



UNIVERSITA' POLITECNICA DELLE MARCHE

FACOLTA' DI INGEGNERIA

Corso di Laurea triennale in Ingegneria Informatica e dell'Automazione

**PROGETTAZIONE DI UN DATABASE PER LA RACCOLTA E L'ANALISI DI DATI RELATIVI
AI TEST DI INGRESSO AI CORSI UNIVERSITARI**

**DATABASE DESIGN FOR THE COLLECTION AND ANALYSIS OF DATA RELATING TO
ENTRANCE TESTS TO UNIVERSITY COURSES**

Relatore:

Prof.ssa Claudia Diamantini

Correlatore:

Dott. Alex Mircoli

Tesi di Laurea di:

Giulio Mezzotero

A.A. 2019 / 2020

Indice

Introduzione	2
1 Analisi dei requisiti	4
1.1 Raccolta informazioni e dati	4
1.2 Analisi dei processi interni	4
1.3 Requisiti espressi in linguaggio naturale	5
1.4 Glossario dei termini	6
1.5 Eliminazione delle ambiguità presenti	7
1.6 Strutturazione dei requisiti	7
1.7 Specifica delle operazioni	7
2 Progettazione del Database	10
2.1 Progettazione Concettuale	10
2.1.1 Diagramma E-R	12
2.1.2 Dizionario dei dati	13
2.1.3 Business Rules	14
2.2 Progettazione Logica	14
2.2.1 Tavole dei volumi e delle operazioni	14
2.2.2 Ristrutturazione dello schema concettuale	16
2.2.3 Schema ristrutturato finale	19
2.2.4 Normalizzazione	19
2.2.5 Traduzione verso il modello relazionale	20
3 Analisi dei dati	21
Conclusioni	51
Bibliografia	52
Ringraziamenti	53

Introduzione

Ogni anno gli studenti delle Scuole Secondarie di II grado che vogliono iscriversi presso la facoltà di Ingegneria dell'Università Politecnica delle Marche devono sostenere un *Test di Verifica delle Conoscenze* al fine di ottenere degli *Obblighi Formativi Aggiuntivi (OFA)*.

Queste prove sono autovalutative e non precludono la possibilità di iscrizione al corso di laurea, tuttavia sono importanti per capire se uno studente è del tutto pronto a sostenere un corso di studi di tipo ingegneristico, se parte con delle carenze o se le conoscenze possedute non sono sufficienti in quanto in esse si valutano gli aspetti basilari e utili ad affrontare tale percorso accademico.

I test sono divisi in trenta domande a loro volta suddivise in sette categorie: Geometria, Algebra, Funzioni, Grafici, Fisica, Logica e Comprensione del testo. Si noti dunque che viene valutata non solo la preparazione scientifica ma anche l'attitudine al ragionamento.

Un'ulteriore suddivisione del test è per tipologia: ad ogni appello viene fornito ad uno studente un quiz di tipo A o di tipo B che presentano le stesse domande ma in ordine diverso.

Il punteggio assegnato a ciascuna domanda è 0.33 se si è risposto correttamente, 0.08 se non si è risposto al quesito e 0.00 se si è risposto erroneamente.

Il problema che sorge è che i dati e i risultati di questi test sono raccolti in file *Excel* di tipo *xlsx* e dunque difficilmente analizzabili.

Lo scopo della presente tesi di laurea è progettare un database per contenere le informazioni relative a queste prove e conferire loro una struttura più facilmente esaminabile così da poter giungere a conclusioni significative riguardo al background degli studenti che li sostengono e alle dinamiche della prova.

Analizzeremo ad esempio quali sono gli istituti da cui provengono principalmente gli studenti, in quali mesi sostengono preferibilmente tale esame, quali sono le categorie e le specifiche domande in cui ottengono i migliori risultati e quali sono quelle in cui ottengono i peggiori, quali sono le scuole che ottengono i migliori punteggi, se ci sono differenze nei risultati ottenuti nelle due tipologie di quiz, quanti si iscrivono effettivamente all'Università dopo aver sostenuto il test e in quale corso di laurea, se ci sono incrementi o decrementi nel numero di iscrizioni nel corso degli anni e quanti studenti maschi e quante studentesse femmine hanno sostenuto il test in generale, nel corso degli anni e come si distribuiscono tra i vari corsi presentati dalla facoltà.

Successivamente all'introduzione verranno illustrati gli strumenti utilizzati e i passaggi effettuati per l'ottenimento del risultato prefissato:

- Nel Capitolo 1 si descrive l'analisi dei requisiti necessari alla progettazione della base di dati: si raccolgono i dati e si strutturano, si studiano a fondo i processi interni allo svolgimento del test, si eliminano le ambiguità e si specificano le operazioni da svolgere.
- Nel Capitolo 2 viene mostrata la progettazione vera e propria. Nella *concettuale* si costruisce il diagramma Entità-Relazione per descrivere la realtà d'interesse in maniera semplice e a prescindere dall'organizzazione dei dati nei calcolatori. Nella *logica* si ristrutturava il diagramma per giungere ad una traduzione verso il Modello Relazionale che ci permette poi l'implementazione del database.

- Nel Capitolo 3 si analizzano i dati attraverso delle queries in linguaggio *SQL* così da ricavare informazioni utili concernenti i test d'ingresso tramite lo studio dei risultati ottenuti.

Capitolo 1

Analisi dei requisiti

1.1 Raccolta informazioni e dati

I dati di partenza riguardanti i Test di Verifica delle Conoscenze sono stati ottenuti in forma anonima dall'Università Politecnica delle Marche. Essi sono contenuti in tre tabelle: la prima ospita i dati concernenti il Test A, la seconda quelli riguardanti il Test B e la terza presenta la Distribuzione dei risultati.

I dati relativi ai test sono lo username assegnato dall'Università allo studente per svolgerlo, la scuola di provenienza dello studente, lo stato del test (che al termine della prova sarà settato a *Completato*), l'orario d'inizio, l'orario di consegna (e quindi il tempo impiegato per svolgerlo), la valutazione complessiva e la valutazione ottenuta in ognuna delle trenta domande e sono appunto divisi in A e B, prove che presentano le stesse domande ma in ordine differente.

I dati relativi alla distribuzione sono l'ID della categoria della domanda (composto dal numero della domanda del test e da un identificativo della categoria, ad es. 01-LOG per la prima domanda, che riguarda la categoria di Logica), la descrizione del tipo di domanda, il numero che tale domanda occupa nel quiz A, il numero che occupa nel quiz B e infine la percentuale delle risposte corrette.

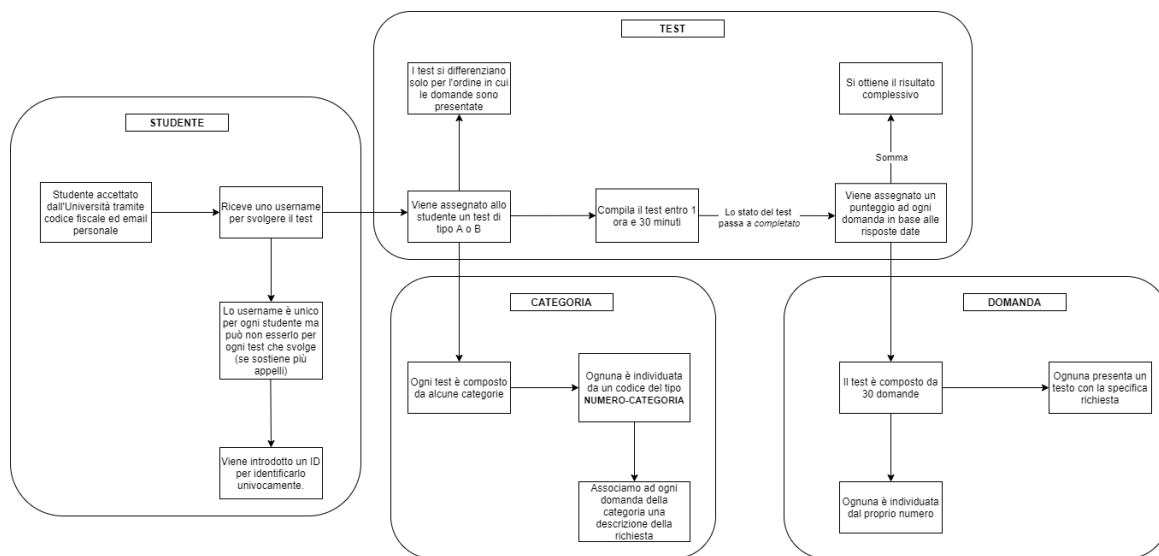
1.2 Analisi dei processi interni

Si passa dunque ad analizzare attentamente i processi interni allo svolgimento di un appello di esami OFA per evidenziare quali sono le entità in gioco e poter così progettare correttamente il database al fine di raggiungere lo scopo prefissato.

Uno studente si iscrive al test fornendo il proprio codice fiscale e la propria e-mail e riceve uno username per svolgerlo. Lo username è unico per ogni studente ma potrebbe non esserlo per ogni prova, in quanto può accadere che uno studente non passi l'esame in una data sessione, debba ripeterlo e riceva nuovamente lo stesso username. Anche se la probabilità che ciò accada è molto bassa, è necessario tenerne conto.

Lo studente riceve un quiz di tipo A o di tipo B ed ha a disposizione 1 ora e 30 minuti per completarlo. Ad ogni domanda è assegnato un punteggio a seconda della correttezza della risposta data.

Ogni quiz è formato da 30 domande divise in 7 categorie (Logica, Geometria, Algebra, Funzioni, Grafici, Fisica, Comprensione del testo).



1.3 Requisiti espressi in linguaggio naturale

Si vuole progettare dunque una base di dati capace di gestire le informazioni riguardanti gli studenti che sostengono il Test OFA presso l'Università e i risultati ottenuti, particolarizzati per ogni tipo di domanda di ogni categoria.

Si dovranno quindi gestire i dati di ogni studente, che ha frequentato o sta ancora frequentando una Scuola Secondaria di II grado, che accede al test in una data d'appello con uno username fornito dall'Università. Lo studente è inoltre riconosciuto attraverso nome, cognome, codice fiscale ed e-mail.

Per quanto riguarda il test svolto da ogni singolo studente, si vuole conoscere il giorno in cui è stato compilato, il tipo, l'orario di inizio e di consegna (e dunque il tempo totale di svolgimento) e il risultato ottenuto, dato dalla somma delle valutazioni ottenute in ciascuna domanda.

I test sono divisi in varie categorie di cui vogliamo conoscere il tipo e la descrizione generale delle domande inerenti ad esse.

La particolare domanda appartenente ad una data categoria associata al test dovrà essere catalogata con il suo numero e la sua richiesta specifica, ovvero il testo.

1.4 Glossario dei termini

Focalizziamoci sui termini che meritano maggior dettaglio in quanto importanti ai fini della nostra applicazione così da darne una spiegazione completa e soddisfacente.

Termine	Descrizione	Sinonimi	Collegamenti
Test	Prova che sostiene ogni studente al fine di valutare la sua preparazione in vista dell'iscrizione all'Università	Prova, esame, quiz	Studente, Categoria, Compilazione, Associazione
Studente	Soggetto frequentante una Scuola Secondaria di II grado volente iscriversi all'Università	Allievo, alunno, scolaro	Test, Compilazione
Categoria	Raggruppamento di domande che presentano la stessa tipologia di argomento richiesto	Classe, genere, specie, divisione, parte, gruppo, famiglia	Test, Domanda, Associazione, Appartenenza
Domanda	Particolare richiesta a cui è assegnato un punteggio a seconda della risposta fornita	Interrogazione, richiesta, quesito, questione	Categoria, Appartenenza
Compilazione	Atto compiuto dallo studente al momento del sostenimento del test	Stesura, redazione	Studente, Test
Associazione	Legame tra il test e una serie di categorie in cui è diviso	Collegamento, corrispondenza	Test, Categoria
Appartenenza	Legame tra una particolare domanda e la categoria di cui fa parte	Pertinenza, attinenza, adesione, inerenza	Categoria, Domanda

1.5 Eliminazione delle ambiguità presenti

- Per motivi di privacy non è reso possibile al tesista accedere a dati sensibili quali nome, cognome, codice fiscale, e-mail. Si lavorerà pertanto, in fase di analisi, sui dati di dominio pubblico. Le analisi sui dati privati saranno solo implementate ed eseguite poi dall'Università, garante di tali informazioni.
- Nell'ID della Categoria, della forma "Numero-Categoria", il numero rappresenta quello della domanda del test quando progettato dal consiglio accademico e non della categoria, quindi ad esempio "10-ALG" non rappresenta la decima domanda della categoria Algebra ma la decima del test, appartenente alla categoria Algebra. Questa domanda avrà poi una posizione, che può anche essere diversa dalla decima, a seconda del test in cui è inserita (se di tipologia A o B).

1.6 Strutturazione dei requisiti

- Frasi di carattere generale: "Si vuole progettare dunque una base di dati capace di gestire le informazioni riguardanti gli studenti che sostengono il Test OFA presso l'Università e i risultati ottenuti, particolarizzati per ogni tipo di domanda di ogni categoria".
- Frasi relative agli studenti: "Si dovranno quindi gestire i dati di ogni studente, che ha frequentato o sta ancora frequentando una Scuola Secondaria di II grado, che accede al test in una data d'appello con uno username fornito dall'Università. Lo studente è inoltre riconosciuto attraverso nome, cognome, codice fiscale ed e-mail".
- Frasi relative al test: "Per quanto riguarda il test svolto da ogni singolo studente, si vuole conoscere il giorno in cui è stato compilato, il tipo, l'orario di inizio e di consegna (e dunque il tempo impiegato nello svolgimento) e il risultato ottenuto, dato dalla somma delle valutazioni ottenute in ciascuna domanda".
- Frasi relative alla categoria: "I test sono divisi in varie categorie di cui vogliamo conoscere il tipo e la descrizione generale delle domande inerenti ad esse".
- Frasi relative alla domanda: "La particolare domanda appartenente ad una data categoria associata al test dovrà essere catalogata con il suo numero e la sua richiesta specifica, ovvero il testo".

1.7 Specifica delle operazioni

1. Inserimento nuovo studente
2. Inserimento nuovo test
3. Inserimento nuova associazione
4. Inserimento nuova categoria
5. Inserimento nuova domanda

6. Cancellazione studente
7. Cancellazione test
8. Cancellazione associazione
9. Cancellazione categoria
10. Cancellazione domanda
11. Modifica scuola provenienza studente
12. Modifica email studente
13. Modifica descrizione categoria
14. Modifica testo domanda
15. Calcolo del voto medio per tipologia di quiz
16. Calcolo del voto medio per tipologia di quiz per sessione
17. Calcolo del voto medio per tipologia di quiz per anno
18. Calcolo del voto medio per scuola di provenienza
19. Calcolo del voto medio per scuola di provenienza per sessione
20. Calcolo del voto medio per scuola di provenienza per anno
21. Calcolo del voto medio per categoria di domanda
22. Calcolo del voto medio per categoria di domanda per sessione
23. Calcolo del voto medio per categoria di domanda per anno
24. Calcolo del voto medio per domanda
25. Calcolo del voto medio per domanda per anno
26. Conteggio del numero di studenti che ha sostenuto il test ogni mese
27. Conteggio del numero di studenti che ha sostenuto il test ogni anno
28. Visualizzazione della tipologia di scuola di provenienza di chi ha sostenuto il test
29. Visualizzazione della tipologia di scuola di provenienza di chi ha sostenuto il test per sessione
30. Visualizzazione della tipologia di scuola di provenienza di chi ha sostenuto il test per anno
31. Conteggio del numero di iscrizioni per anno
32. Conteggio del numero di iscrizioni per anno per scuola di provenienza
33. Conteggio del numero di iscrizioni per corso di studio
34. Conteggio del numero di iscrizioni per corso di studio per anno
35. Conteggio del numero di iscrizioni per corso di studio per scuola di provenienza

36. Conteggio del numero di studenti maschi che hanno sostenuto il test
37. Conteggio del numero di studenti maschi che hanno sostenuto il test per anno
38. Conteggio del numero di studenti maschi iscritti ad ogni corso di studio
39. Conteggio del numero di studenti maschi iscritti ad ogni corso di studio in ogni anno
40. Conteggio del numero di studentesse femmine che hanno sostenuto il test
41. Conteggio del numero di studentesse femmine che hanno sostenuto il test per anno
42. Conteggio del numero di studentesse femmine iscritte ad ogni corso di studio
43. Conteggio del numero di studentesse femmine iscritte ad ogni corso di studio in ogni anno

Capitolo 2

Progettazione del Database

2.1 Progettazione Concettuale

In questa fase, dopo un'attenta analisi dei dati in possesso, si introducono le Entità e Relazioni che andranno a comporre il modello E-R da cui partire per il concepimento del database.

Si è scelto di utilizzare una strategia *mista* per combinare i vantaggi della strategia *bottom-up* con quelli della *top-down*: si dividono i requisiti in componenti separate come nella prima ma allo stesso tempo si definisce uno schema scheletro contenente a livello astratto i concetti principali dell'applicazione come nella seconda.

Le Entità coinvolte sono:

- **Studente**: racchiude tutte le informazioni relative a coloro che sostengono la prova.
 - L'identificatore principale sarà un *ID* (composto dalla data in cui ha sostenuto il test e da un numero progressivo, ad es. 2016-04-17_1) attribuito univocamente a ciascuno studente ad ogni appello di esami OFA così da evitare possibili casi, seppur rari, in cui uno studente sostenga un test in più appelli e riceva lo stesso username.
 - Gli attributi saranno la sua *Scuola di provenienza*, il suo *Codice fiscale*, il suo *Nome*, il suo *Cognome*, la sua *Email* e uno *Username* che riceve il giorno della prova per svolgerla.
- **Test**: racchiude tutte le informazioni relative ad ogni singola prova.
 - L'identificatore principale sarà l'unione dell'*ID* dello studente e della *Data* in cui egli sostiene il test.
 - Gli attributi saranno la *Valutazione* ottenuta nel test, l'*Orario di inizio*, l'*Orario di fine* (che comprendono anche il giorno di svolgimento) e dunque il *Tempo impiegato*, la *Tipologia* (A o B) e lo *Stato di compilazione*, che diventa "Completato" alla consegna o allo scadere del tempo assegnato.
- **Categoria**: racchiude tutte le informazioni relative al tipo di domanda.
 - L'identificatore principale sarà l'*ID* della categoria stessa, espresso nel formato "NumeroDomanda-Categoria".
 - L'unico attributo sarà la *Descrizione* della particolare domanda della categoria in questione.
- **Domanda**: racchiude tutte le informazioni relative alla singola domanda.
 - L'identificatore principale sarà l'*ID* della domanda stessa, ovvero il numero della domanda.

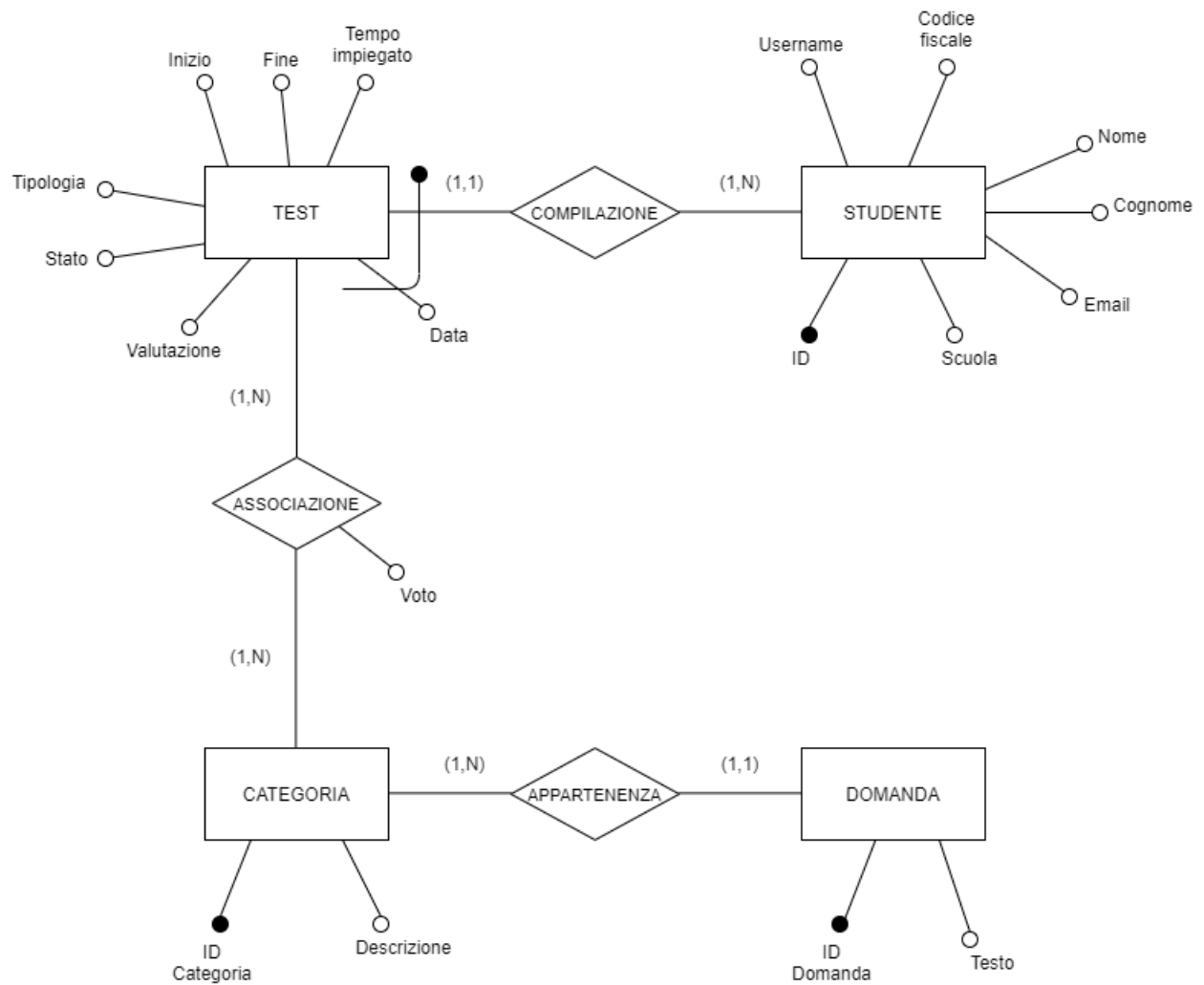
- L'unico attributo sarà il *Testo* della domanda.

In seguito si passa a costruire le Relazioni tra le entità appena introdotte:

- Compilazione: lega Studente e Test.
 - Ciascuno Studente che partecipa all'appello di esami OFA sostiene un solo Test ma può iscriversi a più appelli: (1,N) è la cardinalità che lega Studente a Test.
 - Ciascun Test può essere compilato solo una volta in ogni appello dallo Studente: (1,1) è la cardinalità che lega Test a Studente.
- Associazione: lega Test e Categoria. Questa relazione ha un attributo *Voto* in cui si riporta il voto ottenuto nella singola Domanda di una certa Categoria all'interno del Test.
 - Ciascun Test presenta minimo una Categoria di domande ma solitamente sono in numero maggiore: (1,N) è la cardinalità che lega Test a Categoria.
 - Ciascuna Categoria introdotta è presente minimo in un Test ma può essere presente in più Test: (1,N) è la cardinalità che lega Categoria a Test.
- Appartenenza: lega Categoria e Domanda.
 - Ciascuna Categoria presenta minimo una Domanda ma può averne molte: (1,N) è la cardinalità che lega Categoria a Domanda.
 - Ciascuna Domanda appartiene esclusivamente ad una Categoria: (1,1) è la cardinalità che lega Domanda a Categoria.

2.1.1 Diagramma E-R

L'analisi appena svolta ci permette di elaborare il diagramma E-R, obiettivo di questa prima fase di progettazione.



2.1.2 Dizionario dei dati

Associamo, a questo punto, allo schema E-R i dizionari riguardanti Entità e Relazioni così da facilitarne ulteriormente la comprensione.

- Dizionario delle Entità

Nome Entità	Descrizione	Attributi	Identificatore
Test	Informazioni relative ad ogni singola prova	Valutazione(float), Inizio(stringa), Fine(stringa), Tempo impiegato(stringa), Tipologia(stringa), Stato(stringa), Data(data)	ID Studente(stringa), Data(data)
Studente	Informazioni relative ai soggetti che sostengono la prova	ID(stringa), Username(stringa), Codice Fiscale(stringa), Scuola(stringa), Nome(stringa), Cognome(stringa), Email(stringa)	ID(stringa)
Categoria	Informazioni relative al tipo di domanda	ID Categoria(stringa), Descrizione(stringa)	ID Categoria(stringa)
Domanda	Informazioni relative alla particolare domanda	ID Domanda(int), Testo(stringa)	ID Domanda(int)

- Dizionario delle Relazioni

Nome Relazione	Descrizione	Entità coinvolte	Attributi
Compilazione	Associa lo studente al test svolto	Studente(1,N), Test(1,1)	-
Associazione	Lega il test alle categorie di cui è composto	Test(1,N), Categoria(1,N)	Voto(float)
Appartenenza	Associa ciascuna domanda alla categoria di cui fa parte	Categoria(1,N), Domanda(1,1)	-

2.1.3 Business Rules

1. Ogni studente sostiene in ogni appello per il conseguimento degli OFA un solo test.
2. Ogni studente deve essere obbligatoriamente identificato attraverso nome, cognome, codice fiscale ed e-mail per poter sostenere il test.
3. Ogni test è composto da 30 domande divise in 7 categorie.
4. Lo stato del test, al termine della prova, deve essere settato a "Completato".
5. Ogni domanda deve appartenere esclusivamente ad una categoria.
6. Le domande di Comprensione del testo occupano le prime cinque posizioni del test, qualunque sia la tipologia (A o B).
7. La valutazione del test deve corrispondere alla somma delle valutazioni ottenute in ogni singola domanda di cui è composto.
8. Il tempo impiegato dallo studente per sostenere il test deve corrispondere alla differenza tra l'orario di fine test (consegna) e l'orario di inizio.

2.2 Progettazione Logica

In questa fase si svolgono tutti i passaggi necessari per passare dal diagramma E-R appena costruito al Modello Relazionale, così da poter poi implementare il database vero e proprio.

2.2.1 Tavole dei volumi e delle operazioni

- Tavola dei volumi

Concetto	Tipo	Volume
Studente	E	1000
Test	E	1200
Categoria	E	8400
Domanda	E	36000
Compilazione	R	1200
Associazione	R	8400
Appartenenza	R	36000

Osserviamo i volumi delle relazioni:

- Consideriamo che in media uno studente superi il test al primo tentativo, con una piccola percentuale di casi in cui lo debba sostenere una seconda volta. Su 1000 studenti avremo approssimativamente 1200 compilazioni.
- Consideriamo che ogni test abbia 7 categorie: su 1200 test avremo 8400 associazioni.
- Poiché ogni domanda appartiene ad una categoria e ogni test è composto da 30 domande, su 1200 test avremo 36000 appartenenze di domande a categorie.

- Tavola delle operazioni

Operazione	Frequenza (Annuale)
1	600
2	600
3	600
4	1
5	50
6	1
7	1
8	1
9	1
10	20
11	3
12	3
13	1
14	180
15	6
16	6
17	1
18	6
19	6
20	1
21	6
22	6
23	1
24	6
25	1
26	6
27	1
28	6
29	6
30	1
31	1
32	1
33	6

34	1
35	6
36	6
37	1
38	6
39	1
40	6
41	1
42	6
43	1

2.2.2 Ristrutturazione dello schema concettuale

- Analisi delle ridondanze

Giunti ad un livello di progettazione logica che dovrà portare all'implementazione effettiva della base di dati è necessario svolgere un'analisi delle operazioni che sono state previste e valutare se certe informazioni derivabili dal nostro schema siano particolarmente interessate in termini di numero di accessi e costi computazionali e possa risultare utile introdurre alcune ridondanze o se alcune ridondanze già presenti siano eliminabili al fine di migliorare la funzionalità del nostro sistema.

Attributo *Valutazione* nell'entità *Test*

La valutazione totale di un dato test può essere ricavata dalla somma di tutti gli attributi *Voto* della relazione *Associazione*.

Valutiamo la possibilità di eliminare tale attributo dall'entità *Test*.

- Assenza di ridondanza:

Operazioni 15, 16, 17

Concetto	Costrutto	Accessi	Tipo
Associazione	R	30	L
Associazione	R	1	S
Test	E	1	L

Operazioni 18, 19, 20

Concetto	Costrutto	Accessi	Tipo
Associazione	R	30	L
Associazione	R	1	S
Studente	E	1	L

Operazioni 21, 22, 23

Concetto	Costrutto	Accessi	Tipo
Associazione	R	31	L
Associazione	R	1	S

- Presenza di ridondanza:
Operazioni 15, 16, 17

Concetto	Costrutto	Accessi	Tipo
Test	E	2	L

Operazioni 18, 19, 20

Concetto	Costrutto	Accessi	Tipo
Test	E	1	L
Studente	E	1	L

Operazioni 21, 22, 23

Concetto	Costrutto	Accessi	Tipo
Test	E	1	L
Associazione	E	1	L

Calcoliamo i costi totali:

- Assenza di ridondanza

Operazione	Costo	Frequenza	Totale
15, 16, 18, 19, 21, 22	33	6	1188
17, 20, 23	33	1	99

Totale=1287

- Presenza di ridondanza

Operazione	Costo	Frequenza	Totale
15, 16, 18, 19, 21, 22	2	6	72
17, 20, 23	2	1	6

Totale=78

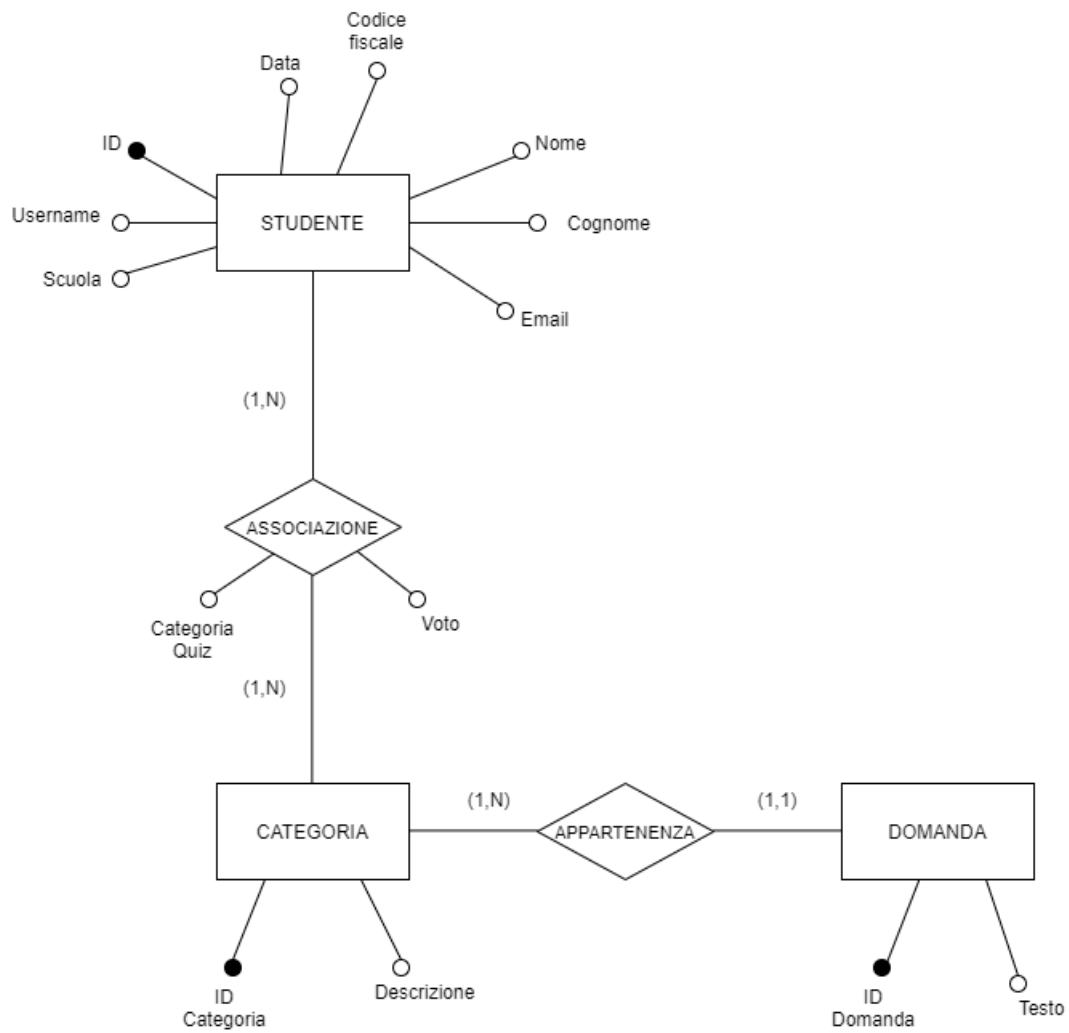
Conviene dunque mantenere l'attributo Valutazione nell'entità Test in quanto, così facendo, si ha un costo computazionale decisamente inferiore.

Osserviamo che nelle operazioni 16, 17, 19, 20, 22, 23, in cui subentra il concetto di raggruppamento per mese e per anno, si deve accedere un'altra volta in lettura ad Associazione (16, 17, 22, 23) o Studente (19, 20) ma ciò non fa variare i casi con o senza ridondanza e quindi queste operazioni sono state studiate insieme a quella più generale di ogni gruppo (ovvero le 15, 18 e 21).

- Eliminazione delle generalizzazioni: non sono presenti generalizzazioni.
- Partizionamento/accorpamento di concetti
Accorpamo l'entità Test in Studente in quanto lo studente è considerato come persona che ha fatto un particolare test (il suo ID è infatti composto da data di sostenimento test+numero progressivo).
Non essendo utili all'analisi (v. par. 1.7), possiamo eliminare gli attributi Inizio, Fine, Tempo Impiegato e Stato. L'attributo Data è unito a Studente mentre l'attributo Tipologia, non adatto ad essere unito al concetto di "studente" ma importante per lo studio (v. par. 1.7), viene inserito nella relazione Associazione.
Anche se l'analisi delle ridondanze ci ha consigliato di mantenere l'attributo Valutazione in Test, accorpando l'intera entità si ha un miglioramento computazionale che bilancia il peggioramento dovuto al calcolare le valutazioni totali dalle somme dei singoli voti.
- Scelta degli identificatori primari
Esplicitiamo la scelta degli identificatori primari di ciascuna entità presente dopo la fase di ristrutturazione.

Entità	Identificatore
Studente	ID
Categoria	ID Categoria
Domanda	ID Domanda

2.2.3 Schema ristrutturato finale



2.2.4 Normalizzazione

Analizzando lo schema ristrutturato si nota che tutte le relazioni presenti sono in forma normale di Boyce e Codd in quanto tutte binarie e che tra gli attributi delle entità non esistono dipendenze non banali.

2.2.5 Traduzione verso il modello relazionale

A questo punto possiamo tradurre ciascuna entità e relazione e mettere in risalto i vincoli di integrità referenziale presenti tra i suoi attributi.

Entità/Relazione	Traduzione	Vincoli di integrità referenziale
Studente	Studente(ID , Scuola, CodiceFiscale, Nome, Cognome, Email, Data, Username)	-
Categoria	Categoria(ID_Categoria , Descrizione)	-
Domanda	Domanda(ID_Domanda , Testo, CategAppartenenza)	CategAppartenenza → Categoria.ID_Categoria
Associazione	Associazione(ID_Test , ID_Categoria , Voto, CategoriaQuiz)	ID_Test → Studente.ID, ID_Categoria → Categoria.ID_Categoria

Osservazioni:

- La relazione Compilazione, con l'accorpamento di Test a Studente, è stata eliminata mentre la relazione Appartenenza, essendo del tipo *uno a molti*, è tradotta come attributo dell'entità che partecipa alla relazione con cardinalità (1,1). L'unica relazione che è tradotta in una tabella a sé stante è Associazione, essendo del tipo *molti a molti*.
- Chiamiamo l'ID dello studente nella tabella Associazione "ID_Test" perché indica lo studente che sostiene un particolare test.

Capitolo 3

Analisi dei dati

Una volta progettato il database lo si è successivamente implementato e popolato garantendo la privacy dei dati sensibili: la tabella *studente* è usufruibile solo dall'ente Università Politecnica delle Marche, mentre è utilizzabile dal tesista come *studentean*, ovvero attraverso una forma anonimizzata. Si è inoltre creata una tabella *iscrizione*, aggiuntiva a quelle a cui si è pervenuto tramite la progettazione, che presenta come attributi l'ID dello studente iscritto (ID), l'ID del corso (CDS_ID), il nome del corso (CDS_DES) e l'anno di iscrizione (AA_ID), per contenere ulteriori dati, esterni a quelli inizialmente a disposizione di cui si è parlato nel par. 1.1, riguardanti l'effettiva immatricolazione dei partecipanti al test.

Osserviamo che ogni studente può essere iscritto solo ad un corso alla volta e volendo cambiarlo deve fare richiesta di variazione in segreteria o se vuole ricominciare il percorso deve effettuare la rinuncia agli studi ed iscriversi nuovamente: possono quindi esserci casi in cui abbiamo stesso ID studente, stesso corso ma anno diverso oppure stesso ID studente ma corso diverso.

Si passa quindi ad analizzare i contenuti della banca dati tramite codice SQL implementato su *PhpMyAdmin* a cui si accede tramite una VPN assegnata dall'Università. I risultati delle queries sono riportati sotto i codici tramite screenshots.

Prima di procedere all'analisi è necessario fare una premessa: i dati con tutte le informazioni necessarie allo studio in questione a disposizione dell'Università provengono, attualmente, solo dagli anni 2016 e 2017 e dunque alcune queries daranno esiti per il momento non troppo significativi ma potranno essere usate per estrapolare importanti risultati all'aumentare della disponibilità dei dati.

- Voto medio per tipologia di quiz

```
select avg(votoTotale) as MediaVoto, categoriaQuiz as TipologiaQuiz
from (select sum(voto) as votoTotale, ID_Test, categoriaQuiz
      from associazione
      group by ID_Test) as Valutazioni
group by categoriaQuiz;
```

MediaVoto	TipologiaQuiz
4.7665059953405144	A
4.668464093933284	B

Osserviamo che non si perviene ad una grossa differenza tra i risultati medi ottenuti nelle due tipologie di quiz: dal momento che essi contengono le stesse domande i voti non variano in maniera significativa in base all'ordine in cui sono presentate.

- Voto medio per tipologia di quiz per sessione

```
select avg(votoTotale) as MediaVoto, categoriaQuiz as TipologiaQuiz,
       substring(ID_Test, 6, 2) as Mese
from (select sum(voto) as votoTotale, ID_Test, categoriaQuiz
      from associazione
      group by ID_Test) as Valutazioni
group by categoriaQuiz, substring(ID_Test, 6, 2);
```

MediaVoto	TipologiaQuiz	Mese
4.227200121879577	A	01
3.9975001104176044	A	04
5.363333515326182	A	05
5.690734137499004	A	07
4.490109724500408	A	09
4.9967827605164565	A	11
4.319091034206477	B	01
4.161951346368324	B	04
5.135937668848783	B	05
5.448105447857004	B	07
4.4234809594168425	B	09
4.874679043889046	B	11

Confermato ciò che si è evidenziato nel caso precedente, dividendo ulteriormente per appello, notiamo che i migliori risultati si ottengono nei mesi di maggio e luglio. Questo è probabilmente dovuto al fatto che gli studenti hanno completato il programma didattico di Matematica e Fisica (materie di cui il test è composto in maggior parte) del quinto anno dell'istituto che frequentano. Si riscontra una buona media anche nell'appello di novembre, effettuato durante le lezioni del primo semestre del primo anno universitario per coloro che non hanno superato il test in precedenza o non l'hanno sostenuto in assoluto e per questo partono con un *debito formativo*. Ciò potrebbe essere dovuto al fatto che gli studenti hanno preso confidenza con tali tematiche frequentando i corsi iniziali di Analisi Matematica e Fisica.

- Voto medio per tipologia di quiz per anno

```
select avg(votoTotale) as MediaVoto, categoriaQuiz as TipologiaQuiz,
       substring(ID_Test, 1, 4) as Anno
from (select sum(voto) as votoTotale, ID_Test, categoriaQuiz
      from associazione
      group by ID_Test) as Valutazioni
group by categoriaQuiz, substring(ID_Test, 1, 4);
```

MediaVoto	TipologiaQuiz	Anno
4.7869652469892605	A	2016
4.227200121879577	A	2017
4.680941703209243	B	2016
4.319091034206477	B	2017

Osserviamo che l'anno in cui sono stati ottenuti i risultati migliori, indipendentemente dalla tipologia di quiz, è il 2016. Da notare che nel 2016 si hanno risultati migliori nel test A, mentre nel 2017 accade il contrario.

Implementare questa query è importante, anche se al momento non fornisce risultati notevoli, perché all'aumentare dei dati provenienti da più anni si possono analizzare le cause di un eventuale miglioramento o peggioramento nei risultati. Può essere interessante, ad esempio, sfruttarla nell'attuale anno accademico per vedere se la chiusura degli istituti (a causa del virus Covid-19) ha influito sulla preparazione dei ragazzi.

- Voto medio per scuola di provenienza

```
select avg(votoTotale) as MediaVoto, count(Username) as NumeroStudenti,
       Scuola
from (select sum(voto) as votoTotale, ID_Test, Scuola, Username
      from associazione, studentean
      where ID_Test=ID
      group by ID_Test) as Valutazioni
group by Scuola;
```

MediaVoto	NumeroStudenti	Scuola
4.104235414021155	85	
4.1879071042288185	43	Altra scuola italiana
4.592000126838684	5	Altra scuola straniera
4.110000133514404	1	Athene Royal Crommelynck
4.597924660961583	53	Classico
4.680000141263008	1	I.I.S Volterra ELia
4.930000156164169	1	i.i.s. volterra elia
3.6800000816583633	1	INDUSTRIALE
6.100000232458115	1	Istituto d'Arte
4.94250014051795	4	Istituto Magistrale
3.580000097552935	3	Istituto Professionale Alberghiero
3.7175000961869955	8	Istituto Professionale Commerciale
4.098077045037196	26	Istituto Professionale Industria e Artigianato
4.113333473602931	3	Istituto tecnico
3.684615488235767	13	Istituto Tecnico Agrario
4.23079376892438	63	Istituto Tecnico Commerciale
4.172391420581203	138	Istituto Tecnico Commerciale per Geometri
4.411460808377141	267	Istituto Tecnico Industriale
5.10000017285347	1	Istituto tecnico Nautico
2.360000029206276	1	Istituto Tecnico per il Turismo
4.895000167191029	2	Istituto Tecnico Sperimentale
4.590000122785568	1	ITET Enrico Fermi Lanciano
3.3000000715255737	11	Liceo Artistico
5.353333509630627	9	liceo scientifico
4.1700000911951065	1	Liceo scientifico "A. Einstein"
4.200909214940938	11	Pedagogico
5.300088046200456	569	Scientifico

Osserviamo che i risultati medi più alti sono ottenuti da studenti provenienti dall'Istituto d'Arte, dall'Istituto Tecnico Nautico e dal Liceo Scientifico. E' stato riportato anche il numero di ragazzi che hanno sostenuto la prova divisi per istituto perché è importante evidenziare che gli studenti che hanno sostenuto il test provenienti dai primi due istituti sono pochi (uno per entrambi) mentre quelli provenienti dallo Scientifico sono in maggior numero (579 in totale) e quindi l'alta media ottenuta da quest'ultima scuola è molto più significativa rispetto alle prime.

I risultati medi più bassi sono invece ottenuti da studenti provenienti dall'Istituto Tecnico per il Turismo, dal Liceo Artistico e dall'Istituto Professionale Alberghiero, probabilmente

perché trattano in modo meno approfondito, nel corso dei cinque anni di studio, gli argomenti di cui è composto il test d'ingresso per la facoltà di Ingegneria. Anche qui bisogna soffermarsi sul numero di studenti che ha sostenuto il test: il basso risultato ottenuto dal Liceo Artistico è più significativo perché ottenuto da un campione formato da più alunni (11) rispetto a quelli degli altri due istituti che hanno avuto invece pochi iscritti (rispettivamente 1 e 3).

- Voto medio per scuola di provenienza per sessione

```

select avg(votoTotale) as MediaVoto, count(Username) as NumeroStudenti,
       Scuola, month(data) as Mese
from (select sum(voto) as votoTotale, ID_Test, Scuola, Username, data
      from associazione, studentean
      where ID_Test=ID
      group by ID_Test) as Valutazioni
group by Scuola, month(data);

```

MediaVoto	NumeroStudenti	Scuola	Mese
4.08074085947908	81		4
5.10000017285347	1		5
4.200000151991844	1		7
4.510000124573708	2		11
4.420000106096268	1	Altra scuola italiana	1
4.335000152389209	6	Altra scuola italiana	7
3.95344839168006	29	Altra scuola italiana	9
5.000000157526562	7	Altra scuola italiana	11
5.160000130534172	1	Altra scuola straniera	1
4.520000144839287	1	Altra scuola straniera	7
4.4100000858306885	1	Altra scuola straniera	9
4.4350001364946365	2	Altra scuola straniera	11
4.110000133514404	1	Athene Royal Crommelynck	9
3.7466667741537094	3	Classico	5
5.204444610410267	9	Classico	7
4.257500122168234	28	Classico	9
5.1076924456999855	13	Classico	11
4.680000141263008	1	I.I.S. Volterra ELia	5
4.930000156164169	1	i.i.s. volterra elia	5
3.6800000816583633	1	INDUSTRIALE	9
6.100000232458115	1	Istituto d'Arte	7
5.4200001657009125	1	Istituto Magistrale	7
5.090000152587891	2	Istituto Magistrale	9
4.1700000911951065	1	Istituto Magistrale	11
4.090000092983246	1	Istituto Professionale Alberghiero	7
3.32500009983778	2	Istituto Professionale Alberghiero	9
3.520000085234642	1	Istituto Professionale Commerciale	1
2.700000062584877	1	Istituto Professionale Commerciale	5
3.3600000888109207	1	Istituto Professionale Commerciale	7
3.470000093181928	3	Istituto Professionale Commerciale	9
4.87500012665987	2	Istituto Professionale Commerciale	11
4.260000109672546	1	Istituto Professionale Industria e Artigianato	1
6.020000234246254	1	Istituto Professionale Industria e Artigianato	5
4.180000111460686	1	Istituto Professionale Industria e Artigianato	7
3.7269231780217242	13	Istituto Professionale Industria e Artigianato	9
4.3640001401305195	10	Istituto Professionale Industria e Artigianato	11
4.610000163316727	2	Istituto tecnico	9
3.1200000941753387	1	Istituto Tecnico	11
3.3000001311302185	1	Istituto Tecnico Agrario	1
3.5350000951439142	8	Istituto Tecnico Agrario	9
4.08000011369586	4	Istituto Tecnico Agrario	11

4.310000129044056	4 Istituto Tecnico Commerciale	1
4.6000001430511475	3 Istituto Tecnico Commerciale	5
4.541666800777118	6 Istituto Tecnico Commerciale	7
3.977222327970796	36 Istituto Tecnico Commerciale	9
4.647857277521065	14 Istituto Tecnico Commerciale	11
4.015384721068235	13 Istituto Tecnico Commerciale per Geometri	1
4.624000126123429	5 Istituto Tecnico Commerciale per Geometri	5
4.848333487908046	12 Istituto Tecnico Commerciale per Geometri	7
3.908000106612841	75 Istituto Tecnico Commerciale per Geometri	9
4.5209092177224885	33 Istituto Tecnico Commerciale per Geometri	11
4.317857276116099	14 Istituto Tecnico Industriale	1
5.057500168681145	12 Istituto Tecnico Industriale	5
4.83090925487605	33 Istituto Tecnico Industriale	7
4.09353753518896	147 Istituto Tecnico Industriale	9
4.84508211251165	61 Istituto Tecnico Industriale	11
5.10000017285347	1 Istituto tecnico Nautico	11
2.360000029206276	1 Istituto Tecnico per il Turismo	9
5.750000178813934	1 Istituto Tecnico Sperimentale	1
4.040000155568123	1 Istituto Tecnico Sperimentale	5
4.590000122785568	1 ITET Enrico Fermi Lanciano	11
3.116666735874282	9 Liceo Artistico	9
4.125000081956387	2 Liceo Artistico	11
4.470000125467777	2 liceo scientifico	5
6.257500223815441	4 Liceo Scientifico	7
4.736666813492775	3 Liceo Scientifico	11
4.1700000911951065	1 Liceo scientifico "A. Einstein"	9
4.590000122785568	1 Pedagogico	1
4.280000150203705	2 Pedagogico	7
4.102000117301941	5 Pedagogico	9
4.183333451549212	3 Pedagogico	11
4.4233334643973246	9 Scientifico	1
5.806129234933084	31 Scientifico	5
6.0376002069711685	125 Scientifico	7
4.9590589812134995	340 Scientifico	9
5.5495314272120595	64 Scientifico	11

Questa query ci permette di analizzare, per ogni scuola, il risultato medio ottenuto nei vari appelli per poter individuare successivamente le possibili cause di un miglioramento o peggioramento. Ad esempio i risultati di chi proviene dal Liceo Classico e dall'Istituto Tecnico Industriale sono migliori a luglio e novembre mentre a settembre, al crescere degli studenti iscritti al test, calano. Si può dedurre quindi che la partecipazione alla prova a ridosso delle lezioni scolastiche o accademiche porti a migliori risultati e che al crescere del numero di partecipanti ci sia qualche risultato particolarmente basso che influenza la media totale.

Per gli alunni provenienti dall'Istituto Tecnico Commerciale per Geometri vale la stessa analisi ma va osservato che presenta un buon numero di partecipanti, rapportandoci a quelli degli altri appelli, anche nel mese di gennaio, ma in questo caso i risultati sono in calo, probabilmente per il fatto di non aver chiaro tutto il programma di studi per poter sostenere con solidità il test.

Analisi differente invece per gli studenti provenienti dal Liceo Scientifico che in ogni appello, indipendentemente dal numero di partecipanti, mantengono buoni risultati, tanto che l'aumentare dei partecipanti nel mese di settembre non causa un significativo crollo della votazione media.

- Voto medio per scuola di provenienza per anno

```
select avg(votoTotale) as MediaVoto, count(Username) as NumeroStudenti,
       Scuola, year(data) as Anno
from (select sum(voto) as votoTotale, ID_Test, Scuola, Username, data
      from associazione, studentean
      where ID_Test=ID
      group by ID_Test) as Valutazioni
group by Scuola, year(data);
```

MediaVoto	NumeroStudenti	Scuola	Anno
4.104235414021155	85		2016
4.182381080374832	42	Altra scuola italiana	2016
4.420000106096268	1	Altra scuola italiana	2017
4.450000125914812	4	Altra scuola straniera	2016
5.160000130534172	1	Altra scuola straniera	2017
4.110000133514404	1	Athene Royal Crommelynck	2016
4.597924660961583	53	Classico	2016
4.680000141263008	1	I.I.S Volterra ELia	2016
4.930000156164169	1	i.i.s. volterra elia	2016
3.6800000816583633	1	INDUSTRIALE	2016
6.100000232458115	1	Istituto d 'Arte	2016
4.94250014051795	4	Istituto Magistrale	2016
3.580000097552935	3	Istituto Professionale Alberghiero	2016
3.745714383465903	7	Istituto Professionale Commerciale	2016
3.520000085234642	1	Istituto Professionale Commerciale	2017
4.091600122451783	25	Istituto Professionale Industria e Artigianato	2016
4.260000109672546	1	Istituto Professionale Industria e Artigianato	2017
4.113333473602931	3	Istituto tecnico	2016
3.7166667679945626	12	Istituto Tecnico Agrario	2016
3.3000001311302185	1	Istituto Tecnico Agrario	2017
4.225423846204402	59	Istituto Tecnico Commerciale	2016
4.310000129044056	4	Istituto Tecnico Commerciale	2017
4.1887201173305515	125	Istituto Tecnico Commerciale per Geometri	2016
4.015384721068235	13	Istituto Tecnico Commerciale per Geometri	2017
4.416640450478543	253	Istituto Tecnico Industriale	2016
4.317857276116099	14	Istituto Tecnico Industriale	2017
5.10000017285347	1	Istituto tecnico Nautico	2016
2.360000029206276	1	Istituto Tecnico per il Turismo	2016
4.040000155568123	1	Istituto Tecnico Sperimentale	2016
5.750000178813934	1	Istituto Tecnico Sperimentale	2017
4.590000122785568	1	ITET Enrico Fermi Lanciano	2016
3.3000000715255737	11	Liceo Artistico	2016
5.353333509630627	9	liceo scientifico	2016
4.1700000911951065	1	Liceo scientifico "A. Einstein"	2016
4.162000124156475	10	Pedagogico	2016
4.590000122785568	1	Pedagogico	2017
5.314178744836577	560	Scientifico	2016
4.4233334643973246	9	Scientifico	2017

Questa query ci permette di esaminare, al crescere del numero dei dati, per ogni scuola, il risultato medio ottenuto nei vari anni per individuare le possibili cause di variazione degli esiti. Anche in questo caso sarebbe interessante studiare i risultati per l'attuale anno accademico per vedere se la chiusura delle scuole abbia determinato un calo generale nelle valutazioni medie e quale tipo di scuola ne abbia risentito maggiormente.

- Voto medio per categoria

```
select avg(votoTotale) as MediaVoto,  
       substring(ID_Categoria, 4, 3) as Categoria  
from (select sum(voto) as votoTotale, ID_Test, ID_Categoria  
      from associazione  
      group by ID_Test, ID_Categoria) as Valutazioni  
group by substring(ID_Categoria, 4, 3);
```

MediaVoto	Categoria
0.13760212191722937	ALG
0.11982299795038942	FIS
0.12442889902318029	FUN
0.14090582808133062	GEO
0.13982602507552291	GRA
0.1635608984588585	LOG
0.27038427679693827	TES

Questa query è utile per studiare i risultati ottenuti nelle varie categorie e individuare le più complicate e le più semplici da affrontare.

Osserviamo che il voto medio più basso è ottenuto nella categoria di Fisica mentre quello più alto è conseguito nella categoria di Comprensione del testo.

- Voto medio per categoria per sessione

```

select avg(votoTotale) as MediaVoto,
       substring(ID_Categoria, 4, 3) as Categoria,
       substring(ID_Test, 6, 2) as Mese
from (select sum(voto) as votoTotale, ID_Test, ID_Categoria
      from associazione
      group by ID_Test, ID_Categoria) as Valutazioni
group by substring(ID_Categoria, 4, 3), substring(ID_Test, 6, 2);

```

	MediaVoto	Categoria	Mese
0.12829787601181802	ALG	01	
0.13864197903945122	ALG	04	
0.12983871371515335	ALG	05	
0.16536765235165754	ALG	07	
0.11796165080952713	ALG	09	
0.1777678624992924	ALG	11	
0.08140425498181202	FIS	01	
0.10103703902827368	FIS	04	
0.1337741976303439	FIS	05	
0.12267647349074776	FIS	07	
0.11983807071281427	FIS	09	
0.12816964516948376	FIS	11	
0.13494681257833827	FUN	01	
0.09996913780492765	FUN	04	
0.15120968180558375	FUN	05	
0.16416054437229155	FUN	07	
0.11233665070242503	FUN	09	
0.12547433364670724	FUN	11	
0.13452127980107956	GEO	01	
0.11462963225296986	GEO	04	
0.15754032723845973	GEO	05	
0.17904412353813065	GEO	07	
0.13433238944377413	GEO	09	
0.13306919955981097	GEO	11	

0.11489361968446285	GRA	01
0.09129629817641811	GRA	04
0.18774194186252932	GRA	05
0.19205882988285786	GRA	07
0.13168324216861615	GRA	09
0.12736607452721468	GRA	11
0.13271277072898885	LOG	01
0.123796300801598	LOG	04
0.15870968312505754	LOG	05
0.20145834105856278	LOG	07
0.1477450335832228	LOG	09
0.2009486678621865	LOG	11
0.24565958406062835	TES	01
0.27244445504965603	TES	04
0.29345162438769495	TES	05
0.28284314832851	TES	07
0.2665085330859504	TES	09
0.2692767962148147	TES	11

Confermiamo ciò che è stato asserito nell'analisi dei voti medi divisi per tipologia di quiz nelle varie sessioni svolta in precedenza, ovvero che i risultati medi delle categorie riguardanti argomenti di Matematica e Fisica sono migliori nei mesi di maggio, luglio e novembre. Questo perché sono i periodi in cui gli studenti hanno già approfondito tali materie: a maggio sono alla conclusione del percorso didattico all'interno dell'istituto che frequentano, a luglio hanno terminato o stanno terminando l'esame di maturità così da avere piena consapevolezza dell'argomento, a novembre si presume che abbiano già iniziato a seguire i corsi universitari, che nel primo anno affrontano nello specifico tali tematiche. Osserviamo inoltre che si hanno alte valutazioni medie nel corso di tutte le sessioni senza distinzione nella categoria di Comprensione del testo, il che avvalorava quanto visto in precedenza su come essa presenti il set di domande più semplice da affrontare.

- Voto medio per categoria per anno

```
select avg(votoTotale) as MediaVoto,
       substring(ID_Categoria, 4, 3) as Categoria,
       substring(ID_Test, 1, 4) as Anno
from (select sum(voto) as votoTotale, ID_Test, ID_Categoria
      from associazione
      group by ID_Test, ID_Categoria) as Valutazioni
group by substring(ID_Categoria, 4, 3), substring(ID_Test, 1, 4);
```

MediaVoto	Categoria	Anno
0.13794510196237003	ALG	2016
0.12829787601181802	ALG	2017
0.12123921827942717	FIS	2016
0.08140425498181202	FIS	2017
0.12404117985683329	FUN	2016
0.13494681257833827	FUN	2017
0.14114118005715165	GEO	2016
0.13452127980107956	GEO	2017
0.14074510198013454	GRA	2016
0.11489361968446285	GRA	2017
0.16469804512811642	LOG	2016
0.13271277072898885	LOG	2017
0.27129569684290417	TES	2016
0.24565958406062835	TES	2017

Questa query ci consente di studiare eventuali miglioramenti o peggioramenti nei risultati di una particolare categoria nel corso degli anni. Va implementata soprattutto nei casi in cui varia un set di domande della categoria o il numero delle domande di cui essa è composta. Sono confermati i risultati dell'analisi generale precedente sulle votazioni medie nelle varie categorie. Notiamo che nel 2017 nella categoria di Fisica si è addirittura scesi sotto lo 0.1 di valutazione media avvalorando quanto osservato precedentemente su come essa sia la categoria più difficoltosa da affrontare nel test.

- Voto medio per domanda

```
select avg(Voto) as MediaVoto, associazione.ID_Categoria as Domanda,
       Descrizione
from associazione, categoria
where associazione.ID_Categoria=categoria.ID_Categoria
group by associazione.ID_Categoria;
```

MediaVoto	Domanda	Descrizione
0.20873016656256047	01-LOG	sillogismi
0.11773998870916287	02-LOG	negazione
0.15006803286994666	03-LOG	implicazioni logiche
0.1772108905056797	04-LOG	problema
0.18419501776733097	05-GEO	piano cartesiano e coordinate cartesiane di un pun...
0.1266213176900117	06-GEO	distanza, aree e volumi
0.14963719249897983	07-GEO	rette nel piano
0.10274376523647899	08-GEO	circonferenze, parabole, iperboli
0.13083900575464816	09-ALG	polinomi
0.14415722277834842	10-ALG	numeri irrazionali, razionali, interi
0.08752078777538136	11-FUN	funzioni composte e funzione inversa
0.15362056438794933	12-FUN	proprietà delle potenze
0.1145653849240964	13-FUN	proprietà del modulo, equazioni e disequazioni con...
0.09656840657430982	14-FUN	proprietà dei logaritmi
0.1090022691793723	15-FUN	angoli e trigonometria
0.19536659784913513	16-FUN	equazioni e disequazioni algebriche intere
0.1700226819411607	17-FUN	equazioni e disequazioni algebriche fratte
0.06801209502479658	18-FUN	equazioni e disequazioni irrazionali
0.13157974650980575	19-GRA	operazioni sui grafici
0.1478609264302416	20-GRA	grafici funzioni pari, dispari, periodiche, confro...
0.140914591396747	21-FIS	fisica
0.1045578252679006	22-FIS	fisica
0.09363567706363632	23-FIS	fisica
0.13866969372004337	24-FIS	fisica
0.12088435650637722	25-FIS	fisica
0.2709221572292094	26-TES	comprensione del testo
0.2687981964419128	27-TES	comprensione del testo
0.2710355358959956	28-TES	comprensione del testo
0.2699319832336695	29-TES	comprensione del testo
0.2702116507273773	30-TES	comprensione del testo

Osserviamo che i voti medi più alti sono stati ottenuti nelle domande di Comprensione del testo, il che avvalorava quanto mostrato nell'analisi precedente, e nella prima domanda (Logica), che riguarda i sillogismi. I voti medi più bassi invece scaturiscono dalle domande 11: funzioni composte e funzioni inverse, 14: proprietà dei logaritmi, 18: equazioni e disequazioni irrazionali (Funzioni) e 23 (Fisica).

Deduciamo quindi che per gli studenti, al momento del test, è più semplice operare su domande di puro ragionamento piuttosto che su quesiti di carattere matematico.

- Voto medio per domanda per anno

```
select avg(Voto) as MediaVoto, ID_Categoria as Domanda,
       substring(ID_Test, 1, 4) as Anno
from associazione
group by ID_Categoria, substring(ID_Test, 1, 4);
```

MediaVoto	Domanda	Anno
0.21222571330290976	01-LOG	2016
0.11382979122882193	01-LOG	2017
0.11794671228266435	02-LOG	2016
0.11212766360729298	02-LOG	2017
0.15307994308329675	03-LOG	2016
0.06829787473729317	03-LOG	2017
0.17502351703128097	04-LOG	2016
0.23659575334254732	04-LOG	2017
0.18524295318182732	05-GEO	2016
0.15574468608866346	05-GEO	2017
0.12460031588109309	06-GEO	2016
0.18148936680022706	06-GEO	2017
0.15066614855653065	07-GEO	2016
0.12170213017057865	07-GEO	2017
0.10361285376885095	08-GEO	2016
0.0791489361448491	08-GEO	2017
0.13265674338211833	09-ALG	2016
0.08148936293226608	09-ALG	2017
0.14301724564926377	10-ALG	2016
0.17510638909136994	10-ALG	2017
0.08795454719784118	11-FUN	2016
0.07574468090179119	11-FUN	2017
0.15463166651213806	12-FUN	2016
0.12617021735678327	12-FUN	2017
0.11402821636592333	13-FUN	2016
0.1291489398225825	13-FUN	2017
0.09420062825784414	14-FUN	2016
0.16085106895325033	14-FUN	2017
0.10547805784217616	15-FUN	2016
0.20468085782325013	15-FUN	2017

0.19664577486980298	16-FUN	2016
0.16063830256462097	16-FUN	2017
0.17273511602119965	17-FUN	2016
0.09638298223627374	17-FUN	2017
0.0658777441397356	18-FUN	2016
0.12595745096815394	18-FUN	2017
0.13251567752243584	19-GRA	2016
0.10617021518818875	19-GRA	2017
0.14875392282971395	20-GRA	2016
0.12361702418073695	20-GRA	2017
0.14306426678892215	21-FIS	2016
0.08255319138790698	21-FIS	2017
0.10569749440892737	22-FIS	2016
0.07361702050300355	22-FIS	2017
0.09333072156648277	23-FIS	2016
0.10191489439061348	23-FIS	2017
0.14119906335788834	24-FIS	2016
0.0699999988713163	24-FIS	2017
0.12242946996817768	25-FIS	2016
0.07893616975621974	25-FIS	2017
0.2715282237805356	26-TES	2016
0.2544680951123542	26-TES	2017
0.26932602932787614	27-TES	2016
0.2544680951123542	27-TES	2017
0.27164577862078493	28-TES	2016
0.2544680951123542	28-TES	2017
0.271410668811828	29-TES	2016
0.2297872428564315	29-TES	2017
0.27150471276110244	30-TES	2016
0.2351063921096477	30-TES	2017

Questa query permette di analizzare eventuali miglioramenti o peggioramenti nei risultati di una particolare domanda col passare degli anni e ciò può risultare utile perché una variazione significativa può essere ricercata ad esempio nella sua modifica e si può valutare così se sia troppo complicata o troppo semplice da rispondere e quindi se valga la pena mantenere tale mutamento.

- Numero studenti che hanno sostenuto il test ogni mese

```
select count(*) as NumeroStudenti, month(data) as Mese
from studentean
group by month(Data);
```

NumeroStudenti	Mese
47	1
81	4
62	5
204	7
704	9
224	11

Osserviamo innanzitutto che gli studenti sostengono principalmente il test negli appelli di luglio, settembre e novembre e dunque dopo l'esame di maturità che viene completato nel periodo che va dalla metà di giugno ai primi di luglio. In particolare, il mese in cui si hanno più iscritti alla prova è settembre, dopo un periodo in cui i ragazzi si sono potuti concentrare sulla preparazione senza altri impegni scolastici. I mesi con il minor numero di partecipanti sono gennaio, perché non si è ancora terminato il programma del quinto anno e quindi non si hanno robuste basi per affrontare un test di questo tipo, e maggio, probabilmente a causa del fatto che gli studenti preferiscono concentrarsi sullo studio in vista del sopracitato esame dell'ultimo anno.

- Numero studenti che hanno sostenuto il test ogni anno

```
select count(*) as NumeroStudenti, year(data) as Anno
from studentean
group by year(data);
```

NumeroStudenti	Anno
1275	2016
47	2017

Questa query è molto importante in quanto, all'aumentare dei dati a disposizione da più anni, permette di studiare se c'è un incremento o un decremento nell'interesse all'iscrizione presso l'Università Politecnica delle Marche e di ottimizzare così l'orientamento e la pubblicità.

- Tipologia scuola di provenienza di chi ha sostenuto il test

```
select count(*) as NumeroStudenti, Scuola
from studentean
group by Scuola;
```

NumeroStudenti	Scuola
85	
43	Altra scuola italiana
5	Altra scuola straniera
1	Athene Royal Crommelynck
53	Classico
1	I.I.S Volterra ELia
1	i.i.s. volterra elia
1	INDUSTRIALE
1	Istituto d 'Arte
4	Istituto Magistrale
3	Istituto Professionale Alberghiero
8	Istituto Professionale Commerciale
26	Istituto Professionale Industria e Artigianato
3	Istituto tecnico
13	Istituto Tecnico Agrario
63	Istituto Tecnico Commerciale
138	Istituto Tecnico Commerciale per Geometri
267	Istituto Tecnico Industriale
1	Istituto tecnico Nautico
1	Istituto Tecnico per il Turismo
2	Istituto Tecnico Sperimentale
1	ITET Enrico Fermi Lanciano
11	Liceo Artistico
9	liceo scientifico
1	Liceo scientifico "A. Einstein"
11	Pedagogico
569	Scientifico

Osserviamo che gli studenti che sostengono il test provengono soprattutto da Liceo Scientifico o Istituto Tecnico Industriale, scuole che preparano meglio a questo tipo di percorso per via delle materie affrontate nel corso dei cinque anni. Si ha comunque una buona affluenza dal Liceo Classico e dagli Istituti Tecnici Commerciali e per Geometri.

Questa query è interessante perché, allacciandosi all'analisi precedente, permette di studiare come poter organizzare la promozione dei corsi presentati dall'Università tra i vari istituti così da poter suscitare l'interesse di nuovi potenziali studenti anche da scuole da cui, storicamente, non provengono molte iscrizioni.

- Tipologia scuola di provenienza di chi ha sostenuto il test per sessione

```
select count(*) as NumeroStudenti, month(data) as Mese, Scuola
from studentean
group by month(data), Scuola;
```

NumeroStudenti	Mese	Scuola
1	1	Altra scuola italiana
1	1	Altra scuola straniera
1	1	Istituto Professionale Commerciale
1	1	Istituto Professionale Industria e Artigianato
1	1	Istituto Tecnico Agrario
4	1	Istituto Tecnico Commerciale
13	1	Istituto Tecnico Commerciale per Geometri
14	1	Istituto Tecnico Industriale
1	1	Istituto Tecnico Sperimentale
1	1	Pedagogico
9	1	Scientifico
81	4	
1	5	
3	5	Classico
1	5	I.I.S. Volterra ELia
1	5	i.i.s. volterra elia
1	5	Istituto Professionale Commerciale
1	5	Istituto Professionale Industria e Artigianato
3	5	Istituto Tecnico Commerciale
5	5	Istituto Tecnico Commerciale per Geometri
12	5	Istituto Tecnico Industriale
1	5	Istituto Tecnico Sperimentale
2	5	liceo scientifico
31	5	Scientifico
1	7	
6	7	Altra scuola italiana
1	7	Altra scuola straniera
9	7	Classico
1	7	Istituto d'Arte
1	7	Istituto Magistrale
1	7	Istituto Professionale Alberghiero
1	7	Istituto Professionale Commerciale
1	7	Istituto Professionale Industria e Artigianato
6	7	Istituto Tecnico Commerciale
12	7	Istituto Tecnico Commerciale per Geometri
33	7	Istituto Tecnico Industriale
4	7	Liceo Scientifico
2	7	Pedagogico

125	7	Scientifico
29	9	Altra scuola italiana
1	9	Altra scuola straniera
1	9	Athene Royal Crommelynck
28	9	Classico
1	9	INDUSTRIALE
2	9	Istituto Magistrale
2	9	Istituto Professionale Alberghiero
3	9	Istituto Professionale Commerciale
13	9	Istituto Professionale Industria e Artigianato
2	9	Istituto tecnico
8	9	Istituto Tecnico Agrario
36	9	Istituto Tecnico Commerciale
75	9	Istituto Tecnico Commerciale per Geometri
147	9	Istituto Tecnico Industriale
1	9	Istituto Tecnico per il Turismo
9	9	Liceo Artistico
1	9	Liceo scientifico "A. Einstein"
5	9	Pedagogico
340	9	Scientifico
2	11	
7	11	Altra scuola italiana
2	11	Altra scuola straniera
13	11	Classico
1	11	Istituto Magistrale
2	11	Istituto Professionale Commerciale
10	11	Istituto Professionale Industria e Artigianato
1	11	Istituto Tecnico
4	11	Istituto Tecnico Agrario
14	11	Istituto Tecnico Commerciale
33	11	Istituto Tecnico Commerciale per Geometri
61	11	Istituto Tecnico Industriale
1	11	Istituto tecnico Nautico
1	11	ITET Enrico Fermi Lanciano
2	11	Liceo Artistico
3	11	Liceo Scientifico
3	11	Pedagogico
64	11	Scientifico

Questa query ci consente di prendere in considerazione i mesi in cui gli studenti dalle diverse scuole preferiscono sostenere il test. A gennaio si ha una prevalenza di studenti provenienti dagli Istituti Tecnici mentre negli altri appelli cedono il primato di partecipazione agli studenti dei Licei Scientifici pur mantenendo sempre una buona percentuale di partecipanti.

- Tipologia scuola di provenienza di chi ha sostenuto il test per anno

```
select count(*) as NumeroStudenti, year(data) as Anno, Scuola
from studentean
group by year(data), Scuola;
```

NumeroStudenti	Anno	Scuola
85	2016	
42	2016	Altra scuola italiana
4	2016	Altra scuola straniera
1	2016	Athene Royal Crommelynck
53	2016	Classico
1	2016	I.I.S Volterra ELia
1	2016	i.i.s. volterra elia
1	2016	INDUSTRIALE
1	2016	Istituto d'Arte
4	2016	Istituto Magistrale
3	2016	Istituto Professionale Alberghiero
7	2016	Istituto Professionale Commerciale
25	2016	Istituto Professionale Industria e Artigianato
3	2016	Istituto tecnico
12	2016	Istituto Tecnico Agrario
59	2016	Istituto Tecnico Commerciale
125	2016	Istituto Tecnico Commerciale per Geometri
253	2016	Istituto Tecnico Industriale
1	2016	Istituto tecnico Nautico
1	2016	Istituto Tecnico per il Turismo
1	2016	Istituto Tecnico Sperimentale
1	2016	ITET Enrico Fermi Lanciano
11	2016	Liceo Artistico
9	2016	liceo scientifico
1	2016	Liceo scientifico "A. Einstein"
10	2016	Pedagogico
560	2016	Scientifico
1	2017	Altra scuola italiana
1	2017	Altra scuola straniera
1	2017	Istituto Professionale Commerciale
1	2017	Istituto Professionale Industria e Artigianato
1	2017	Istituto Tecnico Agrario
4	2017	Istituto Tecnico Commerciale
13	2017	Istituto Tecnico Commerciale per Geometri
14	2017	Istituto Tecnico Industriale
1	2017	Istituto Tecnico Sperimentale
1	2017	Pedagogico
9	2017	Scientifico

Questa query è utile poiché ci permette di esaminare l'interessamento degli studenti provenienti da determinate scuole a frequentare la facoltà di Ingegneria nel corso degli anni così da poter favorire un più vasto orientamento nel caso di un eventuale calo.

- Numero di iscrizioni per anno

```
select count(*) as NumeroIscritti, AA_ID as Anno
from iscrizione
group by AA_ID;
```

NumeroIscritti	Anno
1	2011
3	2014
23	2015
1058	2016
23	2017
6	2018
120	2019
45	2020

Osserviamo innanzitutto che, sebbene i dati a nostra disposizione provengano solo dagli anni 2016 e 2017, abbiamo anni di iscrizione che vanno dal 2011 al 2020: questo perché alcuni studenti hanno sostenuto nuovamente il test dopo essersi iscritti e, in seguito, aver fatto la rinuncia agli studi (iscritti del 2011, 2014, 2015), oppure si sono immatricolati anni dopo aver sostenuto il test (2018, 2019, 2020). Da notare comunque, dal confronto con l'analisi già svolta sul numero di studenti che sostiene il test ogni anno, che c'è un'alta percentuale di iscrizioni: nel 2016 su 1275 ragazzi che hanno sostenuto il test se ne sono iscritti 1058 e nel 2017, su 47 se ne sono iscritti 23. Questo tralasciando inoltre che alcuni dei non iscritti nello stesso anno può essersi iscritto negli anni seguenti.

- Numero di iscrizioni per anno per scuola di provenienza

```
select count(*) as NumeroIscritti, Scuola, AA_ID as Anno
from iscrizione, studentean
where iscrizione.ID=studentean.ID
group by AA_ID, Scuola;
```

Numerolscritti	Scuola	Anno
1	Classico	2011
3	Istituto Tecnico Commerciale per Geometri	2014
2	Altra scuola italiana	2015
1	Classico	2015
1	Istituto Professionale Commerciale	2015
1	Istituto Professionale Industria e Artigianato	2015
1	Istituto Tecnico Commerciale per Geometri	2015
9	Istituto Tecnico Industriale	2015
2	liceo scientifico	2015
6	Scientifico	2015
26		2016
35	Altra scuola italiana	2016
4	Altra scuola straniera	2016
1	Athene Royal Crommelynck	2016
35	Classico	2016
1	i.i.s. volterra elia	2016
1	INDUSTRIALE	2016
1	Istituto d'Arte	2016
3	Istituto Magistrale	2016
1	Istituto Professionale Alberghiero	2016
6	Istituto Professionale Commerciale	2016
23	Istituto Professionale Industria e Artigianato	2016
3	Istituto tecnico	2016
13	Istituto Tecnico Agrario	2016
57	Istituto Tecnico Commerciale	2016
129	Istituto Tecnico Commerciale per Geometri	2016
227	Istituto Tecnico Industriale	2016
1	Istituto tecnico Nautico	2016
1	Istituto Tecnico Sperimentale	2016
1	ITET Enrico Fermi Lanciano	2016
9	Liceo Artistico	2016
6	liceo scientifico	2016
1	Liceo scientifico "A. Einstein"	2016
8	Pedagogico	2016
465	Scientifico	2016

1	Altra scuola italiana	2017
1	Altra scuola straniera	2017
1	Classico	2017
1	Istituto Professionale Industria e Artigianato	2017
3	Istituto Tecnico Commerciale per Geometri	2017
5	Istituto Tecnico Industriale	2017
2	Liceo Artistico	2017
9	Scientifico	2017
1	Istituto Professionale Industria e Artigianato	2018
2	Istituto Tecnico Industriale	2018
3	Scientifico	2018
4		2019
3	Altra scuola italiana	2019
1	Altra scuola straniera	2019
1	Athene Royal Crommelynck	2019
6	Classico	2019
2	Istituto Professionale Industria e Artigianato	2019
9	Istituto Tecnico Commerciale	2019
9	Istituto Tecnico Commerciale per Geometri	2019
19	Istituto Tecnico Industriale	2019
66	Scientifico	2019
4	Altra scuola italiana	2020
2	Istituto Professionale Industria e Artigianato	2020
2	Istituto Tecnico Commerciale	2020
12	Istituto Tecnico Commerciale per Geometri	2020
5	Istituto Tecnico Industriale	2020
20	Scientifico	2020

Questa query ci permette di conoscere gli effettivi iscritti per ogni tipologia di istituto per ogni anno, così da poter apprendere quali basi dovrebbero avere già consolidato nei cinque anni precedenti e capire quali contenuti andranno rafforzati e su quali si può effettuare un semplice ripasso.

Inoltre, dal confronto con la query precedente sulla tipologia di scuola di chi ha sostenuto il test diviso per anni, si può approfondire se ci sono particolari istituti i cui studenti sostengono il test e poi decidono di non iscriversi e analizzare le cause che potrebbero essere, ad esempio, di tipo geografico.

- Numero di iscrizioni per corso di studio

```
select count(*) as NumeroIscritti, CDS_DES as CorsoScelto
from iscrizione
group by CDS_DES;
```

NumeroIscritti	CorsoScelto
1	Architetture, processi e tecnologie Industry 4.0
21	BIOMEDICAL ENGINEERING
1	Computer Music Production
4	COMPUTER MUSIC PRODUCTION & SOUND DESIGN
1	ENVIRONMENTAL ENGINEERING
181	INGEGNERIA BIOMEDICA
10	INGEGNERIA CIVILE
67	INGEGNERIA CIVILE E AMBIENTALE
3	INGEGNERIA DEI MATERIALI COMPOSITI
80	INGEGNERIA EDILE
10	INGEGNERIA EDILE-ARCHITETTURA
98	INGEGNERIA ELETTRONICA
180	INGEGNERIA GESTIONALE
279	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
343	INGEGNERIA MECCANICA

Osserviamo che i corsi di laurea più scelti sono Ingegneria Informatica e dell'Automazione e Ingegneria Meccanica con un buon numero di iscritti anche in Ingegneria Elettronica, Gestionale e Biomedica.

- Numero di iscrizioni per corso di studio per anno

```
select count(*) as NumeroIscritti, AA_ID as Anno, CDS_DES as CorsoScelto
from iscrizione
group by AA_ID, CDS_DES;
```

NumeroIscritti	Anno	CorsoScelto
1	2011	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
1	2014	INGEGNERIA CIVILE E AMBIENTALE
2	2014	INGEGNERIA EDILE-ARCHITETTURA
5	2015	INGEGNERIA BIOMEDICA
4	2015	INGEGNERIA CIVILE E AMBIENTALE
2	2015	INGEGNERIA ELETTRONICA
5	2015	INGEGNERIA GESTIONALE
3	2015	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
4	2015	INGEGNERIA MECCANICA
1	2016	Computer Music Production
170	2016	INGEGNERIA BIOMEDICA
60	2016	INGEGNERIA CIVILE E AMBIENTALE
59	2016	INGEGNERIA EDILE
8	2016	INGEGNERIA EDILE-ARCHITETTURA
88	2016	INGEGNERIA ELETTRONICA
135	2016	INGEGNERIA GESTIONALE
249	2016	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
288	2016	INGEGNERIA MECCANICA
5	2017	INGEGNERIA BIOMEDICA
2	2017	INGEGNERIA CIVILE E AMBIENTALE
3	2017	INGEGNERIA EDILE
1	2017	INGEGNERIA ELETTRONICA
1	2017	INGEGNERIA GESTIONALE
5	2017	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
6	2017	INGEGNERIA MECCANICA
2	2018	COMPUTER MUSIC PRODUCTION & SOUND DESIGN
1	2018	INGEGNERIA ELETTRONICA
2	2018	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
1	2018	INGEGNERIA MECCANICA
1	2019	Architetture, processi e tecnologie Industry 4.0
10	2019	BIOMEDICAL ENGINEERING
2	2019	COMPUTER MUSIC PRODUCTION & SOUND DESIGN
1	2019	ENVIRONMENTAL ENGINEERING
1	2019	INGEGNERIA BIOMEDICA
8	2019	INGEGNERIA CIVILE

3	2019	INGEGNERIA DEI MATERIALI COMPOSITI
8	2019	INGEGNERIA EDILE
6	2019	INGEGNERIA ELETTRONICA
32	2019	INGEGNERIA GESTIONALE
16	2019	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
32	2019	INGEGNERIA MECCANICA
11	2020	BIOMEDICAL ENGINEERING
2	2020	INGEGNERIA CIVILE
10	2020	INGEGNERIA EDILE
7	2020	INGEGNERIA GESTIONALE
3	2020	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
12	2020	INGEGNERIA MECCANICA

Osserviamo che nel corso degli anni sono mantenuti gli andamenti dell'analisi precedente. Questa query è significativa per poter analizzare, con più dati a disposizione, una possibile diminuzione o un eventuale incremento nell'interesse verso un determinato corso e poterne studiare le cause (ad esempio un eventuale calo degli impieghi lavorativi o l'eliminazione di alcuni insegnamenti attraenti al suo interno).

- Numero di iscrizioni per corso di studio per scuola di provenienza

```
select count(*) as NumeroIscritti, CDS_DES as CorsoScelto, Scuola
from iscrizione, studentean
where iscrizione.ID=studentean.ID
group by CDS_DES, Scuola;
```

NumeroIscritti	CorsoScelto	Scuola
1	Architetture, processi e tecnologie Industry 4.0	Classico
4	BIOMEDICAL ENGINEERING	Istituto Tecnico Industriale
17	BIOMEDICAL ENGINEERING	Scientifico
1	Computer Music Production	Scientifico
2	COMPUTER MUSIC PRODUCTION & SOUND DESIGN	Istituto Tecnico Industriale
2	COMPUTER MUSIC PRODUCTION & SOUND DESIGN	Scientifico
1	ENVIRONMENTAL ENGINEERING	Classico
5	INGEGNERIA BIOMEDICA	
6	INGEGNERIA BIOMEDICA	Altra scuola italiana
13	INGEGNERIA BIOMEDICA	Classico
2	INGEGNERIA BIOMEDICA	Istituto Professionale Commerciale
2	INGEGNERIA BIOMEDICA	Istituto Professionale Industria e Artigianato
5	INGEGNERIA BIOMEDICA	Istituto Tecnico Agrario
4	INGEGNERIA BIOMEDICA	Istituto Tecnico Commerciale
2	INGEGNERIA BIOMEDICA	Istituto Tecnico Commerciale per Geometri
25	INGEGNERIA BIOMEDICA	Istituto Tecnico Industriale
1	INGEGNERIA BIOMEDICA	Istituto Tecnico Sperimentale
4	INGEGNERIA BIOMEDICA	Liceo scientifico
1	INGEGNERIA BIOMEDICA	Pedagogico
111	INGEGNERIA BIOMEDICA	Scientifico
8	INGEGNERIA CIVILE	Istituto Tecnico Commerciale per Geometri
2	INGEGNERIA CIVILE	Scientifico
1	INGEGNERIA CIVILE E AMBIENTALE	
1	INGEGNERIA CIVILE E AMBIENTALE	Altra scuola italiana
1	INGEGNERIA CIVILE E AMBIENTALE	Altra scuola straniera
3	INGEGNERIA CIVILE E AMBIENTALE	Istituto Tecnico Agrario
32	INGEGNERIA CIVILE E AMBIENTALE	Istituto Tecnico Commerciale per Geometri
2	INGEGNERIA CIVILE E AMBIENTALE	Istituto Tecnico Industriale
1	INGEGNERIA CIVILE E AMBIENTALE	ITET Enrico Fermi Lanciano
2	INGEGNERIA CIVILE E AMBIENTALE	Liceo Artistico
1	INGEGNERIA CIVILE E AMBIENTALE	Liceo Scientifico
1	INGEGNERIA CIVILE E AMBIENTALE	Pedagogico
22	INGEGNERIA CIVILE E AMBIENTALE	Scientifico
1	INGEGNERIA DEI MATERIALI COMPOSITI	
1	INGEGNERIA DEI MATERIALI COMPOSITI	Istituto Tecnico Industriale
1	INGEGNERIA DEI MATERIALI COMPOSITI	Scientifico
3	INGEGNERIA EDILE	Altra scuola italiana
2	INGEGNERIA EDILE	Classico
2	INGEGNERIA EDILE	Istituto Professionale Industria e Artigianato

2	INGEGNERIA EDILE	Istituto Tecnico Agrario
50	INGEGNERIA EDILE	Istituto Tecnico Commerciale per Geometri
2	INGEGNERIA EDILE	Istituto Tecnico Industriale
6	INGEGNERIA EDILE	Liceo Artistico
13	INGEGNERIA EDILE	Scientifico
1	INGEGNERIA EDILE-ARCHITETTURA	Classico
5	INGEGNERIA EDILE-ARCHITETTURA	Istituto Tecnico Commerciale per Geometri
4	INGEGNERIA EDILE-ARCHITETTURA	Scientifico
8	INGEGNERIA ELETTRONICA	
1	INGEGNERIA ELETTRONICA	Altra scuola italiana
1	INGEGNERIA ELETTRONICA	Altra scuola straniera
1	INGEGNERIA ELETTRONICA	Classico
1	INGEGNERIA ELETTRONICA	Istituto d'Arte
1	INGEGNERIA ELETTRONICA	Istituto Professionale Alberghiero
1	INGEGNERIA ELETTRONICA	Istituto Professionale Commerciale
1	INGEGNERIA ELETTRONICA	Istituto Professionale Industria e Artigianato
3	INGEGNERIA ELETTRONICA	Istituto Tecnico Commerciale
44	INGEGNERIA ELETTRONICA	Istituto Tecnico Industriale
36	INGEGNERIA ELETTRONICA	Scientifico
2	INGEGNERIA GESTIONALE	
10	INGEGNERIA GESTIONALE	Altra scuola italiana
2	INGEGNERIA GESTIONALE	Altra scuola straniera
2	INGEGNERIA GESTIONALE	Athene Royal Crommelynck
9	INGEGNERIA GESTIONALE	Classico
2	INGEGNERIA GESTIONALE	Istituto Professionale Commerciale
4	INGEGNERIA GESTIONALE	Istituto Professionale Industria e Artigianato
30	INGEGNERIA GESTIONALE	Istituto Tecnico Commerciale
15	INGEGNERIA GESTIONALE	Istituto Tecnico Commerciale per Geometri
11	INGEGNERIA GESTIONALE	Istituto Tecnico Industriale
1	INGEGNERIA GESTIONALE	Liceo scientifico "A. Einstein"
3	INGEGNERIA GESTIONALE	Pedagogico
89	INGEGNERIA GESTIONALE	Scientifico
2	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	
4	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Altra scuola italiana
9	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Classico
3	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Istituto Magistrale
1	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Istituto Professionale Commerciale

11	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Istituto Professionale Industria e Artigianato
1	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Istituto Tecnico
1	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Istituto Tecnico Agrario
16	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Istituto Tecnico Commerciale
20	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Istituto Tecnico Commerciale per Geometri
92	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Istituto Tecnico Industriale
2	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Liceo Artistico
3	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Pedagogico
114	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	Scientifico
11	INGEGNERIA MECCANICA	
20	INGEGNERIA MECCANICA	Altra scuola italiana
2	INGEGNERIA MECCANICA	Altra scuola straniera
7	INGEGNERIA MECCANICA	Classico
1	INGEGNERIA MECCANICA	I.i.s. volterra elia
1	INGEGNERIA MECCANICA	INDUSTRIALE
1	INGEGNERIA MECCANICA	Istituto Professionale Commerciale
10	INGEGNERIA MECCANICA	Istituto Professionale Industria e Artigianato
2	INGEGNERIA MECCANICA	Istituto tecnico
2	INGEGNERIA MECCANICA	Istituto Tecnico Agrario
15	INGEGNERIA MECCANICA	Istituto Tecnico Commerciale
25	INGEGNERIA MECCANICA	Istituto Tecnico Commerciale per Geometri
84	INGEGNERIA MECCANICA	Istituto Tecnico Industriale
1	INGEGNERIA MECCANICA	Istituto tecnico Nautico
1	INGEGNERIA MECCANICA	Liceo Artistico
157	INGEGNERIA MECCANICA	Scientifico

Questa query consente di analizzare da che istituti provengono la maggior parte degli studenti che si iscrivono ad un determinato corso di laurea. Ad esempio osserviamo una maggioranza di studenti provenienti dal Liceo Scientifico per i corsi di Ingegneria Biomedica, Gestionale, Informatica e dell'Automazione e Meccanica, una maggioranza di studenti provenienti dall'Istituto Tecnico Commerciale per Geometri per i corsi di Ingegneria Civile e Ambientale, Edile e Edile-Architettura e una maggioranza di studenti provenienti dall'Istituto Tecnico Industriale per i corsi di Ingegneria Elettronica.

- Numero di studenti maschi che hanno sostenuto il test

```
select count(substring(CodiceFiscale, 10, 2)) as M
from studente
where substring(CodiceFiscale, 10, 2)<31;
```

- Numero di studenti maschi che hanno sostenuto il test ogni anno

```
select count(substring(CodiceFiscale, 10, 2)) as M, year(data) as Anno
from studente
where substring(CodiceFiscale, 10, 2)<31
group by year(data);
```

- Numero di studenti maschi iscritti in ogni corso

```
select count(substring(CodiceFiscale, 10, 2)) as M, CDS_DES as Corso
from studente, iscrizione
where studente.ID=iscrizione.ID and substring(CodiceFiscale, 10, 2)<31
group by CDS_DES;
```

- Numero di studenti maschi iscritti in ogni corso in ogni anno

```
select count(substring(CodiceFiscale, 10, 2)) as M, CDS_DES as Corso,
       year(data) as Anno
from studente, iscrizione
where studente.ID=iscrizione.ID and substring(CodiceFiscale, 10, 2)<31
group by CDS_DES, year(data);
```

- Numero di studentesse femmine che hanno sostenuto il test

```
select count(substring(CodiceFiscale, 10, 2)) as W
from studente
where substring(CodiceFiscale, 10, 2)>31;
```

- Numero di studentesse femmine che hanno sostenuto il test ogni anno

```
select count(substring(CodiceFiscale, 10, 2)) as W, year(data) as Anno
from studente
where substring(CodiceFiscale, 10, 2)>31
group by year(data);
```

- Numero di studentesse femmine iscritte in ogni corso

```
select count(substring(CodiceFiscale, 10, 2)) as M, CDS_DES as Corso
from studente, iscrizione
where studente.ID=iscrizione.ID and substring(CodiceFiscale, 10, 2)>31
group by CDS_DES;
```

- Numero di studentesse femmine iscritte in ogni corso in ogni anno

```
select count(substring(CodiceFiscale, 10, 2)) as M, CDS_DES as Corso,
       year(data) as Anno
from studente, iscrizione
where studente.ID=iscrizione.ID and substring(CodiceFiscale, 10, 2)>31
group by CDS_DES, year(data);
```

N.B. Gli screenshots dei risultati delle ultime otto analisi non sono mostrati in quanto le queries non possono essere implementate direttamente dal tesista poiché operano su dati sensibili.

Queste queries sono utili per un fine puramente statistico per osservare la distribuzione di genere negli anni e tra i vari corsi.

Conclusioni

In questa tesi è stata progettata la base di dati necessaria a contenere i dati da studiare riguardanti i test d'ingresso per l'ottenimento degli Obblighi Formativi Aggiuntivi (OFA) sostenuti presso la facoltà di Ingegneria dell'Università Politecnica delle Marche attraverso le fasi concettuale e logica eseguite dopo un'attenta lettura e comprensione degli elementi iniziali.

L'importanza di questa applicazione risiede nel fatto che permette una più semplice e immediata analisi delle informazioni tramite codice SQL al fine di ottenere significativi risultati statistici e numerici riguardanti gli studenti e i risultati da loro ottenuti.

Si è evidenziato quali sono le scuole da cui provengono maggiormente gli alunni e la correlazione con i risultati ottenuti, l'andamento dei voti nel corso degli appelli e degli anni e la loro suddivisione per tipologia di quiz, categoria e singola domanda.

Si è inoltre approfondito quali sono gli appelli in cui gli studenti preferiscono sostenere il test e spiegato le probabili cause, qual è il rapporto tra chi sostiene il test e chi poi si iscrive effettivamente, quali corsi sono scelti principalmente dai ragazzi provenienti da determinati istituti secondari di II grado e quali sono le percentuali di studenti e studentesse che vogliono iscriversi alla facoltà e, in particolare, a quali corsi, sia in generale che nel corso degli anni.

Lo scopo dello sviluppo di questo database prescinde inoltre dal puro ambito tescivo e dagli esiti a cui si è giunti attualmente e si prefissa di essere uno strumento utile, a disposizione dell'Università, per poter inserire ed analizzare nuovi dati provenienti da diversi anni col passare del tempo e svolgere ulteriori tipi di analisi, oltre a quelle già implementate, e ricavare modelli, grafici e conclusioni rilevanti per un eventuale futuro sviluppo dell'Ateneo, usando ad esempio tools di Business Intelligence.

Bibliografia

- [1] S.Abiteboul, R.Hull, V.Vianu, *Foundations of databases*, Addison-Wesley, Reading, Massachussets, 1995.
- [2] A.Albano, *Basi di dati. Strutture e algoritmi*, Addison-Wesley italia, Milano, 1992.
- [3] P.Atzeni, S.Ceri, S.Paraboschi, R.Torlone, *Basi di dati. Modelli e linguaggi di interrogazione*, McGraw-Hill, Milano, 2006.
- [4] B.Thalheim, *Entity-Relationship Modeling: Foundations of Database Technology*, Springer, New York, 2013.
- [5] P.N.Tan, M.Steinbach, V.Kumar, *Introduction to Data Mining*, Pearson, Boston, 2005.

Ringraziamenti

Il mio primo ringraziamento va alla prof.ssa Claudia Diamantini per avermi dato la possibilità di redigere una tesi sullo sviluppo di un database e sull'analisi dei dati, uno degli argomenti che più mi ha entusiasmato in questo percorso triennale, per la preziosa assistenza nella stesura dell'elaborato e per gli insegnamenti ricevuti da lei durante il corso di laurea.

Un grazie particolare è riservato al dott. Alex Mircoli, che mi ha pazientemente seguito durante tutta l'attività sperimentale e di implementazione con grande disponibilità e professionalità, aiutandomi nell'uso degli strumenti necessari all'analisi dei dati.

Infine un ringraziamento speciale va ai miei genitori Francesco e Anna che mi hanno sempre sostenuto durante il percorso di studi e a mio fratello maggiore Andrea, che da sempre è per me fonte di grande ispirazione.