



UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA

Corso di Laurea Triennale in Ingegneria Informatica e
dell'Automazione

**Strumenti di Business Intelligence per
l'Analisi dei Test di Ingresso a
Corsi Universitari**

**Business Intelligence Tools for
the Analysis of Entrance Tests to
University Courses**

Relatore:
**Prof.ssa Claudia
Diamantini**

Tesi di Laurea di:
Lorenzo Iacopini

Correlatore:
Dott. Alex Mircoli

A.A 2019/2020

Sommario

Introduzione	2
CAPITOLO 1: Dati da Analizzare ed Obiettivi	3
1.1 Descrizione DB	3
1.2 Analisi Possibili sui Dati	7
CAPITOLO 2: Strumenti Utilizzati ed Analisi dei Dati.....	9
2.1 Descrizione POWER BI	9
2.2 Connettersi ai Dati	10
2.3 Trasformare e Pulire i Dati e Creare un Modello	12
2.4 Creare Oggetti Visivi.....	13
2.5 Creare un Report.....	14
CAPITOLO 3: Conclusioni e Sviluppi Futuri.....	28
3.1 Conclusioni	28
3.2 Sviluppi Futuri.....	28
Bibliografia	31

Introduzione

Alla base di questo studio vi è l'analisi dei Test di ingresso alla nostra università svolti nell'anno accademico 2016-2017, essi svolgono una verifica preliminare delle conoscenze e competenze degli studenti che hanno intenzione di proseguire il loro percorso formativo successivamente alla scuola secondaria di secondo grado.

I quiz vengono salvati su un software esterno e non vi è la possibilità di analizzarli per andar a ottenere informazioni utili sugli studenti. A tal proposito, l'obiettivo di questa tesi di laurea è quello di fornire tutte le potenziali considerazioni su questi dati e di sviluppare un insieme di strumenti grafici per effettuare un'analisi descrittiva di quest'ultimi. L'elaborato, in questo modo, fornisce delle informazioni utili riguardo gli studenti che affrontano il test di ingresso.

Il resto del lavoro si articola nei seguenti capitoli:

- Nel capitolo 1 verranno presentati i dati su cui sono basate le analisi e verranno illustrate che tipo di conoscenza da questi verrà estratta.
- Nel capitolo 2 verrà descritto in primis Power BI, software che si è rivelato fondamentale per il conseguimento degli obiettivi imposti, poi verrà illustrato il lavoro partendo da come i dati sono stati prelevati fino ad arrivare alla "Data Analysis".
- Nel capitolo 3 verranno mostrate le conclusioni del lavoro svolto e verranno presi in considerazione eventuali sviluppi futuri.

CAPITOLO 1: Dati da Analizzare ed Obiettivi

In questo capitolo vengono presentati sia i dati riguardanti i test di ingresso su cui saranno basate le analisi sia le domande su di essi a cui sarà interessante trovare risposta.

Il test è composto da 30 domande di diversa tipologia: Logica, Geometria, Algebra, Funzioni, Grafici, Fisica e Comprensione del Testo.

Per ogni risposta corretta viene attribuito un punteggio di 0.33, per ogni risposta non data un punteggio di 0.08 e per ogni risposta errata 0. Il punteggio complessivo ottenuto nel test è dato dalla somma dei punteggi delle singole domande.

Il test può appartenere a due categorie: A e B. La differenza fra i due non sta nella tipologia di domande ma semplicemente nell'ordine con cui esse vengono proposte al candidato, ciò per non rendere possibile un confronto immediato fra gli studenti.

1.1 Descrizione DB

Un sistema di gestione di base di dati (in inglese Data Base Management System abbreviato con DBMS) è un sistema software in grado di gestire collezioni di dati che siano grandi, condivise e persistenti, assicurando la loro affidabilità e privatezza e garantendo al tempo stesso efficienza ed efficacia. Un database è una collezione di dati gestita da un DMBS che viene utilizzata per rappresentare le informazioni di interesse per un sistema informatico. [1]

Il modello relazionale dei dati, che è il più diffuso, si basa sui concetti di

relazione e di tabella e consente di organizzare i dati in insiemi di record a struttura fissa. In un database relazionale le righe della tabella rappresentano i record, i quali hanno un ID univoco chiamato *chiave*, mentre le colonne contengono gli attributi dei dati; l'ordine delle righe e delle colonne è sostanzialmente irrilevante.

L'analisi dei dati è partita da un database contenente ben 13 tabelle:

- Associazione;
- Categoria;
- Domanda;
- Iscrizione;
- Studente;
- Studenteam;
- Temp_distribuzione;
- Temp_quiz_a;
- Temp_quiz_b;
- Temp_studente;
- Temp_studente2;
- X;

Essa però si è focalizzata maggiormente solo su una porzione di esse, in particolare su quattro di esse: *Associazione*, *Categoria*, *Iscrizione*, *Studenteam*.

Associazione: in questa tabella vengono registrate tutte le votazioni ottenute dai singoli studenti nelle specifiche domande.

#	Nome	Tipo	Codifica caratteri	Attributi	Null	Predefinito	Commenti	Extra	Azione
<input type="checkbox"/>	1 ID_Test	varchar(22)	latin1_swedish_ci		No	Nessuno			Modifica Elimina Più
<input type="checkbox"/>	2 ID_Categoria	varchar(10)	latin1_swedish_ci		No	Nessuno			Modifica Elimina Più
<input type="checkbox"/>	3 Voto	float			No	Nessuno			Modifica Elimina Più
<input type="checkbox"/>	4 categoriaQuiz	char(1)	latin1_swedish_ci		No	Nessuno			Modifica Elimina Più

ID_TEST: ad ogni singolo studente è stato attribuito un codice identificativo formato dalla data in cui è stato sostenuto il test seguito da un numero casuale, in tal modo ogni domanda risposta dal singolo studente avrà lo stesso “ID_TEST”.

ID_CATEGORIA: indica la tipologia di domanda.

VOTO: indica il punteggio ottenuto nella specifica domanda.

categoriaQuiz: indica a quale categoria appartiene la singola domanda.

Categoria: in questa tabella vengono presentate le differenti tipologie di domande presenti nel test.

#	Nome	Tipo	Codifica caratteri	Attributi	Null	Predefinito	Commenti	Extra	Azione
<input type="checkbox"/>	1 ID_Categoria	varchar(10)	latin1_swedish_ci		No	Nessuno			Modifica Elimina Più
<input type="checkbox"/>	2 Descrizione	varchar(70)	latin1_swedish_ci		Si	NULL			Modifica Elimina Più

ID_Categoria: codice identificativo per la tipologia di domanda.

Descrizione: descrizione sintetica per meglio comprendere la domanda.

ID_Categoria	Descrizione
01-LOG	sillogismi
02-LOG	negazione
03-LOG	implicazioni logiche
04-LOG	problema
05-GEO	piano cartesiano e coordinate cartesiane di un pun...
06-GEO	distanza, aree e volumi
07-GEO	rette nel piano
08-GEO	circonferenze, parabole, iperboli
09-ALG	polinomi
10-ALG	numeri irrazionali, razionali, interi
11-FUN	funzioni composte e funzione inversa
12-FUN	proprietà delle potenze
13-FUN	proprietà del modulo, equazioni e disequazioni con...
14-FUN	proprietà dei logaritmi
15-FUN	angoli e trigonometria
16-FUN	equazioni e disequazioni algebriche intere
17-FUN	equazioni e disequazioni algebriche fratte
18-FUN	equazioni e disequazioni irrazionali
19-GRA	operazioni sui grafici
20-GRA	grafici funzioni pari, dispari, periodiche, confro...
21-FIS	fisica
22-FIS	fisica
23-FIS	fisica
24-FIS	fisica
25-FIS	fisica
26-TES	comprensione del testo
27-TES	comprensione del testo
28-TES	comprensione del testo
29-TES	comprensione del testo
30-TES	comprensione del testo

Iscrizione: in questa tabella vengono registrate le iscrizioni all'università degli studenti che precedentemente avevano preso parte al test.

#	Nome	Tipo	Codifica caratteri	Attributi	Null	Predefinito	Commenti	Extra	Azione
<input type="checkbox"/>	1	ID	varchar(30)	latin1_swedish_ci	No	Nessuno			Modifica Elimina Più
<input type="checkbox"/>	2	CDS_ID	int(11)		Sì	NULL			Modifica Elimina Più
<input type="checkbox"/>	3	CDS_DES	text	latin1_swedish_ci	Sì	NULL			Modifica Elimina Più
<input type="checkbox"/>	4	AA_ID	int(11)		Sì	NULL			Modifica Elimina Più

ID: codice identificativo del singolo studente.

CDS_ID: codice identificativo relativo ad uno specifico Corso di Laurea

CDS_DES: nome del Corso di Laurea.

AA_ID: anno in cui è avvenuto l'immatricolazione al Corso di Laurea.

Studenteanm: in questa tabella vengono registrate le informazioni relative allo studente che partecipa al test.

#	Nome	Tipo	Codifica caratteri	Attributi	Null	Predefinito	Commenti	Extra	Azione
<input type="checkbox"/>	1	Username	varchar(15)	latin1_swedish_ci	No				Modifica Elimina Più
<input type="checkbox"/>	2	Scuola	varchar(50)	latin1_swedish_ci	Sì	NULL			Modifica Elimina Più
<input type="checkbox"/>	3	data	date		No	Nessuno			Modifica Elimina Più
<input type="checkbox"/>	4	ID	varchar(30)	latin1_swedish_ci	No	Nessuno			Modifica Elimina Più

Username: nome che viene associato al singolo studente nel momento in cui prende parte al test.

Scuola: indica la scuola di provenienza dello studente.

Data: indica la data in cui lo studente si è sottoposto al test.

ID: codice identificativo del singolo studente.

1.2 Analisi Possibili sui Dati

Successivamente ad un'accurata analisi ed un approfondito studio dei dati contenuti nelle sopracitate tabelle, sarà possibile rispondere alle seguenti domande:

1. Quanti studenti partecipano al test?
2. Da quali scuole provengono gli studenti che partecipano al test?
3. In quale anno si sono sottoposti al Test?
4. In quale periodo dell'anno hanno svolto il test?
5. In quale periodo e in quale anno hanno svolto il test gli studenti di una particolare scuola?
6. Come è strutturato il test?
7. In quale tipologia di domanda il risultato medio è stato più alto? In quale il più basso?
8. Qual è il risultato medio del test per scuola?
9. Qual è il risultato medio del test in una particolare data?
10. Qual è il risultato medio del test per scuola in una particolare data?

11. Qual è il risultato medio di una scuola nelle varie date in cui si è svolto il test?
12. Quale scuola ha la media più alta? Quale la più bassa?
13. Qual è il trend del voto medio negli anni?
14. Qual è la percentuale di persone che si è iscritta all'università dopo aver sostenuto il test?
15. Qual è il numero di iscritti per facoltà complessivi?
16. Quanti studenti si sono iscritti in un particolare anno? E in quale facoltà?
17. Quando gli iscritti di un determinato corso si sono immatricolati?
18. Qual è la media al test degli iscritti all'università?
19. Qual è il risultato medio di coloro che si sono iscritti ad un particolare corso di laurea?
20. Da quali scuole provengono gli iscritti?
21. Da quali scuole provengono gli iscritti di un particolare corso di laurea?
22. Quali sono i risultati al test dei singoli iscritti ad un particolare corso di laurea?
23. Quali sono i risultati dei singoli iscritti ad un particolare corso di laurea che hanno ottenuto un punteggio più alto della media?

CAPITOLO 2: Strumenti Utilizzati ed Analisi dei Dati

2.1 Descrizione POWER BI

Power BI [2] è un servizio d'analisi aziendale prodotto da Microsoft. Fornisce visualizzazioni interattive e funzionalità di business intelligence con un'interfaccia semplice e intuitiva, per consentire agli utenti di creare report e dashboard visualizzando i dati da analizzare in tutti i modi possibili, ad esempio attraverso grafici a barre, grafici ad anello e grafici combinati. Da sottolineare il fatto che Power BI possa essere utilizzato in maniera differente a seconda dello scopo prefissato o a secondo del ruolo ricoperto in un progetto. Per esempio, può essere utilizzato per andar a creare dei report ex novo, per visualizzare dei report o dashboard create da altri utenti oppure mediante l'app per telefoni fare entrambe le cose dovunque ci si trovi.

Power BI Desktop è un'applicazione gratuita che deve essere installata nel proprio computer e consente di connettersi ai dati per modellarli e analizzarli nel modo che si ritiene opportuno. Utilizzando tale applicazione è possibile connettersi a molte origini diverse di dati e combinarle fra loro ottenendo così un modello di dati che consente di creare oggetti visivi e raccolte di oggetti visivi. Power BI Desktop viene utilizzato solitamente da coloro che lavorano a progetti di business intelligence per creare report, i quali vengono poi condivisi con il servizio *Power BI*. *Power BI Desktop* solitamente viene utilizzato per:

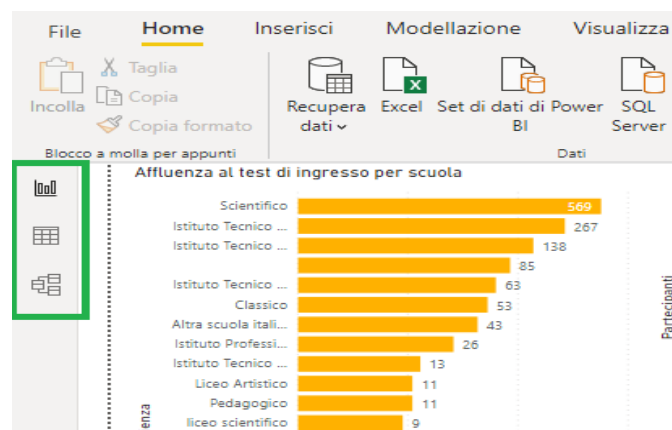
- Connettersi ai dati;

- Trasformare e pulire i dati per elaborare un nuovo modello;
- Creare oggetti visivi per fornire un'analisi descrittiva dei dati;
- Creare report che raccolgono gli oggetti visivi di una o più pagine;
- Condividere i report tramite il servizio Power BI.

In Power BI Desktop sul lato sinistro dell'area di disegno sono presenti tre viste. Le viste, nell'ordine in cui sono visualizzate, sono:

- Report: in questa vista si creano report e oggetti visivi.
- Dati: in questa vista è possibile visualizzare le tabelle ed altri dati usati nel modello di dati associato al report, inoltre è possibile trasformare i dati.
- Modello: in questa vista vengono visualizzate e gestite le relazioni tra le tabelle nel modello dei dati.

Nella figura seguente vengono evidenziate le tre viste sopracitate.

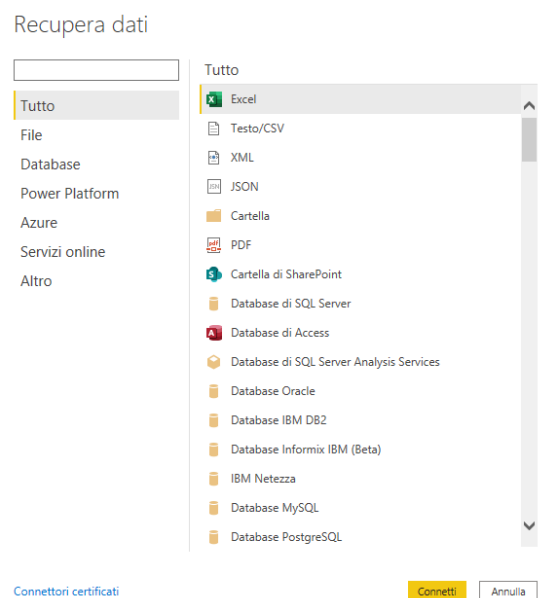


2.2 Connettersi ai Dati

Per iniziare a usare Power BI Desktop, il primo passaggio consiste nel connettersi ai dati. È possibile connettersi a molte origini dati diverse da Power BI Desktop.

Per connettersi ai dati:

1. Nella scheda Home della barra multifunzione cliccare su Recupera dati >Altro. Viene visualizzata la finestra Recupera dati, in cui sono elencate le diverse categorie da cui Power BI Desktop può attingere i dati.



2. nel momento in cui viene selezionata una sorgente dati vengono richieste alcune informazioni necessarie per la connessione. I dati utilizzati, in questo caso, sono stati importati da un Database MySQL.

Database MySQL

Server

Database

Opzioni avanzate

OK

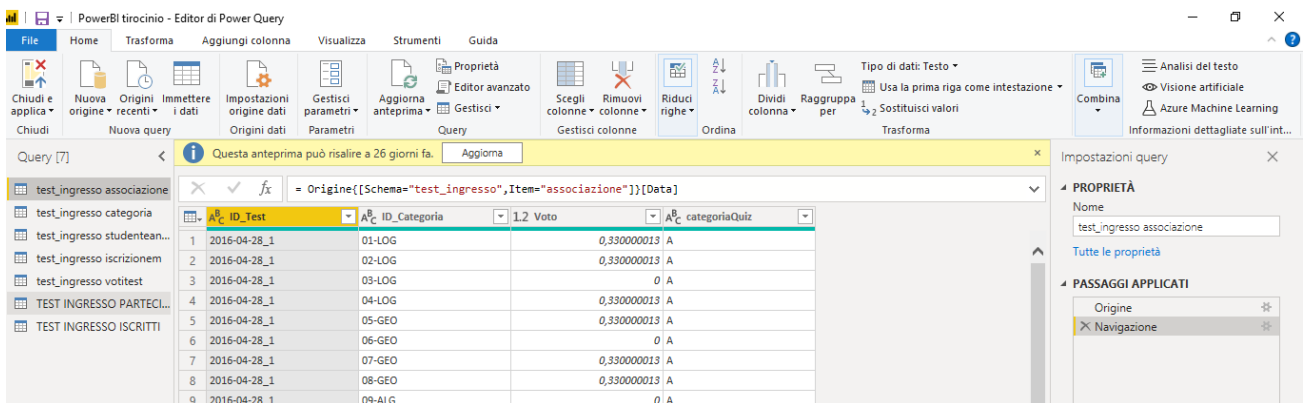
Annulla

3. Dopo essersi connessi a una o più origini dati, è possibile trasformare i dati in modo che siano utili per l'utente.

2.3 Trasformare e Pulire i Dati e Creare un Modello

Una funzionalità molto interessante offerta da Power BI Desktop è quella che consente di pulire e trasformare i dati usando l'editor predefinito Power Query. L'editor di Power Query consente di modificare i dati, ad esempio modificando un tipo di dato, rimuovendo colonne o combinando i dati da più origini. È possibile continuare a modificare i dati fino a quando essi non ottengano esattamente la forma desiderata. Per avviare l'editor di Power Query:

- Cliccare su **Trasforma Dati** > **Trasforma Dati** nella scheda Home della barra multifunzione. Verrà visualizzata la seguente finestra.

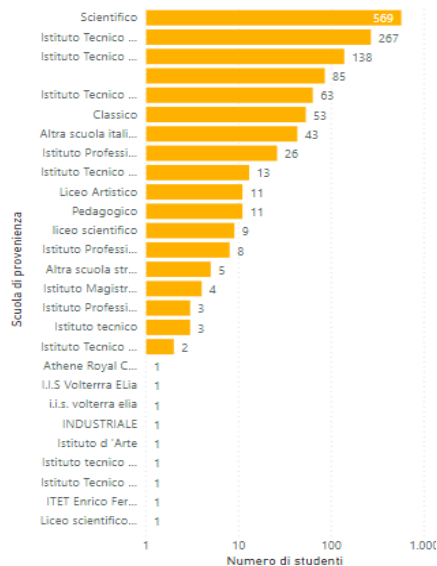


Ogni passaggio effettuato durante la trasformazione dei dati (come ad esempio rinominare una tabella, trasformare un tipo di dati o eliminare una colonna) viene registrato dall'editor di Power Query. Ogni volta che la query si connette all'origine dati, vengono eseguiti questi passaggi in modo che i dati abbiano sempre la forma specificata.

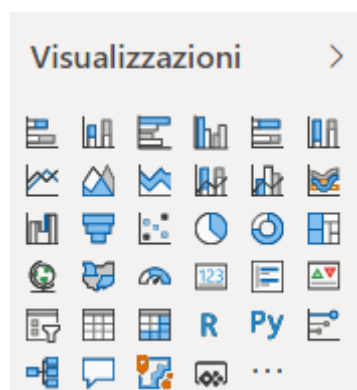
Nel momento in cui i dati raggiungono l'aspetto desiderato, è possibile creare gli oggetti visivi.

2.4 Creare Oggetti Visivi

Successivamente ad aver creato un modello di dati che contenga tutte le informazioni utili per procedere all'analisi, è possibile trascinare i *campi* di tale modello nell'area di disegno del report per creare un nuovo *oggetto visivo*. Un oggetto visivo è una rappresentazione grafica dei dati ed in Power BI ne esistono di molti tipi differenti. Ad esempio, in figura è mostrato un istogramma.



Per creare o modificare un oggetto visivo è necessario selezionare la tipologia di quest'ultimo nel riquadro **Visualizzazioni**.



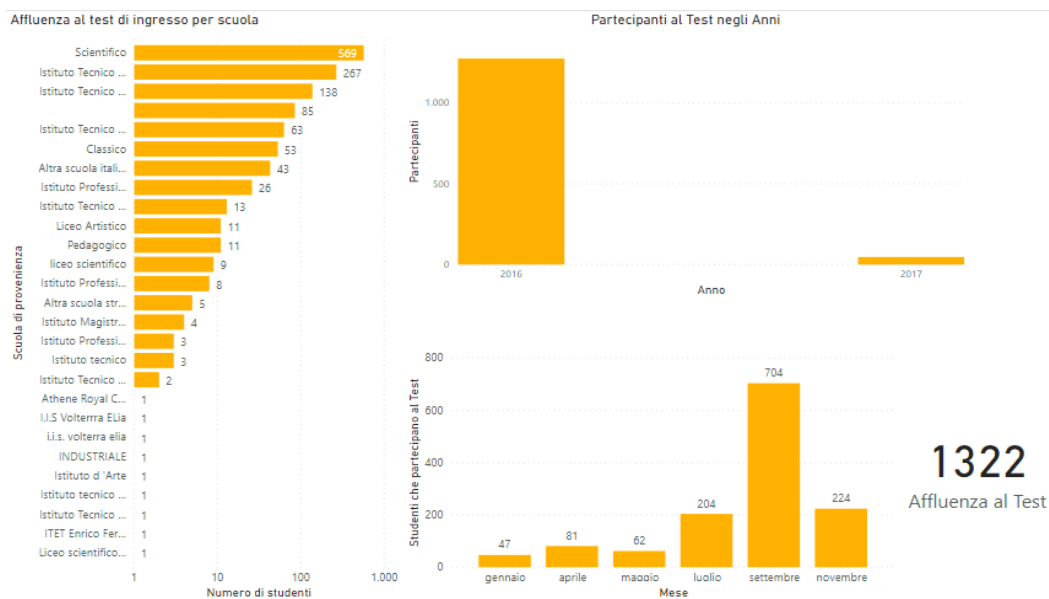
Se nell'area di disegno del report è già stato selezionato un oggetto visivo esso subirà la modifica nel tipo selezionato, in caso contrario ne verrà creato uno nuovo.

2.5 Creare un Report

Molto spesso è necessario creare una raccolta di oggetti visivi che mette in risalto i vari aspetti dei dati usati per creare il modello in Power BI Desktop. Un insieme di oggetti visivi, in un file di Power BI Desktop, prende il nome di *report*. Un report può avere una o più pagine, proprio come un file di Excel può avere uno o più fogli di lavoro. Con Power BI Desktop è possibile creare report complessi, visivamente dettagliati ed interattivi utilizzando dati provenienti da più sorgenti.

Il report creato per analizzare i dati riguardanti i test di ingresso all'università è formato da sette pagine, ognuna delle quali fornisce risposta agli obiettivi preposti.

Nella figura seguente viene visualizzata la prima pagina del report creato su Power BI Desktop denominata “**Partecipanti al Test**”, la quale risponde alle domande numero 1,2,3,4,5 presenti nel paragrafo 1.1.



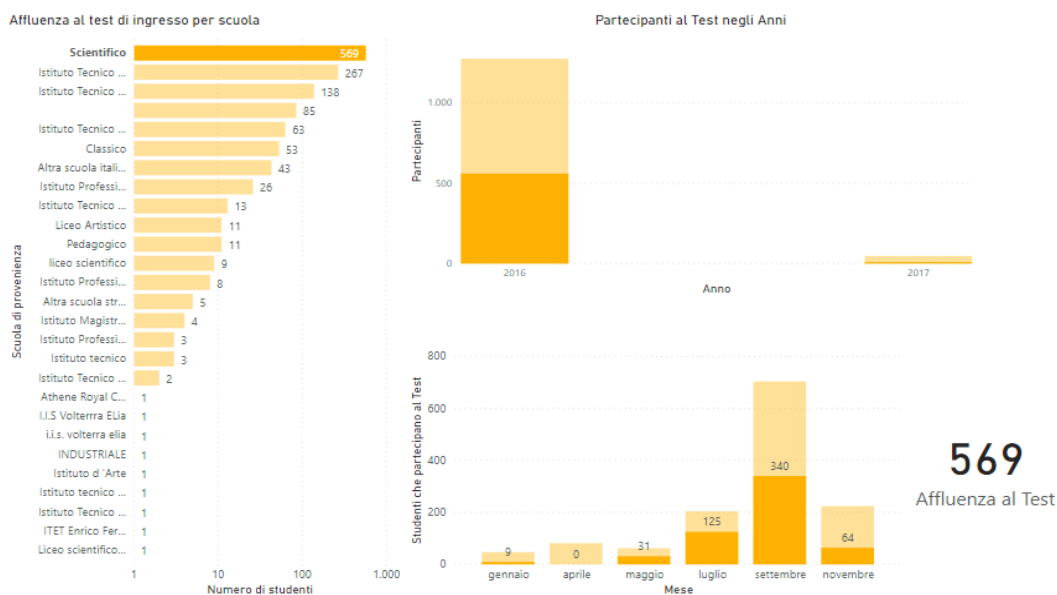
Al test nell'anno accademico 2016/2017 hanno partecipato 1322 studenti, gran parte di essi, però, ha sostenuto il test nell'anno 2016 come si può evincere dal grafico in alto a destra. A guidare la classifica della scuola di provenienza abbiamo, come ci si aspettava, il *Liceo Scientifico* il quale conta quasi la metà dei partecipanti. A seguire abbiamo tre istituti tecnici (*Industriale*, *Commerciale per Geometri e Commerciale*), da segnalare comunque una consistente affluenza dal *Liceo Classico*. Decisamente scarsa la partecipazione di studenti non provenienti dalle scuole sopracitate.

In alto a destra si può osservare un istogramma a colonne raggruppate che indica il numero di studenti che hanno partecipato al test di anno in anno, esso, in seguito a sviluppi futuri con l'inserimento di dati non solo riguardanti l'anno accademico 2016-2017, acquisirà maggior valore.

Il test nella maggioranza dei casi viene svolto una volta terminato la scuola secondaria di secondo grado, quindi a *luglio* e maggiormente a *settembre*. Interessante inoltre una significativa partecipazione (addirittura maggiore

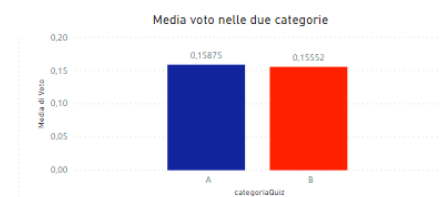
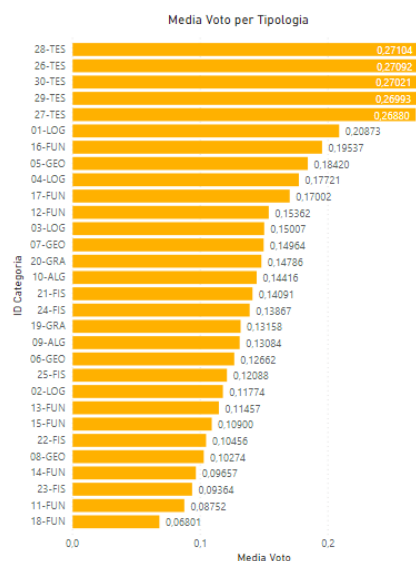
rispetto a quella di *luglio*) al test di *novembre*, probabilmente svolto da studenti con idee ben chiare che si apprestano nell'anno scolastico a terminare le superiori e da studenti in ritardo con il test di ammissione all'università.

Da sottolineare la funzionalità attraverso la quale andando a selezionare nell'istogramma sulla destra una scuola, sarà possibile visualizzare l'affluenza al test degli studenti provenienti da quest'ultima nei mesi e negli anni. Ad esempio, selezionando il "Liceo Scientifico" si può notare che di 569 studenti complessivi, 340 studenti che hanno partecipato al test nel mese di settembre, 125 nel mese di luglio, ecc...Inoltre è palese una maggiore affluenza nell'anno 2016 a discapito del 2017. Tali considerazioni fanno riferimento alla figura sottostante.

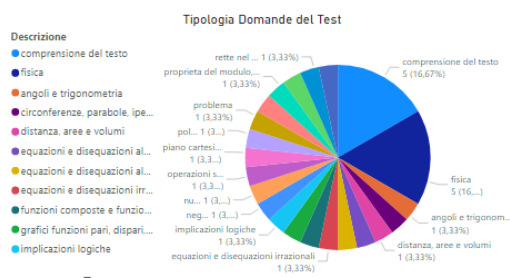


Nella figura seguente viene visualizzata la seconda pagina del report creato su Power BI Desktop denominata **“Medie Test per Categoria e**

Tipologia”, la quale risponde alle domande numero 6 e 7 presenti nel paragrafo 1.1.



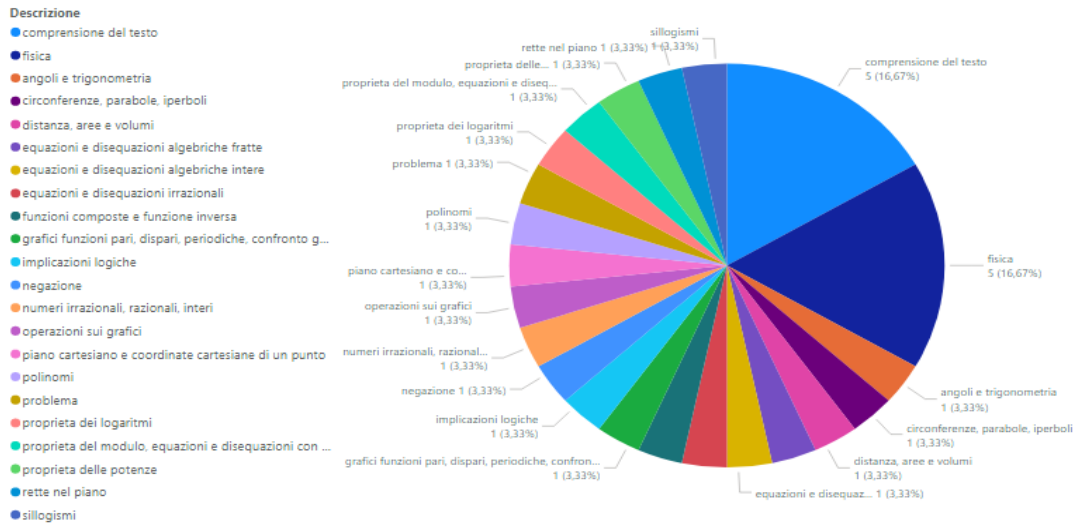
Il test può essere di due categorie: Categoria A e Categoria B. Essi sono composti da domande di diverse tipologie



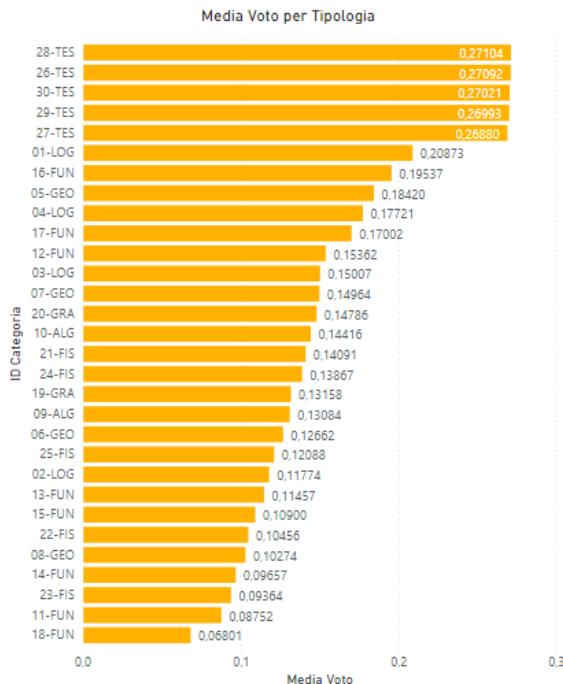
Il test può appartenere a due categorie: Categoria A e Categoria B. La differenza fra i due non sta nella tipologia di domande ma semplicemente nell'ordine con cui esse vengono proposte al candidato, ciò per rendere impossibile un confronto immediato fra gli studenti. Poiché i test non differiscono di composizione è ragionevole pensare che il risultato medio della singola domanda del test di tipologia A e il test di tipologia B non differiscano di molto, infatti ciò è confermato dal grafico che si trova in alto a destra nella pagina.

Il test è composto da 30 domande di diversa tipologia: Logica, Geometria, Algebra, Funzioni, Grafici, Fisica e Comprensione del Testo.

Tipologia Domande del Test



Per ogni risposta corretta viene attribuito un punteggio di 0.33, per ogni risposta non data un punteggio di 0.08 e per ogni risposta errata 0 punti. Come ci si aspettava le domande che risultano più semplici per gli studenti sono quelle di Comprensione del Testo le quali hanno un punteggio medio di circa 0,27, mentre quelle che risultano più ostiche sono quelli riguardanti le Funzione e Fisica le quali presentano un punteggio medio drasticamente inferiore il quale non arriva nemmeno a 0.10.



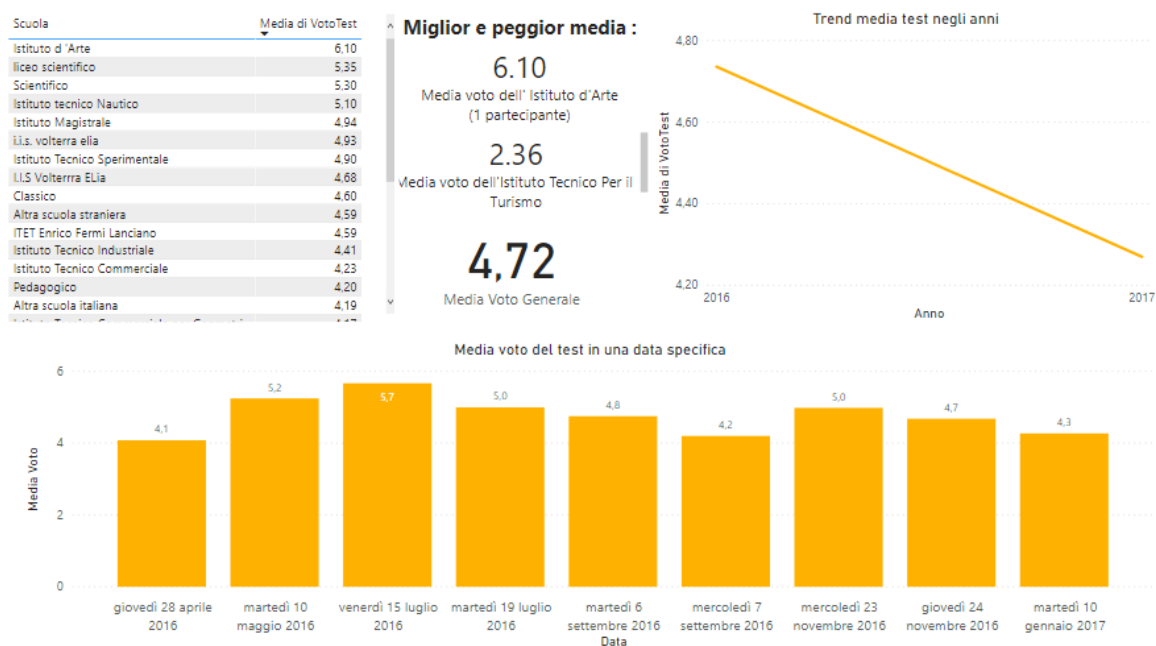
01-LOG	sillogismi
02-LOG	negazione
03-LOG	implicazioni logiche
04-LOG	problema
05-GEO	piano cartesiano e coordinate cartesiane di un pun...
06-GEO	distanza, aree e volumi
07-GEO	rette nel piano
08-GEO	circonferenze, parabole, iperboli
09-ALG	polinomi
10-ALG	numeri irrazionali, razionali, interi
11-FUN	funzioni composte e funzione inversa
12-FUN	proprietà delle potenze
13-FUN	proprietà del modulo, equazioni e disequazioni con...
14-FUN	proprietà dei logaritmi
15-FUN	angoli e trigonometria
16-FUN	equazioni e disequazioni algebriche intere
17-FUN	equazioni e disequazioni algebriche fratte
18-FUN	equazioni e disequazioni irrazionali
19-GRACIA	operazioni sui grafici
20-GRACIA	grafici funzioni pari, dispari, periodiche, confro...
21-FISICA	fisica
22-FISICA	fisica
23-FISICA	fisica
24-FISICA	fisica
25-FISICA	fisica
26-TESTE	comprensione del testo
27-TESTE	comprensione del testo
28-TESTE	comprensione del testo
29-TESTE	comprensione del testo
30-TESTE	comprensione del testo

Nella figura seguente viene visualizzata la terza pagina del report creato su Power BI Desktop denominata **“Medie della Singola Domanda per Tipologia in Relazione alla Scuola”**.

ID Categoria	tecnico Nautico	Istituto Tecnico per il Turismo	Istituto Tecnico Sperimentale	ITET Enrico Fermi Lanciano	Liceo Artistico	liceo scientifico	Liceo scientifico "A. Einstein"	Pedagogico	Scientifico
01-LOG	0,33000	0,00000	0,20500	0,00000	0,18727	0,22889	0,33000	0,12000	0,21903
02-LOG	0,08000	0,00000	0,33000	0,00000	0,04364	0,06333	0,08000	0,13455	0,13123
03-LOG	0,33000	0,00000	0,16500	0,00000	0,06000	0,25667	0,00000	0,09000	0,15685
04-LOG	0,08000	0,08000	0,16500	0,00000	0,15000	0,16444	0,00000	0,21727	0,19842
05-GEO	0,00000	0,33000	0,00000	0,00000	0,12636	0,22889	0,33000	0,11182	0,21511
06-GEO	0,33000	0,00000	0,16500	0,33000	0,05091	0,14556	0,08000	0,06545	0,15761
07-GEO	0,33000	0,08000	0,20500	0,08000	0,05091	0,13667	0,00000	0,13455	0,17960
08-GEO	0,08000	0,08000	0,04000	0,08000	0,08818	0,09000	0,08000	0,08818	0,11717
09-ALG	0,00000	0,08000	0,16500	0,33000	0,08909	0,22000	0,08000	0,11182	0,14504
10-ALG	0,33000	0,08000	0,16500	0,08000	0,04364	0,26556	0,08000	0,09727	0,16654
11-FUN	0,08000	0,00000	0,33000	0,08000	0,05091	0,12778	0,08000	0,06636	0,11849
12-FUN	0,33000	0,08000	0,20500	0,33000	0,01455	0,17333	0,33000	0,09000	0,18074
13-FUN	0,00000	0,00000	0,16500	0,08000	0,09636	0,08111	0,08000	0,13455	0,11729
14-FUN	0,00000	0,33000	0,16500	0,08000	0,07273	0,10000	0,08000	0,11091	0,10513
15-FUN	0,00000	0,08000	0,04000	0,08000	0,08000	0,10000	0,00000	0,11091	0,13728
16-FUN	0,33000	0,08000	0,16500	0,00000	0,09636	0,20111	0,00000	0,18727	0,22810
17-FUN	0,00000	0,00000	0,00000	0,33000	0,08182	0,23778	0,33000	0,19455	0,21517
18-FUN	0,00000	0,00000	0,16500	0,00000	0,08909	0,08989	0,08000	0,05909	0,06882
19-GRA	0,00000	0,08000	0,33000	0,33000	0,06636	0,16444	0,08000	0,15636	0,16265
20-GRA	0,00000	0,08000	0,16500	0,08000	0,19455	0,23778	0,00000	0,06636	0,17392
21-FIS	0,33000	0,08000	0,04000	0,33000	0,08818	0,12667	0,08000	0,09636	0,17070
22-FIS	0,08000	0,00000	0,20500	0,33000	0,06636	0,17333	0,08000	0,08091	0,11445
23-FIS	0,08000	0,00000	0,04000	0,08000	0,14091	0,10889	0,08000	0,11091	0,10353
24-FIS	0,33000	0,08000	0,04000	0,33000	0,10364	0,14556	0,08000	0,08909	0,16506
25-FIS	0,33000	0,08000	0,04000	0,33000	0,06636	0,17333	0,08000	0,09636	0,15343
26-TES	0,00000	0,33000	0,16500	0,08000	0,27000	0,25667	0,33000	0,30000	0,27756
27-TES	0,33000	0,00000	0,16500	0,33000	0,27000	0,25667	0,33000	0,27000	0,27712
28-TES	0,33000	0,33000	0,33000	0,08000	0,18727	0,33000	0,33000	0,30000	0,28452
29-TES	0,33000	0,00000	0,33000	0,08000	0,18727	0,25667	0,33000	0,24000	0,27895
30-TES	0,33000	0,00000	0,20500	0,33000	0,18727	0,29333	0,33000	0,27000	0,28058
Media	0,17000	0,07867	0,16317	0,15300	0,11000	0,17844	0,13900	0,14003	0,17667

La terza pagina del report è interamente dedicata ad una grande tabella in cui, nelle righe troviamo le 30 tipologie di domande a cui è invitato a rispondere colui che si sottopone al test, mentre, nelle colonne troviamo tutte le varie scuole di provenienza degli studenti, inoltre, nelle celle della tabella facendo le dovute intersezioni è possibile visualizzare la media riscontrata degli studenti di una particolare scuola in una determinata tipologia di domanda. Attraverso la consultazione di tale tabella è possibile osservare la preparazione offerta dalla singola scuola, chiaramente valutata in base alle medie ottenute dagli studenti, riguardo uno specifico argomento presente nel test.

Nella figura seguente viene visualizzata la quarta pagina del report creato su Power BI Desktop denominata “**Medie Test per Scuola e per Data**”, la quale risponde alle domande numero 8,9,10,11,12,13 presenti nel paragrafo 1.1.

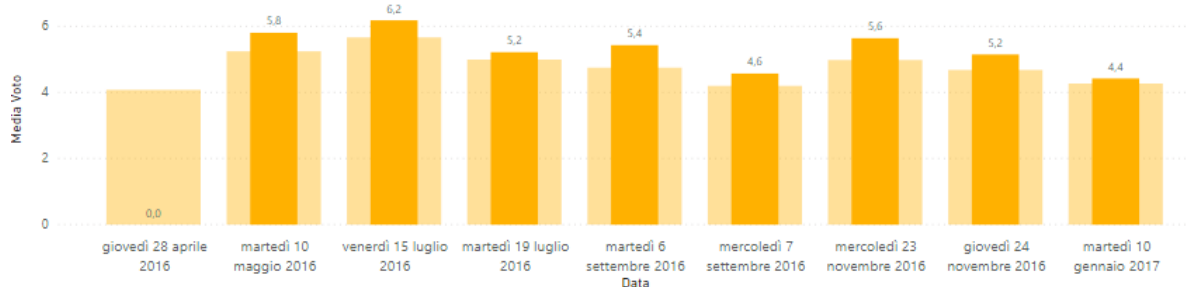


Il punteggio complessivo ottenuto nel test è dato dalla somma dei punteggi delle singole domande. La media generale degli studenti che svolgono il test è di 4,72. La media più alta è ottenuta dagli studenti dell'Istituto d'Arte con 6,10 mentre la più bassa è stata ottenuta dagli studenti dell'Istituto Tecnico per il Turismo con 2.36, esse però non sono molto significative in quanto in entrambi i casi le sopracitate scuole hanno partecipato al test con affluenza pari ad uno. Per tale motivo quindi è opportuno considerare il Liceo Scientifico la scuola non solo con maggior affluenza ma anche quella con il punteggio medio più alto di circa 5.30, più alto rispetto alla media 0.6. La media più bassa invece è stata ottenuta dagli studenti del Liceo Artistico con 3.30, di ben oltre un punto sotto la soglia dei 4.72.

Sulla destra della pagina è presente un grafico a linee il quale mostra il trend della media del test negli anni, nel momento in cui per il momento si prendono in esame solo test svolti durante l'anno accademico 2016-2017, esso non risulta particolarmente significativo ma comunque fa notare una media inferiore per i test svolti nel 2017. Sarà maggiormente interessante osservare tale grafico nel momento in cui, in seguito a sviluppi futuri verranno aggiunti dati al database che spaziano fra diversi anni accademici.

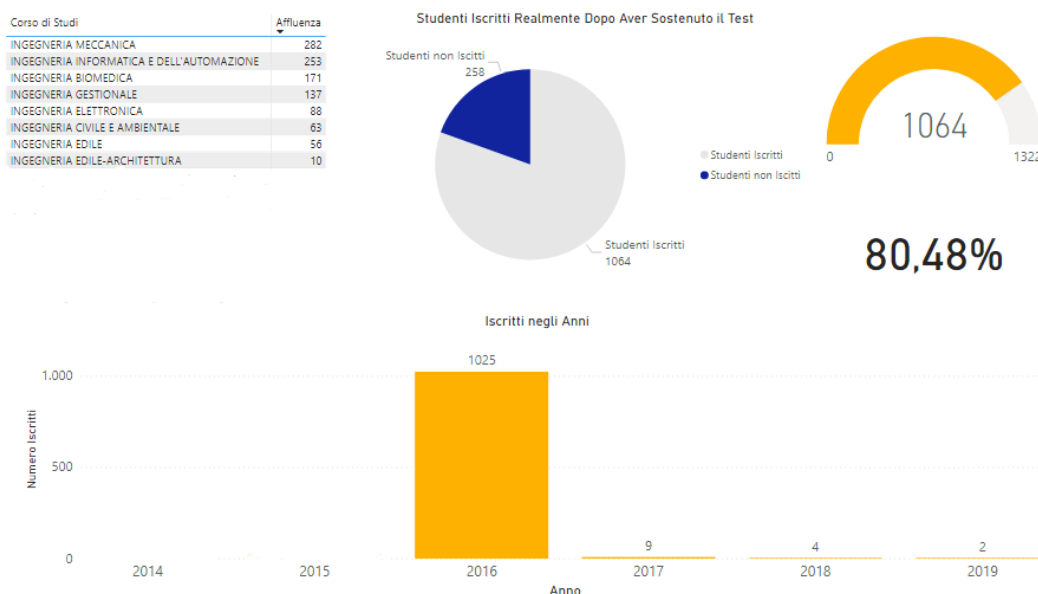
Nella parte inferiore della pagina è possibile osservare la media ottenuta in una specifica data del test, essa ha avuto un picco massimo di 5,7 venerdì 15 luglio 2016 mentre ha avuto un picco minimo di 4,1 giovedì 28 aprile 2016. Si può notare che nei Test di Ingresso svolti appena terminata la scuola superiore o a ridosso della fine di quest'ultima è stato ottenuto un punteggio mediamente più alto e in alcuni casi di anche un punto rispetto alle altre date, e ciò è veramente molto pensando che la media generale si aggira intorno 4.7.

Da sottolineare la funzionalità attraverso la quale selezionando una scuola nella tabella in alto a sinistra, sarà possibile visualizzare nell'istogramma a colonne raggruppate la media degli studenti della scuola selezionata nelle specifiche date in cui è stato svolto il test. Nella figura sottostante è mostrato il risultato medio del test degli studenti del *Liceo Scientifico* nelle date in cui esso è stato svolto.



Si può notare che la media di coloro che appartenevano a questa scuola è sempre stata al di sopra della media complessiva degli studenti, eccezion fatta per il giovedì 28 aprile 2016, data in cui nessuno studente proveniente dal *Liceo Scientifico* ha preso parte al test.

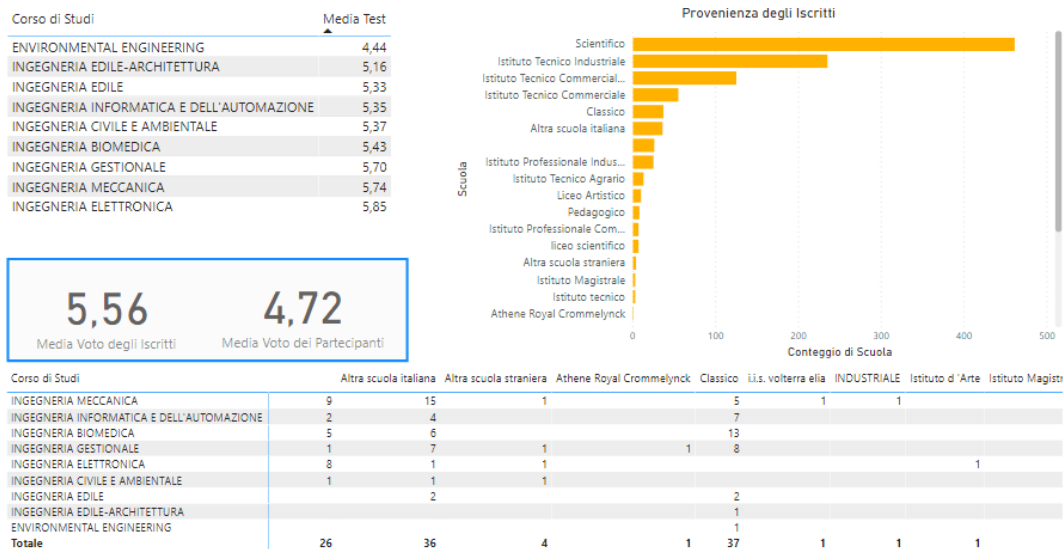
Nella figura seguente viene visualizzata la quinta pagina del report creato su Power BI Desktop denominata **“Iscrizioni”**, la quale risponde alle domande numero 14,15,17,17 presenti nel paragrafo 1.1.



Circa un 80% degli studenti che partecipano al test successivamente si iscrivono realmente alla nostra università, questo vuol dire che nella maggioranza dei casi chi partecipa al test è già ben orientato verso la scelta. Il corso con il maggior numero di iscritti è quello di Ingegneria Meccanica seguito da quello di Ingegneria Informatica e dell'Automazione, Biomedica e Gestionale. Molto meno abbondante è l'affluenza verso i corsi di Ingegneria Elettronica, Civile, Civile e Ambientale, Edile ed Edile-Architettura.

Da sottolineare la funzionalità attraverso la quale selezionando un Corso di Studi nella tabella in alto a sinistra, sarà possibile visualizzare nell'istogramma a colonne raggruppate il numero degli studenti che si sono iscritti al corso selezionato negli anni, tale funzionalità risulterà maggiormente utile nel momento in cui, in seguito ad alcuni sviluppi futuri, verranno inseriti nel database informazioni relative ai test di ingresso ed alle iscrizioni antecedenti e successive al 2016.

Nella figura seguente viene visualizzata la sesta pagina del report creato su Power BI Desktop denominata “**Medie degli Iscritti in Relazione al Corso di Studi**”, la quale risponde alle domande numero 18,19,20,21 presenti nel paragrafo 1.1.



La media voto del test di coloro che poi successivamente si sono iscritti all'Univpm è di 5,56 ed è di quasi un punto superiore rispetto alla media voto ottenuta nel test dalla totalità degli studenti, questo fa capire che coloro che avevano ottenuto un punteggio "basso" non si sono poi iscritti realmente.

Il punteggio medio ottenuto dagli studenti che si sono iscritti ai vari corsi di laurea è molto simile, il più alto è quello di *Ingegneria Elettronica* (5,85) mentre il più basso è quello di *Environmental Engineerig* (4,44).

La provenienza degli iscritti segue il trend dei partecipanti al test, infatti la scuola da cui provengono la maggior parte degli studenti è il liceo scientifico (con ben 461 iscrizioni), seguono l' Istituto Tecnico Industriale (con 235 iscrizioni), l' Istituto Tecnico Commerciale per Geometri (125) e l' Istituto Tecnico Commerciale (55).

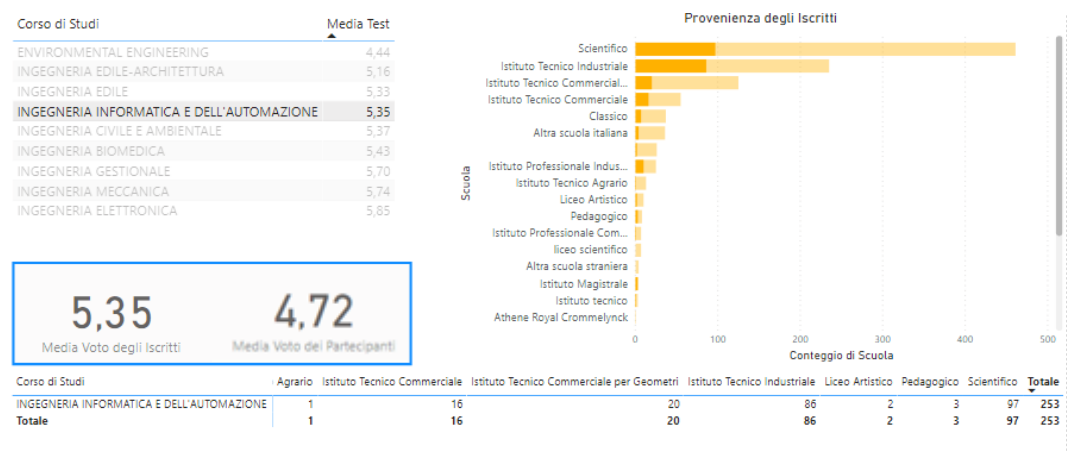
Nella parte bassa della pagina è possibile visualizzare una matrice, la quale riepiloga: il numero di studenti, da dove essi provengono e quale corso di

studio hanno poi scelto.

Corso di Studi	Istituto Tecnico Sperimentale	ITET Enrico Fermi Lanciano	Liceo Artistico	liceo scientifico	Liceo scientifico "A. Einstein"	Pedagogico	Scientifico	Totale
INGEGNERIA MECCANICA	1		1	3			128	282
INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE			2			3	97	253
INGEGNERIA BIOMEDICA		1		3		1	103	171
INGEGNERIA GESTIONALE					1	3	66	137
INGEGNERIA ELETTRONICA							30	88
INGEGNERIA CIVILE E AMBIENTALE			1	1	1	1	20	63
INGEGNERIA EDILE				6			10	56
INGEGNERIA EDILE-ARCHITETTURA							4	10
ENVIRONMENTAL ENGINEERING								1

Nella sovrastante figura è possibile che ben 128 studenti provenienti dal Liceo Scientifico si sono iscritti al corso di *Ingegneria Meccanica*, 97 a quello di *Ingegneria Informatica e dell'Automazione ecc.*

Da sottolineare la funzionalità attraverso la quale, selezionando un Corso di Studi in alto a sinistra sarà possibile osservare la provenienza dei suoi iscritti negli altri due grafici presenti nella pagina. Questa funzionalità è osservabile in figura.



Nella figura seguente viene visualizzata la settima ed ultima pagina del report creato su Power BI Desktop denominata **“Medie degli Iscritti in Relazione al Corso di Studi pt.2”**, la quale risponde alle domande numero 22,23 presenti nel paragrafo 1.1.

Valutazione Test degli Iscritti		
Username Test	Corso di Appartenenza	VotoTest
ofa0001	INGEGNERIA MECCANICA	10,32
ofa0003	INGEGNERIA EDILE	3,26
ofa0005	INGEGNERIA MECCANICA	4,61
ofa0006	INGEGNERIA BIOMEDICA	5,75
ofa0007	INGEGNERIA MECCANICA	4,26
ofa0008	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,29
ofa0009	INGEGNERIA MECCANICA	9,20
ofa0010	INGEGNERIA CIVILE E AMBIENTALE	3,36
ofa0011	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	8,38
ofa0012	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,76
ofa0014	INGEGNERIA MECCANICA	4,51
ofa0015	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,49
ofa0017	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,24
ofa0019	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	8,04
ofa0020	INGEGNERIA EDILE	3,61
ofa0021	INGEGNERIA EDILE	3,67
ofa0022	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,77
ofa0026	INGEGNERIA CIVILE E AMBIENTALE	4,59
ofa0030	INGEGNERIA BIOMEDICA	3,36
ofa0031	INGEGNERIA EDILE	5,09
ofa0032	INGEGNERIA MECCANICA	3,30
ofa0033	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,44
ofa0034	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,84
ofa0035	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,37
ofa0038	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,53
ofa0040	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,50
Media		5,56

Iscritti che hanno conseguito un punteggio sopra la media		
UsernameTest	Corso di Appartenenza	VotoTest
ofa0001	INGEGNERIA MECCANICA	10,32
ofa0006	INGEGNERIA BIOMEDICA	5,75
ofa0009	INGEGNERIA MECCANICA	9,20
ofa0011	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	8,38
ofa0012	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,76
ofa0015	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,49
ofa0017	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,24
ofa0019	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	8,04
ofa0031	INGEGNERIA EDILE	5,09
ofa0034	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,84
ofa0040	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,50
ofa0046	INGEGNERIA GESTIONALE	7,06

In questa pagina, sulla sinistra, viene riportato: lo username che viene associato al singolo studente, il Corso di Laurea a cui egli ha successivamente preso parte e la votazione ottenuta nel test. In basso a destra con una logica molto simile troviamo un analogo della tabella sopra descritta, ma in essa compariranno solamente gli studenti che hanno ottenuto nel test una votazione sopra la media, ossia maggiore di 4,72.

Utilizzando l'apposito filtro dei dati colorato in arancione, spuntando la checkbox relativa ad uno specifico Corso di Laurea, sarà possibile visualizzare, sulla sinistra, il risultato al test dei singoli studenti appartenenti a tale Corso mentre, in basso a destra, verranno evidenziati solamente quelli che avevano conseguito un punteggio superiore alla media. In figura è mostrata la situazione in cui viene spuntata la checkbox relativa ad

Ingegneria informatica e dell'Automazione.

Valutazione Test degli Iscritti		
Username Test	Corso di Appartenenza	VotoTest
ofa0008	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,29
ofa0011	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	8,38
ofa0012	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,76
ofa0015	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,49
ofa0017	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,24
ofa0019	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	8,04
ofa0022	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,77
ofa0033	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,44
ofa0034	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,84
ofa0035	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,37
ofa0038	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,53
ofa0040	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,50
ofa0043	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,95
ofa0045	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,43
ofa5001	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,17
ofa5006	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,33
ofa5021	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	10,36
ofa5023	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,02
ofa5024	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,84
ofa5024	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,84
ofa5024	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,84
ofa5037	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,61
ofa5040	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	3,79
ofa5042	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,59
ofa5043	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,77
Media		5,35

Isctti che hanno conseguito un punteggio sopra la media		
UsernameTest	Corso di Appartenenza	VotoTest
ofa0011	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	8,38
ofa0012	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	4,76
ofa0015	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,49
ofa0017	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,24
ofa0019	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	8,04
ofa0034	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,84
ofa0040	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	5,50
ofa5001	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,17
ofa5006	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,33
ofa5021	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	10,36
ofa5023	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,02
ofa5024	INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE	6,84

SELEZIONARE UN CORSO DI STUDIO PER VISUALIZZARE IL RISULTATO DEI SUOI ISCRITTI

- (Vuoto)
- ENVIRONMENTAL ENGINEERING
- INGEGNERIA BIOMEDICA
- INGEGNERIA CIVILE E AMBIENTALE
- INGEGNERIA EDILE
- INGEGNERIA EDILE-ARCHITETTURA
- INGEGNERIA ELETTRONICA
- INGEGNERIA GESTIONALE
- INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE
- INGEGNERIA MECCANICA

Si può notare che gli studenti che hanno ottenuto un punteggio sotto la media rappresentano una minoranza, ed in ogni caso, il punteggio ottenuto non risulta drasticamente inferiore non superando quasi mai l'unità di scarto.

CAPITOLO 3: Conclusioni e Sviluppi Futuri

3.1 Conclusioni

L'obiettivo di questa tesi di laurea è quello di fornire tutte le potenziali considerazioni sui dati relativi ad i test di ingresso dell'anno accademico 2016-2017 rispondendo ad una serie di domande elencate nel paragrafo 1.2 e di sviluppare un insieme di strumenti grafici per effettuare un'analisi descrittiva di quest'ultimi. Tali obiettivi sono stati raggiunti con successo attraverso un'accurata analisi del database su cui i dati erano memorizzati e grazie alla creazione di un Report utilizzando il software Power BI. In questo modo l'elaborato fornisce delle informazioni utili sia per gli studenti che intendono intraprendere un percorso universitario, rendendoli consci delle possibili difficoltà a cui essi vanno incontro, suggerendo ad essi un percorso formativo di secondo grado che li possa preparare al meglio e mettendo in evidenza le difficoltà che alcuni loro colleghi hanno incontrato in passato, ma anche utile per coloro che all'interno dell'università lavorano, in modo tale che possano modificare e/o integrare il Piano di Studi proposto agli studenti colmando loro eventuali lacune.

3.2 Sviluppi Futuri

È importante tenere presente che questa ricerca si è concentrata nel rispondere ad alcune domande attraverso l'utilizzo delle informazioni ricavabili dai dati contenuti in un software esterno, chiaramente è possibile approfondire questo studio che non è da ritenersi completo in tutti i suoi aspetti.

Un possibile spunto per uno sviluppo futuro potrebbe essere quello di andar a pulire i dati su cui sono state fatte le analisi, andando cioè ad eliminare rappresentazioni differenti per lo stesso dato. Ad esempio, spesse volte, scuole uguali che sono state specificate dallo studente figurano come differenti solamente per il fatto che esse sono state scritte con l'ausilio della lettera maiuscola o minuscola.

Altra idea interessante è quella di utilizzare il Data Mining che è un insieme di tecniche e metodologie in grado di scovare associazioni, anomalie e schemi ricorrenti, quindi in definitiva informazioni utili, da grandi quantità di dati attraverso metodi automatici o semi-automatici.

Infine, una raccomandazione per un ulteriore sviluppo futuro potrebbe essere quella di andare ad integrare i dati con delle informazioni riguardanti la carriera degli studenti che si erano approcciate al test, riuscendo magari a rispondere ad alcune delle seguenti domande:

- C'è correlazione fra un basso punteggio nel test e il tempo impiegato per portar a termine il percorso di studi?
- C'è correlazione fra il punteggio ottenuto in una specifica tipologia di domanda (es. fisica, logica, ecc...) e il tempo impiegato per portar a termine il percorso di studi?
- C'è correlazione fra il punteggio ottenuto nel test ed il voto di Laurea?
- Quanti studenti sono riusciti a completare il percorso di studi in tre anni?

- Quanti studenti dopo essersi iscritti hanno successivamente rinunciato?
- Qual è stato il punteggio medio del test di coloro che sono riusciti a completare il percorso di studi in tre anni?

Bibliografia

[1] P. Atzeni, S. Ceri, P. Fraternali, S. Paraboschi, R Torlone, Basi di Dati,

Mc Graw Hill Education.

[2] Informazioni ricavate dal sito : [https://docs.microsoft.com/it-it/power-](https://docs.microsoft.com/it-it/power-bi/fundamentals/power-bi-overview)

[bi/fundamentals/power-bi-overview](https://docs.microsoft.com/it-it/power-bi/fundamentals/power-bi-overview).