


Corso: Data management and visualization

Quiz: DMV Exam 29 January 2025



PORTALE
ESAMI

	Nunzio Licalzi s344860
Iniziato	29 gennaio 2025, 08:35
Stato	Completato
Terminato	29 gennaio 2025, 10:05
Tempo impiegato	1 ora 30 min.
Valutazione	29,30 su un massimo di 31,00 (95%)
Riepilogo del tentativo	<div><div>1</div><div>2</div><div>3</div><div>4</div><div>5</div><div>6</div><div>7</div><div>8</div><div>9</div><div>10</div><div>11</div><div>12</div><div>13</div><div>14</div><div>15</div><div>16</div><div>17</div><div>18</div></div>

Domanda 1

Risposta
corretta

Punteggio
ottenuto 1,50
su 1,50

What is the primary purpose of the MapReduce programming model?

Scegli un'alternativa:

- ☐ a. To manage distributed file systems
- ☐ b. To process data from both relational and non-relational databases
- ☒ c. To process large datasets in a distributed environment ✓
- ☐ d. To enable the mapping of words into Large Language Models while reducing their computational costs
- ☐ e. To provide a novel and innovative data management solution in two simple functions
- ☐ f. None of the other answers.

Risposta corretta.

La risposta corretta è: To process large datasets in a distributed environment

Domanda 2

Risposta
corretta

Punteggio
ottenuto 1,50
su 1,50

Which of the following is NOT a common characteristic of distributed databases?

Scegli un'alternativa:

- ☐ a. Data is spread across multiple nodes or servers.
- ☐ b. They are designed to improve performance by distributing the workload.
- ☒ c. They are designed to guarantee consistency. ✓
- ☐ d. None of the other answers.
- ☐ e. They are designed to scale horizontally.
- ☐ f. They are designed to provide fault tolerance.

Risposta corretta.

La risposta corretta è: They are designed to guarantee consistency.

Domanda 3

Risposta
corretta

Punteggio
ottenuto 1,00
su 1,00

In data visualization, which of the following statements accurately reflects a best practice for using a double (dual) scale Y-axis?

Scegli un'alternativa:

- ☐ a. They work best when the two metrics being compared are measured in the same unit.
- ☐ b. They can be used interchangeably with a single-axis chart and require no additional labeling.
- ☐ c. They should always be used when comparing two related metrics, regardless of their scales.

- ☒ d. They are appropriate when the two metrics being displayed have distinct scales but a meaningful relationship, requiring simultaneous comparison. ✓
- ☐ e. They only make sense if you have more than two measures in the same chart.

Risposta corretta.

La risposta corretta è: They are appropriate when the two metrics being displayed have distinct scales but a meaningful relationship, requiring simultaneous comparison.

Domanda 4

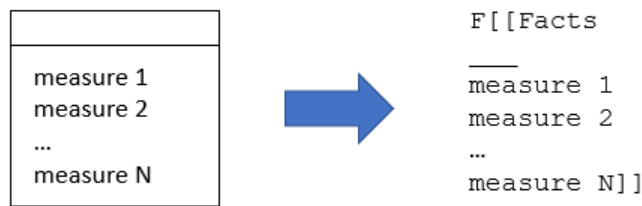
Completo

Punteggio
ottenuto 3,50
su 4,00

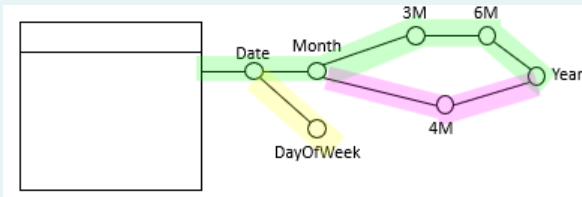
The question starts below.

Conceptual Schema: You must use the formalism discussed during the lectures to solve the conceptual schema.

Declare the fact table as



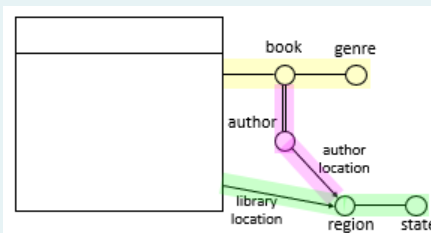
The attributes as arcs starting from the fact table (**F**), using multiple lines for each dimension or path, for instance:



becomes

```
F --- date --- month --- 3M --- 6M --- year
Date --- DayOftheWeek
Month --- 4M --- Year
```

If you need multiple arcs and/or texts for the arcs:



```
F --- book --- genre
book == author -- (author location) -> region
F -- (library location) -> region --- state
```

The European Union (EU) is working to develop a unified voting platform to modernize election systems and ensure consistency, transparency, and efficiency across all member states. This platform will support various types of elections, including government elections at different levels. You are required to design a data warehouse to analyse election outcomes, specifically the number of votes, according to the following specifications.

- Each vote is cast for a specific **election**. Elections are categorized by geographic scope: EU-wide, national, regional, or local. Each election is classified as either a ballot-based vote or not. An election can be either a direct representation (e.g., presidential) or a proportional representation (e.g., parliamentary).
- Each vote is associated with **geographic** attributes regardless of the election. These include, countries, regions, subregions (e.g., provinces), municipalities, and EU electoral districts. Each municipality belongs to one subregion only. EU electoral districts are administrative areas where representatives are elected. In most countries, the entire nation serves as one EU electoral district, while others divide their territory into multiple EU electoral districts. An EU electoral district is always within a single country. You can assume subregions are not split among different EU electoral districts, hence a subregion belongs to one EU electoral district only.
- A **voter** casts each vote. Voters are tracked by age group ("18-35", "35-60", or ">60"), gender, and residence status, in full compliance with data privacy regulations.
- Each vote is cast in support of a **political party**. Each party belongs to one of the EU political groups. Each party is linked to many political issues (e.g., environmental care, freedom, borders, etc.) if explicitly mentioned in its political program. The number of political issues to be tracked is not known in advance and the list of political issues to be tracked can grow over time. Note that the system does not track individual candidate names (representatives) who received votes.
- Each vote is cast at a given **time**. The platform tracks voting periods using multiple time references: month, quarter, year, and the EU Election Cycle. EU Election Cycle is a sequential number representing each term of the EU Commission since its inception, with changes occurring every 4 to 5 years. An EU election cycle starts or ends in a given quarter, for instance, suppose that EU election cycle number 12 starts in Q1-2020 and ends in Q3-2024.

Write the textual formalism to describe the described conceptual schema.

```

F[[
FACTS
-----
Tot_votes
]]

F---ELECTION --- GEOGRAPHIC_SCOPE
ELECTION --- BALLOT
ELECTION ---DIRECT

F--- VOTE --- MUNICIPALITY ---SUBREGION --- REGIONS --- COUNTRY
SUBREGIONS --- EU_DISTRICT --- COUNTRY
F---VOTER_AGE
F---GENDER
F---RESIDENCE_STATUS

F---POLITICAL_PARTY === POLITICAL_ISSUES

F---DATE --- MONTH --- QUARTER --- YEAR
QUARTER--- ELECTION_CYCLE

```

```

F[**VOTES**
number_of_votes
]]
F --- ElectionJunk --- GeographicScope
ElectionJunk --- isBallot
ElectionJunk --- Representation
F --- Municipality --- Subregion --- Region --- Country
Subregion --- EUelectoralDistrict --- Country
F --- PoliticalParty --- EUGroup
PoliticalParty === Ideology
F --- Month --- Quarter --- Year
Quarter --- ElectionCycle
F --- VoterJunk --- AgeGroup
VoterJunk --- Gender
VoterJunk --- ResidenceStatus

```

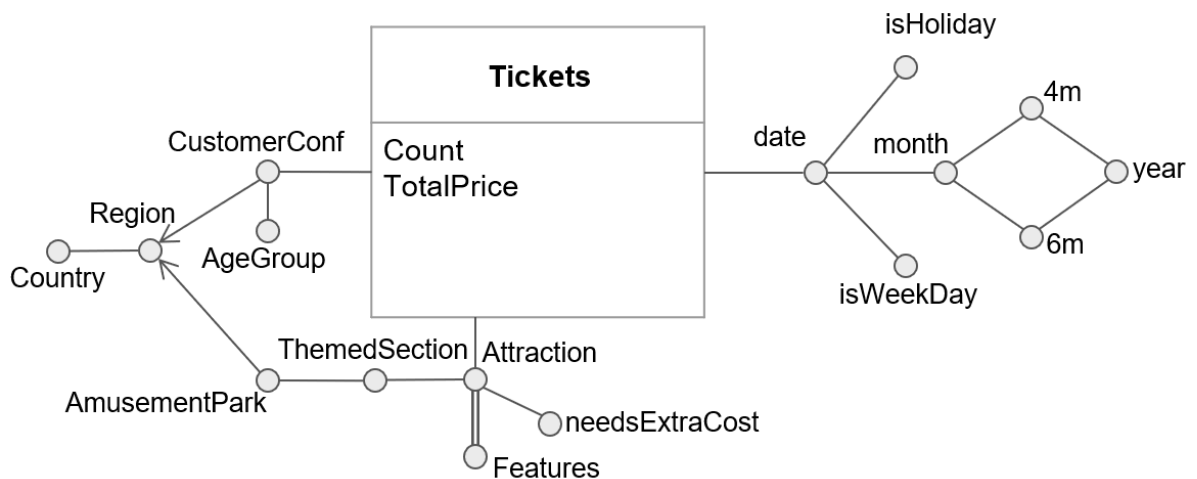
Commento:
Vote entity

Domanda 5

Completo

Punteggio
ottenuto 0,50
su 1,00

Given the following conceptual schema:



- Age group can be either "<3", "3-12", "12-18", "18-65", or ">65"
- Features can assume multiple values such as roller coasters, water rides, and family. The list of possible features is long and may eventually increase over time.

Provide the logical design of the conceptual DW schema indicated in the picture.

Write each table on a new line.

Use the **bold** or the underline for identifying primary-key attributes.

TIME(**DATEID**, date, isHoliday, isWeekDay, month, 4m, 6m, year)

CUSTOMER(**CUSTID**, customerConf, ageGroup, regId)

REGION(**regId**, region, country)

ATTRACTION(**ATT_ID**, attraction, needsExtraCost, ThemedSection, AmusementPark, regId)

ATTRACTION-HAS-FEATURE(**ATT_ID**, **FEATURE**)

TICKETS(**DATEID**, **CUSTID**, **ATT_ID**, count, totPrice)

Tickets(TimeID, CustomerConfID, AttractionID, Count, TotalPrice)
Time(TimeID, Date, isHoliday, isWeekDay, Month, 4m, 6m, Year)
CostumerConf(CostumerConfID, AgeGroup, LocationID)
Attraction(AttractionID, needsExtraCost, FeaturesID, ThemedSection, AmusementPark, LocationID)
Features(FeaturesID, FeatureNames)
Location(LocationID, Region, Country)

Commento:
Error using CUSTOMER as an entity

Domanda 6

Completo

Punteggio
ottenuto 3,50
su 4,00

```
Fact(TimeID, DonorID, ReceiverID, CategoryID, AidAmount, NumberOfTransactions)
Time(TimeID, Date, DayOfWeek, Month, 3m, 6m, Year)
Entity(EntityID, Country, Type, Continent, Association)
Category(CategoryID, HealthCare, Education, Infrastructure, Agriculture)
```

- DonorID and ReceiverID are IDs from Entity table, i.e., EntityID can be a DonorID or a ReceiverID.

Separately for each donor's **country** and **year**, select:

1. The average aid amount per transaction.
2. Assign a rank to each country separately for each year, based on the ratio of its total aid amount to the number of unique entities (rank 1st the highest).
3. The percentage of the total yearly aid amount for each country with respect to the total aid amount across all countries of the same continent during the same period.

Write the corresponding SQL query.

```
SELECT YEAR, COUNTRY, CONTINENT,
       SUM(AIDAMOUNT)/SUM(NumberOfTransactions) AS A,
       RANK() OVER (ORDER BY SUM(AIDAMOUNT)/COUNT(DISTINCT ENTITY.ID) DESC) AS B,
       100*SUM(AIDAMOUNT)/SUM(SUM(AIDAMOUNT)) OVER (PARTITION BY YEAR, CONTINENT) AS C
FROM FACT AS F, TIME AS T, ENTITY AS E
WHERE F.TIMEID=T.TIMEID AND F.DONORID=E.ENTITYID
GROUP BY YEAR, COUNTRY, CONTINENT;
```

```

SELECT country, year,
       SUM(AidAmount) / SUM(NumberOfTransactions) AS A,
       RANK() OVER (
         PARTITION BY year ORDER BY SUM(AidAmount) / COUNT(DISTINCT EntityID) DESC) AS B,
       100 * SUM(AidAmount) / SUM(SUM(AidAmount)) OVER (PARTITION BY year, continent) AS C,
FROM Fact F, Time T, Entity E
WHERE F.TimeID = T.TimeID AND F.DonorID = E.EntityID
GROUP BY country, continent, year;

```

Commento:
Partition error in B

Domanda 7

Completo

Punteggio
ottenuto 4,00
su 4,00

```

Fact(TimeID, DonorID, ReceiverID, CategoryID, AidAmount, NumberOfTransactions)
Time(TimeID, Date, DayOfWeek, Month, 3m, 6m, Year)
Entity(EntityID, Country, Type, Continent, Association)
Category(CategoryID, HealthCare, Education, Infrastructure, Agriculture)

```

- DonorID and ReceiverID are IDs from Entity table, i.e., EntityID can be a DonorID or a ReceiverID.

Separately for each combination of **categories** and **quarter**, select:

- The cumulative aid amount since the beginning of the year.
- The daily average aid amount.
- The percentage of quarterly aid amount for each category combination with respect to the total quarterly aid amount across all categories.

Write the corresponding SQL query.

```

SELECT C.CATEGORYID, 3M, YEAR,
       SUM(SUM(AIDAMOUNT)) OVER (PARTITION BY C.CATEGORYID, YEAR ORDER BY 3M ROWS UNBOUNDED
PRECEDING) AS A,
       SUM(AIDAMOUNT)/COUNT(DISTINCT DATE) AS B,
       100*SUM(AIDAMOUNT)/SUM(SUM(AIDAMOUNT)) OVER (PARTITION BY 3M) AS C
FROM FACT AS F, TIME AS T, CATEGORY AS C
WHERE F.TIMEID=T.TIMEID AND C.CATEGORYID=F.CATEGORYID
GROUP BY 3M, C.CATEGORYID, YEAR;

```

```

SELECT country, year,
       SUM(AidAmount) / SUM(NumberOfTransactions) AS A,
       RANK() OVER (
         PARTITION BY year ORDER BY SUM(AidAmount) / COUNT(DISTINCT EntityID) DESC) AS B,
       100 * SUM(AidAmount) / SUM(SUM(AidAmount)) OVER (PARTITION BY year, continent) AS C,
FROM Fact F, Time T, Entity E
WHERE F.TimeID = T.TimeID AND F.DonorID = E.EntityID
GROUP BY country, continent, year;

```

Commento:

Domanda 8

Completo

Punteggio
ottenuto 4,00
su 4,00

Design a MongoDB database to store the student enrollment to various university degree courses.

For each course, its name (e.g., “Data Science and Engineering”), the university name (e.g., “Politecnico di Torino”), the total number of credits (e.g., 120), the expected standard duration in years (e.g., 2) and some key competences (e.g., programming, database design, SQL queries, data warehousing, NoSQL design, etc.) are recorded.

Students are characterised by their student ID (matricola), their first name, last name, official email address, private email address, and different addresses. For each address, you are required to track the country, city, postal code, street, the type of address (e.g., official residence, fiscal address, temporary stay, etc.), its validity period (e.g., from 31/12/2023 to 31/12/2025, please note that the period end can be absent, i.e., the period is open ended), and its state (e.g., “current”, “past”, “invalid”, “pending”). Multiple addresses of the same type can be stored for a student.

Enrollments of students to courses are characterised by the student ID, the course name, the enrollment date, the enrollment status label, e.g., “active”, “graduated”, etc., the status change date, e.g., active until, graduated on, etc, and the list of exams passed with their grade, number of credits, and date of the exam. Each student can be enrolled in multiple courses at the same time, however most student enroll in a single degree course, and the maximum number is limited and low. Instead, a course can have virtually an unlimited number of students enrolled.

To analyse the student enrollments and their careers, you are required to efficiently retrieve all the exams passed by a specific student, together with the exam name, the exam date, the number of credits, and the grade. Given a student, you are also required to efficiently provide his/her average grade for each degree course he/she is enrolled in. Given a degree course, you are required to efficiently provide the average grade of all the enrolled students.

Write a sample document for each database collection.

Important: In addition to the example documents, explicitly state the design patterns used.

```
COURSE: {
  _id: <>,
  name: <>,
  universityName: <>,
  totCFU: <>,
  expectedDuration: <>,
  competences: [<>, <>, ....],
  avgGradeStudents: <>
}
```

```
STUDENTS: {
  _id: <Matricola>,
  name: <>,
  surname: <>,
  emailOfficial: <>,
  emailPrivate: <>,
  addresses: [{ type: <>, country: <>, city: <>, postalCode: <>, street: <>, validityStart: <>, validityEnd: <>, state: <> }],
  enrollment: [{ course: <>, enrollmentDate: <>, status: <>, changeDate: <>, cid: <> }],
  passedExam: [{ name: <>, grade: <>, weight: <>, date: <> }],
  stats: [{ cid: <>, name: <>, avgGrade: <> }]
}
```

pattern:

- Attribute
- Extended reference
- precomputed

COURSES

```
{
  "_id": "<string> or <ObjectID>",
  "name": "Data Science and Engineering",
  "university": "Politecnico di Torino",
  "credits": 120,
  "duration": 2,
  "competences": ["programming", "database", ...],
  "avg_grade": 26.78 // COMPUTED PATTERN
}
```

```
{
  "_id": 123456,
  "firstname": <string>,
  "lastname": <string>,
  "email":
  [ "official@email.edu.com", // first is the official
    "private@email.me" // second is the private
  ], // this allows a single index on the email field
  "addresses": [ // ATTRIBUTE PATTERN
    { "country": "Italy",
      "city": "Turin",
      "postalcode": "10129",
      "street": "Corso Duca 29",
      "type": "fiscal",
      "state": "current",
      "validity_from": <datetime>,
      "validity_to": <datetime>,
    },
    { // another address...
    },
    ...
  ],
  "enrolled_in": [ // EXTENDED REFERENCE + ATTRIBUTE PATTERN
    { "_id": "course_id",
      "name": "Data Science and Engineering",
      "enrollment_date": <datetime>,
      "status_label": "graduated",
      "status_date": <datetime>,
      "avg_grade": 28.56, // COMPUTED PATTERN
      "exams": [ // ATTRIBUTE PATTERN
        { "name": "Data Management and Visualization",
          "grade": 30,
          "honors": true,
          "date": <datetime>
        },
        { "name": "Data Management and Visualization",
          "grade": 30,
          "honors": true,
          "date": <datetime>
        },
        { "name": "Data Management and Visualization",
          "grade": 30,
          "honors": true,
          "date": <datetime>
        },
        { // another exam
        },
        ...
      ]
    },
    { // another course the student is enrolled in
    },
    ...
  ]
}
```

Commento:

Domanda 9

Completo

Punteggio
ottenuto 2,00
su 2,00

The following document structure represents a university course, with its name, the master degrees where it is taught, the lecture dates, the teachers, and the exam dates.

```
{
  "course_name": "Big Data programming and architectures",
  "master_degrees":
    ["Computer Engineering",
     "Mathematical Engineering",
     "ICT for Smart Societies"],
  "lectures": [
    {
      "teacher": "d12345",
      "date": "2024-09-23",
      "hours": 3.0,
      "type": "theory"},
    {
      "teacher": "d98765",
      "date": "2024-09-24",
      "hours": 3.0,
      "type": "laboratory"} ],
  "teachers" : [
    {
      "id" : "d12345",
      "name" : "John",
      "surname" : "Smith",
      "nationality": "UK"},
    {
      "id" : "d98765",
      "name" : "Jane",
      "surname" : "Brown",
      "nationality": "US"} ],
  "exams" :
    [
      {"date": "2024-09-16",
       "type": "online"},
      {"date": "2025-01-29",
       "type": "onsite"}]
}
```

Find all courses taught in the “Computer Engineering” degree, with at least one teacher from UK, and with at least a lecture of type “laboratory” in 2024 whose length was equal to or greater than 3 hours. Show only the course name and the full list of teachers.

```
db.find(
  {'master_degree': 'Computer Engineering', 'teacher': {'$elemMatch': {'nationality': 'UK'}}}, 'lectures':
  {'$elemMatch': {'type': 'Laboratory',
    'date': {'$gte': new Date("2024-01-01"), '$lte': new Date("2024-12-31")}, 'length': {'$gte': 3}}}},
  {'_id': 0, 'course_name': 1, 'teachers': 1}
)
```

```
db.exam250129.find (
{
  master_degrees: "Computer Engineering",
  "teachers.nationality": "UK",
  lectures: {
    $elemMatch: {
      type: "laboratory",
      date: { $gt: "2024", $lt: "2025" },
      hours: { $gte: 3 }
    }
  }
},
{_id:0, 'name':1, 'teachers':1}
)
```

Commento:

Domanda
10

Completo

Punteggio
ottenuto 3,00
su 3,00

The following document structure represents a university course, with its name, the master degrees where it is taught, the lecture dates, the teachers, and the exam dates.

```
{
  "course_name": "Big Data programming and architectures",
  "master_degrees":
    ["Computer Engineering",
     "Mathematical Engineering",
     "ICT for Smart Societies"],
  "lectures": [
    {
      "teacher": "d12345",
      "date": "2024-09-23",
      "hours": 3.0,
      "type": "theory"},
    {
      "teacher": "d98765",
      "date": "2024-09-24",
      "hours": 3.0,
      "type": "laboratory"} ],
  "teachers" : [
    {
      "id" : "d12345",
      "name" : "John",
      "surname" : "Smith",
      "nationality": "UK"},
    {
      "id" : "d98765",
      "name" : "Jane",
      "surname" : "Brown",
      "nationality": "US"} ],
  "exams" :
    [
      {"date": "2024-09-16",
       "type": "online"},
      {"date": "2025-01-29",
       "type": "onsite"}]
}
```

Considering only courses with exams after 2020, separately for each teacher and type of lecture, calculate the total number of hours taught. Sort the results in descending order by total number of hours.

```
db.aggregate([
  {'$match':{'exams':{'$elemMatch':{'date':{'$gte': new Date('2021-01-01')}} } }},
  {'$unwind':'$lectures'},
  {'$group':{'_id':{'teacher':'$lectures.teacher', 'type':'$lectures.type'}, 'totHours':{'$sum':'$lectures.hours'}}}
  {'$sort':{'totHours':-1}}
])
```

```
[
  {
    $match: {
      "exams.date": {
        $gt: "2020-01-01"
      }
    }
  },
  {
    $unwind: "$lectures"
  },
  {
    $group: {
      _id: {
        teacher: "$lectures.teacher",
        type: "$lectures.type"
      },
      totalHours: {
        $sum: "$lectures.hours"
      }
    }
  },
  {
    $sort: {
      totalHours: -1
    }
  }
]
```

Commento:

**Domanda
11**

Risposta
corretta

Punteggio
ottenuto 0,25
su 0,25



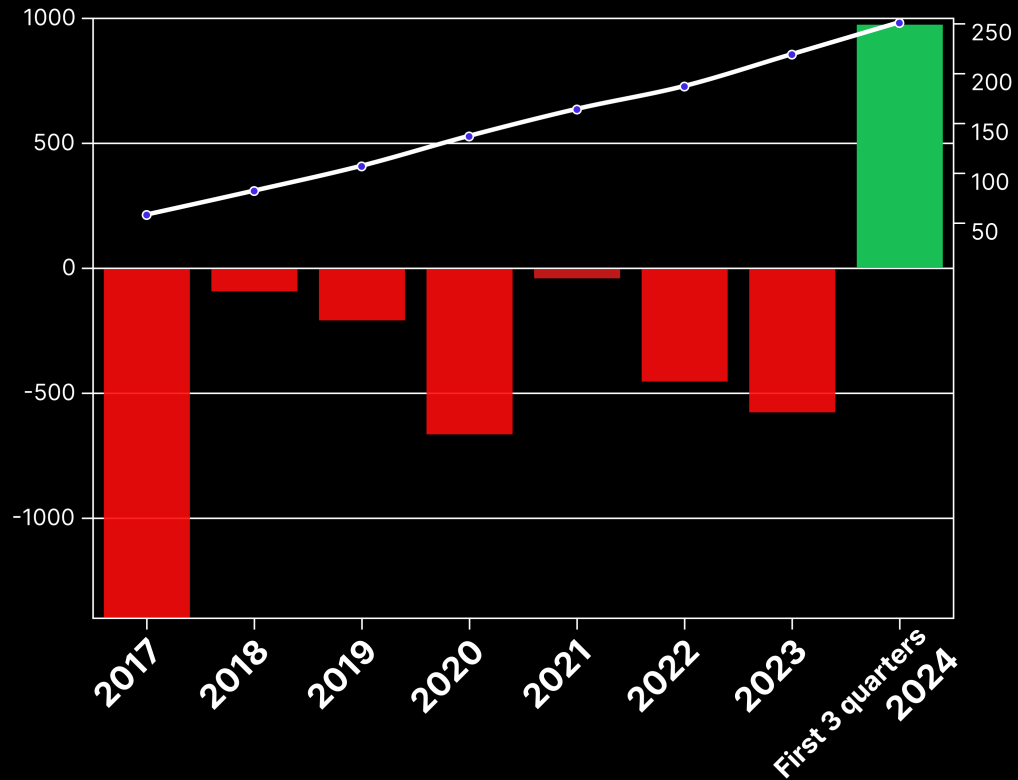
Spotify Races To Its First Profitable Year

Spotify's profit and subscriber numbers have been growing, and the company is on track to be profitable for the full year of 2024

Created by genuine
impact

Profit (€m)

Subscribers (m)



Source: Spotify

More charts: genuineimpact.substack.com

Question

Which one of the following questions represents the purpose of this visualization?

Scegli un'alternativa:

- ☐ a. Does Spotify generate more revenue from advertising or from premium subscriptions?
- ☐ b. How many monthly active users does Spotify have compared to other streaming services in 2023?
- ☐ c. Which streaming platform has the widest global music catalog, and how has its size changed over time?
- ☒ d. How have Spotify's profit and subscriber numbers evolved over time, and when might the company achieve full-year profitability? ✓
- ☐ e. How has Spotify's user base grown specifically in the European market versus North America?

Risposta corretta.

La risposta corretta è:

How have Spotify's profit and subscriber numbers evolved over time, and when might the company achieve full-year profitability?

**Domanda
12**

Risposta
corretta

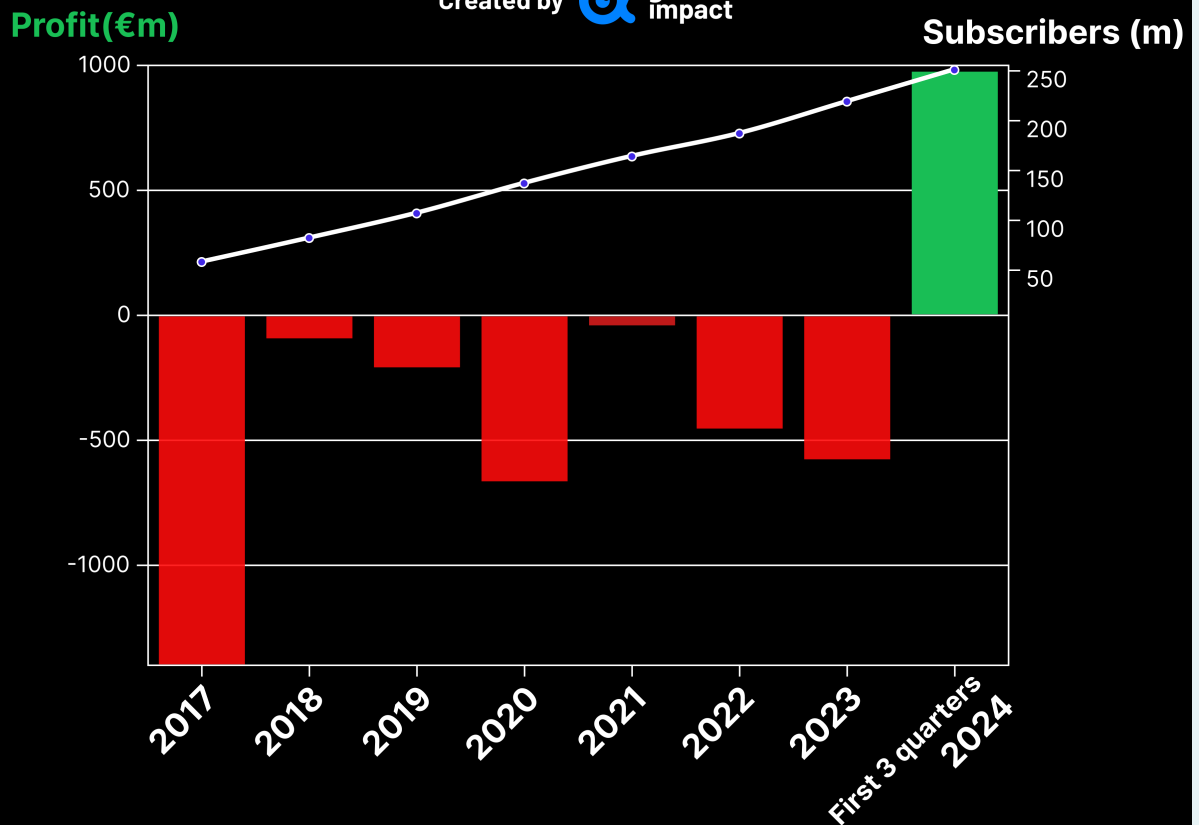
Punteggio
ottenuto 1,25
su 1,25



Spotify Races To Its First Profitable Year

Spotify's profit and subscriber numbers have been growing, and the company is on track to be profitable for the full year of 2024

Created by genuine impact



Source: Spotify

More charts: genuineimpact.substack.com

Data

Is the data quality appropriate? Select true answers only.

Scegli una o più alternative:

- ☐ a. The chart's profit data shows random jumps with no clear timeline, indicating inconsistent reporting.
- ☐ b. The chart lacks any numerical scale, so viewers have no idea what the profit and subscriber figures represent.
- ☐ c. The visualization is incomplete because it doesn't compare Spotify's profit to other streaming platforms.
- ☐ d. Because the chart does not differentiate between free and premium subscribers, the subscriber data is not accurate.
- ☒ e. By labeling the Y-axes with "Profit (€m)" and "Subscribers (m)", the chart demonstrates enough clarity to interpret these measures correctly. ✓ Correct, the units (millions of euros for profit and millions of subscribers) are clearly stated, reducing ambiguity.
- ☒ f. Since the profit figures come from Spotify's own financial reports, the chart's source is credible. ✓ Correct, official company financial statements are generally considered reliable primary sources.

- ☒ g. The chart's time series is consistent from 2017 to 2023, making it straightforward to observe changes in profits and subscriber counts year by year. ✔ Correct, each bar or data point corresponds to a specific year or time period, so there are no gaps or out-of-order data.
- ☐ h. This chart doesn't include data beyond 2020, so it's out of date and lacks currency.
- ☐ i. The data cannot be trusted because Spotify never publicly discloses subscriber figures.
- ☒ j. Including both historical results (2017–2023) and a 2024 forecast makes the chart fairly current, illustrating the latest trajectory of Spotify's performance. ✔ Correct, the visualization extends through the first three quarters of 2023 and includes a projection for 2024, reflecting up-to-date (or near-future) information on Spotify's progress.

Risposta corretta.

Le risposte corrette sono: The chart's time series is consistent from 2017 to 2023, making it straightforward to observe changes in profits and subscriber counts year by year.,

Since the profit figures come from Spotify's own financial reports, the chart's source is credible.,

By labeling the Y-axes with "Profit (€m)" and "Subscribers (m)", the chart demonstrates enough clarity to interpret these measures correctly.,

Including both historical results (2017–2023) and a 2024 forecast makes the chart fairly current, illustrating the latest trajectory of Spotify's performance.

Domanda
13

Completo

Punteggio
ottenuto 0,75
su 0,75



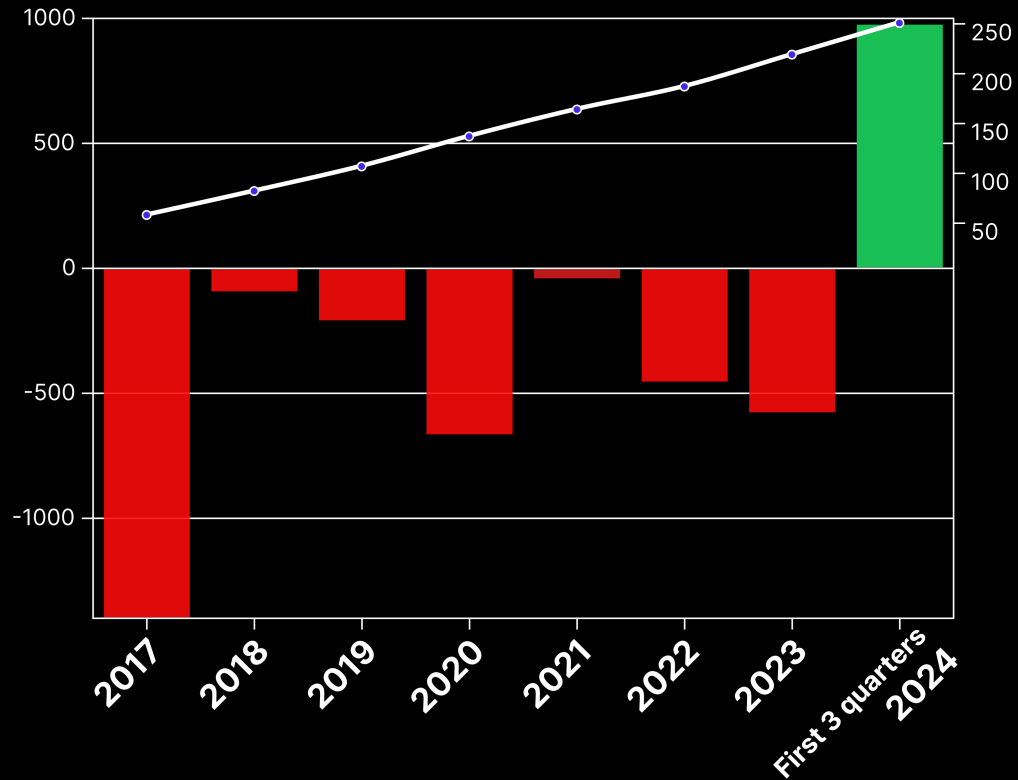
Spotify Races To Its First Profitable Year

Spotify's profit and subscriber numbers have been growing, and the company is on track to be profitable for the full year of 2024

Created by genuine impact

Profit(€m)

Subscribers (m)



Source: Spotify

More charts: genuineimpact.substack.com

Visual Proportionality

Are the values encoded in a uniformly proportional way?

the bar lenght is overall okay and values are belivable, moreover, the axis correctly start from 0 for the bar plot and does not need to for the time series, the unit of measure is overall accurate and the correct number of digits is given. measures can be compared quite easily wrt the time siries and a little less easily with respect to the profit measure, even if't a clarity issue i would asdd val

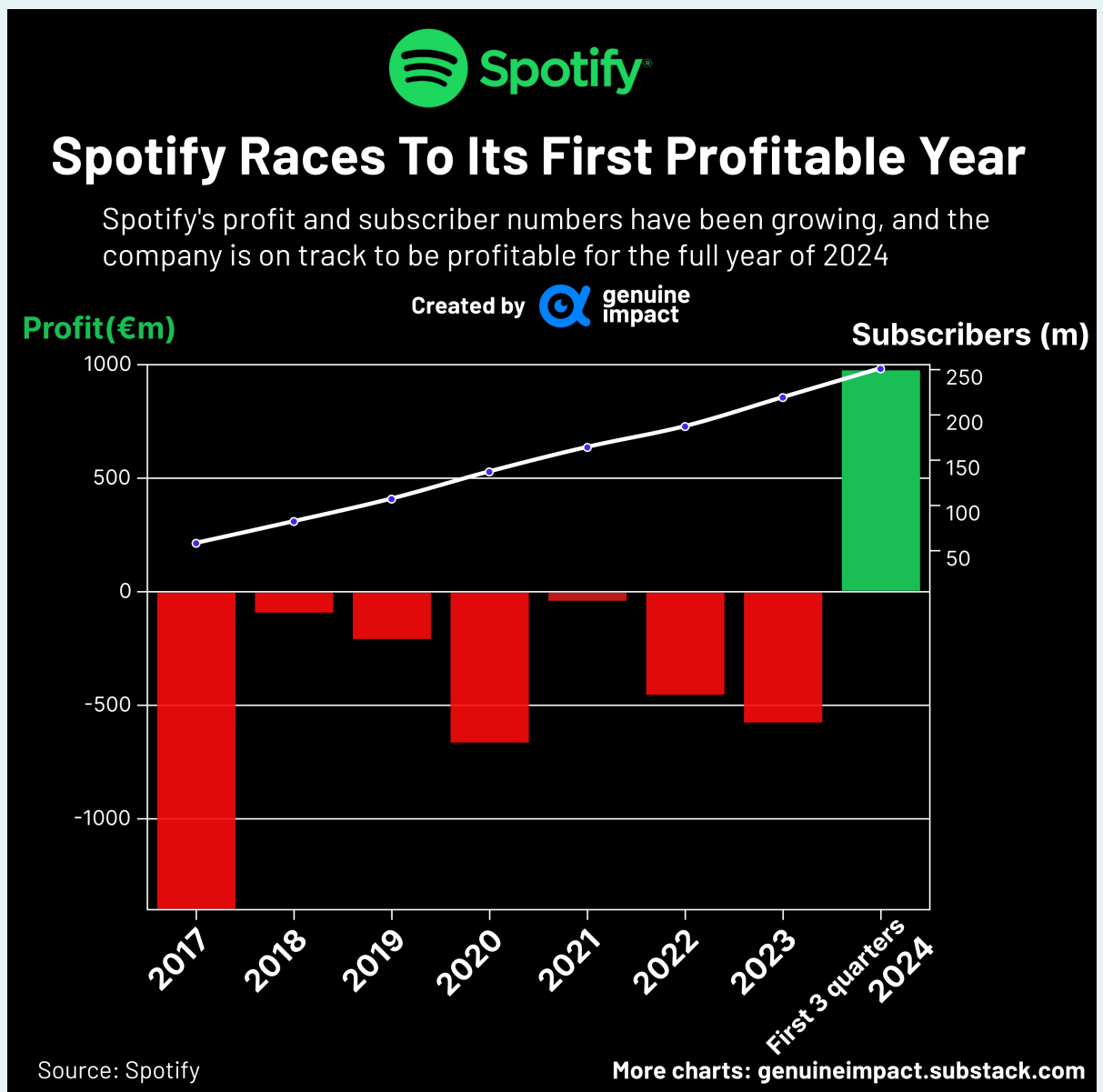
The chart uses two separate vertical axes (one for profit in millions of euros on the left, one for subscribers in millions on the right), and each axis is scaled consistently for its respective metric. The bars (profits) and the line (subscribers) both start at zero or a defined baseline on their own axes, ensuring that within each axis, the height or position of the data is proportional to its value.

Commento:

Domanda
14

Completo

Punteggio
ottenuto 0,75
su 0,75



Visual Utility

All the elements in the graph convey useful information?

the background does not provide visual utility, neither does the more charts text. We can also remove the created by text and logo.

we can also remove spotify logo and green text on top of the chart and the subtitle of the graph can also be removed

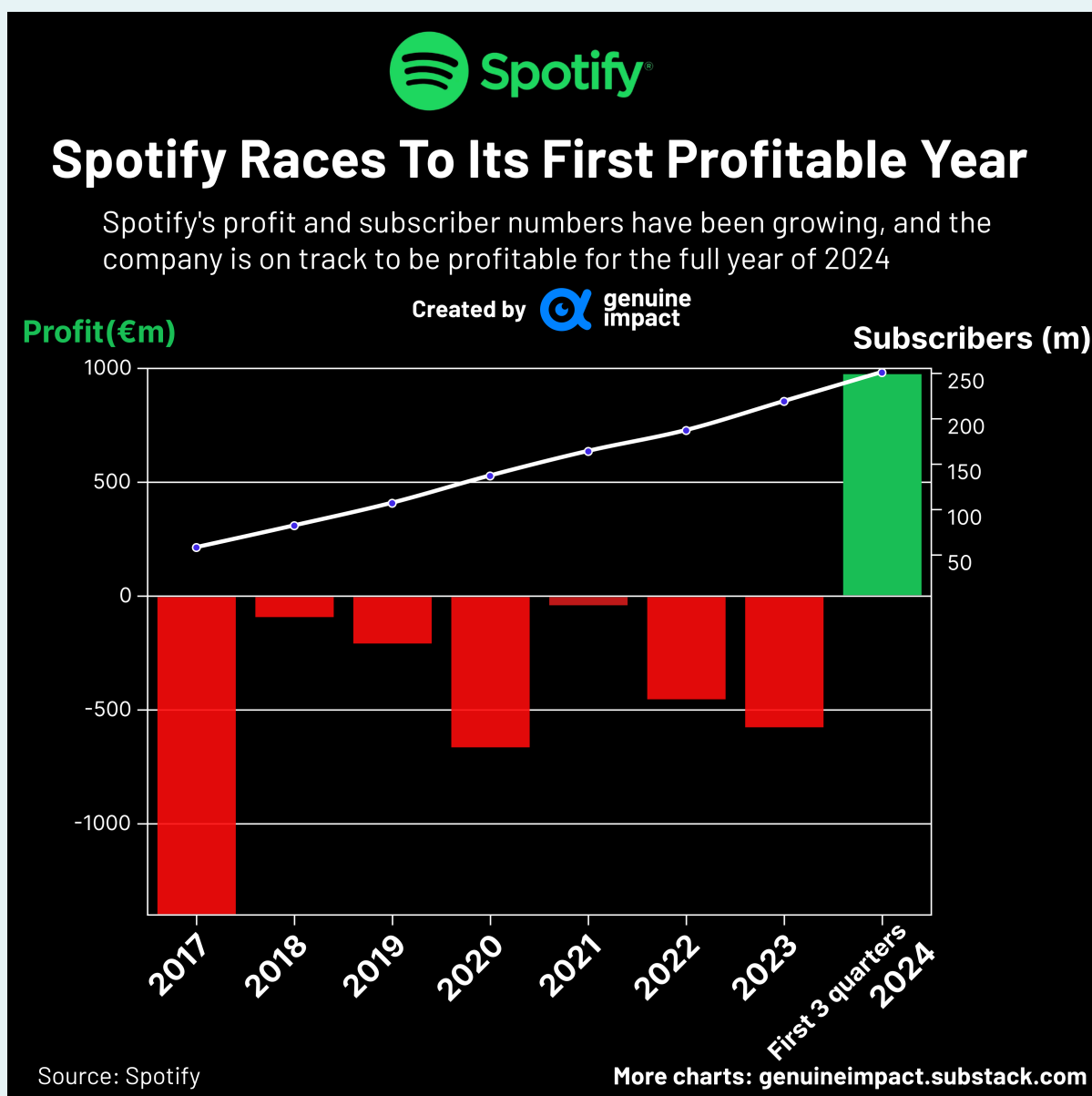
Bars for profit/loss, the line for subscriber counts, labeled axes, and the timeline (2017-2024) are all essential for understanding Spotify's financial performance and subscriber growth. Some visual elements (like brand logos) do add contextual details, but in theory, they can be removed without losing critical data or interpretive value.

Commento:

Domanda
15

Completo

Punteggio
ottenuto 0,30
su 0,50



Visual Clarity

Are the data in the graph clearly identifiable and understandable (properly described)?

the color usage whilst helping with the general trend of the bars is mostly useless and can be removed. The alignment is overall okay,

and the bars are pretty uniform with respect to their size.

Each axis is labeled with metric and unit ("Profit (€m)" vs. "Subscribers (m)"), which helps distinguish the bars from the line. The fact that subscribers cannot be negative also clarifies which axis goes with which series: the bars show profit or loss (which can drop below zero), while the line represents subscribers (always above zero). That said, using two separate y-axes in one chart sometimes requires extra care from viewers to avoid confusion.

Commento:
Double scale

**Domanda
16**

Risposta
corretta

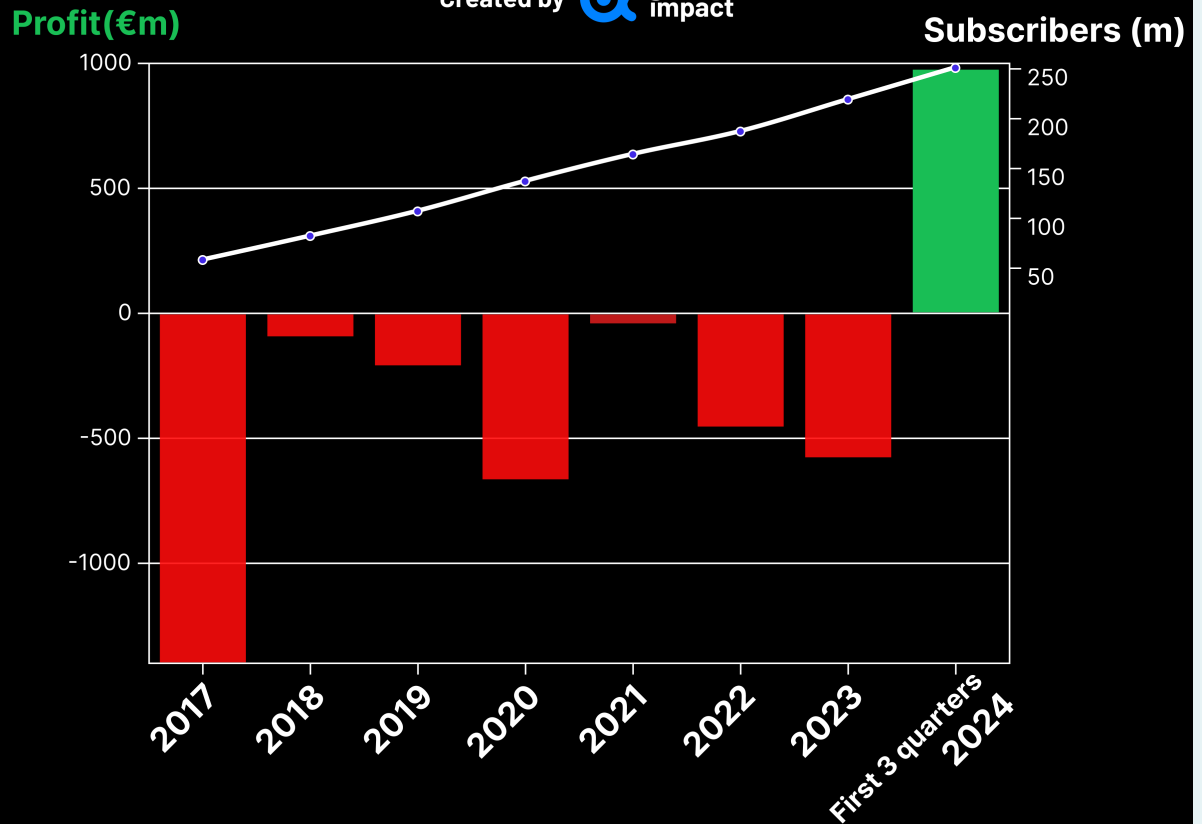
Punteggio
ottenuto 0,25
su 0,25



Spotify Races To Its First Profitable Year

Spotify's profit and subscriber numbers have been growing, and the company is on track to be profitable for the full year of 2024

Created by genuine impact



Source: Spotify

More charts: genuineimpact.substack.com

Design data

Design the visualization based on the following data structure.

PROFIT	Measure	✓
SUBSCRIBERS	Measure	✓
YEAR	Dimension	✓

Risposta corretta.

La risposta corretta è: PROFIT → Measure,
SUBSCRIBERS → Measure, YEAR → Dimension

**Domanda
17**

Completo

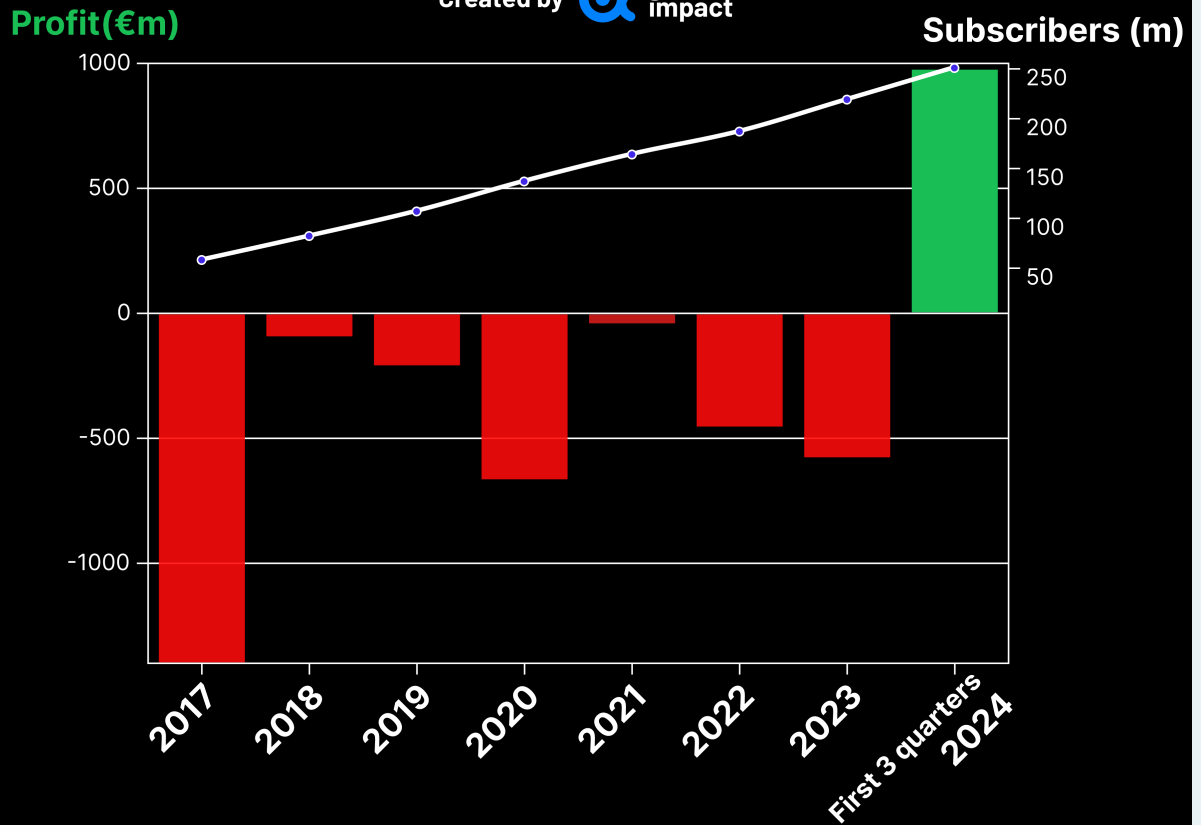
Punteggio
ottenuto 1,25
su 1,25



Spotify Races To Its First Profitable Year

Spotify's profit and subscriber numbers have been growing, and the company is on track to be profitable for the full year of 2024

Created by genuine impact



Source: Spotify

More charts: genuineimpact.substack.com

Design schema & Sketch

Fill in the required schema elements; formulas can be used if required. Then describe in words the design proposal.

Schema	Details
Columns	year(year)
Rows	sum(profit), sum(subscribers)
Graph type	bar
Color	red if profit else blue
Size	default
Label	sum(profit), sum(subscribers)

Design proposal: i would like to do a vertical bar chart where we compare next to one another the number of subscribers with the number of profit, the two bars will be different colors depending if they are one or the other, and a numeric label should appear in both columns. i think that this can be done because even if measuring two different quantities the two scales are similar (both in the million ranges) and as such if we add a legend to tell whether it's euros or people it could work

Schema Details

Columns	YEAR
Rows	SUM(PROFIT), SUM(SUBSCRIBERS)
Graph type	Bar, Line
Color	Orange, Blue
Size	Default
Label	SUM(PROFIT), SUM(SUBSCRIBERS)

Design proposal: The proposed redesign consists of two separate charts: one focused on profit and the other on subscribers. Placing them side by side and sharing a common timeline on the horizontal axis reduces confusion and enables viewers to instantly distinguish between the two data series. Removing unnecessary brand logos further reduces visual clutter. Clear labeling enhances readability, with each chart featuring "Year" on the horizontal axis and either "Profit (€m)" or "Subscribers (m)" on the vertical axis. Distinct color coding, such as orange bars for profit and blue lines for subscribers, also helps viewers associate each chart with its respective metric. Finally, if a forecast for 2024 is included, it should be noted that these figures are projections rather than historical data.

Commento:

Domanda**18**

Risposta non
data

Non valutata

This is a blank question to be used as your personal notepad during the exam.

Anything written here will NOT be evaluated.