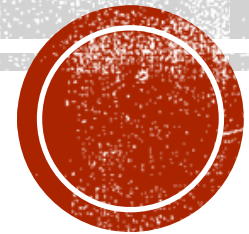


BUSINESS INTELLIGENCE

A.A. 2023-2024



Gianni Eduard Balbin Canchanya - 901609

Alberto Invernizzi - 901611

QUERY 1

Distribuzione del numero degli studenti iscritti nei vari appelli, suddivisa per anni e per corso di laurea

```
SELECT appelli.dtappello,  
       appelli.adcod,  
       ad.ad,  
       appelli.annoAcc,  
       count(studente) AS totstudenti,  
       appelli.appcod  
FROM appelli  
     JOIN iscrizioni ON appelli.appcod = iscrizioni.appcod  
     JOIN ad ON appelli.adcod = ad.adcod  
WHERE dtappello IS NOT NULL  
GROUP BY appelli.appcod  
ORDER BY appelli.dtappello;
```

```
SELECT DtAppello,  
       AdCod,  
       AD,  
       annoAcc,  
       count(studente) AS totstudenti  
FROM bos_denormalizzato  
GROUP BY DtAppello,  
         AdCod  
ORDER BY annoAcc ASC;
```



	dtappello	adcod	ad	annoAcc	totstudenti	appcod
1	2015-09-17	E2401P013	[E2401P013] PSICOLOGIA DINAMICA	2014/2015	1	3459
2	2016-01-20	F5102P015	[F5102P015] PSICOLOGIA DELLE INFLUENZE SOCIALI	2015/2016	1	6860
3	2016-01-25	E2401P103	[E2401P103] PSICOPATOLOGIA GENERALE	2015/2016	1	3605
4	2016-01-25	F5102P016	[F5102P016] ATTEGGIAMENTI E OPINIONI	2015/2016	1	6865
5	2016-01-25	F5102P019	[F5102P019] PSICOLOGIA DELLE CONDOTTE FINANZIARIE	2015/2016	1	6877
6	2016-02-24	F5102P012	[F5102P012] COMUNICAZIONE NELLE ORGANIZZAZIONE E COMUNICAZIONE SOCIALE	2015/2016	1	6847
7	2016-02-26	F5101P139	[F5101P139] PSICOLOGIA DELLA SALUTE E INTERVENTI CLINICI IN AMBITO SANITARIO	2015/2016	1	6774
8	2016-04-26	E2702Q099	[E2702Q099] CHIMICA FISICA APPLICATA	2015/2016	1	3917
9	2016-05-31	E2702Q097	[E2702Q097] CHIMICA FISICA III E LABORATORIO	2015/2016	1	3909
10	2016-06-10	F5102P004	[F5102P004] COGNIZIONE SOCIALE	2015/2016	2	6828

	DtAppello	AdCod	AD	annoAcc	totstudenti
1	2015-09-17	E2401P013	[E2401P013] PSICOLOGIA DINAMICA	2014/2015	1
2	2016-01-08	F5102P019	[F5102P019] PSICOLOGIA DELLE CONDOTTE FINANZIARIE	2015/2016	1
3	2016-01-20	F5102P015	[F5102P015] PSICOLOGIA DELLE INFLUENZE SOCIALI	2015/2016	1
4	2016-01-25	E2401P103	[E2401P103] PSICOPATOLOGIA GENERALE	2015/2016	1
5	2016-01-25	F5102P016	[F5102P016] ATTEGGIAMENTI E OPINIONI	2015/2016	1
6	2016-01-25	F5102P019	[F5102P019] PSICOLOGIA DELLE CONDOTTE FINANZIARIE	2015/2016	1
7	2016-02-05	F5103P002	[F5103P002] METODI DI RICERCA E VALUTAZIONE IN PSICOLOGIA DELLO SVILUPPO	2015/2016	1
8	2016-02-24	F5102P012	[F5102P012] COMUNICAZIONE NELLE ORGANIZZAZIONE E COMUNICAZIONE SOCIALE	2015/2016	1
9	2016-02-26	F5101P139	[F5101P139] PSICOLOGIA DELLA SALUTE E INTERVENTI CLINICI IN AMBITO SANITARIO	2015/2016	1
10	2016-04-26	E2702Q099	[E2702Q099] CHIMICA FISICA APPLICATA	2015/2016	1

For this query, we have decided to add an attribute in both *appelli* and *bos_denormalizzato*. Specifically, we created the academic year column (annoAcc), and based on a condition, a certain value was assigned. For instance, if 'dtappello' is between '2014-09-25' and '2015-09-24,' the value of 'annoAcc' will be '2014/2015,' and so on.

No significant differences were observed, except that in *appelli*, there are some NULL values in the 'dtappello' attribute, which is not found in *bos_denormalizzato*.

Therefore, we are not surprised if *bos_denormalizzato* returns more observations (293 observations).



QUERY 2

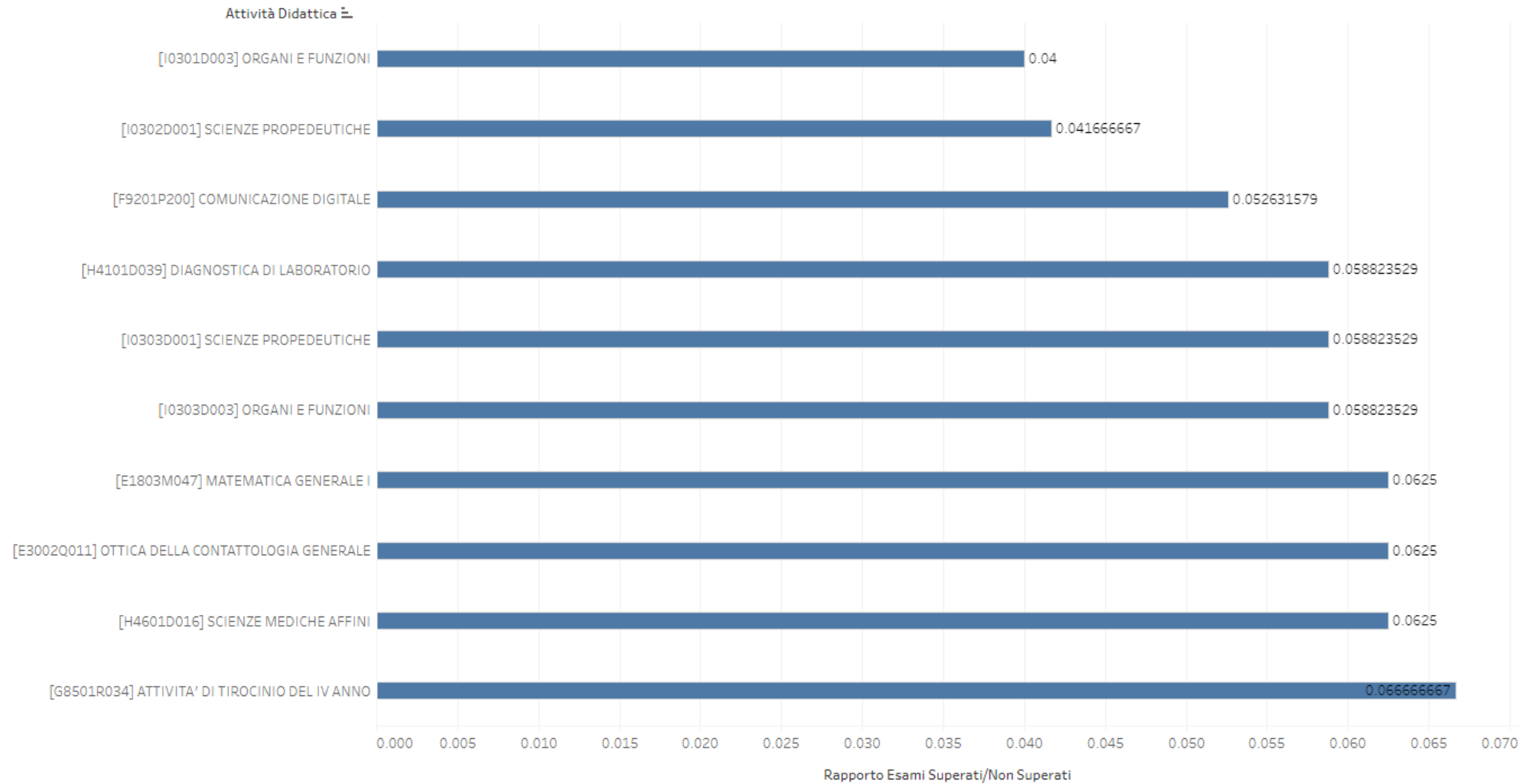
Individuazione della Top-10 degli esami più difficili suddivisi per corso di studi.

```
SELECT t1.appcod,
       t1.ad,
       t1.totSup,
       t2.totNoSup,
       (t1.totSup * 1.0 / (t1.totSup + t2.totNoSup) ) AS ratio
FROM (
  SELECT appelli.appcod,
         ad.ad,
         count(appelli.appcod) AS totSup
  FROM iscrizioni
  JOIN appelli ON iscrizioni.appcod = appelli.appcod
  JOIN ad ON appelli.adcod = ad.adcod
  WHERE assenza = 0 AND Superamento = 1
  GROUP BY appelli.appcod
) AS t1
JOIN
(
  SELECT appelli.appcod,
         ad.ad,
         count(appelli.appcod) AS totNoSup
  FROM iscrizioni
  JOIN appelli ON iscrizioni.appcod = appelli.appcod
  JOIN ad ON appelli.adcod = ad.adcod
  WHERE assenza = 0 AND Superamento = 0
  GROUP BY appelli.appcod
) AS t2 ON t1.appcod = t2.appcod
ORDER BY ratio ASC
LIMIT 10;
```

```
SELECT t1.DtAppello,
       t1.AD,
       t1.totSup,
       t2.totNoSup,
       (t1.totSup * 1.0 / (t1.totSup + t2.totNoSup) ) AS ratio
FROM (
  SELECT AD,
         DtAppello,
         Count(DtAppello) AS totSup
  FROM bos_denormalizzato
  WHERE Assenza = 0 AND Superamento = 1
  GROUP BY DtAppello, AD
  ORDER BY AD ASC
) AS t1
JOIN
(
  SELECT AD,
         DtAppello,
         Count(DtAppello) AS totNoSup
  FROM bos_denormalizzato
  WHERE Assenza = 0 AND Superamento = 0
  GROUP BY DtAppello, AD
  ORDER BY AD ASC
) AS t2 ON t1.DtAppello = t2.DtAppello AND t1.AD = t2.AD
ORDER BY ratio ASC
LIMIT 10;
```



Query 2



There are no relevant differences for both queries.



QUERY 3

Individuazione dei corsi di laurea ad elevato tasso di commitment, ovvero appelli di esami diversi ma del medesimo corso di laurea che si sono svolti nello stesso giorno

```
SELECT cds.cds,  
       appelli.dtappello,  
       count(cds.cds) AS esamiSvolti  
FROM appelli  
JOIN  
     cds ON appelli.cdscod = cds.cdscod  
WHERE appelli.dtappello IS NOT NULL  
GROUP BY appelli.dtappello,  
         appelli.cdscod  
ORDER BY esamiSvolti DESC;
```

```
SELECT Cds,  
       DtAppello,  
       count(DISTINCT AD) AS  
EsamiSvolti  
FROM bos_denormalizzato  
GROUP BY DtAppello,  
         Cds  
ORDER BY EsamiSvolti DESC;
```



	cds	dtappello	esamiSvolti
1	[H4101D] MEDICINA E CHIRURGIA	2017-03-31	42
2	[581] GIURISPRUDENZA	2017-01-25	24
3	[581] GIURISPRUDENZA	2016-11-30	23
4	[581] GIURISPRUDENZA	2017-09-27	23
5	[581] GIURISPRUDENZA	2016-11-28	19
6	[581] GIURISPRUDENZA	2017-07-19	19
7	[581] GIURISPRUDENZA	2017-09-11	19
8	[581] GIURISPRUDENZA	2017-11-27	19
9	[581] GIURISPRUDENZA	2017-11-29	18
10	[I0101D] INFERMIERISTICA (ABILITANTE ALLA PROFESSIONE SANITARIA DI INFERMIERE)	2017-09-22	17

	Cds	DtAppello	EsamiSvolti
1	[H4101D] MEDICINA E CHIRURGIA	2017-03-31	42
2	[581] GIURISPRUDENZA	2017-05-03	30
3	[581] GIURISPRUDENZA	2017-01-25	24
4	[581] GIURISPRUDENZA	2016-11-30	22
5	[581] GIURISPRUDENZA	2017-05-04	21
6	[581] GIURISPRUDENZA	2017-09-27	21
7	[581] GIURISPRUDENZA	2016-11-28	19
8	[581] GIURISPRUDENZA	2017-02-08	19
9	[581] GIURISPRUDENZA	2017-09-11	18
10	[581] GIURISPRUDENZA	2017-11-27	18

As with query 1, for this query, we also observed a slight difference in 'esamiSvolti', probably due to the fact that in the first query, we excluded NULL values from 'dtappello'.



QUERY 4

Individuazione della Top-3 degli esami con media voti maggiore e minore rispettivamente, calcolati per ogni singolo corso di studi

```
SELECT DISTINCT t4.ad, t4.cds, t4.media
FROM (
  SELECT t3.ad, t3.cds, t3.media
  FROM (
    SELECT *, row_number() OVER (PARTITION BY t2.cds ORDER BY t2.cds) AS row_num
    FROM (
      SELECT *
      FROM ( SELECT ad.ad, cds.cds, avg(iscrizioni.Voto) AS media
            FROM iscrizioni
            JOIN appelli ON iscrizioni.appcod = appelli.appcod
            JOIN ad ON appelli.adcod = ad.adcod
            JOIN cds ON appelli.cdscod = cds.cdscod
            WHERE iscrizioni.Voto IS NOT NULL
            GROUP BY ad.ad, cds.cds
            ORDER BY media ASC
          ) AS t1 ORDER BY t1.cds
        ) AS t2 ) AS t3
    WHERE t3.row_num <= 3
  UNION ALL
  SELECT t3.ad, t3.cds, t3.media
  FROM (
    SELECT *, row_number() OVER (PARTITION BY t2.cds ORDER BY t2.cds DESC) AS row_num
    FROM (
      SELECT *
      FROM ( SELECT ad.ad, cds.cds, avg(iscrizioni.Voto) AS media
            FROM iscrizioni
            JOIN appelli ON iscrizioni.appcod = appelli.appcod
            JOIN ad ON appelli.adcod = ad.adcod
            JOIN cds ON appelli.cdscod = cds.cdscod
            WHERE iscrizioni.Voto IS NOT NULL
            GROUP BY ad.ad, cds.cds
            ORDER BY media DESC
          ) AS t1 ORDER BY t1.cds
        ) AS t2 ) AS t3
    WHERE t3.row_num <= 3
  ORDER BY t3.cds
) AS t4
ORDER BY t4.cds, t4.media;
```

```
SELECT DISTINCT t4.AD, t4.CdS, t4.media
FROM (
  SELECT t3.AD, t3.CdS, t3.media
  FROM (
    SELECT *, row_number() OVER (PARTITION BY t2.CdS ORDER BY t2.CdS) AS row_num
    FROM (
      SELECT *
      FROM ( SELECT AD, CdS, avg(Voto) AS media
            FROM bos_denormalizzato
            WHERE Voto IS NOT NULL
            GROUP BY AD, Cds
            ORDER BY media ASC
          ) AS t1 ORDER BY t1.CdS
        ) AS t2 ) AS t3
    WHERE t3.row_num <= 3
  UNION ALL
  SELECT t3.AD, t3.CdS, t3.media
  FROM (
    SELECT *, row_number() OVER (PARTITION BY t2.CdS ORDER BY t2.CdS DESC) AS row_num
    FROM (
      SELECT *
      FROM ( SELECT AD, CdS, avg(Voto) AS media
            FROM bos_denormalizzato
            WHERE Voto IS NOT NULL
            GROUP BY AD, Cds
            ORDER BY media DESC
          ) AS t1 ORDER BY t1.CdS
        ) AS t2 ) AS t3
    WHERE t3.row_num <= 3
  ORDER BY t3.CdS
) AS t4
ORDER BY t4.CdS, t4.media;
```



Query 4

CorsoDiStudi	Attività Didattica	
[524] SCIENZE DELL'EDUCAZIONE	[E1901R086] METODI E TECNICHE DELLA FORMAZIONE MEDIALE	25,00
[541] SCIENZE DEL TURISMO E COMUNITA' LOCALE(C.SO A DISTANZA)	[E1501N009] SOCIOLOGIA DELLA CULTURA E DEL TEMPO LIBERO	25,00
[581] GIURISPRUDENZA	[A5810081] ECONOMIA AZIENDALE	21,00
	[A5810001] DIRITTO DEI CONSUMATORI	23,67
	[A5810130] DIRITTO DELL'UNIONE EUROPEA	25,10
	[A5810139] DIRITTO AMMINISTRATIVO -PARTE SPECIALE	30,00
	[A5810169] EUROPEAN UNION COMMON AND SECURITY DEFENSE POLICY	30,00
	[A5810170] TRANSITIONAL JUSTICE AND STATE BUILDING	30,00
	[E0201Q050] METODOLOGIE BIOCHIMICHE E TECNOLOGIE BIOMOLECOLARI	22,63
	[E1301Q059] ECOLOGIA APPLICATA	23,00
	[E0201Q059] BIOCHIMICA PER LE BIOTECNOLOGIE	23,41

Query 4 Denormalizzato

CorsoDiStudi	Attività Didattica	
[524] SCIENZE DELL'EDUCAZIONE	[E1901R086] METODI E TECNICHE DELLA FORMAZIONE MEDIALE	25,00
[541] SCIENZE DEL TURISMO E COMUNITA' LOCALE(C.SO A DISTANZA)	[E1501N009] SOCIOLOGIA DELLA CULTURA E DEL TEMPO LIBERO	25,00
[581] GIURISPRUDENZA	[A5810081] ECONOMIA AZIENDALE	21,00
	[A5810001] DIRITTO DEI CONSUMATORI	23,67
	[A5810130] DIRITTO DELL'UNIONE EUROPEA	25,10
	[A5810139] DIRITTO AMMINISTRATIVO -PARTE SPECIALE	30,00
	[A5810169] EUROPEAN UNION COMMON AND SECURITY DEFENSE POLICY	30,00
	[A5810170] TRANSITIONAL JUSTICE AND STATE BUILDING	30,00
	[E0201Q050] METODOLOGIE BIOCHIMICHE E TECNOLOGIE BIOMOLECOLARI	22,63
	[E1301Q059] ECOLOGIA APPLICATA	23,00
	[E0201Q059] BIOCHIMICA PER LE BIOTECNOLOGIE	23,41

There are no relevant differences for both queries.



QUERY 5

Calcolare la distribuzione degli studenti “fast&furious” per corso di studi, ovvero studenti con il rapporto “votazione media riportata negli esami superati” su “periodo di attività” maggiore.

```
SELECT t3.cds, count(t3.cds) AS numStud
FROM (
  SELECT t1.*, t2.media,
    (t2.media / CASE WHEN t1.diffGiorni = 0 THEN 1 ELSE t1.diffGiorni END) AS ratio
  FROM (
    SELECT iscrizioni.studente, cds.cds,
      julianday(max(dtappello)) - julianday(min(dtappello)) AS diffGiorni
    FROM iscrizioni
    JOIN appelli ON iscrizioni.appcod = appelli.appcod
    JOIN cds ON appelli.cdscod = cds.cdscod
    WHERE appelli.dtappello IS NOT NULL
    GROUP BY iscrizioni.studente
    ORDER BY cds.cdscod, iscrizioni.studente, appelli.dtappello
  ) AS t1
  JOIN
    ( SELECT iscrizioni.studente, cds.cds, avg(iscrizioni.Voto) AS media
      FROM iscrizioni
      JOIN appelli ON iscrizioni.appcod = appelli.appcod
      JOIN cds ON appelli.cdscod = cds.cdscod
      WHERE appelli.dtappello IS NOT NULL AND iscrizioni.Voto IS NOT NULL
      GROUP BY iscrizioni.studente
      ORDER BY cds.cdscod, iscrizioni.studente, appelli.dtappello
    ) AS t2 ON t1.studente = t2.studente
  WHERE ratio >= 0.94350159714953
  ORDER BY ratio DESC
) AS t3
GROUP BY t3.cds;
```

```
SELECT t3.cds, count(t3.cds) AS numStud
FROM (
  SELECT t1.*, t2.media,
    (t2.media / CASE WHEN t1.diffGiorni = 0 THEN 1 ELSE t1.diffGiorni END) AS ratio
  FROM (
    SELECT Studente, CdS,
      julianday(max(DtAppello)) - julianday(min(DtAppello)) AS diffGiorni
    FROM bos_denormalizzato
    GROUP BY Studente
    ORDER BY CdS, Studente, DtAppello
  ) AS t1
  JOIN
    ( SELECT Studente, CdS, avg(Voto) AS media
      FROM bos_denormalizzato
      WHERE Voto IS NOT NULL
      GROUP BY Studente
      ORDER BY CdS, Studente, DtAppello
    ) AS t2 ON t1.studente = t2.studente
  WHERE ratio >= 0.73341637214748
  ORDER BY ratio DESC
) AS t3
GROUP BY t3.cds;
```



[illegible][illegible]

QUERY 6

Individuazione della Top-3 degli esami “trial&error”, ovvero esami che richiedono il maggior numero di tentativi prima del superamento.

```
SELECT t3.adcod, ad.ad, round(avg(t3.tot) ) AS TrialAndError
FROM ( SELECT *
      FROM ( SELECT *, count(t1.studente) AS tot
            FROM (
                  SELECT iscrizioni.studente,
                        appelli.appcod,
                        appelli.adcod,
                        iscrizioni.Superamento,
                        iscrizioni.Insufficienza,
                        iscrizioni.Ritiro,
                        iscrizioni.Voto
                  FROM iscrizioni
                  JOIN appelli ON iscrizioni.appcod = appelli.appcod
                  WHERE (iscrizioni.Superamento != 0 OR
                        iscrizioni.Insufficienza != 0 OR
                        iscrizioni.Ritiro != 0 OR
                        iscrizioni.Voto IS NOT NULL)
                  ORDER BY appelli.adcod,
                        iscrizioni.studente,
                        appelli.appcod DESC
            ) AS t1
      ) GROUP BY t1.studente, t1.adcod
    ) AS t2
  ORDER BY t2.adcod
) AS t3
JOIN ad ON t3.adcod = ad.adcod
GROUP BY t3.adcod
ORDER BY TrialAndError DESC, ad.ad
LIMIT 3;
```

```
SELECT t3.AdCod, AD, round(avg(t3.tot) ) AS TrialAndError
FROM ( SELECT *
      FROM ( SELECT *, count(t1.Studente) AS tot
            FROM (
                  SELECT Studente,
                        DtAppello,
                        AdCod,
                        AD,
                        Superamento,
                        Insufficienza,
                        Ritiro,
                        Voto
                  FROM bos_denormalizzato
                  WHERE (Superamento != 0 OR
                        Insufficienza != 0 OR
                        Ritiro != 0 OR
                        Voto IS NOT NULL)
                  ORDER BY AdCod,
                        Studente,
                        DtAppello DESC
            ) AS t1
      ) GROUP BY t1.Studente, t1.AdCod
    ) AS t2
  ORDER BY t2.AdCod
) as t3
GROUP BY t3.AdCod
ORDER BY TrialAndError DESC, AD
LIMIT 3;
```



Query 6



There are no relevant differences for both queries.



QUERY 7

Visualizza la media dei voti ottenuti dagli studenti iscritti agli esami per ciascuna attività didattica, considerando solo gli appelli con un tasso di superamento inferiore al 50%. Ordina i risultati in modo decrescente in base alla media dei voti.

```
SELECT t1.ad, t1.dtappello, t1.appcod, t1.Media_Voti,  
       round(CAST (t1.Pass AS REAL) / CAST (t1.Tot AS REAL), 2) AS Rapporto  
FROM (  
    SELECT ad.ad,  
           appelli.dtappello,  
           iscrizioni.appcod,  
           round(avg(iscrizioni.Voto), 2) AS Media_Voti,  
           sum(iscrizioni.Superamento) AS Pass,  
           sum(iscrizioni.Superamento + iscrizioni.Insufficienza + iscrizioni.Ritiro) AS Tot  
    FROM appelli  
    JOIN ad ON appelli.adcod = ad.adcod  
    JOIN iscrizioni ON appelli.appcod = iscrizioni.appcod  
    WHERE iscrizioni.Assenza = 0  
    GROUP BY iscrizioni.appcod  
    ORDER BY ad.ad  
) AS t1  
WHERE Rapporto < 0.5 AND Media_Voti >= 18  
ORDER BY t1.ad, Rapporto DESC;
```

```
SELECT *  
FROM (  
    SELECT AD,  
           DtAppello,  
           round(avg(Voto), 2) AS Media_Voti,  
           round(CAST (sum(Superamento) AS REAL) /  
                 CAST (sum(Superamento + Insufficienza + Ritiro) AS REAL), 2) AS Rapporto  
    FROM bos_denormalizzato  
    WHERE Assenza = 0  
    GROUP BY AD, DtAppello  
    ORDER BY AD  
) AS t1  
WHERE Rapporto < 0.5 AND Media_Voti >= 18  
ORDER BY AD, Rapporto DESC;
```



	ad	dtappello	appcod	Media_Voti	Rapporto
1	[A5810003] SCIENZA DELLE FINANZE	NULL	19	22.25	0.44
2	[A5810003] SCIENZA DELLE FINANZE	2017-07-21	21	23.43	0.44
3	[A5810010] DIRITTO AMMINISTRATIVO	NULL	51	27	0.33
4	[A5810010] DIRITTO AMMINISTRATIVO	NULL	53	25.5	0.25
5	[A5810015] DIRITTO PENALE	2017-07-11	100	25.25	0.42
6	[A5810017] DIRITTO PROCESSUALE CIVILE	NULL	119	25	0.33
7	[A5810019] ISTITUZIONI DI DIRITTO PRIVATO	2017-09-11	148	23.89	0.45
8	[A5810056] DIRITTO FALLIMENTARE	2017-07-19	325	28	0.33
9	[A5810057] DIRITTO INDUSTRIALE	NULL	337	23	0.33
10	[A5810062] DIRITTO PENALE MINORILE	2017-09-26	382	29	0.4

	AD	DtAppello	Media_Voti	Rapporto
1	[A5810003] SCIENZA DELLE FINANZE	2017-06-09	22.25	0.44
2	[A5810010] DIRITTO AMMINISTRATIVO	2017-02-08	27	0.33
3	[A5810010] DIRITTO AMMINISTRATIVO	2017-05-03	25.5	0.25
4	[A5810015] DIRITTO PENALE	2017-07-11	25.25	0.42
5	[A5810056] DIRITTO FALLIMENTARE	2017-07-19	28	0.33
6	[A5810057] DIRITTO INDUSTRIALE	2017-05-04	23	0.33
7	[A5810062] DIRITTO PENALE MINORILE	2017-09-26	29	0.4
8	[E0201Q001] MATEMATICA	2017-02-22	22.56	0.26
9	[E0201Q001] MATEMATICA	2017-09-13	26	0.11
10	[E0201Q005] CHIMICA ORGANICA	2017-10-16	24.2	0.42

In this query, *bos_denormalizzato* returns fewer observations (401 observations) because in the first query, it excludes some observations that, if included, would give me a ratio greater than 0.5. In fact, in the first query with 452 observations, two instances of 'Scienza delle Finanze' are noticeable, one of which has the same 'Media_Voti' and the same 'Rapporto'.

