

# HOMEWORK

---

Bosio Anna

Abete Davide

Strada Tommaso

# 1:who is playing? - Dati

cdscod	cds	anno_appello	numero_studenti
E2401P	[E2401P] SCIENZE E TECNICHE PSICOLOGICHE	2015	1
581	[581] GIURISPRUDENZA	2016	716
E0201Q	[E0201Q] BIOTECNOLOGIE	2016	53
E1301Q	[E1301Q] SCIENZE BIOLOGICHE	2016	16
E1401A	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	2016	94
...	...	...	...
I0301D	[I0301D] IGIENE DENTALE (ABILITANTE ALLA PROFE...	2017	765
I0302D	[I0302D] TECNICHE DI LABORATORIO BIOMEDICO (AB...	2017	799
I0303D	[I0303D] TECNICHE DI RADIOLOGIA MEDICA, PER IM...	2017	572
K0101D	[K0101D] SCIENZE INFERMIERISTICHE E OSTETRICHE	2017	482
E3201Q	[E3201Q] SCIENZE E TECNOLOGIE PER L'AMBIENTE	2018	1

I dati rappresentano la distribuzione degli studenti iscritti nei vari appelli negli anni per ogni corso di laurea. 'numero\_studenti' indica il numero di iscritti negli appelli dell'anno riportato in 'anno\_appello' per il corso di laurea riportato in 'cds'.

```
select cds.cdscod, cds.cds, strftime('%Y',dtappello) as anno_appello, count(*) as numero_studenti
from iscrizioni as isc JOIN appelli as app ON isc.appcod = app.appcod JOIN cds ON app.cdscod = cds.cdscod
group by anno_appello, cds.cdscod
```

Non ci sono differenze tra i risultati ottenuti dal modello relazionale denormalizzato e il modello razionale normalizzato.

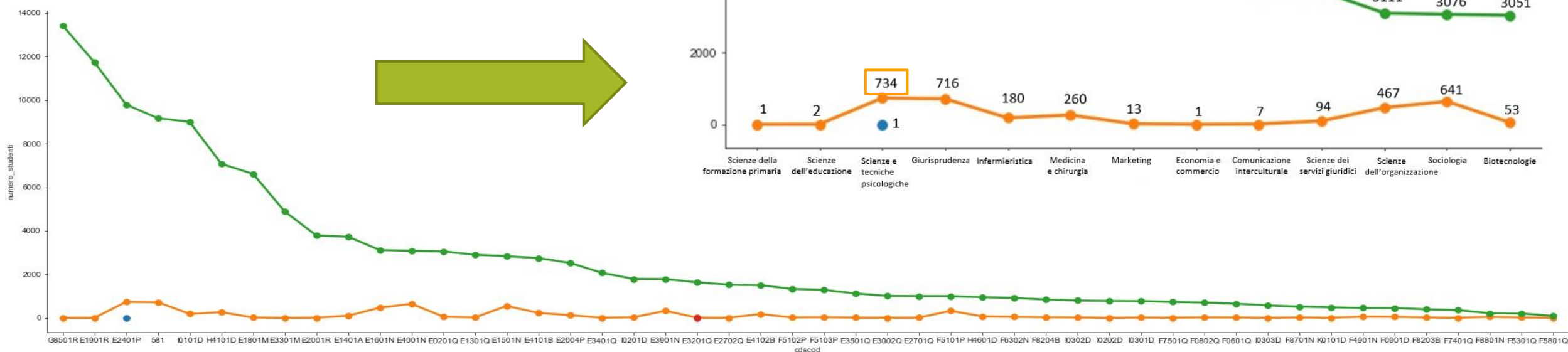
```
select cds, strftime('%Y',dtappello) as anno_appello, count(*) as numero_studenti
from bos_denormalizzato
group by anno_appello, cdscod
```

Dal grafico si vede che nel database sono contenute in prevalenza informazioni del 2017 e del 2016.

Mentre per gli anni 2015 e 2018 è presente una sola iscrizione, nel corso di Scienze e tecniche psicologiche per il 2015 e nel corso di Scienze e tecnologie per l'ambiente per il 2018.

IL corso con più iscritti nei vari appelli è Scienze della formazione primaria con 13415 iscrizioni. Mentre nel 2016 il corso con più iscritti è Scienze e tecniche psicologiche con 734 iscrizioni.

# Query 1: distribuzione

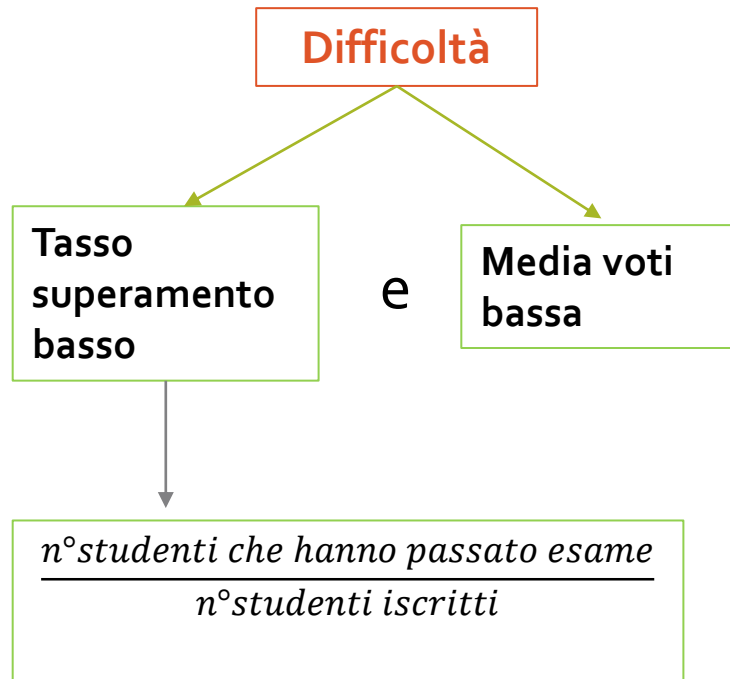


## 2. I'll never graduate

Individuazione della Top-10 degli esami più difficili suddivisi per corso di studi.

SQL

Definiamo la difficoltà:



I dati estratti sono stati raggruppati per 'cdsCod' e 'AdCod' dapprima per calcolare il numero di iscritti a ciascun esame e la votazione media, successivamente per calcolare il n° di studenti che hanno passato l'esame. Infine viene calcolato il Tasso di superamento. *Sono stati considerati solamente gli esami superati da almeno uno studente.* Il risultato è identico per le due query.

Normalizzato

```
SELECT a.CdsCod,
       a.AdCod,
       a.n°iscritti,
       b.n°passati,
       round(CAST (b.n°passati AS FLOAT) / CAST (a.n°iscritti AS FLOAT), 2) AS Tasso_superamento,
       round(1-CAST (b.n°passati AS FLOAT) / CAST (a.n°iscritti AS FLOAT), 2)*100 AS Tasso_bocciatura_percentuale,
       a.voto_medio
FROM (
  SELECT CdsCod,
         AdCod,
         count() AS n°iscritti,
         round(avg(Voto), 2) AS voto_medio
    FROM appelli AS app
   JOIN
   iscrizioni AS iscr ON app.appcod = iscr.appcod
 WHERE Iscrizione = 1
  GROUP BY CdsCod,
           AdCod
)
a
JOIN
(
  SELECT CdsCod,
         AdCod,
         count() AS n°passati
    FROM appelli AS app
   JOIN
   iscrizioni AS iscr ON app.appcod = iscr.appcod
 WHERE Iscrizione = 1 AND
       superamento = 1
  GROUP BY CdsCod,
           AdCod
)
b ON a.CdsCod = b.CdsCod AND
   a.AdCod = b.AdCod
ORDER BY Tasso_superamento ASC
LIMIT 10;
```

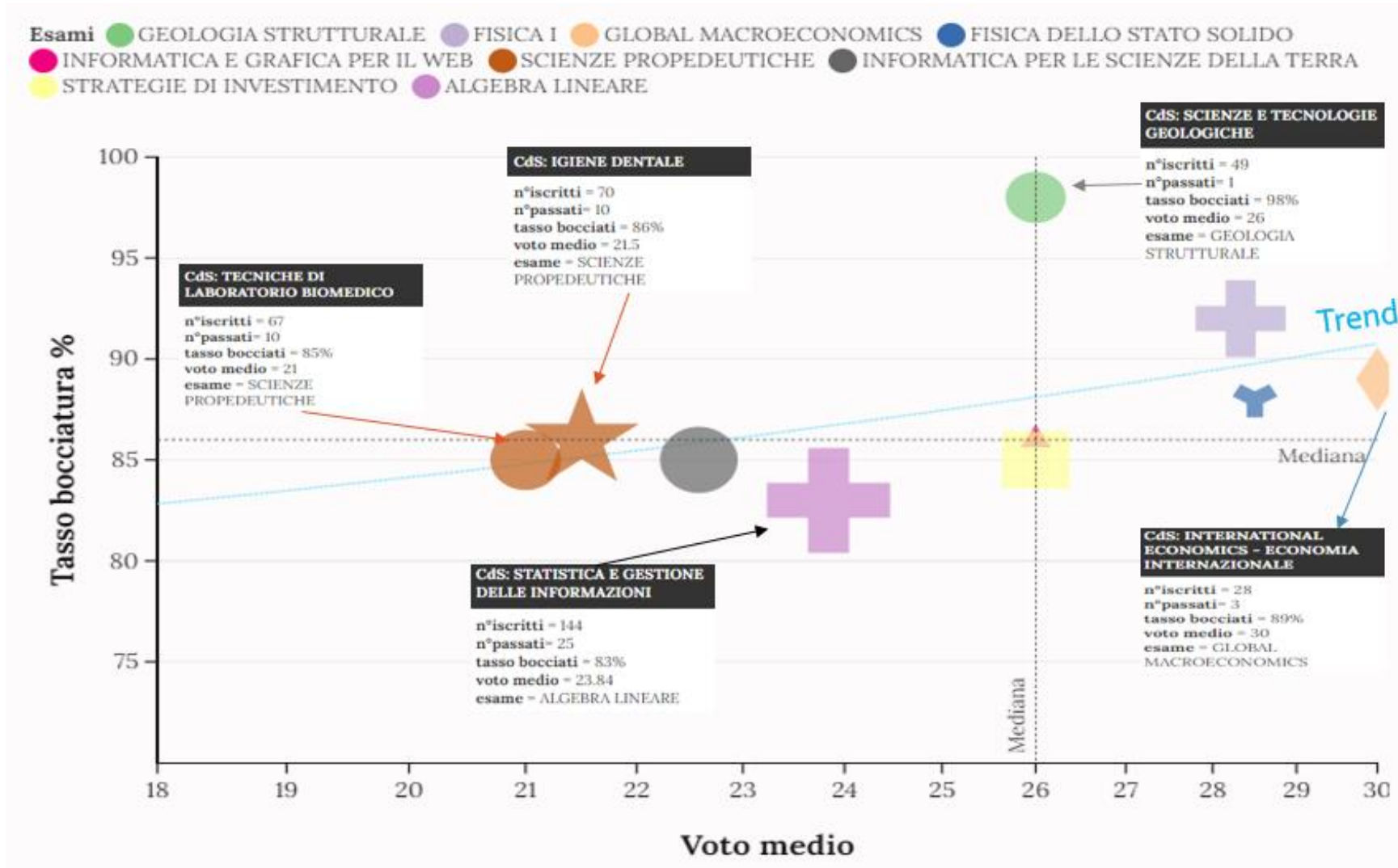
DeNormalizzato

```
SELECT a.CdsCod,
       a.Cds,
       a.AD,
       a.AdCod,
       a.TipoCorso,
       a.n°iscritti,
       b.n°passati,
       round(CAST (b.n°passati AS FLOAT) / CAST (a.n°iscritti AS FLOAT), 2) AS Tasso_superamento,
       round(1-CAST (b.n°passati AS FLOAT) / CAST (a.n°iscritti AS FLOAT), 2)*100 AS Tasso_bocciatura_percentuale,
       a.voto_medio
FROM (
  SELECT CdsCod,
         Cds,
         AD,
         AdCod,
         TipoCorso,
         count() AS n°iscritti,
         round(avg(Voto), 2) AS voto_medio
    FROM bos_denormalizzato
   WHERE Iscrizione = 1
  GROUP BY CdsCod,
           AdCod
)
a
JOIN
(
  SELECT CdsCod,
         AdCod,
         count() AS n°passati
    FROM bos_denormalizzato
   WHERE Iscrizione = 1 AND
       superamento = 1
  GROUP BY CdsCod,
           AdCod
)
b ON a.CdsCod = b.CdsCod AND
   a.AdCod = b.AdCod
ORDER BY Tasso_superamento ASC
LIMIT 10
```

# Query 2: esami più difficili

Gli esami più difficili sono quelli con **alto tasso di bocciatura (1-Tasso superamento)** e **media voto bassa** (Secondo quadrante). Di seguito viene mostrata la **top10** degli esami più difficili suddivisi per **corso di laurea** ed **esame**. La **dimensione** è proporzionale al **n° di iscritti** all'esame.

L'esame più difficile è....  
**Scienze propedeutiche.**  
Notiamo che è sostenuto dagli studenti di due corsi di laurea distinti ma entrambi hanno un **tasso di bocciatura intorno** al 50°percentile e una **media voto inferiore** al 50% delle osservazioni.



**Trend:**  
mediamente  
all'aumentare  
del tasso di  
bocciatura  
aumenta il  
voto medio  
(**correlazione  
lineare  
positiva**)



# 3:Commitment – Dati

Abbiamo raggruppato per *'cdscod'* e per *'dtappello'*, per poi contare quanti esami dello stesso cds sono stati svolti nello stesso giorno, per poi prendere per ogni *cds* il numero massimo di esami nello stesso giorno. Utilizzando il dataset denormalizzato il risultato è diverso, in quanto raggruppando per *cdscod* e *dtappello* restituisce un numero di righe pari a quanti sono gli studenti iscritti.

## Normalizzato

	cds	n°esami stesso giorno	dtappello	tipocorso
1	[H4101D] MEDICINA E CHIRURGIA	42	2017-03-31	[LM6] Laurea Magistrale Ciclo Unico 6 anni
2	[S81] GIURISPRUDENZA	32	2017-05-03	[LM5] Laurea Magistrale Ciclo Unico 5 anni
3	[I0101D] INFERMIERISTICA (ABILITANTE ALLA PROFESSIONE SANITARIA DI INFERMIERE)	17	2017-09-22	[L2] Corso di Laurea - L2
4	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	13	2017-09-27	[L2] Corso di Laurea - L2
5	[E4001N] SOCIOLOGIA	12	2017-07-21	[L2] Corso di Laurea - L2
6	[E1801M] MARKETING, COMUNICAZIONE AZIENDALE E MERCATI GLOBALI	12	2017-04-11	[L2] Corso di Laurea - L2
7	[E1901R] SCIENZE DELL'EDUCAZIONE	11	2017-06-27	[L2] Corso di Laurea - L2
8	[E1803M] ECONOMIA DELLE BANCHE, DELLE ASSICURAZIONI E DEGLI INTERMEDIARI FINANZIARI	11	2017-04-11	[L2] Corso di Laurea - L2
9	[G8501R] SCIENZE DELLA FORMAZIONE PRIMARIA	10	2017-02-28	[LM5] Laurea Magistrale Ciclo Unico 5 anni
10	[E3301M] ECONOMIA E COMMERCIO	10	2017-04-20	[L2] Corso di Laurea - L2
11	[E2001R] COMUNICAZIONE INTERCULTURALE	10	2017-07-11	[L2] Corso di Laurea - L2
12	[E1802M] ECONOMIA E AMMINISTRAZIONE DELLE IMPRESE	10	2017-04-12	[L2] Corso di Laurea - L2
13	[H4601D] ODONTOIATRIA E PROTESI DENTARIA	9	2017-03-31	[LM6] Laurea Magistrale Ciclo Unico 6 anni
14	[F7702M] MARKETING E MERCATI GLOBALI	9	2017-06-21	[LM] Corso di Laurea Magistrale
15	[F1701Q] FISICA	9	2017-07-19	[LM] Corso di Laurea Magistrale

```
SELECT d.cds,
       max(d.conta) AS n°esami_stesso_giorno,
       d.dtappello,
       d.tipocorso
FROM (
  SELECT a.dtappello,
         a.cdscod,
         c.cds,
         count(a.cdscod) AS conta,
         c.tipocorso
  from appelli as a join cds as c on a.cdscod=c.cdscod
  GROUP BY a.cdscod,
           a.dtappello
  ORDER BY count(a.cdscod) DESC
)
GROUP BY d.cds
ORDER BY n°esami_stesso_giorno DESC;
```

```
SELECT d.cds,
       max(d.conta) AS n°esami_stesso_giorno,
       d.dtappello,
       d.tipocorso
FROM (
  SELECT b.dtappello,
         b.cdscod,
         b.cds,
         count(b.cdscod) AS conta,
         b.tipocorso
  FROM bos_denormalizzato AS b
  GROUP BY b.cdscod,
           b.dtappello
  ORDER BY count(b.cdscod) DESC
)
GROUP BY d.cds
ORDER BY n°esami_stesso_giorno DESC;
```

## Denormalizzato

	cds	n°esami stesso giorno	dtappello	tipocorso
1	[G8501R] SCIENZE DELLA FORMAZIONE PRIMARIA	1130	2017-07-25	[LM5] Laurea Magistrale Ciclo Unico 5 anni
2	[E1901R] SCIENZE DELL'EDUCAZIONE	977	2017-07-10	[L2] Corso di Laurea - L2
3	[I0101D] INFERMIERISTICA (ABILITANTE ALLA PROFESSIONE SANITARIA DI INFERMIERE)	595	2017-07-10	[L2] Corso di Laurea - L2
4	[E2401P] SCIENZE E TECNICHE PSICOLOGICHE	499	2017-02-21	[L2] Corso di Laurea - L2
5	[E4001N] SOCIOLOGIA	485	2017-07-21	[L2] Corso di Laurea - L2
6	[E3101Q] INFORMATICA	461	2017-06-16	[L2] Corso di Laurea - L2
7	[F7702M] MARKETING E MERCATI GLOBALI	444	2017-06-21	[LM] Corso di Laurea Magistrale
8	[S81] GIURISPRUDENZA	419	2017-05-03	[LM5] Laurea Magistrale Ciclo Unico 5 anni
9	[H4101D] MEDICINA E CHIRURGIA	410	2017-06-22	[LM6] Laurea Magistrale Ciclo Unico 6 anni
10	[E1802M] ECONOMIA E AMMINISTRAZIONE DELLE IMPRESE	400	2017-06-14	[L2] Corso di Laurea - L2
11	[E1801M] MARKETING, COMUNICAZIONE AZIENDALE E MERCATI GLOBALI	359	2017-06-21	[L2] Corso di Laurea - L2
12	[E1501N] SCIENZE DEL TURISMO E COMUNITÀ? LOCALE	323	2017-04-21	[L2] Corso di Laurea - L2
13	[E1601N] SCIENZE DELL'ORGANIZZAZIONE	314	2016-12-14	[L2] Corso di Laurea - L2
14	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	283	2017-05-02	[L2] Corso di Laurea - L2
15	[E1803M] ECONOMIA DELLE BANCHE, DELLE ASSICURAZIONI E DEGLI INTERMEDIARI FINANZIARI	250	2017-06-20	[L2] Corso di Laurea - L2

```
SELECT CdScod,TipoCorso,AD,studente,DtAppello
FROM bos_denormalizzato
where CdScod='G8501R' and DtAppello='2017-07-25';
```

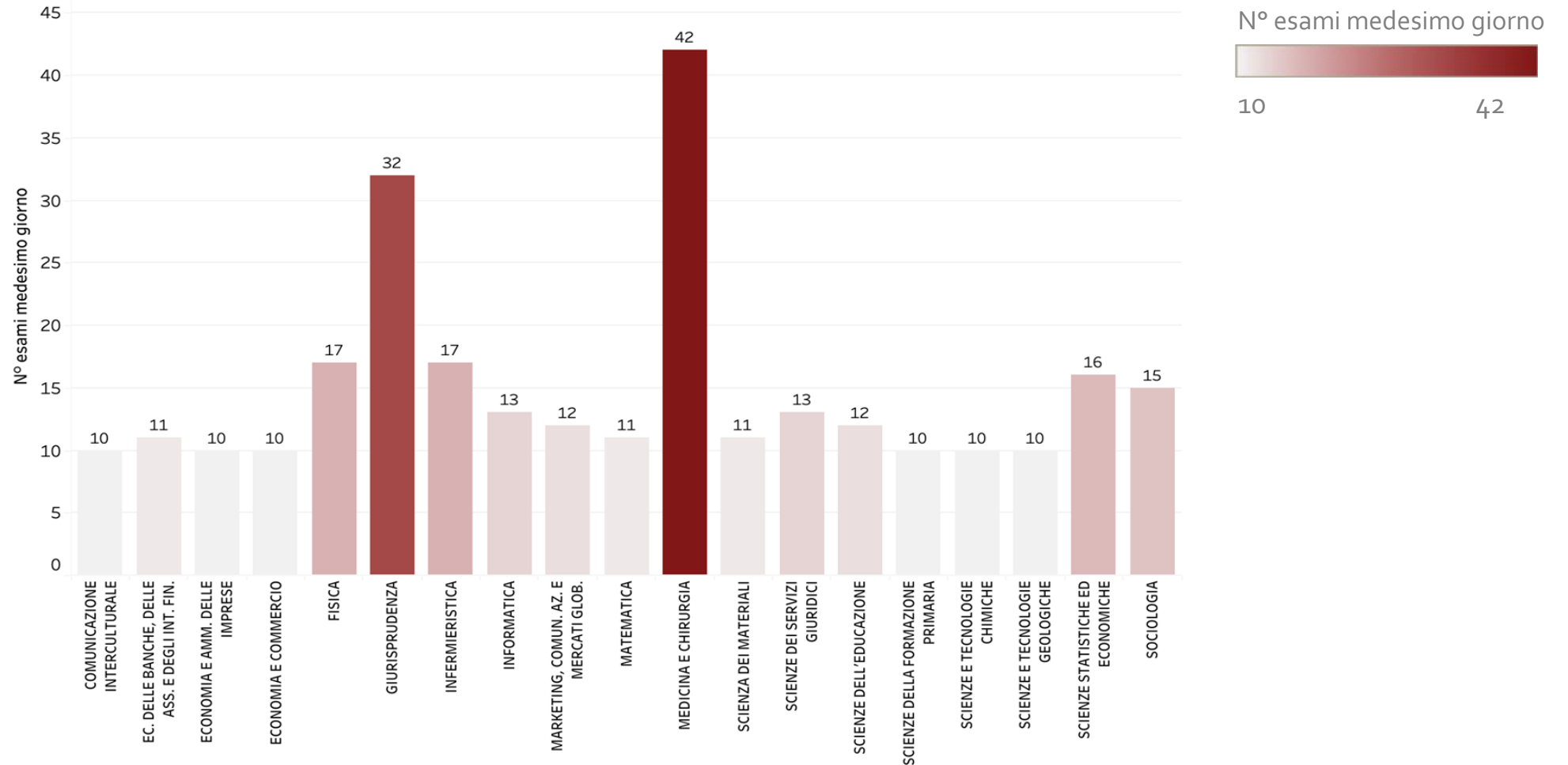
	CdScod	TipoCorso	AD	studente	DtAppello
1	G8501R	[LM5] Laurea Magistrale ...	[G8501R032] ATTIVITA' ...	4131F78F7EDDBF489C4...	2017-07-25
2	G8501R	[LM5] Laurea Magistrale ...	[G8501R032] ATTIVITA' ...	83C66F52128D6E3829E...	2017-07-25
3	G8501R	[LM5] Laurea Magistrale ...	[G8501R032] ATTIVITA' ...	A6BDED2F9F38B4EC42...	2017-07-25
4	G8501R	[LM5] Laurea Magistrale ...	[G8501R032] ATTIVITA' ...	82846641046B04CD0F5...	2017-07-25
5	G8501R	[LM5] Laurea Magistrale ...	[G8501R032] ATTIVITA' ...	E9AC0F01CD8A8DEFD8...	2017-07-25

# Query 3: Corsi di laurea ad elevato tasso di commitment

Sono mostrati i corsi di laurea, sull'asse x, che presentano almeno 10 esami svolti in un giorno.

La gradazione di colore rosso mostra il numero di esami svolti nel medesimo giorno.

Il corso di laurea con il maggiore tasso di commitment risulta essere il corso di **Medicina e chirurgia**, seguito da **Giurisprudenza**, entrambi corsi di Laurea Magistrale a Ciclo Unico.



# Query 4: hardest/easiest exams - Dati

cds	ad	adcod	media_voto	numero_superamenti
[524] SCIENZE DELL'EDUCAZIONE	[E1901R086] METODI E TECNICHE DELLA FORMAZIONE...	E1901R086	25.000000	1
[541] SCIENZE DEL TURISMO E COMUNITA' LOCALE(C...	[E1501N009] SOCIOLOGIA DELLA CULTURA E DEL TEM...	E1501N009	25.000000	1
[581] GIURISPRUDENZA	[A5810234] LA CORTE DI GIUSTIZIA: TECNICHE E S...	A5810234	30.000000	4
[581] GIURISPRUDENZA	[A5810172] INTERNATIONAL CRIMINAL JUSTICE	A5810172	30.000000	1
[581] GIURISPRUDENZA	[A5810171] LAW OF BRICS COUNTRIES AND EU	A5810171	30.000000	8
...	...	...	...	...
[K0101D] SCIENZE INFERMIERISTICHE E OSTETRICHE	[K0101D006] SCIENZE PSICOPEDAGOGICHE	K0101D006	28.423077	26
[K0101D] SCIENZE INFERMIERISTICHE E OSTETRICHE	[K0101D010] STRUTTURE ORGANIZZATIVE E FLUSSI G...	K0101D010	28.076923	13
[K0101D] SCIENZE INFERMIERISTICHE E OSTETRICHE	[K0101D003] PROGETTAZIONE E COORDINAMENTO ASSI...	K0101D003	27.551724	29
[K0101D] SCIENZE INFERMIERISTICHE E OSTETRICHE	[K0101D009] MANAGEMENT INFERMIERISTICO E OSTET...	K0101D009	26.600000	10
[K0101D] SCIENZE INFERMIERISTICHE E OSTETRICHE	[K0101D002] METODOLOGIA DI ANALISI SOCIO-SANIT...	K0101D002	24.607143	28

Per ogni attività didattica è stata calcolata la media dei voti degli studenti che hanno passato l'esame.

Per individuare la top-3 degli esami più facili per ogni corso di studio si prendono le prime tre attività didattiche (ad), mentre per la top-3 degli esami più difficili le ultime tre.

Per rendere il risultato più realistico si è scelto di affiancare la media dei voti con un altro valore, ovvero il numero di studenti che hanno passato l'esame, considerando soltanto le attività didattiche superate da almeno 30 studenti.

Non ci sono differenze tra i risultati ottenuti dal modello relazionale denormalizzato e il modello razionale normalizzato.

```
SELECT cds.cds, ad.adcod, ad.ad, AVG(voto) as media_voto, count(*) AS numero_superamenti
FROM iscrizioni as isc JOIN appelli as app ON iscc.appcod = app.appcod
JOIN cds ON app.cdscod = cds.cdscod
JOIN ad ON app.adcod = ad.adcod
GROUP BY cds.cdscod, ad.ad
ORDER BY cds.cds, media_voto desc
```

```
SELECT cds, ad, AVG(voto) as media_voto, count(*) AS numero_superamenti
FROM bos_denormalizzato
GROUP BY cdscod, ad
ORDER BY cds, media_voto desc
```



# Query 4: commento

Usando il primo metodo di giudizio la top-3 degli esami più facili di giurisprudenza risulta:

[A5810234] LA CORTE DI GIUSTIZIA: TECNICHE E S...

[A5810172] INTERNATIONAL CRIMINAL JUSTICE

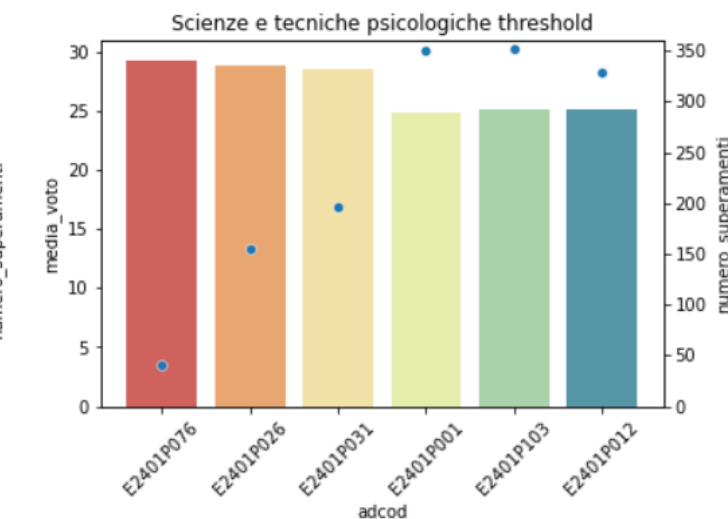
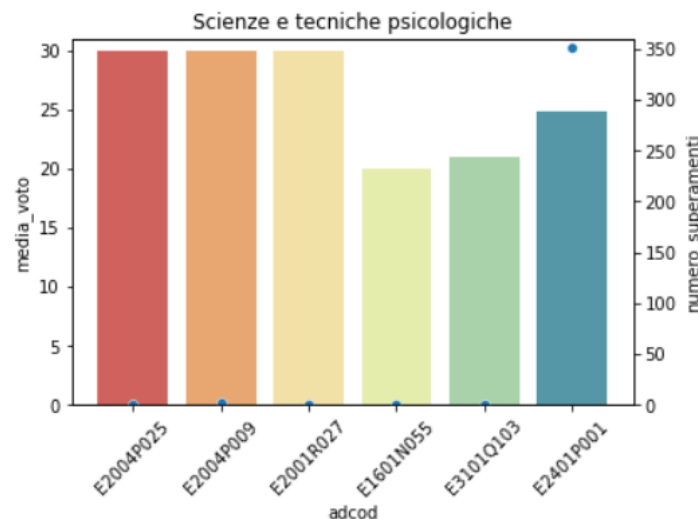
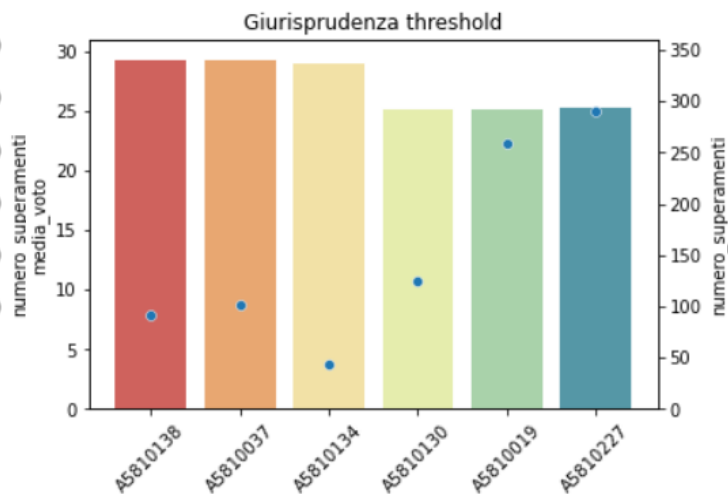
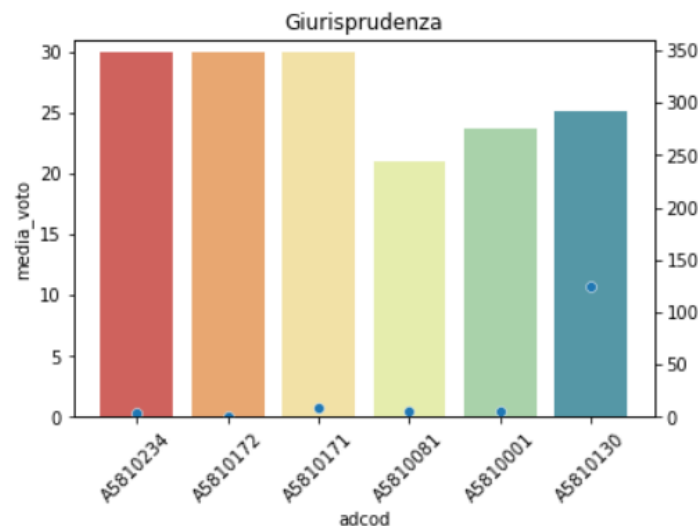
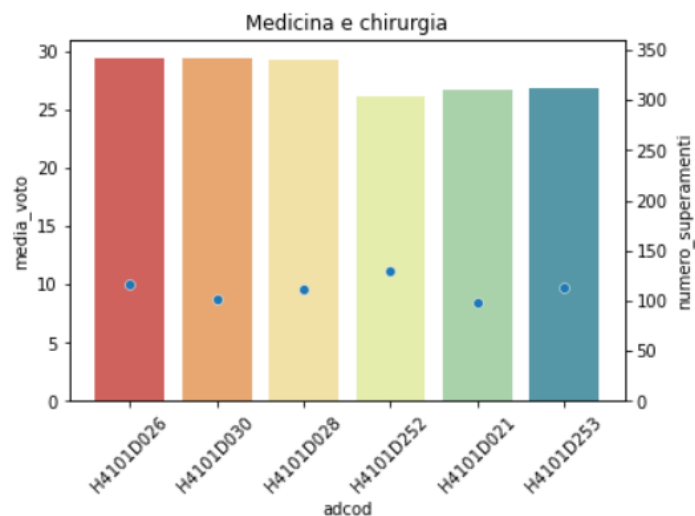
[A5810171] LAW OF BRICS COUNTRIES AND EU

Eliminando le attività didattiche che non sono state superate da almeno 30 studenti la top-3 diventa:

[A5810138] SOCIOLOGIA DEI DIRITTI FONDAMENTALI

[A5810037] STORIA DEL DIRITTO ROMANO

[A5810134] FONDAMENTI ROMANISTICI DEL DIRITTO ...



In alcuni casi il risultato non viene modificato. Un esempio è Medicina e chirurgia, in cui la top-3 degli esami più facili e la top-3 degli esami più difficili è data da esami che sono stati superati da più di 30 studenti.

# 5. Fast & Furious

Calcoliamo la distribuzione degli studenti secondo l'indice Fast & Furious. *Dividiamo* gli studenti in base al numero di esami sostenuti che impattano sulla media (quelli che hanno un voto). Valutiamo infine chi sono i *migliori studenti*, cioè chi ha superato *più esami*, con la *media maggiore* e nel *minor tempo* possibile.

Definiamo chi sono i migliori studenti:

**Studenti Top**

**Fast&Furious  
Alto**

e

**N° esami con  
valutazione  
alti**

$\frac{\text{voto medio}}{\text{periodo attività}}$

Superamento = 1  
AND Voto IS NOT  
Null

Ultimo appello –  
primo appello (gg)

Alcuni esami non  
hanno votazione (es:  
seminari o laboratori)

Normalizzato

```
SELECT a.CdScod,
a.Studente,
a.n°esami superati,
coalesce(b.n°esami superati senza Voto, 0) AS n°esami superati senza Voto,
abs(a.n°esami superati - coalesce(b.n°esami superati senza Voto, 0)) AS n°esami per media,
a.voto_medio,
coalesce(c.n°esami non superati, 0) AS n°esami non superati,
d.primo_appello,
d.ultimo_appello,
d.periodo_attività,
round(a.voto_medio / d.periodo_attività, 4) AS FastAndFurious
FROM (-- quanti superati e con che voto medio
SELECT CdScod,
st.*,
count() AS n°esami superati,
round(avg(Voto), 2) AS voto_medio
FROM
FROM iscrizioni AS iscr
JOIN
appelli AS app ON app.appcod = iscr.appcod
JOIN
studenti AS st ON st.studente = iscr.studente
WHERE superamento = 1
GROUP BY st.studente
) a
LEFT JOIN
(-- Quanti ma senza voto
SELECT CdScod,
st.*,
count() AS n°esami superati senza voto
FROM
FROM iscrizioni AS iscr
JOIN
appelli AS app ON app.appcod = iscr.appcod
JOIN
studenti AS st ON st.studente = iscr.studente
WHERE superamento = 1 AND
voto IS NULL
GROUP BY st.studente
) b ON a.Studente = b.Studente
LEFT JOIN
(-- quanti non superati
SELECT CdScod,
st.*,
count() AS n°esami non superati
FROM
FROM iscrizioni AS iscr
JOIN
appelli AS app ON app.appcod = iscr.appcod
JOIN
studenti AS st ON st.studente = iscr.studente
WHERE Iscrizione = 1 AND
superamento = 0
GROUP BY st.studente
) c ON a.Studente = c.Studente
LEFT JOIN
(-- Periodo di attività
SELECT st.*,
count(),
max(dtappello) AS ultimo_appello,
min(dtappello) AS primo_appello,
(julianday(max(dtappello)) - julianday(min(dtappello))) AS periodo_attività
FROM
FROM iscrizioni AS iscr
JOIN
appelli AS app ON app.appcod = iscr.appcod
JOIN
studenti AS st ON st.studente = iscr.studente
GROUP BY st.studente
) d ON d.studente = a.studente
ORDER BY n°esami per media ASC,
FastAndFurious DESC;
```

SQL

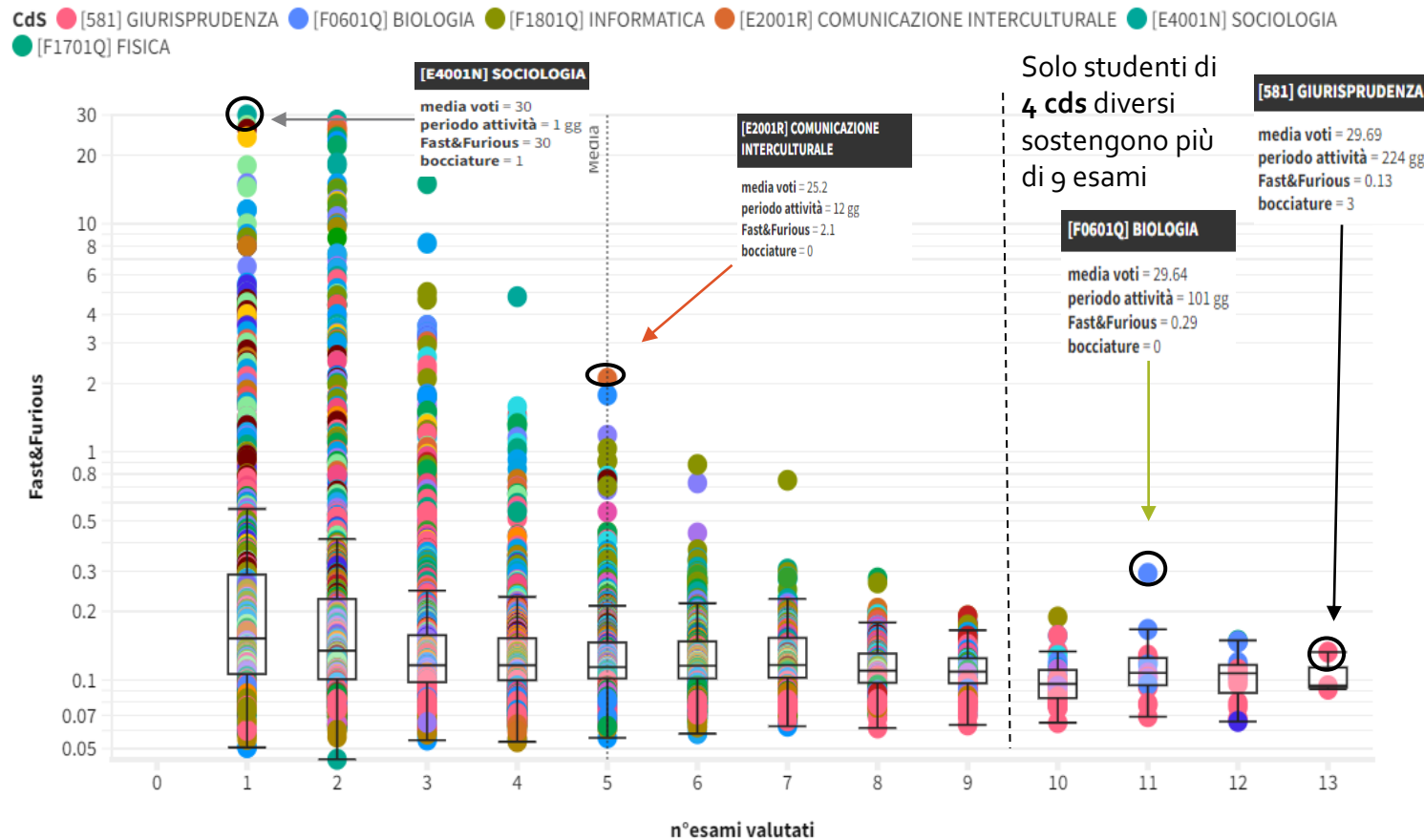
DeNormalizzato

```
SELECT CdS,
Studente,
a.n°esami superati,
coalesce(b.n°esami superati senza Voto, 0) AS n°esami superati senza Voto,
abs(a.n°esami superati - coalesce(b.n°esami superati senza Voto, 0)) AS n°esami per media,
a.voto_medio,
coalesce(c.n°esami non superati, 0) AS n°esami non superati,
d.primo_appello,
d.ultimo_appello,
d.periodo_attività,
round(a.voto_medio / d.periodo_attività, 4) AS FastAndFurious,
a.n°esami rifiutati
FROM (-- quanti superati e con che voto medio
SELECT CdS,
Studente,
count() AS n°esami superati,
round(avg(Voto), 2) AS voto_medio,
(count() - count(DISTINCT AdCod)) AS n°esami rifiutati
FROM bos_denormalizzato
WHERE Superamento = 1
GROUP BY Studente
) a
LEFT JOIN
(-- quanti superati ma senza voto
SELECT
Studente,
count() AS n°esami superati senza Voto
FROM bos_denormalizzato
WHERE Superamento = 1 AND
Voto IS NULL
GROUP BY Studente
ORDER BY Studente ASC
) b ON a.Studente = b.Studente
LEFT JOIN
(-- quanti non superati
SELECT Studente,
count() AS n°esami non superati
FROM bos_denormalizzato
WHERE Iscrizione = 1 AND
Superamento = 0
GROUP BY Studente
) c ON a.Studente = c.Studente
LEFT JOIN
(
SELECT studente,
count(),
max(dtappello) AS ultimo_appello,
min(dtappello) AS primo_appello,
(julianday(max(dtappello)) - julianday(min(dtappello))) AS periodo_attività
FROM bos_denormalizzato
GROUP BY studente
) d ON d.studente = a.studente
ORDER BY n°esami per media ASC,
FastAndFurious desc;
```

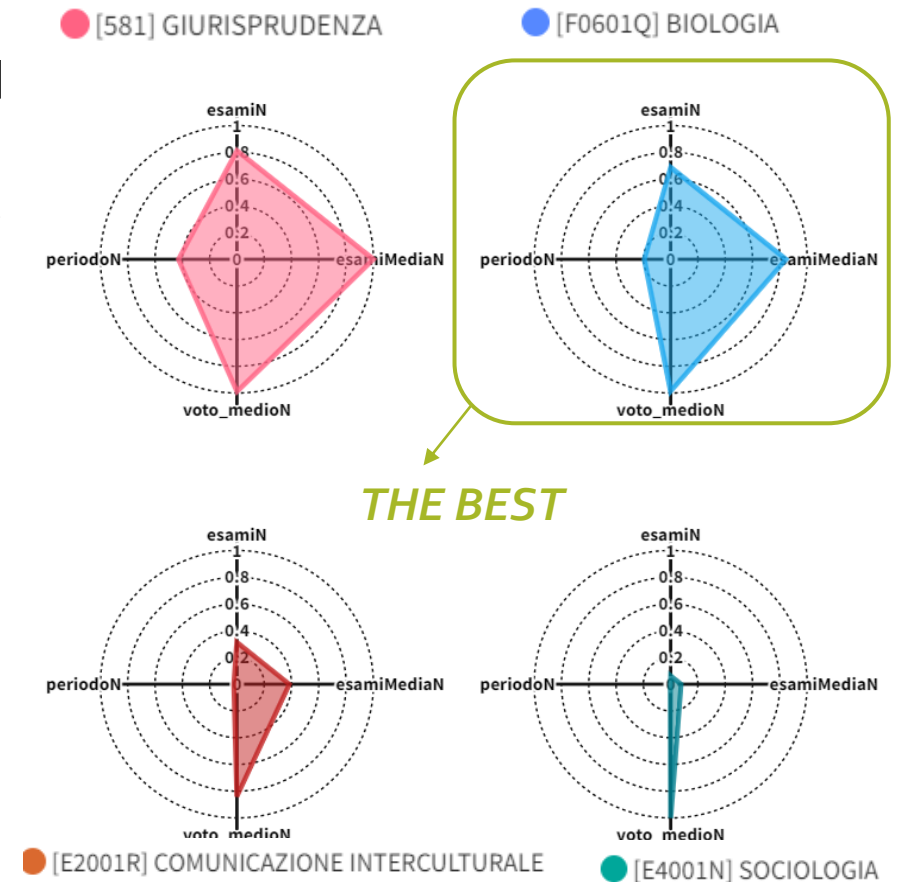
# Query 5: Top studenti fast & furious

Di seguito viene proposta la **distribuzione degli studenti** (sinistra) per cds (colori), n°esami con valutazione (asse x) e indice Fast&Furious (asse y). A destra confrontiamo nel dettaglio (**valori normalizzati**) i più alti indici Fast&Furious dei *due studenti* che hanno sostenuto *più esami* (giurisprudenza e biologia), dello *studente* che rispetta la *media degli esami sostenuti* (comunicazione interculturale), dello *studente* con il *minor numero di esami* valutati (sociologia).

Distribuzione studenti



Confronto con valori normalizzati



## 6: Trial & Error - Dati

Abbiamo raggruppato per 'adcod' e 'studente', al fine di ottenere il conteggio dei non superamenti (*superamento=0*) di ogni studente nei singoli esami. Successivamente, è stata calcolata la colonna 'n°medio\_non\_superamento' come la media del conteggio per avere come risultato la media del numero di tentativi degli studenti in ogni esame, ovvero, il valore «*Trial & Error*» di quest'ultimo. Non ci sono differenze nei risultati ottenuti usando il modello relazionale denormalizzato e il modello relazionale normalizzato.

### Normalizzato

```
SELECT d.ad,
       avg(d.conteggio) AS media
FROM (
  SELECT ad.ad,
         a.dtappello,
         count(i.superamento = 0) AS conteggio
  FROM iscrizioni AS i
  JOIN appelli AS a ON i.appcod = a.appcod
  JOIN ad ON a.adcod = ad.adcod
  WHERE i.superamento = 0
  GROUP BY ad.adcod, i.studente
  ORDER BY conteggio DESC
) d
GROUP BY d.ad
ORDER BY media DESC;
```

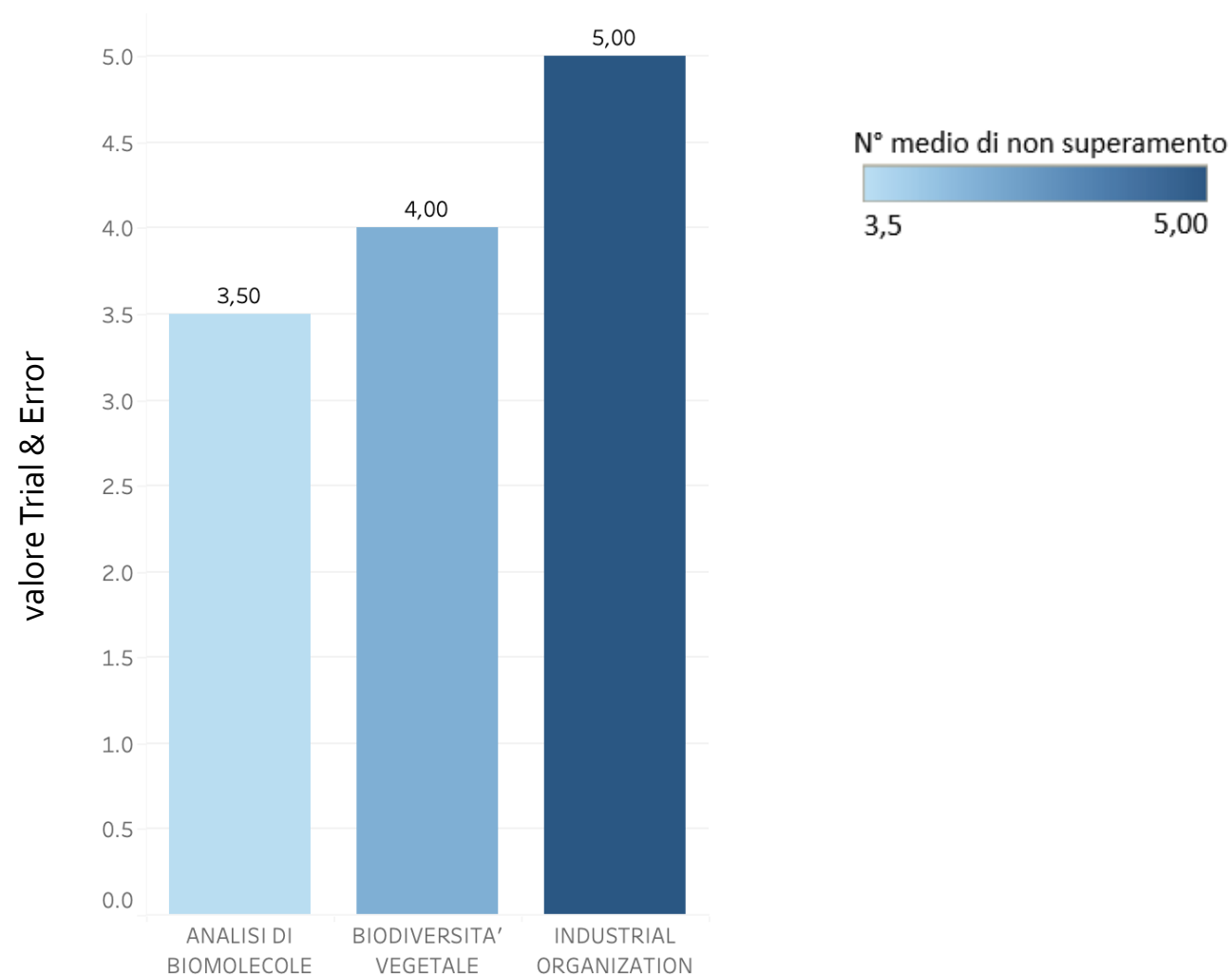
	ad	n°medio non superamento
1	[F5602M014] INDUSTRIAL ORGANIZATION	5
2	[F0601Q058] BIODIVERSITA' VEGETALE	4
3	[F5401Q044] ANALISI DI BIOMOLECOLE	3.5
4	[E3002Q030] STORIA E SVILUPPI DEGLI STRUMENTI OTTICI	3.33333333333333
5	[F0802Q066] CHIMICA ORGANICA FARMACEUTICA	3
6	[E2701Q039] MATEMATICA III	3
7	[E1401A039] DIRITTO DELL'ESECUZIONE CIVILE	3
8	[A5810071] ORGANIZZAZIONE INTERNAZIONALE	3
9	[E2702Q089] CHIMICA FISICA II E LABORATORIO	2.86206896551724
10	[E3002Q006] FISICA I	2.84
11	[E1901R090] GEOGRAFIA CULTURALE	2.8
12	[E3002Q033] OTTICA GEOMETRICA E OFTALMICA CON LABORATORIO	2.67647058823529
13	[E3002Q032] ANATOMIA E ISTOLOGIA UMANA E OCULARE	2.61538461538462
14	[F4901N003] COMUNICAZIONE PER IL TURISMO	2.5
15	[F3301M1831] LINGUA- TEDESCO	2.5

### Denormalizzato

```
SELECT d.ad,
       avg(d.conteggio) AS n°medio_non_superamento
FROM (
  SELECT b.ad,
         b.dtappello,
         count(b.superamento = 0) AS conteggio
  FROM bos_denormalizzato AS b
  WHERE b.superamento = 0
  GROUP BY b.adcod,
           b.studente
  ORDER BY conteggio DESC
) d
GROUP BY d.ad
ORDER BY n°medio_non_superamento DESC;
```

## Query 6: Top 3 esami Trial & Error

La rappresentazione grafica mostra i 3 esami, sull'asse x, con maggiore «*Trial & Error*», in particolare al primo posto abbiamo l'esame **Industrial Organization** che risulta avere la media di non superamenti pari a 5, seguito da **Biodiversità Vegetale** con 4 e **Analisi di Biomolecole** con 3,5.



# 7:Confronto maschi e femmine nei corsi di laurea

Abbiamo raggruppato per 'genere' e 'CdS', al fine di potere calcolare il voto medio dei maschi e delle femmine nei singoli corsi di laurea, e contare il numero di esami superati dai maschi e dalle femmine in ogni CdS.

A questo punto abbiamo potuto confrontare l'andamento dei maschi e delle femmine in termini di *votazione media*.

	genere	CdS	votazione media	esami superati
1	F	[524] SCIENZE DELL'EDUCAZIONE	25	1
2	F	[541] SCIENZE DEL TURISMO E COMUNITA' LOCALE(C.SO A DISTANZA)	25	1
3	F	[581] GIURISPRUDENZA	27.07021604938272	4337
4	M	[581] GIURISPRUDENZA	26.79804028307022	1992
5	F	[E0201Q] BIOTECNOLOGIE	26.59722222222222	1440
6	M	[E0201Q] BIOTECNOLOGIE	26.20260663507109	844
7	M	[E1301Q] SCIENZE BIOLOGICHE	26.43435980551054	617
8	F	[E1301Q] SCIENZE BIOLOGICHE	26.29851909586906	1283
9	F	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	25.57509727626459	1523
10	M	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	24.89928057553957	837
11	F	[E1501N] SCIENZE DEL TURISMO E COMUNITÀ LOCALE	25.33135215453195	2019
12	M	[E1501N] SCIENZE DEL TURISMO E COMUNITÀ LOCALE	24.3274647887324	568
13	F	[E1502N] SCIENZE DEL TURISMO E COMUNITÀ LOCALE (NETTUNO)	19	1
14	F	[E1601N] SCIENZE DELL'ORGANIZZAZIONE	24.79898218829517	1572
15	M	[E1601N] SCIENZE DELL'ORGANIZZAZIONE	24.28792912513843	903

## Normalizzato

```
SELECT s.genere,
       c.CdS,
       avg(i.voto) AS votazione_media,
       count(i.superamento = 1) AS esami_superati
FROM studenti AS s
JOIN
iscrizioni AS i ON s.studente = i.studente
JOIN
appelli AS a ON i.appcod = a.appcod
JOIN
cds AS c ON a.cdscod = c.cdscod
WHERE i.superamento = 1
GROUP BY s.genere,
         c.CdS
ORDER BY c.CdS,
         votazione_media DESC;
```

## Denormalizzato

```
SELECT b.StuGen,
       b.CdS,
       avg(b.voto) AS votazione_media,
       count(b.superamento = 1) AS esami_superati
FROM bos_denormalizzato AS b
WHERE b.superamento = 1
GROUP BY b.StuGen,
         b.CdS
ORDER BY b.CdS,
         votazione_media DESC;
```



## Query 7: Confronto andamento tra maschi e femmine nei diversi corsi di laurea

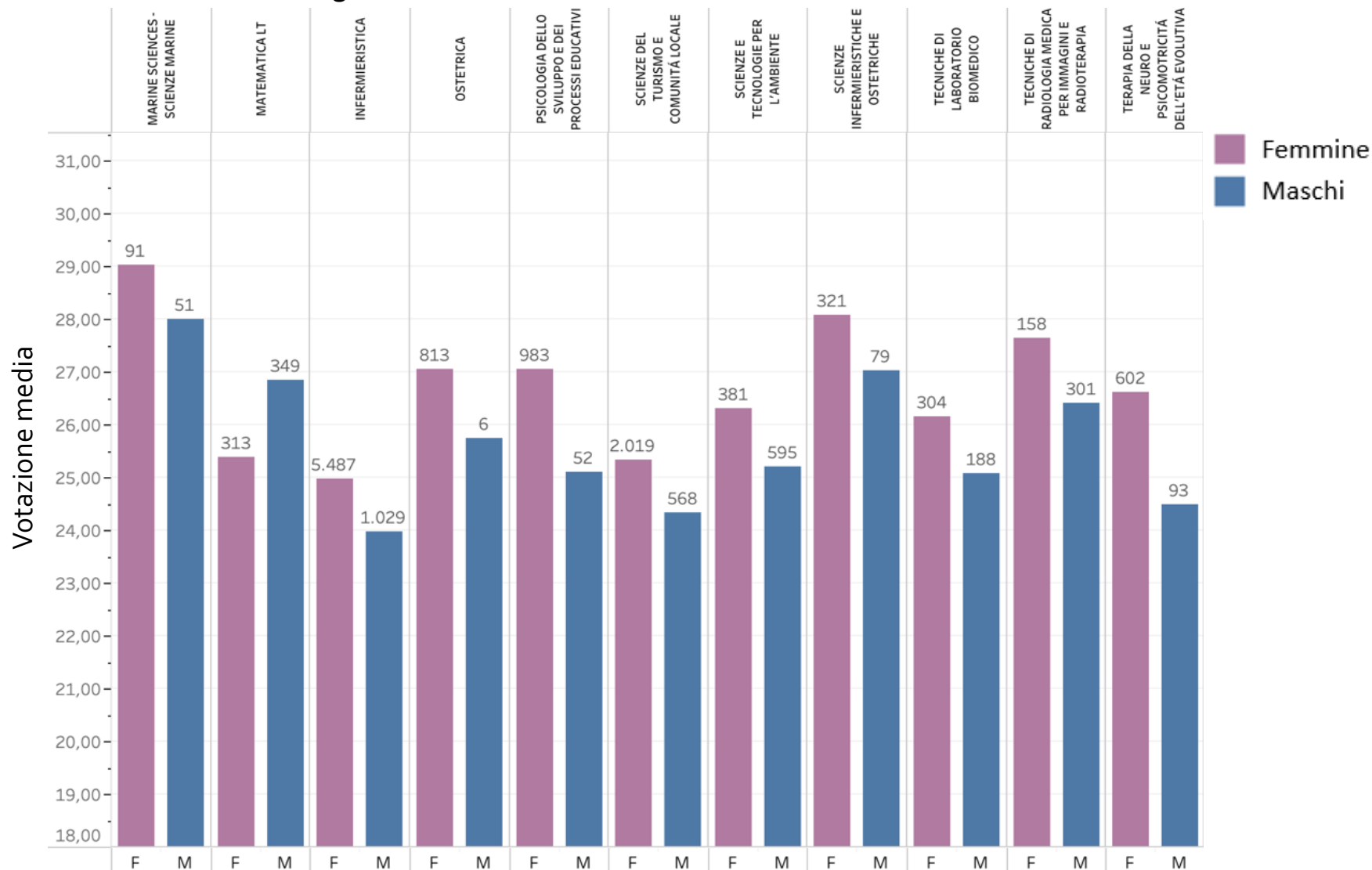
I due colori, **viola** e **blu**, indicano rispettivamente i due generi femmine e maschi.

Gli indicatori vengono etichettati con il numero di esami superati. Sono stati mostrati i dati che presentano una differenza del voto medio tra i due generi di almeno un voto.

Per i corsi di laurea osservati risulta esserci un migliore andamento delle **femmine** in quasi tutti i corsi di laurea, ad eccezione del corso *Matematica LT*.

Generalmente, ad una votazione media più alta corrisponde anche un numero maggiore di esami superati.

In particolare, i corsi con una più grande differenza di voto medio sono *Terapia della neuro e psicomotricità dell'età evolutiva*(+2,15) e *Psicologia dello sviluppo e dei processi educativi*(+1,96).



# 7:Confronto Area di Residenza nei corsi di laurea

L'obbiettivo è quello di valutare *l'impatto che l'area di residenza ha sulla votazione media e sul numero di esami superati mediamente* dagli studenti. Dapprima è stata calcolata la media voti e il numero di esami passati da ogni studente e successivamente ne è stata calcolata la media raggruppando per area di residenza.

## Normalizzato

```
select a.cdscod, b.resarea, round(avg(a.media_voti), 2) as voto_medio, round(avg(b.n°esami_passati),2) as esami_passati_medi
from
  (/*Per ogni studente calcoliamo il n° di esami sostenuti e la relativa media*/
  select st.*, app.cdscod, app.adcod, count() as n°esami_sostenuti, avg(voto) as media_voti
  from iscrizioni as iscr
  join appelli as app on iscr.appcod = app.appcod
  join studenti as st on iscr.studente = st.studente
  join cds on cds.cdscod = app.cdscod
  group by st.studente
  ) a JOIN ( /*Per ogni studente calcoliamo il n° di esami passati*/
  select st.*, app.cdscod, app.adcod, count() as n°esami_passati
  from iscrizioni as iscr
  join appelli as app on iscr.appcod = app.appcod
  join studenti as st on iscr.studente = st.studente
  where superamento = 1
  group by st.studente
  ) b on a.studente= b.studente|
group by a.cdscod, b.resarea
order by a.cdscod, b.resarea;
```

	cdscod	cds	StuResArea	voto medio	esami passati medi
1	524	[524] SCIENZE DELL'EDUCAZIONE	REGIONE	25	1
2	541	[541] SCIENZE DEL TURISMO E COMUNITA' LOCALE(C.SO A DISTANZA)	REGIONE	25	1
3	581	[581] GIURISPRUDENZA	CITTA	26.4	5
4	581	[581] GIURISPRUDENZA	EXTRA-REGIONE	26.32	4
5	581	[581] GIURISPRUDENZA	PROVINCIA	26.45	5
6	581	[581] GIURISPRUDENZA	REGIONE	26.67	5
7	E0201Q	[E0201Q] BIOTECNOLOGIE	CITTA	25.91	5
8	E0201Q	[E0201Q] BIOTECNOLOGIE	EXTRA-REGIONE	25.9	5
9	E0201Q	[E0201Q] BIOTECNOLOGIE	PROVINCIA	25.83	5
10	E0201Q	[E0201Q] BIOTECNOLOGIE	REGIONE	26.27	6
11	E1301Q	[E1301Q] SCIENZE BIOLOGICHE	CITTA	25.48	4
12	E1301Q	[E1301Q] SCIENZE BIOLOGICHE	EXTRA-REGIONE	25.26	4
13	E1301Q	[E1301Q] SCIENZE BIOLOGICHE	PROVINCIA	25.59	4
14	E1301Q	[E1301Q] SCIENZE BIOLOGICHE	REGIONE	26	4
15	E1401A	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	CITTA	24.52	4
16	E1401A	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	EXTRA-REGIONE	25.05	5
17	E1401A	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	PROVINCIA	25.29	5
18	E1401A	[E1401A] SCIENZE DEI SERVIZI GIURIDICI	REGIONE	25	5

## Denormalizzato

```
select a.cdscod, a.cds, b.StuResArea, round(avg(a.media_voti), 2) as voto_medio, round(avg(b.n°esami_passati),0) as esami_passati_medi
from
  (/*Per ogni studente calcoliamo il n° di esami sostenuti e la relativa media*/
  select studente, StuResArea, cdscod, cds, adcod, count() as n°esami_sostenuti, avg(voto) as media_voti
  from bos_denormalizzato
  group by studente
  ) a JOIN ( /*Per ogni studente calcoliamo il n° di esami passati*/
  select studente, StuResArea, cdscod, cds, adcod, count() as n°esami_passati
  from bos_denormalizzato
  where superamento = 1
  group by studente
  ) b on a.studente= b.studente
group by a.cdscod, b.StuResArea
order by a.cdscod, b.StuResArea;
```

## Query 7: Confronto andamento tra maschi e femmine nei diversi corsi di laurea

Il grafico rappresenta i dieci corsi di laurea con le differenze, tra le medie di voti per area di residenza, più elevate. Vengono indicati con appositi label anche il numero di esami medi passati per ogni area di residenza.

Come si può notare non si rilevano delle elevate differenze tra area di residenza. Possiamo quindi dire che l'area di residenza non sembra influenzare la votazione media.

Notiamo che per il corso di laurea 'TECNICHE DI LABORATORIO BIOMEDICO' gli studenti di 'CITTA' superano quasi il doppio degli esami rispetto a studenti di altre aree ma con un voto nettamente più basso

