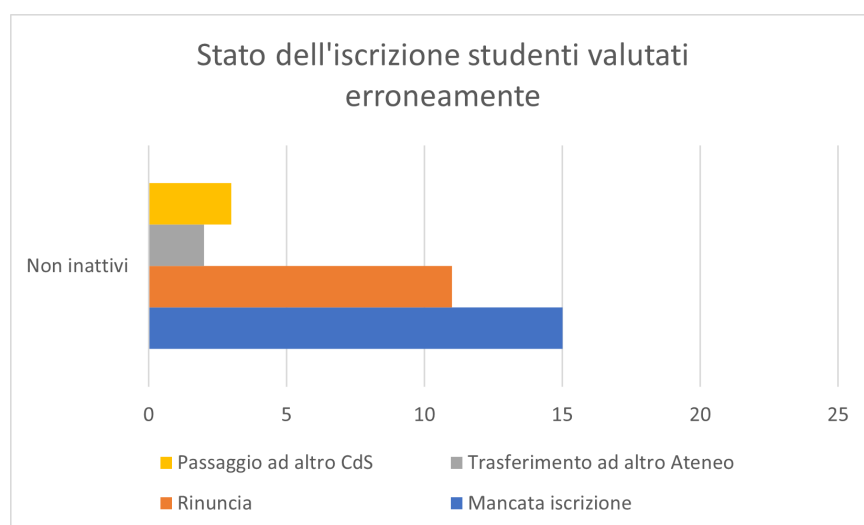


Figura 34: Stato dell'iscrizione degli studenti valutati erroneamente come non inattivi

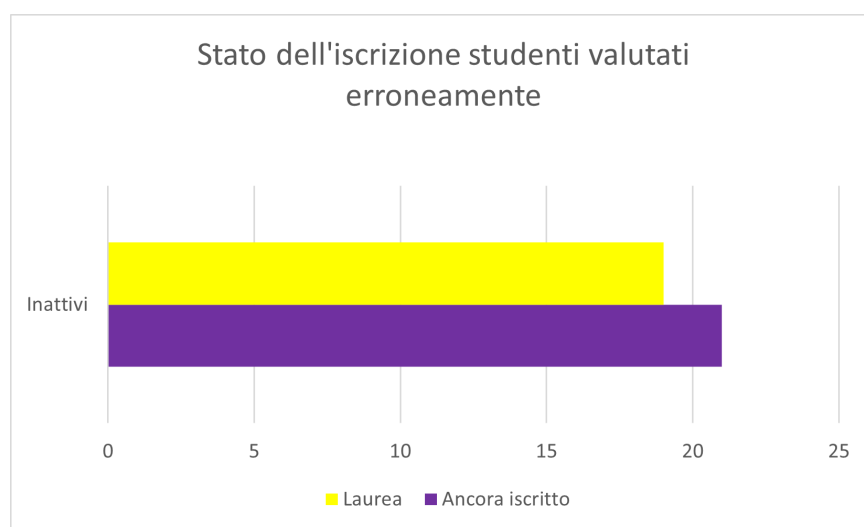


Figura 35: Stato dell'iscrizione degli studenti valutati erroneamente come inattivi

## 7.2 ABBANDONI NELLE COORTI DI ISCRIZIONE 2018 E 2019

Come accennato in precedenza in questa sezione tenteremo di osservare l'andamento degli abbandoni nelle coorti del 2018 e del 2019 che erano state escluse dalla nostra precedente analisi poiché era passato troppo

poco tempo dall'immatricolazione per poter verificare quali studenti fossero effettivamente inattivi.

Per fare ciò abbiamo raggruppato gli studenti delle coorti 2018 e 2019 in base al loro stato di iscrizione con la funzione "group by" di MySQL e li abbiamo contati con la funzione "count". Una volta ottenuto quindi il numero di studenti per ciascuno stato d'iscrizione abbiamo riportato tali dati nel grafico a torta riportato nella Figura 36. Possiamo osservare come circa la metà degli studenti immatricolati nel 2018 o nel 2019 risulti avere ancora l'iscrizione attiva mentre circa un quarto degli studenti ha già presentato una rinuncia agli studi formale. Non è detto, tuttavia, che tutti gli studenti la cui iscrizione risulta essere attiva lo siano anche nel perCorso di Laurea e potrebbero quindi aver deciso di abbandonare il Corso di Laurea. Per questo l'unica cosa che possiamo dire con certezza è che almeno il 47% degli immatricolati nelle coorti 2018 e 2019 ha abbandonato il Corso di Laurea in Informatica.

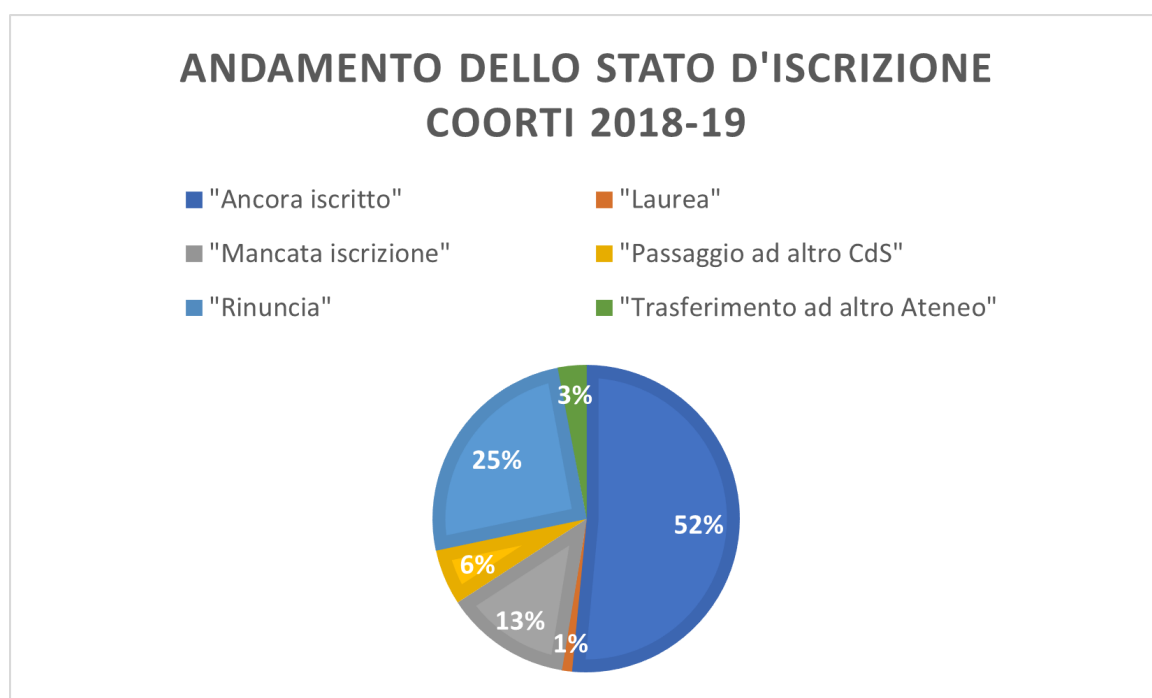


Figura 36: Stato dell'iscrizione degli studenti delle coorti 2018 e 2019 registrato nel 2021

Volendo inoltre studiare i dati anche per ogni singola coorte possiamo suddividere gli studenti in base alla coorte di iscrizione e, ricavando in maniera analoga il numero di studenti per stato d'iscrizione, riportare i

grafici a torta per le due coorti da noi prese in esame. Tali grafici sono riportati nelle figure 37 e 38. Possiamo osservare come le percentuali ottenute siano abbastanza in linea con quelle che abbiamo osservato nel capitolo precedente riguardo all'andamento degli inattivi nel corso degli anni.

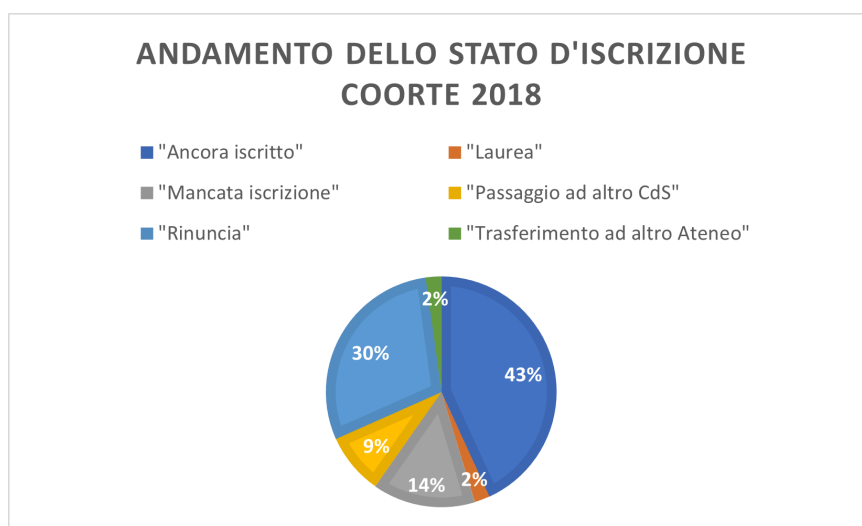


Figura 37: Stato dell'iscrizione degli studenti iscritti nella coorte 2018

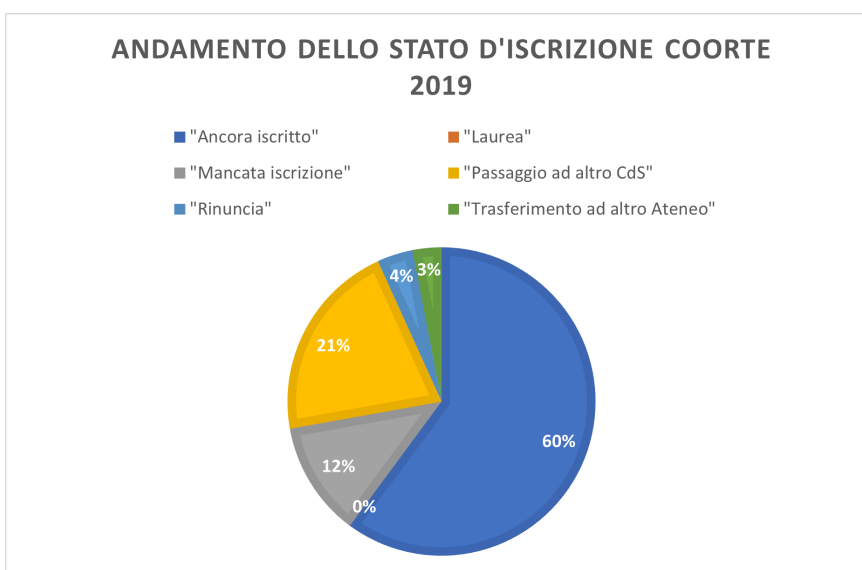


Figura 38: Stato dell'iscrizione degli studenti iscritti nella coorte 2019

### 7.3 STATO D'ISCRIZIONE DI TUTTI GLI STUDENTI

In questa sezione ci soffermiamo invece sull'analisi dello stato d'iscrizione di tutti gli studenti a partire dalla coorte d'iscrizione 2010 fino a quella del 2019. Per poter analizzare tali dati è stato sufficiente raggruppare i dati sullo stato della carriera, grazie alla funzione "group by" di MySQL, e successivamente contare quanti studenti avessero un determinato stato di carriera piuttosto che un altro. I risultati di tale raggruppamento sono stati poi riportati nel grafico a torta della Figura 39.

Possiamo osservare come oltre la metà degli studenti risulti non avere più uno stato d'iscrizione attivo mentre solo il 20% degli studenti risulta essere ancora attivo con appena il 18% degli immatricolati dal 2010 che è riuscito a laurearsi.

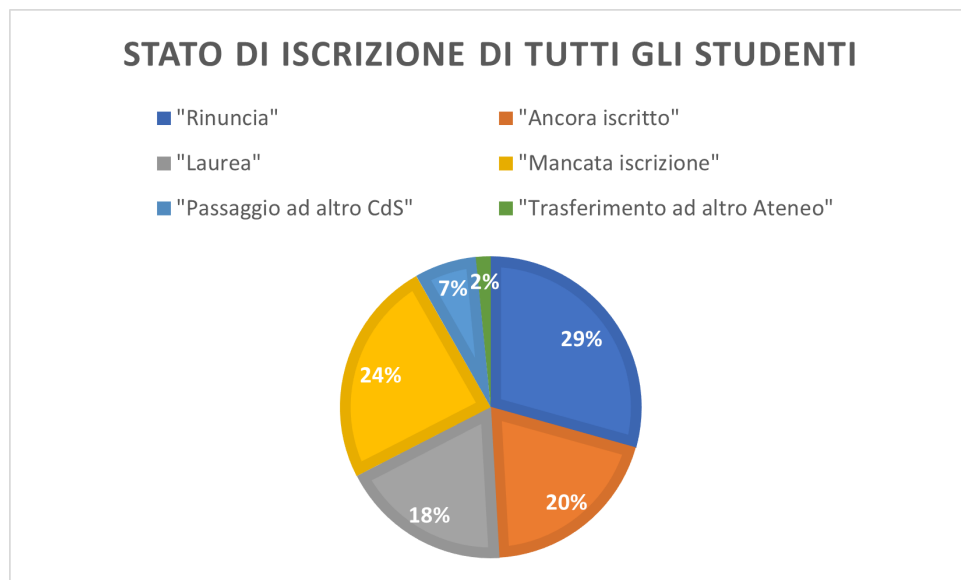


Figura 39: Stato dell'iscrizione degli studenti iscritti di tutte le coorti (2010-2019)



## CONCLUSIONE

---

Da come abbiamo avuto modo di osservare nelle pagine precedenti, i dati in nostro possesso sembrano evidenziare che la maggior parte degli studenti inattivi, che abbiamo stimato aver abbandonato il Corso di Laurea, abbiano in generale relativamente poche caratteristiche che ne permettano una distinzione netta dal resto degli studenti al momento dell'iscrizione. Infatti l'unica caratteristica sufficientemente distintiva, ricavabile al momento dell'iscrizione di uno studente o comunque nei primi mesi della carriera, che abbiamo individuato è la valutazione che uno studente ha conseguito nel test d'ingresso, come abbiamo evidenziato nella sezione 5.3.

Oltre a ciò avevamo anche individuato un fattore di rischio negli studenti che avevano conseguito delle valutazioni molto basse all'esame di maturità (tra 60 e 70), tuttavia in maniera molto meno evidente rispetto alla valutazione conseguita al test d'ingresso (in particolar modo per quanto riguarda il test d'ingresso svolto con la nuova modalità in uso dal 2016). Per questo motivo in questa parte finale abbiamo deciso di approfondire ulteriormente quest'aspetto standardizzando, attraverso una z-score normalization, le valutazioni ottenute nel test d'ingresso come vedremo nella sezione seguente.

### 8.1 ANALISI FINALE DELLA VALUTAZIONE AL TEST D'INGRESSO

Come anticipato in precedenza, per effettuare quest'analisi abbiamo deciso di standardizzare i dati riguardo la valutazione ottenuta al test d'ingresso da uno studente separando però le varie coorti d'iscrizione in quanto, come spiegato nella sezione 5.3, la modalità di valutazione di tale prova è variata nel corso degli anni. Per standardizzare le valutazioni è stato necessario calcolare la loro media e deviazione standard per ogni singola coorte d'iscrizione, grazie alle funzioni "avg" e "std" di MySQL, e successivamente, in base alla coorte d'iscrizione dello studente, sot-

trarre alla valutazione conseguita la media delle valutazioni della coorte di appartenenza e dividere il risultato per la rispettiva deviazione standard della coorte di appartenenza dello studente. Avendo così ricavato le valutazioni standardizzate in base alla coorte di appartenenza di uno studente possiamo tenere di conto nella nostra analisi anche di eventuali differenze di difficoltà nei test d'ingresso che possono essersi presentate nel corso degli anni, in quanto standardizzando la valutazione conseguita da uno studente è come se confrontassimo tale risultato con quello ottenuto da tutti gli altri studenti della stessa coorte a prescindere dal fatto che il test sia stato effettivamente superato o meno.

Un'analisi interessante su questi dati appena ricavati potrebbe essere osservare come la media delle valutazioni standardizzate degli studenti che abbiamo definito come inattivi vari nel corso del tempo. Per effettuare tali analisi è stato sufficiente calcolare la media, coorte per coorte, dei risultati ottenuti in precedenza e visualizzarli nella Figura 40.



Figura 40: Evoluzione della media delle valutazioni standardizzate conseguite da studenti inattivi in base alla coorte

Come è possibile osservare dal grafico, abbiamo che in ogni corte gli studenti inattivi conseguono una valutazione mediamente inferiore nelle valutazioni del test d'ingresso rispetto alla media generale (uguale a 0 visto che stiamo studiando un parametro standardizzato).

Possiamo inoltre notare come ci sia una differenza abbastanza marcata tra le valutazioni ottenute da studenti inattivi con la vecchia modalità (in uso prima del 2016) e con la nuova modalità (dal 2016 in poi). Questa differenza può essere un'altra nota a favore della nuova modalità di



valutazione del test d'ingresso che riesce a distinguere in maniera mediamente più netta e corretta gli studenti che diventeranno inattivi in futuro.

## 8.2 ALTRI PARAMETRI DA TENERE IN CONSIDERAZIONE

In questa sezione evidenziamo invece gli altri parametri che ci permettono di distinguere in maniera abbastanza corretta gli studenti inattivi, che supponiamo aver abbandonato il Corso di Laurea, da quelli attivi o già laureati.

Questi parametri però, a differenza della valutazione conseguita al test d'ingresso, non sono ricavabili al momento dell'iscrizione o comunque all'inizio della carriera universitaria in quanto riguardano i crediti conseguiti nel primo anno da uno studente e la media ponderata delle valutazioni conseguite negli esami superati da uno studente nel corso della propria carriera universitaria, come visto nelle sezioni 5.6 e 5.7. Come anticipato questi valori sono ricavabili dopo un certo lasso di tempo dalla data d'iscrizione di uno studente, tuttavia potrebbero rivelarsi comunque utili per studiare la progressione della carriera degli studenti nel corso del tempo per poter valutare l'andamento dei possibili futuri studenti inattivi.

La media ponderata dei voti conseguiti da uno studente, insieme ai crediti ottenuti, permettono inoltre di evidenziare eventuali criticità occorse nel corso del tempo soprattutto nel caso di evidenti squilibri tra i vari anni, eventualità che non è stata riscontrata nella nostra analisi ma che potrebbe presentarsi in futuro.

## 8.3 CONSIDERAZIONI FINALI

La nostra analisi ha rilevato come più di due terzi degli studenti iscritti nelle coorti da noi analizzate risulti essere inattivo. Avendo stimato in maniera abbastanza corretta gli studenti inattivi come studenti che hanno abbandonato il Corso di Laurea, come visto nel capitolo 7, possiamo affermare con relativa sicurezza che circa due terzi degli studenti iscritti nelle coorti dal 2010 al 2017 ha effettivamente abbandonato il Corso di Laurea in Informatica. Di questi studenti che hanno abbandonato circa il 71% non ha superato nemmeno un esame (escluso al massimo quello di lingua Inglese). Questi dati evidenziano quindi in maniera molto lampante quanto sia importante riuscire a comprendere in maniera il

più chiara possibile le cause di questo fenomeno, per poter porre freno a questa emorragia così impressionante di studenti. Una delle prime possibilità che potrebbe venire in mente, per giustificare un numero così elevato di abbandoni di studenti che non hanno superato alcun esame, è quella che la maggior parte degli studenti che si iscrivono non abbia un'idea ben chiara di cosa andrà ad affrontare, non informandosi probabilmente a sufficienza e trovandosi quindi "impreparati" quando devono sostenere i primi esami.

Tuttavia questa è solo un'ipotesi che non può essere confermata in quanto riguarda ambiti che non si possono studiare con una semplice indagine statistica svolta con i dati in nostro possesso, ma richiederebbe studi come quelli citati in [1] che prendono in considerazione anche dati che riguardano conoscenze pregresse, stato sociale e stato psico-fisico dello studente, che però, come abbiamo visto, sono difficili da ottenere in un'università italiana. Per questo motivo è fondamentale l'attività di orientamento in ingresso per fare in modo che lo studente, intenzionato ad iscriversi al Corso di Laurea in Informatica, sappia effettivamente cosa andrà a studiare e quali esami dovrà sostenere in modo che sia consapevole di ciò che lo attende per evitare che esso abbandoni senza aver superato nemmeno un esame.

## BIBLIOGRAFIA

---

- [1] Catarina Félix de Oliveira, Sónia Rolland Sobral, Maria João Ferreira e Fernando Moreira - *How Does Learning Analytics Contribute to Prevent Students' Dropout in Higher Education: A Systematic Literature Review* (Cited on pages 10 and 64.)
- [2] Francisco Araque, Concepción Roldán, Alberto Salguero - *Factors influencing university drop out rates* - Computers & Education, Volume 53, Issue 3, 2009, Pages 563-574, ISSN 0360-1315, <https://doi.org/10.1016/j.compedu.2009.03.013>. (Cited on page 11.)
- [3] Siri, D. (2015) - *Predicting Students' Dropout at University Using Artificial Neural Networks*. - Italian Journal of Sociology of Education, 7(2), 225-247. Retrieved from <http://journals.padovauniversitypress.it/ijse/content/predicting-students%E2%80%99-dropoutuniversity-using-artificial-neural-networks> (Cited on page 11.)
- [4] Agrusti F., Mezzini M., Bonavolontà G., (2020) - *Deep learning approach for predicting university dropout: a case study at Roma Tre University*. - oJournal of E-Learning and Knowledge Society, 16(1), 44-54. <https://doi.org/10.20368/1971-8829/1135192> (Cited on pages 11 and 46.)
- [5] Paola Perchinunno, Massimo Bilancia, Domenico Vitale - *A Statistical Analysis of Factors Affecting Higher Education Dropouts*. - Soc Indic Res 156, 341–362 (2021). <https://doi.org/10.1007/s11205-019-02249-y> (Cited on pages 12 and 35.)
- [6] Nuankaew, P. (2019). - *Dropout Situation of Business Computer Students, University of Phayao*. - International Journal of Emerging Technologies in Learning. <https://doi.org/10.3991/ijet.v14i19.11177> (Cited on pages 12 and 54.)

- [7] Niurys Lázaro Álvarez, Zoraida Callejas, David Griol - *Predicting Computer Engineering students' dropout in Cuban Higher Education with pre-enrollment and early performance data* - Journal of Technology and Science Education. <http://dx.doi.org/10.3926/jotse.922> (Cited on page 12.)